SUPPLEMENTARY DATA

Cryptic resistance mutations associated to misdiagnoses of Multidrug-resistant tuberculosis

Irving Cancino-Muñoz¹,³, M.Sc., Miguel Moreno-Molina¹, M.Sc., Victoria Furió¹, Ph.D., Galo A. Goig¹, M.Sc., Manuela Torres-Puente¹, Ph.D., Álvaro Chiner-Oms², M.Sc., Luis M. Villamayor¹, B.S., Francisco Sanz⁴, Ph.D., María Remedio Guna-Serrano⁵, Ph.D., and Iñaki Comas¹,⁶,a, Ph.D.

Affiliations:

¹ Instituto de Biomedicina de Valencia (IBV-CSIC), 46010, Valencia, Spain.
² Unidad Mixta “Infección y Salud Pública” (FISABIO-CSISP), 46021, Valencia, Spain.
³ Genomics and Health Unit, FISABIO Public Health, 46021, Valencia, Spain.
⁴ Pulmonology Department, Consorci Hospital General Universitari de València, Valencia, Spain.
⁵ Servicio de Microbiología, Consorci Hospital General Universitari de València, Valencia, Spain.
⁶ CIBER of Epidemiology and Public Health (CIBERESP), Madrid, Spain.

a corresponding author. icomas@ibv.csic.es
# Table of Contents

- Case study: 3
- Supplementary Methods: 5
- Supplementary References: 9
- Supplementary Figures: 11
- Supplementary Tables: 14
  - Supplementary Table 1: 14
  - Supplementary Table 2: 32
  - Supplementary Table 3: 34
  - Supplementary Table 4: 35
  - Supplementary Table 5: 36
Case study

A 43-year old male patient with no tuberculosis (TB) clinically records was admitted to the Hospital Universitario General de Valencia, Spain in April of 2009. The patient did not present any comorbidity or risk factor that supposed a bad therapeutic compliance (e.g HIV negative status). The patient was first diagnosed with pulmonary TB and treated with the standard first-line therapy for active TB (Two months with rifampicin [RIF], isoniazid [INH], ethambutol [EMB] and pyrazinamide [PZA], followed by four months of RIF and INH only), adjusting for body weight. During this period, serial sputum-positive samples was collected and culture in order to 1) confirm the presence of Mycobacterium tuberculosis Complex (MTBC) bacilli and 2) to performed drug susceptibility testing (DST). MTB identification was carried out using the commercial kit BACTEC MGIT Tbc Identification Test (Becton Dickinson and Co, Franklin Lakes, New Jersey). First-line DST was performed using rapid phenotypic BACTEC MGIT 960 system (Becton Dickinson and Co, Franklin Lakes, New Jersey) and genotypic probes (GenoType MTBDR plus, Hain Lifescience, Nehren, Germany). DST results indicated that the isolate was fully-susceptible to all first-line antibiotics. After two months of treatment, the patient became sputum smear and culture negative. Consecutive negative samples for next four months (n=4) confirmed that the patient was cured. During 2010 and 2012 the patient presented negative sputum smear and was discharged.

In September 2013, the patient was readmitted to the hospital because he presented typical TB symptoms suggesting a relapse episode. MTBC was detected in sputum smear samples and confirmed in by rapid culture-based immunoassay (BACTEC MGIT Tbc Identification Test). Relapse disease has been related with the infection by an MDR TB strain. So, phenotypic and genoptypic DST was performed. Again, phenotypic DST and line-probe assays showed no evidence of drug resistance. With this scenario, the patient was treated with the same first-line antituberculous therapy maintained all effective drugs. Six months later (March 2014) the patient became sputum and culture-negative and once again, was considered cured.

Unfortunately, in June 2014, a single sputum specimen was detected as an INH mono-resistant isolate by phenotypic BACTEC MGIT 960 system but no mutation was detected by genotypic Hain line-probes assay. Because the infection strains was considered INH mono-resistant and there were more four effective antituberculous agents, the drug remigen was prolonged another nine months. Notably, during this time, phenotypic DST from all the clinical culture-positive samples tested negative for resistance to first-line drugs. Due to the first episode and the continuum culture-positive sputum status during the relapse, the therapy was extended until December 2015. As a consequence of the long and ineffective TB regimen, the patient developed a great cavity in the right upper lung.

In December 2015, whole genome sequencing (WGS) was performed for a culture derived from a positive sputum sample of the patient isolated in June of 2014. WGS analysis identified mutations in the rpoB and katG genes conferring RIF and INH resistance (rpoB I491F and katG G273R, respectively), which indicated an infection caused by probable a multidrug resistant (MDR) strain. We corroborated and validated RIF and INH resistance by extensive phenotypic DST methods (Proportions method in the case of the RIF and REMA assay for INH, see Supplementary Methods for a detailed description). With this result, the antituberculos therapy was changed, and second-line drugs were added following the recommendations by the WHO guidelines[1]. This regimen included the administration of EMB, Moxifloxacin (MFX), PZA, higher doses of INH (INH+) and the injectable agent Capreomycin (CP). We decided to use CP instead of other aminoglycoside because of the greater ease of intramuscular administration. From this point, WGS was used in retrospective manner to genomic characterized all the stored clinical samples from the patient, since the first TB episode in 2009. In addition, WGS was used prospective manner toghether with hospital phenotypic DST results to guide the antibiotic regimen in the patient.
On June 2016, after six months of the MDR treatment, a new sputum smear and culture-positive appeared. WGS analysis of this sample revealed the presence of mutations conferring EMB, MFX, and high likely CP drug resistance (embB G238Y, gyrB E540D and tlyA L16R, respectively), indicating that TB infecting isolate had evolved to an extensively drug-resistant (XDR) strain. Phenotypic DST performed in the hospital (Sensititre MYCOTB MIC Plate, Trek Diagnostics System, Cleveland, OH) confirmed MFX drug resistance (MIC 4µg/ml) and susceptibility to amikacin (MIC ≤1µg/ml). Based on these findings and following the WHO recommendations [1], we decided to change the drug treatment in the same month adding the second-line amikacin and linezolid agents, INH+ and PZA. In addition, surgical intervention was proposed but there were several cavited foci in both lungs and the patient rejected this option. The last culture-positive sputum was detected in February 2017. A year later, amikacin was removed from the therapy due to the patient had suffered adverse effects such as hearing and sight loss.

It is notable that PZA resistance was no detected by phenotypic (BACTEC MGIT 960 system) and genotypic (WGS analysis) methods during the nine years of the disease.
SUPPLEMENTARY METHODS

Isolate collection and DNA extraction

For this study, we used sixteen serial clinical isolates from a single patient and one epidemiologically related sample, all supplied by the Hospital General de Valencia. The single-patient samples covered a total of 9 years and included a first episode of active TB and a relapse incident years later (Supplementary Table 1). Positive mycobacterial growth indicator tubes (MGIT) were subcultured in Middlebrook 7H11 agar plates supplemented with 10% OADC (Becton-Dickinson) for 5 weeks at 37°C. We scrapped bacteria with a sterile loop four times across the plate to obtain a representative sample of the population and we extracted their DNA using the CTAB protocol [2]. The rest of the bacteria was also scrapped and stored in 1ml of glycerol (20%) at -80°C. Because of the high risk to manage and work with XDR strains, isolates G1257, G1720 and G1928 were directly extracted from the positive MGIT culture.

Routine microbiological diagnostics methods

Standard phenotypic DST for the first-line anti tuberculosis drugs (and linezolid [LZD] in one isolate) was performed in all clinical samples during whole study period using the BACTEC MGIT 960 system (Becton Dickinson and Co, Franklin Lakes, New Jersey) following the manufacturer recommended critical concentrations for each drug; RIF (1.0µg/ml), INH (0.1µg/ml), ethambutol (EMB) (5.0µg/ml), pyrazinamide (PZA) (100µg/ml), streptomycin STR (1.0µg/ml) and LZD (1.0µg/ml). In addition, genotypic line-probes assays for RIF and INH resistance was carried out with the commercial kit GenoType MTBDR plus (Hain Lifescience, Nehren, Germany). Finally, Sensititre MYCOTB MIC Plate (Trek Diagnostics System, Cleveland, OH) was applied within a few clinical samples to determine the minimal inhibitory concentrations (MIC) for some second-line antibiotics such as moxifloxacin (MFX), ethionamide (ETH) and amikacin (AMK). The following range of concentrations were used; MFX, (0.06-8µg/ml); ETH, (0.3-40µg/ml); AMK, (0.12-16µg/ml).

Whole genome sequencing analysis

DNA from diagnostic cultures was extracted as previously described [3]. Sequencing libraries were constructed with Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA) following the manufacturer instructions. WGS was performed on the Illumina MiSeq instrument and the average sequencing depth value per base was 168-fold (range: 103-268). To account for contaminant DNA, we identified reads only belonging to the Mycobacterium tuberculosis Complex (MTBC) using KRAKEN software [4]. Mapping and single nucleotide polymorphism (SNP) calling was performed following a previous pipeline [5]. Briefly, MTBC reads were mapped and aligned to an inferred MTBC most likely common ancestor genome using BWA [6]. Next, we separated variants into INDELS (small insertions and deletions included) and SNPs. Single polymorphisms with at least 10 reads in both strands and a quality score of 20 were selected and classified into two categories based on their frequency in the sample. We considered fixed SNPs those with no less than 90% of frequency and low-frequency SNPs those involving variants whose frequencies range between 10 and 89%. An INDEL was considered whether the mutation was present with a minimum deep coverage of 10x. SNP annotation was performed using H37Rv annotation reference (AL123456.2). Finally, SNPs annotated in regions difficult to map such as repetitive sequences and PPE/PE-PGRS genes were removed from the analysis as well as those detected near INDELS.

Identification of relapse versus re-infection strains
In order to identify whether the isolates from the first and second disease episode were the same (relapse) or coming from two independent infections (re-infection) we applied both a phylogenetic and a SNP threshold approach. Regarding the SNP threshold approach, pairwise genetic distances were compared based on a concatenated SNP alignment obtained from every fixed SNP from every isolate. A genetic distance below 12 SNPs is indicative of clonal diversification from the first episode strain and thus is classified as a relapse case. On the contrary distances beyond 12 SNPs are indicative of exogenous infection with a different strain and thus the case is classified as likely re-infection [7,8]. To differentiate phylogenetically between relapse and re-infection we built a phylogeny including all the isolates from the first and second episode as well as epidemiological related sample (brother). The phylogeny was inferred following maximum likelihood phylogenetic approach using RAxML v8.2 [9] applying the General Time Reversible model of nucleotide substitution with the Gamma distribution (GTRGAMMA).

