

## Original Article

Genotyping of *Mycobacterium Tuberculosis* Isolated from Suspected Patients in Tehran in 2015-2017Saeed Zaker<sup>1\*</sup> Ph.D., Hanie Bagherifard<sup>2</sup> M.Sc.<sup>1</sup>Department of Biology and Biotechnology, Islamic Azad University, Parand branch, Parand, Iran.<sup>2</sup>Department of Microbiology, Islamic Azad University, Damghan branch, Iran.

## A B S T R A C T

**Article history**

Received 9 Jan 2019

Accepted 2 Jul 2019

Available online 31 Aug 2019

**Key words**

Genotyping

*gyrA**katG**Mycobacterium tuberculosis*

**Background and Aims:** Unlike many global efforts to eradicate tuberculosis caused by *Mycobacterium*, it remains as a life-threatening infection with a worldwide incidence of 1.5 million cases each year. However, due to the lack of information about *Mycobacterium tuberculosis* characterization, more studies are required to evaluate strain diversity and epidemiology of tuberculosis to improve the therapeutic approaches. This study aimed to genotype the *Mycobacterium tuberculosis* isolated from suspected patients in Tehran, Iran through 2015-2017.

**Materials and Methods:** In the current study, 30 isolates (sputum, broncho-alveolar lavage and biopsy) were collected from different tuberculosis patients at Massoud Clinical Lab of Tehran from 2015 to 2017. To find the single nucleotide polymorphisms and mutated regions, polymerase chain reaction (PCR) was performed on all the isolates to amplify the *katG* and *gyrA* genes. Then, PCR products were sequenced and analyzed.

**Results:** The majority of isolates were assigned to PGG2 (90%), followed by PGG3 (10%) but no isolate belonging to PGG1 was found.

**Conclusions:** Our findings demonstrate a remarkable epidemiological pattern of *tuberculosis* in Tehran. In group 2, isolates showed a considerably higher frequency compared to isolates in group 3, which is consistent with other findings reported in Iran. However, in contrast to other Iranian studies, no isolated strains were categorized in principal PGG1.

## Introduction

It has been reported that one-third of the world's population is diagnosed to have *Mycobacterium tuberculosis* (*M. tuberculosis*) infections [1]. Tuberculosis is the second-leading cause of death in microbial infectious diseases after the human immunodeficiency virus, which causes nearly 2 million deaths, 50 to 100 million infection report, and 8 to 9 million new cases every year [1, 2]. Tuberculosis is one of the most common infectious diseases in Iran. Based on the World Health Organization (WHO) report in 2012, the incidence rate of tuberculosis was estimated 21 cases per 100,000 in Iran [3]. There are different types of genetic variation among the bacterial population which has influenced the evolutionary history of the microbial community as well as the *M. tuberculosis* complex (MTC) over time [4]. Single nucleotide polymorphism (SNP) is a variation of a particular nucleotide (A, T, C & G) which is the most common type of genetic variation [5, 6]. Duplication, insertion, and deletion are also other genomic errors which can occur at different regions in the bacterial genome. DNA-based genotyping allows the detection of SNPs, genome rearrangements such as deletion or insertion variable-number tandem repeats (VNTRs) and the analysis of target genes for the positive selection [7]. To investigate the variation and epidemiology of *M. tuberculosis* strains, molecular typing analysis of isolated strains has been a valuable strategy which can help researchers to understand the *M. tuberculosis* variations profile more

