Molecular Epidemiology of *Mycobacterium Tuberculosis* Strains in the North-West and West of Iran

Sahebi L, Ansarin K, Hoffner S, Farajnia S, Seyyedi M, Khalili M, Monfaredan A

Department of Tuberculosis and Lung Disease, Tuberculosis and Lung Disease Research Center, Tabriz University of Medical Sciences, Department of Molecular Biology, Drug Applied Research Center, Tabriz University of Medical Sciences, Department of Medical Philosophy and History, Medical Philosophy and History Research Center, Tabriz University of Medical Sciences, Department of Hematology, Faculty of Medicine, Tabriz Branch, Islamic Azad University, Tabriz, Iran, Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet and The Public Health Agency of Sweden, Stockholm, Sweden

**Abstract**

**Background:** Identifying *Mycobacterium tuberculosis* (MTB) transmission type is a key step in the control of this disease. **Aim:** This study aimed to determine the path and transmission type of MTB and the insertion sequence IS6110 band number and verify their relationship to demographic and clinical risk factors. **Subjects and Methods:** In this cross-sectional study, 64 MTB patients from three border provinces of Iran were selected after full clinical history and physical evaluation design. The drug susceptibility testing was carried out using the standard proportion technique on sputum samples. Isolates tested with restriction fragment length polymorphism technique used IS6110. **Results:** Recent transmission of disease was 33/50 (66%) based on clustering rate. The IS6110 band number had a significant relationship with drug resistance detected in proportion method tested by univariate linear regression (*P* < 0.01). Furthermore, the IS6110 band number had association with Bacillus Calmette–Guérin vaccination history (*P* = 0.02), sex (*P* < 0.01), and purified protein derivative (PPD) reaction size (*P* < 0.01) tested by multiple analysis. The risk of recent transmission inferred from the clustering rate was significantly higher in patients from Western provinces compared to those from the North-West province (*P* = 0.048). However, age (*P* = 0.39), gender (*P* = 0.16), vaccination history (*P* = 0.57), drug susceptibility, and PPD (*P* = 0.6) were independent of clustering. The largest cluster of up to six subjects was found in the Western provinces. **Conclusion:** Recent MTB transmission was much more common in the West compared to the North-West of Iran. Large MTB clusters with strong epidemiological links may be reflective of a disease outbreak. Correlation noted between the IS6110 band number and vaccination history; PPD size and female gender necessitates further studies.

**Keywords:** Molecular epidemiology, *Mycobacterium tuberculosis*, Polymorphism, Restriction fragment length

**Introduction**

Tuberculosis (TB) is well-known and has a significant history of affecting human life. Its pathogen was detected by Robert Koch more than 100 years ago, and the disease is still regarded as an important underlying cause for most of the mortalities and morbidities among other chronic infectious diseases. Up to 2010, 8.8 million new cases of TB and 1.4 million TB-related deaths have been reported worldwide; approximately 95% of TB cases and 98% of TB-induced mortality occur in underdeveloped and developing countries. Despite strong control for TB and routine implementation of directly observed treatment, short-course (DOTS) program, 14,000 new cases have been reported annually in Iran. Iran is categorized as a developing country in West Asia. The average prevalence of TB in Iran was estimated at 33/100,000 in 2012 by the World Health Organization (WHO). Reports from different countries indicate that the rate of recent transmission in...
low-income countries is more than in developed countries. Furthermore, it has to be mentioned that the estimation of recent transmission and recurrence rates of TB have a particular application in the TB control program.

