This manuscript was automatically generated on August 27, 2021.

Authors

Yersinia pestis Phylodynamics

The rate variation observed in *Yersinia pestis* (Figure 1) presents a curious case of the time dependency of molecular rates [1]. Rate variation correlates with the sampling time frame, in which populations sampled over several decades have higher rate variation than those sampled over centuries, millennia, etc. This is primarily due to two factors, an exceptionally slow substitution rate in the long-term combined with a high, short-term mutation rate.

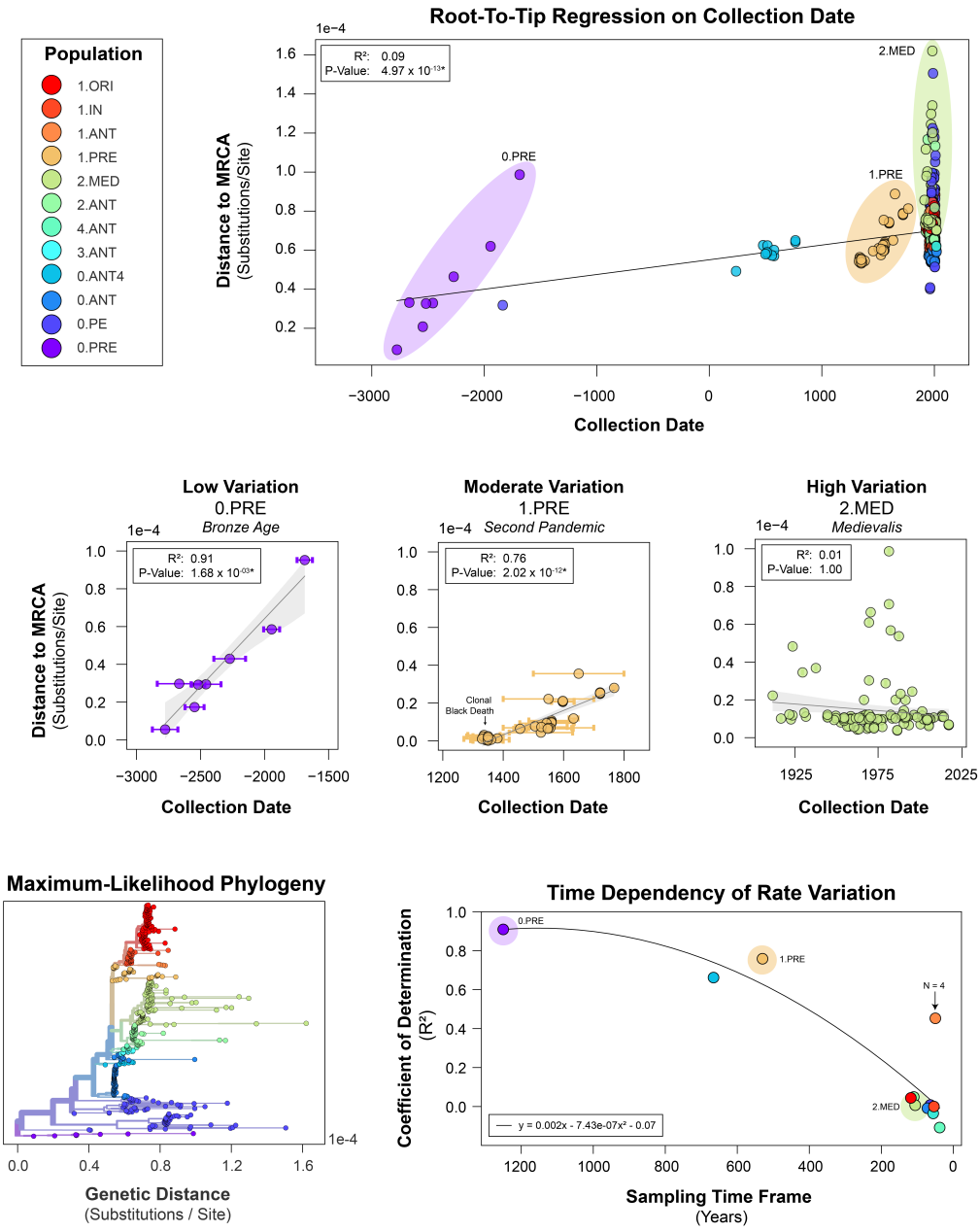

Figure 1: Rate variation in *Yersinia pestis* using root-to-tip regression.

