Supplementary Material for Brown, Tang, et al.

“Genotype-phenotype characterization of serial *Mycobacterium tuberculosis* isolates in bedaquiline-resistant tuberculosis”

**Supplementary Methods**

*Parent study design and inclusion criteria*

PROBeX was an observational cohort study conducted at 3 TB referral hospitals in Gqeberha, Cape Town, and Durban, South Africa between 2016 and 2020 [1]. The study enrolled a total of 195 participants with XDR or pre-XDR pulmonary TB, and individuals with MDR-TB who had toxicity or intolerance with other TB medications, all of whom were eligible for BDQ therapy per national tuberculosis program guidelines. All participants signed written informed consent. We defined XDR and pre-XDR TB using criteria in use at the time of the study (XDR: resistance to at least isoniazid, rifampin, fluoroquinolones, and any injectable medication; pre-XDR: resistance to at least isoniazid, rifampin, and either resistance to fluoroquinolones or an injectable medication). Exclusion criteria included prior treatment with BDQ, prolonged QTc interval on electrocardiogram, and abnormal kidney or liver function testing. Participants in the study were treated with a standardized regimen that typically included BDQ, linezolid, clofazimine, levofloxacin, ethionamide, terizidone, and pyrazinamide. One total participant was excluded for prior BDQ exposure. Total treatment duration was 18-24 month and most participants received BDQ for 6 months. We collected clinical history and reviewed medical records regarding prior episodes of TB and TB therapy. Sputum mycobacterial cultures were collected at least monthly and study staff attempted to obtain the *Mtb* isolates for any positive cultures from the local TB laboratory.

*BDQ minimum inhibitory concentration measurements*

Two-fold serial dilutions of 100x working solutions of BDQ were prepared from a 1 mg/ml stock solution using dimethylsulfoxide (DMSO). Aliquots were stored at -70°C for no longer than 6 months. The 100x working solutions in DMSO were further diluted in the MGIT tubes to obtain final concentrations of 2, 4, 8 μg/mL. The BACTEC™ MGIT™ 960 DST method described by Rüsch-Gerdes et al. was followed, with minor modifications for MIC testing. [2, 3] The MIC is determined to be the lowest drug concentration that tested susceptible (i.e., ≤100 growth units) by automated reading when the control vial turns positive (400 growth units). [quality control procedures]

*Whole genome sequencing*

We conducted whole genome sequencing for Mtb clinical strains using the Illumina NextSeq platform, with read lengths of 150bp and coverage at >100x using the Nextera DNA Flex Library Preparation kit. We trimmed raw sequencing reads using Prinseq v0.20.4 [4], removed non-*Mtb* (contaminant) reads with Kraken [5], and aligned the remaining reads to the H37Rv reference genome (NC000962.3) using BWA-MEMv0.7.15 [6]. After read filtering, all samples had reads covering >99% of the reference genome, and mean read depth across the genome was 356.8 (range: 110.3-700.7). We used Pilon to generate high confidence variant calls from the resulting alignments [7] and excluded variants within 50 base pairs of hypervariable PPE/PE gene families, repeat regions, and mobile elements, similar to prior studies using WGS from *Mtb*.[8] Mean read depth across the resulting set of high-quality SNPs was 163.5 reads and the average number of high quality SNPs per sample (versus the H37Rv reference genome) was 951.1. WGS data is available on the NCBI Sequence Read Archive (BioProject ID: PRJNA1008673).