Identification of known lineage and drug resistance associated mutations

Once we had determined all the polymorphic sites (including fixed and variable SNPs), we compared our SNP and INDEL data with a series of publicly available predictive databases (Phyresee [10], Mykrobe predictor [11], TB-profiler [12], latest accessed on April 2018) to determine lineage and drug resistance mutations present in our samples. Isolates were classified as susceptible, resistant or multidrug resistant (MDR) according to the mutations identified.

Identification of candidate drug resistance variants

Relevant variants during the course of treatment were selected if they appeared at least in one isolate with a minimum frequency of 15% and had at least 15% of difference between any two isolates (n=96, see Supplementary Table 2). Global frequency of the new candidate mutations was evaluated against a database of 4,762 genomes collected world-wide. The same database was used to identify strains with and without “disputed” rpoB mutations as well as new drug resistance determinants to INH in katG gene.

Isolation of single colonies and drug susceptibility testing of new drug resistance mutations

After the identification of a series of novel mutations in genes or genomic regions associated with drug resistance we decided to explicitly test if those mutations conferred resistance by phenotypically and genetically characterizing a series of single-colony isolates. Twenty-two single colonies from six clinical samples with different co-existing haplotypes were isolated from Middlebrook 7H11 agar plates, grown for 10 to 15 days in Middlebrook 7H9 at 37°C and stored in 500µl aliquotes with 20% glycerol at -80°C until used.

Antibiotic stock solutions were prepared at 10 mg/ml either in sterile water (INH) or methanol (RIF, both from Sigma-Aldrich) and stored at -20°C. Nine days before the experiments, 500µl of frozen inoculum from each of the 22 colonies were cultured in 10ml of Middlebrook 7H9 broth supplemented with 10% ADC (Becton-Dickinson) and 0.1% tween80. Exponentially-growing bacteria were adjusted to an optical density of 0.2 (which is equivalent to a 1 McFarcland turbidity standard) in tween-free Middlebrook 7H9 broth immediately prior to use.

We obtained the MIC for INH following the resazurin microtiter assay (REMA) protocol [13]. Briefly, serial two-fold dilutions ranging from 0.125 to 32µg/ml of INH in tween-free 7H9 broth were prepared in a 96-well plate, with a volume of 100µl. We then added 100µl of a 1:100 dilution of a fresh density adjusted bacterial suspension, prepared as above and sealed the plate in a hermetic plastic bag. Because of the slow growth of the samples carried the rpoB I491F mutations, after 9 days of incubation at 37°C, 30µl of 0.02% resazurin was added to each well
and the plate was further incubated at 37ºC for 48 h. A change of color from blue to pink indicates bacterial growth. The MIC was defined as the lowest drug concentration that prevented this color change. Drug and bacteria free wells were used as control.

To determine phenotypic resistance to RIF, we used the proportion method (PM) in Middlebrook 7H11 agar according to standard procedures [14]. The reason for using this particular method is that these strains grow very slowly in RIF-containing broth and give inconsistent results with the REMA. In brief, 50µl of a 10⁻² and a 10⁻⁴ dilutions of a fresh density-adjusted bacterial suspension were cultured in plates containing 0.125 to 64µg/ml of RIF. Drug-free plates were used as a positive control. After 3 weeks of incubation at 37ºC, the number of colonies forming units growing on medium with antibiotic were compared with those on the positive control. The proportion of resistant bacteria was represented as percentage. The MIC was defined as the first drug concentration where there was no growth.

The H37Rv strain was used as a quality control in all experiments. In addition, the isolate from the first episode that is wild-type (G1480) for the mutations was used to measure the increase on MIC in the following isolates. Strains were classified as susceptible or resistant according to the critical concentrations recommended by WHO (guidelines 2012).

**Deep amplicon sequencing: amplification, library construction and sequencing**

In order to confirm the SNPs identified and the possible detection of additional very low-frequency variants (<5%), we performed deep amplicon sequencing following and adjusting the single molecule-overlapping reads (SMOR) approach [15]. Six primer sets were designed to target specific regions of *rpoB*, *katG* and *ahpC* (promoter and CDS included) (see supplementary Table 5), producing amplicons between 285-339bp. We used two additional primers sets targeting phylogenetic lineage-diagnostic SNPs as internal controls. A total of 39 samples were analysed including 15 clinical isolates from the case, 22 from single colonies and two controls (lineage 4 H37Ra and a Lineage 2 Beijing strain). A single PCR amplification step was carried out with parameters: initial denaturation at 95ºC for 3min, 20 cycles of denaturation at 98ºC for 15 sec, annealing at 65ºC for 15s and extension at 72ºC for 30s, with a final extension at 72ºC for 2min. Each reaction contained 12.5µl of 2x KAPA HiFi HotStart ReadyMix (KAPA biosystems), 0.75µl of 10µM forward primer, 0.75µl of 10µM reverse primer, 5 µl of template DNA and 6µl of PCR-grade water. After amplification, reactions were pooled by sample (5µl of each amplicon and 20µl of 10µM Tris-HCl) and purified using 1X NucleoMag NGS Clean-up and Size Select (Macherey-Nagel) up to 50µl final volume.

Amplicon sequencing libraries construction and sequencing were performed as described for whole genome sequencing. Purified libraries were validated on a Bioanalyzer DNA chip (Agilent Technologies) to verify fragment size, and quantified using Qubit 3.0 Fluorometer (Thermo Fisher Scientific). The expected average coverage for this experiment was 50,000 fold per base.

**Data availability**

All genomic data are deposited in the European Nucleotide Archive under the Bioproject numbers PRJEB22237 and PRJEB25887.
SUPPLEMENTARY RESULTS

Non-canonical resistance mutations can lead to under-reporting and treatment of MDR-TB cases

The fact that INH resistance went undetected (likely due to the fitness cost of the RIF mutation on replication [16]) led us to question whether non-canonical (also known as “disputed”) mutations are involved in a systematic under detection of MDR-TB. We pooled together a global SNP database of 4762 strains. A total of 66 strains harbored a non-canonical RIF resistance mutation (7.6% of the strains with a RIF resistance mutation). Of these, only 24 isolates contained a known katG mutation as opposed to 92.4% in strains carrying undisputed RIF resistance mutations (36%, P<0.001, chi-square test). Some of the strains could be RIF mono-resistant, particularly related with relapse cases [17,18]. However, we hypothesized that a percentage of these strains were undetected MDR cases associated to non-canonical mutations. Among strains with a “non-canonical” rpoB mutation we found seven mutations not described before, all of them leading to an aminoacid change. For three of them we had evidence of convergent evolution, a strong predictor of resistance particularly in drug resistance associated regions [19,20]. One, katG V1A have been recently identified as INH associated with resistance [21]. For the other two (V473L and G285V), publicly available phenotypic results are contradictory as the majority of the isolates were susceptible. However, rpoB I491F mutation was present in four out of the six strains indicating a possible parallelism with the phenotyping problems presented in this manuscript.
SUPPLEMENTARY REFERENCES

1. World Health Organization, World Health Organization. Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis. [Internet]. 2014 [cited 2019 Jan 25]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK247420/

2. Soolingen D van, Hermans PW, De Haas PE, Soll DR, Van Embden JD. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J Clin Microbiol. 1991; 29(11):2578–2586.

3. SOOLINGEN DV, HERMANS PWM, DE PEW, SOLL DR, VAN JDA. Occurrence and Stability of Insertion Sequences in *Mycobacterium tuberculosis* Complex Strains: Evaluation of an Insertion Sequence-Dependent DNA Polymorphism as a Tool in the Epidemiology of Tuberculosis. 1991; 29:9.

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

5. Comas I, Coscolla M, Luo T, et al. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. Nat Genet. 2013; 45(10):1176–1182.

6. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv Prepr ArXiv13033997 [Internet]. 2013 [cited 2017 Sep 5]; . Available from: https://arxiv.org/abs/1303.3997

7. Walker TM, Ip CL, Harrell RH, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. Lancet Infect Dis. 2013; 13(2):137–146.

8. Guerra-Assunção JA, Houben RMGJ, Crampin AC, et al. Recurrence due to Relapse or Reinfection With *Mycobacterium tuberculosis*: A Whole-Genome Sequencing Approach in a Large, Population-Based Cohort With a High HIV Infection Prevalence and Active Follow-up. J Infect Dis. 2015; 211(7):1154–1163.

9. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30(9):1312–1313.

10. Feuerriegel S, Schleusener V, Beckert P, et al. PhyResSE: a Web Tool Delineating *Mycobacterium tuberculosis* Antibiotic Resistance and Lineage from Whole-Genome Sequencing Data. Carroll KC, editor. J Clin Microbiol. 2015; 53(6):1908–1914.

11. Bradley P, Gordon NC, Walker TM, et al. Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. Nat Commun [Internet]. 2015 [cited 2018 Jan 30]; 6(1). Available from: http://www.nature.com/articles/ncomms10063

12. Coll F, McNerney R, Preston MD, et al. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. Genome Med [Internet]. 2015 [cited 2018 Jan 30]; 7(1). Available from: http://genomemedicine.com/content/7/1/51

13. Palomino J-C, Martin A, Camacho M, Guerra H, Swings J, Portaels F. Resazurin Microtiter Assay Plate: Simple and Inexpensive Method for Detection of Drug Resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2002; 46(8):2720–2722.
14. Canetti G, Fox W, Khomenko A al, et al. Advances in techniques of testing mycobacterial drug sensitivity, and the use of sensitivity tests in tuberculosis control programmes. Bull World Health Organ. 1969; 41(1):21.

15. Colman RE, Schupp JM, Hicks ND, et al. Detection of Low-Level Mixed-Population Drug Resistance in *Mycobacterium tuberculosis* Using High Fidelity Amplicon Sequencing. Metcalfe JZ, editor. PLOS ONE. 2015; 10(5):e0126626.

16. Miotto P, Cabibbe AM, Borroni E, Degano M, Cirillo DM. Role of disputed mutations in the *rpoB* gene in the interpretation of automated liquid MGIT culture results for rifampicin susceptibility testing of *Mycobacterium tuberculosis*. J Clin Microbiol. 2018; JCM.01599-17.

17. Meyssonnier V, Bui TV, Veziris N, Jarlier V, Robert J. Rifampicin mono-resistant tuberculosis in France: a 2005–2010 retrospective cohort analysis. BMC Infect Dis [Internet]. 2014 [cited 2018 Oct 3]; 14(1). Available from: http://bmcinfectdis.biomedcentral.com/articles/10.1186/1471-2334-14-18

18. Li J, Munsiff SS, Driver CR, Sackoff J. Relapse and Acquired Rifampin Resistance in HIV-Infected Patients with Tuberculosis Treated with Rifampin- or Rifabutin-Based Regimens in New York City, 1997-2000. Clin Infect Dis. 2005; 41(1):83–91.

