Tuberculosis resistance acquisition in space and time: an analysis of globally diverse *M. tuberculosis* whole genome sequences

Yasha Ektefaie[1], Avika Dixit[2,3], Luca Freschi[2], Maha Farhat[2,4]

[1]University of California Berkeley, Berkeley CA, [2]Harvard Medical School, Boston MA, [3]Boston Children's Hospital, Boston MA, [4]Massachusetts General Hospital, Boston, MA

Yasha Ektefaie
Affiliations: University of California Berkeley, Berkeley CA
Address: 120 Sproul Hall, Berkeley CA 94720, USA

Avika Dixit, MBBS MPH MBI
Affiliations: Boston Children's Hospital, Boston MA; Harvard Medical School, Boston MA
Address: 300 Longwood Ave, Mailstop 3103, Boston MA 02115, USA

Luca Freschi, PhD MSc
Affiliations: Harvard Medical School, Boston BA
Address: 10 Shattuck St, Boston MA 02115, USA

Maha R. Farhat, MD MSc
Affiliations: Harvard Medical School, Boston MA; Massachusetts General Hospital, Boston MA
Address: 10 Shattuck St, Boston MA 02115, USA

Corresponding Author:

Maha R. Farhat
10 Shattuck St, Boston MA 02130, USA
Email: maha_farhat@hms.harvard.edu

Research in context

Evidence before this study

Acquisition and spread of drug-resistance by *Mycobacterium tuberculosis* (MTB) varies across countries. Local factors driving evolution of drug resistance in MTB are not well studied.

Added value of this study

We applied molecular dating to 6,099 global MTB patient isolates and found the order of resistance acquisition to be consistent across the countries examined, *i.e.* acquisition of isoniazid resistance first followed by rifampicin and streptomycin followed by resistance to other drugs. In all countries with data available there was evidence for transmission of resistant strains from patient-to-patient and in the majority for extended periods of time (>20 years). Countries with lower gross wealth indices were found to have more recent resistance acquisition to the drug rifampicin. Based on the resistance patterns identified in our study we estimate that commercial diagnostic tests vary considerably in sensitivity for second-line resistance diagnosis by country.

Implications of all available evidence

The longevity of resistant MTB in many parts of the world emphasizes its fitness for transmission and its continued threat to public health. The association between country wealth and recent resistance acquisition emphasizes the need for continued investment in TB care delivery and surveillance programs. Geographically relevant diagnostics that take into account a country's unique distribution of resistance are necessary.

Abstract

Background: *Mycobacterium tuberculosis* (MTB) whole genome sequencing data can provide insights into temporal and geographic trends in resistance acquisition and inform public health interventions.

Methods: We curated a set of clinical MTB isolates with high quality sequencing and culture-based drug susceptibility data spanning four lineages and more than 20 countries. We constructed geographic and lineage specific MTB phylogenies and used Bayesian molecular dating to infer the most-recent-common-susceptible-ancestor age for 4,869 instances of resistance to 10 drugs.

Findings: Of 8,550 isolates curated, 6,099 from 15 countries met criteria for molecular dating. The number of independent resistance acquisition events was lower than the number of resistant isolates across all countries, suggesting ongoing transmission of drug resistance. Ancestral age distributions supported the presence of old resistance, ≥20 years prior, in the majority of countries. A consistent order of resistance acquisition was observed globally starting with resistance to isoniazid, but resistance ancestral age varied by country. We found a direct correlation between country wealth and resistance age (R^2= 0.47, P-value= 0.014). Amplification of fluoroquinolone and second-line injectable resistance among multidrug-resistant isolates is estimated to have occurred very recently (median ancestral age 4.7 years IQR 1.9-9.8 prior to sample collection). We found the sensitivity of commercial molecular diagnostics for second-line resistance to vary significantly by country (P-value <0.0003)

Interpretation: Our results highlight that both resistance transmission and amplification are contributing to disease burden globally but are variable by country. The observation that wealthier nations are more likely to have old resistance suggests that programmatic improvements can reduce resistance amplification, but that fit resistant strains can circulate for decades subsequently.

Funding: This work was supported by the NIH BD2K grant K01 ES026835, a Harvard Institute of Global Health Burke Fellowship (MF), Boston Children's Hospital OFD/BTREC/CTREC Faculty Career Development Fellowship and Bushrod H. Campbell and Adah F. Hall Charity Fund/Charles A. King Trust Postdoctoral Fellowship (AD).

Keywords: tuberculosis, drug resistance, whole genome sequencing

Introduction

Tuberculosis (TB) defines a global epidemic that takes more lives than any other infection due to a single pathogen[1]. The emergence of multidrug-resistant (MDR) and extensively drug-resistant (XDR) resistant TB presents a major hurdle to efforts in accelerating TB decline. Halting the transmission of drug-resistant (DR) TB has been a major focus of studies addressing this hurdle[2]. But the epidemic is ultimately defined by local factors that remain understudied in many parts of the world[3]. The study of geographic and temporal heterogeneity of the DR-TB epidemic can provide insights into these local factors as key drivers of MDR-TB prevalence and persistence in the community, including programmatic and bacterial factors. This understanding is key to future disease control and prevention of antibiotic resistance development.

Over the past decade, increased uptake of molecular and whole genome sequencing (WGS) technologies, and their application to *Mycobacterium tuberculosis* (MTB) clinical isolates has offered novel insights into pathogen biology and diversity in the context of human infection[4–7]. The application of WGS has allowed us to better understand the genetic determinants of drug resistance (DR) within MTB[8]. The detection of these genetic determinants using molecular technologies that include WGS is now increasingly adopted for TB resistance diagnosis in many parts of the world[9] and is beginning to replace the more biohazardous and time consuming culture based drug susceptibility tests (DST). The study of isolates sampled from epidemiological outbreaks or from the same host over time has allowed the estimation of MTB's molecular clock rate, or temporal rate of accumulation of fixed genome-level variation[10,11]. The application of this rate to new WGS data from isolates collected for surveillance has helped improve transmission inference and molecular dating of specific evolutionary events such as resistance acquisition or lineage divergence[10,12-13].

We sought to use a large clinical collection of MTB WGS and resistance phenotype data to study how, when, and where resistance was acquired on a global scale. Using a Bayesian implementation of coalescent theory, we estimate and compare dates of resistance acquisition for MDR/XDR isolates across 15 different countries. We use the recency of resistance acquisition as a measure of fitness of the circulating strains in their respective environments and study the effect of country wealth, as a proxy for TB control programme funding, on the recency of resistance acquisition at a macro level. We also assess the distribution of unexplained MTB phenotypic resistance across 20 countries, to evaluate the accuracy and geographic heterogeneity of molecular detection of common MTB genetic resistance determinants, and discuss implications for DR-TB control.

Methods

Further details available in the supplementary material.

Data and quality control

We compiled a 10,299 MTB WGS dataset with culture based DST (phenotypic) data using public databases (Patric[14], ReSeqTB[15]) and literature curation[11–13,16–26] . A summary table with the phenotypic data is available online at https://github.com/farhat-lab/resdata.

Genomic analysis/variant calling

We used a previously validated genomic analysis pipeline for MTB described by Ezewudo *et al.* [27] with modifications as detailed in the supplement.

Drug resistance definitions

Drugs were labelled as follows: isoniazid (INH), ethambutol (EMB), rifamycins (rifampicin or rifabutin) (RIF), streptomycin (STR), pyrazinamide (PZA), fluoroquinolones (FLQ) (includes moxifloxacin, ciprofloxacin, ofloxacin), second-line injectables (SLIs) (includes kanamycin, amikacin, capreomycin), ethionamide/prothionamide (ETH), and cycloserine (CYS). Para-aminosalicylic acid was not analysed due to the paucity of data. Isolates not tested for susceptibility to both INH and RIF were excluded from the assessment of DR frequency by country and lineage. Isolates resistant to both INH and RIF were labelled MDR. Those resistant to INH, RIF, FLQ and SLIs were labelled XDR.

Estimating resistance acquisition dates

Isolates were separated into 179 groups corresponding to a single drug, lineage and source country, referred to hereafter as a 'group'. Genetic diversity was computed as the average pairwise genetic distance within a group. To accurately date resistance acquisition, a drug-geography-lineage group was analysed only if it consisted of at least 10 isolates, ≥20% of isolates were susceptible and ≥1 isolate was resistant. To exclude isolates that only represent outbreak settings and didn't carry more long-term information about resistance, we excluded groups with a genetic diversity score <1 standard deviation from the mean genetic diversity score measured across all groups. Supplementary methods detail the phylogeny construction and the estimation of the age of the most recent *susceptible* common ancestor (MRSCA) in years prior to isolation of the clinical sample(s).

Distribution of Resistance Mutations

We compared the expected sensitivity and specificity of mutations captured by commercial diagnostics (summarized from the literature in Table S9) and those based on more extensive lists of mutations in DR genes that can be captured using targeted or whole genome sequencing. We used three mutations lists for the latter (1) a set of 267 common resistance-associated mutations that we previously determined using randomForests[28] designated "RF-select WGS test", (2) a set of mutations determined using direct association[29] designated "DA-select WGS test", and (3) any non-synonymous mutation or noncoding mutation in known DR regions (Table S10) in a "all WGS test". We excluded previously described neutral/lineage associated mutations[9].

Code

All code used in the analysis is publically accessible at https://github.com/farhat-lab/geo_dist_tb.

Results

Data and global lineage distribution

Of the 10,299 MTB clinical isolates with WGS and culture-based DST data available, 9,385 passed sequence quality criteria and of these 8,550 had country of origin data (Figure 1). The four major MTB lineages, 1-4, were well represented. A relatively high proportion, 42%, of United Kingdom (UK) isolates (n= 1873) belonged to Lineages-1 & 3 (Figure 2A). Overall, the non-Europe-America-Africa Lineages-1,3, and 2, comprised 40% of European isolates (n=3956) and 7% of North and South American isolates (n=1297).

Phenotypic resistance distribution

Of the 8,550 isolates, 568 isolates lacked either INH or RIF DST data. Out of the remaining 7,909 isolates, 5,022 were pan-susceptible, 2887 were resistant to one or more drugs (DR) and of these 1937 were resistant to INH and RIF (MDR) and 288 were MDR and resistant to an SLI and a FLQ (XDR). The 8,550 isolates originated from 52 countries. Of these, 23 countries were represented by >10 isolates with resistance data, 21/23 were found to have MDR isolates and 9/21 had XDR (Figure 2B). We compared the MDR frequency in our WGS based sample with the WHO reported MDR/RIF resistance (RR) rates for the latest year available[30]. Out of the 21 countries, the confidence interval for the MDR-TB proportion in our sample overlapped with that of the WHO in 4 (19%) countries, was higher in 14 (67%) and lower in 3 (14%) of countries (Table S4). MDR rates by lineage were 3% for Lineage-1 (n=439), 48% for Lineage-2 (n=1085), 4% for Lineage-3 (n=760) and 23% for Lineage-4 (n=3358).

Molecular dating of Resistance Acquisition

Of the 8,550 isolates, 2,451 isolates appeared in groups that did not meet our dating requirements (Methods). The remaining 6,009, included 1,547 isolates resistant to one or more drugs and were grouped into 179 country/lineage/drug combinations. We estimated 4,869 MRSCA dates for 10 drugs across these 179 groups. The number of independent resistance acquisition events *i.e.* unique MRSCA dates, was consistently lower than the total number of dated resistance isolates suggesting ongoing transmission of drug resistant isolates (Table S11). We estimated a lower bound on the burden of resistance due to transmission ranging by country from ≥14% to ≥52% pooled across drugs (Methods, Table S11). The proportion of INH or RIF resistance attributed to transmission was the highest among the 10 drugs at ≥43% and ≥46% respectively pooled across countries (pooled from Table S11).

We examined the relative order of phenotypic resistance acquisition on a global scale. For INH, we found that resistance to INH on average developed before resistance to other drugs (Figure 3A-B). Median MRSCA for INH was 11.4 years prior to isolation (IQR 6.3-16.2) *vs.* 7.6 years (IQR 3.0 – 16.0) for RIF, Wilcoxon P-value $<10^{-14}$. Median MRSCA ages for RIF and STR resistance (7.6 years, IQR 3.0 – 16.0 and 7.7 years, IQR 3.4 – 13.0 respectively) were second oldest and not statistically significant from each other (Wilcoxon P-value 0.31). The dating supported that EMB resistance followed the acquisition of RIF (Wilcoxon P-value $< 10^{-6}$) at a median MRSCA age of 5.0 years prior (IQR 2.1 – 12.5), and that this was followed by resistance to PZA, ETH, FLQ, SLIs, or CYS (Figure 3A), amongst which MRSCA ages did not significantly differ (Figure 3B). We found no significant correlation between the median MRSCA dates and

the drug's date of introduction into clinical use with R^2= 0.04 (F-test with 1 DF P-value= 0.60, Table S12).

We assessed the frequency of recent resistance amplification to PZA, EMB, FLQs and SLIs among MDR, *i.e.* to pre-XDR/XDR, within five years of sample isolation. Among the 11 countries with both MDR and pre-XDR/XDR isolates, we identified four countries (Peru, Russia, Sierra Leone, South Africa) with recent resistance amplification to PZA and EMB (>1% of MDR). The rates of recent amplification ranged from 2% (95% CI 1% - 4%) for PZA in Russia to 33% (95% CI 26% - 41%) for EMB in South Africa (Figure 5). Peru, Romania and South Africa were also measured to have recent resistance amplification to FLQs and SLIs (Figure 5). The median MRSCA age for FLQ or SLI resistance acquisition among MDR isolates was 4.7 years (IQR 1.9-9.8) prior to sample collection.

We found RIF to have the highest proportion of old resistance (MRSCA >20 years prior to isolation) at 17%, 197/1184 out of the total dated RIF resistance acquisition events. Old resistance was well distributed geographically and for RIF occurred in 9 of 12 countries with available dating data (Figure 4). Old FLQ resistance constituted 8% (24/311) of total dated isolates and spanned 6 of the 7 countries with available data.

We compared the geographic distribution of MRSCA ages restricting to four key drug classes, namely INH, RIF, SLIs and FLQs, and the five countries with the largest number of resistant isolates (Figure 4). MRSCA ages did not differ between the UK and China across all four drug classes. These two countries had the oldest median MRSCA across the five countries and four drug classes except for INH. MRSCA ages in the UK were a median of 13.0 years (IQR 10.7-19.0) for RIF, a median of 10.6 years (IQR 7.3-11.2) for SLIs and a median of 8.4 years (IQR 6.7-18.4) for FLQs (Figure 4). South Africa most consistently had the youngest median MRSCA for the four drug classes, but its MRSCA distribution was not significantly different from that of Peru (for FLQs and SLIs) and Russia (for SLIs) (Table S3). A similar geographic/age pattern was observed for the drugs PZA and EMB across these five countries (Table S3).

We examined if the geographic resistance age differences correlated with resources available for TB control programs using the gross domestic product per capita as a proxy. We found GDP to correlate significantly with an older RIF MRSCA date with R^2= 0.47 (F-test with 1 DF P-value= 0.014) (Figure 6 & Table S8).

Distribution of Resistance Mutations:

We assessed the frequency of 267 resistance mutations previously determined to be important for resistance prediction[28] and their geographic distribution among the 8,550 isolates with country of origin and WGS data meeting quality criteria (Figure 1). Resistance mutation prevalence varied significantly by country. The most frequent INH causing mutation[31], *katG* S315T was more frequent among phenotypically INH resistant isolates by DST (pheno-R) in Russia (84%, n=526) than in Peru (67%, n=760) (Fisher P-value $1x10^{-12}$). The second most common INH resistance mutation -15 C>T *fabG1/inhA* promoter was more prevalent among INH pheno-R Peruvian isolates (20%) than in Russian isolates (8%) (Fisher P-value $7x10^{-9}$). Twenty four of the 267 resistance mutations (9%) varied geographically to a larger extent than the mutation *fabG/inhA* promoter -15C>T (standard deviation 11%, frequency range 0-39%, Table S13). The mutation I491F was recently described to be common in Estawini[32] and is not detectable by line-probe or

GeneXpert commercial molecular diagnostics. In our sample that did not contain data from Estawini, we calculated a standard deviation of 1% for the global frequency of I491F (range 0% - 4%) among RIF pheno-R isolates.

We calculated the proportion of pheno-R isolates that can be captured by the Hain Line-probe or GeneXpert commercial molecular diagnostics due to the presence of one or more mutations in their pooled target regions for the drugs INH, RIF, SLIs, and FLQ (Tables 1 and S14). Sensitivity was highest for RIF (90% of 2624) and lowest for FLQs (51% of 854). Specificity was consistently high (lowest for SLIs at 86%) (Table 1). Second-line sensitivity of commercial diagnostics differed significantly across countries (Table S14). FLQ sensitivity in Peru was 38% (n= 121) and 77% in South Africa (n=111) (Fisher P-value $1x10^{-6}$). A similarly low sensitivity for SLI resistance was seen in Peru compared with South Africa (Fisher P-value $3x10^{-4}$) (Table S14).

We examined if expanding the resistance mutation list to variants previously characterized in diverse global MTB genomic datasets using direct association[29] or random forests[28] can improve sensitivity and specificity in a "select WGS test." The select WGS test improved sensitivity slightly for INH and SLIs with relatively preserved specificity (Table 1). In addition select WGS test allowed for prediction of resistance to other drugs not tested by commercial diagnostics: PZA, EMB and Streptomycin. For comparison, we assessed if including any non-silent variant in the resistance regions (excluding a select number of known lineage markers) was indeed inferior to the more informed 'select WGS test" reported previously. We found that this "all WGS test" only modestly improved sensitivity and at the expense of a larger decrease in specificity (Figure 7).

Discussion

Using 8,550 clinical MTB sequences with culture-based DST, we examined geographic trends in the DR-TB epidemic. Geographically, MTB lineages 1-4 were each represented in the continents of Europe, Asia and Africa providing evidence of disease spread across borders, likely driven by human migration[30]. We found MDR-TB in nearly every country represented by more than 10 isolates. XDR isolates were found in half of these countries and spanning all five major continents. Lineage-2 had the highest percentage of MDR isolates in our sample followed by Lineages-4, 3 and 1. Using this diverse sample we dated more than 4869 resistance phenotypes across 4 lineages and 15 countries.

We found a consistent order of resistance acquisition globally among drug classes. The development of INH resistance was previously found to be a sentinel event heralding the development of MDR[33]. Our results corroborate these findings using phenotypic resistance data and across a larger geographically diverse sample. After INH, we find that MTB acquires resistance to RIF/STR then EMB followed by PZA, ETH, FLQ, SLIs, or CYS. We found no correlation between the age of resistance acquisition and the year of clinical introduction of the drug but there may be multiple other causes for the observed order of resistance acquisition. Differences in mutation rates across drug targets or resistance genes have been postulated but shown to be an unlikely explanation for INH resistance arising first[34,35]. Pharmacokinetic difference may result in higher risk for under-dosing[36] for some drugs and earlier resistance acquisition. Bacterial fitness costs are also variable across resistance mutations. For INH

252 resistance, mutations like *katG* S315T carry a low fitness cost and likely contribute to resistance
253 arising earliest for this drug[33,37,38]. The order of drug administration can explain dating
254 differences between first-line (INH, RIF, EMB, PZA) and second (ETH, FLQ, SLIs) or third-line
255 (CYS) resistance, as second-line drugs are usually only administered after resistance to first-
256 line drugs is ascertained. Acquisition of resistance to INH first then RIF may also relate to their
257 use for treatment of latent TB infection, leading to more exposure and selection pressure
258 overall. However, because adoption of INH preventative therapy for latent TB remains low in
259 many parts of the world, we expect it to be a lesser contributor to INH and RIF resistance
260 rates[39]. Lastly, the observation of contemporaneous acquisition of RIF and STR resistance is
261 likely best explained by the effects of Category II TB treatment initially recommended in 1991[40].
262 Category II is no longer recommended by the WHO but consists of adding streptomycin to the
263 first-line drug regimen after treatment failure. Our dating supports that streptomycin resistance
264 amplified among patients failing due to recent RIF resistance and/or MDR acquisition.

265 Published evidence supports that most resistant cases of MTB result from recent resistance
266 acquisition in the host or are related to transmission[41]. Reactivation of resistant MTB disease
267 acquired remotely (>2 years prior) is much less likely[42]. Thus the identification of isolates with
268 old resistance suggests high fitness for continued transmission between human hosts. Most
269 countries with available data had isolates with resistance dated more than 20 years prior. This is
270 also supported by our phylogenetic assessment where we estimate a lower bound of TB
271 resistance due to transmission to range between 14-52% across countries with available data.
272 As our approach cannot distinguish between resistance importation through human migration
273 after transmission outside of the country and new resistance acquisition, these figures are
274 underestimates of the true resistance burden due to transmission. Mathematical models of TB
275 rates have previously predicted transmission to be a major driver of observed resistance rates[43],
276 we present here WGS based evidence of the high burden of resistance transmission.
277 Mathematical models have also emphasized that drug resistant strain fitness is a key parameter
278 that dictates how the resistance epidemic will unfold. Our results support that >14-52% of
279 isolates are fit and successfully transmitting patient-to-patient and in most countries there have
280 been uninterrupted chains of resistance transmission for >20 years. These data emphasize the
281 need to contain resistance transmission through improved diagnosis, treatment and other
282 preventative strategies such as infection control and vaccine development.

283 In addition to transmission, we find evidence for recent resistance amplification, especially to
284 second-line drugs mediating the transition from MDR to XDR-TB. XDR has considerably worse
285 treatment outcome than susceptible TB and incurs more than 25 times the cost[44]. We estimate
286 that half of FLQ and SLI resistant isolates had acquired resistance within 4.7 years of isolation
287 despite the promotion of directly-observed-therapy (DOT) by the WHO since 1994. As most
288 FLQ and SLI resistant isolates are also MDR, our results also emphasize the need for better
289 regimens to treat MDR that can prevent resistance amplification. By country, we found a
290 significant correlation between the estimated age of resistance acquisition and per capita GDP,
291 with more affluent countries having older ages of resistance. This unexpected correlation is
292 likely driven by a combination of factors but the routine use of DST and close patient monitoring
293 in the more well-resourced health systems are likely important contributors. Specifically, we
294 found the UK and China to have the oldest resistance ages across the drugs. The Chinese
295 national TB program budget was increased from $98 million in 2002 to $272 million in 2007[45]

and a new policy for free TB diagnostics tests and drug use was introduced in 2004[58,46]. This increased investment can explain the observed low rates of recent resistance acquisition in China[30,47].

Likely due to geographic differences in MTB lineage, transmission and resistance acquisition rates, we find 10% of assayed resistance mutations to have high geographic variance. We also found commercial diagnostics to vary in sensitivity for second-line drugs. Given recent reports about the accuracy of WGS for confirming susceptibility of MTB[29], we measured improvements in resistance sensitivity offered by including mutations outside of regions targeted by commercial diagnostics through direct association. This offered modest improvements in sensitivity with little to no change in specificity. We found a considerable number of indeterminate mutations in resistance regions, that when included with simple direct association improve sensitivity but at the expense of loss of specificity. The study of these variants through statistical models will likely further inform their diagnostic use in the future[28,48].