*mmpR5 targeted sequencing*

A 1.2 mL heat-killed aliquot of *M. tuberculosis* isolate from the MGIT 960 tube (Becton Dickinson, Sparks, USA) was used for DNA extraction on the NucliSENS easyMAG system (BioMérieux, Marcy-l’Étoile, France). Previously published primers [9] specific for the entire gene length of *mmpR5* were used for amplification using a PCR reaction mix comprising of 10 µL KAPA mastermix (2X), 1 µL forward *mmpR5* primer [10 µM] and 1 µL reverse *mmpR5* primer [10 µM] and 2 µL DNA topped up with PCR-grade water to a final volume of 20 µL. Temperature cycling conditions for amplification were: initial denaturation at 95°C for 5 minutes; 35 cycles of denaturation at 94°C for 1 minute, annealing at 60°C for 1 minute and extension at 72°C for 1 minute; followed by a final extension at 72°C for 10 minutes. The resultant *mmpR5* PCR amplicons were purified using the ExoSAP-IT™ Express PCR Product Cleanup Reagent (ThermoFisher Scientific, Waltham, Massachusetts, USA) as per manufacturer’s instructions. The purified amplicons were sequenced using the using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (ThermoFisher Scientific, Waltham, Massachusetts, USA) and *mmpR5* forward primer [3.2 µM] under the following conditions: initial denaturation at 96°C for 1 minute; 25 cycles of denaturation at 96°C for 10 seconds, annealing at 50°C for 5 seconds and extension at 60°C for 4 minutes. The sequencing products were purified using the BigDye XTerminator™ Purification Kit as per manufacturer’s instructions (ThermoFisher Scientific, Waltham, Massachusetts, USA) and loaded onto the Applied Biosystems® 3500 Series Genetic Analyser (ThermoFisher Scientific, Waltham, Massachusetts, USA). Resequencing analysis and variant detection was performed on CLC Genomics Workbench v11.0.1 (Qiagen, Venlo, the Netherlands) using the NC000962.3 reference genome.

*Study-level phylogenetic tree*

We used filtered variant calls and iqtree v1.6.1011 to infer phylogenetic trees by maximum likelihood [10], employing the ModelFinder function to identify the best-fit nucleotide substitution model for the data and using 1000 bootstrap replicates to evaluate uncertainty for each node in the tree. We annotated the resulting tree with BDQ MIC values, *mmpR5* mutational type, and genotypic antibiotic resistance profiles described above.