Substitution Rate

The substitution rate of *Y. pestis* has previously been estimated to range from 1×10^{-8} to 2×10^{-8} subs/site/year, [2,3] or 1 substitution every 10-25 years. Amongst bacterial pathogens, this is one of the slowest rates observed [4] and means that *Y. pestis* lineages often cannot be differentiated until several decades have passed. This question of how much time must pass before sufficient molecular change occurs is referred to as the phylodynamic threshold [5].

In application, we can see this in the finding that *Y. pestis* isolates dated to the medieval Black Death (1348-1353) are indistinguishable clones, whereas those from subsequent centuries are phylogenetically distinct (Figure 1 “Moderate Variation”). This highlights a significant limitation of *Y. pestis* phylogenetics, as comparisons over short time scale (<10 years) have limited resolution and can be easily biased by noisy mutations.

Mutation Rate

Since it can take decades for a substitution to become fixed in *Y. pestis* populations, rate estimates are highly susceptible to the influence of transient mutations. In whole-genome sequencing, it is common to capture both fixed substitutions in the population and transient mutations found in a single isolate. These transient mutations may arise from “true” biological variation in a wild isolate, or from methodological “artifacts” due to errors in sequencing and genome assembly.

The global phylogeny of *Y. pestis* is heavily impacted by these transient mutations, which manifest as long external branches (Figure 1). These branches constitute 21% of the entire phylogeny (124 / 601 genomes), with the most strongly affected populations being *medievalis* (2.MED), *pestoides* (0.PE), and *orientalis* (1.ORI). Fortunately, samples associated with these long branches have a distinct genomic signature and can be consistently identified based on the ratio of transitions to transversions (TsTv) (Figure 2). We hypothesize that these skewed ratios may derive from sequencing/assembly error (low TsTv) and laboratory adaptation (high TsTv). As we cannot be confident that these outlier samples do not reflect analytical artifacts, we next investigated the impact of their removal.

