



Supplementary Figure 1. Unrooted cgMLST phylogenetic tree of the 728 complete *E. coli* assemblies used in the reference genome databases of Sigma and Sparse. The colors used to mark the different sections of the tree correspond to the colors of the reference genome clusters used in all figures and tables, allowing identify the section of the tree harbouring the assemblies associated with each cluster.