19. Comas I, Borrell S, Roetzer A, et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. Nat Genet. 2012; 44(1):106–110.

20. Farhat MR, Shapiro BJ, Kieser KJ, et al. Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*. Nat Genet. 2013; 45(10):1183–1189.

21. Kandler JL, Mercante AD, Dalton TL, et al. Validation of Novel *Mycobacterium tuberculosis* Isoniazid Resistance Mutations Not Detectable by Common Molecular Tests. Antimicrob Agents Chemother. 2018; 62(10):16.

22. Homolka S, Meyer CG, Hillemann D, et al. Unequal distribution of resistance-conferring mutations among *Mycobacterium tuberculosis* and *Mycobacterium africanum* strains from Ghana. Int J Med Microbiol. 2010; 300(7):489–495.
SUPPLEMENTARY FIGURES
Supplementary Figure 1. Management of the case and role of the genomic and clinical laboratories.
Timeline with the most relevant patient’s management clinical events during 9 years of the infection. The clinical laboratory (red line) refers to the routine BACTEC-MGIT (phenotypic) and Hain Gentotype MTBDRplus (genotypic) assays performed in the hospital. The genomic laboratory (green lines) corresponds to results obtained by whole genome sequencing analysis of the isolates and available from 2015 onwards. Once the relevant mutations were identified using WGS individual clones at different time-points were isolated and tested for RIF and INH resistance using an alternative approach to BACTEC MGIT. WGS was also used in a prospective manner to predict DST to new drugs during MDR-TB treatment.
Supplementary Figure 2. Genetic relationship between all the isolates. Sample G1480 corresponds to the first episode and presented a fully-susceptible resistance profile; G1479 corresponds to the first relapse isolate. A. Inferred maximum Likelihood tree. Isolates G1720, G1721 and G1928 presented an extensively drug-resistant profile. B. Pairwise SNP distanced between all samples.
Supplementary Figure 3. Predicted percentage of the susceptible (grey bar) versus INH resistant populations (blue bar) identified across all isolates.

The percentage of the INH resistant population refers to the sum of the frequencies of the two INH mutations co-existing in the patient (\textit{katG} G273 and G249del).
**SUPPLEMENTARY TABLES**

Supplementary Table 1. List of loci associated with drug resistance used to predict resistance profile.

| Genomic position | Wild-type allele | Mutation allele | Genomic Region | Gene Name | Gene Alias | Gene Direction | Amino acid Change | Codon Change | Antibiotic | SNP Confidence* |
|------------------|------------------|-----------------|----------------|-----------|------------|----------------|------------------|--------------|------------|-----------------|
| 6575             | C                | T               | coding         | Rv0005    | gyrB       | +             | Arg446Cys        | cgt/tgt      | FQ         | Low confidence  |
| 6620             | G                | C               | coding         | Rv0005    | gyrB       | +             | Asp461His        | gac/cac      | FQ         | High confidence |
| 6620             | G                | A               | coding         | Rv0005    | gyrB       | +             | Asp461Asn        | gac/aac      | FQ         | High confidence |
| 6621             | A                | C               | coding         | Rv0005    | gyrB       | +             | Asp461Ala        | gac/gcc      | FQ         | High confidence |
| 6734             | A                | G               | coding         | Rv0005    | gyrB       | +             | Asn499Asp        | aac/gac      | FQ         | High confidence |
| 6735             | A                | C               | coding         | Rv0005    | gyrB       | +             | Asn499Thr        | aac/acc      | FQ         | High confidence |
| 6736             | C                | G               | coding         | Rv0005    | gyrB       | +             | Asn499Lys        | aac/aag      | FQ         | High confidence |
| 6737             | A                | C               | coding         | Rv0005    | gyrB       | +             | Thr500Pro        | acc/ccc      | FQ         | High confidence |
| 6738             | C                | A               | coding         | Rv0005    | gyrB       | +             | Thr500Asn        | acc/aac      | FQ         | High confidence |
| 6741             | A                | T               | coding         | Rv0005    | gyrB       | +             | Glu501Val        | gaa/gta      | FQ         | High confidence |
| 6742             | A                | T               | coding         | Rv0005    | gyrB       | +             | Glu501Asp        | gaa/gat      | FQ         | High confidence |
| 6749             | G                | A               | coding         | Rv0005    | gyrB       | +             | Ala504Thr        | gcg/acg      | FQ         | Low confidence  |
| 6750             | C                | T               | coding         | Rv0005    | gyrB       | +             | Ala504Val        | gcg/gtg      | FQ         | High confidence |
| 7563             | G                | T               | coding         | Rv0006    | gyrA       | +             | Gly88Cys         | ggc/tgc      | FQ         | High confidence |
| 7564             | G                | C               | coding         | Rv0006    | gyrA       | +             | Gly88Ala         | ggc/ggc      | FQ         | High confidence |
| 7566             | G                | A               | coding         | Rv0006    | gyrA       | +             | Asp89Asn         | gac/aac      | FQ         | Low confidence  |
| 7570             | C                | T               | coding         | Rv0006    | gyrA       | +             | Ala90Val         | gcg/gtg      | FQ         | High confidence |
| 7572             | T                | C               | coding         | Rv0006    | gyrA       | +             | Ser91Pro         | tgc/ccg      | FQ         | High confidence |
| 7581             | G                | C               | coding         | Rv0006    | gyrA       | +             | Asp94His         | gac/cac      | FQ         | High confidence |
| 7581             | G                | A               | coding         | Rv0006    | gyrA       | +             | Asp94Asn         | gac/aac      | FQ         | High confidence |
| 7581             | G                | T               | coding         | Rv0006    | gyrA       | +             | Asp94Tyr         | gac/tac      | FQ         | High confidence |
| 7582             | A                | G               | coding         | Rv0006    | gyrA       | +             | Asp94Gly         | gac/ggc      | FQ         | High confidence |
| 7582             | A                | C               | coding         | Rv0006    | gyrA       | +             | Asp94Ala         | gac/gcc      | FQ         | High confidence |
| 7582             | A                | T               | coding         | Rv0006    | gyrA       | +             | Asp94Val         | gac/gtc      | FQ         | Low confidence  |
| Position | DNA Change | Protein Change | Gene | Gene Product | Codon Change | Drug Resistance | Confidence |
|----------|------------|----------------|------|--------------|--------------|----------------|------------|
| 575729   | C T coding | mshA + Gln128STOP cag/tag | Rv0486 | ETH | Low confidence |
| 576164   | C T coding | mshA + Arg273Cys cgc/tgc | Rv0486 | ETH | Low confidence |
| 576242   | G T coding | mshA + Gly299Cys ggc/tgc | Rv0486 | ETH | Low confidence |
| 576338   | C T coding | mshA + Gln331STOP cag/tag | Rv0486 | ETH | Low confidence |
| 576414   | G A coding | mshA + Gly356Asp ggc/gac | Rv0486 | ETH | Low confidence |
| 576429   | A C coding | mshA + Glu361Ala gag/gcg | Rv0486 | ETH | Low confidence |
| 760314   | G T coding | rpoB + Val170Phe gtc/ttc | Rv0667 | RIF | High confidence |
| 761004   | A G coding | rpoB + Thr400Ala acc/gcc | Rv0667 | RIF | Low confidence |
| 761093   | G C coding | rpoB + Gln429His cag/cac | Rv0667 | RIF | Low confidence |
| 761095   | T C coding | rpoB + Leu430Phe gtc/ttc | Rv0667 | RIF | Low confidence |
| 761095   | T G coding | rpoB + Leu430Arg acc/gcc | Rv0667 | RIF | Low confidence |
| 761098   | G T coding | rpoB + Ser431Ile agc/atc | Rv0667 | RIF | Low confidence |
| 761098   | G C coding | rpoB + Ser431Thr agc/acc | Rv0667 | RIF | Low confidence |
| 761100   | C A coding | rpoB + Gln432Lys caa/aaa | Rv0667 | RIF | Low confidence |
| 761101   | A T coding | rpoB + Gln432Leu caa/taa | Rv0667 | RIF | Low confidence |
| 761101   | A C coding | rpoB + Gln432Pro caa/cca | Rv0667 | RIF | Low confidence |
| 761108   | G T coding | rpoB + Met434Ile att/att | Rv0667 | RIF | Low confidence |
| 761109   | G T coding | rpoB + Asp435Glu gac/gag | Rv0667 | RIF | Low confidence |
| 761110   | A G coding | rpoB + Asp435Gly gac/ggc | Rv0667 | RIF | High confidence |
| 761110   | A T coding | rpoB + Asp435Val gag/gtc | Rv0667 | RIF | High confidence |
| 761111   | C G coding | rpoB + Asp435Glu gac/gag | Rv0667 | RIF | Low confidence |
| 761120   | C G coding | rpoB + Asn438Lys aac/aag | Rv0667 | RIF | Low confidence |
| 761128   | C T coding | rpoB + Ser441Leu tgc/ttg | Rv0667 | RIF | Low confidence |
| 761128   | C G coding | rpoB + Ser441Trp tgg/tgg | Rv0667 | RIF | Low confidence |
| 761139   | C A coding | rpoB + His445Asn cac/aaac | Rv0667 | RIF | High confidence |
| 761139   | C G coding | rpoB + His445Asp cac/gac | Rv0667 | RIF | High confidence |
| 761139   | C T coding | rpoB + His445Tyr cac/tac | Rv0667 | RIF | High confidence |
| 761140   | A C coding | rpoB + His445Pro cac/ccc | Rv0667 | RIF | High confidence |
A  G  coding  Rv0667  rpoB  +  His445Arg  cac/cgc  RIF  High confidence
A  T  coding  Rv0667  rpoB  +  His445Leu  cac/ctc  RIF  Low confidence
C  A  coding  Rv0667  rpoB  +  His445Gln  cac/caa  RIF  Low confidence
T  G  coding  Rv0667  rpoB  +  Ser450Ala  tgc/gcg  RIF  Low confidence
C  G  coding  Rv0667  rpoB  +  Ser450Trp  tgc/tgg  RIF  High confidence
C  T  coding  Rv0667  rpoB  +  Ser450Leu  tgc/ttg  RIF  High confidence
T  C  coding  Rv0667  rpoB  +  Ile491Phe  atc/ttc  RIF  High confidence
A  C  coding  Rv0667  rpoB  +  Lys43Arg  aag/agg  RIF  High confidence
A  T  coding  Rv0667  rpsL  +  Lys88Gln  aag/cag  RIF  Low confidence
A  G  coding  Rv0667  rpsL  +  Lys88Arg  aag/agg  RIF  High confidence
T  C  coding  Rv0701  rplC  +  Cys154Arg  tgt/cgt  LZD  Low confidence
C  T  ribosomal  MTB000019  rrs  +  ---  -  STR  Low confidence
C  T  ribosomal  MTB000019  rrs  +  ---  -  STR  Low confidence
A  C  ribosomal  MTB000019  rrs  +  ---  -  STR  Low confidence
C  T  ribosomal  MTB000019  rrs  +  ---  -  STR  Low confidence
C  A  ribosomal  MTB000019  rrs  +  ---  -  STR  Low confidence
A  G  ribosomal  MTB000019  rrs  +  ---  -  STR  Low confidence
A  T  ribosomal  MTB000019  rrs  +  ---  -  STR  Low confidence
A  G  ribosomal  MTB000019  rrs  +  ---  -  AMK, KAN, CM  High confidence
C  T  ribosomal  MTB000019  rrs  +  ---  -  AMK, KAN, CM  High confidence
G  T  ribosomal  MTB000019  rrs  +  ---  -  AMK, KAN, CM  High confidence
G  T  ribosomal  MTB000020  rrl  +  ---  -  LZD  Low confidence
G  T  ribosomal  MTB000020  rrl  +  ---  -  LZD  Low confidence
G  T  intergenic  Rv1483  fabG1  +  ---  -  INH  Low confidence
G  T  intergenic  Rv1483  fabG1  +  ---  -  INH  Low confidence
T  A  intergenic  Rv1483  fabG1  +  ---  -  INH  High confidence
| Gene      | Start | End | Status  | codon       | Amino Acid | Protein | Confidence |
|-----------|-------|-----|---------|-------------|------------|---------|------------|
| fabG1     |       |     | Intergenic |             |            |         |            |
| inhA      |       |     | Coding   |             |            |         | High confidence |
| rpsA      |       |     | Coding   |             |            |         | Low confidence |
| tlyA      |       |     | Coding   |             |            |         | Low confidence |
| ndh       |       |     | Coding   |             |            |         | Low confidence |
| katG      |       |     | Coding   |             |            |         | Low confidence |