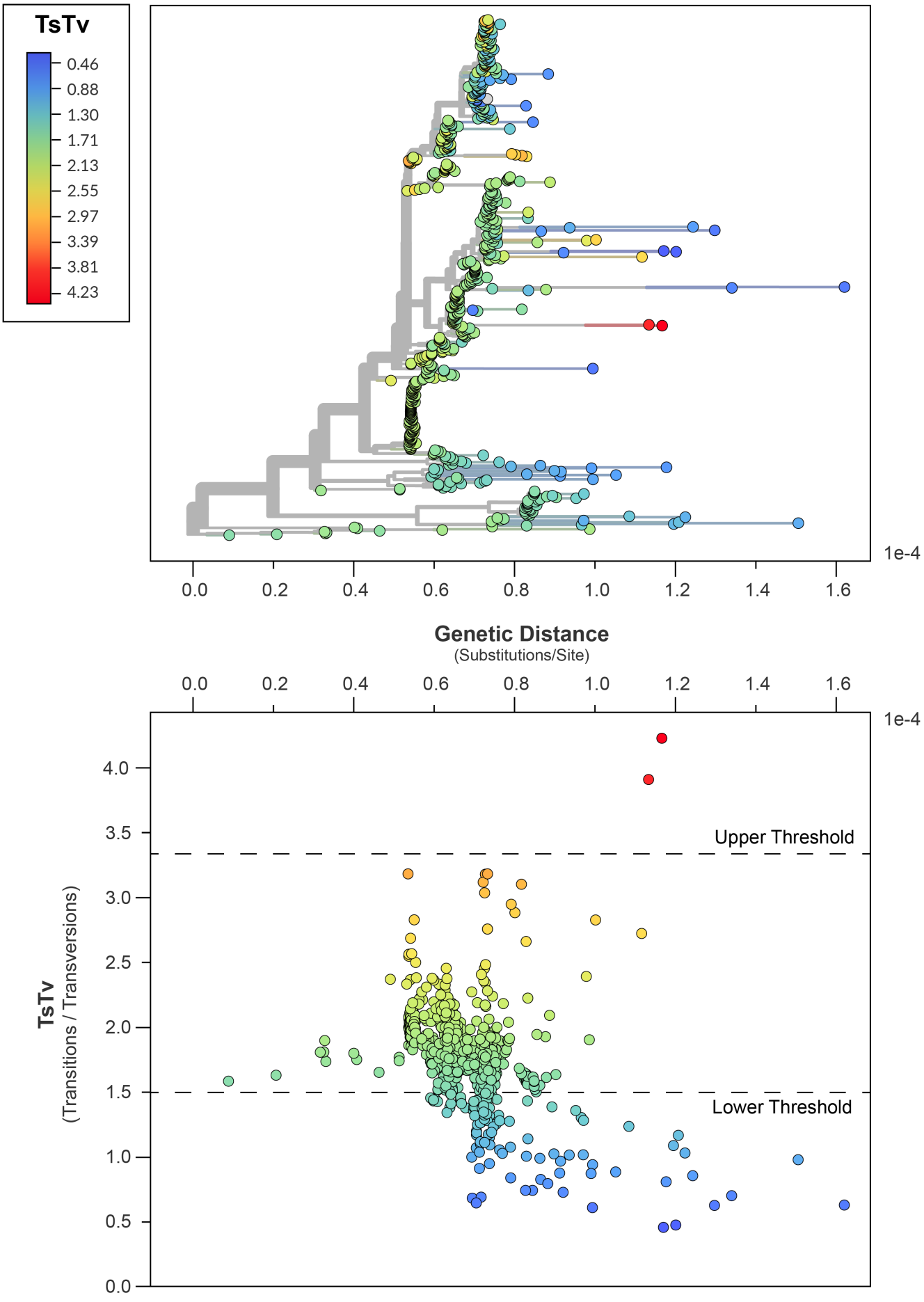
TsTv
0.46
0.88
1.30
1.71
2.13
2.55
2.97
3.39
3.81
4.23
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1e-4
Genetic Distance
(Substitutions/Site)
TsTv
(Transitions / Transversions)
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Upper Threshold
Lower Threshold

Figure 2: Long external branches in the global *Y. pestis* phylogeny are associated with extreme values of the transition/transversion ratios (TsTv).

Filtering the Phylogeny

The removal of samples associated with long, external branches has profound effects on the phylogenetic analyses and subsequent interpretations.

