Sources of multi-drug resistance in patients with previous isoniazid resistant tuberculosis identified using whole genome sequencing: A longitudinal cohort study.

Srinivasan Vijay, *[1,2] Vu T N Ha,*[1] Dao N Vinh,[1] Phan V K Thai,[3] Dang T M Ha,[3] Nguyen H Lan,[3] Hoang T Hai,[1] Timothy M Walker,[2] Do D A Thu,[1] Sarah J Dunstan,[4] Guy E Thwaites,[1,2] Philip M Ashton, [1,2] Maxine Caws, [5] Nguyen T T Thuong[1**]

[1]Oxford University Clinical Research Unit, HCMC, Vietnam [2]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK [3]Pham Ngoc Thach Hospital for Tuberculosis and Lung Disease, HCMC, Vietnam [4]Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, Australia [5]Department of Clinical Sciences, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK.

* These authors contributed equally to this manuscript.

**** Correspondence to**

Dr. Nguyen T T Thuong,

Oxford University Clinical Research Unit,

764 Vo Van Kiet, District 5, HCMC, Vietnam;

thuongntt@oucru.org

Summary. We investigated the sources of MDR-TB in patients with isoniazid-resistant tuberculosis treated with 1st line anti-tuberculosis therapy and show that re-infection with a new MDR-TB strain was just as common as the emergence of rifampicin resistance among these patients.

ABSTRACT

Background. Meta-analysis of patients with isoniazid-resistant tuberculosis given standard first-line anti-tuberculosis treatment indicated an increased risk of multi-drug resistant tuberculosis (MDR-TB) emerging (8%), compared to drug-sensitive tuberculosis (0.3%). Here we use whole genome sequencing (WGS) to investigate whether treatment of patients with pre-existing isoniazid resistant disease with first-line anti-tuberculosis therapy risks selecting for rifampicin resistance, and hence MDR-TB.

Methods. Patients with isoniazid-resistant pulmonary TB were recruited and followed up for 24 months. Drug-susceptibility testing was performed by Microscopic observation drug-susceptibility assay (MODS), Mycobacterial Growth Indicator Tube (MGIT) and by WGS on isolates at first presentation and in the case of re-presentation. Where MDR-TB was diagnosed, WGS was used to determine the genomic relatedness between initial and subsequent isolates. *De novo* emergence of MDR-TB was assumed where the genomic distance was five or fewer single nucleotide polymorphisms (SNPs) whereas reinfection with a different MDR-TB strain was assumed where the distance was 10 or more SNPs.

Results. 239 patients with isoniazid-resistant pulmonary tuberculosis were recruited. Fourteen (14/239, 5.9%) patients were diagnosed with a second episode of tuberculosis that was multi-drug resistant. Six (6/239, 2.5%) were identified as having evolved MDR-TB *de novo* and six as having been re-infected with a different strain. In two cases the genomic distance was between 5-10 SNPs and therefore indeterminate.

Conclusions. In isoniazid-resistant TB, *de novo* emergence and reinfection of MDR-TB strains equally contributed to MDR development. Early diagnosis and optimal treatment of isoniazid resistant TB are urgently needed to avert the *de novo* emergence of MDR-TB during treatment.

Keywords. Tuberculosis; multidrug-resistance; isoniazid-resistance; whole-genome sequencing; rifampicin-resistance.

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, kills more people each year than any other single pathogen [1]. Resistance to first-line anti-TB drug isoniazid is the most common drug resistant TB, with a global prevalence of 10%, and it is associated with increased risk of treatment failure and emergence of MDR-TB with standard first-line TB therapy (11% and 8% respectively) compared to susceptible-TB (1% and 0.3% respectively) [2,3]. Emergence of multi-drug resistant TB (MDR-TB) strains, resistant to both isoniazid and rifampicin, is a major concern with an estimated 600,000 cases of MDR-TB or rifampicin-resistant TB each year [4]. MDR-TB requires longer treatment with more expensive and less effective antibiotics [5]. It is also the precursor for extensively drug-resistant TB (XDR-TB) [6]. Worldwide prevalence of MDR-TB among patients newly diagnosed with TB is approximately 3.4% compared to 18% among patients diagnosed for a subsequent time [1]. Treatment success remains low at about 56% [1]. Vietnam, where this study is set, has been one among the top 20 high TB and MDR-TB burden country in absolute number [1,7].

MDR-TB strains isolated from patients initially with susceptible strains have in some past studies been ascribed to reinfection with a MDR-TB strain [8,9]. However, recent data suggest *de novo* emergence of MDR-TB may be playing a more significant role than previously thought. An analysis of a global data set of *M. tuberculosis* genomes found that isoniazid resistance typically emerges before rifampicin resistance [10], whilst a recent meta-analysis concluded that the treatment of patients with isoniazid-resistant disease with standard first-line drugs risks the emergence of MDR-TB [3]. WGS can be used to distinguish between *de novo* emergence and reinfection of MDR-TB and can provide genomic evidence to assess the source of MDR-TB [11-13].

A recently published study from Vietnam explored the bacterial risk factors for treatment failure among patients with isoniazid-resistant tuberculosis [2]. However that study did not explore whether patients who re-presented with MDR-TB had been reinfected with new strains or whether the original TB strain had evolved resistance *de novo*. Here we whole genome sequenced the longitudinally collected isolates from that study to test the hypothesis that standard first-line treatment of patients with isoniazid-resistant TB risks *de novo* selection for rifampicin-resistant mutations.

