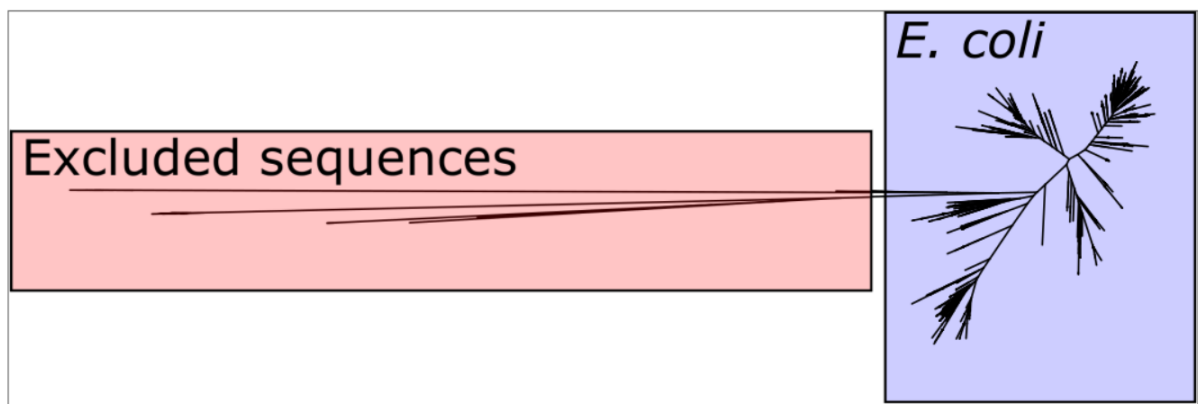


- 1 **Additional File 2. Phylogenetic position of 51 genomes too distantly related to *E. coli* to be**  
2 **included in our analysis**



- 3  
4 A maximum likelihood tree constructed using the concatenated sequence of 201 genes from the  
5 core genome of *E. coli*. Genomes were excluded when they fell distantly from the vast majority of *E.*  
6 *coli* genomes. These sequences are more closely related to other *Escherichia* species.