

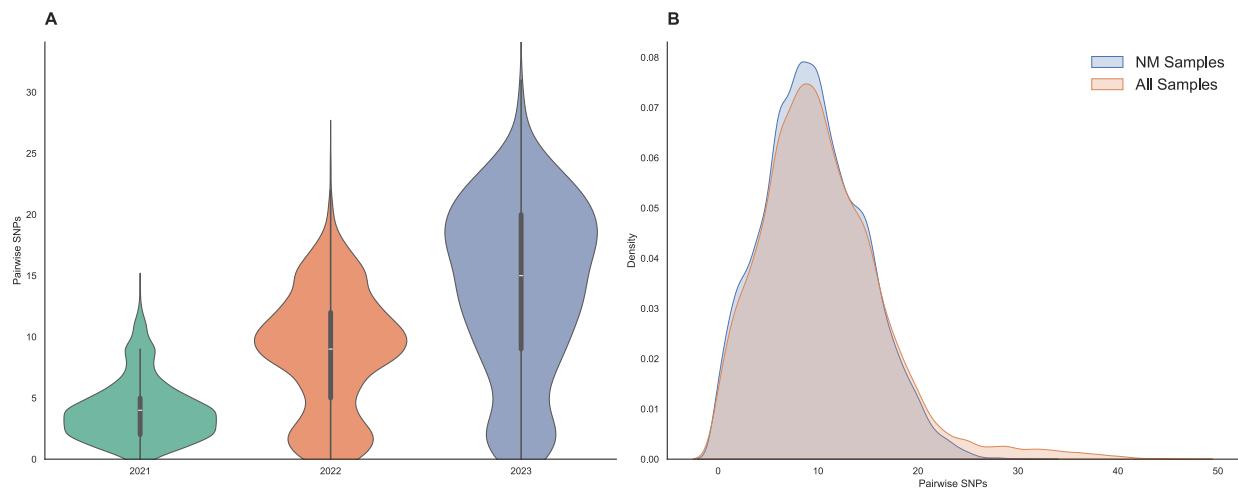
Supplementary Material

Multidrug-resistant *Shigella flexneri* outbreak associated with a high-mortality spillover event into non-human primates

