













# Ancient *Yersinia pestis* from Denmark reveals...

This manuscript ([permalink](#)) was automatically generated from [ktmeaton/obsidian-public@112d8e4a](#) on July 16, 2021.

## Authors

---

- **Katherine Eaton**  [0000-0001-6862-7756](#) ·  [ktmeaton](#)  
McMaster University
- **Ravneet Sidhu**  
McMaster University
- **Jennifer Klunk**  [0000-0002-6521-8516](#)  
Daicel Arbor Biosciences
- **Julia Gamble**  [0000-0001-7486-757X](#)  
University of Manitoba
- **Jesper Boldsen**  [0000-0002-2850-0934](#)  
University of Southern Denmark; ADBOU
- **Ann Carmichael**  
Indian University Bloomington
- **Nükhet Varlık**  [0000-0001-6870-5945](#)  
University of South Carolina
- **Sebastian Duchene**  [0000-0002-2863-0907](#) ·  [sebastianduchene](#)  
University of Melbourne
- **Leo Featherstone**  [0000-0002-8878-1758](#)  
University of Melbourne
- **Vaughan Grimes**  [0000-0002-2177-3147](#)  
Memorial University
- **G. Brian Golding**  [0000-0002-7575-0282](#)  
McMaster University
- **Hendrik N. Poinar**  [0000-0002-0314-4160](#)  
McMaster University

# Introduction

---

*Yersinia pestis*, the causative agent of plague, is the most intensively sequenced ancient pathogen to date. Comparative analyses of ancient *Y. pestis* genomes have significantly advanced our understanding of past pandemics, by shifting discourse on the fluctuating patterns of virulence, geographic dispersal, and local persistence. In particular, genomics research has revealed how historical plagues of the past, such as the First Pandemic (6th-8th century) and the Second Pandemic (14th-19th century), were independent emergences that were capable of persisting for multiple centuries before eventually going extinct.

This curious dynamic of long-term epidemic cycling, followed by centuries of inactivity, is a distinctive feature of plague in Europe, where plague has been identified in both pandemic and pre-pandemic periods. As one of the oldest known strains of plague was discovered in Scandinavia [1], there is great potential for an intensive examination of *Y. pestis* in this region to reveal novel insight into the long-term epidemiology of plague.

Despite the recent boom in ancient *Y. pestis* sequencing, Scandinavia is currently represented by only two genomes from Sweden (~5000 YBP) and Norway (~700 YBP). Due to this limited data, it is currently unknown to what extent plague was a local, endemic disease in this region as compared to novel epidemics with successive re-introduction. In addition, the identification of plague in pre-pandemic Scandinavia suggests the potential for *Y. pestis* to have been present in this region outside of historically documented pandemics. Furthermore, the available historical documentation primarily derives from large commercial centres whereas Scandinavia archives, such as those in the Denmark, have retained limited information about the plague and its impact on society.

In response to this region's historical importance and lack of genomic representation, this study samples and screens skeletal remains from Danish archaeological sites occupied over nearly a millenium, to detect the presence of *Y. pestis*. Following genomic capture of *Y. pestis* from plague-positive individuals, we examine the temporal and geographic structure of Danish plague within a global context.

**Kat's Notes:** To be expanded upon further and refined following co-author discussion.

# Results

## Sites and Samples

325 individuals were sampled across 6 regions from 14 archaeological sites (Table 1, Figure 1). The site occupation dates span from the 10th to 18th centuries which encompasses the Viking Age (8th - 10th century), the Medieval Period (11th - 16th century) and the Early Modern Period (16th - 19th century) in Denmark.

**Table 1:** Summary of archaeological sites sampled in this study.

Region	Site Name	Site Code	Site Occupation	Samples	Plague Positive
Ribe	Ribe Gräbrødre	ASR 1015	1200 - 1560	53	5
	Ribe Lindegården	ASR 2391	900 - 1000	5	0
		ASR 13/13II	900 - 1000	15	0
		ASR 13II	1200 - 1560	28	1
Viby	Nordby	FHM 3970	1050 - 1250	36	0
Horsens	Monastery Church	HOM 1272	1600 - 1800	50	0
	Ole Wormsgade	HOM 1649	1100 - 1500	17	2
	Sejet	HOM 1046	1150 - 1574	25	1
	Tirup	VKH 1201	1150 - 1350	12	1
Hågerup	Hågerup	ØHM 1247	1100 - 1555	7	1
Refshale	Refshale	Refshale	1100 - 1350	19	0
Viborg	Sct. Mikkel	JAH 1-77	1000 - 1529	4	0
	The Catholic Church	VSM 09264	1100 - 1529	6	0
	Sct. Mathias	VSM 855F/906F	1100 - 1529	23	0
	Sct. Drotten	VSM 902F	1100 - 1529	8	0
	Faldborg	VSM 29F	1100 - 1600	17	2
Total				325	13



**Figure 1:** Geographic map of archaeological sites by region. Stars indicate sites where *Yersinia pestis* was detected.

# Plague Detection

13 individuals were identified as plague-positive based on a combination of PCR assays, shotgun sequencing, and targeted enrichment for the *Y. pestis* whole genome (Tables 2 and 3). Of the 13 individuals, 9 had chromosomal coverage sufficient for phylogenetic analyses ( $\geq 3\times$  mean depth).

Overall, plague was detected in 4% (13/325) of all individuals in this study (Table 1). When excluding plague-negative archaeological sites, this estimate rises to 8.2% (13/159) of individuals. *Y. pestis* was observed primarily in the Medieval Period, with no evidence of plague in the Viking Age settlements at Ribe Lindegården or the Early Modern cemetery at Horsens.

## Kat's Notes:

- Absence of plague in Viking/Early Modern sites could easily be a false negative.
- These periods are sparsely sampled, with fewer sites and individuals sampled.

## Questions for Julia Gamble

- Do you have sex and age estimates for these individuals? Some are missing in the database.
- Are G25A and G25B two individuals from the same grave? If so, very interesting!

**Table 2:** Plague positive summary of high coverage genomes. Human (%) and plague (%) derive from shotgun estimates. The mean depth of coverage for the chromosome (CHROM) and the plasmids (pCD1, pMT1, pPCP1) are reported after targeted enrichment.

Arch ID	Project ID	Site	PCR	Human (%)	Plague (%)	CHROM	pCD1	pMT1	pPCP1
G16	D71	Ribe Gräbrødre	6/6	5.95	0.18	22.6	39.4	14.7	4.6
G861 x1035	D75	Ribe Gräbrødre	6/6	4.42	0.23	17.4	40.2	16.5	3.4
G25B x98	R36	Ribe Gräbrødre	6/6	8.41	0.25	24.0	51.8	14.9	5.8
G25A	D62	Ribe Gräbrødre	6/6	1.12	0.10	3.8	10.5	2.5	0.9
G207	D72	Ribe Gräbrødre	6/6	12.94	0.04	6.0	13.5	5.8	2.2
A146 x3011	P187	Sejet	6/6	0.68	0.01	4.9	18.4	6.6	52.2
G371	P212	Tirup	6/6	0.61	0.04	6.7	26.3	8.5	56.6
Gr GC 15	D51	Faldborg	6/6	0.67	0.05	9.0	25.4	8.1	2.0
A1480 x1480	P387	Ole Wormsgade	6/6	0.04	0.01	6.5	21.7	5.0	75.0

**Table 3:** Plague positive summary of low coverage genomes. Human (%) and plague (%) derive from shotgun estimates. The mean depth of coverage for the chromosome (CHROM) and the plasmids (pCD1, pMT1, pPCP1) are reported after targeted enrichment.

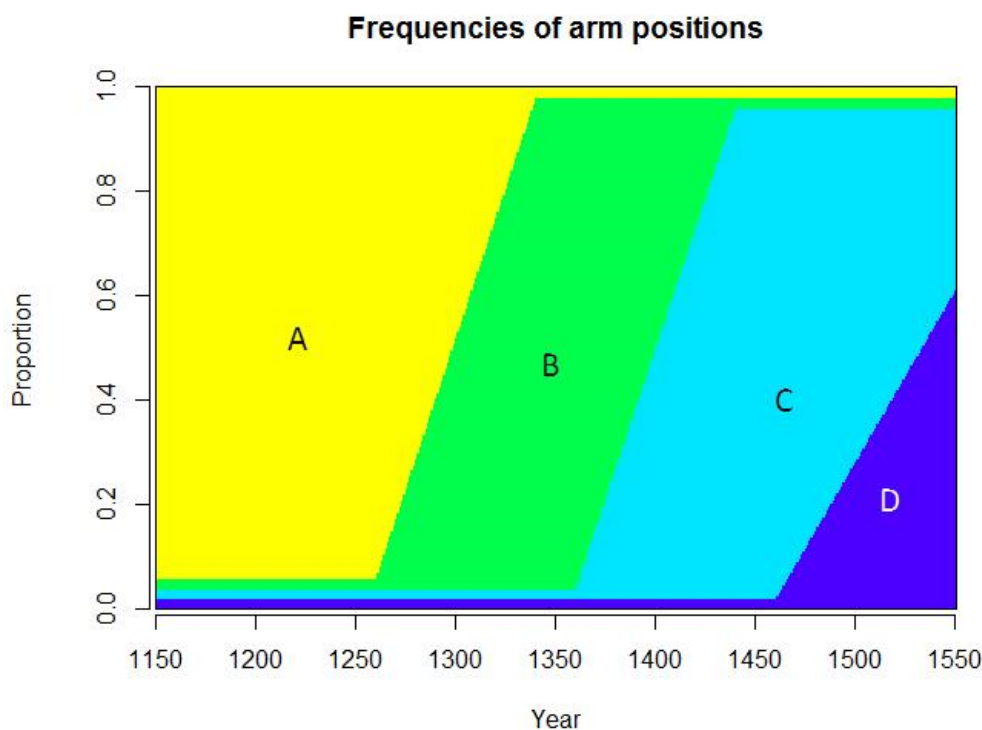
Arch ID	Project ID	Site	PCR	Human (%)	Plague (%)	CHROM	pCD1	pMT1	pPCP1
A1155 x1155	P384	Ole Wormsgade	4/6	0.11	0.01	1.1	4.8	1.4	19.6
Gr ID 319	R21	Faldborg	6/6	0.85	0.01	2.6	3.8	2.3	0.4
A19 X21	D24	Hågerup	6/6	0.55	0.01	2.6	6.1	1.9	0.7
X1265	P246	Ribe Lindegården	6/6	0.03	0.01	0.1	0.1	0.1	3.2

## Skeletal Dating

I'm relying heavily on the discussion in Boldsen (2009) [2], as quoted here:

*"The dating of individual skeletons is a fundamental problem in historical studies like this, and even the period of usage of each cemetery raises some serious problems. However, most cemeteries have at least some documentary sources broadly framing them in time. The most intensely studied skeletal samples, Tirup and Westerhus, are really the only exceptions in being dated solely on archaeological evidence (Kieffer-Olsen et al. 1986, Siv n 2005)."*

*"In medieval graves the position of the arms in relation to the rest of the skeleton in the grave is the only feature that systematically indicates dating of the burial within the temporal frame provided by the period of usage of the cemetery. Arm position dating is primarily based on work by Redin (1976) and Kieffer-Olsen (1993). The successive stages of arm position from A (the arms besides the body) over B (hand joint over the lower part of the abdomen and usually found in the pelvis) and C (the forearms over the upper part of the abdomen and the elbows flexed in an approximately right angle) to D (the hands placed on the shoulders, forearms often crossed over the chest) have primarily been described by Kieffer-Olsen (1993) but Jantzen et al. (1994) have slightly modified the transition dates between the various stages."*



**Figure 2:** Arm position dating.

**Question for Julia Gamble and Jesper Boldsen.** Can I use this diagram to broadly assign time periods to arm positions? If so, I can add greater resolution to Figure 3.