An important factor in the complexity of TB control is the disease’s increasing drug resistance.\(^6\) The prevalence of drug-resistant TB in Iran is 5\% among newly affected patients and 48\% among patients with a recurrence of their previous infection.\(^4\) Although, Iran is ranked moderate in terms of the prevalence of TB and TB (multidrug-resistant tuberculosis [MDR]),\(^5\) being bordered by countries with a high prevalence of TB, including Iraq to the West and the high prevalence of *Mycobacterium tuberculosis* (MTB), and MDR-TB in the Republic of Azerbaijan to the North-West,\(^4\) makes it necessary to track the transmission path of TB and control its prevalence. It can also be said that the identification of disease transmission type, including those endogenous (recurrence of a previous infection) or exogenous (new transmission of the disease) in nature, as well as its severity (response to treatment) are the key factors in the control of TB. At present, several genotyping techniques have been developed which provide invaluable information including the determination of the strain, infection transmission routing, and transmission type (old or new). Among multiple genotyping techniques, *IS6110*-restriction fragment length polymorphism (RFLP) is widely used as a standard method across laboratories worldwide, and its advantage is demonstrated in a recent review by Jagielski et al.\(^7\) Our study aimed to determine the transmission type of MTB in patients with a positive pulmonary smear in three frontier provinces of Iran and analyze the relationship of demographic and clinical characteristics with the transmission type of MTB.

### Subjects and Methods

In this cross-sectional study, from 2012 to 2014, 64 patients with positive pulmonary smears from East Azerbaijan (*n* = 30), Kurdistan (*n* = 17), and Kermanshah (*n* = 17) provinces located in the North-West and West of Iran were investigated. This study was approved by the Medical Ethics Committees of Tabriz Medical University; moreover, patients were selected randomly to avoid selection bias and asked for written informed consent before interviews. The inclusion criteria were: At least two positive pulmonary smears for acid-fast bacilli, or one positive pulmonary smear along with radiographic abnormalities consistent with active pulmonary TB, or one positive smear and one positive MTB culture.\(^8\) MTB was identified using para-nitrobenzoic acid.

### DNA extraction

DNA was extracted from clinical isolates grown on the Löwenstein–Jensen (LJ) medium as described in a study by van Soolingen et al.\(^9\)

### Drug susceptibility testing

Drug susceptibility testing was carried out using the LJ medium according to the proportion method recommended by the WHO and the International Union Against TB and Lung Disease using the following materials:\(^9\) Isoniazid (INH): 0.2 mg/L, rifampicin (RMP): 40 mg/L, ethambutol (EMB): 2 mg/L, and streptomycin (STM): 4 mg/L.

**IS6110-RFLP**

To implement the *IS6110*-RFLP, the PVUII-digested DNA was separated on 1\% agarose gels and capillary-blotted onto a nylon membrane. A 245-base pair internal polymerase chain reaction (PCR) fragment of *IS6110* was amplified by PCR and used as a probe. After labeling, the hybridization of the DNA was performed at 42°C for 24 h following manufacturer (ECL System, GE Healthcare UK Limited, Amersham Place Little Chalfont, Buckinghamshire, UK). The total of PVUII-digested DNA of the MTB reference strain H37RV was used as the control in Southern blot experiments.

### Computer analysis

The *IS6110* fingerprinting pattern of MTB strains were analyzed using Gel Compare II (Windows 7, version 6.6, Applied Math, Kortrijk, Belgium). At first, the autoradiograms were scanned with an optical resolution of 190 dpi, then the position of *IS6110* was normalized by an internal marker, and its accuracy was verified using *IS6110* banding pattern in H37RV strains. The similarity of the matrix was estimated using dice band-based similarity coefficient at a 2\% tolerance-interval, 1\% optimization, and a 10–95\% active zone. Similarity was considered above 70\% when determining clustering. Dendrograms were designed by the hierarchic un-weighted pair clustering group method analysis algorithm.

### Statistical tests

In a pilot study, the validity of the questionnaire was evaluated as confirmed by expert professors, and its reliability was assessed using the statistical reliability test (Cronbach’s alpha = 0.92). A univariate logistic regression test was used to compare geographic, clinical, and demographic variables in the clusters (recent transmission) and unique (recurring transmission) groups. Univariate and multiple linear regression tests were used for analysis of the mean number of strains’ bands per geographical regions, drug resistance, vaccination history, purified protein derivative (PPD) diameter, gender, age, immigration history, and family history of TB using Statistical Package for the Social Sciences (SPSS-18) software (IBM SPSS, Chicago, Illinois, USA).

### Results

There were 31 men, 31/64 (48.4\%) and 33/64 (51.6\%) women involved in our study. The mean age of subjects was 54.4 (23.62) years (range: 15–83 years). They were HIV
negative (ELISA test), and underwent DOTS system during treatment.

