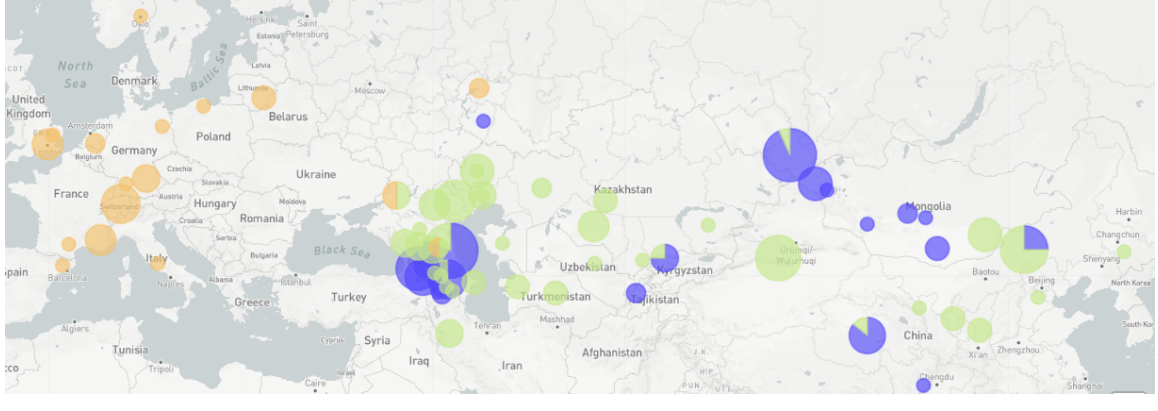
PLAGUED BY A CRYPTIC CLOCK

Insight and issues from the global phylogeny of *Yersinia pestis*

Eaton, K., Featherstone, L., Duchene, S., Carmichael, A., Varlık, N., Golding, G. B., Holmes, E., Poinar, H. N.

BEAP

2021 October 01



INTRODUCTION

CONCEPT 1. DISEASE EXPOSURE

- How do we contract an infectious disease?
- How do we spread it?
- Why does this vary across time and space?

CONCEPT 2. DISEASE EXPERIENCE

- Are the symptoms mild?
- Does it become endemic and part of our everyday life?
- Is there catastrophic mortality?
- Does it radically change society?

ZOONOSES

- Diseases that spread from non-human animals to humans.
- Suddenly appears and disappears in human populations.
- Enigmatic epidemiology that is often poorly understood.
- Difficult to explain, predict, and develop public health policies.
  - COVID-19
  - Malaria
  - Influenza
  - Brucellosis
  - Lyme Disease
  - Plague

PLAGUE

1. **Prolific presence in human populations, at least since the Bronze Age (~3000 YBP).**

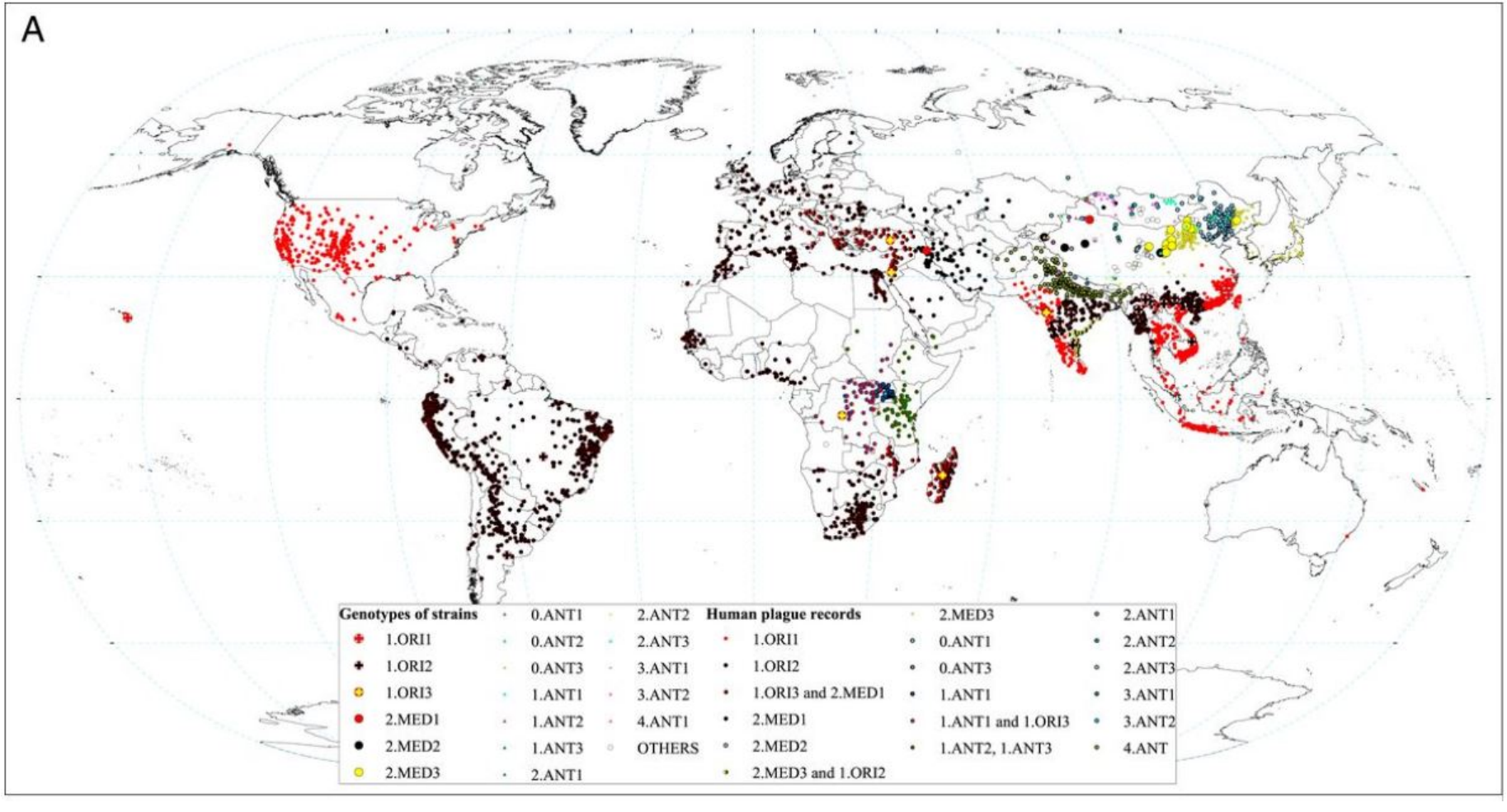


Figure 1: Geographic distribution of human plague cases (1772 - 2014).[1]

PLAGUE

2. **Unnerving virulence and mortality.**
    - The medieval Black Death (1346-1353) killed more than 50% of Europe's population.[2]
    - Untreated pneumonic plague has a case fatality rate of nearly 100%.[3,4]
3. **It keeps coming back!**
    - First Pandemic (6^{th} - 8^{th} Century CE)
    - Second Pandemic (14^{th} - 19^{th} Century CE)
    - Third Pandemic (19^{th} - 20^{th} Century CE)

PLAGUE

4. **A cryptic clock.**
    - Debated whether molecular clock techniques can produce informative rates and dates.
    - An extremely slow evolutionary rate coupled with extreme rate variation.
5. **Genetic data explosion! (2010 - 2020)**
    - 1 ancient genome → 100+
    - 10 modern genomes → 1000+

OBJECTIVES

1. Is there any way to **estimate robust rates and dates** for *Y. pestis*?
    - *When* and *Where*
2. Do I have **anything new to say** about the history and epidemiology of plague?
    - Onset and progression of past pandemics.
    - Plague reservoirs in the inter-pandemic periods.

