

1 **Ancient *Yersinia pestis* genomes provide no evidence for the origins or spread of the**  
2 **Justinianic Plague**

3

4 Marcel Keller<sup>1,2\*</sup>, Maria A. Spyrou<sup>1</sup>, Michael McCormick<sup>3,4</sup>, Kirsten I. Bos<sup>1</sup>, Alexander  
5 Herbig<sup>1\*</sup>, Johannes Krause<sup>1,4\*</sup>

6

7 <sup>1</sup>Department of Archaeogenetics, Max Planck Institute for the Science of Human History,  
8 Jena, Germany

9 <sup>2</sup>Institute of Genomics, University of Tartu, Tartu, Estonia

10 <sup>3</sup>Initiative for the Science of the Human Past, Department of History, Harvard University,  
11 Cambridge, USA

12 <sup>4</sup>Max Planck-Harvard Research Center for the Archaeoscience of the Ancient Mediterranean

13 \*For correspondence: marcel.keller@ut.ee (MK); herbig@shh.mpg.de (AH);  
14 krause@shh.mpg.de (JK)

15 Competing interests: The authors declare that no competing interests exist.

16

17 **Abstract**

18 Along with the publication of 137 ancient human genomes retrieved from archaeological  
19 remains of the Eurasian steppe, Damgaard et al., 2018 identified two individuals infected with  
20 *Yersinia pestis*, yielding one genome with 0.24x average coverage (DA147, 6<sup>th</sup>–9<sup>th</sup> c. AD) and  
21 another with 8.7x (DA101, 2<sup>nd</sup>–3<sup>rd</sup> c. AD). A phylogenetic analysis performed on the latter  
22 placed it in a position ancestral to a 6<sup>th</sup>-century Justinianic genome from Aschheim, Germany.  
23 These results are used to fuel an argument that the Justinianic Plague (541–544 AD) “was  
24 brought to Europe towards the end of the Hunnic period through the Silk Road along the  
25 southern fringes of the steppes” in contrast to the leading hypothesis of introduction via the  
26 Red Sea that is supported by historical accounts. In our reanalysis, we question the contested  
27 historical context of the presented genomes with the Justinianic Plague and show that the lower  
28 coverage genome might be rather related to the Black Death (1346–1353 AD).

## 29 Introduction

30 The recent sequencing of dozens of pathogen genomes reconstructed from ancient DNA  
31 enabled increased-resolution phylogeographic studies on the spread of infectious diseases in  
32 prehistoric and historic times, especially in the context of human migration, mobility and trade  
33 (Andrades Valtueña et al., 2017; Bos et al., 2016; Keller et al., 2019; Namouchi et al., 2018;  
34 Rascovan et al., 2019; Rasmussen et al., 2015; Spyrou et al., 2019, 2016; Vågane et al., 2018).  
35 This is especially true for plague with its long and richly documented history and the abundance  
36 of published ancient genomes of its causative agent, *Yersinia pestis*. Interpretation of  
37 phylogenetic data in the context of human history requires careful assessment of tree  
38 topologies, branch lengths and mutation rates as well as thoughtful consistent approaches in  
39 integrating historical and archaeological data to prevent overly simplistic, deterministic or even  
40 erroneous interpretations.

41 For the Second Pandemic, the geographic origin and the possible persistence within Europe  
42 after the Black Death (1346–1352 AD) are the subject of ongoing scientific and scholarly  
43 discussion (Bos et al., 2016; Namouchi et al., 2018; Schmid et al., 2015; Spyrou et al., 2016;  
44 Spyrou et al., 2019). Whereas the first comprehensive phylogenetic study on *Y. pestis* favoured  
45 an East Asian origin (Cui et al., 2013), other scenarios assume an origin in Central Asia or the  
46 Caucasus (Benedictow, 2004; Namouchi et al., 2018; Sussman, 2011). Similarly, the origin of  
47 the Justinianic Plague (541–544 AD) has long been hypothesized to have originated in Africa  
48 (Achtman et al., 1999; Cui et al., 2008; Sarris, 2002). More recent studies however agree that  
49 the strains causing the First Pandemic (541–750 AD) likely emerged in Central Asia  
50 (Eroshenko et al., 2017; Harper, 2017; Wagner et al., 2014). The fact that the first outbreak of  
51 the Justinianic Plague is reported for Pelusium, Egypt nevertheless raises questions about the  
52 history and itinerary of the causative *Y. pestis* strain prior to this outbreak. The currently  
53 favoured scenario is an introduction via the Red Sea from India (Harper, 2017; Tsiamis et al.,  
54 2009) since there are no historical sources supporting a land route via the Levant or the Arabian  
55 peninsula (Schamiloglu, 2016). As such, discrepancies that arise from different analytical  
56 approaches raise questions about the history of *Y. pestis* prior to the first documented  
57 Justinianic Plague outbreak.

58 In a recent publication, Damgaard et al., 2018 presented two ancient *Y. pestis* genomes: one  
59 from the Tian Shan region (DA101, 2<sup>nd</sup>–3<sup>rd</sup> c. AD, 8.7-fold average coverage), branching  
60 ancestral to the First Pandemic lineage; and one from the Caucasus (DA147, 6<sup>th</sup>–9<sup>th</sup> c. AD,  
61 0.24-fold average coverage) which was not further investigated. Damgaard et al.'s  
62 phylogenetic analysis places DA101 ancestral to the published genome from Aschheim. This

63 positioning is supported by a single SNP shared between the two genomes. An additional five  
64 SNPs are unique to DA101 compared to 95 in Aschheim, which is provisionally consistent  
65 with Aschheim's younger age (reported in the SI and in Extended Data Fig. 9, though the latter  
66 does not present the tree at full resolution). The identified shared ancestry was interpreted as  
67 setting DA101 within the context of the "Justinian plague" (*sic*). The longer branch in the  
68 Aschheim genome is explained by its younger age and a seemingly accelerated substitution  
69 rate, which is supposedly indicative of an epidemic context.