accurately and improve the therapeutic approaches as well [8]. There are various DNA based techniques such as restriction fragment length polymorphism (RFLP), pulsed field gel electrophoresis (PFGE), multiple-locus variable-number tandem repeat analysis (MLVA), random amplified polymorphic DNA (RAPD), and high resolution melting (HRM) to identify the epidemiological characterization of bacterial strains [8]. However, or RFLP method using derived probes from the insertion elements like IS6110 has been considered as the gold standard method for characterization of MTC [9, 10]. RFLP analysis was also applied on the MTC to detect the polymorphic GC-rich repeat sequences. Most of the present MTC genotyping methods rely on the repetitive or mobile DNA elements which are pertinent to the convergent evolution. Genotyping techniques are suboptimal in phylogenetic studies and strain classification [11, 12]. However, SNPs are ideal markers for classification of the MTC into the phylogenetic lineages, as they exhibit very low degrees of homology [13]. Based on the combination of SNPs occurred at the *katG* codon 463 (encoding the catalase-peroxidase enzyme) and *gyrA* codon 95 (encoding the A subunit of DNA gyrase), *M. tuberculosis* strains are divided into three principal genetic groups (PGG1s-PGG3) [4]. DNA sequencing is another effective method to identify the bacterial genomic variation more accurately. In comparison with other SNP genotyping methods, DNA sequencing is particularly suitable for the

identification of multiple SNPs in a small region of DNA with high sensitivity and specificity [14]. The present study is designed to analyze the population structure of *M. tuberculosis* strains in the capital of Iran, Tehran (the most populated city in the country), to evaluate the most commonly circulating strains and to monitor the distribution of PGGs in this region.

## Materials and Methods

### *M. tuberculosis* isolates

From December 2015 to May 2017, *M. tuberculosis* isolates were obtained from sputum, bronchoalveolar lavage and biopsy samples of patients with active pulmonary tuberculosis in Massoud laboratory, Damghan, Iran. All 30 tuberculosis patients had proven registration of clinical diagnostic examinations such as chest X-ray, purified protein derivative, cough, weight loss, gender, etc. The isolates were cultured on Lowenstein-Jensen solid medium and the resulting colonies were identified at the species level using 2-thiophene carboxylic acid and paranitrobenzoic acid (PN99B) selective media, or by standard biochemical procedures [15]. Two sensitive isolates were used as negative controls.

### Susceptibility testing

Anti-microbial drug susceptibility testing (AMST) was performed using the Communicable Disease Center (CDC) standard conventional proportional method. This involved the use of rifampicin (Rif) 40 mg/ml, isoniazid (INH) 0.2 mg/ml, streptomycin (SM) 4 mg/ml, and kanamycin (K) 20 mg/ml in Lowenstein-Jensen medium. In addition, breakpoint concentrations of isoniazid

0.1 mg/ml, and rifampicin 2.0 mg/ml were also used in the BACTEC system. Mutations of the *katG* and *gyrA* genes were identified in 30 rifampicin resistant isolates by DNA sequencing.

### PCR amplification

DNA extraction was carried out using boiling method through which bacterial colonies were suspended in 500 µl of distilled water. PCR was performed using *katG*-463 F: (5'-CGGTCGAACTAGCTGTGAGACAGT-3' R: 5'-AAGCCGAGATTGCCAGCCTTA-3' and *gyrA*-95 F: (5'-CGAGACCATGGGCAACTACCA-3' R: 5'-ATTGCCTGGCGAGCCGAAGT-3' [17].

PCR reaction was performed in a 50 µl reaction mixture containing 50 mM KCl, 10 mM Tris (pH 8.0), 1.5 mM MgCl<sub>2</sub>, 5 mM of deoxynucleoside triphosphates (dNTPs), 1U of Taq polymerase, 20 pmoles of each set of primers and 6 mM of chromosomal DNA. Samples were then subjected to one cycle of denaturation at 94°C for 5 min, followed by 36 cycles at 94°C for 1 min, 57°C for 1 min, 72°C for 1 min and a final cycle at 72°C for 10 min to complete the elongation of the PCR intermediate products. PCR products were then run on 2% agarose gels and examined for the presence of the *katG*-463 and *gyrA*-95 bp band after ethidium bromide staining. The DNA purification was performed using an Amplicon, Viragene kit (cat number=#A180306).

