INTERNATIONAL JOURNAL OF APPLIED AND EXPERIMENTAL BIOLOGY
(ISSN PRINT: 2790-6523; ISSN ONLINE: 2790-6531)
2022, Vol. 1, NO. 2, 49-57.
doi: https://doi.org/10.56612/ijaeb.v1i2.5

Genome-wide analysis to detect multi-drug resistance genes in *Mycobacterium tuberculosis* strains SWLPK and MNPK resourced from Pakistan

Maria Rafique[1], Muhammad Ibrahim[1], Sumaira Kanwal[1], Habib Bokhari[2], Farrukh Jamil[1], Annam Hussain[1], Muhammad Asif Rasheed[1*]

[1]Genomics and Computational Biology Laboratory, Department of Biosciences, COMSATS University Islamabad, Sahiwal Campus, Sahiwal, Pakistan
[2]Kohsar University Murree, KUM Admin/Academic Block, Kashmir Point Campus Murree, Punjab - Pakistan.

Abstract
Development of multidrug-resistant tuberculosis is the after effect of various mutational occasions, which leads to the development of protection from hostile to multiple tuberculosis drugs. In this study, we identified drug resistance genes, their evolutionary analysis, mutational variation, and docking to characterize the drug target potentials of two *Mycobacterium tuberculosis* strains SWLPK and MNPK resourced from Pakistan. For this purpose, we used different bioinformatics tools including the RAST server for annotation, and UniProt, NCBI, BLAST and MUSCLE for data retrieval and analysis. Evolutionary relationships were drawn using MEGA 7. A 3D structure was modelled by I-TASSER, while refinement and minimizations were performed using the UCSF Chimera 1.14.1. Moreover, the Ramachandran plot was used to check the quality of the proteins, while PatchDock was used for the docking analysis. Based on the comparison with the reference genome (*M. tuberculosis* H37RV), the SWLPK and MNPK strains encoded 24 multi-drug resistance genes. The drug resistance genes of nearby strains had developmental relatedness and comparable useful attributes imperative to their ecological specialties. The docking analysis revealed that the proteins accurately bound at their binding region just like the reference protein H37Rv (NuoG). We identified 24 multi-drug resistance genes in the SWLPK and MNPK strains. Moreover, there were a few missing drug resistance genes found in H37Rv, which were not present in the MNPK and SWLPK strains. The 24 genes reported in the MNPK and SWLPK strains may have a major contribution in drug resistance.

HANDLING EDITOR
Muhammad Ashraf

ARTICLE HISTORY
Received: 23 Oct 2021
Accepted: 5 Feb 2022
Published: 15 Mar 2022

KEYWORDS
Antibiotics;
Multi-drug resistance;
Genomics;
In silico analysis;
Strains

SUPPLEMENTARY MATERIAL
The supplementary material included with this manuscript can be downloaded from the journal website.

DISCLAIMER
The information on peer-review and usage of supplementary material can be found on the journal website

Introduction

Antibiotic resistance happens when microorganisms have or build-up the capacity to bypass the components of medications utilized against them. Contaminations brought about by anti-microbial safe microorganisms are typically harder to treat and their backslide capacity can cause critical horribleness and mortality (Van Hoek et al., 2011; Christaki et al., 2020). Protection from anti-toxins is normally the consequence of delayed utilization of anti-toxins, target changes (target substitution, target site transformations, target site enzymatic adjustments, target site assurance, target overproduction or target sidestep), and diminished anti-microbial collection, because of either diminished penetrability or additionally improved efflux (Van Hoek et al., 2011; Christaki et al., 2020). The utilization of antimicrobial

***CONTACT** Muhammad Asif Rasheed, asif.rasheed@cuisahiwal.edu.pk, ☎ +92-40-4305001-3: Department of Biosciences, COMSATS University Islamabad, Sahiwal Campus, 57000, Off GT Road, Sahiwal, Pakistan

TO CITE THIS ARTICLE: Rafique, M., Ibrahim, M., Kanwal, S., Bokhari, H., Jamil, F., Hussain, A., Rasheed, M.A. (2022). Genome-wide analysis to detect multi-drug resistance genes in *Mycobacterium tuberculosis* strains SWLPK and MNPK resourced from Pakistan. *International Journal of Applied and Experimental Biology* 1(2):49-57

specialists has diminished dreariness and mortality of people, and it contributed generously to human's expanded life expectancy. Antibiotics can tweak the protein combination, such as aminoglycoside, chloramphenicol, macrolide, streptothricin, and antibiotic medication or interface with the union of DNA and RNA, e.g., rifampin and quinolone. Other antibacterial gatherings are utilized to restrain the phone divider and energy digestion of the microbial cell, e.g., β-lactam, glycopeptide, sulfonamide and trimethoprim (Komolafe, 2004; Van Hoek et al., 2011; Christaki et al., 2020).

