12-1-2017

Single nucleotide polymorphisms in efflux pumps genes in extensively drug resistant Mycobacterium tuberculosis isolates from Pakistan

Akbar Kanji
Aga Khan University, akbar.kanji@aku.edu

Rumina Hasan
Aga Khan University, rumina.hasan@aku.edu

Asho Ali
King Abdulaziz University, Saudi Arabia

Ambreen Zaver
Aga Khan University

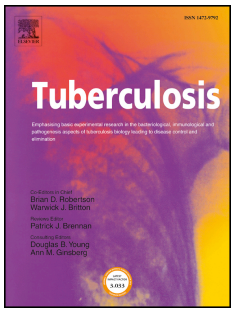
Ying Zhang
Johns Hopkins University, Baltimore, USA.

See next page for additional authors

Authors

Akbar Kanji, Rumina Hasan, Asho Ali, Ambreen Zaver, Ying Zhang, Kehkashan Imtiaz, Wanliang Shi, Taane G. Clark, Ruth McNerney, and Jody Phelan

Accepted Manuscript



Single nucleotide polymorphisms in efflux pumps genes in extensively drug resistant *Mycobacterium tuberculosis* isolates from Pakistan

Akbar Kanji, Rumina Hasan, Asho Ali, Ambreen Zaver, Ying Zhang, Kehkashan Imtiaz, Wanliang Shi, Taane G. Clark, Ruth McNerney, Shoaib Rao, Samreen Shafiq, Zahra Hasan

PII: S1472-9792(16)30553-4

DOI: 10.1016/j.tube.2017.07.012

Reference: YTUBE 1608

To appear in: *Tuberculosis*

Received Date: 16 December 2016

Revised Date: 23 July 2017

Accepted Date: 31 July 2017

Please cite this article as: Kanji A, Hasan R, Ali A, Zaver A, Zhang Y, Imtiaz K, Shi W, Clark TG, McNerney R, Rao S, Shafiq S, Hasan Z, Single nucleotide polymorphisms in efflux pumps genes in extensively drug resistant *Mycobacterium tuberculosis* isolates from Pakistan, *Tuberculosis* (2017), doi: 10.1016/j.tube.2017.07.012.

Single Nucleotide Polymorphisms in Efflux pumps genes in Extensively Drug Resistant *Mycobacterium tuberculosis* isolates from Pakistan

Akbar Kanji[1], Rumina Hasan[1], Asho Ali[6], Ambreen Zaver[1], Ying Zhang[5], Kehkashan Imtiaz[1] , Wanliang Shi[5], Taane G. Clark[2,3], Ruth McNerney[4], Shoaib Rao[1], Samreen Shafiq[1] and Zahra Hasan*[1]

[1]Department of Pathology and Laboratory Medicine, Aga Khan University Hospital, Karachi

[2]Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[3]Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom

[4]Lung Infection and Immunity Unit, Division of Pulmonology, Department of Medicine, University of Cape Town, and UCT Lung Institute, Cape Town, South Africa.

[5]Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205, USA

[6]Faculty of Microbiology, King Abdulaziz University, KSA

Corresponding Author:

Zahra Hasan, PhD

Professor

Department of Pathology and Laboratory Medicine

The Aga Khan University

Stadium Road, P.O.Box 3500

Karachi 74800, Pakistan

Running Title: Resistance related to Efflux Pumps in XDR MTB Strains

Abstract

Background

It is challenging to understand mechanisms of drug resistance in *Mycobacterium tuberculosis* (MTB) due to the large variability in resistance associated genes. Efflux pump genes contribute to drug resistance and thus add to this complexity. Efflux pumps gene protein superfamilies have been characterized by genome analysis of drug resistant strains and through *in vitro* transcriptional studies. However, there is limited information regarding efflux pump genes in extensively drug resistant (XDR) tuberculosis (TB) isolates.

Methods

Whole genome sequencing (WGS) based analysis of 37 extensively drug resistant (XDR) and five drug sensitive (DS) MTB clinical isolates was performed. Single nucleotide polymorphisms (SNPs) in efflux pump genes *Rv0194, Rv1217, Rv1218, drrA, drrB, Rv1258, Rv1634, Rv2688, Rv1273, Rv1819, Rv1458, Rv1877* and *Rv1250* were determined in the clinical isolates as compared with the H37Rv reference strain. Allele frequencies of SNPs identified in XDR strains were compared with DS strains. Gene expression of *Rv0194, Rv2688, Rv1634, drrA* and *drrB* was determined in XDR -TB isolates (n=9), DS-TB strains (n=4) and H37Rv.

Results

We identified SNPs in XDR-TB isolates which were either unique or present at very low frequencies in DS strains; *Rv0194* G170V; *Rv1217* L151R; *Rv1258* P369T and G391R*; Rv1273* S118G and I175T; *Rv1877* I534T*; Rv1250* V318X/A and S333A*,* and *Rv2688* P156T. The expression of *Rv2688* and *drrB* was found to be raised in XDR-TB as compared with DS-TB strains.

Conclusions

We identified some unique SNPs in efflux pump genes which may be associated with increased drug resistance in the isolates. Also increased levels of *Rv2688 and drrB* efflux pump genes in XDR strains even in the absence of antibiotics could make these clinical isolates more refractory to treatment. Further studies are required to directly associate these mutations with increased resistance in MTB.

Keywords

Extensively Drug Resistant, *Mycobacterium tuberculosis*, single nucleotide polymorphism, Efflux Pump

Introduction

Extensively drug resistant (XDR)-tuberculosis (TB) is caused by multidrug-resistant (MDR) *Mycobacterium tuberculosis* (MTB) strains (resistant to rifampicin (RIF) and isoniazid (INH)) which are also resistant to fluoroquinolone (FLQ) and one of the injectable aminoglycosides; amikacin (AMK), kanamycin (KAN) or capreomycin (CAP). In Pakistan, 4-5% of all TB cases are MDR and of these, 4-5% is XDR-TB [1, 2]. Resistance in MTB is mostly attributed to the acquisition of specific mutations in target genes [3-5]. Single nucleotide polymorphisms (SNPs) in *rpo*B, *katG*, *gyrA*, *rrs*, *embB* and *pncA* genes can result in resistance to RIF, INH, FLQ, AMK, ethambutol (EMB) and pyrazinamide (PZA) respectively. However, additional genes have also been shown to confer resistance to each of the above mentioned drugs, adding to the complexity of drug resistance in MTB. Efflux pumps are a likely alternative mechanism of resistance in MTB.

Drug efflux has also been associated with pathogenicity, virulence, biofilm formation and coordination of gene expression in dense bacterial populations [6, 7]. Analysis of the available bacterial genomes has shown that putative drug efflux pumps (EPs) constitute 6% to 18% of all transporters found in any given bacterial cell. MTB presents one of the largest numbers of putative EPs compared to its genome size [8]. The mechanisms for the expression and regulation of these efflux pumps are not yet fully understood. Moreover, their role in drug resistance needs to be understood [9].

Although efflux mechanisms have been studied in several mycobacteria [10-12] mostly, *M. smegmatis* has been used to study heterologous putative efflux pump genes [13-16]. In *M.*

smegmatis as well as in other mycobacteria, the cell wall, rich in mycolic acids, functions as an efficient barrier preventing the access of several molecules including antibiotics. However, this is not enough for explaining the innate drug resistance of these microorganisms [17].

The first efflux pump described in mycobacteria was LfrA; belonging to the Major Facilitator Super-family (MFS) and which confers low-level resistance to fluoroquinolones, ethidium bromide, acridine and some quaternary ammonium compounds[18]. Other Efflux pumps initially characterized in mycobacteria were TetV conferring resistance to tetracycline [19] and *Rv1258c* which confers low-level resistance to aminoglycosides and tetracycline [10]. Adenosine tri phosphate ATP binding cassette (ABC) transporters are found both in prokaryotes and eukaryotes and constitute a large superfamily of multi-subunit permeases that transport amino acids, antibiotics and polysaccharides across biological membranes [13]. The ABC class of efflux pumps are primary transporters and are encoded by 2.5% of the genome of MTB.