METHODS

Ethical approval

The study was approved by the Oxford University Tropical Research Ethics Committee, UK (OxTREC 030-07) and the Institutional Research Board of Pham Ngoc Thach Hospital in Ho Chi Minh city, Vietnam. All participants provided written informed consent.

Patient recruitment

Between December 2008 and June 2011, newly diagnosed patients with smear positive pulmonary TB were recruited in Ho Chi Minh City, Vietnam for a clinical study investigating the bacterial risk factors for treatment failure among patients with isoniazid-resistant TB [2]. Recruitment was restricted to new adult patients (aged $\geq$18) without HIV infection and no prior TB treatment [2]. Initial screening for isoniazid resistance was done using MODS [14] with results later confirmed using MGIT [2]. Follow-up was for 24 months with sputum collected, where this could be produced at 0, 1, 2, 5, 8, 12, 18 and 24 months after diagnosis. The patients were treated by Directly Observed Treatment, Short Course (DOTS) with the then standard first-line regimens according to the Vietnamese Ministry of Health guidelines for susceptible including isoniazid-resistant TB, two months of isoniazid, rifampicin, pyrazinamide and ethambutol followed by six months of isoniazid and ethambutol (2HRZE/6HE) or two months of isoniazid, rifampicin, pyrazinamide and streptomycin followed by six months of isoniazid and ethambutol (2HRZS/6HE) or other individualized treatment regimens (Supplementary Table 1) [2, 15].

Culturing *M. tuberculosis* isolates and Drug Susceptibility Testing

Sputum samples from the patients were used to culture the *M. tuberculosis* isolates in the Pham Ngoc Thach hospital as per the protocol developed from national TB control programme, Vietnam (Supplementary methods).

DNA extraction and whole genome sequencing

M. tuberculosis isolates DNA were extracted using cetyltrimethylammonium bromide method [16]. This genomic DNA was used for library preparation using the Nextera XT kit (Illumina) and 150bp or 300bp paired end sequencing using MiSeq V2 or V3 reagent kits (Illumina) on the MiSeq sequencing platform (Illumina).

Whole genome sequence analysis

FASTQ data generated on the Illumina MiSeq machine were mapped against the $H_{37}Rv$ reference genome (NC_000962.3) using bwa mem [17], SNPs were called using GATK (version 3.8-1-0-gf15c1c3ef) in unified genotyper mode [18]. These steps were carried out using the PHEnix pipeline (https://github.com/phe-bioinformatics/PHEnix) and SnapperDB [19]. Maximum likelihood phylogenetic analysis was performed by IQ-TREE v1.6 [20]. *M. tuberculosis* lineages, sub lineages and genotypic antibiotic resistance were identified by Mykrobe predictor TB platform [21].

Genetic relatedness analysis

Genetic relatedness between the *M. tuberculosis* isolates were analysed by constructing a phylogenetic tree of all the longitudinal isolates with WGS data (n = 368). Phylogenetic location and the SNP distance between the baseline and the MDR-TB isolates emerging in each patient were calculated. From base substitution rate of 0.3-0.5 mutations/per genome/per year in *M. tuberculosis* isolates, SNP difference between the longitudinal isolates from our study and from the published literature [11,12,22], we used equal or less than five SNPs difference as a cut-off for the *de novo* emergence of MDR-TB from the initial isolate and more than ten SNP differences as reinfection with MDR-TB. SNP differences between 5 and 10 were described as indeterminate as it was difficult to differentiate either as *de novo* emergence or reinfection with another strain.

RESULTS

Characteristics of participants in the study

2090 consecutively sampled patients were assessed for entry into the study, 1804 patients samples were culture-positive and provided TB strains (Figure 1) [2]. 392 patients had TB strains with isoniazid resistance on MODS; 50 patients declined to be followed up over 24 months and their results were excluded, 68 patients had MDR-TB and 274 had isolates with resistance to isoniazid and susceptibility to rifampicin. Of these 274, confirmatory phenotypic susceptibility testing by MGIT corroborated the isoniazid resistant result by MODS in 239 cases but reported susceptibility in 35 cases (Figure 1).

Of those patients whose strains were isoniazid-resistant by both MODS and MGIT, 105/239 (43.9%) patients produced at least one more sputum samples, which were culture-positive over the 24 months follow-up, whereas 134 (56.1%) patients had early sputum conversion as their subsequent sputum samples were culture-negative.

Of those patient strains whose MGIT result was susceptible for isoniazid, despite resistance reported by MODS, 15/35 (42.8%) produced subsequent sputum samples which were culture-positive and experienced treatment failure and remaining 20 patients subsequent sputum samples were culture-negative. Similarly, only 43/68 MDR-TB patients had subsequent sputum samples which were culture positive and for remaining 25 MDR-TB patients subsequent *M. tuberculosis* isolates were unavailable. Treatment data for 134 patients with early sputum clearance and 50 patients who declined to participate showed six having unfavorable and rest favorable outcome.

MDR-TB was detected by MGIT during 24 months follow-up in subsequent isolates from 18/105 patients whose baseline isolate was isoniazid-resistant by both MODS and MGIT, and 5/35 patients with baseline isoniazid-susceptible isolates by MGIT, discordant with MODS result. For 3/18 patients with emergence of MDR-TB, the initial isolate was phenotypic rifampicin susceptible by MGIT but WGS detected rifampicin-resistant mutations and WGS data was lacking for isolate from 1/18 patient. These four patients were excluded from analysis (Figure 1). For 163 patients with strains having WGS data, Median age was 41 years, 74.2% were male and 50.6% reported smoking (Supplementary Table 1).