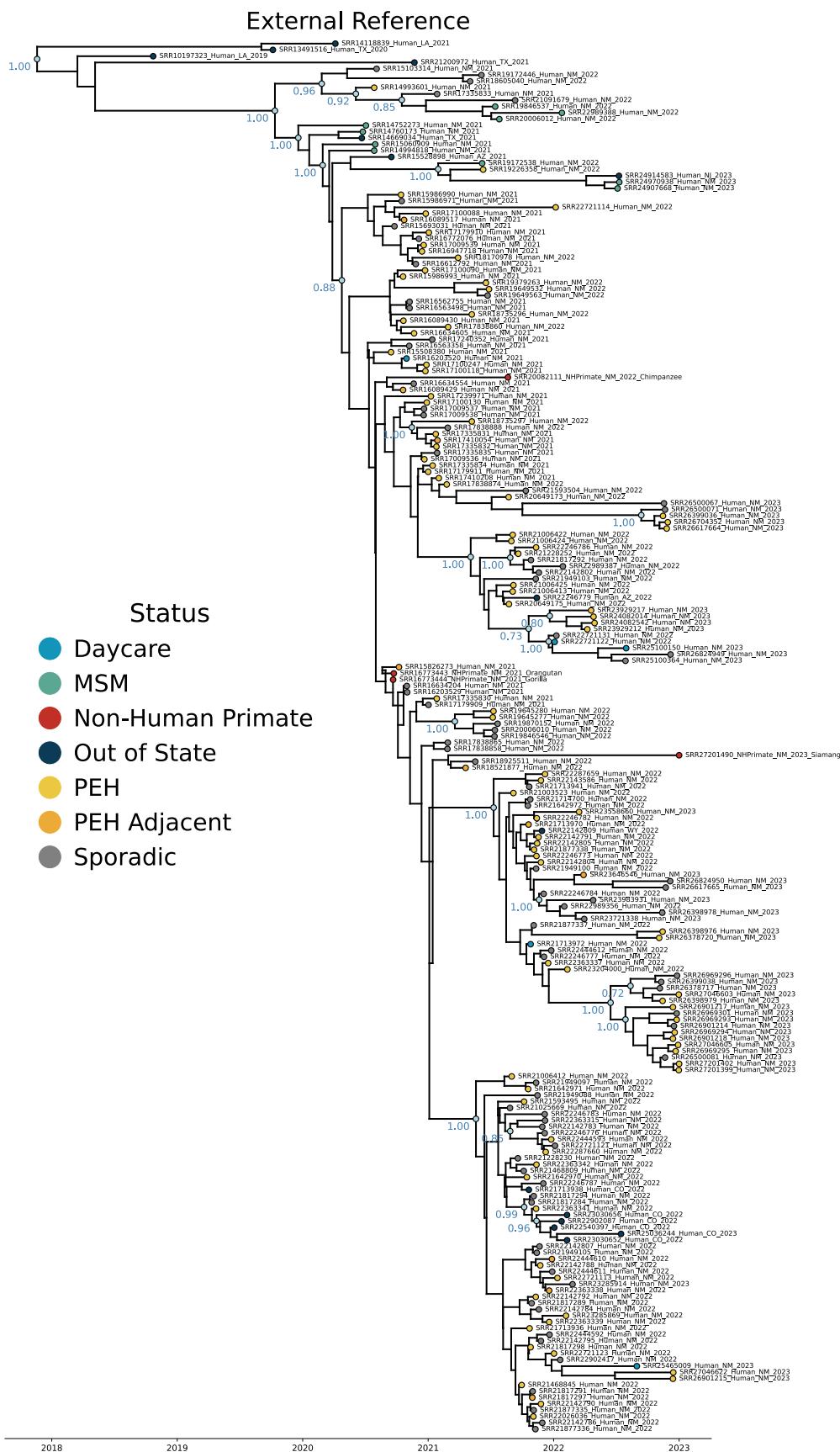
Shrum Davis *et al*

Table of Contents

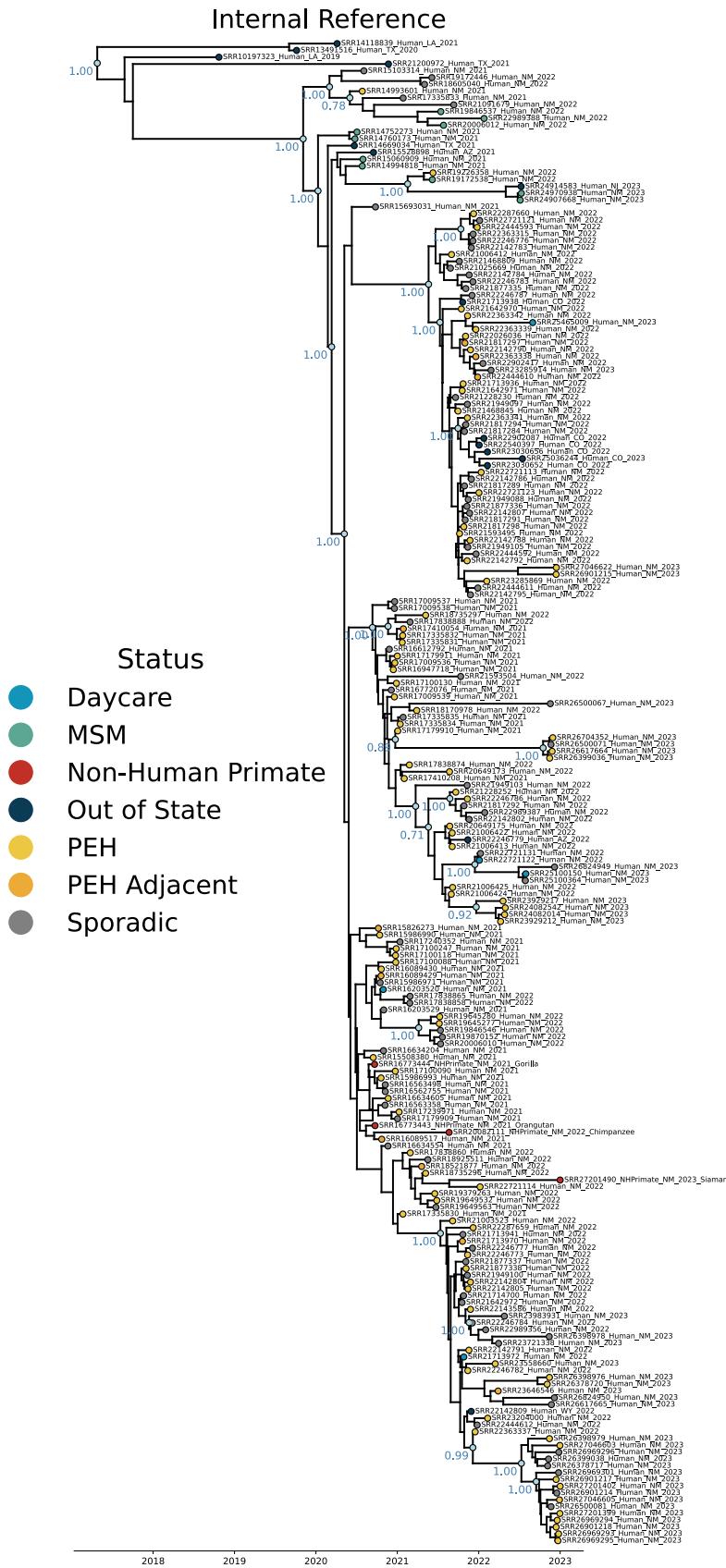
Supplementary Figure 1. Density distribution plot of pairwise SNP comparisons.	2
Supplementary Figure 2. BEAST phylogeny using the alignment against the external reference	3
Supplementary Figure 3. BEAST phylogeny using the alignment against the internal reference	4
Supplementary Figure 4. Tree topology comparisons of BEAST phylogenies.	5
Supplementary Figure 5. Maximum likelihood phylogeny of US contextual strains	6
Supplementary Figure 6. Maximum likelihood phylogeny of 580 contextual strains.....	7
Supplementary Figure 7. Maximum likelihood phylogeny of 28 Nanopore sequenced isolates	8