MATERIALS AND METHODS

MATERIALS AND METHODS

1. **601 publicly available *Y. pestis* genomes.**
    - 540 modern (89.9%): 20th to 21st century CE.
    - 61 ancient (10.1%): 3rd millenium BCE to 18th century CE.
2. **Multiple alignment of the core genome.**
3. **Estimated a maximum-likelihood phylogeny.**
4. **Bayesian Evaluation of Temporal Signal (BETS).**
    - Strict Clock
    - Relaxed Clock
    - Global Clock
    - Population Clocks[5]
5. **Estimated rates and dates for populations with temporal signal.**

POPULATION STRUCTURE

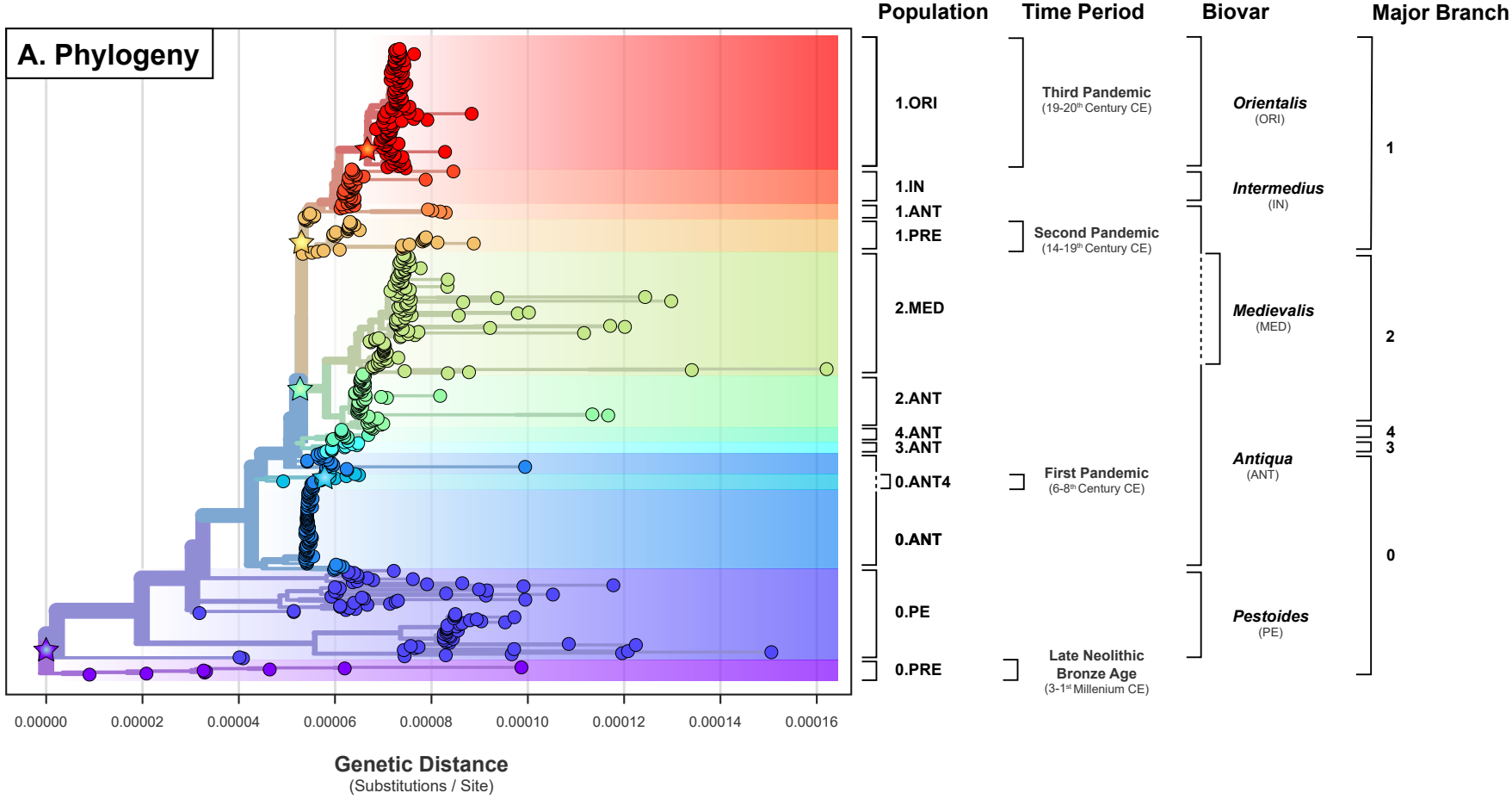


Figure 2: The maximum-likelihood phylogeny of 601 *Yersinia pestis* genomes.

MOLECULAR CLOCKS

- The global clock analysis (all samples) was VERY unstable.

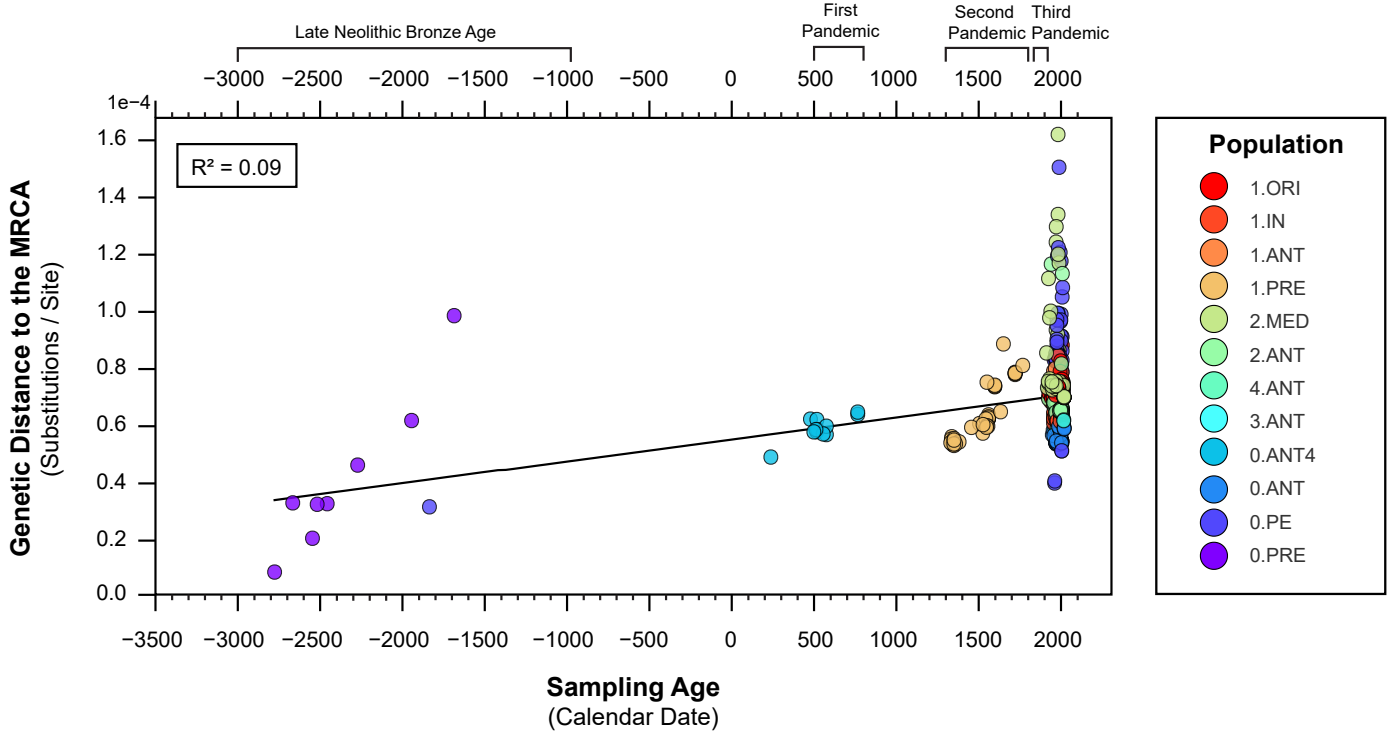


Figure 3: Root-to-tip regression of 601 *Yersinia pestis* genomes using the maximum-likelihood phylogeny.

MOLECULAR CLOCKS

- Tried reducing sources of rate variation:
  - Decreasing the number of samples.
  - Fixing tip dates.
  - Fixing tree topology.

The global diversity of *Y. pestis* is poorly modeled by a single molecular clock.

- Population clocks **recovered temporal signal** in 9 out of 12 populations.
- For all populations, the relaxed clock outerperformed the strict clock.

