ORIGINAL ARTICLE

Genotypic diversity of *Mycobacterium tuberculosis* isolates from North-Central Indian population

Ravi Prakash^{a,b,c}, Rahul Gupta^a, Pragya Sharma^b, Sanjay Jain^d, Devendra Singh Chauhan^b, Vishwa Mohan Katoch^{e,f} and Pramod Kumar Tiwari^a

^aCentre for Genomics, Molecular and Human Genetics, Jiwaji University, Gwalior, India; ^bDepartment of Microbiology and Molecular Biology, National JALMA Institute for Leprosy and other Mycobacterial Diseases, Agra, India; ^cKey Laboratory of Medical Molecular Virology, Fudan University, Shanghai, China; ^dRevised National Tuberculosis Control Programme, District Hospital, Sheopur, India; ^eFormer Secretary, Department of Health Research, Govt. of India and Former Director General, Indian Council of Medical Research, Ansari Nagar, New Delhi, India; ^fRajasthan University of Health Sciences, Jaipur, India

ABSTRACT

Background: Different strains of *Mycobacterium tuberculosis* (MTB) are known to have different epidemiological and clinical characteristics. Some of them are widely distributed and associated with drug resistance, whereas others are locally predominated. Molecular epidemiological investigations have always been beneficial in identifying new strains and studying their transmission dynamics. Sahariya a primitive tribe of North Madhya Pradesh, India, has already been reported to have high prevalence of tuberculosis (TB) than their non-tribal neighbours. However, the information about MTB genotypes prevalent in Sahariya tribe and their non-tribal neighbours is not available.

Methods: A total of 214 clinical isolates representing Sahariya tribe and non-tribes were analyzed by spoligotyping and MIRU-VNTR typing.

Results: The EAI3_IND/SIT11 genotype was observed as major genotype in Sahariya tribe followed by CAS1_Delhi/SIT26 genotype. A 3.04 fold higher risk of getting TB with EAI3_IND/SIT11 genotype was observed in Sahariya as compared to the non-tribal population. The EAI3_IND/SIT11 genotype also found to have more number of MDR-TB cases in Sahariya as well as true and possible transmission links. In Sahariya tribe, 3 clusters (6 isolates) reflected true transmission links, whereas 8 clusters consisted of 26 isolates revealed possible transmission links within the same geographical location or nearby houses.

Conclusion: The present study highlighted the predominance of EAI3_IND/SIT11 genotype in Sahariya tribe followed by CAS1_Delhi/SIT26 genotype. Combined approach of MIRU-VNTR typing and spoligotyping was observed more favourable in discrimination of MTB genotypes. Further, longitudinal studies using whole genome sequencing can provide more insights into genetic diversity, drug resistance and transmission dynamics of these prevalent genotypes.

KEYWORDS

M. tuberculosis; Sahariya tribe; Spoligotyping; MIRU-VNTR typing

Introduction

Tuberculosis (TB) has remained a health threat for centuries and is among one of the top 10 causes of death worldwide. According to World Health Organization (WHO), at present India ranks highest for TB burden. Despite the implementation of Revised National Tuberculosis Control Programme (RNTCP), the burden of TB in India is still very high (accounts for 27% of global incident TB cases in 2017) [1]. The distribution of specific *Mycobacterium tuberculosis* (MTB) clones/genotypes/multi-drug resistant strains at phylogenetic level, regional level, nationwide and globally can be followed by molecular typing methods [2–10]. Differences in the distribution of major MTB genotypes and their sub-genotypes have been observed with regional clustering. For example, LAM11-ZWE genotype is prevalent in South African countries, the Central Asian (CAS) genotype is

predominant in East African countries, the East African Indian genotype has dominance in Asia reflecting possible historic sea trade links and the Cameroon genotype is prevalent in West African countries [2]. Similarly, the Beijing genotype was reported as having significant association with multi-drug resistant tuberculosis (MDR-TB) cases such as in Chongqing, China and Spain [11,12]. Molecular typing has always facilitated in population based monitoring of TB control programmes and understanding the TB epidemics [13]. The most commonly used tools for molecular epidemiological investigations are IS6110-RFLP typing, spoligotyping and MIRU-VNTR typing [14–28]. These markers have been demonstrated successfully in studying micro-evolutionary changes and mixed strain infections in MTB strains [29,30].

Sahariya tribe is an economically deprived, landless community which practice cultivation and labor in

CONTACT Pramod Kumar Tiwari ✉ pk_tiwari@hotmail.com Centre for Genomics, Molecular and Human Genetics, Jiwaji University, Gwalior, India; Devendra Singh Chauhan ✉ devchauhan01@yahoo.co.in Department of Microbiology and Molecular Biology, National JALMA Institute for Leprosy and other Mycobacterial Diseases, Agra, India

Supplemental data for this article can be accessed [here](#).

© 2019 Informa UK Limited, trading as Taylor & Francis Group

lands of their non-tribal neighbours [31]. They face various problems such as poverty, malnutrition, illiteracy, absence of safe drinking water, poor maternal and child health services which contribute to major problems and poor health conditions. Various studies identified several factors such as demography, mycobacterium infection, genetic disorders, mortality rate, life expectancy, nutritional status and health care practices in Sahariya. Those studies concluded that these factors might be the major determinants of the health status of this tribe, mainly inhabiting in North Madhya Pradesh, India [32–36]. One of the studies on Sahariya tribe and their non-tribal neighbours revealed overall high crude prevalence for pulmonary tuberculosis (PTB) among Hindu tribes (29.9%) than in Hindu castes (21.4%) and Muslim groups (15.0%) of the same region [34].