The disclosure and possible acquaintance of antimicrobial specialists with clinical medication was one of the best clinical victories of the 20th century that reformed the treatment of bacterial diseases. Antibiotic resistance has become a global public health concern, because the organisms that cause infections, are becoming resistant to most of prescribed antibiotic treatments, resulting in prolonged illness and greater risk of death. Another worrying aspect is that antibiotic resistance has developed over time, from resistance to single classes of antibiotics to multidrug resistance, and thus extreme drug resistance is increasing the challenge for the development of more effective antibiotics (Komolafe, 2004; Marti et al., 2014). *Mycobacterium tuberculosis* and different individuals from the *M. tuberculosis* complex utilize a few techniques to oppose the activity of antimicrobial agents (Gygli et al., 2017). The genes that are related to these capacities, have been found in *M. tuberculosi*s (Kwon et al., 1995; Cole et al., 1996). Also, hereditary investigations have shown that opposition in *M. tuberculosis* against antimicrobial medications is the outcome of unconstrained changes in genes that encode either the objective of the medication or catalysts that are engaged with drug initiation (Kwon et al., 1995; Cole et al., 1996).

Opposition related point changes, and erasures or additions have been portrayed for all first-line drugs (*isoniazid, rifampin, pyrazinamide, ethambutol,* and *streptomycin*) and more up to date sedates (*ethionamide, fluoroquinolones, para-amino salicylic acid and cycloserine*) (Heym et al., 1994; Stover et al., 2000; Vester et al., 2001). Since MDR strains are the consequence of total transformations, development of *M. tuberculosis* can effectively be controlled in the host by a corresponding treatment with more than one anti-microbium. In this way, treatment regimens that comprise three to four medications are used regularly to treat patients with tuberculosis (Somoskovi et al., 2001). In 2012, there were about 450,000 instances of multidrug safe TB, and 170,000 passings were because of it. Widely drug-safe (XDR) TB alludes to MDR-TB strains that are impervious to fluoroquinolones and second-line injectable drugs (Seung et al., 2015).

Since 2006, presence of significantly more safe strains of *M. tuberculosis* named as XDR-TB has been perceived. These strains are likewise impervious to any fluoroquinolone and to any event of the injectable second-line drugs such as kanamycin, capreomycin, and amikacin. Even more as of late, a really stressing circumstance has arisen with the depiction of *M. tuberculosis* strains that have been discovered impervious to all anti-infection agents accessible for testing (Velayati, et al., 2009; Prasad et al., 2017). As of late, progression of the next generation sequencing has contributed a lot to acquire the genome arrangement of *M. tuberculosis* strains SWLPK and MNPK. A promising proof of MDR genes in Pakistani resourced strains and broad computational investigation with reference strains might provide better information on active components of hostile to TB drugs and the advancement of drug resistance.

Material and Methods

Preprocessing of genome sequence data and annotation

The genome arrangements of two *Mycobacterium tuberculosis* strains, SWLPK and MNPK were refined and utilized for explanation investigation. The explanation was led utilizing the RAST comment worker with default boundaries. The clarified genome of *M. tuberculosis* SWLPK, MNPK and H37Rv were acquired utilizing the SEED watcher of the RAST worker for investigation analysis (Aziz et al., 2008). The data include RNAs, coding sequences (CDS), GC contents, and general feature categories (**Table 1 & 2**).

Identification of MDR genes

The multi-drug resistance genes of *M. tuberculosis* H37Rv were acquired from the NCBI, UniProt and the writing. The genes were utilized as snare groupings in a neighbourhood BLAST search of the RAST worker utilizing H37Rv as a reference strain to look at the multi-drug resistance genes in the genome sequence of *M. tuberculosis* strains SWLPK and MNPK, as described elsewhere (Yar et al., 2018).

Table 1: List of bacterial genomes and their source of separation

Species and strain	Resource	Reference
M. tuberculosis H37Rv	Bacteria	Cumus et al. 2002
M. tuberculosis SWLPK	Bacteria	Asma et al. 2018
M. tuberculosis MNPK	Bacteria	Asma et al. 2018

Evolutionary tree construction

A remarkable set-up of molecular biology and NGS examination instrument MUSCLE, were utilized to adjust the genome sequences to identify the degree of correspondence between them. The Warmth maps in straight geography were additionally acquired to check the degree of rate similarity (Tamura et al., 2011; Van Hoek et al., 2011). Transformative connections among different natural species or different substances attracted the type of a phylogenetic construction tree utilizing the MEGA 7 tool (Tamura et al., 2011; Babicki et al., 2016). To construct the phylogenetic tree of all 24 multi-drug resistance (RpoB, RpoA, RpoC, KatG, InhA, FabG, AhpC, KasA, RpsA, FurA, EmbA, EmbB, EmbC, PncA, GyrA, GidB, PrcA, EthR, ParE1, ThyA, Ddl, CycA, RibD, and FolC) proteins retrieved from three strains of *M. tuberculosis* SWLPK, MNPK, H37Rv, the neighbour-joining method with standard examination inclinations, like phylogeny recreation, factual strategy, and Poisson model, were chosen.