70 Although the Justinianic Plague was previously thought to represent the first major onslaught  
71 of plague in humans (i.e., the First Pandemic), plentiful examples of human infections of *Y.*  
72 *pestis* are surfacing as far back as the Neolithic (Andrades Valtueña et al., 2017; Rascovan et  
73 al., 2019; Spyrou et al., 2018). Here, we present a reanalysis of both DA101 and DA147  
74 genomes which does not seem to support the arguments made by Damgaard et al., 2018.  
75 Instead, the analysis of DA101 suggests it to be yet another example of a pre-Justinianic human  
76 infection. Furthermore, in contrast to its suggested archaeological dating, we show DA147 to  
77 occupy a phylogenetic position much closer to the Black Death (1346–1353 AD) than the  
78 Justinianic Plague (Fig. 1C). As such, neither genome can address the origin of the First  
79 Pandemic.

80

## 81 **Results and Discussion**

82 We reanalysed both presented genomes with a more extensive dataset of published modern and  
83 ancient *Y. pestis* genomes (Fig. 1A, Table S1). We opted to include the genome from  
84 Altenerding (Feldman et al., 2016) as a representative for the Justinianic Plague: though  
85 genetically identical to Aschheim (Wagner et al., 2014), its higher coverage makes it less prone  
86 to false positive SNPs that are common in metagenomic data with high environmental  
87 backgrounds. Of note, the Aschheim genome has been shown to carry a high number of false  
88 positive SNPs (Feldman et al., 2016), which might in part account for its longer branch and  
89 accelerated substitution rate observed by Damgaard et al. (see SI).

90 Analysis of the DA101 genome revealed a minimum of 3 SNPs shared with Altenerding and a  
91 minimum of 9 that are unique. By contrast, Altenerding has 51 unique SNPs (Table S2). This  
92 sets both nodes, i.e., that giving rise to the shared Justinianic/DA101 branch and the one  
93 separating them, deeper in time compared to what is presented in the original publication  
94 (Damgaard et al., 2018).

95 Further to this, we attempted a molecular dating analysis, though the age of individual DA101  
96 proved difficult to determine given discrepancies in the text and SI, ranging from

97 “approximately 180 AD” (main text) to 214–261 calAD/1701 BP (Damgaard et al., 2018  
98 Supplementary Table 2). Ultimately, we opted to use the calibrated radiocarbon interval, which  
99 yielded a mean age of 154 BC (95% HPD: 527 BC to 153 AD) for the emergence of the shared  
100 lineage and 9 BC (95% HPD: 318 BC to 221 AD) for their divergence time (Table S4). For  
101 comparison, dating results without the recently published RT5 genome (Spyrou et al., 2018)  
102 are shown in Table S4. This strongly supports a pre-Justinianic provenience for the DA101  
103 genome. A number of shared or unique SNPs might be undetected for DA101 due to low  
104 coverage, hence the estimated divergence dates are conservative and might be even older.  
105 Regarding the substitution rate, we do not observe a notable acceleration on the Altenerding  
106 branch (mean 2.67E-08) compared to the overall mean (1.48E-08) across the tested dataset  
107 (Fig. S2), particularly since both estimates show overlapping 95 % HPD intervals (Table S5).  
108 Nevertheless, we do observe an overdispersion of substitution rates across different *Y. pestis*  
109 lineages (described previously in Cui et al., 2013 and Spyrou et al., 2019) with the highest  
110 estimate here yielding an 17-fold deviation from the mean (2.46E-07).  
111 Damgaard et al. do not discuss the fact that DA101 predates the onset of the Justinianic Plague  
112 by three centuries according to its radiocarbon date. This fact, however, is incompatible with  
113 their hypothesis of a 6<sup>th</sup>-century pandemic disease introduction to Europe through Hunnic  
114 expansion based on this genome alone, as argued suggestively multiple times in their work: in  
115 the abstract (“Scythians [...] moved westward in about the second or third century BC, forming  
116 the Hun traditions in the fourth–fifth century AD, and carrying with them plague that was basal  
117 to the Justinian plague.”), the subheader (“Origins and spread of the Justinian plague”, p. 372)  
118 and the concluding sentence (“[...], we find provisional support for the hypothesis that the  
119 pandemic was brought to Europe towards the end of the Hunnic period through the Silk Road  
120 along the southern fringes of the steppes.”, p. 373).  
121 Previously published data demonstrating the absence of detectable genetic changes in *Y. pestis*  
122 and its extremely rapid movement during the Black Death in Europe (1347–1353 AD;  
123 (Namouchi et al., 2018; Spyrou et al., 2019, 2016) clearly indicate that this pathogen is able to  
124 travel vast geographic expanses quickly, without accumulating genetic diversity in the process.  
125 As such, the depth of the time interval for the coalescence of DA101 and the Justinianic  
126 genomes offers little to no evidence on the temporal or geographic origin of the Justinianic  
127 Plague (beginning in 541 AD, Fig. 1B). Since individual DA101 comes from a geographical  
128 location that today houses multiple plague foci including modern lineages 0.ANT1, 0.ANT2  
129 and the newly described 0.ANT5 (*sensu* Eroshenko et al., 2017; Fig. 2), it may even be

130 surmised whether the sampled individual fell victim to an epidemic event or a to a sporadic  
131 individual infection.

132 The second *Y. pestis* genome from individual DA147 from North Ossetia, supposedly 6<sup>th</sup>–9<sup>th</sup>  
133 centuries, could substantiate a spread of plague along the “southern fringes of the steppe”  
134 (p. 373), although its phylogenetic placement was not investigated by Damgaard et al. Even  
135 though the coverage is low, our re-analysis of the raw sequence data from this individual and  
136 an assessment of phylogenetically informative positions reveals that it does not share any  
137 derived SNPs with Altenerding or DA101 (Table S2). None of the positions shared between  
138 Altenerding and DA101 are covered in DA147, but 2 out of the 9 unique SNPs of DA101 are  
139 covered and show the ancestral state. Of the unique Altenerding SNPs, 9 are covered in DA147  
140 with 8 showing the ancestral state. The only SNP possibly shared with DA147 is a C>T change  
141 that is potentially caused by DNA damage, as it appears only in a single read.