So far, studies carried out on Sahariya tribe and non-tribes emphasized only on host genetic factors [37–40]. But, no investigation was carried out on the prevalence of mycobacterium strains in Sahariya tribe and non-tribes. One of the recent articles on TB research in tribal areas also indicated that there is a lacuna of data on MTB genotypes prevalent in Indian tribal populations [41]. Sahariya tribe is a migratory population living in forest areas but, they do share environments with their non-tribal neighbours while communicating, working in their fields and in other labor class work. Therefore, the present study was aimed to identify the predominant MTB genotypes responsible for TB and their cluster rate in Sahariya tribe and their non-tribal neighbours using spoligotyping and MIRU-VNTR typing methods.

Materials and methods

Study population and patient selection

Health camps were organized at different intervals over a period from September 2010 to February 2014 in Sheopur and Gwalior districts of Madhya Pradesh, India. Individuals representing Sahariya tribe and non-tribes were screened for TB symptoms. A clinician from the district hospital accompanied the sampling team for detailed diagnosis of TB patients. Individuals were observed positive for TB on the basis of Ziehl-Neelsen (ZN) staining and other symptoms were enrolled in the study. A well informed and written consent was signed by all the participants. The methods and protocols employed in the study had been approved by Institutional Ethics Committee, Jiwaji University, Gwalior. The effective sample size for combined analysis of Sahariya tribe and non-tribes was calculated on the basis of previous studies carried out on prevalence of tuberculosis [34,42], through rollin's Power or Sample Size Calculator (www.stat.ubc.ca/~rollin/stats/ssize/caco.html). The power of the study was calculated

to be >90% for 100 culture positive samples from Sahariya tribe and non-tribes.

Sputum sampling and MTB culture

Sputum samples were collected, decontaminated and cultured on Lowenstein-Jensen (LJ) media by modified Petroff's method as per RNTCP protocol (www.rntcp.org.in). A total number of sputum as well as culture positive samples from Sahariya tribe (N = 111) and non-tribes (N = 103) were included in the study. In Sahariya tribe, out of 111 isolates 80 isolates were new smear positive/Category I cases (individuals diagnosed first time for TB), whereas 31 isolates belonged to Category II [Retreated/Treatment After Default(TAD)] cases. Similarly in non-tribes, out of 103 isolates 80 isolates belonged to Cat I and 23 isolates to Cat II cases. Mycobacterium growth was observed up to 8 weeks and MTB was confirmed using biochemical assays [43].

DNA isolation and spoligotyping

The genomic DNA of MTB isolates was isolated as per standard protocol [44]. Spoligotyping was conducted, results obtained were coded in binary format and converted into octal code as described earlier [45].

MIRU-VNTR typing

MIRU-VNTR typing was performed to discriminate MTB isolates using 12 loci pattern [46]. The corresponding repeat number for each locus of all the isolates was determined by using allele calling table. The allele assignment was confirmed by comparing with MTB H37Rv as a laboratory control [47].

Data analysis

The spoligotyping and MIRU-VNTRs data was analysed using online tool MIRU-VNTRplus (www.miru-vntrplus.org) [48,49]. The major genotypes observed by MIRU-VNTRplus were also compared by using online SITVITWEB database (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) [50]. Strains which matched the pattern in the database were assigned a Spoligotype International Type (SIT) and those did not match were considered as 'orphan'. Spoligotyping was used to describe MTB families predominant in the area, whereas the data after combined analysis of spoligotyping and MIRU-VNTR typing was used to describe clustering and transmission status of strains. Strains with identical SIT profiles and identical MIRU-VNTRs were classified as 'clusters', whereas others were termed as 'unique'. A true transmission link refers to the cluster of two or more isolates consisting of a relative, friend or colleague and possible transmission link was considered

within the same geographical location. Chi-square analysis was conducted to identify any association between the major MTB genotypes as well as clustered and non-clustered isolates of both populations [Graph Pad prism software (www.graphpad.com)]. The Hunter-Gaston Diversity Index (HGDI) was calculated as per the given formula [51].

Results

Spoligotyping

When analyzed through MIRU-VNTRplus database 106/111 isolates from Sahariya tribe were classified into 22 SITs while the remaining 5 could not be assigned to any genotype and were named as orphan/unknown. The major genotypes observed were EAI3_IND/SIT11 (N = 51, 46%) followed by CAS1_Delhi/SIT26 (N = 21, 19%), whereas others were fewer in number (Table S1). Not a single Beijing genotype could be observed in Sahariya tribe. Similarly, in non-tribes, 96/103 isolates were observed in 21 different SITs and 7 remained orphan. In non-tribes, the major genotypes observed were CAS1_Delhi/SIT26 (N = 30, 29%) and EAI3_IND/SIT11 (N = 29, 28%) while others were fewer in number. Only two Beijing and one Beijing like genotypes were found in non-tribes (Table S1 and Figure 1).