In silico analysis of protein domains and functional characterization

The recognizable proof of domains can give the knowledge for their capacity and affiliation. The NCBI's Conserved Domain Database (CDD), InterPro and UniProt are assets for the explanation of protein sequences with the area of conserved domains. The useful locales were gathered from the space footprints (Apweiler et al., 2001; Bateman et al., 2015).

The amino acid sequences of *M. tuberculosis* drug impediment genes were transferred in the FASTA format in the CD pursuit apparatus of the NCBI CDD, InterPro, and UniProt databases (Sussman et al., 1998). An inquiry was performed against the position-specific score matrices (PSSMs) for the quick distinguishing proof of well-preserved domains in protein groupings (Apweiler et al., 2001).

Protein structure retrieval

The RAST server was used for annotation and protein sequence retrieval of both genomes (SWLPK and MNPK), and the reference sequence *M. tuberculosis* H37Rv was collected from NCBI. The conversions of both *M. tuberculosis* (SWLPK and MNPK) protein sequences into PDBs structures using available template structure (PDB ID: C1II0 and C5TW1), respectively, were also accomplished via the threading method PHYRE2 online webserver (Bateman et al., 2015). Twenty four (24) MDR protein sequences were retrieved using the UniProt database. Of 24 MDR proteins, 17 protein 3D structures (RpoA, RpoC, KatG, InhA, AhpC, KasA, RpsA, EmbA, EmbB, PncA, GidB, PrcA, EthR, GyrA, ThyA, Ddl, and FolC) (as mentioned in Table S4) were identified using the RCSB Protein Data Bank (PDB). While 7 (FabG, FurA, ParE1, CycA, RibD, EmbC, and FurA) protein structures were found on the PDB database, so, their 3D structure prediction was performed using the online webserver threading tool, PHYRE2. To verify all these results, we further retrieved one protein (NuoG) NADH-quinone oxidoreductase subunit G from the reference protein *M. tuberculosis* H37Rv and selected to authenticate, validate and identify our molecular docking results, their binding regions, and binding energy values, using the Uniprot database with ID: P9WIV9; their 3D structure was modeled with an accessible template on database (PDB ID: 6TG9) using the threading method I-TASSER webserver *(http://zhanglab. ccmb. med. umich. edu/I-TASSER)* online webserver database. Refinement and minimizations of *M. tuberculosis* strains SWLPK and MNPK, and H37Rv (NuoG) were performed using the UCSF Chimera 1.14.1 (Sussman et al., 1998).

Protein structure validation and minimization

Different online tools such as MolProbity for good/bad angles were used for structure assessment and for validation of 17 multi-drug resistance proteins. The MolProbity results suggested that the structures were quite reliable (Chen et al., 2016). The RAMPAGE tool for poor/favored rotamers (Wang et al., 2016), and physicochemical and stereo-chemical properties (rotamers, outliers, stability index, and bonding patterns) were calculated to validate the predicted structures. The UCSF Chimera 1.14.1 was used for the minimization and refinement of 17 multi-drug resistance proteins (Schneidman-Duhovny et al., 2005).

Table 2: Genome and their annotated features isolated from *M. tuberculosis*

Genome	MNPK	SWLPK	H37Rv
Genome size (bp)	4,409,295	4,391,906	4,411,532
GC content	65.6	65.6	65.6
RNAs	47	48	50
No. of contigs	20	12	1
No. of coding sequences	4309	4295	4367

Molecular docking and protein residues interaction analysis

Protein-protein docking was performed to check the binding regions of 17 multi-drug resistance proteins target (RpoB, RpoA, RpoC, KatG, InhA, AhpC, KasA, RpsA, EmbA, EmbB,, PncA, GyrA, PrcA, EthR, ThyA, Ddl, and FolC) and against 3D protein structures of all three *M. tuberculosis* strains, H37Rv (NuoG, Uniprot ID: P9WIV9), SWLPK, and MNPK, to get the ideal local conformation for predicting a sufficient interaction. The reference protein *M. tuberculosis* H37Rv NuoG (NADH-quinone oxidoreductase subunit G) was performed using PatchDock (Wallace et al., 1995; Schneidman-Duhovny et al., 2005). LigPlus was used for molecular docking analysis, to identify binding interactions of *M. tuberculosis* SWLPK , MNPK, and H37RV (NuoG) and 17 targets multi-drug resistance proteins.

Results and Discussion

Preprocessing of genome sequence data and annotation

High quality functional annotation of *M. tuberculosis* strains, SWLPK and MNPK, was obtained from the RAST server. The results contained the information on genome size, domain, GC contents, number of contigs, number of coding sequences, and number of RNAs. The genome size of *M. tuberculosis* SWLPK was lower (4,391,906 bp) than that of the reference genome H37RV (4,411,532 bp), while the genome of *M. tuberculosis* MNPK (4,409,295 bp) was equal to the reference genome. The GC content of both *M. tuberculosis,* SWLPK and MNPK strains, was 65.6%. The numbers of contigs in *M. tuberculosis* SWLPK and MNPK were 12 and 20, respectively. Similarly, numbers of coding sequences of *M. tuberculosis* strains SWLPK and MNPK were 4295 and 4309, respectively, while the numbers of RNAs of both genomes were 48 and 47, respectively (Migliori et al., 2012).

