RESEARCH ARTICLE

Genetic Diversity of *Mycobacterium tuberculosis* 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

Kangjam Rekha Devi, Rinchenla Bhutia, Shovonlal Bhowmick, Kaustab Mukherjee, Jagadish Mahanta, Kanwar Narain*

Regional Medical Research Centre, N.E. Region (Indian Council of Medical Research), Post Box #105, Dibrugarh 786 001, Assam, India

* kanwar_narain@hotmail.com

Abstract

Tuberculosis (TB) is one of the major public health concerns in Assam, a remote state located in the northeastern (NE) region of India. The present study was undertaken to explore the circulating genotypes of *Mycobacterium tuberculosis* complex (MTBC) in this region. A total of 189 MTBC strains were collected from smear positive pulmonary tuberculosis cases from different designated microscopy centres (DMC) from various localities of Assam. All MTBC isolates were cultured on Lowenstein-Jensen (LJ) media and subsequently genotyped using spoligotyping and 24-loci mycobacterial interspersed repetitive units-variable number of tandem repeats (MIRU-VNTR) typing. Spoligotyping of MTBC isolates revealed 89 distinct spoligo patterns. The most dominant MTBC strain belonged to Beijing lineage and was represented by 35.45% (n = 67) of total isolates, followed by MTBC strains belonging to Central Asian-Delhi (CAS/Delhi) lineage and East African Indian (EAI5) lineage. In addition, in the present study 43 unknown spoligo patterns were detected. The discriminatory power of spoligotyping was found to be 0.8637 based on Hunter Gaston Discriminatory Index (HGDI). On the other hand, 24-loci MIRU-VNTR typing revealed that out of total 189 MTBC isolates from Assam 185 (97.9%) isolates had unique MIRU-VNTR profiles and 4 isolates grouped into 2 clusters. Phylogenetic analysis of 67 Beijing isolates based on 24-loci MIRU-VNTR typing revealed that Beijing isolates from Assam represent two major groups, each comprising of several subgroups. Neighbour-Joining (NJ) phylogenetic tree analysis based on combined spoligotyping and 24-loci MIRU-VNTR data of 78 Non-Beijing isolates was carried out for strain lineage identification as implemented by MIRU-VNTR*plus* database. The important lineages of MTBC identified were CAS/CAS1_Delhi (41.02%, n = 78) and East-African-Indian (EAI, 33.33%). Interestingly, phylogenetic analysis of orphan (23.28%) MTBC spoligotypes revealed that majority of these orphan isolates from Assam represent two new sub-clades Assam/EAI

and Assam/CAS. The prevalence of multidrug resistance (MDR) in Beijing and Non-Beijing strains was found to be 10.44% and 9.01% respectively. In conclusion, the present study has shown the predominance of Beijing isolates in Assam which is a matter of great concern because Beijing strains are considered to be ecologically more fit enabling wider dissemination of *M. tuberculosis*. Other interesting finding of the present study is the discovery of two new clades of MTBC isolates circulating in Assam. More elaborate longitudinal studies are required to be undertaken in this region to understand the transmission dynamics of MTBC.

Introduction

In spite of the implementation of the Revised National Tuberculosis Control Program (RNTCP), the burden of tuberculosis (TB) in India is still the highest which is an important public health concern [1]. India being the second highest populated country in the world, alone accounted for 26% of the world's tuberculosis cases [2]. According to the World Health Organization (WHO) estimate [2] the largest number of incident TB cases (2.0–2.3 million) are in India.

In order to understand the epidemiology of TB globally as well as locally, the use of molecular typing methods has become imperative not only for understanding genetic diversity and population structure of *Mycobacterium tuberculosis* complex (MTBC) but also for carrying out supervision and monitoring of TB control programs and for understanding TB epidemics [3]. In recent epidemiological studies, most currently used state-of-the-art molecular genotyping techniques are spoligotyping and mycobacterial interspersed repetitive units-variable number of tandem repeats (MIRU-VNTR) typing [4]. Numerous studies from India [5–20] have already reported characterization of MTBC isolates using spoligotyping which is PCR-based reverse hybridization blotting technique based on polymorphism in the presence or absence of "43 spacers" in the Direct Repeat (DR) locus of MTBC genome. Although spoligotyping is quick and convenient genotyping method it has some limitations due to frequent homoplasy and recent studies have pointed out that sometimes phylogenetically unrelated MTBC strains tend to have same spoligotype patterns as a result of convergent evolution or due to independent mutational changes [21]. On the other hand, MIRU-VNTR typing method is more discriminatory and is less prone to homoplasy and has thus been increasingly used along with spoligotyping as a better tool for assessment of tuberculosis progression among infected associates [22]. Therefore, we combined these two typing methods in the present study. After the introduction of an optimized 12-loci and 15-loci MIRU-VNTR typing panels, currently, 24-loci MIRU-VNTR set is being used for best possible discrimination of strongly related strains [23]. A number of studies have been reported from India based on 12-loci MIRU-VNTR typing method [13, 14, 20, 24, 25], however, only a few studies are available from India which have used 24-loci MIRU-VNTR typing method [11, 26]. Till date there is no published report of molecular genotyping study of MTBC based on combined analysis of spoligotyping and MIRU-VNTR typing from the state of Assam located in the northeastern (NE) region of India. In this study 189 MTBC isolates from the state of Assam were studied using spoligotyping along with 24-loci MIRU-VNTR typing to better understand the diversity of MTBC isolates and to find the value of different MIRU-VNTR loci for their ability and usefulness to characterize MTBC isolates from this region.

Materials and Methods

Ethical statement

This study was approved by the Ethical Committee of Regional Medical Research Centre, Dibrugarh, (Indian Council of Medical Research, NE Region). Written informed consent was

sought from all the participants who donated sputum prior to their inclusion in the study. On behalf of the children, consent was taken from their parents or guardians. Patients whose sputum samples were AFB positive were referred to the nearest DOTs (Directly Observed Treatment-Short course centre) centre for treatment.

Bacterial isolation, culture and identification

In this study a total of 189 isolates of *M. tuberculosis* complex were selected for genotyping. In order to be representative of MTBC strains circulating in Assam we collected AFB positive sputum samples over a period of 3 years (September, 2010 to December, 2013) from 14 different designated microscopy centre (DMC) of Assam. 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 were identified and confirmed by means of standard microbiological and biochemical methods [27].

DNA extraction

DNA was extracted from fresh sub-cultures of 189 MTBC isolates using cetyl-trimethyl ammonium bromide (CTAB) method [28]. DNA was eluted in TE buffer (pH 8.0) and stored at -80°C until used. DNA purity and concentration was determined spectrophotometrically [29].

Spoligotyping and MIRU-VNTR typing

Spoligotyping was carried out by amplifying the whole Direct Repeat (DR) region using the commercially available kit (Isogen Biosciences®, BV, Maarsen the Netherlands now Ocimum Biosolutions) according to a standardized method described [30] using the designated primers pairs of DRa and DRb (DRa, 5'-GGTTTTGGGTCTGACGAC-3' (biotinylated 5' end) and DRb, 5'-CCGAGAGGGGACGGAAAC-3'). GeneAmp® PCR system 9700 of Applied Biosystems was used for DNA amplification. The amplified PCR products were hybridized with nitrocellulose membrane having covalently linked 43 spacer oligonucleotides following the manufacturer's instructions. The hybridized fragments were identified using enhanced chemiluminescence system (GE Healthcare, UK Ltd., Buckinghamshire, UK). The spoligotypes representation was initially reported as 43 digits binary representation of 43 spacers, one (1) was scored for positive hybridization and zero (0) for negative hybridization.

MIRU-VNTR genotyping was performed by amplifying a standard panel of 24- MIRU-VNTR markers, comprised of 12 mycobacterial interspersed repeat units (MIRU2, MIRU4, MIRU10, MIRU16, MIRU20, MIRU23, MIRU24, MIRU26, MIRU27, MIRU31, MIRU39 and MIRU40,) 3 exact tandem repeats (ETR-A, ETR-B and ETR-C) and MIRU-VNTR loci identified by Queen's University Belfast group (QUB26, QUB11b and QUB4156) and Mtub group (Mtub04, Mtub21, Mtub29, Mtub30, Mtub34 and Mtub39) [23, 31, 32]. Polymerase chain reaction (PCR) condition was followed as previously described [33]. All PCR products of the MTBC strains were analyzed in Lab Chip (Caliper life sciences Inc., USA) or conventional submerged agarose gel electrophoresis. After sizing of the PCR fragments using appropriate molecular size marker the copy number of various alleles was assigned according to the reference table as given by Supply *et al.* [23]. The PCR experiment was repeated for confirmation in some ambiguous cases. For quality control of spoligotyping and MIRU-VNTR typing experiments DNA samples of *M. tuberculosis* H37Rv, *M. bovis* BCG, a known Beijing isolate were included as positive controls and sterile distilled water was used as a negative control.

Drug sensitivity testing (DST)

DST was carried out using conventional proportion method [34] for all first line anti-TB drugs like rifampicin (RIF), isoniazid (INH), streptomycin (STR), ethambutol (EMB), pyrazinamide (PZA) for all 189 MTBC strains.

Genotype analysis and comparison with databases

SITVITWEB database freely available online (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) [35] was used for identification of spoligotypes based on presence or absence of 43 spacers in the format of June 12, 2015. We also analyzed our data using MIRU-VNTR*plus* web server available at *www.miru-vntrplus.org* [36, 37] to characterize the MTBC isolates based on spoligotyping and 24-loci MIRU-VNTR typing profile. Spoligotypes were compared with other reference strains present in the above mentioned databases (SITVITWEB and MIRU-VNTR*plus*) for assignment of Shared International Types (SIT) and assigning specific genotypic lineages. Lineage assignment of MTBC isolates was accomplished using Jaccard's distance matrix and NJ trees based on combined spoligotyping and MIRU-VNTR typing data. For determining the discriminatory power of spoligotyping and MIRU-VNTR typing we used Hunter Gaston Discriminatory Index (HGDI) which was calculated by the formula given below [38]:

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{s} x_j(x_j - 1)$$

where D is the index of discriminatory power, N the number of unrelated strains tested, '*s*' the number of different types, and x_j the number of strains belonging to the jth type. HGDI was also used to calculate genetic diversity of individual MIRU-VNTR loci.