**Supplementary Tables**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | WGS | Sanger | Curated WGS |  | ID | WGS | Sanger | Curated WGS |
| 509-1 | 198insG  [I67fs] | 198insG  [I67fs] | 198insG  [I67fs] |  | **8036-1** | G203A  [S68N] |  | G203A  [S68N] |
| 1003-1 | T425G  [L142R] | T425G  [L142R] | T425G  [L142R] |  | **8036-2** | G203A  [S68N] |  | G203A  [S68N] |
| 7002-1 | WT |  | WT |  | **9002-1** | WT |  | WT |
| 7002-2 | WT |  | WT |  | **9002-2** | WT |  | WT |
| 7002-3 | WT |  | WT |  | **9002-3** | WT |  | WT |
| 7002-4 | WT | WT | WT |  | **9002-4** | 349insC  [L117fs] | 349insC  [L117fs] | 349insC  [L117fs] |
| 7002-5 | WT |  | WT |  | **9017-1** | WT | WT | WT |
| 7033-1 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9019-1** | G188A [S63N] 198insG [I67fs] | 198insG[I67fs] | G188A[S63N] 198insG[I67fs] |
| 7033-2 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9019-2** | WT |  | WT |
| 7034-1 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9019-3** | WT |  | WT |
| 7034-2 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9020-1** | 139\_142insGATC  [P48fs] |  | 139\_142insGATC  [P48fs] |
| 7034-3 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9020-2** | 139\_142insGATC  [P48fs] | 139\_142insGATC  [P48fs] | 139\_142insGATC  [P48fs] |
| 7034-4 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9022-1** | 493insG  [D165fs] | 493insG  [D165fs] | 493insG  [D165fs] |
| 7034-5 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9022-2** | 493insG  [D165fs] | 493insG  [D165fs] | 493insG  [D165fs] |
| 7034-6 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9022-3** | 493insG  [D165fs] | 493insG  [D165fs] | 493insG  [D165fs] |
| 7034-7 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9024-1** | WT |  | WT |
| 7034-8 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9024-2** | WT |  | WT |
| 7034-9 | 144insC  [E49fs] | 144insC  [E49fs] | 144insC  [E49fs] |  | **9024-3** | WT | WT | WT |
| 7043-1 | T437C  [M146T] | T437C  [M146T] | T437C  [M146T] |  | **9024-4** | WT |  | WT |
| 7046-1 | 139insG[D47fs]  144insC[E49fs] | 139insG[D47fs]  144insC[E49fs] | 139insG[D47fs]  144insC[E49fs] |  | **9024-5** | WT |  | WT |
| 7046-2 | 139insG  [D47fs] | 139insG  [D47fs] | 139insG  [D47fs] |  | **9024-1** | 144insC[E49fs] | 144insC[E49fs] | 144insC[E49fs] |
| 7071-1 | T277C | T277C | T277C |  | **9027-1** | WT |  | WT |
| 7072-1 | WT |  | WT |  | **9027-2** | WT |  | WT |
| 7072-2 | WT |  | WT |  | **9027-3** | WT |  | WT |
| 7072-3 | WT | WT | WT |  | **9027-4** | WT |  | WT |
| 7072-4 | WT |  | WT |  | **9027-5** | WT |  | WT |
| 7080-1 | 144insC[E49fs] | 144insC[E49fs] | 144insC[E49fs] |  | **9027-6** | A202G[S68G] | A202G[S68G] | A202G[S68G] |
| 7080-2 | 144insC[E49fs] | 144insC[E49fs] | 144insC[E49fs] |  | **9027-7** | A202G[S68G] |  | A202G[S68G] |
| 7080-3 | 144insC[E49fs] | 144insC[E49fs] | 144insC[E49fs] |  | **9027-8** | A202G[S68G] |  | A202G[S68G] |
| 7080-4 | 144insC[E49fs] | 144insC[E49fs] | 144insC[E49fs] |  | **9028-1** | WT |  | WT |
| 7080-5 | 144insC[E49fs] | 144insC[E49fs] | 144insC[E49fs] |  | **9028-2** | WT |  | WT |
| 7080-6 | 198insG[I67fs] | 198insG[I67fs] | 198insG[I67fs] |  | **9028-3** | 139insG[D47fs] | 139insG[D47fs] | 139insG[D47fs] |
| 7080-7 | A202G[S68R] |  | A202G[S68R] |  | **9028-4** | WT |  | WT |
| 8012-1 | WT |  | WT |  | **9028-5** | WT |  | WT |
| 8012-2 | 144insC[E49fs] |  | 144insC[E49fs] |  | **9030-1** | WT |  | WT |
| 8012-3 | 137insG[C46Y]  C446T[R156\*] |  | 137insG[C46Y]  144insC[E49fs]  C446T[R156\*] |  | **9030-2** | WT |  | WT |
| 8012-4 | 144insC[E49fs]  C446T[R156\*] |  | 137insG[C46Y]  144insC[E49fs]  C446T[R156\*] |  | **9030-3** | WT |  | WT |
| 8012-5 | C446T[R156\*] |  | 137insG[C46Y]  144insC[E49fs]  C446T[R156\*] |  | **9030-4** | WT |  | WT |
| 8012-6 | WT |  | WT |  | **9030-5** | WT | WT | WT |
| 8012-7 | 144insC[E49fs] |  | 144insC[E49fs] |  | **9030-6** | WT |  | WT |
| 8012-8 | 144insC[E49fs] |  | 144insC[E49fs] |  | **9030-7** | WT |  | WT |
| 8012-9 | 144insC[E49fs] |  | 144insC[E49fs] |  | **9030-8** | WT |  | WT |
| 8014-1 | WT |  | WT |  | **9030-9** | WT |  | WT |
| 8014-6 | 198delG[I67fs] |  | G137A[C46Y] 198delG[I67fs] |  | **9030-10** | WT |  | WT |
| 8014-2 | WT |  | WT |  | **9031-1** | WT |  | WT |
| 8014-3 | WT |  | WT |  | **9031-2** | WT |  | WT |
| 8014-4 | 198delG[I67fs] |  | 198delG[I67fs] |  | **9031-3** | 379delG[D127fs] | 379delG[D127fs] | 379delG[D127fs] |
| 8014-5 | G137A[C46Y] 198delG[I67fs] |  | G137A[C46Y] 198delG[I67fs] |  | **9040-1** | WT |  | WT |
| 8014-7 | 198delG[I67fs] |  | G137A[C46Y] 198delG[I67fs] |  | **9040-2** | A202C[S68R] | A202C[S68R] | A202C[S68R] |

