SUPPLEMENTARY INFORMATION

The role of epistasis in amikacin, kanamycin, bedaquiline, and clofazimine resistance in *Mycobacterium tuberculosis* complex

Roger Vargas Jr, Luca Freschi, Andrea Spitaleri, Sabira Tahseen, Ivan Barilar, Stefan Niemann, Paolo Miotto, Daniela Maria Cirillo, Claudio U. Köser, and Maha R. Farhat

Supplementary Figures

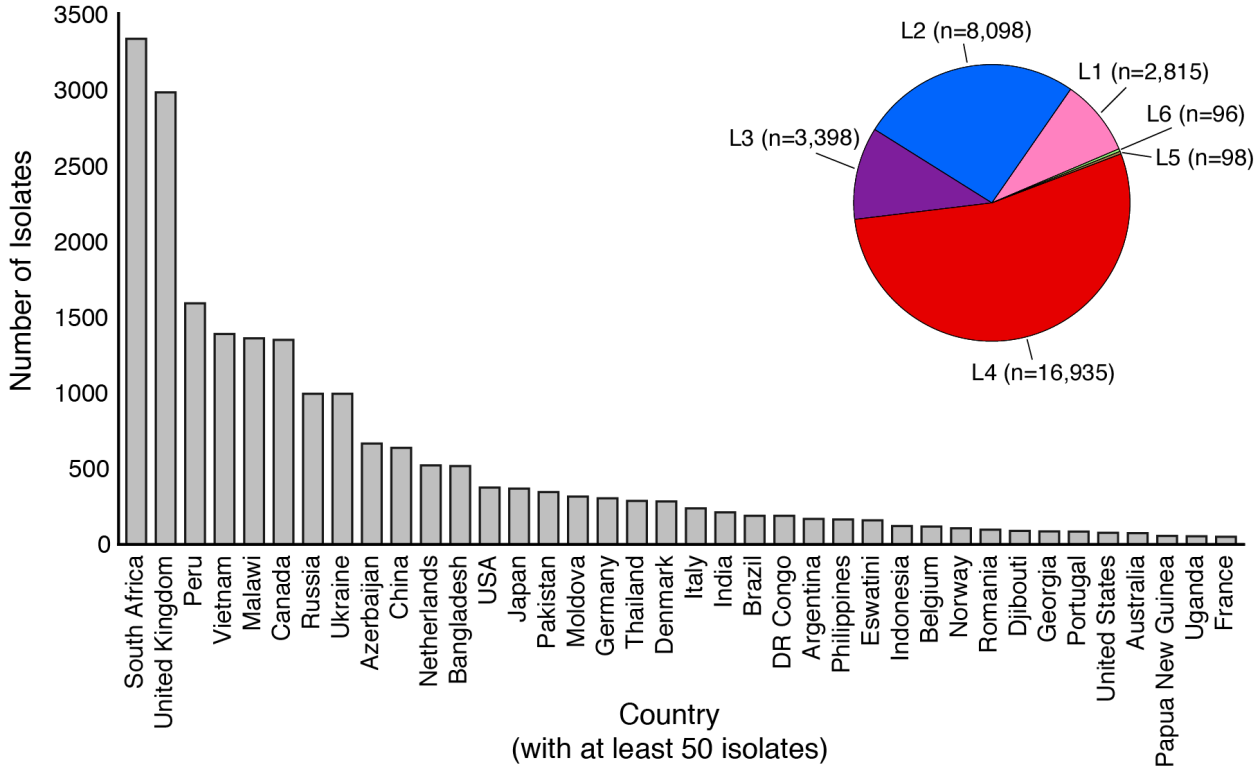

Fig. S1. Summary Characteristics of Sample. A breakdown of the geographical origin and global lineage for isolates in our sample. Of the 31,440 isolates in our sample, we had geographic information for 21,512 isolates. Of these 80 countries, 38/80 countries had more than 50 isolates. Our sample represented six global MTBC lineages (breakdown shown above).

Supplementary Table Descriptions

Table S1. Mutations detected in a global sample of MTBC clinical isolates. A full list of mutations that occur within our sample of 31,440 clinical isolates within the *mmpL5*, *mmpS5*, *mmpR*, *ahpC*, *eis*, *whiB7* coding sequences and *oxyR-ahpC*, *eis*-Rv2417c, *whiB7-uvrD2* intergenic regions. (XLSX file)

Table S2. Mixed indels in the *mmpR-mmpL5-mmpS5* chromosomal region. A list of frameshift indels that were detected at an intermediate allele frequency between 10% and 75% in *mmpR*, *mmpS5*, or *mmpL5* within our sample of 31,428 isolates (excludes the set of 12 added isolates, see **Methods**). (XLSX file)

Table S3. Co-occurrence of regulator resistance mutations and regulon LoF mutations. A more detailed version of **Table 2**. (XLSX file)

Table S4. Binary resistance phenotypes for MTBC sub-lineage 4.11 isolates. A table of binary resistance phenotype (STR, INH, RIF, EMB, PZA, AMK & KAN) data for a subset isolates that belong to sub-lineage 4.11 (**Fig. 2**), curated from multiple studies (1). (XLSX file)

Table S5. Count of isolates with *eis* promoter mutations and no coinciding *rrs* AG resistance mutations. The count of isolates with *eis* promoter mutations (G-10A, C-12T, C-14T, G-37T) that coincide with any AG resistance mutations in *rrs* (A1401G, C1402T, G1484T). (XLSX file)

Table S6. KAN and AMK resistance details for strains with MICs and strains with double *eis* promoter SNP & *eis* LoF mutations. A more detailed version of **Table 3** with binary resistance phenotype (STR, INH, RIF, EMB, PZA, AMK & KAN) data for a subset of isolates (1). (XLSX file)

Table S7. Accession IDs for genomes in sample. A table with the IDs for all 31,440 genomes analyzed in this study and information on the source (ie repository that they can be accessed from). (XLSX file)

References

1. Groschel MI, Owens M, Freschi L, Vargas R, Marin MG, Phelan J, Iqbal Z, Dixit A, Farhat MR. 2021. GenTB: A user-friendly genome-based predictor for tuberculosis resistance powered by machine learning. bioRxiv.