Principal Component Analysis

Spoligotyping data from MTBC isolates from Assam along with similar data from Kanpur, Mumbai and other parts of mainland India, Pakistan, Vietnam, Thailand, China, Korea, Japan, Sri Lanka, Uganda, South Brazil, Brazil, Colombia, Turkey, Germany and Ethiopia were subjected to principal component analysis (PCA) to find geographical areas having similar overall spoligopatterns in the reduced multivariate data space without loss of information due to dimensionality reduction. In other words, geographical localities which cluster together will tend to have similar spoligopatterns. Spoligotype data from other regions used in the above analysis were obtained from various sources including the data was also obtained from SIT-VITWEB database [9, 13, 35, 39–45]. For this analysis a correlation coefficient matrix of the selected geographical areas was generated based on spoligotype frequency data. This correlation matrix was subjected to PCA. In this analysis only few extracted principal components (PCs) explain most of the data variance, thereby eliminating redundant variables and reducing dataset complexity. The first two PCs which explain most of the variation were used to plot the geographical areas in a two dimensional scatter plot. Similarly, PCA was also used to explore clustering pattern of MTBC isolates from Assam and other geographical localities based on the HGDI of individual standard 12-loci MIRU-VNTR alleles. We used a panel of only 12-MIRU-VNTRs for this analysis because of availability of the limited data on 24-MIRU-VNTR alleles from some regions. Varimax rotated first three PCs were used to make a three dimensional scatter plot. Places that will tend to cluster together in the 3D scatter plot will have similarity in the discriminatory power of 12 corresponding MIRU-VNTR alleles. All the statistical analysis was done in SPSS® version 17.

Results

Spoligotyping

Spoligotyping of 189 clinical *M. tuberculosis* isolates revealed 89 distinct spoligopatterns (Fig 1). Seventy-six (76) out of 189 clinical isolates were represented by unique patterns whereas remaining 113 (59.78%) isolates were grouped into 13 different clusters. Lineages were identified and assigned using international database MIRU-VNTR*plus* (www.miru-vntrplus.org) and SITVITWEB (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) online tool. In the present study, 44 (23.28%) MTBC isolates had 43 unknown spoligopatterns (i.e. the patterns not yet reported in international database MIRU-VNTR*plus* and SITVITWEB). Out of these 43 unknown spoligo patterns 42 were unique and were therefore designated as orphan strains (ASM 1–23 and ASM 25–43). One unknown spoligotype was found to be represented by two isolates and was therefore considered as new shared type tentatively named ASM24. Table 1 gives the description of spoligotyping defined lineages/sublineages among *M. tuberculosis* strains isolated from Assam and their frequencies in study based on SITVITWEB database search. The Beijing lineage (SIT1) representing 35.45% (n = 189) of the total MTBC isolates was found to be the most dominant spoligotype followed by Central Asian-Delhi (CAS1_Delhi) (11.64%). One isolate which was found to belong to SIT1979 could not be assigned to any lineage based on SITVITWEB search. In the present study the discriminatory power of spoligotyping for 189 MTBC isolates from Assam was found to be 0.8637 based on HGDI.

The results of the PCA is given in Fig 2. The 2-D scatter plot based on PC 1 and 2 explained 61.83% of the total variation and revealed that based on overall pattern of spoligotypes, MTBC isolates from Assam were more similar to the MTBC isolates reported from Southeast Asian countries but are quite dissimilar to MTBC isolates from mainland states of India.

MIRU-VNTR typing of Beijing and Non-Beijing isolates

Allelic profiles and discrimination. Table 2 summarizes the allelic profiles and HGDI of the 24-loci MIRU-VNTR in MTBC isolates from Assam. The HGDI of 24-loci MIRU-VNTR revealed that out of 24-loci, twelve loci (MIRU10, MIRU16, MIRU26, MIRU31, MIRU39, Mtub04, Mtub21, Mtub39, QUB11b, QUB26, QUB4156 and ETRA) showed high discriminatory power (HGDI>0.6). Whereas nine loci (MIRU04, MIRU23, MIRU24, MIRU40, Mtub29, Mtub30, Mtub34, ETRB and ETRC) were found to be moderately discriminative (HGDI varies from 0.3 to 0.6), and three loci (MIRU02, MIRU20 and MIRU27) were poorly discriminative (HGDI<0.3).

The HGDI of individual alleles of 24-loci MIRU-VNTRs in Assam based on stratified analysis of Beijing and Non-Beijing strains revealed that discriminatory power of various MIRU-VNTR alleles was lower in Beijing isolates as compared to Non-Beijing isolates. However, seven alleles such as MIRU02, MIRU20, MIRU27, MIRU40, Mtub04, Mtub30 and Mtub34 showed slightly higher discriminatory power in Beijing MTBC isolates from Assam as compared to Non-Beijing isolates from Assam (Table 2).

Out of total 189 MTBC isolates from Assam 185 isolates were found to have unique MIRU-VNTR profiles i.e. there was absence of any clustering of 185 isolates. Four isolates were grouped into two clusters. The discriminatory power of combined 24-loci MIRU-VNTR panel and combined 15-loci MIRU-VNTR panel was same (HGDI: 0.9999 for both). However, for 12-loci MIRU-VNTR panel the HGDI was 0.9985 (Table 3).

Phylogenetic analysis. In this study we also determined the genetic relatedness among 67 Beijing isolates of MTBC based on Neighbour-joining (NJ) tree constructed using 24-loci MIRU-VNTR data based on Jaccard's distance matrix (Fig 3). Phylogenetic analysis clearly shows

Order	Clade	Type[a]	n[b]	Rules(s)[c]	Binary (d)
1	CAS	SIT2102	1	Δ4-7,Δ15-36	
2	Orphan	ASM1	1	Δ4-7,Δ17-34	
3	Orphan	ASM2	1	Δ4-7,Δ17-34,Δ19-34	
4	Orphan	ASM3	1	Δ4-7,Δ13-21,Δ23-34	
5	EAI5	SIT763	2	Δ13-18,Δ29-32,Δ34,Δ40-43	
6	Orphan	ASM4	1	Δ12-20,Δ29-32,Δ34,Δ40-43	
7	Orphan	ASM5	1	Δ4,Δ9,Δ13,Δ15-17,Δ29-32,Δ34,Δ40-43	
8	Orphan	ASM6	1	Δ1-7,Δ12-18,Δ29-32,Δ34,Δ30-43	
9	Orphan	ASM7	1	Δ1,Δ9-23,Δ29-32,Δ34,Δ40-43	
10	Orphan	ASM8	1	Δ2-3,Δ6-7,Δ9-14,Δ29-32,Δ34	
11	Orphan	ASM9	1	Δ10-17,Δ19,Δ24-37,Δ40	
12	Orphan	ASM10	1	Δ10-17,Δ19,Δ24-36	
13	CAS	SIT2419	1	Δ4-7,Δ10-37	
14	Orphan	ASM11	1	Δ1,Δ4-36	
15	CAS	SIT1962	2	Δ4-34	
16	BEIJING	SIT1	67	Δ1-34	
17	EAI5	SIT1495	1	Δ22,Δ29-32,Δ34	
18	EAII_SOM	SIT48	3	Δ29-32,Δ34,Δ40	
19	Orphan	ASM12	1	Δ29-32,Δ34,Δ36,Δ40	
20	Orphan	ASM13	1	Δ1,Δ29-32,Δ34,Δ40-43	
21	EAI5	SIT138	3	Δ29-32,Δ34,Δ40-43	
22	Orphan	ASM14	1	Δ26,Δ29-32,Δ34,Δ40-43	
23	Orphan	ASM15	1	Δ23,Δ29-32,Δ34,Δ40-43	
24	EAI5	SIT517	1	Δ29-32,Δ34,Δ40-41	
25	EAII_SOM	SIT477	1	Δ2,Δ29-32,Δ34,Δ40	
26	EAII_SOM	SIT1389	1	Δ1-2,Δ29-32,Δ34,Δ40	
27	EAI5	SIT380	1	Δ2,Δ29-32,Δ34	
28	EAI5	SIT236	1	Δ29-32,Δ34	
29	EAI6_BGD1	SIT292	1	Δ23,Δ29-32,Δ34,Δ37	
30	EAI6_BGD1	SIT591	1	Δ23,Δ29-32,Δ34	
31	EAII_SOM	SIT806	2	Δ5,Δ29-32,Δ34,Δ40	
32	EAII_SOM	SIT1316	2	Δ4-5,Δ29-32,Δ34,Δ40	
33	Orphan	ASM16	1	Δ1,Δ8,Δ23,Δ29-35	
34	H4	SIT597	1	Δ2,Δ6,Δ29-31,Δ33-36	
35	H4	SIT127	1	Δ2,Δ29-31,Δ33-36	
36	?	SIT1979	1	Δ29-36	
37	Orphan	ASM17	1	Δ4,Δ27-36	
38	Orphan	ASM18	1	Δ27-31	
39	Orphan	ASM19	1	Δ7-9,Δ17,Δ29-32,Δ34,Δ40-43	
40	Orphan	ASM20	1	Δ5-9,Δ29-34,Δ36,Δ38-43	
41	Orphan	ASM21	1	Δ6-8,Δ29-34,Δ38-43	
42	Orphan	ASM22	1	Δ2-4,Δ8-10,Δ29-32,Δ34,Δ38-43	
43	Orphan	ASM23	1	Δ8-13,Δ22,Δ33-36,Δ40-43	
44	New shared type	ASM24	2	Δ1-13,Δ28-43	
45	EAI5	SIT340	4	Δ2-3,Δ8-10,Δ29-32,Δ34	
46	Orphan	ASM25	1	Δ12-13,Δ16-17,Δ21,Δ31,Δ33-36	
47	T1	SIT53	1	Δ33-36	
48	T1	SIT628	1	Δ33-36,Δ42-43	
49	T1	SIT7	1	Δ1,Δ33-36	
50	T1	SIT2723	1	Δ2,Δ33-37	
51	T1	SIT1252	1	Δ19-20,Δ33-36	
52	H3	SIT50	2	Δ31,Δ33-36	
53	U	SIT124	1	Δ31-36	
54	H3	SIT655	1	Δ2-3,Δ31,Δ33-36	
55	Orphan	ASM26	1	Δ2-3,Δ31,Δ33-34	
56	Orphan	ASM27	1	Δ20-21,Δ33-34	
57	MANU1	SIT100	1	Δ34	
58	MANU2	SIT54	1	Δ33-34	
59	Orphan	ASM28	1	Δ7,Δ33-34	
60	LAM9	SIT891	1	Δ21-24,Δ30,Δ33-36	
61	LAM5	SIT725	1	Δ13,Δ21-24,Δ33-36,Δ43	
62	LAM11_ZWE	SIT59	1	Δ21-24,Δ27-30,Δ33-36	
63	Orphan	ASM29	1	Δ2,Δ21,Δ24-25,Δ29,Δ33-36	
64	U(CAS_ANCESTOR0)	SIT27	2	Δ4-7,Δ23-24,Δ34-37	
65	Orphan	ASM30	1	Δ4-7,Δ9,Δ23,Δ34-36	
66	Orphan	ASM31	1	Δ3-6,Δ13,Δ19,Δ33-36	
67	Orphan	ASM32	1	Δ22-34,Δ37-39,Δ43	
68	EAI	SIT1391	2	Δ22-34,Δ37	
69	U	SIT1374	1	Δ22-34	
70	Orphan	ASM33	1	Δ23,Δ25-32,Δ34,Δ37	
71	Orphan	ASM34	1	Δ1-3,Δ23,Δ25-27,Δ29-34,Δ37	
72	Orphan	ASM35	1	Δ26-29,Δ31	
73	Orphan	ASM36	1	Δ22-29,Δ32	
74	Orphan	ASM37	1	Δ5-10,Δ23-24,Δ29-34	
75	Orphan	ASM38	1	Δ4-10,Δ29-34	
76	CAS	SIT22	1	Δ4-7,Δ20-35	
77	CAS	SIT357	1	Δ4-7,Δ23-36	
78	CAS1_DELHI	SIT1091	1	Δ4-7,Δ23-34,Δ36	
79	CAS	SIT2726	1	Δ4-7,Δ21-36	
80	CAS	SIT1093	1	Δ4-7,Δ21-34	
81	CAS1_DELHI	SIT26	20	Δ4-7,Δ23-34	
82	Orphan	ASM39	1	Δ1,Δ4-7,Δ12,Δ23-34,Δ37-38	
83	CAS1_DELHI	SIT428	1	Δ4-7,Δ23-34,Δ37	
84	CAS	SIT1151	1	Δ4-10,Δ23-35	
85	Orphan	ASM40	1	Δ4-7,Δ9-10,Δ23-36	
86	CAS2	SIT288	1	Δ4-10,Δ23-34	
87	Orphan	ASM41	1	Δ4-9,Δ24-34	
88	Orphan	ASM42	1	Δ2,Δ6-9,Δ22-36	
89	Orphan	ASM43	1	Δ1-7,Δ20-34	