Our study has several limitations including the oversampling of DR isolates as evidenced by our comparison with WHO reported MDR rates. We tried to control for this by dating only in countries with at least 20% susceptible isolates and limiting dating of low diversity samples that represent unique outbreak settings and lack long term information about resistance acquisition. This may have resulted in underestimation of rates of recent resistance acquisition but despite this we were able to document recent resistance acquisition in many countries. Molecular dating is also reliant on the accurate estimation of the phylogenetic tree of MTB isolates and the molecular clock assumption. We thus used a rigorous approach to phylogenetic estimation and dating despite its computational and time cost[49]. Our analysis also assumes the accuracy of culture based phenotypic DST, even though test to test variability is known to exist. We justify this as our data was curated from ReseqTB[15] and studies were phenotypic testing was performed in national or supranational laboratories with rigorous quality control.

Overall, our results support that DR rates are fuelled by both recent resistance acquisition and ongoing transmission, and suggest the need for better detection, treatment and health system investment. In the future, reassessment of these patterns will be enabled by the sharing of systematically collected isolate data, data that is increasingly generated as by-products of TB surveillance and resistance diagnosis[29].

Figures and Tables

Figure 1: Flow diagram showing process of identification and exclusion of genomic data included in the study. WGS: Whole genome sequencing, QC: Quality control.

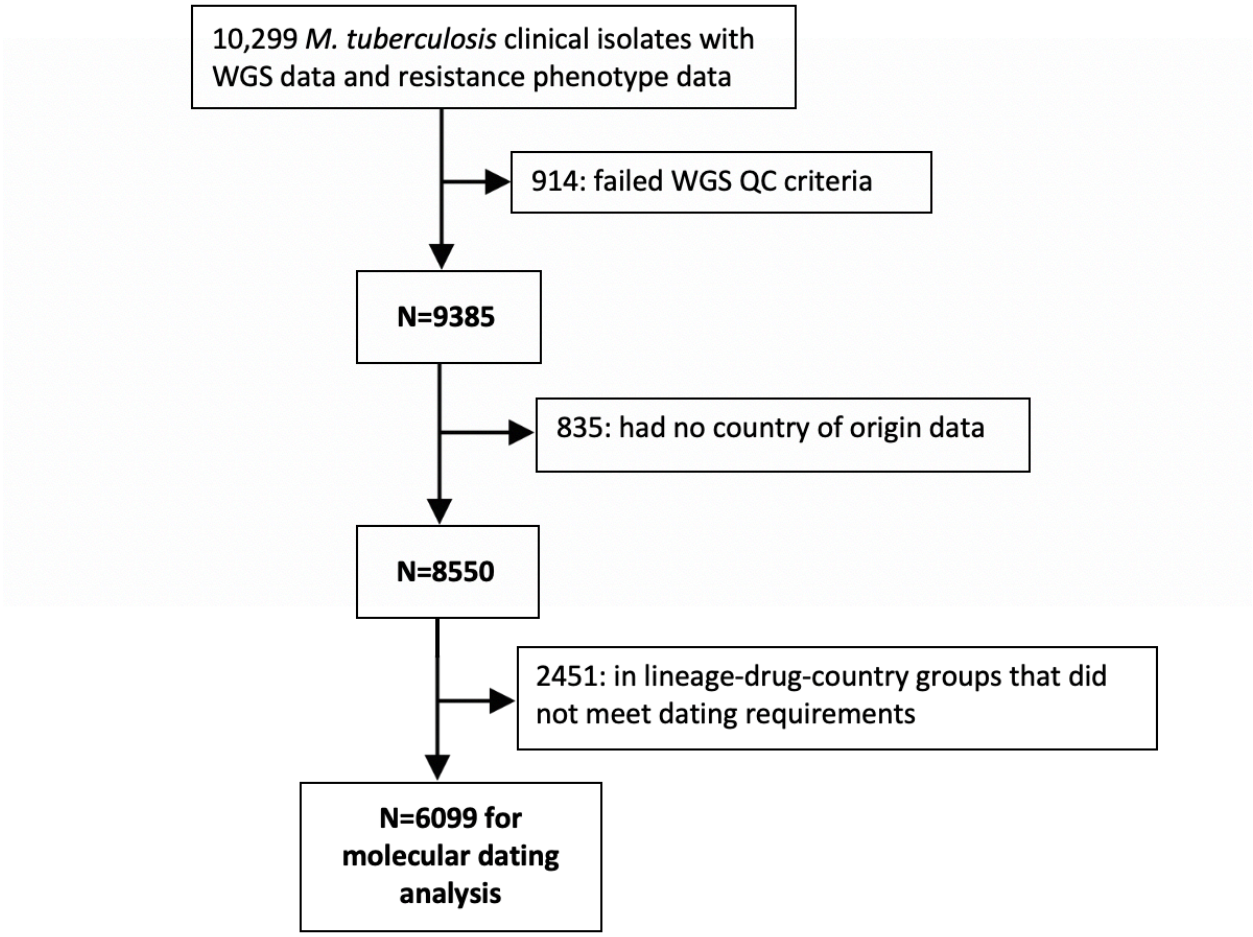
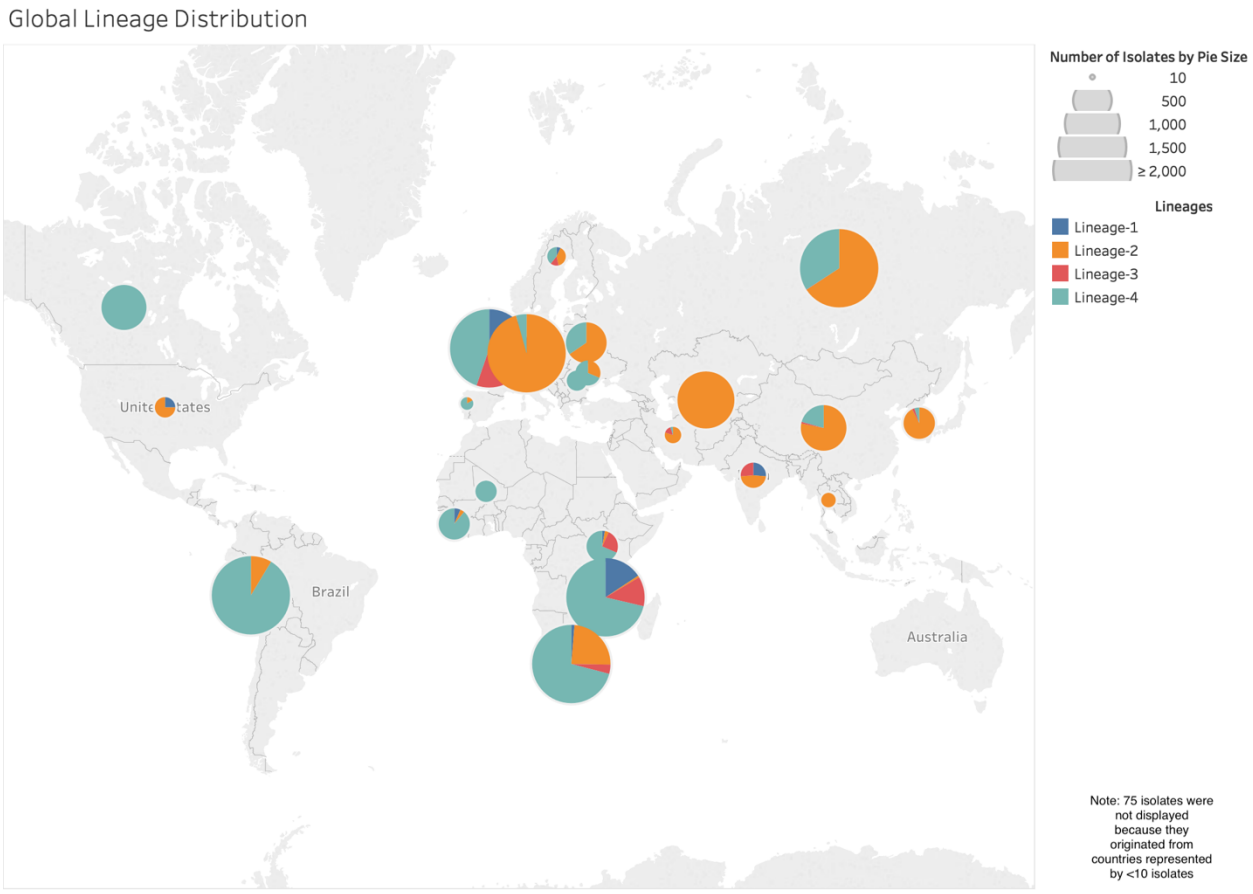
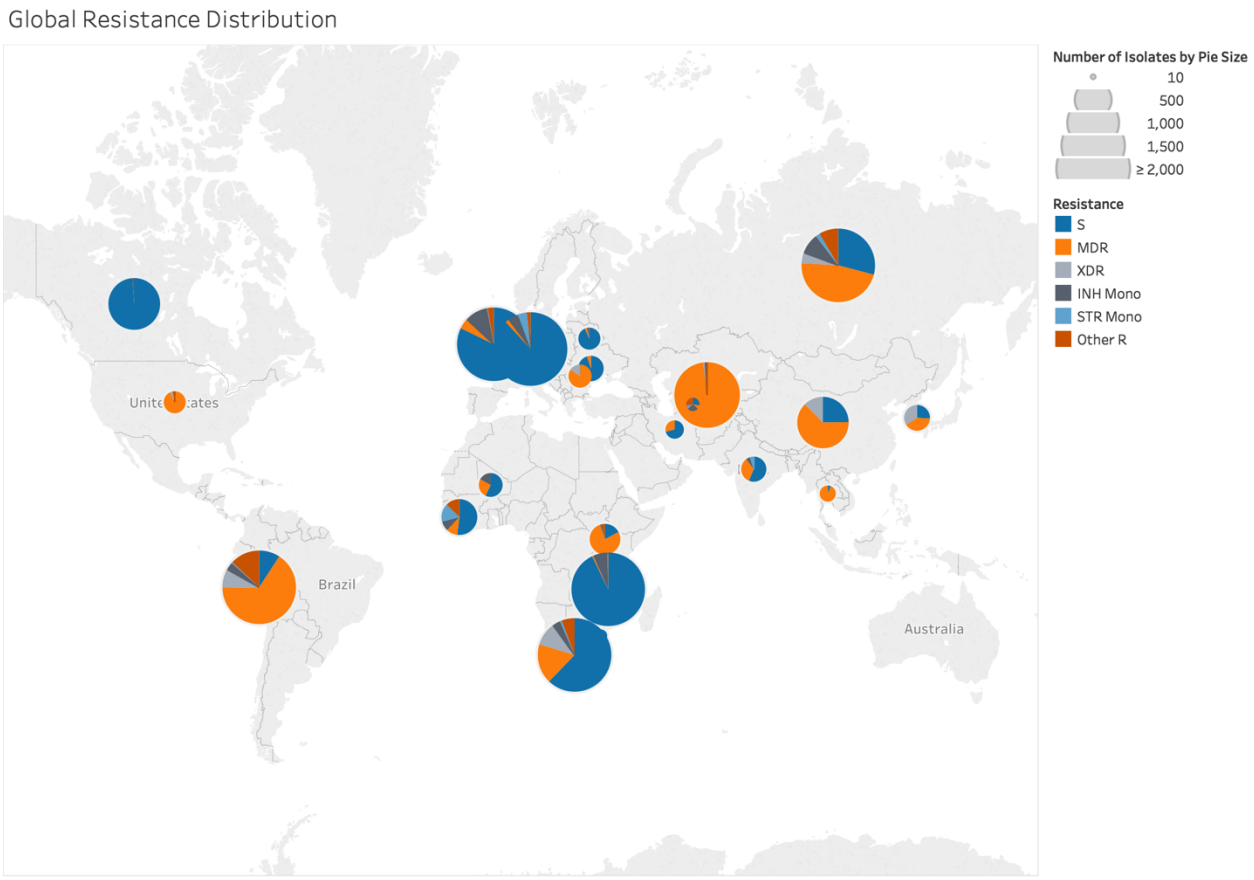


Figure 2: Global Distribution of *M. tuberculosis* in the study sample. Counts from countries represented by fewer than 10 isolates (n=75) not shown. **A:** Lineage distribution (n=8477). Pie charts represent the proportion of each lineage among isolates available from each country. Size of the pie is proportional to the number of isolates from each country. Detailed counts are in Table S1. **B:** Drug resistance distribution (n=7834). Pie charts provide the distribution of resistance patterns (S: Susceptible, MDR: Multidrug-resistant, XDR: Extensively drug resistant, INH Mono: mono resistant to isoniazid, STR Mono: mono resistant to streptomycin, Other R: resistance other than defined categories) by country. 75 isolates originated Pie size is proportional to the number of isolates in each country (Table S2).

2A:



345 2B:



346

347

348

Figure 3A: MRSCA distribution by drug (n=4844). Boxplots showing range of MRSCA distribution globally for nine anti-tubercular drugs (MRSCA: Most recent susceptible common ancestor, INH: isoniazid, RIF: rifampicin, EMB: ethambutol, PZA: pyrazinamide, STR: streptomycin, FLQ: fluoroquinolones, SLIs: second-line injectables, ETH: ethionamide, CYS: cycloserine)

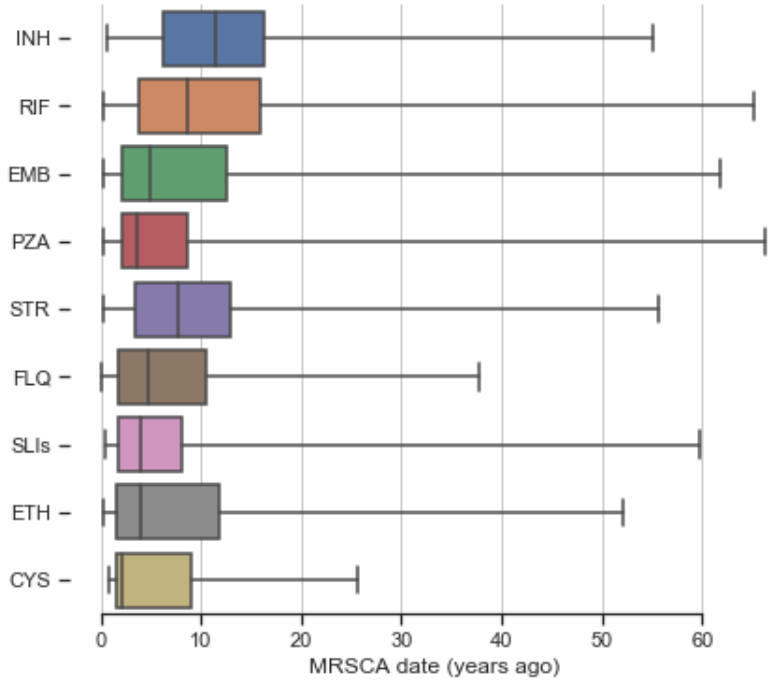
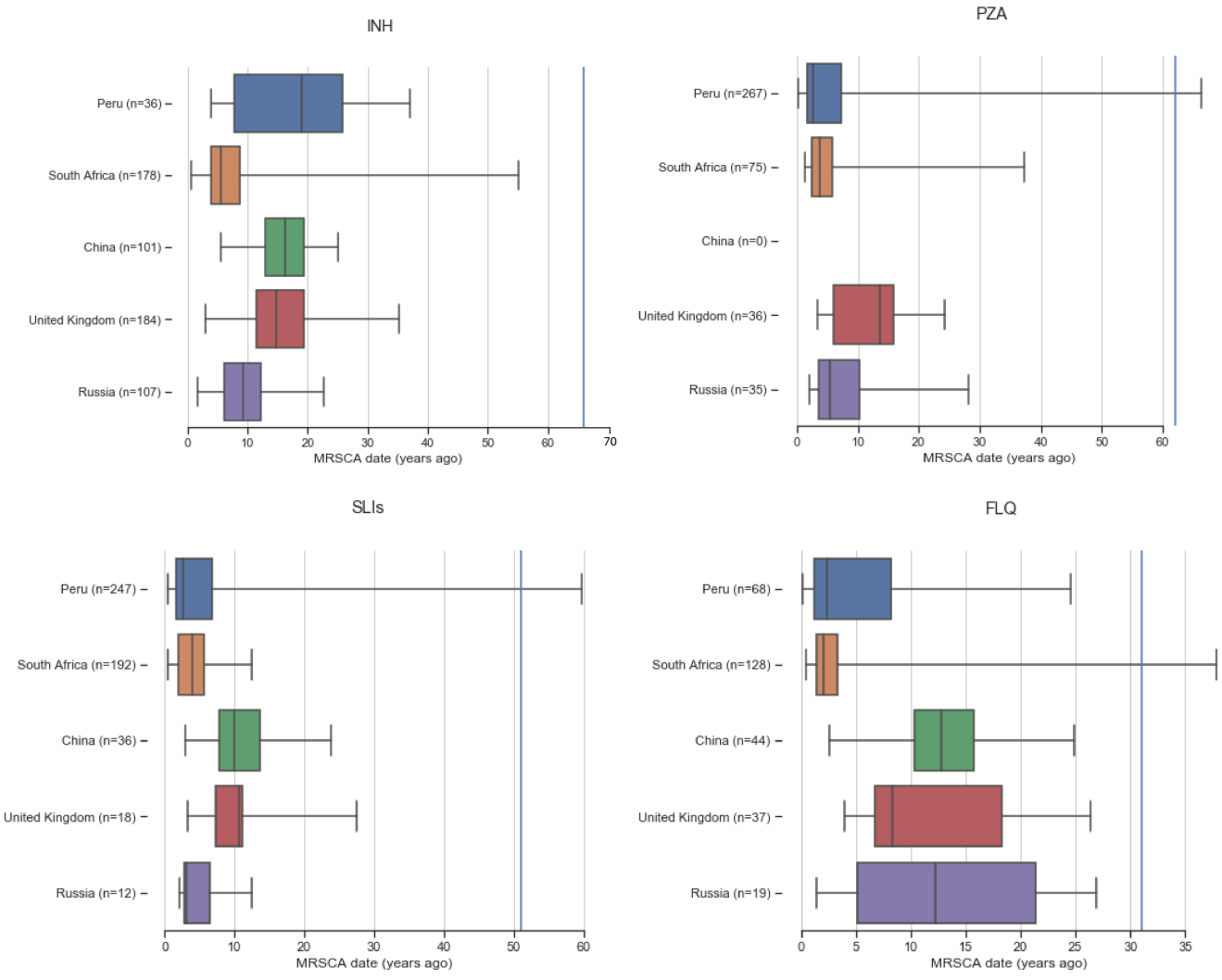


Figure 3B: Pairwise Wilcoxon rank sum tests comparing MRSCA ages by drug category. Dark red indicates P-value <0.001 (Bonferroni threshold); pink indicates P-value <0.01; white indicates P-value ≥0.01.

	INH	RIF	EMB	PZA	STR	FLQ	SLIs	ETH	CYS
INH		5 E-16	2 E-29	3 E-46	1 E-17	1 E-31	1 E-59	1 E-23	5 E-6
RIF			2 E-6	1 E-15	0.3	4 E-11	9 E-21	2 E-8	9 E-4
EMB				1 E-3	5 E-4	0.01	1 E-4	0.02	0.03
PZA					5 E-12	0.9	0.8	0.9	0.08
STR						5 E-8	2 E-16	6 E-6	3 E-3
FLQ							0.4	0.9	0.2
SLIs								0.5	0.1
ETH									0.2
CYS									

361 **Figure 4: Box and whisker plots summarizing the MRSCA distribution per country.** (Blue
362 vertical line indicates year when drug was introduced (Table S5))

363



36

365
366

Figure 5: Proportion of MDR isolates with recent amplification of resistance to ethambutol (EMB), pyrazinamide (PZA), fluoroquinolones (FLQ) or second-line injectables (SLIs) by country (MRSCA age estimate <5 years ago). The legend lists the number of multidrug-resistant isolates analysed from each country. Error bars indicate 95% confidence intervals. Full data given in Supplementary Table 3. Four countries displayed a measurable proportion of recent FLQ and SLI amplification (95% CI does not include 0) – South Africa, Peru, Romania and China.

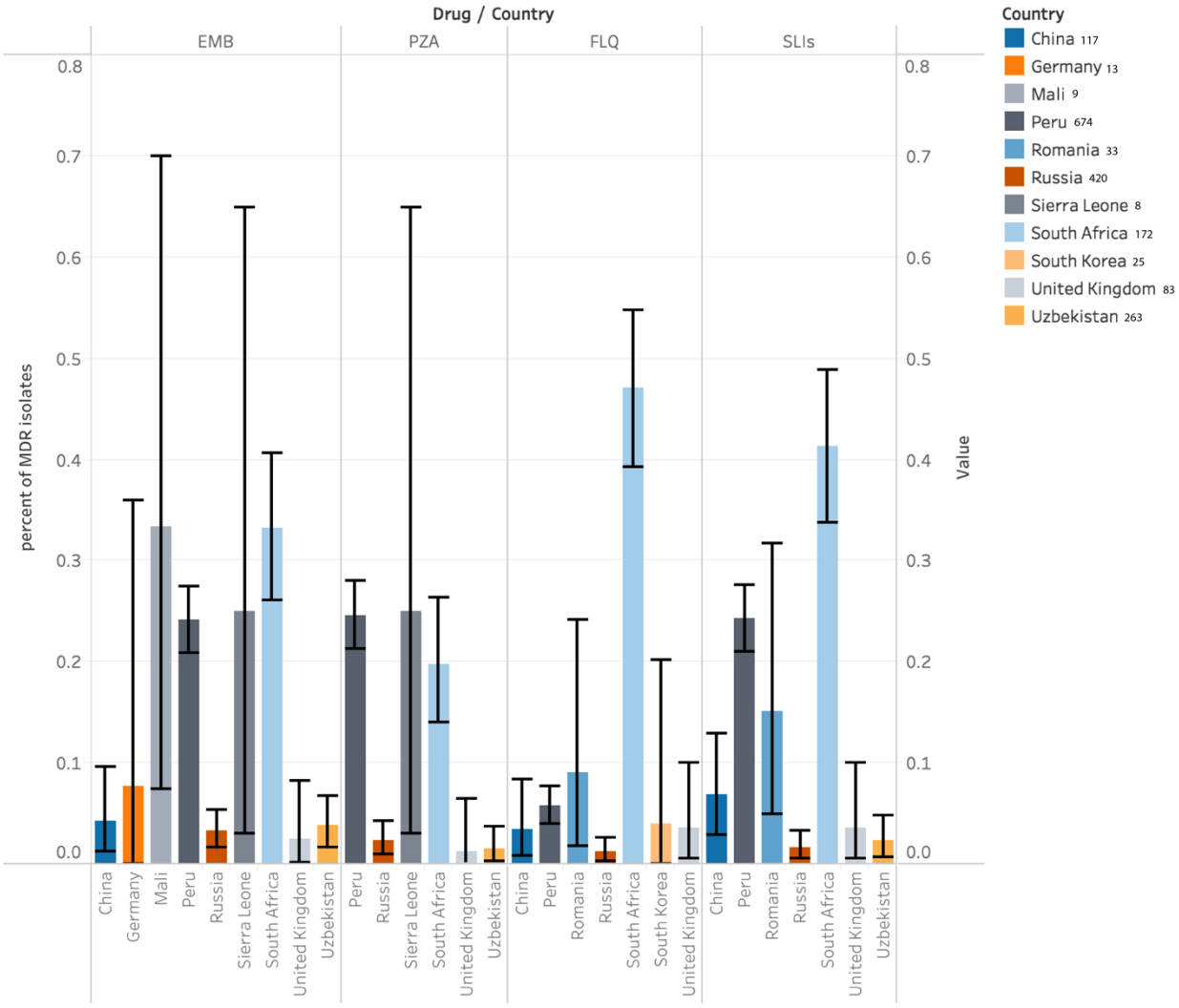


Figure 6: Median Rifamycin (RIF) most-recent-common-susceptible-ancestor (MRSCA) date vs Gross Domestic Product (GDP) per capita for 12 countries. Data plotted is provided in Supplementary Table 7 and includes other drugs than RIF. (R^2 = 0.47, F-test P-value (1 DF) =0.014).