- Position A: 1150 - 1325
- Position B: 1250 - 1425
- Position C: 1350 - 1550
- Position D: 1450 - 1550

# Molecular Dating

A Bayesian Evaluation of Temporal Signal (BETS) revealed decisive support for temporal signal during the Second Pandemic, with the relaxed clock model having the highest likelihood (Table 4). After running the Monte Carlo Markov Chain (MCMC) analysis for a total of 150,000,000 generations, the estimated sample size (ESS) scores were sufficiently high (>200) for all parameters except for the tip-date of Russian strain Azov38, which could not be refined beyond a 95% highest posterior density (HPD) interval of 1553 to 1686 (Figure 12).

**Table 4:** Summary of clock model comparisons using a Bayesian Evaluation of Temporal Signal (BETS) analysis.

Model	Dates	Likelihood	Dates vs. No Dates	Relaxed Clock vs. Strict Clock
Relaxed Clock	Yes	<b>-5947948</b>	715	140
	No	-5948663	-	174
Strict Clock	Yes	-5948088	749	-
	No	-5948837	-	-

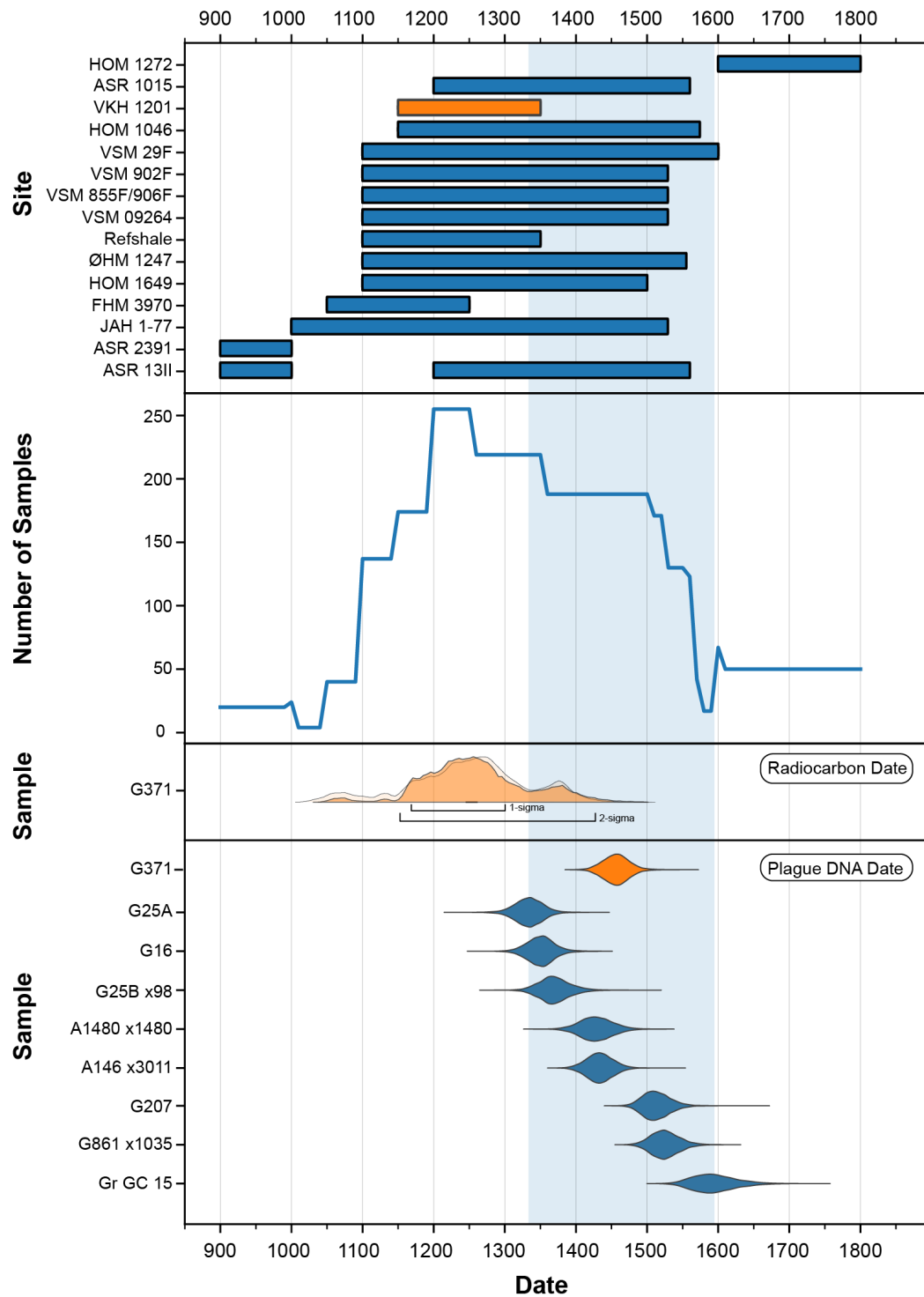
The skeletal and molecular dates of the 9 high coverage *Y. pestis* genomes are presented in Table 5 and Figure 3. Almost all molecular dates had overlap with the archaeological dates as determined by the site occupation period and burial patterns. The exception to this pattern was individual G371 from the site of Tirup.

To investigate the dating disparity for G371, we performed 14C radiocarbon dating. The radiocarbon estimate largely agrees with the archaeological dates, with an estimated mean date of 1260 CE (+/- 75 yrs with 1 sigma). Thus there is robust evidence that the individual lived and died sometime between the late 12th and early 14th century. However, the associated *Y. pestis* genome is dated to the 15th century. While the 2 sigma distribution of the radiocarbon date does extend into the 15th century, there remains a substantial conflict between the date of the host and the associated pathogen. We then critically examined the phylogeny of *Y. pestis* to assess whether individual G371 was an outlier compared to the phylogenetic position of other samples from Horsens.

**Table 5:** Summary of the *Y. pestis* molecular dates. The estimated tip date reflects the 95% highest posterior density.

ID	Region	Site	Site Occupation	Arm Position	Skeletal Date	Tip Date
G16	Ribe	Ribe Gräbrødre	1200 - 1560	C	1350 - 1550	1310 - 1388
G861 x1035	Ribe	Ribe Gräbrødre	1200 - 1560	C	1350 - 1550	1489 - 1567
G25B x98	Ribe	Ribe Gräbrødre	1200 - 1560	C	1350 - 1550	1327 - 1414
G25A	Ribe	Ribe Gräbrødre	1200 - 1560	C	1350 - 1550	1295 - 1375
G207	Ribe	Ribe Gräbrødre	1200 - 1560	C	1350 - 1550	1477 - 1551
A146 x3011	Horsens	Sejet	1150 - 1574	B	1250 - 1425	1397 - 1470
A1480 x1480	Horsens	Ole Wormsgade	1100 - 1500	?	?	1384 - 1473
G371	Horsens	Tirup	1150 - 1350	B	1250 - 1425	1419 - 1490
Gr GC 15	Viborg	Faldborg	1100 - 1600	C	1350 - 1550	1539 - 1655