a = Shared International Type strains are designated as in the SITVIT WEB database and MIRU-VNTR*plus* web tool and orphan types have been designated as ASM1-23, ASM25-43 and one new shared type named ASM24.

b = Frequency of *Mycobacterium tuberculosis* complex (MTBC) isolates, c = Rules as identified by absence of spacers,

d = Dark boxes represented positive hybridization and empty boxes represents no hybridization i.e. absence of spacers.

?=Undefined clade

Fig 1. Spoligotype patterns of MTBC circulating in Assam.

doi:10.1371/journal.pone.0145860.g001

Table 1. Distribution of different lineages of MTBC Spoligotypes isolated from Assam in the study based on the classification by SITVITWEB.

Lineage (Family/Sub-family)	No. of isolates	(Prevalence in %, n = 189)
BEIJING	67	35.45
CAS1_DELHI	22	11.64
EAI5	13	6.88
EAI1_SOM	9	4.76
CAS	9	4.76
T1	5	2.65
H3	3	1.59
EAI6_BGD1	2	1.06
H4	2	1.06
U(CAS_ANCESTORO)	2	1.06
EAI	2	1.06
U	2	1.06
MANU1	1	0.53
MANU2	1	0.53
LAM9	1	0.53
LAM5	1	0.53
LAM11_ZWE	1	0.53
CAS2	1	0.53
Undefined	1	0.53
Orphan	44	23.28

doi:10.1371/journal.pone.0145860.t001
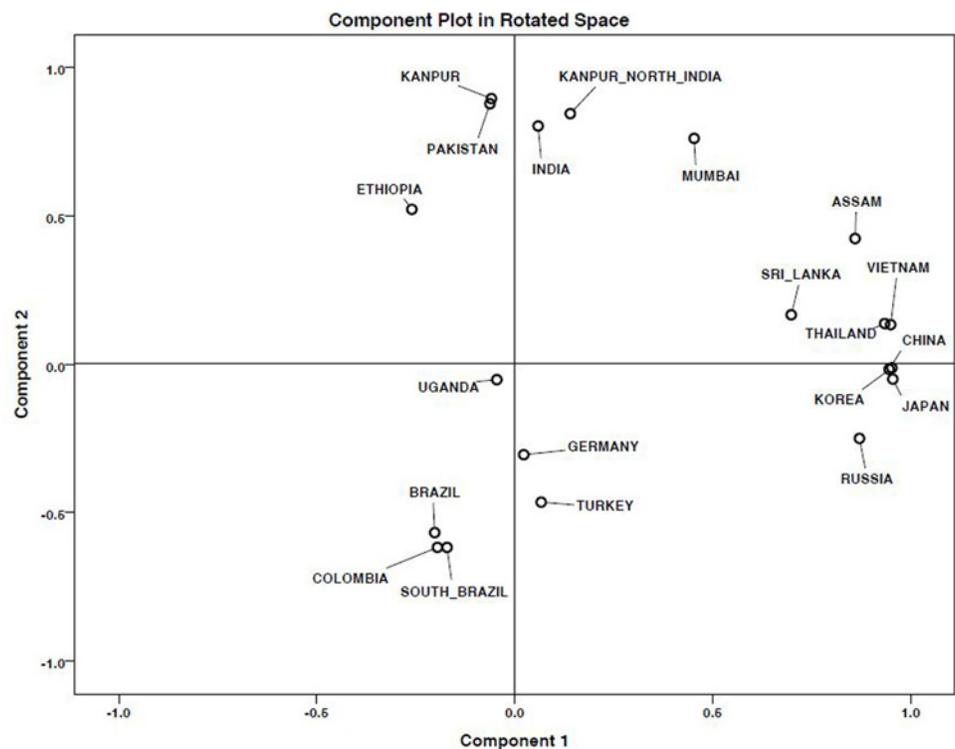
Fig 2. A two dimensional scatter plot based on principal component analysis (PCA). Geographical regions having similarity in patterns of spoligotypes tend to clusters together.

doi:10.1371/journal.pone.0145860.g002

Table 2. The allelic profiles and Hunter Gaston Discriminatory Index (HGDI) of each of the 24 MIRU-VNTR loci in *Mycobacterium tuberculosis* isolates (n = 189) from Assam.