Temporal dynamics of emergence of MDR-TB in patients

Of the 14 patients who initially had an isoniazid resistant strain and developed MDR-TB, 11 did so within the first five months of treatment whereas three were diagnosed with MDR-TB 12 or 24 months after completing initial treatment. Of the five patients who developed MDR-TB with baseline susceptible strain by MGIT, four did so within five months of starting treatment and one was diagnosed with MDR-TB at 12 months (Figure 2). 161/162 patients received only two or three months of rifampicin during the intensive phase, whilst 1/162 patient received rifampicin for six months during the treatment (Figure 2, Supplementary Table 1).

Genetic relatedness between the initial and the first MDR-TB isolate in the same patients

To help assess genomic links between isolates and potentially explain MDR-TB acquisition, all longitudinally collected whole genome sequenced isolates were assessed for genomic relatedness (n=368 isolates) (Figure 1). In 6 (43%) out of 14 patients with initial isoniazid-resistant disease, the subsequent MDR-TB isolates were within five SNPs of their original isolates, and not closely related to any other sequenced strains, indicating *de novo* emergence (Figure 3A, B and Supplementary Table 2). One patient appeared to have no SNPs separating the initial isoniazid-resistant and subsequent MDR-TB isolate. However, on closer inspection a mixed-call was

detected in *rpoB* at codon 445 with a His to Tyr substitution accounting for 70% of sequencing reads, below the 90% cut-off used for SNP calling (Table 1, Pt080). In two cases SNP difference between the initial isoniazid-resistant and the MDR-TB isolates were six and seven SNPs respectively, thus not clearly distinguishing *de novo* acquisition from reinfection. In the remaining four patients the initial isoniazid-resistant and MDR-TB isolates were separated by 19, 43, 896 and 1036 SNPs respectively, indicating reinfection (Figure 3A, B and Supplementary Table 2), whilst for two patients WGS indicated a mixture of strains in their second clinical isolate, with at least one of the strains in each mixture being MDR. The initial isoniazid-resistant isolate was not present at the later time-point in either sample. Six of 14 patients were therefore deemed to have been reinfected with MDR-TB (43%).

Of the five patients who initially had susceptible disease and were later diagnosed with MDR-TB, SNP distances between paired isolates ranged from 69 to 1077, indicating reinfection in each instance. Overall, we therefore found that MDR-TB emerged *de novo* in 6/239 (2.5%) patients who were diagnosed with isoniazid-resistant TB by MODS and MGIT, and in 0/35 patients whose strains initially tested isoniazid-resistant by MODS only (Figure 3A, B and Supplementary Table 2).

De novo emergence or selection of *M. tuberculosis* variant with rifampicin resistant mutations during the emergence of MDR-TB

For 5/6 patients with *de novo* emergence of MDR-TB, mutations known to confer resistance to isoniazid (*katG* S315T (in 4/6 cases, 66.66%) and *fabG1* C-15T (in 1/6 case, 16.66%)) and to streptomycin (*rpsL* K43R and K88R) were detected in the original isolates (Table 1). In the remaining patient, the isoniazid and streptomycin phenotypic resistant isolate had pre-existing known pyrazinamide-resistant mutations in the genes *rpsA* and *pncA*, but lacked any known isoniazid or streptomycin resistant mutations, so was probably a resistant phenotype linked to unknown genetic variants. One patient also had an *embB* mutation at the outset, although the ethambutol phenotype was susceptible (Pt072, Table 1). In 155 patients without emergence of MDR-TB, 111 had *katG* S315T (71.61%), 6 had *fabG1* C-15T (3.87%) and rest lacked any known isoniazid-resistant mutations. No significant difference of these mutation frequencies from the strains in which *de novo* MDR-TB emerged ($p = 0.25$, Fisher's exact test). In each of the six *de novo* MDR-TB cases, known rifampicin-resistant mutations emerged in subsequent isolates (S450L, H445Y and D435V) (Table 1). The proportion of sequencing reads containing either the relevant *rpoB* mutation or wildtype could be assessed at different time intervals in the

six patients. One month into treatment the resistant allele accounted for as few as 10% of reads in one patient and for over 90% of reads in another patient's isolate, although in the former case the phenotype did not convert to 'resistant' until the number of resistant alleles had grown further to 90% at 12 months (Table 1). For four patients the resistant *rpoB* allele accounted for between 66% – 76% of reads by eight months, below the 90% cut-off used for the variant calling, but sufficient to impact the phenotype and be detected by Mykrobe analysis (Table 1). The emergence of an *embB* mutation resulting in resistance to ethambutol could also be observed in one case (Pt108) after eight months of treatment (Table 1A). Three other non-synonymous mutations also emerged, in hypothetical protein *Rv1444c* (M109V) and *Rv3806c*/*ubiA* (I162L) in Pt078 and hypothetical protein Rv2472 (C84R) in Pt102 (Table 1). *ubiA* has previously been linked to ethambutol resistance, although it did not result in a phenotypic change on this occasion [23].

