



THE AGA KHAN UNIVERSITY

eCommons@AKU

Department of Pathology and Laboratory Medicine

Medical College, Pakistan

March 2015

Characterization of genomic variations in SNPs of PE_PGRS genes reveals deletions and insertions in extensively drug resistant (XDR) *M. tuberculosis* strains from Pakistan.

Akber Kanji

Aga Khan University, akbar.kanji@aku.edu

Zahra Hasan

Aga Khan University, zahra.hasan@aku.edu

Ruth Mcnerney

London School of Hygiene and Tropical Medicine

Mridul Nair

Sydney Medical School, University of Sydney

Asho Ali

Aga Khan University

Follow this and additional works at: http://ecommons.aku.edu/pakistan_fhs_mc_pathol_microbiol



Part of the [Microbiology Commons](#), and the [Pathology Commons](#)

Recommended Citation

Kanji, A., Hasan, Z., Mcnerney, R., Nair, M., Ali, A., Mallard, K., Clark, T. G., Zaver, A., Jafri, S., Hasan, R. (2015). Characterization of genomic variations in SNPs of PE_PGRS genes reveals deletions and insertions in extensively drug resistant (XDR) *M. tuberculosis* strains from Pakistan.. *Int J Mycobacteriol*, 4(1), 73-79.

Available at: http://ecommons.aku.edu/pakistan_fhs_mc_pathol_microbiol/511

Authors

Akber Kanji; Zahra Hasan; Ruth Mcnerney; Mridul Nair; Asho Ali; Kim Mallard; Taane G. Clark; Ambreen Zaver; Sana Jafri,; and Rumina Hasan

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/271274731>

Characterization of genomic variations in SNPs of PE_PGRS genes reveals deletions and insertions in extensively drug resistant...

Article in *International Journal of Mycobacteriology* · January 2015

DOI: 10.1016/j.ijmyco.2014.11.049

CITATIONS

0

READS

66

12 authors, including:



Akbar Kanji

Aga Khan University, Pakistan

14 PUBLICATIONS 188 CITATIONS

[SEE PROFILE](#)



Zahra Hasan

Aga Khan University, Pakistan

88 PUBLICATIONS 1,249 CITATIONS

[SEE PROFILE](#)



Ruth Mcnerney

London School of Hygiene and Tropical Medi...

137 PUBLICATIONS 2,856 CITATIONS

[SEE PROFILE](#)



Mridul Nair

Biopharmaceutical New Technologies

37 PUBLICATIONS 220 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



DRS pAKISTAN [View project](#)



Molecular mechanisms of drug resistance in Mycobacterium tuberculosis [View project](#)

All content following this page was uploaded by [Grant A Hill-Cawthorne](#) on 29 January 2015.

The user has requested enhancement of the downloaded file. All in-text references [underlined in blue](#) are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

Accepted Manuscript

Characterization of genomic variations in SNPs of PE_PGRS genes reveals deletions and insertions in extensively drug resistant (XDR) *M. tuberculosis* strains from Pakistan

Akbar Kanji, Zahra Hasan, Asho Ali, Ruth McNerney, Kim Mallard, Francesc Coll, Grant Hill-Cawthorne, Mridul Nair, Taane G Clark, Ambreen Zaver, Sana Jafri, Rumina Hasan

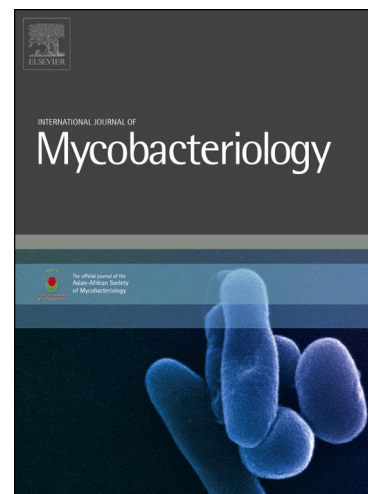
PII: S2212-5531(15)00012-6
DOI: <http://dx.doi.org/10.1016/j.ijmyco.2014.11.049>
Reference: IJMYCO 277