**Question for Vaughan Grimes.** Is the interpretation/visual of the radiocarbon date appropriate?



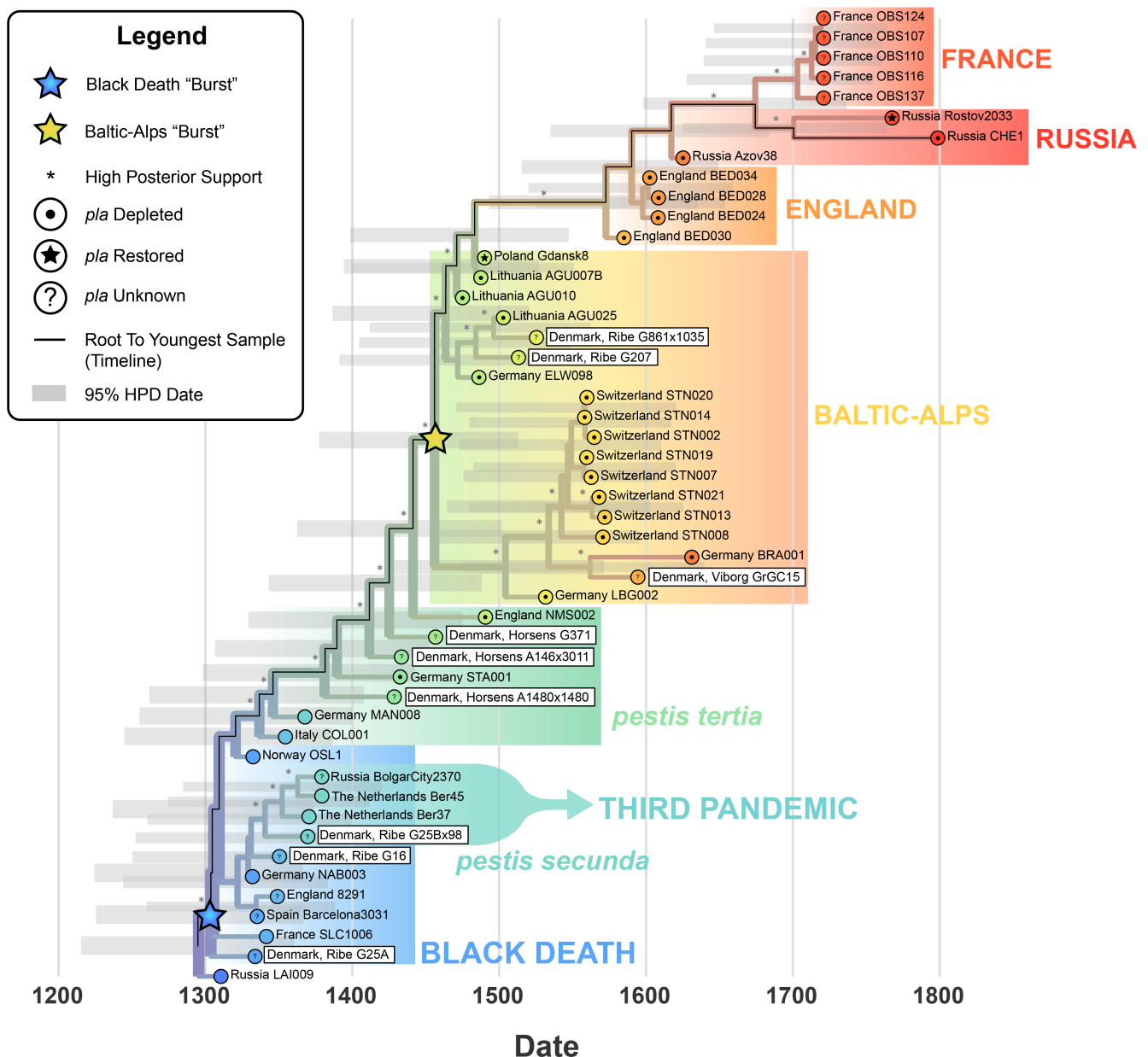
**Figure 3:** Timeline of archaeological sites and plague-positive individuals. The blue shaded range spans the highest probability period from the oldest to the youngest sample. The color orange indicates the sample and site with disparate dates.



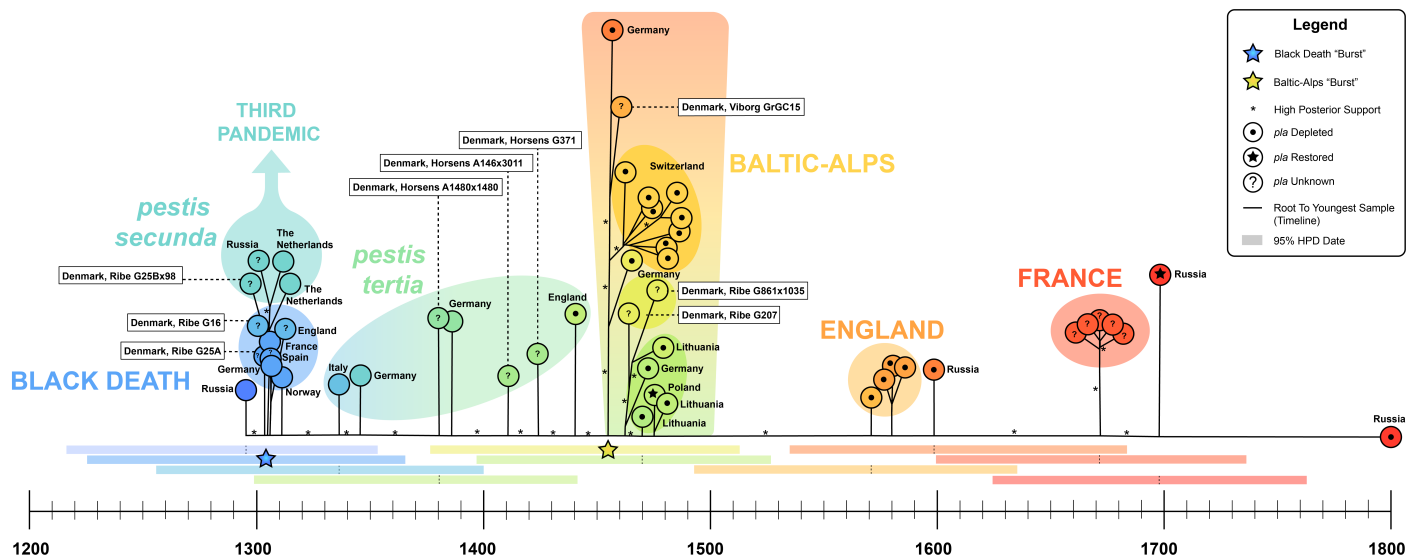
## Phylogeny

A time-scaled phylogeny was estimated to compare the 9 Danish *Y. pestis* genomes to 40 previously published Second Pandemic samples (Figure 4). The temporal structure of the Second Pandemic is also visualized as a timeline, which re-orientates the time-scaled phylogeny to trace a path from the root to the most recently collected sample (Figure 5). The geographic distribution of *Y. pestis* genomes used in the phylogenetic analysis is displayed in Figure 6.

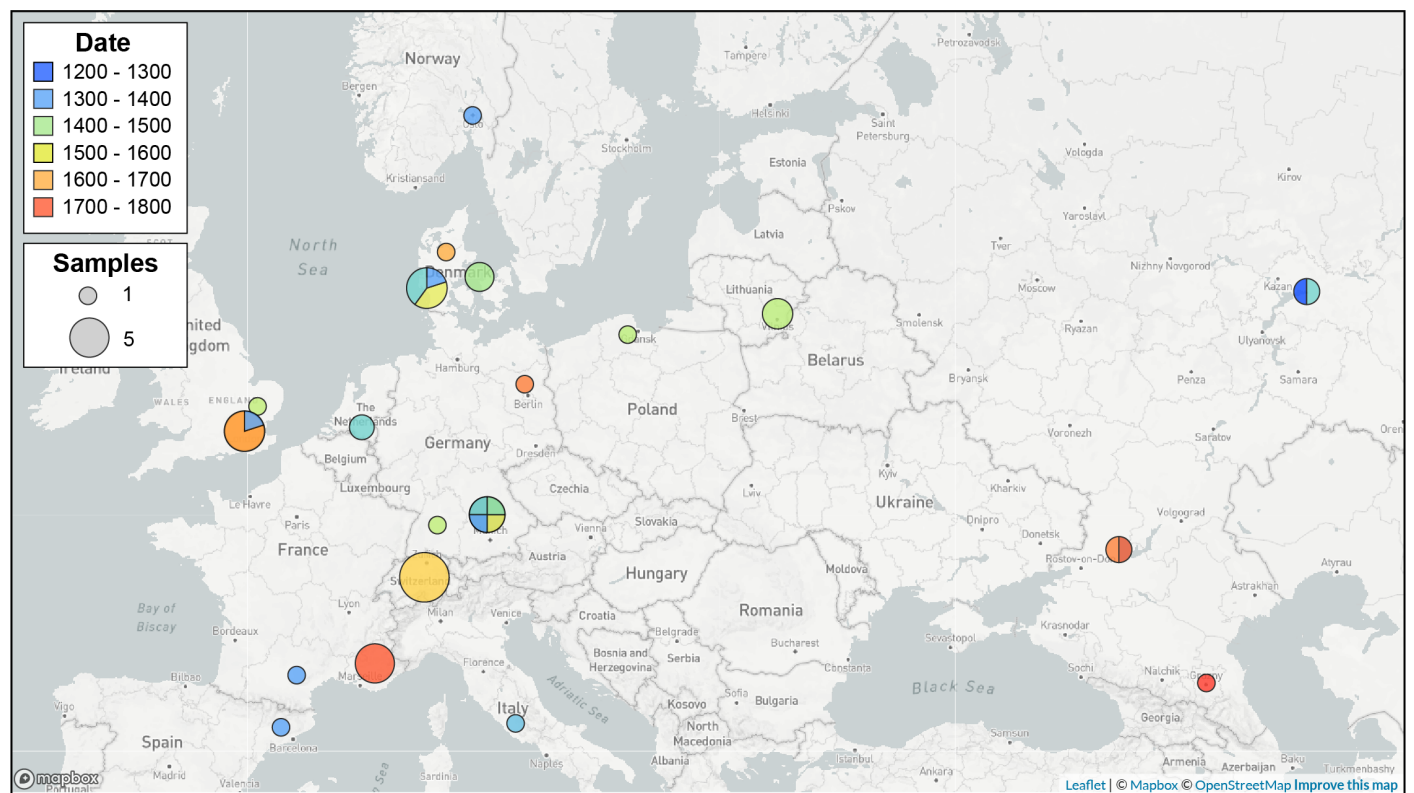
*Y. pestis* genomes from Denmark do not form a single, geographically restricted clade. Instead, Danish plague is distributed throughout the phylogeny, in a similar fashion to the genomes retrieved from Germany and England. This distribution aligns closely with historical documentation describing the multiple waves of ‘pestilence’ that affected medieval Europe [3], and thus we use this theoretical framework to contextualized the observed genetic diversity in Denmark.



**Figure 4:** A time-scaled phylogeny of the Second Plague Pandemic.



**Figure 5:** A timeline phylogeny of the Second Plague Pandemic.



**Figure 6:** Geographic distribution of Second Pandemic samples used in the *Yersinia pestis* phylogeny.

## The First Phase: Black Death

Two individuals from Ribe, G25A and G16, cluster with *Y. pestis* strains dated to the 14th century (Table [10](#)). This cluster is thought to be associated with the Black Death (1346-1353), as the genomes are nearly identical and have been isolated from all across Europe, including France, Spain, England, Germany, and Norway. These genomes mark the first observation of *Y. pestis* in Denmark and are congruent with historical records that document the arrival of the Black Death in Ribe in 1349 [\[4\]](#).

Similar to the rest of Western Europe, *Y. pestis* appears suddenly in Denmark in the 14th century. No evidence of *Y. pestis* was found in Denmark during the preceding centuries, suggesting plague was a relatively new disease for medieval Danish populations. Unfortunately, the high degree of genetic similarity means that the branching patterns, and thus dispersal of *Y. pestis*, cannot be resolved during this period. Our understanding of the Black Death clade is that of a “burst” radiation, with limited genetic diversity spreading rapidly over a vast geographic area. This pattern is typical of epidemic spread, and has also been observed at the advent of the Third Plague Pandemic [???].