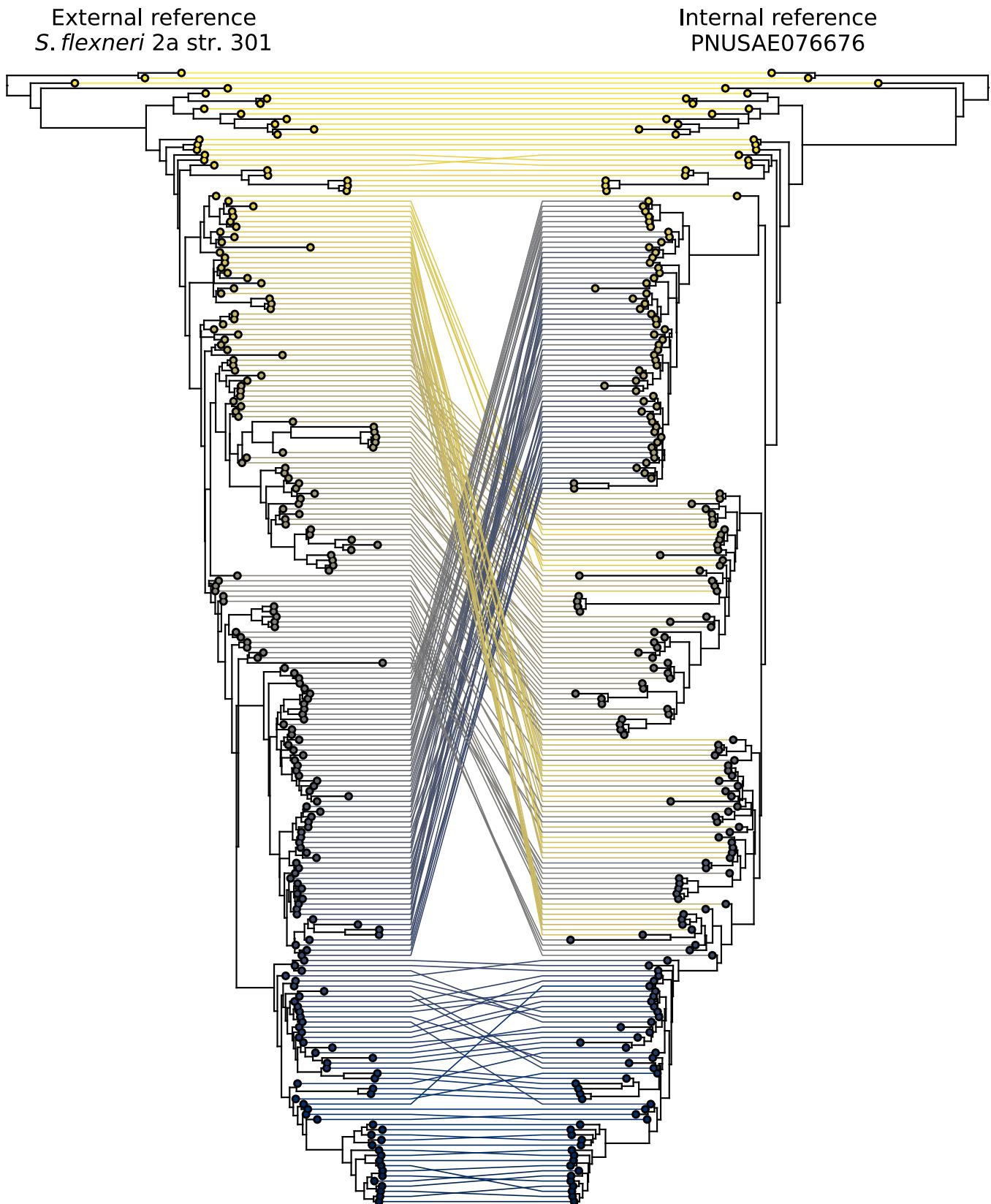
Supplementary Figure 1. Pairwise SNP comparisons from alignments against the reference genome *S. flexneri* 2a str. 301. **A.** Pairwise SNP distributions by year for sample from New Mexico. **B.** Density distribution plot of pairwise SNP comparisons. The NM outbreak samples consist only of pairwise comparisons between samples collected within the state. The Out of state samples are all comparisons where one or both samples were collected outside of New Mexico and includes the pre-cursor strains to the NM outbreak.