To appear in: *International Journal of Mycobacteriology*

Received Date: 27 October 2014
Accepted Date: 2 November 2014

Please cite this article as: A. Kanji, Z. Hasan, A. Ali, R. McNerney, K. Mallard, F. Coll, G. Hill-Cawthorne, M. Nair, T.G. Clark, A. Zaver, S. Jafri, R. Hasan, Characterization of genomic variations in SNPs of PE_PGRS genes reveals deletions and insertions in extensively drug resistant (XDR) *M. tuberculosis* strains from Pakistan, *International Journal of Mycobacteriology* (2015), doi: <http://dx.doi.org/10.1016/j.ijmyco.2014.11.049>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Characterization of genomic variations in SNPs of PE_PGRS genes reveals deletions and insertions in extensively drug resistant (XDR) *M. tuberculosis* strains from Pakistan

Akbar Kanji¹, Zahra Hasan^{1#}, [Asho Ali¹](#), Ruth McNerney², Kim Mallard², Francisc Coll², Grant Hill-Cawthorne^{3,4}, Mridul Nair³, Taane G Clark², Ambreen Zaver¹, Sana Jafri¹, Rumina Hasan¹

¹Aga Khan University, Karachi Pakistan

²London School of Hygiene and Tropical Medicine (LSHTM)

³Pathogen Genomics Laboratory, King Abdullah University of Science and Technology (KAUST), Thuwal, Kingdom of Saudi Arabia

⁴Sydney Emerging Infections and Biosecurity Institute and School of Public Health, Sydney Medical School, University of Sydney, NSW 2006, Australia

Running Title: Genomic variations in SNPs of PE_PGRS genes in XDR MTB strains

#Corresponding author
Zahra Hasan, PhD
Professor
Department of Pathology and Microbiology
The Aga Khan University
Karachi, Pakistan
Email.zahra.hasan@aku.edu
Tel: + 92 21 3486 4527

ABSTRACT

Background: *Mycobacterium tuberculosis* (MTB) PE_PGRS genes belong to the PE multigene family. Although the function of PE-PGRS genes is unknown, it is hypothesized that the PE_PGRS genes may be associated with antigenic variability in MTB.

Material and methods: Whole Genome Sequencing analysis was performed on (n=37) extensively drug-resistant (XDR) MTB strains from Pakistan, which included Lineage 1 (East African Indian, n=2); Other lineage 1 (n=3); Lineage 3 (Central Asian, n=24); Other lineage 3 (n=4); Lineage 4 (X3, n=1) and T group (n=3) MTB strains.

Results: There were 107 SNPs identified from the analysis of 42 PE_PGRS genes; of these, 13 were non-synonymous SNPs (nsSNPs). The nsSNPs identified in PE_PGRS genes–6, 9 and 10– were common in all EAI, CAS, Other lineages (1 and 3), T1 and X3. Deletions (DELs) in PE_PGRS genes–3 and 19– were observed in 17 (80.9%) CAS1 and 6 (85.7%) in Other lineage (1 and 3) XDR MTB strains, while DELs in the PE_PGRS49 were observed in all CAS1, CAS, CAS2 and Other lineage (1 and 3) XDR MTB strains. All CAS, EAI and Other lineage (1 and 3) strains showed insertions (INS) in PE_PGRS6 gene, while INS in the PE_PGRS genes 19 and 33 were observed in 20 (95.2%) CAS1, all CAS, CAS2, EAI and Other lineage (1 and 3) XDR MTB strains.

Conclusion: Genetic diversity in PE_PGRS genes contributes to antigenic variability and may result in increased immunogenicity of strains. This is the first study identifying variations in nsSNPs and INDELs in the PE_PGRS genes of XDR-TB strains from Pakistan. It highlights common genetic variations which may contribute to persistence.

1. Introduction

Mycobacterium tuberculosis (MTB), the causative agents of tuberculosis (TB) kills approximately 2 million people worldwide each year [1]. The control of TB requires a better understanding of the mechanisms that allow MTB to evade the immune system and remain persistent with the host. The sequencing of the genomes of MTB strains has provided important insights into possible mechanisms of persistence within the host, along with the discovery of ~60 genes named PE_PGRS (Proline glutamic acid polymorphic GC rich repetitive sequence), which belong to the PE multigene family [2]. These multigene families represent approximately 10% of the coding capacity of the MTB genome and are characterized by their high GC content and repetitive sequences. This may be due to the high frequency of gene duplication, recombination and strand slippage mechanisms [2, 3].

