



HAL
open science

[Report] Paratyphoid fever and relapsing fever in 1812 Napoleon's devastated army

Rémi Barbieri, Julien Fumey, Helja Kabral, Christiana Lyn Scheib, Michel Signoli,
Caroline Costedoat, Nicolás Rascovan

► **To cite this version:**

Rémi Barbieri, Julien Fumey, Helja Kabral, Christiana Lyn Scheib, Michel Signoli, et al.. [Report] Paratyphoid fever and relapsing fever in 1812 Napoleon's devastated army. *Current Biology*, 2025, 35, pp.1-8. <10.1016/j.cub.2025.09.047>. <pasteur-05331040v2>

HAL Id: pasteur-05331040

<https://pasteur.hal.science/pasteur-05331040v2>

Submitted on 27 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

Paratyphoid Fever and Relapsing Fever in 1812 Napoleon's Devastated Army

Rémi Barbieri¹, Julien Fumey^{1,2}, Helja Kabral³, Christiana Lyn Scheib^{3,4}, Michel Signoli⁵, Caroline Costedoat⁵, Nicolás Rascovan^{1,6*}.

1. Institut Pasteur, Université de Paris Cité, CNRS UMR 2000, Microbial Paleogenomics Unit, F-75015 Paris, France

2. Institut Pasteur, Université Paris Cité, Bioinformatics and Biostatistics Hub, F-75015 Paris, France

3. Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu 51010, Estonia

4. Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, United Kingdom

5. Aix Marseille Université, CNRS, EFS, ADES, Pierre Dramard Boulevard, 13015 Marseille, France

6. Lead contact

*Correspondence : nicolas.rascovan@pasteur.fr

Summary

During Napoleon's retreat from Russia in 1812¹, countless soldiers of the French army succumbed to infectious diseases, but the responsible pathogen or pathogens remain debated²⁻⁵. We recovered and sequenced ancient DNA from the teeth of 13 soldiers who, based on historical records, likely died from infectious diseases, aiming to identify the pathogens responsible for their deaths⁶. Our results confirmed the presence of *Salmonella enterica* subsp *enterica* belonging to the lineage Para C, the causative agent of paratyphoid fever⁷; and *Borrelia recurrentis*, responsible for relapsing fever transmitted by body lice⁸. We were not able to detect *Rickettsia prowazekii* (the agent of typhus) and *Bartonella quintana* (the cause of trench fever), which had previously been associated with this deadly event, based on PCR results and historical symptom descriptions³. The presence of these previously unsuspected pathogens in these soldiers reveals that they could have contributed to the devastation of Napoleon's Grande Armée during its disastrous retreat in 1812.

Results and Discussion

In June of 1812, Napoleon I, the French emperor, assembled a military force of about 500,000 to 600,000 soldiers to invade Russia⁹. After arriving in Moscow without decisively defeating the Russian army, the Napoleonic forces, finding themselves isolated in a ruined city, opted to initiate a retreat and to establish winter encampments along the border with Poland in October that year. The retreat from Russia spanned from October 19th to December 14th 1812¹⁰ and resulted in the loss of nearly the entire Napoleonic army. According to historians, it wasn't the harassment from the Russian army that claimed the lives of about 300,000 men², but rather the harsh cold of the Russian winter, coupled with hunger and diseases. A doctor during the Russian campaign, J.R.L. de Kirckhoff, authored a book detailing the illnesses that afflicted soldiers in 1812. Specifically, he documented the prevalence of typhus, diarrhea, dysentery, fevers, pneumonia, and jaundice⁵. Other physicians¹¹, as well as officers¹², made similar observations about the illnesses affecting soldiers. Different infectious diseases, such as typhus, have been described in French regiments even before the start of the Russian Campaign¹³. Typhus, in particular, which is commonly referred to as camp fever due to its frequent association with armies, has long been suspected of being the main infectious cause of the demise of the Grande Armée in 1812. This assumption was fueled by the discovery of body lice, the main vector of typhus, among the remains of Napoleonic soldiers who perished during the Great Retreat from Russia in December 1812 in Vilnius, Lithuania, as well as by the alleged identification of *R. prowazekii* and *B. quintana* sequences amplified by nested PCR in some of these individuals³. However, this earlier study was limited by the technologies available at the time and relied solely on the amplification of two short DNA fragments (192 and 429 base pairs long), which did not offer sufficient resolution to provide unambiguous evidence for the presence of these pathogens in Napoleon's army. Several years later, another study successfully detected *Anelloviridae* viral ancient DNA (aDNA) in Napoleonic soldiers recovered in Kaliningrad in 1812⁴, but these viruses are ubiquitous and asymptomatic in human populations, and therefore unrelated to the fatal fate of these soldiers. Using *state-of-the-art* aDNA methodologies, we reanalyzed samples from Napoleonic soldiers who died in Vilnius and identified pathogen-specific genetic material, providing direct evidence of infectious agents that may have contributed to the army's collapse.

Sequencing data from 13 teeth from Vilnius (Figure 1), each comprising approximately 20 million reads, were first taxonomically classified using KrakenUniq¹⁴ against the full microbial NCBI database. We then screened for known human pathogens by comparing the identified TaxIDs against a curated list of 535 TaxIDs representing 185 bacterial species known to be pathogenic to humans, retrieved from the PATRIC database (see Document S2)¹⁵. From this initial analysis, we generated a list of taxa with more than 200 uniquely assigned reads to known pathogens in at least one sample, a threshold established as optimal for ancient metagenomic pathogen detection in Pochon et al. 2023¹⁶. This criterion yielded 14 candidate taxa (Table S1). These were further evaluated based on their *k-mer* counts, duplication rates and coverage estimated by KrakenUniq, and prioritized according to their known epidemiological relevance and their plausibility given the early 19th-century European military context and historically documented symptomatology (e.g., fever, diarrhea, jaundice). These steps identified two candidate pathogens with consistent signals: *Salmonella enterica* in sample YYY087A (268 hits) and *Borrelia recurrentis* in sample YYY093A (239 hits) (Tables S1-S3). Notably, after all screening and authentication steps (see STAR Methods), we did not identify any reliable reads attributable to *R. prowazekii* (the agent of typhus) or *B. quintana* (the cause of trench fever), two pathogens previously reported in individuals from this site using PCR-based methods³. However, the absence of authenticated reads from these two species in our dataset does not preclude their presence at this site or during this historical event, especially given the degraded nature of ancient DNA and potential variability in pathogen load among individuals.

