**Figure S1. Distribution of sample numbers in three TB databases**

Number of isolates that are distributed in the three TB databases: PATRIC, ReSeqTB, and a set generated from a PUBMED literature review which contributed an additional 455 isolates not present in PATRIC and ReSeqTB. The numbers displayed here represent the number of samples with both WGS data and DST phenotypes for at least one drug.