The single nucleotide polymorphisms (SNPs) have been shown to be the most common form of genetic variation in MTB complex (MTBC). The SNPs have been shown to be of low frequency and limited horizontal gene transfer (HGT), which have resulted in low levels of independent occurrence of the same SNP in phylogenetically unrelated strains [4, 5]. Depending on their position in the genome, SNPs can be either coding or non-coding. The coding density in the MTBC has been reported to be 90%-96%, with most of the SNPs in the MTBC in the coding region of the genome [6]. SNPs which results in an amino acid change are non-synonymous (nsSNPs) while those which do not cause a change in amino acid coding are synonymous (sSNPs).

While the function of the members of the PE_PGRS multigene family is not known, it is hypothesized that the PE_PGRS genes may be associated with antigenic and genetic variability

as well as virulence [7-10]. Previous studies have been shown that some members of the PE_PGRS family are expressed on the cell surface during MTB infection and recognized by the host immune system [7, 11, 12]. It has been suggested that MTB PE_PGRS genes are variably expressed during infection [13-15]. The roles of PE_PGRS16 and PE_PGRS26 in causing infection have previously been studied in mouse models; the expression of PE_PGRS16 was significantly up-regulated as compared with PE_PGRS26, suggesting that these two PE_PGRS genes may serve as a marker of latent TB infection [14]. In addition, PE_PGRS33 gene has been shown to encode for a surface expressed protein, while transposon mutagenesis-based studies have shown that the MTB PE_PGRS33 gene plays an important role in interactions with other mycobacteria as well as with macrophages [16].

The sequence variations such as SNPs, insertions and deletions (INDELs) in MTB clinical strains have also been investigated for PE_PGRS16, PE_PGRS26 and PE_PGRS33 [17, 18]. Moreover, in a population-based study of 649 clinical MTB strains, patients infected with MTB strains with a significant change in the PE_PGRS33 protein were more likely to belong to a cluster of TB cases as compared with MTB strains with minimal changes to the PE_PGRS33 protein suggests that PE_PGRS33 may play an important role in MTB transmission [19].

The extensively drug-resistant tuberculosis (XDR-TB) has emerged worldwide as one of the biggest threats to public health and TB control programs. XDR-TB is defined as TB caused by the MTB strain that is resistant to at least rifampin (RIF) and isoniazid (INH) among the first-line anti-TB drugs and resistant to fluoroquinolones and to at least one of the three injectable second-line drugs [20-22].

Previous studies have reported SNPs in MTBC as drug resistance-conferring mutations, with 1,447 mutations significant for most anti-TB drugs [23-26]. However, no data are available on sequence variations such as SNPs and INDELs in the PE_PGRS genes of XDR MTB isolates. This study analyzed 37 XDR MTB strains for the presence of SNPs, INDELs by whole genome sequencing (WGS) analysis method. The aim of this study is to specifically investigate SNPs and INDELs in XDR MTB strains to understand how genetic diversity contributes to antigenic variability.

2. Methodology

2.1 Strain selection

The MTB strains were obtained from the strain bank of Aga Khan University Clinical Microbiology Laboratory, Pakistan. The 37 XDR MTB strains were from the period of 2004-2009, which had previously been spoligotyped [27-29], and were randomly selected for the WGS analysis. The study samples selected for the WGS analysis belonged to all three Principal Genetic Groups: PGG1, PGG2 and PGG3, which were comprised of PGG1: 21 CAS1, 2 CAS, 1 CAS2, 2 EAI and 7 Other lineages (1 and 3) (89.2%); PGG2: 1 X3 (2.7%); and PGG3: 3 T (8.1%) groups, respectively.

2.2 Culture and antibiotic susceptibility testing

The XDR MTB strains used had been isolated from the specimens using Lowenstein-Jensen media and MGIT (Becton Dickinson, Franklin Lakes, NJ, and USA). MTB was identified using BACTEC NAP TB differentiation test (Becton Dickinson), growth on para-nitrobenzoic acid

containing media, nitrate reduction, and niacin accumulation [30, 31]. The drug susceptibility testing (DST) of these isolates was previously performed using an agar proportion method on enriched Middlebrook 7H10 medium (BBL Microbiology Systems, Cockeysville, MD, USA) at the following concentrations: rifampicin 1 µg/mL, isoniazid 0.2 µg/mL, streptomycin 2 µg/mL and 10 µg/mL, and ethambutol 5 µg/mL. Pyrazinamide sensitivity was determined by using BACTEC 7H12 medium, pH 6.0, at 100 µg/mL (BACTEC PZA test medium, Becton Dickinson). MDR TB strains were further tested with capreomycin 10 µg/mL, ciprofloxacin 2 µg/mL, ethionamide 5 µg/mL, amikacin 5µg/mL, and kanamycin 6µg/mL. Reference strain MTB H37Rv was used as a control with each susceptibility testing batch [32].

