JIDC | THE JOURNAL OF INFECTION IN DEVELOPING COUNTRIES

Original Article

Molecular diversity of *Mycobacterium tuberculosis* isolates in Treatment-experienced Patients from Andhra Pradesh, India

Anagoni Srikar[1], Banda Venkataramana[1], Alladi Mohan[2], PVGK Sarma[3], Kangjam Rekha Devi[4], Kanwar Narain[4], Abhijit Chaudhury[1]

[1] Department of Microbiology, Sri Venkateswara Institute of Medical Sciences, Tirupati, Andhra Pradesh, India
[2] Department of Medicine, Sri Venkateswara Institute of Medical Sciences, Tirupati, Andhra Pradesh, India
[3] Department of Biotechnology, Sri Venkateswara Institute of Medical Sciences, Tirupati, Andhra Pradesh, India
[4] Department of Enteric diseases, Regional Medical Research Centre, Dibrugarh, Assam, India

Abstract

Introduction: To get a comprehensive idea about the transmission and epidemiology of TB globally and locally, the use of molecular typing methods has become imperative not only for understanding genetic diversity but also the population structure of *Mycobacterium tuberculosis* complex (MTBC). We aimed to investigate the drug resistance pattern and genetic diversity of MTBC among previously treated patients with sputum smear-positive pulmonary tuberculosis in a South Indian population.

Methodology: 104 patients with sputum smear positivity and who had previously undergone treatment were selected. Drug susceptibility testing, Spoligotyping, MIRU-VNTR, and SNP typing were performed.

Results: Mono-resistance to isoniazid 16 (15.38%) was the highest among all drugs. Out of 104 isolates, 24 (23%) isolates were classified as MDR strains. The distributions of most common lineages were: EAI3-Ind–20 (19.23%), EAI5-13 (12.50%), Beijing-12 (11.54%), CAS1-Delhi- 9 (8.65%), and 7 (6.73%) each of T-H37rv, Unknown and Orphan types. MIRU-VNTR-based analysis revealed that there are two major groups: CAS1-Delhi and Beijing groups. Out of 104 isolates, 82 belonged to well-defined lineages and 6 clusters, and the remaining 22 were singletons. SNP analysis showed no mutations associated with five sets of genes in 33 strains.

Conclusions: The occurrence of 11.54% Beijing strains in South India is an important finding. High frequency of Isoniazid mono resistance noticed. Spoligotyping along with MIRU-VNTR and SNP typing is the best approach to the identification of strain lineages. No mutation with Antigen85C gene represents, can be used for vaccine candidates.

Key words: Drug resistance; genetic diversity; MIRU-VNTR typing; *Mycobacterium tuberculosis*; Spoligo-international type (SIT); SNP typing.

J Infect Dev Ctries 2023; 17(8):1114-1124. doi:10.3855/jidc.17757

(Received 07 December 2022 – Accepted 18 March 2023)

Introduction

Tuberculosis (TB) is a communicable disease that is a major cause of morbidity; one of the top 10 causes of death worldwide; and the leading cause of death from a single infectious agent, ranking above HIV/AIDS. Globally, an estimated 10 million (range, 9.0–11.1 million) people fell ill with TB in 2018, a number that has been relatively stable in recent years. Geographically, most TB cases in 2018 were in the WHO regions of South-East Asia (44%), Africa (24%), and the Western Pacific (18%). Eight Countries accounted for two-thirds of the global total with India (27%) leading the table [1].

The situation is made more alarming by the serious problem of drug resistance in *M. tuberculosis*, especially among previously treated patients. The First National Anti-Tuberculosis Drug Resistance survey conducted by the Indian Government showed that among patients previously treated for TB, there were high levels of resistance to first-line drugs tested-isoniazid (any 25%, mono-resistance 8%) followed by resistance to streptomycin (any 13%, mono-resistance 2%), pyrazinamide (any 9%, mono-resistance 4%) and ethambutol (any 7%, mono-resistance 0.21%) [2]. As per the latest data from India, overall drug resistance (MDR/RR-TB) is estimated to be 2.8% (2.3-3.5) in new cases is and 14% (11-14) in previously treated patients in 2018 [1].

Molecular Epidemiology helps in improving our understanding of the pathogenesis of the disease and it also provides unique insights into the international dissemination of tuberculosis by the geographic

comparison and evolutionary analysis of this highly widespread pathogen population [3]. Genotyping is used to track the specific isolates of *M. tuberculosis* in a community.

The techniques which have been found to be most useful include IS*6110* based fingerprinting (mostly by Restriction Fragment Length Polymorphism, RFLP); Spacer oligonucleotide typing (Spoligotyping), interspersed repetitive-unit-variable-number tandem-repeats (MIRU-VNTR) typing and Single nucleotide polymorphism (SNPs) typing.

Spoligotyping, which has been used in a number of studies, is based on the polymorphism in the direct repeat (DR) locus, and the presence and absence of spacers will result in different polymorphisms. On the other hand, MIRU-VNTR typing method has more discriminatory power compared to spoligotyping [4]. Presently, the 24 locus MIRU-VNTR set typing is being mostly used, which increases the number of analysed loci thereby improving the discriminatory power and better strain identification [5].

SNP typing is another method that can provide clearer identification compared to spoligotyping and MIRU-VNTR typing [6,7]. Homolka *et al*. developed an SNP-based algorithm for the identification of 17 MTB lineage with high resolution [8] where he used the five most variable genes.

The few studies which are available from India have revealed the apparent preponderance of certain clades or lineages in defined geographic regions like the Northern and Western regions of the country. But the number of isolates included in most of the studies is only a few. Occurrences of major genogroups such as the ancestral (or the TbD1+ type or the East African Indian type and MANU), the Central Asian (CAS), or Delhi type have been previously reported in one study from this area which lies in the Southern part of the country [9]. Three additional recent reports are available from Southern India from the states of Tamilnadu [10], Puducherry [11], and Karnataka [12]. Only one study from Delhi used the 15 locus MIRU-VNTR in combination with spoligotyping and nine SNP markers [13].

Spoligotyping with MIRU-VNTR typing and SNP typing gives more accurate strain identification with high discrimination [4,8]. In addition, to find out Lineage-specific drug resistance, we combined all three methods with drug susceptibility testing in this study. There is no study available from this part of India where all three typing methods for strain identification have been used among previously treated patients with MTB.

Methodology

The study was conducted at Sri Venkateswara Institute of Medical Sciences, Tirupati in Andhra Pradesh, India with the *Mycobacterium tuberculosis* complex (MTBC) strain collected from the accredited Revised National Tuberculosis Control Programme (RNTCP) Culture and Drug Sensitivity Testing laboratory at our institute. These strains have been isolated by culture of sputum smear positive pulmonary tuberculosis patients attending this hospital as well as from the other hospitals and health care facilities situated in the districts of Chittoor, Kadapa and Nellore in Andhra Pradesh, India.

Staining

Ziehl-Neelsen staining was performed with the sputum and scored according to the RNTCP guidelines.

Bacterial isolation, culture, and identification:

In this study, a total of 104 isolates of *M. tuberculosis* complex were selected for Drug susceptibility and genotyping. In order to be representative of MTBC strains circulating in south India, we collected a total of 104 AFB positive sputum samples over a period of one year (September 2018 to August 2019). All clinical samples were decontaminated using modified Petroff's method and cultured in conventional Lowenstein-Jensen (LJ) media at 37 °C for 6 to 8 weeks. MTBC was identified and confirmed by means of standard microbiological and biochemical methods [14].

Drug Sensitivity Testing

The sensitivity of the isolates to the first line of anti-tuberculosis drugs was done by proportion method for rifampicin (RIF, 40.0 μg/mL), ethambutol (ETH, 2.0 μg/mL), streptomycin (STR, 4.0 μg/mL), and isoniazid (INH, 0.2 μg/mL) according to the protocol of National Institute for Research in Tuberculosis (previously, Tuberculosis Research Centre), Chennai, India [15]. If a strain was found to be resistant to Isoniazid and Rifampicin with or without resistance to other drugs, it was defined as Multi-drug resistant (MDR) [16].

DNA extraction

DNA was extracted from fresh sub-cultures of the 104 MTBC isolates using cetyl-trimethyl ammonium bromide (CTAB) method [14]. DNA was eluted in TE buffer (pH 8.0) and stored at -80 °C until used. DNA purity and concentration were determined spectrophotometrically [17].

Spoligotyping

Spoligotyping was carried out by amplifying the whole Direct Repeat (DR) region using the commercially available kit (Ocimum BioSolutions, India) according to the standardized method [6] using the designated primer pairs of DRa and DRb (DRa, 5'-GGTTTTGGGTCTGACGAC-3' (biotinylated 5' end) and DRb, 5'-CCGAGAGGGGACGGAAAC-3'). After DNA amplification, amplified PCR products were hybridized with covalently linked 43 spacer oligonucleotides on nitrocellulose membrane as per the manufacturer's instructions. The hybridized fragments were identified using enhanced chemiluminescence system. The spoligotype representation was initially reported as 43 digits binary representation of 43 spacers and one (1) was scored for positive hybridization and zero (0) for negative hybridization.

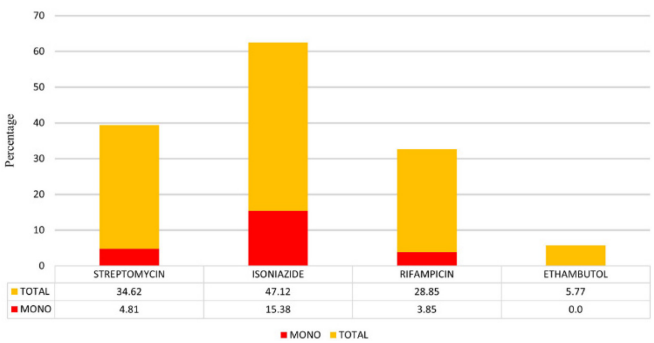
MIRU-VNTR typing

24 locus set MIRU-VNTR typing markers, comprised of 12 MIRU units, 3 Exact tandem repeats (ETRs), 6 Mtub group, and 3 Queen's University Belfast group (QUB) [18–20]. PCR condition and composition were followed as previously described [19]. Gel electrophoresis was performed for all PCR products, and molecular weight was measured using GelQuant Express software. Allelic copy number was assigned according to the standard protocol proposed by Supply *et al* [19].

SNP typing

Total 33 unique isolates from Spoligotyping and MIRU-VNTR typing was selected for SNP typing. For DNA sequence analysis 5 genes (RV0129C, RV1009, RV0557, RV1811, RV2628) were amplified by PCR. Primer sequences and amplification targets are summarized in Supplementary Table 1. Sanger dideoxy sequencing method was used for sequencing with both forward and reverse sequences.

Figure 1. Comparison of total and mono resistant among 104 isolates. All values are expressed in percentages (%).