### Kat's Note:

- G25A has the lowest sequencing depth, with a mean chromosomal depth of 3.8 molecules per nucleotide. This might mean that diagnostic positions for the post-Black Death clade are missing as “false” negatives. This would make it seem earlier/less-derived if those mutations are ‘missing’. I will investigate!

## The Second Phase: *pestis secunda*

Interestingly, the *Y. pestis* genome associated with individual G25Bx98 (who was found in the same grave as G25A?) is genetically distinct from the earlier strains from Ribe. This isolate falls within a clade of high epidemiological significance, which is the ancestral group giving rise to the Third Pandemic of plague and the rest of Branch 1. Historically, this clade has been linked to the *pestis secunda*, a wave of plague outbreaks dated between 1357 to 1366 [\[5,6\]](#).

The phylogenetic position of samples within this clade has been hypothesized to reflect a “backward” migration of plague from Northern Europe into Asia. The *Y. pestis* genome retrieved from G25Bx98 tentatively supports this hypothesis, as it falls basal to the more derived strains from The Netherlands and Russia. However, directionality cannot be robustly inferred from four samples alone, particularly given the strong Western European sampling bias of Second Pandemic samples. It will be an important avenue of future research to further develop the relationship between Northern European plague and the only Second Pandemic lineage that is known to persist until the present.

## The Tertiary Phases: *pestis tertia*

In contrast to the genetic homogeneity observed across Europe during the Black Death period, isolates of plague in the post-Black Death period are easily distinguished. Three genomes collected from three different sites near Horsens derive from independent emergences, despite having temporal overlap. This has also been observed in Germany during this period, as samples collected in relatively close proximity are genetically distinct. A product of this increased genetic diversity means that branching patterns are well resolved in the post-Black Death period. Unfortunately, the geographic origins and dispersal of plague are still challenging to reconstruct, as the number of genomes (N=7) is sparsely sampled relative to the minimum number of countries (N=4) that are implicated.

The pattern of independent emergence is the defining dynamic of plague during this period as all *Y. pestis* collected after the Black Death, but prior to the Early Modern Period, are unique lineages. The transition captured here, from little genetic diversity spread across a continent to significant diversity accumulating within a country, may indicate “boom-bust” dynamics [7].

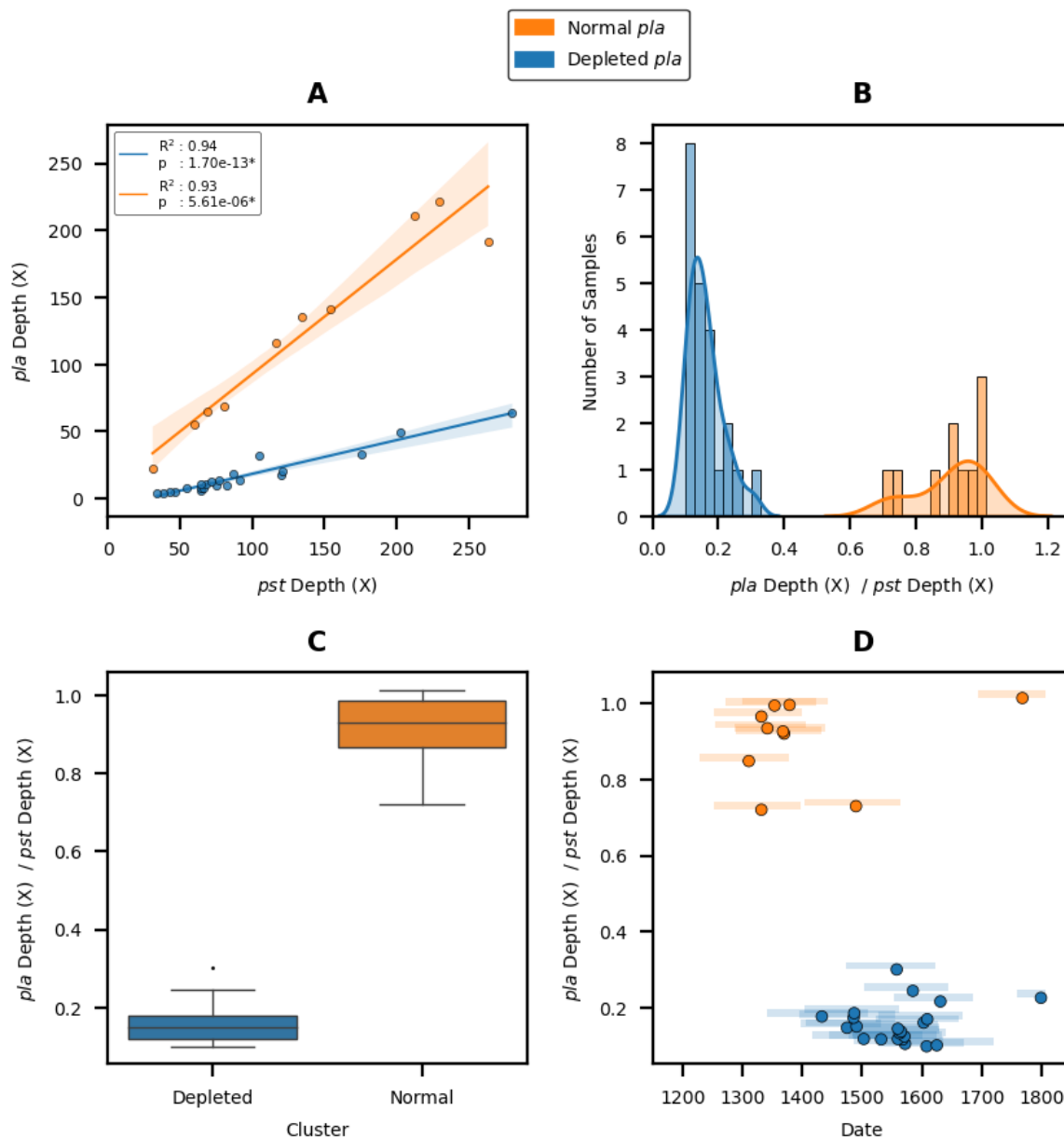
■ The epidemiological interpretation of this transition requires more thought.

Another defining characteristic of the post-Black Death period is depletion of a key virulence factor, the *plasminogen activator* (*pla*). Figure 7 compares the sequencing depth of the *pla* gene to its corresponding plasmid pPCP1 across samples from the Second Pandemic. Two linear trends are observed separating samples with ‘normal’ gene-to-plasmid ratios from samples that have a ‘depleted’ gene-to-plasmid ratio. This event was previously observed [Susat 2020 *Yersinia Pestis* Strains | [8]], and found in samples from as early as the 15th century.

The three Danish genomes from Horsens have temporal overlap with the 15th century and may potentially capture the transition from to a *pla*-depleted state. Unfortunately, the sequencing depth of the pPCP1 plasmid in these samples is insufficient for statistical analysis. The results of a targeted enrichment for the pPCP1 plasmid will be the subject of a forthcoming publication.

■ Ravneet has completed this experiment and has exciting results, stay tuned!

## Relative depletion of the plasminogen activator (*pla*) virulence factor



**Figure 7:** Relative Depletion of the plasminogen activator (*pla*) virulence factor. A: A linear regression of the depleted gene (*pla*) depth on the pPCP1 baseline gene (*pst*) depth. B: The distribution of *pla* depletion ratios. C: Variation within and between the normal and depleted *pla* clusters. D: The relationship between collection date and *pla* depletion. Colored bars represent the 95% HPD on estimated tip date.

## Early Modern Plague

Following the Medieval Period, the evolutionary dynamic of *Y. pestis* changes yet again and a second “Burst” event occurs. Similar to the Black Death “Burst”, this event begins with a series of divergences in quick succession, producing highly similar genomes in contemporaneous individuals from Germany, Lithuania, and Poland. Following these divergences, two major lineages emerge.

The first lineage, termed the Baltic-Alps, has been previously observed in the Baltic countries (Lithuania, Poland, Germany) and the Alps (Switzerland). Three Danish genomes fall within the Baltic-Alps clade, including two samples from coastal Ribe and one sample from inland Viborg. The addition of these samples strengthens the geographic ties of this lineage to the Baltic Region. The

epidemiological significance of this localization has been speculated to indicate the formation of a novel plague reservoir within Europe.

The second lineage that emerges in this period gives rise to later epidemics of plague in Russia, England, and France, including the Great Plague of Marseille (1720-1721). No Danish plague in this study is associated with this lineage. As sampling efforts intensify, it will be interesting to see if these two lineages remain geographically distinct, and to investigate what might be obstructing the flow of plague between them.

## Conclusion

---

### Results

**The earliest evidence of Danish *Y. pestis* is found in Ribe and dates to the mid-14th century.** This suggests that plague may have been a relatively new disease for Danish populations at the time, and was unlikely to have affected generations immediately prior. However, the Viking Age (900-1000), which is the earliest time period examined, is under-sampled and thus this absence of evidence does not prove the absence of plague in the Viking Age or earlier.

***Y. pestis* in medieval Denmark (14th-15th century) reveals population structure that closely aligns with historical records of plague.** Specifically, temporal and geographic patterns are observed that follow the documentation of successive epidemics sweeping across Europe, such as the primary (Black Death), secondary (*pestis secunda*), and tertiary (*pestis tertia*) waves.

***Y. pestis* in pre-modern Denmark (16th century) forms a geographically-restricted lineage with other samples collected from the Baltic countries and the Alps.** This unique geographic structure supports the theory concerning formation of a novel plague reservoir.

**Three sequential genomes from the Horsens region capture a key virulence change, in which the plasminogen activator (*pla*) becomes depleted for the remainder of the Second Pandemic.** Restoration of this virulence factor is not observed again in Denmark, although it is observed to be intermittently restored in Poland and Russia.

***Y. pestis* is detected across diverse individuals and mortuary practices.** Plague was identified in both adults and sub-adults, rural and urban settlements, and in single and multiple burials. This finding contributes to discourse on epidemic mortuary practices, which have primarily focused on the practice of mass graves and plague pits such as those observed in large centres such as London.

### Significance

- This study is the most intensive longitudinal study of plague in a single region, both in terms of time span and geographic sampling.
- Additional evidence that *Y. pestis* was a novel pathogen for European populations, will help guide immunity-related work seeking to identify changes.

First, there have been relatively few studies that explore the genetics of plague in Scandinavia across time and geography, particularly in Denmark. Second, this paper contributes to a larger body of epidemiological literature that considers the patterns and mechanisms by which diseases emerge, propagate, and go extinct.