### DNA Sequencing

A 463bp - 95 fragments of the *katG* and *gyrA* genes were amplified by PCR using two primers: *katG*-F(5'-CGGTCGAACTAGCTGTGAGACAGT-3') *katG* -R (5'- AAGCCGAGATTGCCAGCCTTA -3') and *gyrA*-95(5'-CGAGACCATGGGCAACTACCA-3' forward primer, and reverse primer

5'-ATTGCCTGGCGAGCCGAAGT-3'. PCR was carried out in a 8 ml reaction mixture containing 0.25 ml of DNA polymerase in 0.9 ml of buffer (PCR), 2 ml of a mixture of dNTP and dNNTP (dATP, dTTP, dCTP, dGTP), 0.5 ml of each primer (2.5 pmoles), 1 ml of DNA and 3.35 ml of H<sub>2</sub>O (Molecular Biology grade). Amplification was carried out for 33 cycles, with the following program: denaturation at 94°C for 30 sec; primer annealing at 54°C for 30 sec; extension at 72°C for 90 sec. A 463bp - 95 fragments of the *katG-gyrA* genes extracted from *M. tuberculosis* strains were sequenced by the Amersham auto sequencer and Amersham Pharmacia DYEnamic ET Terminator Cycle Sequencing Premix Kits. Alignment of the DNA fragments (*katG-463* and *gyrA-95*) was performed using the MEGA 6 software.

#### Data analysis of DNA sequences

Alignment of the DNA fragments (*katG-gyrA*) was carried out using MEGA 6 DNAMAN software and compared with standard strains CDC1551, H37RV and *M. tuberculosis* strain 210. The BLAST 2 sequences program was used for DNA sequence comparisons (<http://www.ncbi.nlm.nih.gov/BLAST/>).

## Results

#### Bacterial strains and drug susceptibility

**assay:** All samples were cultured and identified as *M. tuberculosis* by the PCR method. All the 30 isolates examined were resistant to drugs. But 11 (34%), 28 (90%) and 10 (31%) of the isolates were found to be resistant to isoniazid,

streptomycin and etambutol, respectively. In this study we found four strains to be mono-resistant to rifampicin. From 30 resistant isolates, 12 (35%) were isolated from sputum and 18 isolated from bronchoalveolar lavage and biopsy of patients with primary infection and 22 (65%) isolates were obtained from secondary infections.

#### Definitions in this study

Primary infection is attributed to the patient who has no previous history of tuberculosis nor medical treatment. Secondary infection demonstrates a previous history of tuberculosis in the patient's medical records.

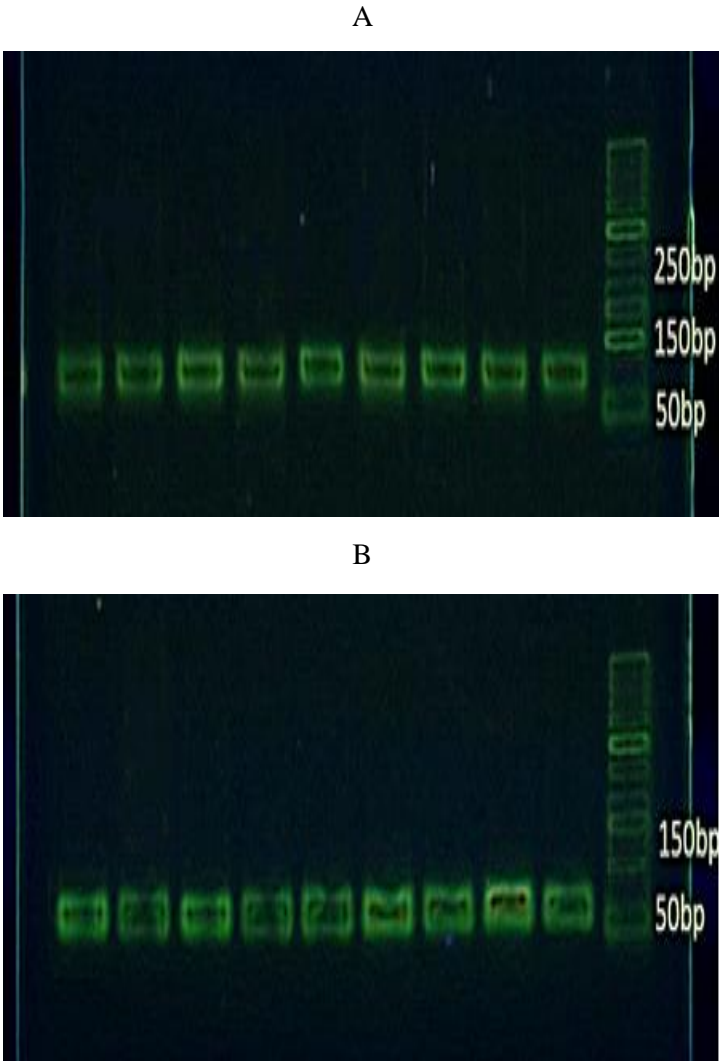
#### PCR amplification and DNA sequencing