Statistical analysis

The Drug resistance and genotypes were arranged in Excel spreadsheets. All spoligotype data were arranged in binary and 15-digit octal code by using SPOLTOOL. Spoligotyping and MIRU-VNTR data were submitted to the international databases MIRU-VNTRplus (http://www.miru-vntrplus.org) and SITVIT 2 Database (SPOLDB 4.0). Neighbour-joining (NJ) tree constructed using 24 loci MIRU-VNTR data. Cluster analysis was performed and a dendrogram was generated in BioNumerics software (Version 7.6; Applied maths) using the Dice similarity coefficient and UPGMA (Unweighted Pair Group Method with Arithmetic Mean) coefficient. Minimum spanning tree were calculated with spoligotype results from SITVIT Database by using BioNumerics software. Hunter Gaston Discriminatory Index (HGDI) was calculated for the determination of Discriminatory power for MIRU-VNTR typing by using an online tool (http://insilico.ehu.es/mini_tools/discriminatory_power/index.php). Sequenced data analysed by using SeqScape software from applied by systems.

Results

A total of 104 patients with sputum smear positive pulmonary tuberculosis were included in this study. The 104 patients belonged to all age groups (10-90 years) with a mean age of 50 years; the most common being in 61-70 years range (27.88%). Females were 41 (39.42%) and males 63 (60.57%).

Drug susceptibility testing

Among the 104 patients with TB, the overall proportion of drug resistance to INH, SM, RIF, and EMB were 49 (47.12%), 36 (34.62%), 30 (28.85%), and 6 (5.77%) respectively. The proportion of mono resistance to INH, SM, RIF, and EMB was 16 (15.38%), 5 (4.81%), 4 (3.85%), and 0 (0%) respectively (Figure 1). A significant difference ($p < 0.0001$) was found between rifampicin and isoniazid resistant pattern. There were 2 isolates resistant to all four first-line antituberculosis drugs (INH, RIF, SM, and EMB), 4 isolates were resistant to INH, RIF, and EMB while 20 isolates were resistant to SM, RIF, and INH. Out of 104 isolates, 24 (23%) isolates were classified as MDR strains, being resistant to both INH and rifampicin, the two most powerful anti-TB drugs [16].

Spoligotyping

Spoligotyping results showed 33 different spoligotyping patterns. Out of 104 isolates, 97 belonged

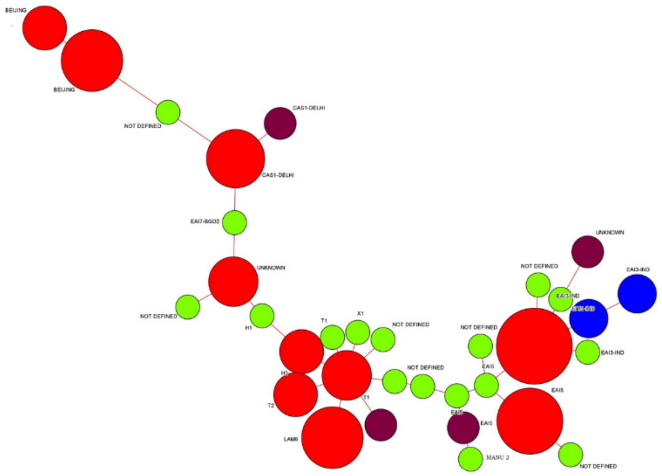
to well-defined lineages, while the remaining 7 were not assigned to any lineages i.e., orphan strains. Lineages were identified and assigned using the international database MIRU-VNTRplus (www.miru-vntrplus.org) and SITVIT2 WEB (http://www.pasteur-uadeloupe.fr:8081/SITVIT2_ONLINE/) online tool. Two lineages (MANU2 & EAI3-IND) were not assigned to any SIT number by SITVIT database; probably they were newly recognised lineages. The distributions of the most common lineages were: EAI3-Ind:-20 (19.23%), EAI5- 13 (12.50%), Beijing-12 (11.54%), CAS1-Delhi- 9 (8.65%) and 7 (6.73%) each of T-H37rv, Unknown, and Orphan. The highest among all lineages was EAI family accounting for 33%. Among the EAI family, sub lineages were EAI3-Ind, EAI5 and EAI7-BGD2 comprising 20 (19.23%), 13 (12.50%), and 1 (1%) strains respectively. There were 9 different spoligotypes found among EAI family and all spoligotypes were assigned by Spoligotype database (SIT11, SIT126, SIT236, SIT340, SIT458, SIT473, SIT1391, SIT1680 and SIT1948).

Cluster analysis showed that among 104 isolates, 82 isolates belonged to 6 defined clusters while the remaining 22 were singletons (Figure 2). Cluster 1 shared 31 isolates divided into 7 lineages, Cluster 2 shared 17 isolates into 1 lineage, Cluster 3 shared 5 isolates into 2 lineages, Cluster 4 shared 12 isolates into 1 lineage, Cluster 5 shared 10 isolates into 2 lineages, and Cluster 6 shared 6 isolates into 2 lineages.

MIRU-VNTR typing

Allelic profile and discrimination were summarized in Table 1. HGDI of 24-loci MIRU-VNTR revealed that out of 24-loci, twelve loci (MIRU10, MIRU16, MIRU23, MIRU26, MIRU31, MIRU39, Mtub21, Mtub30, Mtub39, QUB11b, QUB26, QUB4156 and ETRA) showed high discriminatory power (HGDI > 0.7). Eleven loci (MIRU04, MIRU23, MIRU24, MIRU20, MIRU27, MIRU40, Mtub29, Mtub30, Mtub34, ETRB, and ETRC) were found to be moderately discriminative (HGDI varies from 0.3 to 0.6), and one locus (MIRU02) were poorly discriminative (HGDI < 0.3).

Figure 2. Minimum spanning tree based on the diversity of Spoligotyping data. The size of each circle is proportional to the number of Spoligotypes belonging to a particular complex.



LAM: Latin American Mediterranean; CAS: Central Asian Strain-; EAI: East African Indian EAI-BGD2: East African Indian- Bangladesh variant; H: Haarlem.

Table 1. The allelic profiles and Hunter Gaston Discriminatory index (HGDI) of each of the 24 MIRU-VNTR loci in MTB isolates (n = 104).

			Copy number of Tandem repeats: EAI / NON EAI family																	
S.NO	**ALIAS**	**LOCUS**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**EAI FAMILY**	**NON-EAI FAMILY**	**TOTAL**
1	**MIRU 02**	**154**	0	0/2	33/68	0	1/0	0	0	0	0	0	0	0	0	0	0	0	0.0588	0.0563	0.0569
2	**MTBU 04**	**424**	0	1/2	31/38	0/9	2/14	0/7	0	0	0	0	0	0	0	0	0	0	0.1693	0.6472	0.5284
3	**ETR-C**	**577**	0	0/1	1/9	4/18	29/40	0	0/2	0	0	0	0	0	0	0	0	0	0.2656	0.5983	0.5103
4	**MIRU 04**	**580**	0	0	2/44	0/6	2/13	29/5	0/1	1/0	0	0/1	0	0	0	0	0	0	0.2727	0.5656	0.6796
5	**MIRU 40**	**802**	0	0	14/26	17/29	3/5	0/8	0	0/2	0	0	0	0	0	0	0	0	0.5900	0.6812	0.6505
6	**MIRU 10**	**960**	0/2	0/2	2/2	4/15	11/21	5/21	9/6	1/1	2/0	0	0	0	0	0	0	0	0.8057	0.7752	0.7933
7	**MIRU 16**	**1644**	0	1/12	3/7	20/26	9/14	1/11	0	0	0	0	0	0	0	0	0	0	0.5918	0.7689	0.7242
8	**MTUB 21**	**1955**	0	0/7	2/14	0/15	0/6	12/12	8/7	1/2	0	0/5	1/0	0	0	0	0/2	10/0	0.7504	0.8629	0.8691
9	**MIRU 20**	**2059**	0	0/3	20/54	14/13	0	0	0	0	0	0	0	0	0	0	0	0	0.4991	0.3739	0.4296
10	**QUB 11B**	**2163b**	0	7/4	8/12	14/16	1/10	2/1	2/19	0/2	0/6	0	0	0	0	0	0	0	0.7469	0.8244	0.8200
11	**ETR-A**	**2165**	0	0	0/14	2/12	0/31	0/1	21/5	8/1	1/0	0/1	2/1	0/4	0	0	0	0	0.5722	0.7358	0.8101
12	**MTUB 29**	**2347**	0	0	0/5	12/8	22/26	0/31	0	0	0	0	0	0	0	0	0	0	0.4706	0.6571	0.6652
13	**MTUB-30**	**2401**	0	18/9	2/30	0/8	12/23	0	0	2/0	0	0	0	0	0	0	0	0	0.6061	0.6886	0.7254
14	**ETR-B**	**2461**	0	7/3	8/46	2/8	5/12	12/1	0	0	0	0	0	0	0	0	0	0	0.7754	0.5313	0.6761
15	**MIRU 23**	**2531**	0	0	0	0/5	2/8	0/41	22/14	0	0	0	10/2	0	0	0	0	0	0.5062	0.6066	0.7067
16	**MIRU 24**	**2687**	0/4	12/43	22/10	0/13	0	0	0	0	0	0	0	0	0	0	0	0	0.4706	0.5727	0.6145
17	**MIRU 26**	**2996**	0	0/5	18/14	13/0	0	3/15	0/24	0/10	0/2	0	0	0	0	0	0	0	0.5829	0.7814	0.8023
18	**MIRU 27**	**3007**	0	2/1	13/14	9/28	10/27	0	0	0	0	0	0	0	0	0	0	0	0.7148	0.6605	0.6852
19	**MTUB-34**	**3171**	0	0/2	3/11	21/40	10/10	0	0	0	0	0	0	0	0	0	0/7	0	0.5401	0.6265	0.6018
20	**MIRU 31**	**3192**	0	0	0/10	12/20	7/21	5/12	10/7	0	0	0	0	0	0	0	0	0	0.7469	0.7797	0.7776
21	**MTUB-39**	**3690**	0	0/8	3/3	15/20	3/24	9/9	3/6	1/0	0	0	0	0	0	0	0	0	0.7326	0.7731	0.7801
22	**QUB 26**	**4052**	0	0/2	2/4	1/8	3/0	3/3	7/11	6/9	8/20	4/13	0	0	0	0	0	0	0.8627	0.8356	0.8428
23	**QUB 4156**	**4156**	3/1	20/18	6/31	5/16	0/4	0	0	0	0	0	0	0	0	0	0	0	0.6114	0.6919	0.7029
24	**MIRU 39**	**4348**	0/2	0/2	1/9	21/51	11/6	0	1/0	0	0	0	0	0	0	0	0	0	0.5276	0.4501	0.4886

We analysed HGDI of individual alleles of 24 loci MIRU-VNTRs for EAI family and Non EAI family strains. The discriminatory power of various MIRU-VNTR alleles was lower in EAI family strains as compared to Non EAI family strains. The seven MIRU-VNTRs alleles showed lower discriminatory power such as MIRU04, MIRU16, MIRU26, Mtub04, Mtub29, ETRA, ETRC. Among seven alleles MIRU04, Mtub04, ETRC showed very lowest discriminatory power in EAI strains as compared to Non EAI strains.