# Methods

---

## DNA Extraction

Ancient DNA laboratory work was performed in dedicated clean-room facilities at the McMaster Ancient DNA Centre (Hamilton, ON). A single root of each individual's molar was sectioned into two subsamples (50-100 mg) using a circular Dremel. Demineralisation and digestion were performed as previously described [9] and DNA extraction was conducted using a specialized protocol designed for ancient DNA [10]. Reagent blanks were introduced as negative controls to monitor DNA contamination in subsequent steps.

## *Yersinia pestis* PCR screening

An initial plague-screening PCR was performed in duplicate on 1:10 extract dilutions using a *pla* assay [11]. In brief, the PCR primers used in this study target the 3'UTR of the *pla* gene which has reduced sequence similarity in non-*Yersinia* species. The forward primer used is thus far known to be identical to *Yersinia pestis* only. A second round of the *pla* PCR assay was performed for all extracts that amplified in the initial PCR, using the 1:10 dilution in duplicate and the original concentration in duplicate. In total, 6 PCR replicates were performed for each plague-positive sample.

## Shotgun Sequencing

The extracted DNA of plague-positive individuals was converted into Illumina sequencing libraries using a modified protocol [12] and quantified using an Illumina library qPCR assay. Sample libraries were then pooled at equimolar concentrations while negative controls were sequenced at maximum volume input to maximize detection of contaminant organisms. Paired-end sequencing was performed on an Illumina HiSeq 1500 platform (Farncombe Family Digestive Health Research Institute, Hamilton, ON).

## Targeted Sequencing

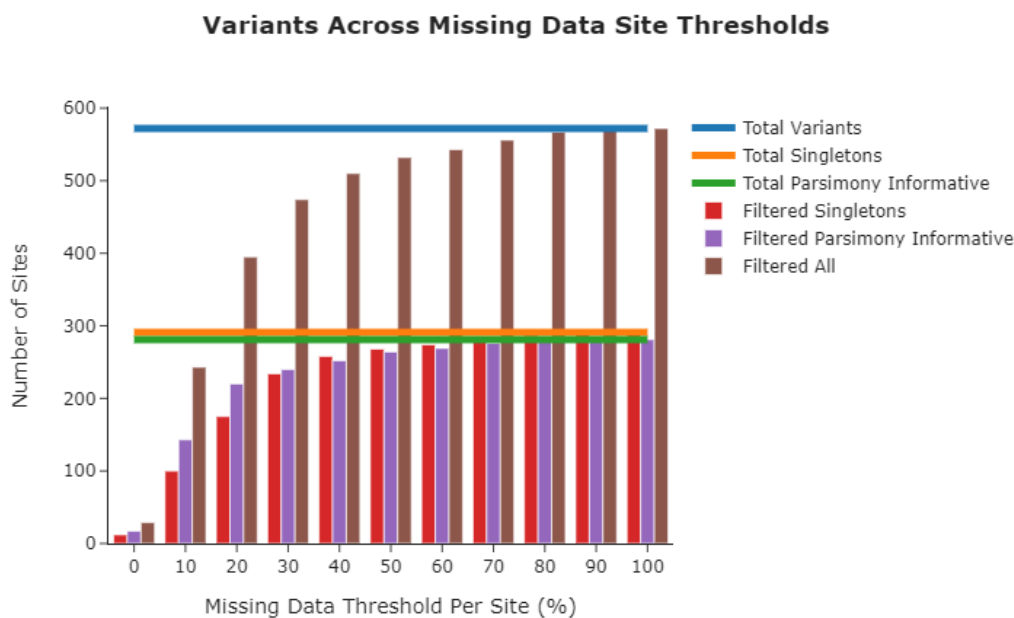
In-solution enrichment for the pan-genome of *Yersinia pestis* was performed with a previously designed bait-set [11] and using the myBaits v4 protocol. The following modifications were incorporated to improve recovery of degraded and divergent DNA sequences: 5uL library input, 100ng bait concentration, hybridisation at 60°C, 16–24h hybridisation capture, and two rounds of enrichment. The enriched libraries were quantified using an Illumina library quantification qPCR assay and pooled at maximum input volume (13uL) due to low concentration. Following pooling, libraries were size-selected on an agarose gel to retain 150–500bp fragments which corresponds to molecule lengths of approximately 15–365bp without the adapter sequences. Paired-end sequencing was performed on an Illumina HiSeq 1500 platform at the Farncombe Metagenomics Facility (Hamilton, ON).

# Comparative Genomes

Ancient unassembled genomes from the Second Plague Pandemic were identified using NCBI meta [14] and downloaded from the SRA database in FASTQ format using the SRA Toolkit [15] (Table 7). *Y. pestis* strain CO92 was used as the reference genome for sequence alignment and annotation (Table 8). All assembled genomes belonging to phylogenetic branch 0.ANT3 were downloaded and used as an outgroup to root the maximum-likelihood phylogeny (Table 9).

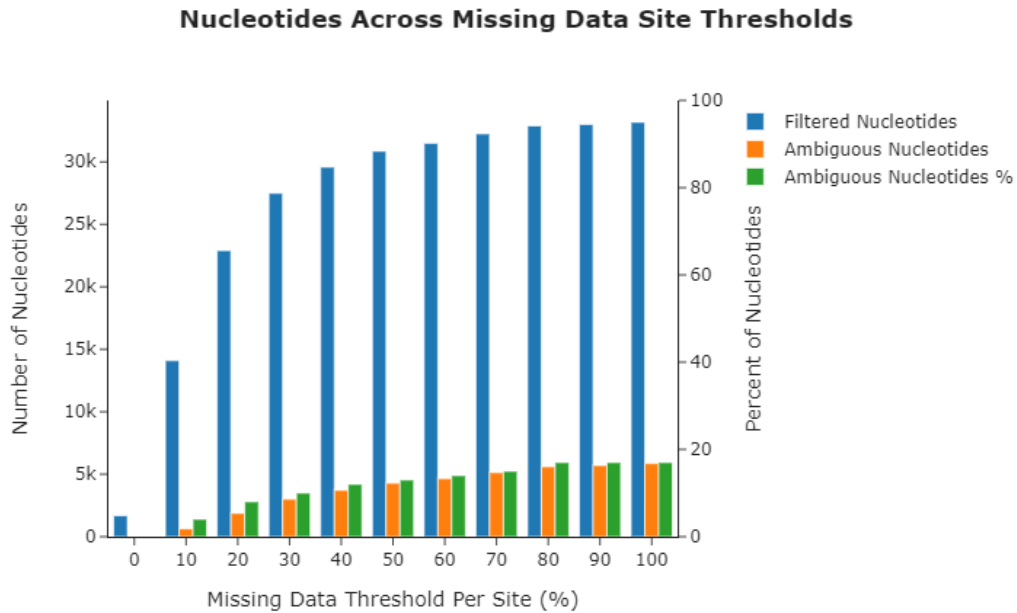
## Genomic Alignment

Pre-processing and alignment to the reference genome was performed using the *nf-core/eager* pipeline, a reproducible workflow for ancient genome reconstruction [16]. A multiple sequence alignment was constructed using the Snippy Core module of the Snippy pipeline [17]. The output alignment was filtered to only include chromosomal variants and to exclude sites that had no more than 30% missing data (ie. no more than 30% of samples having an ambiguous nucleotide). The value of 30% was selected as the most permissive threshold where the number of shared, parsimony-informative sites (240) was still larger than the number of singleton sites (234) which are observed in a single genome (Figure 8). Furthermore, a 30% threshold was selected as ambiguous nucleotides made up no more than 10% of the alignment (9). The filtered chromosomal alignment contained 4,289,810 constant sites, as well as 474 variants when the outgroup clade (0.ANT3) was included and 356 variants when the outgroup was excluded.



**Figure 8:** The number of variant positions used in the multiple alignment according to different missing data thresholds.





**Figure 9:** The number of ambiguous nucleotides used in the multiple alignment according to different missing data thresholds.

## Phylogeny

Model selection was performed using Modelfinder which identified the K3Pu+F+I model as the optimal choice based on the Bayesian Information Criterion (BIC) [18]. A maximum-likelihood phylogeny was then estimated across 10 independent runs of IQTREE [19]. Branch support was evaluated using 1000 iterations of the ultrafast bootstrap approximation [20], with a threshold of 95% required for strong support. The outgroup clade (0.ANT3) was used to estimate the root position and was subsequently pruned from the phylogeny for downstream analysis and visualization.