The prevalence of any drug resistance and MDR was 9/64 (14.1%) and 2/64 (3.1%) among TB patients, respectively. And the frequency of resistance to RMP was 3/64 (4.7%), to INH was 3/64 (4.7%), to STM was 5/64 (7.8%), and to EMB was 2/64 (3.1%). The average number of bands was 9.2 (standard deviation [SD] =3.2) and the minimum and maximum number of isolate bands were 3 and 16, respectively. In this study, 14/64 (21.88%) isolates had less than seven bands, which were excluded from similarity analysis. After the exclusion of bands less than seven, 17/50 (34%) isolates were unique and 33/50 (66%) isolates were clustered. Composition of clusters of isolates were as follows, seven clusters with two isolates, tree clusters with three isolates, one cluster with four isolates, and one cluster with six isolates [Figure 1].

In a comparison of classical and molecular epidemiology results, significant geographical and gender matching existed within different clusters [Table 1]. Typically, an increased number of isolates in each cluster decreases the probability of matching between classic and molecular data. Meanwhile, the largest cluster of six members had a 100% geographical connection; all six members were from the Western provinces [Table 1]. In one of the dual clusters, no classical connection existed between them and given that these two samples were always together during the experiment, there was a possibly of cross-contamination, although no definite comment can be made in this case.

In terms of formed clusters, 9/33 (27.27%) of East Azerbaijan strains, 12/33 (36.36%) of Kermanshah strains, and 12/33 (36.36%) of Kurdistan strains were involved in clustering. Accordingly, isolates from Kurdistan and Kermanshah were more involved in the creation of large population clusters.

In terms of drug resistance, there was a significant relationship between any drug resistance and the band number of the strains in univariate linear regression analysis (β =2.5, P = 0.04).

The average number of strains’ bands in isolates from the North-West and West geographical regions differed significantly from each other so that the mean number of bands was 7.2 (SD = 2.5) in East Azerbaijan strains, and 10.8 (SD = 2.8) in Kurdistan and Kermanshah (P < 0.001). Interestingly, there was a statistically significant relationship between the band number of strains and vaccination history, PPD diameter, and gender in multiple linear regression. Therefore, in female patients without vaccination history, the number of bands in TB strains were more than in patients with vaccination history and males (P < 0.01 and P < 0.001, respectively) and negative correlation existed between PPD diameter and the number of bands (P < 0.01) [Table 2].

Comparative analysis of the North-West and West provinces based on the odds of clustering showed that isolates from Western provinces had a significantly (6 times) higher chance to form clusters than those from East Azerbaijan (odds ratio [OR] =6.0, confidence interval [CI] 0.95%: 1.4–24.7, P = 0.013).

No significant statistical relationship was found when comparing gender, age, vaccination history, PPD, and drug susceptibility in unique and clustered groups. Sex (female): OR = 3.3, CI 95%, 0.6–18.0, P = 0.16; age (year): β = –0.019, SE (β) =0.023, P =0.39; vaccination history (yes): OR = 0.57, CI 95%, 0.055–5.8, P = 0.63; PPD (mm): β =0.27, SE (β) =0.56, P = 0.63 any drug resistance (yes): OR = 1.9, CI 95%, 0.02–7.7, P = 0.60.

**Discussion**

The participants of the present study were native to East Azerbaijan (North-West of Iran), Kurdistan, and

![Figure 1: IS6110 DNA fingerprint patterns of the 64 Mycobacterium tuberculosis strains and the corresponding dendrogram banding patterns are ordered by similarity. The position of each IS6110 band is normalized so that all banding patterns of all strains are mutually comparable. Scale depicts similarity of patterns calculated as described in materials and methods](downloaded from http://www.amhsr.org)
Kermanshah Provinces (West of Iran). Among 64 strains in different patients, the mean number of bands of IS6110 was 9.1 (SD = 3.2). In the present study, the frequency of strains with less than seven bands was 14 (21.88% of the isolates). The relative frequency of low-bands strains in various regions was highly diverse. For example, in Abbadi’s study in Egypt, none of the models were low-copy (less than five bands),[19] whereas in the study by Siddiqi in India, approximately 42% of isolates had less than five bands.[20] There is also no consistency in the relative frequency of an IS6110 copy-number in different regions of Iran[9,12-14]; for example, the percentage of isolates with less than six copies was 3.2% in Fars Province (South of Iran),[13] 38% in Khorasan Province (East of Iran),[11] and 22.75% in Markazi Province (Central Iran).[10] It seems that the number of bands is influenced by different geographic regions. However, the mean number of bands varied significantly in different geographical regions (P < 0.001), but in a closer examination and modeling based on multiple analysis, this association was not maintained. It is possible that the unique characteristics of each region may affect the variation in band number and in this regard, the region may play a role as a mediating variable.