The clustering rate of Spoligotyping and MIRU-VNTR according to MIRU VNTRplus database is 0.68 and 0.17 respectively. Cluster rate was low in MIRU-VNTR typing because 100 isolates have a unique pattern. The clustering rate was almost similar when compared to 24 locus MIRU-VNTR with 15 locus MIRU-VNTRs typing (0.14). However, for 12 locus MIRU-VNTRs typing was low (0.36).

Phylogenetic analysis

Genetic relatedness among 104 isolates was determined by using Neighbour-joining (NJ) tree constructed using 24-loci MIRU-VNTRs typing (Figure 3). Analysis revealed that there are two major groups, one compromising CAS1-Delhi and Beijing lineages and the other group containing the remaining strains (Family of T, EAI, Haarlem, Unknown, etc.).

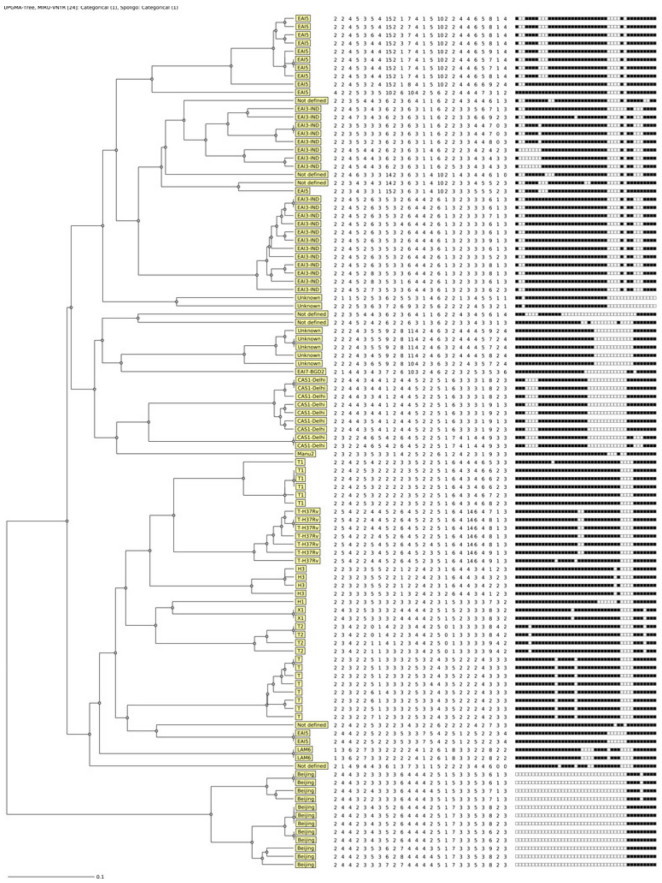
Novel strains

Among 104 isolates, 11 isolates showed 10 different spoligotyping patterns that have not been reported previously in spoligotyping database (i.e. Orphan or new). Among these 11 isolates, 4 were clustered isolates, while the remaining 7 were unique and singletons (non-clustered isolates) (Table 2). Based on 24 Loci MIRU-VNTRs NJ tree analysis, out of 11 orphan or Novel isolates in our study, 10 strains were similar to EAI family while one strain was related to the Beijing type.

SNP typing

A total of 33 unique isolates were chosen for sequencing of five selected Genes. These included Beijing -12, T-H37RV -7, Unknown-7, and Not defined-7. The genes include Rv0129c (1023 bp), Rv0557 (1137 bp), Rv1009 (1089 bp), Rv1811 (705 bp), Rv2628 (363 bp). No mutation was observed

Figure 3. Neighbour-Joining (NJ) tree showing genetic diversity of 104 MTB isolates based on 24 locus MIRU-VNTR loci data corresponding Spoligotypes.



Filled boxes represent positive hybridization while empty boxes represent absence of spacers.

Table 2. Orphan strains that are newly found in our study. Some lineages assigned by Spol DB 4.0. Some lineages not assigned.

Strain number	SIT	Octal code	Pattern	Lineage	Cluster
003	Orphan or new	477477777413711		Not defined	Singleton
004	Orphan or new	703760000000371		Not defined	Singleton
009	Orphan or new	777777777733771		Not defined	Cluster 7
021	Orphan or new	777777620020771		Not defined	Singleton
033	Orphan or new	477077777413071		Not defined	Singleton
041	Orphan or new	777777777423771		Manu2	Cluster 3
048	Orphan or new	477777377413071		EAI3-IND	Cluster 2
052	Orphan or new	777757347760771		Not defined	Singleton
055	Orphan or new	474377717413771		Not defined	Cluster 5
060 & 090	Orphan or new	677777777400000		Unknown	Singleton

SIT: Spoligo international type. Filled boxes represent positive hybridization while empty boxes represent absence of spacers.

among 33 samples after both forward and reverse sequencing.

Lineages and Drug resistance

There were 2 Beijing strains resistant to all four first line antituberculosis drugs. The distribution of lineages among 24 MDR strains were: EAI3-IND - 6 (25%), Beijing- 5 (20.83%), T1 -5 (20.83%), EAI5- 4 (16.67%), CAS1-Delhi- 3 (12.50%) and Manu1- 1 (4.17%). In the remaining i.e., T, T2, X1, H1, H3, EAI7-BGD2, LAM6, Unknown and Orphan lineages, no MDR strains were found. Among Isoniazid mono-resistant strains, the distribution was found to be T–5 (31.25%), EAI3-IND- 4 (25%), Orphan -2(12.5%), T-H37rv- 2 (12.5%) and EAI5, LAM6, T1 lineages- 1 (6.25%) each. Mono resistance in Rifampicin was noted in only 4 (3.84%). Interestingly all mono resistant rifampicin isolates belonged to H3 Lineage in our study (Table 3). All lineages with spoligotyping, MIRU-VNTR typing, and antibiotic susceptibility pattern is given in Supplementary Table 2.

Discussion

MDR-TB is caused by *M. tuberculosis* and is defined as TB that is resistant to at least isoniazid and rifampicin, the two most effective first-line, anti-TB drugs [1]. As reported by WHO [1], a global total of 1,86,772 cases of MDR/RR-TB were detected and notified in 2018, up from 1,60,684 in 2017. The situation is alarming in India with an incidence rate of 9.6 (5.7–15)/100,000 population. In India, the estimated proportion of TB cases with RR/MDR in 2018 was 14%. The present study showed the prevalence of drug-resistant TB in previously treated patients to be 58%. Among them, Multidrug-resistant TB (MDR TB) was 23%. These results are comparable to some previous studies in India. A study from Gujarat in the year 2009 reported 17.4% MDR-TB [21]. A study from Madhya Pradesh., in the year 2015 reported it to be 8.2% [22]. A study from Delhi in 2011 reported a higher prevalence of 33% MDR-TB in previously treated patients [23]. High levels of MDR-TB among previously treated patients can be due to inadequate understanding of the end-user about the drug regimens, poor patient adherence to treatment, or substandard quality of the drug.

The detection of 33 different spoligotype lineages reflects the high rate of secondary or acquired resistance in this population. East African Indian (EAI), the most ancient *M. tuberculosis* lineage, was first described in Guinea-Bissau [24], which is prevalent in the African and Indian subcontinent. In our study, EAI family prevalence was 33%, of which EAI3-Ind were 20 (19.23%) EAI5- 13 (12.50%) and EAI7-BGD2-1(1%). A similar study in North India in 2017 reported 19.10% EAI prevalence [25]. Among the 34 (35%) EAI family isolates, 24 (70.6%) were from males and 10 (29.4%) were from females. A similar result was arrived in the studies from Vellore and Delhi, India [10, 26] where EAI family was more common among the male population compared to females. This may be due to the greater proportion of males in the study population. Among 34 EAI family isolates, 15 (44.1%) isolates were sensitive to all drugs while the remaining 19 (55.9%) were drug-resistant strains.

M. tuberculosis Beijing strain is of East Asian lineage which originated in China this strain is now present in most parts of world and is associated with drug resistance as reported in many studies [27,28]. But one study from China reported that Beijing strains may

Table 3. Comparison of Drug susceptible results among different lineages.

Lineages	No. Of isolates	Streptomycin		Isoniazid		Rifampicin		Ethambutol		MDR
		Sensitive	Resistant	Sensitive	Resistant	Sensitive	Resistant	Sensitive	Resistant	
Beijing	12 (11.54%)	8 (66.7%)	4 (33.3%)	7 (58.3%)	5 (41.7%)	7 (58.3%)	5 (41.7%)	7 (58.3%)	5 (41.7%)	5 (20.83%)
CAS1-Delhi	9 (8.65%)	6 (66.7%)	3 (33.3%)	6 (66.7%)	3 (33.3%)	6 (66.7%)	3 (33.3%)	9 (100%)	0 (0%)	3 (12.50%)
EAI3-IND	20 (19.23%)	10 (50%)	10 (50%)	7 (35%)	14 (70%)	14 (70%)	6 (30%)	20 (100%)	0 (0%)	6 (25%)
EAI5	13 (12.50%)	10 (76.9%)	3 (23.1%)	8 (61.5%)	5 (38.5%)	9 (69.2)	4 (30.8%)	12 (92.3%)	1 (7.7%)	4 (16.67%)
EAI7-BGD2	1 (0.96%)	1 (100%	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	0 (0%)
H1	1 (0.96%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	0 (0%)
H3	4 (3.85%)	4 (100%)	0 (0%)	4 (100%)	0 (0%)	0 (0%)	4 (100%)	4 (100%)	0 (0%)	0 (0%)
LAM6	2 (1.92%)	2 (100%)	0 (0%)	1 (50%)	1 (50%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)	0 (0%)
Manu2	1 (0.96%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	0 (0%)	1 (100%)	1 (100%)	0 (0%)	1 (4.17%)
T	8 (7.69%)	5 (62.5%)	3 (37.5%)	0 (0%)	8 (100%)	8 (100%)	0 (0%)	8 (100%)	0 (0%)	0 (0%)
T1	6 (5.77%)	1 (16.7%)	5 (83.3%)	0 (0%)	5 (83.3%)	1 (16.7%)	5 (83.3%)	6 (100%)	0 (0%)	5 (20.83%)
T2	4 (3.85%)	4 (100%)	0 (0%)	4 (100%)	0 (0%)	4 (100%)	0 (0%)	4 (100%)	0 (0%)	0 (0%)
UNKNOWN	7 (6.73%)	2 (28.6%)	5 (71.4%)	5 (71.4%)	2 (28.6%)	5 (71.4%)	2 (28.6%)	7 (100%)	0 (0%)	0 (0%)
X1	2 (1.92%)	0 (0%)	2 (100%)	0 (0%)	2 (100%)	2 (100%)	0 (0%)	2 (100%)	0 (0%)	0 (0%)
T-H37rv	7 (6.73%)	7 (100%)	0 (0%)	5 (71.4%)	2 (28.6%)	7 (100%)	0 (0%)	7 (100%)	0 (0%)	0 (0%)
NOT DEFINED	7 (7.21%)	7 (100%)	0 (0%)	6 (875.7%)	1 (14.3%)	7 (100%)	0 (0%)	7 (100%)	0 (0%)	0 (0%)

LAM: Latin American Mediterranean; **CAS:** Central Asian Strain-; **EAI:** East African Indian; **EAI-BGD2:** East African Indian- Bangladesh variant; **H: Harleem**; ND: Not Defined; MDR: Multi drug resistant.

favor the transmission of disease but not drug resistance [29]. The prevalence of Beijing strains reported from earlier studies has ranged from 2.4 to 14.9% [30]. Low prevalence has been found in Tamil Nadu in South India (3%) [31], and also in North India (3.28%) [25]. Another study from the same region of North India (Uttar Pradesh) reported a prevalence of 10.8% [32]. Higher prevalence has been reported from the Western part of India in a study from Mumbai (18.8%) [33] and from Assam in North-East India (35.4%) [34]. In our study occurrence of the Beijing strain was 11.5% which is higher than that reported from the neighbouring state of Tamil Nadu [31]. There were two different Beijing lineages found in our study, SIT001 and SIT255 with only a single spacer difference (spacer 39). The Beijing isolates belonged to the 4th large Cluster in this study. Five (41.7%) of these isolates were sensitive to all drugs, while the remaining 7 (58.3%) isolates were drug-resistant strains. Although Beijing strains are associated with drug resistance, some studies have not found any association between the Beijing family and drug resistance [35, 36] and in one Indian study, some of the Beijing strains have also been found to be pan-susceptible [37] similar to our finding.