THE PHYLODYNAMIC THRESHOLD

"The point in time at which sufficient molecular evolutionary change has accumulated in available genome samples to obtain robust phylodynamic estimates."

"Before the phylodynamic threshold is reached, genomic variation is so low that even large amounts of genome sequences may be insufficient to estimate the virus's evolutionary rate and the time scale of an outbreak."

- Duchene et al. (2020).[6]

RATES OF EVOLUTIONARY CHANGE

- *Y. pestis* has one of the slowest substitution rates among bacterial pathogens.[7]
- *Y. pestis* has substantial rate variation.[5]

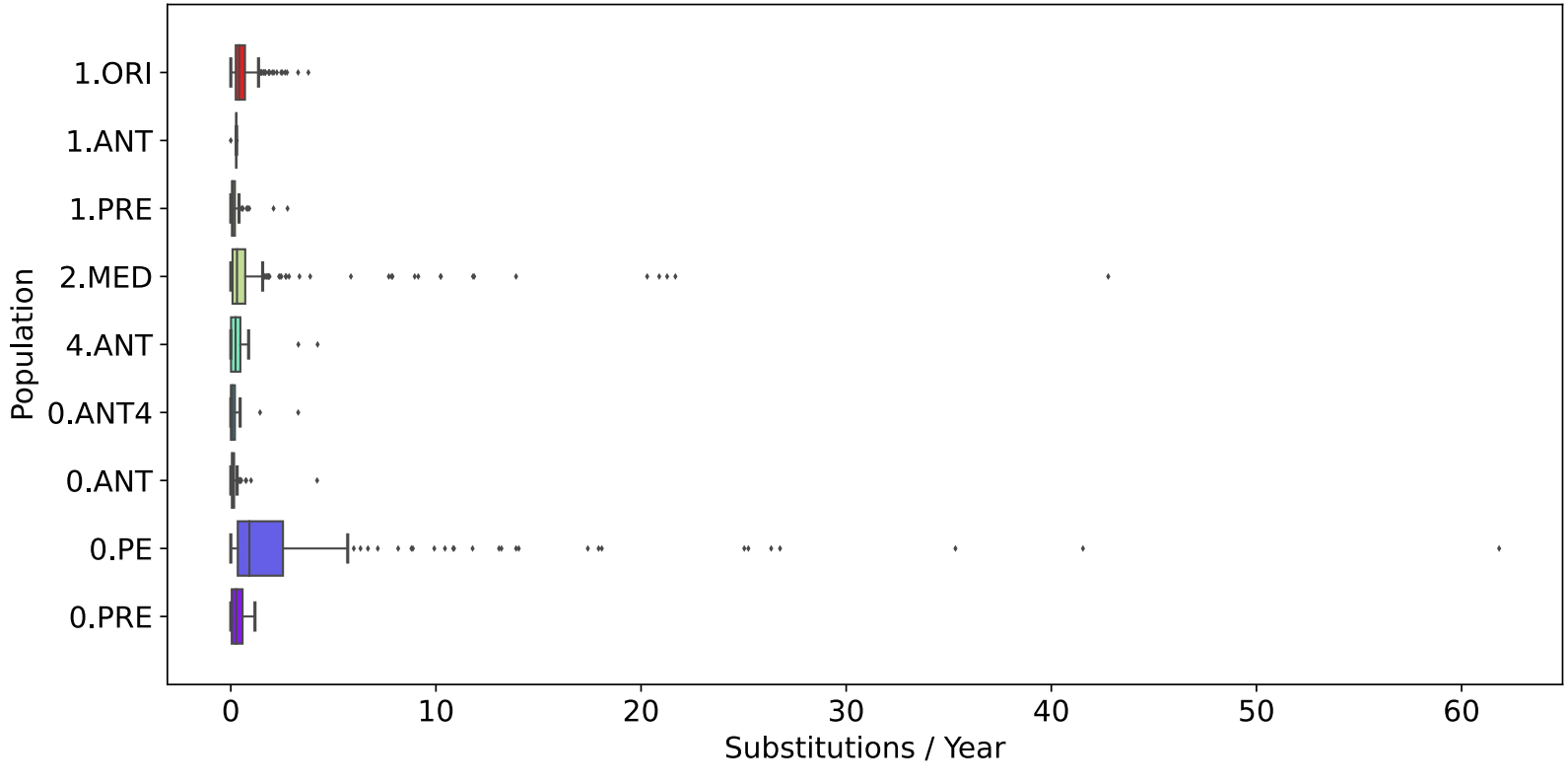


Figure 4: Substitution rates across the maximum-clade credibility (MCC) trees.

RATES OF EVOLUTIONARY CHANGE

- *Y. pestis* has one of the slowest substitution rates among bacterial pathogens.[7]
- *Y. pestis* has substantial rate variation.[5]

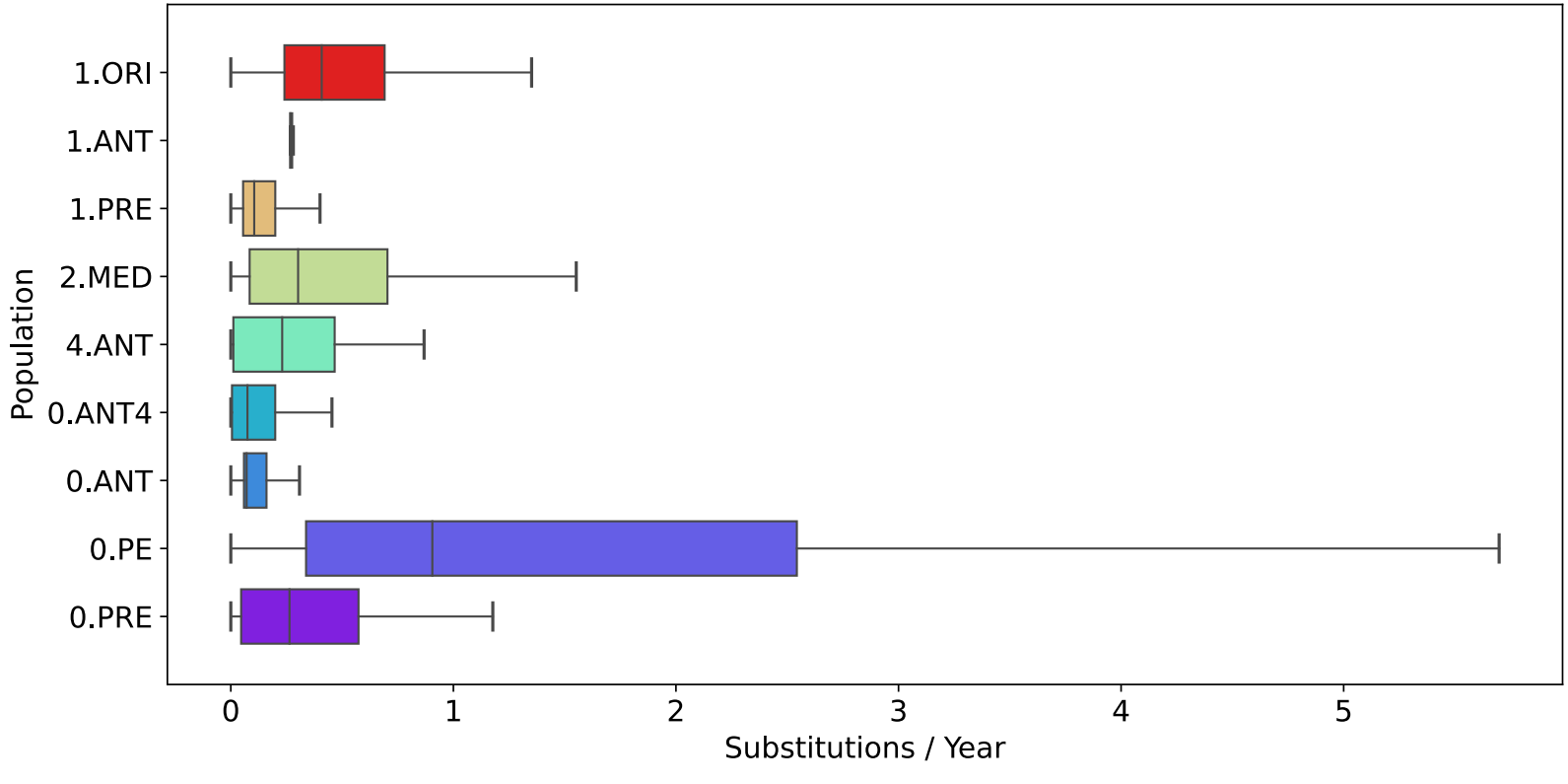


Figure 5: Substitution rates across the maximum-clade credibility (MCC) trees with outliers removed.