When all of these isolates were analyzed through SITVITWEB database 92/111 isolates were separated into 25 SITs, whereas 19 remained orphan in Sahariya tribe. Out of these 25 SITs, the EAI3_IND/SIT11 and CAS1_Delhi/SIT26 genotypes were again observed as major genotypes while other genotypes such as EAI5/SIT126, CAS1_Delhi/SIT25, EAI3_IND/SIT2656, CAS1_Delhi/SIT428, SIT510, CAS1_Delhi/SIT1199, X2/SIT1342, CAS1_Delhi/SIT1343, EAI5/SIT138, CAS1_Delhi/SIT1405, EAI5/SIT1427, EAI3_IND/SIT1680, EAI5/SIT1970, EAI5/SIT236, EAI3_IND/SIT2457, EAI3_IND/SIT2726, CAS2/SIT288, CAS1_Delhi/SIT289, CAS1_Delhi/SIT381, EAI1_SOM/SIT48, EAI5/SIT625, EAI3_IND/SIT654, EAI5/SIT934 were fewer in number. Similarly, in non-tribes 19 SITs comprised of 71 isolates were observed, whereas 32 isolates remained orphan. The major genotypes observed were CAS1_Delhi/SIT26 and EAI3_IND/SIT11 genotypes and others such as CAS1_Delhi/SIT25, CAS2/SIT288, H3/SIT625, CAS/SIT1789, CAS1_Delhi/SIT22, EAI5/SIT236, EAI3_IND/SIT2656, CAS1_Delhi/SIT1092, EAI5/SIT26, CAS1_Delhi/SIT1314, CAS1_Delhi/SIT1343, EAI5/SIT138, CAS1_Delhi/SIT142, CAS1_Delhi/SIT1965, CAS1_Delhi/SIT357, CAS1_Delhi/SIT428, CAS1_Delhi/SIT429, SIT1/Beijing and SIT269/Beijing like genotypes were fewer in number.

MIRU-VNTR typing

The 12-loci MIRU-VNTR typing was performed on all the isolates from Sahariya tribe and non-tribes. In Sahariya tribe, the HGDI for 12 loci MIRU-VNTRs revealed locus 10 (0.73) as most polymorphic followed by loci 4 (0.71), 31 (0.69), 23 (0.67), 16 (0.66) and 40 (0.66). In non-tribes also, the locus 10 (0.74) was found most polymorphic followed by loci 26 (0.72), 16 (0.70) and 31 (0.65) (Table 1). Among isolates from Sahariya tribe, allele 2 was more commonly observed in MIRU loci 2, 20, 24 & 26, allele 3 in MIRU loci 16, 27 & 39, allele 4 in MIRU locus 10, allele 5 in MIRU loci 10, 4, 23 & 31 and allele 6 in MIRU locus 23 only. However, in case of non-tribes, allele 2 was found more commonly in MIRU loci 2, 4 & 20, allele 1 in MIRU locus 24, allele 3 in MIRU loci 16, 27, 39 & 40, allele 4 in MIRU loci 10 & 16 and allele 5 in MIRU loci 23 & 31.

Spoligotyping and MIRU-VNTR typing

Through spoligotyping alone, in Sahariya tribe 17 clusters consisted of 96 isolates were observed. A few of them were in range of 2–42 isolates per cluster (HGDI = 0.792). In case of non-tribes, 84 isolates were found in 23 clusters and few of them were in range of 2–25 isolates per cluster (HGDI = 0.826). When analyzed through spoligotyping and MIRU-VNTR typing methods combined, two large clusters consisted of 18 & 42 isolates in Sahariya tribe as well as 17 & 25 isolates in non-tribe were further split apart into smaller ones. The HGDI was also found to be increased 0.988 and 0.992 for Sahariya tribe and non-tribe, respectively. This signified that both the methods collectively divided large clusters into smaller ones and helped in discriminating MTB genotypes (Table 2).

MTB genotypes vs cluster rate

The chi-square analysis for the distribution of the major genotypes EAI3_IND/SIT11 and CAS1_Delhi/SIT26 between Sahariya tribe and non-tribes revealed significant difference (OR = 3.045, 95% CI = 1.392–6.663, P = 0.0084) (Table 3). Furthermore, the major genotypes EAI3_IND/SIT11 and CAS1_Delhi/SIT26 were compared for cluster formation in two population groups. The EAI3_IND/SIT11 as well as CAS1_Delhi/SIT26 both genotypes were observed with an inclination to develop maximum number of clusters in Sahariya (80%, N = 36 for EAI3_IND/SIT11 and 63%, N = 12 for CAS1_Delhi/SIT26) as compared to non-tribes (71%, N = 15 for EAI3_IND/SIT11 and 59%, N = 16 for CAS1_Delhi/SIT26). However, this could not

Figure 1. Dendrogram generated using UPGMA algorithm for 12 loci MIRU-VNTR typing and spoligotyping in Sahariya tribe and non-tribes through miru-vntrplus.org.

Table 1. The Hunter Gaston Diversity Index for 12 MIRU-VNTR loci in MTB isolates from Sahariya tribe and non-tribes.

S. No.	Alias	Locus	Sahariya tribe	Non-tribe
1	MIRU 02	154	−0.01	0.01
2	MIRU 04	580	0.71	0.57
3	MIRU 40	802	0.66	0.61
4	MIRU 10	960	0.73	0.74
5	MIRU 16	1644	0.66	0.7
6	MIRU 20	2059	0.04	0.1
7	MIRU 23	2531	0.67	0.6
8	MIRU 24	2687	0.48	0.5
9	MIRU 26	2996	0.38	0.72
10	MIRU 27	3007	0.03	0.16
11	MIRU 31	3192	0.69	0.65
12	MIRU 39	4348	0.58	0.46

Table 2. Distribution of clustered and non-clustered isolates in Sahariya tribe and non-tribe along with their HGDI for spoligotyping alone and combined approach of both the methods.