CAS1-Delhi (SIT 26 & SIT 381) was the third-largest lineage in the present study. CAS1-Delhi lineage is supposed to have originated from India and later disseminated to regions such as Saudi Arabia, Kenya, South Africa, Malaysia, Myanmar, Australia, the USA, and parts of Europe through frequent migration [37,38].

A new, ancient clade of strains, called 'MANU' was previously identified in India which belongs to the ancestral family of principle genetic group - 1 and is heavily concentrated in Mumbai [33] but in our study prevalence of MANU was significantly less (1%).

H37rv is the reference strain used for spoligotyping in many studies. But this H37rv strain itself has been reported from patients with Tuberculosis disease in many studies. Totally above 180 strains have been reported all over the world till 2010 (Spoligotyping database) (Table 4). There were 7 different spoligotyping patterns (SIT1647, SIT451, SIT824, SIT568, SIT3840, SIT305 & orphan) found related to T-H37rv according to SPTBDB 4.0. One study from Ireland reported 44 (19%) H37rv strains in patients with tuberculosis, in a span of two years (2010 & 2011) [39]. A larger study conducted in New York City from 1996 to 2004, reported 122 H37rv strains which is the highest number of T- H37rv reported all over the world [40]. From India, 3 isolates of H37rv have been reported according to spoligotyping database (Table 4). In our study 7 (6.73%) T-H37rv strains were found, and 2 strains were resistant to isoniazid while the remaining 5 strains were sensitive to all four first anti-tuberculosis drugs. All T-H37rv strains from our study belonged to a single spoligotype pattern (SIT 451).

Among 104 isolates, 82 defined lineages were clustered into 6 groups. Three large clusters were identified comprising 31, 17, and 12 patient isolates. Cluster size varied between 5 and 31. Among 11 orphan strains (10 unique spoligo patterns), EAI3-IND, MANU2, and Unknown lineages are being reported for the first time from India in our study (Table 2).

High degrees of variation in the allelic diversity of MIRU-VNTRs are reported among MTBC strains in different geographical locations [41–44]. In the present study MIRU alleles (MIRU10, MIRU26, Mtub21, QUB11b, and QUB26) showed high discriminatory power (HGDI > 0.7) (Table 1). On the other hand, the lowest discriminatory power was found in alleles

Table 4. Prevalence comparison of different lineages in India vs worldwide.

Lineages	No of strains (in database)[a]	No of strains (reported from India)[a]	India contribution globally (%)[a]	India prevalence (%)[a]	Our study		
					No. Of isolates	No. Of spoligotypes	Prevalence (%)
Beijing	10850	285	2.63	10.77	12	2	11.54
Eai3-ind	982	298	30.65	11.26	20	5	8.65
Unknown	5002	188	3.76	7.1	7	2	19.23
Eai5	1826	186	10.19	7.03	13	4	12.50
Eai7-bgd2	104	2	1.92	0.08	1	1	0.96
Lam6	635	4	0.63	0.15	2	1	0.96
Cas1-delhi	3220	731	22.7	27.62	9	2	3.85
Manu2	584	73	12.5	2.76	1	1	1.92
X1	1733	20	1.15	0.76	2	1	0.96
T	812	11	1.35	0.42	8	1	7.69
T1	11831	168	1.42	6.35	6	2	5.77
T2	1849	15	0.81	0.57	4	1	3.85
T-h37rv	180	3	1.66	0.11	7	1	6.73
H1	3001	13	0.43	0.49	1	1	1.92
H3	6065	20	0.33	0.76	4	1	6.73
Orphan or ND	-	-	-	-	7	7	6.73

A: data according to spoligotyping database 4.0 (SITVIT2). LAM: Latin American Mediterranean; CAS: Central Asian Strain-; EAI: East African Indian; EAI-BGD2: East African Indian- Bangladesh variant; H: Harleem; ND: Not Defined.

MIRU02 and MIRU27 (HGDI < 0.3). An earlier study has shown that MIRU26 and MIRU10 have high HGDI [45]. A recent study from Mumbai has also reported high discriminatory power for the allele MIRU26 and low discriminatory power for alleles MIRU02 and MIRU27. The above study from Mumbai also reported low discriminatory power for MIRU20 (HGDI: 0.04), however, in our samples MIRU20 showed moderate discriminatory power (HGDI: 0.2514).

Mono-resistance to isoniazid was very high among the previously treated patients. In our study resistance to isoniazid was 47.12% and mono resistance was 15.38%. A study by Paramasivan *et al.* from Chennai [46] reported overall isoniazid resistance in 67.5% and mono resistance in only 5.6%. In a study from North India, the corresponding figures were 30.3% and 7.1% respectively [47]. Twenty-four (23%) multi-drug resistant *M. Tuberculosis* were identified in this study. Among these 24 MDR isolates, 6 belonged to EAI3-Ind, 5 each to Beijing and T1, 4 to EAI5, 3 to CAS1-Delhi and one belonged to MANU2 lineages. We found no relationship between different clusters and MDR TB isolates.

The sequencing results displayed no mutation associated with these 33 isolates: the results corresponding to the study by Homolka *et al* [8]. He showed that RV0557 321t > c mutation will categorize the Euro-American and Non-Euro-American strains. But in our study, we selected only a selected group of lineages to categorize them more precisely with SNPs. On the other hand, the 5 sets of genes play a key role in MTB infection. RV0129c for mycobacterial cell wall biosynthesis [48], Rv0557-Replication of MTB within the human macrophages [49], Rv1811-Intracellular survival [50], Rv2628-Adaptation to low oxygen and induction of latency [51], Rv1009-Re-activation of disease from dormancy [52]. No mutations associated with these important five set genes in these novel strains may qualify them to be potential vaccine candidates, particularly Antigen 85c (Rv0129c).

India is endemic for tuberculosis and moreover, our samples being from one particular region, it is to be expected that only a few types of strains will be in circulation. However, the presence of multiple lineages and sub-lineages in our study population goes against this accepted dictum. EAI family accounts for 33% in our study and EAI3-Ind (19.23%) sub-lineages were the most predominant lineage in this region. A previous study carried out with 540 strains from all over India arrived at the conclusion that spoligotypes belonging to CAS family predominated in the North, whereas the EAI family was more common in Southern India [53]. The occurrence of 11.54% Beijing strains in South India is an important finding in our study together with a high frequency of Isoniazid mono resistance among previously treated patients. A similar study shows Both the modern and ancestral *M. tuberculosis* strains are prevailing in this region with north-south compartmentalization, respectively, and the isolates show a high degree of spoligotype signature diversity [9].

Several studies have been conducted to find out the method with high discriminatory potential for strain typing of *M. tuberculosis* [8,54]. Hence based on our study for better strain discrimination, finding epidemiological links, and to identify the rate of evolution, all three typing methods are useful genotyping tools.

Ancestral genotypes of *M. tuberculosis* are the predominant strains circulating in most parts of India, which are supposed to be of low virulence and with lesser power for dissemination. But the gradual increase of newer and more virulent and aggressive lineages like the Beijing genotype poses a significant risk since they may gradually replace the ancestral types in the future [55]. The available data from global reference collection in all databases as well as from the population-based strain collection indicates that the MTBC strain diversity is undergoing rapid changes and divergence and the emergence of new and orphan (Novel) strains is being underestimated. Novel genomic tools may be necessary in the future to correctly identify and discriminate these new emerging strains.

Conclusions

There is an increased isoniazid mono-resistance among already treated cases. Occurrence of 11.54 % Beijing strains is the most important finding in this study. The number of isolates belonging to EAI family is high among already treated cases which seem to be increased and also EAI family accounting 42% among total MDR cases. 11 unknown or orphan strains (Novel) not been reported anywhere in the world previously, this evolution especially in previously treated patients is another important finding. The combination of spoligotyping with MIRU-VNTR and/or SNP typing can give a much clearer and unambiguous picture of the molecular epidemiology of *M. tuberculosis* in a given region. The antigen 85C is the most important antigen for latency in tuberculosis and it is clear that all orphan (Novel) and already existing lineages have shown no mutation; hence it may be very useful for *M. tuberculosis* diagnosis as well as a vaccine candidate.

Acknowledgements

The authors acknowledge the help by the Sri Balaji Arogya Vara Prasadhini Scheme, SVIMS, Tirupati for partly funding the work and the RMRC, Dibrugarh in carrying out spoligotyping of the isolates. State level virus diagnostics laboratory for carryout out MIRU-VNTR typing.

Authors' contributions

Conception and design of the study: Dr. Anagoni Srikar, Dr. Abhijit Chaudhury; Acquisition of data: Dr. Anagoni srikar; Analysis and/or interpretation of data: Dr. Anagoni Srikar, Dr. Abhijit Chaudhury, Dr. Rekha Devi, Dr. Kanwar Narain; Drafting the manuscript: Dr. Alladi Mohan, Dr. Venkataramana B, Dr PVGK Sarma; Revising the manuscript critically for important intellectual content: Dr. Abhijit Chaudhury; Approval of the version of the manuscript to be published: Dr. Abhijit Chaudhury.

Ethical clearance

This study was approved by institutional ethical committee (Roc.No.AS/11/IEC/SVIMS/2017, Date- 05.09.2018).