WHY DO WE CARE?

- Rates and dates of short-term events will be meaningless.
  - Ex. Epidemics, recently emerged populations.

But what does “short-term” mean?

NO TEMPORAL SIGNAL

- Three populations have no temporal signal (Figure 6).
- Shortest sampling time frames.
- Node density is highest close to the root.

These populations have rates and dates that are **non-informative**.
The **phylodynamic threshold** of *Y. pestis* is greater than 50 years.
Comparisons over shorter time scales will have **limited resolving power**.

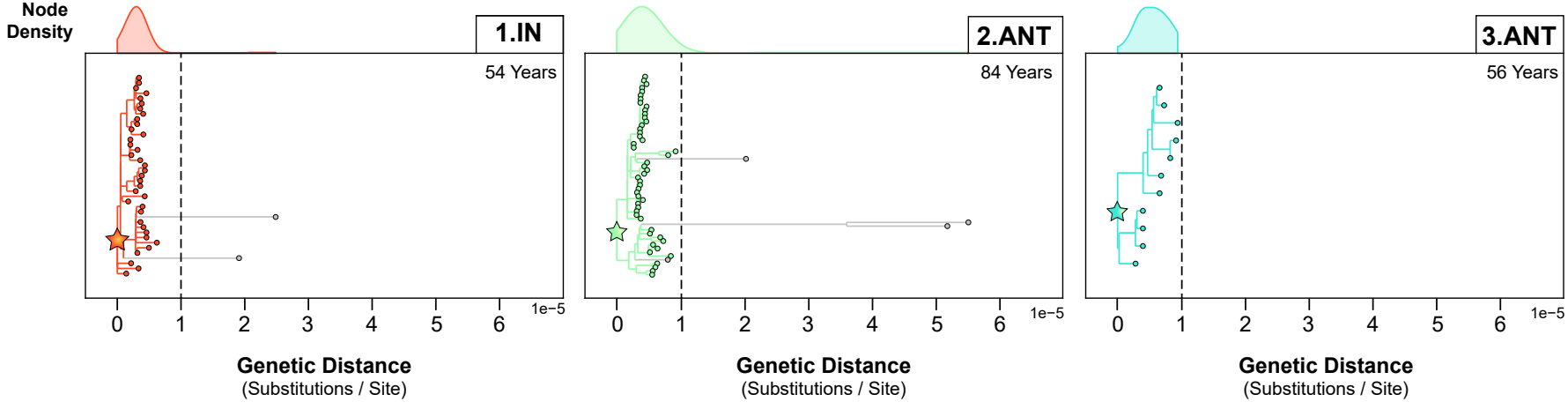


Figure 6: Maximum-likelihood phylogenies of *Yersinia pestis* populations with no detectable temporal signal. Grey branches indicate outliers.

EPIDEMIC ANALYSIS

Intra-epidemic diversity is poorly resolved by a core genome SNP phylogeny

- Black Death (1346 - 1353) *Y. pestis* isolates are indistinguishable clones.[8]

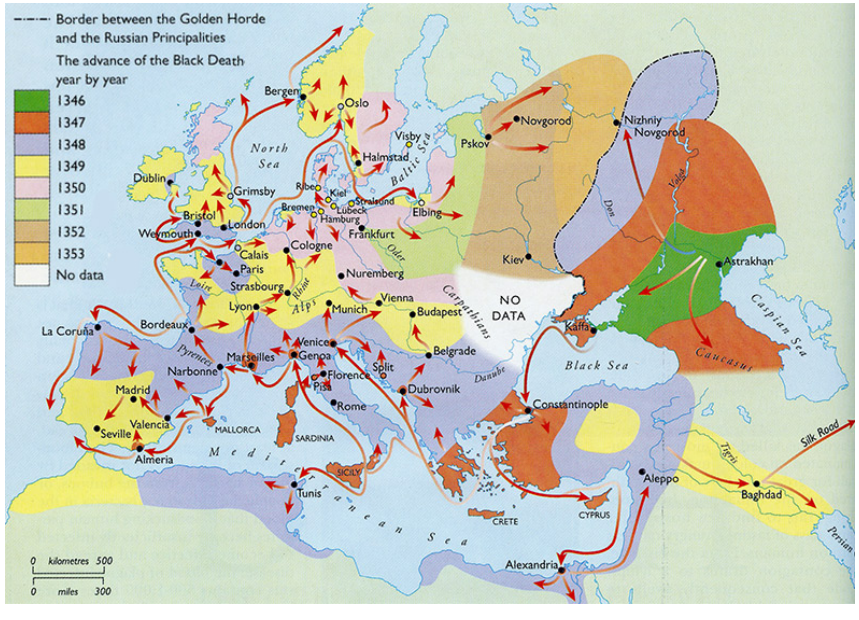


Figure 7: Spread map of the Black Death.[9]

THE THIRD PANDEMIC

- The events leading up to the Third Pandemic (1.ORI) are highly uncertain.
- When and how, did the epidemiology of plague change from localized epidemics in Yunnan Province, China to a global pandemic?

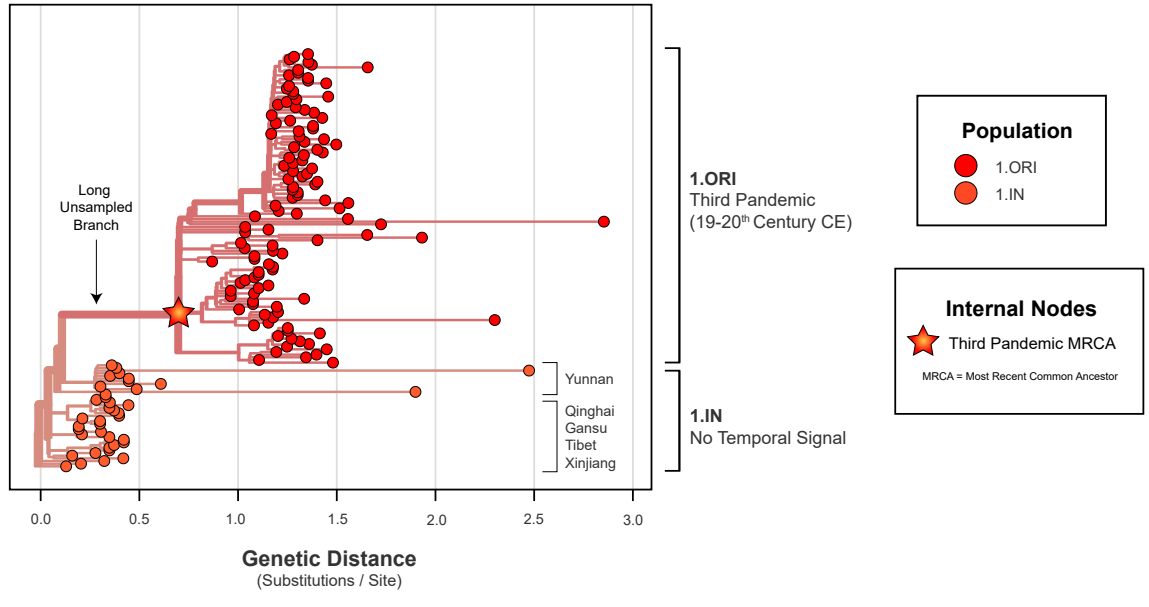


Figure 8: Third Pandemic Origins

ANCIENT DNA

TEMPORAL SIGNAL WITH ANCIENT DNA

All (3) ancient *Y. pestis* populations had temporal signal:

