Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance

World Health Organization

Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance

World Health Organization

Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance

ISBN 978-92-4-002817-3 (electronic version)

ISBN 978-92-4-002818-0 (print version)

Suggested citation. Catalogue of mutations in *Mycobacterium tuberculosis* complex and their association with drug resistance. Geneva: World Health Organization; 2021. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at http://apps.who.int/iris.

Sales, rights and licensing. To purchase WHO publications, see http://apps.who.int/bookorders. To submit requests for commercial use and queries on rights and licensing, see http://www.who.int/about/licensing.

Contents

Acknowledgements

Development of this document was led by Nazir Ahmed Ismail, Carl-Michael Nathanson and Alexei Korobitsyn, with support from Matteo Zignol and under the overall direction of Tereza Kasaeva, Director of the WHO Global TB Programme. The document is based on systematic analysis of a large collection of isolates of *Mycobacterium tuberculosis* complex with whole-genome sequencing and phenotypic drug-susceptibility testing. Data analysis, development of the catalogue and preparation of the report were conducted by four people, **Timothy Rodwell**, Foundation for Innovative New Diagnostics (FIND), Geneva, Switzerland and University of California, San Diego, USA; **Paolo Miotto**, San Raffaele Scientific Institute, Milan, Italy; **Claudio Köser**, University of Cambridge, United Kingdom; **Timothy Walker**, Oxford University, United Kingdom. Further support was provided by: **Philip Fowler**, Oxford University, United Kingdom; **Jeff Knaggs**, Oxford University, United Kingdom; **Zamin Iqbal**, European Bioinformatics Institute, London, United Kingdom; **Martin Hunt** European Bioinformatics Institute, London, United Kingdom; **Leonid Chindelevitch**, Imperial College, London, United Kingdom; **Maha Farhat**, Harvard Medical School, Boston (MA), USA; **Daniela Cirillo**, San Raffaele Scientific Institute, Milan, Italy; **Derrick Crook**, Oxford University, United Kingdom; **Iñaki Comas**, Biomedicine Institute of Valencia, Spain**; Jamie Posey**, Centers for Disease Control and Prevention, Atlanta (GA), USA; **Shaheed Vally Omar**, National Institute for Communicable Diseases, Pretoria, South Africa; **Tim Peto**, Oxford University, United Kingdom; **Sarah Walker,** Oxford University, United Kingdom. Reviewing and editing of drafts was provided by **Anita Suresh**, **Swapna Uplekar**, **Sacha Laurent** and **Rebecca Colman,** FIND, Geneva, Switzerland. This work was funded by grants provided by Unitaid and the United States Agency for International Development.

WHO is grateful to the following people who provided technical input:

Heidi Albert, FIND, Cape Town, South Africa; **Emmanuel André**, University Hospitals Leuven, Belgium; **Arnold Bainomugisa**, Queensland Department of Health, Brisbane, Australia; **Marie Ballif**, University of Bern, Switzerland; **Ivan Barilar**, Forschungszentrum Borstel, Germany; **Draurio Barreira Cravo Neto**, Unitaid, Geneva, Switzerland; **Alain Baulard**, Institut Pasteur de Lille, France; **Erik Böttger**, University of Zürich, Switzerland; **Francesc Coll**, London School of Hygiene & Tropical Medicine, London, United Kingdom; **Teresa Cortes**, London School of Hygiene and Tropical Medicines, United Kingdom; **Chris Coulter**, Queensland Department of Health, Brisbane, Australia; **Alan Cristoffels**, University of the Western Cape, Bellville, South Africa; **James Dawson**, Centers for Disease Control and Prevention, Atlanta (GA), USA; **Matthias Egger**, University of Bern, Switzerland; **Victoria Furió**, Universitat de València, Spain; **Sebastien Gagneux**, Swiss Tropical and Public Health Institute, Basel, Switzerland; **Wayne van Gemert**, Stop TB Partnership, Geneva, Switzerland; **Patricia Hall**, Centers for Disease Control and Prevention, Atlanta (GA), USA; **Zahra Hasan**, The Aga Khan University, Karachi, Pakistan; **Harald Hoffmann**,

Institute of Microbiology and Laboratory Medicine, Gauting, Germany; **Babak Javid**, University of California, San Francisco (CA), USA; **Moses Joloba**, National Reference Laboratory of the National TB and Leprosy Programme, Kampala, Uganda; **George Kasule**, National Reference Laboratory of the National TB and Leprosy Programme, Kampala, Uganda; **Peter Keller**, University of Bern, Switzerland; **Ellis Kelly**, University of Cambridge, United Kingdom; **Marguerite Massinga Loembé**, Africa CDC, Libreville, Gabon; **Alberto Mendoza**, Ministry of Health, Lima, Peru; **Matthias Merker**, Forschungszentrum Borstel, Germany; **John Metcalf**, University of California San Francisco (CA), USA; **Satoshi Mitarai**, Research Institute of Tuberculosis, Tokyo, Japan; **Stefan Niemann**, Forschungszentrum Borstel, Germany; **Sushil Pandey**, Queensland Department of Health, Brisbane, Australia; **Amy Piatek**, US Agency for International Development (USAID), Washington DC, USA; **Paola Rancoita**, Vita-Salute San Raffaele University, Milan, Italy; **Leen Rigouts**, Institute of Tropical Medicine, Antwerp, Belgium; **Jaime Robledo**, Corporación Para Investigaciones Biológicas, Medellín, Colombia; **Camilla Rodrigues**, P.D. Hinduja National Hospital and Medical Research Centre, Mumbai, India; **Michael Whitfield**, Institute of Tropical Medicine, Antwerp, Belgium; **Zhao Yanlin**, National Tuberculosis Reference Laboratory, China CDC, Beijing, China; **Danila Zimenkov**, Engelhardt Institute of Molecular Biology, Moscow, Russian Federation. Technical editing by Elisabeth Heseltine and design by Inis Communication.

Abbreviations and acronyms

7H10	Middlebrook 7H10
7H11	Middlebrook 7H11
BMD	broth microdilution
bp	base pairs
CC	critical concentration
CI	confidence interval
CLSI	Clinical and Laboratory Standards Institute
DST	drug-susceptibility testing
FDR	false discovery rate
FRS	fraction of supporting reads
GPI	genotype–phenotype intersection
indel	insertion/deletion
LoF	loss of function
MGIT	BACTEC™ Mycobacterial Growth Indicator Tube™ 960
MIC	minimum inhibitory concentration
MODS	microscopic observation drug-susceptibility assay
MTBC	*Mycobacterium tuberculosis* complex
OR	odds ratio
OR SOLO	odds ratio of solo mutations
PPV	positive predictive value
R	resistant or resistance
RRDR	rifampicin resistance-determining region
S	susceptible or susceptibility
SNP	single nucleotide polymorphism
TB	tuberculosis
U	uncertain
VCF	variant calling file
WGS	whole-genome sequencing

Drugs

AMK	amikacin
BDQ	bedaquiline
CAP	capreomycin
CFZ	clofazimine
DLM	delamanid
EMB	ethambutol
ETO	ethionamide
FQ	fluoroquinolone
INH	isoniazid
KAN	kanamycin
LFX	levofloxacin
LZD	linezolid
MFX	moxifloxacin
OFX	ofloxacin
PTO	prothionamide
PZA	pyrazinamide
RIF	rifampicin
STM	streptomycin

The terms and abbreviations used in the tables are listed in Table 3.

Introduction

A total of 1.4 million people died of tuberculosis (TB) in 2019, and approximately 10 million more developed active TB disease due to *Mycobacterium tuberculosis* complex (MTBC). Of the latter 10 million, an estimated 500 000 had TB resistant to rifampicin (RIF), and one million had TB susceptible to RIF but resistant to isoniazid (INH). Collectively, resistance to either of these two first-line anti-TB drugs is approximately 15%. This must be detected rapidly and accurately to initiate appropriate alternative treatment *(1)*.

The detection of RIF resistance has improved significantly with the introduction of rapid diagnostic tools that require less complex infrastructure and are simpler to perform *(2)*. Only 7% of patients with bacteriologically confirmed TB globally were tested for RIF resistance in 2012, while 61% were tested in 2019. Commensurate with the increased detection, the number of individuals started on multi-drug- or RIF-resistant TB treatment increased by 129% over the same period, from 77 321 to 177 099, highlighting the central role of diagnostics in the TB response *(1, 3)*. The molecular basis of RIF resistance is primarily mutations in the RIF resistance-determining region (RRDR), an 81-base-pair fragment of the *rpoB* gene *(4)*. This knowledge and research and development of new molecular tools have been important in the delivery of solutions for policy and impact *(5)*.

WHO recommends routine testing for resistance to RIF and INH of all TB patients, while fluoroquinolones (FQs) should be tested for use in RIF- and INH resistant TB *(5)*. The mechanisms of resistance to INH and FQs are well understood, and molecular tools for their detection are commercially available *(6)*; however, the genotypic drug-susceptibility testing (DST) assays for these drugs are less sensitive than those for resistance to RIF, and additional phenotypic testing is necessary to detect resistance missed by genotypic DST *(5)*.

After a long period of stagnation in TB treatment, the introduction of new drugs and repurposing of existing antimicrobial agents for the treatment of TB have generated optimism for improved TB therapies. As resistance to these new and repurposed drugs in the community gradually increases, however, concern has been raised about the lack of options for rapid detection of resistance. Novel rapid molecular tools are needed; however, the molecular basis of resistance to these drugs is still poorly understood, and therefore phenotypic DST methods are currently recommended to detect clinically relevant resistance *(7)*. In order to advance genotypic DST for the new drugs, matched phenotypic DST and sequencing data are necessary to identify and understand the mutations associated with resistance phenotypes. Thus, while detection of resistance to RIF requires molecular interrogation of just 81 nucleotides of the RRDR region of one gene (*rpoB*), detection of resistance to all the existing, repurposed and new drugs would require molecular interrogation of over 90 genes in the full genome of MTBC, which is over 4 million bases long *(7, 8)*.

Novel molecular assays for comprehensive genotypic DST, particularly next-generation sequencing, are potential solutions *(7)*. These include whole-genome sequencing (WGS), which is usually performed on culture isolates, as direct testing of clinical samples by WGS results in sequencing all the genetic material, including vast amounts of human DNA and that of other commensal organisms. In contrast, targeted next-generation sequencing can be applied directly to clinical samples. This is similar to current molecular testing in that it amplifies only the genetic targets of interest *(2)*. Instead of probes to detect variants, however, deep sequencing of the amplified fragments is performed, providing nucleotide-level detail as well as high-resolution detection of minor variants in mixed populations.

A major limitation to the development and diagnostic utility of sequence-based technologies and of next-generation molecular diagnostics for comprehensive genotypic DST is the lack of a standardized, comprehensive catalogue of mutations and their association with drug resistance for use by test developers and end users. Continuing technical uncertainty about the number, identity and clinical interpretation of genomic resistance-determining regions has limited broad uptake and the clinical relevance of these tests, especially for new and repurposed drugs *(2, 9)*. A high-quality, comprehensive catalogue of confidence-graded MTBC genetic markers of phenotypic resistance is necessary to distinguish clinically significant resistant variants from those that are not associated with resistance or those for which there are already sufficient data.

Although *in vitro* allelic exchange experiments are the reference standard for demonstrating that a specific mutation is both necessary and sufficient to confer phenotypic resistance, these approaches are expensive, slow and technically demanding and cannot be used to identify novel resistance genes or to identify genome-wide variants of clinical importance *(10)*. Therefore, association studies with WGS and associated standardized phenotypic DST data are indispensable for comprehensive investigation of resistance-associated mutations, particularly in non-essential genes, where hundreds of loss-of-function (LoF) mutations can result in resistance *(7, 11)*.

Previous work to create such a catalogue was based on reviews of published data, but that approach is limited by lack of standardization of genotyping methods in data sources and by the scope and scale of the available data *(6, 12)*. For the analysis presented here, we assembled WGS and phenotype data on the largest collection of multinational MTBC isolates to date (> 38 000) to establish the basis for the first WHO-endorsed catalogue of resistance-associated genetic variants for predicting clinically relevant resistance phenotypes from genetic data. This Mutation catalogue provides a common, standardized reference for interpretation of resistance to all first-line drugs (RIF, INH, ethambutol [EMB] and pyrazinamide [PZA]) and also to second-line drugs in group A (levofloxacin [LFX], moxifloxacin [MFX], bedaquiline [BDQ] and linezolid [LZD]), group B (clofazimine [CFZ]) and group C (delamanid [DLM], amikacin [AMK], streptomycin [STM], ethionamide [ETO] and prothionamide [PTO]). Kanamycin (KAN) and capreomycin (CAP) are no longer endorsed for TB treatment. They are included for historical interest and also because KAN provides useful insights for interpretation of some mutations that confer resistance to AMK *(13, 14)*. This report describes the methods used, the mutations identified and summaries of the key findings for each drug. Areas for future research are also outlined.

Development of the mutation catalogue

A detailed description of the methods used to generate the catalogue is provided in the section "Detailed methods". A summary is provided here. Data were collected from many contributors globally, who are listed under "Data contributors". The data included those from phenotypic DST and from WGS of cultured MTBC isolates.

Four primary components were essential for developing the catalogue:

1. **High-quality phenotypic DST**

   The phenotypic DST results were curated to ensure use of the best phenotypic reference. As different methods were used over the long period covered and critical concentrations (CCs) changed over time, the phenotypic methods were rank ordered. For specific isolates with multiple phenotypes, a hierarchy of selection was used, in which the most recent WHO-endorsed DST methods were ranked highest and older or non-WHO methods were ranked lower. A large dataset on broth microdilution (BMD) DST was included, which contributed strongly to data on new and repurposed drugs. As this method and criteria for its interpretation have not been reviewed or endorsed by WHO, findings based on these data alone were classified as "interim" associations in the catalogue.

2. **High-quality, standardized WGS for generating unbiased raw sequence data**

   Only WGS data that were generated with Illumina instruments were included. This platform is currently the most widely used, and ensured that the sequencing data provided were standardized. The output file of raw sequencing reads was used as the starting point for the bioinformatics analysis.

3. **A standardized bioinformatics pipeline for variant detection and annotation**

   To ensure that mutations were identified uniformly in all the data sources, a standardized bioinformatic pipeline was used to remove non-TB reads, process the data through quality checks, align the reads to the H37Rv reference genome and, finally, to identify mutations. The pipeline was designed to maximize variant detection of both single nucleotide polymorphisms (SNPs) and small insertions/deletions (indels) with a combination of tools. Heteroresistant isolates with a major variant of < 90% were excluded at this stage.

4. **A standardized, validated methodological approach to associating variants with phenotypes**

   The matched and curated phenotype and genotype data were then processed in an algorithm to identify "solo" mutations, i.e. single mutations within a set of genes of interest that best explain the observed drug-resistance phenotype, once neutral mutations (not associated with resistance) have been excluded. The final data for all mutations and phenotypes were then evaluated statistically for confidence grading. This included determining odds ratios (ORs) for the association of the variant with resistance and the positive predictive value (PPV). Because the dataset was large and heterogeneous, strict criteria were applied, and, when applicable, the bounds of the 95% confidence interval (CI) were used.

In the process outlined above, each mutation was classified as "associated with resistance", "not associated with resistance" or of "uncertain significance"; interim categories were included for "associated with resistance" and "not associated with resistance", resulting in a total of five groups. The interim categories are intended to reflect uncertainty in some of the observed associations and should be considered subject to change over time and with new evidence. Mutations in the interim categories consisted primarily of those found in scenarios in which less strict statistical grading thresholds were applied when appropriate (e.g. for PZA), when information from previous WHO guidance was used (e.g. for some *inhA* promoter mutations for INH), when expert rules were applied (when justified) and in other specific scenarios detailed later in the document.

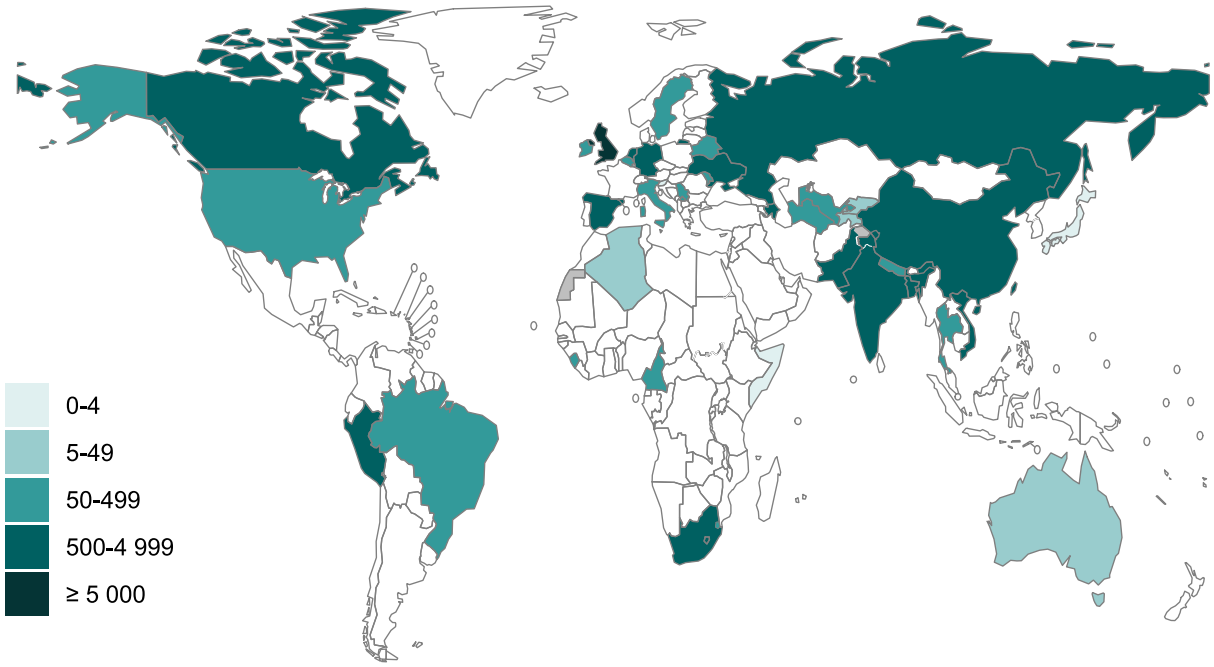
Overview of the mutation catalogue

A total of 50 396 MTBC isolates with phenotypic DST results were collated. Matching WGS data were available for 41 137 isolates. After additional quality control steps, 2922 were dropped from further consideration, leaving a total of 38 215 isolates for downstream analyses.

Geographical distribution of data

Forty-one countries contributed data on five or more MTBC isolates, 32 counties on ≥ 50 and 17 on ≥ 500. Only one (United Kingdom) contributed ≥ 5000 isolates (Fig. 1).

Fig. 1. Global distribution of sources of data on MTBC isolates used in this catalogue of genotype–phenotype associations



The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate borderlines for which there may not yet be full agreement.

Summary of phenotypic DST data

The number of phenotypic DST results available per isolate varied, results for the four first-line drugs (RIF, INH, EMB and PZA) being most common. The numbers of phenotypic DST results and of resistant phenotypes for each drug are listed in Table 1. The ALL dataset comprised phenotypic DST based on both WHO-endorsed methods (WHO dataset) and methods not endorsed by WHO. The latter dominated for BDQ, CFZ, DLM and LZD, but the prevalence of resistance to these drugs was ≤ 1.2%. More than 10 000 isolates were available in the ALL dataset for all drugs except BDQ and DLM, and there were more than 30 000 for some drugs. The prevalence of resistance to RIF and INH was 24–36%, whereas that for FQs was 14–20%, depending on the drug and dataset. Over 15 000 isolates were available for PZA, of which 14.6% were resistant by WHO-endorsed methods. The prevalence of resistance in the WHO dataset ranged from 0.6% for CFZ to 40.5% for ETO. In the ALL data, the range was from 0.9% for BDQ to 35.4% for INH.

Table 1. Phenotypic drug susceptibility results stratified by drug and dataset

Drug	Dataset[a]	Total no. of isolates	Resistant isolates No.	Percentage (95% CI)
Rifampicin	WHO	27 063	6 736	24.9 (24.4–25.4)
	ALL	34 375	9 868	28.7 (28.2–29.2)
Isoniazid	WHO	26 727	8 440	31.6 (31.0–32.1)
	ALL	34 437	12 199	35.4 (34.9–35.9)
Ethambutol	WHO	23 706	3 615	15.2 (14.8–15.7)
	ALL	30 708	4 900	16.0 (15.5–16.4)
Pyrazinamide	WHO	15 903	2 329	14.6 (14.1–15.2)
	ALL	15 902	2 329	14.6 (14.1–15.2)
Levofloxacin	WHO	10 305	2 019	19.6 (18.8–20.4)
	ALL	18 277	3 108	17.0 (16.5–17.6)
Moxifloxacin	WHO	6 904	1 094	15.8 (15.0–16.7)
	ALL	13 351	1 869	14.0 (13.4–14.6)
Bedaquiline	WHO	88	3	3.4 (0.7–9.6)
	ALL	8 321	73	0.9 (0.7–1.1)
Linezolid	WHO	1 131	9	0.8 (0.4–1.5)
	ALL	11 018	123	1.1 (0.9–1.3)
Clofazimine	WHO	3 635	23	0.6 (0.4–0.9)
	ALL	10 179	125	1.2 (1.0–1.5)
Delamanid	WHO	89	2	2.2 (0.3–7.9)
	ALL	7 778	82	1.1 (0.8–1.3)
Amikacin	WHO	8 040	664	8.3 (7.7–8.9)
	ALL	16 978	1 288	7.6 (7.2–8.0)
Streptomycin	WHO	9 043	2 562	28.3 (27.4–29.3)
	ALL	13 984	4 635	33.1 (32.4–33.9)
Ethionamide	WHO	2 184	884	40.5 (38.4–42.6)
	ALL	13 918	2 965	21.3 (20.6–22.0)
Kanamycin[a]	WHO	7 381	688	9.3 (8.7–10.0)
	ALL	16 154	1 481	9.2 (8.7–9.6)
Capreomycin[a]	WHO	9 103	702	7.7 (7.2–8.3)
	ALL	11 526	970	8.4 (7.9–8.9)

[a] Phenotypic DST dataset: The WHO dataset comprises phenotypic DST results obtained according to current or previous WHO guidelines, whereas the ALL dataset includes results obtained by additional methods not endorsed by WHO.

[b] No longer endorsed for TB treatment.

Performance of confidence-graded mutations for predicting phenotypic drug susceptibility

The sensitivity, specificity and PPV of confidence-graded mutations as predictors of phenotypic drug susceptibility are presented in Table 2. The performance metrics are presented by level of association (group 1, associated with R; group 2, associated with R – interim) as well as in combination, to indicate the contribution of the different confidence-graded mutations to overall performance. Important limitations to the estimated performance of mutations for predicting phenotypic drug susceptibility are due to imperfect reference standards, CCs that have changed over time and the noise inherent in phenotypic DST data collected from hundreds of different laboratories. These limitations are discussed in detail for each drug later in the report. KAN and CAP are included only for information, as these drugs are no longer endorsed for the treatment of TB.

Table 2. Sensitivity, specificity and PPV of the confidence-graded mutations as predictors of phenotypic drug susceptibility based on the ALL dataset

		1) Associated with R	2) Associated with R – Interim	3) Uncertain significance	4) NOT Associated with R – Interim	5) NOT Associated with R
RIF	# of variants identified	24	111	1550	0	28
	sens, spec, PPV (% 95% CI)	92.3 (91.8-92.8), 98.3 (98.1-98.5), 95.6 (95.2-96.0)	3.5 (3.2-3.9), 99.8 (99.8-99.9), 88.7 (85.2-91.7)			
	Combined Performance	**93.8 (93.3-94.2), 98.2 (98.0-98.3), 95.4 (94.9-95.8)**				
INH	# of variants identified	5	118	2252	6 (1)	22
	sens, spec, PPV (% 95% CI)	90.0 (89.4-90.5), 98.5 (98.4-98.7), 97.1 (96.8-97.4)	2.8 (2.5-3.1), 99.9 (99.8-99.9), 91.8 (88.6-94.4)			
	Combined Performance	**91.2 (90.7-91.7), 98.4 (98.2-98.6), 96.9 (96.6-97.2)**				
EMB	# of variants identified	14	1	2641	5	39
	sens, spec, PPV (% 95% CI)	86.3 (85.3-87.3), 93.3 (93.0-93.6), 71.1 (69.9-72.2)	0.4 (0.2-0.6), 100.0 (99.9-100.0), 72.0 (50.6-87.9)			
	Combined Performance	**86.7 (85.7-87.6), 93.3 (93.0-93.6), 71.1 (69.9-72.2)**				

		1) Associated with R	2) Associated with R – Interim	3) Uncertain significance	4) NOT Associated with R – Interim	5) NOT Associated with R
PZA	# of variants identified	105	233 (5)	775	16 (1)	15
	sens, spec, PPV (% 95% CI)	56.8 (54.8-58.8), 99.1 (99.0-99.3), 91.9 (90.4-93.3)	15.6 (14.2-17.2), 99.6 (99.5-99.7), 88.1 (84.6-91.1)			
	Combined Performance	**72.3 (70.5-74.2), 98.8 (98.6-99.0), 91.1 (89.7-92.3)**				
LFX	# of variants identified	8	6	672	2	14
	sens, spec, PPV (% 95% CI)	83.1 (81.7-84.4), 98.7 (98.5-98.9), 93.0 (92.0-93.9)	2.4 (1.9-3.0), 99.6 (99.5-99.7), 54.8 (46.0-63.4)			
	Combined Performance	**84.4 (83.1-85.7), 98.3 (98.1-98.5), 91.1 (90.0-92.2)**				
MFX	# of variants identified	9	5	552	2	11
	sens, spec, PPV (% 95% CI)	87.3 (85.7-88.8), 91.8 (91.3-92.3), 63.6 (61.7-65.4)	1.3 (0.8-1.9), 99.7 (99.6-99.8), 40.0 (27.6-53.5)			
	Combined Performance	**87.7 (86.2-89.2), 91.6 (91.0-92.1), 62.9 (61.0-64.7)**				
BDQ	# of variants identified	0	0	533	0	4
LZD	variants	1	0	403	0	2
	sens, spec, PPV (% 95% CI)	38.2 (29.6-47.4), 99.8 (99.8-99.9), 73.4 (60.9-83.7)				
CFZ	# of variants identified	0	0	601	0	7
DLM	variants	0	1	359	0	0
	sens, spec, PPV (% 95% CI)		6.1 (2.0-13.7), 100.0 (99.9-100.0), 83.3 (35.9-99.6)			
AMK	# of variants identified	2	2	1543	0	24
	sens, spec, PPV (% 95% CI)	76.4 (74.0-78.7), 99.1 (98.9-99.2), 87.4 (85.3-89.3)	0.9 (0.4-1.5), 99.9 (99.9-100.0), 47.8 (26.8-69.4)			
	Combined Performance	**77.3 (74.9-79.5), 99.0 (98.9-99.2), 86.6 (84.5-88.5)**				

		1) Associated with R	2) Associated with R – Interim	3) Uncertain significance	4) NOT Associated with R – Interim	5) NOT Associated with R
STM	# of variants identified	12	166	1327	1 (1)	15
	sens, spec, PPV (% 95% CI)	75.2 (73.9-76.4), 98.0 (97.6-98.2), 94.8 (94.0-95.5)	7.8 (7.0-8.6), 97.5 (97.1-97.8), 60.5 (56.4-64.4)			
	Combined Performance	**82.4 (81.3-83.5), 95.4 (95.0-95.8), 89.9 (89.0-90.8)**				
ETO	# of variants identified	4	327	1099	0	0
	sens, spec, PPV (% 95% CI)	47.2 (45.3-49.0), 95.8 (95.4-96.2), 75.2 (73.2-77.2)	34.3 (32.6-36.0), 95.5 (95.1-95.9), 67.3 (64.9-69.7)			
	Combined Performance	**75.7 (74.1-77.3), 91.4 (90.8-91.9), 70.4 (68.8-72.0)**				
***KAN**	# of variants identified	7	1	594	1	12
	sens, spec, PPV (% 95% CI)	72.9 (70.6-75.2), 98.4 (98.2-98.6), 82.4 (80.3-84.5)	0.3 (0.1-0.7), 100.0 (100.0-100.0), 66.7 (22.3-95.7)			
	Combined Performance	**73.2 (70.9-75.4), 98.4 (98.2-98.6), 82.4 (80.2-84.4)**				
***CAP**	# of variants identified	5	33	1009	0	20
	sens, spec, PPV (% 95% CI)	64.3 (61.2-67.3), 98.5 (98.2-98.7), 79.4 (76.4-82.2)	5.1 (3.8-6.6), 99.8 (99.8-99.9), 75.4 (63.1-85.2)			
	Combined Performance	**69.4 (66.4-72.3), 98.3 (98.0-98.6), 79.1 (76.2-81.8)**				

* Drugs no longer endorsed for TB treatment

The numbers in parentheses for INH, PZA and STM were excluded from these calculations, as these represent mutations that were not found in the isolates analysed but were included on the basis of previous WHO guidance or the literature.

