The mutation rate of *rpoB* gene showed an upward trend with the increase of MIRU10, MIRU39 and QUB4156 repetitive number

Fan Su[1†], Lei Cao[1†], Xia Ren[1], Jian Hu[1], Grace Tavengana[1], Huan Wu[2], Yumei Zhou[1], Yuhan Fu[1], Mingfei Jiang[3] and Yufeng Wen[1*]

Abstract

Background Mycobacterial interspersed repetitive unit-variable number tandem repeat (MIRU-VNTR) is a frequently used typing method for identifying the Beijing genotype of *Mycobacterium tuberculosis* (*Mtb*), which is easily transformed into rifampicin (RIF) resistance. The RIF resistance of *Mtb* is considered to be highly related with the mutation of *rpoB* gene. Therefore, this study aimed to analyze the relationship between the repetitive number of MIRU loci and the mutation of *rpoB* gene.

Methods An open-source whole-genome sequencing data of *Mtb* was used to detect the mutation of *rpoB* gene and the repetitive number of MIRU loci by bioinformatics methods. Cochran-Armitage analysis was performed to analyze the trend of the *rpoB* gene mutation rate and the repetitive number of MIRU loci.

Results Among 357 rifampicin-resistant tuberculosis (RR-TB), 304 strains with mutated *rpoB* genes were detected, and 6 of 67 rifampicin susceptible strains were detected mutations. The *rpoB* gene mutational rate showed an upward trend with the increase of MIRU10, MIRU39, QUB4156 and MIRU16 repetitive number, but only the repetitive number of MIRU10, MRIU39 and QUB4156 were risk factors for *rpoB* gene mutation. The Hunter-Gaston discriminatory index (HGDI) of MIRU10 (0.65) and QUB4156 (0.62) was high in the overall sample, while MIRU39 (0.39) and MIRU16 (0.43) showed a moderate discriminatory Power.

Conclusion The mutation rate of *rpoB* gene increases with the addition of repetitive numbers of MIRU10, QUB4156 and MIRU39 loci.

Keywords RpoB gene, Mutation, MIRU10, QUB4156, MIRU39, Repetitive number

[†]Fan Su and Lei Cao contributed equally to this work.

*Correspondence:
Yufeng Wen
wyf2015w@sina.com
[1] School of Public Health, Wannan Medical College, Wuhu, Anhui Province, China
[2] School of Laboratory Medicine, Wannan Medical College, Wuhu, Anhui Province, China
[3] School of Clinical Medicine, Wannan Medical College, Wuhu, Anhui Province, China

Introduction

In recent years, the global tuberculosis (TB) epidemic continues to be serious. Drug-resistant tuberculosis, especially those resistant to rifampicin (RR-TB), has become one of the major obstacles to achieve the goal of TB elimination [1]. According to the Global Tuberculosis Report 2020 of World Health Organization (WHO), there were an estimated 465,000 (range, 400,000–535,000) incident cases of RR-TB, and China accounts for 14% of them [2].

Su *et al. BMC Genomics*   (2023) 24:26

Page 2 of 6

More than 95% of rifampicin resistance is associated with mutations in the *rpoB* gene of *Mycobacterium tuberculosis* (*Mtb*), with 97% of mutations occurring within the 81 bp rifampicin-resistant determining region (RRDR) of this gene [3]. Besides, it has been proved that sequence mutation out of RRDR may be involved in the formation of rifampicin cross resistance [4]. In China, Beijing genotype tuberculosis occupies a dominant position of *Mtb*, Uddin MKM et al. [5] had proved the mutation of *rpoB* gene was a risk factor of rifampicin resistance for Beijing genotype TB.

In recent years, mycobacterial interspersed repetitive units-variable number tandem repeats (MIRU-VNTR) had been widely used in the typing of TB. Combined with Spoligotyping, MIRU-VNTR typing can distinguish Beijing family genotype with other genotype strains by cluster analysis [6]. Besides, different MIRU loci showed different discriminatory power for Beijing and non-Beijing genotype strains and significant differences were found in mutation of the *rpoB* gene between two genotype [7, 8]. On this basis, we hypothesized that there may be a correlation between the mutation of *rpoB* gene and the repetitive number of MIRU loci.

With the popularization of whole-genome sequencing technology, the mutation of known drug resistance genes and MIRU-VNTR information can be obtained based on the analysis of *Mtb* Illumina, Pacific Biosciences or Oxford Nanopore sequencing data [9, 10]. Therefore, we conducted this study to explore the relationship of *rpoB* gene mutation and MIRU loci with sequencing data.