2.3 DNA extraction and spoligotyping

Spoligotyping was performed as described previously [33] and inferred *in silico*; the spoligotypes were also confirmed from the sequencing reads using SpolPred software [34].

2.4 Whole genome sequencing of MTB strains

DNA was extracted by the Cetyl-trimethyl ammonium bromide (CTAB) method [35]. All samples (n=37) underwent WGS with 76-base paired end fragment sizes using Illumina paired end HiSeq2000 technology. Briefly, the raw sequence data were mapped distinctively to the H37Rv reference genome. The resulting alignments allowed SNPs and small INDELS to be called using SAMtools/BCFtools (<http://samtools.sourceforge.net>), as well as larger INDELS using a consensus from paired end mapping distance or split read approaches 21-24.

Only those variants of high quality (at least Q30, equating to 1 error per 1000) and supported by bi-directional reads were retained. Spoligotypes were inferred from the sequencing reads SpolPred software [34]. The raw data files were analyzed using R to determine the presence of PE_PGRS INDELS in genomes of 37 XDR MTB strains.

3. Results

3.1 Detection of number of copies of PE_PGRS genes in the XDR MTB strains

Forty-two PE_PGRS genes were analyzed in 37 XDR MTB strains. The numbers of copies of PE_PGRS genes were variable and show no significant differences amongst the CAS1, CAS, CAS2, EAI, Other (lineages 1 and 3), T and X3 XDR MTB strains ($P>0.05$) analyzed by Pearson's Chi-square test (**Table 1**).

3.2 Detection of SNPs in PE_PGRS genes of XDR MTB strains

A total of 107 SNPs were observed, including 13 nsSNPs in 37 XDR MTB strains. All CAS1, CAS, CAS2, EAI, Other (lineages 1 and 3) and X3 XDR MTB strains showed nsSNPs in the PGRS genes 6, 7, 9 and 10. Whereas 20 (95%) CAS1 XDR MTB strains showed nsSNPs in PE_PGRS17 and PE_PGRS37; 18 (85.7%) in PE_PGRS55 genes. However, all CAS, CAS2, EAI, Other (lineages 1 and 3), T1 and X3 XDR MTB strains showed nsSNPs in PE_PGRS17 and 37 genes. Additionally, all CAS2, EAI, T1 and X3 XDR MTB strains revealed nsSNPs in the PE_PGRS47 gene (**Table 2**).

3.3 Detection of deletions in the PE_PGRS genes of XDR MTB strains

Thirty-nine PE_PGRS genes (92.8%) were observed to have variable DELs in the CAS1, CAS, CAS2, EAI, Orphan, T1 and X3 XDR MTB strains. However, all CAS1, CAS, CAS2 and Other (lineages 1 and 3) XDR MTB strains were observed to have 1-36bp DEL in PE_PGRS49, while all CAS, CAS2 and Other (lineages 1 and 3) XDR MTB strains showed 1-18bp DEL in PE_PGRS50 gene sequences. The 6-36 bps DEL in the PE_PGRS19 gene was observed in 17

(81.0%) CAS1, 6 (85.7%) Other lineage (1 and 3), all CAS, CAS2 and EAI XDR MTB strains. The DELs in the PE_PGRS5 and PE_PGRS49 genes were significantly more ($p=0.0002$, $p=0.0216$) in XDR CAS1 compared with XDR other lineages (T1 and X3), respectively, analyzed by Pearson's Chi-square test (**Table 3**).

3.4 Detection of insertions in the PE_PGRS genes of XDR MTB strains

Thirty-nine PE_PGRS genes (92.8%) were observed to have variable INS in the CAS1, CAS, CAS2, EAI, Other lineage (1 and 3), T1 and X3 XDR MTB strains. The INS in the PE_PGRS51 gene was present in 35 (94.6%) with a 9bp INS, while PE_PGRS53 INS was present in 36 (97.3%) of XDR MTB strains with INS size range of 9-18bps. The INS in the PE_PGRS genes 6, 19, 28, 30, 33, 50, 57 and 61 were variably present in the XDR MTB strains with INS size ranging from 1-21bps.