References

1. World Health Organisation (2019) Global tuberculosis report 2019. Available: https://www.who.int/publications-detail-redirect/9789241565714. Accessed: 14 Aug 2020.
2. Ministry of Health and Family Welfare GOI (2016) Report of the first national anti- tuberculosis drug resistance survey, 2014-16. Available: https://tbcindia.gov.in/showfile.php?lid=3315. Accessed: 22 August 2020.
3. Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C (2001) Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. J Clin Microbiol 39: 3563–3571. doi: 10.1128/JCM.39.10.3563-3571.2001.
4. Oelemann MC, Diel R, Vatin V, Haas W, Rüsch-Gerdes S, Locht C, Niemann S, Supply P (2007) Assessment of an optimized mycobacterial interspersed repetitive- unit-variable-number tandem-repeat typing system combined with spoligotyping for population-based molecular epidemiology studies of tuberculosis. J Clin Microbiol 45: 691–697. doi: 10.1128/JCM.01393-06.
5. Jagielski T, Minias A, van Ingen J, Rastogi N, Brzostek A, Żaczek A, Dziadek J (2016) Methodological and clinical aspects of the molecular epidemiology of *Mycobacterium tuberculosis* and other mycobacteria. Clin Microbiol Rev 29: 239–290. doi: 10.1128/CMR.00055-15.
6. Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbón MH, Bobadilla del Valle M, Fyfe J, García-García L, Rastogi N, Sola C, Zozio T, Guerrero MI, León CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendón A, Sifuentes-Osornio J, Ponce de León A, Cave MD, Fleischmann R, Whittam TS, Alland D (2006) Global phylogeny of Mycobacterium tuberculosis based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. J Bacteriol 188: 759–772. doi: 10.1128/JB.188.2.759-772.2006.
7. Dou H-Y, Lin C-H, Chen Y-Y, Yang S-J, Chang J-R, Wu K-M, Chen Y-T, Chin P-J, Liu Y-M, Su I-J, Tsai S-F (2017) Lineage-specific SNPs for genotyping of *Mycobacterium tuberculosis* clinical isolates. Sci Rep 7: 1425. doi: 10.1038/s41598-017-01580-z.
8. Homolka S, Projahn M, Feuerriegel S, Ubben T, Diel R, Nübel U, Niemann S (2012) High resolution discrimination of clinical *Mycobacterium tuberculosis* complex strains based on single nucleotide polymorphisms. PLOS ONE 7: e39855. doi: 10.1371/journal.pone.0039855.
9. Thomas SK, Iravatham CC, Moni BH, Kumar A, Archana BV, Majid M, Priyadarshini Y, Rani PS, Valluri V, Hasnain SE, Ahmed N (2011) modern and ancestral genotypes of *Mycobacterium tuberculosis* from Andhra Pradesh, India. PLoS ONE 6: e27584. doi: 10.1371/journal.pone.0027584.
10. Suzana S, Shanmugam S, Uma Devi KR, Swarna Latha PN, Michael JS (2017) Spoligotyping of *Mycobacterium tuberculosis* isolates at a tertiary care hospital in India. Trop Med Int Health TM IH 22: 703–707. doi: 10.1111/tmi.12875.
11. Kandhakumari G, Stephen S, Sivakumar S, Narayanan S (2015) Spoligotype patterns of *Mycobacterium tuberculosis* isolated from extra pulmonary tuberculosis patients in Puducherry, India. Indian J Med Microbiol 33: 267–270. doi: 10.4103/0255-0857.154871.
12. Chandramuki A, Khanna N, Shashkina E, Kurepina N, Mathema B, Kreiswirth BN, Venkataswamy MM (2017) Genotypic characterisation of *Mycobacterium tuberculosis* isolates from tuberculous meningitis patients at a tertiary neurocare centre in southern india. Indian J Med Microbiol 35: 211–216. doi: 10.4103/ijmm.IJMM_16_166.
13. Varma-Basil M, Narang A, Chakravorty S, Garima K, Gupta S, Kumar Sharma N, Giri A, Zozio T, Couvin D, Hanif M, Bhatnagar A, Menon B, Niemann S, Rastogi N, Alland D, Bose M (2016) A snapshot of the predominant single nucleotide polymorphism cluster groups of *Mycobacterium tuberculosis* clinical isolates in Delhi, India. Tuberc Edinb Scotl 100: 72–81. doi: 10.1016/j.tube.2016.07.007.
14. Kent PT, Kubica GP (1985) Public health mycobacteriology: a guide for the level iii laboratory. Centers for Disease Control, Atlanta, GA.
15. Tuberculosis Research Centre, Chennai (2010) Standard Operating Procedures for Mycobacteriology Laboratory. Available: https://nirt.res.in/pdf/bact/SOP.pdf. Accessed: 05 September 2019.
16. Millard J, Ugarte-Gil C, Moore DAJ (2015) Multidrug resistant tuberculosis. BMJ 350: h882. doi: 10.1136/bmj.h882.
17. Noguti EN, Leite CQF, Malaspina AC, Santos ACB, Hirata RDC, Hirata MH, Mamizuka EM, Cardoso RF (2010) Genotyping of *Mycobacterium tuberculosis* isolates from a low-endemic setting in north-western state of Paraná in Southern Brazil. Mem Inst Oswaldo Cruz 105: 779–785. doi: 10.1590/s0074-02762010000600008.
18. Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C (2001) Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. J Clin Microbiol 39: 3563–3571. doi: 10.1128/JCM.39.10.3563-3571.2001.
19. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi

N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Locht C, van Soolingen D (2006) Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J Clin Microbiol 44: 4498–4510. doi: 10.1128/JCM.01392-06.

20. Le Flèche P, Fabre M, Denoeud F, Koeck J-L, Vergnaud G (2002) High resolution, on-line identification of strains from the *Mycobacterium tuberculosis* complex based on tandem repeat typing. BMC Microbiol 2: 37. doi: 10.1186/1471-2180-2-37.
21. Ramachandran R, Nalini S, Chandrasekar V, Dave PV, Sanghvi AS, Wares F, Paramasivan CN, Narayanan PR, Sahu S, Parmar M, Chadha S, Dewan P, Chauhan LS (2009) Surveillance of drug-resistant tuberculosis in the state of Gujarat, India. Int J Tuberc Lung Dis 13: 1154–1160.
22. Bhat J, Rao VG, Yadav R, Muniyandi M, Sharma R, Karfarma C, Luke C (2015) Situation of drug resistant tuberculosis in Saharia tribe of central India. Indian J Med Res 141: 636–639. doi: 10.4103/0971-5916.159562.
23. Sharma SK, Kumar S, Saha PK, George N, Arora SK, Gupta D, Singh U, Hanif M, Vashisht RP (2011) Prevalence of multidrug-resistant tuberculosis among Category II pulmonary tuberculosis patients. Indian J Med Res 133: 312–315.
24. Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, Hilty M, Hopewell PC, Small PM (2006) Variable host-pathogen compatibility in *Mycobacterium tuberculosis.* Proc Natl Acad Sci U S A 103: 2869–2873. doi: 10.1073/pnas.0511240103.
25. Sharma P, Katoch K, Chandra S, Chauhan DS, Sharma VD, Couvin D, Rastogi N, Katoch VM (2017) Comparative study of genotypes of *Mycobacterium tuberculosis* from a Northern Indian setting with strains reported from other parts of India and neighboring countries. Tuberc Edinb Scotl 105: 60–72. doi: 10.1016/j.tube.2017.04.003.
26. Singh UB, Suresh N, Bhanu NV, Arora J, Pant H, Sinha S, Aggarwal RC, Singh S, Pande JN, Sola C, Rastogi N, Seth P (2004) Predominant tuberculosis spoligotypes, Delhi, India. Emerg Infect Dis 10: 1138-42. doi: 10.3201/eid1006.030575.
27. Rad ME, Bifani P, Martin C, Kremer K, Samper S, Rauzier J, Kreiswirth B, Blazquez J, Jouan M, van Soolingen D, Gicquel B (2003) Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the w-Beijing family. Emerg Infect Dis 9: 838–845. doi: 10.3201/eid0907.020803.
28. European Concerted Action on New Generation Genetic Markers and Techniques for the Epidemiology and Control of Tuberculosis (2006) Beijing/W genotype *Mycobacterium tuberculosis* and drug resistance. Emerg Infect Dis 12: 736–743. doi: 10.3201/eid1205.050400.
29. Yang C, Luo T, Sun G, Qiao K, Sun G, DeRiemer K, Mei J, Gao Q (2012) *Mycobacterium tuberculosis* Beijing strains favor transmission but not drug resistance in China. Clin Infect Dis Off Publ Infect Dis Soc Am 55: 1179–1187. doi: 10.1093/cid/cis670.
30. Mathuria JP, Sharma P, Prakash P, Samaria JK, Katoch VM, Anupurba S (2008) Role of spoligotyping and IS6110-RFLP in assessing genetic diversity of *Mycobacterium tuberculosis* in India. Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis 8: 346–351. doi: 10.1016/j.meegid.2008.02.005.
31. Shanmugam S, Selvakumar N, Narayanan S (2011) Drug resistance among different genotypes of *Mycobacterium tuberculosis* isolated from patients from Tiruvallur, South India. Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis 11: 980–986. doi: 10.1016/j.meegid.2011.03.011.
32. Mathuria JP, Sharma P, Prakash P, Samaria JK, Katoch VM, Anupurba S (2016) Assessing genetic diversity of *Mycobacterium tuberculosis* by spoligotyping and IS6110-based restriction fragment length polymorphism in North India. Int J Mycobacteriology 5 Suppl 1: S174–S175. doi: 10.1016/j.ijmyco.2016.11.015.
33. Chatterjee A, D'Souza D, Vira T, Bamne A, Ambe GT, Nicol MP, Wilkinson RJ, Mistry N (2010) Strains of *Mycobacterium tuberculosis* from western Maharashtra, India, exhibit a high degree of diversity and strain-specific associations with drug resistance, cavitary disease, and treatment failure. J Clin Microbiol 48: 3593–3599. doi: 10.1128/JCM.00430-10.
34. Devi KR, Bhutia R, Bhowmick S, Mukherjee K, Mahanta J, Narain K (2015) Genetic diversity of *Mycobacterium tuberculosi*s isolates from Assam, India: dominance of Beijing family and discovery of two new clades related to CAS1_Delhi and EAI family based on spoligotyping and MIRU-VNTR Typing. PloS One 10: e0145860. doi: 10.1371/journal.pone.0145860.
35. Asiimwe BB, Ghebremichael S, Kallenius G, Koivula T, Joloba ML (2008) *Mycobacterium tuberculosis* spoligotypes and drug susceptibility pattern of isolates from tuberculosis patients in peri-urban Kampala, Uganda. BMC Infect Dis 8: 101. doi: 10.1186/1471-2334-8-101.
36. Glynn JR, Crampin AC, Traore H, Yates MD, Mwaungulu FD, Ngwira BM, Chaguluka SD, Mwafulirwa DT, Floyd S, Murphy C, Drobniewski FA, Fine PEM (2005) *Mycobacterium tuberculosis* Beijing genotype, northern Malawi. Emerg Infect Dis 11: 150–153. doi: 10.3201/eid1101.040869.
37. Gupta A, Kulkarni S, Rastogi N, Anupurba S (2014) A study of *Mycobacterium tuberculosis* genotypic diversity & drug resistance mutations in Varanasi, north India. Indian J Med Res 139: 892–902.
38. Arora J, Bhalla M, Sidiq Z, Lal P, Behera D, Rastogi N, Myneedu VP (2013) Predominance of Beijing genotype in extensively drug resistant *Mycobacterium tuberculosis* isolates from a tertiary care hospital in New Delhi, India. Int J Mycobacteriology 2: 109–113. doi: 10.1016/j.ijmyco.2013.03.001.
39. Fitzgibbon MM, Gibbons N, Roycroft E, Jackson S, O'Donnell J, O'Flanagan D, Rogers TR (2013) A snapshot of genetic lineages of *Mycobacterium tuberculosis* in Ireland over a two-year period, 2010 and 2011. Euro Surveill Bull Eur Sur Mal Transm Eur Commun Dis Bull 18: 20367.
40. Vitol I, Driscoll J, Kreiswirth B, Kurepina N, Bennett KP (2006) Identifying *Mycobacterium tuberculosis* complex strain families using spoligotypes. Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis 6: 491–504. doi: 10.1016/j.meegid.2006.03.003.
41. Sharma P, Chauhan DS, Upadhyay P, Faujdar J, Lavania M, Sachan S, Katoch K, Katoch VM (2008) Molecular typing of *Mycobacterium tuberculosis* isolates from a rural area of Kanpur by spoligotyping and mycobacterial interspersed repetitive units (MIRUs) typing. Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis 8: 621–626. doi: 10.1016/j.meegid.2008.05.002.
42. Chatterjee A, Mistry N (2013) MIRU-VNTR profiles of three major *Mycobacterium tuberculosis* spoligotypes found in western India. Tuberc Edinb Scotl 93: 250–256. doi: 10.1016/j.tube.2012.10.004.