For most drugs, groups 1 and 2 mutations combined had a sensitivity of ≥ 75% and a specificity ≥ 95%. The 24 group-1 mutations for resistance to RIF had a sensitivity of 92.3% and a PPV of 95.6%, while the five group-1 mutations for INH had a combined sensitivity of 90.0% and PPV of 97.1%. The nine group-1 mutations had a sensitivity of 87.3% for moxifloxacin, and the 14 mutations for EMB had a sensitivity of 86.3%. Because few isolates were resistant to new and repurposed drugs, resistance association to these drugs could not be clearly determined. Some drugs had lower specificity than expected, which might be explained by issues with the phenotypic reference or the interpretive criteria used, which are discussed later where relevant. The interim category (group 2) had no large effect on the predicted performance of most drugs; however, the sensitivity for PZA and ETO increased from 56.8% to 72.3% and from 47.2% to 75.7%, respectively.

The main drivers of resistance are often only a few key mutations. This knowledge has facilitated the development of rapid nucleic acid amplification tests, and the new information presented here should facilitate further improvements. It is also apparent that comprehensive molecular DST would require a new approach, ideally to cover multiple genes. Furthermore, while each mutation may not always contribute a large proportion of the observed resistance, their cumulative effect is likely to be important. As more data and further evidence are generated, it is expected that some of the interim mutations may be moved into group 1, and some of the many mutations of uncertain significance may be found to be associated with resistance or neutrality, moving the field further forwards.

Even though overall predictive performance may be suboptimal for some drugs, this should be considered in the context of the prevalence of resistance to the drug in the dataset and in the community being tested. Thus, in practice, when the pre-test probability is higher (e.g. in high-burden settings or among people starting with RIF-resistant TB), the sensitivity of these mutations as predictors of phenotypic DST results will probably be different.

Identification of mutations that are not associated with resistance is also important. A conservative approach was used to classify such mutations, and the available data indicated that there are clearly many mutations in this category that had not been identified previously *(8)*. This is another important advance and should aid test developers and users of sequencing data in making better-informed interpretations of individual mutations. Additionally, manufacturers of technologies on the market are encouraged to update the interpretation of their tests to avoid systematic false results for resistance, where relevant *(15)*.

Future updates of the Mutation catalogue, with additional data, will probably address some of the shortcomings of this analysis. We encourage those with data, particularly on isolates resistant to new and repurposed drugs and on mutations not well represented in this catalogue, to contribute to this global initiative.

Mutation catalogue

Reading the tables

The terms and abbreviations used in the tables are listed in Table 3.

Table 3. Terms used in the report and their description

Term used in the report	Description
ALL only	Information only from the ALL dataset
Assoc w R	Associated with resistance
Assoc w RI	Associated with resistance – interim
FQ X-R	Fluoroquinolone cross-resistance
Inf	Infinity
Lit	Information from the literature
Not Assoc w R	Not associated with resistance
Not Assoc w RI	Not associated with resistance – interim
Prev. WHO	Previous WHO guidance
Uncert. Sig.	Uncertain significance
WHO-end. gDST	WHO-endorsed genotypic drug susceptibility testing assay
Drug	Name of drug
Variant (common name)	Mutation, with common name where relevant
Present_S	Number of susceptible isolates with the mutation
Absent_S	Number of susceptible isolates without the mutation
Present_R	Number of resistant isolates with the mutation
Absent_R	Number of resistant isolates without the mutation
Sensitivity	True positive rate of mutation
Specificity	True negative rate of mutation
PPV	Positive predictive value of mutation
PPV|SOLO*	Positive predictive value conditional on being solo
OR SOLO	Odds ratio of solo mutation
Initial confidence grading	Initial grouping of mutation
Dataset(s)	Dataset(s) used to derive the initial confidence grading
Additional grading criteria	Criterion for changing the initial confidence grading (e.g. previous WHO guidance or WHO-endorsed genotypic DST assays) to yield the final confidence grading
Final confidence grading	Final grouping of mutation after additional grading criteria were applied (if relevant)

Additional variables shown in full catalogue	Description
Tier	A-priori grouping of genomic regions; tiers 1 and 2
Algorithm pass	Algorithm pass during which mutation was classified; 0, prior to algorithm (i.e. neutral mutation); 1, first pass; 2, second pass
Genome position	Genomic position in H37Rv for indels, inter-genetic and ribosomal mutations
Present_SOLO_R	Resistant isolates with the solo mutation
Present_SOLO_SR	Sum of resistant and susceptible isolates with the solo mutation
Sensitivity*	True positive rate of mutation
Specificity*	True negative rate of mutation
PPV*	Positive predictive value of mutation
LR+*	Positive likelihood ratio of mutation
LR-*	Negative likelihood ratio of mutation
OR*	Odds ratio of mutation
OR SOLO*	Odds ratio of solo mutation
OR SOLO_FE-sig	Fisher's exact test for the false discovery rate (FDR)-corrected *P* for the OR SOLO; TRUE = FDR-corrected $P \leq 0.05$, FALSE = FDR-corrected $P > 0.05$
Neutral masked	0 = not masked; 1 = masked
Previous WHO guidance	NGS Guide 2018, Level of resistance to INH or MFX, RIF CC guide 2021, Miotto et al. (PubMed identifier 29284687) *(4, 6, 12, 13, 16, 17)*
WHO-endorsed genotypic DST assays	Hain GenoType MTBDR*plus* V2.0, Nipro Genoscholar NTM+MDRTB II, Cepheid Xpert MTB/RIF, Cepheid Xpert MTB/RIF Ultra, Hain GenoType MTBDR*s*l V2.0, Cepheid Xpert MTB/XDR, Nipro Genoscholar PZA-TB II, Truenat MTB-RIF Dx is WHO-endorsed but Molbio has not disclosed precisely which part of *rpoB* is interrogated *(4)*.

* The lower bound (lb) and upper bound (up) of the 95% CI for these figures are provided in additional columns.

The tables in this report were simplified and abridged to fit the page space (i.e. all group-3 mutations and some group-2 mutations are not shown, as stated in the footnotes to the relevant tables). If both datasets supported the initial confidence grading, the values for the ALL dataset are shown. The complete list of graded mutations is available as supplementary material (WHO-UCN-GTB-PCI-2021.7-eng.xlsx). The raw datasets are also available on request to the Prevention, Diagnosis, Treatment, Care and Innovation unit, Global TB department, WHO. The mutations in each table were rank ordered according to the final confidence grading and by sensitivity and specificity.

The thresholds used to define the initial confidence grading are listed below; if they were met, the entry is highlighted in the colour shown in parentheses, below):

Group 1: Associated with resistance

Mutations that met five criteria:

1. sum of resistant and susceptible isolates with the solo mutation (Present_SOLO_SR) ≥ 5 (red)
2. lower bound of 95% CI of PPV conditional on being solo (PPV|SOLO_lb) ≥ 25% (red)
3. OR > 1, which always applies if criterion 4 is met (red)
4. OR SOLO > 1 (red)
5. statistical significance of OR SOLO (OR SOLO_FE-sig) with Fisher exact FDR-corrected $P \leq 0.05$ (red)

Criteria 4 and 5 are merged in the "OR SOLO" column of the simplified tables in this report and are shown in red if both criteria were met.

Group 2: Associated with resistance – interim

Mutations that met "relaxed" criteria for *pncA*:

1. resistant isolates with the solo mutation (Present_SOLO_R) ≥ 2 (yellow)
2. PPV ≥ 50% (yellow)

Group 3: Uncertain significance

Mutations that did not meet the criteria for inclusion in group 1, 2, 4 or 5.

Group 4: Not associated with resistance – Interim

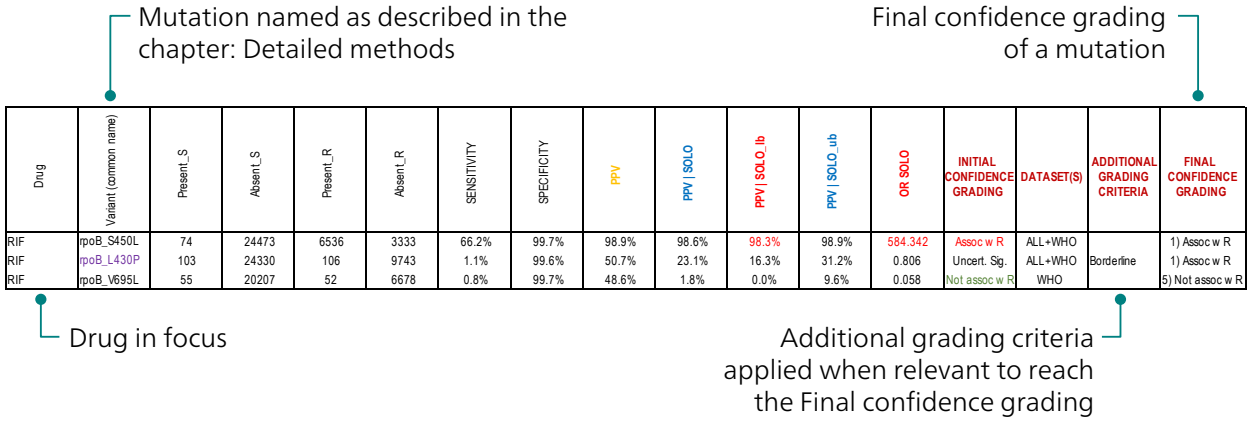
Mutations that met "relaxed" criteria for *pncA*:

1. PPV conditional on being solo (PPV|SOLO) < 40% (blue)
2. Upper bound of 95% CI of PPV conditional on being solo (PPV|SOLO_ub) < 75% (blue)

Group 5: Not associated with resistance

Neutral mutations that were masked before use of the algorithm (see "Initial identification of neutral mutations" under "Association studies")

An illustrative example



Drug	Variant (common name)	Present_S	Absent_S	Present_R	Absent_R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	DATASET(S)	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
RIF	rpoB_S450L	74	24473	6536	3333	66.2%	99.7%	98.9%	98.6%	98.3%	98.9%	584.342	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_L430P	103	24330	106	9743	1.1%	99.6%	50.7%	23.1%	16.3%	31.2%	0.806	Uncert. Sig.	ALL+WHO	Borderline	1) Assoc w R
RIF	rpoB_V695L	55	20207	52	6678	0.8%	99.7%	48.6%	1.8%	0.0%	9.6%	0.058	Not assoc w R	WHO		5) Not assoc w R

In the first example above, the drug considered is RIF. The variant is in the *rpoB* gene, the amino acid change is at codon 450, and the change is from serine to leucine (this corresponds to codon 531 in the old *Escherichia coli* nomenclature *(4, 18)*). This variant was found in 74 phenotypically susceptible isolates and in 6536 resistant isolates. The mutation was not found in 24 473 susceptible isolates or in another 3333 resistant isolates.

The sensitivity, specificity and PPV represent the performance of this mutation in the dataset. The next four columns indicate the statistical performance of this mutation when it occurs solo in the genomic regions selected when assessing RIF resistance. The values given are the midpoint PPV

and the corresponding lower bound (lb) and upper bound (ub), and the odds ratio for the solo mutation (OR SOLO).

The initial confidence grading for *rpoB* S450L was group 1 (Assoc w R) because:

- Present_SOLO_SR (see Mutation catalogue) was 5399 and, consequently, ≥ 5.
- PPV|SOLO_lb of 98.3% was ≥ 25%.
- OR SOLO of 584.342 was > 1 and statistically significant.

As the initial confidence grading for the WHO and for the ALL datasets concorded for this mutation, the figures shown are for the ALL dataset. Additional grading criteria were not applied to this mutation, and, therefore, the final confidence grading was unchanged. In contrast, the initial confidence grading for *rpoB* L430P was revised according to the expert rule related to borderline RIF resistance mutations. More details can be found in the section "Confidence grading" of the "Detailed methods".

Rifampicin

Only mutations in *rpoB* were found to be associated with RIF resistance. The 24 group-1 mutations (Assoc w R), which included two mutations outside the RRDR (V170F and I491F), and the remaining six borderline RIF-resistance mutations (L430P, D435Y, H445L, H445N, H445S and L452P) yielded a sensitivity of 92.3% (95% CI, 91.8–92.8%) for predicting phenotypic drug susceptibility in the ALL dataset. The 117 group-2 mutations (Assoc w RI) were all in the RRDR and had a combined sensitivity of only 3.5% (95% CI 3.2–3.9%). The vast majority of these RRDR mutations were too rare to meet the criteria for definitive classification into group 1 or 5. Instead, they were classified according to the expert rule that any RRDR mutation, with the exception of synonymous mutations, should be assumed to confer RIF resistance. This expert rule was first introduced by WHO in 2018 and reaffirmed in 2021 *(4, 19, 20)*.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
RIF	rpoB_S450L	74	24473	6536	3333	66.2%	99.7%	98.9%	98.6%	98.3%	98.9%	584.342	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_D435V	9	24424	732	9117	7.4%	100.0%	98.8%	98.7%	97.6%	99.4%	236.417	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_H445Y	4	24429	347	9502	3.5%	100.0%	98.9%	98.7%	96.7%	99.6%	392.067	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_H445D	3	24430	288	9561	2.9%	100.0%	99.0%	98.9%	96.9%	99.8%	234.224	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_D435Y	44	24389	162	9687	1.6%	99.8%	78.6%	58.9%	49.0%	68.3%	4.287	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_S450W	5	24428	151	9698	1.5%	100.0%	96.8%	96.2%	91.4%	98.8%	63.979	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_L452P	53	24380	121	9728	1.2%	99.8%	69.5%	59.5%	50.6%	68.0%	3.910	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_H445L	8	24425	115	9734	1.2%	100.0%	93.5%	92.9%	86.5%	96.9%	32.934	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_S450F	0	24433	112	9737	1.1%	100.0%	100.0%	100.0%	96.5%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_L430P	103	24330	106	9743	1.1%	99.6%	50.7%	23.1%	16.3%	31.2%	0.806	Uncert. Sig.	ALL+WHO	Borderline	1) Assoc w R
RIF	rpoB_H445R	2	24431	79	9770	0.8%	100.0%	97.5%	97.0%	89.5%	99.6%	80.020	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_V170F	0	24433	71	9778	0.7%	100.0%	100.0%	100.0%	90.0%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_I491F	57	24376	54	9795	0.5%	99.8%	48.6%	44.1%	34.3%	54.3%	2.113	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_H445N	39	24394	46	9803	0.5%	99.8%	54.1%	23.5%	12.8%	37.5%	0.786	Uncert. Sig.	ALL+WHO	Borderline	1) Assoc w R
RIF	rpoB_D435F	3	24430	39	9810	0.4%	100.0%	92.9%	92.1%	78.6%	98.3%	29.054	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_H445C	3	24430	36	9813	0.4%	100.0%	92.3%	91.4%	76.9%	98.2%	26.555	Assoc w R	ALL+WHO		1) Assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
RIF	rpoB_Q432K	1	24432	34	9815	0.3%	100.0%	97.1%	94.7%	74.0%	99.9%	44.807	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_Q432P	1	24432	29	9820	0.3%	100.0%	96.7%	95.2%	76.2%	99.9%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_S441L	0	24433	26	9823	0.3%	100.0%	100.0%	100.0%	79.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_Q432L	1	24432	21	9828	0.2%	100.0%	95.5%	94.1%	71.3%	99.9%	39.775	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_H445S	2	20260	12	6718	0.2%	100.0%	85.7%	80.0%	44.4%	97.5%	12.063	Assoc w R	WHO		1) Assoc w R
RIF	rpoB_S441Q	0	24433	11	9838	0.1%	100.0%	100.0%	100.0%	71.5%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_1296_ins_3_a_attc	0	24433	9	9840	0.1%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_1328_ins_3_t_tgac	0	24433	9	9840	0.1%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
RIF	rpoB_D435G	9	24424	76	9773	0.8%	100.0%	89.4%	35.7%	12.8%	64.9%	2.083	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
RIF	rpoB_L430R	1	24432	21	9828	0.2%	100.0%	95.5%	0.0%	0.0%	97.5%	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_H445Q	1	24432	16	9833	0.2%	100.0%	94.1%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_M434I	3	24430	16	9833	0.2%	100.0%	84.2%	25.0%	0.6%	80.6%	2.484	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_N437D	2	24431	13	9836	0.1%	100.0%	86.7%	0.0%	0.0%	84.2%	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_L449M	0	24433	9	9840	0.1%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_Q429L	0	24433	9	9840	0.1%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S450Q	0	24433	8	9841	0.1%	100.0%	100.0%	100.0%	54.1%	100.0%	Inf	Assoc w R	ALL	ALL only	2) Assoc w RI
RIF	rpoB_K446Q	0	24433	7	9842	0.1%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_H445G	3	24430	7	9842	0.1%	100.0%	70.0%	57.1%	18.4%	90.1%	3.310	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
RIF	rpoB_A451V	4	24429	7	9842	0.1%	100.0%	63.6%	0.0%	0.0%	60.2%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_R448Q	0	24433	6	9843	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S431G	0	24433	6	9843	0.1%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_Q429H	1	24432	6	9843	0.1%	100.0%	85.7%	0.0%	0.0%	97.5%	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_D435A	2	24431	6	9843	0.1%	100.0%	75.0%	0.0%	0.0%	84.2%	0.000	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
RIF	rpoB_D435E	0	24433	5	9844	0.1%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_H445P	0	24433	5	9844	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S428R	0	24433	5	9844	0.1%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
RIF	rpoB_S450M	0	24433	5	9844	0.1%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_T427I	0	24433	4	9845	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_M434V	3	24430	4	9845	0.0%	100.0%	57.1%	0.0%	0.0%	70.8%	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_D435H	0	24433	3	9846	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_H445T	0	24433	3	9846	0.0%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_N437H	0	24433	3	9846	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_Q432H	0	24433	3	9846	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S428T	0	24433	3	9846	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S431R	0	24433	3	9846	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_A451G	0	24433	2	9847	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_D435L	0	24433	2	9847	0.0%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_K446R	0	24433	2	9847	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_R448K	0	24433	2	9847	0.0%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S441M	0	24433	2	9847	0.0%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S441V	0	24433	2	9847	0.0%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S450Y	0	24433	2	9847	0.0%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_T444I	0	24433	2	9847	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_L443F	1	24432	2	9847	0.0%	100.0%	66.7%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_N437S	1	24432	2	9847	0.0%	100.0%	66.7%	0.0%	0.0%	97.5%	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S450V	1	24432	2	9847	0.0%	100.0%	66.7%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_T427A	1	24432	2	9847	0.0%	100.0%	66.7%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_G442E	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_H445F	0	24433	1	9848	0.0%	100.0%	100.0%	100.0%	2.5%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_K446E	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_K446T	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_L452V	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
RIF	rpoB_N437I	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_P439S	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_Q429P	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_Q432E	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_Q432N	0	24433	1	9848	0.0%	100.0%	100.0%	100.0%	2.5%	100.0%	Inf	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_Q436P	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S428G	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S428I	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S431T	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S441A	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S450A	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S450G	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_T427S	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_T444S	0	24433	1	9848	0.0%	100.0%	100.0%	NA	NA	NA	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_S450C	1	24432	1	9848	0.0%	100.0%	50.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_G426S	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_L452M	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_L452Q	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_M434L	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_N437Y	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_P439L	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_Q436N	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	NA	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_T427G	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoB_T427N	1	24432	0	9849	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	RRDR	2) Assoc w RI
RIF	rpoC_G594E	3795	16467	26	186	12.3%	81.3%	0.7%	0.6%	0.4%	0.9%	0.573	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoC_A172V	2637	17625	22	190	10.4%	87.0%	0.8%	0.3%	0.2%	0.6%	0.768	Not assoc w R	WHO		5) Not assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
RIF	rpoC_E1092D	540	19722	21	191	9.9%	97.3%	3.7%	3.4%	2.1%	5.3%	3.854	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_c-61t	3870	16392	635	6095	9.4%	80.9%	14.1%	0.6%	0.4%	0.9%	0.016	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoC_P601L	1389	18873	12	200	5.7%	93.1%	0.9%	0.0%	0.0%	0.3%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_a-261g	523	19739	5	207	2.4%	97.4%	0.9%	0.0%	0.0%	0.7%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoC_A621T	513	19749	2	210	0.9%	97.5%	0.4%	0.0%	0.0%	0.7%	NA	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_V695L	55	20207	52	6678	0.8%	99.7%	48.6%	1.8%	0.0%	9.6%	0.058	Not assoc w R	WHO		5) Not assoc w R
RIF	Rv2752c_G161S	65	20197	1	211	0.5%	99.7%	1.5%	0.0%	0.0%	5.5%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	Rv2752c_M31I	90	20172	1	211	0.5%	99.6%	1.1%	1.1%	0.0%	6.0%	1.062	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoA_E319K	388	19874	1	211	0.5%	98.1%	0.3%	0.0%	0.0%	0.9%	NA	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_E250G	71	20191	6	6724	0.1%	99.6%	7.8%	1.4%	0.0%	7.5%	0.043	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_I925V	36	20226	5	6725	0.1%	99.8%	12.2%	0.0%	0.0%	9.7%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_E639D	52	20210	5	6725	0.1%	99.7%	8.8%	0.0%	0.0%	6.8%	NA	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_S388L	53	20209	5	6725	0.1%	99.7%	8.6%	0.0%	0.0%	6.7%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_K944E	54	20208	1	6729	0.0%	99.7%	1.8%	0.0%	0.0%	6.6%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoC_E784Q	53	20209	0	212	0.0%	99.7%	0.0%	0.0%	0.0%	6.7%	NA	Not assoc w R	WHO		5) Not assoc w R
RIF	Rv2752c_K435E	54	20208	0	212	0.0%	99.7%	0.0%	0.0%	0.0%	6.6%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_2546_ins_6_g_gcgagga	55	20207	0	6730	0.0%	99.7%	0.0%	0.0%	0.0%	6.5%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoB_N381H	74	20188	0	6730	0.0%	99.6%	0.0%	0.0%	0.0%	4.9%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoC_R69P	80	20182	0	212	0.0%	99.6%	0.0%	0.0%	0.0%	4.5%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	Rv2752c_V300A	87	20175	0	212	0.0%	99.6%	0.0%	0.0%	0.0%	4.2%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoA_t-310a (rpsD_150)	103	20159	0	212	0.0%	99.5%	0.0%	0.0%	0.0%	3.5%	NA	Not assoc w R	WHO		5) Not assoc w R
RIF	Rv2752c_A273V	107	20155	0	212	0.0%	99.5%	0.0%	0.0%	0.0%	3.4%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	Rv2752c_A296V	123	20139	0	212	0.0%	99.4%	0.0%	0.0%	0.0%	3.0%	NA	Not assoc w R	WHO		5) Not assoc w R
RIF	Rv2752c_P123L	151	20111	0	212	0.0%	99.3%	0.0%	0.0%	0.0%	2.4%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoC_P906A	165	20097	0	212	0.0%	99.2%	0.0%	0.0%	0.0%	2.2%	0.000	Not assoc w R	WHO		5) Not assoc w R
RIF	rpoC_D271G	235	20027	0	212	0.0%	98.8%	0.0%	0.0%	0.0%	1.6%	0.000	Not assoc w R	WHO		5) Not assoc w R

Indels in the RRDR of *rpoB* classified as "Assoc w RI" in accordance with the associated expert rule (see "Additional criteria for final confidence grading") are not listed in this table but can be found in the Mutation catalogue. Borderline RIF resistance mutations are shown in purple and in-frame indels in gold.

Isoniazid

The combined sensitivity of groups 1 and 2 mutations for predicting phenotypic INH resistance in the ALL dataset was 91.2% (95% CI, 90.7–91.7%). They included four promoter mutations upstream of the *fabG1-inhA* operon and the g-154a change upstream of *inhA* (i.e. g609a in codon 203 of *fabG1*), which is known to confer low-level resistance by creating an alternative promoter for *inhA (17, 21)*. Only the three *katG* mutations (S315T, S315N and W328L) met the criteria for group 1. All *katG* mutations in group 2 were premature stop codons and indels that, in accordance with the newly endorsed expert rule, were assumed to result in a LoF phenotype and, consequently, high-level INH resistance *(4, 17)*. Other non-synonymous mutations in *katG* or upstream of *katG* can confer INH resistance but were rare in this dataset *(22, 23)*. *ahpC* promoter mutations were either too rare in this dataset to meet the criteria for markers of resistance or had a low PPV|SOLO. This is probably because these mutations function mostly or only as compensatory mutations (i.e. they typically coincide with *katG* mutations responsible for INH resistance *(4)*).