142 Such initial results motivated a further exploration of DA147’s possible phylogenetic position.  
143 For this, we used MultiVCFAnalyzer v0.85 for a comparative SNP analysis against our dataset  
144 of ancient and modern *Y. pestis* genomes (Table S1), while omitting all private calls in DA147  
145 since their vast majority will represent DNA damage and sequencing errors due to the  
146 genome’s low coverage. The remaining SNPs forming the branch of DA147 in Fig. S1 (red)  
147 are an artefact caused by homoplastic or triallelic sites. We computed a maximum likelihood  
148 phylogenetic tree that, unexpectedly, placed DA147 closest to the previously described  
149 polytomy of Branches 1–4 (Fig. S1). The genomes’s placement was further investigated by  
150 visual inspection of all diagnostic SNPs separating Branches 1, 2, 3&4 and Branch 0 (see Table  
151 S3). Our analysis reveals several potential placements for DA147: (1) it is one SNP ancestral  
152 to the polytomy but derived with respect to the 0.ANT3 node, (2) it is directly on the polytomy,  
153 (3) it is one SNP ancestral to the Black Death strain (Bos et al., 2011) on Branch 1, or (4) it is  
154 one to 16 SNPs basal on Branch 2 (Fig. 1C; Table S3). The third scenario is of particular  
155 interest in the context of a recently discovered genome from Laishevo, Russia (Spyrou et al.,  
156 2019) which could be identical to DA147. Therefore, DA147 might instead offer currently  
157 unexplored insights into the origin of the Black Death.

158 Furthermore, this finding raises doubts about the precision in the archaeological dating of this  
159 specimen (6<sup>th</sup>–9<sup>th</sup> centuries; Damgaard et al., 2018). Unfortunately, the provenience of this  
160 genome cannot be further investigated since metadata from this individual are absent in Table  
161 S2 in Damgaard et al., 2018. Based on our molecular dating analysis, the node giving rise to  
162 0.ANT3, which is basal to all possible placements of DA147, is dated to a mean age of 1030

163 AD (95% HPD: 732 AD to 1274 AD), thus placing this low coverage genome within the  
164 diversity that has accumulated within the last millennium.

165 Finally, we would like to correct two inaccuracies in nomenclature in the study: First, the label  
166 “0.ANT5” has already been given to a modern clade of *Y. pestis* strains reported by Eroshenko  
167 et al., 2017. In general, we recommend against applying nomenclature combining phylogenetic  
168 and metabolic features to ancient genomes (Achtman, 2016), since their metabolic profile has  
169 not yet been characterized. Second, the “Justinianic Plague” is named after the Roman emperor  
170 Justinian I (c. 482–565 AD) who reigned during the onset of this pandemic (Little et al., 2007).  
171 The term “Justinian Plague” as used by the authors is misleading, since it suggests a connection  
172 to either Justin I or Justin II of the Justinianic dynasty.

173 Overall, we argue that the two presented *Y. pestis* genomes cannot contribute to our  
174 understanding of the Justinianic Plague that began in 541 AD in the southeast Mediterranean  
175 basin due to their phylogenetic, temporal and geographical distance. Moreover, these genomes  
176 offer no support for a connection between the Justinianic Plague and the Hunnic expansion, or  
177 for a spread through the southern steppe, both of which are also in conflict with the leading,  
178 document-based hypothesis of a plague introduction via trade routes linking India to the Red  
179 Sea (Harper, 2017; Fig. 2). The low coverage genome might rather hold clues for the onset of  
180 the Black Death or on the origins of Branch 2. We suggest a redirected focus here, especially  
181 if higher coverage data from this or a similar archaeological sample becomes available in the  
182 future.

## 183 184 **Materials and Methods**

185 Sequencing data for the samples DA101 and DA147 were retrieved from ENA with the  
186 provided accession numbers (Damgaard et al., 2018) and processed with the EAGER pipeline  
187 (Peltzer et al., 2016), including Illumina adapter removal, sequencing quality filtering  
188 (minimum base quality of 20) and length filtering (minimum length of 30 bp).

189 For the DA101 sample with higher coverage, reads were clipped on both ends by 3 bases to  
190 remove the majority of damaged sites and subsequently filtered again for length using the same  
191 parameter. Mapping against the CO92 reference genome (chromosome NC\_003143.1) was  
192 done with BWA (-l 32, -n 0.1, -q 37), reads with low mapping quality (-q 37) were removed  
193 with Samtools and duplicates were removed with MarkDuplicates.

194 SNP calling was performed with the UnifiedGenotyper within the Genome Analysis Toolkit  
195 (GATK) using the ‘EMIT\_ALL\_SITES’ option to generate a call for every position in the  
196 reference genome.

