

3 Supplementary Figures

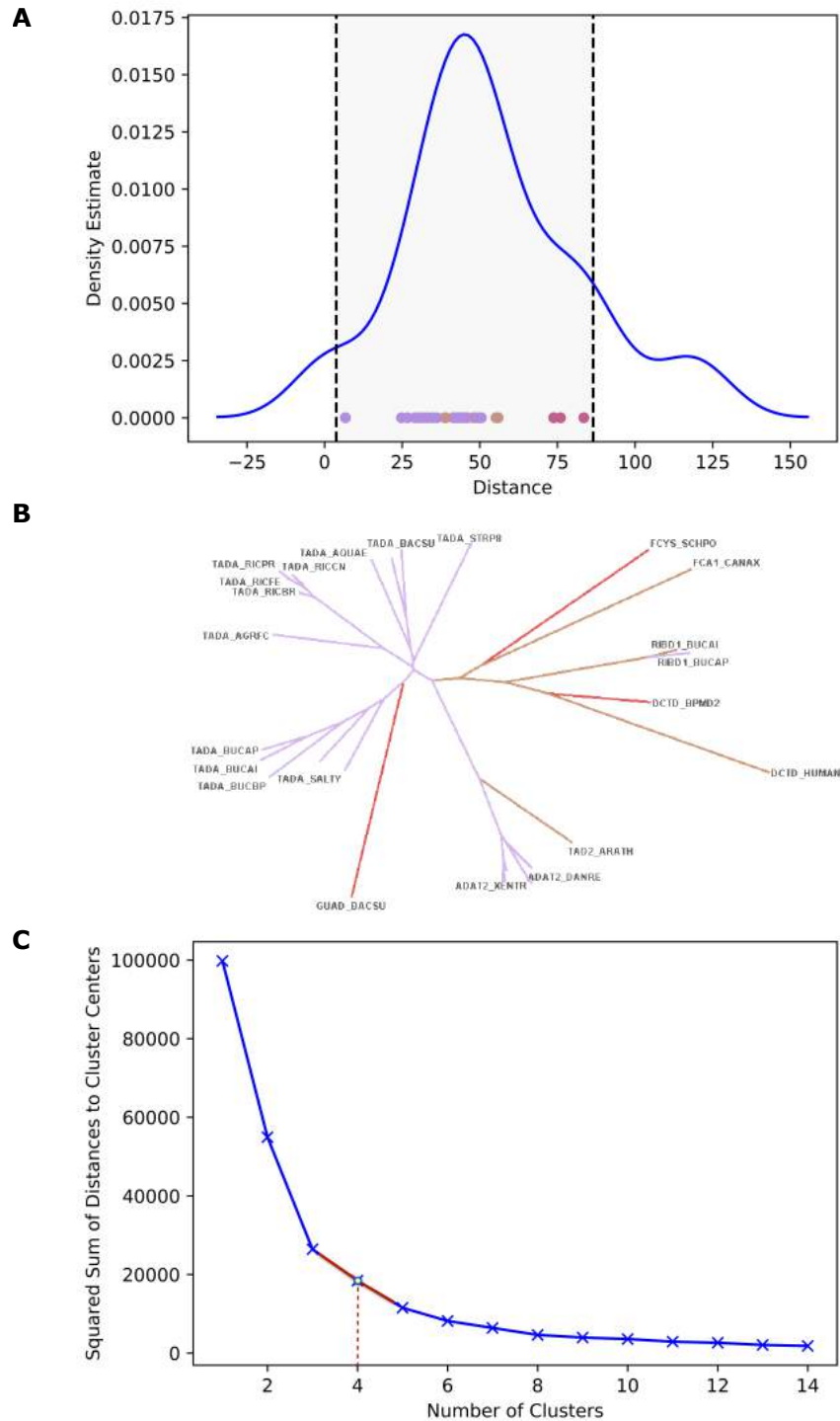


Figure S1: (A) Kernel density estimate (KDE) showing the distance of sequences from *E. coli* wtTadA from the first 2 principal components. (B) Phylogenetic analysis of the alignment for *E. coli* wtTadA. (C) Elbow plot to determine the optimal number of clusters to be used in k-means clustering.