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
INH	katG_S315T	101	22144	9483	2717	77.7%	99.5%	98.9%	98.6%	98.3%	98.9%	718.422	Assoc w R	ALL+WHO		1) Assoc w R
INH	inhA_c-777t (fabG1_c-15t)	172	21965	2290	9882	18.8%	99.2%	93.0%	81.4%	78.8%	83.9%	10.689	Assoc w R	ALL+WHO		1) Assoc w R
INH	inhA_g-154a (fabG1_L203L)	50	22087	387	11785	3.2%	99.8%	88.6%	66.9%	58.8%	74.3%	3.944	Assoc w R	ALL+WHO		1) Assoc w R
INH	katG_S315N	2	22135	117	12055	1.0%	100.0%	98.3%	97.7%	91.9%	99.7%	78.037	Assoc w R	ALL+WHO		1) Assoc w R
INH	katG_W328L	0	22137	15	12157	0.1%	100.0%	100.0%	100.0%	78.2%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
INH	inhA_t-770c (fabG1_t-8c)	14	22123	127	12045	1.0%	99.9%	90.1%	33.3%	14.6%	57.0%	0.989	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
INH	inhA_t-770a (fabG1_t-8a)	5	22132	86	12086	0.7%	100.0%	94.5%	50.0%	18.7%	81.3%	2.289	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
INH	inhA_a-778g (fabG1_a-16g)	1	22136	0	12172	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	NA	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
INH	ndh_R268H	112	18089	10	343	2.8%	99.4%	8.2%	6.7%	2.9%	12.7%	3.767	Not assoc w R	WHO	Lit. (PMID 32143680)	4) Not assoc w RI
INH	katG_g-278c (furA_L68F)	137	22000	2	478	0.4%	99.4%	1.4%	1.4%	0.2%	5.1%	0.672	Uncert. Sig.	ALL+WHO	Prev. WHO	4) Not assoc w RI
INH	inhA_g-864a (Rv1482c_c-39t, fabG1_g-102a)	43	18158	17	8408	0.2%	99.8%	28.3%	0.0%	0.0%	8.2%	NA	Not assoc w R	WHO	Prev. WHO	4) Not assoc w RI
INH	inhA_V78A	49	18152	8	8417	0.1%	99.7%	14.0%	0.0%	0.0%	7.3%	NA	Not assoc w R	WHO	Lit. (PMID 32143680)	4) Not assoc w RI
INH	ndh_g-70t	43	18158	0	353	0.0%	99.8%	0.0%	0.0%	0.0%	8.2%	0.000	Not assoc w R	WHO	Lit. (PMID 32143680)	4) Not assoc w RI
INH	inhA_T4I														Prev. WHO	4) Not assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
INH	katG_R463L	8477	9724	5122	3303	60.8%	53.4%	37.7%	1.6%	1.4%	1.9%	0.100	Not assoc w R	WHO		5) Not assoc w R
INH	Rv1258c_581_ins_1_t_tg	2079	16122	81	272	22.9%	88.6%	3.8%	0.3%	0.1%	0.6%	3.175	Not assoc w R	WHO		5) Not assoc w R
INH	mshA_A187V	1976	16225	74	279	21.0%	89.1%	3.6%	0.0%	0.0%	0.2%	0.000	Not assoc w R	WHO		5) Not assoc w R
INH	ahpC_g-88a	3592	14609	941	7484	11.2%	80.3%	20.8%	0.1%	0.0%	0.2%	0.781	Not assoc w R	WHO		5) Not assoc w R
INH	mshA_N111S	2061	16140	18	335	5.1%	88.7%	0.9%	0.8%	0.5%	1.3%	0.398	Not assoc w R	WHO		5) Not assoc w R
INH	katG_c-85t (furA_L133L)	271	17930	11	342	3.1%	98.5%	3.9%	3.9%	2.0%	6.9%	2.128	Not assoc w R	WHO		5) Not assoc w R
INH	ndh_V18A	592	17609	10	343	2.8%	96.7%	1.7%	1.5%	0.7%	2.8%	0.787	Not assoc w R	WHO		5) Not assoc w R
INH	inhA_c-40t (fabG1_G241G)	509	17692	121	8304	1.4%	97.2%	19.2%	0.0%	0.0%	0.7%	0.000	Not assoc w R	WHO		5) Not assoc w R
INH	mshA_D218A	172	18029	5	348	1.4%	99.1%	2.8%	1.1%	0.1%	4.1%	1.502	Not assoc w R	WHO		5) Not assoc w R
INH	Rv2752c_P123L	127	18074	4	349	1.1%	99.3%	3.1%	3.1%	0.8%	7.6%	1.657	Not assoc w R	WHO		5) Not assoc w R
INH	Rv2752c_A296V	103	18098	3	350	0.8%	99.4%	2.8%	0.0%	0.0%	3.5%	NA	Not assoc w R	WHO		5) Not assoc w R
INH	Rv2752c_M31I	124	18077	3	350	0.8%	99.3%	2.4%	2.4%	0.5%	6.7%	1.260	Not assoc w R	WHO		5) Not assoc w R
INH	Rv2752c_K435E	54	18147	1	352	0.3%	99.7%	1.8%	0.0%	0.0%	6.6%	0.000	Not assoc w R	WHO		5) Not assoc w R
INH	ndh_G313R	71	18130	1	352	0.3%	99.6%	1.4%	0.0%	0.0%	5.1%	NA	Not assoc w R	WHO		5) Not assoc w R
INH	katG_c-354t (furA_A43V)	75	18126	1	352	0.3%	99.6%	1.3%	0.0%	0.0%	4.8%	0.000	Not assoc w R	WHO		5) Not assoc w R
INH	mshA_R443H	46	18155	0	353	0.0%	99.7%	0.0%	0.0%	0.0%	7.7%	0.000	Not assoc w R	WHO		5) Not assoc w R
INH	inhA_c-522g (fabG1_P81A)	53	18148	0	8425	0.0%	99.7%	0.0%	0.0%	0.0%	6.7%	NA	Not assoc w R	WHO		5) Not assoc w R
INH	katG_V469L	54	18147	0	8425	0.0%	99.7%	0.0%	0.0%	0.0%	6.6%	NA	Not assoc w R	WHO		5) Not assoc w R
INH	Rv1258c_E243A	54	18147	0	353	0.0%	99.7%	0.0%	0.0%	0.0%	6.6%	0.000	Not assoc w R	WHO		5) Not assoc w R
INH	Rv2752c_V300A	75	18126	0	353	0.0%	99.6%	0.0%	0.0%	0.0%	4.8%	0.000	Not assoc w R	WHO		5) Not assoc w R
INH	Rv2752c_A273V	107	18094	0	353	0.0%	99.4%	0.0%	0.0%	0.0%	3.4%	0.000	Not assoc w R	WHO		5) Not assoc w R
INH	Rv1258c_P414S	129	18072	0	353	0.0%	99.3%	0.0%	0.0%	0.0%	2.8%	0.000	Not assoc w R	WHO		5) Not assoc w R

Premature stop codons or indels in the coding regions of *katG* classified as "Assoc w RI" because of the associated expert rule (see "Additional criteria for final confidence grading") are not listed in this table but can be found in the Mutation catalogue.

Ethambutol

Only 14 non-synonymous mutations in *embB* and one inter-genic mutation upstream of *embA* met the criteria for group 1 or 2. The resulting sensitivity of 86.7% (95% CI, 85.7–87.6%) was good, but the specificity of 93.3% (95% CI, 93.0–93.6%) and the PPV of 71.1% (95% CI, 69.9–72.2%) were relatively low, because many *embB* mutations confer minimum inhibitory concentrations (MICs) close to the CC, resulting in poor categorical agreement with phenotypic DST *(10, 24–26).* Moreover, it is not clear whether the currently used CCs correspond well to the epidemiological cut-off values. Inappropriately high breakpoints may exacerbate the rate of misclassification of *embB* mutations, as was the case with borderline RIF resistance mutations *(20)*.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
EMB	embB_M306V	367	25441	1801	3099	36.8%	98.6%	83.1%	80.8%	79.0%	82.5%	36.528	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_M306I	625	25183	984	3916	20.1%	97.6%	61.2%	52.7%	50.0%	55.4%	7.638	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_Q497R	111	25697	497	4403	10.1%	99.6%	81.7%	71.2%	66.4%	75.7%	20.844	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embA_c-12t	99	25709	238	4662	4.9%	99.6%	70.6%	45.9%	38.5%	53.4%	5.939	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_D354A	153	25655	236	4664	4.8%	99.4%	60.7%	55.3%	49.8%	60.6%	7.426	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_G406A	120	25688	176	4724	3.6%	99.5%	59.5%	50.4%	43.9%	56.9%	6.031	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_Y319S	24	25784	121	4779	2.5%	99.9%	83.4%	83.3%	76.2%	89.0%	29.429	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_G406D	112	25696	117	4783	2.4%	99.6%	51.1%	40.7%	33.7%	48.1%	4.096	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_G406S	46	25762	79	4821	1.6%	99.8%	63.2%	42.5%	31.5%	54.1%	5.191	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_M306L	18	25790	68	4832	1.4%	99.9%	79.1%	73.5%	61.4%	83.5%	14.826	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_Q497K	25	25783	68	4832	1.4%	99.9%	73.1%	57.6%	44.1%	70.4%	8.246	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_D328Y	3	25805	24	4876	0.5%	100.0%	88.9%	81.3%	54.4%	96.0%	22.933	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_G406C	20	25788	24	4876	0.5%	99.9%	54.5%	42.9%	26.3%	60.6%	3.967	Assoc w R	ALL+WHO		1) Assoc w R
EMB	embB_Y319C	9	20082	14	3601	0.4%	100.0%	60.9%	52.6%	28.9%	75.6%	6.971	Assoc w R	WHO		1) Assoc w R
EMB	embB_L74R	7	25801	18	4882	0.4%	100.0%	72.0%	72.0%	50.6%	87.9%	13.590	Assoc w R	ALL+WHO	Pass 2	2) Assoc w RI
EMB	embB_T1082A	96	19995	99	3516	2.7%	99.5%	50.8%	0.0%	0.0%	3.8%	0.000	Not assoc w R	WHO	Lit. (PMID 32143680)	4) Not assoc w RI
EMB	embB_G156C	16	20075	31	3584	0.9%	99.9%	66.0%	0.0%	0.0%	20.6%	NA	Not assoc w R	WHO	Lit. (PMID 32143680)	4) Not assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
EMB	embA_A813G	16	20075	30	3585	0.8%	99.9%	65.2%	0.0%	0.0%	20.6%	NA	Not assoc w R	WHO	Lit. (PMID 32143680)	4) Not assoc w RI
EMB	embA_E951D	37	20054	4	3611	0.1%	99.8%	9.8%	0.0%	0.0%	9.5%	0.000	Not assoc w R	WHO	Lit. (PMID 32143680)	4) Not assoc w RI
EMB	embB_V668I	27	20064	0	3615	0.0%	99.9%	0.0%	0.0%	0.0%	12.8%	NA	Not assoc w R	WHO	Lit. (PMID 32143680)	4) Not assoc w RI
EMB	embC_R738Q	4049	16042	304	3311	8.4%	79.8%	7.0%	0.6%	0.4%	0.8%	0.032	Not assoc w R	WHO		5) Not assoc w R
EMB	ubiA_E149D	2950	17141	14	155	8.3%	85.3%	0.5%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embR_c-207g	2951	17140	14	155	8.3%	85.3%	0.5%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_V981L	3380	16711	256	3359	7.1%	83.2%	7.0%	0.5%	0.3%	0.8%	0.025	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_c-565t (aftA_456)	1936	18155	10	159	5.9%	90.4%	0.5%	0.5%	0.2%	0.9%	0.591	Not assoc w R	WHO		5) Not assoc w R
EMB	embR_C110Y	2556	17535	10	159	5.9%	87.3%	0.4%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embB_E378A	2947	17144	150	3465	4.1%	85.3%	4.8%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_T270I	2950	17141	150	3465	4.1%	85.3%	4.8%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embA_P913S	2556	17535	136	3479	3.8%	87.3%	5.1%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_N394D	2555	17536	135	3480	3.7%	87.3%	5.0%	0.0%	0.0%	0.1%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_c-1520a (aftA_138)	78	20013	3	166	1.8%	99.6%	3.7%	0.0%	0.0%	4.6%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	ubiA_V49I	79	20012	3	166	1.8%	99.6%	3.7%	0.0%	0.0%	4.6%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_g-1743a (aftA_63)	478	19613	3	166	1.8%	97.6%	0.6%	0.0%	0.0%	0.8%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embA_V206M	1330	18761	59	3556	1.6%	93.4%	4.2%	0.0%	0.0%	0.3%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_c-900t (aftA_344)	741	19350	2	167	1.2%	96.3%	0.3%	0.3%	0.0%	1.0%	0.340	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_A774S	95	19996	32	3583	0.9%	99.5%	25.2%	1.0%	0.0%	5.7%	0.063	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_a-20c (aftA_638)	55	20036	1	168	0.6%	99.7%	1.8%	0.0%	0.0%	6.5%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_c-1193t (aftA_247)	84	20007	1	168	0.6%	99.6%	1.2%	1.2%	0.0%	6.4%	1.418	Not assoc w R	WHO		5) Not assoc w R
EMB	ubiA_G268D	163	19928	1	168	0.6%	99.2%	0.6%	0.0%	0.0%	2.2%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embA_P639S	55	20036	8	3607	0.2%	99.7%	12.7%	0.0%	0.0%	6.5%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embA_V468A	55	20036	8	3607	0.2%	99.7%	12.7%	0.0%	0.0%	6.5%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_V104M	257	19834	7	3608	0.2%	98.7%	2.7%	0.0%	0.0%	1.4%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_R567H	330	19761	6	3609	0.2%	98.4%	1.8%	0.0%	0.0%	1.1%	NA	Not assoc w R	WHO		5) Not assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
EMB	embB_N13S	214	19877	5	3610	0.1%	98.9%	2.3%	0.0%	0.0%	1.7%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embB_S1054P	37	20054	1	3614	0.0%	99.8%	2.6%	0.0%	0.0%	9.5%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embA_P383S	56	20035	1	3614	0.0%	99.7%	1.8%	0.0%	0.0%	6.4%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_L661I	57	20034	1	3614	0.0%	99.7%	1.7%	0.0%	0.0%	6.3%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embA_A201T	94	19997	1	3614	0.0%	99.5%	1.1%	0.0%	0.0%	3.8%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embB_Q139H	97	19994	1	3614	0.0%	99.5%	1.0%	0.0%	0.0%	3.7%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_g-1419a (aftA_171)	41	20050	0	169	0.0%	99.8%	0.0%	0.0%	0.0%	8.6%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_-587_ins_18_a_atggccgccgccatcgggt (aftA_1345)	46	20045	0	169	0.0%	99.8%	0.0%	0.0%	0.0%	7.7%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embA_T308A	48	20043	0	3615	0.0%	99.8%	0.0%	0.0%	0.0%	7.4%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embR_A70S	55	20036	0	169	0.0%	99.7%	0.0%	0.0%	0.0%	6.5%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_c-589g (aftA_448)	59	20032	0	169	0.0%	99.7%	0.0%	0.0%	0.0%	6.1%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	ubiA_a-3741c (glf_a-99c)	60	20031	0	169	0.0%	99.7%	0.0%	0.0%	0.0%	6.0%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_g-270a (aftA_554)	64	20027	0	169	0.0%	99.7%	0.0%	0.0%	0.0%	5.6%	NA	Not assoc w R	WHO		5) Not assoc w R
EMB	embB_R213Q	117	19974	0	3615	0.0%	99.4%	0.0%	0.0%	0.0%	3.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_c-4586g (Rv3789_c-97g)	117	19974	0	169	0.0%	99.4%	0.0%	0.0%	0.0%	3.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
EMB	embC_c-100t (aftA_611)	214	19877	0	169	0.0%	98.9%	0.0%	0.0%	0.0%	1.7%	0.000	Not assoc w R	WHO		5) Not assoc w R

Pyrazinamide

Because of the large number of potential resistance mutations in *pncA* and its promoter, mutations found to be associated with phenotypic PZA resistance (according to the criteria employed for group 1) had a sensitivity of only 56.8% (95% CI, 54.8–58.8%). The sensitivity increased to 72.3% (95% CI, 70.5–74.2%) when group 2 (Assoc w RI) mutations were included. The vast majority of group-2 mutations were included according to the expert rule that any nonsense mutation and indel in the coding region of *pncA* should be assumed to confer an LoF phenotype and, therefore, PZA resistance *(27)*. This expert rule was applied irrespective of whether the isolate in question was resistant to any other drug; this represents a change from previous WHO guidance, which restricted this expert rule to nonsense mutations and frameshifts *(6, 12)*. Notably, *pncA* T47A and I31T, which were previously considered neutral by WHO, were reclassified here as definitive and interim markers for resistance, respectively, on the basis of the data used in this analysis or evidence from the literature *(6, 12, 28)*. Both mutations were originally misclassified because their MIC distributions were close to the CC *(28)*.

Sixteen *pncA* mutations were found to be neutral according to the interim criteria (group 4). Given that 13 of these mutations fall into the *pncA* region interrogated by the Nipro Genoscholar PZA-TB II, they might result in systematic false resistance in this assay *(29)*. No *pncA* mutations met the definitive criteria for neutrality (group 5), as such *pncA* mutations are rare *(28, 30)*. Modlin et al. *(31)* raised the possibility that *clpC1* V63A, which appears to be a marker for lineage 1 of MTBC, is responsible for elevated PZA MICs of this lineage compared with other MTBC genotypes (except *M. bovis* and *M. canettii*) and that the current CC of 100 mg/L may divide the upper end of the MIC distribution of lineage 1, resulting in a higher rate of PZA mono-resistance for this lineage. This hypothesis was not contradicted by the fact that, in this analysis, *clpC1* V63A was found to be a group-5 neutral mutation because of a PPV of only 4.6% (95% CI, 3.2–6.3%) in set B. (See Fig. 4a in "Association studies" under "Detailed methods".)

M. canettii, which is rarely found outside of the Horn of Africa, is intrinsically resistant to PZA *(32, 33)*. The genetic basis of this phenotype is unclear (more than one mechanism may be involved), and most of the isolates described to date do not have non-synonymous mutations or indels in *pncA (34, 35)*. Therefore, this intrinsic resistance must be diagnosed indirectly through a phylogenetically informative surrogate (e.g. *M. canettii* has a synonymous mutation at codon 46 of *pncA* [a138g]).

Although the number of mutations classified into groups 1, 2, 4 and 5 represented a significant advance over previous WHO classifications, it is acknowledged that they comprise only a small proportion of possible non-synonymous mutations in *pncA (28)*. Therefore, an additional expert rule for *pncA* is proposed, which is not directly applied to the tables but suggested for routine practice to complement the final confidence grading. Specifically, any novel, non-synonymous mutation (i.e. not already classified into group 1, 2, 4 or 5) and all non-synonymous mutations in group 3 should be assumed to confer PZA resistance if they occur in an isolate that is resistant to RIF, given that the pre-test probability of PZA resistance is high under these circumstances *(36–38)*.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_H57D	3	13570	157	2172	6.7%	100.0%	98.1%	98.1%	94.5%	99.6%	318.633	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_a-11g	3	13570	85	2244	3.6%	100.0%	96.6%	96.6%	90.4%	99.3%	171.338	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Q10P	1	13572	58	2271	2.5%	100.0%	98.3%	98.3%	90.9%	100.0%	346.621	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_H51D	0	13573	33	2296	1.4%	100.0%	100.0%	100.0%	89.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_T76P	3	13570	30	2299	1.3%	100.0%	90.9%	90.9%	75.7%	98.1%	59.026	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_G97D	4	13569	28	2301	1.2%	100.0%	87.5%	73.3%	44.9%	92.2%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V7G	1	13572	27	2302	1.2%	100.0%	96.4%	96.4%	81.7%	99.9%	159.185	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_C14R	1	13572	24	2305	1.0%	100.0%	96.0%	96.0%	79.6%	99.9%	141.314	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_D49G	1	13572	24	2305	1.0%	100.0%	96.0%	96.0%	79.6%	99.9%	141.314	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_H57R	1	13572	24	2305	1.0%	100.0%	96.0%	96.0%	79.6%	99.9%	141.314	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Q141P	1	13572	24	2305	1.0%	100.0%	96.0%	96.0%	79.6%	99.9%	141.314	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L4S	3	13570	24	2305	1.0%	100.0%	88.9%	88.9%	70.8%	97.6%	47.098	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_I133T	13	13560	20	2309	0.9%	99.9%	60.6%	60.6%	42.1%	77.1%	9.035	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_H71Y	1	13572	19	2310	0.8%	100.0%	95.0%	95.0%	75.1%	99.9%	111.631	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_T135P	1	13572	19	2310	0.8%	100.0%	95.0%	95.0%	75.1%	99.9%	111.631	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_H51R	1	13572	17	2312	0.7%	100.0%	94.4%	94.4%	72.7%	99.9%	99.794	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_I31S	2	13571	17	2312	0.7%	100.0%	89.5%	89.5%	66.9%	98.7%	49.893	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V139A	3	13570	17	2312	0.7%	100.0%	85.0%	83.3%	58.6%	96.4%	29.347	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_D8G	2	13571	16	2313	0.7%	100.0%	88.9%	88.9%	65.3%	98.6%	46.938	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_A146V	0	13573	15	2314	0.6%	100.0%	100.0%	100.0%	78.2%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_W68C	0	13573	14	2315	0.6%	100.0%	100.0%	100.0%	76.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_H71R	1	13572	14	2315	0.6%	100.0%	93.3%	92.3%	64.0%	99.8%	70.352	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_K96T	0	13573	13	2316	0.6%	100.0%	100.0%	100.0%	75.3%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Q10R	0	13573	13	2316	0.6%	100.0%	100.0%	100.0%	71.5%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Y103!	1	13572	13	2316	0.6%	100.0%	92.9%	92.9%	66.1%	99.8%	76.181	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_T47A	4	13569	13	2316	0.6%	100.0%	76.5%	76.5%	50.1%	93.2%	19.041	Assoc w R	ALL+WHO		1) Assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_533_ins_1_gcggtgcgcatctcctccagcgcggcgacggtgg_gcggtgcgcatctcctcccagcgcggcgacggtgg	0	13573	12	2317	0.5%	100.0%	100.0%	100.0%	73.5%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_A134V	1	13572	12	2317	0.5%	100.0%	92.3%	92.3%	64.0%	99.8%	70.291	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_P54L	2	13571	12	2317	0.5%	100.0%	85.7%	85.7%	57.2%	98.2%	35.143	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_I6T	0	13573	11	2318	0.5%	100.0%	100.0%	100.0%	71.5%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_S59P	0	13573	11	2318	0.5%	100.0%	100.0%	100.0%	71.5%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_S67P	0	13573	11	2318	0.5%	100.0%	100.0%	100.0%	71.5%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_G97C	1	13572	11	2318	0.5%	100.0%	91.7%	91.7%	61.5%	99.8%	64.406	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L151S	4	13569	11	2318	0.5%	100.0%	73.3%	73.3%	44.9%	92.2%	16.098	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_a-11c	0	13573	10	2319	0.4%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V7L	0	13573	10	2319	0.4%	100.0%	100.0%	100.0%	69.2%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_W68G	0	13573	10	2319	0.4%	100.0%	100.0%	100.0%	69.2%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_W68R	0	13573	10	2319	0.4%	100.0%	100.0%	100.0%	69.2%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_G97S	1	13572	10	2319	0.4%	100.0%	90.9%	90.9%	58.7%	99.8%	58.525	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_R154G	4	13569	10	2319	0.4%	100.0%	71.4%	71.4%	41.9%	91.6%	14.628	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_D12A	5	13568	10	2319	0.4%	100.0%	66.7%	66.7%	38.4%	88.2%	11.702	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_F94C	0	13573	9	2320	0.4%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_G132S	0	13573	9	2320	0.4%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L172P	0	13573	9	2320	0.4%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L85P	0	13573	9	2320	0.4%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_T160P	0	13573	9	2320	0.4%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Y34D	0	13573	9	2320	0.4%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_-4_del_1_tc_t	1	13572	9	2320	0.4%	100.0%	90.0%	90.0%	55.5%	99.7%	52.650	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_F58L	3	13570	9	2320	0.4%	100.0%	75.0%	75.0%	42.8%	94.5%	17.547	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L182S	4	13569	9	2320	0.4%	100.0%	69.2%	69.2%	38.6%	90.9%	13.160	Assoc w R	ALL+WHO		1) Assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_390_del_4_cacat_c	0	13573	8	2321	0.3%	100.0%	100.0%	100.0%	63.1%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_G97R	0	13573	8	2321	0.3%	100.0%	100.0%	100.0%	59.0%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_H57Y	0	13573	8	2321	0.3%	100.0%	100.0%	100.0%	63.1%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L120P	0	13573	8	2321	0.3%	100.0%	100.0%	100.0%	63.1%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L27P	1	13572	8	2321	0.3%	100.0%	88.9%	88.9%	51.8%	99.7%	46.780	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_318_del_1_ga_g	2	13571	8	2321	0.3%	100.0%	80.0%	80.0%	44.4%	97.5%	23.388	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_D49A	0	13573	7	2322	0.3%	100.0%	100.0%	100.0%	59.0%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L85R	0	13573	7	2322	0.3%	100.0%	100.0%	100.0%	59.0%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Q122!	0	13573	7	2322	0.3%	100.0%	100.0%	100.0%	59.0%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V180F	0	13573	7	2322	0.3%	100.0%	100.0%	100.0%	59.0%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_D8N	1	13572	7	2322	0.3%	100.0%	87.5%	87.5%	47.3%	99.7%	40.915	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_P69L	1	13572	7	2322	0.3%	100.0%	87.5%	87.5%	47.3%	99.7%	40.915	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_T142M	1	13572	7	2322	0.3%	100.0%	87.5%	87.5%	47.3%	99.7%	40.915	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_A146T	2	13571	7	2322	0.3%	100.0%	77.8%	77.8%	40.0%	97.2%	20.456	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_I90S	3	13570	7	2322	0.3%	100.0%	70.0%	70.0%	34.8%	93.3%	13.636	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_M175V	3	13570	7	2322	0.3%	100.0%	70.0%	66.7%	29.9%	92.5%	11.688	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Y103C	3	13570	7	2322	0.3%	100.0%	70.0%	70.0%	34.8%	93.3%	13.636	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_518_ins_1_t_tc	0	13573	6	2323	0.3%	100.0%	100.0%	100.0%	54.1%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L159R	0	13573	6	2323	0.3%	100.0%	100.0%	100.0%	54.1%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_P62L	0	13573	6	2323	0.3%	100.0%	100.0%	100.0%	54.1%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V180G	0	13573	6	2323	0.3%	100.0%	100.0%	100.0%	54.1%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_A102P	1	13572	6	2323	0.3%	100.0%	85.7%	83.3%	35.9%	99.6%	29.212	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_D12G	1	13572	6	2323	0.3%	100.0%	85.7%	85.7%	42.1%	99.6%	35.055	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_G132A	1	13572	6	2323	0.3%	100.0%	85.7%	85.7%	42.1%	99.6%	35.055	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_H51Q	1	13572	6	2323	0.3%	100.0%	85.7%	85.7%	42.1%	99.6%	35.055	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V155G	1	13572	6	2323	0.3%	100.0%	85.7%	85.7%	42.1%	99.6%	35.055	Assoc w R	ALL+WHO		1) Assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_D63A	2	13571	6	2323	0.3%	100.0%	75.0%	71.4%	29.0%	96.3%	14.605	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_F94L	2	13571	6	2323	0.3%	100.0%	75.0%	75.0%	34.9%	96.8%	17.526	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V125F	3	13570	6	2323	0.3%	100.0%	66.7%	66.7%	29.9%	92.5%	11.683	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_389_del_9_acatcgacct_a	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_392_ins_1_a_ac	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_392_ins_2_a_acc	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_395_del_9_ccgaccacat_c	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_A143G	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_C138R	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_D12E	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_G105V	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_H71P	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_L4W	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_M1T	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Q141!	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V128G	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V130G	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V139G	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_Y103H	0	13573	5	2324	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_467_ins_1_agcaccctggtggccaa_agcaccctggtgggccaa	1	13572	5	2324	0.2%	100.0%	83.3%	83.3%	35.9%	99.6%	29.200	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_G24D	1	13572	5	2324	0.2%	100.0%	83.3%	83.3%	35.9%	99.6%	29.200	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_K96E	1	13572	5	2324	0.2%	100.0%	83.3%	83.3%	35.9%	99.6%	29.200	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_S164P	1	13572	5	2324	0.2%	100.0%	83.3%	83.3%	35.9%	99.6%	29.200	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_T142A	1	13572	5	2324	0.2%	100.0%	83.3%	83.3%	35.9%	99.6%	29.200	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_T177P	1	13572	5	2324	0.2%	100.0%	83.3%	83.3%	35.9%	99.6%	29.200	Assoc w R	ALL+WHO		1) Assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV| SOLO	PPV| SOLO_lb	PPV| SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_465_ins_1_c_ca	1	13572	4	2325	0.2%	100.0%	80.0%	80.0%	28.4%	99.5%	23.350	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_K96R	1	13572	4	2325	0.2%	100.0%	80.0%	80.0%	28.4%	99.5%	23.350	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_V7A	1	13572	4	2325	0.2%	100.0%	80.0%	80.0%	28.4%	99.5%	23.350	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_W119C	1	13572	4	2325	0.2%	100.0%	80.0%	80.0%	28.4%	99.5%	23.350	Assoc w R	ALL+WHO		1) Assoc w R
PZA	pncA_K48T	7	13566	6	2323	0.3%	99.9%	46.2%	46.2%	19.2%	74.9%	5.006	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
PZA	pncA_F13L	1	13572	5	2324	0.2%	100.0%	83.3%	75.0%	19.4%	99.4%	17.520	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_458_del_1_gt_g	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_a-11t	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_C72R	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_G105D	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_G162D	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_H51P	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_I5S	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_L120Q	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_L120R	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_L182W	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_Q10!	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_t-12c	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_V131F	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_V21G	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_V9G	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_W119!	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_W68!	0	13573	4	2325	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_35_ins_1_t_tc	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_398_ins_2_ataccgaccac atcgacctcatcgacgccgcgttgccgca _ataccgaccccacatcgacctcatcga cgccgcgttgccgca	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_457_ins_1_t_tg	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_A3E	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_F13I	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_G97V	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_H82R	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_M175T	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_P62T	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_Q10H	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_S164!	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_T47P	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_W119R	0	13573	3	2326	0.1%	100.0%	100.0%	100.0%	29.2%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_A46V	1	13572	3	2326	0.1%	100.0%	75.0%	75.0%	19.4%	99.4%	17.505	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_G17D	1	13572	3	2326	0.1%	100.0%	75.0%	75.0%	19.4%	99.4%	17.505	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_L19P	1	13572	3	2326	0.1%	100.0%	75.0%	75.0%	19.4%	99.4%	17.505	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_S104R	1	13572	3	2326	0.1%	100.0%	75.0%	75.0%	19.4%	99.4%	17.505	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_V155M	1	13572	3	2326	0.1%	100.0%	75.0%	75.0%	19.4%	99.4%	17.505	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_L116P	2	13571	3	2326	0.1%	100.0%	60.0%	60.0%	14.7%	94.7%	8.752	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_S18!	3	13570	3	2326	0.1%	100.0%	50.0%	50.0%	11.8%	88.2%	5.834	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_162_ins_1_c_cg	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_193_del_10_aatagtccgg tg_aa	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_288_ins_33_c_caccccgtc cacgatgacattcgagatcgccagg	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_294_del_1_ggcaccct_gg cccct	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_-3_del_1_gtc_gt	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_319_del_1_cga_cg	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_356_ins_1_c_ca	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_390_del_1_ca_c	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_398_del_9_ataccgaccacatcgacctcatcgacgccgcgttgccgca_ataccgacctcatcgacgccgcgttgccgca	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_408_ins_1_a_at	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_418_ins_1_g_gc	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_450_ins_1_g_gc	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_60_ins_1_c_cg	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_65_del_1_gt_g	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_8_del_1_gc_g	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_A171E	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_A46E	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_C72Y	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_D49E	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_D49N	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_D8A	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_E91!	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_F106S	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_F94S	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_G132D	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_H51Y	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_I90T	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_L172R	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_P54Q	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_S104I	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_T142K	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_T168P	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_T76I	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_t-7c	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_V131G	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_V180A	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_V44G	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_W119G	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_Y103D	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_Y103S	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_Y34!	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_Y99!	0	13573	2	2327	0.1%	100.0%	100.0%	100.0%	15.8%	100.0%	Inf	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_D8E	1	13572	2	2327	0.1%	100.0%	66.7%	66.7%	9.4%	99.2%	11.665	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_G108R	1	13572	2	2327	0.1%	100.0%	66.7%	66.7%	9.4%	99.2%	11.665	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_K48E	1	13572	2	2327	0.1%	100.0%	66.7%	66.7%	9.4%	99.2%	11.665	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_K96Q	1	13572	2	2327	0.1%	100.0%	66.7%	66.7%	9.4%	99.2%	11.665	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_P62S	1	13572	2	2327	0.1%	100.0%	66.7%	66.7%	9.4%	99.2%	11.665	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_D63G	2	13571	2	2327	0.1%	100.0%	50.0%	50.0%	6.8%	93.2%	5.832	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_F81V	2	13571	2	2327	0.1%	100.0%	50.0%	50.0%	6.8%	93.2%	5.832	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_Y64D	2	13571	2	2327	0.1%	100.0%	50.0%	50.0%	6.8%	93.2%	5.832	Assoc w RI	ALL+WHO		2) Assoc w RI
PZA	pncA_D12N	0	13573	1	2328	0.0%	100.0%	100.0%	100.0%	2.5%	100.0%	Inf	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
PZA	pncA_V125G	0	13573	1	2328	0.0%	100.0%	100.0%	100.0%	2.5%	100.0%	Inf	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
PZA	pncA_V139L	1	13572	0	2329	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
PZA	pncA_I31T														Lit. (PMID 32571824)	2) Assoc w RI
PZA	pncA_442_ins_12_cgc_cgcgacgcggtacgc														Prev. WHO	2) Assoc w RI
PZA	pncA_H71D														Prev. WHO	2) Assoc w RI
PZA	pncA_L116R														Prev. WHO	2) Assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV| SOLO	PPV| SOLO_lb	PPV| SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	pncA_T135N														Prev. WHO	2) Assoc w RI
PZA	pncA_D136N	4	13569	1	2328	0.0%	100.0%	20.0%	20.0%	0.5%	71.6%	1.457	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_A171V	6	13567	1	2328	0.0%	100.0%	14.3%	14.3%	0.4%	57.9%	0.971	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_I6L	6	13567	1	2328	0.0%	100.0%	14.3%	0.0%	0.0%	45.9%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_V21A	6	13567	1	2328	0.0%	100.0%	14.3%	0.0%	0.0%	45.9%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_L35R	23	13550	1	2328	0.0%	99.8%	4.2%	4.2%	0.1%	21.1%	0.253	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_A79T	3	13570	0	2329	0.0%	100.0%	0.0%	0.0%	0.0%	70.8%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_E15G	3	13570	0	2329	0.0%	100.0%	0.0%	0.0%	0.0%	70.8%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_T168I	3	13570	0	2329	0.0%	100.0%	0.0%	0.0%	0.0%	70.8%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_t-60g (Rv2044c_100)	3	13570	0	458	0.0%	100.0%	0.0%	0.0%	0.0%	70.8%	NA	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_A79V	4	13569	0	2329	0.0%	100.0%	0.0%	0.0%	0.0%	60.2%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_g-33a	5	13568	0	2329	0.0%	100.0%	0.0%	0.0%	0.0%	52.2%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_T114M	5	13568	0	2329	0.0%	100.0%	0.0%	0.0%	0.0%	52.2%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_S66L	8	13565	0	2329	0.0%	99.9%	0.0%	0.0%	0.0%	36.9%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_-2_ins_1_c_cg	10	13563	0	2329	0.0%	99.9%	0.0%	0.0%	0.0%	30.8%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_T87M	17	13556	0	2329	0.0%	99.9%	0.0%	0.0%	0.0%	19.5%	0.000	Not assoc w RI	ALL+WHO		4) Not assoc w RI
PZA	pncA_-125_del_1_cc_c														Prev. WHO	4) Not assoc w RI
PZA	PPE35_L896S	5996	7578	320	138	69.9%	55.8%	5.1%	0.2%	0.1%	0.4%	0.274	Not assoc w R	WHO		5) Not assoc w R
PZA	Rv3236c_T102A	2049	11525	202	256	44.1%	84.9%	9.0%	0.0%	0.0%	0.2%	NA	Not assoc w R	WHO		5) Not assoc w R
PZA	Rv1258c_581_ins_1_t_tg	2010	11564	201	257	43.9%	85.2%	9.1%	0.0%	0.0%	0.3%	14.999	Not assoc w R	WHO		5) Not assoc w R
PZA	clpC1_V63A	1146	12428	178	2151	7.6%	91.6%	13.4%	8.8%	7.3%	10.5%	0.579	Not assoc w R	WHO		5) Not assoc w R
PZA	PPE35_P822S	817	12757	23	435	5.0%	94.0%	2.7%	1.8%	1.0%	3.0%	1.134	Not assoc w R	WHO		5) Not assoc w R
PZA	Rv3236c_A370T	460	13114	8	450	1.7%	96.6%	1.7%	1.5%	0.6%	3.1%	0.445	Not assoc w R	WHO		5) Not assoc w R
PZA	Rv3236c_a-520g (Rv3237c_a-33g)	397	13177	7	451	1.5%	97.1%	1.7%	0.0%	0.0%	0.9%	NA	Not assoc w R	WHO		5) Not assoc w R
PZA	Rv3236c_V151A	379	13195	5	453	1.1%	97.2%	1.3%	1.3%	0.4%	3.0%	0.515	Not assoc w R	WHO		5) Not assoc w R
PZA	clpC1_P796L	153	13421	6	2323	0.3%	98.9%	3.8%	0.6%	0.0%	3.6%	0.038	Not assoc w R	WHO		5) Not assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
PZA	PPE35_P670L	299	13275	1	457	0.2%	97.8%	0.3%	0.0%	0.0%	1.2%	NA	Not assoc w R	WHO		5) Not assoc w R
PZA	Rv1258c_262_del_1_cg_c	38	13536	0	458	0.0%	99.7%	0.0%	0.0%	0.0%	9.3%	0.000	Not assoc w R	WHO		5) Not assoc w R
PZA	Rv1258c_E243A	43	13531	0	458	0.0%	99.7%	0.0%	0.0%	0.0%	8.2%	NA	Not assoc w R	WHO		5) Not assoc w R
PZA	panD_c-1937t (Rv3603c_c-100t)	74	13500	0	458	0.0%	99.5%	0.0%	0.0%	0.0%	4.9%	NA	Not assoc w R	WHO		5) Not assoc w R
PZA	PPE35_T712P	97	13477	0	458	0.0%	99.3%	0.0%	0.0%	0.0%	3.7%	0.000	Not assoc w R	WHO		5) Not assoc w R
PZA	Rv1258c_P414S	108	13466	0	458	0.0%	99.2%	0.0%	0.0%	0.0%	3.4%	0.000	Not assoc w R	WHO		5) Not assoc w R