68
69 To assess the presence of *S. enterica*, we examined both the number of reads and their alignment profiles across
70 different serovars. In four individuals (YYY087A, YYY092B, YYY095A, YYY097B), between 34 and 968 unique reads
71 (i.e., after duplicates removal) mapped most closely and in greater numbers to the *Paratyphi C* serovar (strain RKS4594), as
72 opposed to *Typhi* or *Typhimurium* (Table 1, Figure 2A). The corresponding edit distance distributions displayed the expected
73 declining pattern consistent with a genuine ancient pathogen signal, and a negative difference proportion of 1.0 (NDP, Table
74 S2)¹⁷. In contrast, other samples with low read counts mapping to *S. enterica* exhibited flatter or irregular edit distance
75 profiles, and low NDP values and were considered negative hits, likely originating from environmental related species or
76 conserved genomic regions with limited discriminatory power between taxa (Table S2). A similar analysis was performed for
77 *Borrelia* species. In individual YYY093A, 4,062 unique reads mapped to *B. recurrentis*, compared to 1,556 and 1,441 reads
78 mapping to *B. duttonii* and *B. crociduræ*, respectively (Table S1, Figure 3A). Reads mapped across all eight fragments of *B.*
79 *recurrentis* genome (the chromosome and seven plasmids), and the edit distance profiles again showed a consistent declining
80 pattern and a NDP=1 (Figure 3A, Table S3)¹⁷. A second individual (YYY092B), who also tested positive for *S. enterica*,
81 yielded 322 unique reads mapping to *B. recurrentis* and an edit distance distribution similarly consistent with a true hit
82 (NDP=1 in Table S3, Figure 3A). All edit distance profiles and downstream analyses were conducted using deduplicated and
83 high-mapping-quality reads (MAPQ > 30).
84

85 Given the low number of reads, the breadth of coverage ranged from 0.0003 to 0.009 for *S. enterica* and from 0.002
86 to 0.14 for *B. recurrentis*. The read-length distribution, with an average of 56 bp for *B. recurrentis* and 43 bp for *S. enterica*,
87 indicated extensive DNA fragmentation, consistent with aDNA degradation patterns¹⁸. Analysis of post-mortem damage in
88 the 4,062 reads from sample YYY093A that aligned to the modern *B. recurrentis* reference genome (strain A1) revealed that
89 approximately 5% displayed cytosine deamination at read termini (Figure S1C). This is slightly lower than the ~8% observed
90 in human DNA from the same sample (Figure S1D), but within the same range as other individuals from the site (Figure
91 S1E-H) or aDNA samples from the same period¹⁹, and is thus consistent with authentic bacterial aDNA. It is worth noting
92 that lower deamination rates in bacterial compared to human aDNA within the same host remains have been reported
93 previously, suggesting that bacterial DNA can degrade at a different rate and often less extensively than human DNA²⁰⁻²².
94 Unfortunately, for *S. enterica*, the number of mapping reads was too low to reliably assess deamination patterns (Table 1), as
95 several hundred aligned reads are typically needed for accurate estimations.
96

97 To further authenticate the reads assigned to both pathogens, we re-analyzed the deduplicated mapping reads
98 against the full NCBI nt database using BLASTN, followed by taxonomic classification with the LCA algorithm
99 implemented in MEGAN²³ (Figures S1A-B; Tables S2 and S3). This analysis confirmed the initial assignments, with a large
100 proportion of reads attributed to the expected pathogens, while also filtering out those that could not be unambiguously
101 assigned to the corresponding taxa. These results reinforced the identification of four positive samples for *S. enterica*
102 (YYY087A, YYY092B, YYY095A, YYY097B) and one (YYY093A), possibly two (YYY092B, with only 18 on-target hits
103 out of 322 mapped reads), for *B. recurrentis* (Tables S2 and S3).
104

105 As a final authentication step, we analyzed the on-target reads identified via BLASTN and MEGAN (B+M) using a
106 phylogenetic placement strategy previously applied to low-coverage ancient genomes^{24,25}. To establish a robust reference
107 framework, we first reconstructed whole-genome phylogenies of *S. enterica* spp. and *B. recurrentis*, including both ancient
108 and modern strains representative of the known diversity for each species²⁶⁻²⁹. The methodologies used were similar to those
109 of the original studies, and followed standard practices in ancient pathogen genomics (see STAR Methods), successfully
110 reproducing previously published tree topologies with high bootstrap support. We then applied epa-ng³⁰ to evaluate the most
111 likely placement of the ancient strains, based on the positions covered by B+M on-target reads across the multiple sequence
112 alignment (MSAs) used to build the reference trees (Tables 1 and S4). EPA-ng places low-coverage genomes onto a fixed
113 reference phylogeny and multiple sequence alignment (MSA) by computing the most likely position for each sample based
114 on its covered genomic sites. This approach is well suited to ancient DNA, where only a small fraction of the genome is
115 typically recovered, and often enables confident lineage assignment even under sparse coverage.
116

117 For *B. recurrentis*, on-target reads from sample YYY093A aligned to 125,514 positions in the MSA (of which
118 5,854 were variant sites), while sample YYY092B aligned to 584 positions covering 21 variant sites, out of a total of
119 1,032,378 positions in the MSA. For *S. enterica*, the on-target reads from YYY087A, YYY092B, YYY095A, and YYY097B
120 aligned to 26,116, 4,864, 1,236 and 1,382 positions, respectively, and covered 800, 140, 44 and 42 variant sites out of a total
121 MSA length of 4,083,601 positions (see Table 1). Our final *S. enterica* phylogeny consisted of 10 ancient and 20 modern
122 published genomes representative of the known diversity of the species, in which we were able to place the sequences from
123 the four positive Vilnius individuals (YYY087A, YYY092B, YYY095A, YYY097B) within the paratyphi C serovar, a clade
124 that also includes three ancient genomes dating from the 14th to 16th centuries³¹⁻³⁶ (Figures 2B and Figure S2). Together,
125 phylogenetic placements and the genotypes at variant positions (Figure S4A-D) add a much stronger support to our initial
126 serovar estimations. However, despite the consistent placement of all strains within the same serovar, the resolution of our
127 data was not sufficient to determine a stronger affiliation with a specific ancient or modern branch within *Paratyphi C*.

128 For *B. recurrentis*, we reconstructed a reference phylogenetic tree using four ancient and eleven modern publicly
129 available genomes from across the *Borrelia* genus. Our 19th-century strain YYY093A was confidently placed (LWR = 0.89)
130 in a basal position relative to two other ancient strains from the 14th–15th centuries (C10976 and OSL9)^{27,29} as well as all
131 modern *B. recurrentis* genomes (Figures 3B, Figure S3A-B, Figure S4E-F). To further explore *B. recurrentis* diversity, we
132 applied our phylogenetic placement approach to two previously published ancient strains that had not been analyzed with this
133 technique. Sample C11907, recovered from England, dated to 1288–1461 calCE and covering ~32% of the genome (Table

134 S4), was also placed in a basal position relative to C10976 and OSL9 (LWR = 0.81), consistent with its original placement by
135 Swali et al.²⁹ where it clustered near the Iron Age genome C10416 (300–100 BCE). Strikingly, YYY093A exhibited an
136 almost identical phylogenetic placement to C11907 (Figures 3B, Figure S3A-B), and when placed on a tree that included the
137 C11907 partial genome, YYY093A clustered at the base of this strain’s branch with very high confidence (LWR = 0.99,
138 Figure S3C-D).

139 To further investigate the phylogenetic positioning of YYY093A, we built a *de novo* phylogenetic tree restricted to
140 the 144,470 genome positions covered by reads classified as *Borrelia* genus by B+M (Figure S3E-F). In this analysis,
141 YYY093A and C11907 formed a monophyletic clade much closer to the Iron Age genome C10416 than to C10976 or OSL9.
142 Since the bootstrap support for this topology was relatively low (58%), we decided to examine the source of this uncertainty
143 using SplitsTree³⁷. By visualizing the 100 bootstrapped trees as a consensus network, we observed that the topological
144 ambiguity was constrained to the subtree encompassing C10416, C11907, and YYY093A, rather than being spread across the
145 entire phylogeny (Figure 3C). This result suggests that YYY093A and C11907 consistently fall within the same evolutionary
146 neighborhood as the Iron Age strain, supporting that these genomes represent a distinct and relatively ancient lineage or
147 differentiated group of *B. recurrentis* that persisted in Europe for at least two millennia and remained in circulation into at
148 least the early 1800s.

