



**Figure S18.** Frequency of ST lineages across bacterial species and programs using two random sample sizes. Pairwise frequency-based (y-axis = count) distribution of unique ST lineages identified by either mlst or stringMLST across bacterial species using a random sampling of 100 or 1,000 genomes. Each couple of pairwise bar counts (both programs) represents a unique ST number or lineage. Highlighted with black rectangles are the ST lineages uniquely found by the stringMLST program. Moreover, the STs unique for stringMLST are listed in a text box in each plot.