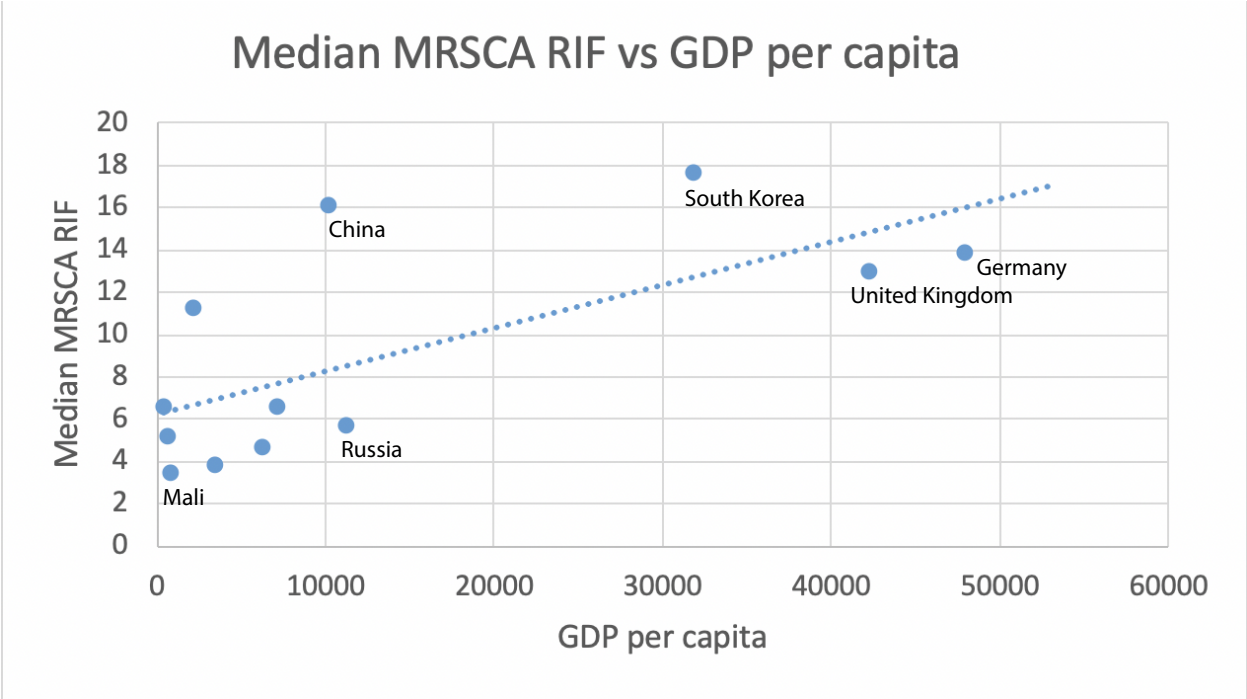


Table 1: Sensitivity and Specificity of commercial and WGS based tests for resistance diagnosis. Abbreviations defined in "drug resistance definitions" section of methods.

Drug	commercial test		RF-select WGS test		DA-select WGS test		all WGS test*	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivty	Specificty
INH	83% (2759/3306)	93% (4834/5201)	88% (2900/3306)	92% (4780/5201)	89% (2956/3306)	92% (4776/5201)	92% (3029/3306)	64% (3314/5201)
RIF	90% (2354/2624)	93% (5361/5786)	91% (2385/2624)	92% (5341/5786)	91% (2395/2624)	92% (5338/5786)	92% (2405/2624)	89% (5178/5786)
FLQ	53% (452/854)	94% (2626/2790)	51% (439/854)	95% (2637/2790)	51% (440/854)	95% (2641/2790)	57% (488/854)	86% (2406/2790)
SLI	56% (517/921)	86% (2179/2547)	58% (535/921)	84% (2136/2547)	57% (524/921)	86% (2185/2547)	64% (594/921)	80% (2027/2547)
PZA			65% (862/1324)	95% (4660/4907)	75% (996/1324)	91% (4485/4907)	78% (1030/1324)	89% (4384/4907)
EMB			79% (1476/1863)	86% (4702/5441)	75% (1389/1863)	88% (4772/5441)	85% (1589/1863)	74% (4042/5441)
STR			76% (1601/2112)	86% (3204/3726)	77% (1627/2112)	87% (3223/3726)	91% (1928/2112)	71% (2657/3726)
Legend								
*	classifying FLQ and SLI as susceptible if no INH and no RIF resistance mutation found							
Sensitivity	Percent of resistant isolates classified as resistant							
Specificity	Percent of susceptible isolates classified as susceptible							

392 **Authors' contributions:**

393 Yasha Ektefaie conducted the data analysis, drafted and revised the manuscript.
394 All authors provided key edits to the manuscript.
395 Additionally:
396 Luca Freschi contributed to the data analysis.
397 Avika Dixit contributed to the data analysis.
398 Maha Farhat conceptualized the study, supervised the data analysis, reviewed, wrote and
399 edited the manuscript.
400
401 **Declaration of interests:**
402
403 Yasha Ektefaie: None
404 Luca Freschi: None
405 Avika Dixit: None
406 Maha Farhat: None

References:

1. *World Health Organization. Global Tuberculosis Report 2016*. (World Health Organization, 2016).
2. Churchyard, G. *et al.* What We Know About Tuberculosis Transmission: An Overview. *J Infect Dis* **216**, S629–S635 (2017).
3. Mathema, B. *et al.* Drivers of Tuberculosis Transmission. *J Infect Dis* **216**, S644–S653 (2017).
4. Farhat, M. R. *et al.* Genomic analysis identifies targets of convergent positive selection in drug-resistant Mycobacterium tuberculosis. *Nat. Genet.* (2013) doi:10.1038/ng.2747.
5. Farhat, M. R. *et al.* Genome wide association with quantitative resistance phenotypes in Mycobacterium tuberculosis reveals novel resistance genes and regulatory regions. *bioRxiv* 429159 (2018) doi:10.1101/429159.
6. Coll, F. *et al.* Genome-wide analysis of multi- and extensively drug-resistant Mycobacterium tuberculosis. *Nature Genetics* **50**, 307–316 (2018).
7. Nebenzahl-Guimaraes, H. *et al.* Transmissible Mycobacterium tuberculosis Strains Share Genetic Markers and Immune Phenotypes. *Am J Respir Crit Care Med* **195**, 1519–1527 (2016).
8. The CRyPTIC Consortium and the 100, 000 Genomes Project. Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *New England Journal of Medicine* (2018) doi:10.1056/NEJMoa1800474.
9. Walker, T. M. *et al.* Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect Dis* (2015) doi:10.1016/S1473-3099(15)00062-6.

10. Dixit, A. *et al.* Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. *Scientific Reports* **9**, 5602 (2019).
11. Walker, T. M. *et al.* Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *Lancet Infect Dis* **13**, 137–146 (2013).
12. Gardy, J. L. *et al.* Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak. *New England Journal of Medicine* **364**, 730–739 (2011).
13. Merker, M. *et al.* Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nature Genetics* **47**, 242–249 (2015).
14. Wattam, A. R. *et al.* Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res.* **45**, D535–D542 (2017).
15. Ezewudo, M. *et al.* Integrating standardized whole genome sequence analysis with a global Mycobacterium tuberculosis antibiotic resistance knowledgebase. *Sci Rep* **8**, 15382 (2018).
16. Chatterjee, A., Nilgiriwala, K., Saranath, D., Rodrigues, C. & Mistry, N. Whole genome sequencing of clinical strains of Mycobacterium tuberculosis from Mumbai, India: A potential tool for determining drug-resistance and strain lineage. *Tuberculosis* **107**, 63–72 (2017).
17. Manson, A. L. *et al.* Mycobacterium tuberculosis Whole Genome Sequences From Southern India Suggest Novel Resistance Mechanisms and the Need for Region-Specific Diagnostics. *Clin. Infect. Dis.* **64**, 1494–1501 (2017).
18. Walker, T. M. *et al.* Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. *Lancet Infect. Dis.* **15**, 1193–1202 (2015).

19. Bryant, J. M. *et al.* Inferring patient to patient transmission of Mycobacterium tuberculosis from whole genome sequencing data. *BMC Infect. Dis.* **13**, 110 (2013).

20. Casali, N. *et al.* Evolution and transmission of drug-resistant tuberculosis in a Russian population. *Nat. Genet.* **46**, 279–286 (2014).

21. Clark, T. G. *et al.* Elucidating emergence and transmission of multidrug-resistant tuberculosis in treatment experienced patients by whole genome sequencing. *PLoS One* **8**, e83012 (2013).

22. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans | Nature Genetics. https://www.nature.com/articles/ng.2744#s1.

23. Perdigão, J. *et al.* Unraveling Mycobacterium tuberculosis genomic diversity and evolution in Lisbon, Portugal, a highly drug resistant setting. *BMC Genomics* **15**, 991 (2014).

24. Blouin, Y. *et al.* Significance of the identification in the Horn of Africa of an exceptionally deep branching Mycobacterium tuberculosis clade. *PLoS One* **7**, e52841 (2012).

25. Cohen, K. A. Evolution of Extensively Drug-Resistant Tuberculosis over Four Decades: Whole Genome Sequencing and Dating Analysis of Mycobacterium tuberculosis Isolates from KwaZulu-Natal. https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001880.

26. Zhang, H. *et al.* Genome sequencing of 161 Mycobacterium tuberculosis isolates from China identifies genes and intergenic regions associated with drug resistance. *Nat. Genet.* **45**, 1255–1260 (2013).

27. Ezewudo, M. *et al.* Integrating standardized whole genome sequence analysis with a global Mycobacterium tuberculosis antibiotic resistance knowledgebase. *Scientific Reports* **8**, 15382 (2018).

476 28. Farhat, M. R. *et al.* Genetic Determinants of Drug Resistance in Mycobacterium tuberculosis
477 and Their Diagnostic Value. *Am J Respir Crit Care Med* **194**, 621–630 (2016).
478 29. Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. *New*
479 *England Journal of Medicine* **379**, 1403–1415 (2018).
480 30. WHO | Tuberculosis country profiles. *WHO* http://www.who.int/tb/country/data/profiles/en/.
481 31. Nebenzahl-Guimaraes, H., Jacobson, K. R., Farhat, M. R. & Murray, M. B. Systematic
482 review of allelic exchange experiments aimed at identifying mutations that confer drug
483 resistance in Mycobacterium tuberculosis. *J Antimicrob Chemother* **69**, 331–342 (2014).
484 32. André, E. *et al.* Novel rapid PCR for the detection of Ile491Phe rpoB mutation of
485 Mycobacterium tuberculosis, a rifampicin-resistance-conferring mutation undetected by
486 commercial assays. *Clin. Microbiol. Infect.* **23**, 267.e5-267.e7 (2017).
487 33. Manson, A. L. *et al.* Genomic analysis of globally diverse *Mycobacterium tuberculosis*
488 strains provides insights into the emergence and spread of multidrug resistance. *Nature*
489 *Genetics* **49**, 395–402 (2017).
490 34. Bergval, I. L., Schuitema, A. R. J., Klatser, P. R. & Anthony, R. M. Resistant mutants of
491 Mycobacterium tuberculosis selected in vitro do not reflect the in vivo mechanism of
492 isoniazid resistance. *J. Antimicrob. Chemother.* **64**, 515–523 (2009).
493 35. Ford, C. B. *et al.* Mycobacterium tuberculosis mutation rate estimates from different lineages
494 predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat. Genet.*
495 **45**, 784–790 (2013).
496 36. Jutte, P. C., Rutgers, S. R., Van Altena, R., Uges, D. R. & Van Horn, J. R. Penetration of
497 isoniazid, rifampicin and pyrazinamide in tuberculous pleural effusion and psoas abscess.
498 *Int. J. Tuberc. Lung Dis.* **8**, 1368–1372 (2004).

499 37. Ordway, D. J., Sonnenberg, M. G., Donahue, S. A., Belisle, J. T. & Orme, I. M. Drug-
500 resistant strains of Mycobacterium tuberculosis exhibit a range of virulence for mice. *Infect*
501 *Immun* **63**, 741–743 (1995).
502 38. Pym, A. S., Saint-Joanis, B. & Cole, S. T. Effect of katG mutations on the virulence of
503 Mycobacterium tuberculosis and the implication for transmission in humans. *Infect. Immun.*
504 **70**, 4955–4960 (2002).
505 39. Alsdurf, H., Hill, P. C., Matteelli, A., Getahun, H. & Menzies, D. The cascade of care in
506 diagnosis and treatment of latent tuberculosis infection: a systematic review and meta-
507 analysis. *Lancet Infect Dis* **16**, 1269–1278 (2016).
508 40. WHO | Eliminating the category II retreatment regimen from national tuberculosis
509 programme guidelines: the Georgian experience. *WHO*
510 https://www.who.int/bulletin/volumes/90/1/11-092320/en/.
511 41. Kendall, E. A., Fofana, M. O. & Dowdy, D. W. Burden of transmitted multidrug resistance
512 in epidemics of tuberculosis: a transmission modelling analysis. *Lancet Respir Med* **3**, 963–
513 972 (2015).
514 42. Behr, M. A., Edelstein, P. H. & Ramakrishnan, L. Revisiting the timetable of tuberculosis.
515 *BMJ* **362**, k2738 (2018).
516 43. Kendall, E. A., Fojo, A. T. & Dowdy, D. W. Expected effects of adopting a 9 month regimen
517 for multidrug-resistant tuberculosis: a population modelling analysis. *Lancet Respir Med* **5**,
518 191–199 (2017).
519 44. Marks, S. M. *et al.* Treatment Practices, Outcomes, and Costs of Multidrug-Resistant and
520 Extensively Drug-Resistant Tuberculosis, United States, 2005–2007 - Volume 20, Number
521 5—May 2014 - Emerging Infectious Diseases journal - CDC. doi:10.3201/eid2005.131037.

45. WHO | Global tuberculosis reports (from 1997). *WHO* http://www.who.int/tb/publications/global_report/archive/en/.

46. Health reform and development report. http://www.gov.cn/ztzl/2005-10/20/content_80720.htm.

47. Zhao, Y. *et al.* National survey of drug-resistant tuberculosis in China. *N. Engl. J. Med.* **366**, 2161–2170 (2012).

48. Chen, M. L. *et al.* Beyond multidrug resistance: Leveraging rare variants with machine and statistical learning models in Mycobacterium tuberculosis resistance prediction. *EBioMedicine* **43**, 356–369 (2019).

49. Menardo, F., Duchêne, S., Brites, D. & Gagneux, S. The molecular clock of Mycobacterium tuberculosis. *PLoS Pathog.* **15**, e1008067 (2019).

Supplementary Methods:

Data curation and quality control:

The isolate metadata including their geographic locations were downloaded using metatools_ncbi (https://github.com/farhat-lab/metatools_ncbi). We also performed literature curation to fill the gaps in the NCBI geographic location data. The resulting table of the geographic locations of the isolates is available in Supplementary File 1. We excluded isolates that did not meet WGS quality control criteria as detailed below, had no geographic information or were not tested for phenotypic resistance to one or more drugs.

Genomic analysis/variant calling:

Briefly, reads were trimmed using PRINSEQ[1] setting average phred score threshold to 20. Raw read data was confirmed to belong to MTB complex using Kraken[2]. Isolates with <90% mapping were excluded. Reads were aligned to H37Rv (GenBank NC000962.3) reference genome using BWA MEM[3]. Duplicate reads were removed using PICARD[4]. We excluded any isolates with coverage <95% of known drug resistance regions (*katG*, *inhA* & its promoter, *rpoB*, *embA, embB, embC* & *embB* promoter, *ethA*, *gyrA* and *gyrB*, *rrs*, *rpsL*, *gid*, *pncA*, *rpsA*, *eis* promoter) at 10x or higher. Variants were called using Pilon[5] that uses local assembly to increase indel (insertions and deletions) call accuracy. This deviates from Ezewudo *et al.* that uses Samtools for variant calling[6]. The reference allele was implied if allele frequency was <75% or the Pilon filter was not PASS. Low confidence coordinates were filtered from all strains if >95% of strains did not have coverage of (trimmed reads) at least 10x at that site. Isolate lineage belonging to one of the seven main MTB lineages was confirmed using the Coll *et al.* SNP barcode[7].

Drug resistance definitions and comparison with WHO reported resistance proportions:

The 'Mono' resistant designation was given to isolates that were resistant to only one specific drug and susceptible to all others that were tested, with the exception of the INH-mono resistant category that encompassed any isolates that were resistant to INH and/or STR but not to others that were tested. The 'Other-R' category was reserved for isolates that were resistant to some drugs but were neither INH or STR mono resistant, nor MDR or XDR. Isolates were labelled susceptible if they were susceptible to all drugs tested.

To compute exact binomial confidence intervals for MDR proportions by country we used the python library statsmodel[8]. To assess overlap with World Health Organization (WHO) estimates we determined whether the confidence intervals of our proportion intersected with that of the WHO. We labelled our estimate as high if our confidence interval was higher than the WHO, low if it was lower, and the same if they intersected.

Estimating resistance acquisition dates and lower bounds of resistance transmission:

A maximum likelihood tree was generated for each group via RAxML 8.2.11[9] with H37Rv (NC000962.3) as the outgroup, starting from a neighbour-joining seed tree and assuming a generalized time reversible (GTR) nucleotide substitution model with the Γ distribution used to model site rate heterogeneity[10]. We bootstrapped the maximum likelihood tree 1000 times. The maximum likelihood tree was dated using BEAST v1.10.4[11] assuming a relaxed molecular clock with a log normal distribution and a mean rate of 0.5 SNP per genome per year based on prior published data[12]. Sumtrees.py from the DendroPy library[13] was then used to combine the output from the bootstrap analysis and that of BEAST to get our final dated phylogenetic tree with nodal bootstrap support.

We dated the most recent common ancestor between all the resistant isolates and their most closely related susceptible isolate. Accordingly, the dates of resistance acquisition will be referred to as the estimated age of the most recent *susceptible* common ancestor (MRSCA) in years prior to isolation of the clinical sample(s) throughout the text. We excluded resistant isolates with MRSCAs inferred at nodes with less than 50 bootstrap support.

We calculated the number of phylogenetically inferred resistance acquisition events (N_{aq}) per country and lineage as the number of unique MRSCAs identified. This was compared with the total resistant isolates that could be dated (N_{td}). Phylogenetically inferred unique resistance acquisition events for a particular country may be related to either *in host* evolution of new resistance or due to human migration/importation from another country with the latter still possibly related to transmission elsewhere. Thus, the following quantity represents a lower bound on the burden of resistance due to transmission for a particular country:

$$(N_{td} - N_{aq})/N_{td}$$

To estimate the order of resistance acquisition for different drugs we pooled the MRSCA dates by drug across countries and lineages. We compared the medians of the MRSCA distributions and performed pairwise Wilcoxon Rank Sum tests to assess for statistical significance, correcting for multiple testing using the Bonferroni approach. Interquartile ranges were calculated using the python package numpy[14].

We correlated the median MRSCA date per drug pooled across countries against the date of drug introduction using linear regression as implemented in Microsoft Excel version 16.25.

Distribution of resistance mutations

We measured resistance mutations' geographic variance by calculating the proportion of resistant isolates with the resistance mutation per country, excluding countries with less than 10 resistant isolates for a particular drug. We computed the standard deviation of this distribution of proportions across countries. To test cross-country differences in the proportion of INH phenotypically resistant isolates that contained specific mutations, we used a Fisher-test using the python library scipy[15].

GDP/Programmatic Spending

To correlate countries' gross domestic product (GDP) per capita against resistance acquisition dates, GDP per capita data for 2019 was gathered from the International Monetary Fund (IMF)[16] and plotted against the median MRSCA date for RIF using Microsoft Excel version 16.25. F-value and significance was calculated via Anova[17] in Excel.