149 By contrast, two lower-coverage strains (YYY092B (this study) and Las Gobas (8th–11th century CE, Spain²⁸))
150 were placed within the clade of the near-identical modern *B. recurrentis* genomes (Figure 3B, Figure S3A,B,D, Figure S4G).
151 While the placement of Las Gobas genome appears well supported (50,657 mapped positions; 2,357 variant sites, Table S4),
152 YYY092B aligned to only 584 positions and covered 21 variants (Table 1), limiting placement resolution. However, the
153 consistency in placement across both strains and the relatively robust signal from Las Gobas suggest that these genomes may
154 indeed represent members of the clade grouping the available modern genomes. Importantly, while the low coverage of
155 YYY092B precludes definitive sub-lineage assignment, it nonetheless supports its classification at the species level.
156 Altogether, our findings support the existence of multiple distinct *B. recurrentis* lineages circulating in Europe during the
157 past centuries.

158 Recent aDNA work has revealed that paratyphoid fevers have been present in Europe for millennia^{26,31–36}, and it
159 was already well known and documented by 1812. The disease is transmitted to man through food or water contaminated
160 with infected feces, and symptoms include fever, headache, rash, weakness, loss of appetite, diarrhea, constipation, stomach
161 pain and vomiting³⁸. Throughout Napoleon's Russian campaign, paratyphoid or typhoid fever was not mentioned in any
162 historical sources of our knowledge, likely due to these nonspecific and varied symptoms. Yet, an 1812 report from J.R.L. de
163 Kirckhoff, a physician serving in Napoleon's army, contains key information about the events that could potentially explain
164 the origins of an epidemic. In this document, he specifies that soldiers suffered from typhus, dysentery, and diarrhea on their
165 arrival in Vilnius. Insisting on this last aspect, he wrote “Diarrhea was common among us in Lithuania. One powerful
166 contributing factor to this illness was that we encountered in almost every house, from Orcha to Wilna, large barrels of salted
167 beets (*buraki kwaszone*), which we ate and drank the juice of when we were thirsty, greatly upsetting us and strongly
168 irritating the intestinal tract”⁵. This description could be consistent with both the characteristics of a paratyphoid fever
169 infection caused by contaminated food, and the digestive symptoms typically associated with the disease, although we
170 acknowledge that they can also match various other diseases that were common in the 19th-century in Europe. Furthermore,
171 even today, two centuries later, it would still be impossible to perform a differential diagnosis between typhus, typhoid or
172 paratyphoid fever based solely on the symptoms or the testimonies of survivors.

173
174 From a molecular perspective, our findings provide strong support that the soldiers were infected with paratyphoid
175 fever caused by *S. enterica* Paratyphi C. Although we did not recover sufficient genome coverage to determine the specific
176 phylogenetic positioning within the known diversity of this lineage, our thorough authentication workflow resulted in a solid
177 proof of its presence. Our study thus provides the first direct evidence that paratyphoid fever contributed to the deaths of
178 Napoleonic soldiers during their catastrophic retreat from Russia. However, the limited number of samples that were
179 processed (n=13), in relation to the large number of reported bodies in this site (over 3000), is not sufficient evidence to
180 conclude that this pathogen alone contributed to all the deaths at the site. Considering the extreme and harsh conditions that
181 characterized this retreat, the presence of multiple overlapping infections is highly plausible. Typhus has long been reported
182 to have affected Russia in this period, but the existing evidence remains inconclusive to support a role of this disease in the
183 devastation of Napoleon’s Army.

184
185 In light of our results, a reasonable scenario for the deaths of these soldiers would be a combination of fatigue, cold,
186 and several diseases, including paratyphoid fever and louse-borne relapsing fever. While not necessarily fatal, the louse borne
187 relapsing fever could significantly weaken an already exhausted individual. Our study confirms the presence of two
188 previously undocumented pathogens, but the analysis of a larger number of samples will be necessary to fully understand the
189 spectrum of epidemic diseases that impacted the Napoleonic army during the Russian retreat. Our work demonstrates that
190 high-throughput sequencing of ancient DNA is a powerful approach for investigating historical disease dynamics and
191 underscores its capacity to accurately identify ancient pathogens, even when only limited genomic data are available.

192
193 **Resource availability**

194
195 **Lead contact**

196

197 Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Nicolás
198 Rascovan.

199
200
201 **Material availability**

202 This study did not generate any new reagents

203
204
205 **Data and code availability**

- 206
- 207 • Raw sequencing data from the 13 sequenced individuals have been deposited at the SRA Archive under the
208 bioproject PRJNA1188378.
 - 209 • The complete source code used in this study is available from GitHub
210 (<https://github.com/Metapaleo/Napoleon1812>).
 - 211 • Any additional information required to reanalyze the data reported in this paper is available from the lead contact
212 upon request.
- 213

214 **Acknowledgments**

215 We want to thank the HPC Core Facility from Institut Pasteur for their support for computational analyses and Marc Monot
216 and Laurence Motreff from the Biomix Platform (supported by France Génomique (ANR-10-INBS-09-09 and IBISA) for
217 their assistance in sample sequencing. This project was made possible through the following funding sources: ERC-2020-
218 STG - PaleoMetAmerica – 948800 (NR), Institut Pasteur and CNRS UMR funding (NR) and INCEPTION program
219 (Investissement d’Avenir grant ANR-16-CONV-0005, NR).

220
221 **Author’s contribution**

222 RB: conceptualization, methodology, validation, investigation, data curation, original draft preparation, writing, review and
223 editing, visualization, project administration; JF: methodology, validation, investigation, data curation, writing, review and
224 editing, visualization; HK: methodology, validation, investigation, writing, review and editing, visualization; CLS:
225 methodology, validation, investigation, writing, review and editing, visualization; MS: conceptualization, methodology,
226 validation, investigation, resources, data curation, original draft preparation, writing, review and editing, visualization,
227 supervision, project administration, funding acquisition; CC: conceptualization, methodology, validation, investigation,
228 resources, data curation, original draft preparation, writing, review and editing, visualization, supervision, project
229 administration; NR: conceptualization, methodology, validation, investigation, writing, review and editing, visualization,
230 supervision, project administration, funding acquisition;

231 All authors have read and agreed to the published version of the manuscript.

232
233 **Declaration of interests:** The authors declare no competing interests.

234
235
236 **Figure titles and legends**

237
238 **Figure 1. Historical, geographical and archeological context of the study.** (A) Painting dating from 1851 entitled
239 “Napoleon’s retreat from Moscow” by Adolph Northen, depicting the conditions of the retreat of Napoleon’s army. (B)
240 Geographical map of Europe showing the location and dating of the archaeological site in Vilnius, Lithuania, from which the
241 samples in our study were collected. (C) In situ photographs taken during the excavation of the trenches containing the
242 bodies of Napoleonic soldiers. The top photograph shows the discovery of an imperial-type button of uniform in the mass
243 grave. The lower photograph shows a general view of the mass grave (credit photo, Michel Signoli).