43. Vadwai V, Shetty A, Supply P, Rodrigues C (2012) Evaluation of 24-locus MIRU-VNTR in extrapulmonary specimens: study from a tertiary centre in Mumbai. Tuberc Edinb Scotl 92: 264–272. doi: 10.1016/j.tube.2012.01.002.
44. Refaya AK, Sivakumar S, Sundararaman B, Narayanan S (2012) Polymorphism in the RD1 locus and its effect on downstream genes among South Indian clinical isolates of *Mycobacterium tuberculosis*. J Med Microbiol 61: 1352–1359. doi: 10.1099/jmm.0.044453-0.
45. Mokrousov I, Narvskaya O, Limeschenko E, Vyazovaya A, Otten T, Vyshnevskiy B (2004) Analysis of the allelic diversity of the mycobacterial interspersed repetitive units in *Mycobacterium tuberculosi*s strains of the Beijing family: practical implications and evolutionary considerations. J Clin Microbiol 42: 2438–2444. doi: 10.1128/JCM.42.6.2438-2444.2004.
46. Paramasivan CN, Rehman F, Wares F, Sundar Mohan N, Sundar S, Devi S, Narayanan PR (2010) First- and second-line drug resistance patterns among previously treated tuberculosis patients in India. Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis 14: 243–246.
47. Maurya AK, Kant S, Nag VL, Kushwaha R a. S, Dhole TN (2012) Trends of anti-tuberculosis drug resistance pattern in new cases and previously treated cases of extrapulmonary tuberculosis cases in referral hospitals in northern India. J Postgrad Med 58: 185–189. doi: 10.4103/0022-3859.101379.
48. Jackson M, McNeil MR, Brennan PJ (2013) Progress in targeting cell envelope biogenesis in *Mycobacterium tuberculosis*. Future Microbiol 8:855–875. doi: 10.2217/fmb.13.52.
49. Besra GS, Brennan PJ (1997) The mycobacterial cell wall: biosynthesis of arabinogalactan and lipoarabinomannan. Biochem Soc Trans 25: 845–850. doi: 10.1042/bst0250845.
50. Downing KJ, Betts JC, Young DI, McAdam RA, Kelly F, Young M, Mizrahi V (2004) Global expression profiling of strains harbouring null mutations reveals that the five rpf-like genes of *Mycobacterium tuberculosis* show functional redundancy. Tuberc Edinb Scotl 84: 167–179. doi: 10.1016/j.tube.2003.12.004.
51. Buchmeier N, Blanc-Potard A, Ehrt S, Piddington D, Riley L, Groisman EA (2000) A parallel intraphagosomal survival strategy shared by *Mycobacterium tuberculosis* and *Salmonella enterica*. Mol Microbiol 35: 1375–1382. doi: 10.1046/j.1365-2958.2000.01797.x.
52. Sherman DR, Voskuil M, Schnappinger D, Liao R, Harrell MI, Schoolnik GK (2001) Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha -crystallin. Proc Natl Acad Sci U S A 98: 7534–7539. doi: 10.1073/pnas.121172498.
53. Desikan P, Chauhan DS, Sharma P, Panwalkar N, Chourey M, Patidar ML, Yadav P, Chandrasekaran V, Ohri BS (2016) Genetic diversity of *Mycobacterium tuberculosis* isolates from central India. Indian J Med Res 143: 481–486. doi: 10.4103/0971-5916.184287.
54. Pitondo-Silva A, Santos ACB, Jolley KA, Leite CQF, Darini AL da C (2013) Comparison of three molecular typing methods to assess genetic diversity for *Mycobacterium tuberculosis*. J Microbiol Methods 93: 42–48. doi: 10.1016/j.mimet.2013.01.020.
55. Ahmed N, Ehtesham NZ, Hasnain SE (2009) Ancestral *Mycobacterium tuberculosis* genotypes in India: implications for TB control programmes. Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis 9: 142–146. doi: 10.1016/j.meegid.2008.10.001.

Corresponding author

Abhijit Chaudhury M.D, DNB
Department of Microbiology,
Sri Venkateswara Institute of Medical Sciences,
Tirupati, Andhra Pradesh, India.
Tel: +919492570422
Email: ach1964@rediffmail.com.

Conflict of interests: No conflict of interests is declared.

Annex – Supplementary Items

Supplementary Table 1. Selected five genes primer sequences (forward and reverse).

Primer	Sequence	Position	Gene length	PCR product
Rv0129c - F	5`- GAA CCT CCA CGC CCG CAA C -3`	- 175	1023 bp	1304 bp
Rv0129c - F	5`- GCG CTG CGG CCA CGA CAT TC-3`	+ 1129		
Rv0557 - F	5`- CTG GAC AAG CGG TTG GAA C -3`	- 112	1137 bp	1340 bp
Rv0557 - R	5`- GTC ATA CTT GCG GGC GAC G -3`	+ 1210		
Rv1009 - F	5`- GGC CCA TTT TGC TTT TTG TT -3`	- 105	1089 bp	1360 bp
Rv1009 - R	5`- GGC CCG ACC TCC AAA ACC AG -3`	+ 1236		
Rv1811 - F	5`- CGC CTA GGC TCA AAC TGC TG -3`	- 78	705 bp	880 bp
Rv1811 - R	5`- CAA TAC CCG GCG GAT CTA CC -3`	+ 783		
Rv2628 - F	5`- GGC GCG ACC GGG CAC ATC -3`	- 83	363 bp	521 bp
Rv2628 - R	5`- GCG GGA AGG CAT AGG GAC CAA AGT -3`	+ 415		

Supplementary Table 2. Detailed genotyping (24-MIRU-VNTR typing and spoligotyping) and drug-susceptibility testing results of 104 Mycobacterium tuberculosis strains isolated from Andhrapradesh, India.