Identification of MDR genes

The *M. tuberculosis* strains SWLPK and MNPK encoded 24 multi-drug resistance genes against notorious drugs such as rifampicin, isoniazid, ethionamide, streptomycin, ethambutol, pyrazinamide, fluoroquinolones and para-amino salicylic acid (**Table S1**). The virulence determinants of the bacteria are based on molecular features or cellular localization such as complex lipids, lipoglycans and mycolic acids associated pathogenesis, etc. (Blanchard, 1996). Nonetheless, genetic studies have described that the obstruction of tuberculosis to antimicrobial medications is the result of unconstrained transformations in the genes that encode either the objective of the medication, or chemicals that are associated with drug initiation. Obstruction related point changes, erasures, or additions have been depicted for all the first-line drugs including isoniazid, rifampin, pyrazinamide, ethambutol, and streptomycin. Additionally, a few second-line and more current medications including ethionamide, cycloserine and fluoroquinolones are likewise remembered for setting of opposition (**Table S1 & S2**).

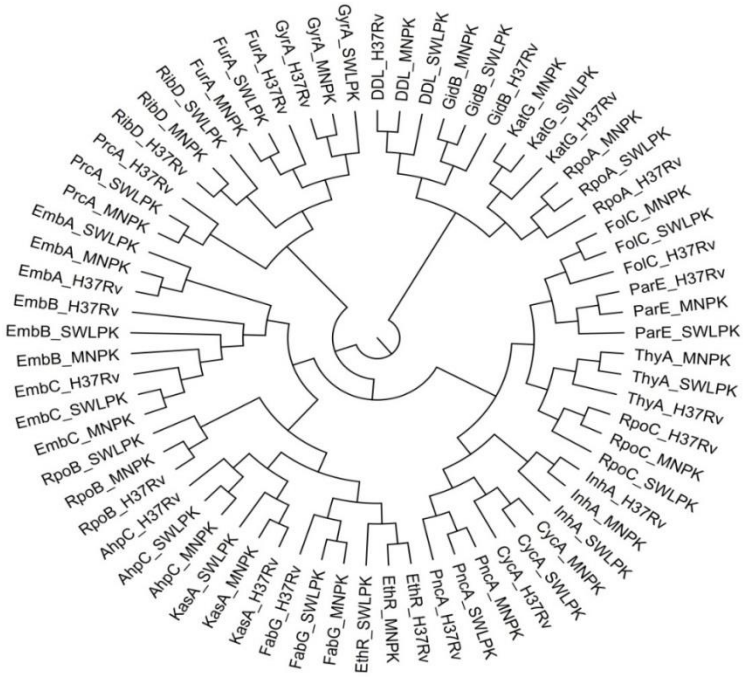

Figure 1. An evolutionary tree of *M. tuberculosis* H37Rv, *M. tuberculosis* MNPK and *M. tuberculosis* SWLPK showing similarities and differences with 24 multi-drug resistance proteins

In silico analysis of protein domains and functional characterization

Domains are diverse structural and/or functional units in a protein. Protein sequences were submitted to InterPro (Apweiler et al., 2001). The resistance related proteins against rifampin antibiotic include rpoB and rpoB, rpoA, and rpoC belonging to the same family DNA-directed RNA polymerase, and they function as β-subunit of RNA polymerase pyrazinamidase (Smith et al., 2012; Palomino et al.,2014; Cui et al., 2016). Moreover, the proteins against ethambutol including EmbA, EmbB, and EmbC belong to arabinose_trans_C family and code for arabinosyltransferase, while the proteins KatG, inhA, FabG, ahpC, kasA, and furA were found to be related to resistance against isoniazid and RpsA and PncA against pyrazinamide. Furthermore, the proteins ParC and ParE1, showed resistance against fluoroquinolones. The protein EthR mediates resistance to ethionamide. Moreover, the proteins Ddl and CycA take part in resistance against cycloserine, and CycA encodes a D-alanine transporter, which was believed to be partially responsible for resistance to cycloserine in *M. bovis* BCG. While GyrA was found to be against ofloxacin, GidB against streptomycin drug, and ThyA, and FolC against para-amino salicylic acid (Smith et al., 2012; Palomino et al., 2014; Cui et al., 2016). Although MDR proteins of SWLPK, MNPK and the reference strain H37Rv, their domains and families are similar and show that they may have association in drug resistance and virulence as H37Rv, the existence of domain leads to a variety of biological functions, such as similar domains may occur in proteins with distinct functions (Mulder et al., 2010)