197 For the DA147 sample with low coverage, mapping was performed without prior damage  
198 clipping and with less stringent parameters in BWA (-l 16, -n 0.01, -q 37) to retrieve a  
199 maximum of coverage. Reads with low mapping quality were removed with Samtools (-q 37)  
200 and duplicates were removed with MarkDuplicates. For a phylogenetic analysis of the low  
201 coverage DA147 genome (0.24-fold), the bam-file was converted into a fastq-file using  
202 bedtools, multiplied by 5 and mapped again with identical parameters but without duplicate  
203 removal to reach the necessary coverage of positions for SNP calling. SNP calling was  
204 performed with the UnifiedGenotyper within the Genome Analysis Toolkit using  
205 'EMIT\_ALL\_SITES' to generate calls for all positions in the reference genome.  
206 For the phylogenetic analyses, we used 166 previously published modern *Y. pestis* genomes  
207 (Cui et al., 2013; Eroshenko et al., 2017; Kislichkina et al., 2015; Zhgenti et al., 2015), a *Y.*  
208 *pseudotuberculosis* reference genome (IP32953; Chain et al., 2004) as an outgroup and the  
209 following ancient genomes: nine genomes from Neolithic/Bronze Age contexts (Andrades  
210 Valtueña et al., 2017; Rasmussen et al., 2015; Spyrou et al., 2018), one genome of the  
211 Justinianic Plague (Altenerding; Feldman et al., 2016), one genome representing Black Death  
212 (8291-11972-8124; (Bos et al., 2011), and six genomes of the subsequent second plague  
213 pandemic (Observance OBS116, OBS137, OBS110, OBS107, OBS124; Bos et al., 2016);  
214 Bolgar, (Spyrou et al., 2016)). A complete list of all *Y. pestis* genomes used is given in Table  
215 S1. Previously identified problematic regions as well as regions annotated as repeat regions,  
216 rRNAs, tRNAs and tmRNAs were excluded for all subsequent analyses (Cui et al., 2013;  
217 Morelli et al., 2010). MultiVCFAnalyzer v0.85 (Bos et al., 2014) was used for generating a  
218 SNP table with the following settings: Minimal coverage for base call of 5 with a minimum  
219 genotyping quality of 30 for homozygous positions, minimum support of 90% for calling the  
220 dominant nucleotide in a 'heterozygous' position. The sample DA147 was processed in the  
221 outgroup mode in MultiVCFAnalyzer to remove all singletons, for the most part representing  
222 damaged sites called due to the prior multiplication of reads. The unfiltered SNP alignment  
223 produced by MultiVCFAnalyzer was used for all following analyses. Additionally, all  
224 phylogenetically informative positions were visually inspected in IGV v2.4 (Thorvaldsdóttir et  
225 al., 2013). Maximum likelihood trees were generated with RAxML v8 (Stamatakis, 2014)  
226 using the GTR substitution model based on a partial deletion (95 %) SNP alignment for DA101  
227 (3673 SNPs) and a full SNP alignment for DA147 (3885 SNPs). Robustness of all trees was  
228 tested by the bootstrap methods using 1000 pseudo-replicates.  
229

230 For the estimation of divergence times and substitution rates with BEAST 1.10 (Drummond  
231 and Rambaut, 2007), we used the coalescent Bayesian skyline model with a setup identical to  
232 that published in Spyrou et al., 2018 with the following modifications: Integration of the  
233 DA101 sample (Damgaard et al., 2018), 95% partial deletion SNP alignment and 800,000,000  
234 states as chain length. A second run was performed without the recently published Bronze Age  
235 genome RT5 (Spyrou et al., 2018) to investigate whether the results are affected by previously  
236 unavailable data. The recently published genomes by Eroshenko et al., 2017 were not included  
237 in the dating analysis due to exceptionally long branches and unavailability of raw data to  
238 address potential mismapping. An MCC tree was produced using TreeAnnotator of BEAST  
239 v1.10, showing the relative mean substitution rates (Fig. S2). All trees were visualized in  
240 FigTree v1.4.3 (<http://tree.bio.ed.ac.uk/software/figtree/>).

241 For Fig. 2, we used the coordinates of DA146 and DA160, since they are identical and frame  
242 the sample DA147 which is not part of the Table S2 of Damgaard et al., 2018. The first  
243 outbreaks of the Justinianic Plague come from Harper, 2017 and (Stathakopoulos, 2004). For  
244 the trade routes, we used the “Indian and Persian trade routes with the West 50 BCE - 300 CE”  
245 and “Silk Road routes 1–1400 CE” from OWTRAD (<http://www.ciolek.com/owtrad.html>).  
246 Important trade centres are adopted from Harper, 2017.

247

#### 248 Author Contributions

249 M.K., A.H., K.I.B and J.K. planned and designed the study. M.K. performed data processing  
250 and phylogenetic analyses; M.A.S. and M.K. performed dating analyses. M.M. provided and  
251 reviewed historical context information. M.K. wrote the manuscript with contributions from  
252 M.A.S., K.I.B. and A.H. and edits from all co-authors.



