

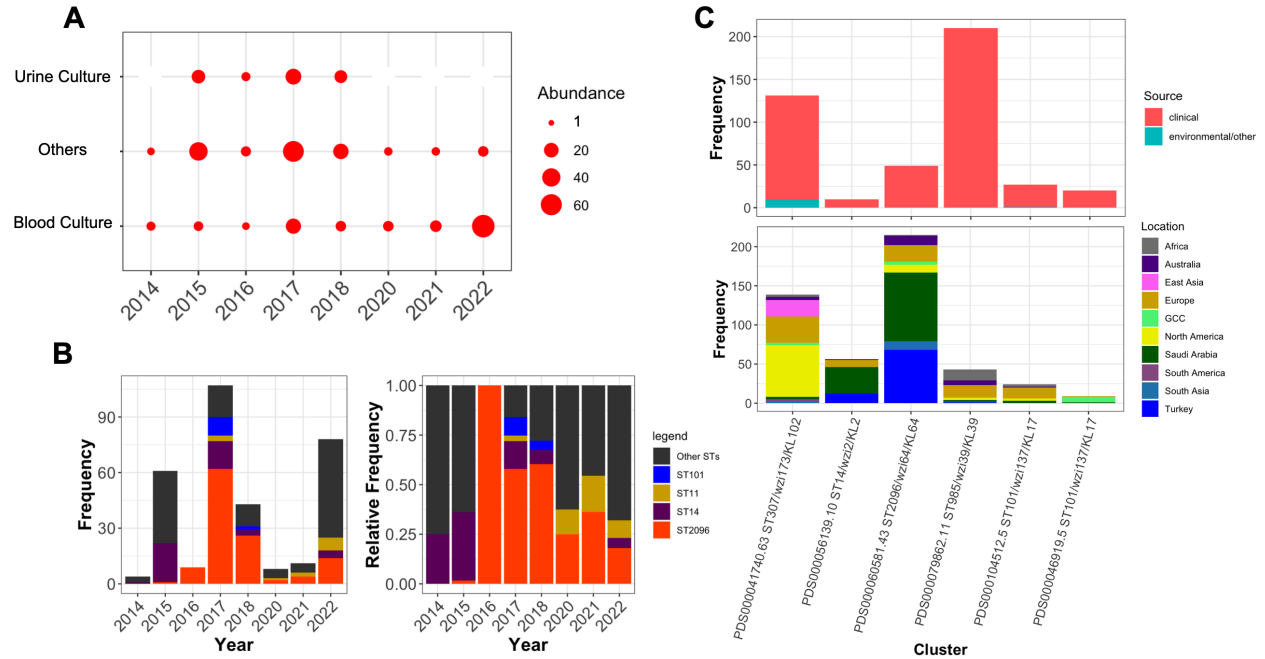


Fig S1 Frequency of source and STs in the collection. A) The distribution of sources of isolates over eight years. B) Absolute and relative frequency of STs over eight years. Other STs have STs with fewer than ten representative samples in the collection. The relative frequency corresponds to the count of STs. In a year, divided by the total number of isolates from the same year. C) The distribution of strains that fell under the same SNP cluster, as defined by Pathogen Detection, as one sample in our collection. Distributions for the location of isolation and sample types are illustrated.