Evolutionary analysis and association

Twenty-four MDR proteins that encode in H37Rv and the MDR proteins identified in the strains, SWLPK and MNPK, were aligned using Heatmap. Drug resistance genes of the reference genome (H37Rv) including *ahpC, embA, ethR, folC, inhA, parE, prcA, rpoA, rpoC* and *GyrA* were completely (100%) similar to those of *M. tuberculosis* SWLPK and MNPK. Moreover, genes *thyA, rpoB, ribD, RpsA, KatG*, *furA* and *kasA* were 99.5% identical. However, the multi-drug resistance genes *gidB*, *pncA*, and *FabG*, showed 94.9%, 82.3%, and 19.1% similarity, respectively, to the reference genome *H37Rv*. The gene *pncA* was 50% similar to *parE*, while *folC* was 33.3% similar to *parE1,* and 17.4% to *GyrA* in sequence alignment (Verhaak et al., 2006; López et al., 2007; Velmurugan et al., 2007). These results specifically indicate that multi-drug resistance proteins' genes have evolutionary association pathogenic properties with those present in both *M. tuberculosis* strains SWLPK and MNPK (**Table S1**).

Similarly, the evolutionary tree analysis was also conducted using the alignment of sequences obtained through different parameters of the tree building method using MEGA10 to check similarity between sequences compared with the reference genome H37RV (**Figure 1**). It consisted of 3 main clusters showing multiple clades of organisms defined by their common biological ancestor, irrespective of how closely they are related or not related to each other. The tips of the tree represent groups of descendent taxa (often species or genes), while the nodes on the tree represent the common ancestors of those descendants. In a tree construction, two descendants that split from the same node are called sister groups. Moreover, a taxon outside the cluster of interest is called out-group. All the members of the group of interest are more closely related to each other than to the out-group. The results of the evolutionary analysis reveal that comparatively all MDR associated genes of MNPK and SWLPK have a close association with H37Rv.

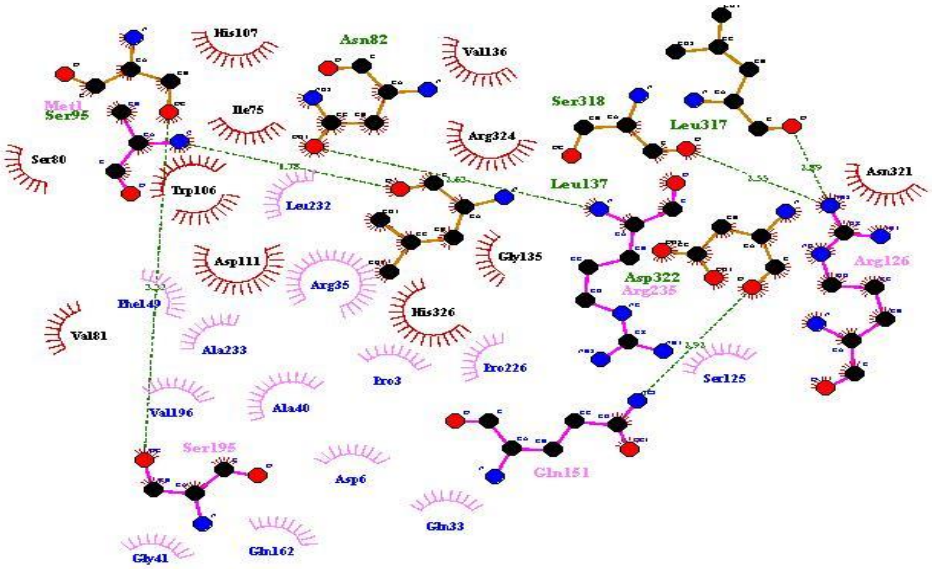

Figure 2. Molecular docking results of ThyA MDR protein with those of *M. tuberculosis* SWLPK

Our phylogenetic approach made it possible to test if the MNPK and SWLPK strains were likely to be associated with drug resistance or not. We detected strong convergent-evolution signals for almost all drug resistance genes.

Protein structures validation and minimization

Out of 24 multi-drug resistance proteins, the 3D structures of 17 proteins were present on the Protein Data Bank (PDB) in complexes. Their 3D structures were separated from complexes and used for further study (**Table S3**). Protein structures were assessed using the MolProbity-Rampage tool, and minimized and refined via the UCSF Chimera. The MolProbity results suggested that the structures were quite reliable. The Rampage tool showed poor/favored rotamers. Physicochemical and stereo-chemical properties (rotamers, outliers, stability index, and bonding patterns) were calculated to validate the predicted structures. The assessment results suggested that the structures were quite reliable for molecular docking.

Protein binding interaction analysis

Molecular docking results analysis showed that multi-drug resistance gene *RpoA* had a binding energy of -589.61 kJ/mol for *M. tuberculosis* MNPK, *ThyA* with a binding score of -52.40 kJ/mol for *M. tuberculosis* SWLPK and *RpoB* with a binding energy of -54.72 kJ/mol for *M. tuberculosis* H37Rv, respectively (**Table S3**). For binding interaction analysis, different tools were used such as LigPlus and PDBSum database. Multi-drug resistance protein residues with a best interaction were observed and compared with each other such as *M. tuberculosis* SWLPK, *M. tuberculosis* MNPK, and the reference protein NuoG of *M. tuberculosis* H37Rv (**Table S4**). Top 3 (ThyA-SWLPK, RpoA-MNPK, and RpoB-H37Rv) MDR proteins were selected based on their highest binding energy score, and protein-protein interactions were observed in Figures using LigPlot. These results showed their binding sites, binding distance, and hydrophobic residues as shown in all 3 Figures (**Figures 2, 3 and 4**) (Palomino and Martin, 2005)**.**