## 253 References

- 254 Achtman M. 2016. How old are bacterial pathogens? *Proc Biol Sci* **283**:e1003471–e1003471.
- 255 Achtman M, Zurth K, Morelli G, Torrea G, Guiyoule A, Carniel E. 1999. *Yersinia pestis*, the  
256 cause of plague, is a recently emerged clone of *Yersinia pseudotuberculosis*. *Proc Natl*  
257 *Acad Sci U S A* **96**:14043–14048.
- 258 Andrades Valtueña A, Mittnik A, Key FM, Haak W, Allmäe R, Belinskij A, Daubaras M,  
259 Feldman M, Jankauskas R, Janković I, Massy K, Novak M, Pfrengle S, Reinhold S,  
260 Šlaus M, Spyrou MA, Szécsényi-Nagy A, Törv M, Hansen S, Bos KI, Stockhammer  
261 PW, Herbig A, Krause J. 2017. The Stone Age Plague and Its Persistence in Eurasia.  
262 *Curr Biol* **27**:3683–3691.e8.
- 263 Benedictow OJ. 2004. *The Black Death 1346–1353: The Complete History*. Woodsbridge:  
264 Boydell & Brewer.
- 265 Bos KI, Harkins KM, Herbig A, Coscolla M, Weber N, Comas I, Forrest SA, Bryant JM,  
266 Harris SR, Schuenemann VJ, Campbell TJ, Majander K, Wilbur AK, Guichon RA,  
267 Wolfe Steadman DL, Cook DC, Niemann S, Behr MA, Zumarraga M, Bastida R, Huson  
268 D, Nieselt K, Young D, Parkhill J, Buikstra JE, Gagneux S, Stone AC, Krause J. 2014.  
269 Pre-Columbian mycobacterial genomes reveal seals as a source of New World human  
270 tuberculosis. *Nature* **514**:494–497.
- 271 Bos KI, Herbig A, Sahl J, Waglechner N, Fourment M, Forrest SA, Klunk J, Schuenemann  
272 VJ, Poinar D, Kuch M, Golding GB, Dutour O, Keim P, Wagner DM, Holmes EC,  
273 Krause J, Poinar HN. 2016. Eighteenth century *Yersinia pestis* genomes reveal the long-  
274 term persistence of an historical plague focus. *Elife* **5**:e12994.
- 275 Bos KI, Schuenemann VJ, Golding GB, Burbano HA, Waglechner N, Coombes BK, McPhee  
276 JB, DeWitte SN, Meyer M, Schmedes S, Wood J, Earn DJD, Herring DA, Bauer P,  
277 Poinar HN, Krause J. 2011. A draft genome of *Yersinia pestis* from victims of the Black  
278 Death. *Nature* **478**:506–510.
- 279 Chain PSG, Carniel E, Larimer FW, Lamerdin J, Stoutland PO, Regala WM, Georgescu AM,  
280 Vergez LM, Land ML, Motin VL, Brubaker RR, Fowler J, Hinnebusch J, Marceau M,  
281 Medigue C, Simonet M, Chenal-Francisque V, Souza B, Dacheux D, Elliott JM, Derbise  
282 A, Hauser LJ, Garcia E. 2004. Insights into the evolution of *Yersinia pestis* through  
283 whole-genome comparison with *Yersinia pseudotuberculosis*. *Proceedings of the*  
284 *National Academy of Sciences* **101**:13826–13831.
- 285 Cui Y, Li Y, Gorgé O, Platonov ME, Yan Y, Guo Z, Pourcel C, Dentovskaya SV,

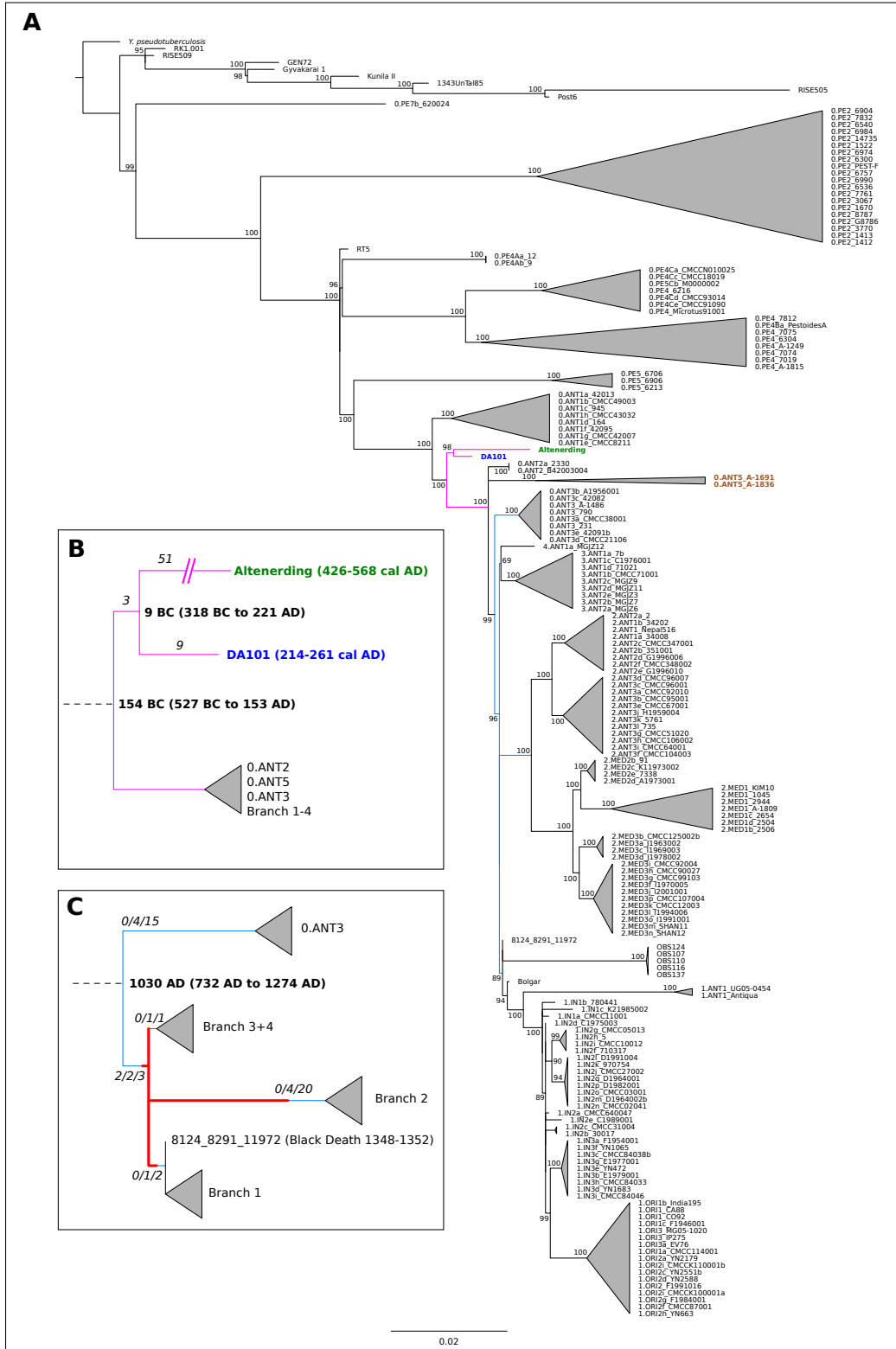
- 286 Balakhonov SV, Wang X, Song Y, Anisimov AP, Vergnaud G, Yang R. 2008. Insight  
287 into microevolution of *Yersinia pestis* by clustered regularly interspaced short  
288 palindromic repeats. *PLoS One* **3**:e2652.
- 289 Cui Y, Yu C, Yan Y, Li D, Li Y, Jombart T, Weinert L a., Wang Z, Guo Z, Xu L, Zhang Y,  
290 Zheng H, Qin N, Xiao X, Wu M, Wang X, Zhou D, Qi Z, Du Z, Wu H, Yang X, Cao H,  
291 Wang H, Wang J, Yao S, Rakin A, Li Y, Falush D, Balloux F, Achtman M, Song Y,  
292 Wang J, Yang R. 2013. Historical variations in mutation rate in an epidemic pathogen,  
293 *Yersinia pestis*. *Proc Natl Acad Sci U S A* **110**:577–582.
- 294 Damgaard P de B, Marchi N, Rasmussen S, Peyrot M, Renaud G, Korneliussen T, Moreno-  
295 Mayar JV, Pedersen MW, Goldberg A, Usmanova E, Baimukhanov N, Loman V,  
296 Hedeager L, Pedersen AG, Nielsen K, Afanasiev G, Akmatov K, Aldashev A, Alpaslan  
297 A, Baimbetov G, Bazaliiskii VI, Beisenov A, Boldbaatar B, Boldgiv B, Dorzhu C,  
298 Ellingvag S, Erdenebaatar D, Dajani R, Dmitriev E, Evdokimov V, Frei KM, Gromov A,  
299 Goryachev A, Hakonarson H, Hegay T, Khachatryan Z, Khaskhanov R, Kitov E,  
300 Kolbina A, Kubatbek T, Kukushkin A, Kukushkin I, Lau N, Margaryan A, Merkyte I,  
301 Mertz IV, Mertz VK, Mijiddorj E, Moiyesev V, Mukhtarova G, Nurmukhanbetov B,  
302 Orozbekova Z, Panyushkina I, Pieta K, Smrčka V, Shevnina I, Logvin A, Sjögren K-G,  
303 Štolcová T, Taravella AM, Tashbaeva K, Tkachev A, Tulegenov T, Voyakin D,  
304 Yepiskoposyan L, Undrakhbold S, Varfolomeev V, Weber A, Wilson Sayres MA,  
305 Kradin N, Allentoft ME, Orlando L, Nielsen R, Sikora M, Heyer E, Kristiansen K,  
306 Willerslev E. 2018. 137 ancient human genomes from across the Eurasian steppes.  
307 *Nature* **557**:369–374.
- 308 Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling  
309 trees. *BMC Evol Biol* **8**. doi:10.1186/1471-2148-7-214
- 310 Eroshenko GA, Nosov NY, Krasnov YM, Oglodin YG, Kukleva LM, Guseva NP, Kuznetsov  
311 AA, Abdikarimov T, Dzhaparova AK, Kuttyrev VV. 2017. *Yersinia pestis* strains of  
312 ancient phylogenetic branch 0.ANT are widely spread in the high- mountain plague foci  
313 of Kyrgyzstan. *PLoS One* **12**:e0187230–e0187230.
- 314 Feldman M, Harbeck M, Keller M, Spyrou MA, Rott A, Trautmann B, Scholz HC, Pääffgen  
315 B, Peters J, McCormick M, Bos K, Herbig A, Krause J. 2016. A High-Coverage *Yersinia*  
316 *pestis* Genome from a Sixth-Century Justinianic Plague Victim. *Mol Biol Evol* **33**:2911–  
317 2923.
- 318 Harper K. 2017. *The Fate of Rome: Climate, Disease, and the End of an Empire*. Princeton:  
319 Princeton University Press.