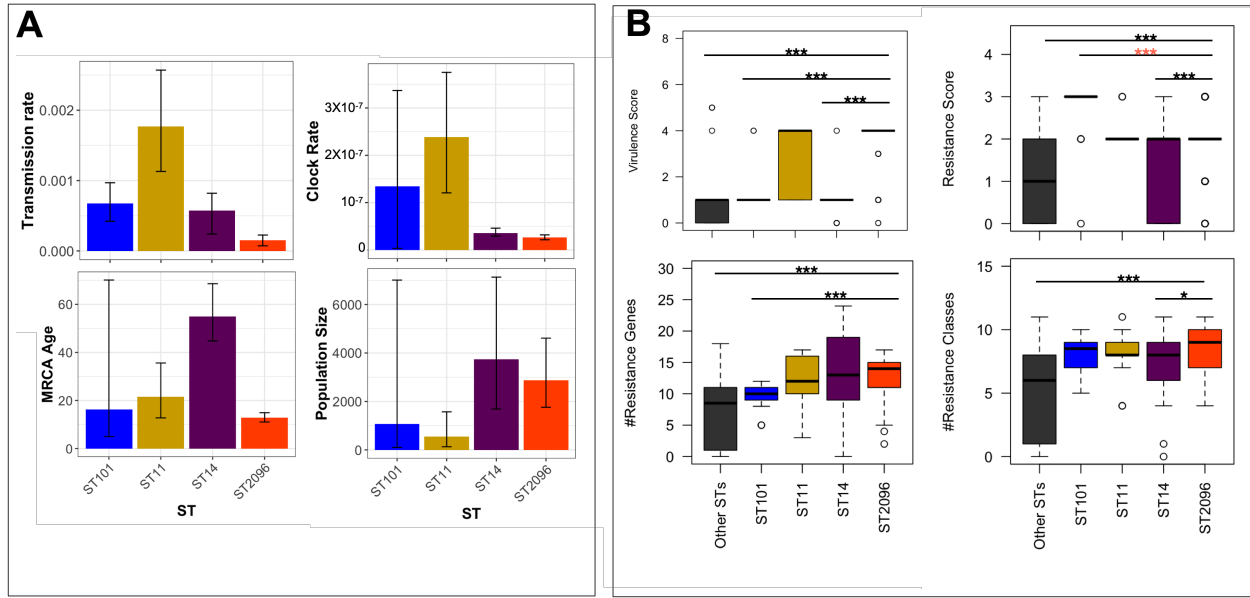
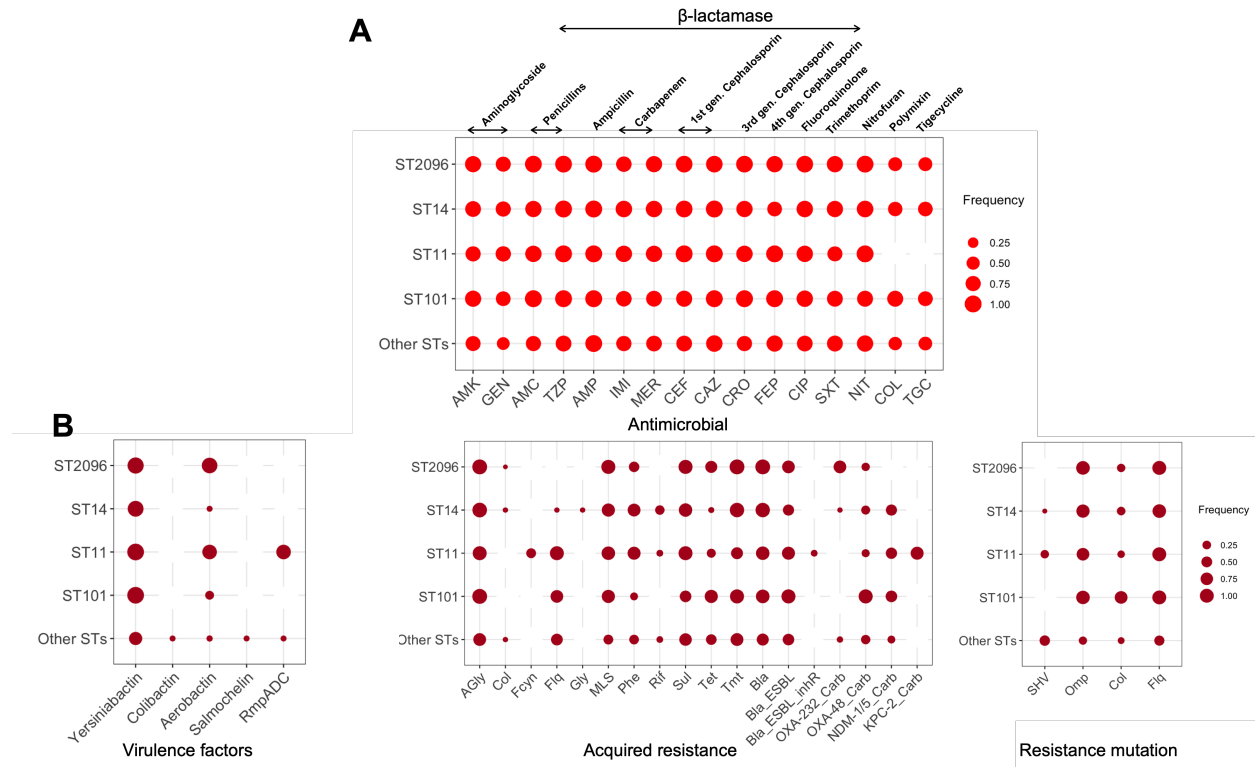