SL. No.	Alias	Locus	Copy Numbers of Tandem Repeats: Beijing/Non-Beijing																HGDI		
			0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Beijing	Non-Beijing	Total
1	MIRU02	154	0/0	0/0	65/119	0/0	2/3	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.0588	0.0484	0.0518
2	Mtub04	424	1/1	3/5	29/74	1/14	24/17	8/4	1/5	0/1	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0.6775	0.5998	0.6461
3	ETRC	577	0/0	0/0	17/75	4/6	46/41	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.4677	0.5109	0.5513
4	MIRU04	580	1/1	0/0	59/73	0/2	1/13	6/33	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.2194	0.5617	0.4664
5	MIRU40	802	0/0	5/7	10/16	43/78	6/15	3/6	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.5586	0.5578	0.5555
6	MIRU10	960	0/0	0/0	2/6	33/14	17/27	3/33	5/15	6/12	1/10	0/5	0/0	0/0	0/0	0/0	0/0	0/0	0.6866	0.8359	0.8259
7	MIRU16	1644	0/0	0/0	5/9	7/9	19/37	29/45	7/19	0/2	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.7155	0.7426	0.7311
8	Mtub21	1955	0/0	0/1	2/8	5/17	18/36	31/13	7/11	2/13	0/1	0/12	1/3	1/3	0/0	0/1	0/0	0/3	0.7056	0.8537	0.8316
9	MIRU20	2059	0/0	0/1	56/107	4/5	0/1	3/3	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.2967	0.2279	0.2514
10	QUB11b	2163	0/0	1/12	13/40	4/10	5/8	16/12	24/13	0/2	2/13	1/10	1/4	0/1	0/0	1/0	0/0	0/0	0.7779	0.8392	0.8407
11	ETRA	2165	0/0	1/4	2/6	11/21	46/47	0/0	0/0	1/8	0/4	1/5	1/4	3/22	1/1	0/0	0/0	0/0	0.5052	0.7842	0.7084
12	Mtub29	2347	0/1	0/0	5/3	0/3	14/54	48/60	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.4441	0.5655	0.5448
13	Mtub30	2401	0/0	2/11	23/82	0/9	39/17	2/2	1/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.5495	0.5192	0.5992
14	ETRB	2461	0/0	2/9	53/69	5/4	2/4	5/32	0/3	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.3668	0.6080	0.5409
15	MIRU23	2531	0/0	0/0	0/4	1/3	0/1	48/59	17/44	1/5	0/1	0/0	0/5	0/0	0/0	0/0	0/0	0/0	0.4283	0.6361	0.5757
16	MIRU24	2687	0/0	55/74	9/16	3/31	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.3107	0.5548	0.4868
17	MIRU26	2996	0/0	2/3	9/44	0/4	2/9	3/4	27/23	21/29	2/2	1/0	0/3	0/1	0/0	0/0	0/0	0/0	0.7273	0.7751	0.7788
18	MIRU27	3007	0/0	0/0	1/5	9/11	56/104	0/1	0/0	0/0	1/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.2872	0.2655	0.2725
19	Mtub34	3171	0/0	3/3	0/3	13/22	48/89	3/3	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0.4518	0.4369	0.4403
20	MIRU31	3192	0/0	0/0	1/0	11/21	5/18	9/13	41/67	0/3	0/0	0/0	0/0	0/0	0/0	0/0	0/1	0/0	0.5834	0.6403	0.6194
21	Mtub39	3690	0/0	0/6	3/8	6/8	49/51	0/11	6/24	3/10	0/1	0/1	0/1	0/0	0/1	0/0	0/0	0/0	0.4518	0.7667	0.6803
22	QUb26	4052	0/0	2/14	2/7	4/10	4/17	1/8	3/12	10/12	11/16	15/15	15/10	0/1	0/0	0/0	0/0	0/0	0.8521	0.9020	0.8922
23	QUB4156	4156	4/1	10/49	35/26	16/29	0/11	2/5	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.6531	0.7328	0.7401
24	MIRU39	4348	0/0	1/3	11/36	16/27	38/52	1/3	0/0	0/0	0/1	0/0	0/0	0/0	0/0	0/0	0/0	0/0	0.6029	0.6866	0.6622

doi:10.1371/journal.pone.0145860.t002

Table 3. Hunter Gaston Discriminatory Index (HGDI) and cluster results based on MIRU-VNTR loci analysis of 189 *Mycobacterium tuberculosis* complex (MTBC) isolates from Assam.

Typing Method	Total No. of patterns	No. of uniquetypes	Total no.of clusters	Total no. of isolates in clusters (%)	Maximum no. of isolates in a cluster	HGDI
24 loci MIRU-VNTR	187(N = 189)	185	2	4(2.11)	2	0.9999
15 loci MIRU-VNTR	187(N = 189)	185	2	4(2.11)	2	0.9999
12 loci MIRU-VNTR	172(N = 189)	161	11	28(14.81)	5	0.9985

doi:10.1371/journal.pone.0145860.t003

that Beijing isolates from Assam represent two major groups (Group A and Group B) each comprising of several sub-groups.

We also analyzed 78 Non-Beijing isolates (excluding orphan isolates) obtained in the present study for strain lineage identification based on 24-loci MIRU-VNTR and spoligotyping.

Fig 3. Neighbour-Joining (NJ) tree showing genetic diversity of 67 Beijing isolates from Assam based on 24 MIRU-VNTR loci. Two major clusters (Group A and Group B) and their sub-lineages can be seen from this figure.

doi:10.1371/journal.pone.0145860.g003

MTBC lineage distribution was determined by constructing a NJ tree based on 24-loci MIR-U-VNTR and spoligotyping. This analysis was achieved by submitting the MTBC isolate geno-types to the MIRU-VNTR*plus* database which in turn carried lineage prediction of submitted isolates by comparing to MIRU-VNTR and spoligotyping profiles of their reference strains (Fig 4). This analysis revealed that 32 (41.02%) out of 78 isolates belong to Central-Asian (CAS) lineage, out of which 22(68.8%; n = 32) belong to CAS1-Delhi. Another major lineage in this group was East-African-Indian (EAI), 33.33% (n = 78). Among EAI lineage the EAI5 (50%, n = 26) was most prominent. Other minor lineages include: T1 (5), H3 (3), H4 (2), U (2), U(CAS_ANCESTORO) (2), MANU1 (1), MANU2 (1), LAM5 (1), LAM9 (1), LAM11_ZWE (1) and undefined (1).

In this study we also compared phylogeographic diversity of MTBC isolates belonging to Beijing family from Assam with similar strains from Singapore, Cambodia, Hong Kong, China, Taiwan, South Africa and rest of India where these strains are also reported. The results of this analysis are shown as a NJ radiation tree (Fig 5) which was constructed using 24-MIRU-VNTR alleles. Beijing isolates from Assam grouped into two distinct clusters, one major cluster formed a distinct lineage whereas the Beijing strains in the minor cluster were close to some strains from South Africa and Taiwan.

In order to explore the similarity in the MTBC strains based on the discriminatory power of individual MIRU-VNTR alleles we carried out PCA analysis. The results are given in the Fig 6. The first three PCs explained about 84.2% of the total variance in the dataset. The 3D scatter plot based on the first three PCs revealed three major clusters of geographical localities. From this figure it is clear that MTBC isolates from Assam cluster with MTBC isolates from other Asian countries. MTBC isolates from Cameroon formed a separate cluster. Similarly, MTBC isolates from South American countries (Brazil and Colombia) and European countries (Bulgaria and Morocco) formed a separate cluster.

MIRU-VNTR typing of orphan strains

In order to understand the genetic affiliation of MTBC orphan spoligotypes discovered in the present study we carried out a phylogenetic analysis including all the 44 MTBC orphan isolates with reference MTBC strains present in the MIRU-VNTR*plus* database [36]. The NJ tree was constructed using combined dataset of 24-loci MIRU-VNTR and spoligotypes as implemented by MIRU-VNTR*plus* tree based identification. The result of the above analysis is given in Fig 7, which clearly shows that majority of these orphan MTBC isolates from Assam represent two new sub-clades namely sub-clade one: Assam/EAI which is related to EAI-Clade and sub-clade two: Assam/CAS which is related to Delhi/CAS clade. One orphan strain clustered with the LAM family of the reference MTBC strain.

Status of Multidrug-resistance TB (MDR-TB)