- 1.PRE | Second Pandemic (14^{th} - 19^{th} century CE)
- 0.ANT4 | First Pandemic (6^{th} - 8^{th} century CE)
- 0.PRE | Bronze Age (3^{rd} - 1^{st} Millennia BCE)

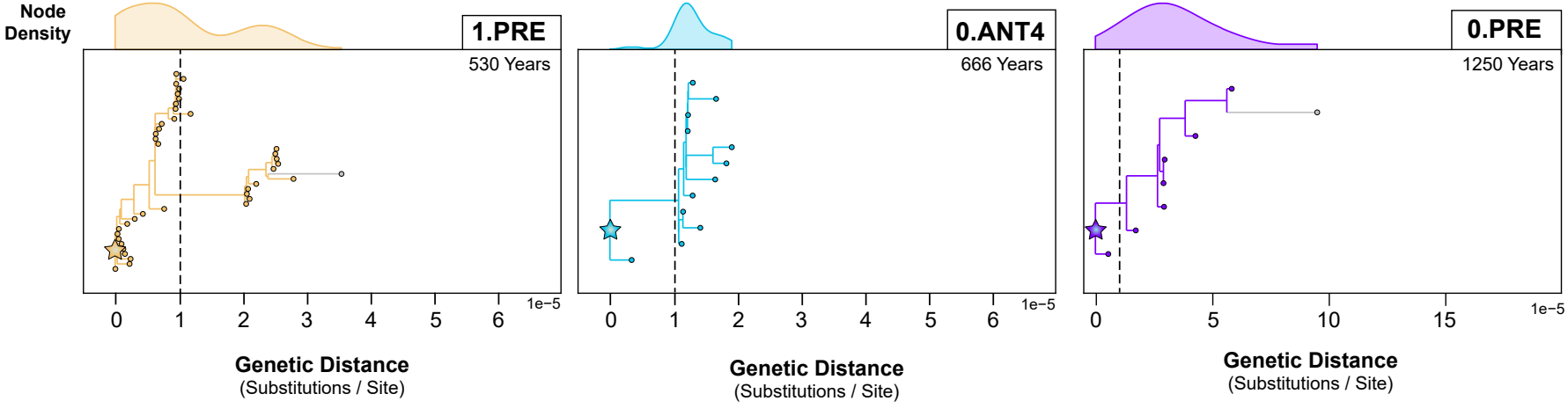


Figure 9: Maximum-likelihood phylogenies of ancient *Y. pestis* populations with detectable temporal signal. Grey branches indicate outliers.

WHAT CAN WE LEARN?

- Informative rates and dates can be obtained when **the sampling time frame is >500 years**.
- Node age **uncertainty can be 100+ years** (95% HPD).
- Genetic evidence suggests that the **plague pandemics may have earlier origins**.
- Most evidence of historical plague comes from European sources.
- New scholarship of non-European sources is rewriting these narratives.[10]

Table 1: Node dating of key historical plague events.

Population	Event	tMRCA (Lower)	tMRCA (Upper)	Historical Start
1.PRE	Second Pandemic	1214	1315	1346
0.ANT4	First Pandemic	272	466	541
0.PRE	Bronze Age	-3098	-2786	–

MODERN DNA

TEMPORAL SIGNAL WITH MODERN DNA

- Two modern populations had temporal signal, despite having no ancient DNA calibrations.
  - The Third Pandemic (1.ORI), global.
  - *Medievalis* (2.MED), Asia.
- Have high epidemiological significance:
  - High mortality epidemics.
  - Extensive and rapid geographic spread.

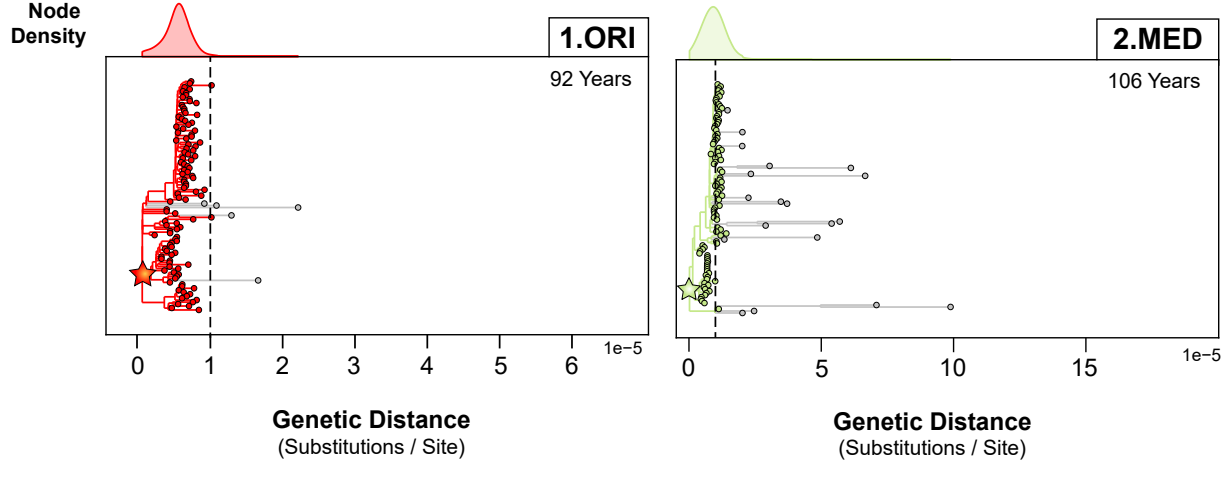


Figure 10: Maximum-likelihood phylogenies of modern *Y. pestis* populations with detectable temporal signal. Grey branches indicate outliers.

TEMPORAL SIGNAL WITH MODERN DNA

- As more data becomes available, the **node dates are becoming younger.**[12]

Table 2: Node dating of modern *Yersinia pestis* populations.

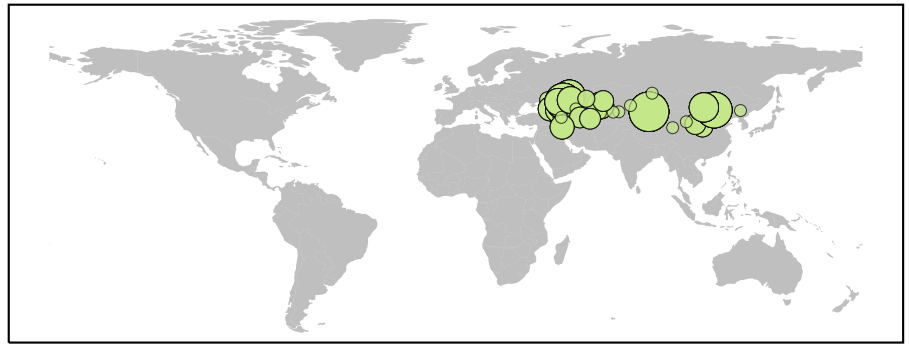
Population	Name	Study	Samples	tMRCA (Lower)	tMRCA (Upper)
1.ORI	Third Pandemic	[11]	6	-381	1738
		[5]	17	1735	1863
		[12]	93	1742	1842
		–	117	1806	1901
2.MED	*Medievalis*	[11]	2	-638	1715
		[5]	25	1450	1750
		[12]	68	1298	1582
		–	116	1560	1845

TIMELINE OF THE THIRD PANDEMIC

- We estimate that the Third Pandemic (1.ORI) **emerged between 1806 and 1901 (1885)**.
- This **aligns with historical evidence** from Chinese epidemiologists[13] [14].
- 1772 - 1800: Human plague cases are isolated to Yunnan, China.
- 1880 - 1900: Plague spreads to the south-eastern provinces.
- 1894: Plague spreads worldwide out of Hong Kong.

MEDIEVALIS

- We estimate that *Medievalis* (2.MED) **emerged between 1560 and 1845 (1796)**.
- This **aligns with historical evidence.**[15–17]
- Late 1798: Human plague cases are observed in the Caucasus.
- 1800s: Plague cases throughout the Caspian Sea region.
- 1900s-2000s: Widespread dispersal across Asia.



The geographic distribution of *Medievalis* (2.MED).

PLAGUE NEXUS

- Caucasus: modern boundary between Europe and Asia.
- Nexus of “European” and “Asian” *Y. pestis* populations.