Population	Number of isolates (Number of clusters)	Non-clustered /unique isolates N	HGDI
Spoligotyping			
Sahariya tribe	96 (17)	15	0.792
Non-tribe	84 (23)	19	0.826
Spoligotyping and MIRU-VNTR typing			
Sahariya tribe	80 (34)	31	0.988
Non-tribe	73 (36)	30	0.992

Table 3. Chi-square analysis for the distribution of major genotypes between Sahariya tribe and non-tribe by spoligotyping and MIRU-VNTR typing methods combined.

Tribe/Non-Tribe N	EAI3_IND/ SIT11 N (%)	CAS1_Delhi/ SIT26 N (%)	OR#	95% CI	P-value
Sahariya tribe (N = 72)	45 (41)	19 (17)	3.045	1.392–6.663	0.0084*
Non-tribe (N = 59)	21 (20)	27 (26)			

* $P < 0.05$ considered statistically significant, # Odd Ratio, 95% CI – Confidence Interval

reveal any significant difference (Table 4). Further stratification of clustered isolates from EAI3_IND/SIT11 and CAS1_Delhi/SIT26 genotypes into Cat I, Cat II cases as well as on the basis of gender in Sahariya tribe and non-tribes too did not reveal any significant difference.

MTB genotypes vs MDR-TB clusters

The major MTB genotypes were compared for cluster formation in MDR-TB isolates from our own earlier

published study [52]. In Sahariya tribe, 4 MDR-TB isolates were observed in different clusters, whereas in non-tribes 3 MDR-TB isolates were found in different clusters and others were unique. The EAI3_IND/SIT11 genotype was observed predominating in MDR-TB cases from Sahariya tribe followed by CAS1_Delhi/SIT26 while in non-tribes, none of the genotype could observe predominant in MDR-TB cases.

Relationship between geographical location and family or relative contact

Although, the present study's purpose was not a contact investigation for TB patients, however likely possibilities of transmission links/events were probed after identification of clustered isolates. In Sahariya tribe, 3 clusters (comprised of 6 isolates) reflected true transmission links, whereas 8 clusters consisted of 26 isolates revealed possible transmission links within the same geographical location or nearby houses. In non-tribes, only one true transmission link could be observed and only one cluster consisting of 2 isolates revealed possible transmission links. This indicated that the probability/frequency of transmission of MTB genotypes in Sahariya tribe is greater than their non-tribal neighbours.

Discussion

Only a few studies on TB research from different tribal populations of India are available [53–57]. Of these, Sahariya tribe has been most widely studied in context of TB prevalence, genetic susceptibility to TB and associated risk factors [41]. But, the

Table 4. Chi-square analysis for the distribution of clustered and non-clustered isolates of EAI3_IND/SIT11 and CAS1_Delhi/SIT26 genotypes in Sahariya tribe and non-tribe by spoligotyping and MIRU-VNTR typing methods combined.

MTB Genotypes	Clutered Isolates N (%)	Non-clustered/Unique isolates N (%)	OR#	95% CI	P-value
Sahariya tribe					
EAI3_IND/SIT11 N (%)	36 (80%)	9 (20%)	2.333	0.7137–7.628	0.2688
CAS1_Delhi/SIT26 N (%)	12 (63%)	7 (37%)			
Non-tribe					
EAI3_IND/SIT11 N (%)	15 (71%)	6 (29%)	1.719	0.5078–5.817	0.5684
CAS1_Delhi/SIT26 N (%)	16 (59%)	11 (41%)			

* $P < 0.05$ considered statistically significant, # Odd Ratio, 95% CI – Confidence Interval

information on MTB genotypes in Sahariya tribe and their neighbouring non-tribes is still not available. To our knowledge, this is the first molecular epidemiological investigation on MTB genotypes among Sahariya tribe and their non-tribal neighbours of North-Central India, using spoligotyping and MIRU-VNTR typing methods. Indian population is diverse in its ethnicity and geography which reflects differential predominance of MTB genotypes in different regions of the country. The CAS1_Delhi genotype is predominant in North India [7,22,23,25,58], EAI3_IND in South Indian states [7,19,22,25,59,60], whereas North East Indian regions such as Assam and Kolkata are mainly dominated by Beijing genotype [7,17]

The present study revealed EAI3_IND/SIT11 as one of the major genotypes in Sahariya tribe followed by CAS1_Delhi/SIT26 while in non-tribes, EAI3_IND/SIT11 and CAS1_Delhi/SIT26 both genotypes were observed to a similar degree. A 3.04 fold higher risk of infection with EAI3_IND/SIT11 genotype was observed in Sahariya tribe as compared to non-tribes. This indicated that EAI3_IND/SIT11 genotype is mainly predominating and circulating/transmitting in Sahariya tribe. However, other studies have reported that this genotype is mainly distributed in South India [19,59,60]. A likely possibility for the predominance of EAI3_IND/SIT11 genotype in Sahariya tribe is that it is a primitive tribal group. The tribe might have carried EAI3_IND/SIT11 genotype along with them during their migration in ancient times which signifies the possibility of host pathogen co-evolution in them and thus, there is a predominance of EAI3_IND/SIT11 genotype. It is also suggested that MTB diversity can be affected by geography, demography as well as human migrations [58].