- 320 Keller M, Spyrou MA, Scheib CL, Neumann GU, Kröpelin A, Haas-Gebhard B, Pfüffgen B,  
321 Haberstroh J, Ribera I Lacomba A, Raynaud C, Cessford C, Durand R, Stadler P, Nägele  
322 K, Bates JS, Trautmann B, Inskip SA, Peters J, Robb JE, Kivisild T, Castex D,  
323 McCormick M, Bos KI, Harbeck M, Herbig A, Krause J. 2019. Ancient *Yersinia pestis*  
324 genomes from across Western Europe reveal early diversification during the First  
325 Pandemic (541–750). *Proc Natl Acad Sci U S A* **116**:12363–12372.
- 326 Kislichkina AA, Bogun AG, Kadnikova LA, Maiskaya NV, Platonov ME, Anisimov NV,  
327 Galkina EV, Dentovskaya SV, Anisimov AP. 2015. Nineteen Whole-Genome  
328 Assemblies of *Yersinia pestis* subsp. *microtus*, Including Representatives of Biovars  
329 *caucasica*, *talassica*, *hissarica*, *altaica*, *xilingolensis*, and *ulegeica*: TABLE 1. *Genome*  
330 *Announc* **3**:e01342–15.
- 331 Little LK, Hays JN, Morony MG, Kennedy HN, Stathakopoulos D, Sarris P, Stoclet AJ,  
332 Kulikowski M, Maddicott J, Dooley A, Sallares R, McCormick M. 2007. Plague and the  
333 End of Antiquity: The Pandemic of 541–750. Cambridge University Press.
- 334 Morelli G, Song Y, Mazzoni CJ, Eppinger M, Roumagnac P, Wagner DM, Feldkamp M,  
335 Kusecek B, Vogler AJ, Li Y, Cui Y, Thomson NR, Jombart T, Leblois R, Lichtner P,  
336 Rahalison L, Petersen JM, Balloux F, Keim P, Wirth T, Ravel J, Yang R, Carniel E,  
337 Achtman M. 2010. *Yersinia pestis* genome sequencing identifies patterns of global  
338 phylogenetic diversity. *Nat Genet* **42**:1140–1143.
- 339 Namouchi A, Guellil M, Kersten O, Hänsch S, Ottoni C, Schmid BV, Pacciani E, Quaglia L,  
340 Vermunt M, Bauer EL, Derrick M, Jensen AØ, Kacki S, Cohn SK, Stenseth NC,  
341 Bramanti B. 2018. Integrative approach using *Yersinia pestis* genomes to revisit the  
342 historical landscape of plague during the Medieval Period. *Proceedings of the National*  
343 *Academy of Sciences* **115**:E11790-E11797.
- 344 Peltzer A, Jäger G, Herbig A, Seitz A, Kniep C, Krause J, Nieselt K. 2016. EAGER: efficient  
345 ancient genome reconstruction. *Genome Biol* **17**:60–60.
- 346 Rascovan N, Sjögren K-G, Kristiansen K, Nielsen R, Willerslev E, Desnues C, Rasmussen S.  
347 2019. Emergence and Spread of Basal Lineages of *Yersinia pestis* during the Neolithic  
348 Decline Article Emergence and Spread of Basal Lineages of *Yersinia pestis* during the  
349 Neolithic Decline. *Cell* **176**:295–305.e10.
- 350 Rasmussen S, Allentoft ME, Nielsen K, Orlando L, Sikora M, Sjögren K-G, Pedersen AG,  
351 Schubert M, Van Dam A, Kapel CMO, Nielsen HB, Brunak S, Avetisyan P, Epimakhov  
352 A, Khalyapin MV, Gnuni A, Kriiska A, Lasak I, Metspalu M, Moiseyev V, Gromov A,  
353 Pokutta D, Saag L, Varul L, Yepiskoposyan L, Sicheritz-Pontén T, Foley RA, Lahr MM,