Premature stop codons or indels in the coding regions of *pncA* classified as "Assoc w RI" according to the associated expert rule (see "Additional criteria for final confidence grading") are not listed in this table but can be found in the Mutation catalogue. In-frame indels are shown in gold.

Levofloxacin and moxifloxacin

In this analysis, an expert rule was applied to ensure that any *gyrA* or *gyrB* mutation that met the criteria for LFX resistance was also considered to confer resistance to MFX and vice versa. Fourteen mutations were therefore classified as either definitive or interim resistance markers for both FQs. Because of this rule, which is consistent with published MIC data and, where available, direct enzymatic measurements, *gyrA* G88A, *gyrB* D461N and *gyrB* A504V were classified as interim markers for MFX resistance *(13, 39)*. The expert rule was also applied to upgrade *gyrB* E501D to a group-2 mutation for LFX resistance, as it met the criteria for group 1 in MFX-resistant isolates. This mutation is known to have a more marked effect on MFX resistance than on LFX resistance *(13, 40, 41)*.

The sensitivity and specificity for prediction of MFX- and LFX-resistant phenotypes calculated in this study are probably not representative of the actual performance of these 14 FQ resistance mutations in clinical laboratory settings. First, only mutations with an allele frequency ≥ 90% in mixed population samples were included in the analysed dataset (see "Heteroresistance"), which probably resulted in an underestimate of the sensitivity, given that heteroresistance plays an important role in FQ resistance *(42)*. Secondly, the WHO CCs for MFX and, to a lesser extent, LFX used between 2014 and 2018 were too high, resulting in poor concordance between mutations and phenotypic resistance for isolates tested with those CCs *(13)*. Additionally, the MFX CC for the CRyPTIC BMD plates (Table 1) used in this study is undergoing peer review and has not yet been reviewed by WHO *(43)*. Thirdly, some mutations confer MICs close to the epidemiological cut-off value, resulting in a lower categorical agreement with phenotypic DST *(13, 41)*.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
LFX	gyrA_D94G	51	15118	1190	1918	38.3%	99.7%	95.9%	95.7%	94.4%	96.8%	186.052	Assoc w R	ALL+WHO		1) Assoc w R
LFX	gyrA_A90V	72	15097	679	2429	21.8%	99.5%	90.4%	89.1%	86.5%	91.4%	53.927	Assoc w R	ALL+WHO		1) Assoc w R
LFX	gyrA_D94A	29	15140	228	2880	7.3%	99.8%	88.7%	86.2%	80.8%	90.6%	36.596	Assoc w R	ALL+WHO		1) Assoc w R
LFX	gyrA_D94N	11	15158	216	2892	6.9%	99.9%	95.2%	94.9%	91.0%	97.4%	97.203	Assoc w R	ALL+WHO		1) Assoc w R
LFX	gyrA_S91P	20	15149	127	2981	4.1%	99.9%	86.4%	84.8%	77.6%	90.5%	28.458	Assoc w R	ALL+WHO		1) Assoc w R
LFX	gyrA_D94Y	10	15159	106	3002	3.4%	99.9%	91.4%	91.1%	84.2%	95.6%	51.506	Assoc w R	ALL+WHO		1) Assoc w R
LFX	gyrA_D94H	3	15166	48	3060	1.5%	100.0%	94.1%	94.0%	83.5%	98.7%	116.471	Assoc w R	ALL+WHO		1) Assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
LFX	gyrA_G88C	0	15169	28	3080	0.9%	100.0%	100.0%	100.0%	87.2%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
LFX	gyrB_E501D	24	15145	28	3080	0.9%	99.8%	53.8%	38.5%	23.4%	55.4%	3.353	Uncert. Sig.	ALL+WHO	FQ X-R	2) Assoc w RI
LFX	gyrB_D461N	17	15152	24	3084	0.8%	99.9%	58.5%	45.2%	27.3%	64.0%	4.299	Assoc w R	ALL	ALL only	2) Assoc w RI
LFX	gyrA_G88A	8	15161	8	3100	0.3%	99.9%	50.0%	20.0%	2.5%	55.6%	1.223	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
LFX	gyrB_A504V	11	15158	8	3100	0.3%	99.9%	42.1%	0.0%	0.0%	28.5%	0.000	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
LFX	gyrB_N499D	0	15169	6	3102	0.2%	100.0%	100.0%	100.0%	39.8%	100.0%	Inf	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
LFX	gyrB_E501V	1	15168	0	3108	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
LFX	gyrA_A90G	3	15166	0	3108	0.0%	100.0%	0.0%	0.0%	0.0%	70.8%	NA	Uncert. Sig.	ALL+WHO	Lit. (PMID 28137812)	4) Not assoc w RI
LFX	gyrA_T80A	16	15153	0	3108	0.0%	99.9%	0.0%	0.0%	0.0%	20.6%	0.000	Uncert. Sig.	ALL+WHO	Prev. WHO	4) Not assoc w RI
LFX	gyrA_E21Q	8282	4	2019	0	100.0%	0.0%	19.6%	0.1%	0.1%	0.2%	Inf	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrA_G668D	7701	585	1932	87	95.7%	7.1%	20.1%	0.0%	0.0%	0.0%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrA_S95T	7698	588	1929	90	95.5%	7.1%	20.0%	0.0%	0.0%	0.0%	0.000	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrB_c-165t	367	7919	7	129	5.1%	95.6%	1.9%	1.9%	0.8%	3.8%	1.171	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrA_G247S	447	7839	75	1944	3.7%	94.6%	14.4%	0.0%	0.0%	0.8%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrA_A384V	1127	7159	56	1963	2.8%	86.4%	4.7%	0.0%	0.0%	0.3%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrB_M291I	1115	7171	55	1964	2.7%	86.5%	4.7%	0.0%	0.0%	0.3%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrA_A463S	98	8188	28	1991	1.4%	98.8%	22.2%	0.0%	0.0%	3.7%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrA_c-34t	325	7961	22	1997	1.1%	96.1%	6.3%	0.0%	0.0%	1.1%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrA_Q613E	119	8167	5	2014	0.2%	98.6%	4.0%	0.0%	0.0%	3.1%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrB_P94L	223	8063	3	2016	0.1%	97.3%	1.3%	0.0%	0.0%	1.6%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrA_R252L	41	8245	2	2017	0.1%	99.5%	4.7%	0.0%	0.0%	8.6%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrB_A403S	40	8246	0	2019	0.0%	99.5%	0.0%	0.0%	0.0%	8.8%	NA	Not assoc w R	WHO		5) Not assoc w R
LFX	gyrB_V301L	83	8203	0	2019	0.0%	99.0%	0.0%	0.0%	0.0%	4.3%	NA	Not assoc w R	WHO		5) Not assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
MXF	gyrA_D94G	327	11155	815	1054	43.6%	97.2%	71.4%	70.1%	67.3%	72.8%	25.210	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrA_A90V	353	11129	323	1546	17.3%	96.9%	47.8%	41.7%	37.8%	45.8%	5.279	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrA_D94N	41	11441	162	1707	8.7%	99.6%	79.8%	78.2%	71.6%	83.9%	24.631	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrA_D94A	105	11377	125	1744	6.7%	99.1%	54.3%	45.9%	38.7%	53.2%	5.924	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrA_S91P	56	11426	90	1779	4.8%	99.5%	61.6%	58.8%	50.1%	67.2%	9.342	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrA_D94Y	34	11448	77	1792	4.1%	99.7%	69.4%	68.2%	58.5%	76.9%	13.716	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrB_E501D	8	11474	33	1836	1.8%	99.9%	80.5%	73.3%	54.1%	87.7%	17.186	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrA_G88C	2	11480	29	1840	1.6%	100.0%	93.5%	93.3%	77.9%	99.2%	87.348	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrA_D94H	15	11467	26	1843	1.4%	99.9%	63.4%	62.5%	45.8%	77.3%	10.370	Assoc w R	ALL+WHO		1) Assoc w R
MXF	gyrB_D461N	21	11461	9	1860	0.5%	99.8%	30.0%	12.5%	2.7%	32.4%	0.924	Uncert. Sig.	ALL+WHO	FQ X-R	2) Assoc w RI
MXF	gyrB_A504V	4	11478	8	1861	0.4%	100.0%	66.7%	0.0%	0.0%	60.2%	0.000	Uncert. Sig.	ALL+WHO	FQ X-R	2) Assoc w RI
MXF	gyrA_G88A	9	11473	4	1865	0.2%	99.9%	30.8%	10.0%	0.3%	44.5%	0.769	Uncert. Sig.	ALL+WHO	FQ X-R	2) Assoc w RI
MXF	gyrB_N499D	1	11481	3	1866	0.2%	100.0%	75.0%	66.7%	9.4%	99.2%	12.305	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
MXF	gyrB_E501V	1	11481	0	1869	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	0.000	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
MXF	gyrA_A90G	3	11479	0	1869	0.0%	100.0%	0.0%	0.0%	0.0%	70.8%	NA	Uncert. Sig.	ALL+WHO	Lit. (PMID 28137812)	4) Not assoc w RI
MXF	gyrA_T80A	9	11473	0	1869	0.0%	99.9%	0.0%	0.0%	0.0%	33.6%	0.000	Uncert. Sig.	ALL+WHO	Lit. (PMID 28137812)	4) Not assoc w RI
MXF	gyrA_E21Q	5807	3	1094	0	100.0%	0.1%	15.9%	0.1%	0.0%	0.2%	Inf	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrA_S95T	5439	371	1045	49	95.5%	6.4%	16.1%	0.0%	0.0%	0.1%	NA	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrA_G668D	5445	365	1045	49	95.5%	6.3%	16.1%	0.0%	0.0%	0.1%	NA	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrB_c-165t	331	5479	2	52	3.7%	94.3%	0.6%	0.6%	0.1%	2.2%	0.637	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrA_G247S	247	5563	28	1066	2.6%	95.7%	10.2%	0.0%	0.0%	1.5%	NA	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrB_M291I	1081	4729	23	1071	2.1%	81.4%	2.1%	0.0%	0.0%	0.3%	NA	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrA_A384V	1093	4717	23	1071	2.1%	81.2%	2.1%	0.0%	0.0%	0.3%	NA	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrA_A463S	69	5741	17	1077	1.6%	98.8%	19.8%	0.0%	0.0%	5.2%	NA	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrA_c-34t	298	5512	11	1083	1.0%	94.9%	3.6%	0.0%	0.0%	1.2%	NA	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrB_P94L	72	5738	2	1092	0.2%	98.8%	2.7%	0.0%	0.0%	5.0%	NA	Not assoc w R	WHO		5) Not assoc w R
MXF	gyrB_V301L	41	5769	0	1094	0.0%	99.3%	0.0%	0.0%	0.0%	8.6%	NA	Not assoc w R	WHO		5) Not assoc w R

Bedaquiline and clofazimine

In the data available for this analysis, no mutations met the criteria for association with BDQ- or CFZ-resistant phenotypes. This does not contradict previous studies that found *atpE* and *Rv0678* to be the key resistance genes for one or both agents *(7, 13, 44, 45)*. Instead, the results of this analysis were probably due to the following limitations. First, most mutations in *Rv0678* and *atpE* were rare in this dataset, even though some *Rv0678* variants are frequent in certain settings *(46)*. Secondly, *Rv0678* mutations are more likely to be heteroresistant with < 90% resistant alleles and may therefore have been excluded from this analysis *(44)*. Thirdly, some *Rv0678* mutations result in MICs that are close to the CC, resulting in inconsistent categorical phenotypic DST results *(13, 46, 47)*. Fourthly, most data on BDQ phenotypic DST were from the CRyPTIC BMD plates, which are based on CCs that are undergoing peer review, have not yet been reviewed by WHO and might be revised up or down. Finally, epistasis may have confounded the classification of some *Rv0678* mutations, as LoF mutations in this gene can confer resistance only if the efflux pump encoded by *mmpS5-mmpL5* is active *(7, 8)*[1].

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
BDQ	mmpL5_D767N	49	36	3	0	100.0%	42.4%	5.8%	0.0%	0.0%	7.3%	NA	Not assoc w R	WHO		5) Not assoc w R
BDQ	mmpL5_T794I	58	27	3	0	100.0%	31.8%	4.9%	0.0%	0.0%	6.2%	NA	Not assoc w R	WHO		5) Not assoc w R
BDQ	mmpL5_I948V	85	0	3	0	100.0%	0.0%	3.4%	0.0%	0.0%	4.2%	NA	Not assoc w R	WHO		5) Not assoc w R
BDQ	Rv1979c_a-129g	85	0	2	0	100.0%	0.0%	2.3%	2.3%	0.3%	8.1%	NA	Not assoc w R	WHO		5) Not assoc w R

[1] Vargas R, Freschi L, Spitaleri A, Tahseen S, Barilar I, Niemann S et al. The role of epistasis in amikacin, kanamycin, bedaquiline, and clofazimine resistance in *Mycobacterium tuberculosis* complex. bioRxiv 2021.05.07.443178.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
CFZ	mmpL5_I948V	3601	11	23	0	100.0%	0.3%	0.6%	0.1%	0.0%	0.2%	Inf	Not assoc w R	WHO		5) Not assoc w R
CFZ	Rv1979c_a-129g	3601	11	22	0	100.0%	0.3%	0.6%	0.6%	0.4%	0.9%	Inf	Not assoc w R	WHO		5) Not assoc w R
CFZ	mmpL5_T794I	3063	549	20	3	87.0%	15.2%	0.6%	0.0%	0.0%	0.1%	NA	Not assoc w R	WHO		5) Not assoc w R
CFZ	mmpL5_D767N	1102	2510	16	7	69.6%	69.5%	1.4%	0.0%	0.0%	0.3%	NA	Not assoc w R	WHO		5) Not assoc w R
CFZ	Rv1979c_D286G	611	3001	1	21	4.5%	83.1%	0.2%	0.0%	0.0%	0.6%	NA	Not assoc w R	WHO		5) Not assoc w R
CFZ	mmpL5_F696L	210	3402	1	22	4.3%	94.2%	0.5%	0.0%	0.0%	1.7%	NA	Not assoc w R	WHO		5) Not assoc w R
CFZ	Rv1979c_c-389a (mpt64_159)	226	3386	0	22	0.0%	93.7%	0.0%	0.0%	0.0%	1.6%	NA	Not assoc w R	WHO		5) Not assoc w R

Linezolid

Only *rplC* C154R was found to be a marker for resistance (group 1), resulting in a sensitivity of 38.2% (95% CI, 29.6–47.4%). This may be an underestimate, as the PPV of phenotypic DST is unlikely to be high, with a prevalence of resistance of only 1.1% in this dataset (95% CI, 0.9–1.3%). This finding is consistent with earlier findings that this is the dominant LZD resistance mutation *in vitro* and in clinical isolates *(7)*. In fact, the PPV of this mutation (73%; 95% CI, 61–84%) was comparable to the PPVs of *rrl* g2270t and g2814t, two other well-documented LZD resistance mutations in MTBC and other bacteria (i.e. 70%; 95% CI, 35–94%; and 75%; 95% CI, 48–93%, respectively *(7)*). As resistance mutations in *rrl* were rarer and more diverse, however, the PPV|SOLOs for these *rrl* mutations did not meet the criteria used in this analysis.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
LZD	rplC_C154R	17	10878	47	76	38.2%	99.8%	73.4%	71.2%	57.9%	82.2%	375.720	Assoc w R	ALL+WHO		1) Assoc w R
LZD	rrs_c-187t	1122	0	1	0	100.0%	0.0%	0.1%	0.1%	0.0%	0.5%	NA	Not assoc w R	WHO		5) Not assoc w R
LZD	rrl_c344t	57	1065	0	9	0.0%	94.9%	0.0%	0.0%	0.0%	6.3%	0.000	Not assoc w R	WHO		5) Not assoc w R

Delamanid

With one exception, mutations associated with phenotypic resistance to DLM were found to be of uncertain significance (group 3). *ddn* L49P, however, was found to be a group-2 mutation, as it was detected in the ALL dataset only. This mutation is known to be transmitted between patients and was selected for pretomanid monotherapy in mice *(48)*. As for BDQ and CFZ, these findings do not contradict the strong experimental evidence that other *ddn* mutations or mutations in *fbiA*, *fbiB*, *fbiC*, *fgd1* and *Rv2983* confer DLM resistance *(13, 49)*. In fact, DLM resistance mutations typically result in large increases in MIC *(7, 13)*; however, because the aforementioned genes are non-essential and span approximately 7500 base pairs (bp), including promoters, a large spectrum of rare resistance mutations is possible *(2, 50)*. This also means that the sensitivity of 6.1% (95% CI, 2.1–13.7%) for *ddn* L49P is probably not representative.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
DLM	ddn_L49P	1	7695	5	77	6.1%	100.0%	83.3%	83.3%	35.9%	99.6%	499.675	Assoc w R	ALL	ALL only	2) Assoc w RI

Amikacin

Previously, only *rrs* a1401t and *rrs* g1484t were recognized as resistance mutations for AMK *(13)*; however, the 2018 WHO systematic review of the CCs for AMK revealed that *rrs* c1402t and *eis* c-14t also modestly increase the MIC close to the newly endorsed CC of 2 mg/L for 7H10, although they were not formally recognized as markers for AMK resistance at that time *(13)*. In this analysis, *eis* c-14t was classified as a definitive marker for resistance (group 1), which is consistent with the interpretation of this mutation in the Cepheid Xpert MTB/XDR assay *(51)*[2]. It should be noted, however, that epistasis complicates interpretation of *eis* c-14t for genotypic prediction of AMK resistance, given that it can only result in overexpression of *eis* and, therefore, resistance if *eis* is functional. Thus, interpreting this promoter mutation without considering LoF mutations could result in overestimating resistance to AMK.[3] On the basis of the phenotypic DST results in this dataset, *rrs* c1402t was classified as uncertain significance (group 3) in initial confidence grading. In light of the aforementioned systematic review, however, the fact that the Xpert MTB/XDR assay already interprets this mutation as a resistance mutation for AMK and to err on the side of caution, it was decided to recognize *rrs* c1402t as a group-2 mutation (Assoc w RI) in the final confidence grading *(51)*.[2] The official instructions for use of the Hain GenoType MTBDR*sl* V2.0 will probably have to be updated accordingly for consistency *(17, 52)*.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
AMK	rrs_a1401g	50	15640	939	349	72.9%	99.7%	94.9%	94.8%	93.2%	96.1%	857.063	Assoc w R	ALL+WHO		1) Assoc w R
AMK	eis_c-14t	51	7325	32	632	4.8%	99.3%	38.6%	35.4%	25.0%	47.0%	6.623	Assoc w R	WHO		1) Assoc w R
AMK	rrs_g1484t	2	15688	6	1282	0.5%	100.0%	75.0%	71.4%	29.0%	96.3%	30.593	Assoc w R	ALL	ALL only	2) Assoc w RI
AMK	rrs_c1402t	10	15680	5	1283	0.4%	99.9%	33.3%	28.6%	8.4%	58.1%	5.432	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
AMK	rrs_c-187t	7328	48	85	1	98.8%	0.7%	1.1%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
AMK	aftB_D397G	2730	4646	52	34	60.5%	63.0%	1.9%	0.0%	0.0%	0.1%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	ccsA_I245M	2692	4684	51	35	59.3%	63.5%	1.9%	0.0%	0.0%	0.1%	NA	Not assoc w R	WHO		5) Not assoc w R

[2] S. Chakravorty. Personal communication. 2021.