Despite inhabiting the same geographical location, a significant difference was observed for the distribution of EAI3_IND/SIT11 and CAS1_Delhi/SIT26 genotypes in between Sahariya tribe and non-tribe which reflects ethnic differences in the prevalence of MTB genotypes. Similarly, significant differences between

CAS and Manu1 genotypes were also observed in a study from Mumbai, South West of India. Isolates with CAS genotypes were found originating from North India (migrated for work) but, were living in Mumbai while isolates with Manu1 genotype were observed in residents of Mumbai (already living in Mumbai). For MTB isolates from permanent residents of Mumbai, Manu1 and Manu2 genotypes were observed more in number when compared to CAS1_Delhi and EAI genotypes [61]. In another study, Manu1 genotype was also found to be the major MTB genotype predominating in Satna, Madhya Pradesh, India followed by Manu2 genotype [62]. All of this information indicates varied geographical distribution of MTB genotypes in the Indian population. The distribution of EAI3_IND/SIT11 and CAS1_Delhi/SIT26 genotypes in various regions/parts of India is shown in Table 5.

We observed MIRU locus 10 as the most polymorphic locus in both population groups but, the MIRU locus 4, which was the second highest polymorphic locus in Sahariya tribe (0.71), was found less discriminatory in non-tribes (0.58). Similar to our observations, MIRU locus 10 was observed highly polymorphic in a study on MDR-TB isolates from Chongqing, China [12]. The MIRU locus 31 appeared third and fourth most polymorphic in our population groups, whereas in a study from Delhi, India, MIRU locus 31 was observed to have maximum discriminatory power [21]. In another study from Delhi, India, locus QUB 26 was observed more discriminatory followed by MIRU 26, MIRU 10, Mtub 04, MIRU 31 and Mtub 30 loci [15]. The HGDI for spoligotyping in Sahariya tribe and non-tribe observed 0.792 and 0.826, respectively. But, when both the methods (spoligotyping and MIRU-VNTR typing) were combined the resolution power for discriminating MTB genotypes was increased (0.988 for Sahariya tribe & 0.992 for non-tribe). One of the studies suggested that the use of 6 hyper variable MIRU-VNTR loci is more successful in discriminating MTB genotypes in developing countries like India [23]. Allix-Beguec and

Table 5. Distribution of EAI3_IND/SIT11 and CAS1_Delhi/SIT26 genotypes in various parts of India.

S. No.	Tribe/Non-tribe	EAI3_IND/SIT11	CAS1_Delhi/SIT26	Region/State	Reference
1.	Sahariya tribe	46%	19%	Gwalior and Sheopur, M. P.	Present study
2.	Non-tribe	28%	29%		
3.	Non-tribe	22%	53%	Delhi	[25]
		15%	22%	Lucknow, U. P.	
		5%	19%	Pune, Maharashtra	
		16%	6%	Chennai, Tamilnadu	
		43%	–	Trivandrum, Kerala	
4.	Non-tribe	11%	37%	Kanpur, U. P.	[23]
5.	Non-tribe	41%	4%	Thiruvallur, Chennai	[19]
6.	Non-tribe	28.6%	4.16%	Kerala	[60]
7.	Non-tribe	–	11%	Assam	[17]
8.	Non-tribe	6.32%	32.37%	Delhi	[15]
9.	Non-tribe	25%	16%	Vellore, Tamilnadu	[19,60]
10.	Non-tribe	19%	59.1%	Kanpur, U. P.	[58]
11.	Non-tribe	8%	12%	Udupi, Karnatka	[63]

colleagues also proposed the addition of four hyper-variable loci (1982, 3820, 3232 and 4120) into 24-VNTR spectrum to increase resolution power in the regions mainly dominated by Beijing genotypes [64]. However, these set of combinations may vary in different geographical settings. Therefore, it is recommended that the combined approach of spoligotyping and MIRU-VNTR typing methods in molecular epidemiological investigations of TB may provide better insights on the clustering and transmission rate of MTB genotypes in a population or region [58]. The present study could not reveal any significant differences between clustered isolates from EAI3_IND/SIT11 and CAS1_Delhi/SIT26 genotypes into Cat I and Cat II cases (data not shown). This is also similar to the findings from two earlier published studies for clustering percentage between new cases and treatment failures [61,65].

In Sahariya tribe, the maximum number of MDR-TB isolates belonged to EAI3_IND/SIT11 genotype. Albanna and co-workers reported that EAI genotypes form less clustered TB cases, whereas in our population group EAI3_IND/SIT11 genotype developed large number of clusters when compared to CAS1_Delhi/SIT26 genotype [66]. This indicates that the EAI3_IND/SIT11 genotype is mainly circulating in this region/population and may also be associated with drug resistance but, further investigation is required in close contacts and families. Sahariya tribe has already been found to have 1.95 fold higher risks to develop drug resistance than non-tribes [52]. EAI genotype has also been observed prevalent in 3 tropical Asian countries such as Cambodia (60%), Singapore (25.62%) and Taiwan (21.85%) with high rate of MDR-TB cases [5].

In the present study, 4 true epidemiological/transmission links as well as 8 possible transmission links were observed in Sahariya tribe and non-tribes. A total of 3–4 clusters of MDR-TB isolates were also found associated to geographical locations and familial history. We observed that EAI3_IND/SIT11 genotype is mainly responsible for the transmission in familial and geographical settings in Sahariya tribe which suggests plausible transmission of a particular strain in a geographical location in the near future. However, in non-tribes not a single genotype could be observed as responsible for transmission. The likely reason for differences in true and possible transmission links between Sahariya tribe and non-tribe may be the population density. The population density of Sahariya tribe found to be 415 persons per square km, whereas for non-tribes, it is 103 persons per square km. People from Sahariya tribal community live in small, poor ventilated houses, in close clusters/proximity to each other and mostly in remote areas far away from the general population. They are also affected by poverty, low literacy rate, malnutrition, unsafe drinking water, poor

maternal and child health services and lack of sanitary conditions. They too keep sharing cigarette butts in the same environment as one another. A 4.46 fold higher risk for developing MDR-TB among smokers of Sahariya tribe has already been reported [52]. However, non-tribal people live in a healthier and more education environment than Sahariya tribe. They use agricultural practices, farming, maintain hygienic practices, are educated and well aware about associated health problems.