**Supplementary Table S1.** Sanger- and WGS-based *mmpR5* genotypesfor 98 *Mtb* isolates. Curated WGS genotypes, including discordant genotypes resolved via manual inspection of WGS read pileups, are shown in the third column.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Collection timepoint | Lineages | Proportion variant reads (lineage-defining SNPs) | Rv0678 | Proportion variant reads (mmpR5 SNPs) |
| 8012-3 | Month 9 | Lineage 4.4.1.1 | 1 | 137insG [C46fs] | 0.37 |
| 144insC [E49fs] | 0.18 |
| C466T [R156\*] | 0.39 |
| 8012-4 | Month 10 | Lineage 4.4.1.1 | 1 | 137insG [C46fs] | 0.02 |
| 144insC [E49fs] | 0.36 |
| C466T [R156\*] | 0.58 |
| 8012-5 | Month 11 | Lineage 4.4.1.1 | 1 | 137insG [C46fs] | 0.11 |
| 144insC [E49fs] | 0.12 |
| C466T [R156\*] | 0.72 |
| 8012-7 | Month 27 | Lineage 4.4.1.1 | 0.3 | 144insC [E49fs] | 0.41 |
| *M. bovis* | 0.66 |
| 8014-4 | Week 2 | Lineage 4.3.2.1 | 0.85 | 198delG [I67fs] | 0.11  0.89 |
| Lineage 2.2.1 | 0.13 |
| 8014-5 | Week 6 | Lineage 2.2.1 | 0.72 | G137A [C46Y] | 0.35 |
| Lineage 4.3.2.1 | 0.3 | 198delG [I67fs] | 0.41 |
| 8014-6 | Week 10 | Lineage 2.2.1 | 1 | G137A [C46Y] | 0.07 |
| 198delG [I67fs]  WT/other | 0.91  0.02 |
| 8014-7 | Month 3 | Lineage 2.2.1 | 1 | G137A [C46Y] | 0.17 |
| 198delG [I67fs] | 0.88 |
| 7046-1 | Month 6 | 4.8 | 1 | 139insG [D47fs] | 0.46 |
| 144insC [E49fs] | 0.50 |
| 9019-1 | Month 1 | Lineage 2.2.2 | 1 | G188A [S63N] | 0.31 |
| 198insG [I67fs] | 0.68 |

**Supplementary Table S2.** Proportion of reads supporting variant alleles for SNPs at lineage-defining positions and in *mmpr5* for samples with evidence of mixed *mmpR5* genotypes. Variant allele frequencies are calculated using the Pilon-reported depth of high-quality sequencing reads at either lineage defining SNP positions (per Coll et al. [cite]) or *mmpR* SNP positions, divided by the total depth of all high-quality reads at each position.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Any isolate with BDQ MIC >1 (n=24) |  | No isolate with BDQ MIC >1 (n=123) |  | p-value |
| *Demographics* |  |  |  |  |  |
| Female | 13 (0.54) |  | 71 (0.58) |  |  |
| Black | 21 (0.88) |  | 97 (0.79) |  |  |
| Mixed race | 3 (0.12) |  | 24 (0.20) |  |  |
| White | 0 (0) |  | 1 (0.01) |  |  |
|  |  |  |  |  | 0.780 |
| *HIV status* |  |  |  |  |  |
| HIV infected | 14 (0.58) |  | 78 (0.63) |  |  |
| HIV not on ART | 2 (0. 08) |  | 14 (0.11) |  |  |
| CD4 < 200 | 7 (0. 21) |  | 33 (0.27) |  |  |
|  |  |  |  |  | 0.942 |
| *TB history* |  |  |  |  |  |
| Prior TB | 17 (0.71) |  | 90 (0.73) |  |  |
| Prior DR-TB | 8 (0.33) |  | 47 (0.38) |  |  |
| Prior clofazimine | 3 (0.13) |  | 6 (0.05) |  |  |
|  |  |  |  |  | 0.332 |
| *TB drug resistance* |  |  |  |  |  |
| Pre-XDR | 7 (0.29) |  | 43 (0.35) |  |  |
| XDR | 10 (0.42) |  | 52 (0.42) |  |  |
|  |  |  |  |  | 0.797 |
|  |  |  |  |  |  |