Sequence based analysis of MTB genomes has identified at least 12 putative ABC transporters which include: *Rv0194, Rv1218c-Rv1217c, drrA-drrB-drrC, Rv1273c-Rv1272c, Rv2688c-Rv2687c-Rv2686c, Rv1348-Rv1349, Rv1456c-Rv1457c-Rv1458c, Rv1473, Rv1667c-Rv1668c, Rv1686c-Rv1687c, Rv1819, Rv2477*[20]. The common feature of all ABC transporters is the presence of two hydrophobic membrane spanning domains (MSDs) associated with two cytoplasmic nucleotide binding domains (NBDs). The MFS, small multidrug resistance (SMR), Resistance-Nodulation-Division (RND) and Multidrug and toxic compounds extrusion (MATE) members are secondary transporters, characteristically energized by the proton motive force (H+ or Na+)[21]. *Rv1218c,* an ABC class Efflux pump and *Rv3065,* of the SMR class have been shown to mediate the efflux of different chemical classes and antibiotics in *M smegmatis*[22]. Efflux pumps encoded by *Rv0849* and *Rv1258c* also help extrude these compounds but to a

lesser extent in MTB[23]. The *Rv1217c* (MT1255)–*1218c* (MT1256): tetronasin-transport membrane protein ABC transporter, is composed of two copies of the *Rv1218c*, active as the nucleotide binding domain (NBD), and one copy of the *Rv1217c*, fused by two membrane-spanning domains (MSDs)[21].

When overexpressed in *M smegmatis*, the *Rv2686c-Rv2687c-Rv2688c* ABC transporter has been shown to confer resistance to fluoroquinolones such as ciprofloxacin, with a 8 fold increase in the MIC for wild type MTB strains [15]. The Daunorubicin- phthiocerol dimycocerosates (DIMs)-transport ABC transporter operon comprises *drrA* (*Rv2936*) which is the NBD and *drr*B (*Rv2937*) which encodes the membrane integral component [24]. The *drr* operon plays a significant role in the virulence of MTB [25], it is an important drug target and exports a surface antigenic lipid known as DIMs to the cell surface [26]. *Rv0194* is an ABC multidrug efflux pump which can increase resistance in *M bovis* BCG to several other antibiotics [27].

Based on sequence analysis, at least 16 putative MFS efflux pump genes have been identified in MTB which include: *Rv0037c*, *Rv0191*, *Rv0783c*, *Rv0849*, *Rv1250*, *Rv1258c*, *Rv1410c*, *Rv1634*, *Rv1877*, *Rv2333c*, *Rv2456c*, *Rv2459*, *Rv2846c (efpA)*, *Rv28994*, *Rv3239c* and *Rv3728* [19].

Whilst many studies describe efflux pump genes and their over-expression in response to anti-tuberculous drugs, there is limited information regarding mutations in the genes. Garima *et* al. identified SNPs in ten efflux pump genes (*Rv1877*, *Rv1634*, *Rv1250*, *Rv0194*, *Rv1273*, *Rv1458*, *Rv1819*, *Rv0507*, *Rv3832*, *Rv0676*) by comparing a MDR TB strain sequence with H37Rv and showed the same genes to be over-expressed in response to INH and RIF treatment [28]. Li *et* al, compared the gene expression of 20 putative efflux pump genes between MDR and pan sensitive clinical isolates, to find that in MDR isolates the basal levels of some of these efflux pumps

genes were higher even without drug inducement [29]. To date there is limited information describing efflux pump genes in XDR TB strains. In this study, we analysed whole genome sequence (WGS) data of XDR MTB clinical strains for non-synonymous (nsSNPs) in efflux pump genes belonging to mainly the ABC and MFS super families; *Rv0194, Rv1217, Rv1218, Rv1258, drrA, drrB, Rv1634, Rv2688, Rv1273, Rv1819, Rv1458, Rv1877 and Rv1250*. These were genes in which SNPs had been previously described and/ or had been shown to be over-expressed in response to antibiotic treatment. We also investigated the transcription levels of efflux pump genes *Rv0194*, *Rv1634*, *Rv2688*, *drrA* and *drrB* in XDR isolates as compared with pan sensitive clinical isolates and H37Rv.

Materials and Methods

This study was approved by the Ethical Review Committee, The Aga Khan University.

XDR-TB isolates

XDR-TB strains (n=37) and drug sensitive strains (DS) (n=5) were obtained from the Aga Khan University Clinical Microbiology laboratory strain bank and have been described previously [30]. The XDR strains comprised mainly CAS family isolates: CAS1-Delhi (n=20), CAS (n=2) and CAS2 (n=1); others were EAI3-IND (n=2), X3 (n=1), T1 (n=3) and other lineage 1 and other lineage 3 isolates (n=8) lineages [29]. The DS clinical strains belonged to the CAS1-Delhi lineage.

Whole gene sequencing analysis

Briefly, all samples underwent WGS with 76-base paired end fragment sizes, using Illumina paired end HiSeq2000 technology, and the raw sequence data is available in the European nucleotide archive (http://www.ebi.ac.uk/ena/data/view/ PRJEB7798). XDR genome data was mapped to H37Rv reference genome (Genbank accession: AL123456.3) using *bwa-mem* software and variants called using SAMtools (http://samtools.sourceforge.net)[31, 32].

Resistance associated genes previously evaluated in the 37 XDR and 5 DS-Pak strains were; *rpo*B, *rpo*C, *rpo*A, *kat*G, *inh*A, *fbp*C, *Rv1392c*, *ndh*, *Rv2242*, *fab*D, *kas*A, *acc*D, *oxy*R, *fad*E24, *ndh*, *Rv1772*, *gyr*A, *gyr*B, *rrs*, *tly*A, *eth*A, *fab*G1, *rps*L, *gid*B, *pnc*A, *emb*B, *emb*A, *emb*C, *Rv3124*, *rml*D, *ini*A, *ini*B, *ini*C, *PPE49* and *man*B [29]. In addition, we investigated *emb*A_pro, *ubi*A, *fab*G1_pro, *eis*_pro mutations. SNPs identified in MTB isolates are listed in Supplementary Table 2.

Efflux pump genes annotations are as per the MTB genome [33] available at www.sanger.ac.uk/Projects/M_tuberculosis/Gene_list/functional_classes/III.A.6.shtml.
Previously published Artemis Comparison Tool (ACT)[34] was used to identify nsSNPs in the open reading frames of target genes. A maximum likelihood phylogenetic tree was created with RAxML using all SNPs present in the XDR isolates.

Allele frequencies of the SNPs identified in efflux pump genes of XDR strains were determined in a cohort of 1,526 drug susceptible isolates from Karonga, Malawi [35]. The data set is publically available at the EBI European Nucleotide Archive (http://www.ebi.ac.uk/ena/data/view/ERP000436). Sample genotypes were called using the majority allele (minimum frequency 75%) in positions supported by at least 20-fold coverage; otherwise we classified them as missing (thus ignoring heterozygous calls).

Culture and determination of minimum inhibitory concentrations (MICs)

MTB strains were isolated from 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 [36]. Drug Susceptibility Testing (DST) of these isolates was performed using the agar proportion method on enriched Middlebrook 7H10 medium (BBL Microbiology Systems, Cockeysville, MD, USA) as per Global TB Programme (WHO) guidelines for first-line and second-line DST [37]. Concentrations used were: RIF 1 µg/mL, INH 0.2 µg/ml, streptomycin (STR) 2 µg/ml, and EMB 5 µg/ml. Pyrazinamide (PZA) 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 ofloxacin (OFX) 2ug/ml, CAP 4µg/ml, AMK 4µg/ml, KAN 5µg/ml and ethionamide (ETH) 5µg/ml. The reference strain MTB H37Rv was used as a control with each susceptibility testing batch[38].