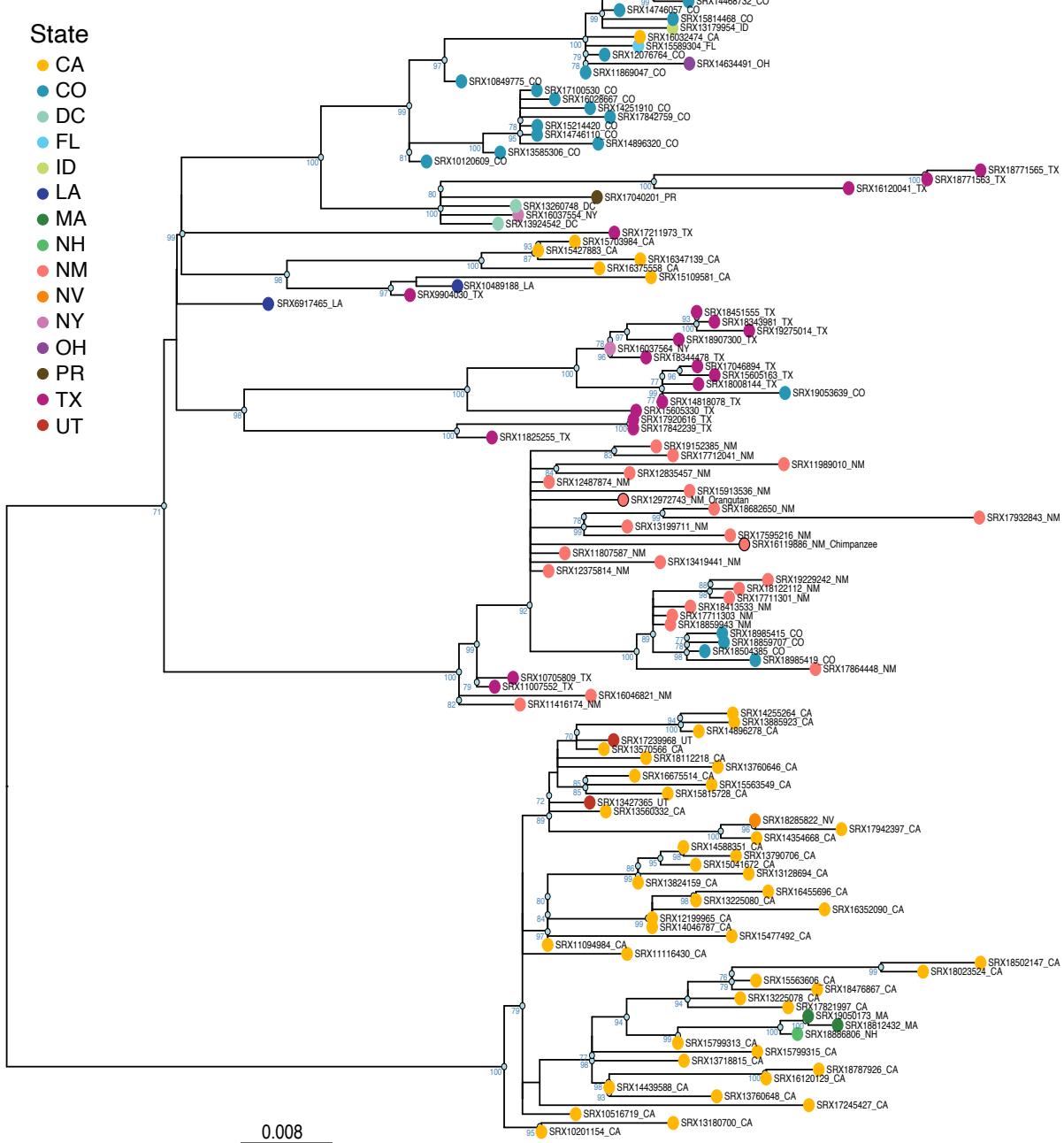
Supplementary Figure 2. BEAST phylogeny using the alignment against the external reference *S. flexneri* 2a str. 301 The phylogeny was inferred under the GTR+GAMMA site heterogeneity model and with a constant molecular clock model. Posterior probability values are shown for nodes with more than three descendants with values >0.7 . Nodes are colored by demographic status.