Fig S2 Comparative analysis of STs resistance, virulence and dynamics. A) Epidemiological features from phylodynamic analysis by SCOTTI for STs. The units for MRCA age, transmission rate and clock rate are years, transmission per individual per year and substitution per site per year, respectively. The error bars correspond to 95% credible interval. B) resistance and virulence score and resistance genes/classes count for the major STs and minor STs combined. The boxes show the distribution of the feature in the isolates belonging to the ST group, with dots showing the outliers. For a definition of virulence and resistance score see Methods. The * and ** sign correspond to <0.05 and <0.01 significance level from the Wilcoxon signed-rank test, respectively. The orange and black colors for the asterisks correspond to cases where the mean value (in A and B) and relative frequency (in C) for ST2096 are greater and smaller than the other group, respectively.



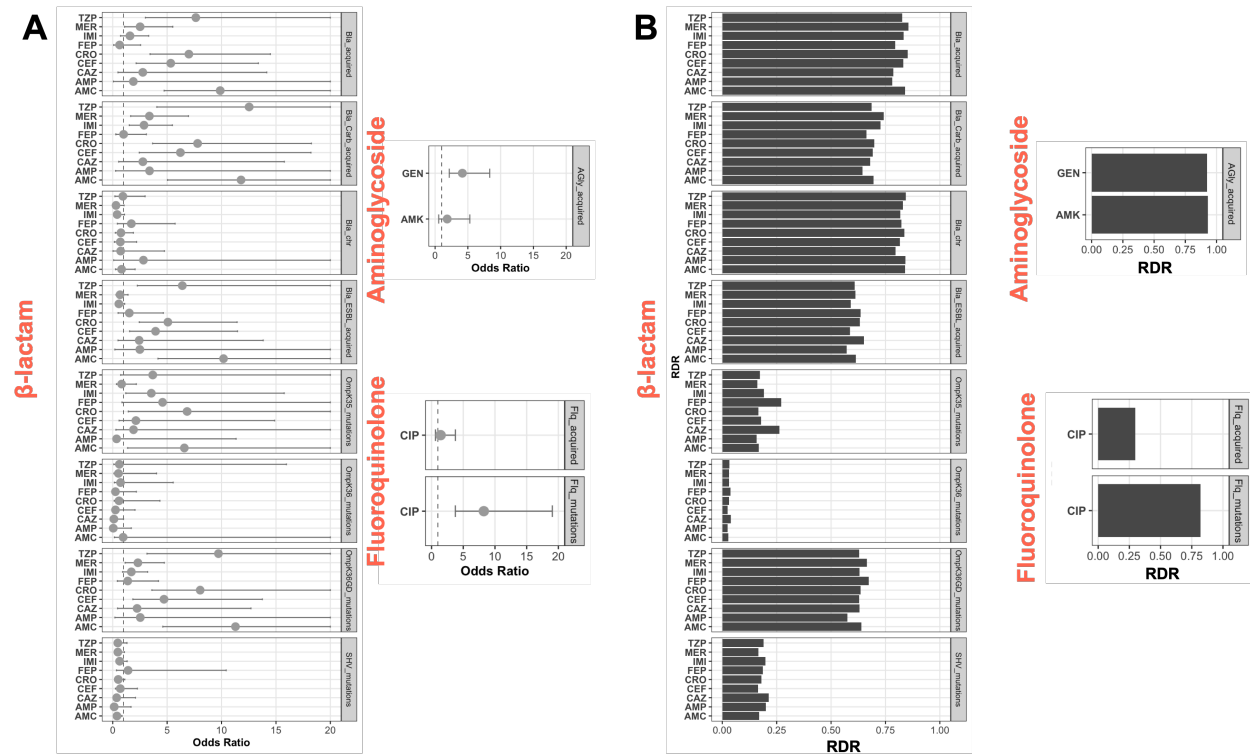


Fig S4 Integration of phenotypic resistance data and known resistance determinants. A) Odds ratio of resistance for different resistance determinants in three classes of antimicrobials. The error bars correspond to a 95% confidence interval. B) Resistance detection rate (RDR) for different resistance determinants in three classes of antimicrobials. The rate corresponds to the number of resistant strains that carry the determinant divided by the total number of resistant strains for each antimicrobial. Abbreviations in genes are detailed in the caption of Figure 1.