For the two patients with intermediate SNP distances between their first and subsequent isolates, known rifampicin-resistant mutations emerged, and in one case an ethambutol resistant mutation also emerged along with a corresponding resistant phenotype (Pt079). Two different rifampicin-resistant variants were observed in patient (Pt079) (Table 1).

Out of eight patients with *de novo* MDR-TB emergence or an intermediate SNP distance between isolates, five received 2SHZR/6HE, two received 2RHZE/6HE and one received 2SRHZ/1RHZ/5HE as treatment regimens (Table 1).

For 9/11 patients with MDR-TB reinfection but no mixed reads in their MDR-TB isolates, all reinfections were of lineage 2.2.1 with mutation in EsxW-Thr2Ala. This was the same lineage as the initial infection for five patients whereas the other four were initially infected with strains from lineages 1.1.1.1, 4.8, 4.1.2 and 4.5 (Table 2). The overall prevalence of lineage 2.2.1 among MDR-TB isolates was 79% and 71% among isoniazid-resistant and susceptible isolates.

There were no instances where rifampicin-resistant alleles were detected in the initial *M. tuberculosis* isolates of either patients who later went on to evolve MDR-TB *de novo* or due to reinfection at sequencing depth of 30x.

DISCUSSION

Here we provide genetic evidence for the *de novo* emergence of MDR-TB among patients treated with first-line drugs for isoniazid-resistant TB. Contrary to previous studies that found MDR-TB to be the consequence of reinfection [8, 9], *de novo* emergence of MDR-TB was equally common to reinfection with a separate MDR-TB strain among patients with pre-existing isoniazid-resistant TB.

Our findings support the conclusions from recent studies indicating the risk of prior isoniazid resistance in the evolution of rifampicin resistance [3, 10]. We observed 6/239 (2.5%) patients with initial isoniazid-resistant TB acquiring MDR-TB *de novo* and 8/239 (3.3%) patients who were either reinfected with a new strain that was MDR, or for whom the results were indeterminate. There was no significant difference in clinical presentations between patients with and without emergence of MDR-TB except for drinking alcohol (Supplementary Table 3).

The isolates from patients in Vietnam are not routinely screened for isoniazid-resistance [2]. This is also true for patients in many other low and middle-income countries. Rapid molecular diagnosis methods are available or under development to improve the detection of antibiotic-resistant TB such as Xpert MTB/RIF Ultra for rifampicin resistance and DNA line-probe assays such as the AID TB Resistance LPA and GenoType MTBDRplus VER2.0 for isoniazid and rifampicin resistance detection [24]. It is well understood that sub-optimal antibiotic regimens can select for resistant mutations in the *M. tuberculosis* population [25]. All the six patients with *de novo* emergence of MDR-TB as well as the two patients with intermediate SNP distances separating their longitudinal isolates were already resistant to streptomycin as well as isoniazid. Two also had mutations conferring resistance to ethambutol leaving rifampicin almost entirely unprotected during the intensive phase, exposing it to selection pressure driving the emergence of rifampicin-resistant variants in the population.

Although treatment regimens for isoniazid-resistant TB have changed to 2RHZE/4RHE since this study recruited, the emergence of rifampicin resistance during the intensive phase of treatment among our study patients is a major concern. In today's regimens it is protected only by ethambutol in the continuation phase in patients with isoniazid resistance. Our findings clearly underscore the need for rapid, comprehensive DST testing and implementation of new World Health Organization guidelines for treating isoniazid-resistant TB with six months of rifampicin, ethambutol, pyrazinamide and levofloxacin [26].

TB endemic countries have a higher risk of mixed or reinfection [27]. Mixed infection is harder to diagnose and patients risk being treated with regimens that select for resistant bacterial populations [28]. Reinfection with MDR-TB is a major concern especially where hospitalization, visit to out-patient departments and DOTS clinic risks exposure to other TB patients [29].

Standard culture based WGS on *M. tuberculosis* isolates cannot rule out the presence of minor-resistant alleles prior to treatment [30]. The early detection of emergence of MDR-TB minor-variants in the patient can help clinicians to appropriately change the treatment regimen [31].

The Beijing sub-lineage 2.2.1 was responsible for each patient who was secondarily infected with MDR-TB, consistent with the high prevalence and observation that Beijing sub-lineage 2.2.1 is involved in enhanced transmission among the host population in Vietnam [32].

There are some limitations in our study. Most importantly, we have only focused on the old eight-month TB treatment regimen that lacks rifampicin in the continuation phase. This was because the strains from a previous study were readily available to us to investigate this important question [2]. This may have decreased the frequency of *de novo* emergence of MDR-TB from isoniazid resistant TB, as there was no rifampicin selection pressure after initial two months of treatment. However, observing resistance emerge during the intensive phase when rifampicin is supposedly protected by more drugs than in the continuation phase is sobering. MTB/RIF Xpert remains the assay of choice in many low and middle-income settings but would no more pick up the resistance to ethambutol, pyrazinamide or second-line injectable drugs now than it would have then. The risks associated with incomplete diagnostics are therefore apparent. A separate weakness is we cannot rule out the possibility of MDR-TB reinfection with an isolate that is related genetically to the initial isolate, for example from a household contact. We also lacked follow up data for the patients whose initial MODS screening result was isoniazid susceptible. This may have underestimated the *de novo* emergence of MDR-TB in patients with a susceptible *M. tuberculosis* isolate.