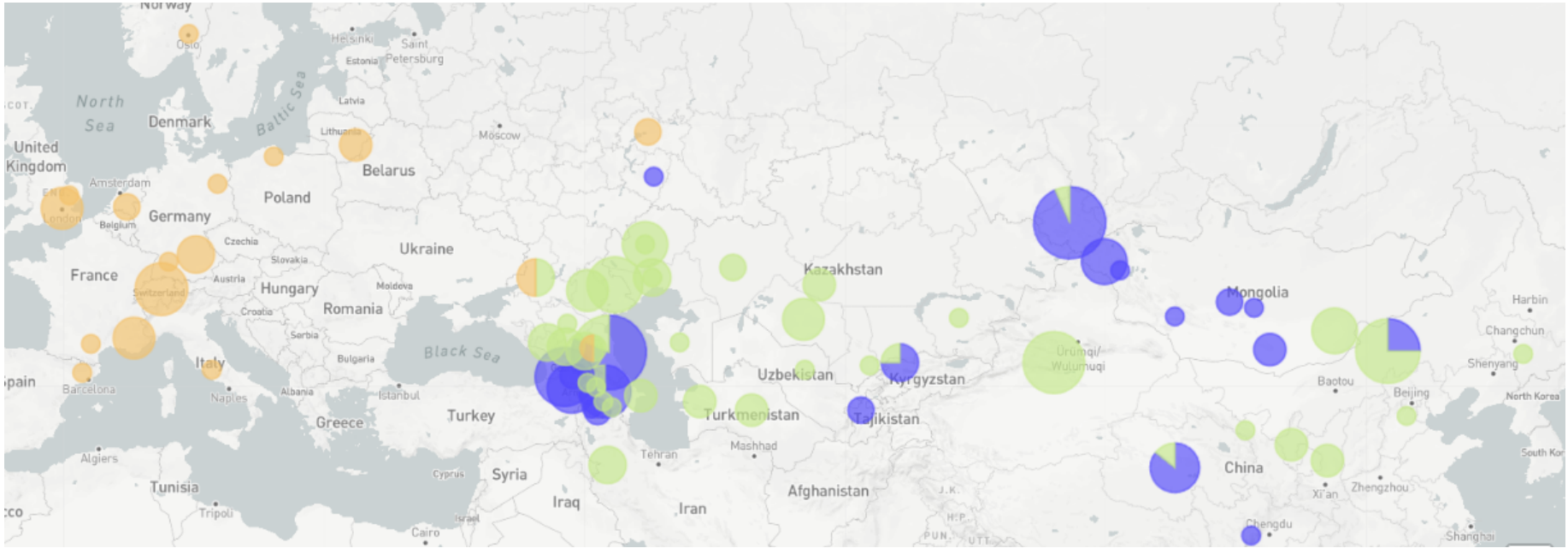


Figure 11: The geographic distribution of “European” *Y. pestis* (1.PRE) and “Asian” *Y. pestis* (0.PE, 2.MED).

GOOD TEMPORAL SIGNAL
≠
GOOD NODE DATES

NODE DATE CONFLICTS

- Several populations have conflicting dates for their tMRCA.
- These populations are not monophyletic.
- Ex. *Antiqua* (0.ANT) is ancestral to the First Pandemic (0.ANT4).
- The tMRCA of *Antiqua* (0.ANT) incorrectly post-dates the First Pandemic (0.ANT4).

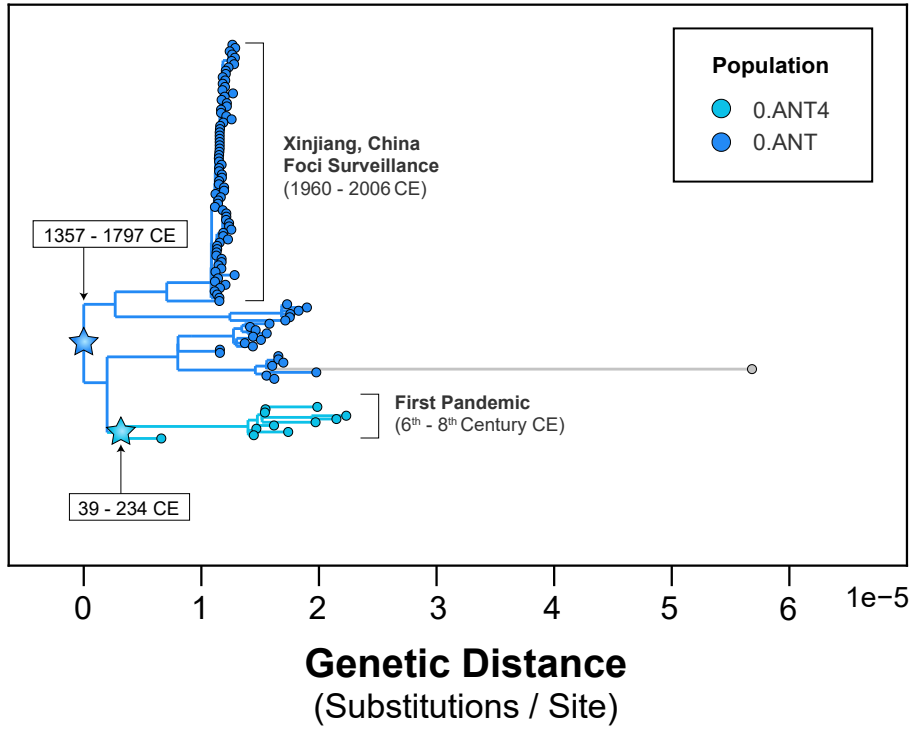


Figure 12: Node dating conflicts between the ancestral *Antiqua* (0.ANT) and the First Pandemic (0.ANT4).

NODE DATE CONFLICTS

- Node density is low close to the root.
- Node dates are underestimated (too young). Can be off by a 1000 years!

Insufficient sample sizes (1.ANT).
Insufficient geographic sampling (0.ANT).
Extensive rate variation (0.PE) with no ancient DNA calibrations.

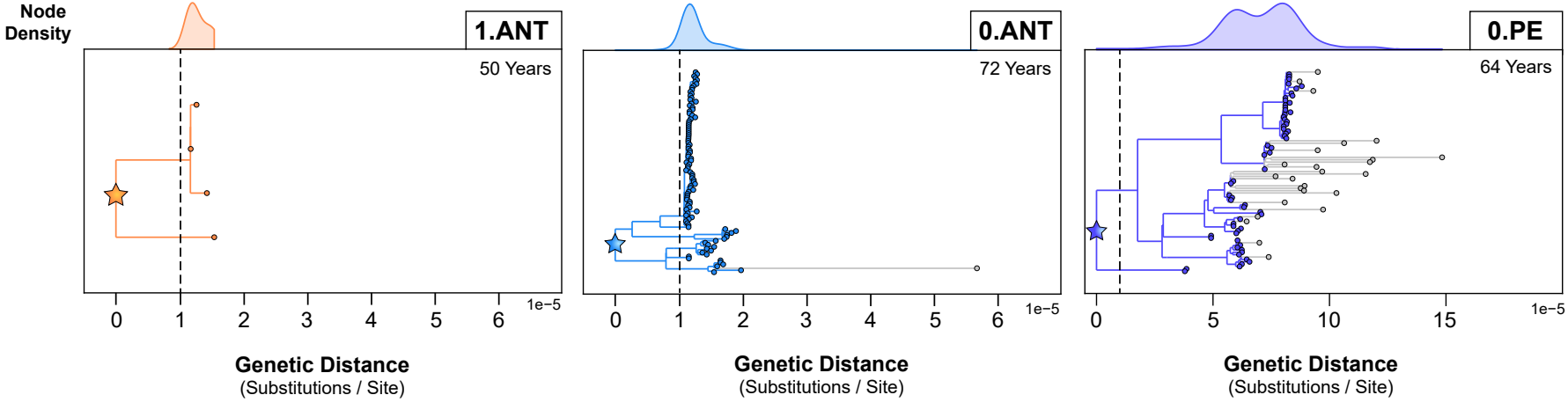


Figure 13: Maximum-likelihood phylogenies of modern *Y. pestis* populations lacking internal calibrations. Grey branches indicate outliers.