Methods

Data source
Sample information was acquired from one study of the Chinese Center for Disease Control and Prevention (Chinese CDC) [11], including the phenotypic drug resistance of each strain, type of patient from which the strain originated, etc. Whole-genome sequencing raw data were deposited at NCBI Sequence Read Archive (SRP134826) and Genome Sequence Archive (CRA000786) (https://ngdc.cncb.ac.cn/search/?dbId=gsa&q=CRA000786)

RpoB gene mutation determination
In the first step, the sequencing data was submitted to remove linker and low-quality base treatment (filtering the bases with Phred <20) using Fastp (https://github.com/OpenGene/fastp) software. Secondly, BWA (http://bio-bwa.sourceforge.net/bwa.shtml) software was used to compare the above sequence data with the genome template sequence of *Mycobacterium tuberculosis* standard strain (H37Rv) (obtained from the gene sequence database GenBank access: NC 000962.3 maintained by the National Institutes of Health). In the third step, according

to the comparison results, sequencing data samples were screened that the sequencing depth is more than $10\times$ and the genome coverage is more than 95%. Finally, SNPs of each strain compared with H37Rv were identified using Samtools (https://github.com/samtools/samtools/issues), and the lowest value of comparison quality was set to 30. Then VarScan 2 (http://varscan.sourceforge.net) software was used to further identify and screen SNP fixed mutations with a frequency of more than 75% and supported by at least 10 sequences. The whole genome SNPs detected in this study were compared with known *rpoB* gene mutations (obtained from GenBank gene database) to obtain the mutation information of *rpoB* gene of each strain [12], and only the non-synonymous mutations were recorded.

MIRU loci repetitive number determination
The sequencing data outputted from Fastp were assessed by FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) to guarantee good reads quality. Spades (https://github.com/ablab/spades) was carried to assemble second generation sequencing data to long sequence, the finally assembled data were assessed by QUAST (http://bioinf.spbau.ru/quast) and BUSCO (https://busco.ezlab.org). MIRUReader (https://github.com/phglab/MIRUReader) was used to get the repetitive number of 24 MIRU loci (MIRU02, MTUB04, ETRC, MIRU04, MIRU40, MIRU10, MIRU16, MTUB21, MIRU20, QUB11B, ETRA, Mtub29, Mtub30, ETRB, MIRU23, MIRU24, MIRU26, MIRU27, Mtub34, MIRU31, Mtub39, QUB26, QUB4156, MIRU39) directly from long sequence reads [9].

HGDI calculation

$$HGDI = 1 - \frac{1}{N(N-1)}\sum_{j=1}^{s} nj(j-1)$$

N stands for the total number of strains, nj is the number of strains with the jth genotype, and s is the number of different genotypes at the MIRU-VNTR loci.

Statistical analysis
IBM SPSS 18.0 and GraphPad 7 were implemented for statistical analysis. Chi-square test or t-test was conducted to compare the differences in variables of general characteristics between TB groups with mutational and non-mutational *rpoB* gene. All variables with a *P*-value <0.10 on Chi-Square test and t-test were included in a multivariate conditional logistic regression model to investigate the relationship of the mutation of *rpoB* gene and the repetitive number of the MIRU loci. Besides, Cochran-Armitage analysis was conducted to determine