244
245 **Figure 2. Authentication of *S. enterica* aDNA data.** (A) Distribution of edit distances of the YYY087A, YYY092B,
246 YYY095A and YYY097B reads, with the associated *negative difference proportion* ($-\Delta\%$)¹⁷ (B) Maximum likelihood
247 phylogeny of 31 *S. enterica* genomes representative of the known diversity of the species, calculated on 4,083,601 aligned
248 genome positions (including 91,636 variant sites). The YYY087A sample is represented by red stars, YYY092B by yellow
249 circles, YYY095A by blue triangles and YYY097B by green squares. The phylogenetic placements of these samples were
250 done using the epa-ng software³⁰ and are based on the partial genome sequence reconstructed from reads assigned to the *S.*
251 *enterica* species by MEGAN (see STAR methods). The previously published ancient genomes are in blue, and modern
252 genomes in black. The lineages are color-coded as follows: purple represents the paratyphi A lineage, blue corresponds to the
253 typhi lineage, green stands for the typhimurium lineage, yellow represents the Birkenhead lineage, and red indicates the para
254 C lineage. The purple dots indicate a bootstrap value of 1. Bootstrap symbols are displayed on the branch leading to each
255 supported node, following iTOL’s default parameters. The multiple symbol size represents different possible placements of a
256 single sample on the phylogenetic tree. Each symbol’s location indicates a probable placement, with varying sizes determined
257 by the like weight ratio and distal length values. See also Figure S1, Figure S2, Figure S4 and Table S4.

258
259 **Figure 3. Authentication of *B. recurrentis* aDNA data.** (A) Distribution of edit distances of YYY093A and YYY092B
260 mapping reads on three closely related *Borrelia* species, with the associated *negative difference proportion* ($-\Delta\%$)¹⁷ (B)
261 Maximum likelihood phylogeny of 15 *Borrelia* sp. genomes calculated on 1,032,378 aligned genome positions (including
262 39,217 variant sites). The phylogenetic placement done by epa-ng³⁰ is based on reads classified under the genus *Borrelia* by
263 MEGAN. The YYY093A and YYY092B are represented by a blue triangle and a red star respectively. Due to their low

264 coverage, the two previously published ancient genomes, C11907 and Las Gobas, have been incorporated in this plot by
 265 phylogenetic placement, similarly to our study samples, and are indicated by a yellow circle and a green square. The symbol
 266 sizes are determined by the like weight ratio and distal length values. The previously published ancient genome is written in
 267 blue font, and modern genomes in black. Numerical values correspond to branch bootstrap supports. Bootstrap values are
 268 displayed on the branch leading to each supported node, following iTOL's default parameters C) Consensus network
 269 constructed using SplitsTree³⁹ based on 100 bootstrapped trees built from an MSA of 144,470 sites that are covered by
 270 YYY093A on-target reads (5,854 variant sites), to assess the nature of the topological uncertainty in the placement of
 271 YYY093A (Figure S3). The network reveals that although bootstrap values are low for the YYY093A–C11907 clade (Figure
 272 S3E-F), together with C10416 the three strains are consistently grouped within a distinct cluster across the 100 bootstrapped
 273 trees. See also Figure S1, Figure S4 and Table S4.

274 **TABLE**

Pathogen	Sample ID	Unique mapped reads	Unique authenticated reads	Avg. read length (bp)	Total mapped positions	Non-N positions in epa-ng MSA	Variant sites covered in epa-ng MSA
<i>S. enterica</i>	YYY087A	968	874	45	29,725	26,116	800
	YYY092B	225	209	42	6,471	4,864	140
	YYY095A	34	29	58	1,374	1,236	44
	YYY097B	74	54	40	1,568	1,382	42
<i>B. recurrentis</i>	YYY093A	4,062	3,239	60	144,470	125,514	5,854
	YYY092B	322	21	45	718	584	21
	Las Gobas (*)	999	760	96	60,432	50,657	2,357
	C11907 (*)	14,027	13,235	44	312,604	270,532	13,392

277 **Table 1. Summary of sequencing, authentication, and phylogenetic placement statistics for ancient samples positive**
 278 **for *Salmonella enterica* and *Borrelia recurrentis*.** Four individuals yielded authenticated *S. enterica* DNA and two yielded
 279 *B. recurrentis* DNA. For the latter, we additionally reanalyzed two previously published low-coverage strains (*) using our
 280 authentication workflow. “Unique mapped reads” refers to the number of deduplicated reads aligned to the respective
 281 reference genome. “Authenticated reads” are the subset of unique reads taxonomically assigned to the target pathogen using
 282 BLASTN+MEGAN. Average read length of authenticated reads (in base pairs) reflects the extent of DNA fragmentation
 283 characteristic of ancient DNA. “Total mapped positions” denotes the number of genomic positions covered by at least one
 284 read. For phylogenetic placement, we report the number of positions covered by authenticated reads (i.e., non-N) in the
 285 multiple sequence alignment (MSA), and the number of variant (i.e., phylogenetically informative) sites covered. The MSAs
 286 used included 4,083,601 positions for *S. enterica* and 1,032,378 for *B. recurrentis*. To optimize taxonomic specificity,
 287 authenticated reads for *S. enterica* were limited to those classified at the species level, since public databases include a broad
 288 representation of the species’ genomic diversity and lowering the classification threshold would risk capturing environmental
 289 *Salmonella*. In contrast, *B. recurrentis* authenticated reads were collected at the genus level (*Borrelia*), given the narrow
 290 representation of species-level diversity in current databases and the potential for ancient strains to fall outside the reference-
 291 defined clade. Additional mapping and authentication metrics –including read duplication rates, total mapped reads, and
 292 species-level assignments for *B. recurrentis*– are provided in Supplementary Tables S3 and S4. See also Table S1 and Table
 293 S2.

294
 295
 296

297 STAR Methods

298 299 Experimental model and subject details

300 301 Archeological context of samples

302 In autumn 2001, a mass grave was discovered during construction work in the northern sector of Vilnius. The mass grave
303 was excavated in several stages. A minimum of 3,269 individuals were exhumed from the site⁶. Preliminary observations
304 linked the mass grave to Napoleon I's Grande Armée, and specifically to the Russian retreat of December 1812. During that
305 year, a few days after crossing the Berezina (November 26-29), some 70,000 men arrived in Wilna (December 3). The
306 extreme cold and poor sanitary conditions of the army (famine, exhaustion, disease, low morale, etc.) led to a large number of
307 deaths in the town^{54,55}. The presence of epidemic diseases among these soldiers is widely described in the *memoirs* of many
308 of them: soldiers or officers^{12,55-57} and medical staff alike¹¹. Observations made during the excavation testify to the context of
309 demographic and health crisis: (i) the simultaneous burial of horses and men; (ii) the burial of bodies still dressed in their
310 military uniforms: gaiters on the shins; boot or shoe soles in contact with the foot bones; a soldier's shako in contact with a
311 soldier's skull; (iii) the absence of weapons attesting either to the disarmament of the bodies, but more likely still to the loss
312 of weapons, helmets and cuirasses during the retreat⁵⁸. Observations of the demographic structure (98.52% male, 85.61%
313 aged between 20 and 50 years) of this sample, combined with artefacts (buttons, remains of uniforms), confirmed the military
314 origin of the mass grave⁵⁷.

315
316 Unlike other burial sites from this period, knives or firearm traumas were not observed on any of the skeletons at the site.
317 This finding is in line with historical data (there was no fighting in Wilna in December 1812) and confirms that these soldiers
318 died from the combined effects of cold, hunger, exhaustion, and disease⁵⁹. In this study, we collected 13 intact teeth, each
319 corresponding to a different individual. The individuals examined in this study differ from those tested in the previous study
320 by Raoult et al³. All the five individuals that were positive for a pathogen were males, three in the 20-29 age bracket
321 (YYY095A, YYY097B and YYY093A) and one (YYY087A) in the 30-39 age bracket. We were unable to estimate the age
322 of skeleton YYY092B at the time of death, as it was too poorly preserved⁵⁷.