ARRIVAL OF PLAGUE IN AFRICA

- Africa accounts for more than 90% of all human plague cases[18] .
- But only 1.5% of *Y. pestis* genomes are from Africa.
- Heavily debated when plague first arrived in Central and East Africa (1.ANT).

Population	Name	Study	Samples	tMRCA (Lower)	tMRCA (Upper)
1.ANT	African *Antiqua*	[11]	2	-4964	1322
		[5]	2	1377	1650
		–	4	1655	1835

Genetic dates for the tMRCA of 1.ANT in Africa are non-informative!
They should not be used as a basis for historical or archaeological interpretation.[19]

CONCLUSIONS

METHODOLOGY

Poor performance of a global molecular clock.
- Intensive global sampling → More genetic diversity → More rate variation.
- Partitioning the data by population stabilizes clock estimates.

The phylodynamic threshold is very important!
- A short sampling time frame (<50 years) led to non-informative rates and dates.
- Longer sampling time frames (>90 years) led to detectable temporal signal.

Temporal signal does not mean *all* dates will be informative.
- Insufficient sampling produces conflicting node dates.
- Can be detected in trees with few nodes close to the root.

The phylogeographic threshold!
- If we can estimate “when” plague was present in the past, what about “where?”

PLAGUE HISTORY

The tMRCAs of **ancient plague pandemics are older** than previously thought.
- Insight into disease events that pre-date textual evidence.
- Aligns with recent historical work focused on non-European plague history.
- Challenges Eurocentric views of plague history.

The tMRCAs of **modern plague are younger** than previously thought.
- Aligns with historical and epidemiological evidence.
- New insight into “nexus” regions, where multiple *Y. pestis* populations co-exist.

REFERENCES

1. Xu, L., Stige, L. C., Leirs, H., Neerinckx, S., Gage, K. L., Yang, R., Liu, Q., Bramanti, B., Dean, K. R., Tang, H., Sun, Z., Stenseth, N. C., & Zhang, Z. (2019). Historical and genomic data reveal the influencing factors on global transmission velocity of plague during the Third Pandemic. *Proceedings of the National Academy of Sciences*, *116*(24), 11833–11838. https://doi.org/10.1073/pnas.1901366116
2. Benedictow, O. J. (2021). *The Complete History of the Black Death*. Boydell Press.
3. Prentice, M. B., & Rahalison, L. (2007). Plague. *The Lancet*, *369*(9568), 1196–1207. https://doi.org/10.1016/S0140-6736(07)60566-2
4. Salam, A. P., Rojek, A., Cai, E., Raberahona, M., & Horby, P. (2020). Deaths Associated with Pneumonic Plague, 1946 - Volume 26, Number 10 2020 - Emerging Infectious Diseases journal - CDC. *Emerging Infectious Diseases*, *26*(10), 2432–2434. https://doi.org/10.3201/eid2610.191270
5. Cui, Y., Yu, C., Yan, Y., Li, D., Li, Y., Jombart, T., Weinert, L. A., Wang, Z., Guo, Z., Xu, L., Zhang, Y., Zheng, H., Qin, N., Xiao, X., Wu, M., Wang, X., Zhou, D., Qi, Z., Du, Z., … Yang, R. (2013). Historical variations in mutation rate in an epidemic pathogen, *Yersinia Pestis*. *Proceedings of the National Academy of Sciences*, *110*(2), 577–582. https://doi.org/10.1073/pnas.1205750110
6. Duchene, S., Featherstone, L., Haritopoulou-Sinanidou, M., Rambaut, A., Lemey, P., & Baele, G. (2020). Temporal signal and the phylodynamic threshold of SARS-CoV-2. *Virus Evolution*, *6*(2). https://doi.org/10.1093/ve/veaa061
7. Duchêne, S., Holt, K. E., Weill, F.-X., Le Hello, S., Hawkey, J., Edwards, D. J., Fourment, M., & Holmes, E. C. (2016). Genome-scale rates of evolutionary change in bacteria. *Microbial Genomics*, *2*(11). https://doi.org/10.1099/mgen.0.000094
8. Spyrou, M. A., Tukhbatova, R. I., Feldman, M., Drath, J., Kacki, S., Beltrán de Heredia, J., Arnold, S., Sitdikov, A. G., Castex, D., Wahl, J., Gazimzyanov, I. R., Nurgaliev, D. K., Herbig, A., Bos, K. I., & Krause, J. (2016). Historical *Y. Pestis* Genomes Reveal the European Black Death as the Source of Ancient and Modern Plague Pandemics. *Cell Host & Microbe*, *19*(6), 874–881. https://doi.org/10.1016/j.chom.2016.05.012
9. Benedictow, O. J. (2004). *The Black Death, 1346-1353: The Complete History*. Boydell Press.
10. Varlık, N. (2020). The plague that never left: Restoring the Second Pandemic to Ottoman and Turkish history in the time of COVID-19. *New Perspectives on Turkey*, *63*, 176–189. https://doi.org/10.1017/npt.2020.27
11. Morelli, G., Song, Y., Mazzoni, C. J., Eppinger, M., Roumagnac, P., Wagner, D. M., Feldkamp, M., Kusecek, B., Vogler, A. J., Li, Y., Cui, Y., Thomson, N. R., Jombart, T., Leblois, R., Lichtner, P., Rahalison, L., Petersen, J. M., Balloux, F., Keim, P., … Achtman, M. (2010). *Yersinia Pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nature Genetics*, *42*(12), 1140–1143. https://doi.org/10.1038/ng.705
12. Pisarenko, S. V., Evchenko, A. Yu., Kovalev, D. A., Evchenko, Y., Bobrysheva, O. V., Shapakov, N. A., Volynkina, A. S., & Kulichenko, A. N. (2021). Yersinia pestis strains isolated in natural plague foci of Caucasus and Transcaucasia in the context of the global evolution of species. *Genomics*, *113*(4), 1952–1961. https://doi.org/10.1016/j.ygeno.2021.04.021
13. Xu, L., Stige, L. Chr., Kausrud, K. L., Ben Ari, T., Wang, S., Fang, X., Schmid, B. V., Liu, Q., Stenseth, N. Chr., & Zhang, Z. (2014). Wet climate and transportation routes accelerate spread of human plague. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1780), 20133159. https://doi.org/10.1098/rspb.2013.3159
14. Tan, J., Liu, Y., Shen, E., Zhu, W., Wang, W., Li, R., & Yang, L. (2002). Towards the atlas of plague and its environment in the People's Republic of China: idea, principle and methodology of design and research results. *Huan Jing Ke Xue= Huanjing Kexue*, *23*(3), 1–8.

15. Kahan, A. (1979). Social Aspects of the Plague Epidemics in Eighteenth-Century Russia. *Economic Development and Cultural Change*, *27*(2), 255–266.
16. Hashemi Shahraki, A., Carniel, E., & Mostafavi, E. (2016). Plague in Iran: Its history and current status. *Epidemiology and Health*, *38*, e2016033. https://doi.org/10.4178/epih.e2016033
17. Eroshenko, G. A., Popov, N. V., Alkhova, Zh. V., Balykova, A. N., Kukleva, L. M., Chervyakova, N. S., Maykanov, N. S., Sarmuldina, A. Kh., & Kutyrev, V. V. (2019). Circulation of *Yersinia Pestis* in the Volga-Ural sandy focus: Spatiotemporal analysis. *Problems of Particularly Dangerous Infections*, *3*, 51–57. https://doi.org/10.21055/0370-1069-2019-3-51-57
18. Munyenyiwa, A., Zimba, M., Nhiwatiwa, T., & Barson, M. (2019). Plague in Zimbabwe from 1974 to 2018: A review article. *PLOS Neglected Tropical Diseases*, *13*(11), e0007761. https://doi.org/10.1371/journal.pntd.0007761
19. Green, M. H. (2018). Putting Africa on the Black Death map: Narratives from genetics and history. *Afriques. Débats, Méthodes Et Terrains d'histoire*, *09*. https://doi.org/10.4000/afriques.2125