The drug susceptibility testing study revealed that 18 (9.52%) isolates were multidrug-resistant (MDR-TB) i.e. these isolates were resistant to at least rifampicin (RIF) and isoniazid (INH) first line anti-tubercular drugs. Eighty-two (43.38%) isolates were found to be resistant to at least one or more anti-TB drugs. Among the 189 MTBC isolates, 63 (33.33%) isolates were resistant to streptomycin (STR). The lowest resistance was observed against pyrazinamide (PZA) with 11(5.82%) of the total isolates. Prevalence of MDR in Beijing and Non-Beijing strains was 10.44% and 9.01%, respectively. A detailed table of the results for each strain: spoli-gotyping, clade, VNTR typing, origin region, and results of antibiotic susceptibility is given in S1 Table.
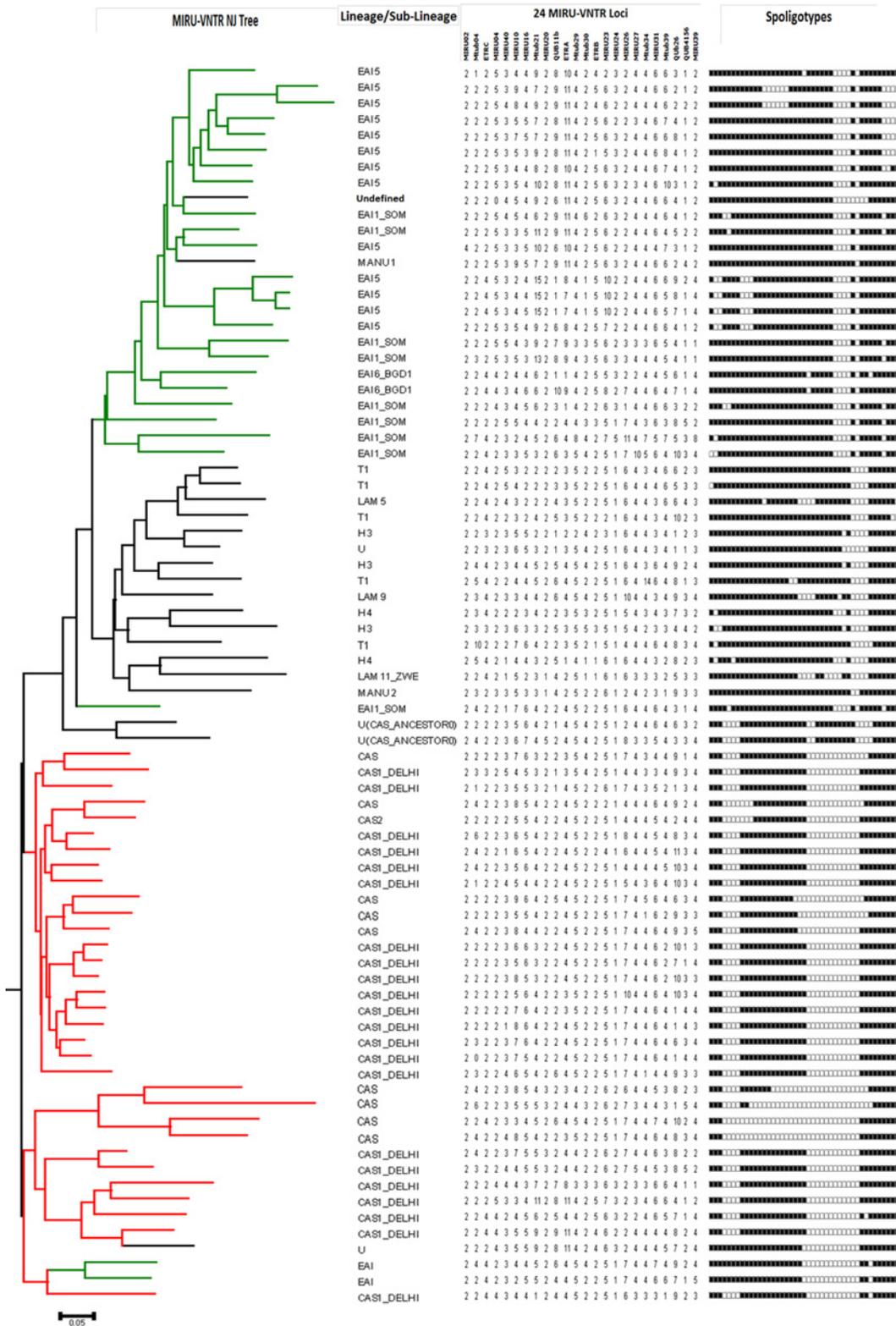
Fig 4. Genotypic diversity of 78 Non-Beijing and Non-orphan MTBC isolates from Assam. A MIRU-VNTR based dendrogram was generated using NJ tree analysis. *M. tuberculosis* strain lineages were identified using MIRU-VNTR*plus* identification database. 24-loci MIRU-VNTR alleles and spoligotypes from 78 isolates are also represented along with MIRU-VNTR NJ tree.
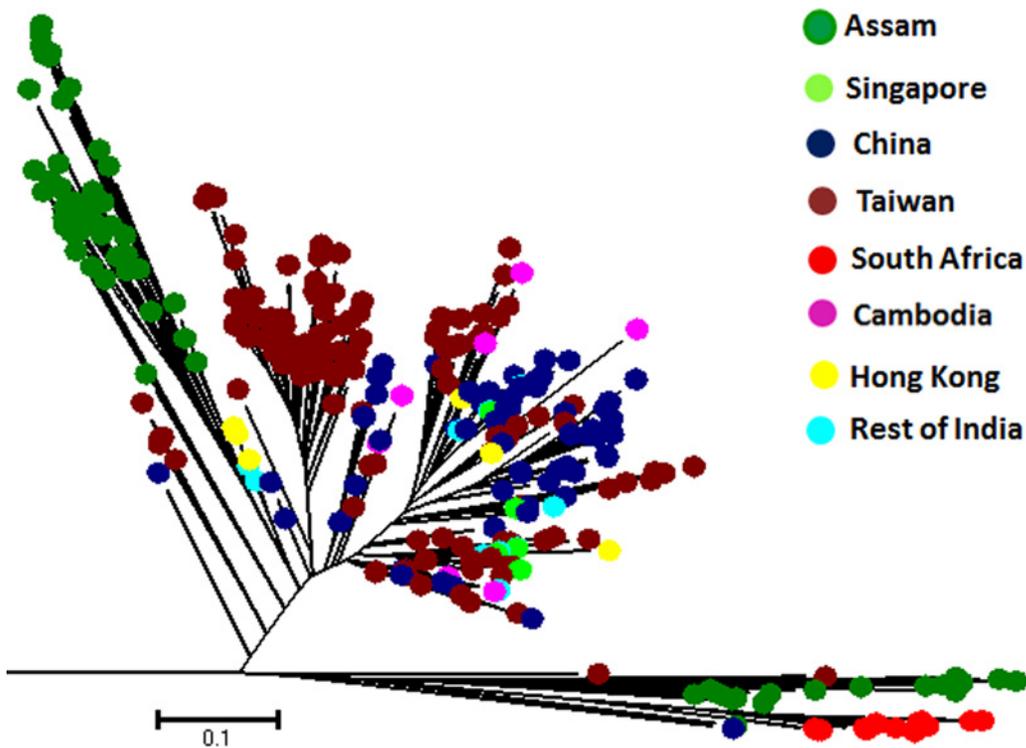
doi:10.1371/journal.pone.0145860.g004

Fig 5. Comparison of Beijing MTBC isolates from Assam with Beijing isolates from other countries (Singapore, Cambodia, Hong Kong, China, Taiwan, South Africa and rest of India) where Beijing isolates are also reported. The radiation tree was constructed based on 24-MIRU-VNTR loci data using Neighbour-Joining method. Beijing isolates from Assam group into two well defined clusters, one major MTBC cluster which is close to some Beijing strains from Hong Kong, China, Taiwan and rest of India whereas the Beijing strains in the minor cluster were phylogenetically more close to a few Beijing strains from South Africa and Taiwan.

doi:10.1371/journal.pone.0145860.g005

Discussion

The present study is the first study to give an insight into the genetic diversity of MTBC strains from Assam using simultaneously two classical genotyping methodologies viz., spoligotyping and 24-MIRU-VNTR typing. The state Assam is located in the NE region of India and shares international borders with Bhutan and Bangladesh. Assam has sub-tropical monsoon climate and experiences heavy rainfall and high humidity. India being a vast country having great variety in ethnicity and geographical diversity shows differential predominance of MTBC lineages in northern, southern, western and central regions. In northern part of India CAS1_Delhi is predominant strain whereas in Southern and Central India the dominant strain is EAI1 [13, 46–49]. In a recent study from Mumbai, western part of India, a high percentage of Beijing isolates (23%) was reported from a hospital based study which was associated with high prevalence of MDR-TB status [46]. In the present study, we have found 35.45% clinical isolates of MTBC from Assam belonged to Beijing family followed by CAS1_Delhi (11.64%), EAI5 (6.88%), EAI1_SOM (4.76%), CAS (4.76%) and T1 (2.65%) families. Apart from this, very few isolates belonged to H3, EAI6_BGD1, H4, U, U(CAS_ANCESTORO), EAI, LAM5, LAM9, MANU1, MANU2, LAM11_ZWE and CAS2 families. Several recent studies [50–53] have also reported a significant increase in the frequency of Beijing genotypes and this genotype has also been shown to be associated with increased anti-tubercular drugs resistance [51, 53, 54] moreover these strains have been found to have high adaptability to the host intracellular environment in comparison to other lineages of MTBC [55, 56]. A recent study conducted in Dhaka,
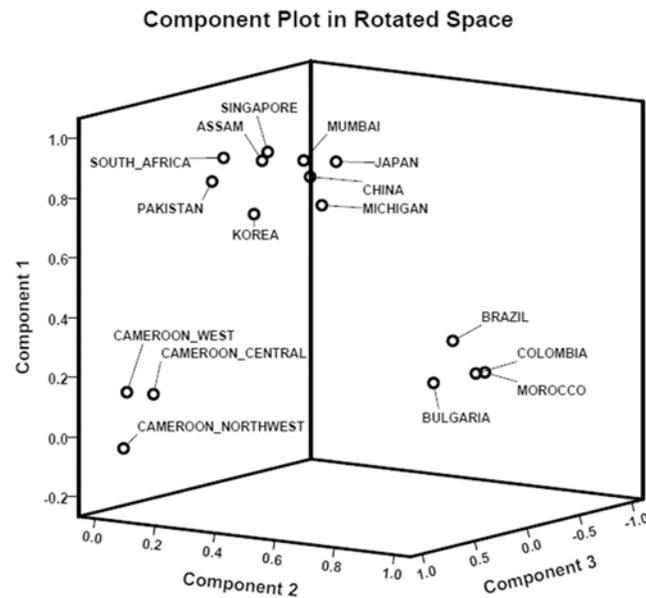
Fig 6. A three dimensional scatter plot based on principal component analysis (PCA) showing relatedness of different geographical regions based on their similarity in patterns of 12-loci MIRU-VNTR diversity in MTBC from Assam and other geographical localities. Different geographical areas were clustered in three groups and the state of Assam clustered with other Asian countries.

doi:10.1371/journal.pone.0145860.g006

Bangladesh which is not far away from our study area shown increase in multi drug resistant MTBC to be associated with Beijing strains [57]. In the present study genetic diversity of MTBC strains belonging to the Beijing family isolated from Assam were further characterized by 24-loci MIRU-VNTR profiling and it was found that the Beijing strains from Assam showed two major lineages and five sub-lineages based on NJ phylogenetic tree based analysis. Which indicates highly diverse strains of Beijing family are circulating in Assam and the present finding supports the hypothesis that Beijing isolates are undergoing continuous evolution and active transmission in Assam. The apparent success and predominance of Beijing strains in Assam may be due to wide variety of reasons including human population movements and proximity of Assam to geographical areas where predominance of Beijing strains is high. The present study also revealed that 44 (23.28%) spoligotypes from Assam were found to be orphan which advocates that a significant number of the MTBC isolates prevalent in India are still not well recognized. Interestingly, the NJ based phylogenetic tree constructed revealed that these orphan isolates in fact represent two new clades Assam/EAI and Assam/CAS related to EAI and Delhi/CAS, respectively.