The INS were also observed to be significantly more in PE_PGRS6 ($p=0.05$) and PE_PGRS30 ($p=0.0447$) in XDR CAS1 as compared with XDR other lineages strains (T1 and X3) groups analyzed by Pearson's Chi-square test (**Table 4**).

3.5 Genetic diversity amongst the XDR MTB Principal Genetic Groups (PGG)

Of the 42 PE_PGRS genes investigated for detection of 37 XDR MTB strains by WGS, 34 (91.90%), 1 (2.7%) and 3 (8.1%) of the 37 MTB strains belonged to PGG1, PGG2 and PGG3, respectively. CAS1, CAS CAS2, EAI and Orphan belong to PGG1, while X3 and T1 belong to the PGG2 and PGG3 groups, respectively. There were significantly more nsSNPs present in

PGG1 strains (n=34) as compared with PGG2 (n=1) and PGG3 (n=2) strains (**Figure 1**). However, given the larger proportion of PGG1 strains in this sample size of XDR strains as compared with others, it is difficult to get a complete picture for this comparison.

4. Discussion

The PE_PGRS genes have been shown to be limited to mycobacterial species; they have specially expanded within the genomes of pathogenic mycobacteria, probably through widespread gene duplication events and genetic divergence during their adaptation to the intramacrophage environment [9].

It has been hypothesized that gene conversion may have contributed to the evolution of members of the PE_PGRS subfamily and may have participated in the generation of antigenic variation in their members [2, 7, 17]. A previous study has shown nucleotide substitution of A to G in the PE_PGRS17 genes of CAS, EAI and LAM strains. In addition, SNPs of G to A, C to A and G to C nucleotide substitutions were observed in PE_PGRS18 genes of CAS, EAI, LAM and Haarlem MTB strains [36].

The PGG1 is considered to be the most ancient of the MTB lineages, which has also been reported for other regions of Asia [37]. In this study, it was observed that all XDR MTB strains belonging to the PGG1 group had nsSNPs in the PE_PGRS6, 7, 9, 10, 13 and 32 genes, whereas it was variable in the PGG2 and PGG3 XDR MTB strains. However, the large proportion of PGG1 strains in this study as compared with PGG2 and PGG3 strains makes it difficult to get a complete picture of the proportion of SNPs. In this study, the nucleotide substitution of A to G and T to C in the PE_PGRS17 genes of CAS1, CAS, CAS2, EAI, T, X and orphan (lineages 1

and 3) XDR MTB strains was also observed. Moreover, nucleotide substitution of A to G, C to T and A to C in the PE_PGRS18 genes of CAS1, CAS, CAS2, EAI3, T, X and Other lineages (1 and 3) XDR MTB strains was also observed. However, the function of these PE_PGRS genes are yet unknown [36].

In a population-based study of 649 MTB clinical isolates it was shown that in PE_PGRS33 gene, Thr47Ile was found in 1 isolate, but Thr1172Ile was found in 11 (1.7%) isolates, respectively [38]. However, non-synonymous SNPs (nsSNPs) in the PE_PGRS33 genes were not observed in the XDR MTB strains in the present study. Also, previously, it has been observed that a Gly686Asp in PE_PGRS26 genes was observed in 20 (10%) MTB clinical strains belonging to the principal genetic group (PGG2), while Gly1122Gly in the PE_PGRS26 genes was observed in 1 (0.5%) of the MTB clinical strains also belonging to PGG2 [18]. However, this study did not observe nsSNPs in the PE_PGRS26 genes of XDR MTB strains.