- 354 Nielsen R, Kristiansen K, Willerslev E. 2015. Early divergent strains of *Yersinia pestis*  
355 in Eurasia 5,000 years ago. *Cell* **163**:571–582.
- 356 Sarris P. 2002. The Justinianic plague: origins and effects. *Contin Chang* **17**:169–182.
- 357 Schamiloglu U. 2016. The Plague in the Time of Justinian and Central Eurasian History: An  
358 agenda for Research In: Zimonyi I, Karatay O, editors. Wiesbaden: Harrasowitz Verlag.
- 359 Schmid BV, Büntgen U, Easterday WR, Ginzler C, Walløe L, Bramanti B, Stenseth NC.  
360 2015. Climate-driven introduction of the Black Death and successive plague  
361 reintroductions into Europe. *Proc Natl Acad Sci U S A* **112**:3020–3025.
- 362 Spyrou MA, Keller M, Tukhbatova RI, Scheib CL, Nelson EA, Andrades Valtueña A,  
363 Neumann GU, Walker D, Alterauge A, Carty N, Cessford C, Fetz H, Gourvenec M,  
364 Hartle R, Henderson M, von Heyking K, Inskip SA, Kacki S, Key FM, Knox EL, Later  
365 C, Maheshwari-Aplin P, Peters J, Robb JE, Schreiber J, Kivisild T, Castex D, Lösch S,  
366 Harbeck M, Herbig A, Bos KI, Krause J. 2019. Phylogeography of the second plague  
367 pandemic revealed through analysis of historical *Yersinia pestis* genomes. *Nat Commun*  
368 **10**:4470.
- 369 Spyrou MA, Tukhbatova RI, Feldman M, Drath J, Kacki S, Beltrán de Heredia J, Arnold S,  
370 Sitdikov AG, Castex D, Wahl J, Gazimzyanov IR, Nurgaliev DK, Herbig A, Bos KI,  
371 Krause J. 2016. Historical *Y. pestis* Genomes Reveal the European Black Death as the  
372 Source of Ancient and Modern Plague Pandemics. *Cell Host Microbe* **19**:874–881.
- 373 Spyrou MA, Tukhbatova RI, Wang C-C, Andrades Valtueña A, Lankapalli AK, Kondrashin  
374 VV, Tsybin VA, Khokhlov A, Kühnert D, Herbig A, Bos KI, Krause J. 2018. Analysis  
375 of 3800-year-old *Yersinia pestis* genomes suggests Bronze Age origin for bubonic  
376 plague. *Nat Commun* **9**:2234.
- 377 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of  
378 large phylogenies. *Bioinformatics* **30**:1312–1313.
- 379 Stathakopoulos D. 2004. *Famine and Pestilence in the Late Roman and Early Byzantine*  
380 *Empire: A Systematic Survey of Subsistence Crises and Epidemics*. Aldershot: Ashgate.
- 381 Sussman GD. 2011. Was the Black Death in India and China? *Bull Hist Med* **85**:319–355.
- 382 Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV):  
383 high-performance genomics data visualization and exploration. *Brief Bioinform* **14**:178–  
384 192.
- 385 Tsiamis C, Poulakou-Rebelakou E, Petridou E. 2009. The Red Sea and the Port of Clysma. A  
386 Possible Gate of Justinian's Plague. *Gesnerus* **66**:209–217.
- 387 Vågane ÅJ, Herbig A, Campana MG, Robles García NM, Warinner C, Sabin S, Spyrou MA,

- 388           Andrades Valtueña A, Huson D, Tuross N, Bos KI, Krause J. 2018. Salmonella enterica  
389           genomes from victims of a major sixteenth-century epidemic in Mexico. *Nature Ecology  
390           and Evolution* **2**:520–528.
- 391       Wagner DM, Klunk J, Harbeck M, Devault A, Waglechner N, Sahl JW, Enk J, Birdsell DN,  
392           Kuch M, Lumibao C, Poinar D, Pearson T, Fourment M, Golding B, Riehm JM, Earn  
393           DJD, Dewitte S, Rouillard J-M, Grupe G, Wiechmann I, Bliska JB, Keim PS, Scholz  
394           HC, Holmes EC, Poinar H. 2014. *Yersinia pestis* and the plague of Justinian 541–543  
395           AD: a genomic analysis. *Lancet Infect Dis* **14**:319–326.
- 396       Zhgenti E, Johnson SL, Davenport KW, Chanturia G, Daligault HE, Chain PS, Nikolich MP.  
397           2015. Genome Assemblies for 11 *Yersinia pestis* Strains Isolated in the Caucasus  
398           Region. *Genome Announc* **3**:1030–1015.