In our study, it was observed through spoligotyping based lineage identification that five MTBC isolates belonged to T1 family, characterized by the deletion of spacers 33–36. This 'T1' ancestor is believed to be the prevailing genotype in Delhi region [13]. It was also interesting to find MTBC isolates belonging to CAS and EAI1_SOM family (4.76%), and is the fourth dominant MTBC strain circulating in Assam, NE region of India. In Somalia, this EAI1_SOM family was first reported, and according to the SITVITWEB database (http://www.pasteur-guadeloupe.fr/tb/bd_myco.html) it was also reported from Europe, Asia, Middle East and South Africa [58]. Other clinically important MTBC isolates found in Assam belonged to MANU family but their prevalence was low. The isolate MANU1 (SIT100) is characterized by the absence of only one spacer (spacer 34), whereas, MANU2 (SIT54) strains have two missing spacers (spacer 33 and 34). MANU strains are more frequently reported from mainland India
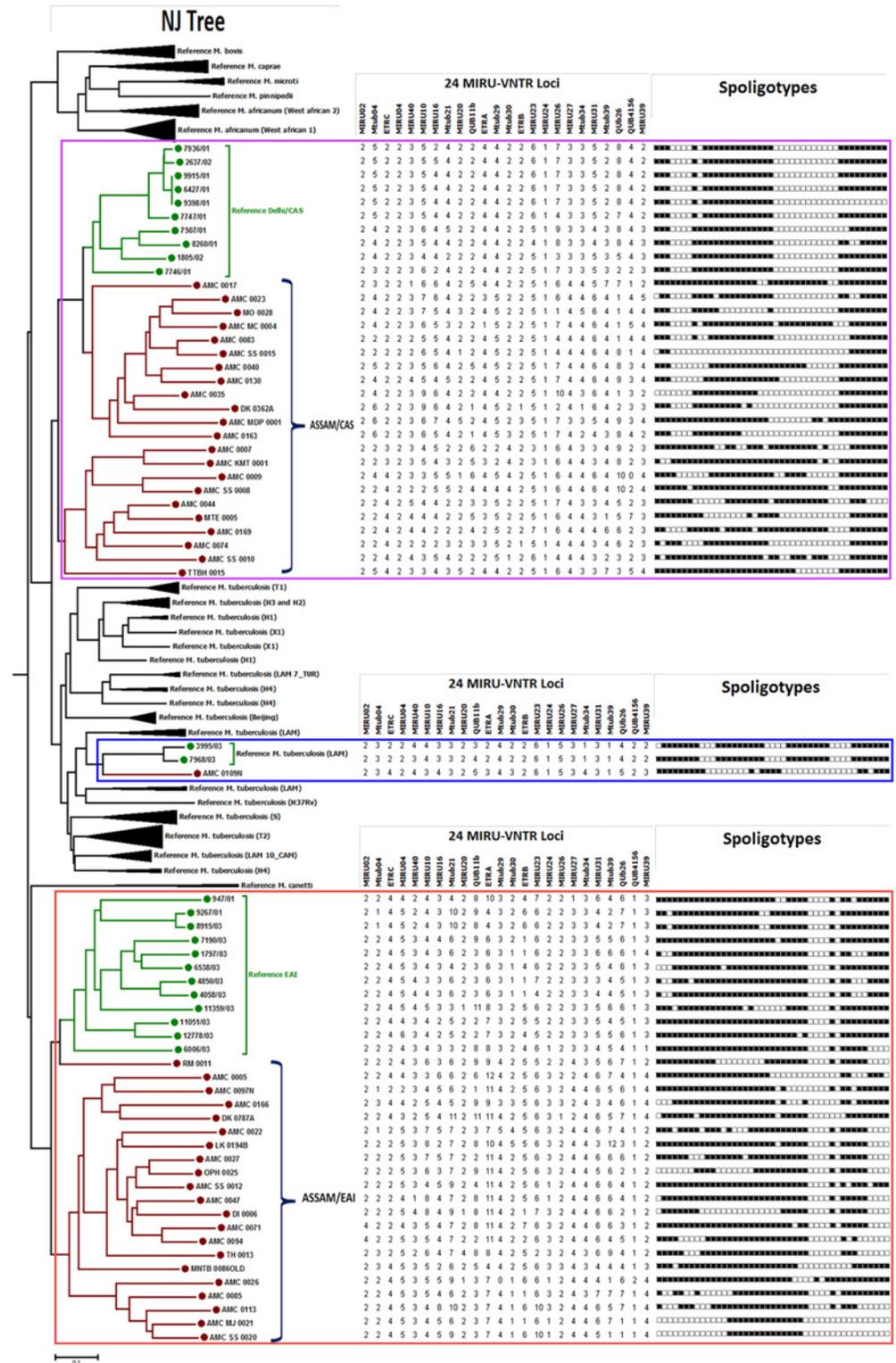
Fig 7. Fig 7 showing phylogenetic relationship of orphan strains of MTBC from Assam with reference MTBC isolates available at the database MIRU-VNTR*plus*. The NJ tree was constructed using 24-loci MIRU-VNTR data. 24-loci MIRU-VNTR alleles and spoligotypes from 44 isolates are also represented along with MIRU-VNTR NJ tree. This phylogenetic tree clearly shows that the orphan strains of MTBC from Assam represent two new well defined sub-clades namely Assam/EAI and Assam/CAS.

[59] and it is believed that MANU lineage has originated from India [16]. The PCA unambiguously established that the overall spoligotyping patterns of MTBC isolates from Assam are more or less like MTBC isolates reported from Southeast Asian countries rather than the pattern found in mainland states of India. In rest of India Beijing/Beijing-like strains of MTBC are less prevalent, whereas CAS and EAI are more dominant strains of MTBC. On the other hand, in Assam (NE region, India) as in other Southeast Asian countries Beijing isolates of MTBC are more predominant.

High degrees of variation in the allelic diversity of MIRU-VNTRs are reported among MTBC strains in different geographical locations [13, 25, 26, 60]. In the present study MIRU alleles (MIRU10, MIRU26, Mtub21, QUB11b, and QUB26) showed high discriminatory power (HGDI>0.7). On the other hand, lowest discriminatory power was found in alleles MIRU02 and MIRU27 (HGDI<0.3). An earlier study (Mokrousov *et al.*, 2004) has shown that MIRU26 and MIRU10 have high HGDI [61]. 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 slightly moderate discriminatory power (HGDI: 0.2514). Moderate discriminatory power for MIRU20 has also been reported from Pakistan (HGDI: 0.5080). From the PCA analysis of MIRU-VNTR alleles it was found that MTBC isolates of Assam cluster with MTBC isolates from other Asian countries. This clustering may be due to high genetic diversity of MTBC isolates from Asian countries.

Recent studies have shown that Beijing lineage of MTBC shows genetic heterogeneity [62] and often various branches can be separated into modern (typical) and ancient (atypical) sub-lineages. Ancient sublineages of Beijing strains of MTBC are predominantly reported from Asian countries like northeastern China, Korea and Japan [63]. It can be presumed that most MTBC isolates of Beijing lineages from Assam may belong to ancient Beijing sub-lineage because in NJ tree (S1 Fig) based on 24-loci MIRU-VNTR analysis Beijing isolates from Assam clustered with some Beijing isolates from Korea belonging to ancient sub-lineage. In the present study Beijing strains from Assam were found to form two distinct clusters on the basis of 24-MIRU-VNTR analysis. Recently [21], it has been suggested that some strains which are traditionally referred to as Beijing strains might actually belong to different phylogenetic lineage, which they have referred to as "Pseudo-Beijing" strains. Thus more elaborate studies are required to characterize such strains keeping in view the mutation dynamics within the DR locus [21].

Our DST results revealed that prevalence of MDR was slightly more among Beijing isolates (10.44%) as compared to non-Beijing isolates. However, this difference in prevalence of MDR-TB was not statistically significant.

Limitations of the study

The small sample size in the present study could present a potential bias regarding representativeness of the strains of MTBC circulating in Assam. In spite of this limitation the present study showed high diversity of MTBC spoligotypes (89 spoligotypes) circulating in Assam. However, future studies with more sample size will throw more light on the diversity of MTBC in Assam.

The second limitation which needs to be pointed out is regarding usefulness of spoligotyping and MIRU-VNTR typing in phylogenetic analysis. Numerous authors have questioned the reliability of spoligotyping for making phylogenetic inference due to homoplasy [64–66]. However, it has been shown that MIRU-VNTR typing is more reliable because of low rate of homoplasy and moreover it has been suggested that only deep phylogenetic inferences may be

unreliable and use of more loci of VNTRs will increase reliability of phylogenetic inference [67]. Thus in the present study we have used 24-loci MIRU-VNTR analysis to maximize the phylogenetic reliability of MIRU-VNTR analysis. None the less, the phylogenetic inference drawn by spoligotyping and MIRU-VNTR typing need to be viewed with caution and further studies like whole genome sequencing of MTBC will be required as corroborative data.

In conclusion, the present study has shown the predominance of Beijing isolates in Assam which is a matter of great concern from the public health point of view because it is widely known that Beijing strains of MTBC have high degree of transmission potential in addition to their association with spread of MDR tuberculosis. Other interesting finding of the present study is the discovery of two new clades of MTBC isolates circulating in Assam. More elaborate longitudinal studies are required to be undertaken in this region to understand the transmission dynamics of MTBC. The information on genetic diversity of MTBC isolates obtained from the present study may prove to be useful baseline information for future epidemiological studies on tuberculosis in Assam and will also be helpful for evaluating TB control programs in this region.

Supporting Information

S1 Fig. Comparison of Beijing MTBC isolates from Assam with Beijing isolates from other countries (Japan, Korea, Singapore, Cambodia, Hong Kong, China, Taiwan, South Africa and rest of India) where Beijing isolates are also reported. The radiation tree was constructed based on 24-MIRU-VNTR loci data using Neighbour-Joining method. Beijing isolates from Assam group into two well defined clusters, one major MTBC cluster which is close to some ancient Beijing strains from Korea.
(TIF)

S1 Table. Detailed genotyping (Spoligotyping and 24-MIRU-VNTR typing) and drug-susceptibility testing results of 189 *Mycobacterium tuberculosis* strains isolated from Assam, North-eastern state of India.
(XLS)

Acknowledgments

Author Contributions

Conceived and designed the experiments: KRD RB KN JM. Performed the experiments: KRD RB KM. Analyzed the data: KRD KN SB. Contributed reagents/materials/analysis tools: KRD KN. Wrote the paper: KRD RB SB KM KN JM.