In addition to SNPs, deletions have also been shown in the PE_PGRS genes 16, 26 and 33. In the XDR strains studied here, a 9bp DEL was observed in the PE_PGRS16 gene sequence of 2 (25%) XDR Other lineage 1 MTB strains belonging to PGG1 strains. Previously, in the PE_PGRS16 gene sequence, a 3bp deletion in PGG3 group strains and a 252bp deletion in three PGG2 group strains have been reported [18]. Also, in the PE_PGRS26 gene, both 3bp and 150 bp DELs have been reported in PGG2 group strains and PGG1 group strains [18]. In this study, 2 (9.5%) CAS1 MTB strains showed 9bp DELs in the PE_PGRS26 gene belonging to the PGG1 group. It was observed that 1-9bp DELs were present in the PE_PGRS33 gene in EAI (100%)

and Other lineage 1 (28.57%). This correlates with previous work which has 3 bp, 9 bp and 273 bp, respectively, in the PE_PGRS33 gene of MTB strains [38].

In this study, the variable 18bp INS in PE_PGRS26 and 9bp INS PE_PGRS33 was observed. INS of 9bp and 18bp have been reported in the PE_PGRS26 gene [18].

Overall, this work demonstrates the variability in the genome of MTB clinical strains. As these PE_PGRS genes contribute to immunogenicity, the variations in XDR strains may be associated with their ability to persist within the host.

Conclusion: Multiple DELs and INS were observed in the PE_PGRS genes in CAS1 XDR–MTB strains. As these strains are predominant in this endemic region, it indicates a need to further understand the functions of the PE_PGRS genes and its persistence.

Acknowledgement:

We thank the Department of Pathology and Microbiology, Clinical Microbiology, Aga Khan University Hospital for providing support for collection of samples used in the study. Our thanks to Pathogen Genomics Laboratory, King Abdullah University of Science and Technology (KAUST) for whole genome sequencing of MTB strains.

Competing interest: Authors declare that they have no competing interests.

Ethical approval: This work received approval from the Ethical Review Committee of the Aga Khan University.