[3] Vargas R, Freschi L, Spitaleri A, Tahseen S, Barilar I, Niemann S et al. The role of epistasis in amikacin, kanamycin, bedaquiline, and clofazimine resistance in *Mycobacterium tuberculosis* complex. bioRxiv 2021.05.07.443178.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
AMK	whiB6_-66_del_1_agctccgagctctagt_agctccgagtctagt	3226	4150	38	48	44.2%	56.3%	1.2%	0.0%	0.0%	0.2%	4.803	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB6_-74_del_1_gc_g	3436	3940	36	50	41.9%	53.4%	1.0%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
AMK	rrs_a514c	180	7196	65	599	9.8%	97.6%	26.5%	0.0%	0.0%	2.0%	0.000	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB6_-73_del_1_agc_ag	232	7144	7	79	8.1%	96.9%	2.9%	0.0%	0.0%	1.6%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB6_T51P	343	7033	7	79	8.1%	95.3%	2.0%	0.0%	0.0%	1.1%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	fprA_N385D	747	6629	5	81	5.8%	89.9%	0.7%	0.0%	0.0%	0.5%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB6_-74_del_1_gct_gt	110	7266	4	82	4.7%	98.5%	3.5%	0.0%	0.0%	3.3%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB6_c-82t	568	6808	4	82	4.7%	92.3%	0.7%	0.0%	0.0%	0.6%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	eis_V163I	211	7165	26	638	3.9%	97.1%	11.0%	1.4%	0.3%	4.0%	0.164	Not assoc w R	WHO		5) Not assoc w R
AMK	rrs_c517t	163	7213	21	643	3.2%	97.8%	11.4%	0.0%	0.0%	2.2%	0.000	Not assoc w R	WHO		5) Not assoc w R
AMK	eis_c-12t	288	7088	20	644	3.0%	96.1%	6.5%	2.7%	1.2%	5.3%	0.337	Not assoc w R	WHO		5) Not assoc w R
AMK	ccsA_V27I	138	7238	2	84	2.3%	98.1%	1.4%	0.0%	0.0%	2.6%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB6_R107C	270	7106	2	84	2.3%	96.3%	0.7%	0.0%	0.0%	1.4%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	rrs_c492t	212	7164	4	660	0.6%	97.1%	1.9%	1.9%	0.5%	4.7%	0.209	Not assoc w R	WHO		5) Not assoc w R
AMK	eis_807_del_2_ggt_g	42	7334	0	664	0.0%	99.4%	0.0%	0.0%	0.0%	8.4%	0.000	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB7_192_del_1_gc_g	44	7332	0	664	0.0%	99.4%	0.0%	0.0%	0.0%	8.0%	0.000	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB6_R54Q	53	7323	0	86	0.0%	99.3%	0.0%	0.0%	0.0%	6.7%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB7_c-178t	94	7282	0	86	0.0%	98.7%	0.0%	0.0%	0.0%	3.8%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB7_t-100c	104	7272	0	664	0.0%	98.6%	0.0%	0.0%	0.0%	3.5%	0.000	Not assoc w R	WHO		5) Not assoc w R
AMK	whiB6_-74_del_1_gctctagtg_gtctagta	126	7250	0	86	0.0%	98.3%	0.0%	0.0%	0.0%	2.9%	NA	Not assoc w R	WHO		5) Not assoc w R
AMK	aftB_V293M	151	7225	0	86	0.0%	98.0%	0.0%	0.0%	0.0%	2.4%	NA	Not assoc w R	WHO		5) Not assoc w R

Streptomycin

The 12 group-1 mutations identified had a sensitivity of 75.2% (95% CI, 73.9–76.4%) and a specificity of 98.0% (95% CI, 97.6–98.2%). Most of those mutations (i.e. the six *gid* mutations, *rpsL* K43R, K88M and K88R, *rrs* a514c and c517t) are functionally well understood and documented as associated with resistance *(6, 8, 26, 53–56)*. In contrast, Malinga et al. *(57)* implicated *rrs* g878a, the 12th mutation, in CAP resistance, although they did not test STM. The evidence from other studies is also not conclusive *(54, 58–60).* In this analysis, however, the PPV|SOLO values for AMK, CAP and KAN were much lower (i.e. 0.0% [95% CI 0.0–8.0%], 0.0% [95% CI 0.0–21.8%] and 0.0% [95% CI 0.0–21.8%], respectively) than 85.7% (95% CI, 63.7–97.0%) for STR. In fact, this mutation, which corresponds to position 885 in the *E. coli* numbering, was described in one STR-resistant chloroplast mutant of *Nicotiana plumbaginifolia*, a species of tobacco plant *(61)*. The wild-type nucleotide at this position usually forms a Watson-Crick base pair with position 905 (912 in *E. coli*) of 16s rRNA, the disruption of which provides a plausible mechanism for STR resistance *(62, 63)*.

The 166 group-2 mutations, of which 162 were classified on the assumption that any nonsense or indel in *gid* should cause STM resistance, increased the sensitivity to 82.4% (95% CI, 81.3–83.5%) but lowered the specificity to 95.4% (95% CI, 95.0–95.8%). The relatively low PPV for group-2 mutations of 60.5% (95% CI, 56.4–64.4%) was probably due to the fact that even LoF mutations in *gid* only confer small increases in MIC, so that the current CCs divide the resulting MIC distribution at its lower end *(25, 26, 54, 64, 65)*. As with EMB, it is also not clear whether the currently used CCs correspond well to the epidemiological cut-off values, and this may exacerbate the very major phenotypic DST error rate for this resistance mechanism. Three mutations (*rrs* c492t and gid L16R and E92D) that are known to be frequent because they are deeply rooted in the MTBC phylogeny were classified into group 5 as definitively neutral, thereby confirming previous findings *(6, 8, 66)*.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
STM	rpsL_K43R	55	9294	2245	2390	48.4%	99.4%	97.6%	96.9%	96.0%	97.7%	132.521	Assoc w R	ALL+WHO		1) Assoc w R
STM	rpsL_K88R	21	9328	501	4134	10.8%	99.8%	96.0%	95.2%	92.7%	97.0%	52.023	Assoc w R	ALL+WHO		1) Assoc w R
STM	rrs_c517t	49	9300	329	4306	7.1%	99.5%	87.0%	85.2%	80.9%	88.8%	12.689	Assoc w R	ALL+WHO		1) Assoc w R
STM	rrs_a514c	12	9337	258	4377	5.6%	99.9%	95.6%	93.1%	88.2%	96.4%	42.931	Assoc w R	ALL+WHO		1) Assoc w R
STM	gid_103_del_1_gc_g	31	9318	55	4580	1.2%	99.7%	64.0%	58.1%	46.1%	69.5%	2.916	Assoc w R	ALL+WHO		1) Assoc w R
STM	rrs_g878a	3	9346	29	4606	0.6%	100.0%	90.6%	85.7%	63.7%	97.0%	12.175	Assoc w R	ALL+WHO		1) Assoc w R
STM	gid_352_del_1_gc_g	14	9335	28	4607	0.6%	99.9%	66.7%	63.2%	46.0%	78.2%	3.741	Assoc w R	ALL+WHO		1) Assoc w R
STM	gid_Q125!	1	6480	13	2549	0.5%	100.0%	92.9%	90.9%	58.7%	99.8%	25.422	Assoc w R	WHO		1) Assoc w R
STM	rpsL_K88M	0	9349	11	4624	0.2%	100.0%	100.0%	100.0%	66.4%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
STM	gid_G69D	2	9347	11	4624	0.2%	100.0%	84.6%	83.3%	51.6%	97.9%	10.107	Assoc w R	ALL+WHO		1) Assoc w R
STM	gid_A134E	0	9349	10	4625	0.2%	100.0%	100.0%	100.0%	69.2%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
STM	gid_P75R	0	9349	9	4626	0.2%	100.0%	100.0%	100.0%	63.1%	100.0%	Inf	Assoc w R	ALL+WHO		1) Assoc w R
STM	gid_P84L	13	9336	49	4586	1.1%	99.9%	79.0%	76.8%	63.6%	87.0%	8.754	Assoc w R	ALL	ALL only	2) Assoc w RI
STM	gid_116_del_1_cg_c	22	9327	39	4596	0.8%	99.8%	63.9%	56.9%	42.2%	70.7%	2.675	Assoc w R	ALL	ALL only	2) Assoc w RI
STM	gid_354_del_1_gcgcccccgcacgatctcaacggcca_gcgccccgcacgatctcaacggcca	6	9343	24	4611	0.5%	99.9%	80.0%	75.0%	53.3%	90.2%	6.079	Assoc w R	ALL	ALL only	2) Assoc w RI
STM	gid_G73A	8	9341	22	4613	0.5%	99.9%	73.3%	68.0%	46.5%	85.1%	6.885	Assoc w R	ALL	ALL only	2) Assoc w RI
STM	rpsL_K88Q	4	9345	7	4628	0.2%	100.0%	63.6%	55.6%	21.2%	86.3%	2.524	Uncert. Sig.	ALL+WHO	Prev. WHO	2) Assoc w RI
STM	gid_V110G														Prev. WHO	4) Not assoc w RI
STM	rpsL_t-165c	6411	70	85	0	100.0%	1.1%	1.3%	0.0%	0.0%	0.1%	NA	Not assoc w R	WHO		5) Not assoc w R
STM	rrs_c-187t	6474	7	84	1	98.8%	0.1%	1.3%	0.0%	0.0%	0.1%	0.000	Not assoc w R	WHO		5) Not assoc w R
STM	whiB6_-74_del_1_gc_g	4923	1558	51	34	60.0%	24.0%	1.0%	0.0%	0.0%	0.1%	NA	Not assoc w R	WHO		5) Not assoc w R
STM	gid_E92D	1041	5440	1390	1172	54.3%	83.9%	57.2%	0.0%	0.0%	0.4%	0.000	Not assoc w R	WHO		5) Not assoc w R
STM	Rv1258c_581_ins_1_t_tg	1000	5481	1353	1209	52.8%	84.6%	57.5%	0.0%	0.0%	0.4%	NA	Not assoc w R	WHO		5) Not assoc w R
STM	gid_L16R	949	5532	222	2340	8.7%	85.4%	19.0%	0.4%	0.1%	1.1%	0.017	Not assoc w R	WHO		5) Not assoc w R
STM	whiB6_-74_del_1_gct_gt	215	6266	3	82	3.5%	96.7%	1.4%	0.0%	0.0%	1.7%	NA	Not assoc w R	WHO		5) Not assoc w R

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
STM	whiB6_-74_del_1_gctctagtg_gtctagta	476	6005	3	82	3.5%	92.7%	0.6%	0.0%	0.0%	0.8%	NA	Not assoc w R	WHO		5) Not assoc w R
STM	whiB6_R54Q	307	6174	2	83	2.4%	95.3%	0.6%	0.0%	0.0%	1.2%	NA	Not assoc w R	WHO		5) Not assoc w R
STM	whiB7_t-100c	73	6408	42	2520	1.6%	98.9%	36.5%	1.4%	0.0%	7.3%	0.035	Not assoc w R	WHO		5) Not assoc w R
STM	gid_Y195H	139	6342	34	2528	1.3%	97.9%	19.7%	0.0%	0.0%	2.6%	NA	Not assoc w R	WHO		5) Not assoc w R
STM	whiB7_192_del_1_gc_g	285	6196	20	2542	0.8%	95.6%	6.6%	0.7%	0.1%	2.5%	0.018	Not assoc w R	WHO		5) Not assoc w R
STM	rrs_c492t	359	6122	19	2543	0.7%	94.5%	5.0%	0.0%	0.0%	1.0%	NA	Not assoc w R	WHO		5) Not assoc w R
STM	rpsL_c-259t (Rv0681_194)	50	6431	0	85	0.0%	99.2%	0.0%	0.0%	0.0%	7.1%	NA	Not assoc w R	WHO		5) Not assoc w R
STM	whiB6_-68_del_1_ctccgagc_cgccgag	63	6418	0	85	0.0%	99.0%	0.0%	0.0%	0.0%	5.7%	NA	Not assoc w R	WHO		5) Not assoc w R

Premature stop codons or indels in the coding regions of *gid* classified as "Assoc w RI" according to the associated expert rule (see "Additional criteria for final confidence grading") are not listed in this table but can be found in the Mutation catalogue.

Ethionamide

Numerous resistance mechanisms to ETO exist, some of which include a large spectrum of potential resistance mutations, as they are non-essential *(67, 68)*. Four group-1 and 327 group-2 mutations yielded a combined sensitivity of 75.7% (95% CI, 74.1–77.3%) but an associated specificity of only 91.4% (95% CI, 90.8–91.9 %) and PPV of 70.4% (95% CI, 68.8–72.0%). Of those mutations, 304 were classified in group 2 according to the expert rule that any premature stop codon and indel in *ethA* should be assumed to confer ETO resistance. Notably, the aforementioned g-154a mutation upstream of *inhA* (i.e. g609a in codon 203 of *fabG1*) was classified as a group-2 mutation, which is consistent with published data on allelic exchange *(21)*. In contrast, this mutation is interpreted as a marker for INH resistance only by the Xpert MTB/XDR assay, which will probably have to be updated accordingly for consistency *(69)*.

The low PPV is probably due mainly to the fact that many ETO resistance mutations confer modest increases in MIC, resulting in a considerable overlap with the MIC distribution of susceptible isolates, if these mutations occur alone *(25, 26, 70)*. Indeed, the development of ETO resistance may be similar to that to EMB, whereby resistance evolves in a stepwise manner, given that it is not uncommon for isolates to harbour multiple mechanisms with presumably additive effects *(24)*. For example, the most frequent group-1 mutation (c-15t upstream of the *fabG1-inhA* operon) can coincide with *ethA* mutations or the group 2 *inhA* S94A, which is known to confer ETO and INH cross-resistance in allelic exchange results *(25, 56, 71–73)*.

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
ETO	inhA_c-777t (fabG1_c-15t)	409	10544	1297	1668	43.7%	96.3%	76.0%	64.4%	61.5%	67.1%	15.266	Assoc w R	ALL+WHO		1) Assoc w R
ETO	ethA_111_del_1_ct_c	23	10930	44	2921	1.5%	99.8%	65.7%	63.5%	50.4%	75.3%	7.127	Assoc w R	ALL+WHO		1) Assoc w R
ETO	ethA_R207G	12	10941	30	2935	1.0%	99.9%	71.4%	71.4%	55.4%	84.3%	9.319	Assoc w R	ALL+WHO		1) Assoc w R
ETO	ethA_M1R	17	10936	27	2938	0.9%	99.8%	61.4%	61.4%	45.5%	75.6%	5.912	Assoc w R	ALL+WHO		1) Assoc w R
ETO	inhA_g-154a (fabG1_L203L)	47	10906	175	2790	5.9%	99.6%	78.8%	56.5%	46.6%	66.0%	5.961	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	inhA_t-770c (fabG1_t-8c)	36	10917	66	2899	2.2%	99.7%	64.7%	20.0%	9.6%	34.6%	2.607	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI
ETO	inhA_S94A	10	10943	62	2903	2.1%	99.9%	86.1%	56.5%	34.5%	76.8%	4.900	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	inhA_t-770a (fabG1_t-8a)	18	10935	42	2923	1.4%	99.8%	70.0%	21.7%	7.5%	43.7%	3.118	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI

Drug	Variant (common name)	MUT present_pheno S	MUT absent_pheno S	MUT present_pheno R	MUT absent_pheno R	SENSITIVITY	SPECIFICITY	PPV	PPV | SOLO	PPV | SOLO_lb	PPV | SOLO_ub	OR SOLO	INITIAL CONFIDENCE GRADING	SUPPORTING DATASET	ADDITIONAL GRADING CRITERIA	FINAL CONFIDENCE GRADING
ETO	ethA_t-7c	39	10914	35	2930	1.2%	99.6%	47.3%	45.1%	33.2%	57.3%	3.406	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_1243_del_1_ca_c	4	10949	27	2938	0.9%	100.0%	87.1%	85.2%	66.3%	95.8%	21.428	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_P378L	31	10922	25	2940	0.8%	99.7%	44.6%	42.6%	29.2%	56.8%	2.946	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_M1T	6	10947	24	2941	0.8%	99.9%	80.0%	78.6%	59.0%	91.7%	13.648	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_S390F	4	10949	19	2946	0.6%	100.0%	82.6%	82.6%	61.2%	95.0%	17.654	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_753_ins_1_g_gc	6	10947	19	2946	0.6%	99.9%	76.0%	64.7%	38.3%	85.8%	8.175	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_342_del_1_gt_g	11	10942	19	2946	0.6%	99.9%	63.3%	56.0%	34.9%	75.6%	5.778	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_A341V	2	10951	18	2947	0.6%	100.0%	90.0%	89.5%	66.9%	98.7%	31.586	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_Y147!	6	10947	12	2953	0.4%	99.9%	66.7%	66.7%	41.0%	86.7%	7.414	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_S57Y	1	10952	11	2954	0.4%	100.0%	91.7%	90.9%	58.7%	99.8%	37.075	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_704_del_1_ta_t	5	10948	11	2954	0.4%	100.0%	68.8%	68.8%	41.3%	89.0%	8.154	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_Y32D	6	10947	10	2955	0.3%	99.9%	62.5%	62.5%	35.4%	84.8%	6.174	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_367_ins_1_ggctttggatgtgaac_ggcttttggatgtgaac	0	10953	9	2956	0.3%	100.0%	100.0%	100.0%	63.1%	100.0%	Inf	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_T88I	1	10952	8	2957	0.3%	100.0%	88.9%	88.9%	51.8%	99.7%	29.630	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_Q359!	2	10951	8	2957	0.3%	100.0%	80.0%	80.0%	44.4%	97.5%	14.814	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_G11V	0	10953	7	2958	0.2%	100.0%	100.0%	100.0%	54.1%	100.0%	Inf	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_441_ins_1_g_ga	0	10953	6	2959	0.2%	100.0%	100.0%	100.0%	47.8%	100.0%	Inf	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	ethA_1392_ins_2_gt_gtct	1	10952	6	2959	0.2%	100.0%	85.7%	85.7%	42.1%	99.6%	22.208	Assoc w R	ALL	ALL only	2) Assoc w RI
ETO	inhA_a-778g (fabG1_a-16g)	1	10952	0	2965	0.0%	100.0%	0.0%	0.0%	0.0%	97.5%	NA	Uncert. Sig.	ALL+WHO	WHO-end. gDST	2) Assoc w RI

Premature stop codons or indels in the coding regions of *ethA* classified as "Assoc w RI" according to the associated expert rule (see "Additional criteria for final confidence grading") are not listed in this table but can be found in the Mutation catalogue.

Future research priorities

This Mutation catalogue will be updated and revised regularly. The following research areas were identified as priorities for future revisions.

1. Types of data to be analysed:
    a. Data on MIC and allelic exchange and direct measurements of enzymatic activity:
        - Current WHO treatment guidelines indicate that the level of resistance to INH and MFX has notable treatment implications (e.g. high-dose INH may be useful for low-level resistant isolates, and isolates with high-level resistance mutations to MFX cannot be treated with high-dose MFX, even as part of a longer, individualized regimen *(5)*). In this analysis, 131 group-1 and -2 mutations for these drugs were flagged as low and high-level resistance based on precedents in WHO guidelines *(17)*. Additional MIC data will be required for the five mutations that could not be stratified (e.g. *katG* W328L and *gyrA* G88A).
        - The validity of the breakpoints used should be assessed, particularly if the dosing of antibiotics changes (e.g. if a higher dose of RIF is endorsed *(4)*).
        - Some mutations result in only modest increases in MIC that are difficult to classify using categorical phenotypic DST data, as may be the case with the potential lineage effect linked to *clpC1* V63A *(31)*. Ideally, the shape and, in particular, the mode of individual mutations should be analysed to identify potential borderline resistance mechanisms *(20)*. Moreover, "areas of technical uncertainty" as defined by the European Committee on Antimicrobial Susceptibility Testing, may be necessary to minimize very major phenotypic DST errors *(4, 46, 47)*.
    b. A more strategic approach to collecting data, commissioning additional testing and interpreting the findings is necessary to maximize the utility of genotypic DST:
        - Group-3 mutations that may yield the largest potential gain in sensitivity should be a priority, particularly if they are homoplasic and, consequently, probably involved in resistance (e.g. specific *rrl* mutations for LZD *(7, 28, 74)*).
        - Exceptions to the expert rules that may result in significant harm in some settings should be studied. For example, *rpoB* T427A may not confer RIF resistance despite being in the RRDR *(20)*.
2. Grading criteria:
    a. Given the large spectrum of rare resistance mutations in non-essential genes for several key drugs (e.g. BDQ and DLM), "relaxed" grading criteria and/or new expert rules for LoF mutations might have to be adopted in order to grade mutations associated with resistance to these drugs, similar to the strategy used for PZA and *pncA (7)*.

b. Alternative approaches will be required to classify compensatory mechanisms, such as in *ahpC*, as, by definition, they have a low PPV|SOLO.
c. The selection of genes and corresponding regulatory regions will have to be revised in relation to the latest scientific evidence. For example, the role of *Rv1979c* in CFZ and BDQ resistance has been questioned *(7, 8)*.
d. The *a priori* assumption that synonymous mutations are not associated with resistance, unless they abolish a start codon or occur in either *fabG1* or *Rv3793*, should be examined and, depending on the findings, potentially handled differently in future revisions *(21, 24)*.[4] The same applies to the assumption that different nucleotide changes that result in the same amino acid substitution have the same effect *(28)*.

3. Bioinformatics pipeline:
   a. The potential contribution of isolates with heteroresistant alleles that make up < 90% of all variant calls at a locus should be evaluated.
   b. Large indels are not identified in the current algorithm. This results in some mutations falsely appearing as SOLO mutations, which is a particular problem for *ahpC* promoter mutations (e.g. the c-57t mutation). Similarly, the abrogation of stop codons was not considered in the current analysis.
   c. The exact effects of null calls and filter fails should be explored (see "Association studies").
4. Guidance should be developed on what, if any, confirmatory testing should be conducted if a marker for resistance is found and how discordant results should be resolved if an isolate tests susceptible by phenotypic DST (i.e. the extent to which a composite reference standard should be endorsed for individual patient treatment *(20)*). The relative contributions of the following factors should be considered for each mechanism and/or mutation:
   a. the reproducibility of phenotypic DST, the accuracy of the breakpoint used and the prevalence of resistance;
   b. whether a mutation results in MICs close to the breakpoint;
   c. the classification of some mutations according to expert rules or to previous WHO decisions, which may be incorrect. For example, a nonsense mutation one codon before the actual stop codon of *ethA*, *gid*, *katG*, or *pncA* is unlikely to confer resistance. Such exceptions, which could be identified experimentally or by structural modelling, could be excluded from the expert rules by adding them to group 3, 4 or 5, depending on the quality of the evidence.
   d. whether epistasis can confound the interpretation of a mutation (e.g. if *mymA* (*Rv3083*) is naturally overexpressed in some isolates, it could counteract a LoF mutation in *ethA*, although this has not been described to date *(68)*).

[4] Vargas R, Freschi L, Spitaleri A, Tahseen S, Barilar I, Niemann S et al. The role of epistasis in amikacin, kanamycin, bedaquiline, and clofazimine resistance in *Mycobacterium tuberculosis* complex. bioRxiv 2021.05.07.443178.

Detailed methods

Overview

A genotype-phenotype association study for predicting phenotypic DST results for MTBC has four components:

- high-quality phenotypic DST results derived with WHO-endorsed methods and CCs as the reference standard;
- high-quality, standardized WGS for generating unbiased raw sequence data;
- a standardized bioinformatics pipeline for variant detection and annotation; and
- a standardized, validated methodological approach for identifying variants associated with resistance phenotypes, generating statistics on the strength of the associations and confidence grading-associated variants

For the purposes of this analysis, only FASTQ files from Illumina-based sequencing instruments were considered, which, although they differ by instrument model, are all relatively standardized with regard to workflows and error profiles. The bioinformatics workflow for identifying variants and the algorithms for identifying variants "Associated with" and "Not associated with" resistant phenotypes, were adapted from approaches developed by the multinational CRyPTIC Consortium *(43, 75)*, and the confidence grading method was developed in the *Seq & Treat* project *(6, 78)*.

The final methods for these analyses were applied after a series of meetings with multinational experts in sequencing, bioinformatics, biostatistics and mycobacteriology. Methods were proposed, adapted and finalized via webinars and e-mail communications because of travel restrictions due to the COVID-19 pandemic.

Data sources

Raw FASTQ WGS files and associated anonymous metadata, including from phenotypic DST, and limited other clinical and demographic data were collected from:

- the CRyPTIC Consortium,
- ReSeqTB,
- contributors to the WHO surveillance programme,
- multinational TB researchers and
- public health bodies

See "Data contributors".

Curation of phenotypic DST data

All phenotypic DST results associated with the MTBC isolates for which there were also WGS data were collected and analysed. Categorical (resistant, susceptible or intermediate) and/or MIC phenotypic DST data were considered. Intermediate categorical phenotypic DST results were converted to binary results (resistant or susceptible) or excluded according to expert rules. MIC data were converted into categorical binary results (resistant or susceptible) according to CCs appropriate to the phenotypic DST method, as described below. We then stratified the phenotypic DST data into four categories according to the level of WHO endorsement for the method.

Category 1. Phenotypic DST methods currently endorsed by WHO (WHO CURRENT)

Categorical phenotypic DST results for Löwenstein-Jensen, 7H10, 7H11 and BACTEC™ Mycobacterial Growth Indicator Tube™ 960 (MGIT) were regarded as "current" if the CCs in the latest published WHO DST manual were used, with some exceptions *(19)*. We used the recently updated RIF CCs *(4)*. Results for OFX and KAN were regarded as current if they were based on the CCs in the 2018 WHO technical report *(13)*. KAN was still considered, even though it is no longer recommended for TB treatment, as it provides useful insights into the effect of *eis* promoter mutations and *rrs* c1402t on AMK *(13)*. Similarly, although WHO no longer recommends testing of OFX to ensure that it is not used clinically, testing of OFX at concentration *x* is equivalent to testing LFX at *x*/2, given that OFX consists of equal amounts of the active L-isomer of OFX (i.e. LFX) and the largely inactive D-isomer, as reflected in the CCs for both drugs *(13)*. For this reason and because OFX was more widely tested in WHO-endorsed media, phenotypic DST results for OFX and LFX were pooled and reported as LFX in the mutation tables in this report. Similarly, phenotypic DST results for ETO and PTO were pooled and reported as ETO. In some cases, we had to assume that phenotypic DST was conducted by the proportion method with the correct critical proportion, as this information was not collected systematically from all sources. In practice, this assumption was probably correct for the majority of results (i.e. only a minority of testing on Löwenstein-Jensen may have been done with the resistance ratio or absolute concentration method). Microscopic observation drug susceptibility (MODS) results with a CC of 1 mg/L RIF or 0.4 mg/L INH were also considered to be "current" *(77, 78)*.

Category 2. Phenotypic DST methods previously endorsed by WHO (WHO PAST)

This category included phenotypic DST results for Löwenstein-Jensen 7H10, 7H11, MGIT or BACTEC™ 460 obtained either with outdated WHO CCs or simply reported to have relied on WHO CCs without providing the concentration tested, in which case it was not clear which WHO CC was followed *(79–81)*. Again, it was assumed that the proportion method with the correct critical proportion was used.

Category 3. Other phenotypic DST methods

This category consisted primarily of a very large genotypic and phenotypic DST dataset from the CRyPTIC Consortium, which used novel BMD plates manufactured by Thermo Fisher for phenotypic

DST. Although CCs of the Clinical and Laboratory Standards Institute (CLSI) for RIF, INH and EMB exist for the MYCOTB BMD plate, which is also manufactured by Thermo Fisher, use of the novel CRyPTIC BMD plates and CRyPTIC determined CCs described below has not been reviewed or endorsed by WHO. Two different CRyPTIC plates were used to produce raw MIC data, namely the UKMYC5 and UKMYC6 plates. MIC data were translated into binary resistant or susceptible results with the CCs listed in Table 4. The rationale and derivation of these CCs can be found in Fowler et al. *(43)*. Category 3 also included phenotypic DST results derived with methods for which it was unclear whether they met either current or previous WHO guidelines, either because no information on the CC was provided or a CC was available but it was not clear whether 7H10 or 7H11 had been used.

Table 4. Drugs tested with two CRyPTIC BMD plate designs and corresponding CC used to translate MIC data into binary resistant or susceptible. Only drugs considered in this Mutation catalogue analysis are included.

Group	Drug	Concentration range (mg/L)		CC used for interpretation (mg/L)[a]
		UKMYC5	UKMYC6	
First-line	RIF	0.06–4	0.03–8	0.5[b]
	INH	0.025–1.6	0.025–12.8	0.1[c]
	EMB	0.125–8	0.25–32	4[d]
Group A	LFX	0.125–8	0.12–8	1
	MFX	0.6–4	0.06–4	1
	BDQ	0.016–2	0.008–1	0.25
	LZD	0.03–2	0.06–4	1
Group B	CFZ	0.06–4	0.03–2	0.25
Group C	DLM	0.016–1	0.008–0.5	0.125
	AMK	0.25–8	0.25–16	1
	ETO	0.25–8	0.25–8	4
Other[e]	KAN	1–16	1–16	4

[a] Neither method nor CCs endorsed by WHO

[b] Lower than the CLSI CC of 1 mg/L for the MYCOTB plate *(20, 82)*

[c] Equivalent to the CLSI CC of 0.12 mg/L for the MYCOTB plate *(82)*

[d] Identical to the CLSI CC of 4 mg/L for the MYCOTB plate, assuming that the "inconclusive" concentration of 4 mg/L corresponds to an area of technical uncertainty, as defined by the European Committee on Antimicrobial Susceptibility Testing *(82)*

[e] Drug no longer endorsed for TB treatment

Category 4. Excluded phenotypic DST Results

This category included all phenotypic DST results that did not fit into categories 1–3, which were excluded from the analysis.