Su *et al. BMC Genomics*   (2023) 24:26

Page 3 of 6
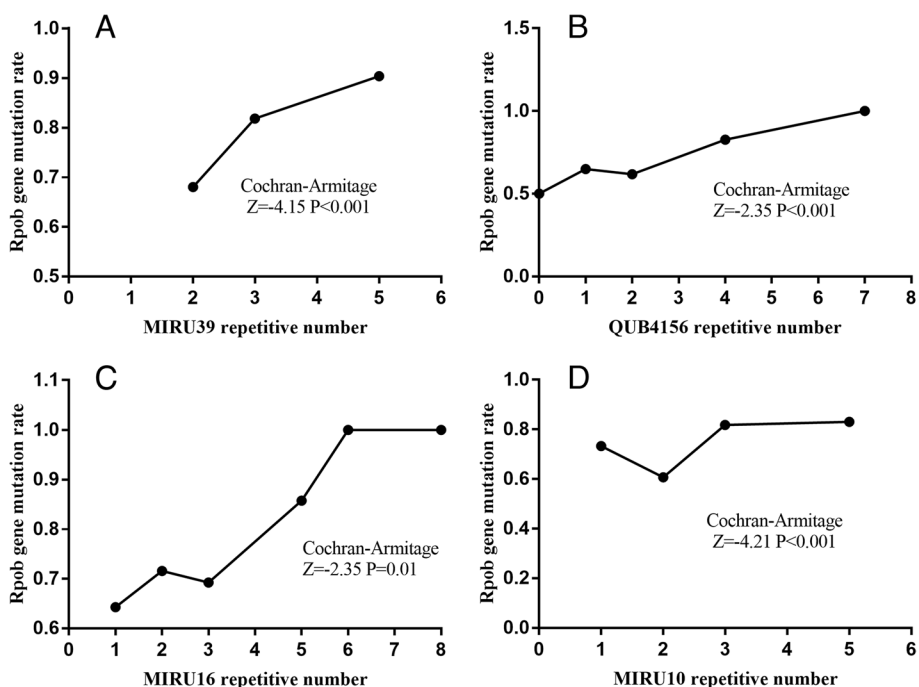
Fig. 1 Cochran-Armitage analysis of *rpoB* gene mutation rate and MIRU repeated numbers. The abscissa of the black dot on the broken line represents the repetitive number of different MIRU loci, and the ordinate represents the corresponding *rpoB* gene mutation rate. The broken line trend reflects whether the *rpoB* mutation rate increases or decreases with the increase of MIRU loci repetitive number. When *P* < 0.05, there was a significant overall trend between them

the trend of *rpoB* gene mutation rate and the repetitive number of MIRU loci.

Results

The rpoB mutation results of the study samples

There were 424 TB samples included in our study, 357 (84.2%) strains extracted from them were RR-TB, and 67 (15.8%) were rifampicin sensitive strains. Among RR-TB, *rpoB* genes of 304 strains were detected mutations, and 6 of 67 susceptible strains were detected mutations. *RpoB* gene mutational rate between strains of retreated cases (83.41%) and new cases (62.32%) showed a significant difference ($\chi^2 = 24.0$ $P < 0.05$).

Relation between rpoB gene and 24-loci MIRU-VNTR

The mutation rate of *rpoB* gene showed an upward trend with the increase of MIRU10, MIRU39, QUB4156 and MIRU16 repetitive number after the Cochran-Armitage analysis (Fig. 1). However, only the repetitive number of MIRU10, MRIU39 and QUB4156 were risk factors for *rpoB* gene mutation after adjusted by category (retreated or new cases) and MIRU23 (Table 1).

Allelic diversity of the MIRU loci

As shown in the Table 2, two loci (MIRU10, QUB4156) were highly discriminative (Hunter-Gaston discriminatory index, HGDI > 0.6), two loci (MIRU39, MIRU16) were moderately discriminative (HGDI > 0.3) among all

Table 1 Logistic regression analysis of the *rpoB* gene mutation and the repetitive number of 24 MIRU loci

Variables	Unadjusted			Adjusted*		
	OR	*95%CI*	*P*	*OR*	*95%CI*	*P*
MIRU10	1.409	1.197–1.658	<.001	1.332	1.114–1.591	0.002
MIRU16	1.314	1.042–1.658	0.021	1.101	0.846–1.433	0.475
QUB4156	1.403	1.195–1.648	<.001	1.400	1.057–1.856	0.019
MIRU39	1.656	1.284–2.137	<.001	1.257	1.052–1.501	0.012

Note*adjusted for category (new or retreated case), MIRU23

Su *et al. BMC Genomics*   (2023) 24:26

Page 4 of 6

Table 2 Allelic diversity of four loci in *rpoB* gene mutation and non-mutation isolates

Loci	HGDI	Alleles		Most frequently repeat		Alleles diversity	
		M	N-M	M	N-M	M	N-M
MIRU10	0.65	4	4	5	2	0.65	0.56
MIRU39	0.39	3	3	2	2	0.45	0.20
QUB4156	0.62	5	4	4	2	0.59	0.66
MIRU16	0.43	6	4	2	2	0.45	0.38

Notes: *M rpoB* gene mutational strains, *N-M rpoB* gene non-mutational strains, *HGDI* Hunter-Gaston discriminatory index

24 loci studied. The allelic diversity of the 4 loci were different between the *rpoB* gene mutational strains and non-mutational strains. It was worth noting that MIRU39 showed a moderately discriminablility in *rpoB* gene mutational strains, while a low discriminablility in *rpoB* gene non-mutational strains.