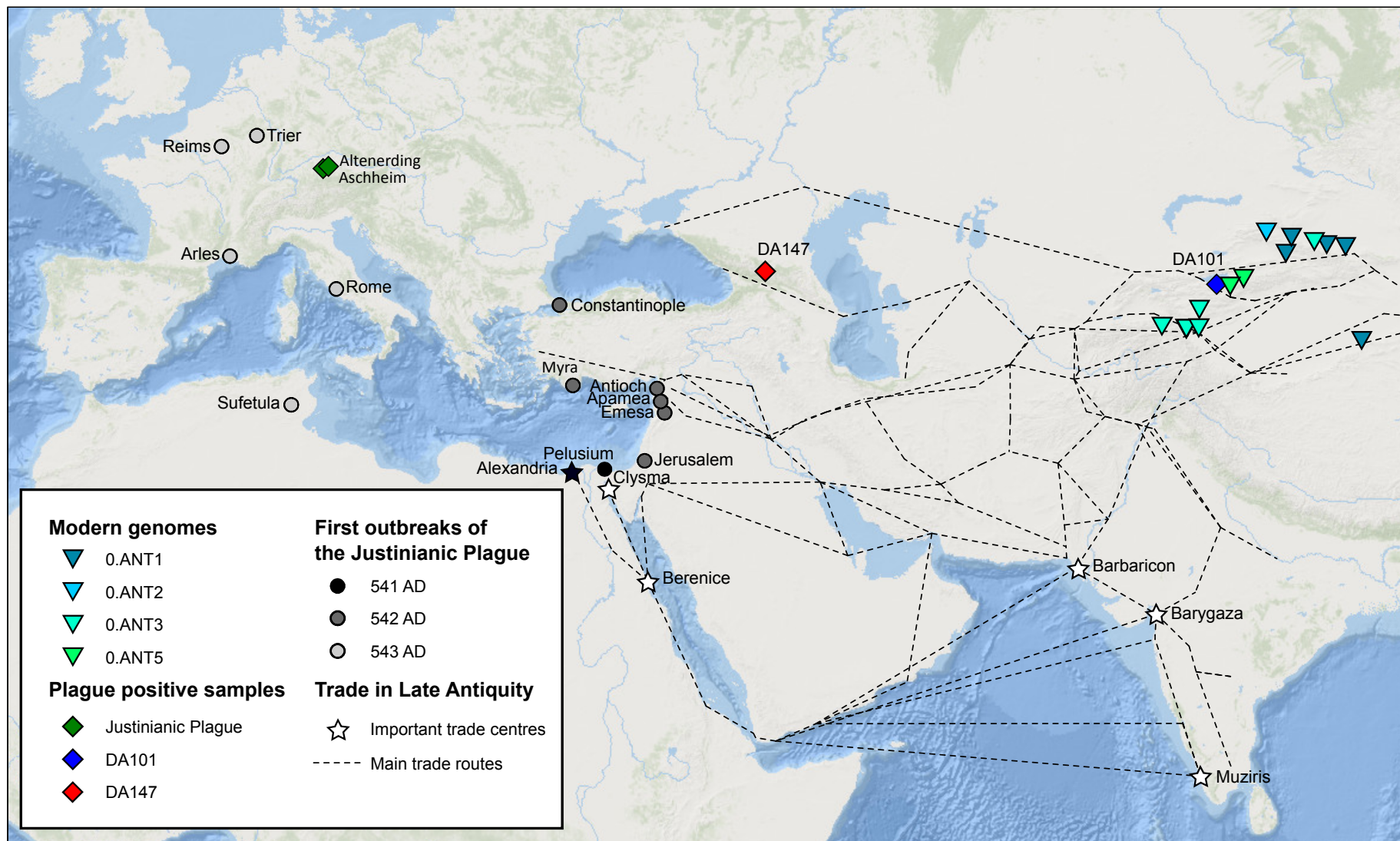
399 **Fig. 1: A.** Maximum likelihood tree based on 3673 SNPs of 167 modern and 18 ancient genomes. Main branches are collapsed for clarity.  
 400 Numbers on nodes indicate bootstrap support. Highlighted are the Justinianic genome from Altenerding (green), the investigated Tian Shan  
 401 genome DA101 (blue) and the recently characterized modern strains of clade 0.ANT5 (brown). Relevant parts of the tree are highlighted in  
 402 pink (Fig. 1B) and light blue (Fig. 1C). **B.** Schematic tree of the common branch of Altenerding and DA101 showing the minimum number  
 403 of SNPs (italics) and divergence dates (bold, mean and 95 % HPD). **C.** Schematic tree of the highlighted part of A (light blue) for the  
 404 positioning of DA147 with numbers of SNPs (italics, # of diverged SNPs in DA147/# of covered SNPs in DA147/# of total SNPs on  
 405 branch) and estimated divergence date of clade 0.ANT3 (bold, mean and 95 % HPD) based on the dating analysis with BEAST 1.10 (see  
 406 Materials and Methods and SI). Possible placements of DA147 in the phylogeny are highlighted in red.



407

408  
409  
410

**Fig. 2:** Map of the origin of selected modern and ancient *Y. pestis* genomes in relation to the first outbreaks of the Justinianic Plague as well as important trade centres and routes. An introduction of plague via the Indian Ocean and the Red Sea is supported by well-established sea communications connecting Egyptian and Indian trade centres. The hypothesis of an overland transport through the steppe is in conflict with the directionality of the first outbreaks of the Justinianic Plague and lacks historical support. Although DA147 presumably dates later than 750 AD, it is shown for comparison.



411