In this study, a total of 30 *M. tuberculosis* strains were classified in the three so-called principal genotypic groups. The *katG* and *gyrA* genes have been detected through the conventional PCR in all the infected samples (Fig. 1). Allele C and G were found in 27 and 3 cases at *gyrA* locus 95, respectively and all the isolates showed allele G at *katG-463* locus.

Our results indicated the majority of the isolates belonging to group 2 (90%), followed by group 3 (10%) and no isolate belonged to group 1. Patients, characteristics including age and sex were analyzed against principal genetic groups based on polymorphisms at *katG-463* and *gyrA-95*, the results of which indicated that neither age nor sex was significantly associated with *M. tuberculosis* genotypes (Table 1).

**Table 1.** Comparison of patient characteristics according to *katG*-463 and *gyrA*-95 genotype of the corresponding isolate

Characteristics	Group 2 (N=27)	Group 3 (N=3)	P-value
No. (%) of subjects			
Female	13 (48.1)	1 (33.3)	0.626
Male	14 (51.9)	2 (66.7)	
Mean (SD) age (yr)	50.59 (20.35)	52 (16.64)	0.9



**Fig. 1.** Electrophoresis analysis of the *katG* and *gyrA* gene PCR on 1.5% agarose gel. 130-bp products amplified from the *gyrA* gene of *M. tuberculosis* isolates (A) 72-bp products amplified from the *katG* gene of *M. tuberculosis* isolates (B).

Discussion

In a relatively similar study, whole-genome sequencing of *M. tuberculosis* showed 110 SNPs unique to each isolated strain. In this

study, a subset of 32 SNPs were used to type *M. Tuberculosis* isolates into the ancient Beijing, modern Beijing, East African–Indian,

and Latin American group. However, there was high genetic homozygosity within the Euro-American lineage, which included spoligotype-classified Haarlem and T strains. Whole-genome sequencing of 12 representative Euro-American isolates divided the multiple subtype-specific SNPs into two major branches within the Euro-American lineage [19].

SNPs have been reported as robust and stable markers to evaluate the genetic variation phylogenetically. Based on Francisc Coll study, 92k SNP was detected among a global collection of 1,601 genomes. Of the 7k strain-specific SNPs, 62 markers are well studied which are proposed to distinguish known circulating strains. These SNP-based barcodes are the first reliable markers to cover all the main lineages and classify a higher number of sublineages than current alternatives [20].

Two complementary SNP-based genotyping methods have been well established to classify strains into six main human-associated lineages of MTC, the “Beijing” sublineage, and clade comprising *Mycobacterium bovis* and *Mycobacterium caprae*. Phylogenetically informative SNPs were acquired from 22 MTC whole-genome sequences. The first evaluation, referred to as MOL-PCR, is a ligation-dependent PCR with signal detection by fluorescent microspheres and a Luminex flow cytometer, which simultaneously interrogates eight SNPs. The second assay is based on six individual TaqMan real-time PCR assays for singleplex SNP-typing. Stucki et al. compared MOL-PCR and TaqMan results in two panels of clinical MTC isolates. Both methods fully

agreed when assigning 36 strains into the main phylogenetic lineages.