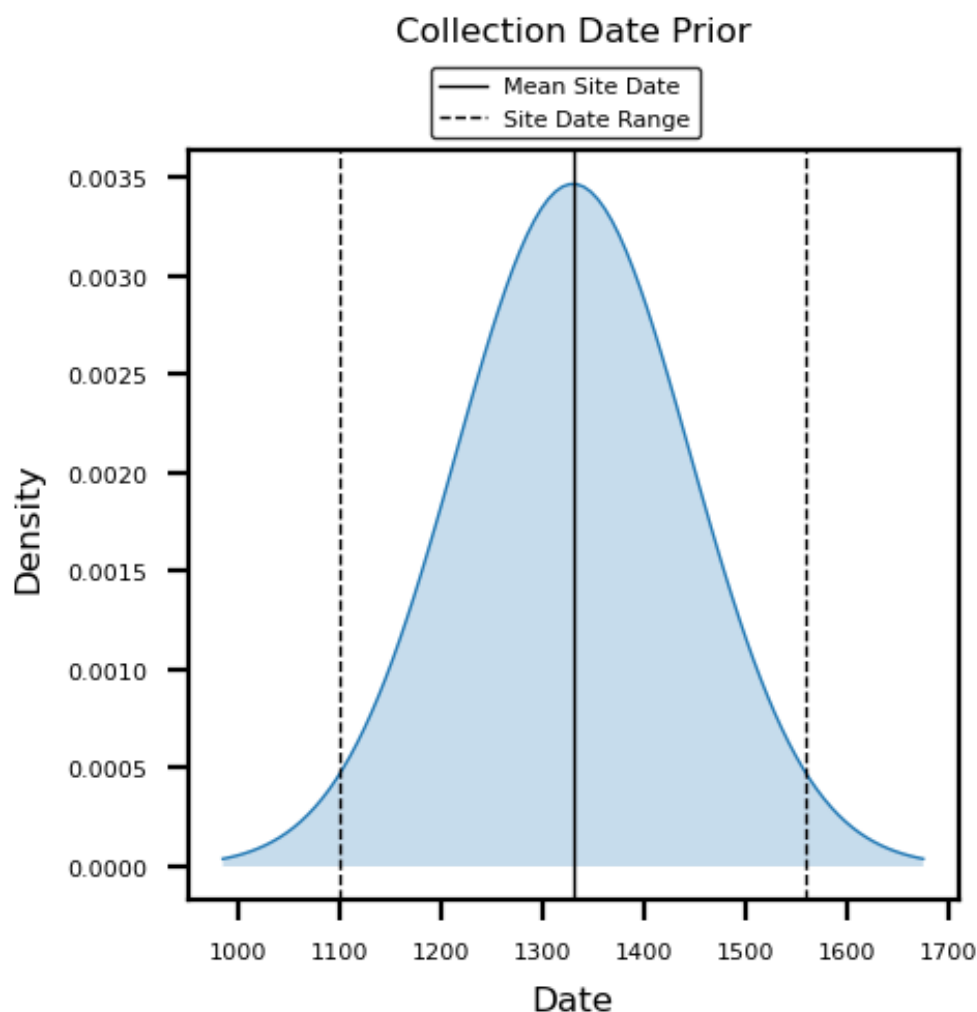
## Phylodynamics

To evaluate the degree of temporal signal present, a Bayesian Evaluation of Temporal Signal (BETS) was first performed using BEAST v2.6.2. As *Y. pestis* exhibits significant rate variation between clades [21], the chromosomal alignment that excludes the outgroup clade (0.ANT3) was used. To robustly estimate the root position, strain LAI009 from Russia was specified as the outgroup as this sample falls basal to all other Second Pandemic genomes in the maximum-likelihood phylogeny (Figure 11) as well as in previously published analyses [22].

Briefly, each candidate model was tested using the correct collection dates of all samples and then compared to the same model with all collection dates assumed to be contemporaneous. A constant population size was assumed as the coalescent prior to minimize additional parameter variation. Two candidate models were tested: a strict clock and an uncorrelated relaxed lognormal clock with default priors and starting positions.

For the models that incorporate the correct collection date, individual priors were constructed for all samples. Strain 8291 from England and all OBS strains from France were assumed to have fixed dates of 1349 and 1721 respectively, as their collection date uncertainty spans only 2 years (1348-1350 and 1720-1722). For all other previously published genomes, a diffuse normal prior was constructed using the mean radiocarbon/mortuary date and half the uncertainty as the standard deviation. Individual

priors for the new Danish samples were similarly constructed using the widest possible occupation dates of plague-positive sites (Figure 10).



**Figure 10:** Individual tip-date prior used to estimate the collection date of all Danish *Y. pestis* genomes.

Bayes factors (BF) were calculated by comparing the marginal likelihoods of each model, as estimated with a generalized stepping stone (GSS) computation across 100 chains each sampled over 1,000,000 generations. The model with the highest marginal likelihood was then run fully for 150,000,000 generations to ensure the effective sample size (ESS) of all relevant parameters was greater than 200.

# References

---

1. **Emergence and spread of basal lineages of *Yersinia pestis* during the Neolithic Decline**  
Nicolás Rascovan, Karl-Göran Sjögren, Kristian Kristiansen, Rasmus Nielsen, Eske Willerslev, Christelle Desnues, Simon Rasmussen  
*Cell* (2019-01-10) [https://www.cell.com/cell/abstract/S0092-8674\(18\)31464-8](https://www.cell.com/cell/abstract/S0092-8674(18)31464-8)  
DOI: [10.1016/j.cell.2018.11.005](https://doi.org/10.1016/j.cell.2018.11.005) · PMID: [30528431](https://pubmed.ncbi.nlm.nih.gov/30528431/)
2. **Leprosy in Medieval Denmark — Osteological and epidemiological analyses**  
Jesper L. Boldsen  
*Anthropologischer Anzeiger* (2009) <http://www.jstor.org/stable/29543069>
3. **Out of the West: Formation of a Permanent Plague Reservoir in South-Central Germany (1349–1356) and its Implications\***  
Philip Slavin  
*Past & Present* (2021-01-25) <https://doi.org/10.1093/pastj/gtaa028>  
DOI: [10.1093/pastj/gtaa028](https://doi.org/10.1093/pastj/gtaa028)
4. **The Black Death**  
Kristina Lenz, Nils Hybel  
*Scandinavian Journal of History* (2016) [https://journals.scholarsportal.info/details/03468755/v41i0001/54\\_tbd.xml](https://journals.scholarsportal.info/details/03468755/v41i0001/54_tbd.xml)  
DOI: [10.1080/03468755.2015.1110533](https://doi.org/10.1080/03468755.2015.1110533)
5. **A genomic and historical synthesis of plague in 18th century Eurasia**  
Meriam Guellil, Oliver Kersten, Amine Namouchi, Stefania Luciani, Isolina Marota, Caroline A. Arcini, Elisabeth Iregren, Robert A. Lindemann, Gunnar Warfvinge, Lela Bakanidze, ... Barbara Bramanti  
*Proceedings of the National Academy of Sciences* (2020-11-10) <https://www.pnas.org/content/117/45/28328>  
DOI: [10.1073/pnas.2009677117](https://doi.org/10.1073/pnas.2009677117) · PMID: [33106412](https://pubmed.ncbi.nlm.nih.gov/33106412/)
6. **Integrative approach using *Yersinia pestis* genomes to revisit the historical landscape of plague during the Medieval Period**  
Amine Namouchi, Meriam Guellil, Oliver Kersten, Stephanie Hänsch, Claudio Ottoni, Boris V. Schmid, Elsa Pacciani, Luisa Quaglia, Marco Vermunt, Egil L. Bauer, ... Barbara Bramanti  
*Proceedings of the National Academy of Sciences* (2018-12-11) <http://www.pnas.org/lookup/doi/10.1073/pnas.1812865115>  
DOI: [10.1073/pnas.1812865115](https://doi.org/10.1073/pnas.1812865115)
7. **Boom-bust dynamics in biological invasions: towards an improved application of the concept**  
David L. Strayer, Carla M. D'Antonio, Franz Essl, Mike S. Fowler, Juergen Geist, Sabine Hilt, Ivan Jarić, Klaus Jöhnk, Clive G. Jones, Xavier Lambin, ... Jonathan M. Jeschke  
*Ecology Letters* (2017-10)  
DOI: [10.1111/ele.12822](https://doi.org/10.1111/ele.12822) · PMID: [28834087](https://pubmed.ncbi.nlm.nih.gov/28834087/)
8. ***Yersinia pestis* strains from Latvia show depletion of the *pla* virulence gene at the end of the second plague pandemic**  
Julian Susat, Joanna H. Bonczarowska, Elīna Pētersone-Gordina, Alexander Immel, Almut Nebel, Guntis Gerhards, Ben Krause-Kyora

*Scientific Reports* (2020-09-03) <https://www.nature.com/articles/s41598-020-71530-9>  
DOI: [10.1038/s41598-020-71530-9](https://doi.org/10.1038/s41598-020-71530-9)

9. **New insights from old bones: DNA preservation and degradation in permafrost preserved mammoth remains**  
Carsten Schwarz, Regis Debruyne, Melanie Kuch, Elizabeth McNally, Henry Schwarcz, Andrew D. Aubrey, Jeffrey Bada, Hendrik Poinar  
*Nucleic Acids Research* (2009-06-01) <https://doi.org/10.1093/nar/gkp159>  
DOI: [10.1093/nar/gkp159](https://doi.org/10.1093/nar/gkp159)
10. **Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments**  
J. Dabney, M. Knapp, I. Glocke, M.-T. Gansauge, A. Weihmann, B. Nickel, C. Valdiosera, N. Garcia, S. Paabo, J.-L. Arsuaga, M. Meyer  
*Proceedings of the National Academy of Sciences* (2013-09-24) <http://www.pnas.org/cgi/doi/10.1073/pnas.1314445110>  
DOI: [10.1073/pnas.1314445110](https://doi.org/10.1073/pnas.1314445110)
11. ***Yersinia pestis* and the Plague of Justinian 541–543 AD: a genomic analysis**  
David M Wagner, Jennifer Klunk, Michaela Harbeck, Alison Devault, Nicholas Waglechner, Jason W Sahl, Jacob Enk, Dawn N Birdsell, Melanie Kuch, Candice Lumibao, ... Hendrik Poinar  
*The Lancet Infectious Diseases* (2014-04) <https://linkinghub.elsevier.com/retrieve/pii/S1473309913703232>  
DOI: [10.1016/s1473-3099\(13\)70323-2](https://doi.org/10.1016/s1473-3099(13)70323-2)
12. **Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing**  
M. Meyer, M. Kircher  
*Cold Spring Harbor Protocols* (2010-06-01) <http://www.cshprotocols.org/cgi/doi/10.1101/pdb.prot5448>  
DOI: [10.1101/pdb.prot5448](https://doi.org/10.1101/pdb.prot5448)
13. **Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform**  
Martin Kircher, Susanna Sawyer, Matthias Meyer  
*Nucleic Acids Research* (2012-01-01) <https://academic.oup.com/nar/article/40/1/e3/1287690>  
DOI: [10.1093/nar/gkr771](https://doi.org/10.1093/nar/gkr771)
14. **NCBImeta: efficient and comprehensive metadata retrieval from NCBI databases**  
Katherine Eaton  
*Journal of Open Source Software* (2020) <https://doi.org/10.21105/joss.01990>  
DOI: [10.21105/joss.01990](https://doi.org/10.21105/joss.01990)
15. **ncbi/sra-tools**  
NCBI - National Center for Biotechnology Information/NLM/NIH  
(2021-05-18) <https://github.com/ncbi/sra-tools>
16. **Reproducible, portable, and efficient ancient genome reconstruction with nf-core/eager**  
James A. Fellows Yates, Thisseas C. Lamnidis, Maxime Borry, Aida Andrades Valtueña, Zandra Fagernäs, Stephen Clayton, Maxime U. Garcia, Judith Neukamm, Alexander Peltzer  
*PeerJ* (2021-03-16) <https://peerj.com/articles/10947>  
DOI: [10.7717/peerj.10947](https://doi.org/10.7717/peerj.10947)