SPECIES NAME, LINEAGE AND SIT NUMBER				24-MIRU-VNTR typing																						
Sample id	**Lineage**	**SIT NUMBER**	**Country of Isolation**	**MIRU02**	**Mtub04**	**ETRC**	**MIRU04**	**MIRU40**	**MIRU10**	**MIRU16**	**Mtub21**	**MIRU20**	**QUB11b**	**ETRA**	**Mtub29**	**Mtub30**	**ETRB**	**MIRU23**	**MIRU24**	**MIRU26**	**MIRU27**	**Mtub34**	**MIRU31**	**Mtub39**	**QUb26**	**QUB4156**	**MIRU39**
1	EAI5	340	INDIA	2	2	4	5	3	5	4	15	2	1	7	4	1	5	10	2	2	4	4	6	5	8	1	4
2	T1	498	INDIA	2	2	4	2	5	4	2	2	2	3	3	5	2	2	5	1	6	4	4	4	6	5	3	3
3	Not defined	Orphan or New	INDIA	2	2	3	5	4	4	3	6	2	3	6	4	1	1	6	2	2	4	3	4	4	6	1	3
4	Not defined	Orphan or New	INDIA	2	2	3	5	4	4	3	6	2	3	6	4	1	1	6	2	2	4	3	4	5	6	1	4
5	EAI5	126	INDIA	2	2	4	5	3	2	4	15	2	1	8	4	1	5	10	2	2	4	4	6	6	9	2	4
6	T	442	INDIA	2	2	3	2	2	5	1	3	3	3	2	5	3	2	4	3	5	2	2	2	4	3	3	3
7	EAI3-IND	11	INDIA	2	2	4	5	2	6	3	5	3	2	6	4	4	2	6	1	3	2	3	3	3	6	1	3
8	T1	53	INDIA	2	2	4	2	5	3	2	2	2	2	3	5	2	2	5	1	6	4	3	4	6	6	2	3
9	Not defined	Orphan or New	INDIA	2	2	4	2	2	5	3	2	2	3	4	3	2	2	6	2	2	2	2	4	2	7	3	3
10	T1	53	INDIA	2	2	4	2	5	3	2	2	2	2	3	5	2	2	5	1	6	4	3	4	6	6	2	3
11	Beijing	255	INDIA	2	4	4	3	2	3	3	3	3	6	4	4	4	2	5	1	5	3	3	5	3	6	1	3
12	EAI3-IND	11	INDIA	2	2	4	5	2	6	3	5	3	2	6	4	4	2	6	1	3	2	3	3	3	6	1	3
13	EAI3-IND	11	INDIA	2	2	4	5	2	6	3	5	3	2	6	4	4	2	6	1	3	2	3	3	3	7	1	3
14	EAI3-IND	11	INDIA	2	2	4	5	2	7	3	5	3	3	6	4	4	3	6	1	3	2	3	3	3	6	2	3
15	EAI5	340	INDIA	2	2	4	5	3	4	4	15	2	1	7	4	1	5	10	2	2	4	4	6	5	9	1	4
16	T	442	INDIA	2	2	3	2	2	5	1	2	3	3	2	5	3	2	4	3	5	2	2	2	4	3	3	3
17	T1	53	INDIA	2	2	4	2	5	3	2	2	2	2	3	5	2	2	5	1	6	4	3	4	6	7	2	3
18	CAS1-Delhi	26	INDIA	2	2	4	4	3	4	4	1	2	4	4	5	2	2	5	1	6	3	3	3	1	8	2	3
19	T-H37Rv	451	INDIA	2	5	4	2	2	4	4	5	2	6	4	5	2	2	5	1	6	4	14	6	4	7	1	3
20	T	442	INDIA	2	2	3	2	2	6	1	3	3	3	2	5	3	3	4	3	5	2	2	2	4	2	3	3
21	Not defined	Orphan or New	INDIA	2	2	4	5	2	4	2	6	2	2	6	3	1	3	6	2	2	3	3	4	3	3	1	3
22	Unknown	602	INDIA	2	2	2	4	3	5	5	9	2	8	11	4	2	4	6	3	2	4	4	4	5	9	2	4
23	EAI5	340	INDIA	2	2	4	5	3	5	4	15	2	2	7	4	1	5	10	2	2	4	4	6	5	8	1	4
24	T	442	INDIA	2	2	3	2	2	5	1	3	3	3	2	5	3	2	4	3	5	2	2	2	4	3	3	3
25	EAI3-IND	11	INDIA	2	2	4	5	2	6	3	5	3	3	6	4	4	2	6	1	3	2	3	3	3	8	1	3
26	EAI3-IND	1948	INDIA	2	2	4	5	3	4	3	6	2	3	6	3	1	1	6	2	2	3	3	5	6	7	1	3
27	LAM6	64	INDIA	1	3	6	2	7	3	3	2	2	2	2	4	1	2	6	1	8	3	3	2	2	8	2	2
28	EAI5	340	INDIA	2	2	4	5	3	4	4	15	2	1	7	4	1	5	10	2	2	4	4	6	5	7	1	4
29	T1	53	INDIA	2	2	4	2	5	3	2	2	2	2	3	5	2	2	5	1	6	4	3	4	6	6	2	3
30	H3	50	INDIA	2	2	3	2	3	5	5	2	2	1	2	2	4	2	3	1	6	4	4	3	4	1	2	3
31	EAI5	458	INDIA	2	4	4	2	2	5	2	2	3	5	3	3	7	5	4	2	5	1	2	5	2	2	3	4
32	EAI5	458	INDIA	2	4	4	2	2	5	2	2	3	5	3	3	7	5	4	2	5	1	2	5	2	2	3	4
33	Not defined	Orphan or New	INDIA	2	2	4	6	3	3	3	14	2	3	6	3	1	4	10	2	1	4	3	4	4	6	1	0
34	T	442	INDIA	2	2	3	2	2	7	1	2	3	3	2	5	3	2	4	3	5	2	2	2	4	2	3	3
35	EAI3-IND	11	INDIA	2	2	4	5	2	5	3	5	3	2	6	4	4	3	6	1	3	2	3	3	3	6	1	3
36	EAI3-IND	11	INDIA	2	2	4	5	2	6	3	5	3	4	6	4	4	2	6	1	3	2	3	3	3	6	1	3
37	Beijing	255	INDIA	2	4	4	3	2	4	3	3	3	6	4	4	4	4	5	1	5	3	3	5	3	7	1	3
38	EAI3-IND	11	INDIA	2	2	4	5	2	6	3	5	3	2	6	4	4	2	6	1	3	2	3	3	3	5	2	3
39	T-H37Rv	451	INDIA	2	5	4	2	2	4	4	5	2	6	4	5	2	2	5	1	6	4	14	6	4	8	1	3
40	H1	47	INDIA	2	2	3	2	3	5	3	3	2	3	3	2	4	2	3	1	5	3	3	3	3	7	3	2
41	Manu2	Orphan or New	INDIA	2	3	2	3	3	5	3	3	1	4	2	5	2	2	6	1	2	4	2	3	1	9	3	3
42	T	442	INDIA	2	2	3	2	2	5	1	3	3	3	2	5	3	4	4	3	5	2	2	2	4	3	3	3
43	H3	50	INDIA	2	2	3	3	3	5	5	3	2	1	3	2	4	2	3	2	6	4	4	3	4	1	2	3
44	EAI5	340	INDIA	2	2	4	5	3	4	4	15	2	1	7	4	1	5	10	2	2	4	4	6	5	8	1	4
45	Unknown	602	INDIA	2	2	2	4	3	5	5	9	2	8	11	4	2	4	6	3	2	4	4	4	5	7	2	4
46	EAI3-IND	473	INDIA	2	2	4	5	4	4	2	6	2	3	6	3	1	4	6	2	2	2	3	4	2	4	2	3
47	EAI3-IND	1680	INDIA	2	2	3	5	3	3	3	6	2	3	6	3	1	1	6	2	2	3	3	4	4	7	0	3
48	EAI3-IND	Orphan or New	INDIA	2	2	4	7	3	4	3	6	2	3	6	3	1	1	6	2	2	3	3	6	6	9	2	3
49	EAI5	340	INDIA	2	2	4	5	3	6	4	15	2	3	7	4	1	5	10	2	2	4	4	6	5	8	1	4
50	T	442	INDIA	2	2	3	2	2	6	1	4	3	3	2	5	3	2	4	3	5	2	2	2	4	3	3	3
51	EAI5	340	INDIA	2	2	4	5	3	4	4	15	2	1	7	4	1	5	10	2	2	4	4	6	5	7	1	4
52	Not defined	Orphan or New	INDIA	2	1	4	9	4	4	3	6	1	3	7	3	1	1	5	2	2	2	3	4	4	6	0	0
53	Beijing	1	INDIA	2	4	4	2	3	4	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	8	2	3
54	Beijing	255	INDIA	2	4	4	3	2	2	3	3	3	6	4	4	4	3	5	1	5	3	3	5	3	7	1	3
55	Not defined	Orphan or New	INDIA	2	2	3	4	3	4	3	14	2	3	6	3	1	4	10	2	2	3	3	4	5	5	2	3

56	Beijing	1	INDIA	2	4	4	2	3	4	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	8	2	3
57	T2	153	INDIA	2	3	4	2	2	0	1	4	2	2	3	4	4	2	5	0	1	3	3	3	3	8	4	2
58	T1	53	INDIA	2	2	4	2	5	3	2	2	2	2	3	5	2	2	5	1	6	4	3	4	6	8	2	3
59	CAS1-Delhi	26	INDIA	2	2	4	4	3	4	4	1	2	4	4	5	2	2	5	1	6	3	3	3	1	9	1	3
60	Unknown	Orphan or New	INDIA	2	1	1	5	2	5	3	6	2	5	5	3	1	4	6	2	2	1	3	4	5	5	1	1
61	Unknown	602	INDIA	2	2	2	4	3	6	5	9	2	8	10	4	2	3	6	3	2	2	4	3	5	7	2	4
62	CAS1-Delhi	26	INDIA	2	2	4	4	3	3	4	1	2	4	4	5	2	2	5	1	6	3	3	3	1	8	2	3
63	LAM6	64	INDIA	1	3	6	2	7	3	3	2	2	2	2	4	1	2	6	1	8	3	3	2	2	8	2	2
64	EAI3-IND	11	INDIA	2	2	4	5	2	8	3	5	3	1	6	4	4	2	6	1	3	2	3	3	3	6	1	3
65	EAI3-IND	1680	INDIA	2	2	3	5	3	2	3	6	2	3	6	3	1	1	6	2	2	3	3	4	4	8	0	3
66	CAS1-Delhi	26	INDIA	2	2	4	4	3	4	4	1	2	4	4	5	2	2	5	1	6	3	3	3	1	9	2	3
67	CAS1-Delhi	26	INDIA	2	2	4	4	3	5	4	1	2	4	4	5	2	2	5	1	6	3	3	3	1	9	1	3
68	EAI3-IND	473	INDIA	2	2	4	5	4	4	3	6	2	3	6	3	1	1	6	2	2	3	3	4	3	4	3	3
69	EAI5	236	INDIA	4	2	2	5	3	3	5	10	2	6	10	4	2	5	6	2	2	4	4	4	7	3	1	2
70	T	442	INDIA	2	2	3	2	2	5	1	3	3	3	2	5	3	3	4	3	5	2	2	2	4	2	3	3
71	H3	50	INDIA	2	2	3	2	3	5	5	2	2	1	2	2	4	2	3	1	6	4	4	3	4	3	2	3
72	EAI3-IND	473	INDIA	2	2	4	5	4	4	3	6	2	3	6	3	1	1	6	2	5	3	3	4	3	4	3	3
73	EAI3-IND	11	INDIA	2	2	4	5	2	6	3	5	3	2	6	4	4	4	6	1	3	2	3	3	3	6	1	3
74	T-H37Rv	451	INDIA	2	5	4	2	2	5	4	5	2	6	4	5	2	4	5	1	6	4	14	6	4	8	1	3
75	CAS1-Delhi	26	INDIA	2	2	4	4	3	4	4	1	2	4	4	5	2	2	5	1	6	3	3	3	1	8	2	3
76	CAS1-Delhi	381	INDIA	2	3	2	2	4	6	5	4	2	6	4	5	2	2	5	1	7	4	1	4	4	9	3	3
77	EAI3-IND	11	INDIA	2	2	4	5	2	8	3	5	3	3	6	4	4	2	6	1	3	2	3	3	3	8	1	3
78	CAS1-Delhi	381	INDIA	2	3	2	2	4	6	5	4	2	6	4	5	2	2	5	1	7	4	1	4	4	9	3	3
79	CAS1-Delhi	26	INDIA	2	2	4	4	3	5	4	1	2	4	4	5	2	2	5	1	6	3	3	3	1	9	2	3
80	EAI3-IND	1680	INDIA	2	2	3	5	3	3	3	6	2	3	6	3	1	1	6	2	2	3	3	4	4	7	0	3
81	T2	153	INDIA	2	3	4	2	2	1	1	3	3	2	3	3	4	2	5	0	1	3	3	3	3	9	4	2
82	T2	153	INDIA	2	3	4	2	2	0	1	4	2	2	3	4	4	2	5	0	1	3	3	3	3	8	4	2
83	X1	336	INDIA	2	4	3	2	5	3	3	3	2	4	4	4	4	2	5	1	5	2	3	3	3	8	3	2
84	Beijing	1	INDIA	2	4	4	2	3	5	3	6	2	7	4	4	4	3	5	1	7	3	3	5	3	9	2	3
85	Unknown	602	INDIA	2	2	2	4	3	4	5	9	2	8	11	4	2	4	6	3	2	4	4	4	5	8	2	4
86	H3	50	INDIA	2	2	3	2	3	5	5	2	2	1	2	2	4	2	3	1	6	4	4	3	4	2	2	3
87	EAI5	340	INDIA	2	2	3	4	3	3	1	15	2	3	6	3	1	4	10	2	3	3	3	5	5	5	2	3
88	T-H37Rv	451	INDIA	2	5	4	2	2	4	4	5	2	6	4	5	2	3	5	1	6	4	14	6	4	9	1	3
89	Unknown	Orphan or New	INDIA	2	2	2	5	3	6	3	7	2	6	9	3	2	5	6	2	2	2	2	4	5	3	2	1
90	Beijing	255	INDIA	2	4	4	3	2	3	3	3	3	6	4	4	4	2	5	1	5	3	3	5	3	6	1	3
91	EAI5	340	INDIA	2	2	4	5	3	4	4	15	2	3	7	4	1	5	10	2	2	4	4	6	5	8	1	4
92	Beijing	1	INDIA	2	4	4	2	3	4	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	8	2	3
93	EAI3-IND	11	INDIA	2	2	4	5	2	6	3	5	3	2	6	4	4	4	6	1	3	2	3	3	3	9	1	3
94	Beijing	1	INDIA	2	4	4	2	3	5	3	6	2	8	4	4	4	4	5	1	7	3	3	5	3	8	2	3
95	T-H37Rv	451	INDIA	2	5	4	2	2	4	4	5	2	6	4	5	2	2	5	1	6	4	14	6	4	8	1	3
96	Unknown	602	INDIA	2	2	2	4	3	5	5	9	2	8	11	4	2	4	6	3	2	4	4	4	5	7	2	4
97	T-H37Rv	451	INDIA	2	5	4	2	2	3	4	5	2	6	4	5	2	3	5	1	6	4	14	6	4	9	1	3
98	Beijing	1	INDIA	2	4	4	2	3	4	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	6	2	3
99	T-H37Rv	451	INDIA	2	5	4	2	2	2	4	5	2	6	4	5	2	2	5	1	6	4	14	6	4	8	1	3
100	Beijing	1	INDIA	2	4	4	2	3	3	3	7	2	7	4	4	4	4	5	1	7	3	3	5	3	8	2	3
101	T2	153	INDIA	2	3	4	2	2	1	1	4	1	2	3	4	4	2	5	0	1	3	3	3	3	9	4	2
102	EAI7-BGD2	1391	INDIA	2	1	4	4	3	4	3	7	2	6	10	3	2	4	6	2	2	3	2	5	3	5	3	6
103	X1	336	INDIA	2	4	3	2	5	3	3	3	2	4	4	4	4	2	5	1	5	2	3	3	3	8	3	2
104	Beijing	1	INDIA	2	4	4	2	3	4	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	6	2	3