SUPPLEMENTARY SLIDES

BRANCH LENGTH DISTRIBUTION

PHYLOGEOGRAPHY

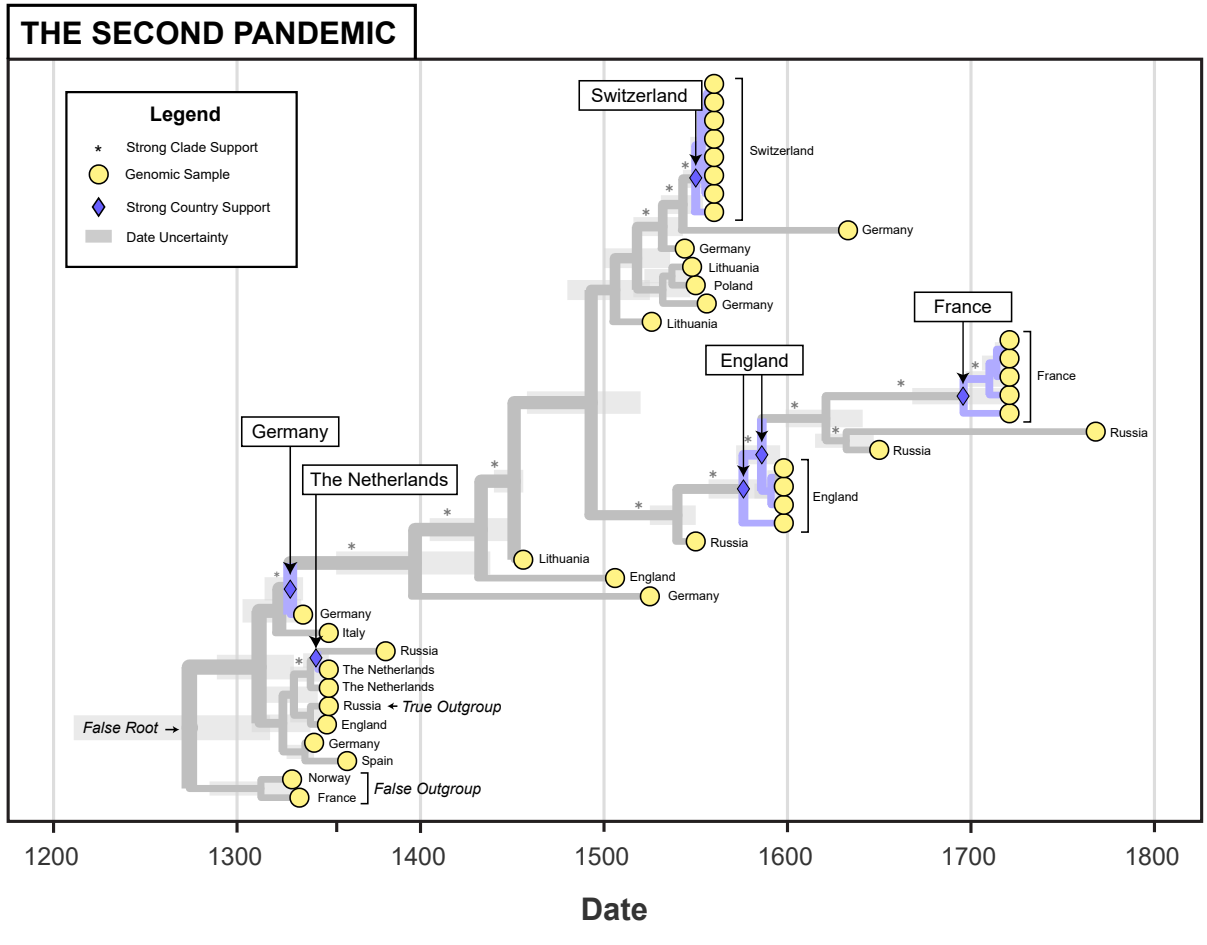


Figure 14: Phylogeography of the Second Plague Pandemic (1.PRE).

PHYLOGEOGRAPHY

Inter-country geogrpahy….

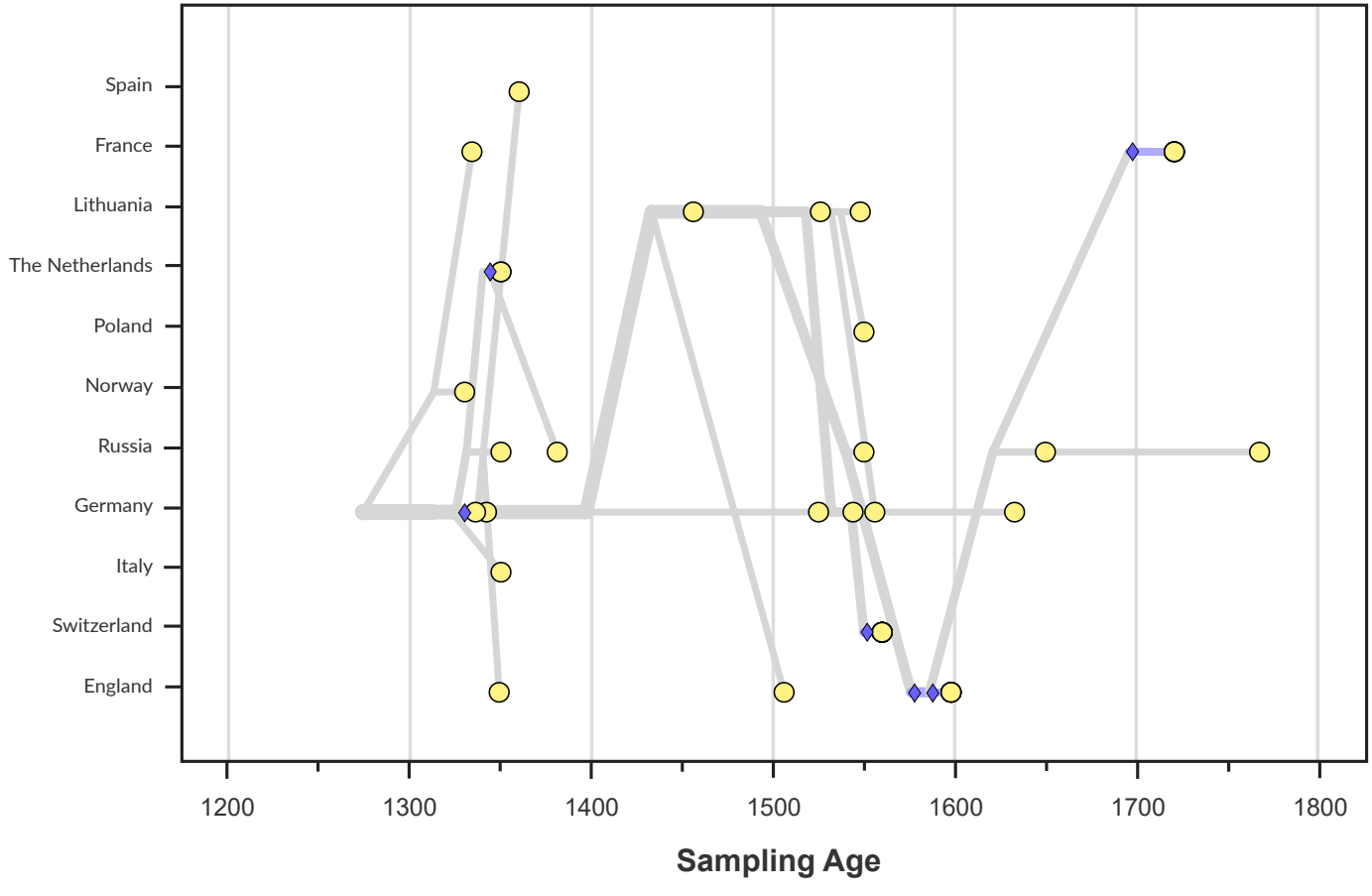


Figure 15: Phylogeography of the Second Plague Pandemic (1.PRE). Blue diamonds indicate nodes where the ancestral location was inferred to be in a single country with a probability greater than 0.95 .

NOTES

- Confidence Interval Distribution

// remote reveal.js plugins // local reveal.js plugins