Phylogeography

One measure of signal

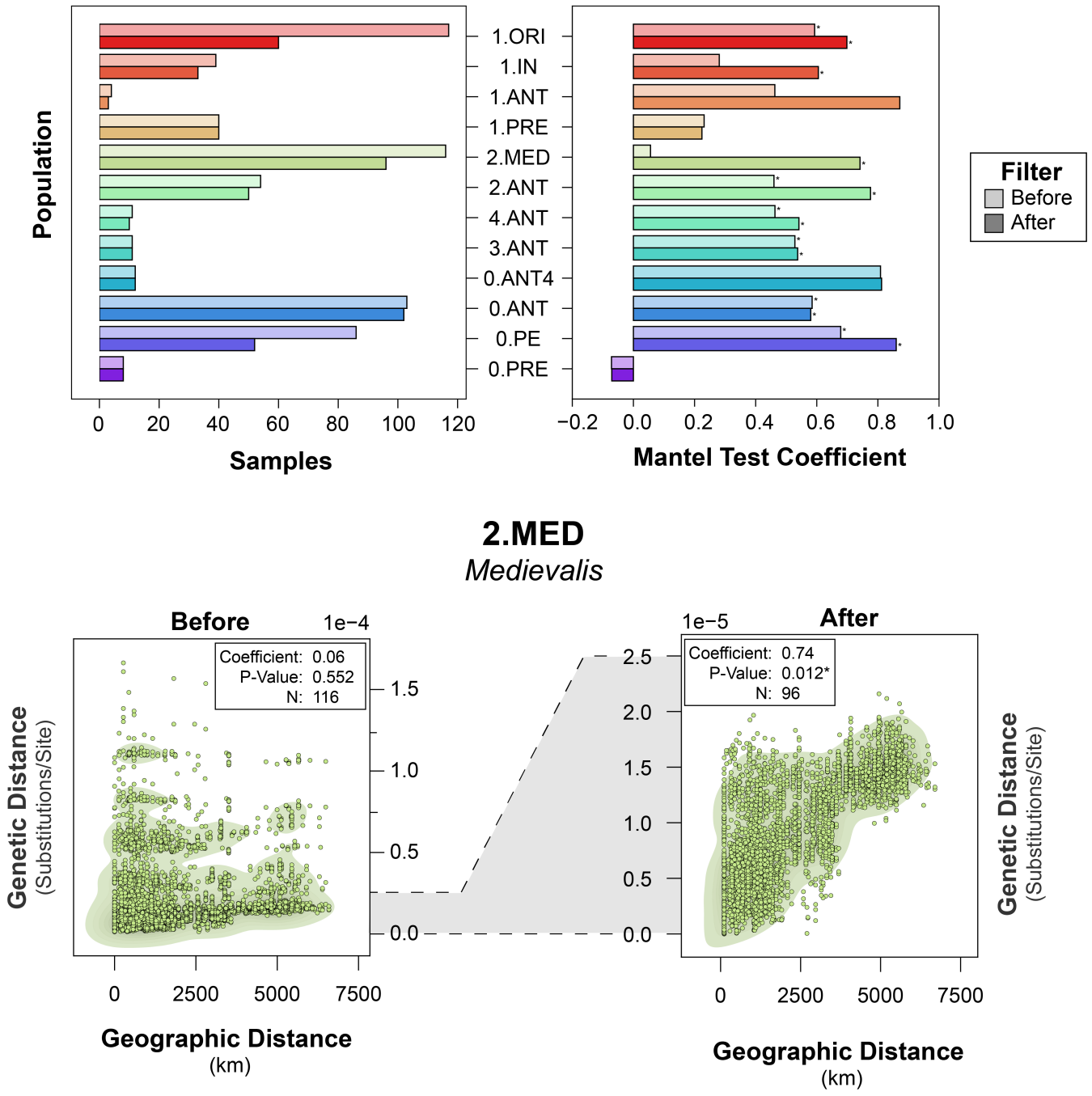

Figure 3: Isolation by distance

Phylodynamics

Stablizing the chain.

1. **Time Dependency of Molecular Rate Estimates and Systematic Overestimation of Recent Divergence Times**
   Simon Y. W. Ho, Matthew J. Phillips, Alan Cooper, Alexei J. Drummond
   Molecular Biology and Evolution (2005-07-01) https://doi.org/10.1093/molbev/msi145
   DOI: 10.1093/molbev/msi145

2. **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

3. **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

4. **Genome-scale rates of evolutionary change in bacteria**
   Sebastian Duchêne, Kathryn E. Holt, François-Xavier Weill, Simon Le Hello, Jane Hawkey, David J. Edwards, Mathieu Fourment, Edward C. Holmes
   Microbial Genomics (2016-11-30) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5320706/
   DOI: 10.1099/mgen.0.000094 · PMID: 28348834 · PMCID: PMC5320706

5. **Temporal signal and the phylodynamic threshold of SARS-CoV-2**
   Sebastian Duchene, Leo Featherstone, Melina Haritopoulou-Sinanidou, Andrew Rambaut, Philippe Lemey, Guy Baele
   Virus Evolution (2020-07-01) https://doi.org/10.1093/ve/veaa061
   DOI: 10.1093/ve/veaa061