Molecular docking analysis

To identify protein binding sites of *M. tuberculosis* SWLPK, MNPK, and H37Rv (NuoG) with 17 multi-drug resistance proteins targets such as RpoB, RpoA, RpoC, KatG, InhA, AhpC, KasA, RpsA, EmbA, EmbB, PncA, GyrA, PrcA, EthR, ThyA, Ddl, and FolC 3D structures, molecular docking analysis was performed by employing PatchDock with default parameters. The docked complexes were passed through the FireDock server to pick the best models from the short-list of 10 complexes. FireDock resulted in 10 models for the *M. tuberculosis* SWLPK, MNPK, and H37RV (NuoG) proteins, and 17 multi-drug resistance protein complexes and structures with best binding scores or binding energy. The reference H37Rv (NuoG) protein was selected based on its role in inhibition of bacterial virulence. Different researchers have reported that the expression of *M. tuberculosis* NuoG in nonpathogenic

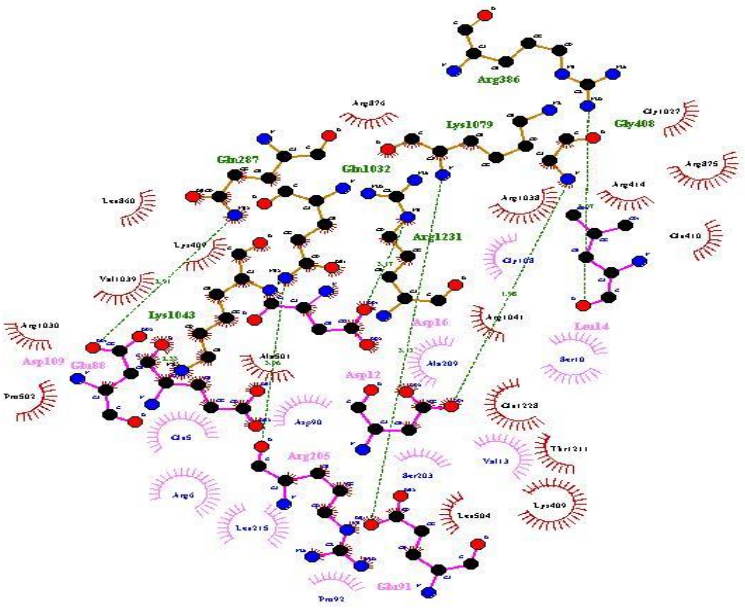

Figure 3. Molecular docking results of RpoA MDR protein with those of *M. tuberculosis* MNPK

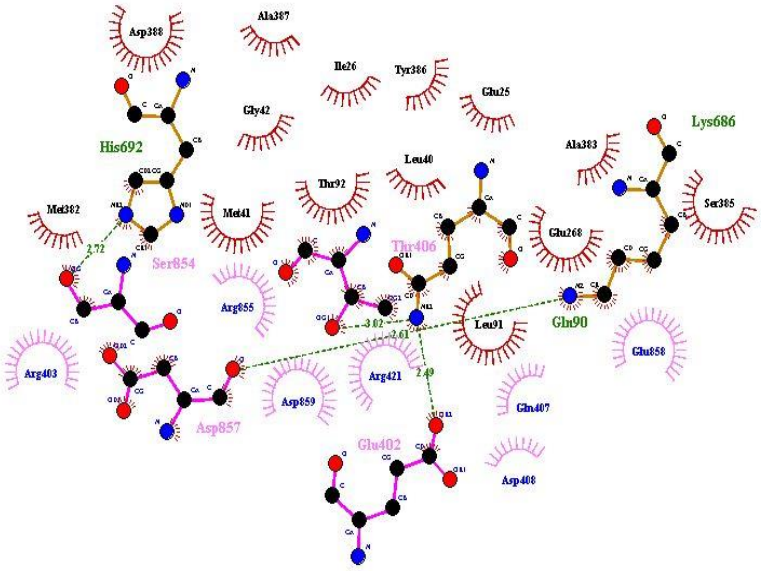

Figure 4. Molecular docking results of RpoB MDR protein with those of *M. tuberculosis* H37Rv

mycobacteria endowed them with the ability to inhibit apoptosis of infected human or mouse macrophages, and showed increased virulence in a SCID mouse model (Velmurugan et al., 2007; Kopez et al., 2007). Conversely, deletion of NuoG in *M. tuberculosis* ablated its ability to inhibit macrophage apoptosis and significantly reduced its virulence in mice. These results identify a key component of the genetic basis for an important virulence trait of *M. tuberculosis* and support a direct causal relationship between virulence of pathogenic mycobacteria and their ability to inhibit macrophage apoptosis. Moreover, NuoG of *M. tuberculosis*, encodes a subunit of the type-I NADH dehydrogenase complex, a critical bacterial gene for inhibition of host cell death. Molecular docking results of the MDR proteins were observed and their binding energy score analysis was performed (as mentioned in Table S5).