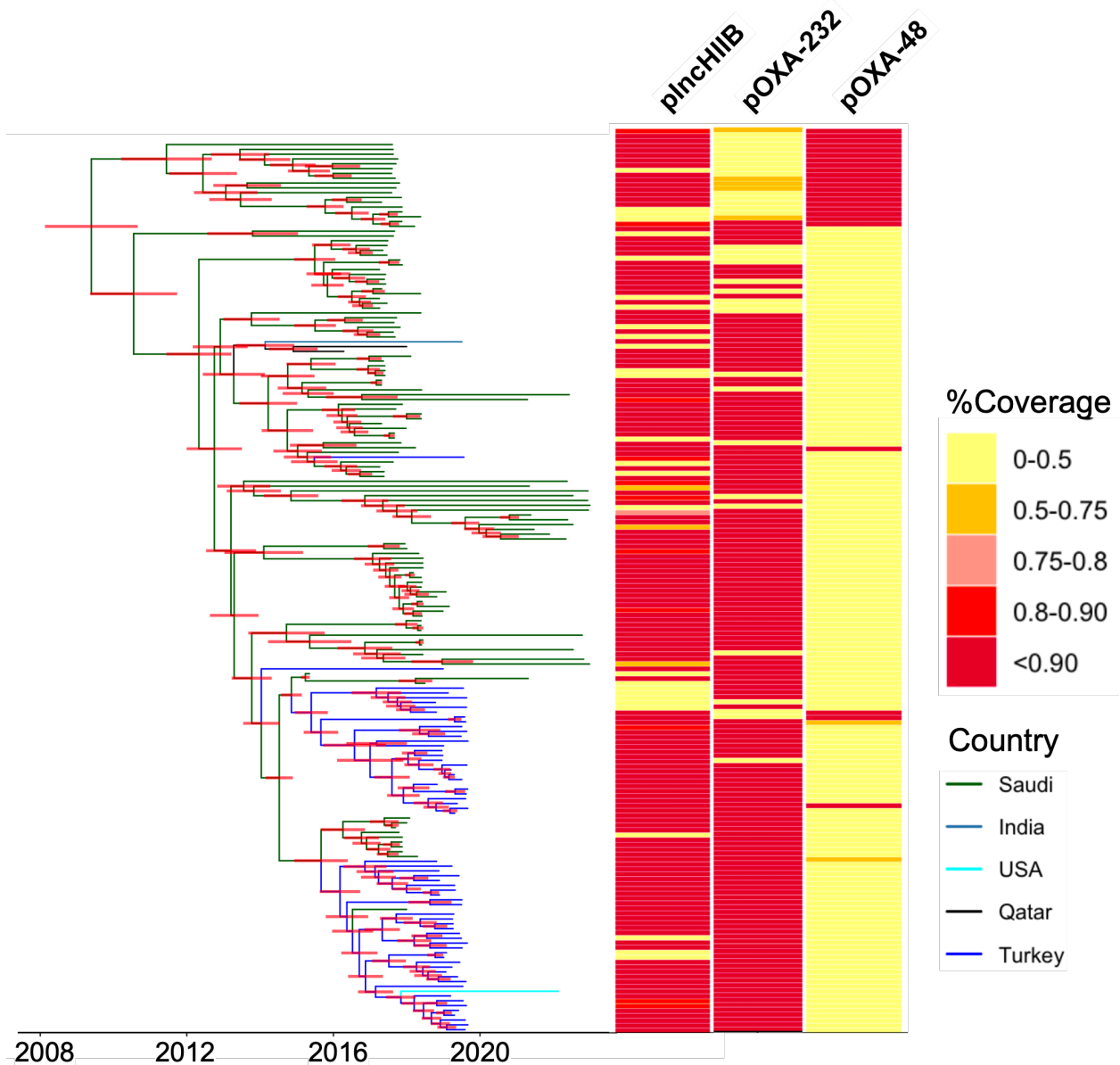


Fig S5: phylogeographic analysis for ST2096 strains in Saudi Arabia and globally. Only strains with an available date of collection were included. The tree is a dated Bayesian tree resulting from 1961 SNP sites for 185 samples in total. The horizontal red bars on the tree show the 95% Highest Posterior Density (HPD) for the age of the node. The colour of ancestral branches represents the most likely location (country) of the ancestors. The heatmap strips display the percentage coverage of plasmids recovered from long-read sequencing data for the entire collection.

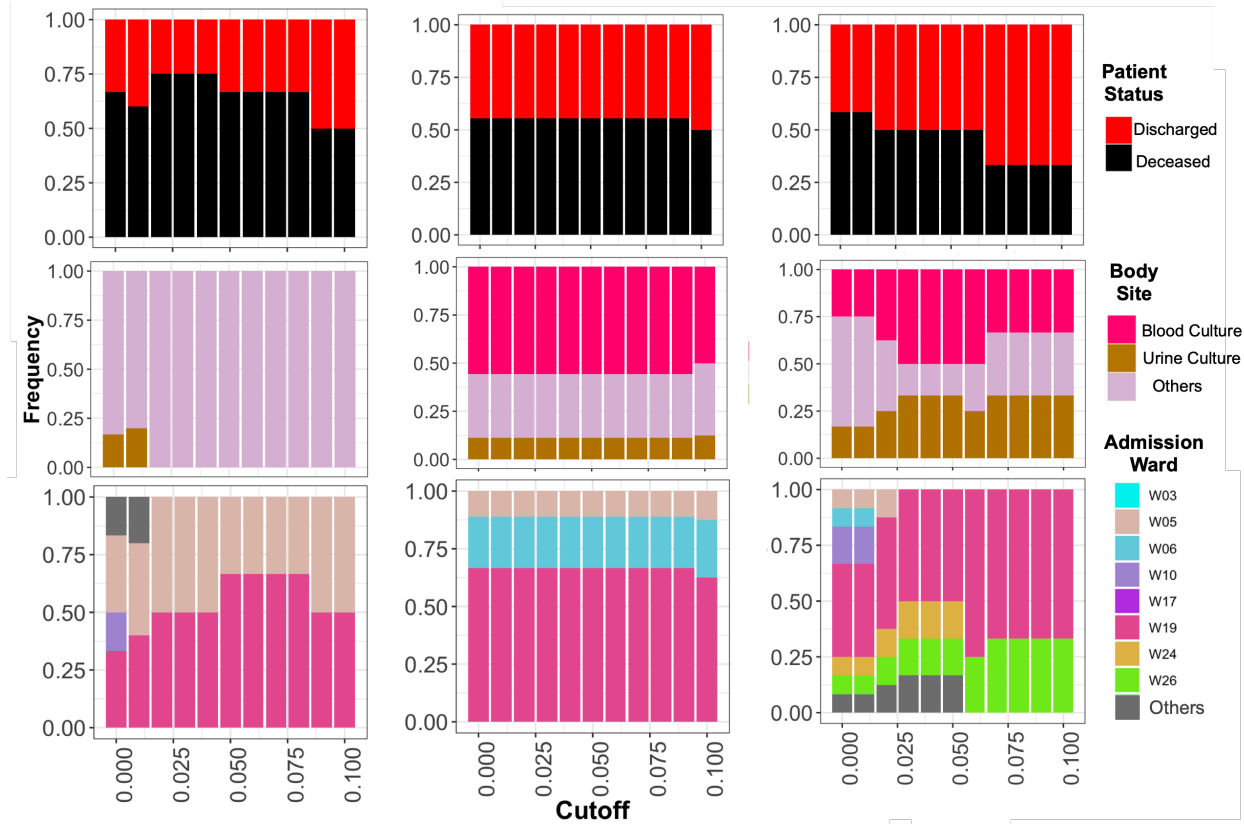


Fig S6 Transmission dynamics of other major STs. The distribution of patient status, body site, and admission hospital ward in networks with different cut-offs for the transmission routes for major STs other than ST2096, i.e., ST14, ST11 and ST101.