Prioritization of phenotypic DST results

Given that the phenotypic DST method and the CCs used to evaluate genotype/phenotype associations may be important, we completed separate association analyses with data based on phenotypic DST in category 1 (WHO CURRENT), categories 1 and 2 (WHO CURRENT and PAST) and categories 1, 2 and 3 (ALL). We then compared and contrasted the results for association

according to different categories of phenotypic DST results and our findings, ultimately using the phenotypic DST results from ALL and WHO (WHO CURRENT and PAST) to classify mutations.

Additionally, many MTBC isolates had phenotypic DST results obtained with more than one DST method. In these cases, we formulated a priority order to include only the phenotypic DST data most recently endorsed by WHO. The following hierarchies were used when phenotypic DST results for the same drug obtained with several methods were available for the same isolate:

1. Category 1 > category 2 > category 3.
2. Within the same category, solid media results were deemed more important than liquid media results, because liquid media are more likely to miss key, clinically relevant *rpoB* mutations *(83)*.
3. Of liquid methods within the same category, MGIT > MODS > BACTEC™ 460 > CRyPTIC as BACTEC™ 460 are no longer used, and MGIT has undergone more validation than MODS.

This hierarchical organization enabled the creation of standardized datasets of isolates that had been tested with only the same methods, e.g. with only currently WHO-endorsed methods or any "reasonable" DST method (ALL). While there was probably more variation in the latter dataset, the number of isolates was significantly larger.

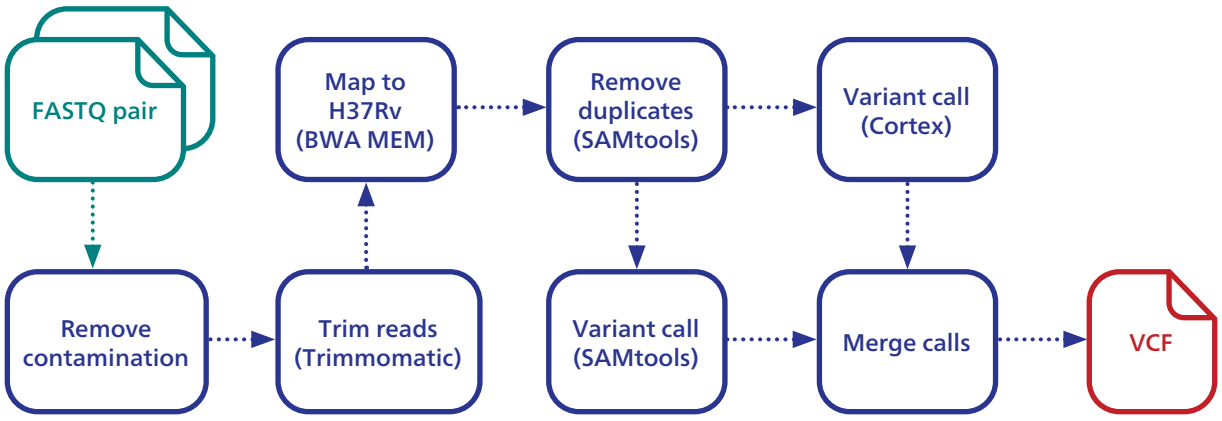
Variant analysis

Bioinformatics pipeline

All raw WGS data were processed with the Clockwork pipeline, originally developed for the CRyPTIC Consortium by a research team at the European Bioinformatics Institute. The full pipeline is available for review at https://github.com/iqbal-lab-org/clockwork.

While there are many variant-callers for Illumina sequencing data, each with different strengths and weaknesses, Clockwork provides a statistically robust means of combining the results of two "callers" to produce a result that is better than either one individually. An overview of the pipeline is shown in Fig. 2. FASTQ file pairs are indexed in a relational database. The pipeline verifies that read file MD5 sums are not duplicated between isolates. Additional details are provided below.

Fig. 2. Scheme of the "Clockwork" bioinformatics pipeline used to process raw WGS data for each isolate. See text for details.



Removal of human and HIV reads

Human and HIV reads were removed with the "remove contamination" pipeline of Clockwork (version 0.1.7). Reads were mapped with BWA MEM to a reference genome containing the *M. tuberculosis* reference genome H37Rv (NCBI Nucleotide database accession ID: NC_000962.3), the human reference genome GRCh38, the HIV genome NC_001802.1 and nasopharyngeal flora genomes from the NIH human microbiome project *(84)*.

A read pair was retained if either read-mapped to H37Rv or if neither read was mapped. A read pair was removed if neither read was mapped to H37Rv and one or both reads were mapped to the HIV or human genome. To avoid incorrect removal of reads, the number of reads mapped to nontuberculous mycobacteria was counted, but this mapping was not used to decide whether or not to remove the reads. Isolates were removed at a later stage according to the proportion of reads mapped to nontuberculous mycobacteria genomes (described later).

Removal of low-quality or non-MTBC isolates

Isolates were removed from consideration for either of the following reasons:

- average depth of paired reads was ≤ 15 when aligned with the H37Rv reference or
- > 5% of the reads mapped to nontuberculous mycobacteria genomes.

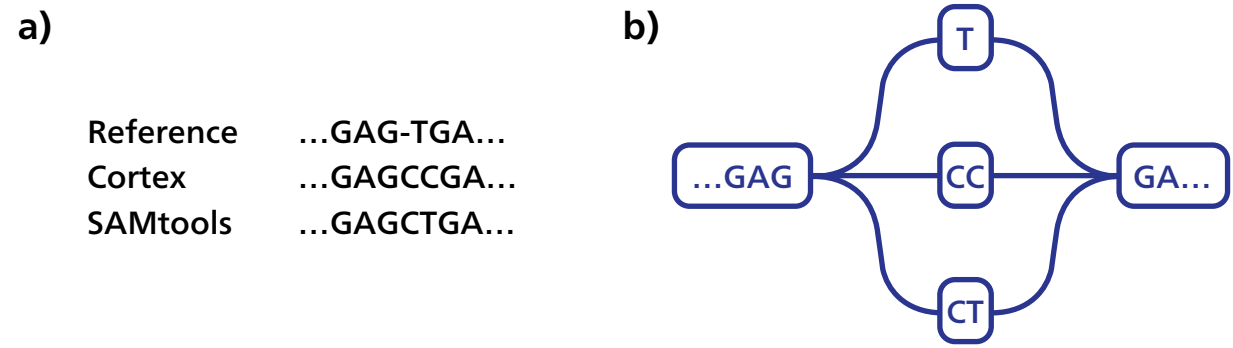
At the end of the "Remove contamination" step, a very rapid quality control process was run, mapping the reads with BWA and calling variants with SAMtools, but with no filtering. From those raw, unfiltered calls, the number of heterozygous SNPs was counted; a count > 100 000 was considered to signal contamination. This is an ineffective filter when applied to *M. tuberculosis* reads, as no two isolates of MTBC differ by that many SNPs; however, it will catch contamination by other species. Note that this is a second line of filtering after the "Remove contamination" step.

Variant calling

Variant calls were made with the "variant call" pipeline of Clockwork (version 0.8.3). Only reads that remained after removal of contamination were used. The stages of the pipeline were as follows.

1. Trim adapters and low-quality ends from the reads with Trimmomatic *(85)*.
2. Map the reads to H37Rv with BWA MEM, and remove duplicates with SAMtools rmdup *(86)*.
3. Call variants independently with Cortex and SAMtools mpileup *(87)*.
4. Merge the call sets from Cortex and SAMtools into a final call set with minos (https://github.com/iqbal-lab-org/minos). Briefly, minos remaps reads to both alternatives when the Cortex and SAMtools disagree and "compares pileups" to find the better call (see Fig. 3).

Fig. 3. Method used to adjudicate a variant site. a) Cortex and samtools calls aligned to the reference sequence. Cortex called a change from T to CC, whereas samtools called a change from T to CT. b) Graph constructed from the reference, cortex and samtools calls. Reads are mapped to this graph, and the allele best supported by the reads is chosen.



The output of the pipeline was a "per-sample" variant calling file (VCF) for each sample. These contained the variants identified in a sample from the reference genome H37Rv. They do not contain information at every position in the genome. In particular, the positions of the variants differed between isolates. These calls included both SNPs and indels (which may be kilobases in length).

Re-genotyping or "joint genotyping" genotype–phenotype intersection isolates

The set of isolates that had both phenotypes and WGS was named the genotype–phenotype intersection (GPI). The GPI sample set was processed further to produce a unified call set comprising the same positions in all of the isolates. Because of technical limitations, this analysis includes only variants shorter than 50 bp. This process is equivalent to the GATK joint genotyping process but consumes less memory.

A graph was generated that showed the union of all variants (shorter than 50 bp) found in the per-sample call sets. This was used as a reference to re-genotype at every variant in every sample, producing a VCF file for each sample. As the VCF files contain the same variant positions, we have information about every sample at every position for which there was evidence for a call. In addition to a re-genotyped VCF for each sample, we also produced a single "wide" VCF file containing all isolates.

The default filters used in Clockwork are a minimum depth of 5x and a fraction of supporting reads (FRS) of 0.9 (at a SNP, ≥ 90% of reads must agree with the genotype).

Genome mask and accessibility filter

Two filters were added to the VCF files:

1. the existing Oxford/PHE/COMPASS mask generated by identifying regions with self-blast matches, which comprises 324 971 bp (or just over 7%) of the reference genome; and

2. positions in the GPI dataset that repeatedly failed our filters. A position was filtered if < 90% of the isolates passed the default variant call filter used by Clockwork, which comprises 95 703 bp of the genome, of which 55 980 bp intersect with the COMPASS mask. These positions are marked in the FILTER column of the VCF as "inaccessible".

Content and uses of the two types of VCF

The per-sample VCF generated by Clockwork records positions at which the sample differs from the reference. These comprise a few hundred to a few thousand SNPs, and each sample generally includes at least one multi-kilobase indel. Like any variant caller, Clockwork has < 100% discovery power, so that some SNPs or indels will be missed, particularly if a sample has lower-than-expected coverage of some region. With just 3x depth, for example, it might refuse to call a SNP. As we are co-analysing thousands of isolates, Clockwork will inevitably detect the same missing SNP in one or many other isolates.

In the "joint genotyping" process, all the variants seen in all isolates are collected and listed in a de-duplicated list. Then, all the isolates are reviewed to confirm whether they are REF or ALT at that position. This has two benefits: the genotypes for low-coverage isolates and positions are recovered, and the results are in uniform VCFs with the same number of lines and the same SNPs and indels. It is therefore possible to make a simple binary matrix that indicates which isolates have which SNPs, which is useful in some analyses.

There is one caveat. For technical reasons, we exclude indels > 50 bp in this process. Therefore, to look at the larger indels in the data, the per-sample VCFs must be consulted.

Heteroresistance

For each SNP recorded, the proportion of reads that supported the called allele was used to generate the FRS. For example, five reads that support A and five that support T would result in an FRS of 0.5. All SNPs with their associated FRS values were recorded in the VCF. Mutations with FRS < 0.9 were flagged and excluded from downstream algorithmic association analyses (see "Association analyses" for details).

Validation of variant calls

We used 17 highly characterized MTBC WGS files supplied by Iñaki Comas to determine the accuracy of the Clockwork variant calls. Each sample had multiple PacBio and Illumina WGS reads. We considered the polished PacBio assembly of each sample to be the reference "truth" (with caveats below). We "spiked" the Illumina reads into the joint genotyping process to use them for evaluating error rates.

The following methods are included in a new "verifier" tool available at https://github.com/iqbal-lab-org/varifier.

Precision

To decide whether a variant call was correct, a "probe" sequence was generated, consisting of the called allele plus the 1000 bp of the reference genome flanking the allele. This probe sequence was mapped to the truth genome (i.e. the PacBio assembly of a sample). If the probe mapped and the allele sequence had no mismatches, the variant call was considered to be a true positive; otherwise it was called a false positive. Note that mismatches were still allowed in the flanking sequence mapping, without affecting whether the call was considered a true positive.

Recall

A "truth" set of variants was determined as follows. Two assembly alignment methods, Minimap2 plus paftools *(88)* and the show-snps command from MUMmer *(89)*, were used to collect a putative list of "truth variants" between a PacBio assembly and H37Rv. The union of these two call sets was taken and false positives removed with the probe mapping method described under Precision. This resulted in a truth call set consisting of all the variants that should be found in the PacBio "truth" assembly and H37Rv.

Next, the variant calls from Clockwork were applied to H37Rv to form a "mutated" genome. To determine recall, we must determine whether every variant in the truth call set (as described in the previous paragraph) is found in the mutated genome. The same probe mapping method as for precision was used, but this time the VCF to be tested is the truth call set, and the mutated genome is used as the "truth" genome.

Results of Clockwork pipeline validation

The precision and recall results are shown in Table 5. "All" denotes that all unfiltered calls from the Clockwork VCF files were used. "Filter" indicates that calls were filtered with the default filter from Clockwork, the COMPASS mask and the inaccessible Filter described earlier. "Filter recall" indicates the recall, after exclusion of true variants which fall in the COMPASS mask or fail the inaccessibility filter.

Table 5. Precision and recall results after validation in the Clockwork pipeline of WGS data from highly characterized isolates

Isolate ID	All precision	All recall	Filter precision	Filter recall
N0004	0.980	0.804	0.999	0.940
N0031	0.977	0.831	1.000	0.935
N0052	0.968	0.703	0.999	0.920
N0054	0.980	0.810	0.999	0.932
N0072	0.979	0.811	1.000	0.970
N0091	0.987	0.888	1.000	0.961
N0136	0.969	0.787	1.000	0.957
N0145	0.983	0.735	0.998	0.967
N0153	0.972	0.744	0.998	0.978
N0155	0.976	0.722	0.998	0.948
N0157	0.988	0.848	0.999	0.959

Isolate ID	All precision	All recall	Filter precision	Filter recall
N1176	0.993	0.845	0.998	0.942
N1177	0.987	0.834	1.000	0.958
N1202	0.961	0.879	0.978	0.955
N1216	0.962	0.778	1.000	0.919
N1272	0.994	0.771	0.999	0.965
N1283	0.978	0.684	1.000	0.885

Annotation of variants

The GenBank NC_000962.3 version of the sequence and annotation of *M. tuberculosis* H37Rv was used as the reference genome. The standard single-letter code for nucleotides and amino acids was employed, with the exception that "!" was used instead of "*" for stop codons. The genomic position was included for all upstream changes and indels.

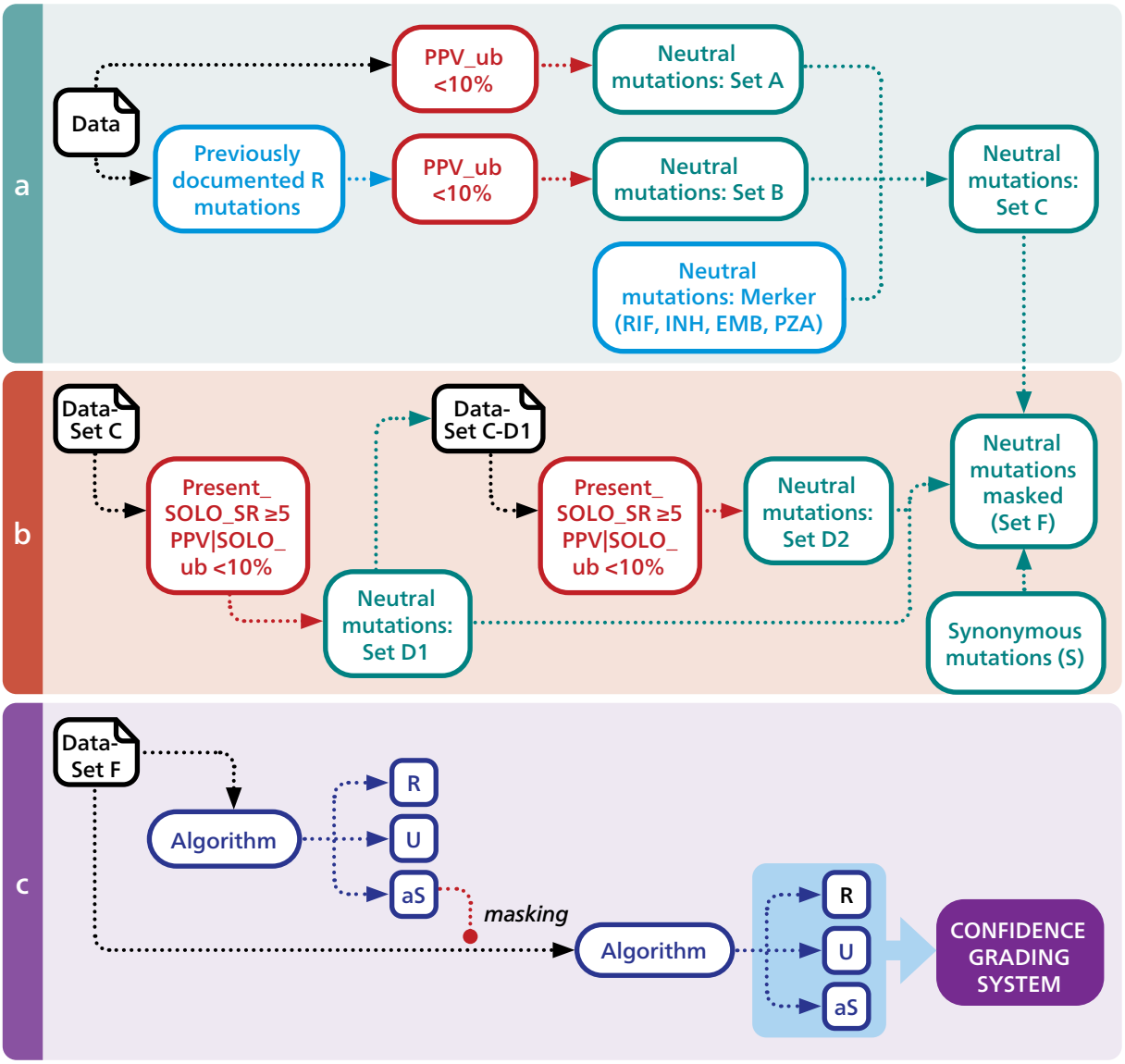
Association studies

An algorithmic method for identifying "solo" mutations

The Mutation catalogue of individual mutations associated independently with resistance was produced with a method first applied to MTBC in 2015 *(11)*. The approach closely resembles the "definite defectives" algorithm used in the field of "group testing", pioneered in 1943 *(92, 93)*. When applied to MTBC genomic data, it served to characterize the effects of specific mutations in a range of genes and promoters that are considered highly probably linked to phenotypic resistance. Deviations from the published method *(11)* are detailed below.

The algorithm, outlined in Fig. 4, characterizes non-synonymous mutations and indels as resistance-determining ("R") or as consistent with susceptibility ("S"). Synonymous mutations were all assumed a priori to be neutral ("S") and masked, unless they abolished a start codon, in which case they were assessed algorithmically with other types of mutations (because they were effectively non-synonymous). Mutations that abrogated the stop codons were not considered non-synonymous in the pipeline and were therefore not accounted for in the analysis but will be considered in the next version of the Mutation catalogue. Any non-synonymous mutation or indel identified as the only such mutation in the set of gene and promoter regions examined (i.e. a solo mutation) in a phenotypically resistant isolate was classified as "R". Any non-synonymous mutation or indel featured only in phenotypically susceptible isolates or always associated with a susceptible phenotype when it is the only non-synonymous mutation or indel in the set of examined gene and promoter regions, was algorithmically classified as a mutation consistent with "susceptibility" ("algorithmic S" or "aS"). Any mutation that could not be characterized as either "R" or "aS", such any mutation that never appeared as a solo mutation and was n ever seen in phenotypically susceptible strains, was classified as uncertain ("U"). The resulting solo mutations were further classified in a series of grading criteria described below.

Fig. 4. Steps in the algorithm for determining neutral mutations (a and b) and quantifying associations between genotype and phenotype (c)



Before running the algorithm, several preliminary steps were taken to prepare the raw data for analysis.

1. Selection of candidate genes and promoters

The success of the algorithmic approach requires exploration of relatively short genetic sequences. In the method published in 2015, this range was 1–8 genes and up to an arbitrarily defined 100 bp upstream, which was curtailed if it ran into an adjacent coding sequence. As several additional candidate genes have come to light since 2015 and as most of the promotor sequences identified for this analysis were longer than the arbitrary 100 bp, for this analysis, we divided the candidate genomic regions into two tiers to avoid simultaneous excessive genomic variation. Tier 1 comprised gene sequences considered most probably to contain resistance mutations (Table 6). Tier 2 included genes with a reasonable pre-test probability of containing resistance

mutations with the additional, bespoke promotor sequences not already included in the 100-bp rule defined in the original method.

Several candidate genes were selected for inclusion in the analysis, each with support in the literature for a high probability of containing genetic variation associated with resistance to one or more drugs. Associated promotor regions were also identified from the published literature, some of which run into adjacent coding sequences, others of which are operon promotors upstream of other genes and thus not necessarily contiguous with the coding sequence of interest.

Table 6. Candidate genes determined to be associated with phenotypic resistance to the listed TB drugs

Drug	Tier 1	Tier 2	References
INH	*ahpC, inhA, katG*	*mshA, ndh, Rv1258c, Rv2752c*	*6, 21, 92–96*
RIF	*rpoB*	*rpoA, rpoC, Rv2752c*	*6, 95*
EMB	*embA, embB, embC*	*embR, ubiA*	*24, 97, 99*
PZA	*pncA, clpC1, panD*	*Rv1258c, PPE35, Rv3236c*	*95, 96, 99, 100*
FQ	*gyrA, gyrB*	None	*6*
BDQ	*pepQ, Rv0678, mmpL5, mmpS5, atpE*	*Rv1979c*	*7*
LZD	*rplC, rrl*	None	*7*
CFZ	*pepQ, Rv0678, mmpL5, mmpS5*	*Rv1979c*	*7*
DLM	*fgd1, ddn, fbiA, fbiB, fbiC, Rv2983*	None	*50*
AMK	*rrs, eis, whiB7*	*whiB6, ccsA, fprA, aftB*	*6, 95, 101*
STM	*rrs, rpsL, gid, whiB7, Rv1258c*	*whiB6*	*6, 95, 96, 101*
ETO	*inhA, ethA*	*ethR, mshA, Rv3083, ndh*	*6, 21, 68, 93, 94, 102*
KAN	*rrs, eis, whiB7*	None	*6, 95, 101*
CAP	*rrs, tlyA*	*whiB6, ccsA, fprA, aftB*	*6, 95*

Tier 1 comprises genes considered most likely to contain the resistance mutations. Tier 2 comprises genes that are reasonably likely to contain resistance mutations, with the additional, literature-defined promotor sequences. Tier 2 genomic regions were investigated only for isolates in which the presence of a tier 1 mutation did not preclude its interpretation. The exact genomic coordinates used are available on request.

2. Nucleotide null calls and calls that fail pipeline filters

Wild-type sequence was assumed for any position for which the Clockwork pipeline did not generate an entry in the VCF. When Clockwork reported a "null call", i.e. there were either no data available for a nucleotide position or there was the same degree of support for two competing nucleotides, the position was masked to downstream analysis and was effectively read as a wild-type sequence. Clockwork reported a "filter-fail" for genomic loci with evidence for more than one nucleotide, such as may be observed in mixed infections (i.e. FRS < 0.9; see "Heteroresistance" above for details on FRS flagging). These calls were also masked unless the variant existed in a non-mixed form in at least one other isolate, suggesting that it was a biologically plausible variant rather than a sequencing artefact. This avoided misclassification of compensatory mutations as "resistant", as these would otherwise have appeared as solo mutations (e.g. *rpoC* mutations accompanying mixed (heteroresistant) or null calls at *rpoB* S450L).

3. Quality control

Sample mislabelling is relatively common and can lead to spurious results in algorithm-based analyses such as these. One way of mitigating such error is to exclude all isolates with a previously well-established resistance mutation but a susceptible phenotype in the dataset, i.e. isolates in which a susceptible phenotype is not credible. With this logic, MTBC isolates that had a *katG* S315T mutation for INH resistance or a *rpoB* S450L mutation for RIF resistance but were recorded as having "susceptible" phenotypes for the corresponding drug in the data were excluded from further consideration, on the assumption that these mutations were best explained by sample mislabelling. Additionally, isolates with phenotypes for which the corresponding candidate genomic regions contained an excessive number of non-variant, non-wild type calls (i.e. null calls or positions at which a variant filter failed) were also excluded. "Excessive" was defined as < 1% probability of at least that many non-variant, non-wild type calls for any set of candidate genomic regions, as computed by Poisson distribution.

4. Initial identification of neutral mutations

As the aim of the algorithmic approach is to identify specific lone mutations associated with resistance ("solos"), a preliminary step should optimally be taken to identify as many "neutral" (i.e. not associated with resistance) mutations as possible and to mask them to downstream analyses. Isolates with categories 1 and 2 phenotypic DST results were analysed separately in four ways to identify neutral mutations. In each case, the quality control steps (see "Quality control steps" above) were implemented, and uncertain base calls were masked (see "Handling of null calls" above). Any mutation with a PPV or PPV|SOLO < 10% with 95% confidence was considered neutral. To ensure a conservative approach, phenotypes unique to category 3 (i.e. not WHO-endorsed) were not used to identify neutral mutations.

Neutral mutations were identified in a stepwise fashion described below and shown in Fig. 4a and 4b.

(i) $\text{PPV} = \frac{\text{Present\_R}}{\text{Present\_R} + \text{Present\_S}}$ was computed for each mutation.

(ii) The PPV was computed for each mutation after removal of isolates that contained one of a number of previously documented resistance mutations (see Table 7).

(iii) $\text{PPV|SOLO} = \frac{\text{Present\_SOLO\_R}}{\text{Present\_SOLO\_R} + \text{Present\_S}}$ was computed for each mutation that remained as the only (SOLO) mutation in a set of candidate genes after the masking of the neutral mutations identified in steps (i) and (ii).

(iv) PPV|SOLO was computed for each mutation that remained as the only (SOLO) mutation in a set of candidate genes after masking of the neutral mutations identified in steps (i), (ii) and (iii).

A list of mutations identified specifically as neutral by Merker et al. *(8)* was then appended to the results, and all neutral mutations were masked from all downstream analyses.

Table 7. Previously documented resistance mutations used in the algorithm for identifying neutral mutations

INH	*fabG1* (a-16g, c-15t, t-8a, t-8c, or g609a L203L), *inhA* S94A, or any amino acid substitution or indel at *katG* 315	*16, 21, 71*
RIF	*rpoB* (V170F or I491F) or any amino acid substitution/indel between *rpoB* codons 426 and 452 (i.e. RRDR)	*4*
EMB	*embB* (M306I, M306L, M306V, D354A, G406A, G406C, G406D, G406S, or Q497R)	*10, 24*
PZA	*pncA* a-11g, or any amino acid substitution/indel in *pncA*, except I6L and L35R	*28*
FQ	*gyrA* (G88A, G88C, D89N, A90V, S91P, D94A, D94G, D94H, D94N, or D94Y), *gyrB* A504V or plus any amino acid substitution/indel between *gyrB* codons 497 and 502	*6, 13, 16*
AMK	*rrs* (a1401g, c1402t, or g1484t) or *eis* c-14t	*13*
STM	*rrs* (a514c or c517t) or *rpsL* (K43R, K88Q, or K88R)	*6*
ETO	*fabG1* (a-16g, c-15t, t-8a, t-8c, or g609a L203L) *inhA* S94A	*16, 21, 71*
KAN	*rrs* (a1401g, c1402t, or g1484t) or *eis* (g-37t, c-14t, c-12t, or g-10a)	*13*
CAP	*rrs* (a1401g, c1402t, or g1484t) or *tlyA* N236K	*6, 13, 103*

Isolates containing these mutations were removed during step (ii) of the process for identifying neutral mutations.

Running the algorithm to identify mutations associated with resistance

The algorithm was run separately for each phenotypic data type (categories 1–3 above). In each instance, the quality control steps outlined above were implemented first, then the neutral mutations identified in Fig. 4a and 4b were masked, with synonymous mutations (unless these abolish start codons), before the final algorithm was run for a total of two iterations (Fig. 4c).

