

## **Description of Additional Supplementary Files**

File Name: Supplementary Data 1

Description: Metadata, accession numbers and assembly statistics for 752 ST58 genomes used in this study

File Name: Supplementary Data 2

Description: Gene screening data for 752 ST58 genomes used in this study

File Name: Supplementary Data 3

Description: Scoary output of genes identified to be significantly over-or under-represented in BAP clusters

File Name: Supplementary Data 4

Description: Metadata and ColV gene screening data for 34,364 *E. coli* genome assemblies from Enterobase