Conclusion

In general, progressions in next-generation sequencing advances have furnished a comprehensive genomics information with the accessibility of numerous new bacterial genome successions. The computational forecast of drug resistance genes in *M. tuberculosis* strain SWLPK genomes will be valuable for recognizing new pathogenicity-related genes and their components inside the genome of microbes that contain multi-drug resistance genes. Our results show that the identification of multi-drug resistance genes (MDR) and their sequence similarity may indicate similar roles of these proteins in novel molecular docking interaction analysis and their evolutionary association. Theses novel results of molecular docking of *M. tuberculosis* SWLPK, MNPK, and H37Rv (NuoG) with 17 multi-drug resistance proteins and phylogenetic tree will be helpful for more advanced studies. A few drug resistance genes of *M tuberculosis* previously revealed in the reference strain *M. tuberculosis* H37Rv were not revealed; for example, *iniA* and *faDE24* were linked with isoniazid MDR genes *embR, rmlD*, and *iniA*, which are *ethambutol* related genes, *RRS* and *RpsL*. These are associated with streptomycin. In addition, the missing MDR genes may affect the viability of antimicrobial specialists, especially when their essence has not been diagnosed. Thus, changes in *RpsL* and *RRS* are the significant components of protection from streptomycin. Among the transformations announced in *RpsL*, a replacement in codon 43 from lysine to arginine has been the most regularly revealed one. This change delivers a significant level of protection from streptomycin. In *RRS,* the most widely recognized transformations were found to be happened around the nucleotides 530 and 915. There stayed a significant level of strains impervious to streptomycin that need changes in both of these two genes, suggesting extra-mechanisms of resistance. However, the absence of multi-drug resistance genes like *embR, FadE24, RmlD, Rrs, RpsL IniA,* in *M. tuberculosis* SWLPK and *M. tuberculosis* MNPK demonstrate that these genes may have been autonomously obtained, and various parts identified with pathogenicity, have a range assurance or potential specialty variation. The presence of these notable MDR genes in Pakistani strains will be useful for additional investigation on ID and portrayals, and will advance the tuberculosis hereditary information sources. The missing of *RmlD, rpoR, faDE24, rpsL iniA,* and *rrs* genes demonstrates that further examinations are required and these genes may not be utilized as a marker for examining Pakistan-resourced strains. In this way, contrasting outcomes from genome and developing a waitlist of

normal hits might be the best method to think about bacterial genome sequences. Although some genes were absent, due to the presence of key drug resistance genes, we have to carefully design a medicine avoiding further spread of drug resistance genes.

Acknowledgement

This research was supported by the Higher Education Commission of Pakistan through an NRPU project titled "Sequencing and mapping of *Mycobacterium tuberculosis* from patients in Pakistan using next-generation sequencing (No-20-3658/R&D/HEC/14).

References

Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Birney, E., Biswas, M. (2001). The interPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Research* 29:37–40.

Aziz, R.K., Bartels D. (2008). Rapid annotation using subsystem technology rapid annotation using subsystem technology. *BMC Genomics* 9:75.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., Wishart, D.S. (2016). Heatmapper: web–enabled heat mapping for all. *Nucleic Acids Research* 8:44(W1):W147–53.

Bateman, A., Martin, M.J., O'Donovan, C., Magrane, M., Apweiler, R. (2015). UniProt: A hub for protein information. *Nucleic Acids Research* 43:D204–D212.

Blanchard, J.S. (1996). Molecular mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Annual Review of Biochemistry* 65:215–239.

Camus, J.C., Pryor, M.J., Médigue, C., Cole, S.T. (2002). Re–annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* 148:2967–2973.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C. (2010). MolProbity: All–atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography* 66:12–21.

Christaki, E., Marcou, M., Tofarides, A. (2020). Antimicrobial resistance in bacteria: Mechanisms, evolution, and persistence. *Journal of Molecular Evolution* 88:26–40.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393:537–544.

Cui, Z.J., Yang, Q.Y., Zhang, H.Y., Zhu, Q., Zhang, Q.Y. (2016). Bioinformatics identification of drug resistance–associated gene pairs in *Mycobacterium tuberculosis*. *International Journal of Molecular Science* 17:1417.

Gygli, S.M., Borrell, S., Trauner, A., Gagneux, S. (2017). Antimicrobial resistance in *Mycobacterium tuberculosis*: Mechanistic and evolutionary perspectives. *FEMS Microbiology Reviews* 41:354–373.

Heym, B., Honoré, N., Schurra, C., Cole, S.T. (1994). Implications of multidrug resistance for the future of short–course chemotherapy of tuberculosis: A molecular study. *The Lancet* 344:293–298.

Komolafe, O. (2004). Antibiotic resistance in bacteria – an emerging public health problem. *Malawi Medical Journal* 2:63–67.