Supplementary Material

Supplementary Table 1 Global Lineage Distribution

	Lineage-1	Lineage-2	Lineage-3	Lineage-4	Lineage-5	Lineage-6	Lineage-7	Total
Belarus	0	87	0	46	0	0	0	135
Canada	0	0	0	17	0	0	0	165
China	0	132	2	35	0	0	0	170
Germany	0	22	0	1	0	0	0	857
India	5	9	5	0	0	0	0	52
Iran	0	16	3	1	0	0	0	22
Malawi	208	10	162	945	0	0	0	1427
Mali	0	0	0	37	0	0	0	37
Moldova	0	16	0	35	0	0	0	51
Netherlands	7	19	0	70	0	0	0	98
Peru	0	82	0	869	0	0	0	1098
Portugal	0	2	0	10	0	0	0	13
Romania	0	0	0	33	0	0	0	33
Russia	0	525	1	273	0	0	0	868
Sierra Leone	3	2	0	41	4	10	0	79
South Africa	8	144	23	427	0	0	0	974
South Korea	0	37	1	2	0	0	0	80
Swaziland	0	9	0	0	0	0	0	10
Sweden	2	11	4	11	0	0	0	28
Thailand	0	2	0	0	0	0	0	17
Turkmenistan	0	10	0	0	0	0	0	11
Uganda	2	3	18	50	0	0	0	80
United Kingdom	222	102	560	716	9	4	1	1873
United States of America	1	3	0	0	0	0	0	34
Uzbekistan	0	20	0	0	0	0	0	265
Unknown	37	81	35	792	1	0	0	1749
Total	495	1344	814	4411	14	14	1	10226

Supplementary Table 2 Global Resistance Distribution and Other Resistance Count

Abbreviations: Susceptible (S), Multi-drug Resistant (MDR), Extensively Drug Resistant (XDR), Isoniazid Mono Resistant (INH_MONO), Streptomycin Mono Resistant (STR_MONO), Other Resistant (OTHER_R), Sum With Data (SUM_W_DATA)

country	S	MDR	XDR	INH_MONO	STR_MONO	OTHER_R	NO_DATA	SUM	SUM_W_DATA	MDR/OTHER_R
Belarus	28	1	0	1	0	0	105	135	30	NA
Canada	163	0	0	2	0	0	0	165	165	NA
China	46	117	23	0	0	0	7	170	163	NA
Germany	749	13	0	40	34	14	7	857	850	0.928571429
India	22	13	0	2	2	0	13	52	39	NA
Iran	15	6	0	0	0	0	1	22	21	NA
Malawi	1302	7	0	90	1	3	24	1427	1403	2.333333333
Mali	20	9	0	6	0	0	2	37	35	NA
Moldova	38	2	0	0	0	0	11	51	40	NA
Netherlands	68	0	0	5	20	0	5	98	93	NA
Not Provided	179	646	138	90	6	109	719	1749	1030	5.926605505
Peru	94	674	86	34	5	134	157	1098	941	5.029850746
Romania	0	33	6	0	0	0	0	33	33	NA
Russia	261	420	42	80	17	76	14	868	854	5.526315789
Sierra Leone	41	8	0	7	13	10	0	79	79	0.8
South Africa	612	172	102	34	8	56	92	974	882	3.071428571
South Korea	16	25	21	0	0	0	39	80	41	NA
Swaziland	8	1	0	0	0	1	0	10	10	1
Thailand	1	15	0	0	0	0	1	17	16	NA
Turkmenistan	3	1	0	3	1	3	0	11	11	0.333333333
Uganda	10	45	0	1	0	2	22	80	58	22.5
United Kingdom	1525	83	4	187	5	50	23	1873	1850	1.66
ıited States of Ameri	0	29	1	0	0	1	4	34	30	29
Uzbekistan	0	263	3	1	0	1	0	265	265	263
Total	5201	2583	426	583	112	460	1246	10185	8939	5.615217391

Abbreviations: RIF (Rifamycin), ETH (Ethionamide), FQ (Fluoroquinolones), SLIS (Second Line Injectable), "_" (AND—EX: RIF_ISONIAZID_ETH = Isolate resistant to Rifamycin, Isoniazid, and Ethionamide)

Other Resistance Categories	Count
RIF_	78
PYRAZINAMIDE_	65
ETHAMBUTOL_	44
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_	29
ISONIAZID_ETHAMBUTOL_	28
ISONIAZID_PYRAZINAMIDE_	21
ISONIAZID_PYRAZINAMIDE_STREPTOMYCIN_	16
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_	11
RIF_ETHAMBUTOL_	11
RIF_STREPTOMYCIN_	10
ISONIAZID_STREPTOMYCIN_ETH_	9
ISONIAZID_ETHAMBUTOL_FQ_	9
SLIS_	8
ISONIAZID_STREPTOMYCIN_FQ_	8
ETHAMBUTOL_FQ_	7
FQ_	7
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_FQ_	7
ETHAMBUTOL_STREPTOMYCIN_	6
ISONIAZID_ETH_	6
STREPTOMYCIN_FQ_	5
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_ETH_	5
ISONIAZID_FQ_	4
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_ETH_	4
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_SLIS_ETH_	3
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_SLIS_	3
ISONIAZID_PYRAZINAMIDE_ETH_	3
ETHAMBUTOL_PYRAZINAMIDE_	3
RIF_PYRAZINAMIDE_STREPTOMYCIN_	2
RIF_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_FQ_	2
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_ETH_	2
PYRAZINAMIDE_STREPTOMYCIN_	2
RIF_ETHAMBUTOL_STREPTOMYCIN_	2
ISONIAZID_SLIS_	2
PYRAZINAMIDE_SLIS_	2
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_SLIS_ETH_	2
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_	2
PYRAZINAMIDE_FQ_	1
ISONIAZID_PYRAZINAMIDE_FQ_SLIS_	1
STREPTOMYCIN_SLIS_	1
RIF_ETHAMBUTOL_FQ_	1
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_SLIS_ETH_	1
ISONIAZID_ETHAMBUTOL_STREPTOMYCIN_SLIS_	1
RIF_ETHAMBUTOL_STREPTOMYCIN_SLIS_ETH_	1
ETH_	1
RIF_PYRAZINAMIDE_FQ_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_FQ_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_ETH_	1
RIF_PYRAZINAMIDE_	1
ETHAMBUTOL_ETH_	1
ISONIAZID_ETHAMBUTOL_ETH_	1
RIF_ETHAMBUTOL_SLIS_	1
ISONIAZID_PYRAZINAMIDE_STREPTOMYCIN_ETH_	1
RIF_PYRAZINAMIDE_SLIS_ETH_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_STREPTOMYCIN_	1
ISONIAZID_ETHAMBUTOL_FQ_SLIS_	1
ETHAMBUTOL_STREPTOMYCIN_FQ_	1
PYRAZINAMIDE_SLIS_ETH_PAS_	1
ISONIAZID_STREPTOMYCIN_FQ_SLIS_ETH_PAS_	1
ISONIAZID_STREPTOMYCIN_SLIS_PAS_	1
ISONIAZID_ETHAMBUTOL_SLIS_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_FQ_SLIS_ETH_	1
RIF_FQ_SLIS_	1
ISONIAZID_PYRAZINAMIDE_SLIS_	1
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_SLIS_ETH_PAS_	1
RIF_ETHAMBUTOL_PYRAZINAMIDE_SLIS_	1
RIF_PYRAZINAMIDE_STREPTOMYCIN_FQ_SLIS_	1
ISONIAZID_ETHAMBUTOL_PYRAZINAMIDE_SLIS_	1
RIF_SLIS_	1
RIF_FQ_	1
	462

Supplementary Table 3 Pairwise Country MRSCA date comparison

Dark red indicates P-value <0.001 (Bonferroni threshold); pink indicates P-value <0.01; white indicates P-value ≥0.01. *No PZA resistance phenotypes was available on Chinese isolates.

INH	Peru	South Africa	China	UK	Russia
Peru		6 E-9	0.1	0.2	9 E-5
South Africa			7 E-29	2 E-32	3 E-7
China				0.2	4 E-17
UK					5 E-15
Russia					

FLQ	Peru	South Africa	China	UK	Russia
Peru		0.3	6 E-10	1 E-6	1 E-4
South Africa			2 E-19	1 E-15	2 E-7
China				0.04	0.6
UK					0.7
Russia					

SLIs	Peru	South Africa	China	UK	Russia
Peru		0.4	1 E-8	2 E-5	0.2
South Africa			5 E-12	9 E-8	0.5
China				0.8	2 E-4
UK					1 E-3
Russia					

RIF	Peru	South Africa	China	UK	Russia
Peru		3 E-9	1 E-14	2 E-8	0.9
South Africa			1 E-31	4 E-21	1 E-3
China				0.02	3 E-17
UK					3 E-11
Russia					

PZA*	Peru	South Africa	China	UK	Russia
Peru		0.01		8 E-9	1 E-4
South Africa				2 E-11	4 E-3
China					
UK					3 E-4
Russia					

EMB	Peru	South Africa	China	UK	Russia
Peru		2 E-9	1 E-7	7 E-6	0.9
South Africa			2 E-17	1 E-12	1 E-6
China				0.4	3 E-7
UK					3 E-6
Russia					

Supplementary Table 4 MDR Frequency Comparison:

Comparison of MDR Frequency between our estimate based on WGS data (WGS MDR Frequency) and estimate based on WHO data (WHO MDR Frequency). Number of isolates with resistance data per country is provided in the "Total Number of Isolates" column.

	WGS MDR Frequency	WHO MDR Frequency	Total Number of Isolates
Belarus	0.03±0.07	0.71(0.67-0.75)	30
Canada	0	0.01(0.01-0.02)	165
China	0.72±0.07	0.08(0.07-0.09)	163
Germany	0.02±0.01	0.03(0.02-0.05)	850
India	0.33±0.15	0.05(0.04-0.06)	39
Iran	0.29±0.19	0.02(0.01-0.02)	21
Malawi	0.01±0.01	0.01(0.01-0.02)	1403
Mali	0.26±0.14	0.03(0.02-0.04)	35
Moldova	0.05±0.07	0.34(0.31-0.36)	40
Netherlands	0	0.02(<0.01-0.01)	93
Not Provided	0.63±0.03		1030
Peru	0.70±0.03	0.09 (0.09-0.1)	941
Romania	1	0.05(0.05-0.06)	33
Russia	0.49±0.03	0.43(0.43-0.44)	854
Sierra Leone	0.10±0.07	0.03(0.02-0.04)	79
South Africa	0.20±0.03	0.04(0.04-0.05)	882
South Korea	0.61±0.15	0.04(0.03-0.04)	41
Swaziland	0.10±0.19	0.1(0.09-0.11)	10
Thailand	0.94±0.12	0.03(0.03-0.04)	16
Turkmenistan	0.10±0.17	0.22(0.19-0.24)	11
Uganda	0.78±0.11	0.02(0.02-0.03)	58
United Kingdom	0.04±0.01	0.02(0.01-0.02)	1850
United States of America	0.97±0.06	0.02(0.01-0.02)	30
Uzbekistan	0.99±0.01	0.33(0.29-0.34)	265

Supplementary Table 5 Dates Drugs Introduced

Drug	Year Introduced
Isoniazid[18]	1952
Rifampicin[18]	1965
Pyrazinamide[19]	1952
Streptomycin[18]	1944
Ethambutol[18]	1965
Ethionamide[18]	1960
Kanamycin/amikacin[18]	1958
Cycloserine[18]	1955
Capreomycin[18]	1967
PAS[18]	1944
Ofloxacin[20]	1985
Moxifloxacin[20]	1999
Ciprofloxacin[20]	1986

Supplementary Table 6 Resistance Mutation Counts Per Country and Per Isolate Phenotype

Amikacin (AMK) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to AMK, susceptible = how many isolates had this mutation but were phenotypically susceptible to AMK, resistant = how many isolates had this mutation but were resistant to AMK

country	AMK_SNP_N_1472359_A514C_rrs	AMK_SNP_N_1473246_A1401G_rrs	AMK_othersnp	AMK_rrs_A906G_u	AMK_rrs_C517T_u	AMK_rrs_G1484T_u	AMK_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	20	51	0	0	0	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	10	14	0	0	0	0	0
Colombia	0	1	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	3	0	0	0	0	0
Germany	3	3	0	0	0	0	1
Guinea	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	1	3	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	1	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	1	0	0	0	0	0	0
Netherlands	1	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	39	160	5	0	8	4	100
Pakistan	0	0	0	0	0	0	0
Peru	14	110	6	3	0	0	48
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	2	12	0	0	0	0	0
Russia	17	35	0	0	14	1	15
Rwanda	0	2	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	83	164	0	0	0	0	11
South Korea	2	26	0	1	0	0	23
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	1	0	0	0	0	0
United Kingdom	16	8	0	0	0	0	0
Uzbekistan	13	46	1	0	0	0	15
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
AMK_SNP_N_1473246_A1401G_rrs	639	208	45	386
AMK_SNP_N_1472359_A514C_rrs	223	84	38	101
AMK_unknown	213	0	0	213
AMK_rrs_C517T_u	22	0	0	22
AMK_rrs_G1484T_u	5	0	0	5
AMK_rrs_A906G_u	4	0	0	4
AMK_rrs_C513T_u	4	0	0	4
AMK_rrs_A1461G_u	1	0	0	1
AMK_rrs_A908C_u	1	0	0	1
AMK_rrs_C1105G_u	1	0	0	1
AMK_rrs_C1402T_u	1	0	0	1
AMK_rrs_C799A_u	1	0	0	1
AMK_rrs_C905A_u	1	0	0	1
AMK_rrs_C905G_u	1	0	0	1
AMK_rrs_G878A_u	1	0	0	1

Capreomycin (CAP) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to CAP, susceptible = how many isolates had this mutation but were phenotypically susceptible to CAP, resistant = how many isolates had this mutation but were resistant to CAP

country	CAP_SNP_N_1472753_A908C_rrs	CAP_SNP_N_1473246_A1401G_rrs	CAP_othersnp	CAP_rrs_A514C_u	CAP_rrs_C517T_u	CAP_rrs_G1484T_u	CAP_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	0	51	0	0	0	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	0	14	0	1	1	0	9
Colombia	0	1	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	3	0	0	0	0	0
Germany	0	3	0	0	0	0	1
Guinea	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	3	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	0	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	2	160	7	1	0	7	53
Pakistan	0	0	0	0	0	0	0
Peru	3	110	13	4	0	0	87
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	1	12	0	0	0	0	0
Russia	1	35	0	0	21	1	20
Rwanda	0	2	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	0	164	0	1	0	0	15
South Korea	0	26	1	0	0	0	9
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	4	1	0	0	0	0	0
United Kingdom	0	8	0	0	0	0	1
Uzbekistan	0	46	2	1	0	0	13
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
CAP_SNP_N_1473246_A1401G_rrs	639	167	127	345
CAP_unknown	208	0	0	208
CAP_rrs_C517T_u	22	0	0	22
CAP_SNP_N_1472753_A908C_rrs	11	4	7	0
CAP_rrs_A514C_u	8	0	0	8
CAP_rrs_G1484T_u	8	0	0	8
CAP_rrs_A906G_u	5	0	0	5
CAP_rrs_C513T_u	5	0	0	5
CAP_rrs_C1402T_u	4	0	0	4
CAP_rrs_C1105G_u	1	0	0	1
CAP_rrs_C1390T_u	1	0	0	1
CAP_rrs_C1402A_u	1	0	0	1
CAP_rrs_C239A_u	1	0	0	1
CAP_rrs_C905A_u	1	0	0	1
CAP_rrs_T1322G_u	1	0	0	1
CAP_rrs_T16A_u	1	0	0	1
CAP_rrs_T16C_u	1	0	0	1
CAP_rrs_T556C_u	1	0	0	1
CAP_SNP_N_1473109_T1264G_rrs	0	0	0	0
CAP_SNP_N_1473160_G1315A_rrs	0	0	0	0
CAP_SNP_N_1473343_G1498T_rrs	0	0	0	0

Ciprofloxacin (CIP) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to CIP, susceptible = how many isolates had this mutation but were phenotypically susceptible to CIP, resistant = how many isolates had this mutation but were resistant to CIP

country	CIP_SNP_CN_7570_CT_gyrA_A90V	CIP_SNP_CN_7572_TC_gyrA_S91P	CIP_SNP_CN_7581_GT_gyrA_D94Y	CIP_SNP_CN_7582_AC_gyrA_D94A	CIP_SNP_CN_7582_AG_gyrA_D94G	CIP_othersnp	CIP_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	23	6	3	11	21	2	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	1	0	0	0
Canada	0	0	0	0	0	0	0
China	11	4	4	7	18	0	0
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	1	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	2	0	1	2	0	0
Guinea	0	0	0	0	0	0	0
India	2	0	1	0	2	0	0
Indonesia	0	0	0	0	0	0	0
Iran	3	0	0	1	6	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	1	0	0	0	0	0	0
Moldova	0	0	0	0	0	1	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	1	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	57	36	13	25	86	11	6
Pakistan	1	0	0	0	0	0	0
Peru	20	1	2	9	29	7	0
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	2	0	1	3	1	0	0
Russia	10	5	3	12	31	0	0
Rwanda	1	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	78	4	6	21	44	1	0
South Korea	13	3	1	2	14	4	22
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	1	0	0	0	0	0
United Kingdom	7	7	0	1	30	2	0
Uzbekistan	0	0	0	0	0	0	0
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
CIP_SNP_CN_7582_AG_gyrA_D94G	286	249	1	36
CIP_SNP_CN_7570_CT_gyrA_A90V	229	201	12	16
CIP_SNP_CN_7582_AC_gyrA_D94A	94	92	2	0
CIP_SNP_CN_7572_TC_gyrA_S91P	69	63	2	4
CIP_SNP_CN_7581_GT_gyrA_D94Y	34	28	0	6
CIP_unknown	28	0	0	28
CIP_SNP_CN_6735_AC_gyrB_N538T	10	10	0	0
CIP_SNP_CN_7566_GA_gyrA_D89N	6	6	0	0
CIP_gyrA_A74S_u	2	0	0	2
CIP_gyrA_D94N_u	2	0	0	2
CIP_gyrB_A471V_u	2	0	0	2
CIP_gyrB_S486F_u	2	0	0	2
CIP_gyrB_T539N_u	2	0	0	2
CIP_gyrA_A288D_u	1	0	0	1
CIP_gyrA_D94H_u	1	0	0	1

Ethambutol (EMB) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to EMB, susceptible = how many isolates had this mutation but were phenotypically susceptible to EMB, resistant = how many isolates had this mutation but were resistant to EMB

country	EMB_SNP_CN_4242182_GT_embC_A774S	EMB_SNP_CN_4247429_AG_embB_M306V	EMB_SNP_CN_4247431_GA_embB_M306I	EMB_SNP_CN_4247431_GC_embB_M306I	EMB_SNP_CN_4247431_GT_embB_M306I	EMB_othersnp	EMB_unknown
Azerbaijan	0	0	0	0	0	1	0
Bangladesh	0	0	0	0	0	1	2
Belarus	0	16	46	46	46	30	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	1	1	1	1	0
Canada	0	0	0	0	0	0	0
China	0	28	17	17	17	45	5
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	2	0	0	0	0	0
Georgia	0	0	0	0	0	4	0
Germany	3	4	4	4	4	12	1
Guinea	0	1	0	0	0	0	0
India	0	4	0	0	0	4	0
Indonesia	0	0	0	0	0	0	0
Iran	0	4	2	2	2	6	1
Kazakhstan	0	2	0	0	0	0	0
Malawi	2	1	0	0	0	5	0
Mali	0	8	4	4	4	3	0
Moldova	0	5	5	5	5	11	0
Morocco	0	0	0	0	0	1	0
Nepal	0	0	0	0	0	0	1
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	1	1	1	0	0
Not Provided	54	203	135	135	135	280	34
Pakistan	0	1	0	0	0	0	0
Peru	220	139	178	178	178	367	121
Philippines	0	0	0	0	0	0	1
Portugal	0	0	1	1	1	0	0
Romania	0	6	11	11	11	17	0
Russia	0	93	46	46	46	232	46
Rwanda	0	1	0	0	0	1	1
Sierra Leone	1	1	4	4	4	7	2
South Africa	0	139	132	132	132	56	6
South Korea	0	11	12	12	12	16	4
Spain	0	1	0	0	0	0	0
Swaziland	0	0	0	0	0	1	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	1	1	1	1	0	0
Uganda	1	7	18	18	18	18	1
United Kingdom	10	21	45	45	45	31	1
Uzbekistan	0	115	56	56	56	60	7
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
EMB_SNP_CN_4247429_AG_embB_M306V	814	123	185	506
EMB_SNP_CN_4247431_GA_embB_M306I	719	122	185	412
EMB_SNP_CN_4247431_GC_embB_M306I	719	122	185	412
EMB_SNP_CN_4247431_GT_embB_M306I	719	122	185	412
EMB_SNP_CN_4242182_GT_embC_A774S	291	33	89	169
EMB_SNP_CN_4247574_AC_embB_D354A	236	6	108	122
EMB_unknown	234	0	0	234
EMB_SNP_CN_4248003_AG_embB_Q497R	164	42	40	82
EMB_SNP_CN_4247730_GC_embB_G406A	129	20	33	76
EMB_SNP_CN_409569_GA_iniB_A70T	117	19	25	73
EMB_SNP_CN_4247729_GA_embB_G406S	71	9	14	48
EMB_SNP_CN_4249583_GA_embB_D1024N	56	7	17	32
EMB_SNP_CN_4243392_AG_embA_N54D	50	8	1	41
EMB_SNP_P_4243222_CA.11_embA.embB	41	11	7	23
EMB_SNP_CN_4247429_AC_embB_M306L	34	7	8	19
EMB_embB_G406D_u	31	0	0	31
EMB_SNP_P_4243225_CT.8_embA.embB	28	4	3	21
EMB_upstream_intergenic-embA-embB_C16G_u	22	0	0	22
EMB_SNP_CN_4249518_AG_embB_H1002R	19	0	4	15
EMB_SNP_CN_4247495_GT_embB_D328Y	17	3	2	12
EMB_upstream_intergenic-embA-embB_C12T_u	17	0	0	17
EMB_embB_D328G_u	12	0	0	12
EMB_embB_L74R_u	12	0	0	12
EMB_embB_Q497K_u	12	0	0	12
EMB_upstream_intergenic-embA-embB_C16T_u	10	0	0	10
EMB_embB_E405D_u	8	0	0	8
EMB_embB_N296H_u	8	0	0	8
EMB_embB_S297A_u	7	0	0	7
EMB_upstream_intergenic-embA-embB_G43C_u	7	0	0	7
EMB_SNP_CN_4247729_GT_embB_G406C	6	3	1	2
EMB_embB_A409P_u	5	0	0	5
EMB_upstream_intergenic-embA-embB_C16A_u	5	0	0	5
EMB_embB_Q445R_u	4	0	0	4
EMB_embB_Q497P_u	4	0	0	4
EMB_embB_Y319C_u	4	0	0	4
EMB_embB_Y319S_u	4	0	0	4
EMB_embB_G603R_u	3	0	0	3
EMB_embB_I450L_u	3	0	0	3
EMB_embB_T1082A_u	3	0	0	3
EMB_embB_T642A_u	3	0	0	3
EMB_embA_D808A_u	2	0	0	2
EMB_embA_I905V_u	2	0	0	2
EMB_embA_L535V_u	2	0	0	2
EMB_embA_S654L_u	2	0	0	2
EMB_embA_V534G_u	2	0	0	2
EMB_embB_E504D_u	2	0	0	2
EMB_embB_S347I_u	2	0	0	2
EMB_embB_V50A_u	2	0	0	2
EMB_embC_M351T_u	2	0	0	2
EMB_iniB_A182V_u	2	0	0	2
EMB_upstream_intergenic-embA-embB_G76C_u	2	0	0	2
EMB_embA_G554D_u	1	0	0	1
EMB_embA_G884D_u	1	0	0	1
EMB_embA_T652R_u	1	0	0	1
EMB_embA_V31I_u	1	0	0	1
EMB_embA_V391L_u	1	0	0	1
EMB_embA_V479M_u	1	0	0	1
EMB_embB_C361Y_u	1	0	0	1
EMB_embB_D1056A_u	1	0	0	1
EMB_embB_D300G_u	1	0	0	1
EMB_embB_F323L_u	1	0	0	1
EMB_embB_I1009L_u	1	0	0	1
EMB_embB_I72L_u	1	0	0	1
EMB_embB_I72S_u	1	0	0	1
EMB_embB_L402V_u	1	0	0	1
EMB_embB_M306T_u	1	0	0	1
EMB_embB_M396I_u	1	0	0	1
EMB_embB_N399T_u	1	0	0	1
EMB_embB_N644I_u	1	0	0	1
EMB_embB_P404S_u	1	0	0	1
EMB_embB_P731L_u	1	0	0	1
EMB_embB_S347T_u	1	0	0	1
EMB_embB_S412P_u	1	0	0	1
EMB_embB_S538P_u	1	0	0	1
EMB_embB_T546I_u	1	0	0	1
EMB_embB_T581A_u	1	0	0	1
EMB_embB_T643I_u	1	0	0	1
EMB_embB_W332R_u	1	0	0	1
EMB_embB_Y319D_u	1	0	0	1
EMB_embC_A68T_u	1	0	0	1
EMB_embC_C411Y_u	1	0	0	1
EMB_embC_I10T_u	1	0	0	1
EMB_iniB_G386D_u	1	0	0	1
EMB_upstream_intergenic-embA-embB_C15G_u	1	0	0	1
EMB_upstream_intergenic-embA-embB_d42CAT_u	1	0	0	1
EMB_upstream_intergenic-embA-embB_d43G_u	1	0	0	1
EMB_upstream_intergenic-embA-embB_i.14-5TACCATCGAG_u	1	0	0	1