In conclusion, our study found that *de novo* emergence of MDR-TB in patients with isoniazid-resistant TB occurred equally frequently to reinfection with MDR-TB in this cohort. It is not routine for drugs other than rifampicin to be screened for resistance at diagnosis. This study provides genetic evidence that such a narrow diagnostic focus risks selection for MDR-TB.

Notes

Authors' contributions. SV, VTNH, DNV and NTTT designed the study; PVKT, DTMH, and NHL supervised patients recruitment and sample collection; SV, VTNH, HTH and DDAT carried out experiments; NTTT, SV, VTNH, DNV, PMA, TMW, SJD, GET, and MC analysed and interpreted the results; All authors contributed to manuscript preparation.

Acknowledgments. Authors acknowledge all participants in the study. We would like to thank the physicians and nurses at the District tuberculosis Units, Pham Ngoc Thach Hospital and Oxford University Clinical Research Unit who contributed to the study.

Disclaimer. The study funding agency did not have any role in study design; data collection, analysis, or interpretation; or writing of the manuscript. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit the manuscript for publication.

Funding. This work was supported by the Wellcome Trust (Research Fellowship to MC (081814/Z/06/Z), Wellcome Trust Intermediate Fellowship in Public Health and Tropical Medicine to NTTT (206724/Z/17/Z), and Wellcome Trust Major Overseas Program Funding to GT (106680/B/14/Z) and the Agency for Science, Technology and Research, Biomedical Research Council and the National Health and Medical Research Council (A*STAR/NHMRC grant APP1056689 to SJD).

Potential conflicts of interest. All the authors declare no potential conflicts of interest.

REFERENCES

1. World Health Organization. Global Tuberculosis Report. 2019. Geneva: World Health Organization, 2019.

2. Thai PVK, Ha DTM, Hanh NT, et al. Bacterial risk factors for treatment failure and relapse among patients with isoniazid resistant tuberculosis. BMC Infect Dis 2018; 18(1):112. doi:10.1186/s12879-018-3033-9

3. Gegia M, Winters N, Benedetti A, et al. Treatment of isoniazid-resistant tuberculosis with first-line drugs: a systematic review and meta-analysis. Lancet Infect Dis 2017; 17(2):223-34. doi:10.1016/S1473-3099(16)30407-8

4. Floyd K, Glaziou P, Zumla A, et al. The global tuberculosis epidemic and progress in care, prevention, and research: an overview in year 3 of the End TB era. Lancet Respir Med 2018; 6(4):299-314. doi: 10.1016/S2213-2600(18)30057-2

5. Caminero JA, Sotgiu G, Zumla A, et al. Best drug treatment for multidrug-resistant and extensively drug-resistant tuberculosis. Lancet Infect Dis 2010; 10(9): 621-9. doi: 10.1016/S2213-2600(18)30057-2

6. Cohen KA, Abeel T, Manson McGuire A, et al. Evolution of Extensively Drug-Resistant Tuberculosis over Four Decades: Whole Genome Sequencing and Dating Analysis of *Mycobacterium tuberculosis* Isolates from KwaZulu-Natal. PLoS Med 2015; 12(9):e1001880. doi: 10.1371/journal.pmed.1001880

7. Brigden G, Nhung NV, Skrahina A, et al. Advances in clinical trial design for development of new TB treatments-Translating international tuberculosis treatment guidelines into national strategic plans: Experiences from Belarus, South Africa, and Vietnam. PLoS Med. 2019;16(10):e1002896. doi: 10.1371/journal.pmed.1002896

8. Nsofor CA, Jiang Q, Wu J, et al. Transmission is a Noticeable Cause of Resistance Among Treated Tuberculosis Patients in Shanghai, China.Sci Rep 2017; 7(1):7691. doi: 10.1038/s41598-017-08061-3

9. Andrews JR, Gandhi NR, Moodley P, et al. Exogenous reinfection as a cause of multidrug-resistant and extensively drug-resistant tuberculosis in rural South Africa. J Infect Dis 2008; 198(11):1582-9. doi: 10.1086/592991

10. Manson AL, Cohen KA, Abeel T, et al. Genomic analysis of globally diverse *Mycobacterium tuberculosis* strains provides insights into the emergence and spread of multidrug resistance. Nat Genet 2017; 49(3): 395-402. doi: 10.1038/ng.3767

11. Witney AA, Bateson AL, Jindani A, et al. Use of whole-genome sequencing to distinguish relapse from reinfection in a completed tuberculosis clinical trial. BMC Med 2017; 15(1):71. doi: 10.1186/s12916-017-0834-4

12. Bryant JM, Harris SR, Parkhill J, et al. Whole-genome sequencing to establish relapse or re-infection with *Mycobacterium tuberculosis*: a retrospective observational study. Lancet Respir Med 2013; 1(10):786-92. doi.org/10.1016/S2213-2600(13)70231-5

13. Merker M, Kohl TA, Roetzer A, et al. Whole genome sequencing reveals complex evolution patterns of multidrug-resistant *Mycobacterium tuberculosis* Beijing strains in patients. PLoS One 2013; 8(12):e82551. doi: 10.1371/journal.pone.0082551

14. Moore DA,Evans CA,Gilman RH, et al. Microscopic-observation drug-susceptibility assay for the diagnosis of TB. N Engl J Med 2006; 355(15):1539-50. doi: 10.1056/NEJMoa055524

15. WHO. Implementing the WHO Stop TB Strategy: a handbook for national tuberculosis control programmes. Geneva: World Health Organization 2008.