In conformity with Green’s study in South Africa,[13] a significant relationship was found between drug resistance and band number in univariate linear regression, but this association was not maintained in multiple analyses. In our study, patients with no history of vaccination were at higher odds of infection by MTB isolates with a greater number of IS6110 bands. Based on some studies, vaccination can prevent the development of more severe forms of the disease in children.[16] In addition, vaccination strongly protects against extra-pulmonary TB.[17] According to the study by Lei in China, the odds of infection with resistant TB in individuals with a vaccination history was low.[18] Given the possible relationship between strain type and drug resistance on one hand[19] and relationship of the strain type and the number of IS6110 bands in various strains[20] on the other, it seems that there is a relationship between vaccination history and the number of bands. This result may indicate the crucial assumption that vaccination creates more immunity against MTB strains with low band number, which necessitates further research in this area. Another finding about different band numbers of MTB isolates was PPD diameter so that with increased PPD diameter, the likelihood of a higher number of bands in isolates was low. Based on the results of a study by Shahemabadi et al. in 2007, obvious changes occur in immunological responses, hence PPD in patients with MDR-TB.[21] However, no study is available about strain type or isolates band number and immunological response in the patients.

Another important result is that women were more likely to have high-band isolates than men. Studies show that genetic characteristics highly affect the manifestation, severity, and drug resistance of MTB.[22] Therefore, the IS6110 band number of strains may also be related to gender. However, in general, the mechanism(s) by which the IS6110 copy number affects the characteristic of strains is are unclear, and further studies need to be made to address this issue.

In the present study, the relative frequency of clustering groups was 66%. This is higher than the frequency reported in another study in Iran in which the clustering rate was 43%.[23] In a study conducted in Arkansas, the clustering percentage was 52%.[24] Abbadi reported the clustering rate as 18% in Egypt[19] and Borges reported 29% in Brazil.[25]

In the present study, the geographic region had a significant relationship with clustering and hence the risk of new transmission. Because of this, the majority of patients in the Western provinces (Kurdistan and Kermanshah) were infected due to the recent transmission of the disease. The largest cluster of up to six subjects was found in the Western provinces (which can be a serious warning of a potential outbreak in the specific area). This outbreak may originate from Iraqi states adjacent to these provinces. To consolidate this hypothesis, further investigation is required to determine and compare the common strains with those in Iraq and the Azerbaijan Republic. In the present study, no statistically significant relationship was found between age and gender with clustering in TB-related isolates. Although a positive relationship was achieved between males and clustering in a study in France,[26] a number of studies have reported age and gender independent of clustering.[27,28] In our study, the epidemiological link based on area, vaccination history, and gender were obtained 59.53; 40.47, and 44.4%, respectively, for patients with the recent transmission. In a study by Borger, the epidemiological link based on residential proximity and the family reaction was 37.5%,[29] whereas it was 30.2% in Hong Kong.[28] Due to a strong correlation between band number of IS6110 and no vaccination history, PPD diameter, and female gender, it is
necessary to conduct further studies to determine the causal relationship.

In this study, recent transmission of the disease was much more common in the West compared to North-West of Iran. Large MTB clusters with strong epidemiological links may be reflective of a disease outbreak. Correlation noted between the IS6110 band number and vaccination history, PPD size, and female gender necessitates further studies.

Acknowledgment

This research was a section of Ph.D. Thesis. We wish to thank Mrs. Lesley Carson, who helpfully edited the article.

References


Source of Support: Nil. Conflict of Interest: None declared.