17. **Snippy: Rapid haploid variant calling and core genome alignment.**  
Torsten Seemann  
(2020-03-08) <https://github.com/tseemann/snippy>
18. **ModelFinder: fast model selection for accurate phylogenetic estimates**  
Subha Kalyaanamoorthy, Bui Quang Minh, Thomas K. F. Wong, Arndt von Haeseler, Lars S. Jermiin  
*Nature Methods* (2017-06) <http://www.nature.com/articles/nmeth.4285>  
DOI: [10.1038/nmeth.4285](https://doi.org/10.1038/nmeth.4285)
19. **IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era**  
Bui Quang Minh, Heiko A. Schmidt, Olga Chernomor, Dominik Schrempf, Michael D. Woodhams, Arndt von Haeseler, Robert Lanfear  
*Molecular Biology and Evolution* (2020-05-01) <https://academic.oup.com/mbe/article/37/5/1530/5721363>  
DOI: [10.1093/molbev/msaa015](https://doi.org/10.1093/molbev/msaa015)
20. **UFBoot2: Improving the Ultrafast Bootstrap Approximation**  
Diep Thi Hoang, Olga Chernomor, Arndt von Haeseler, Bui Quang Minh, Le Sy Vinh  
*Molecular Biology and Evolution* (2018-02-01) <https://academic.oup.com/mbe/article/35/2/518/4565479>  
DOI: [10.1093/molbev/msx281](https://doi.org/10.1093/molbev/msx281)
21. **Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis***  
Y. Cui, C. Yu, Y. Yan, D. Li, Y. Li, T. Jombart, L. A. Weinert, Z. Wang, Z. Guo, L. Xu, ... R. Yang  
*Proceedings of the National Academy of Sciences* (2013-01-08) <http://www.pnas.org/cgi/doi/10.1073/pnas.1205750110>  
DOI: [10.1073/pnas.1205750110](https://doi.org/10.1073/pnas.1205750110)
22. **Phylogeography of the second plague pandemic revealed through analysis of historical *Yersinia pestis* genomes**  
Maria A. Spyrou, Marcel Keller, Rezeda I. Tukhbatova, Christiana L. Scheib, Elizabeth A. Nelson, Aida Andrades Valtueña, Gunnar U. Neumann, Don Walker, Amelie Alterauge, Niamh Carty, ... Johannes Krause  
*Nature Communications* (2019-10-02) <https://www.nature.com/articles/s41467-019-12154-0>  
DOI: [10.1038/s41467-019-12154-0](https://doi.org/10.1038/s41467-019-12154-0)
23. **Black Death**  
Robert Gottfried  
(1985-03-01) <https://www.simonandschuster.com/books/Black-Death/Robert-S-Gottfried/9780029123706>  
ISBN: [978-0-02-912370-6](https://www.simonandschuster.com/books/Black-Death/Robert-S-Gottfried/9780029123706)

# Appendix

## Plague Detection

Table 6: Plague false positive summary.

Arch ID	Project ID	Site Code	PCR	Human (%)	Plague (%)	Chrom	pCD1	pMT1	pPCP1
G70 x212	R44	ASR 1015	3/6	1.48	0.00	0.1	0.1	0.1	0.0
G860	R39	ASR 1015	5/6	0.09	?	?	?	?	?
G364	R43	ASR 1015	4/6	?	?	?	?	?	?
K1167 x1167	P235	ASR 13 II	3/6	?	?	?	?	?	?
A21 x23	D25	ØHM 1247	4/6	0.01	0.00	0.05	0.1	0.0	0.0
G260 K539 x876	R27	VSM 09264	3/6	?	?	?	?	?	?

## Comparative Genomes

Table 7: Second Pandemic genomes metadata.

Strain	Date	Branch	Country	Accession
STN021	1485 - 1635	1.PRE1	Switzerland	SAMEA5818830
STN020	1485 - 1635	1.PRE1	Switzerland	SAMEA5818829
STN019	1485 - 1635	1.PRE1	Switzerland	SAMEA5818828
STN014	1485 - 1635	1.PRE1	Switzerland	SAMEA5818826
STN013	1485 - 1635	1.PRE1	Switzerland	SAMEA5818825
STN008	1485 - 1635	1.PRE1	Switzerland	SAMEA5818822
STN007	1485 - 1635	1.PRE1	Switzerland	SAMEA5818821
STN002	1485 - 1635	1.PRE1	Switzerland	SAMEA5818818
STA001	1420 - 1630	1.PRE1	Germany	SAMEA5818817
NMS002	1475 - 1536	1.PRE1	England	SAMEA5818815
NAB003	1292 - 1392	1.PRE1	Germany	SAMEA5818811
MAN008	1283 - 1390	1.PRE1	Germany	SAMEA5818809
LBG002	1455 - 1632	1.PRE1	Germany	SAMEA5818808
LAI009	1300 - 1400	1.PRE0	Russia	SAMEA5818806
ELW098	1485 - 1627	1.PRE1	Germany	SAMEA5818805
BRA001	1618 - 1648	1.PRE1	Germany	SAMEA5818803
BED034	1560 - 1635	1.PRE1	England	SAMEA5818801
BED030	1560 - 1635	1.PRE1	England	SAMEA5818800
BED028	1560 - 1635	1.PRE1	England	SAMEA5818799
BED024	1560 - 1635	1.PRE1	England	SAMEA5818798

Strain	Date	Branch	Country	Accession
SLC1006	1279 - 1389	1.PRE1	France	SAMEA5054093
OSL1	1270 - 1390	1.PRE1	Norway	SAMEA5054092
Ber45	1300 - 1400	1.PRE2	The Netherlands	SAMEA5054090
Ber37	1300 - 1400	1.PRE2	The Netherlands	SAMEA5054089
BolgarCity2370	1362 - 1400	1.PRE3	Russia	SAMEA3937654
Barcelona3031	1300 - 1420	1.PRE1	Spain	SAMEA3937653
OBS137	1720 - 1720	1.PRE1	France	SAMEA3713715
OBS124	1720 - 1720	1.PRE1	France	SAMEA3713714
OBS116	1720 - 1720	1.PRE1	France	SAMEA3713713
OBS110	1720 - 1720	1.PRE1	France	SAMEA3713712
OBS107	1720 - 1720	1.PRE1	France	SAMEA3713711
8291	1348-1350	1.PRE1	England	SAMN00715800
COL001	1300 - 1400	1.PRE1	Italy	SAMEA7293136
CHE1	1500 - 1800	1.PRE1	Russia	SAMEA7293135
Rostov2033	1762 - 1773	1.PRE1	Russia	SAMEA7313236_38
Azov38	1400 - 1700	1.PRE1	Russia	SAMEA7313243_45
Gdansk8	1400 - 1700	1.PRE1	Poland	SAMEA7313246_49
AGU010	1435 - 1477	1.PRE1	Lithuania	SAMEA6651390
AGU025	1441 - 1612	1.PRE1	Lithuania	SAMEA6637004
AGU007B	1463 - 1632	1.PRE1	Lithuania	SAMEA6637002

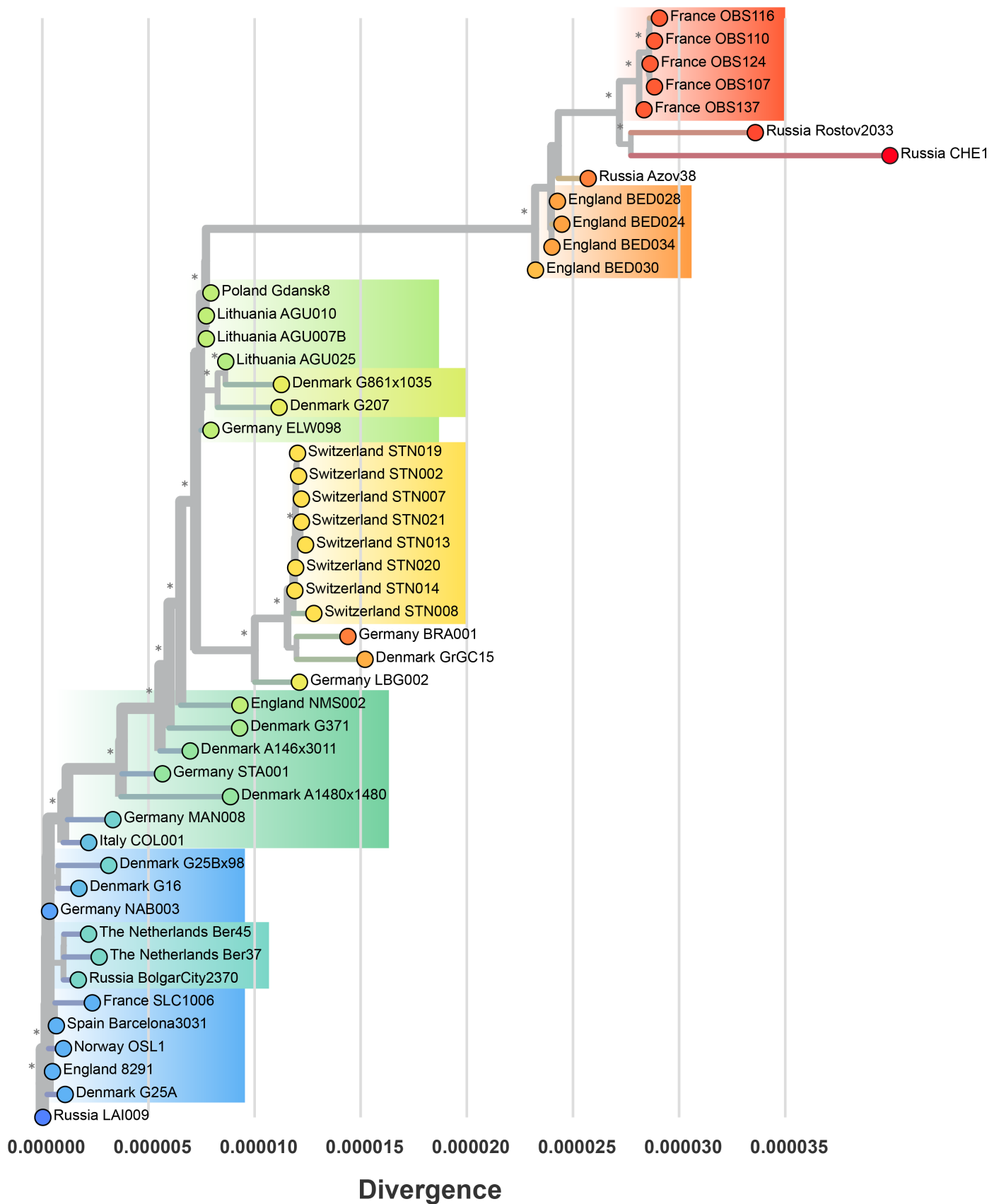
**Table 8:** Reference genome metadata.