16. Vijay S, Vinh DN, Hai HT, et al. Influence of Stress and Antibiotic Resistance on Cell-Length Distribution in *Mycobacterium tuberculosis* Clinical Isolates. Front Microbiol 2017; 8:2296. doi: 10.3389/fmicb.2017.02296

17. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:13033997 2013.

18. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010; 20(9):1297-303. doi: 10.1101/gr.107524.110

19. Dallman T, Ashton P, Schafer U, et al. SnapperDB: a database solution for routine sequencing analysis of bacterial isolates. Bioinformatics 2018; 34(17):3028-29. doi: 10.1093/bioinformatics/bty212

20. Nguyen LT, Schmidt HA, von Haeseler A, et al. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015; 32(1):268-74. doi: 10.1093/molbev/msu300

21. Bradley P, Gordon NC, Walker TM, et al. Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. Nat Commun 2015; 6:10063. doi: 10.1038/ncomms10063

22. Ford CB, Shah RR, Maeda MK, et al. *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. Nat Genet 2013; 45(7):784-90. doi: 10.1038/ng.2656

23. Safi H, Lingaraju S, Amin A, et al. Evolution of high-level ethambutol-resistant tuberculosis through interacting mutations in decaprenylphosphoryl-beta-D-arabinose biosynthetic and utilization pathway genes. Nat Genet 2013; 45(10):1190-7. doi: 10.1038/ng.2743

24. Nguyen TNA, Anton-Le Berre V, Banuls AL, et al. Molecular Diagnosis of Drug-Resistant Tuberculosis; A Literature Review. Front Microbiol 2019; 10:794. doi: 10.3389/fmicb.2019.00794

25. Trauner A, Liu Q, Via LE, et al. The within-host population dynamics of *Mycobacterium tuberculosis* vary with treatment efficacy. Genome Biol 2017; 18(1):71. doi: 10.1186/s13059-017-1196-0

26. World Health Organization. WHO treatment guidelines for isoniazid-resistant tuberculosis: supplement to the WHO treatment guidelines for drug-resistant tuberculosis. Geneva: World Health Organization; 2018. Available from: https://www.who.int/tb/publications/2018/ WHO_guidelines_isoniazid_resistant_TB/en/.

27. Tarashi S, Fateh A, Mirsaeidi M, et al. Mixed infections in tuberculosis: The missing part in a puzzle. Tuberculosis (Edinb) 2017; 107:168-74. doi: 10.1016/j.tube.2017.09.004

28. Hingley-Wilson SM, Casey R, Connell D, et al. Undetected multidrug-resistant tuberculosis amplified by first-line therapy in mixed infection. Emerg Infect Dis 2013; 19(7):1138-41. doi: 10.3201/1907.130313

29. Crudu V, Merker M, Lange C, et al. Nosocomial transmission of multidrug-resistant tuberculosis. Int J Tuberc Lung Dis 2015; 19(12):1520-3. doi: 10.5588/ijtld.15.0327

30. Nimmo C, Shaw LP, Doyle R, et al. Whole genome sequencing *Mycobacterium tuberculosis* directly from sputum identifies more genetic diversity than sequencing from culture. BMC Genomics 2019; 20(1):389. doi: 10.1186/s12864-019-5782-2

31. Colman RE, Schupp JM, Hicks ND, et al. Detection of Low-Level Mixed-Population Drug Resistance in *Mycobacterium tuberculosis* Using High Fidelity Amplicon Sequencing. PLoS One 2015; 10(5):e0126626. doi: 10.1371/journal.pone.0126626

32. Holt KE, McAdam P, Thai PVK, et al. Frequent transmission of the *Mycobacterium tuberculosis* Beijing lineage and positive selection for the EsxW Beijing variant in Vietnam. Nat Genet 2018; 50(6):849-56. doi: 10.1038/s41588-018-0117-9

Figure legends

Figure 1. Study flow diagram.

Figure 2. Emergence of MDR-TB during treatment in patients. Mapping of phenotypic drug susceptibility testing of longitudinal *M. tuberculosis* isolates at different months (0M, 1M, 2M, 5M, 8M, 12M, 18M and 24M) during treatment or recurrence post-treatment from 101 patients initially with isoniazid-resistant TB and five patients with susceptible TB. MDR-TB emergence grouped at the bottom, confirmed based on phenotypic and genotypic DST. Color code indicates antibiotic susceptibility and no isolate (time points lacking positive *M. tuberculosis* cultures from the patients). 99 patients initially with isoniazid resistant TB had DST results for more than one isolate, whereas two patients had DST for only initial 0M isolate as later isolates failed to revive during sub-culture.