References

1. John TJ. Tuberculosis control in India: why are we failing? Indian pediatrics. 2014, 51(7):523–527. | PMID: 25031127

2. WHO. WHO Global tuberculosis report 2014. Available: http://www.who.int/tb/publications/global_report/en/.

3. Banu S, Rahman MT, Uddin MK, Khatun R, Khan MS, Rahman MM, et al. Effect of active case finding on prevalence and transmission of pulmonary tuberculosis in Dhaka central jail, Bangladesh. PloS one. 2015, 10(5):e0124976. doi: 10.1371/journal.pone.0124976 PMID: 25933377

4. Brudey K, Filliol I, Ferdinand S, Guernier V, Duval P, Maubert B, et al. Long-term population-based genotyping study of *Mycobacterium tuberculosis* complex isolates in the French departments of the Americas. J Clin Microbiol. 2006, 44(1):183–191. doi: 44/1/183 | PMID: 16390968

5. Ahmed N, Hasnain SE. Molecular epidemiology of tuberculosis in India: moving forward with a systems biology approach. Tuberculosis (Edinb). 2011, 91(5):407–413. doi: 10.1016/j.tube.2011.03.006 PMID: 21514230

6. Chatterjee A, D'Souza D, Vira T, Bamne A, Ambe GT, Nicol MP, et al. 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. 2010, 48(10):3593–3599. doi: 10.1128/JCM.00430-10 PMID: 20720028

7. Desikan P, Chauhan DS, Sharma P, Panwalkar N, Gautam S, Katoch VM. A pilot study to determine genetic polymorphism in *Mycobacterium tuberculosis* isolates in Central India. Indian J Med Microbiol. 2012, 30(4):470–473. doi: 10.4103/0255-0857.103774 PMID: 23183476

8. Gupta A, Kulkarni S, Rastogi N, Anupurba S. A study of *Mycobacterium tuberculosis* genotypic diversity & drug resistance mutations in Varanasi, north India. The Indian journal of medical research. 2014, 139(6):892–902. | PMID: 25109724

9. Kulkarni S, Sola C, Filliol I, Rastogi N, Kadival G. Spoligotyping of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mumbai, India. Res Microbiol. 2005, 156(4):588–596. doi: S0923-2508(05)00016 | PMID: 15862459

10. Mathuria JP, Sharma P, Prakash P, Samaria JK, Katoch VM, Anupurba S. Role of spoligotyping and IS6110-RFLP in assessing genetic diversity of *Mycobacterium tuberculosis* in India. Infect Genet Evol. 2008, 8(3):346–351. doi: 10.1016/j.meegid.2008.02.005 PMID: 18372222

11. Sankar MM, Singh J, Diana SC, Singh S. Molecular characterization of *Mycobacterium tuberculosis* isolates from North Indian patients with extrapulmonary tuberculosis. Tuberculosis (Edinb). 2013, 93(1):75–83. doi: 10.1016/j.tube.2012.10.005 PMID: 23140853

12. Shanmugam S, Selvakumar N, Narayanan S. Drug resistance among different genotypes of *Mycobacterium tuberculosis* isolated from patients from Tiruvallur, South India. Infect Genet Evol. 2011, 11(5):980–986. doi: 10.1016/j.meegid.2011.03.011 PMID: 21453793

13. Sharma P, Chauhan DS, Upadhyay P, Faujdar J, Lavania M, Sachan S, et al. Molecular typing of *Mycobacterium tuberculosis* isolates from a rural area of Kanpur by spoligotyping and mycobacterial interspersed repetitive units (MIRUs) typing. Infect Genet Evol. 2008, 8(5):621–626. doi: 10.1016/j.meegid.2008.05.002 PMID: 18567544

14. Sharma R, Kalyani M, Kajal NC, Sarin BC, Sehajpal PK. Genetic diversity in clinical *Mycobacterium tuberculosis* isolates from Punjab. Int J Tuberc Lung Dis. 2008, 12(10):1122–1127 | PMID: 18812040

15. Singh UB, Arora J, Suresh N, Pant H, Rana T, Sola C, et al. Genetic biodiversity of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in India. Infect Genet Evol. 2007, 7(4):441–448. doi: S1567-1348(07)00004-4 | PMID: 17317334

16. Singh UB, Suresh N, Bhanu NV, Arora J, Pant H, Sinha S, et al. Predominant tuberculosis spoligotypes, Delhi, India. Emerg Infect Dis. 2004, 10(6):1138–1142. doi: 10.3201/eid1006.030575 PMID: 15207071

17. Suresh N, Arora J, Pant H, Rana T, Singh UB. Spoligotyping of *Mycobacterium tuberculosis* DNA from Archival Ziehl-Neelsen-stained sputum smears. J Microbiol Methods. 2007, 68(2):291–295. doi: S0167-7012(06)00266-1 | PMID: 17070946

18. Thomas SK, Iravatham CC, Moni BH, Kumar A, Archana BV, Majid M, et al. Modern and ancestral genotypes of *Mycobacterium tuberculosis* from Andhra Pradesh, India. PloS one. 2011, 6(11):e27584. doi: 10.1371/journal.pone.0027584 PMID: 22114678

19. Vadwai V, Daver G, Udwadia Z, Sadani M, Shetty A, Rodrigues C. Clonal population of *Mycobacterium tuberculosis* strains reside within multiple lung cavities. PloS one. 2011, 6(9):e24770. doi: 10.1371/journal.pone.0024770 PMID: 21935462

20. Varma-Basil M, Kumar S, Arora J, Angrup A, Zozio T, Banavaliker JN, et al. Comparison of spoligotyping, mycobacterial interspersed repetitive units typing and IS6110-RFLP in a study of genotypic diversity of *Mycobacterium tuberculosis* in Delhi, North India. Mem Inst Oswaldo Cruz. 2011, 106(5):524–535. doi: S0074-02762011000500002 | PMID: 21894371

21. Fenner L, Malla B, Ninet B, Dubuis O, Stucki D, Borrell S, et al. "Pseudo-Beijing": evidence for convergent evolution in the direct repeat region of *Mycobacterium tuberculosis*. PloS one. 2011, 6(9): e24737. doi: 10.1371/journal.pone.0024737 PMID: 21935448

22. Oelemann MC, Diel R, Vatin V, Haas W, Rusch-Gerdes S, Locht C, et al. 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. 2007, 45(3):691–697. doi: JCM.01393-06 | PMID: 17192416

23. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J Clin Microbiol. 2006, 44(12):4498–4510. doi: 10.1128/JCM. 01392-06 PMID: 17005759

24. Stavrum R, Myneedu VP, Arora VK, Ahmed N, Grewal HM. In-depth molecular characterization of *Mycobacterium tuberculosis* from New Delhi—predominance of drug resistant isolates of the 'modern' (TbD1) type. PloS one. 2009, 4(2):e4540. doi: 10.1371/journal.pone.0004540 PMID: 19234602

25. Chatterjee A, Mistry N. MIRU-VNTR profiles of three major *Mycobacterium tuberculosis* spoligotypes found in western India. Tuberculosis (Edinb). 2013, 93(2):250–256. doi: 10.1016/j.tube.2012.10.004 PMID: 23219234

26. Vadwai V, Shetty A, Supply P, Rodrigues C. Evaluation of 24-locus MIRU-VNTR in extrapulmonary specimens: study from a tertiary centre in Mumbai. Tuberculosis (Edinb). 2012, 92(3):264–272. doi: 10.1016/j.tube.2012.01.002 PMID: 22321730

27. Kent PT. Public health mycobacteriology: a guide for the level III laboratory. Edited by Kubica GP. Atlanta, Ga.:: U.S. Dept. of Health and Human Services, Public Health Service, Centers for Disease Control; 1985.

28. Somerville W, Thibert L, Schwartzman K, Behr MA. Extraction of *Mycobacterium tuberculosis* DNA: a question of containment. J Clin Microbiol. 2005, 43(6):2996–2997. doi: 10.1128/JCM.43.6.2996–2997. 2005 PMID: 15956443

29. Noguti EN, Leite CQ, Malaspina AC, Santos AC, Hirata RD, Hirata MH, et al. Genotyping of *Mycobacterium tuberculosis* isolates from a low-endemic setting in northwestern state of Parana in Southern Brazil. Mem Inst Oswaldo Cruz. 2010, 105(6):779–785. | PMID: 20944992

30. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol. 1997, 35(4):907–914. | PMID: 9157152

31. Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. J Clin Microbiol. 2001, 39(10):3563–3571. doi: 10.1128/JCM.39.10.3563–3571.2001 PMID: 11574573

32. Le Fleche P, Fabre M, Denoeud F, Koeck JL, Vergnaud G. High resolution, on-line identification of strains from the *Mycobacterium tuberculosis* complex based on tandem repeat typing. BMC microbiology. 2002, 2:37. | PMID: 12456266

33. Dymova MA, Liashenko OO, Poteiko PI, Krutko VS, Khrapov EA, Filipenko ML. Genetic variation of *Mycobacterium tuberculosis* circulating in Kharkiv Oblast, Ukraine. BMC infectious diseases. 2011, 11:77. doi: 10.1186/1471-2334-11-77 PMID: 21439097

34. RNTCP: Revised National TB Control Programme: Training Manual for *Mycobacterium tuberculosis* Culture & Drug susceptibility testing. Central TB Division Directorate General of Health Services, Ministry of Health and Family Welfare, Nirman Bhawan, New Delhi 110011 2009:1–76.