The strength of our study is that it involves two distinct populations; one, which is genetically highly homogenous (Sahariya tribe) and the other, which exhibits varying ethnicity (non-tribes). A few limitations are that Sahariya tribal people are always migratory in search of work which poses a hurdle to keep a strict vigil on their clinical records and history of taking anti-tuberculosis treatment (ATT). Second, we used 12 loci MIRU-VNTR typing to differentiate MTB genotypes which could have been more helpful by using 15/24 loci MIRU-VNTRs.

Conclusion

The present investigation revealed predominance of EAI3_IND/SIT11 genotype in Sahariya tribe followed by CAS1_Delhi/SIT26 genotype. The EAI_IND/SIT11 genotype was observed to have a greater number of MDR-TB cases as well as increased transmission links associated with familial and geographical settings in Sahariya tribe, which is a major concern. The use of MIRU-VNTRs along with spoligotyping was found to be more beneficial in discriminating MTB genotypes. Therefore, the establishment of suitable local VNTR combination as per the characteristics of the local circulating clinical strains might be needed in this population or geographical location. This will increase the resolution power as well as helping to identifying/correlating the transmission events. The use of whole genome sequencing and single colony cultures of strains in such populations or geographical settings will also play an important role in identifying the contact cases as well as transmission events with more accuracy in the near future. Hence, the information generated will be beneficial for future investigations and in evaluating the performance of TB control programmes in this region.

Competing Interests

All authors have declared that they have no competing interests.

Consent for Publication

Not Applicable.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The financial support for the study was provided by Indian Council of Medical Research (ICMR), New Delhi, India, through a research grant [R/P No. Tribal/ 37/ 2008-ECD-II].

Ethical Approval and consent to participate

The study and protocols were approved by Institutional Ethics Committee, Jiwaji University, Gwalior.

References

- [1] World Health Organization. Global tuberculosis report 2018. Switzerland: World Health Organization Geneva; 2018.
- [2] Chihota VN, Niehaus A, Streicher EM, et al. Geospatial distribution of *Mycobacterium tuberculosis* genotypes in Africa. *PLoS One*. 2018;13(8):e0200632.
- [3] Monteserin J, Paul R, Gravina E, et al. Genotypic diversity of *Mycobacterium tuberculosis* in Buenos Aires, Argentina. *Inf Gen Evol*. 2018;62:1–7.
- [4] Lopez-Avalos G, Gonzalez-Palomar G, Lopez-Rodriguez M, et al. Genetic diversity of *Mycobacterium tuberculosis* and transmission associated with first-line drug resistance: a first analysis in Jalisco, Mexico. *J Global Antimicrob Res*. 2017;11:90–97.
- [5] Chen YY, Chang JR, Huang WF, et al. Genetic diversity of the *Mycobacterium tuberculosis* East African–Indian family in three tropical Asian countries. *J Microbiol Immunol Infect*. 2017;50(6):886–892.
- [6] Ejo M, Gehre F, Barry MD, et al. First insights into circulating *Mycobacterium tuberculosis* complex lineages and drug resistance in Guinea. *Inf Gen Evol*. 2015;33:314–319.
- [7] Singh J, Sankar MM, Kumar P, et al. Genetic diversity and drug susceptibility profile of *Mycobacterium tuberculosis* isolated from different regions of India. *J Inf*. 2015;71(2):207–2019.
- [8] Gibson AL, Huard RC, van Pittius NCG, et al. Application of sensitive and specific molecular methods to uncover global dissemination of the major RD Rio Sublineage of the Latin American-Mediterranean *Mycobacterium tuberculosis* spoligotype family. *J Clin Microbiol*. 2008;46(4):1259–1267.
- [9] Filliol I, Motiwala AS, Cavatore M, et al. 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*. 2006;188(2):759–772.
- [10] Mathema B, Kurepina NE, Bifani PJ, et al. Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev*. 2006;19(4):658–685.
- [11] Samper S, Gavín P, Millán-Lou MI, et al. Spanish working group on MDR-TB. *Mycobacterium tuberculosis* genotypes and predominant clones among the multidrug-resistant isolates in Spain 1998–2005. *Inf Gen Evol*. 2017;55:117–126.
- [12] Zhang D, An J, Wang Y, et al. Genetic diversity of multidrug-resistant tuberculosis in a resource-limited region of China. *Int J Infect Dis*. 2014;29:7–11.
- [13] Banu S, Rahman MT, Uddin MKM, 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.
- [14] Medeiros TF, Nogueira CL, Prim RI, et al. Molecular epidemiology of *Mycobacterium tuberculosis* strains from prison populations in Santa Catarina, Southern Brazil. *Inf Gen Evol*. 2018;58:34–39.
- [15] Varma-Basil M, Narang A, Chakravorty S, et al. A snapshot of the predominant single nucleotide polymorphism cluster groups of *Mycobacterium tuberculosis* clinical isolates in Delhi, India. *Tuberculosis*. 2016;100:72–81.
- [16] Pichat C, Couvin D, Carret G, et al. Combined genotypic, phylogenetic, and epidemiologic analyses of *Mycobacterium tuberculosis* genetic diversity in the Rhône Alpes region, France. *PLoS One*. 2016;11(4):e0153580.
- [17] Devi KR, Bhutia R, Bhowmick S, et al. 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. *PLoS One*. 2015;10(12):e0145860.
- [18] Gupta A, Kulkarni S, Rastogi N, et al. A study of *Mycobacterium tuberculosis* genotypic diversity & drug resistance mutations in Varanasi, North India. *Ind J Med Res*. 2014;139(6):892.
- [19] Joseph BV, Soman S, Radhakrishnan I, et al. Molecular epidemiology of *Mycobacterium tuberculosis* isolates from Kerala, India using IS6110-RFLP, spoligotyping and MIRU-VNTRs. *Inf Gen Evol*. 2013;16:157–164.
- [20] Sougakoff W. Molecular epidemiology of multidrug-resistant strains of *Mycobacterium tuberculosis*. *Clin Microbiol Infect*. 2011;17(6):800–805.
- [21] Varma-Basil M, Kumar S, Arora J, 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. *Memorias Do Instituto Oswaldo Cruz*. 2011;106(5):524–535.
- [22] Arora J, Singh UB, Suresh N, et al. Characterization of predominant *Mycobacterium tuberculosis* strains from different subpopulations of India. *Inf Gen Evol*. 2009;9(5):832–839.
- [23] Sharma P, Chauhan DS, Upadhyay P, et al. Molecular typing of *Mycobacterium tuberculosis* isolates from a rural area of Kanpur by spoligotyping and mycobacterial interspersed repetitive units (MIRUs) typing. *Inf Gen Evol*. 2008;8(5):621–626.
- [24] Chauhan DS, Sharma VD, Parashar D, et al. Molecular typing of *Mycobacterium tuberculosis* isolates from different parts of India based on IS6110 element polymorphism using RFLP analysis. *Ind J Med Res*. 2007;125(4):577–581.
- [25] Singh UB, Arora J, Suresh N, et al. Genetic biodiversity of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in India. *Inf Gen Evol*. 2007;7:441–448.
- [26] Chauhan A, Chauhan DS, Parashar D, et al. DNA fingerprinting of *Mycobacterium tuberculosis* isolates from Agra region by is 6110 probe. *Ind J Med Microbiol*. 2004;22(4):238–240.