Strain	Date	Branch	Country	Accession
CO92	1992	1.ORI1	United States of America	SAMEA1705942

**Table 9:** Outgroup genomes metadata.

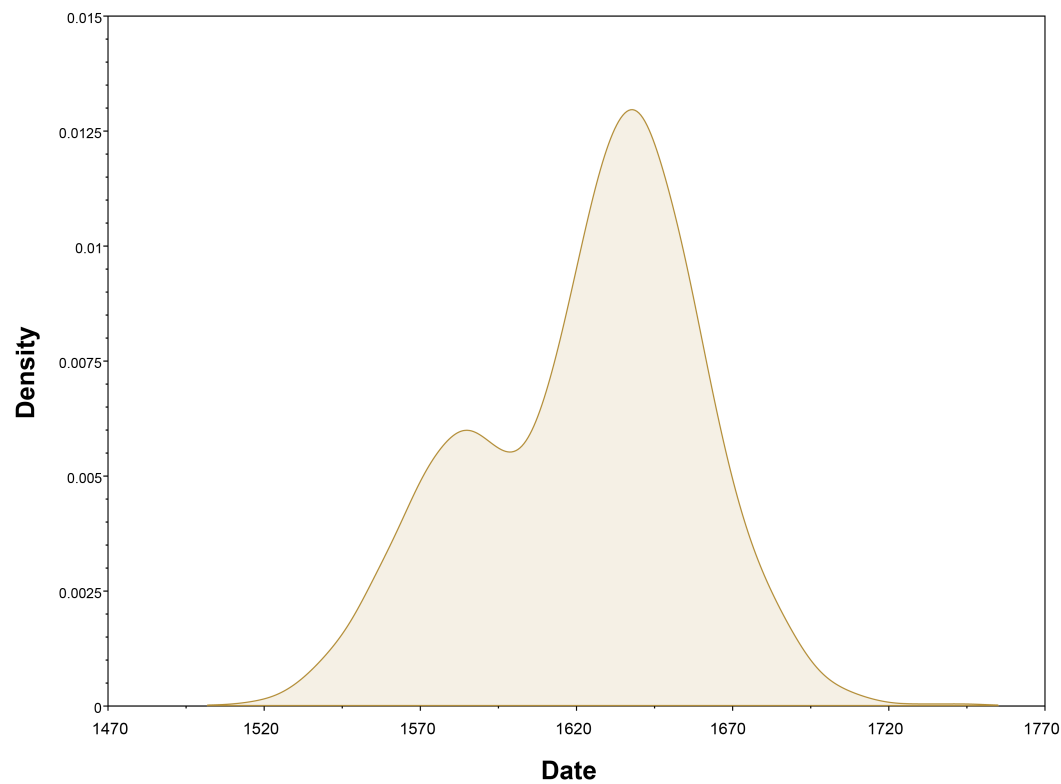
Strain	Date	Branch	Country	Accession
231	1947	0.ANT3	Kyrgyzstan	SAMN02777961
A-1486	1966	0.ANT3	Kyrgyzstan	SAMN05149973
790	1961 - 1976	0.ANT3	Kyrgyzstan	SAMN02769799
CMCC38001	1979	0.ANT3	China	SAMN02403043
CMCC21106	2001	0.ANT3	China	SAMN02403038
A1956001	1956	0.ANT3	China	SAMN02403019
42091	1999	0.ANT3	China	SAMN02403004
42082	1995	0.ANT3	China	SAMN02403003

# Phylogeny





# Phylodynamics



**Figure 12:** Tip-dating uncertainty for strain Azov38.

**Table 10:** Estimated tip-dates for the Black Death clade. Strain 8291 was assumed to have a fixed date of 1349, given the archaeological dates of 1348-1350.

Country	Strain	95% HPD Date
Denmark	G25A	1295 - 1375
Norway	OSL1	1300 - 1364
Germany	NAB003	1304 - 1359
Spain	Barcelona3031	1304 - 1364
France	SLC1006	1308 - 1376
Denmark	G16	1310 - 1388
England	8291	-

**Table 11:** Estimated tip-dates for the *pestis secunda* clade.

Country	Strain	95% HPD Date
Denmark	G25Bx98	1327 - 1414
The Netherlands	Ber37	1342 - 1398
The Netherlands	Ber45	1354 - 1404
Russia	BolgarCity2370	1363 - 1396

**Table 12:** Estimated tip-dates for the *pestis tertia* clade.

Country	Strain	95% HPD Date
Italy	COL001	1322 - 1386
Germany	MAN008	1334 - 1401
Denmark	A1480x1480	1384 - 1473
Germany	STA001	1390 - 1476
Denmark	A146x3011	1397 - 1470
Denmark	G371	1419 - 1490
England	NMS002	1464 - 1518

**Table 13:** Estimated tip-dates for the Baltic-Alps clade.

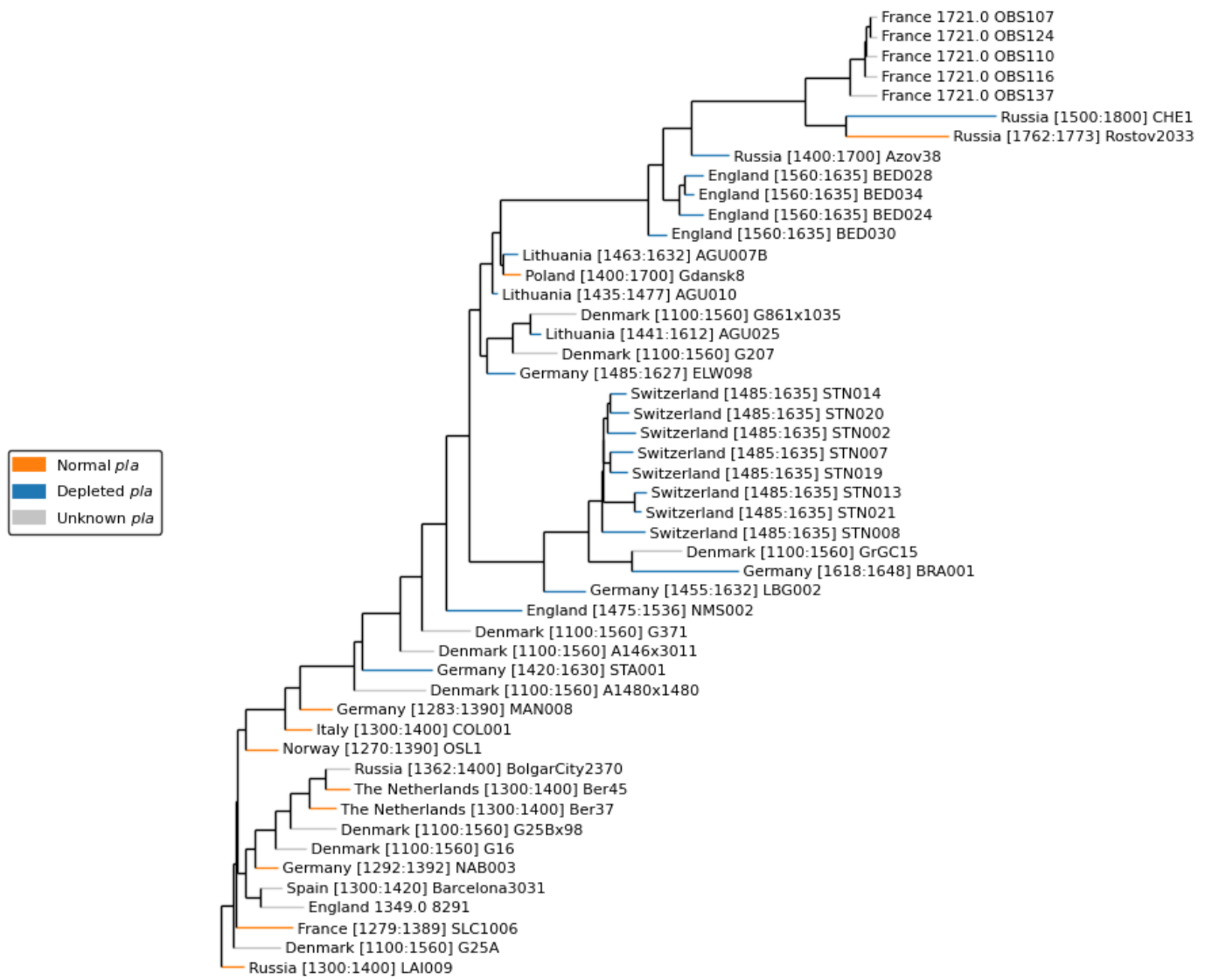
Country	Strain	95% HPD Date
Germany	LGB002	
Germany	ELW098	
Denmark	Gr GC 15	
Germany	BRA001	
Switzerland	STN008	
Switzerland	STN013	
Switzerland	STN021	
Switzerland	STN007	
Switzerland	STN019	
Switzerland	STN002	
Switzerland	STN014	
Switzerland	STN020	
Germany	ELW098	

**Table 14:** Estimated tip-dates for the England-France-Russia clade.

Country	Strain	95% HPD Date
England	BED030	
England	BED024	
England	BED028	
England	BED034	
France	OBS124	
France	OBS107	
France	OBS110	
France	OBS116	
France	OBS137	
Russia	Azov38	

Country	Strain	95% HPD Date
Russia	Rostov2033	
Russia	CHE1	

## pla Depletion



**Figure 13:** Relative Depletion of the plasminogen activator (*pla*) virulence factor on the Second Pandemic phylogeny.

## Notes

- pestis segunda* mortality at 20% [3].
- pestis segunda* was more deadly to children than the Black Death [3].
- Hotspot in Central Germany (Hesse), outbreak in 1356 not preceded by anything [3].
- pestis tertia* (1364-1376) was followed by at least eight subsequent quasi-global plague waves , in the course of the late fourteenth and the fifteenth centuries. Again points back to Central germany as an origin point.
- Timeline, spend a year or two confined in Germany, spread to the north and south Germany territories, then spread in all four directions.
- Diminishing mortality rates: 50-60% Black Death [23].

- *"After 1369, the most important feature to the second plague pandemic was not the death rate in any given epidemic, but rather, the frequency with which those epidemics occurred."* [23](#) p.131], plague entered a cycle in which it recurred every 5-12 years.