Notes:
- **Intergenic** indicates that the gene is located between two other genes.
- **Coding** indicates that the gene codes for a protein.
- **INH** indicates that the mutation is recognized by INH.
- **ETH** indicates that the mutation is recognized by ETH.
- **CM** indicates that the mutation is recognized by CM.
| Gene   | Position | Change | Amino Acid Change | Confidence | Antimicrobial | 
|--------|----------|--------|-------------------|------------|---------------| 
| katG   | 1908c    | G > C  | Ser302Arg         | Low        | INH           | 
| katG   | 1908c    | G > C  | Trp300Cys         | Low        | INH           | 
| katG   | 1908c    | G > C  | Trp300Gly         | High       | INH           | 
| katG   | 1908c    | A > C  | Gly297Val         | Low        | INH           | 
| katG   | 1908c    | A > C  | Thr275Pro         | Low        | INH           | 
| katG   | 1908c    | A > C  | Asn138Ser         | Low        | INH           | 
| pncA   | 2043c    | T > C  | STOP187Arg        | High       | PZA           | 
| pncA   | 2043c    | T > C  | Val180Gly         | High       | PZA           | 
| pncA   | 2043c    | T > C  | Val180Phe         | Low        | PZA           | 
| pncA   | 2043c    | T > C  | Met175Thr         | High       | PZA           | 
| pncA   | 2043c    | A > G  | Leu172Pro         | Low        | PZA           | 
| pncA   | 2043c    | A > G  | Val175Val         | High       | PZA           | 
| pncA   | 2043c    | A > G  | Met175Thr         | High       | PZA           | 
| pncA   | 2043c    | T > C  | Leu172Pro         | Low        | PZA           | 
| pncA   | 2043c    | T > C  | Ala171Val         | Low        | PZA           | 
| pncA   | 2043c    | T > C  | Thr168Pro         | High       | PZA           | 
| pncA   | 2043c    | T > C  | Ser164Pro         | High       | PZA           | 
| pncA   | 2043c    | T > C  | Val163Ala         | High       | PZA           | 
| pncA   | 2043c    | G > A  | Gly162Asp         | Low        | PZA           | 
| pncA   | 2043c    | G > A  | Ala161Val         | High       | PZA           | 
| pncA   | 2043c    | A > C  | Thr160Pro         | High       | PZA           | 
| pncA   | 2043c    | T > G  | Leu159Arg         | Low        | PZA           | 
| pncA   | 2043c    | T > G  | Val157Ala         | High       | PZA           | 
| pncA   | 2043c    | T > G  | Val157Gly         | High       | PZA           | 
| pncA   | 2043c    | T > G  | Val155Gly         | High       | PZA           | 
| pncA   | 2043c    | G > A  | Arg154Gly         | High       | PZA           | 
| pncA   | 2043c    | C > T  | Ala146Val         | High       | PZA           |
| Position | Gene | Codon Change | Amino Acid Change | Codon | Protein Effect | Confidence |
|----------|------|--------------|------------------|-------|----------------|------------|
| 2288805 | C    | A            | coding Rv2043c pmcA | - | Ala146Glu | gcg/gag | PZA | Low confidence |
| 2288806 | G    | C            | coding Rv2043c pmcA | - | Ala146Pro | gcg/ccg | PZA | High confidence |
| 2288806 | G    | A            | coding Rv2043c pmcA | - | Ala146Thr | gcg/acg | PZA | Low confidence |
| 2288817 | C    | A            | coding Rv2043c pmcA | - | Thr142Lys | agc/aag | PZA | High confidence |
| 2288817 | C    | T            | coding Rv2043c pmcA | - | Thr142Met | acg/atg | PZA | High confidence |
| 2288818 | A    | G            | coding Rv2043c pmcA | - | Thr142Ala | agc/ccg | PZA | High confidence |
| 2288820 | A    | C            | coding Rv2043c pmcA | - | Gln141Pro | cag/ccg | PZA | Low confidence |
| 2288821 | C    | T            | coding Rv2043c pmcA | - | Gln141STOP | cag/tag | PZA | Low confidence |
| 2288823 | G    | C            | coding Rv2043c pmcA | - | Arg140Pro | cgc/ccc | PZA | High confidence |
| 2288826 | T    | G            | coding Rv2043c pmcA | - | Val139Gly | gtg/ggg | PZA | High confidence |
| 2288827 | G    | C            | coding Rv2043c pmcA | - | Val139Leu | gtt/ctg | PZA | High confidence |
| 2288828 | T    | G            | coding Rv2043c pmcA | - | Cys138Trp | tgt/tgg | PZA | High confidence |
| 2288830 | T    | C            | coding Rv2043c pmcA | - | Cys138Arg | tgt/cgt | PZA | High confidence |
| 2288832 | A    | G            | coding Rv2043c pmcA | - | His137Arg | cat/cgt | PZA | High confidence |
| 2288832 | A    | C            | coding Rv2043c pmcA | - | His137Pro | cat/ctg | PZA | High confidence |
| 2288833 | C    | G            | coding Rv2043c pmcA | - | His137Asp | cat/gat | PZA | High confidence |
| 2288836 | G    | A            | coding Rv2043c pmcA | - | Asp136Asn | gat/aat | PZA | Low confidence |
| 2288836 | G    | T            | coding Rv2043c pmcA | - | Asp136Tyr | gat/tat | PZA | Low confidence |
| 2288838 | C    | A            | coding Rv2043c pmcA | - | Thr135Asn | acc/aac | PZA | High confidence |
| 2288839 | A    | C            | coding Rv2043c pmcA | - | Thr135Pro | acc/ccc | PZA | High confidence |
| 2288841 | C    | T            | coding Rv2043c pmcA | - | Ala134Val | gcc/gtc | PZA | High confidence |
| 2288844 | T    | C            | coding Rv2043c pmcA | - | Ile133Thr | att/act | PZA | Low confidence |
| 2288847 | G    | C            | coding Rv2043c pmcA | - | Gly132Ala | ggt/gtc | PZA | High confidence |
| 2288847 | G    | A            | coding Rv2043c pmcA | - | Gly132Asp | ggt/gat | PZA | High confidence |
| 2288848 | G    | T            | coding Rv2043c pmcA | - | Gly132Cys | ggt/tgt | PZA | High confidence |
| 2288848 | G    | A            | coding Rv2043c pmcA | - | Gly132Ser | ggt/agt | PZA | Low confidence |
| 2288853 | T    | C            | coding Rv2043c pmcA | - | Val130Ala | gtt/ccg | PZA | High confidence |
| 2288853 | T    | G            | coding Rv2043c pmcA | - | Val130Gly | gtt/ggg | PZA | High confidence |
| Position | Chrom | Codon | Gene | Codon Position | Amino Acid Change | Nucleotide Change | Antimicrobial | Confidence |
|----------|-------|-------|------|----------------|------------------|------------------|--------------|------------|
| 2288857  | G     | T     | coding | Rv2043c | pncA | Asp129Tyr | gat/tat | PZA | High confidence |
| 2288859  | T     | G     | coding | Rv2043c | pncA | Val128Gly | gtc/ggc | PZA | Low confidence |
| 2288868  | T     | G     | coding | Rv2043c | pncA | Val125Gly | gtc/ggc | PZA | High confidence |
| 2288869  | G     | T     | coding | Rv2043c | pncA | Val125Phe | gtc/ttc | PZA | High confidence |
| 2288874  | G     | C     | coding | Rv2043c | pncA | Arg123Pro | cgc/ccc | PZA | High confidence |
| 2288880  | G     | C     | coding | Rv2043c | pncA | Arg121Pro | cgg/ccg | PZA | High confidence |
| 2288883  | T     | A     | coding | Rv2043c | pncA | Leu120Gln | ctg/cag | PZA | High confidence |
| 2288883  | T     | G     | coding | Rv2043c | pncA | Leu120Pro | ctg/ccg | PZA | High confidence |
| 2288888  | G     | A     | coding | Rv2043c | pncA | Trp119STOP | tgg/tga | PZA | Low confidence |
| 2288888  | G     | C     | coding | Rv2043c | pncA | Trp119Ser | tgg/tcg | PZA | High confidence |
| 2288888  | G     | A     | coding | Rv2043c | pncA | Trp119STOP | tgg/tag | PZA | High confidence |
| 2288887  | T     | C     | coding | Rv2043c | pncA | Trp119Arg | tgg/cgg | PZA | High confidence |
| 2288887  | T     | G     | coding | Rv2043c | pncA | Trp119Gly | tgg/ggg | PZA | High confidence |
| 2288895  | T     | G     | coding | Rv2043c | pncA | Leu116Arg | ctg/cgg | PZA | High confidence |
| 2288895  | T     | C     | coding | Rv2043c | pncA | Leu116Pro | ctg/ccg | PZA | High confidence |
| 2288902  | A     | C     | coding | Rv2043c | pncA | Thr114Pro | acg/ccg | PZA | Low confidence |
| 2288920  | G     | A     | coding | Rv2043c | pncA | Gly108Arg | gga/aga | PZA | High confidence |
| 2288920  | G     | C     | coding | Rv2043c | pncA | Gly108Arg | gga/cga | PZA | High confidence |
| 2288928  | G     | A     | coding | Rv2043c | pncA | Gly105Asp | ggc/gac | PZA | High confidence |
| 2288930  | C     | A     | coding | Rv2043c | pncA | Ser104Arg | agc/aga | PZA | High confidence |
| 2288931  | G     | T     | coding | Rv2043c | pncA | Ser104Ile | agc/atc | PZA | High confidence |
| 2288933  | C     | G     | coding | Rv2043c | pncA | Tyr103STOP | tac/tag | PZA | High confidence |
| 2288934  | A     | G     | coding | Rv2043c | pncA | Tyr103Cys | tac/tgc | PZA | High confidence |
| 2288934  | A     | C     | coding | Rv2043c | pncA | Tyr103Ser | tac/tcc | PZA | Low confidence |
| 2288935  | T     | G     | coding | Rv2043c | pncA | Tyr103Asp | tac/gac | PZA | High confidence |
| 2288938  | G     | A     | coding | Rv2043c | pncA | Ala102Thr | gec/ace | PZA | High confidence |
| 2288944  | A     | G     | coding | Rv2043c | pncA | Thr100Ala | acc/gcc | PZA | High confidence |
| Job # | A | C | coding | Rv2043c | pncA | - | Thr100Pro | acc/ccc | PZA | High confidence |
|-------|---|---|---------|---------|------|---|-----------|--------|-----|----------------|
| 2288944 | C | A | coding | Rv2043c | pncA | - | Tyr99STOP | tac/taa | PZA | High confidence |
| 2288945 | G | C | coding | Rv2043c | pncA | - | Gly97Ala | ggt/gct | PZA | High confidence |
| 2288952 | G | A | coding | Rv2043c | pncA | - | Gly97Asp | ggt/gat | PZA | High confidence |
| 2288953 | G | A | coding | Rv2043c | pncA | - | Gly97Ser | ggt/agt | PZA | Low confidence |
| 2288954 | G | A | coding | Rv2043c | pncA | - | Lys96Asn | aag/aac | PZA | High confidence |
| 2288955 | A | C | coding | Rv2043c | pncA | - | Lys96Arg | aag/agg | PZA | High confidence |
| 2288956 | A | C | coding | Rv2043c | pncA | - | Lys96Thr | aag/acg | PZA | High confidence |