DNA extraction

MTB isolates grown on Middlebrook 7H10 medium were inoculated in 500ul of 1X Tris EDTA (TE) buffer. The MTB cell suspension was heat killed at 85^0C for 30mins in a sonicating water bath. The sample was centrifuged at 13,000 g for 5 min, and the supernatant was discarded. Cellular pellets were processed for DNA extraction was performed using the cetyl-trimethyl ammonium bromide (CTAB) method [39].

RNA extraction and reverse transcription

MTB isolates were grown to log phase in Middlebrook medium. MTB was heat killed at 85^0C for 30mins in a sonicating water bath. The sample was centrifuged at 13,000 g for 5 min, and the supernatant was discarded. RNA was extracted using Trizol reagent (Invitrogen, USA) as per the

manufacturer's recommendations. One microgram of RNA was reverse transcribed using Moloney Murine Leukemia Virus (M-MuLV, MMLV) Reverse Transcriptase (Invitrogen, USA) as described previously[40].

The cDNA was stored in aliquots at -35^{0}C and 2µl of cDNA were used in each PCR reaction.

Real time PCR was performed in duplicate 20µl reactions containing Platinum SYBR Green qPCR Supermix-UDG (Invitrogen). Forward and Reverse primers at a final concentration of 10pmole/µl. The forward and reverse primers for SYBR real time PCR for *Rv1634*, *drrA* and *drrB* efflux pump genes were taken from Li *et al*, PLoS One, 2015[29]; *16S RNA* from Sarathy *et al*, PLoS One, 2013[41]. The primer sequences for *Rv2688* (Forward: CACGGTGTGGGACAAAGA, Reverse: AGAAGCGCAGGTTCTCATAC); *Rv0194*: (Forward: GTCCACTATTCCTACCGAACAC, Reverse: TTGATCAGGGTGGATTTCCC).

All PCRs were run at 94°C for 15 min; 45 cycles at 95^{0}C for 30s, annealing at 60^{0}C for (*drrA* and *drrB*); 57^{0}C (*Rv0194*); 56^{0}C (*Rv1634*); 58^{0}C (*Rv2688*) followed by extension at 72^{0}C for 30s. Target amplification was determined by melting curve analysis.

Quantification of gene expression using SYBR Green real-time quantitative PCR (qPCR)

Efflux pump gene expression was compared between nine XDR strains and five drug susceptible MTB strains (H37Rv and four clinical isolates). Fold change was calculated by 2-ΔΔCT method of Livak and Schmittgen [42]. *16S RNA* was used as used as a control gene to calculate the ΔCT values for individual samples. Differences in threshold cycle (ΔCT) values for each gene were obtained by subtracting the CT value of *16S RNA* in each case.

PCR was performed in duplicates for each gene. The relative amount of gene transcripts was calculated using the 2-ΔΔCT method as described [42]. These values were then used to calculate the relative expression of efflux pump mRNA in each of the samples tested [42].

Data analysis

GraphPad Prism 5.0 was used to perform Mann-Whitney test, and the difference was considered to be statistically significant when $P < 0.05$.

Results

We first analysed the genome sequence data obtained for 37 XDR-TB and 5 drug sensitive strains. According to lineage characterization, the representation of XDR-TB strains was largely from Lineage 3 (75.7 %) followed by Lineage 1 (13.5%) and Lineage 4 (10.8%) isolates (Supplementary Table 1). The relatedness between the XDR strain lineages as per their SNPs is depicted in a phylogenetic tree (Figure 1). The SNPs in the XDR strains were largely associated with their lineage groups.

Further we focused on SNPs in putative efflux pump genes belonging to two superfamilies of genes found in *Mycobacterium sp.,* ABC transporters and MFS proteins. These were; *Rv0194, Rv1217, Rv1218, Rv1258, drr*A, *drr*B, *Rv1634, Rv2688, Rv1273, Rv1819, Rv1458, Rv1877* and *Rv1250*. SNPs identified in the candidate efflux pump genes in XDR strains were compared with our WGS data from drug susceptible clinical isolates, see Table 1.

To further investigate the relevance of SNPs identified in target efflux pumps from the XDR analysis, data of 1,526 drug susceptible isolates from Karonga, Malawi was also analysed for the SNPs of interest identified from the XDR analysis [35]. The isolates from Malawi comprised

16% Lineage of 1, 5% Lineage of 2, 11% of Lineage 3 and 68% of Lineage 4 strains. The frequency of the SNPs compared with the larger data set of Karonga isolates is depicted in Table 2 with 0.1 as a cut-off for common alleles [35].

***Rv0194* gene**

Sequence analysis of *Rv0194* showed that the SNP M74T was present in all XDR strains and the 4 DS-TB (Pakistan, Pak) strains (Table 1). M74T was also present at a frequency of 0.95 in Karonga strains (Table 2). However, one XDR strain (Lineage 1, X17) had a G170V change which was not present in the DS-TB (Pak) strains. *Rv0194* G170V was present at a low (0.04) frequency in the DS-Malawi set (Table 2).

***Rv1217* efflux pump gene**

The WGS comparisons of *Rv1217* revealed that one XDR EAI-3-IND strain (X32) had a 41 bp deletion at nt+1361325. Another XDR strain (Lineage 1, X49) had *Rv1217* L151R which was not present in the DS-Pak strains and was only found at 0.03 frequency in the DS-Malawi set. X49 also had the *drr*A SNP H309D (Table 1).

***Rv1218* efflux pump gene**

Thirty-three (89%) of XDR and all four CAS DS-Pak strains had *Rv1218* Q243R. Of the four XDR strains that did not have Q243R, three belonged to T1 and one to the CAS1-Delhi lineage. *Rv1218* SNP Q243R was found at frequency of 0.39 in the DS-Malawi data set.

Polymorphism in *Rv1258* efflux pump gene

One XDR strain (Lineage 1, X57) had *Rv1258* G391R, *drr*A H309D and was resistant to eight drugs (Table1). Another XDR strain (X47) with *Rv1258* P369T was resistant to all ten drugs tested. Neither *Rv1258* G391R nor P369T were found in DS-Pak or DS-Malawi isolates (Table 2).

drrA efflux pump gene

Five XDR Lineage 1 strains (X5, X17, X49, X32 and X57) had *drr*A H309D. Two of these XDR strains belonged to EAI-IND lineage and three were 'Other Lineage 1' isolates. This SNP was present at a frequency of 0.16 in DR-Malawi isolates. It is likely to be a lineage associated SNP as the 5 XDR strains we describe here clustered within a two nodes of relatedness (Fig 1).

***Rv1634* efflux pump gene**

Analysis of *Rv1634* revealed I47V in three of above described Lineage 1 strains (X49, X32 and X5) (Table 1) which clustered together (Fig. 1). This SNP was present at a frequency of 0.16 in DRS isolates.

Rv2688

We found the *Rv2688* SNP P156T in each of the XDR strains we studied. This SNP was not present in the DS-Pak isolates but was found at a frequency of 0.08 in the DS-Malawi isolates.

Rv1273

The *Rv1273* SNP G462K was present in 29 XDR strains, in all five DS-Pak strains and in 11 % of DS-Malawi strains. A 'lineage 1' XDR isolate (X17) had SNPs *Rv1273* S118G and I175T. Another 'lineage 1' XDR strain (X57) had *Rv1273* I175T. *Rv1273* S118G was also observed in

a CAS1-Delhi XDR strain (X55). Neither *Rv1273* S118G nor I175T were observed in the DS-Pak and DS-Malawi data sets. There appeared to be no phylogenetic relationship between X55 and X57 (Fig 1) although both had *Rv1273* S118G.