Figure 3. Genetic distance between initial isoniazid-resistant (INH-R) or susceptible isolate and MDR-TB isolates. (A) Phylogenetic tree of longitudinal *M. tuberculosis* isolates. Emergences of MDR-TB in the phylogenetic tree are indicated in the adjacent panel by patient code, location number in the phylogenetic tree and collection time points (in months). Patients are grouped based on SNPs difference between initial and MDR-TB isolates, equal or less than five SNPs (*de novo*), 6 to 10 SNPs (intermediate) and more than 10 SNPs (reinfection) of emergence of MDR-TB from patients initially with isoniazid or susceptible TB (color code indicates antibiotic susceptibility). Genetically related isolates from the same patient at different time points are indicated by blue bars in the phylogenetic tree at the respective location number and blue square highlighting the respective collection time points, genetically unrelated isolates at different time points from the same patient are indicated by red bars in the phylogenetic tree at respective location number. Location numbers for isolates from a patient follow the order of collection time point, related isolates from the same patient are given single location number. Outer ring around the phylogenetic tree indicates different *M. tuberculosis* lineages by color code. * Patient with 19 SNPs difference between initial isoniazid resistant and MDR-TB isolates, ** Patients with mixed infection removed from phylogenetic tree but analysed manually. (B) SNP distance or difference between the initial and the first MDR-TB isolate pair in patients initially with susceptible (SNPs range 69 - 1077) or isoniazid resistant isolate (INH-R, SNPs range 1 to 1036). **Note : one patient had zero SNP difference between the initial isoniazid resistant and MDR-TB isolate and that data point is not shown in the graph**. Black line indicates five SNPs cut-off.

Case ID	Preexisting antibiotic resistant mutations (WGS)	Preexisting antibiotic resistant phenotype (MGIT)	Antibiotic resistant mutations emerged in MDR-TB (Month, % genetic variant)	Emerging antibiotic resistant phenotype (Month)	Other mutations (Month, % genetic variant)	*M. tuberculosis* sub-lineage	Lineage specific SNP	Treatment regimen
Pt072	***katG S315T*** ***rpsL K43R*** ***embB M306I***	**INH STR**	***rpoB H445Y*** **(1M, 2M > 90%, 5M = 74%)**	**RIF (1M, 2M, 5M)**		2.2.1.1	*embB (D534D)*	2RHZE/6HE
Pt078	***rpsA*** **V260I** ***pncA*** **C14R**	**INH STR**	***rpoB*** **H445Y (5M = 66%, 8M > 90%)**	**RIF (5M, 8M)**	*Rv1444c* (M109V): hypothetical protein, (5M = 62%, 8M > 90%) *Rv3806c* (I162L), *ubiA** (8M > 90%)	1.1	*Rv3915* (L352L)	2SRHZ/RHZ/5HE
Pt080	***katG*** **S315T** ***rpsL*** **K88R**	**INH STR**	***rpoB*** **H445Y (8M = 70%)**	**RIF (8M)**		2.2.1	*Rv0697* (L268L)	2SHRZ/6HE
Pt102	***katG*** **S315T** ***rpsL*** **K43R**	**INH STR**	***rpoB*** **S450L (1M =10%, 12M, 18M > 90%)**	**RIF (12M, 18M)**	*Rv2472* (C84R) Hypothetical protein, (0M = 73%, 12M, 18M >90%)	2.2.1.1	*embB* (D534D)	2RHZE/6HE
Pt108	***katG*** **S315T** ***rpsL*** **K43R**	**INH STR**	***rpoB*** **D435V (2M = 76%, 8M,12M, 18M, 24M > 90%)** ***embB*** **M306V (8M,12M, 18M , 24M >90%)**	**RIF (2M, 8M, 12M, 18M, 24M) EMB (8M, 12M)**		2.2.1	*Rv0697* (L268L)	2SRHZ/6HE
Pt152	***fabG1*** **C-15T** ***rpsL*** **K88R**	**INH STR**	***rpoB*** **D435V (24M > 90%)**	**RIF (24M)**		4.5	*Rv1524* (P344P)	2SHZR/6HE
Patients with intermediate SNPs difference								
Pt061	***katG S315T*** ***embB M306I*** ***rpsL K43R***	**INH STR**	***rpoB*** **S450L (2M = 20%)**	**RIF (2M)**	NADH pyrophosphatase *nudC P239R* (0M= 80%, 2M >90%)	2.2	*Rv2231c* (A205A)	2SHRZ/6HE
Pt079	***katG S315T*** ***rpsL K43R***	**INH STR**	***rpoB H445P*** **(1M = 77%)** ***rpoB S450L*** **(8M > 90%),** ***embB Q497R*** **(8M = 88%)**	**RIF (1M, 8M) EMB (8M)**		2.2.1.1	*embB (D534D)*	2SRHZ/6HE

Table 1. Emergence of genetic variants in *de novo* and intermediate emergence of MDR-TB isolates.

*ubiA – gene involved in *M. tuberculosis* cell wall biosynthesis and ethambutol resistance