Ethionamide (ETH) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to ETH, susceptible = how many isolates had this mutation but were phenotypically susceptible to ETH, resistant = how many isolates had this mutation but were resistant to ETH

country	ETH_SNP_CN_1674263_TC_inhA_I21T	ETH_SNP_CN_1674481_TG_inhA_S94A	ETH_SNP_CN_4326333_CG_ethA_A381P	ETH_SNP_P_1673423_GT.17_fabG1.inhA	ETH_SNP_P_1673425_CT.15_fabG1.inhA	ETH_othersnp	ETH_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	0	0	0	0	30	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	0	1	0	0	18	13	15
Colombia	0	0	0	1	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	0	0	0	1	0	0
Germany	0	3	0	0	21	0	0
Guinea	0	0	0	0	2	0	0
India	0	0	0	0	1	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	0	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	1	11	0	0	15	0	0
Mali	0	1	0	0	2	0	0
Moldova	0	0	0	0	25	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	12	98	12	14	573	51	46
Pakistan	0	0	0	0	1	0	0
Peru	33	20	1	27	178	100	76
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	0	0	0	0	7	0	0
Russia	0	4	0	2	50	0	3
Rwanda	0	0	0	0	1	0	0
Sierra Leone	0	1	0	1	1	0	0
South Africa	8	2	76	37	113	11	6
South Korea	0	2	0	0	8	6	10
Spain	0	1	0	0	1	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	0	0	0	4	0	1
United Kingdom	2	4	1	6	76	1	0
Uzbekistan	1	0	0	1	9	0	1
Vietnam	0	0	0	0	1	0	0

resistance mutation	counts	no_data	susceptible	resistant
ETH_SNP_P_1673425_CT.15_fabG1.inhA	1138	716	157	265
ETH_unknown	158	0	0	158
ETH_SNP_CN_1674481_TG_inhA_S94A	148	52	13	83
ETH_SNP_CN_4326333_CG_ethA_A381P	90	21	35	34
ETH_SNP_P_1673423_GT.17_fabG1.inhA	89	32	38	19
ETH_SNP_CN_1674263_TC_inhA_I21T	57	18	10	29
ETH_SNP_CN_4326116_GA_ethA_T453I	35	6	14	15
ETH_SNP_CN_4327380_AC_ethA_Y32D	15	2	9	4
ETH_upstream_intergenic-fabG1-inhA_T8A_u	14	0	0	14
ETH_DEL_F_4326184_d1289G_ethA	12	5	4	3
ETH_SNP_CN_4327416_CA_ethA_A20S	12	5	4	3
ETH_ethA_P378L_u	9	0	0	9
ETH_INS_F_4326722_i751C_ethA	6	2	1	3
ETH_SNP_CN_4326305_GA_ethA_S390F	6	2	1	3
ETH_SNP_CN_4326713_TG_ethA_Q254P	6	1	0	5
ETH_upstream_intergenic-fabG1-inhA_T8C_u	6	0	0	6
ETH_upstream_intergenic-fabG1-inhA_T8G_u	4	0	0	4
ETH_ethA_S266R_u	3	0	0	3
ETH_ethA_C403R_u	2	0	0	2
ETH_ethA_E400D_u	2	0	0	2
ETH_ethA_F302L_u	2	0	0	2
ETH_ethA_S110W_u	2	0	0	2
ETH_ethA_Y50C_u	2	0	0	2
ETH_SNP_CN_4327311_AG_ethA_S55P	1	0	1	0
ETH_ethA_A76V_u	1	0	0	1
ETH_ethA_F320S_u	1	0	0	1
ETH_ethA_F431V_u	1	0	0	1
ETH_ethA_G139C_u	1	0	0	1
ETH_ethA_G139S_u	1	0	0	1
ETH_ethA_G139V_u	1	0	0	1
ETH_ethA_G182S_u	1	0	0	1
ETH_ethA_G423R_u	1	0	0	1
ETH_ethA_G43C_u	1	0	0	1
ETH_ethA_G450D_u	1	0	0	1
ETH_ethA_H166P_u	1	0	0	1
ETH_ethA_L134P_u	1	0	0	1
ETH_ethA_L194P_u	1	0	0	1
ETH_ethA_L205P_u	1	0	0	1
ETH_ethA_L374R_u	1	0	0	1
ETH_ethA_L397R_u	1	0	0	1
ETH_ethA_N379D_u	1	0	0	1
ETH_ethA_P149S_u	1	0	0	1
ETH_ethA_P257A_u	1	0	0	1
ETH_ethA_P257S_u	1	0	0	1
ETH_ethA_P51H_u	1	0	0	1
ETH_ethA_P51L_u	1	0	0	1
ETH_ethA_P51S_u	1	0	0	1
ETH_ethA_Q206E_u	1	0	0	1
ETH_ethA_Q246R_u	1	0	0	1
ETH_ethA_R38P_u	1	0	0	1
ETH_ethA_R404L_u	1	0	0	1
ETH_ethA_R54S_u	1	0	0	1
ETH_ethA_S18R_u	1	0	0	1
ETH_ethA_S208P_u	1	0	0	1
ETH_ethA_S57F_u	1	0	0	1
ETH_ethA_T189K_u	1	0	0	1
ETH_ethA_T342K_u	1	0	0	1
ETH_ethA_T383P_u	1	0	0	1
ETH_ethA_T392R_u	1	0	0	1
ETH_ethA_T44P_u	1	0	0	1
ETH_ethA_T61M_u	1	0	0	1
ETH_ethA_Y235D_u	1	0	0	1
ETH_fabG1_L85V_u	1	0	0	1
ETH_inhA_I194T_u	1	0	0	1
ETH_inhA_I21V_u	1	0	0	1
ETH_inhA_I95L_u	1	0	0	1
ETH_upstream_intergenic-fabG1-inhA_C34T_u	1	0	0	1
ETH_INS_F_4326141_i1332C_ethA	0	0	0	0
ETH_SNP_CN_1673449_AC_fabG1_T4P	0	0	0	0
ETH_SNP_CN_1674434_TG_inhA_V78G	0	0	0	0
ETH_SNP_CN_4326278_GT_ethA_S399.	0	0	0	0
ETH_SNP_CN_4326600_GA_ethA_R292.	0	0	0	0
ETH_SNP_CN_4326714_GA_ethA_Q254.	0	0	0	0
ETH_SNP_CN_4327148_CT_ethA_W109.	0	0	0	0

Isoniazid (INH) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to INH, susceptible = how many isolates had this mutation but were phenotypically susceptible to INH, resistant = how many isolates had this mutation but were resistant to INH

country	INH_SNP_CN_1674481_TG_inhA_S94A	INH_SNP_CN_2155168_CG_katG_S315T	INH_SNP_CN_2518919_GA_kasA_G269S	INH_SNP_CN_409569_GA_iniB_A70T	INH_SNP_P_1673425_CT.15_fabG1.inhA	INH_othersnp	INH_unknown
Azerbaijan	0	1	0	0	0	0	0
Bangladesh	0	1	0	0	0	0	2
Belarus	0	101	23	0	30	43	1
Brazil	0	0	0	0	0	3	0
Burma	0	1	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	1	65	0	0	18	24	4
Colombia	0	0	0	0	0	1	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	2	0	0	0	0	0
Georgia	0	5	0	0	1	0	0
Germany	3	43	24	2	21	4	4
Guinea	0	0	0	0	2	0	0
India	0	8	0	0	1	1	0
Indonesia	0	0	0	0	0	0	0
Iran	0	11	0	0	0	0	1
Kazakhstan	0	2	0	0	0	0	0
Malawi	11	50	24	0	15	4	12
Mali	1	23	0	0	2	5	0
Moldova	0	43	3	0	25	0	0
Morocco	0	1	0	0	0	0	0
Nepal	0	4	0	0	0	0	0
Netherlands	0	3	0	0	0	0	1
Nigeria	0	1	0	0	0	0	0
Not Provided	98	539	120	27	573	102	82
Pakistan	0	0	0	0	1	0	0
Peru	20	593	299	88	178	80	40
Philippines	0	0	0	0	0	0	2
Portugal	0	1	0	0	0	0	0
Romania	0	30	1	0	7	5	0
Russia	4	470	61	0	50	20	49
Rwanda	0	1	0	0	1	1	2
Sierra Leone	1	16	8	0	1	3	2
South Africa	2	299	98	0	113	57	13
South Korea	2	35	0	0	8	3	6
Spain	1	1	0	0	1	0	0
Swaziland	0	0	0	0	0	1	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	6	0	0	0	0	0
Uganda	0	35	1	0	4	3	2
United Kingdom	4	173	34	0	76	26	13
Uzbekistan	0	226	11	0	9	24	5
Vietnam	0	0	0	0	1	0	0

resistance mutation	counts	no_data	susceptible	resistant
INH_SNP_CN_2155168_CG_katG_S315T	2790	210	310	2270
INH_SNP_P_1673425_CT.15_fabG1.inhA	1138	424	100	614
INH_SNP_CN_2518919_GA_kasA_G269S	707	58	248	401
INH_unknown	241	0	0	241
INH_SNP_CN_1674481_TG_inhA_S94A	148	25	10	113
INH_SNP_CN_409569_GA_iniB_A70T	117	19	9	89
INH_SNP_P_1673432_TC.8_fabG1.inhA	102	33	36	33
INH_SNP_P_1673423_GT.17_fabG1.inhA	89	8	12	69
INH_SNP_CN_2155168_CT_katG_S315N	58	5	5	48
INH_SNP_CN_2155167_GT_katG_S315R	56	3	5	48
INH_SNP_P_1673432_TG.8_fabG1.inhA	10	0	1	9
INH_inhA_I21T_u	3	0	0	3
INH_katG_G279D_u	3	0	0	3
INH_katG_R104Q_u	3	0	0	3
INH_katG_S315G_u	3	0	0	3
INH_katG_W161R_u	3	0	0	3
INH_katG_D735Y_u	2	0	0	2
INH_katG_L141F_u	2	0	0	2
INH_katG_L141S_u	2	0	0	2
INH_katG_N138H_u	2	0	0	2
INH_katG_W191R_u	2	0	0	2
INH_katG_Y337C_u	2	0	0	2
INH_katG_Y413C_u	2	0	0	2
INH_ahpC_P44R_u	1	0	0	1
INH_inhA_K118E_u	1	0	0	1
INH_iniB_A222T_u	1	0	0	1
INH_iniB_G131R_u	1	0	0	1
INH_iniB_G171D_u	1	0	0	1
INH_iniB_G386D_u	1	0	0	1
INH_iniB_S249R_u	1	0	0	1
INH_kasA_A45S_u	1	0	0	1
INH_kasA_R161S_u	1	0	0	1
INH_katG_A109V_u	1	0	0	1
INH_katG_A139P_u	1	0	0	1
INH_katG_A162E_u	1	0	0	1
INH_katG_A614E_u	1	0	0	1
INH_katG_D142G_u	1	0	0	1
INH_katG_D163N_u	1	0	0	1
INH_katG_D189A_u	1	0	0	1
INH_katG_D259Y_u	1	0	0	1
INH_katG_D282G_u	1	0	0	1
INH_katG_D419H_u	1	0	0	1
INH_katG_D487N_u	1	0	0	1
INH_katG_D656A_u	1	0	0	1
INH_katG_D695A_u	1	0	0	1
INH_katG_D94G_u	1	0	0	1
INH_katG_F408L_u	1	0	0	1
INH_katG_G120S_u	1	0	0	1
INH_katG_G182R_u	1	0	0	1
INH_katG_G273S_u	1	0	0	1
INH_katG_G285S_u	1	0	0	1
INH_katG_G297V_u	1	0	0	1
INH_katG_G299S_u	1	0	0	1
INH_katG_G630R_u	1	0	0	1
INH_katG_G699V_u	1	0	0	1
INH_katG_G99R_u	1	0	0	1
INH_katG_K537E_u	1	0	0	1
INH_katG_L159F_u	1	0	0	1
INH_katG_L159P_u	1	0	0	1
INH_katG_L598F_u	1	0	0	1
INH_katG_L76P_u	1	0	0	1
INH_katG_N138D_u	1	0	0	1
INH_katG_P232R_u	1	0	0	1
INH_katG_P325S_u	1	0	0	1
INH_katG_Q461P_u	1	0	0	1
INH_katG_R571H_u	1	0	0	1
INH_katG_S302R_u	1	0	0	1
INH_katG_S315I_u	1	0	0	1
INH_katG_T180K_u	1	0	0	1
INH_katG_T271I_u	1	0	0	1
INH_katG_T475I_u	1	0	0	1
INH_katG_V320L_u	1	0	0	1
INH_katG_V423I_u	1	0	0	1
INH_katG_V442A_u	1	0	0	1
INH_katG_V626E_u	1	0	0	1
INH_katG_V633A_u	1	0	0	1
INH_katG_W191G_u	1	0	0	1
INH_katG_W300C_u	1	0	0	1
INH_katG_W300R_u	1	0	0	1
INH_katG_W300S_u	1	0	0	1
INH_katG_W328L_u	1	0	0	1
INH_katG_W341G_u	1	0	0	1
INH_katG_W438G_u	1	0	0	1
INH_katG_W90R_u	1	0	0	1
INH_katG_Y413H_u	1	0	0	1
INH_katG_Y608F_u	1	0	0	1
INH_katG_Y98C_u	1	0	0	1
INH_upstream_intergenic-fabG1-inhA_G102A_u	1	0	0	1
INH_upstream_intergenic-fabG1-inhA_T8A_u	1	0	0	1
INH_SNP_CN_2726338_TG_ahpC_V49G	0	0	0	0

Kanamycin (KAN) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to KAN, susceptible = how many isolates had this mutation but were phenotypically susceptible to KAN, resistant = how many isolates had this mutation but were resistant to KAN

country	KAN_SNP_N_1473246_A1401G_rrs	KAN_othersnp	KAN_rrs_A514C_u	KAN_rrs_A906G_u	KAN_rrs_C513T_u	KAN_rrs_G1484T_u	KAN_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	51	0	0	0	0	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	14	1	2	0	0	0	6
Colombia	1	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	3	0	0	0	0	0	0
Germany	3	0	0	0	0	0	0
Guinea	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	3	0	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	0	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	160	6	3	0	3	7	53
Pakistan	0	0	0	0	0	0	0
Peru	110	4	1	5	4	0	59
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	12	0	0	0	0	0	2
Russia	35	0	0	0	0	0	0
Rwanda	2	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	164	2	2	0	0	0	17
South Korea	26	0	0	1	0	0	24
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	1	0	0	0	0	0	0
United Kingdom	8	0	0	0	0	0	4
Uzbekistan	46	0	0	0	0	0	0
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
KAN_SNP_N_1473246_A1401G_rrs	639	208	83	348
KAN_unknown	165	0	0	165
KAN_rrs_A514C_u	8	0	0	8
KAN_rrs_C513T_u	7	0	0	7
KAN_rrs_G1484T_u	7	0	0	7
KAN_rrs_A906G_u	6	0	0	6
KAN_rrs_C517T_u	4	0	0	4
KAN_rrs_C905A_u	2	0	0	2
KAN_rrs_A365G_u	1	0	0	1
KAN_rrs_A514T_u	1	0	0	1
KAN_rrs_A908C_u	1	0	0	1
KAN_rrs_A908G_u	1	0	0	1
KAN_rrs_C1105G_u	1	0	0	1
KAN_rrs_C1402A_u	1	0	0	1
KAN_rrs_G878A_u	1	0	0	1
KAN_SNP_CN_1918745_AG_tlyA_.269W	0	0	0	0

Levofloxacin (LEVO) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to LEVO, susceptible = how many isolates had this mutation but were phenotypically susceptible to LEVO, resistant = how many isolates had this mutation but were resistant to LEVO

country	LEVO_SNP_CN_7570_CT_gyrA_A90V	LEVO_SNP_CN_7572_TC_gyrA_S91P	LEVO_SNP_CN_7581_GA_gyrA_D94N	LEVO_SNP_CN_7582_AC_gyrA_D94A	LEVO_SNP_CN_7582_AG_gyrA_D94G	LEVO_othersnp	LEVO_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	23	6	2	11	21	5	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	1	0	0	0
Canada	0	0	0	0	0	0	0
China	11	4	1	7	18	4	0
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	1	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	2	1	1	2	0	0
Guinea	0	0	0	0	0	0	0
India	2	0	0	0	2	1	0
Indonesia	0	0	0	0	0	0	0
Iran	3	0	0	1	6	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	1	0	0	0	0	0	0
Moldova	0	0	0	0	0	1	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	1	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	57	36	12	25	86	19	10
Pakistan	1	0	0	0	0	0	0
Peru	20	1	8	9	29	10	3
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	2	0	0	3	1	1	0
Russia	10	5	7	12	31	4	0
Rwanda	1	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	78	4	6	21	44	12	0
South Korea	13	3	2	2	14	1	0
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	1	0	0	0	0	0
United Kingdom	7	7	2	1	30	1	0
Uzbekistan	0	0	0	0	0	0	0
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
LEVO_SNP_CN_7582_AG_gyrA_D94G	286	270	0	16
LEVO_SNP_CN_7570_CT_gyrA_A90V	229	214	2	13
LEVO_SNP_CN_7582_AC_gyrA_D94A	94	91	0	3
LEVO_SNP_CN_7572_TC_gyrA_S91P	69	66	0	3
LEVO_SNP_CN_7581_GA_gyrA_D94N	41	39	0	2
LEVO_SNP_CN_7581_GT_gyrA_D94Y	34	33	0	1
LEVO_unknown	13	0	0	13
LEVO_SNP_CN_6735_AC_gyrB_N538T	10	8	1	1
LEVO_SNP_CN_7563_GT_gyrA_G88C	8	7	0	1
LEVO_SNP_CN_7566_GA_gyrA_D89N	6	5	1	0
LEVO_gyrA_R128K_u	1	0	0	1

Ofloxacin (OFLX) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to OFLX, susceptible = how many isolates had this mutation but were phenotypically susceptible to OFLX, resistant = how many isolates had this mutation but were resistant to OFLX

country	OFLX_SNP_CN_7570_CT_gyrA_A90V	OFLX_SNP_CN_7581_GA_gyrA_D94N	OFLX_SNP_CN_7582_AC_gyrA_D94A	OFLX_SNP_CN_7582_AG_gyrA_D94G	OFLX_gyrA_S91P_u	OFLX_othersnp	OFLX_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	23	2	11	21	0	3	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	1	0	0	0	0
Canada	0	0	0	0	0	0	0
China	11	1	7	18	3	6	13
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	1	0	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	1	1	2	0	1	0
Guinea	0	0	0	0	0	0	0
India	2	0	0	2	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	3	0	1	6	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	1	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	1	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	57	12	25	86	26	49	97
Pakistan	1	0	0	0	0	0	0
Peru	20	8	9	29	0	3	3
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	2	0	3	1	0	4	2
Russia	10	7	12	31	5	11	57
Rwanda	1	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	78	6	21	44	1	22	16
South Korea	13	2	2	14	0	0	0
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	0	0	0	1	1	0
United Kingdom	7	2	1	30	3	1	0
Uzbekistan	0	0	0	0	0	0	7
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
OFLX_SNP_CN_7582_AG_gyrA_D94G	286	89	30	167
OFLX_SNP_CN_7570_CT_gyrA_A90V	229	62	75	92
OFLX_unknown	195	0	0	195
OFLX_SNP_CN_7582_AC_gyrA_D94A	94	19	23	52
OFLX_SNP_CN_7581_GA_gyrA_D94N	41	14	5	22
OFLX_gyrA_S91P_u	39	0	0	39
OFLX_gyrA_D94Y_u	22	0	0	22
OFLX_SNP_CN_7581_GC_gyrA_D94H	14	3	5	6
OFLX_SNP_CN_6735_AC_gyrB_N538T	10	6	1	3
OFLX_gyrA_G88C_u	4	0	0	4
OFLX_gyrB_D500N_u	4	0	0	4
OFLX_gyrB_V340L_u	4	0	0	4
OFLX_gyrA_Q613E_u	3	0	0	3
OFLX_gyrB_D500H_u	3	0	0	3
OFLX_gyrA_A74S_u	2	0	0	2
OFLX_gyrA_N282K_u	2	0	0	2
OFLX_gyrA_T267I_u	2	0	0	2
OFLX_gyrB_E540D_u	2	0	0	2
OFLX_gyrB_P133L_u	2	0	0	2
OFLX_gyrB_S486F_u	2	0	0	2
OFLX_gyrA_A288D_u	1	0	0	1
OFLX_gyrA_A463S_u	1	0	0	1
OFLX_gyrA_A90G_u	1	0	0	1
OFLX_gyrA_D829E_u	1	0	0	1
OFLX_gyrA_D89N_u	1	0	0	1
OFLX_gyrA_D94V_u	1	0	0	1
OFLX_gyrA_E214D_u	1	0	0	1
OFLX_gyrA_G88A_u	1	0	0	1
OFLX_gyrA_P472S_u	1	0	0	1
OFLX_gyrA_R292G_u	1	0	0	1
OFLX_gyrA_R382L_u	1	0	0	1
OFLX_gyrA_R448H_u	1	0	0	1
OFLX_gyrA_R592S_u	1	0	0	1
OFLX_gyrA_T135S_u	1	0	0	1
OFLX_gyrB_A471V_u	1	0	0	1
OFLX_gyrB_A543T_u	1	0	0	1
OFLX_gyrB_H350Y_u	1	0	0	1
OFLX_gyrB_I39V_u	1	0	0	1
OFLX_gyrB_N538D_u	1	0	0	1
OFLX_gyrB_N538Y_u	1	0	0	1
OFLX_gyrB_P439R_u	1	0	0	1
OFLX_gyrB_R485C_u	1	0	0	1
OFLX_gyrB_T539N_u	1	0	0	1
OFLX_gyrB_V496L_u	1	0	0	1
OFLX_gyrB_V600A_u	1	0	0	1

Para-aminosalicylic acid (PAS) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to PAS, susceptible = how many isolates had this mutation but were phenotypically susceptible to PAS, resistant = how many isolates had this mutation but were resistant to PAS

country	PAS_SNP_CN_3073852_TC_thyA_H207R	PAS_SNP_CN_3074182_TC_thyA_Q97R	PAS_SNP_CN_3074449_AT_thyA_L8Q	PAS_inter-thyX-hsdS.1_G228A_u	PAS_othersnp	PAS_thyA_T202A_u	PAS_unknown
Azerbaijan	0	0	0	0	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	0	0	0	0	0	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	1	0	0	0	0	0	0
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	0	0	0	0	0	0
Guinea	0	0	0	0	0	0	0
India	0	0	0	0	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	0	0	0	0	0	0
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	0	0	0
Mali	0	0	0	0	0	0	0
Moldova	0	0	0	0	0	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	5	2	4	3	15	18	5
Pakistan	0	0	0	0	0	0	0
Peru	10	0	4	0	1	0	0
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	0	0	0	0	0	0	0
Russia	0	1	0	0	1	0	0
Rwanda	0	0	0	0	0	0	0
Sierra Leone	0	0	0	0	0	0	0
South Africa	0	0	0	1	2	0	3
South Korea	0	0	0	6	6	0	3
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	0	0
Uganda	0	0	0	0	0	0	0
United Kingdom	0	1	0	0	0	0	0
Uzbekistan	0	0	0	0	0	0	0
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
PAS_thyA_T202A_u	18	0	0	18
PAS_SNP_CN_3073852_TC_thyA_H207R	16	14	0	2
PAS_unknown	11	0	0	11
PAS_upstream_intergenic-thyX-hsdS.1_G228A_u	10	0	0	10
PAS_SNP_CN_3074449_AT_thyA_L8Q	8	7	0	1
PAS_SNP_CN_3074182_TC_thyA_Q97R	4	3	0	1
PAS_folC_R49Q_u	2	0	0	2
PAS_folC_R49W_u	2	0	0	2
PAS_upstream_intergenic-thyX-hsdS.1_C235T_u	2	0	0	2
PAS_upstream_intergenic-thyA_G24A_u	2	0	0	2
PAS_folC_E153G_u	1	0	0	1
PAS_folC_I43S_u	1	0	0	1
PAS_folC_I43T_u	1	0	0	1
PAS_folC_I43V_u	1	0	0	1
PAS_folC_S98G_u	1	0	0	1
PAS_upstream_intergenic-thyX-hsdS.1_C226A_u	1	0	0	1
PAS_upstream_intergenic-thyX-hsdS.1_G240A_u	1	0	0	1
PAS_upstream_intergenicr-thyA_T117C_u	1	0	0	1
PAS_upstream_intergenic-thyA_d49CCGCAGCGACTCGCCGCCAAACAAACCCAGCGGGCGATCGCAAGCGCGGCGAAGCCG_u	1	0	0	1
PAS_thyA_A262V_u	1	0	0	1
PAS_thyA_D6G_u	1	0	0	1
PAS_thyA_D81A_u	1	0	0	1
PAS_thyA_E137G_u	1	0	0	1
PAS_thyA_E58V_u	1	0	0	1
PAS_thyA_F152V_u	1	0	0	1
PAS_thyA_P145L_u	1	0	0	1
PAS_thyA_T22P_u	1	0	0	1
PAS_SNP_P_3074479_AG.157_thyA	0	0	0	0