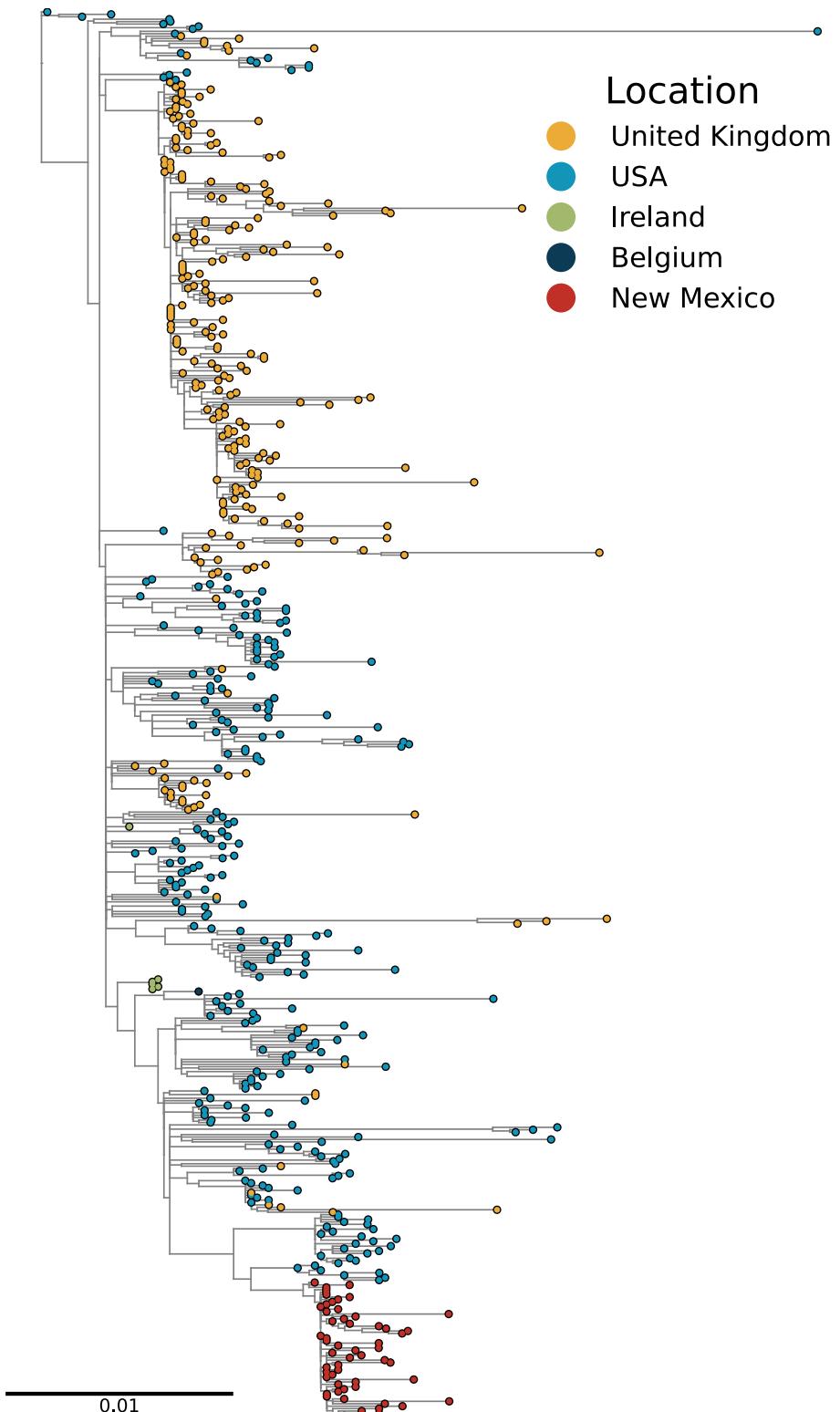
Supplementary Figure 3. BEAST phylogeny using the alignment against the internal reference PNUSA076676. The phylogeny was inferred under the GTR+GAMMA site heterogeneity model and with a constant molecular clock model. Posterior probability values are shown for nodes with more than three descendants with values >0.7. Nodes are colored by demographic status.