323 324 Sample information and ethical statement

325 All skeletal elements were sampled during the archaeological excavation with permissions from the representative
326 bodies/host institutions (Michel Signoli, director of the French team and Rimantas Jankauskas director of the Lithuanian
327 team). Samples were taken and processed to maximize research value and minimize destructive sampling. Molars were
328 prioritized due to their bigger roots and larger mass, but premolars and canine were also sampled.

329 330 Method details

331 332 Sampling, aDNA extraction and library preparation

333 Inside a class IIB hood in the dedicated aDNA facility of the University of Tartu, Institute of Genomics, root portions of teeth
334 were removed with a sterile drill wheel in parallel with an extraction blank. All samples and blank were decontaminated and
335 aDNA was extracted according to a publicly available online protocol that was extensively used in previous aDNA
336 studies^{60,61}. Library preparation was conducted using a protocol modified from the manufacturer's instructions included in
337 the NEBNext® Library Preparation Kit for 454 (E6070S, New England Biolabs, Ipswich, MA) that is also detailed in a
338 second publicly available online protocol⁶².

339 340 aDNA sequencing

341 aDNA libraries were sequenced using the Illumina NextSeq500/550 High Output paired-end 75-cycle kit at the University of
342 Tartu, Institute of Genomics DNA Sequencing Facility. The library of individual YYY093A that was positive for *B.*
343 *recurrentis* was sequenced at higher depth by paired-end 2x150 using an Illumina NovaSeq 6000 instrument through the
344 Biomix Platform of Institut Pasteur.

345 346 Preliminary metagenomic screening and pathogen detection

347 Raw sequencing reads were quality filtered, the remaining adapter sequences were removed and overlapping R1 and R2
348 reads collapsed, using AdapterRemoval v2.3.1⁴¹ (minimum length of 30 bp, allowing up to 3 mismatches for adapter
349 detection, and trimming bases with quality scores below 2.). Quality trimmed reads were initially analyzed with KrakenUniq
350 v0.5.8¹⁴ using the full microbial database from NCBI, that was built following author's recommendations. Taxonomic
351 classification reports were screened to identify potential hits over 535 TaxIDs representing 185 human-pathogenic bacterial
352 species, retrieved from the PATRIC database (see Document S2)¹⁵. A manual assessment of KrakenUniq hits was also
353 conducted for eukaryotes and DNA viruses. A pathogen species was considered a potential hit when at least 200 reads were
354 assigned to its TaxID in at least one sample, and ensuring that both the number and complexity of *k-mers* (i.e., distinct *k-mers*
355 and duplication estimates), as well as the observed/estimated vs. expected coverage, were consistent with expectations for a
356 true positive result (a criterion well described by Pochon et al. 2023¹⁶) (Table S1). Among the identified pathogens, only
357 *Salmonella enterica* subsp. *enterica* and *Borrelia recurrentis*, were selected as potential candidates due to their classification
358 as highly infectious and transmissible and also for being common pathogens in humans that could potentially explain the
359 symptoms documented in the historical records. *Yersinia pseudotuberculosis* was an additional potential hit that was
360 discarded during the authentication process.

361 362 Authentication step 1: mapping on reference genomes with strict parameters

363 The presence of the two pathogens was more carefully investigated in all samples using a more rigorous strategy. Quality-
364 trimmed reads from all the 13 samples were mapped to the reference genomes of different serovars of *S. enterica* and the
365 closest species to *B. recurrentis* using bwa aln (v0.7.18 with option -l 1024) and BAM files were filtered with samtools⁴³
366 (v1.21 with option -q 30). The targeted reference genomes included *S. enterica subsp. enterica serovar paratyphi C* (str.
367 RKS4594), *S. enterica subsp. enterica ser. typhimurium* (str. LT2), *S. enterica subsp. enterica ser. typhi* (strain
368 343077_213147), and the chromosomal assemblies of *B. recurrentis* (str. A1), *B. duttonii* (str. CR2A), and *B. crocidurae* (str.
369 Achema) in addition to the pl23, pl33, pl35, pl37, pl53, pl124, pl6, plasmids of *B. recurrentis* (str. A1) and the pla1 and pla2
370 plasmids of *B. crocidurae* (str. Achema). The edit distances of the mapping reads from each sample to each of these target
371 genomes were then evaluated to determine the closest genome, and the negative difference proportion ($-\Delta\%$, as defined by
372 Hübler et al.¹⁷) was computed as the ratio between the sum of declining steps in the edit distance histogram, and the total sum
373 of all steps, providing a quantitative measure of whether the read mismatch distribution decreased monotonically as expected
374 for authentic ancient DNA. The distribution of mapped reads across the closest reference (i.e., observed vs. expected breadth
375 of coverage) was also used as authentication criteria.

377 **Authentication step 2: filtering out off-target mapping reads with BlastN+MEGAN**

378 Reads mapping to the closest target were then blasted against the full NT NCBI database (Blast+ v2.16.0 with the following
379 parameters : blastn -task blastn-short -max_target_seqs 25 -evalue 1E-05), a much slower, but much more sensitive method
380 than KrakenUniq, and results were processed and visualized using MEGAN software⁴⁵ (version 6.24.6, built on 16 Nov
381 2022), retaining only reads that were summarized at the expected taxon. For *Salmonella enterica*, we tested assignments at
382 both the species level (*S. enterica*) and the subspecies/serovar level (*S. enterica subsp. enterica*), to account for varying
383 classification resolution and potential database biases. For *Borrelia recurrentis*, we compared assignments at both the species
384 level and the genus level (*Borrelia*). Due to the very poor genomic representation of *B. recurrentis* in public databases
385 (limited to a handful of nearly identical strains) reads from ancient or divergent lineages may fail to meet species-level
386 assignment thresholds and therefore be erroneously excluded. This can result in artifactual results in downstream
387 phylogenetic analyses, with an attraction toward modern lineages. To mitigate this bias and better capture uncharacterized
388 ancient diversity, we used genus-level (*Borrelia*) summarized reads for downstream analyses. The proportion of pathogen
389 reads over total raw reads was too low to possibly obtain high genome coverage by simply sequencing libraries at higher
390 depth. Targeted enrichment was not attempted due to the extremely high costs of the necessary kits.

392 **Building reference phylogenetic trees of *S. enterica* and *Borrelia* species**

393 Phylogenetic analyses including a selected number of ancient and modern strains of *S. enterica subsp. enterica* and *Borrelia*
394 sp., representative of the known diversity of each species, was run using our in-house Nextflow pipeline that follows the
395 standard methodology used in ancient pathogen genomics⁶³. The pipeline takes as input raw fastq files, which are trimmed
396 with AdapterRemoval v2.3.1⁴¹ and aligned to a reference genome with BWA v0.7.17⁴² (-l 1024 -t 28, q >30) against the
397 chromosomal and plasmid assemblies of the detected pathogen species. In case of having multiple libraries from the same
398 sample, the BAM files of all libraries are merged with Samtools v1.11⁴³. Duplicates are removed from sample's BAM files
399 with Picard v2.23.8⁴⁶. Mapping reads in ancient samples are rescaled using mapDamage to mask transitions that are due to
400 deamination processes during DNA decay. Variant calling and genotyping steps are performed with GATK v4.1.9.0⁴⁸, with
401 a minimum threshold of at least 4 supporting reads to call positions in both ancient and modern strains. These variants are
402 taken by BCFtools commit f27f849⁴⁵ to reconstruct a consensus sequence (i.e., the reference genome with the ALT alleles
403 placed in the respective position), and the genes of interest are extracted with BEDtools v2.29.2⁴⁹ based on a GFF provided
404 by the user. Once coding genes have been extracted, some filters are applied. If a gene has more than 10% Ns, it will be
405 discarded from that sample and replaced by Ns, and if a gene is discarded in more than 50% of the samples, it will be
406 discarded from all samples. The genes that passed these filters are concatenated with GoAlign v0.3.4⁵⁰. For *S. enterica*, we
407 used the whole genome instead of concatenated genes, to increase covered positions by our ultra-low coverage strains. The
408 resulting alignments were used to compute a maximum likelihood phylogenetic tree using RAxML⁵¹ using the
409 GTR+FO+G4+I model and 100 bootstrap replicates.