Rv1819

Rv1819 I603V has previously been identified by Garima et al in MDR strains [28]. The SNP was also present at a frequency of 0.16 in DS-Malawi strains. We found the combination of *Rv1819* I603V and *drrA* H309D SNPs to be present in the five Lineage 1 (clustered) XDR strains described earlier suggesting, that these SNPs may be associated with a common lineage.

Rv1458

We did not identify any SNPs in the *Rv1458* in the XDR or DS-Pak isolates studied.

Rv1877

The *Rv1877* I534T SNP was present in two XDR strains belonging to 'lineage 1'. This SNP was not present in DS-Pak strains and had a 0.04 allele frequency in the DS-Malawi isolates.

Rv1250

Four different SNPs were observed in *Rv1250*. We found R278G to be present in 5 XDR strains. The R278 SNP has been mentioned by Garima et al to be a high confidence mutation [28]. However, we also found it at a frequency of 0.16 in the DS-Malawi set. Three XDR strains had unique *Rv1250* SNPs A196T, V318X/A and S333A which were not present in any DS-Pak isolate studied.

Associating Lineage specific SNPs with resistance genotypes

To summarise, we found a combination of SNPs; *Rv1634* I47V, *drr*A H309D and *Rv1819* I603V in five Lineage 1 XDR isolates (X5, X17, X32, X49 and X57). We also observed that the *ubi*A E149D SNP was also present in the abovementioned strains. As these XDR strains clustered together in a phylogenetic tree (Fig. 1) the abovementioned SNPs can be considered to be related to a common lineage.

Further, we looked for particular genotypes related to efflux pumps which could possibly be associated with dug resistance. We analysed the XDR strains for common drug resistance conferring mutations related to flouroquinolone, aminoglycoside and streptomycin. Eight of the XDR strains (X5, X11, X16, X32, X37, X40, X49 and X58) did not have any *gyr*A mutations, Supplementary Table 2. Five XDR strains (X32, X42, X43, X49 and X56) resistant to AMK/KAN did not present with any *rrs* (1400 region) mutations. Fourteen XDR strains were phenotypically resistant to STR but did not display *rrs* (500 region) or *rps*L mutations.

In the context of unique or low frequency efflux pump SNPs, nine XDR strains had one or more SNP in an efflux pump gene which has been associated with increased resistance to anti-tuberculous drugs (Table 3). As all of these strains presented with a combination of efflux pump gene mutations (Table 1) it is quite possible that these SNPs have contributed to resistance in the XDR isolates.

Quantification of basal mRNA expression levels of *Rv1634, Rv2688, drrA* and *drrB and Rv0194* efflux pump genes in XDR MTB strains

Nine XDR-TB strains from the overall XDR strains were randomly selected for transcriptional studies. We were able to re-grow four of the five drug susceptible strains (for which WGS data

was available) thus, these four strains and H37Rv were included as pan sensitive isolates for the gene expression study.

We investigated particular efflux pump genes by determining their expression levels in XDR-TB isolates as compared with drug susceptible clinical isolates, in the absence of antibiotic selection pressure. Efflux pump genes *Rv0194, Rv1634*, *Rv2688*, *drrA* and *drrB* was chosen as these belonged to ABC-type transporter and MFS superfamilies which have previously been shown to be upregulated in drug resistant MTB strains. We studied nine XDR strains across different lineages; X5 was from EAI-IND, X9, X10, X42, X55 and X61 were from CAS1-Delhi and X49, X57 and X60 were from Orphan lineages. XDR strains were compared with five drug susceptible strains, 4 pan-susceptible CAS1-Delhi strains and H37Rv.

When the gene expression in XDR and drug sensitive strains was compared it was apparent that mean basal mRNA expression levels of *Rv2688c* (p=0.0108) and *drrB* (p=0.0166) were significantly higher in the XDR as compared with drug susceptible MTB strains (Figure 2). However, no significant differences were observed in the mRNA levels of *Rv0194* (p=0.1643), *Rv1634* (p=0.1891) and *drrA* (p=0.1891) respectively between XDR and DS MTB strains (Figure 2).

Discussion

We focused on efflux pumps belonging to the larger families of transport proteins; ABC and MFS types and those which have been well studied and shown to contribute to antibiotic resistance in MTB. By determining nsSNPs in selected efflux pump genes in XDR strains and comparison with DS strains we have identified a number of SNPs which could potentially be

associated with drug resistance. We first identified SNPs in efflux pump genes *Rv0194*, *Rv1217*, *Rv1218*, *Rv1258*, *drrA*, *drrB*, *Rv1273*, *Rv1634*, *Rv1819*, *Rv2688*, *Rv1273*, *Rv1819*, *Rv1877* Rv1250 and *Rv1458*. Through the comparison of data sets we identified alleles which were either unique to XDR strains or were present at a very low frequency in DS-TB strains. The Pakistan strains comprised mainly of lineage 1 and 3 strains. The African MTB strains comprised a different proportion of lineage isolates but as the data set was fairly large (n=1526), there was substantial representation of strains including, lineage 1 and 3 strains. Through this work we excluded a number of SNPs previously been labeled as high confidence mutation by comparison between H37Rv and a reference MDR-TB sequence [28].

Firstly, we identified SNPs in efflux pump genes *Rv0194*, *Rv1217*, *Rv1218*, *Rv1258*, *drrA*, *drrB*, *Rv1273*, *Rv1634*, *Rv1819*, *Rv2688*, *Rv1273*, *Rv1819*, *Rv1877* and *Rv1250* and but not in *drrB* or *Rv1458*. Further comparison with DS-Pak and DS-Malawi isolates revealed a sub-set of unique and low frequency alleles in seven of the efflux pump genes; *Rv0194*, *Rv1217*, *Rv1258*, *Rv2688*, *Rv1273*, *Rv1877* and *Rv1250* respectively.

The *Rv0194* SNP G170V was present in a single XDR strain belonging to the 'Other Lineage 1' family. The role of *Rv0194* in efflux of beta lactams such as vancomycin and tetracycline [27] has been shown but as yet this efflux pump has not been associated with rifampicin, isoniazid or any other second line anti-TB drugs in MDR or XDR-TB strains.

One 'Other lineage 1' isolate had *Rv1217* SNP L151R whilst another had a 40 nucleotide deletion in *Rv1217*. *Rv1217-Rv1218* operon is ABC transporter shown to play a role in mediating antibiotic resistance in *M. smegmatis* [22]. *Rv1218* and *Rv1273* genes have been shown to be upregulated in MDR isolates which did not have *rpo*B or *kat*G mutations [43]

suggesting, it plays a role in MDR resistance. Two XDR strains had *Rv1273* S118G which has been reported previously in association with STR and EMB drug resistance [28].

The *Rv1258* SNP G391R was present in an Orphan strain and P369T in a CAS1 strain. *Rv1258* encodes a tetracycline/aminoglycoside resistance -like efflux pump [10, 44] and is found to be upregulated in MDR-TB under rifampicin pressure [45].

Rv1634 belongs to the Major facilitator superfamily (MFS) and has been reported to decrease susceptibility to various fluoroquinolones when overexpressed in *M. smegmatis* and is involved in norfloxacin and ciprofloxacin efflux [46]. *Rv1634* efflux pump gene has also been shown to be upregulated in MTB strains after exposure to RIF [29].

Drr proteins of MTB are suggested to be involved in the export of complex lipids dimycocerosates (DIM) to the cell surface in MTB [47]. Previously, it was reported that *drrA* was overexpressed in MDR isolates induced by INH and zero MDR isolates induced by RIF, suggesting that *drrA* was one of the factors for INH resistance in MTB [29].

The *Rv1877* SNP I534T has been reported previously in resistant strains and *Rv1877* gene expression is shown to be upregulated in drug resistant MTB [28].