Prioritizing gene targets

As the algorithm approach is hierarchical, tier-1 gene sequences (with a higher probability of association with phenotypic resistance) were investigated first. For a resistant phenotype linked to an "R" or "U" mutation, no additional sequence was analysed to explain that phenotype. Tier-2 gene sequences were later investigated for the remaining data. Two passes of the algorithm were performed for tier-1 sequences (Fig. 4c), whereby mutations characterized as "aS" in the first pass were masked in the second pass in order to characterize further, now solitary mutations as "aS" or "R". Tier-2 sequences were assessed only after the second pass of tier-1 sequences. Only one pass of the algorithm was performed on tier-2 sequences because of the lower prior probability of finding resistance-associated mutations in those targets.

Strengths and limitations of the algorithmic approach

The strengths of the algorithmic approach are that it is known to perform well for building catalogues *(11)*. Its theoretical underpinning in the field of DST provides additional confidence *(90)*. By focusing on a set of candidate sequences with a high probability of being associated with phenotypic resistance, the chances of correctly identifying solo "R" mutations is enhanced, although the risk of not interrogating additional relevant sequences remains. In the original method, all phylogenetically deep-rooted mutations ("PhyloSNPs") were masked, while, in this

analysis, these were masked only if there was sufficient evidence of their neutrality. This analysis may therefore be more vulnerable to confounding due to population structure, but may be simultaneously less prone to the problem of arbitrary masking of the wrong mutations.

A further potential weakness is that null calls and many filter-fail calls were masked under the implicit assumption that these are most likely to be wild-type. For some null and filter-fail calls, this may not be the case. Similarly, mutations that abolish stop codons were not considered. In addition, as it is phenotypes and not isolates that are assigned to analytical categories 1–3, phenotypes from an isolate may feature in category 1 or 2, but the entire isolate might be removed from category 3 if a RIF or INH phenotype introduced in category 3 fails quality control (e.g. a *katG* S315T mutation in a susceptible phenotype). This situation was observed once in the entire dataset and was therefore tolerated for consistency.

Statistical support for resistance mutations

The numbers of resistant and susceptible isolates with and without a mutation were collated into a 2 x 2 contingency table from which ORs were computed in a Fisher's exact test, with corresponding CIs and *P* values according to the hypergeometric distribution. To control for multiple testing, a Benjamini-Hochberg correction procedure was used, with an FDR of 5%. This correction was applied to a number of tests equal to the number of R, aS and U mutations identified by the algorithm for a given drug, plus the number of mutations that were algorithmically identified as neutral in the preliminary steps. As they were excluded before the algorithm was run, the number of tests used for the correction did not include those of mutations labelled as neutral solely on the basis of Merker et al. *(8)*. The same statistical procedure was applied to the ORs for solo mutations (OR SOLO), whereby only isolates with a solo mutation were counted instead of all isolates with a mutation and compared with the corresponding numbers of isolates without the mutation:

$$\text{OR SOLO} = \frac{\text{Present\_SOLO\_R}}{\text{Absent\_R}} \Big/ \frac{\text{Present\_SOLO\_S}}{\text{Absent\_S}}$$

The PPV|SOLO was calculated for all mutations as the number of times the mutation was observed as a solo with a resistant phenotype divided by the sum of that number and the number of times the mutation was observed with a susceptible phenotype; 95% CIs were obtained by the Clopper-Pearson method. These statistical metrics were then used to stratify and prioritize mutations by the confidence grading approach described below, and the algorithmic labels "R", "aS", "S" and "U" were replaced according to the grading criteria.

Because of the quality control steps taken before these analyses, mutations at *katG* S315T and at *rpoB* S450L would necessarily have appeared to be perfectly associated with resistance to INH and RIF, respectively. To generate real-world data for these mutations, a separate analysis was performed in which those to INH and RIF were reanalysed without removing probably mislabelled isolates. Only the results for *katG* S315T and *rpoB* S450L were kept from that analysis and substituted for the results for those mutations in the main analysis.

Confidence grading

Once the algorithm had identified variants associated with and not associated with resistant phenotypes, and relevant association statistics were generated as described in "Association studies" above, a set of consensus statistical thresholds and expert rules for confidence grading and ranking the observed MTBC mutations were applied to stratify the data into five groups according to the strength of the evidence for a genotype–phenotype association and the level of support for the phenotypic method used (see "Prioritization of phenotypic DST results" above). According to these grading criteria, we stratified the final mutation associations into five groups:

1. Associated with R
2. Associated with R – Interim
3. Uncertain significance
4. Not associated with R – Interim
5. Not associated with R

Grading criteria (see specific criteria below) were applied equally to all mutations for all drugs, as shown in Fig. 5. As individual mutations associated with PZA resistance are both found less frequently and distributed more broadly among genes such as *pncA* than other resistance mutations (such as those in *rpoB*), they required special consideration. Therefore, we applied "relaxed" grading criteria with less stringent thresholds to identify additional, infrequent mutations associated and not associated with resistance to PZA in the *pncA* gene only (Fig. 6).

Fig. 5. Grading of mutations that confer resistance to drugs

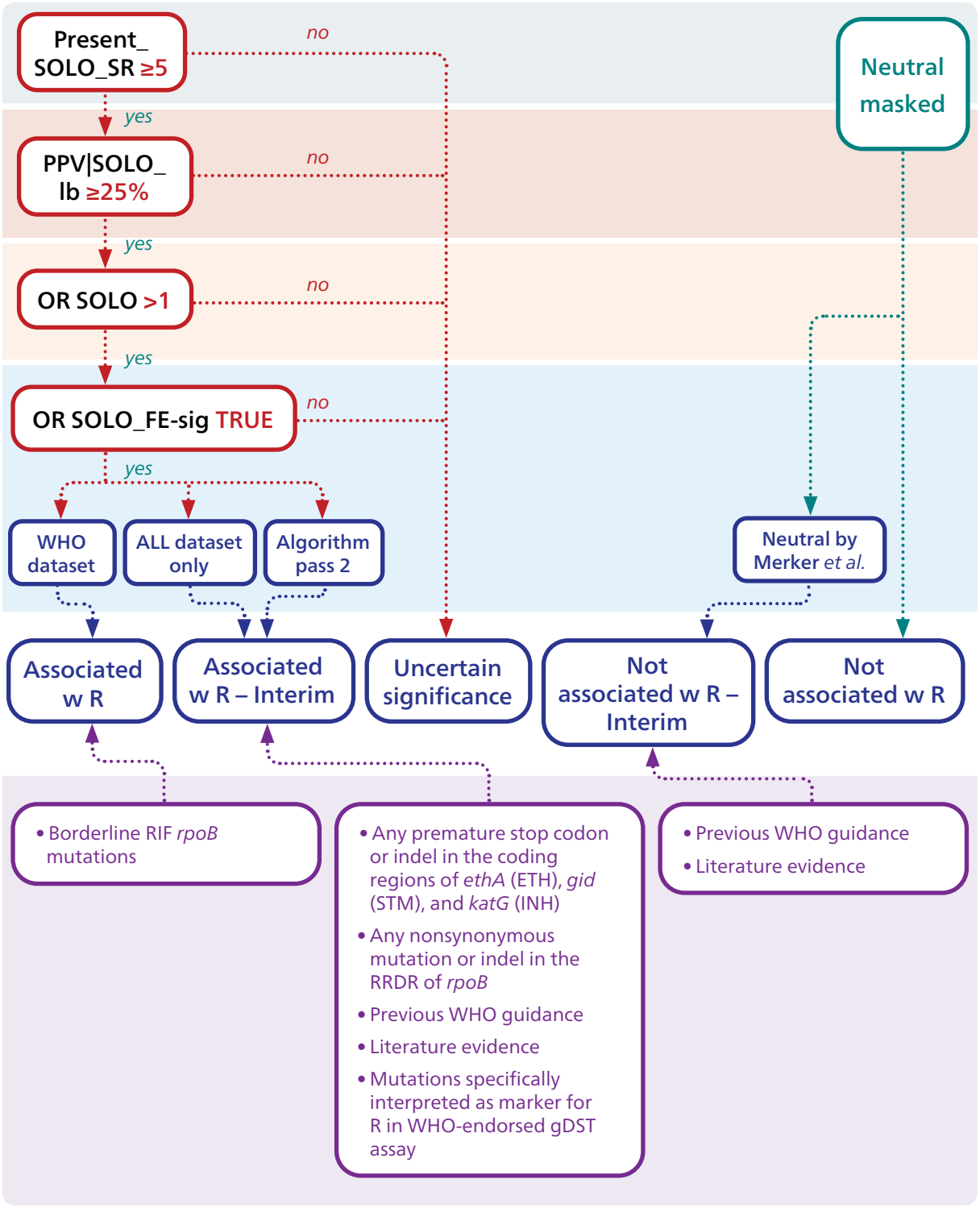
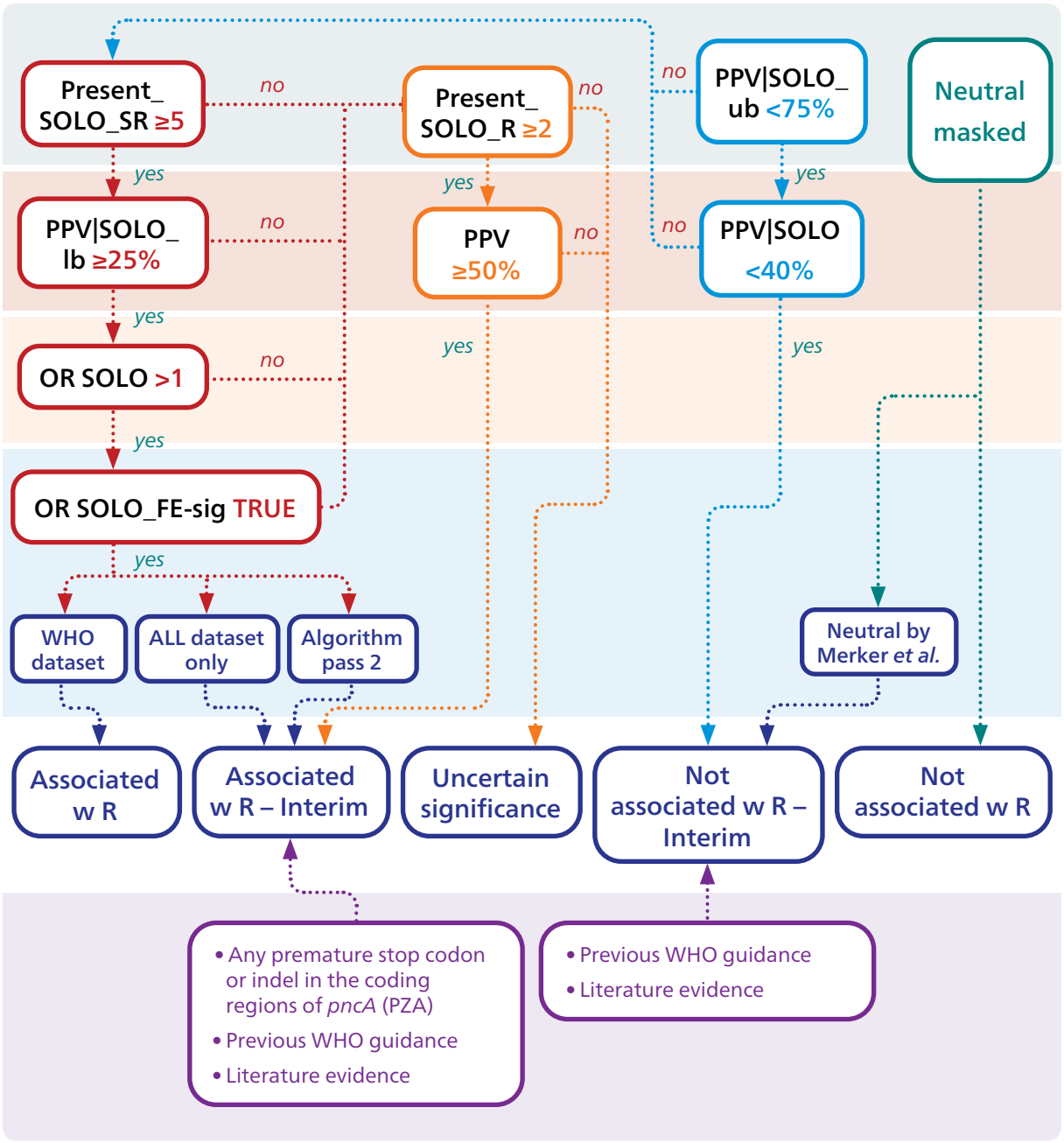


Fig. 6. Grading of mutations associated with PZA resistance



General principles for the grading approach and presentation of mutation tables

- While most mutations were identified from the data set we analysed, some assigned to drug susceptibility 1 (Assoc w R), drug susceptibility 2 (Assoc w RI) or drug susceptibility 4 (Not assoc w RI) were identified from the literature and published WHO documents, according to expert rules. Any mutations included in the tables from these sources are clearly marked as such. All mutations not flagged as such were processed in the algorithm and then classified according to the grading criteria detailed below. It should be noted that the expert rules were based on the available evidence and will be revised as new evidence emerges.
- The grading criteria used to stratify mutations into drug susceptibility 2 or 4 (Associated or not associated with R – Interim) were generally more permissive than those for drug susceptibility 1 or 5 (Associated/Not associated with R) or based solely on phenotypic DST methods that have not been endorsed by WHO.
- Drug susceptibility 3 mutations (Uncertain significance) could not be classified with the available data and comprise all the mutations that do not fall into group 1, 2, 4 or 5.
- While the MICs of the mutations analysed in this dataset were not evaluated to determine whether the observed mutations were associated with high or low MICs to specific drugs, WHO has previously endorsed specific mutations associated with "high" or "low" MICs to INH and MFX *(4, 13, 16, 17)*. These mutations are therefore flagged as "high" or "low" in the "WHO-UCN-GTB-PCI-2021.7-eng.xlsx" but were not flagged as such in the tables in this report.

Criteria for initial confidence grading

Group 1: Associated with R

Mutations that met five criteria:

1. sum of resistant and susceptible isolates with the solo mutation (Present_SOLO_SR) ≥ 5
2. lower bound of 95% CI of PPV conditional on being solo (PPV|SOLO_lb) ≥ 25%
3. OR > 1 (always applies if criterion 4 is met)
4. OR SOLO mutation > 1
5. statistical significance of OR for solo mutation (OR SOLO_FE-sig) with Fisher exact FDR test, corrected for FDR, $P \leq 0.05$

Group 2: Associated with R – Interim

Mutations that met "relaxed" criteria for *pncA*:

1. resistant isolates with the solo mutation (Present_SOLO_R) ≥ 2
2. PPV ≥ 50%

Group 3: Uncertain significance

All mutations that did not meet the criteria for inclusion in group 1, 2, 4 or 5.

Group 4: Not associated with R – Interim

Mutations that met "relaxed" criteria for *pncA*:

1. PPV conditional on being solo (PPV|SOLO) < 40%
2. upper bound of 95% CI of PPV conditional on being solo (PPV|SOLO_ub) < 75%

Group 5: Not associated with R

Neutral mutations that were masked before use of the algorithm (see "Initial identification of neutral mutations")

Additional criteria for final confidence grading

When the initial confidence grades of mutations selected according to statistical thresholds were in conflict with strong evidence before this analysis, expert rules and established precedent were applied. As far as possible, the grading was based on the available data, and such changes were kept to a minimum. When additional criteria were used to override the initial confidence grading, the mutation in question was annotated with the specific criterion applied in the mutation tables. The basis for these changes and the abbreviations for each criterion are described below.

Group 1 (Associated with R):

1. Two "borderline" *rpoB* resistance mutations (L430P and H445N) were included in group 1 even though they fell into group 3 during initial confidence grading. Previous WHO documents explicitly state that these mutations are valid markers of resistance, do not require confirmation by phenotypic DST and that their detection supersedes a susceptible phenotypic DST result *(4)*.

Group 2 (Associated with R – Interim):

1. All only: Tier-1 mutations that met the criteria for group 1 during initial confidence grading only because the ALL dataset contained methods that were not endorsed by WHO
2. Pass 2: Tier-1 mutations that met the criteria for group 1 only during initial confidence grading in pass 2 of the algorithmic method (i.e. after masking mutations classified as neutral during pass 1).
3. Tier 2: Tier-2 mutations that met the criteria for group 1 during the initial confidence grading (i.e. mutations that rely on algorithmic decisions about tier 1-mutations). This did not apply to any of the mutations in this analysis.
4. Previous WHO guidance ("prev. WHO"): all mutations listed as associated with resistance by Miotto et al. after correction of *P* values for CAP, LFX and STM. Miotto et al. formed the basis of the 2018 WHO next-generation sequencing guide, which was used to analyse WGS data from surveillance studies *(6, 12, 36)*.
5. WHO-endorsed genotypic DST assays ("WHO-end. gDST"): any mutations specifically interpreted as markers of resistance in a WHO-endorsed assay. This included *gyrB* N499D, which is specifically detected by a mutant probe in the GenoType MTBDR*s*l V2.0, and *rrs* c1402t,

which is specifically interpreted as a marker for AMK resistance by Xpert MTB/XDR *(4, 17, 29)*[5]. This rule was not used to upgrade mutations that did not meet the grading criteria when they were inferred and not specifically detected in a WHO-endorsed assay probe (e.g. any mutation in *gyrB* codons 497–502 is inferred by lack of probe binding). Given that the binding regions of some probes in Hain assays have not been disclosed, we were unable to mark some mutations that are inferred (e.g. some *eis* promoter mutations *(17)*). Because Molbio has not published this information for the Truenat MTB-RIF Dx, this assay could not be included *(4).*

6. RRDR expert rule ("RRDR"): Any non-synonymous mutation or indel in the RRDR of *rpoB (4)*.
7. FQ cross-resistance expert rule ("FQ X-R"): Any *gyrA* or *gyrB* mutation for LFX group 1 or 2 during initial confidence grading that fell into group 3 for MFX during initial confidence grading was "upgraded" to group 2 or vice versa.
8. LoF expert rule ("LoF"): Any premature stop codon (i.e. nonsense mutation) or indel in the coding regions of *ethA* (ETO), *gid* (STM), *katG* (INH) and *pncA* (PZA) was considered a group 2 mutation for that drug. The mechanisms by which LoF mutations in these genes confer resistance is well understood, and no epistatic interaction has been observed that could render an isolate with a LoF mutation in one of these genes susceptible *(23, 27, 28, 54, 64, 68, 104, 105)*.
9. Literature evidence expert rule ("Lit."): This rule was used to support classification of the *pncA* I31T mutation into drug susceptibility 2 for PZA *(28)*.

Group 4 (Not associated with R – Interim):

1. Previous WHO guidance: mutations previously documented as "Not associated with R" *(6, 12)* that did not meet the criteria for group 1, 2 or 5.
2. Literature evidence expert rule ("Lit."): mutations previously documented as "Not associated with R" that are frequent in some settings were placed in group 4:
   a. Mutations that could not be classified as neutral in the CURRENT or CURRENT + PAST datasets but that were classified as neutral by Merker et al. *(8)* according to phenotypic DST data and consequently masked before use of the algorithmic method.
   b. *gyrA* T80A and A90G, because these are frequent in the Uganda genotype *(6, 15)*.

[5] S. Chakravorty. Personal communication, 2021.

Data contributors

Solange Alves Vinhas, Federal University of Espírito Santo, Vitória, Brazil; **Jason Andrews**, Stanford University School of Medicine, Stanford (CA), USA; **Simone Battaglia**, San Raffaele Scientific Institute, Milan, Italy; **Marcel Behr**, McGill University Health Centre, Montréal, Canada; **Emanuele Borroni**, San Raffaele Scientific Institute, Milan, Italy; **Andrea Cabibbe**, San Raffaele Scientific Institute, Milan, Italy; **Joshua Carter**, Oxford University, United Kingdom; **Angkana Chaiprasert**, Mahidol University, Bangkok, Thailand; **Alvaro Chiner-Oms**, Biomedicine Institute of Valencia, Spain; **Joconiah Chirenda**, University of Zimbabwe, Harare, Zimbabwe; **Daniela Cirillo**, San Raffaele Scientific Institute, Milan, Italy; **Pauline Claxton**, Scottish Mycobacteria Reference Laboratory, Edinburgh, United Kingdom; **David Clifton**, Oxford University, United Kingdom; **Ted Cohen**, Yale University, New Haven (CT), USA; **Iñaki Comas**, Biomedicine Institute of Valencia, Spain; **Tom Connor**, Public Health Wales, Cardiff, United Kingdom ; **Marcelo Cordeiro dos Santos**, Doctor Heitor Vieira Dourado Foundation of Tropical Medicine, Manaus, Brazil; **Jorge Coronel**, London School of Hygiene & Tropical Medicine and Cayetano Heredia Peruvian University, Lima, Peru; **Julio Croda**, Federal University of Mato Grosso do Sul, Campo Grande, Brazil; **Derrick Crook**, Oxford University, United Kingdom; **Keertan Dheda,** University of Cape Town, Cape Town, South Africa; **Sarah Dunstan**, University of Melbourne, Australia; **Tatiana Dutra**, Federal University of Rio de Janeiro, Brazil; **Sarah Earle**, Oxford University, United Kingdom; **Vincent Escuyer**, New York State Department of Health, Albany (NY), USA; **Kiatichi Faksri**, Khon Kaen University, Khon Kaen, Thailand; **Maha Farhat**, Harvard Medical School, Boston (MA), USA; **Lucilaine Ferrazoli,** Institute Adolfo Lutz, São Paulo, Brazil; **Philip Fowler**, Oxford University, United Kingdom; **Sebastien Gagneux**, Swiss Tropical and Public Health Institute, Basel, Switzerland; **Neel Gandhi**, Emory University, Atlanta (GA), USA; **George Gao,** China CDC, Beijing, China; **Jennifer Gardy**, University of British Columbia, Vancouver, Canada; **Amy Gaskin**, Public Health Wales, Cardiff, United Kingdom; **Arash Ghodousi**, San Raffaele Scientific Institute, Milan, Italy; **Ana Gibertoni Cruz**, Oxford University, United Kingdom; **Simon Grandjean Lapierre**, Institut Pasteur de Madagascar, Antananarivo, Madagascar; **Jennifer Guthrie**, Public Health Ontario, Toronto, Canada; **Tanya Halse**, New York State Department of Health, Albany, USA; **Guangxue He**, China CDC, Beijing, China; **Harald Hoffmann**, Institute of Microbiology and Laboratory Medicine, Gauting, Germany; **Kathryn Holt**, Monash University, Melbourne, Australia; **Sarah Hoosdally**, Oxford University, United Kingdom; **Martin Hunt**, European Bioinformatics Institute, London, United Kingdom; **Stella Huo**, University of California San Francisco (CA), USA; **Zamin Iqbal**, European Bioinformatics Institute, London, United Kingdom; **Frances Jamieson**, Public Health Ontario, Toronto, Canada (retired); **Lisa Jarrett**, Public Health England, Birmingham, United Kingdom; **Bouke de Jong**, Institute of Tropical Medicine, Antwerp, Belgium; **Lavania Joseph**, National Institute for Communicable Diseases, Pretoria, South Africa; **Ruwen Jou**, CDC Taiwan, Taipei, Taiwan (China); **Priti Kambli**, P.D. Hinduja National Hospital and Medical Research

Centre, Mumbai, India; **Rukhsar Khot**, P.D. Hinduja National Hospital and Medical Research Centre, Mumbai, India; **HeeJin Kim**, Korean Institute of Tuberculosis, Seoul, Republic of Korea; **Jeff Knaggs**, European Bioinformatics Institute, London, United Kingdom; **Anastasia Koch**, University of Cape Town, South Africa; **Thomas Kohl**, Forschungszentrum Borstel, Germany; **Samaneh Kouchaki**, Oxford University, United Kingdom; **Afrânio Kritski**, Federal University of Rio de Janeiro, Brazil; **Julianna Kus**, Public Health Ontario, Toronto, Canada; **Alexander Lachapelle**, Oxford University, United Kingdom; **Ajit Lalvani,** Imperial College London, United Kingdom; **Pascal Lapierre**, New York State Department of Health, Albany (NY), USA; **Ian Laurenson**, Scottish Mycobacteria Reference Laboratory, Edinburgh, United Kingdom; **Brice Letcher,** European Bioinformatics Institute, London, United Kingdom; **Wan-Hsuan Lin**, CDC Taiwan, Taipei, Taiwan (China); **Chunfa Liu,** China CDC, Beijing, China; **Maria Lucia Rossetti**, Centre for Scientific and Technological Development, Porto Alegre, Brazil; **Maria Luiza Bazzo**, Federal University of Santa Catarina, Florianópolis, Brazil; **Aijing Ma,** China CDC, Beijing, China; **Juliana Maira Watanabe Pinhata**, Institute Adolfo Lutz, São Paulo, Brazil; **Kerri Malone**, European Bioinformatics Institute, London, United Kingdom; **Ayan Mandal**, Foundation for Medical Research, Mumbai, India; **Flora Martinez**, Federal University of Grande Dourados, Brazil; **Neil Martinson**, University of the Witwatersrand, Johannesburg, South Africa; **Daniela Matias**, Public Health England, Birmingham, United Kingdom; **Vidya Mave**, BJ Medical College, Pune, India; **Conor Meehan**, University of Bradford, Bradford, United Kingdom; **Fabrizio Menardo**, Swiss Tropical and Public Health Institute, Basel, Switzerland; **Carl Mendel**, TB Alliance, London, United Kingdom (retired); **Matthias Merker**, Forschungszentrum Borstel, Germany; **James Millard,** Africa Health Research Institute, Durban, South Africa; **Paolo Miotto**, San Raffaele Scientific Institute, Milan, Italy; **Nerges Mistry**, Foundation for Medical Research, Mumbai, India; **David Moore**, London School of Hygiene & Tropical Medicine and Cayetano Heredia Peruvian University, Lima, Peru; **Kimberlee Musser**, New York State Department of Health, Albany (NY), USA; **Han de Neeling**, National Institute for Public Health and the Environment, Bilthoven, Netherlands; **Isabela Neves de Almeida**, Federal University of Minas Gerais, Belo Horizonte, Brazil; **Dumisani Ngcamu**, National Institute for Communicable Diseases, Pretoria, South Africa; **Hoang Ngoc Nhung**, Oxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam; **Stefan Niemann**, Forschungszentrum Borstel, Germany; **Kayzad Nilgiriwala**, Foundation for Medical Research, Mumbai, India; **Camus Nimmo**, University College London, United Kingdom; **Nana Okozi**, National Institute for Communicable Diseases, Pretoria, South Africa; **Shaheed Vally Omar**, National Institute for Communicable Diseases, Pretoria, South Africa; **Rick Ong Twee Hee**, National University of Singapore, Singapore; **Moises Palaci**, Federal University of Espírito Santo, Vitória, Brazil; **Nicholas Paton**, National University of Singapore, Singapore; **Tim Peto**, Oxford University, United Kingdom; **Sara Plesnik**, Institute of Microbiology and Laboratory Medicine, Gauting, Germany; **Jamie Posey**, Centers for Disease Control and Prevention, Atlanta (GA), USA; **Marie-Sylvianne Rabodoarivelo**, Institut Pasteur de Madagascar, Antananarivo, Madagascar; **Niaina Rakotosamimanana**, Institut Pasteur de Madagascar, Antananarivo, Madagascar; **Priti Rathod**, Public Health England, Birmingham, United Kingdom; **Elis Regina Dalla Costa**, Centre for Scientific and Technolical Development, Porto Alegre, Brazil; **Camilla Rodrigues**, P.D. Hinduja National Hospital and Medical Research Centre, Mumbai, India; **Tom Rogers**, Trinity College Dublin, Ireland; **Alex Rosenthal**, National Institutes of Health, Bethesda (MD), USA; **Emma Roycroft**, St James' Hospital, Dublin, Ireland; **Richard Salvato**, Centre for Scientific and Technolical

Development, Porto Alegre, Brazil; **Marco Schito**, Critical Path Institute, Tuscon (AZ), USA; **Sanchi Shah**, Foundation for Medical Research, Mumbai, India; **Joseph Shea**, New York State Department of Health, Albany (NY), USA; **Soyoun Shin**, Korean Institute of Tuberculosis, Seoul, Republic of Korea; **Elisangela Silva**, Federal University of Rio de Janeiro, Brazil; **Vitali Sintchenko**, University of Sydney, Australia; **Rosangela Siqueira de Oliveira**, Institute Adolfo Lutz, São Paulo, Brazil; **Grace Smith**, Public Health England, Birmingham, United Kingdom; **Walter Solano**, London School of Hygiene & Tropical Medicine and Cayetano Heredia Peruvian University, Lima, Peru; **Dick van Soolingen**, National Institute for Public Health and the Environment, Bilthoven, Netherlands; **Mel Spigelman**,TB Alliance, New York City (NY), USA; **Silvana Spindola de Miranda**, Federal University of Minas Gerais, Belo Horizonte, Brazil; **Andrea Spitaleri**, San Raffaele Scientific Institute, Milan, Italy; **Prapaporn Srilohasin**, Mahidol University, Bangkok, Thailand; **Vijay Srinivasan**, Oxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam; **Angela Starks**, Centers for Disease Control and Prevention, Atlanta (GA), USA; **Utkarsha Surve**, P.D. Hinduja National Hospital and Medical Research Centre, Mumbai, India; **Sabira Tahseen**, National Tuberculosis Control Programme, Islamabad, Pakistan; **Jill Taylor**, Association of Public Health Laboratories, Silver Spring (MD), USA; **Grant Theron**, Stellenbosch University, Cape Town, South Africa; **Nguyen Thuy Thuong Thuong**, Oxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam; **Guy Thwaites**, Oxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam; **Katharina Todt**, Institute of Microbiology and Laboratory Medicine, Gauting, Germany; **Alberto Trovato**, San Raffaele Scientific Institute, Milan, Italy; **Annelies Van Rie**, University of Antwerp, Belgium; **Mirela Verza**, Federal University of Santa Catarina, Florianópolis, Brazil; **Timothy Walker**, Oxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam; **Sarah Walker**, Oxford University, United Kingdom; **Rob Warren**, University of Stellenbosch, Cape Town, South Africa; **Jim Werngren**, Public Health Agency, Stockholm, Sweden; **Robert Wilkinson**, University of Cape Town, Imperial College London and Francis Crick Institute, South Africa; **Daniel Wilson,** Oxford University, United Kingdom; **Penelope Wintringer,** European Bioinformatics Institute, London, United Kingdom; **Yu-Xin Xiao**, CDC Taiwan, Taipei, Taiwan (China); **Yang Yang**, Oxford University, United Kingdom; **Zhao Yanlin**, National Tuberculosis Reference Laboratory, China CDC, China; **Shen Yuan Yao**, National Institute for Communicable Diseases, Pretoria, South Africa; **Baoli Zhu**, Institute of Microbiology, CAS, Beijing, China.