411 **Authentication step 3: phylogenetic validation of ultra-low-coverage strains**

412 To authenticate the presence of *Salmonella enterica* and *Borrelia recurrentis* in samples that yielded ultra-low genome
413 coverage, we implemented a phylogeny-informed approach based on phylogenetic placement. Since our ancient pathogen-
414 positive samples did not reach sufficient depth to be included in standard *de novo* phylogenetic reconstructions (Tables 1 and
415 S4), we used the evolutionary placement algorithm (EPA) implemented in epa-ng v0.3.8³⁰, a method previously used in
416 aDNA studies^{25,26,64,65}. We first compared the effect of using different subsets of reads for each pathogen, namely, i) all
417 mapping reads, and ii) those obtained from the previous BLAST+MEGAN filtering step (see Authentication Step 2). For *S.*
418 *enterica*, we tested reads taxonomically summarized at either the species level (*S. enterica*) or the subspecies/serovar level (*S.*
419 *enterica subsp. enterica*). For *B. recurrentis*, we compared reads summarized at the species level (*B. recurrentis*) versus the
420 genus level (*Borrelia*), the latter being preferable due to the limited diversity in the available reference genomes (Figure S3B
421 illustrates this issue). For each sample and pathogen, we retrieved the corresponding subset of reads from the rescaled BAM
422 files containing all mapping reads. We then genotyped all mapped positions, regardless of read depth, using GATK v4.1.9.0
423 (i.e., positions with at least one supporting read $-DP \geq 1$ were retained). The list of genomic positions recovered for each
424 sample and subset of reads is provided in Table S4. Next, we reconstructed partial genome sequences by inserting the
425 observed alleles at all positions covered by at least one read (i.e., REF or ALT alleles) onto the reference genome backbone,
426 while filling all uncovered positions with Ns, as explained above for the processing of high-coverage genomes. The resulting
427 FASTA sequences were then also processed identically to high-coverage strains to generate a multiple sequence alignment
428 (MSA) matching the length and genomic coordinates of the reference MSA used to build the *S. enterica* and *Borrelia*
429 phylogenetic trees. The partial genomes were then placed onto the reference trees using epa-ng, which estimates the most

430 likely phylogenetic position of each query based on its observed alleles and the topological and branch length information of
431 the tree.

432
433 **Authentication Step 4: Additional phylogenetic validation of the ancient *B. recurrentis* strain YYY093A**

434 To further investigate the phylogenetic position of YYY093A, we performed an additional analysis leveraging its relatively
435 high genome recovery (144,470 covered positions). We reconstructed a *de novo* phylogenetic tree including the same 16
436 strains as in the reference tree, but using only the positions covered in YYY093A, thereby restricting the analysis to genomic
437 sites where our ancient strain provided information. This tree recapitulated the overall phylogeny of the species, although
438 bootstrap values were low for the clade including YYY093A and the previously published ancient strain C11907. To better
439 resolve this uncertainty, we built a consensus phylogenetic network using SplitsTree³⁷ from all 100 bootstrap trees. This
440 graphical representation allowed us to visualize regions of the tree with topological ambiguities, which supported the close
441 affinity between YYY093A and C11907.

442
443 **SUPPLEMENTAL INFORMATION TITLES AND LEGENDS**

444 **Document S1, Figure S1-S4, Tables S1-S4 and supplemental references.**

445 **Document S2. Curated list of 535 TaxIDs representing 185 bacterial species known to be pathogenic to humans,**
446 **retrieved from the PATRIC database.**

447 **REFERENCES**

- 448
- 449 1. Lieven, D. (2010). *Russia against Napoleon: The true story of the campaigns of war and peace* (Penguin).
- 450
- 451 2. Talty, S. (2009). *The Illustrious Dead: The Terrifying Story of How Typhus Killed Napoleon's Greatest Army*
- 452 (Crown).
- 453
- 454 3. Raoult, D., Dutour, O., Houhamdi, L., Jankauskas, R., Fournier, P.-E., Ardagna, Y., Drancourt, M., Signoli, M., La
- 455 V.D., Macia, Y., et al. (2006). Evidence for louse-transmitted diseases in soldiers of Napoleon's Grand Army in Vilnius. *J.*
- 456 *Infect. Dis.* *193*, 112–120. <https://doi.org/10.1086/498534>.
- 457
- 458 4. Bédarida, S., Dutour, O., Buzhilova, A.P., de Micco, P., and Biagini, P. (2011). Identification of viral DNA
- 459 (Anelloviridae) in a 200-year-old dental pulp sample (Napoleon's Great Army, Kaliningrad, 1812). *Infect. Genet. Evol. J.*
- 460 *Mol. Epidemiol. Evol. Genet. Infect. Dis.* *11*, 358–362. <https://doi.org/10.1016/j.meegid.2010.11.007>.
- 461
- 462 5. De Kirckhoff, J.R.L.D. (1836). *Histoire des maladies observées à la Grande Armée française: pendant les*
- 463 *campagnes de Russie en 1812 et d'Allemagne en 1813* (Janssens).
- 464
- 465 6. Signoli, M., Ardagna, Y., Adalian, P., Devriendt, W., Lalys, L., Rigeade, C., Vette, T., Kuncevicus, A., Poskiene,
- 466 J., Barkus, A., et al. (2004). Discovery of a mass grave of Napoleonic period in Lithuania (1812, Vilnius). *Comptes Rendus*
- 467 *Palevol* *3*, 219–227. <https://doi.org/10.1016/j.crpv.2004.02.003>.
- 468
- 469 7. Paratyphoid fever <https://www.who.int/teams/immunization-vaccines-and-biologicals/diseases/paratyphoid-fever>.
- 470
- 471 8. Warrell, D.A. (2019). Louse-borne relapsing fever (*Borrelia recurrentis* infection). *Epidemiol. Infect.* *147*, e106.
- 472
- 473 9. Chandler, D.G. (2009). *The Campaigns of Napoleon* (Simon and Schuster).
- 474
- 475 10. Austin, P.B. (2012). *1812: The Great Retreat* (Frontline Books).
- 476
- 477 11. Larrey, D.J.B. (1812). *Mémoires de chirurgie militaire et campagne* (chez J. Smith).
- 478
- 479 12. Lejeune, L.-F. (1896). *Mémoires du général Lejeune* (Librairie de Firmin-Didot et Cie, imprimeurs de l'Institut).
- 480
- 481 13. Boulangier, V.A. (1820). *Dissertation sur le typhus contagieux, observé à Auxerre à la fin de 1811, et au*
- 482 *commencement de 1812, etc.*
- 483
- 484 14. Breitwieser, F.P., Baker, D.N., and Salzberg, S.L. (2018). KrakenUniq: confident and fast metagenomics
- 485 classification using unique k-mer counts. *Genome Biol.* *19*, 198. <https://doi.org/10.1186/s13059-018-1568-0>.
- 486
- 487 15. Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., Conrad, N., Dietrich, E.M., Disz, T., and
- 488 Gabbard, J.L. (2017). Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center.
- 489 *Nucleic Acids Res.* *45*, D535–D542.
- 490
- 491 16. Pochon, Z., Bergfeldt, N., Kırdök, E., Vicente, M., Naidoo, T., Van Der Valk, T., Altınışık, N.E., Krzewińska, M.,
- 492 Dalén, L., Götherström, A., et al. (2023). aMeta: an accurate and memory-efficient ancient metagenomic profiling workflow.
- 493 *Genome Biol.* *24*, 242. <https://doi.org/10.1186/s13059-023-03083-9>.
- 494
- 495 17. Hübler, R., Key, F.M., Warinner, C., Bos, K.I., Krause, J., and Herbig, A. (2019). HOPS: automated detection and
- 496 authentication of pathogen DNA in archaeological remains. *Genome Biol.* *20*, 280. [https://doi.org/10.1186/s13059-019-1903-](https://doi.org/10.1186/s13059-019-1903-0)
- 497 *0*.
- 498
- 499 18. Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., De
- 500 Filippo, C., et al. (2012). A High-Coverage Genome Sequence from an Archaic Denisovan Individual. *Science* *338*, 222–226.
- 501 <https://doi.org/10.1126/science.1224344>.
- 502
- 503 19. Sandoval-Velasco, M., Jagadeesan, A., Ramos-Madriral, J., Ávila-Arcos, M.C., Fortes-Lima, C.A., Watson, J.,
- 504 Johannesdóttir, E., Cruz-Dávalos, D.I., Gopalakrishnan, S., and Moreno-Mayar, J.V. (2023). The ancestry and geographical
- 505 origins of St Helena's liberated Africans. *Am. J. Hum. Genet.* *110*, 1590–1599.