We described five Lineage 1 XDR strains with a combination of *drr*A SNP H309D, *Rv1634* SNP I74V and *ubiA* E149D. Ethambutol resistance has previously been associated with *ubi*A gene mutations and it has been previously shown that E149D was present with resistant isolates only [48]. Association of the above mentioned SNPs with lineages was demonstrated by the clustering of lineage 1 XDR strains which had a unique subset of SNPs.

Streptomycin resistance is thought to be associated with *rpsL*, *rrs* (500 region) mutations. *gidB* polymorphism is associated with low level of streptomycin resistance in MTB [49]. We found that 11 isolates were STR resistant but did not have *rpsL*, *rrs* (500 region) mutations. Also, there were eight OFX resistant isolates which did not have mutations in *gyr*A genes and there were five KAN/AMK resistant isolates without mutations in the *rrs* (1400 region). It may be that the combination of the efflux pump gene associated SNPs may have an additive effect on drug resistance of the XDR isolates, so that they become resistant to a range of anti-tuberculous drugs.

mRNA expression analysis of the *Rv0194, Rv2688, Rv1634*, *drrA* and *drrB* basal mRNA chosen efflux pump genes revealed that *Rv2688* and *drr*B gene expression was increased in XDR strains as compared with drug susceptible isolates. In the nine XDR strains studied, none had SNPs in the *drrB* gene. However, all the XDR strains had the *Rv2688* SNP P168T. There were no *drr*B or *Rv2688* SNPs present in the drug susceptible isolates studied. We found significantly increased ($p<0.05$) basal level mRNA expression levels of expression levels in XDR MTB strains as compared with the susceptible clinical isolates and the H37Rv laboratory strain.

The upregulation of Rv*2688* has been associated with flouroquinolone resistance as it was observed that the *Rv2688* along with *Rv2686* and *Rv2687* actively pumped out ciprofloxacin[15].

Previous reports have shown reduced mRNA expression of *Rv1877, Rv2265* and *MmpL13a* in MDR MTB strains as compared with pan sensitive MTB strains, with the suggestion that these efflux pumps regulate intracellular levels of nutrients and cofactors but not pump out anti-TB drugs [29].

We have identified novel SNPs in efflux pump genes that could be associated with drug resistance however, further work needs to be done so that this association can be confirmed. We were not able to perform gene expression analysis of all the efflux pump genes in which SNPs were identified and this could be done in the future. This data provides important insights into the variability of efflux pump related SNPs and their potential role in increasing drug resistance in MTB.

Conclusion

Overall, our data identifies unique SNPs in efflux pump genes *Rv0194, Rv1217, Rv1250, Rv1877* and *Rv1273.* As these were not found in drug susceptible strains these may be associated with alternate mechanism of drug resistance in XDR-TB strains. The upregulation of efflux pump genes in the absence of antibiotic pressure further points to the important role of efflux pump proteins in the acquisition and maintenance of drug resistance in MTB isolates.

Acknowledgements

We thank the Aga Khan University Hospital Clinical Laboratory for samples used in this study. This work was made possible through Dr. Arnab Pain through the whole genome sequencing data provided by KAUST, Saudi Arabia.

References

1. Ayaz, A., et al., *Characterizing Mycobacterium tuberculosis isolates from Karachi, Pakistan: drug resistance and genotypes.* Int J Infect Dis, 2012. **16**(4): p. e303-9.
2. Hasan, R., et al., *Extensively drug-resistant tuberculosis, Pakistan.* Emerg Infect Dis, 2010. **16**(9): p. 1473-5.
3. Takiff, H.E., et al., *Cloning and nucleotide sequence of Mycobacterium tuberculosis gyrA and gyrB genes and detection of quinolone resistance mutations.* Antimicrob Agents Chemother, 1994. **38**(4): p. 773-80.
4. Taniguchi, H., et al., *Rifampicin resistance and mutation of the rpoB gene in Mycobacterium tuberculosis.* FEMS Microbiol Lett, 1996. **144**(1): p. 103-8.
5. Taniguchi, H., et al., *Molecular analysis of kanamycin and viomycin resistance in Mycobacterium smegmatis by use of the conjugation system.* J Bacteriol, 1997. **179**(15): p. 4795-801.
6. Piddock, L.J., K.J. Williams, and V. Ricci, *Accumulation of rifampicin by Mycobacterium aurum, Mycobacterium smegmatis and Mycobacterium tuberculosis.* J Antimicrob Chemother, 2000. **45**(2): p. 159-65.
7. Bina, X.R., J.A. Philippart, and J.E. Bina, *Effect of the efflux inhibitors 1-(1-naphthylmethyl)-piperazine and phenyl-arginine-beta-naphthylamide on antimicrobial susceptibility and virulence factor production in Vibrio cholerae.* J Antimicrob Chemother, 2009. **63**(1): p. 103-8.
8. Paulsen, I.T., et al., *Comparative genomics of microbial drug efflux systems.* J Mol Microbiol Biotechnol, 2001. **3**(2): p. 145-50.
9. Poelarends, G.J., et al., *An ABC-type multidrug transporter of Lactococcus lactis possesses an exceptionally broad substrate specificity.* Drug Resist Updat, 2000. **3**(6): p. 330-334.
10. Ainsa, J.A., et al., *Molecular cloning and characterization of Tap, a putative multidrug efflux pump present in Mycobacterium fortuitum and Mycobacterium tuberculosis.* J Bacteriol, 1998. **180**(22): p. 5836-43.
11. Ramon-Garcia, S., et al., *Role of the Mycobacterium tuberculosis P55 efflux pump in intrinsic drug resistance, oxidative stress responses, and growth.* Antimicrob Agents Chemother, 2009. **53**(9): p. 3675-82.
12. Nomura, K., et al., *Antibiotic susceptibility of glutaraldehyde-tolerant Mycobacterium chelonae from bronchoscope washing machines.* Am J Infect Control, 2004. **32**(4): p. 185-8.
13. Liu, J., H.E. Takiff, and H. Nikaido, *Active efflux of fluoroquinolones in Mycobacterium smegmatis mediated by LfrA, a multidrug efflux pump.* J Bacteriol, 1996. **178**(13): p. 3791-5.
14. De Rossi, E., et al., *mmr, a Mycobacterium tuberculosis gene conferring resistance to small cationic dyes and inhibitors.* J Bacteriol, 1998. **180**(22): p. 6068-71.
15. Pasca, M.R., et al., *Rv2686c-Rv2687c-Rv2688c, an ABC fluoroquinolone efflux pump in Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 2004. **48**(8): p. 3175-8.
16. Kim, S.Y., et al., *Identification of novel metronidazole-inducible genes in Mycobacterium smegmatis using a customized amplification library.* FEMS Microbiol Lett, 2008. **282**(2): p. 282-9.
17. Mendez, S., et al., *Role for CD4(+) CD25(+) regulatory T cells in reactivation of persistent leishmaniasis and control of concomitant immunity.* J Exp Med, 2004. **200**: p. 201.
18. Takiff, H.E., et al., *Efflux pump of the proton antiporter family confers low-level fluoroquinolone resistance in Mycobacterium smegmatis.* Proc Natl Acad Sci U S A, 1996. **93**(1): p. 362-6.
19. De Rossi, E., et al., *The multidrug transporters belonging to major facilitator superfamily in Mycobacterium tuberculosis.* Mol Med, 2002. **8**(11): p. 714-24.