References

1. Global tuberculosis report 2020. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf, accessed 10 January 2021).
2. Mohamed S, Köser CU, Salfinger M, Sougakoff W, Heysell SK. Targeted next-generation sequencing: a Swiss army knife for mycobacterial diagnostics? Eur Respir J. 2021;57:2004077.
3. Global tuberculosis report 2013. Geneva: World Health Organization; 2013 (https://apps.who.int/iris/bitstream/handle/10665/91355/9789241564656_eng.pdf, accessed 12 April 2021).
4. Technical report on critical concentrations for drug susceptibility testing of isoniazid and the rifamycins (rifampicin, rifabutin and rifapentine). Geneva: World Health Organization; 2021 (https://apps.who.int/iris/rest/bitstreams/1330649/retrieve, accessed 5 February 2021).
5. WHO operational handbook on tuberculosis. Module 3: Diagnosis – rapid diagnostics for tuberculosis detection. Geneva: World Health Organization; 2020 https://apps.who.int/iris/rest/bitstreams/1284635/retrieve (accessed 30 June 2020).
6. Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C et al. A standardised method for interpreting the association between mutations and phenotypic drug resistance in *Mycobacterium tuberculosis*. Eur Respir J. 2017;50:1701354.
7. Kadura S, King N, Nakhoul M, Zhu H, Theron G, Köser CU et al. Systematic review of mutations associated with resistance to the new and repurposed *Mycobacterium tuberculosis* drugs bedaquiline, clofazimine, linezolid, delamanid and pretomanid. J Antimicrob Chemother. 2020;75:2031–43.
8. Merker M, Kohl TA, Barilar I, Andres S, Fowler PW, Chryssanthou E et al. Phylogenetically informative mutations in genes implicated in antibiotic resistance in *Mycobacterium tuberculosis* complex. Genome Med. 2020;12:27.
9. Schön T, Miotto P, Köser CU, Viveiros M, Böttger E, Cambau E. *Mycobacterium tuberculosis* drug-resistance testing: challenges, recent developments and perspectives. Clin Microbiol Infect. 2017;23:154–60.
10. Nebenzahl-Guimaraes H, Jacobson KR, Farhat MR, Murray MB. Systematic review of allelic exchange experiments aimed at identifying mutations that confer drug resistance in *Mycobacterium tuberculosis*. J Antimicrob Chemother. 2014;69:331–42.
11. Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P et al. Modernizing Medical Microbiology Informatics G. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. Lancet Infect Dis. 2015;15:1193–202.
12. The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex: technical guide (WHO/CDS/TB/2018.19). Geneva: World Health Organization; 2018 (https://apps.who.int/iris/bitstream/handle/10665/274443/WHO-CDS-TB-2018.19-eng.pdf, accessed 20 September 2018).
13. Technical report on critical concentrations for drug susceptibility testing of medicines used in the treatment of drug-resistant tuberculosis (WHO/CDS/TB/2018.5). Geneva: World Health Organization; 2018 (http://apps.who.int/iris/bitstream/10665/260470/1/WHO-CDS-TB-2018.5-eng.pdf, accessed 10 January 2021).

14. Meeting report of the WHO expert consultation on the definition of extensively drug-resistant tuberculosis, 27–29 October 2020. Geneva: World Health Organization; 2021 (https://apps.who.int/iris/rest/bitstreams/1327986/retrieve, accessed 26 January 2021).

15. Ajileye A, Alvarez N, Merker M, Walker TM, Akter S, Brown K et al. Some synonymous and nonsynonymous *gyrA* mutations in *Mycobacterium tuberculosis* lead to systematic false-positive fluoroquinolone resistance results with the Hain GenoType MTBDR*sl* assays. Antimicrob Agents Chemother. 2017;61:e02169-16.

16. Line probe assays for drug-resistant tuberculosis detection. Interpretation and reporting guide for laboratory staff and clinicians. Geneva: Global Laboratory Initiative, World Health organization; 2018 (http://www.stoptb.org/wg/gli/assets/documents/LPA_test_web_ready.pdf (accessed 2 November 2018).

17. Interpretation guide for GenoType MTBDR*plus* VER 2.0 and GenoType MTBDR*sl* VER 2.0. A technical guidance document developed by the European Laboratory Initiative. Version 1.0. Copenhagen: World Health Organization Regional Office for Europe; 2020 (https://openwho.org/courses/multi-drug-resistant-tb, accessed 9 May 2020).

18. André E, Goeminne L, Cabibbe A, Beckert P, Kabamba Mukadi B, Mathys V et al. Consensus numbering system for the rifampicin resistance-associated *rpoB* gene mutations in pathogenic mycobacteria. Clin Microbiol Infect. 2017;23:167–72.

19. Technical manual for drug susceptibility testing of medicines used in the treatment of tuberculosis (WHO/CDS/TB/2018.24). Geneva: World Health Organization; 2018 (http://apps.who.int/iris/bitstream/handle/10665/275469/9789241514842-eng.pdf, accessed 17 November 2018).

20. Köser CU, Georghiou SB, Schön T, Salfinger M. On the consequences of poorly defined breakpoints for rifampin susceptibility testing of *Mycobacterium tuberculosis* complex. J Clin Microbiol. 2021;59:e02328-20.

21. Ando H, Miyoshi-Akiyama T, Watanabe S, Kirikae T. A silent mutation in *mabA* confers isoniazid resistance on *Mycobacterium tuberculosis*. Mol Microbiol. 2014;91:538–47.

22. Ando H, Kondo Y, Suetake T, Toyota E, Kato S, Mori T et al. Identification of *katG* mutations associated with high-level isoniazid resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2010;54:1793–9.

23. Ando H, Kitao T, Miyoshi-Akiyama T, Kato S, Mori T, Kirikae T. Downregulation of *katG* expression is associated with isoniazid resistance in *Mycobacterium tuberculosis*. Mol Microbiol. 2011;79:1615–28.

24. Safi H, Lingaraju S, Amin A, Kim S, Jones M, Holmes M et al. Evolution of high-level ethambutol-resistant tuberculosis through interacting mutations in decaprenylphosphoryl-β-D-arabinose biosynthetic and utilization pathway genes. Nat Genet. 2013;45:1190–7.

25. Gygli SM, Keller PM, Ballif M, Blochliger N, Homke R, Reinhard M et al. Whole-genome sequencing for drug resistance profile prediction in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2019;63:e02175-18.

26. Nonghanphithak D, Kaewprasert O, Chaiyachat P, Reechaipichitkul W, Chaiprasert A, Faksri K. Whole-genome sequence analysis and comparisons between drug-resistance mutations and minimum inhibitory concentrations of *Mycobacterium tuberculosis* isolates causing M/XDR-TB. PLoS One. 2020;15:e0244829.

27. Lamont EA, Dillon NA, Baughn AD. The bewildering antitubercular action of pyrazinamide. Microbiol Mol Biol Rev. 2020;84:e00070-19.

28. Köser CU, Cirillo DM, Miotto P. How to optimally combine genotypic and phenotypic drug susceptibility testing methods for pyrazinamide. Antimicrob Agents Chemother. 2020;64:e01003-20.

29. Willby MJ, Wijkander M, Havumaki J, Johnson K, Werngren J, Hoffner S et al. Detection of *Mycobacterium tuberculosis pncA* mutations by the Nipro Genoscholar PZA-TB II assay compared to conventional sequencing. Antimicrob Agents Chemother. 2018;62:e01871-17.

30. Whitfield MG, Warren RM, Streicher EM, Sampson SL, Sirgel FA, van Helden PD et al. *Mycobacterium tuberculosis pncA* polymorphisms that do not confer pyrazinamide resistance at a breakpoint concentration of 100 micrograms per milliliter in MGIT. J Clin Microbiol. 2015;53:3633–5.

31. Modlin SJ, Marbach T, Werngren J, Mansjo M, Hoffner SE, Valafar F. Atypical genetic basis of pyrazinamide resistance in mono-resistant *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2021;65:e01916-20.

32. Tagliani E, Hassan MO, Waberi Y, De Filippo MR, Falzon D, Dean A et al. Culture and next-generation sequencing-based drug susceptibility testing unveil high levels of drug-resistant-TB in Djibouti: results from the first national survey. Sci Rep. 2017;7:17672.

33. Briquet A, Vong R, Roseau JB, Javelle E, Cazes N, Rivière F et al. Clinical features of *Mycobacterium canettii* infection: a retrospective study of 20 cases among French soldiers and relatives. Clin Infect Dis. 2019;69:2003–10.

34. Feuerriegel S, Köser CU, Richter E, Niemann S. *Mycobacterium canettii* is intrinsically resistant to both pyrazinamide and pyrazinoic acid. J Antimicrob Chemother. 2013;68:1439–40.

35. Loiseau C, Brites D, Moser I, Coll F, Pourcel C, Robbe-Austerman S et al. Revised interpretation of the Hain Lifescience GenoType MTBC to differentiate *Mycobacterium canettii* and members of the *Mycobacterium tuberculosis* complex. Antimicrob Agents Chemother. 2019;63:e00159-19.

36. Zignol M, Cabibbe AM, Dean AS, Glaziou P, Alikhanova N, Ama C et al. Genetic sequencing for surveillance of drug resistance in tuberculosis in highly endemic countries: a multi-country population-based surveillance study. Lancet Infect Dis. 2018;18:675–83.

37. CRyPTIC Consortium and the 100000 Genomes Project, Allix-Beguec C, Arandjelovic I, Bi L, Beckert P, Bonnet M, Bradley P et al. Prediction of susceptibility to first-line tuberculosis drugs by DNA sequencing. N Engl J Med. 2018;379:1403–15.

38. Ektefaie Y, Dixit A, Freschi L, Farhat MR. Globally diverse *Mycobacterium tuberculosis* resistance acquisition: a retrospective geographical and temporal analysis of whole genome sequences. Lancet Microbiol. 2021;2:e96-e104.

39. Matrat S, Veziris N, Mayer C, Jarlier V, Truffot-Pernot C, Camuset J et al. Functional analysis of DNA gyrase mutant enzymes carrying mutations at position 88 in the A subunit found in clinical strains of *Mycobacterium tuberculosis* resistant to fluoroquinolones. Antimicrob Agents Chemother. 2006;50:4170–3.

40. Maruri F, Sterling TR, Kaiga AW, Blackman A, van der Heijden YF, Mayer C et al. A systematic review of gyrase mutations associated with fluoroquinolone-resistant *Mycobacterium tuberculosis* and a proposed gyrase numbering system. J Antimicrob Chemother. 2012;67:819–31.

41. Disratthakit A, Prammananan T, Tribuddharat C, Thaipisuttikul I, Doi N, Leechawengwongs M et al. Role of *gyrB* mutations in pre-extensively and extensively drug-resistant tuberculosis in Thai clinical isolates. Antimicrob Agents Chemother. 2016;60:5189–97.

42. Köser CU, Ellington MJ, Peacock SJ. Whole-genome sequencing to control antimicrobial resistance. Trends Genet. 2014;30:401–7.

43. CRyPTIC Consortium, Fowler PW. Epidemiological cutoff values for a 96-well broth microdilution plate for high-throughput research antibiotic susceptibility testing of M. tuberculosis. medRxiv 2021 (https://doi.org/10.1101/2021.02.24.21252386, accessed 9 March 2021).

44. Nimmo C, Brien K, Millard J, Grant AD, Padayatchi N, Pym AS et al. Dynamics of within-host *Mycobacterium tuberculosis* diversity and heteroresistance during treatment. EBioMedicine. 2020;55:102747.

45. Mokrousov I, Akhmedova G, Molchanov V, Fundovnaya E, Kozlova E, Ostankova Y et al. Frequent acquisition of bedaquiline resistance by epidemic extensively drug-resistant *Mycobacterium tuberculosis* strains in Russia during long-term treatment. Clin Microbiol Infect. 2021;27:478–80.

46. Beckert P, Sanchez-Padilla E, Merker M, Dreyer V, Kohl TA, Utpatel C et al. MDR *M. tuberculosis* outbreak clone in Eswatini missed by Xpert has elevated bedaquiline resistance dated to the pre-treatment era. Genome Med. 2020;12:104.

47. Nimmo C, Millard J, Brien K, Moodley S, van Dorp L, Lutchminarain K et al. Bedaquiline resistance in drug-resistant tuberculosis HIV co-infected patients. Eur Respir J. 2020;55:1902383.

48. Lee BM, Harold LK, Almeida DV, Afriat-Jurnou L, Aung HL, Forde BM et al. Predicting nitroimidazole antibiotic resistance mutations in *Mycobacterium tuberculosis* with protein engineering. PLoS Pathog. 2020;16:e1008287.

49. Rifat D, Li SY, Ioerger T, Shah K, Lanoix JP, Lee J et al. Mutations in *fbiD* (*Rv2983*) as a novel determinant of resistance to pretomanid and delamanid in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2020;65:e01948-20.

50. Battaglia S, Spitaleri A, Cabibbe AM, Meehan CJ, Utpatel C, Ismail N et al. Characterization of genomic variants associated with resistance to bedaquiline and delamanid in naive *Mycobacterium tuberculosis* clinical strains. J Clin Microbiol. 2020;58:e01304-20.

51. Cao Y, Parmar H, Gaur RL, Lieu D, Raghunath S, Via N et al. Xpert MTB/XDR: A ten-color reflex assay suitable for point of care settings to detect isoniazid-, fluoroquinolone-, and second-line injectable drug-resistance directly from *Mycobacterium tuberculosis* positive sputum. J Clin Microbiol. 2021; 59: e02314-20.

52. GenoType MTBDRsl VER 2.0. Instructions for use. IFU-317A-04. Tübingen: Hain Lifescience; 2017 (https://www.hain-lifescience.de/include_datei/kundenmodule/packungsbeilage/download.php?id=2020, accessed 28 February 2021).

53. Katsukawa C, Tamaru A, Miyata Y, Abe C, Makino M, Suzuki Y. Characterization of the *rpsL* and *rrs* genes of streptomycin-resistant clinical isolates of *Mycobacterium tuberculosis* in Japan. J Appl Microbiol. 1997;83:634–40.

54. Okamoto S, Tamaru A, Nakajima C, Nishimura K, Tanaka Y, Tokuyama S et al. Loss of a conserved 7-methylguanosine modification in 16S rRNA confers low-level streptomycin resistance in bacteria. Mol Microbiol. 2007;63:1096–106.

55. Feuerriegel S, Oberhauser B, George AG, Dafae F, Richter E, Rüsch-Gerdes S et al. Sequence analysis for detection of first-line drug resistance in *Mycobacterium tuberculosis* strains from a high-incidence setting. BMC Microbiol. 2012;12:90.

56. Cambau E, Viveiros M, Machado D, Raskine L, Ritter C, Tortoli E et al. Revisiting susceptibility testing in MDR-TB by a standardized quantitative phenotypic assessment in a European multicentre study. J Antimicrob Chemother. 2015;70:686–96.

57. Malinga L, Brand J, Olorunju S, Stoltz A, van der Walt M. Molecular analysis of genetic mutations among cross-resistant second-line injectable drugs reveals a new resistant mutation in *Mycobacterium tuberculosis*. Diagn Microbiol Infect Dis. 2016;85:433–7.

58. Daum LT, Rodriguez JD, Worthy SA, Ismail NA, Omar SV, Dreyer AW et al. Next-generation Ion Torrent sequencing of drug resistance mutations in *Mycobacterium tuberculosis* strains. J Clin Microbiol. 2012;50:3831–7.

59. Ahmad S, Mokaddas E, Al-Mutairi N, Eldeen HS, Mohammadi S. Discordance across phenotypic and molecular methods for drug susceptibility testing of drug-resistant *Mycobacterium tuberculosis* isolates in a low TB incidence country. PLoS One. 2016;11:e0153563.

60. Umar FF, Husain DR, Hatta MM, Natzir RR, Sjahril RS, Dwiyanti RR et al. Molecular characterisation of mutations associated with resistance to first- and second-line drugs among Indonesian patients with tuberculosis. J Taibah Univ Med Sci. 2020;15:54–8.

61. Yeh KC, To KY, Sun SW, Wu MC, Lin TY, Chen CC. Point mutations in the chloroplast 16s rRNA gene confer streptomycin resistance in *Nicotiana plumbaginifolia*. Curr Genet. 1994;26:132–5.

62. Gregory ST, Dahlberg AE. Genetic and structural analysis of base substitutions in the central pseudoknot of *Thermus thermophilus* 16S ribosomal RNA. RNA. 2009;15:215–23.

63. Demirci H, Murphy FV, Murphy EL, Connetti JL, Dahlberg AE, Jogl G et al. Structural analysis of base substitutions in *Thermus thermophilus* 16S rRNA conferring streptomycin resistance. Antimicrob Agents Chemother. 2014;58:4308–17.

64. Wong SY, Lee JS, Kwak HK, Via LE, Boshoff HI, Barry CE. Mutations in *gidB* confer low-level streptomycin resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2011;55:2515–22.

65. Cohen KA, Stott KE, Munsamy V, Manson AL, Earl AM, Pym AS. Evidence for expanding the role of streptomycin in the management of drug-resistant *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2020;64:e00860-20.

66. Victor T, van Rie A, Jordaan A, Richardson M, van Der Spuy G, Beyers N et al. Sequence polymorphism in the rrs gene of *Mycobacterium tuberculosis* is deeply rooted within an evolutionary clade and is not associated with streptomycin resistance. J Clin Microbiol. 2001;39:4184–6.

67. Vilchèze C, Jacobs WR Jr. Resistance to isoniazid and ethionamide in *Mycobacterium tuberculosis*: Genes, mutations, and causalities. Microbiol Spectr. 2014;2:MGM2–0014–2013.

68. Grant SS, Wellington S, Kawate T, Desjardins CA, Silvis MR, Wivagg C et al. Baeyer-Villiger monooxygenases EthA and MymA are required for activation of replicating and non-replicating *Mycobacterium tuberculosis* inhibitors. Cell Chem Biol. 2016;23:666–77.

69. Xpert® MTB/XDR. GXMTB/XDR-10. 302–3514, Rev. B. Sunnyvale (CA): Cepheid; 2020.

70. Lefford MJ, Mitchison DA. Comparison of methods for testing the sensitivity of *Mycobacterium tuberculosis* to ethionamide. Tubercle. 1966;47:250–61.

71. Vilchèze C, Wang F, Arai M, Hazbon MH, Colangeli R, Kremer L et al. Transfer of a point mutation in *Mycobacterium tuberculosis* inhA resolves the target of isoniazid. Nat Med. 2006;12:1027–9.

72. Rueda J, Realpe T, Mejia GI, Zapata E, Rozo JC, Ferro BE et al. Genotypic analysis of genes associated with independent resistance and cross-resistance to ioniazid and ethionamide in *Mycobacterium tuberculosis* clinical isolates. Antimicrob Agents Chemother. 2015;59:7805–10.

73. Ghodousi A, Tagliani E, Karunaratne E, Niemann S, Perera J, Köser CU et al. Isoniazid resistance in *Mycobacterium tuberculosis* is a heterogeneous phenotype composed of overlapping MIC distributions with different underlying resistance mechanisms. Antimicrob Agents Chemother. 2019;63:e00092-19.

74. Nakatani Y, Opel-Reading HK, Merker M, Machado D, Andres S, Kumar SS et al. Role of alanine racemase mutations in *Mycobacterium tuberculosis* D-cycloserine resistance. Antimicrob Agents Chemother. 2017;61:e01575-17.

75. Rancoita PMV, Cugnata F, Gibertoni Cruz AL, Borroni E, Hoosdally SJ, Walker TM et al. Validating a 14-drug microtitre plate containing bedaquiline and delamanid for large-scale research susceptibility testing of *Mycobacterium tuberculosis*. Antimicrob Agents Chemother. 2018;62:e00344-18.

76. Seq&Treat. Geneva: Foundation for Innovative New Diagnostics; 2021 (https://www.finddx.org/tb/seq-treat, accessed 13 April 2021).

77. Noncommercial culture and drug-susceptibility testing methods for screening patients at risk for multidrug-resistant tuberculosis. Policy statement. Geneva: World Health Organization; 2011 (https://apps.who.int/iris/bitstream/handle/10665/44601/9789241501620_eng.pdf, accessed 20 March 2021).

78. Martin L, Coronel J, Faulx D, Valdez M, Metzler M, Crudder C et al. A field evaluation of the Hardy TB MODS Kit™ for the rapid phenotypic diagnosis of tuberculosis and multi-drug resistant tuberculosis. PLoS One. 2014;9:e107258.

79. Guidelines for drug susceptibility testing for second-line anti-tuberculosis drugs for dots-plus (WHO/CDS/TB/2001.288). Geneva: World Health Organization; 2001 (https://apps.who.int/iris/bitstream/handle/10665/83807/WHO_CDS_TB_2001.288_eng.pdf, accessed 15.2.2020).

80. Policy guidance on drug-susceptibility testing (DST) of second-line antituberculosis drugs (WHO/HTM/TB/2008.392). Geneva: World Health Organization; 2008 (http://apps.who.int/iris/bitstream/handle/10665/70500/WHO_HTM_TB_2008.392_eng.pdf, accessed 10 January 2021).

81. Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis (WHO/HTM/TB/2014.11). Geneva: World Health Organization; 2014 (http://apps.who.int/iris/bitstream/handle/10665/130918/9789241548809_eng.pdf, accessed 12 January 2020).

82. Performance standards for susceptibility testing of mycobacteria, *Nocardia* spp., and other aerobic actinomycetes (CLSI supplement M62). Wayne (PA): Clinical and Laboratory Standards Institute; 2018.

83. Torrea G, Ng KCS, Van Deun A, André E, Kaisergruber J, Ssengooba W et al. Variable ability of rapid tests to detect M*ycobacterium tuberculosis* rpoB mutations conferring phenotypically occult rifampicin resistance. Sci Rep. 2019;9:11826.

84. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv. 2013:1303.3997.

85. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

86. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.

87. Iqbal Z, Turner I, McVean G. High-throughput microbial population genomics using the Cortex variation assembler. Bioinformatics. 2013;29:275–6.

88. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34:3094–100.

89. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: A fast and versatile genome alignment system. PLoS Comput Biol. 2018;14:e1005944.

90. Dorfman R. The detection of defective members of large populations. Ann Math Stat. 1943;14:436–40.

91. Aldridge M, Baldassini L, Johnson O. Group testing algorithms: bounds and simulations. IEEE Trans Inf Theory. 2014;60:3671–87.

92. Zhang Y, Dhandayuthapani S, Deretic V. Molecular basis for the exquisite sensitivity of *Mycobacterium tuberculosis* to isoniazid. Proc Natl Acad Sci U S A. 1996;93:13212–6.

93. Vilchèze C, Weisbrod T, Chen B, Kremer L, Hazbón M, Wang F et al. Altered NADH/NAD+ ratio mediates coresistance to isoniazid and ethionamide in mycobacteria. Antimicrob Agents Chemother. 2005;49:708–20.

94. Vilchèze C, Av-Gay Y, Barnes SW, Larsen MH, Walker JR, Glynne RJ et al. Coresistance to isoniazid and ethionamide maps to mycothiol biosynthetic genes in *Mycobacterium bovis*. Antimicrob Agents Chemother. 2011;55:4422–3.

95. Farhat MR, Freschi L, Calderon R, Ioerger T, Snyder M, Meehan CJ et al. GWAS for quantitative resistance phenotypes in *Mycobacterium tuberculosis* reveals resistance genes and regulatory regions. Nat Commun. 2019;10:2128.

96. Liu J, Shi W, Zhang S, Hao X, Maslov DA, Shur KV et al. Mutations in efflux pump Rv1258c (Tap) cause resistance to pyrazinamide, isoniazid, and streptomycin in *M. tuberculosis*. Front Microbiol. 2019;10:216.

97. Srivastava S, Ayyagari A, Dhole T, Nyati K, Dwivedi S. *emb* nucleotide polymorphisms and the role of embB306 mutations in *Mycobacterium tuberculosis* resistance to ethambutol. Int J Med Microbiol. 2009;299:269–80.

98. Plinke C, Cox HS, Zarkua N, Karimovich HA, Braker K, Diel R et al. *embCAB* sequence variation among ethambutol-resistant *Mycobacterium tuberculosis* isolates without *embB*306 mutation. J Antimicrob Chemother. 2010;65:1359–67.

99. Scorpio A, Zhang Y. Mutations in pncA, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. Nat Med. 1996;2:662–7.

100. Gopal P, Sarathy JP, Yee M, Ragunathan P, Shin J, Bhushan S et al. Pyrazinamide triggers degradation of its target aspartate decarboxylase. Nat Commun. 2020;11:1661.

101. Reeves AZ, Campbell PJ, Sultana R, Malik S, Murray M, Plikaytis BB et al. Aminoglycoside cross-resistance in *Mycobacterium tuberculosis* due to mutations in the 5´ untranslated region of *whiB7*. Antimicrob Agents Chemother. 2013;57:1857–65.

102. Dover LG, Alahari A, Gratraud P, Gomes JM, Bhowruth V, Reynolds RC et al. EthA, a common activator of thiocarbamide-containing drugs acting on different mycobacterial targets. Antimicrob Agents Chemother. 2007;51:1055–63.

103. Walker TM, Merker M, Knoblauch AM, Helbling P, Schoch OD, van der Werf MJ et al. A cluster of multidrug-resistant *Mycobacterium tuberculosis* among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. Lancet Infect Dis. 2018;18:431–40.

104. Zhang Y, Heym B, Allen B, Young D, Cole S. The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. Nature. 1992;358:591–3.

105. Wong SY, Javid B, Addepalli B, Piszczek G, Strader MB, Limbach PA et al. Functional role of methylation of G518 of the 16S rRNA 530 loop by GidB in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother, 2013;57:6311–8.

World Health Organization

For further information, please contact:
World Health Organization
20, Avenue Appia CH-1211 Geneva 27 Switzeland
Global TB Programme
Web site: www.who.int/tb

9789240028173