Discussion

In this study, we assessed the associated risk factors for *rpoB* gene mutation in data sourced areas. The *rpoB* gene mutation rate of retreated TB patients (83.41%) was higher than that of new cases (62.32%), it has been proved that the RIF resistance rate of retreated tuberculosis is higher than that of new cases in previous studies [13, 14], the higher rate may since that patients with retreated pulmonary tuberculosis often fail in the initial treatment due to unreasonable or irregular anti-tuberculosis treatment, resulting in the dominant growth of drug-resistant tuberculosis bacteria, and it's drug resistance mechanism is related to the mutation of *rpoB* gene which coding RNA polymerase β-subunit [3]. Notably, we found no *rpoB* gene mutation in partial RR-TB strains, but mutations in sensitive strains, the inconsistency between gene resistance and phenotype resistance may be caused by heterogeneity of *Mtb*. The presence of low-frequency RR-TB and the predominance of sensitive *Mtb* in the specimen may result in ineffective extraction of drug-resistant DNA if the specimen is not handled properly, while the proportional method of drug sensitivity suggested that it was RR-TB [15]. Patients may have been treated with multiple anti-tuberculosis drugs before sputum specimens were sent for testing, resulting in multiple *Mtb* states in sputum specimens, which can also lead to this result [16]. And mutations in the *rpoB* gene leading to low levels of rifampicin resistance may be the reason that these strains with mutations in the *rpoB* gene were detected as sensitive [17].

Different VNTR loci always has different discrimination ability between Beijing and non-Beijing genotype Mtb [18]. In our study, MIRU10, MIRU39, QUB4156 and MIRU16 all showed a difference in allellic diversity between the Beijing and non-Beijing genotype strains, but only MIRU39 showed remarkable difference ($\triangle$HGDI > 0.2).

VNTR is a highly polymorphic and highly repetitive DNA fragment, which is characterized by variety and wide distribution. The distribution of VNTR in *Mtb* showed high individual specificity [19]. In recent years, MIRU-VNTR had been widely used in the typing of tuberculosis, some loci, such as MIRU10, MIRU39 and QUB4156 could genotype *Mtb* with high discriminatory power [20–22]. In this study, we found that strains with high MIRU10, MIRU39 QUB4156 or MIRU16 repetitive numbers may often have a high *rpoB* gene mutation rate, but only the repetitive number of MIRU10, MRIU39 and QUB4156 were risk factors for *rpoB* gene mutation after adjusting by category (retreated or new cases) and MIRU23.

MIRU loci are located in the spacer of DNA coding genes, and their specific functions are not clear. Some scholars [23–25] believed that the difference in the copy number of MIRU sites upstream of the coding gene will lead to the difference in the number of ribosomal binding sites (RBS), thus affecting the transcription and expression level of the gene. The coding product encoded by the *fadB* gene downstream of MIRU10 is an oxidoreductase that binds to flavin adenine dinucleotide (FAD), the oxidative stress response induced by this gene may be one of the mechanisms of anti-tuberculosis drugs killing bacteria [26]. With the increase of MIRU10 loci repetitive number, it may increase the inhibition of *fadB* gene expression [27], finally resulting in RIF resistance. *EccCa1* gene which downstream of MIRU39 is part of the ESX-1 specialized secretion system, which delivers several virulence factors to host cells during infection, including the key virulence factors ESAT-6 and CFP-10 [28, 29]. The increase of MIRU39 repetitive number may target up-regulation of *eccCa1* gene expression, resulting in increased bacterial virulence. The coding product encoded by the murT gene downstream of QUB4156 is involved in the pathway peptidoglycan biosynthesis, which is part of cell wall biogenesis [30]. The increase of QUB4156 repetitive number may enhance the virulence of *Mtb* by promoting the synthesis of cell wall.

Su *et al. BMC Genomics*   (2023) 24:26

Page 5 of 6

The mutation rate of *rpoB* gene increased with the addition of MIRU10, MRIU39 and QUB4156 repetitive numbers, we speculated that these MIRU loci caused the RIF resistance in Mtb respectively through different ways. Repetitive numbers of MIRU loci are relatively easy to detect in the laboratory [31]. We hope that this upward trend can deepen the understanding of the function of MIRU10, QUB4156 and MIRU39 loci and the mechanism of RIF resistance. However, this experiment is limited to the research characteristics of molecular epidemiology, which need to be further verified by experimental research.