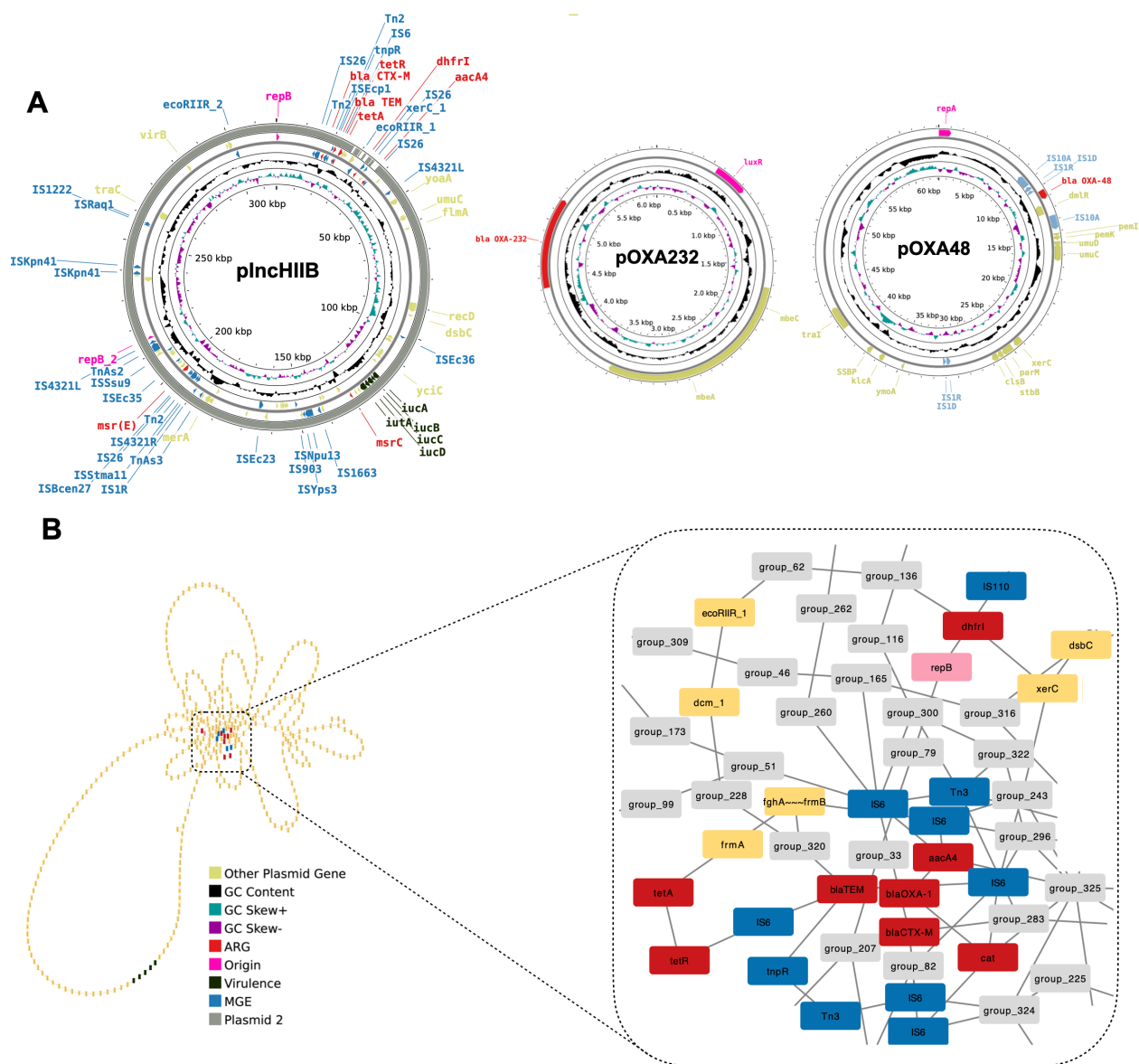


Fig S7 Plasmid structures and pangenome A) Genomic map of ESBL (pIncHIIB) and Carbapenemase plasmids in the ST2096 clones. For each plasmid, we selected the longest plasmid contig in the collection. For the H1-1B plasmid, we did not show the mobile genetic elements to improve presentation. Detailed information on the plasmids and genes is provided in the Supplementary Material. B) Pangenome graph for the pIncHIIB plasmid. Each node corresponds to a gene family. Edges denote connections between the two nodes if the nodes are neighbors on a contig in any sample from the population. The annotation of the genes is provided in the GitHub directory for the project.