Kwon, H.H., Tomioka, H., Saito, H. (1995). Distribution and characterization of β–lactamases of mycobacteria and related organisms. *Tubercle and Lung Disease* 76:141–148.

López, G., Valencia A., Tress, M.T. (2007). Firestar—prediction of functionally important residues using structural templates and alignment reliability. *Nucleic Acids Research* 35:W573–W577.

Marti, E., Variatza, E., Balcazar, J.L. (2014). The role of aquatic ecosystems as reservoirs of antibiotic resistance. *Trends in Microbiology* 22:36–41.

Migliori, G.B., Centis, R., D'Ambrosio, L., Spanevello, A., Borroni, E., Cirillo, D.M., Sotgiu, G. (2012). Totally drug–resistant and extremely drug–resistant tuberculosis: The same disease? *Clinical Infectious Diseases* 54:1379–1380.

Palomino, J.C., Martin, A. 2014. Drug resistance mechanisms in *Mycobacterium tuberculosis*. *Antibiotics* 3:317–340.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E. (2004). UCSF chimera–A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25:1605–1612.

Prasad, R., Singh, A., Balasubramanian, V., Gupta, N. (2017). Extensively drug–resistant tuberculosis in India: Current evidence on diagnosis and management. *Indian Journal of Medical Research* 145:271–293.

Schneidman–Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Research* 33:363–367.

Seung, K.J., Keshavjee, S., Rich, M.L. (2015). Multidrug–resistant tuberculosis and extensively drug–resistant tuberculosis. *Cold Spring Harbor Perspectives in Medicine* 5:a017863.

Smith, T., Wolff, K.A., Nguyen, L. (2012). Molecular biology of drug resistance in *Mycobacterium tuberculosis*. *Current Topics in Microbiology and Immunology* 374:53–80.

Somoskovi, A., Parsons, L.M., Salfinger, M. (2001). The molecular basis of resistance to isoniazid, rifampin, and pyrazinamide in *Mycobacterium tuberculosis*. *Respiratory Research* 2:164–168.

Stover, C.K., Warrener, P., VanDevanter, D.R., Sherman, D.R., Arain, T. (2000). A small–molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature* 405:962–966.

Sussman, J.L., Lin, D., Jiang, J., Manning, N.O., Prilusky, J., Ritter, O. Abola, E.E. (1998). Protein Data Bank (PDB): Database of three–dimensional structural information of biological macromolecules. *Acta Crystallographica Section D: Structural Biology* 54:1078–1084.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28:2731–2739.

van Hoek, A.H.A.M., Mevius, D., Guerra, B., Mullany, P., Roberts, A.P., Aarts, H.J.M. (2011). Acquired antibiotic resistance genes: an overview. *Frontiers in Microbiology* 2:203.

Velayati, A.A., Masjedi, M.R., Farnia, P., Tabarsi, P., Ghanavi, J., ZiaZarifi, A.H., Hoffner, S.E. (2009). Emergence of new forms of totally drug–resistant *Tuberculosis bacilli*: super extensively drug–resistant tuberculosis or totally drug–resistant strains in Iran. *Chest* 136:420–425.

Velmurugan, K., Chen, B., Miller, J.L., Azogue, S., Gurses, S., Hsu, T., Glickman, M., Jacobs, W.R., Porcelli, S.A., Briken, V. (2007). *Mycobacterium tuberculosis* NuoG is a virulence gene that inhibits apoptosis of infected host cells. *PLoS Pathogens* 3:e110.

Verhaak, R.G.W., Sanders, M.A., Bijl, M.A., Delwel, R., Horsman, S., Moorhouse, M.J., van der Spek, P.J., Löwenberg, B., Valk, P.J.M. (2006). HeatMapper: powerful combined visualization of gene expression profile correlations, genotypes, phenotypes and sample characteristics. *BMC Bioinformatics* 7:337.

Vester, B., Douthwaite, S. (2001). Macrolide resistance conferred by base substitutions in 23S rRNA. *Antimicrobial Agents and Chemotherapy* 45:1–12.

Wallace, A.C., Laskowski, R.A., Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Engineering, Design and Selection* 8:127–134.

Wang, W., Xia, M., Chen, J., Deng, F., Yuan, R., Zhang, X., Shen, F. (2016). Data set for phylogenetic tree and RAMPAGE Ramachandran plot analysis of SODs in *Gossypium raimondii* and *G. arboreum*. *Data Brief* 9:345–348.

Yang, J., Zhang, Y., Yang, J., Zhang, Y. (2015). I–TASSER Server: new development for protein structure and function predictions. *Nucleic Acids Research* 43:W174–W181.

Yar, A.M., Zaman, G., Hussain, A., Changhui, Y., Rasul, A., Hussain, A., Bo, Z., Bokhari, H., Ibrahim, M. (2018). Comparative genome analysis of 2 *Mycobacterium tuberculosis* strains from Pakistan: insights globally into drug resistance, virulence, and niche adaptation. *Evolutionary Biology* 14:1-9.