Unlike sensitivity and specificity of MOL-PCR and TaqMan, the former seems ideal for classification of isolates with no previous information while the latter requires less time for confirmation. However, both methods are flexible and comparably inexpensive [13].

These new molecular-based technologies can provide tuberculosis evolution, identification of resistant pathogens and, more importantly, the possible correlation of patient origin and the tuberculosis epidemiology [21].

Interestingly, Bouakaze et al. developed an innovative strategy based on two multiplex allele-specific mini-sequencing assays that allowed the detection of eight species-specific and eight lineage-specific SNPs [17]. Each test consisted of eightplex PCR amplification, followed by an eightplex mini-sequencing reaction with the SNaPshot multiplex kit (Applied Biosystems) and the analysis of the extension products by capillary electrophoresis. The whole strategy was developed with a panel of 56 MTC strains and 15 negative controls [17].

Except for one *M. africanum* strain, other isolates were classified to the species level, and all *M. tuberculosis* isolates were further genotyped successfully in the study mentioned above.

This two-step strategy based on SNaPshot mini-sequencing allows the simultaneous differentiation of closely related members of the MTC, the distinction between principal genetic groups, and the characterization of *M. tuberculosis* isolates into one of the seven prominent SNP cluster groups and could be a



useful tool for diagnostic and epidemiological purposes [17].

## Conclusion

Our findings demonstrate a remarkable epidemiological pattern of tuberculosis in Tehran and confirm the limitation of MTC isolates to evolutionary modern principle genetic group (PGG)2/3 strains. However, the group 2 isolates showed a considerably higher frequency than isolates belonging to group 3, which is consistent with the other findings [22]. However, in contrast to other studies in Iran, we found no isolate in PGG1 [22, 23]. Moreover, no correlation was found between the age and gender of the patients and the genetic grouping of *M. tuberculosis*.

This study gives an outline of the *M. tuberculosis* strains circulating in Tehran and describes the distribution of the significant principal genetic groups. It contributes to a better understanding of the current trend of tuberculosis in the capital city of a high-incidence middle-east country which can improve therapeutic approaches, treatment, and vaccination policy.

## Conflict of Interest

The authors declare that there are no conflicts of interest associated with this manuscript.

## Acknowledgments

We appreciate our colleagues in Massoud clinical Lab of Microbiology department. Also we thank our colleague from the Microbiology department of Massoud Lab (Reza Azizyan) for providing clinical strains as well as his valuable recommendations.

## References

- [1]. Frieden TR, Sterling TR, Munsiff SS, Watt CJ, Dye C. Tuberculosis. Lancet 2003; 362(9387): 887-99.
- [2]. Lari N, Rindi L, Sola C, Bonanni D, Rastogi N, Tortoli E, et al. Genetic diversity, determined on the basis of katG463 and gyrA95 polymorphisms, Spoligotyping, and IS6110 typing, of Mycobacterium tuberculosis complex isolates from Italy. J clin microbiol. 2005; 43(4): 1617-624.
- [3]. World Health Organization. Global health observatory data repository 2013. Available at: who.int/gho/data/node/main A. 2013; 364.
- [4]. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, et al. Restricted structural gene polymorphism in the Myco-bacterium tuberculosis complex indicates evolutionarily recent global dissemination. Proc Natl Acad Sci USA 1997; 94(18): 9869-874.
- [5]. Kim S, Misra A. SNP genotyping: technologies and biomedical applications. Annu Rev Biomed Eng. 2007; 9(1): 289-320.
- [6]. Royo JL, Galan JJ. Pyrosequencing for SNP genotyping. Methods in molecular biology 2009; 578: 123-33.
- [7]. Li R, Li Y, Fang X, Yang H, Wang J, Kristiansen K, et al. SNP detection for massively parallel whole-genome resequencing. Genome Res. 2009; 19(6): 1124-132.
- [8]. Ravansalar H, Tadayon K, Ghazvini K. Molecular typing methods used in studies of Mycobacterium tuberculosis in Iran: a systematic review. Iran J Med Microbiol. 2016; 8(5): 338-46.
- [9]. McEvoy CR, van Pittius NCG, Victor TC, van Helden PD, Warren RM. The role of IS6110 in the evolution of Mycobacterium tuberculosis. Tuberculosis 2007; 87(5): 393-404.
- [10]. Roychowdhury T, Mandal S, Bhattacharya A. Analysis of IS6110 insertion sites provide a glimpse into genome evolution of Mycobacterium tuberculosis. Sci Rep. 2015; 5(1): 12567.
- [11]. Poulet S, Cole ST. Repeated DNA sequences in mycobacteria. Arch Microbiol. 1995; 163(2): 79-86.
- [12]. Poulet S, Cole ST. Characterization of the highly abundant polymorphic GC-rich-repetitive sequence (PGRS) present in Mycobacterium