- [27] Chander A, Yadava R, Ghosh S, et al. Strain differentiation among Indian clinical isolates of *M. tuberculosis* by RFLP analysis. *Mol Biol Today*. 2000;1(2):53–55.
- [28] Das S, Paramasivan CN, Lowrie DB, et al. IS6110 restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in Madras, south India. *Tubercle Lung Dis*. 1995;76(6):550–554.
- [29] Jajou R, Kamst M, van Hunen R, et al. The occurrence and nature of double alleles in VNTR patterns of more than 8,000 *Mycobacterium tuberculosis* complex isolates in The Netherlands. *J Clin Microbiol*. 2018;56(2):e00761–17.
- [30] Streit E, Millet J, Rastogi N. *Mycobacterium tuberculosis* polyclonal infections and microevolution identified by MIRU-VNTRs in an epidemiological study. *Int J Mycobacteriol*. 2015;4(3):222–227.
- [31] Mohanty PK. *Encyclopaedia of primitive tribes in India*. Vol. 2. Kalpaz Publications, Delhi, India; 2003.
- [32] Bhat J, Rao VG, Yadav R, et al. Situation of drug resistant tuberculosis in Saharia tribe of central India. *Indian J Med Res*. 2015;141:636–639.
- [33] Ghosh-Jerath S, Singh A, Bhattacharya A, et al. Dimensions of nutritional vulnerability: assessment of women and children in Sahariya tribal community of Madhya Pradesh in India. *Ind J Pub Health*. 2013;57(4):260–267.
- [34] Raj P, Prakash R, Mishra G, et al. Prevalence of smear-positive pulmonary tuberculosis in different ethnic groups in India: evaluation of public health. *Pub Health*. 2012;126(4):295–299.
- [35] Sharma PR, Jain S, Bamezai RNK, et al. Utility of serum LDH and its isoforms in the assessment of *Mycobacterium tuberculosis* induced pathology in TB patients of Sahariya tribe. *Ind J Clin Biochem*. 2010;25(1):57–63.
- [36] Rao KM, Kumar RH, Venkaiah K, et al. Nutritional status of Saharia - A primitive tribe of Rajasthan. *J Hum Ecol*. 2006;19(2):117–123.
- [37] Mishra G, Poojary SS, Jain S, et al. Genotype-phenotype relationship of CCL5 in pulmonary tuberculosis infection in Sahariya tribe: a pilot study. *Ind J Med Res*. 2017;146(6):768.
- [38] Mishra G, Kumar N, Kaur G, et al. Distribution of HLA-A, B and DRB1 alleles in Sahariya tribe of North Central India: an association with pulmonary tuberculosis. *Inf Gen Evol*. 2014;22:175–182.
- [39] Mishra G, Poojary SS, Raj P, et al. Genetic polymorphisms of CCL2, CCL5, CCR2 and CCR5 genes in Sahariya tribe of North Central India: an association study with pulmonary tuberculosis. *Inf Gen Evol*. 2012;12(5):1120–1127.
- [40] Sharma PR, Singh S, Jena M, et al. Coding and non-coding polymorphisms in VDR gene and susceptibility to pulmonary tuberculosis in tribes, castes and Muslims of Central India. *Inf Gen Evol*. 2011;11(6):1456–1461.
- [41] Rao VG, Muniyandi M, Bhat J, et al. Research on tuberculosis in tribal areas in India: a systematic review. *Ind J Tubercul*. 2018;65:8–14.
- [42] Rao VG, Gopi PG, Bhat J, et al. Pulmonary tuberculosis: a public health problem amongst the Saharia, a primitive tribe of Madhya Pradesh, Central India. *Int J Inf Dis*. 2010;14(8):e713–e716.
- [43] Vestal AL. *Procedures for isolation and identification of Mycobacteria*. Atlanta: US Dep Health, Ed Welfare Publication; 1977. (NO. CDC 77-8230).
- [44] van Soolingen D, Hermans PW, de Haas PE, et al. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol*. 1991;11:2578–2586.
- [45] Kamerbeek J, Schouls L, Kolk A, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol*. 1997;35(4):907–914.
- [46] Supply P, Lesjean S, Savine E, et al. 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.
- [47] Supply P. Multilocus variable number tandem repeat genotyping of *Mycobacterium tuberculosis*. *Tech Guide*. 2005;6:1–74.
- [48] Weniger T, Krawczyk J, Supply P, et al. MIRU-VNTR plus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. *Nucleic Acid Res*. 2010;38(2):W326–W331.
- [49] Allix-Beguec C, Harmsen D, Weniger T, et al. 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.
- [50] Demay C, Liens B, Burguière T, et al. SITVITWEB—a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. *Inf Gen Evol*. 2012;12(4):755–766.
- [51] 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.
- [52] Prakash R, Kumar D, Gupta VK, et al. Status of multi-drug resistant tuberculosis (MDR-TB) among the Sahariya tribe of North Central India. *J Inf Pub Health*. 2016;9(3):289–297.
- [53] Thomas BE, Adinarayanan S, Manogaran C, et al. Pulmonary tuberculosis among tribals in India: a systematic review & meta-analysis. *Ind J Med Res*. 2015;141(5):614.
- [54] Yadav R, Rao VG, Bhat J, et al. Prevalence of pulmonary tuberculosis amongst the Saharia A primitive tribe of Madhya Pradesh, Central India. *Ind J Tub*. 2010;57(2):114–116.
- [55] Bhat J, Rao V, Gopi PG, et al. Prevalence of pulmonary tuberculosis amongst the tribal population of Madhya Pradesh, Central India. *Int J Epidemiol*. 2009;38:1026–1032.
- [56] Narang P, Tyagi NK, Mendiratta DK, et al. Prevalence of sputum-positive pulmonary tuberculosis in tribal and non-tribal populations of the Ashta and Karanja tahsils in Wardha district, Maharashtra State, India. *Int J Tub Lung Dis*. 1999;3(6):478–482.
- [57] Chakma T, Rao PV, Pill S, et al. Survey of Pulmonary Tuberculosis in a Primitive Tribe of Madhya Pradesh. *Ind J Tub*. 1996;43:85–90.
- [58] Sharma P, Katoch K, Chandra S, et al. Comparative study of genotypes of *Mycobacterium tuberculosis* from a Northern Indian setting with strains reported from other parts of India and neighbouring countries. *Tuberculosis*. 2017;105:60–72.