- 506 20. Lopopolo, M., Avanzi, C., Duchene, S., Luisi, P., De Flamingh, A., Ponce-Soto, G.Y., Tressieres, G., Neumeyer,
507 S., Lemoine, F., Nelson, E.A., et al. (2025). Pre-European contact leprosy in the Americas and its current persistence.
508 Science, eadu7144. <https://doi.org/10.1126/science.adu7144>.
- 509
- 510 21. Nguyen-Hieu, T., Aboudharam, G., and Drancourt, M. (2012). Heat degradation of eukaryotic and bacterial DNA:
511 an experimental model for paleomicrobiology. BMC Res. Notes 5, 528. <https://doi.org/10.1186/1756-0500-5-528>.
- 512
- 513 22. Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jäger, G., Bos, K.I., Herbig, A., Economou, C.,
514 Benjak, A., Busso, P., et al. (2013). Genome-Wide Comparison of Medieval and Modern *Mycobacterium leprae*. Science
515 341, 179–183. <https://doi.org/10.1126/science.1238286>.
- 516
- 517 23. Huson, D.H., Beier, S., Flade, I., Górská, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.-J., and Tappu, R. (2016).
518 MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. PLOS
519 Comput. Biol. 12, e1004957. <https://doi.org/10.1371/journal.pcbi.1004957>.
- 520
- 521 24. Seersholm, F.V., Sjögren, K.-G., Koelman, J., Blank, M., Svensson, E.M., Staring, J., Fraser, M., Pinotti, T.,
522 McColl, H., Gaunitz, C., et al. (2024). Repeated plague infections across six generations of Neolithic Farmers. Nature 632,
523 114–121. <https://doi.org/10.1038/s41586-024-07651-2>.
- 524
- 525 25. Mühlemann, B., Vinner, L., Margaryan, A., Wilhelmson, H., de la Fuente Castro, C., Allentoft, M.E., de Barros
526 Damgaard, P., Hansen, A.J., Holtsmark Nielsen, S., Strand, L.M., et al. (2020). Diverse variola virus (smallpox) strains were
527 widespread in northern Europe in the Viking Age. Science 369, eaaw8977. <https://doi.org/10.1126/science.aaw8977>.
- 528
- 529 26. Neumann, G.U., Skourtanioti, E., Burri, M., Nelson, E.A., Michel, M., Hiss, A.N., McGeorge, P.J.P., Betancourt,
530 P.P., Spyrou, M.A., Krause, J., et al. (2022). Ancient Yersinia pestis and Salmonella enterica genomes from Bronze Age
531 Crete. Curr. Biol. CB 32, 3641–3649.e8. <https://doi.org/10.1016/j.cub.2022.06.094>.
- 532
- 533 27. Guellil, M., Kersten, O., Namouchi, A., Bauer, E.L., Derrick, M., Jensen, A.Ø., Stenseth, N.C., and Bramanti, B.
534 (2018). Genomic blueprint of a relapsing fever pathogen in 15th century Scandinavia. Proc. Natl. Acad. Sci. U. S. A. 115,
535 10422–10427. <https://doi.org/10.1073/pnas.1807266115>.
- 536
- 537 28. Rodríguez-Varela, R., Yaka, R., Pochon, Z., Sanchez-Pinto, I., Solaun, J.L., Naidoo, T., Guinet, B., Pérez-Ramallo,
538 P., Lagerholm, V.K., de Anca Prado, V., et al. (2024). Five centuries of consanguinity, isolation, health, and conflict in Las
539 Gobas: A Northern Medieval Iberian necropolis. Sci. Adv. 10, eadp8625. <https://doi.org/10.1126/sciadv.adp8625>.
- 540
- 541 29. Swali, P., Booth, T., Tan, C.C.S., McCabe, J., Anastasiadou, K., Barrington, C., Borrini, M., Bricking, A.,
542 Buckberry, J., Büster, L., et al. (2024). Ancient Borrelia genomes document the evolutionary history of louse-borne relapsing
543 fever. Preprint at bioRxiv, <https://doi.org/10.1101/2024.07.18.603748> <https://doi.org/10.1101/2024.07.18.603748>.
- 544
- 545 30. Barbera, P., Kozlov, A.M., Czech, L., Morel, B., Darriba, D., Flouri, T., and Stamatakis, A. (2019). EPA-ng:
546 Massively Parallel Evolutionary Placement of Genetic Sequences. Syst. Biol. 68, 365–369.
547 <https://doi.org/10.1093/sysbio/syy054>.
- 548
- 549 31. Vågene, Å.J., Herbig, A., Campana, M.G., Robles García, N.M., Warinner, C., Sabin, S., Spyrou, M.A., Andrades
550 Valtueña, A., Huson, D., Tuross, N., et al. (2018). Salmonella enterica genomes from victims of a major sixteenth-century
551 epidemic in Mexico. Nat. Ecol. Evol. 2, 520–528. <https://doi.org/10.1038/s41559-017-0446-6>.
- 552
- 553 32. Haller, M., Callan, K., Susat, J., Flux, A.L., Immel, A., Franke, A., Herbig, A., Krause, J., Kupczok, A., Fouquet,
554 G., et al. (2021). Mass burial genomics reveals outbreak of enteric paratyphoid fever in the Late Medieval trade city Lübeck.
555 iScience 24, 102419. <https://doi.org/10.1016/j.isci.2021.102419>.
- 556
- 557 33. Zhou, Z., Lundstrøm, I., Tran-Dien, A., Duchêne, S., Alikhan, N.-F., Sergeant, M.J., Langridge, G., Fotakis, A.K.,
558 Nair, S., Stenøien, H.K., et al. (2018). Pan-genome Analysis of Ancient and Modern Salmonella enterica Demonstrates
559 Genomic Stability of the Invasive Para C Lineage for Millennia. Curr. Biol. CB 28, 2420–2428.e10.
560 <https://doi.org/10.1016/j.cub.2018.05.058>.
- 561
- 562 34. Key, F.M., Posth, C., Esquivel-Gomez, L.R., Hübner, R., Spyrou, M.A., Neumann, G.U., Furtwängler, A., Sabin,
563 S., Burri, M., Wissgott, A., et al. (2020). Emergence of human-adapted Salmonella enterica is linked to the Neolithization
564 process. Nat. Ecol. Evol. 4, 324–333. <https://doi.org/10.1038/s41559-020-1106-9>.