**Supplementary Table S3.** Demographic characteristics of participants with (n=24) and without (n=123) any isolate with BDQ MIC >1. Omnibus p-values for the distribution of counts within each subcategory (demographics, HIV status, TB history, and TB drug resistance) are calculated using Fisher’s exact test in R.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | BDQ resistant or intermediate (n=24) |  | All isolates BDQ susceptible (n=123) | OR (95%CI) |
|  |  |  |  |  |
| Cured | 14 (0.58) |  | 76 (0.62) | 0.87 (0.36-2.11) |
| Death | 3 (0.13) |  | 14 (0.11) | 1.10 (0.29-4.17) |
| Treatment Complete | 2 (0.08) |  | 7 (0.06) | 1.49 (0.29-7.67) |
| Treatment Default | 2 (0.08) |  | 12 (0.10) | 0.83 (0.17-3.99) |
| Treatment Failure | 3 (0.13) |  | 4 (0.03) | 4.21 (0.88-20.20) |

**Supplementary Table S4.** Clinical outcomes in participants with resistance or intermediate susceptibility to BDQ (n=24) versus those with only BDQ-susceptible isolates (n=123). Omnibus p-value from Fischer’s exact test for distributions of counts across all outcomes: 0.42

**Supplementary Figures**

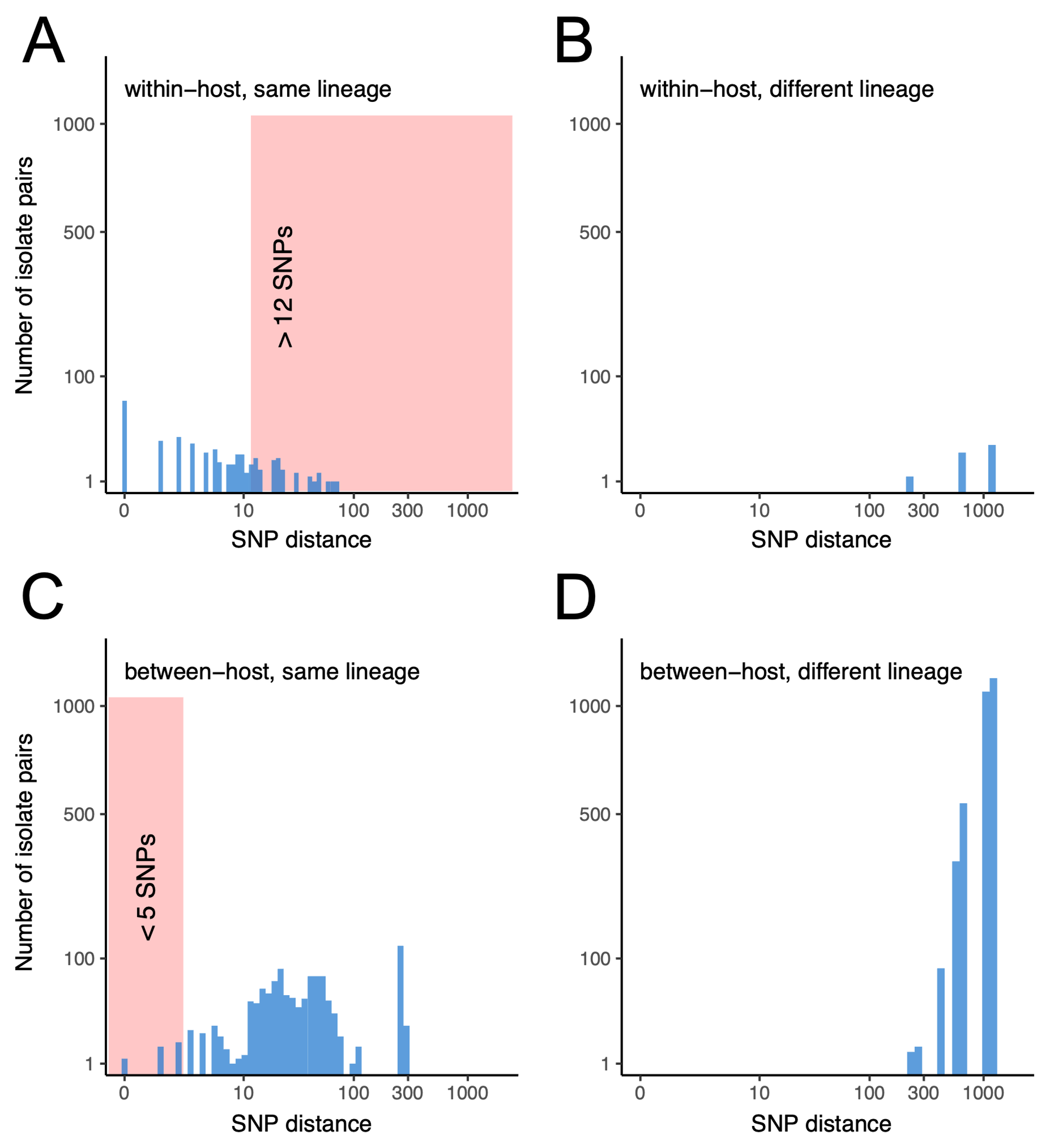
A picture containing text, screenshot, line, plot