Supplementary Table 3. Spoligotyping and Drug Susceptibility Testing data.

Sample id	Lineage	SIT NUMBER	Spoligotyping (43-digit Binary)	Drug Susceptibility testing			
				INH	SM	EMB	RMP
1	EAI5	340	10011110001111111111111111111000010111111111	s	s	s	s
2	T1	498	11111111110111111111111111111111100001111111	r	r	s	r
3	Not defined	Orphan or New	10011111110011111111111111111000010111110011	s	s	s	s
4	Not defined	Orphan or New	11100001111111000000000000000000000000111111	r	s	s	s
5	EAI5	126	10011111111111111111111111111000010111111111	s	s	s	s
6	T	442	11111111111011111011111111111100001111111	r	r	s	s
7	EAI3-IND	11	10011111111111111111111111111000010110001111	s	s	s	s
8	T1	53	11111111111111111111111111111111000011111111	r	r	s	r
9	Not defined	Orphan or New	11111111111111111111111111111110110111111111	s	s	s	s
10	T1	53	11111111111111111111111111111111000011111111	r	s	s	s
11	Beijing	255	00000000000000000000000000000000000111101111	s	s	s	s
12	EAI3-IND	11	10011111111111111111111111111000010110001111	r	r	s	r
13	EAI3-IND	11	10011111111111111111111111111000010110001111	s	s	s	s
14	EAI3-IND	11	10011111111111111111111111111000010110001111	r	s	s	s
15	EAI5	340	10011110001111111111111111111000010111111111	s	s	s	s
16	T	442	11111111111011111011111111111100001111111	r	s	s	s
17	T1	53	11111111111111111111111111111111000011111111	r	r	s	r
18	CAS1-Delhi	26	11100001111111111111110000000000000111111111	r	r	s	r
19	T-H37Rv	451	11111111111111111100111111111111000011111111	s	s	s	s
20	T	442	11111111111011111011111111111100001111111	r	s	s	s
21	Not defined	Orphan or New	11111111111111111110010000000010000111111	s	s	s	s
22	Unknown	602	11111111111111111111111000000000000111111	s	r	s	r
23	EAI5	340	10011110001111111111111111111000010111111111	s	s	s	s
24	T	442	11111111111011111011111111111100001111111	r	s	s	s
25	EAI3-IND	11	10011111111111111111111111111000010110001111	s	s	s	s

26	EAI3-IND	1948	10011111111111111111111111111000000110001111	r	r	s	r
27	LAM6	64	11111111111111111111000011110111000011111111	r	s	s	s
28	EAI5	340	10011110001111111111111111111000010111111111	s	s	s	s
29	T1	53	11111111111111111111111111111111000011111111	r	r	s	r
30	H3	50	11111111111111111111111111111010000111111111	s	s	s	r
31	EAI5	458	11111111111111111111111111111000000111111111	r	r	s	r
32	EAI5	458	11111111111111111111111111111000000111111111	r	r	s	r
33	Not defined	Orphan or New	10011111100011111111111111111000010110001111	s	s	s	s
34	T	442	11111111111011111011111111111110000111111111	r	r	s	s
35	EAI3-IND	11	10011111111111111111111111111000010110001111	r	s	s	s
36	EAI3-IND	11	10011111111111111111111111111000010110001111	r	r	s	s
37	Beijing	255	00000000000000000000000000000000000111101111	r	s	r	r
38	EAI3-IND	11	10011111111111111111111111111000010110001111	r	r	s	r
39	T-H37Rv	451	11111111111111111100111111111110000111111111	s	s	s	s
40	H1	47	11111111111111111111111111000000100001111111	s	s	s	s
41	Manu2	Orphan or New	11111111111111111111111111000100111111111111	r	r	s	r
42	T	442	11111111111011111011111111111110000111111111	r	s	s	s
43	H3	50	11111111111111111111111111111010000111111111	s	s	s	r
44	EAI5	340	10011110001111111111111111111000010111111111	s	s	s	s
45	Unknown	602	11111111111111111111111111000000000000111111	s	r	s	r
46	EAI3-IND	473	10000000111111111111111111111000010110001111	r	r	s	s
47	EAI3-IND	1680	10011110111111111111111111111000010110001111	s	s	s	s
48	EAI3-IND	Orphan or New	10011111111111111101111111111000010110001111	s	s	s	s
49	EAI5	340	10011110001111111111111111111000010111111111	r	s	r	r
50	T	442	11111111111011111011111111111110000111111111	r	r	s	s
51	EAI5	340	10011110001111111111111111111000010111111111	r	s	s	s
52	Not defined	Orphan or New	11111111111101111011001111111000011111111	s	s	s	s
53	Beijing	1	00000000000000000000000000000000000111111111	s	s	s	s
54	Beijing	255	00000000000000000000000000000000000111101111	r	s	r	r
55	Not defined	Orphan or New	10011110001111111111001111100001011111111	s	s	s	s
56	Beijing	1	00000000000000000000000000000000000111111111	s	r	s	s
57	T2	153	11110111111111111111111111111111000011110111	s	s	s	s
58	T1	53	11111111111111111111111111111111000011111111	r	r	s	r
59	CAS1-Delhi	26	11100001111111111111100000000000011111111111	s	s	s	s
60	Unknown	Orphan or New	11011111111111111111111111111000000000000000	r	s	s	s
61	Unknown	602	11111111111111111111111100000000000001111111	s	r	s	s
62	CAS1-Delhi	26	11100001111111111111111000000000000011111111	r	r	s	r
63	LAM6	64	11111111111111111110000111101100001111111	s	s	s	s
64	EAI3-IND	11	10011111111111111111111111111000010110001111	r	s	s	s
65	EAI3-IND	1680	10011110111111111111111111111000010110001111	s	s	s	s
66	CAS1-Delhi	26	11100001111111111111111000000000000011111111	s	s	s	s
67	CAS1-Delhi	26	11100001111111111111100000000000011111111	s	s	s	s
68	EAI3-IND	473	10000000111111111111111111111000010110001111	r	r	s	s
69	EAI5	236	11111111111111111111111111111000010111111111	r	r	s	r
70	T	442	11111111111011111011111111111110000111111111	r	s	s	s
71	H3	50	11111111111111111111111111111010000111111111	s	s	s	r
72	EAI3-IND	473	10000000111111111111111111111000010110001111	r	r	s	s
73	EAI3-IND	11	10011111111111111111111111111000010110001111	r	r	s	r
74	T-H37Rv	451	11111111111111111100111111111110000111111111	s	s	s	s
75	CAS1-Delhi	26	11100001111111111111111000000000000011111111	s	s	s	s
76	CAS1-Delhi	381	11100001111111111111100000000000011000111111	s	s	s	s
77	EAI3-IND	11	10011111111111111111111111111000010110001111	r	r	s	r
78	CAS1-Delhi	381	11100001111111111111100000000000011000111111	s	s	s	s
79	CAS1-Delhi	26	11100001111111111111111000000000000011111111	r	r	s	r
80	EAI3-IND	1680	10011110111111111111111111111000010110001111	s	s	s	s
81	T2	153	11110111111111111111111111111111000011110111	s	s	s	s
82	T2	153	11110111111111111111111111111111000011110111	s	s	s	s
83	X1	336	11111111111111111101111111111111000011101111	r	r	s	s
84	Beijing	1	00000000000000000000000000000000000111111111	r	r	r	r
85	Unknown	602	11111111111111111111111100000000000001111111	s	r	s	s
86	H3	50	11111111111111111111111111111010000111111111	s	s	s	r
87	EAI5	340	10011110001111111111111111111000010111111111	s	s	s	s
88	T-H37Rv	451	11111111111011111011111111111110000111111111	r	s	s	s
89	Unknown	Orphan or New	11011111111111111111111111111000000000000000	r	s	s	s
90	Beijing	255	00000000000000000000000000000000000111101111	r	s	r	r
91	EAI5	340	10011110001111111111111111111000010111111111	s	s	s	s
92	Beijing	1	00000000000000000000000000000000000111111111	s	s	s	s
93	EAI3-IND	11	10011111111111111111111111111000010110001111	r	r	s	r
94	Beijing	1	00000000000000000000000000000000000111111111	r	r	r	r
95	T-H37Rv	451	11111111111111111100111111111110000111111111	s	s	s	s
96	Unknown	602	11111111111111111111111100000000000001111111	s	r	s	s
97	T-H37Rv	451	11111111111111111100111111111110000111111111	s	s	s	s
98	Beijing	1	00000000000000000000000000000000000111111111	s	s	s	s
99	T-H37Rv	451	11111111111111111100111111111110000111111111	r	s	s	s
100	Beijing	1	00000000000000000000000000000000000111111111	s	r	s	s
101	T2	153	11110111111111111111111111111111000011110111	s	s	s	s
102	EAI7-BGD2	1391	11111111111111111111000000000000110111111	s	s	s	s
103	X1	336	11111111111111111011111111111110000111101111	r	r	s	s
104	Beijing	1	00000000000000000000000000000000000111111111	s	s	s	s