Pyrazinamide (PZA) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to PZA, susceptible = how many isolates had this mutation but were phenotypically susceptible to PZA, resistant = how many isolates had this mutation but were resistant to PZA

country	PZA_SNP_CN_2289090_TC_pncA_H51R	PZA_SNP_CN_2289099_TG_pncA_K48T	PZA_SNP_CN_2289213_TC_pncA_Q10R	PZA_SNP_CN_2289213_TG_pncA_Q10P	PZA_othersnp	PZA_promoter-pncA_T11C_u	PZA_unknown
Azerbaijan	0	0	0	1	0	0	0
Bangladesh	0	0	0	0	0	0	0
Belarus	0	0	2	0	56	0	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	0	1	0	0
Canada	0	0	0	0	0	0	0
China	0	0	0	2	22	0	0
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	0	0	0	0
Georgia	0	0	0	0	0	0	0
Germany	0	0	0	0	14	0	5
Guinea	0	0	0	0	1	0	0
India	0	0	0	0	2	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	0	0	0	4	0	1
Kazakhstan	0	0	0	0	0	0	0
Malawi	0	0	0	0	4	0	0
Mali	0	0	0	0	3	0	0
Moldova	0	0	0	2	6	0	0
Morocco	0	0	0	0	0	0	0
Nepal	0	0	0	0	0	0	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	0	0	0	0
Not Provided	10	12	6	6	260	15	78
Pakistan	0	0	0	0	0	0	0
Peru	79	35	45	9	296	8	150
Philippines	0	0	0	0	0	0	0
Portugal	0	0	0	0	0	0	0
Romania	0	0	0	0	20	0	0
Russia	2	0	0	1	121	6	71
Rwanda	0	0	0	0	0	0	1
Sierra Leone	0	0	0	0	4	1	3
South Africa	0	0	0	3	112	0	66
South Korea	0	0	0	0	23	0	24
Spain	0	0	0	0	0	0	0
Swaziland	0	0	0	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	0	0	1	2
Uganda	0	0	0	0	12	1	10
United Kingdom	0	0	0	0	41	2	19
Uzbekistan	1	0	2	39	69	17	40
Vietnam	0	0	0	0	0	0	0

resistance mutation	count	no data	susceptible	resistant
PZA_unknown	470	0	0	470
PZA_SNP_CN_2289090_TC_pncA_H51R	92	12	9	71
PZA_SNP_CN_2289213_TG_pncA_Q10P	63	11	4	48
PZA_SNP_CN_2289213_TC_pncA_Q10R	55	4	5	46
PZA_upstream_intergenic-pncA_T11C_u	51	0	0	51
PZA_SNP_CN_2289099_TG_pncA_K48T	47	9	22	16
PZA_SNP_CN_2289202_AG_pncA_C14R	46	12	0	34
PZA_SNP_CN_2288844_AG_pncA_I133T	40	1	16	23
PZA_SNP_CN_2289016_TG_pncA_T76P	38	23	2	13
PZA_SNP_CN_2288883_AG_pncA_L120P	35	15	0	20
PZA_SNP_CN_2288839_TG_pncA_T135P	34	12	2	20
PZA_SNP_CN_2289040_AC_pncA_W68G	31	22	5	4
PZA_SNP_CN_2289073_GC_pncA_H57D	26	5	0	21
PZA_SNP_CN_2289231_AG_pncA_L4S	25	8	0	17
PZA_SNP_CN_2288826_AG_pncA_V139A	23	5	7	11
PZA_SNP_CN_2289142_AC_pncA_Y34D	22	19	1	2
PZA_SNP_CN_2288820_TG_pncA_Q141P	20	7	2	11
PZA_SNP_CN_2289081_GA_pncA_P54L	20	4	4	12
PZA_SNP_CN_2289180_AC_pncA_V21G	20	5	4	11
PZA_pncA_V125G_u	20	0	0	20
PZA_SNP_CN_2289207_TG_pncA_D12A	19	6	4	9
PZA_SNP_CN_2288727_AG_pncA_L172P	17	11	2	4
PZA_SNP_CN_2289097_CT_pncA_D49N	16	3	1	12
PZA_SNP_CN_2288952_CT_pncA_G97D	15	9	0	6
PZA_SNP_CN_2288953_CT_pncA_G97S	15	3	1	11
PZA_SNP_CN_2289040_AG_pncA_W68R	15	4	2	9
PZA_DEL_F_2288939_d302TCCGGTGTAG_pncA	13	9	0	4
PZA_SNP_CN_2288805_GA_pncA_A146V	13	9	0	4
PZA_SNP_CN_2289054_TG_pncA_D63A	12	2	3	7
PZA_SNP_CN_2289070_AG_pncA_F58L	12	4	3	5
PZA_SNP_CN_2288805_GT_pncA_A146E	11	0	4	7
PZA_SNP_CN_2289207_TC_pncA_D12G	11	1	2	8
PZA_pncA_I6L_u	11	0	0	11
PZA_DEL_F_2289069_d172A_pncA_F58L	10	6	0	4
PZA_SNP_CN_2288703_AC_pncA_V180G	10	2	0	8
PZA_SNP_CN_2288764_TC_pncA_T160A	10	3	0	7
PZA_SNP_CN_2288827_CT_pncA_V139M	10	3	1	6
PZA_SNP_CN_2289220_CT_pncA_D8N	10	6	0	4
PZA_SNP_CN_2288704_CA_pncA_V180F	9	0	2	7
PZA_INS_F_2289009_i232C_pncA_G78G	8	3	1	4
PZA_INS_F_2289050_i191T_pncA_Y64.	8	3	1	4
PZA_SNP_CN_2288850_AC_pncA_V131G	8	3	1	4
PZA_SNP_CN_2288859_AC_pncA_V128G	8	1	2	5
PZA_upstream_intergenic-pncA_T11G_u	8	0	0	8
PZA_SNP_CN_2288778_AC_pncA_V155G	7	3	1	3
PZA_SNP_CN_2288943_GA_pncA_T100I	7	0	1	6
PZA_SNP_CN_2288988_AG_pncA_L85P	7	2	1	4
PZA_SNP_CN_2289043_AG_pncA_S67P	7	1	0	6
PZA_SNP_CN_2289072_TC_pncA_H57R	7	2	0	5
PZA_SNP_CN_2288696_CA_pncA_L182F	6	0	5	1
PZA_SNP_CN_2288835_TC_pncA_D136G	6	3	1	2
PZA_SNP_CN_2288965_CA_pncA_V93L	6	0	1	5
PZA_SNP_CN_2289072_TA_pncA_H57L	6	0	1	5
PZA_SNP_CN_2289103_TC_pncA_T47A	6	2	3	1
PZA_DEL_F_2288923_d318C_pncA	5	2	1	2
PZA_INS_F_2288825_i416C_pncA	5	1	0	4
PZA_INS_F_2288851_i390CC_pncA	5	1	1	3
PZA_INS_F_2288942_i299T_pncA	5	2	1	2
PZA_SNP_CN_2288697_AG_pncA_L182S	5	2	2	1
PZA_SNP_CN_2288718_AG_pncA_M175T	5	2	0	3
PZA_SNP_CN_2288938_CG_pncA_A102P	5	2	2	1
PZA_SNP_CN_2288944_TG_pncA_T100P	5	1	0	4
PZA_SNP_CN_2288956_TC_pncA_K96E	5	1	0	4
PZA_SNP_CN_2288973_AG_pncA_I90T	5	1	1	3
PZA_SNP_CN_2289028_AG_pncA_C72R	5	2	1	2
PZA_SNP_CN_2289030_TC_pncA_H71R	5	0	3	2
PZA_SNP_CN_2289162_AG_pncA_L27P	5	1	2	2
PZA_SNP_CN_2289219_TC_pncA_D8G	5	2	1	2
PZA_pncA_C138R_u	5	0	0	5
PZA_pncA_D8E_u	5	0	0	5
PZA_pncA_S164P_u	5	0	0	5
PZA_INS_F_2288851_i390C_pncA	4	2	0	2
PZA_INS_F_2288887_i354A_pncA	4	2	0	2
PZA_SNP_CN_2288766_AC_pncA_L159R	4	0	0	4
PZA_SNP_CN_2288818_TC_pncA_T142A	4	2	0	2
PZA_SNP_CN_2288841_GA_pncA_A134V	4	2	0	2
PZA_SNP_CN_2288955_TG_pncA_K96T	4	1	0	3
PZA_SNP_CN_2289037_GA_pncA_P69S	4	0	4	0
PZA_pncA_F13L_u	4	0	0	4
PZA_pncA_M1T_u	4	0	0	4
PZA_pncA_R140P_u	4	0	0	4
PZA_pncA_V7G_u	4	0	0	4
PZA_DEL_F_2288776_d465GCACCCTG_pncA	3	1	1	1
PZA_INS_F_2288725_i516C_pncA	3	1	0	2
PZA_SNP_CN_2288775_AG_pncA_L156P	3	2	0	1
PZA_SNP_CN_2288883_AC_pncA_L120R	3	1	0	2
PZA_SNP_CN_2289001_AC_pncA_F81V	3	0	2	1
PZA_SNP_CN_2289015_GA_pncA_T76I	3	1	1	1
PZA_SNP_CN_2289054_TC_pncA_D63G	3	0	1	2
PZA_SNP_CN_2289073_GA_pncA_H57Y	3	0	0	3
PZA_pncA_C14G_u	3	0	0	3
PZA_pncA_D136Y_u	3	0	0	3
PZA_pncA_H51Q_u	3	0	0	3
PZA_pncA_H82R_u	3	0	0	3
PZA_pncA_L116R_u	3	0	0	3
PZA_pncA_L4W_u	3	0	0	3
PZA_pncA_P54Q_u	3	0	0	3
PZA_pncA_P62S_u	3	0	0	3
PZA_upstream_intergenic-pncA_A12G_u	3	0	0	3

PZA_SNP_CN_2288718_AC_pncA_M175R	2	1	0	1
PZA_SNP_CN_2288742_GA_pncA_T167I	2	1	1	1
PZA_SNP_CN_2288818_TG_pncA_T142P	2	1	0	1
PZA_SNP_CN_2288826_AC_pncA_V139G	2	0	0	2
PZA_SNP_CN_2288869_CA_pncA_V125F	2	0	1	1
PZA_SNP_CN_2288887_AC_pncA_W119G	2	0	0	2
PZA_SNP_CN_2288919_CT_pncA_G108E	2	0	1	1
PZA_SNP_CN_2288925_AG_pncA_F106S	2	1	0	1
PZA_SNP_CN_2288964_AC_pncA_V93G	2	0	1	1
PZA_SNP_CN_2289009_CA_pncA_G78V	2	0	0	2
PZA_SNP_CN_2289030_TG_pncA_H71P	2	1	0	1
PZA_SNP_CN_2289046_AG_pncA_S66P	2	1	0	1
PZA_SNP_CN_2289090_TG_pncA_H51P	2	0	1	1
PZA_SNP_CN_2289091_GA_pncA_H51Y	2	2	0	0
PZA_SNP_CN_2289095_GC_pncA_D49E	2	1	0	1
PZA_SNP_CN_2289212_CG_pncA_Q10H	2	0	0	2
PZA_SNP_CN_2289216_AC_pncA_V9G	2	1	0	1
PZA_SNP_CN_2289228_AG_pncA_I5T	2	0	0	2
PZA_pncA_A46V_u	2	0	0	2
PZA_pncA_D63H_u	2	0	0	2
PZA_pncA_D8A_u	2	0	0	2
PZA_pncA_F94L_u	2	0	0	2
PZA_pncA_G97C_u	2	0	0	2
PZA_pncA_H71Q_u	2	0	0	2
PZA_pncA_H71Y_u	2	0	0	2
PZA_pncA_H82D_u	2	0	0	2
PZA_pncA_I6M_u	2	0	0	2
PZA_pncA_K48E_u	2	0	0	2
PZA_pncA_L35R_u	2	0	0	2
PZA_pncA_M175V_u	2	0	0	2
PZA_pncA_P62L_u	2	0	0	2
PZA_pncA_P69R_u	2	0	0	2
PZA_pncA_S104R_u	2	0	0	2
PZA_pncA_V130M_u	2	0	0	2
PZA_pncA_V155M_u	2	0	0	2
PZA_pncA_V44G_u	2	0	0	2
PZA_pncA_Y103C_u	2	0	0	2
PZA_pncA_Y95S_u	2	0	0	2
PZA_upstream_intergenic-pncA_A7G_u	2	0	0	2
PZA_INS_F_2288835_i406T_pncA	1	0	0	1
PZA_SNP_CN_2288697_AC_pncA_L182W	1	1	0	0
PZA_SNP_CN_2288730_GA_pncA_A171V	1	1	0	0
PZA_SNP_CN_2288817_GA_pncA_T142M	1	1	0	0
PZA_SNP_CN_2288847_CT_pncA_G132D	1	0	0	1
PZA_SNP_CN_2288935_AG_pncA_Y103H	1	0	0	1
PZA_SNP_CN_2289042_GC_pncA_S67W	1	0	0	1
PZA_SNP_CN_2289069_AC_pncA_F58C	1	0	0	1
PZA_SNP_CN_2289186_AG_pncA_L19P	1	0	0	1
PZA_pncA_A102T_u	1	0	0	1
PZA_pncA_A146P_u	1	0	0	1
PZA_pncA_A171E_u	1	0	0	1
PZA_pncA_A171T_u	1	0	0	1
PZA_pncA_A3E_u	1	0	0	1
PZA_pncA_A46E_u	1	0	0	1
PZA_pncA_A46P_u	1	0	0	1
PZA_pncA_C14W_u	1	0	0	1
PZA_pncA_D129Y_u	1	0	0	1
PZA_pncA_D12N_u	1	0	0	1
PZA_pncA_D136N_u	1	0	0	1
PZA_pncA_D49G_u	1	0	0	1
PZA_pncA_F13I_u	1	0	0	1
PZA_pncA_F58S_u	1	0	0	1
PZA_pncA_F81C_u	1	0	0	1
PZA_pncA_G105D_u	1	0	0	1
PZA_pncA_G108R_u	1	0	0	1
PZA_pncA_G132A_u	1	0	0	1
PZA_pncA_G162D_u	1	0	0	1
PZA_pncA_G17D_u	1	0	0	1
PZA_pncA_G24D_u	1	0	0	1
PZA_pncA_G78C_u	1	0	0	1
PZA_pncA_G78S_u	1	0	0	1
PZA_pncA_G97R_u	1	0	0	1
PZA_pncA_H43P_u	1	0	0	1
PZA_pncA_I133S_u	1	0	0	1
PZA_pncA_I6T_u	1	0	0	1
PZA_pncA_K96R_u	1	0	0	1
PZA_pncA_L151S_u	1	0	0	1
PZA_pncA_L156Q_u	1	0	0	1
PZA_pncA_L159V_u	1	0	0	1
PZA_pncA_L85R_u	1	0	0	1
PZA_pncA_M175I_u	1	0	0	1
PZA_pncA_M175K_u	1	0	0	1
PZA_pncA_M1I_u	1	0	0	1
PZA_pncA_P62R_u	1	0	0	1
PZA_pncA_R154G_u	1	0	0	1
PZA_pncA_R2W_u	1	0	0	1
PZA_pncA_S104G_u	1	0	0	1
PZA_pncA_S32I_u	1	0	0	1
PZA_pncA_T100A_u	1	0	0	1
PZA_pncA_T114P_u	1	0	0	1
PZA_pncA_T160K_u	1	0	0	1
PZA_pncA_T160P_u	1	0	0	1
PZA_pncA_T168P_u	1	0	0	1
PZA_pncA_V139L_u	1	0	0	1
PZA_pncA_V155A_u	1	0	0	1
PZA_pncA_V163A_u	1	0	0	1
PZA_pncA_V7F_u	1	0	0	1
PZA_pncA_V7L_u	1	0	0	1
PZA_pncA_W119R_u	1	0	0	1
PZA_pncA_W68C_u	1	0	0	1
PZA_pncA_W68L_u	1	0	0	1
PZA_pncA_Y103D_u	1	0	0	1
PZA_pncA_Y64D_u	1	0	0	1
PZA_upstream_intergenic-pncA_T11A_u	1	0	0	1
PZA_upstream_intergenic-pncA_T15C_u	1	0	0	1
PZA_upstream_intergenic-pncA_T16C_u	1	0	0	1
PZA_upstream_intergenic-pncA_d10A_u	1	0	0	1
PZA_upstream_intergenic-pncA_i14T_u	1	0	0	1
PZA_DEL_F_2288697_d544AACT_pncA	0	0	0	0
PZA_DEL_F_2289060_d181GTGCCGGA_pncA	0	0	0	0
PZA_DEL_N_2288942_d299GGTGTA_pncA	0	0	0	0
PZA_SNP_CN_2288784_GT_pncA_T153N	0	0	0	0
PZA_SNP_CN_2288848_CT_pncA_G132S	0	0	0	0
PZA_SNP_CN_2288853_AC_pncA_V130G	0	0	0	0
PZA_SNP_CN_2288853_AT_pncA_V130E	0	0	0	0
PZA_SNP_CN_2288878_GA_pncA_Q122.	0	0	0	0
PZA_SNP_CN_2288933_GC_pncA_Y103.	0	0	0	0
PZA_SNP_CN_2288956_TG_pncA_K96Q	0	0	0	0
PZA_SNP_CN_2289042_GT_pncA_S67.	0	0	0	0
PZA_SNP_CN_2289050_AT_pncA_Y64.	0	0	0	0
PZA_SNP_CN_2289150_AC_pncA_I31S	0	0	0	0
PZA_SNP_CN_2289206_GC_pncA_D12E	0	0	0	0
PZA_SNP_CN_2289214_GA_pncA_Q10.	0	0	0	0
PZA_SNP_P_2289245_TA.37_pncA	0	0	0	0
PZA_SNP_P_2289251_AC.31_pncA	0	0	0	0
PZA_SNP_P_2289252_TC.30_pncA	0	0	0	0
PZA_SNP_P_2289252_TG.30_pncA	0	0	0	0

Rifamycin (RIF) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to RIF, susceptible = how many isolates had this mutation but were phenotypically susceptible to RIF, resistant = how many isolates had this mutation but were resistant to RIF

country	RIF_SNP_CN_761109_GT_rpoB_D435Y	RIF_SNP_CN_761110_AT_rpoB_D435V	RIF_SNP_CN_761139_CT_rpoB_H445Y	RIF_SNP_CN_761155_CT_rpoB_S450L	RIF_SNP_CN_761161_TC_rpoB_L452P	RIF_othersnp	RIF_unknown
Azerbaijan	0	0	0	1	0	0	0
Bangladesh	0	0	0	0	0	0	1
Belarus	3	0	1	77	0	20	0
Brazil	0	0	0	0	0	0	0
Burma	0	0	0	1	0	0	0
Canada	0	0	0	0	0	0	0
China	3	4	11	57	9	36	5
Colombia	0	0	0	0	0	0	1
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	0	0	2	0	0	0
Georgia	0	0	0	4	1	0	0
Germany	0	1	1	13	1	5	1
Guinea	0	0	0	1	0	0	0
India	0	2	0	8	0	0	0
Indonesia	0	0	0	0	0	0	0
Iran	0	0	0	10	0	2	1
Kazakhstan	0	0	0	2	0	0	0
Malawi	0	1	0	4	0	1	1
Mali	0	8	0	4	3	5	0
Moldova	1	0	0	34	0	4	0
Morocco	0	0	0	0	0	1	0
Nepal	0	0	0	1	0	1	0
Netherlands	0	0	0	0	0	0	0
Nigeria	0	0	0	1	0	0	0
Not Provided	42	73	49	515	18	123	57
Pakistan	0	0	0	1	0	0	0
Peru	22	134	28	491	10	110	57
Philippines	0	0	0	1	0	0	1
Portugal	0	0	0	1	0	0	0
Romania	6	4	2	19	0	3	0
Russia	6	7	3	350	6	28	74
Rwanda	0	1	0	3	0	0	1
Sierra Leone	2	0	3	6	1	2	1
South Africa	10	55	15	188	62	36	8
South Korea	4	3	2	35	1	1	3
Spain	0	0	0	1	1	0	0
Swaziland	0	0	0	2	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	0	0	1	0	0	0
Uganda	1	3	2	25	0	14	4
United Kingdom	3	5	7	54	6	65	3
Uzbekistan	3	8	7	213	4	18	13
Vietnam	0	0	0	0	0	0	0

resistance mutation	counts	no_data	susceptible	resistant
RIF_SNP_CN_761155_CT_rpoB_S450L	2126	232	228	1666
RIF_SNP_CN_761110_AT_rpoB_D435V	309	19	27	263
RIF_unknown	232	0	0	232
RIF_SNP_CN_761139_CT_rpoB_H445Y	131	16	10	105
RIF_SNP_CN_761161_TC_rpoB_L452P	123	5	75	43
RIF_SNP_CN_761109_GT_rpoB_D435Y	106	18	22	66
RIF_SNP_CN_761139_CG_rpoB_H445D	99	21	10	68
RIF_SNP_CN_761140_AG_rpoB_H445R	59	9	3	47
RIF_SNP_CN_761155_CG_rpoB_S450W	50	3	15	32
RIF_SNP_CN_761140_AT_rpoB_H445L	40	7	4	29
RIF_SNP_CN_761095_TC_rpoB_L430P	37	3	20	14
RIF_SNP_CN_761140_AT_rpoB_I491F	36	8	18	10
RIF_SNP_CN_761277_AT_rpoB_I491F	36	8	18	10
RIF_SNP_CN_760314_GT_rpoB_V170F	29	13	1	15
RIF_rpoB_S450P_u	9	0	0	9
RIF_rpoB_Q432P_u	7	0	0	7
RIF_rpoB_S441L_u	7	0	0	7
RIF_rpoB_D435G_u	6	0	0	6
RIF_rpoB_Q432K_u	6	0	0	6
RIF_rpoB_R448Q_u	6	0	0	6
RIF_rpoB_E250G_u	4	0	0	4
RIF_rpoB_H445N_u	4	0	0	4
RIF_rpoB_V695L_u	4	0	0	4
RIF_rpoB_P454L_u	3	0	0	3
RIF_rpoB_Q432L_u	3	0	0	3
RIF_rpoB_D435N_u	2	0	0	2
RIF_rpoB_H445P_u	2	0	0	2
RIF_rpoB_L494P_u	2	0	0	2
RIF_rpoB_M434I_u	2	0	0	2
RIF_rpoB_Q432E_u	2	0	0	2
RIF_rpoB_A451G_u	1	0	0	1
RIF_rpoB_D435A_u	1	0	0	1
RIF_rpoB_G675D_u	1	0	0	1
RIF_rpoB_G981D_u	1	0	0	1
RIF_rpoB_H835Q_u	1	0	0	1
RIF_rpoB_H835R_u	1	0	0	1
RIF_rpoB_I925V_u	1	0	0	1
RIF_rpoB_K446Q_u	1	0	0	1
RIF_rpoB_L430R_u	1	0	0	1
RIF_rpoB_L443W_u	1	0	0	1
RIF_rpoB_M707T_u	1	0	0	1
RIF_rpoB_P280L_u	1	0	0	1
RIF_rpoB_P439S_u	1	0	0	1
RIF_rpoB_S428R_u	1	0	0	1
RIF_rpoB_S441T_u	1	0	0	1
RIF_rpoB_T400A_u	1	0	0	1
RIF_rpoB_T400P_u	1	0	0	1
RIF_rpoB_T444P_u	1	0	0	1
RIF_rpoB_V113I_u	1	0	0	1
RIF_rpoB_V469L_u	1	0	0	1
RIF_SNP_CN_761102_AC_rpoB_Q432H	0	0	0	0
RIF_SNP_CN_761155_CA_rpoB_S450.	0	0	0	0