- tuberculosis. *Arch Microbiol.* 1995; 163(2): 87-95.
- [13]. Stucki D, Malla B, Hostettler S, Huna T, Feldmann J, Yeboah-Manu D, et al. Two new rapid SNP-typing methods for classifying *Mycobacterium tuberculosis* complex into the main phylogenetic lineages. *PloS One* 2012; 7(7): 41253.
- [14]. Ford CB, Lin PL, Chase MR, Shah RR, Iartchouk O, Galagan J, et al. Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat Genet.* 2011; 43(5): 482-86.
- [15]. Petroff S. A new and rapid method for the isolation and cultivation of tubercle bacilli directly from the sputum and feces. *J Exp Med.* 1915; 21(1): 38-42.
- [16]. Sambrook J, Fritsch EF, Maniatis T. *Molecular cloning: a laboratory manual.* 2<sup>nd</sup> ed, NY: Cold spring harbor laboratory Cold Spring Harbor; 1989.
- [17]. Bouakaze C, Keyser C, De Martino S, Sougakoff W, Veziris N, Dabernat H, et al. Identification and genotyping of *Mycobacterium tuberculosis* complex species by use of a SNaPshot Minisequencing-based assay. *J Clin Microbiol* 2010; 48(5): 1758-766.
- [18]. The NCBI handbook. National Library of Medicine (US), N.C.f.B.I.O.C., The Reference Sequence (RefSeq) Project [Internet]. Available at: <http://www.ncbi.nlm.nih.gov/nuccore/448814763/?report=genbank>.
- [19]. Dou HY, Lin CH, Chen YY, Yang SJ, Chang JR, Wu KM, et al. Lineage-specific SNPs for genotyping of *Mycobacterium tuberculosis* clinical isolates. *Sci Rep.* 2017; 7(1): 1425.
- [20]. Coll F, McNerney R, Guerra-Assuncao JA, Glynn JR, Perdigao J, Viveiros M, et al. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nat commun.* 2014; 5(2): 4812.
- [21]. de Viedma DG, Mokrousov I, Rastogi N. Innovations in the molecular epidemiology of tuberculosis. *Enferm Infecc Microbiol Clin.* 2011; 29(1): 8-13.
- [22]. Eshghinejad A, Farazi A, Eshrati B, Khalili H, Shojapour M, Ahmadi A, et al. Detection of major genetic groups of mycobacterium tuberculosis isolated from tuberculosis patients in Markazi province by polymorphism determination in kat G and gyr A genes. *J Arak Uni Med Sci.* 2012; 15 (5): 26-34.
- [23]. Zaker Bostanabad S, Titov LP, Slizen VV, Taghikhani M, Bahrmand A. katG mutations in isoniazid-resistant strains of *Mycobacterium tuberculosis* isolates from Belarusian patients. *Tuberkuloz ve toraks.* 2007; 55(3): 231-37.