Conclusion
The mutation rate of *rpoB* gene increased with the addition of the number of repeats at MIRU10, QUB4156 and MIRU39 loci.

Declarations

References
1. Caminero Luna JA, Pérez Mendoza G, Rodríguez de Castro F. Multi-drug resistant tuberculosis, ten years later. Tuberculosis multirresistentediez años después. Med Clin (Barc). 2021;156(8):393–401. https://doi.org/10.1016/j.medcli.2020.08.018.
2. WHO, Global Tuberculosis Report 2020. (https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf)，20220603.
3. Zaw MT, Emran NA, Lin Z. Mutations inside rifampicin-resistance determining region of rpoB gene associated with rifampicin-resistance in mycobacterium tuberculosis. J Infect Public Health. 2018;11(5):605–10. https://doi.org/10.1016/j.jiph.2018.04.005.
4. Portelli S, Myung Y, Furnham N, Vedithi SC, Pires DEV, Ascher DB. Prediction of rifampicin resistance beyond the RRDR using structure-based machine learning approaches. Sci Rep. 2020;10(1):18120. https://doi.org/10.1038/s41598-020-74648-y.
5. Uddin MKM, Rahman A, Ather MF, et al. Distribution and frequency of rpoB mutations detected by Xpert MTB/RIF assay among Beijing and non-Beijing rifampicin resistant mycobacterium tuberculosis isolates in Bangladesh. Infect Drug Resist. 2020;10(13):789–97. https://doi.org/10.2147/IDR.S240408.
6. Chee CBE, Lim LKY, Ong RTH, et al. Whole genome sequencing analysis of multidrug-resistant tuberculosis in Singapore, 2006-2018. Eur J Clin Microbiol Infect Dis. 2021;40(5):1079–83. https://doi.org/10.1007/s10096-020-04100-6.
7. Shi J, Zheng D, Zhu Y, et al. Role of MIRU-VNTR and spoligotyping in assessing the genetic diversity of mycobacterium tuberculosis in Henan Province, China. BMC Infect Dis. 2018;18(1):447. https://doi.org/10.1186/s12879-018-3351-y.
8. Gupta A, Sinha P, Nema V, et al. Detection of Beijing strains of MDR M. tuberculosis and their association with drug resistance mutations in katG, rpoB, and embB genes. BMC Infect Dis. 2020;20(1):752. https://doi.org/10.1186/s12879-020-05479-5.
9. Tang CY, Ong RT. MIRUReader: MIRU-VNTR typing directly from long sequencing reads. Bioinformatics. 2020;36(5):1625–6. https://doi.org/10.1093/bioinformatics/btz771.
10. Maladan Y, Krismawati H, Wahyuni T, et al. The whole-genome sequencing in predicting mycobacterium tuberculosis drug susceptibility and resistance in Papua, Indonesia. BMC Genomics. 2021;22(1):844. https://doi.org/10.1186/s12864-021-08139-3.
11. Huang H, Ding N, Yang T, et al. Cross-sectional whole-genome sequencing and epidemiological study of multidrug-resistant mycobacterium tuberculosis in China. Clin Infect Dis. 2019;69(3):405–13. https://doi.org/10.1093/cid/ciy883.
12. Walker TM, Kohl TA, Omar SV, et al. Whole-genome sequencing for prediction of mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. Lancet Infect Dis. 2015;15(10):1193–202. https://doi.org/10.1016/S1473-3099(15)00062-6.
13. Liu L, Zhao X, Wu X, et al. Prevalence and molecular characteristics of drug-resistant mycobacterium tuberculosis in Hainan, China: from 2014 to 2019. BMC Microbiol. 2021;21(1):185. https://doi.org/10.1186/s12866-021-02246-7.
14. Li D, Song Y, Yang P, Li X, Zhang AM, Xia X. Genetic diversity and drug resistance of mycobacterium tuberculosis in Yunnan, China. J Clin Lab Anal. 2019;33(5):e22884. https://doi.org/10.1002/jcla.22884.
15. Zheng C, Li S, Luo Z, et al. Mixed infections and rifampin Heteroresistance among mycobacterium tuberculosis clinical isolates. J Clin Microbiol. 2015;53(7):2138–47. https://doi.org/10.1128/JCM.03507-14.
16. Sun G, Luo T, Yang C, Dong X, Li J, Zhu Y, et al. Dynamic population changes in mycobacterium tuberculosis during acquisition and fixation of drug resistance in patients. J Infect Dis. 2012;206(11):1724–33. https://doi.org/10.1093/infdis/jis601.
17. Gopie FA, Commiesie E, Baldi S, et al. Should treatment of low-level rifampicin mono-resistant tuberculosis be different? J Clin Tuberc Other Mycobact Dis. 2021;23:100222. https://doi.org/10.1016/j.jctube.2021.100222.
18. Li Y, Pang Y, Zhang T, Xian X, Yang J, Wang R, et al. Genotypes of mycobacterium tuberculosis isolates circulating in Shaanxi Province, China. PLoS One. 2020;15(12):e0242971. https://doi.org/10.1371/journal.pone.0242971.
19. Al-Mahrouqi S, Ahmed R, Al-Azri S, Al-Hamidhi S, Balkhair AA, Al-Jardani A, et al. Dynamics of mycobacterium tuberculosis lineages in Oman, 2009 to 2018. Pathogens. 2022;11(5):541. https://doi.org/10.3390/pathogens11050541.
20. Elsayed MSAE. Applicability of using 15 MIRU-VNTR loci for genotyping of Mycobacterium avium subsp. paratuberculosis from two cattle farms