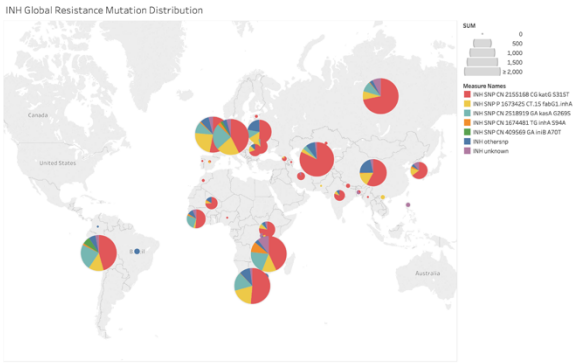
Streptomycin (STR) Data, Abbreviations: counts = how many isolates had this mutation, no_data = how many isolates had this mutation but no data on whether they were phenotypically resistant or not to RIF, susceptible = how many isolates had this mutation but were phenotypically susceptible to STR, resistant = how many isolates had this mutation but were resistant to STR

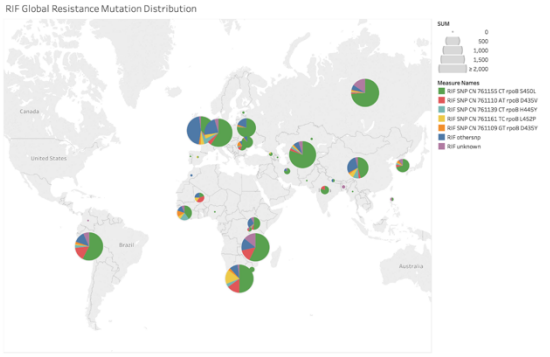
country	STR_SNP_CN_4407967_AG_gid_L79S	STR_SNP_CN_781687_AG_rpsL_K43R	STR_SNP_CN_781822_AG_rpsL_K88R	STR_SNP_N_1472359_A514C_rrs	STR_SNP_N_1472362_C517T_rrs	STR_othersnp	STR_unknown
Azerbaijan	0	0	0	0	1	0	0
Bangladesh	0	0	1	0	0	0	1
Belarus	0	63	6	20	10	1	0
Brazil	0	1	0	0	0	0	1
Burma	0	1	0	0	0	0	0
Canada	0	0	0	0	0	0	0
China	0	46	16	10	2	4	7
Colombia	0	0	0	0	0	0	0
Democratic Republic of the Congo	0	0	0	0	0	0	0
Denmark	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0
Dominican Republic	0	0	0	0	0	0	0
Estonia	0	2	0	0	0	0	0
Georgia	0	5	0	0	0	0	0
Germany	3	29	3	3	2	38	11
Guinea	0	0	0	0	1	0	0
India	0	9	1	0	1	2	2
Indonesia	0	0	0	0	0	0	0
Iran	0	8	0	1	3	0	1
Kazakhstan	0	2	0	0	0	0	0
Malawi	0	4	6	0	0	18	3
Mali	17	0	1	1	0	1	1
Moldova	0	13	27	0	1	2	0
Morocco	0	0	0	0	0	0	1
Nepal	0	2	0	1	0	0	0
Netherlands	0	9	1	1	7	0	0
Nigeria	0	0	0	0	0	0	1
Not Provided	15	239	29	39	20	244	50
Pakistan	0	0	0	0	1	0	0
Peru	2	71	8	14	7	340	96
Philippines	0	0	0	0	0	0	1
Portugal	0	1	0	0	0	0	0
Romania	0	9	2	2	0	6	5
Russia	0	221	43	17	184	19	47
Rwanda	0	0	0	0	0	0	1
Sierra Leone	0	9	7	0	0	13	4
South Africa	74	72	15	83	7	9	15
South Korea	0	12	1	2	0	3	3
Spain	0	0	0	0	0	0	1
Swaziland	0	0	1	0	0	0	0
Thailand	0	0	0	0	0	0	0
Turkmenistan	0	5	1	0	0	0	0
Uganda	0	2	3	0	2	13	9
United Kingdom	5	53	10	16	5	55	5
Uzbekistan	0	212	12	13	3	13	5
Vietnam	0	2	0	0	0	0	0

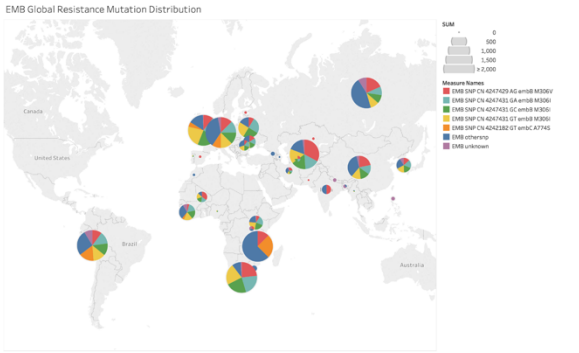
resistance mutation	counts	no_data	susceptible	resistant
STR_SNP_CN_781687_AG_rpsL_K43R	1102	172	112	818
STR_unknown	271	0	0	271
STR_SNP_N_1472362_C517T_rrs	257	14	54	189
STR_SNP_N_1472359_A514C_rrs	223	54	25	144
STR_SNP_CN_781822_AG_rpsL_K88R	194	22	44	128
STR_SNP_CN_4407967_AG_gid_L79S	116	27	13	76
STR_SNP_CN_4407967_AC_gid_L79W	73	15	26	32
STR_SNP_CN_781822_AC_rpsL_K88T	55	8	6	41
STR_SNP_CN_4407934_AC_gid_L90R	35	6	10	19
STR_SNP_CN_4407768_CA_gid_L145F	34	3	22	9
STR_SNP_N_1472358_C513T_rrs	34	6	2	26
STR_SNP_N_1472751_A906G_rrs	29	3	8	18
STR_DEL_F_4407640_d562A_gid	28	0	17	11
STR_SNP_CN_4408094_CT_gid_G37R	28	0	17	11
STR_DEL_F_4407852_d350C_gid	27	19	5	3
STR_SNP_CN_4407995_TG_gid_S70R	27	19	5	3
STR_SNP_CN_4407985_CG_gid_G73A	23	0	8	15
STR_SNP_CN_4407809_CA_gid_D132Y	21	5	1	15
STR_SNP_CN_4407916_CA_gid_R96L	16	3	5	8
STR_gid_L50R_u	14	0	0	14
STR_SNP_N_1472753_A908C_rrs	11	0	3	8
STR_SNP_CN_4407748_AG_gid_L152S	10	1	5	4
STR_gid_P84L_u	10	0	0	10
STR_SNP_CN_4408064_GA_gid_R47W	9	2	5	2
STR_rrs_A514T_u	8	0	0	8
STR_SNP_CN_4407832_AG_gid_V124A	7	0	2	5
STR_SNP_CN_781822_AT_rpsL_K88M	7	1	2	4
STR_gid_G76C_u	7	0	0	7
STR_gid_A138V_u	6	0	0	6
STR_gid_D67G_u	6	0	0	6
STR_gid_L26F_u	6	0	0	6
STR_SNP_CN_4408148_CG_gid_A19P	5	0	2	3
STR_gid_H48Y_u	5	0	0	5
STR_gid_P75R_u	5	0	0	5
STR_gid_R137P_u	5	0	0	5
STR_gid_V77G_u	5	0	0	5
STR_SNP_CN_4408102_CT_gid_G34E	4	2	0	2
STR_gid_A10P_u	4	0	0	4
STR_gid_S149R_u	4	0	0	4
STR_SNP_CN_4407947_GA_gid_L86F	3	0	3	0
STR_SNP_CN_4408060_TG_gid_H48P	3	1	1	1
STR_gid_A134E_u	3	0	0	3
STR_gid_A138E_u	3	0	0	3
STR_gid_A80P_u	3	0	0	3
STR_gid_D85A_u	3	0	0	3
STR_gid_E92Q_u	3	0	0	3
STR_gid_G69D_u	3	0	0	3
STR_gid_L86P_u	3	0	0	3
STR_gid_P75S_u	3	0	0	3
STR_gid_R118S_u	3	0	0	3
STR_gid_T146K_u	3	0	0	3
STR_gid_V77A_u	3	0	0	3
STR_rrs_C905A_u	3	0	0	3
STR_SNP_N_1473167_T1322G_rrs	2	0	0	2
STR_gid_A200E_u	2	0	0	2
STR_gid_E170G_u	2	0	0	2
STR_gid_E40A_u	2	0	0	2
STR_gid_G117E_u	2	0	0	2
STR_gid_G117V_u	2	0	0	2
STR_gid_G30R_u	2	0	0	2
STR_gid_G34A_u	2	0	0	2
STR_gid_G34V_u	2	0	0	2
STR_gid_G37E_u	2	0	0	2
STR_gid_H48N_u	2	0	0	2
STR_gid_L101F_u	2	0	0	2
STR_gid_L108R_u	2	0	0	2
STR_gid_L18H_u	2	0	0	2
STR_gid_L59R_u	2	0	0	2
STR_gid_P75L_u	2	0	0	2
STR_gid_R137W_u	2	0	0	2
STR_gid_R83G_u	2	0	0	2
STR_gid_R83W_u	2	0	0	2
STR_gid_V110A_u	2	0	0	2
STR_gid_V65G_u	2	0	0	2
STR_gid_W45S_u	2	0	0	2
STR_rpsL_K88Q_u	2	0	0	2
STR_SNP_CN_4408138_TC_gid_Y22C	1	0	0	1
STR_gid_A10V_u	1	0	0	1
STR_gid_A119D_u	1	0	0	1
STR_gid_A133P_u	1	0	0	1
STR_gid_A138P_u	1	0	0	1
STR_gid_A140S_u	1	0	0	1
STR_gid_A19G_u	1	0	0	1
STR_gid_A205E_u	1	0	0	1
STR_gid_A205T_u	1	0	0	1
STR_gid_A27P_u	1	0	0	1
STR_gid_A72V_u	1	0	0	1
STR_gid_A82E_u	1	0	0	1
STR_gid_C52F_u	1	0	0	1
STR_gid_D67A_u	1	0	0	1
STR_gid_E120K_u	1	0	0	1
STR_gid_E121D_u	1	0	0	1
STR_gid_G117R_u	1	0	0	1
STR_gid_G117W_u	1	0	0	1
STR_gid_G130A_u	1	0	0	1
STR_gid_G182R_u	1	0	0	1
STR_gid_G30D_u	1	0	0	1
STR_gid_G30V_u	1	0	0	1
STR_gid_G34R_u	1	0	0	1

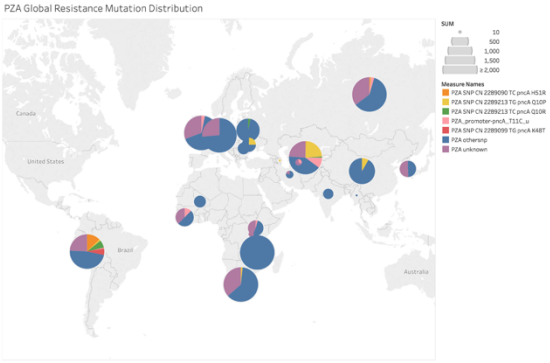
STR_gid_G34W_u	0	0	0	1
STR_gid_G71R_u	1	0	0	1
STR_gid_H48D_u	1	0	0	1
STR_gid_H48Q_u	1	0	0	1
STR_gid_H48R_u	1	0	0	1
STR_gid_I11N_u	1	0	0	1
STR_gid_I162S_u	1	0	0	1
STR_gid_I81R_u	1	0	0	1
STR_gid_I81T_u	1	0	0	1
STR_gid_L108P_u	1	0	0	1
STR_gid_L142W_u	1	0	0	1
STR_gid_L160P_u	1	0	0	1
STR_gid_L79F_u	1	0	0	1
STR_gid_L86R_u	1	0	0	1
STR_gid_L91P_u	1	0	0	1
STR_gid_M150R_u	1	0	0	1
STR_gid_P78L_u	1	0	0	1
STR_gid_P93L_u	1	0	0	1
STR_gid_P93Q_u	1	0	0	1
STR_gid_P93R_u	1	0	0	1
STR_gid_R116W_u	1	0	0	1
STR_gid_R118C_u	1	0	0	1
STR_gid_R118L_u	1	0	0	1
STR_gid_R137L_u	1	0	0	1
STR_gid_R175P_u	1	0	0	1
STR_gid_R20P_u	1	0	0	1
STR_gid_R21W_u	1	0	0	1
STR_gid_R33P_u	1	0	0	1
STR_gid_R47G_u	1	0	0	1
STR_gid_R47P_u	1	0	0	1
STR_gid_R64W_u	1	0	0	1
STR_gid_R83P_u	1	0	0	1
STR_gid_R96C_u	1	0	0	1
STR_gid_S131T_u	1	0	0	1
STR_gid_S70N_u	1	0	0	1
STR_gid_T98I_u	1	0	0	1
STR_gid_V105E_u	1	0	0	1
STR_gid_V115G_u	1	0	0	1
STR_gid_V203L_u	1	0	0	1
STR_gid_V41I_u	1	0	0	1
STR_gid_V66A_u	1	0	0	1
STR_gid_V88A_u	1	0	0	1
STR_gid_V89G_u	1	0	0	1
STR_gid_Y195C_u	1	0	0	1
STR_gid_Y22S_u	1	0	0	1
STR_rpsL_E70A_u	1	0	0	1
STR_rpsL_E76A_u	1	0	0	1
STR_rrs_A1012G_u	1	0	0	1
STR_rrs_A1223G_u	1	0	0	1
STR_rrs_A1325C_u	1	0	0	1
STR_rrs_A170G_u	1	0	0	1
STR_rrs_A504C_u	1	0	0	1
STR_rrs_A700T_u	1	0	0	1
STR_rrs_A703G_u	1	0	0	1
STR_rrs_A753T_u	1	0	0	1
STR_rrs_A907C_u	1	0	0	1
STR_rrs_A907T_u	1	0	0	1
STR_rrs_A908G_u	1	0	0	1
STR_rrs_A948T_u	1	0	0	1
STR_rrs_C1067T_u	1	0	0	1
STR_rrs_C1199G_u	1	0	0	1
STR_rrs_C1390T_u	1	0	0	1
STR_rrs_C1402T_u	1	0	0	1
STR_rrs_C239A_u	1	0	0	1
STR_rrs_C244T_u	1	0	0	1
STR_rrs_C270T_u	1	0	0	1
STR_rrs_C332T_u	1	0	0	1
STR_rrs_C397T_u	1	0	0	1
STR_rrs_C414A_u	1	0	0	1
STR_rrs_C499T_u	1	0	0	1
STR_rrs_C546T_u	1	0	0	1
STR_rrs_C692T_u	1	0	0	1
STR_rrs_C708T_u	1	0	0	1
STR_rrs_C897T_u	1	0	0	1
STR_rrs_C905G_u	1	0	0	1
STR_rrs_C936T_u	1	0	0	1
STR_rrs_G1042C_u	1	0	0	1
STR_rrs_G1072A_u	1	0	0	1
STR_rrs_G1415T_u	1	0	0	1
STR_rrs_G319A_u	1	0	0	1
STR_rrs_G395C_u	1	0	0	1
STR_rrs_G407A_u	1	0	0	1
STR_rrs_G408T_u	1	0	0	1
STR_rrs_G537A_u	1	0	0	1
STR_rrs_G544A_u	1	0	0	1
STR_rrs_G685A_u	1	0	0	1
STR_rrs_G749C_u	1	0	0	1
STR_rrs_G771A_u	1	0	0	1
STR_rrs_G878A_u	1	0	0	1
STR_rrs_G887T_u	1	0	0	1
STR_rrs_G909T_u	1	0	0	1
STR_rrs_T1206C_u	1	0	0	1
STR_rrs_T1208A_u	1	0	0	1
STR_rrs_T1217A_u	1	0	0	1
STR_rrs_T16C_u	1	0	0	1
STR_rrs_T327C_u	1	0	0	1
STR_rrs_T411A_u	1	0	0	1
STR_rrs_T529G_u	1	0	0	1
STR_rrs_T545A_u	1	0	0	1
STR_rrs_T580C_u	1	0	0	1
STR_rrs_T672A_u	1	0	0	1
STR_rrs_T696G_u	1	0	0	1
STR_DEL_F_4408023_d179T_gid	0	0	0	0
STR_DEL_F_4408116_d86G_gid	0	0	0	0
STR_SNP_CN_4408091_GT_gid_P38T	0	0	0	0
STR_SNP_I_1473637_A.21_rrs.rrl	0	0	0	0
STR_SNP_N_1473109_T1264G_rrs	0	0	0	0
STR_SNP_N_1473343_G1498T_rrs	0	0	0	0

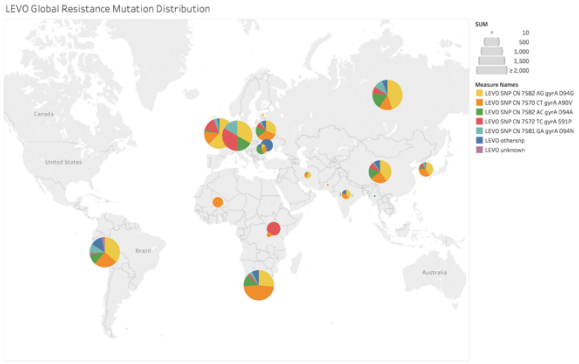
Supplementary Figure 7 Distribution of resistance mutations for six drugs (n=9385). Pie chart size is proportional isolate number from each country. Mutations listed in order of frequency. Full data provided in Supplementary Table 6.

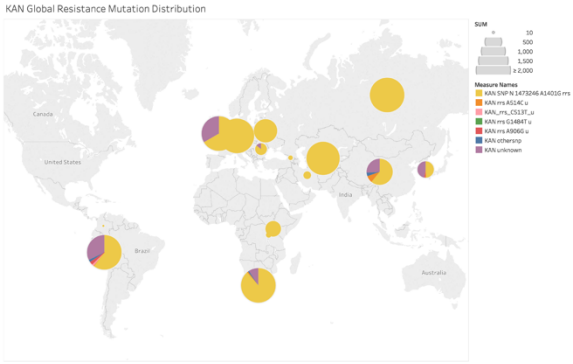






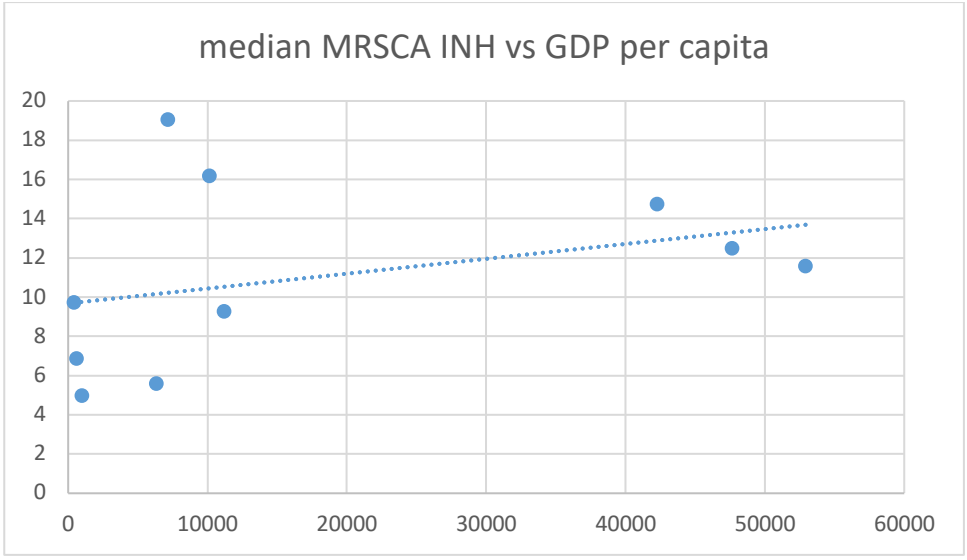


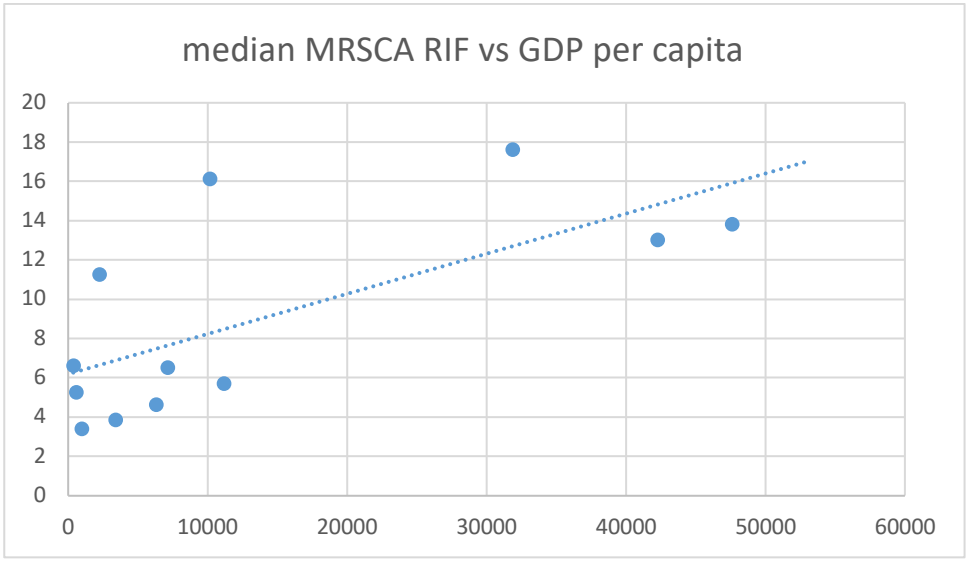


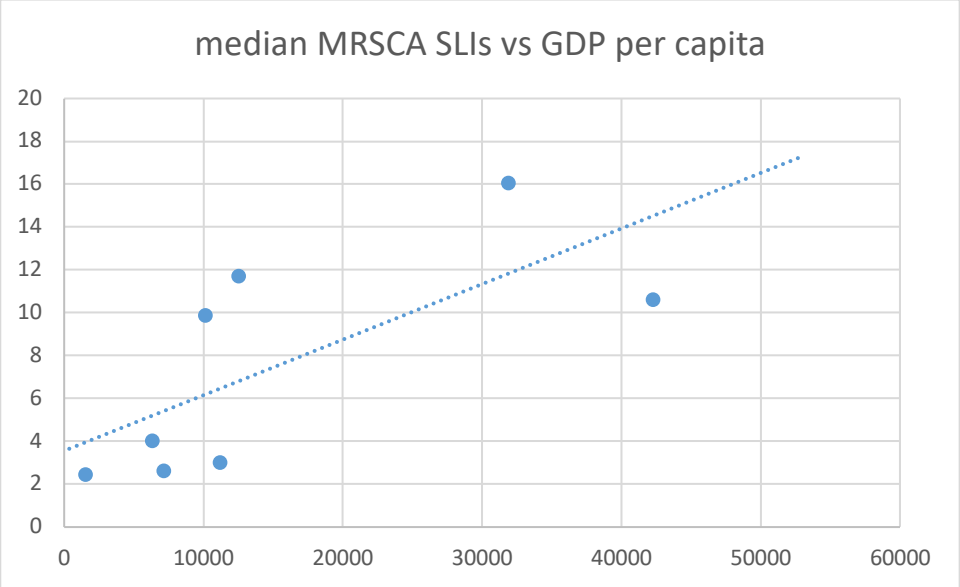


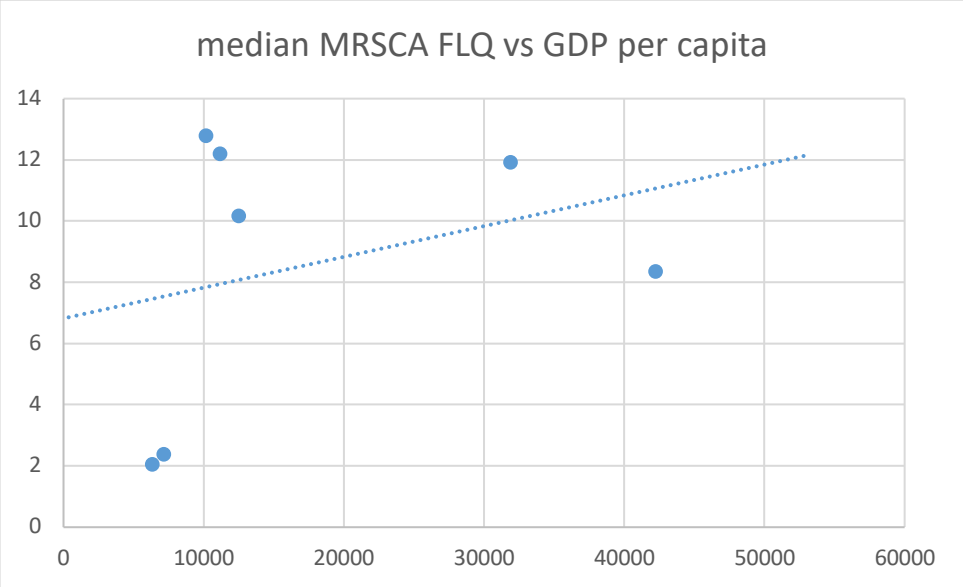
Supplementary Table 8 GDP Per Capita Verses Median MRSCA date for INH, RIF, SLIs, and FLQ.