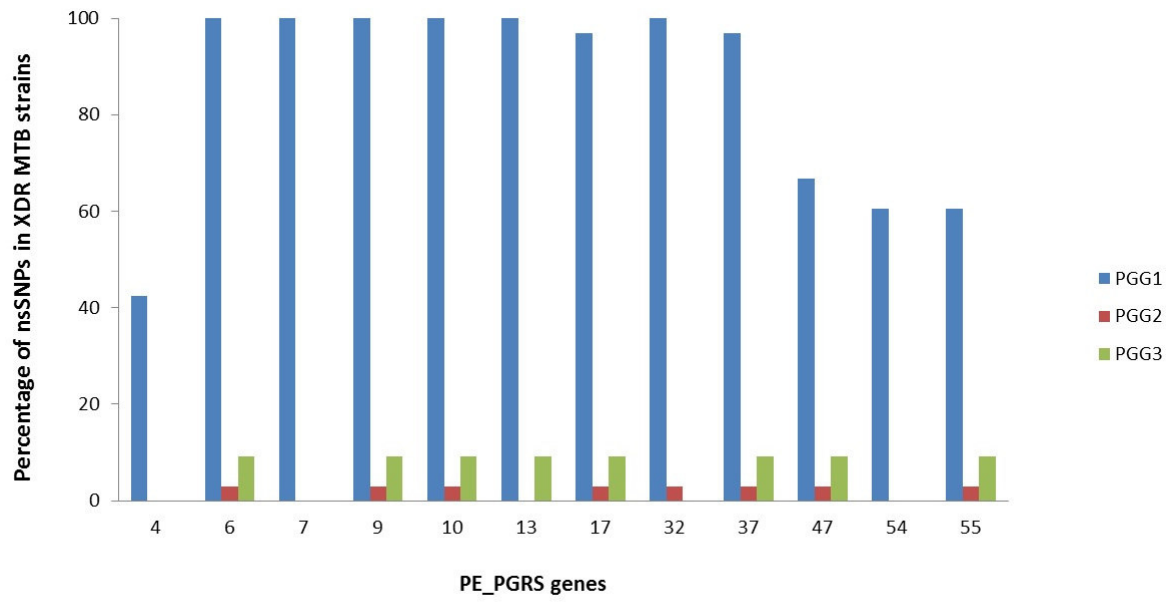
References

1. WHO, The global plan to stop TB. 2006-2015.
2. Cole S.T., et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, 1998. **393**(6685): p. 537-44.
3. Lamichhane G., et al. A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A*, 2003. **100**(12): p. 7213-8.
4. Comas I., et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet*, 2010. **42**(6): p. 498-503.
5. Hershberg R., et al. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol*, 2008. **6**(12): p. e311.
6. Namouchi A., et al. After the bottleneck: Genome-wide diversification of the *Mycobacterium tuberculosis* complex by mutation, recombination, and natural selection. *Genome Res*, 2012. **22**(4): p. 721-34.
7. Banu S., et al. Are the PE-PGRS proteins of *Mycobacterium tuberculosis* variable surface antigens? *Mol Microbiol*, 2002. **44**(1): p. 9-19.
8. Delogu G., et al. Rv1818c-encoded PE_PGRS protein of *Mycobacterium tuberculosis* is surface exposed and influences bacterial cell structure. *Mol Microbiol*, 2004. **52**(3): p. 725-33.
9. Brennan M.J., Delogu G. The PE multigene family: a 'molecular mantra' for mycobacteria. *Trends Microbiol*, 2002. **10**(5): p. 246-9.
10. Ramakrishnan L., Federspiel N.A., Falkow S. Granuloma-specific expression of *Mycobacterium* virulence proteins from the glycine-rich PE-PGRS family. *Science*, 2000. **288**(5470): p. 1436-9.
11. Singh K.K., et al. Antigens of *Mycobacterium tuberculosis* expressed during preclinical tuberculosis: serological immunodominance of proteins with repetitive amino acid sequences. *Infect Immun*, 2001. **69**(6): p. 4185-91.
12. Delogu G., Brennan M.J. Comparative immune response to PE and PE_PGRS antigens of *Mycobacterium tuberculosis*. *Infect Immun*, 2001. **69**(9): p. 5606-11.
13. Voskuil M.I., et al. Regulation of the *Mycobacterium tuberculosis* PE/PPE genes. *Tuberculosis (Edinb)*, 2004. **84**(3-4): p. 256-62.
14. Dheenadhayalan V., et al. Variable expression patterns of *Mycobacterium tuberculosis* PE_PGRS genes: evidence that PE_PGRS16 and PE_PGRS26 are inversely regulated in vivo. *J Bacteriol*, 2006. **188**(10): p. 3721-5.
15. Delogu G., et al. PE_PGRS proteins are differentially expressed by *Mycobacterium tuberculosis* in host tissues. *Microbes Infect*, 2006. **8**(8): p. 2061-7.
16. Brennan M.J., et al. Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells. *Infect Immun*, 2001. **69**(12): p. 7326-33.
17. Talarico S., et al. Variation of the *Mycobacterium tuberculosis* PE_PGRS 33 gene among clinical isolates. *J Clin Microbiol*, 2005. **43**(10): p. 4954-60.
18. Talarico S., et al. *Mycobacterium tuberculosis* PE_PGRS16 and PE_PGRS26 genetic polymorphism among clinical isolates. *Tuberculosis (Edinb)*, 2008. **88**(4): p. 283-94.
19. Chaves F., et al. Usefulness of the secondary probe pTBN12 in DNA fingerprinting of *Mycobacterium tuberculosis*. *J Clin Microbiol*, 1996. **34**(5): p. 1118-23.