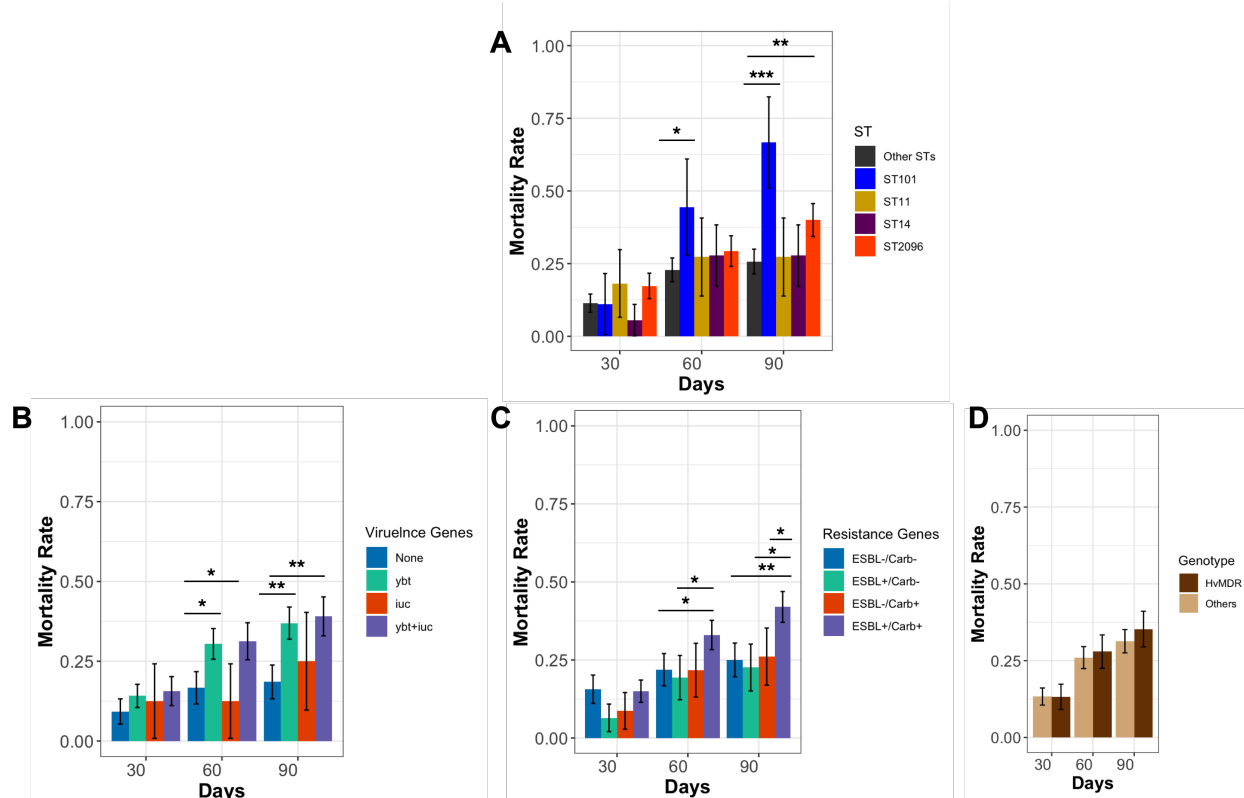


Fig S9 The overall mortality rate identified in the collection for patients carrying different A) STs, B) key virulence factor genes, C) key beta-lactamase gene profiles and D) hypervirulent-resistant as defined in Figure 3. The mortality rate is the reported in-hospital mortality rate and does not include the deaths occurred after the patient's discharge. The stars correspond the significance of the one-sides proportion test, with *, ** and *** corresponding to significance level 0.05, 0.01 and 0.001 respectively. The hospital bars show the bars compared.