20. Braibant, M., P. Gilot, and J. Content, *The ATP binding cassette (ABC) transport systems of Mycobacterium tuberculosis.* FEMS Microbiol Rev, 2000. **24**(4): p. 449-67.
21. Pal, R., Fatima, Z.,Hameed, S,, *Efflux Pumps in drug resistance of Mycobacterium tuberculosis.* International Journal of Current Microbiology and Applied Sciences, 2014. **3**(8).
22. Balganesh, M., et al., *Rv1218c, an ABC transporter of Mycobacterium tuberculosis with implications in drug discovery.* Antimicrob Agents Chemother, 2010. **54**(12): p. 5167-72.
23. Wang, K., et al, *The expression of ABC efflux pump, Rv1217c-Rv1218c, and its association with multidrug resistance of Mycobacterium tuberculosis in China.* Curr Microbiol, 2013. **66**(3): p. 222-6.
24. Tseng, T.T., et al., *The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins.* J Mol Microbiol Biotechnol, 1999. **1**(1): p. 107-25.
25. Tamang, D.G., et al., *Comprehensive analyses of transport proteins encoded within the genome of "Aromatoleum aromaticum" strain EbN1.* J Membr Biol, 2009. **229**(2): p. 53-90.
26. Cattoir, V., *[Efflux-mediated antibiotics resistance in bacteria].* Pathol Biol (Paris), 2004. **52**(10): p. 607-16.
27. Danilchanka, O., C. Mailaender, and M. Niederweis, *Identification of a novel multidrug efflux pump of Mycobacterium tuberculosis.* Antimicrob Agents Chemother, 2008. **52**(7): p. 2503-11.
28. Garima, K., et al., *Differential expression of efflux pump genes of Mycobacterium tuberculosis in response to varied subinhibitory concentrations of antituberculosis agents.* Tuberculosis (Edinb), 2015. **95**(2): p. 155-61.
29. Ali, A., et al., *Whole genome sequencing based characterization of extensively drug-resistant Mycobacterium tuberculosis isolates from Pakistan.* PLoS One, 2015. **10**(2): p. e0117771.
30. Tanveer, M., et al., *Genotyping and drug resistance patterns of M. tuberculosis strains in Pakistan.* BMC Infect Dis, 2008. **8**: p. 171.
31. Phelan, J., et al., *Mycobacterium tuberculosis whole genome sequencing and protein structure modelling provides insights into anti-tuberculosis drug resistance.* BMC Med, 2016. **14**: p. 31.
32. Coll, F., et al., *A robust SNP barcode for typing Mycobacterium tuberculosis complex strains.* Nat Commun, 2014. **5**: p. 4812.
33. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**(6685): p. 537-44.
34. Carver, T.J., et al., *ACT: the Artemis Comparison Tool.* Bioinformatics, 2005. **21**(16): p. 3422-3.
35. Guerra-Assuncao, J.A., et al., *Large-scale whole genome sequencing of M. tuberculosis provides insights into transmission in a high prevalence area.* Elife, 2015. **4**.
36. Woods GL B-EB, D.E., Hall GS, Heifets L, Pfyffer GE, et al, *Susceptibility testing of mycobacteria nocardiae, and other aerobic actinomycetes approved standard* P. Wayne, Editor. 2003.
37. WHO, *Updated critical concentrations for first-line and second-line DST (as of May 2012)*. 2012, Global TB Programme.
38. Woods GL B-EB, D.E., Hall GS, Heifets L, Pfyffer GE, et al, *Susceptibility testing of mycobacteria nocardiae, and other aerobic actinomycetes approved standard Wayne, PA: NCCLS.* 2003.
39. Honore-Bouakline S, V.J., Giacuzzo V, Lagrange PH, Herrmann JL *Rapid diagnosis of extrapulmonary tuberculosis by PCR: impact of sample preparation and DNA extraction.* J Clin Microbiol : , 2003. **41**: p. 2323-2329.
40. Dheda, K., et al., *Validation of housekeeping genes for normalizing RNA expression in real-time PCR.* Biotechniques, 2004. **37**(1): p. 112-4, 116, 118-9.
41. Sarathy, J.P., E. Lee, and V. Dartois, *Polyamines inhibit porin-mediated fluoroquinolone uptake in mycobacteria.* PLoS One. **8**(6): p. e65806.

42. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
43. Yamchi, J.K., Haeili, M., Feyisa, S.G., Kazemian, H., Shahraki, A.H., Zahednamazi, F., Fooladi, A.A.I. and Feizabadi, M.M.,, *Evaluation of efflux pump gene expression among drug susceptible and drug resistant strains of Mycobacterium tuberculosis from Iran.* Infection, Genetics and Evolution, 2015(36): p. 23-26.
44. Siddiqi, N., et al., *Mycobacterium tuberculosis isolate with a distinct genomic identity overexpresses a tap-like efflux pump.* Infection, 2004. **32**(2): p. 109-11.
45. Jiang, X., et al., *Assessment of efflux pump gene expression in a clinical isolate Mycobacterium tuberculosis by real-time reverse transcription PCR.* Microb Drug Resist, 2008. **14**(1): p. 7-11.
46. De Rossi, E., J.A. Ainsa, and G. Riccardi, *Role of mycobacterial efflux transporters in drug resistance: an unresolved question.* FEMS Microbiol Rev, 2006. **30**(1): p. 36-52.
47. Camacho, L.R., et al., *Analysis of the phthiocerol dimycocerosate locus of Mycobacterium tuberculosis. Evidence that this lipid is involved in the cell wall permeability barrier.* J Biol Chem, 2001. **276**(23): p. 19845-54.
48. Xu, Y., et al., *Mutations Found in embCAB, embR, and ubiA Genes of Ethambutol-Sensitive and -Resistant Mycobacterium tuberculosis Clinical Isolates from China.* Biomed Res Int, 2015. **2015**: p. 951706.
49. Okamoto, S., et al., *Loss of a conserved 7-methylguanosine modification in 16S rRNA confers low-level streptomycin resistance in bacteria.* Mol Microbiol, 2007. **63**(4): p. 1096-106.

Figure Legends

Figure 1. Phylogenetic tree of XDR isolates. A maximum likelihood phylogenetic tree was created with RAxML using all SNPs in our isolates. XDR-TB strains (n=37) are identified by numbers and also their lineages. Central Asian Strain (CAS), East African Indian (EAI), Other Lineage (Ot. Lin.) for those which had in silico profiles similar to primary lineage families.

Figure 2. Gene expression of efflux pumps in DR and drug susceptible (DS) MTB strains. The figure depicts the relative mRNA expression of *Rv0194*, *Rv1634*, *Rv2688c*, *drrA* and *drrB* genes in XDR MTB strains (n=9) as compared with DS MTB strains (n=4). The results from real-time PCR showed no significant differences ($p>0.05$) in mean basal level mRNA expression levels of (A) *Rv0194* (B) *Rv1634* and (D) *drrA* efflux pump genes. The differences in mean basal level mRNA expression levels for (C) *Rv2688c* and (E) *drrB* were significantly higher ($p<0.05$)

Table1: An Overview of Non-synonymous SNPS (nsSNPs) in efflux pump genes of XDR MTB Isolates From Pakistan