Su *et al. BMC Genomics*   (2023) 24:26

Page 6 of 6

in Egypt. Mol Biol Rep. 2019;46(6):6253–62. https://doi.org/10.1007/s11033-019-05065-2.

21. Antonenko PB, Kresyun VI, Antonenko KO. Clusters of mycobacterium tuberculosis genotypes in Odesa region. Mikrobiol Z. 2016;78(2):103–10. Russian 30141603.

22. Ghavidel M, Tadayon K, Mosavari N, et al. Introducing the best six loci in mycobacterial interspersed repetitive unit-variable-number tandem repeat (MIRU-VNTR) typing for mycobacterium tuberculosis genotyping. Rep Biochem Mol Biol. 2019;8(3):335–46. https://doi.org/10.1186/s12879-020-05479-5.

23. 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):1352–9. https://doi.org/10.1099/jmm.0.044453-0.

24. Yu-Feng W, Chao J, Xian-Feng C. Drug-resistant tuberculosis can be predicted by mycobacterial interspersed repetitive unit locus. Front Microbiol. 2015;6:147. https://doi.org/10.3389/fmicb.2015.00147.

25. Cheng XF, Jiang C, Zhang M, et al. Mycobacterial interspersed repetitive unit can predict drug resistance of mycobacterium tuberculosis in China. Front Microbiol. 2016;7:378. https://doi.org/10.3389/fmicb.2016.00378.

26. Cox JAG, Taylor RC, Brown AK, Attoe S, Besra GS, Fütterer K. Crystal structure of mycobacterium tuberculosis FadB2 implicated in mycobacterial β-oxidation. Acta Crystallogr D Struct Biol. 2019;75(Pt 1):101–8. https://doi.org/10.1107/S2059798318017242.

27. Pérez-Lago L, Navarro Y, Herranz M, Bouza E, García-de-Viedma D. Differences in gene expression between clonal variants of mycobacterium tuberculosis emerging as a result of microevolution. Int J Med Microbiol. 2013;303(8):674–7. https://doi.org/10.1016/j.ijmm.2013.09.010.

28. Clemmensen HS, Knudsen NPH, Rasmussen EM, et al. An attenuated mycobacterium tuberculosis clinical strain with a defect in ESX-1 secretion induces minimal host immune responses and pathology. Sci Rep. 2017;7:46666. https://doi.org/10.1038/srep46666.

29. Soler-Arnedo P, Sala C, Zhang M, Cole ST, Piton J. Polarly localized EccE1 is required for ESX-1 function and stabilization of ESX-1 membrane proteins in mycobacterium tuberculosis. J Bacteriol. 2020;202(5):e00662–19. https://doi.org/10.1128/JB.00662-19.

30. Maitra A, Nukala S, Dickman R, et al. Characterization of the MurT/GatD complex in mycobacterium tuberculosis towards validating a novel antitubercular drug target. JAC Antimicrob Resist. 2021;3(1):dlab028. https://doi.org/10.1093/jacamr/dlab028.

31. Shafipour M, Shirzad-Aski H, Ghaemi EA, et al. Mycobacterium tuberculosis typing using allele-specific oligonucleotide multiplex PCR (ASO-PCR) method. Curr Microbiol. 2021;78(12):4009–13. https://doi.org/10.1007/s00284-021-02659-7.

Publisher's Note