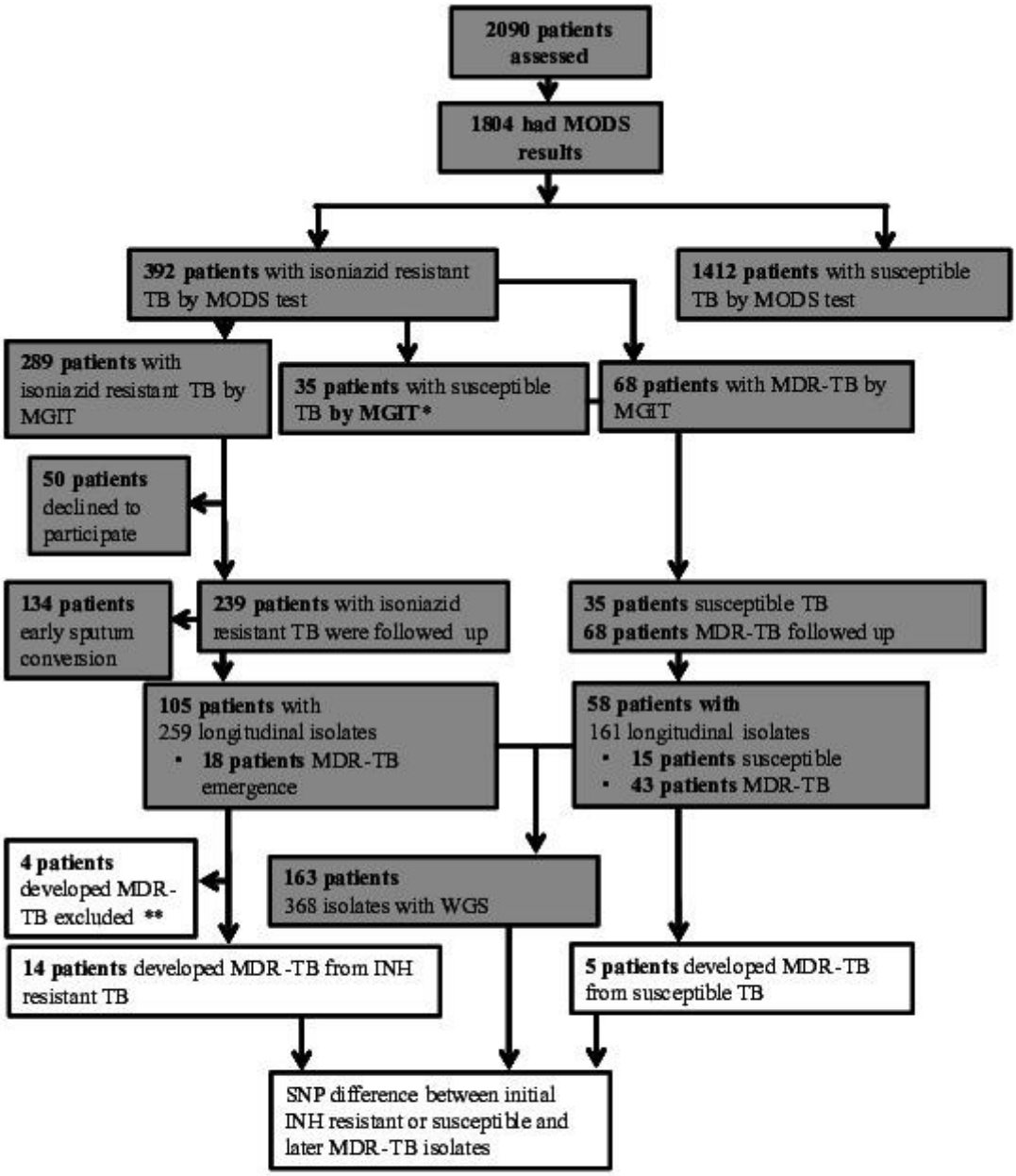
% - Percentage of reads with genetic variant compared to wild type reference.

Bold sections highlight patient code, preexisting and emerging antibiotic resistant variants and phenotypes.

Case ID	sub-lineage of initial *M.tb* isolate (Lineage specific SNPs)	sub-lineage of MDR-TB isolate (Month)
Pt006	2.2.1 (*Rv0697* (L268L))	2.2.1 (2M)
Pt007	2.2.1 (2M)	2.2.1 (5M)
Pt008	2.2.1	2.2.1 (2M , 5M)
Pt010	2.2.1	2.2.1 (1M , 2M, 5M)
Pt012	4.8 (*Rv3417c* (D51D))	2.2.1 (5M , 8M)
Pt013	1.1.1.1 (*Rv2907c*(V113V))	2.2.1 (12 M)
Pt070	4.1.2 (*Rv0798c* (L172L))	2.2.1 (5M)
Pt093	4.5 (*Rv1524*(P344P))	2.2.1 (12M)
Pt151*	2.2.1 (1M)	2.2.1 (5M)

Table 2. Sub-lineages of initial (0M, 1M and 2M) and MDR-TB isolates from secondary infection.
* only 19 SNPs difference between initial and MDR-TB isolate.

Figure 1



*Patients with isoniazid susceptible TB (n = 35) by MGIT were followed-up due to discordance between MODS and MGIT.

Shaded boxes indicate classification of *M. tuberculosis* isolates based only on phenotypic DST, later classification was based on both phenotypic and genotypic DST concordance.

** MDR-TB excluded due to isolates with known rifampicin resistant mutations in gene *rpoB* but rifampicin susceptible MGIT results (n = 3) and lack of WGS for further analysis (n = 1).

Figure 2

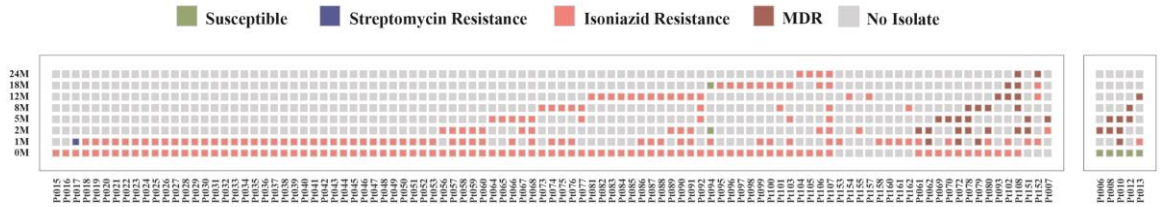


Figure 3