Isolates ID	Spoligo	Lineage	*Rv0194*		*Rv1217*		*Rv1218*	*Rv1258*		*drrA*	Rv1634	Rv2688	Rv1273			Rv1819	Rv1877	Rv1250				Rv1458
			Position		Position			Position		Position	Position	Position	Position			Position	Position	Position				
			227098	227386	1361325	1361350	1362006	1406170	1406236	3273138	1839306	3005185	1422667	1423527	1423699	2062922	2127504	1394764	1395010	1395131	1395175	
EP1	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	-	G462K	-	-	-	-	-	-	-	-	-
EP2	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	-	G462K	-	-	-	-	-	-	-	-	-
S4	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	-	G462K	-	-	-	-	-	-	-	-	-
S5	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	-	G462K	-	-	-	-	-	-	-	-	-
C3	CAS1-Delhi	Lineage3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X1	Other lineage 3	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X4	CAS2	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X5	EAI3-IND	Lineage 1	M74T	-	-	-	Q243R	-	-	H309D	I47V	P156T	-	-	-	I603V	-	-	R278G	-	-	-
X6	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X7	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X8	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X9	Other lineage 3	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X10	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X11	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X12	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X13	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	V318X	-	-
X14	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X16	CAS1-Delhi	Lineage 3	M74T	-	-	-	-	-	-	-	-	P156T	-	-	-	-	-	-	-	-	-	-
X17	Other lineage 1	Lineage 1	M74T	G170V	-	-	Q243R	-	-	H309D	-	P156T	-	I175T	S118G	I603V	I534T	-	R278G	-	-	-
X18	CAS	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	-	-	-	-	-	A196T	-	-	-	-
X21	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X22	CAS	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	A196T	-	-	-	-
X32	EAI3-IND	Lineage 1	M74T	-	Δ	-	Q243R	-	-	H309D	I47V	P156T	-	-	-	I603V	-	-	R278G	-	-	-
X33	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	V318A	-	-
X37	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	S333A	S333A
X39	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X40	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X41	T1	Lineage 4	M74T	-	-	-	-	-	-	-	-	P156T	-	-	-	-	-	-	-	-	-	-
X42	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X43	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X44	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X45	Other lineage 3	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X46	X3	Lineage 4	M74T	-	-	-	Q243R	-	-	-	-	P156T		-	-	-	-	-	-	-	-	-
X47	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	P369T	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X48	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X49	Other Lineage 1	Lineage 1	M74T	-	-	L151R	Q243R	-	-	H309D	I47V	P156T	-	-	-	I603V	-	-	R278G	-	-	-
X55	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	S118G	-	-	-	-	-	-	-
X56	T1	Lineage 4	M74T	-	-	-	-	-	-	-	-	P156T	-	-	-	-	-	-	-	-	-	-
X57	Other Lineage 1	Lineage 1	M74T	-	-	-	Q243R	G391R		H309D	-	P156T	-	I175T	-	I603V	I534T	-	R278G	-	-	-
X58	T1	Lineage 4	M74T	-	-	-	-	-	-	-	-	P156T	-	-	-	-	-	-	-	-	-	-
X60	Other lineage 3	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-
X61	CAS1-Delhi	Lineage 3	M74T	-	-	-	Q243R	-	-	-	-	P156T	G462K	-	-	-	-	-	-	-	-	-

CAS1: Central Asian Strain1;EAI: East African Indian; OPN: Orphan; T1: Undefined Group; X-family;Δ denotes deletion; X denotes unknown change

Table 2. Frequency of efflux pump gene polymorphisms in susceptible *M. tuberculosis* isolates

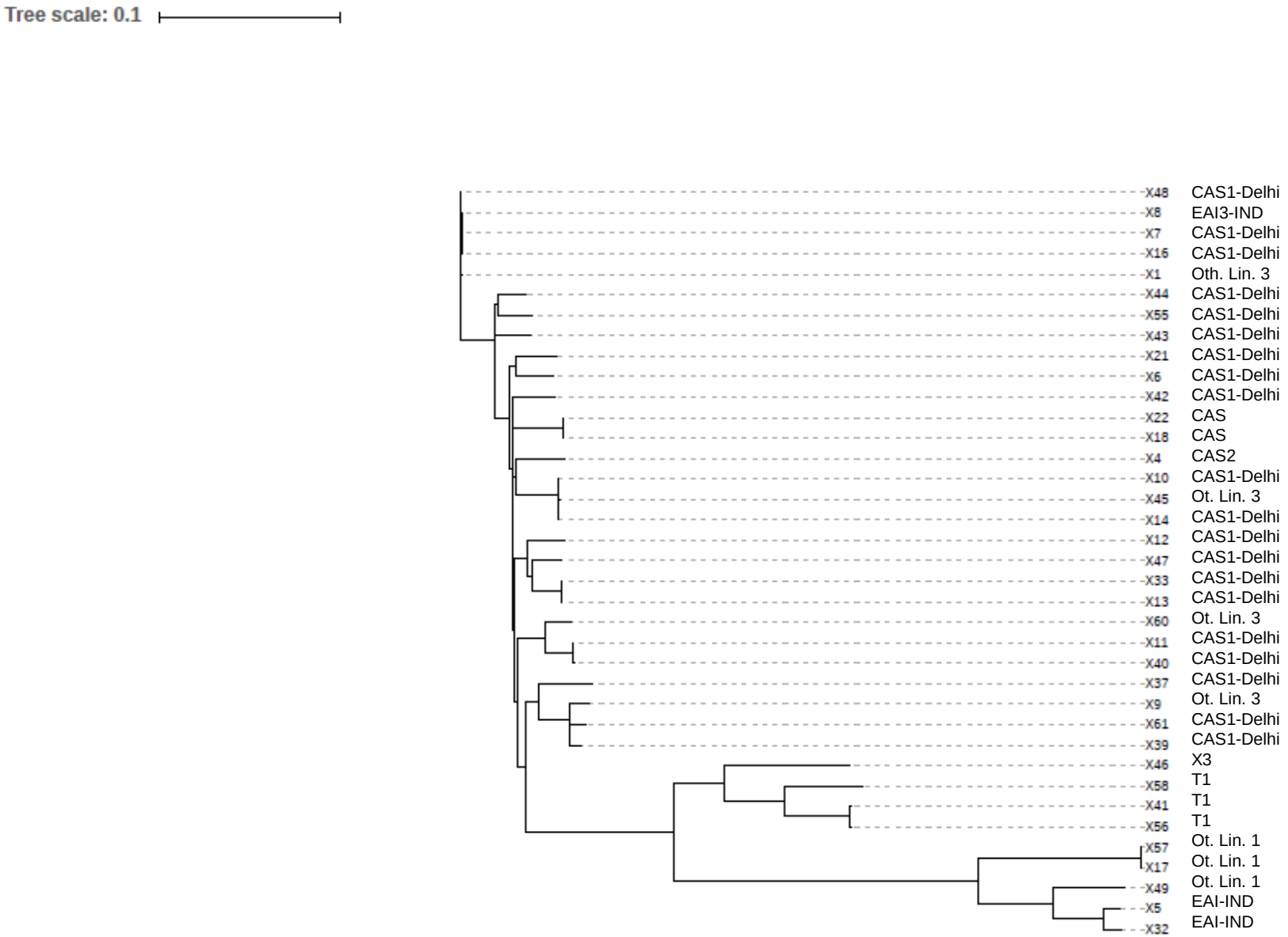
Gene name	Chr pos	codon_position	nucleotide_alleles	amino_acid_alleles	frequencies	
Rv0194	227098	74	C,T	T,M	0.95,0.05	
Rv0194	***227386***	***170***	***T,G***	***V,G***	***0.04,0.96***	
Rv1217c	***1361350***	***151***	***A,C***	***L,R***	***0.97,0.03***	
Rv1218c	1362006	243	C,T	R,Q	0.39,0.61	
Rv1258c	***1406170***	***391***	***C***	***G***	***1***	
Rv1258c	***1406236***	***369***	***G***	***P***	***1***	
Rv2936 (drrA)	3273138	309	C,G	H,D	0.84,0.16	
Rv1634	1839306	47	A,G	I,V	0.88,0.12	
Rv2688c	***3005185***	***156***	***T,G***	***T,P***	***0.92,0.08***	
Rv1273	1422666	462	C,T	G,E/R	0.89,0.11	
Rv1273	***1423527***	***175***	***A***	***I***		***1***
Rv1273	***1423699***	***118***	***C,T***	***G,S***	***0.02,0.98***	
Rv1819	2062922	603	C,T	V,I	0.16,0.84	
Rv1877	***2127504***	***534***	***C,T***	***T,I***	***0.04,0.96***	
Rv1250	***1394764***	***196***	***G***	***A***		***1***
Rv1250	1395010	278	A,G	R,G	0.84,0.16	
Rv1250	***1395131***	***318***	***T***	***V***		***1***
Rv1250	***1395175***	***333***	***T***	***S***		***1***

The frequencies are those found in a dataset of 1526 drug susceptible isolates from Karonga, Malawi [35]. Significant allele frequencies (in bold) are based on a cut-off of 0.1.