| 2288956 | A | G | coding | Rv2043c | pncA | - | Lys96Gln | aag/cag | PZA | High confidence |
| 2288957 | C | G | coding | Rv2043c | pncA | - | Tyr95STOP | tac/tcg | PZA | High confidence |
| 2288960 | C | A | coding | Rv2043c | pncA | - | Phe94Leu | ttc/tta | PZA | High confidence |
| 2288961 | T | G | coding | Rv2043c | pncA | - | Phe94Cys | ttc/tgc | PZA | High confidence |
| 2288961 | T | C | coding | Rv2043c | pncA | - | Phe94Ser | ttc/tcc | PZA | High confidence |
| 2288962 | T | C | coding | Rv2043c | pncA | - | Phe94Leu | ttc/tcg | PZA | High confidence |
| 2288971 | G | T | coding | Rv2043c | pncA | - | Glu91STOP | gag/tag | PZA | High confidence |
| 2288982 | C | T | coding | Rv2043c | pncA | - | Thr87Met | acg/atg | PZA | Low confidence |
| 2288988 | T | G | coding | Rv2043c | pncA | - | Leu85Arg | ctc/cgg | PZA | High confidence |
| 2288988 | T | C | coding | Rv2043c | pncA | - | Leu85Pro | ctc/ccg | PZA | Low confidence |
| 2288997 | A | G | coding | Rv2043c | pncA | - | His82Arg | cat/cgt | PZA | High confidence |
| 2288998 | C | G | coding | Rv2043c | pncA | - | His82Asp | cat/gat | PZA | High confidence |
| 2289000 | T | G | coding | Rv2043c | pncA | - | Phe81Val | ttc/tgc | PZA | High confidence |
| 2289000 | T | C | coding | Rv2043c | pncA | - | Phe81Val | ttc/tcc | PZA | High confidence |
| 2289001 | T | G | coding | Rv2043c | pncA | - | Phe81Val | ttc/tgc | PZA | High confidence |
| 2289016 | A | C | coding | Rv2043c | pncA | - | Thr76Pro | act/cct | PZA | Low confidence |
| 2289028 | T | C | coding | Rv2043c | pncA | - | Cys72Arg | ttc/cgc | PZA | High confidence |
| 2289029 | T | A | coding | Rv2043c | pncA | - | His71Gln | cat/caa | PZA | High confidence |
| Gene  | Codon | Amino Acid 1 | Amino Acid 2 | Amino Acid 3 | Amino Acid 4 |
|-------|-------|-------------|-------------|-------------|-------------|
| pncA  | Rv2043c | His71Arg   | cat/cgt    | PZA         | High confidence |
| pncA  | Rv2043c | His71Tyr   | cat/tat    | PZA         | Low confidence |
| pncA  | Rv2043c | Pro69Leu   | cca/cta    | PZA         | Low confidence |
| pncA  | Rv2043c | Trp68Cys   | tgg/tgc    | PZA         | High confidence |
| pncA  | Rv2043c | Trp68Cys   | tgg/tgt    | PZA         | High confidence |
| pncA  | Rv2043c | Trp68STOP  | tgg/tag    | PZA         | High confidence |
| pncA  | Rv2043c | Trp68Ser   | tgg/ctg    | PZA         | Low confidence |
| pncA  | Rv2043c | Trp68Gly   | tgg/ggg    | PZA         | Low confidence |
| pncA  | Rv2043c | Ser67Pro   | tgg/ccg    | PZA         | High confidence |
| pncA  | Rv2043c | Tyr64STOP  | tat/tag    | PZA         | High confidence |
| pncA  | Rv2043c | Tyr64Asp   | tat/gat    | PZA         | Low confidence |
| pncA  | Rv2043c | Asp63Gly   | cac/ggc    | PZA         | Low confidence |
| pncA  | Rv2043c | Pro62Leu   | ccc/ctg    | PZA         | High confidence |
| pncA  | Rv2043c | Thr61Pro   | aca/cca    | PZA         | Low confidence |
| pncA  | Rv2043c | Phe58Leu   | ttc/ctg    | PZA         | High confidence |
| pncA  | Rv2043c | His57Gln   | cac/ag     | PZA         | High confidence |
| pncA  | Rv2043c | His57Arg   | cac/cgc    | PZA         | High confidence |
| pncA  | Rv2043c | His57Pro   | cac/ccc    | PZA         | High confidence |
| pncA  | Rv2043c | His57Asp   | cac/gac    | PZA         | Low confidence |
| pncA  | Rv2043c | His57Tyr   | cac/tac    | PZA         | Low confidence |
| pncA  | Rv2043c | Pro54Arg   | ccc/ctg    | PZA         | High confidence |
| pncA  | Rv2043c | Pro54Gln   | ccc/cag    | PZA         | High confidence |
| pncA  | Rv2043c | Pro54Leu   | ccc/ctg    | PZA         | Low confidence |
| pncA  | Rv2043c | Pro54Ser   | ccc/tcg    | PZA         | High confidence |
| pncA  | Rv2043c | His51Gln   | cac/caa    | PZA         | High confidence |
| Position | Nucleotide | Codon | Protein Feature | Amino Acid Change | Confidence |
|----------|------------|-------|----------------|-------------------|------------|
| 2289090 | A          | G     | coding         | Rv2043c pncA      | -          | His51Arg cac/cgc PZA High confidence |
| 2289090 | A          | C     | coding         | Rv2043c pncA      | -          | His51Pro cac/ccc PZA High confidence |
| 2289091 | C          | T     | coding         | Rv2043c pncA      | -          | His51Tyr cac/tac PZA High confidence |
| 2289096 | A          | C     | coding         | Rv2043c pncA      | -          | Asp49Ala gac/gcc PZA Low confidence |
| 2289096 | A          | G     | coding         | Rv2043c pncA      | -          | Asp49Gly gac/ggc PZA Low confidence |
| 2289097 | G          | A     | coding         | Rv2043c pncA      | -          | Asp49Asn cac/tac PZA High confidence |
| 2289100 | A          | G     | coding         | Rv2043c pncA      | -          | Lys48STOP aag/tac PZA High confidence |
| 2289103 | A          | C     | coding         | Rv2043c pncA      | -          | Thr47Pro acc/ccc PZA High confidence |
| 2289103 | A          | G     | coding         | Rv2043c pncA      | -          | Thr47Ala acc/gcc PZA Low confidence |
| 2289108 | T          | G     | coding         | Rv2043c pncA      | -          | Val45Gly gta/ggg PZA Low confidence |
| 2289111 | T          | G     | coding         | Rv2043c pncA      | -          | Val44Gly gta/gga PZA High confidence |
| 2289133 | G          | T     | coding         | Rv2043c pncA      | -          | Glu37STOP gaa/taa PZA High confidence |
| 2289138 | T          | C     | coding         | Rv2043c pncA      | -          | Leu35Pro ctc/ccg PZA High confidence |
| 2289140 | C          | G     | coding         | Rv2043c pncA      | -          | Tyr34STOP tac/tag PZA High confidence |
| 2289150 | T          | G     | coding         | Rv2043c pncA      | -          | Ile31Ser atc/agc PZA High confidence |
| 2289159 | C          | A     | coding         | Rv2043c pncA      | -          | Ala28Asp gcc/gac PZA Low confidence |
| 2289162 | T          | C     | coding         | Rv2043c pncA      | -          | Leu27Pro ctc/ccg PZA High confidence |
| 2289171 | G          | A     | coding         | Rv2043c pncA      | -          | Gly24Asp ggc/gac PZA High confidence |
| 2289180 | T          | G     | coding         | Rv2043c pncA      | -          | Val21Gly gta/gga PZA High confidence |
| 2289186 | T          | C     | coding         | Rv2043c pncA      | -          | Leu19Pro ctc/ccg PZA High confidence |
| 2289193 | G          | A     | coding         | Rv2043c pncA      | -          | Gly17Ser ggc/ggc PZA High confidence |
| 2289200 | C          | A     | coding         | Rv2043c pncA      | -          | Cys14STOP tgc/tga PZA High confidence |
| 2289201 | G          | A     | coding         | Rv2043c pncA      | -          | Cys14Tyr tgc/tac PZA Low confidence |
| 2289202 | T          | C     | coding         | Rv2043c pncA      | -          | Cys14Arg tgc/ccg PZA High confidence |
| 2289203 | C          | G     | coding         | Rv2043c pncA      | -          | Phe13Leu ttc/ttg PZA High confidence |
| 2289204 | T          | C     | coding         | Rv2043c pncA      | -          | Phe13Ser ttc/tcc PZA Low confidence |
| 2289206 | C          | G     | coding         | Rv2043c pncA      | -          | Asp12Glu gac/gag PZA High confidence |
| 2289207 | A | C | coding | Rv2043c | pncA | - | Asp12Ala | gac/gcc | PZA | High confidence |
| 2289208 | G | A | coding | Rv2043c | pncA | - | Asp12Asn | gac/aac | PZA | Low confidence |
| 2289213 | A | G | coding | Rv2043c | pncA | - | Gln10Arg | cag/cgg | PZA | Low confidence |
| 2289213 | A | C | coding | Rv2043c | pncA | - | Gln10Pro | cag/ccg | PZA | Low confidence |
| 2289214 | C | A | coding | Rv2043c | pncA | - | Gln10Lys | cag/aag | PZA | High confidence |
| 2289216 | T | G | coding | Rv2043c | pncA | - | Val9Ala | gtt/gcg | PZA | Low confidence |
| 2289216 | T | G | coding | Rv2043c | pncA | - | Val9Gly | gtt/ggg | PZA | Low confidence |
| 2289218 | C | A | coding | Rv2043c | pncA | - | Asp8Glu | gac/gaa | PZA | High confidence |
| 2289219 | A | C | coding | Rv2043c | pncA | - | Gln10Lys | cag/aag | PZA | High confidence |
| 2289219 | A | G | coding | Rv2043c | pncA | - | Gln10Pro | cag/ccg | PZA | Low confidence |
| 2289222 | T | G | coding | Rv2043c | pncA | - | Val7Gly | gtc/ggc | PZA | High confidence |
| 2289222 | T | A | coding | Rv2043c | pncA | - | Val7Asp | gtc/gac | PZA | Low confidence |
| 2289225 | T | G | coding | Rv2043c | pncA | - | Val7Phe | gtc/ttc | PZA | High confidence |
| 2289231 | T | G | coding | Rv2043c | pncA | - | Ile6Thr | atc/acc | PZA | High confidence |
| 2289231 | T | G | coding | Rv2043c | pncA | - | Leu4Ser | ttg/tcg | PZA | High confidence |
| 2289234 | C | A | coding | Rv2043c | pncA | - | Ala3Glu | gcg/gag | PZA | High confidence |
| 2289235 | G | C | coding | Rv2043c | pncA | - | Ala3Pro | gcg/ccg | PZA | Low confidence |
| 2289239 | G | A | coding | Rv2043c | pncA | - | Met1Ile | atg/ata | PZA | High confidence |
| 2289239 | G | T | coding | Rv2043c | pncA | - | Met1Ile | atg/att | PZA | Low confidence |
| 2289240 | T | A | coding | Rv2043c | pncA | - | Met1Lys | atg/aag | PZA | High confidence |
| 2289240 | T | C | coding | Rv2043c | pncA | - | Met1Thr | atg/aeg | PZA | High confidence |
| 2289248 | T | C | coding | Rv2043c | pncA | - | --- | --- | PZA | High confidence |
| 2289248 | T | G | coding | Rv2043c | pncA | - | --- | --- | PZA | High confidence |
| 2289252 | A | G | coding | Rv2043c | pncA | - | --- | --- | PZA | High confidence |
| 2289252 | A | C | coding | Rv2043c | pncA | - | --- | --- | PZA | High confidence |
| 2289252 | A | T | coding | Rv2043c | pncA | - | --- | --- | PZA | High confidence |