20. World Health Organization (WHO). Case definition for extensively drug-resistant tuberculosis, in *Wkly. Epidemiol. Rec* 2006, World Health Organization.
21. [Hasan R. Drug resistant tuberculosis: Challenges of urbanization. *International Journal of Mycobacteriology* 2014. **3**\(2\): p. 79-81.](#)
22. [Macedo R., Antunes A.F. Villar M., Portugal I. Multidrug and extensively drug-resistant tuberculosis in Lisbon and Vale do Tejo, Portugal, from 2008 to 2010, *International Journal of Mycobacteriology*, 2014. **1**\(3\): p. 131–136.](#)
23. [Riska P.F., Jacobs W.R. Jr., Alland D. Molecular determinants of drug resistance in tuberculosis. *Int J Tuberc Lung Dis*, 2000. **4**\(2 Suppl 1\): p. S4-10.](#)
24. [Musser J.M. Antimicrobial agent resistance in mycobacteria: molecular genetic insights. *Clin Microbiol Rev*, 1995. **8**\(4\): p. 496-514.](#)
25. [Telenti A. Genetics of drug resistance in tuberculosis. *Clin Chest Med*, 1997. **18**\(1\): p. 55-64.](#)
26. [Ramaswamy S., Musser J.M. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tuber Lung Dis*, 1998. **79**\(1\): p. 3-29.](#)
27. [Kanji A., et al. Occurrence of RD149 and RD152 deletions in *Mycobacterium tuberculosis* strains from Pakistan. *J Infect Dev Ctries*, 2011. **5**\(2\): p. 106-13.](#)
28. [Ali A., et al. Characterization of mutations conferring extensive drug resistance to *Mycobacterium tuberculosis* isolates in Pakistan. *Antimicrob Agents Chemother*, 2011. **55**\(12\): p. 5654-9.](#)
29. [Asho Ali Z.H., Jafri S., Inayat R., Hasan R. *Mycobacterium tuberculosis* Central Asian Strain \(CAS\) lineage strains in Pakistan reveal lower diversity of MIRU loci than other strains. *International Journal of Mycobacteriology*, 2014. **3**\(2\): p. 108-116.](#)
30. Koneman E.W., Allen S.D., Dowell V.R., Sommers H.M. *Color Atlas and Textbook of Diagnostic Microbiology*. 1983: Publisher; JB Lippincott Company, Philadelphia.
31. [Hasan M., Munshi S.K., Banu Momi M.S., Rahman F., Noor R. Evaluation of the effectiveness of BACTEC MGIT 960 for the detection of mycobacteria in Bangladesh. *International Journal of Mycobacteriology*, 2014. **2**\(4\): p. 214-219.](#)
32. [Woods G.L., Brown-Elliott BA., Conville P.S., Desmond E.P., Hall G.S., Lin G., et al. Susceptibility testing of mycobacteria, nocardiae, and other aerobic actinomycetes; approved standard. Wayne, PA: NCCLS; 2003. Document No M24-A. 2003.](#)
33. [Kamerbeek J., et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol*, 1997. **35**\(4\): p. 907-14.](#)
34. [Coll F., et al. SpolPred: rapid and accurate prediction of *Mycobacterium tuberculosis* spoligotypes from short genomic sequences. *Bioinformatics*, 2012. **28**\(22\): p. 2991-3.](#)
35. [Honore-Bouakline S., et al. Rapid diagnosis of extrapulmonary tuberculosis by PCR: impact of sample preparation and DNA extraction. *J Clin Microbiol*, 2003. **41**\(6\): p. 2323-9.](#)
36. [Karboul A., et al. Insights into the evolutionary history of tubercle bacilli as disclosed by genetic rearrangements within a PE_PGRS duplicated gene pair. *BMC Evol Biol*, 2006. **6**: p. 107.](#)
37. [Arjomandzadegan M., et al. Determination of principal genotypic groups among susceptible, MDR and XDR clinical isolates of *Mycobacterium tuberculosis* in Belarus and Iran. *Tuberk Toraks*. **60**\(2\): p. 153-9.](#)

38. Talarico S., et al. Association of Mycobacterium tuberculosis PE PGRS33 polymorphism with clinical and epidemiological characteristics. Tuberculosis (Edinb), 2007. **87**(4): p. 338-46.

ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT

Table 3. The detection of deletions in PE_PGRS genes of XDR *M. tuberculosis* strains

Strains	PE_PGRS genes																																							
	2	3	4	5	6	7	9	10	13	16	19	21	23	25	26	27	28	29	30	32	33	36	38	42	43	44	46	48	49	50	51	52	53	54	55	56	57	58		
Lineage 1																																								
EAI3																																								
X5																																								
X32																																								
Other lineage 1																																								
X17																																								
X49																																								
X57																																								
Lineage 3																																								
CAS1																																								
X6																																								
X7																																								
X8																																								
X10																																								
X11																																								
X12																																								
X13																																								
X14																																								
X16																																								
X21																																								
X33																																								
X37																																								
X39																																								
X40																																								
X42																																								
X43																																								
X44																																								
X47																																								
X48																																								
X55																																								
X61																																								
CAS																																								
X18																																								
X22																																								
CAS2																																								
X4																																								
Other lineage 3																																								
X1																																								
X9																																								
X45																																								
X60																																								
Lineage 4																																								
T1																																								
X41																																								
X56																																								
X58																																								
X3																																								
X46																																								

CAS1: Central Asian Strain 1, EAI: East African Indian, T1: undefined group, X3: X family

*The shaded portion denotes deletions in the PE_PGRS genes

#The unshaded portion denotes absence of deletions in the PE_PGRS genes