Description automatically generated

**Supplementary Figure S1.** SNP non-reference allele frequency spectrum for raw (unfiltered) SNPs, filtered SNPs, and SNPs at lineage defining SNPs for WGS data from 98 samples in this study. Red lines: Samples with evidence of with-sample admixture, measured by the number of variants with highly mixed collection of reads (for both reference and non-reference alleles). Grey lines: All other samples.



**Supplementary Figure S2.** Participants with highly discrepant BDQ MIC measurements for within-lineage *Mtb* samples collected less than < 56 days apart that do not have interval changes in *mmpR5* genotype. Each panel represents all samples from each of 4 participants. Point color and shape indicate *Mtb* lineage and *mmpR5* genotype, respectively. Lines between points show pairwise SNP distances between samples, with darker blue for smaller SNP distances. Arrows indicate samples with ≥ 4-fold change in MIC, collected < 56 days apart, and with no identified differences in *mmpR5* genotype between samples.



**Supplementary Figure S3: *Mtb* within-host genetic diversity and between-host genetic similarity.** (A) Histogram of SNP distances for within-host isolate pairs from the same *Mtb* lineage. (B) Histogram of SNP distances for within-host isolate pairs from different *Mtb* lineages. (C) Histogram of SNP distances for between-host isolate pairs from the same *Mtb* lineage. (D) Histogram of SNP distances for between-host isolate pairs from different *Mtb* lineages. The pink shaded region in panel C shows between-host isolate pairs that differ by less than 5 total SNPs.



**Supplementary Figure S4: Phylogenetic reconstruction for 95 *Mtb* isolates from 24 individual study participants.** Annotations include BDQ MIC, *mmpR5* mutation type, genotypic resistance profile, and SNP-based phylogeographic lineage classification (following Coll et al.[11]). The clade containing isolates from the BDQ-resistant *mmpR5* 144insC [E49fs] mutant Strain A are highlighted in red.

**References**

1. Brust JCM, Gandhi NR, Wasserman S, et al. Effectiveness and Cardiac Safety of Bedaquiline-Based Therapy for Drug-Resistant Tuberculosis: A Prospective Cohort Study. Clin Infect Dis **2021**; 73(11): 2083-92.

2. Kaniga K, Aono A, Borroni E, et al. Validation of Bedaquiline Phenotypic Drug Susceptibility Testing Methods and Breakpoints: a Multilaboratory, Multicountry Study. J Clin Microbiol **2020**; 58(4).

3. Rusch-Gerdes S, Domehl C, Nardi G, Gismondo MR, Welscher HM, Pfyffer GE. Multicenter evaluation of the mycobacteria growth indicator tube for testing susceptibility of Mycobacterium tuberculosis to first-line drugs. J Clin Microbiol **1999**; 37(1): 45-8.

4. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics **2011**; 27(6): 863-4.

5. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol **2014**; 15(3): R46.

6. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **2009**; 25(14): 1754-60.

7. Walker BJ, Abeel T, Shea T, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One **2014**; 9(11): e112963.

8. Eldholm V, Monteserin J, Rieux A, et al. Four decades of transmission of a multidrug-resistant Mycobacterium tuberculosis outbreak strain. Nat Commun **2015**; 6: 7119.

9. Zimenkov DV, Nosova EY, Kulagina EV, et al. Examination of bedaquiline- and linezolid-resistant Mycobacterium tuberculosis isolates from the Moscow region. J Antimicrob Chemother **2017**; 72(7): 1901-6.

10. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol **2015**; 32(1): 268-74.

11. Coll F, McNerney R, Guerra-Assuncao JA, et al. A robust SNP barcode for typing Mycobacterium tuberculosis complex strains. Nat Commun **2014**; 5: 4812.