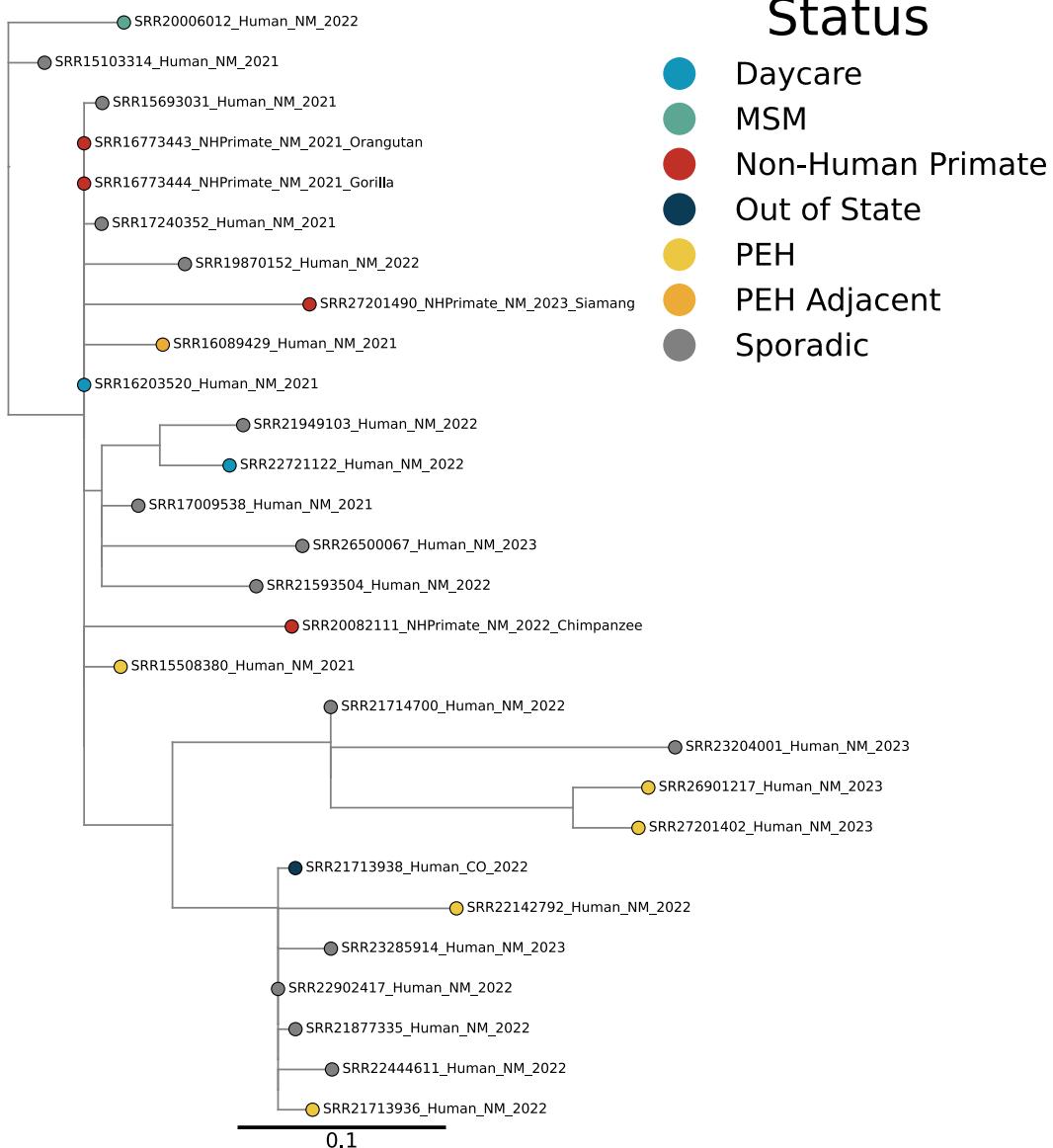
Supplementary Figure 4. Tree topology comparisons of BEAST phylogenies using the external reference *Shigella flexneri* 2a str. 301 and an internal reference strain PNUSA076676.



Supplementary Figure 5. Maximum likelihood phylogeny of 128 *S. flexneri* 2a genomes circulating within the U.S. Sequences were identified by CDC's PulseNet program using the alignment against the external reference *S. flexneri* 2a str. 301 under the GTR+GAMMA site heterogeneity model. Bootstrap values are shown for nodes with values >0.7. Nodes are colored by state.



Supplementary Figure 6. Maximum likelihood core-genome phylogeny of 580 *S. flexneri* 2a genomes that are within 13 cgMLST alleles from the New Mexico isolates. Data was pulled from the PathogenWatch database. There are 56 genomes represented in the phylogeny from our NM outbreak.



Status

- Daycare
- MSM
- Non-Human Primate
- Out of State
- PEH
- PEH Adjacent
- Sporadic

Supplementary Figure 7. Maximum likelihood phylogeny of 28 *S. flexneri* 2a genomes sequenced via Nanopore and Illumina to create complete genomes. The resulting genomes were compared using a reference-free method using *ska2* and the resulting alignment passed to IQ-Tree to generate a maximum likelihood phylogeny under the GTR+GAMMA model. Nodes are colored by demographic status.