565 35. Wu, X., Ning, C., Key, F.M., Andrades Valtueña, A., Lankapalli, A.K., Gao, S., Yang, X., Zhang, F., Liu, L., Nie,
566 Z., et al. (2021). A 3,000-year-old, basal *S. enterica* lineage from Bronze Age Xinjiang suggests spread along the Proto-Silk
567 Road. *PLoS Pathog.* *17*, e1009886. <https://doi.org/10.1371/journal.ppat.1009886>.

568 36. de-Dios, T., Carrión, P., Olalde, I., Llovera Nadal, L., Lizano, E., Pàmies, D., Marques-Bonet, T., Balloux, F., van
569 Dorp, L., and Lalueza-Fox, C. (2021). *Salmonella enterica* from a soldier from the 1652 siege of Barcelona (Spain) supports
570 historical transatlantic epidemic contacts. *iScience* *24*, 103021. <https://doi.org/10.1016/j.isci.2021.103021>.

571 37. Huson, D.H., and Bryant, D. (2024). The SplitsTree App: interactive analysis and visualization using phylogenetic
572 trees and networks. *Nat. Methods* *21*, 1773–1774.

573 38. Bhan, M.K., Bahl, R., and Bhatnagar, S. (2005). Typhoid and paratyphoid fever. *Lancet Lond. Engl.* *366*, 749–762.
574 [https://doi.org/10.1016/S0140-6736\(05\)67181-4](https://doi.org/10.1016/S0140-6736(05)67181-4).

575 39. Huson, D.H., and Bryant, D. (2024). The SplitsTree App: interactive analysis and visualization using phylogenetic
576 trees and networks. *Nat. Methods* *21*, 1773–1774.

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

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

581 42. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform.
582 *Bioinforma. Oxf. Engl.* *25*, 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.

583 43. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy,
584 S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. *GigaScience* *10*, giab008.
585 <https://doi.org/10.1093/gigascience/giab008>.

586 44. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009).
587 BLAST+: architecture and applications. *BMC Bioinformatics* *10*, 421. <https://doi.org/10.1186/1471-2105-10-421>.

588 45. Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007). MEGAN analysis of metagenomic data. *Genome Res.*
589 *17*, 377–386. <https://doi.org/10.1101/gr.5969107>.

590 46. Picard Tools - By Broad Institute <https://broadinstitute.github.io/picard/>.

591 47. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L.F., and Orlando, L. (2013). mapDamage2.0: fast
592 approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* *29*, 1682–1684.
593 <https://doi.org/10.1093/bioinformatics/btt193>.

594 48. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytzky, A., Garimella, K., Altshuler, D.,
595 Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation
596 DNA sequencing data. *Genome Res.* *20*, 1297–1303. <https://doi.org/10.1101/gr.107524.110>.

597 49. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features.
598 *Bioinforma. Oxf. Engl.* *26*, 841–842. <https://doi.org/10.1093/bioinformatics/btq033>.

599 50. Lemoine, F., and Gascuel, O. (2021). Gotree/Goalign: toolkit and Go API to facilitate the development of
600 phylogenetic workflows. *NAR Genomics Bioinforma.* *3*, lqab075. <https://doi.org/10.1093/nargab/lqab075>.

601 51. Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAXML-NG: a fast, scalable and user-
602 friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* *35*, 4453–4455.
603 <https://doi.org/10.1093/bioinformatics/btz305>.

604 52. Team, R.C. (2020). R A language and environment for statistical computing, R Foundation for Statistical
605 Computing.

606 53. Racine, J.S. (2012). Rstudio: A Platform-Independent Ide for R and Sweave. *J. Appl. Econom.* *27*, 167–172.

- 625 54. Tranié, J., and Carmigniani, J.C. (1997). *Napoléon : 1812, la campagne de Russie* (Pygmalion/G. Watelet).
- 626
- 627 55. Coignet, J.-R. (2022). *Les cahiers du Capitaine Coignet* (DigiCat).
- 628
- 629 56. Marbot, baron de, Jean-Baptiste-Antoine-Marcelin, Dutourd, J., and Garnier, J. (1983). *Mémoires du général*
- 630 *Baron de Marbot* (Mercure de France).
- 631
- 632 57. Signoli, M., and Vette, T. (2013). *Mémoires d'os : la fouille et l'étude d'un charnier de soldats de la Grande*
- 633 *Armée*. In (Musée de l'Armée/Somogy Editions d'Art), p. 60.
- 634
- 635 58. Signoli, M., Vette, T., Dutour, O., and Ardagna, Y. (2008). *Les Oubliés de la Retraite de Russie; Vilna 1812-*
- 636 *Vilnius 2002*. (Librairie Historique Teissedre Paris).
- 637
- 638 59. Palubeckaitė, Ž., Jankauskas, R., Ardagna, Y., Macia, Y., Rigeade, C., Signoli, M., and Dutour, O. (2006). Dental
- 639 status of Napoleon's Great Army's (1812) mass burial of soldiers in Vilnius: childhood peculiarities and adult dietary habits.
- 640 *Int. J. Osteoarchaeol.* *16*, 355–365. <https://doi.org/10.1002/oa.846>.
- 641
- 642 60. Keller, M., and Scheib, C.L. (2023). Decontamination of tooth roots/petrous bone cores for ancient DNA
- 643 extraction.
- 644
- 645 61. Keller, M., and Scheib, C.L. (2023). Ancient DNA extract purification (chunk samples/high volume).
- 646
- 647 62. Keller, M., Scheib, C.L., and Bonucci, B. (2023). Library preparation (dsDNA double indexing, non-UDG, 2x
- 648 split).
- 649
- 650 63. Spyrou, M.A., Bos, K.I., Herbig, A., and Krause, J. (2019). Ancient pathogen genomics as an emerging tool for
- 651 infectious disease research. *Nat. Rev. Genet.* *20*, 323–340.
- 652
- 653 64. Bettisworth, B., Psonis, N., Poulakakis, N., Pavlidis, P., and Stamatakis, A. (2025). Read Length Dominates
- 654 Phylogenetic Placement Accuracy of Ancient DNA Reads. *Mol. Biol. Evol.*, msaf006.
- 655
- 656 65. Martiniano, R., De Sanctis, B., Hallast, P., and Durbin, R. (2022). Placing ancient DNA sequences into reference
- 657 phylogenies. *Mol. Biol. Evol.* *39*, msac017.
- 658
- 659