- [59] Suzana S, Shanmugam S, Devi KRU, et al. Spoligotyping of *Mycobacterium tuberculosis* isolates at a tertiary care hospital in India. *Trop Med Int Health*. 2017;22(6):703–707.
- [60] Narayanan S, Gagneux S, Hari L, et al. Genomic interrogation of ancestral *Mycobacterium tuberculosis* from South India. *Inf Gen Evol*. 2008;8:474–483.
- [61] Chatterjee A, D'souza D, Vira T, et al. Strains of *Mycobacterium tuberculosis* from Western Maharashtra, India, Exhibit a high degree of diversity and strain specific associations with drug resistance, cavitory disease and treatment failure. *J Clin Microbiol*. 2010;48(10):3593–3599.
- [62] Swarnkar P, Katoch VM, Soni H, et al. Spoligotyping of *Mycobacterium tuberculosis* strains from Ziehl-Neelsen stained sputum slides from DTC, Satna (MP). *Int J Res Pharm BioMed Sci*. 2010;1(2):97–101.
- [63] Chawla K, Kumar A, Shenoy VP, et al. Genetic diversity of *Mycobacterium tuberculosis* in South coastal Karnataka, India using spoligotyping. *Ind J Med Res*. 2018;147:278–286.
- [64] Allix-Beguec C, Wahl C, Hanekom M, et al. Proposal of a consensus set of hypervariable mycobacterial interspersed repetitive-unit-variable-number tandem-repeat loci for subtyping of *Mycobacterium tuberculosis* Beijing isolates. *J Clin Microbiol*. 2014;52(1):164–172.
- [65] Singh UB, Suresh N, Bhanu NV, et al. Predominant Tuberculosis Spoligotypes, Delhi, India. *Emerg Inf Dis*. 2004;10(6):1138–1142.
- [66] Albanna AS, Reed MB, Kotar KV, et al. Reduced transmissibility of East African Indian strains of *Mycobacterium tuberculosis*. *PloS One*. 2011;6(9):e25075.