**Columns:**
- **Gene:** Rv2043c
- **Coding Region:** pncA
- **Codon:**
- **Amino Acid Change:**
- **Database:** PZA
- **Confidence Level:**
  - High confidence
  - Low confidence
| Gene | Type | Start | End | Allele | Change | Function | Confidence |
|------|------|-------|-----|--------|---------|----------|------------|
| Rv2416c | intergenic | 2715342 | 2715346 | C>T | intergenic | eis | - | KAN | Low confidence |
| Rv2428 | coding | 2726136 | 2726145 | G>A | coding | ahpC | + | INH | Low confidence |
| Rv2428 | coding | 3073808 | 3073808 | C>G | coding | thyA | Arg222Gly | PAS | Low confidence |
| Rv3793 | coding | 4241078 | 4241078 | A>G | coding | embC | Ile406Val | EMB | Low confidence |
| Rv3794 | coding | 4243221 | 4243221 | C>A | coding | embA | Asp4Asn | EMB | Low confidence |
| Rv3794 | coding | 4243242 | 4243242 | G>A | coding | embA | Gly5Ser | EMB | Low confidence |
| Rv3794 | coding | 4243245 | 4243245 | G>A | coding | embA | Ala201Thr | EMB | Low confidence |
| Rv3794 | coding | 4243254 | 4243254 | G>A | coding | embA | Gly321Ser | EMB | Low confidence |
| Rv3794 | coding | 4243261 | 4243261 | G>A | coding | embA | Gly350Asp | EMB | Low confidence |
| Rv3794 | coding | 4243833 | 4243833 | G>A | coding | embA | Ala462Val | EMB | Low confidence |
| Rv3795 | coding | 4247402 | 4247402 | T>G | coding | embB | Ser297Ala | EMB | Low confidence |
| Rv3795 | coding | 4247429 | 4247429 | A>C | coding | embB | Met306Leu | EMB | High confidence |
| Rv3795 | coding | 4247430 | 4247430 | T>C | coding | embB | Met306Thr | EMB | High confidence |
| Rv3795 | coding | 4247471 | 4247471 | A>C | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247478 | 4247478 | G>T | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247485 | 4247485 | A>G | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247506 | 4247506 | T>C | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247512 | 4247512 | G>T | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247522 | 4247522 | A>G | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247523 | 4247523 | G>T | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247524 | 4247524 | A>G | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247525 | 4247525 | G>T | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247526 | 4247526 | A>G | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247527 | 4247527 | G>T | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247528 | 4247528 | A>G | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247529 | 4247529 | G>T | coding | embB | Met306Ile | EMB | High confidence |
| Rv3795 | coding | 4247530 | 4247530 | A>C | coding | embB | Met306Ile | EMB | High confidence |
| Chromosome | Position | Gene | Codon Change | Amino Acid Change | Confidence |
|------------|----------|------|--------------|-------------------|------------|
| 4247573    | 26       | G    A  | coding       | Rv3795 embB +    | Asp354Asn  | gac/aac   | EMB       | Low confidence |
| 4247717    | 402      | C    G  | coding       | Rv3795 embB +    | Leu402Val  | ctg/gtg   | EMB       | Low confidence |
| 4247723    | 404      | C    T  | coding       | Rv3795 embB +    | Pro404Ser  | ccg/tcg   | EMB       | Low confidence |
| 4247729    | 406      | G    A  | coding       | Rv3795 embB +    | Gly406Ser  | ggc/agc   | EMB       | High confidence |
| 4247729    | 406      | G    T  | coding       | Rv3795 embB +    | Gly406Cys  | ggc/tgc   | EMB       | High confidence |
| 4247730    | 406      | G    C  | coding       | Rv3795 embB +    | Gly406Ala  | ggc/gcc   | EMB       | High confidence |
| 4247730    | 406      | G    A  | coding       | Rv3795 embB +    | Gly406Cys  | ggc/tgc   | EMB       | High confidence |
| 4247863    | 450      | C    G  | coding       | Rv3795 embB +    | Ile450Met  | atc/atg   | EMB       | Low confidence |
| 4248002    | 454      | C    G  | coding       | Rv3795 embB +    | Ala454Thr  | gcg/acg   | EMB       | Low confidence |
| 4248003    | 497      | A    G  | coding       | Rv3795 embB +    | Gln497Lys  | cag/aag   | EMB       | Low confidence |
| 4248747    | 497      | C    A  | coding       | Rv3795 embB +    | Gln497Arg  | cag/cgg   | EMB       | High confidence |
| 4249518    | 500      | A    G  | coding       | Rv3795 embB +    | His1002Arg | cac/cgc   | EMB       | Low confidence |
| 4326087    | 463      | C    A  | coding       | Rv3854c ethA -   | Arg463Ser  | cgt/agt   | ETH       | Low confidence |
| 4326236    | 413      | G    A  | coding       | Rv3854c ethA -   | Gly413Asp  | ggt/gat   | ETH       | Low confidence |
| 4326320    | 392      | G    A  | coding       | Rv3854c ethA -   | Thr392Ala  | acg/gcg   | ETH       | Low confidence |
| 4326333    | 381      | G    C  | coding       | Rv3854c ethA -   | Ala381Pro  | gcc/ccc   | ETH       | Low confidence |
| 4326449    | 342      | C    A  | coding       | Rv3854c ethA -   | Thr342Lys  | acg/aag   | ETH       | Low confidence |
| 4326641    | 338      | T    G  | coding       | Rv3854c ethA -   | Ile338Ser  | atc/agc   | ETH       | Low confidence |
| 4326738    | 246      | C    T  | coding       | Rv3854c ethA -   | Gln246STOP | cag/tag   | ETH       | Low confidence |
| 4326807    | 223      | G    A  | coding       | Rv3854c ethA -   | Glu223Lys  | gac/aac   | ETH       | Low confidence |
| 4326927    | 186      | C    A  | coding       | Rv3854c ethA -   | Thr186Lys  | acg/aag   | ETH       | Low confidence |
| 4327224    | 84       | T    G  | coding       | Rv3854c ethA -   | Tyr84Asp   | tac/gac   | ETH       | Low confidence |
| 4327301    | 56       | A    C  | coding       | Rv3854c ethA -   | Asp56Asp   | gac/gcc   | ETH       | Low confidence |
| 4327307    | 56       | A    C  | coding       | Rv3854c ethA -   | Asp56Asp   | gac/gcc   | ETH       | Low confidence |
| 4327322    | 51       | C    T  | coding       | Rv3854c ethA -   | Pro51Leu   | ccc/ctc   | ETH       | Low confidence |
| 4327346    | 43       | G    A  | coding       | Rv3854c ethA -   | Gly43Asp   | ggc/gac   | ETH       | Low confidence |
| SeqID   | Location | Gene | Codon | Amino Acid Change   | Allele | Ref Allele | Resistance | Confidence |
|---------|----------|------|-------|--------------------|--------|------------|------------|------------|
| 4327347 | T        | Rv3854c | ethA  | Gly43Cys           | ggc    | tgc        | ETH        | Low        |
| 4407604 | C        | Rv3919c | gidB  | Ala200Glu          | gcg    | gag        | STR        | Low        |
| 4407790 | C        | Rv3919c | gidB  | Ala138Val          | gcg    | gtg        | STR        | Low        |
| 4407824 | C        | Rv3919c | gidB  | Gln127STOP         | caa    | taa        | STR        | Low        |
| 4407931 | T        | Rv3919c | gidB  | Leu91Pro           | ctc    | cca        | STR        | Low        |
| 4407940 | T        | Rv3919c | gidB  | Val88Ala           | gta    | gca        | STR        | Low        |
| 4407992 | G        | Rv3919c | gidB  | Gly71Arg           | gga    | aga        | STR        | Low        |
| 4408009 | T        | Rv3919c | gidB  | Val65Gly           | gtc    | ggc        | STR        | Low        |