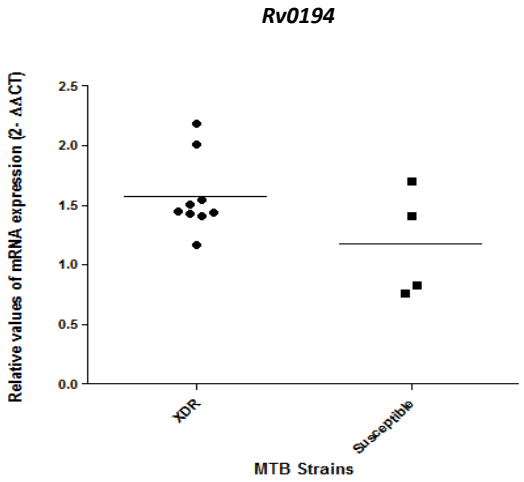
Table 3. XDR TB strains with unique SNPs in efflux pump genes

XDR strain	Lineage	No of ATT drugs to which it is Resistant	Efflux gene SNP profile
X13	CAS1-Delhi	I R P S E F K Ak	*Rv1250* V318X
X17	Other Lineage 1	I R P S E F K Ak	*Rv0194* G170V; *Rv1273* S118G, I175T; *Rv1877* I534T
X32	EAI3-IND	I R P E F Cap K Ak Em	*Rv1217* del 41nt
X33	CAS1-Delhi	I R P S E F Cap K Ak	*Rv1250* V318A
X37	CAS1-Delhi	I R P S E F K Ak	*Rv1250* S333A
X47	CAS1-Delhi	I R P S E F Cap K Ak Em*	*Rv1258* P369T
X49	Other Lineage 1	I R P F K Ak	*Rv1217* L151R
X55	CAS1-Delhi	I R P S E F K Ak	*Rv1273* S118G
X57	Other Lineage 1	I R P S E F K Ak	*Rv1258* G391R; *Rv1273* I175T; *Rv1877* I534T

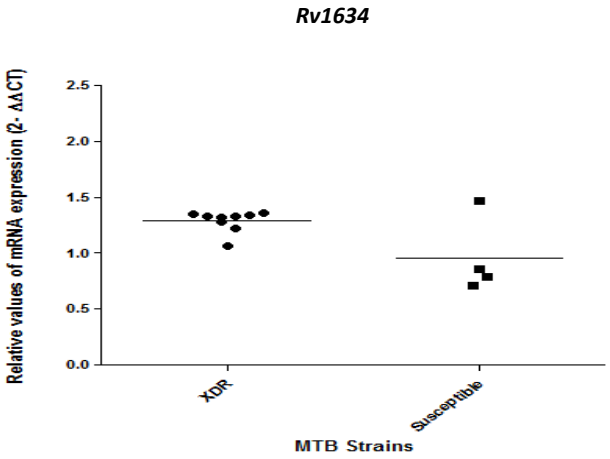
I=Isoniazid, R= Rifampicin, P= Pyrazinamide, S= Streptomycin, E= Ethambutol, F= Fluoroquinolones, Cap= Capreomycin, K= Kanamycin, Ak= Amikacin, Em= Ethionamide; * Total Drug Resistant


Tree scale: 0.1
X48 CAS1-Delhi
X8 EAI3-IND
X7 CAS1-Delhi
X16 CAS1-Delhi
X1 Oth. Lin. 3
X44 CAS1-Delhi
X55 CAS1-Delhi
X43 CAS1-Delhi
X21 CAS1-Delhi
X6 CAS1-Delhi
X42 CAS1-Delhi
X22 CAS
X18 CAS
X4 CAS2
X10 CAS1-Delhi
X45 Ot. Lin. 3
X14 CAS1-Delhi
X12 CAS1-Delhi
X47 CAS1-Delhi
X33 CAS1-Delhi
X13 CAS1-Delhi
X60 Ot. Lin. 3
X11 CAS1-Delhi
X40 CAS1-Delhi
X37 CAS1-Delhi
X9 Ot. Lin. 3
X61 CAS1-Delhi
X39 CAS1-Delhi
X46 X3
X58 T1
X41 T1
X56 T1
X57 Ot. Lin. 1
X17 Ot. Lin. 1
X49 Ot. Lin. 1
X5 EAI-IND
X32 EAI-IND

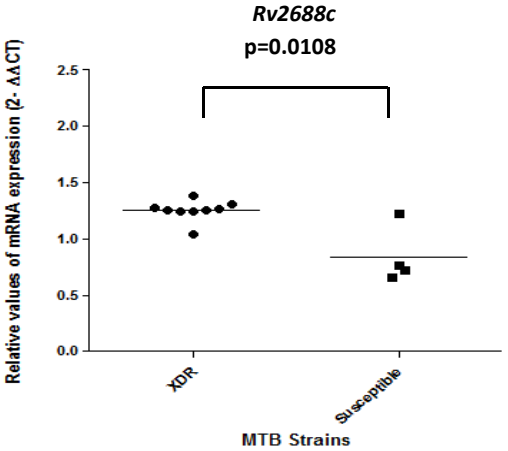
A


Rv0194
Relative values of mRNA expression (2- ΔΔCT)
2.5
2.0
1.5
1.0
0.5
0.0
XDR
Susceptible
MTB Strains

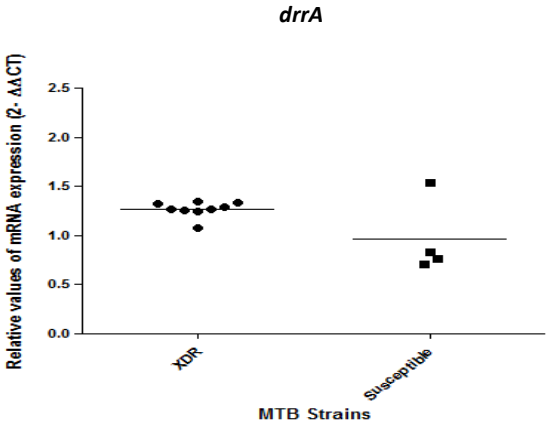
B


Rv1634
Relative values of mRNA expression (2- ΔΔCT)
2.5
2.0
1.5
1.0
0.5
0.0
XDR
Susceptible
MTB Strains

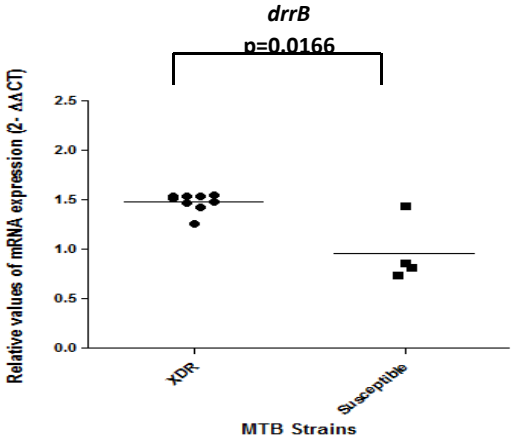
C


Rv2688c
p=0.0108
Relative values of mRNA expression (2- ΔΔCT)
2.5
2.0
1.5
1.0
0.5
0.0
XDR
Susceptible
MTB Strains

D


drrA
Relative values of mRNA expression (2- ΔΔCT)
2.5
2.0
1.5
1.0
0.5
0.0
XDR
Susceptible
MTB Strains

E


drrB
p=0.0166
Relative values of mRNA expression (2- ΔΔCT)
2.5
2.0
1.5
1.0
0.5
0.0
XDR
Susceptible
MTB Strains