Countries	GDP per capita	Median MRSCA Date INH (Years Ago)	Median MRSCA Date RIF (Years Ago)	Median MRSCA Date SLIs (Years Ago)	Median MRSCA Date FLQ (Years Ago)
China	10150	16.18	16.11	9.88	12.8
Peru	7140	19.04	6.52	2.61	2.39
Russia	11190	9.26	5.71	3.02	12.22
South Africa	6330	5.58	4.63	4.01	2.06
United Kingdom	42310	14.76	13.02	10.61	8.36
Uzbekistan	1480			2.45	
Romania	12510			11.74	10.19
Mali	933.88	4.97	3.4		
Netherlands	53020	11.59			
Moldova	3400		3.83		
Malawi	366.53	9.72	6.62		
South Korea	31940		17.66	16.08	11.95
Sierra Leone	516.76	6.85	5.24		
Germany	47790	12.48	13.86		
India	2200		11.27		









Supplementary Table 9 Commercial Diagnostics[21,22,23,24,25]:

Drug	Gene	Location
Isoniazid	katG codon	315
Isoniazid	inhA promoter	-15,-16,-8
Rifamycin	rpoB codons	424-454
Aminoglycosides	rrs	1401,1402
Aminoglycosides	eis	-10 to -14, -37
Fluoroquinolones	gyrA	89-94
Fluoroquinolones	gyrB	500-541

*E-coli 505-534

Supplementary Table 10 Genes Searched Per Drug.

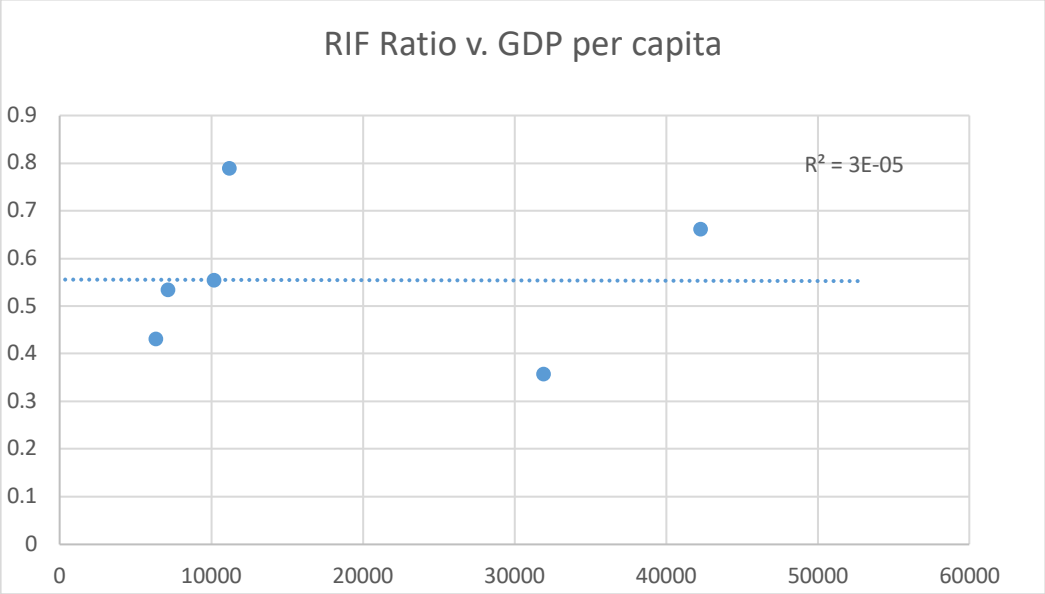
Genes searched for resistant mutations in "All WGS Test" (AMK: amikacin, PAS: para-aminosalicylic acid, EMB: ethambutol, CAP: capreomycin, KAN: kanamycin, CIP: ciprofloxacin, INH: isoniazid, STR: streptomycin, RIF: rifampicin, LEVO: levofloxacin, ETH: ethionamide, OFLX: ofloxacin, PZA: pyrazinamide)

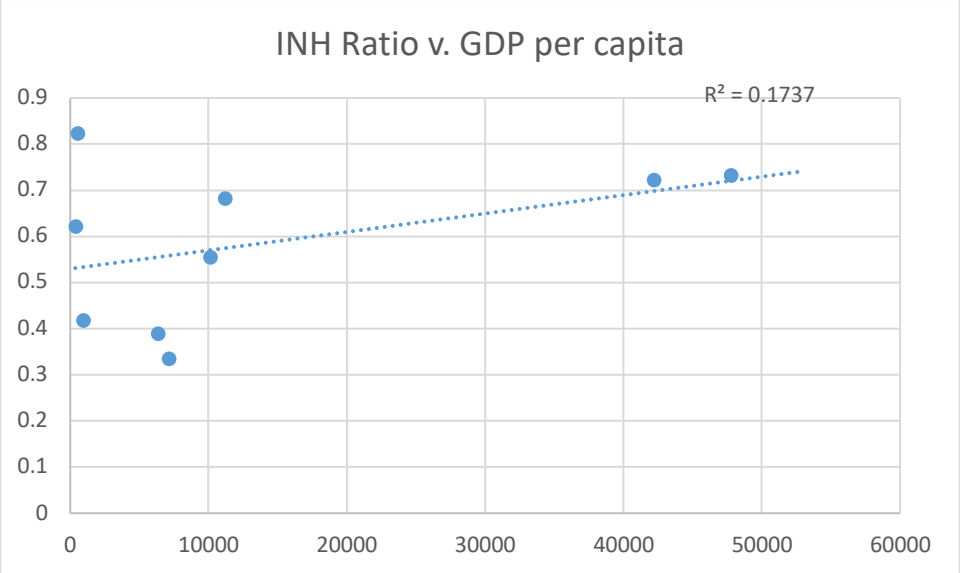
Drug	Genes Searched
AMK	rrs
PAS	thyA, inter-thyA-Rv2765, folC, inter-thyX-hsdS.1
EMB	embA, embB, embC, iniB, inter-embC-embA
CAP	rrs, tlyA
KAN	rrs, inter-eis-Rv2417c

CIP	gyrB, gyrA
INH	inhA, iniB, embB, inter-Rv1482c-fabG1, ahpC, inter-oxyR'-ahpC, inter-embC-embA, kasA, katG, fabG1
STR	gid, rpsL, rrs, inter-rrs-rrl
RIF	rpoB
LEVO	gyrB, gyrA
ETH	inhA, inter-Rv1482c-fabG1, ethA, inter-ethA-ethR
OFLX	gyrB, gyrA
PZA	inter-pncA-Rv2044c, pncA, rpsA

Supplementary Table 11 Transmission rate estimates and association with GDP by country. Transmission rate X/Y indicates X unique MRSCA dates/Y number of MRSCA dates for a specific country and drug. Legend: Unique = number of unique MRSCA dates for drug and country, total = total number of dated resistant isolates for drug and country, Lower Bound (LB) = (total-unique)/total, Note: Data for country fewer than 10 resistant isolates per drug not shown.

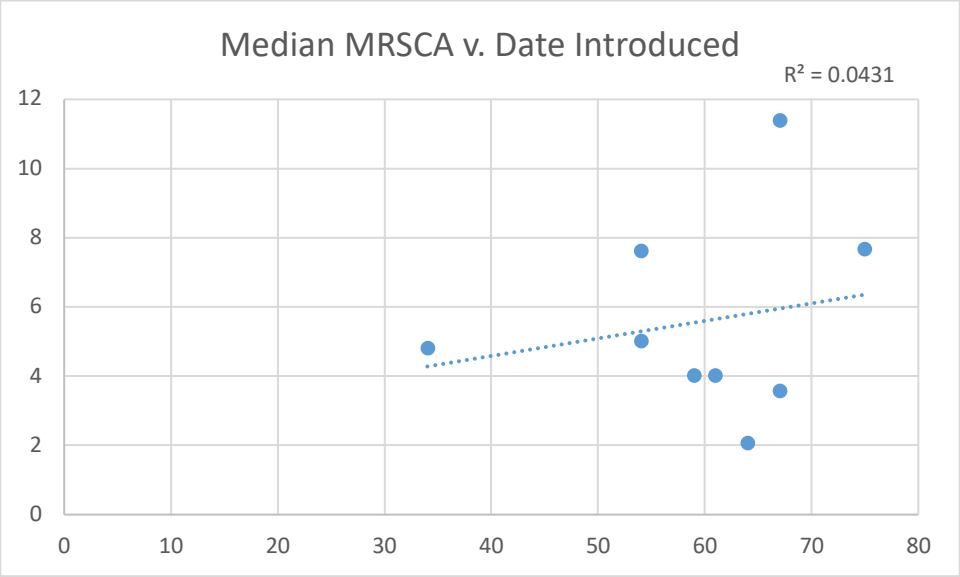
	Pooled			INH			RIF			PZA			FLQ			STR			SLIs			EMB		
Country*	unique	total	LB Transmission	unique	total	LB Transmission	unique	total	LB Transmission	unique	total	LB Transmission	unique	total	LB Transmission	unique	total	LB Transmission	unique	total	LB Transmission	unique	total	LB Transmission
China	296	441	33%	56	101	45%	56	101	45%				32	44	27%	48	71	32%	33	36	8%	52	58	10%
Germany	90	113	20%	30	41	27%										40	50	20%						
India	13	20	35%																					
Malawi	44	68	35%	23	37	38%										15	23	35%						
Mali	18	30	40%	5	12	58%																		
Moldova	8	13	38%																					
Netherlands																								
Peru	1260	2116	40%	12	36	67%	382	715	47%	177	268	34%	59	68	13%	170	247	31%	194	251	23%	239	395	39%
Romania	29	41	29%													9	18	50%	13	15	13%			
Russia	302	394	23%	73	107	32%	56	71	21%	32	35	9%	19	19	0%	69	92	25%	10	12	17%	31	34	9%
Sierra Leone	48	64	25%	14	17	18%										16	27	41%						
South Africa	491	1015	52%	69	178	61%	72	167	57%	34	75	55%	79	128	38%	65	120	46%	86	192	55%	54	86	37%
South Korea	43	79	46%				5	14	64%										14	30	53%			
United Kingdom	299	416	28%	133	184	28%	49	74	34%	27	36	25%	27	37	27%	28	33	15%	14	17	18%	21	35	40%
Uzbekistan	31	36	14%																9	11	18%	14	17	18%
Total	2972	4846	39%	415	713	42%	620	1142	46%	270	414	35%	216	296	27%	460	681	32%	373	564	34%	411	625	34%





Supplementary Table 12 MRSCA age is not associated with the earliest date of drug introduction into clinical use.

Drug	Median MRSCA	Date Introduced Years ago
INH	11.4	67
RIF	7.61	54
PZA	3.56	67
FLQ	4.81	34
STR	7.67	75
SLIs	4.02	61
EMB	5.01	54
ETH	4.01	59
CYS	2.05	64



Supplementary Table 13 Geographic variance of resistance mutations

A: *katG S315T* mutation prevalence among INH resistant isolates in Russia verses Peru (Fisher P-value $1*10^{-12}$)

	Russia	Peru
Mutation	444	510
No Mutation	82	250

B: *fabG/inhA*-15 C>T mutation prevalence among INH resistant isolates in Russia verses Peru (Fisher P-value $7*10^{-9}$)

	Russia	Peru
Mutation	43	149
No Mutation	483	611

C: 25 mutations that varied geographically to a larger extent than the mutation *fabG/inhA* promoter -15C>T (SD 10%, Range 0%-39%) , their standard deviations, and ranges.

Mutation	Standard Deviation	Range of observed frequency across countries
INH_SNP_CN_2155168_CG_katG_S315T	12.49%	54.55%-93.75%
INH_SNP_CN_4247429_AG_embB_M306V	12.03%	0.0%-43.56%
INH_SNP_CN_4247431_GA_embB_M306I	11.88%	0.0%-42.63%
INH_SNP_CN_4247431_GC_embB_M306I	11.88%	0.0%-42.63%
PZA_SNP_CN_2289073_GC_pncA_H57D	15.07%	0.0%-46.67%
EMB_SNP_CN_4247429_AG_embB_M306V	12.48%	8.33%-53.37%
EMB_SNP_CN_4247431_GA_embB_M306I	11.95%	9.02%-49.58%
EMB_SNP_CN_4247431_GC_embB_M306I	11.95%	9.02%-49.58%

EMB_SNP_CN_4247431_GT_embB_M306I	11.95%	9.02%-49.58%
STR_SNP_CN_781687_AG_rpsL_K43R	21.03%	0.0%-81.85%
STR_SNP_N_1472359_A514C_rrs	12.00%	0.0%-50.4%
STR_SNP_N_1473246_A1401G_rrs	17.90%	0.0%-60.0%
STR_SNP_CN_4407927_TG_gid_E92D	31.17%	0.0%-87.95%
STR_SNP_N_1472362_C517T_rrs	11.32%	0.0%-38.89%
STR_SNP_CN_4407967_AG_gid_L79S	20.81%	0.0%-72.73%
ETH_SNP_P_1673425_CT.15_fabG1.inhA	14.48%	27.78%-62.5%
ETH_SNP_CN_4326333_CG_ethA_A381P	19.24%	0.0%-48.44%
ETH_SNP_CN_1674481_TG_inhA_S94A	11.89%	0.0%-31.46%
KAN_SNP_N_1473246_A1401G_rrs	14.83%	45.65%-85.71%
CAP_SNP_N_1473246_A1401G_rrs	19.62%	23.08%-85.84%
AMK_SNP_N_1473246_A1401G_rrs	20.11%	46.43%-100.0%
AMK_SNP_N_1472359_A514C_rrs	20.90%	0.0%-64.0%
CIP_SNP_CN_7582_AG_gyrA_D94G	26.57%	16.67%-78.26%
OFLX_SNP_CN_7582_AG_gyrA_D94G	12.60%	8.33%-50.0%
OFLX_SNP_CN_7570_CT_gyrA_A90V	11.83%	5.45%-46.15%

D: Six mutations that varied by lineage to a larger extent than the mutation *fabG/inhA* promoter -15C>T (SD 9.3%, Range 8.8%-33%), their standard deviations, and ranges.

Mutation	Standard Deviation	Range of observed frequency across lineages
INH_SNP_CN_2155168_CG_katG_S315T	16.00%	40.28%-84.23%
INH_SNP_CN_2518919_GA_kasA_G269S	11.11%	0.0%-25.66%
PZA_SNP_CN_2288952_CT_pncA_G97D	11.74%	0.0%-27.27%
EMB_SNP_CN_4247429_AG_embB_M306V	13.65%	18.87%-53.85%
STR_SNP_CN_781687_AG_rpsL_K43R	20.23%	4.76%-57.36%
STR_SNP_CN_4407927_TG_gid_E92D	43.23%	0.0%-99.83%

Supplementary Table 14 Sensitivity and Specificity of commercial tests for INH, RIF, SLIS, and FLQ in five countries with the largest number of phenotyped strained: Russia, South Africa, Peru, Uzbekistan, and United Kingdom

Legend: Sensitivity=Percent of resistant isolates classified as resistant, Specificity=Percent of susceptible isolates classified as susceptible

Drug	commercial test-Peru		commercial test-Russia		commercial test-South Africa		commercial test-Uzbekistan		commercial test-UnitedKingdom	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
INH	84% (641/760)	90% (151/168)	87% (459/526)	91% (281/310)	88% (168/190)	72% (444/619)	90% (238/264)	NA 0/0	87% (245/283)	99% (1559/1567)
RIF	90% (623/692)	89% (209/236)	82% (347/425)	88% (363/412)	95% (158/166)	70% (433/616)	95% (248/262)	50% (1/2)	91% (86/95)	99% (1742/1750)
FLQ	38% (46/121)	98% (788/801)	38% (51/133)	94% (275/292)	77% (86/111)	87% (510/585)	0% (0/7)	100% (213/213)	90% (26/29)	98% (295/301)
SLI	47% (101/214)	98%(692/708)	77% (63/82)	48%(163/343)	77% (101/131)	91%(486/536)	79% (44/56)	66% (109/164)	90% (9/10)	100%(83/83)

Supplementary References:

1. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27(6):863-864. doi:10.1093/bioinformatics/btr026

2. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*. 2014;15(3):R46. doi:10.1186/gb-2014-15-3-r46

3. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:13033997 [q-bio]*. March 2013. http://arxiv.org/abs/1303.3997. Accessed April 11, 2019.

4. *Picard Toolkit*. Broad Institue, GitHub repository: Broad Institute; 2019. http://broadinstitute.github.io/picard/.

5. Walker BJ, Abeel T, Shea T, et al. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLOS ONE*. 2014;9(11):e112963. doi:10.1371/journal.pone.0112963

6. Ezewudo M, Borens A, Chiner-Oms Á, et al. Integrating standardized whole genome sequence analysis with a global Mycobacterium tuberculosis antibiotic resistance knowledgebase. *Scientific Reports*. 2018;8(1):15382. doi:10.1038/s41598-018-33731-1

7. Coll F, McNerney R, Guerra-Assunção JA, et al. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nature Communications*. 2014;5:4812. doi:10.1038/ncomms5812

8. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. 2010:5.

9. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312-1313. doi:10.1093/bioinformatics/btu033

10. Yang Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J Mol Evol*. 1994;39(3):306-314. doi:10.1007/BF00160154

11. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol*. 2018;4(1). doi:10.1093/ve/vey016

12. Dixit A, Freschi L, Vargas R, et al. Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. *Scientific Reports*. 2019;9(1):5602. doi:10.1038/s41598-019-41967-8

13. Sukumaran J, Holder MT. DendroPy: a Python library for phylogenetic computing. *Bioinformatics*. 2010;26(12):1569-1571. doi:10.1093/bioinformatics/btq228

14. Oliphant T. *NumPy: A Guide to NumPy*. USA: Trelgol Publishing; 2006. http://www.numpy.org/.

15. Jones E, Oliphant T, Peterson P. SciPy: Open Source Scientific Tools for Python. January 2001.

16. World Economic Outlook, April 2019: Growth Slowdown, Precarious Recovery. IMF. https://www.imf.org/en/Publications/WEO/Issues/2019/03/28/world-economic-outlook-april-2019. Accessed April 11, 2019.

17. Kaufmann J, Schering AG. Analysis of Variance ANOVA. In: *Wiley StatsRef: Statistics Reference Online*. American Cancer Society; 2014. doi:10.1002/9781118445112.stat06938

18. Murray JF, Schraufnagel DE, Hopewell PC. Treatment of Tuberculosis. A Historical Perspective. *Annals ATS*. 2015;12(12):1749-1759. doi:10.1513/AnnalsATS.201509-632PS

19. Zhang Y, Shi W, Zhang W, Mitchison D. Mechanisms of Pyrazinamide Action and Resistance. *Microbiology Spectrum*. 2014;2(4). doi:10.1128/microbiolspec.MGM2-0023-2013

20. Alapi EM, Fischer J. Table of Selected Analogue Classes. In: *Analogue-Based Drug Discovery*. John Wiley & Sons, Ltd; 2006:441-552. doi:10.1002/3527608001.ch23

21. Rahman A, Sahrin M, Afrin S, et al. Comparison of Xpert MTB/RIF Assay and GenoType MTBDRplus DNA Probes for Detection of Mutations Associated with Rifampicin Resistance in Mycobacterium tuberculosis. *PLOS ONE*. 2016;11(4):e0152694. doi:10.1371/journal.pone.0152694

22. policy_statement.pdf. https://www.who.int/tb/features_archive/policy_statement.pdf. Accessed May 25, 2019.

23. Chen H-Y, Yu M-C, Huang W-L, et al. Molecular Detection of Rifabutin-Susceptible Mycobacterium tuberculosis. *Journal of Clinical Microbiology*. 2012;50(6):2085-2088. doi:10.1128/JCM.00652-12

24. Tagliani E, Cabibbe AM, Miotto P, et al. Diagnostic Performance of the New Version (v2.0) of GenoType MTBDRsl Assay for Detection of Resistance to Fluoroquinolones and Second-Line Injectable Drugs: a Multicenter Study. *Journal of Clinical Microbiology*. 2015;53(9):2961-2969. doi:10.1128/JCM.01257-15

25. Xie YL, Chakravorty S, Armstrong DT, et al. Evaluation of a Rapid Molecular Drug-Susceptibility Test for Tuberculosis. *New England Journal of Medicine*. 2017;377(11):1043-1054. doi:10.1056/NEJMoa1614915