Confidence of the variant. High confidence means that variant has a statistical significant value, while Low confidence does not. The list was obtained by PhyResSe available database[10] (latest accessed in April 2018).

Abbreviations: AMK, amikacin; CM, capreomycin; FQ, fluoroquinolones (ciprofloxacin, levofloxacin, moxifloxacin); EMB, ethambutol; ETH, ethionamide; INH, isoniazid; KAN, kanamycin; LZD, linezolid; STR, streptomycin; PAS, para-aminosalicylic acid; PZA, pyrazinamide.
## Supplementary Table 2. Significant variant frequencies across all isolates

| Genomic Position | Wild-type Allele | Variant Allele | Jan 2014 | March 2014 | June 2014 | Aug 2014 | Oct 2014 | Dec 2014 | Jan 2015 | April 2015 | June 2015 | Nov 2015 | Dec 2015 | June 2016 | Oct 2016 | Jan 2017 |
|------------------|------------------|----------------|----------|------------|-----------|----------|---------|----------|----------|-----------|----------|----------|----------|----------|----------|----------|
| 6742             | A                | C              | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 100%     | 100%     | 100%     |
| 36471            | C                | G              | 0%       | 0%         | 57.14%    | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 85674            | C                | G              | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 14.58%   | 0%       | 0%       | 0%       |
| 130660           | T                | A              | 100%     | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 208321           | C                | Ga             | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 93.33%   | 85.35%   | 28.22%   | 0%       | 0%       | 0%       |
| 232974           | A                | C              | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 93.33%   | 85.35%   | 28.22%   | 0%       | 0%       | 0%       |
| 580797           | A                | G              | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 623273           | G                | A              | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 650379           | C                | T              | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 699980           | G                | C              | 0%       | 0%         | 0%        | 15%      | 0%      | 55.56%   | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 725200           | C                | T              | 0%       | 20.93%     | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 761152           | T                | A              | 0%       | 0%         | 0%        | 0%       | 11.82%  | 0%       | 7.25%    | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 761277           | A                | T              | 100%     | 100%       | 99.24%    | 99.56%   | 99.07%  | 99.32%   | 100%     | 100%      | 98.63%   | 98.97%   | 100%     | 98.97%   | 100%     | 99.32%   |
| 851731           | T                | G              | 0%       | 0%         | 0%        | 0%       | 0%      | 23.48%   | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 852638           | C                | G              | 0%       | 0%         | 0%        | 0%       | 0%      | 50.81%   | 71.91%   | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 854253           | C                | G              | 0%       | 0%         | 0%        | 30.77%   | 0%      | 0%       | 0%       | 0%        | 27.27%   | 0%       | 0%       | 0%       | 0%       | 0%       |
| 916546           | G                | T              | 0%       | 0%         | 0%        | 0%       | 0%      | 19.05%   | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 939197           | G                | C              | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 1002273          | G                | A              | 100%     | 100%       | 100%      | 100%     | 100%    | 100%     | 100%     | 100%      | 100%     | 100%     | 100%     | 100%     | 100%     | 100%     |
| 1131770          | C                | T              | 0%       | 0%         | 0%        | 0%       | 14.97%  | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 1157076          | C                | A              | 100%     | 100%       | 99.25%    | 98.97%   | 100%    | 99.07%   | 100%     | 99.35%    | 100%     | 100%     | 100%     | 100%     | 100%     | 100%     |
| 1230842          | T                | C              | 0%       | 0%         | 0%        | 0%       | 0%      | 0%       | 0%       | 0%        | 0%       | 0%       | 0%       | 0%       | 0%       | 0%       |
| 1474752          | T                | C              | 100%     | 100%       | 100%      | 98.68%   | 95.83%  | 97.83%   | 100%     | 95.65%    | 100%     | 98.04%   | 100%     | 100%     | 99.49%   | 99.49%   |
| Sample ID   | A | G | C | T |
|-------------|---|---|---|---|
| 1502314     |   |   | C | T |
|             |   |   |   | 99.48% |
| 1514788     | G | T |   |   |
|             | 0% | 25.66% | 0% | 0% |
| 1519823     | A | G |   |   |
|             | 0% | 0% | 0% | 0% |
| 1524571     | A | T |   |   |
|             | 0% | 0% | 100% | 0% |
| 1543413     | C | G |   |   |
|             | 0% | 0% | 0% | 0% |
| 1917986     | T | G |   |   |
|             | 0% | 0% | 0% | 0% |
| 1960078     | C | A |   |   |
|             | 0% | 0% | 0% | 0% |
| 1997457     | G | C |   |   |
|             | 0% | 0% | 0% | 36.36% |
| 2074547     | G | T |   |   |
|             | 0% | 0% | 0% | 0% |
| 2094917     | T | A |   |   |
|             | 0% | 0% | 0% | 0% |
| 2154827     | C |   |   | 99.00% |
|             | 0% | 0% | 0% | 67% |
| 2155295     | C | G |   |   |
|             | 0% | 98.06% | 40.74% | 97.14% |
| 2180818     | T | C |   |   |
|             | 0% | 0% | 0% | 0% |
| 2183360     | C | T |   |   |
|             | 0% | 0% | 0% | 0% |
| 2262857     | C | T |   |   |
|             | 100% | 100% | 100% | 100% |
| 2304456     | G | A |   |   |
|             | 0% | 0% | 0% | 0% |
| 2304044     | T | A |   |   |
|             | 0% | 0% | 0% | 0% |
| 2314650     | T | C |   |   |
|             | 0% | 0% | 0% | 0% |
| 2339605     | A | G |   |   |
|             | 100% | 100% | 90.91% | 100% |
| 2363682     | C | A |   |   |
|             | 0% | 0% | 0% | 100% |
| 2370345     | C | T |   |   |
|             | 0% | 0% | 0% | 0% |
| 2482657     | G | C |   |   |
|             | 0% | 0% | 21.05% | 36.17% |
| 2525723     | G | C |   |   |
|             | 20% | 0% | 0% | 0% |
| 2527757     | C | T |   |   |
|             | 0% | 0% | 23.27% | 36.18% |
| 2534563     | G | C |   |   |
|             | 30% | 0% | 0% | 0% |
| 2536917     | T | C |   |   |
|             | 0% | 16.39% | 0% | 0% |
|     | C   | T   | 0%  | 0%  | 0%  | 0%  | 60% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 2542966 | C   | T   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2551675 | A   | C   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2726112 | C   | T   | 0%  | 0%  | 19.77% | 0%  | 35.34% | 0%  | 25.83% | 0%  | 96.72% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2726139 | C   | T   | 0%  | 0%  | 13.79% | 0%  | 67.91% | 0%  | 26%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2726141 | C   | A/T | 0%  | 5.71% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 45.19% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2726145 | G   | A   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2726153 | G   | A   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2726767 | A   | C   | 94.94% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2786787 | T   | G   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 2794585 | C   | A   | 100% | 100% | 100% | 100% | 99.32% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% |
| 2879778 | G   | T   | 100% | 100% | 97.22% | 99.02% | 100% | 98.41% | 100% | 100% | 100% | 100% | 100% | 100% | 98.15% | 92.86% | 98.91% | 97.87% |
| 2895066 | C   | G   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 96.30% | 82.96% | 30.65% | 0%  | 0%  | 0%  | 0%  |
| 2895750 | G   | A   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 75%  | 40.58% | 41.73% | 0%  | 16.34% | 66.83% | 97.47% | 100% | 100% | 100% |
| 3007115 | C   | A   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 3007143 | T   | A   | 0%  | 0%  | 100% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 100% | 0%  | 0%  |
| 3008407 | C   | G   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 3008414 | C   | T   | 15.85% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 300839  | G   | T   | 0%  | 34.78% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 3045144 | C   | T   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 3096287 | G   | C   | 100% | 100% | 0%  | 100% | 0%  | 100% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 3096295 | G   | C   | 0%  | 0%  | 100% | 100% | 0%  | 0%  | 0%  | 75%  | 0%  | 80%  | 100% | 75%  | 0%  | 100% | 0%  |
| 3160000 | A   | C   | 71.43% | 0%  | 60%  | 87.50% | 0%  | 100% | 80%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 3173107 | G   | A   | 100% | 100% | 100% | 100% | 98.23% | 100% | 100% | 98.65% | 99.40% | 100% | 100% | 100% | 100% | 100% | 100% |
| 3247866 | C   | A   | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 3247867 | A   | G   | 80%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
| 3247868 | A   | G   | 100% | 0%  | 0%  | 0%  | 0%  | 0%  | 42.86% | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  | 0%  |
|       |   |   |   |   |   |   |   |   |   |   |   |
|-------|---|---|---|---|---|---|---|---|---|---|---|
|       | A | G | 100% | 0% | 0% | 0% | 0% | 66.67% | 0% | 0% | 60% | 0% | 0% | 0% | 0% |
| 3247869 | T | C | 0% | 0% | 0% | 0% | 16.28% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 3349257 | G | A | 0% | 0% | 0% | 0% | 0% | 0% | 22.22% | 0% | 0% | 16.13% | 0% | 0% | 30.77% | 13.16% |
| 3415194 | A | G | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 3498811 | T | C | 62.50% | 0% | 0% | 0% | 50% | 50% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 3580637 | G | A | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 3596631 | A | G | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 30% | 0% |
| 3664945 | C | T | 99.17% | 100% | 99.40% | 100% | 98.90% | 99.12% | 100% | 96.84% | 100% | 100% | 99.37% | 100% | 100% | 100% |
| 3854063 | G | T | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 30% | 0% |
| 3862473 | A | G | 0% | 0% | 0% | 0% | 71.43% | 77.78% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 3878567 | G | C | 0% | 28.57% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 16.67% |
| 3915436 | C | T | 0% | 0% | 0% | 0% | 0% | 57.14% | 61.90% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 4026874 | A | G | 0% | 0% | 0% | 0% | 71.43% | 77.78% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 4117169 | C | G | 0% | 60% | 0% | 0% | 50% | 50% | 66.67% | 0% | 37.50% | 0% | 54.55% | 55.56% | 50% | 0% |
| 4247429 | A | G | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 10.26% | 0% | 0% |
| 4247495 | G | T | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 4255385 | C | G | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 4269341 | C | T | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 97.03% | 86.79% | 23.53% | 0% | 0% |
| 4338596 | T | G | 0% | 0% | 0% | 0% | 50% | 40% | 0% | 0% | 50% | 0% | 0% | 0% | 0% | 0% |
| 4353414 | G | C | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 4358702 | A | G | 0% | 0% | 0% | 0% | 100% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% |
| 4373008 | T | C | 0% | 0% | 0% | 0% | 0% | 42.86% | 0% | 0% | 0% | 0% | 60% | 0% | 0% | 33.3% | 0% |
### Supplementary Table 3. Isolate sequencing results and main drug resistance associated mutations