35. Demay C, Liens B, Burguiere T, Hill V, Couvin D, Millet J, et al. SITVITWEB—a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. Infect Genet Evol. 2012, 12(4):755–766. doi: 10.1016/j.meegid.2012.02.004 PMID: 22365971

36. Allix-Beguec C, Harmsen D, Weniger T, Supply P, Niemann S. Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. J Clin Microbiol. 2008, 46(8):2692–2699. doi: 10.1128/JCM.00540-08 PMID: 18550737

37. Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D. MIRU-VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. Nucleic Acids Res. 2010, 38(Web Server issue):W326–331. doi: 10.1093/nar/gkq351gkq351 PMID: 20457747

38. Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J Clin Microbiol. 1988, 26(11):2465–2466. | PMID: 3069867

39. Rajapaksa US, Victor TC, Perera AJ, Warren RM, Senevirathne SM. Molecular diversity of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Sri Lanka. Trans R Soc Trop Med Hyg. 2008, 102(10):997–1002. doi: 10.1016/j.trstmh.2008.04.025 PMID: 18513770

40. Purwar S, Chaudhari S, Katoch VM, Sampath A, Sharma P, Upadhyay P, et al. Determination of drug susceptibility patterns and genotypes of *Mycobacterium tuberculosis* isolates from Kanpur district, North India. Infect Genet Evol. 2011, 11(2):469–475. doi: 10.1016/j.meegid.2010.12.010 PMID: 21237291

41. Hasan Z, Tanveer M, Kanji A, Hasan Q, Ghebremichael S, Hasan R. Spoligotyping of *Mycobacterium tuberculosis* isolates from Pakistan reveals predominance of Central Asian Strain 1 and Beijing

isolates. J Clin Microbiol. 2006, 44(5):1763–1768. doi: 10.1128/JCM.44.5.1763–1768.2006 PMID: 16672404

42. Cerezo I, Jimenez Y, Hernandez J, Zozio T, Murcia MI, Rastogi N. A first insight on the population structure of *Mycobacterium tuberculosis* complex as studied by spoligotyping and MIRU-VNTRs in Bogota, Colombia. Infect Genet Evol. 2012, 12(4):657–663. doi: 10.1016/j.meegid.2011.07.006 PMID: 21784173

43. Diriba B, Berkessa T, Mamo G, Tedla Y, Ameni G. Spoligotyping of multidrug-resistant *Mycobacterium tuberculosis* isolates in Ethiopia. Int J Tuberc Lung Dis. 2013, 17(2):246–250. doi: 10.5588/ijtld.12.0195 PMID: 23317962

44. Kisa O, Tarhan G, Gunal S, Albay A, Durmaz R, Saribas Z, et al. Distribution of spoligotyping defined genotypic lineages among drug-resistant *Mycobacterium tuberculosis* complex clinical isolates in Ankara, Turkey. PloS one. 2012, 7(1):e30331. doi: 10.1371/journal.pone.0030331 PMID: 22279583

45. Roetzer A, Schuback S, Diel R, Gasau F, Ubben T, di Nauta A, et al. Evaluation of *Mycobacterium tuberculosis* typing methods in a 4-year study in Schleswig-Holstein, Northern Germany. J Clin Microbiol. 2011, 49(12):4173–4178. doi: 10.1128/JCM.05293-11 PMID: 21998434

46. Arora J, Singh UB, Suresh N, Rana T, Porwal C, Kaushik A, et al. Characterization of predominant *Mycobacterium tuberculosis* strains from different subpopulations of India. Infect Genet Evol. 2009, 9(5):832–839. doi: 10.1016/j.meegid.2009.05.008 PMID: 19463970

47. Narayanan S, Gagneux S, Hari L, Tsolaki AG, Rajasekhar S, Narayanan PR, et al. Genomic interrogation of ancestral *Mycobacterium tuberculosis* from south India. Infect Genet Evol. 2008, 8(4):474–483. doi: S1567-1348(07)00139-6| PMID: 18024233

48. Rao KR, Ahmed N, Srinivas S, Sechi LA, Hasnain SE. Rapid identification of *Mycobacterium tuberculosis* Beijing genotypes on the basis of the mycobacterial interspersed repetitive unit locus 26 signature. J Clin Microbiol. 2006, 44(1):274–277. doi: 10.1128/JCM.44.1.274–277.2006 PMID: 16390989

49. Bhanu NV, van Soolingen D, van Embden JD, Dar L, Pandey RM, Seth P. Predominace of a novel *Mycobacterium tuberculosis* genotype in the Delhi region of India. Tuberculosis (Edinb). 2002, 82(2–3):105–112. | PMID: 12356462

50. Chang JR, Chen YY, Huang TS, Huang WF, Kuo SC, Tseng FC, et al. Clonal expansion of both modern and ancient genotypes of *Mycobacterium tuberculosis* in southern Taiwan. PloS one. 2012, 7(8):e43018. doi: 10.1371/journal.pone.0043018 PMID: 22937008

51. Buu TN, van Soolingen D, Huyen MN, Lan NT, Quy HT, Tiemersma EW, et al. Increased transmission of *Mycobacterium tuberculosis* Beijing genotype strains associated with resistance to streptomycin: a population-based study. PloS one. 2012, 7(8):e42323. doi: 10.1371/journal.pone.0042323 PMID: 22912700

52. Liu Y, Tian M, Wang X, Wei R, Xing Q, Ma T, et al. Genotypic diversity analysis of *Mycobacterium tuberculosis* strains collected from Beijing in 2009, using spoligotyping and VNTR typing. PloS one. 2014, 9(9):e106787. doi: 10.1371/journal.pone.0106787 PMID: 25237849

53. Almeida D, Rodrigues C, Ashavaid TF, Lalvani A, Udwadia ZF, Mehta A. High incidence of the Beijing genotype among multidrug-resistant isolates of *Mycobacterium tuberculosis* in a tertiary care center in Mumbai, India. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America. 2005, 40(6):881–886. doi: 10.1086/427940 PMID: 15736024

54. de Steenwinkel JE, ten Kate MT, de Knegt GJ, Kremer K, Aarnoutse RE, Boeree MJ, et al. Drug susceptibility of *Mycobacterium tuberculosis* Beijing genotype and association with MDR TB. Emerg Infect Dis. 2012, 18(4):660–663. doi: 10.3201/eid1804.110912 PMID: 22469099

55. Kaplan G, Post FA, Moreira AL, Wainwright H, Kreiswirth BN, Tanverdi M, et al. *Mycobacterium tuberculosis* growth at the cavity surface: a microenvironment with failed immunity. Infection and immunity. 2003, 71(12):7099–7108. | PMID: 14638800

56. Wang C, Peyron P, Mestre O, Kaplan G, van Soolingen D, Gao Q, et al. Innate immune response to *Mycobacterium tuberculosis* Beijing and other genotypes. PloS one. 2010, 5(10):e13594. doi: 10.1371/journal.pone.0013594 PMID: 21049036

57. Banu S, Gordon SV, Palmer S, Islam MR, Ahmed S, Alam KM, et al. Genotypic analysis of *Mycobacterium tuberculosis* in Bangladesh and prevalence of the Beijing strain. J Clin Microbiol. 2004, 42(2):674–682. | PMID: 14766836

58. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, et al. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC microbiology. 2006, 6:23. doi: 1471-2180-6-23 | PMID: 16519816

59. Joseph BV, Soman S, Radhakrishnan I, Hill V, Dhanasooraj D, Kumar RA, et al. Molecular epidemiology of *Mycobacterium tuberculosis* isolates from Kerala, India using IS6110-RFLP, spoligotyping and

MIRU-VNTRs. Infect Genet Evol. 2013, 16:157–164. doi: 10.1016/j.meegid.2013.01.012 PMID: 23395649

60. Refaya AK, Sivakumar S, Sundararaman B, Narayanan S. Polymorphism in the RD1 locus and its effect on downstream genes among South Indian clinical isolates of *Mycobacterium tuberculosis*. J Med Microbiol. 2012, 61(Pt 10):1352–1359. doi: 10.1099/jmm.0.044453–0 PMID: 22723257

61. Mokrousov I, Narvskaya O, Limeschenko E, Vyazovaya A, Otten T, Vyshnevskiy B. Analysis of the allelic diversity of the mycobacterial interspersed repetitive units in *Mycobacterium tuberculosis* strains of the Beijing family: practical implications and evolutionary considerations. J Clin Microbiol. 2004, 42 (6):2438–2444. doi: 10.1128/JCM.42.6.2438–2444.2004 PMID: 15184416

62. Ribeiro SC, Gomes LL, Amaral EP, Andrade MR, Almeida FM, Rezende AL, et al. *Mycobacterium tuberculosis* strains of the modern sublineage of the Beijing family are more likely to display increased virulence than strains of the ancient sublineage. J Clin Microbiol. 2014, 52(7):2615–2624. doi: 10. 1128/JCM.00498-14 PMID: 24829250

63. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, et al. Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. Nature genetics. 2015, 47(3):242–249. doi: 10.1038/ng.3195 PMID: 25599400

64. Vasconcellos SE, Acosta CC, Gomes LL, Conceicao EC, Lima KV, de Araujo MI, et al. Strain classification of *Mycobacterium tuberculosis* isolates in Brazil based on genotypes obtained by spoligotyping, mycobacterial interspersed repetitive unit typing and the presence of large sequence and single nucleotide polymorphism. PloS one. 2014, 9(10):e107747. doi: 10.1371/journal.pone.0107747 PMID: 25314118

65. Nakanishi N, Wada T, Arikawa K, Millet J, Rastogi N, Iwamoto T. Evolutionary robust SNPs reveal the misclassification of *Mycobacterium tuberculosis* Beijing family strains into sublineages. Infect Genet Evol. 2013, 16:174–177. doi: 10.1016/j.meegid.2013.02.007 PMID: 23438651

66. Rindi L, Lari N, Garzelli C. Large Sequence Polymorphisms of the Euro-American lineage of *Mycobacterium tuberculosis*: a phylogenetic reconstruction and evidence for convergent evolution in the DR locus. Infect Genet Evol. 2012, 12(7):1551–1557. doi: 10.1016/j.meegid.2012.06.008 PMID: 22728915

67. Reyes JF, Chan CH, Tanaka MM. Impact of homoplasy on variable numbers of tandem repeats and spoligotypes in *Mycobacterium tuberculosis*. Infect Genet Evol. 2012, 12(4):811–818. doi: 10.1016/j. meegid.2011.05.018 PMID: 21683165