| Isolate | Date       | Accession number | Mean coverage | MTBC (%)<sup>a</sup> | Predicted resistance profile | Genetic changes associated to resistance mutations | Genetic changes associated to resistance mutations |
|----------|------------|------------------|---------------|-----------------------|------------------------------|---------------------------------------------------|---------------------------------------------------|
| G1480    | 27/04/2009 | ERS1880079       | 171.36        | 99.95                 | DS                           | WT                                               | WT                                               |
| G1479    | 20/09/2013 | ERS1880078       | 134.25        | 99.87                 | RR                           | I491F                                          | WT                                               |
| G516     | 21/01/2014 | ERS1880086       | 212.10        | 99.94                 | MDR                          | I491F                                          | P429del                                          |
| G520     | 17/03/2014 | ERS1880087       | 134.72        | 99.96                 | MDR                          | I491F                                          | G273R                                           |
| G249     | 27/06/2014 | ERS1880084       | 184.85        | 99.91                 | MDR                          | I491F                                          | G273R                                           |
| G252     | 11/08/2014 | ERS1880085       | 206.85        | 99.97                 | MDR                          | I491F                                          | G273R                                           |
| G535     | 13/10/2014 | ERS1880088       | 122.67        | 99.89                 | MDR                          | I491F                                          | P429del                                          |
| G841     | 09/12/2014 | ERS1880089       | 162.37        | 99.93                 | MDR                          | I491F                                          | P429del                                          |
| G842     | 27/01/2015 | ERS1880090       | 115.82        | 99.97                 | MDR                          | I491F                                          | G273R                                           |
| G993     | 15/04/2015 | ERS1880091       | 137.62        | 99.77                 | MDR                          | I491F                                          | G273R                                           |
| G1003    | 11/06/2015 | ERS1880075       | 186.61        | 99.95                 | MDR                          | I491F                                          | G273R                                           |
| G1257    | 09/11/2015 | ERS1880076       | 268.63        | 99.90                 | MDR                          | I491F                                          | G273R                                           |
| G1478    | 11/12/2015 | ERS1880077       | 151.24        | 99.91                 | MDR                          | I491F                                          | G273R                                           |

<sup>a</sup> Mean coverage for MTBC (%).
| Sample | Date       | Accession | MTBC% | Read % | Resistance | Mutations | Drug Sensitivity | Other Mutations |
|--------|------------|-----------|-------|--------|-------------|-----------|-----------------|-----------------|
| G1720  | 02/06/2016 | ERS1880081| 118.75| 99.97  | XDR         | I491F     | P429del         | G-9A            |
|        |            |           |       |        |             |           |                 | C-10T C-39T     |
| G1721  | 14/10/2016 | ERS1880082| 249.21| 99.75  | XDR         | I491F     | P429del         | WT              |
|        |            |           |       |        |             |           |                 | WT              |
| G1928  | 09/01/2017 | ERS1880083| 201.00| 99.66  | XDR         | L449Q     | P429del         | C-39T           |
|        |            |           |       |        |             |           |                 | WT              |

*Percentage of reads belonging to MTBC.

**Promoter mutations between oxyR and ahpc genes.

Abbreviations: DS, drug susceptible; MDR, multidrug resistance; MTBC, *Mycobacterium tuberculosis* Complex; XDR, extensively drug-resistant.*
| Drug associated | Genomic position | Wild type allele | Mutant allele | Gene alias | Gene name | Gene type | Mutation type | Nucleotide change | Aminoacid change |
|-----------------|------------------|-----------------|---------------|------------|-----------|-----------|---------------|------------------|-----------------|
| Moxifloxacin    | 6742             | A               | C             | gyrB       | Rv0005    | essential | nonsynonymous | A1620C           | E540D           |
| Rifampicin      | 761152           | T               | A             | rpoB       | Rv0667    | essential | nonsynonymous | A1346T           | L449Q           |
| Rifampicin      | 761277           | A               | T             | rpoB       | Rv0667    | essential | nonsynonymous | A1471T           | I491F           |
| Capreomycin     | 1917986          | T               | G             | tlyA       | Rv1694    | nonessential | nonsynonymous | T47G             | L16R            |
| Isoniazid       | 2154827          | C               | -             | katG       | Rv1908c   | nonessential | nonsynonymous | C1285Δ           | G429del         |
| Isoniazid       | 2155295          | C               | G             | katG       | Rv1908c   | nonessential | nonsynonymous | G817C            | G273R           |
| Isoniazid       | 2726112          | C               | T             |            |           |           | ahpC promoter mutation |               |                 |
| Isoniazid       | 2726139          | C               | T             |            |           |           | ahpC promoter mutation |               |                 |
| Isoniazid       | 2726141          | C               | A/T           |            |           |           | ahpC promoter mutation |               |                 |
| Isoniazid       | 2726145          | G               | A             |            |           |           | ahpC promoter mutation |               |                 |
| Isoniazid       | 2726153          | G               | A             |            |           |           | ahpC promoter mutation |               |                 |
| Isoniazid       | 2726767          | A               | C             | ahpC       | Rv2428    | nonessential | nonsynonymous | A575C            | K192T           |
| Ethambutol      | 4247429          | A               | G             | embB       | Rv3795    | essential | nonsynonymous | A916G            | M306V           |
| Ethambutol      | 4247495          | G               | T             | embB       | Rv3795    | essential | nonsynonymous | G982T            | D328Y           |
| Ethambutol      | 4269341          | C               | T             | ubiA       | Rv3806c   | essential | nonsynonymous | G493A            | G165S           |
**Supplementary Table 5. Drug resistance profile predicted by publicly databases**

| Isolate | Date       | Predicted resistance type (this study)a | PhyResSe[10] | Mykrobe predictor[11] | TB profiler[12] |
|---------|------------|----------------------------------------|--------------|------------------------|-----------------|
| G1480   | 27/04/2009 | DS                                     | +/-          | +/-                    | +/-             |
| G1479   | 20/09/2013 | RR                                     | +/-          | +/-                    | +/-             |
| G516    | 21/01/2014 | MDR                                    | +/-          | +/-                    | +/-             |
| G520    | 17/03/2014 | MDR                                    | +/-          | +/-                    | +/-             |
| G249    | 27/06/2014 | MDR                                    | +/-          | +/-                    | +/-             |
| G252    | 11/08/2014 | MDR                                    | +/-          | +/-                    | +/-             |
| G535    | 13/10/2014 | MDR                                    | +/-          | +/-                    | +/-             |
| G841    | 09/12/2014 | MDR                                    | +/-          | +/-                    | +/-             |
| G842    | 27/01/2015 | MDR                                    | +/-          | +/-                    | +/-             |
| G993    | 15/04/2015 | MDR                                    | +/-          | +/-                    | +/-             |
| G1003   | 11/06/2015 | MDR                                    | +/-          | +/-                    | +/-             |
| G1257   | 09/11/2015 | MDR                                    | +/-          | +/-                    | +/-             |
| G1478   | 11/12/2015 | MDR                                    | +/-          | +/-                    | +/-             |
| G1720   | 02/06/2016 | XDR                                    | +/-          | +/-                    | +/-             |
| G1721   | 14/10/2016 | XDR                                    | +/-          | +/-                    | +/-             |
| G1928   | 09/01/2017 | XDR                                    | +/-          | +/-                    | +/-             |

a Based on WGS data.

Predicted whole-genome-based DST (WGS-DST) using publicly available software and after this study

Abbreviations: DS, drug susceptible; MDR, multidrug resistance; RR, rifampicin resistant; XDR, extensively drug-resistance.
### Supplementary Table 6. Primers used for amplicon sequencing of relevant rpoB, katG and ahpC regions

| Amplicon target* | Primer sequences | Product length (bp)** | Reference |
|------------------|------------------|-----------------------|-----------|
| rpoB RRDR (761020-761233) | 5-CGATCACACCCGCAGACGTT-3 5-GTTTCGATCGGGGACATCC-3 | 232 | [15] |
| rpoB 491 (761126-761362) | 5-GTCGGGGTGGACCCACAAG-3 5-CAGGTACACGATCTCGTCGC-3 | 256 | This study |
| katG 315 (2155074-2155345) | 5-CCATGAACGACGTCGAAACAG-3 5-GCTCTTCGTCAGCTCCCACCTC-3 | 272 | [15] |
| katG 429del (2154941-2154723) | 5-AGACAGTCAATCCCGATGCC-3 5-GCGGGTGGATCCGATCTATG-3 | 257 | This study |
| oxyR-ahpC promoter (2726015-2726251) | 5-ACCACTGCTTTTGCCGCCACC-3 5-CGATGAGAGCGGTGAGCTG-3 | 236 | [22] |
| ahpC 192 (2726608-2726856) | 5-ACCCCAACACGAGATCCAG-3 5-GATGTCTTTGGCGTACTCGG-3 | 218 | This study |

*Positions correspond to H37Rv genome. NCBI Reference Sequence NC_000962.3

**Length includes Illumina adapter sequences

Abbreviations: RRDR, rifampicin-resistance-determining-region
