Analyses of Efflux pump proteins involved in *M. tuberculosis* and determination of their drug target potential

Enrollment No: 101507

Name of student: Pankaj Kapoor

Name of supervisor: Dr. Chittaranjan Rout

Bachelor of Technology

in

Bioinformatics

DEPARTMENT OF BIOTECHNOLOGY AND BIOINFORMATICS

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

LIST OF CONTENT

CERTIFICATE

This is to certify that the work titled "**Analysis of proteins involved in the efflux mechanism of Mycobacterium tuberculosis And their drug target potential.**", submitted by Pankaj Kumar (101507) for the reward of Degree of Bachelor's Of Technology (Bioinformatics) from **JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT** has been carried out under my supervision. The work has not been submitted partially or wholly to any other university or institute for the award of this year or any other degree and diploma.

Dr. Chittaranjan Rout

Assistant professor

Dept. of Biotechnology and bioinformatics

Jaypee University of Information Technology

Waknaghat , Solan

Himachal Pradesh-173234

AKNOWLEDGEMENT

We take this opportunity to express our profound gratitude and deep regards to our guide, **Dr. C. Rout** for her exemplary guidance, monitoring and constant encouragement throughout the course of this project . The blessing, help and guidance given by her time to time shall carry us a long way in the journey of life on which we are about to embark.

We are also thankful to our Head of Department, **Prof. R.S. Chauhan** for his constant encouragement and suggestions for motivating us to put in our best effort.

We also take this opportunity to express a deep sense of gratitude to our Institution, college faculty and staff members for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

Name: Pankaj Kumar

Enrollment No. : 101507

Date :

Place : JUIT, Waknaghat

List of tables

List of figures

Abbreviations:

M.tb	Mycobacterium tuberculosis
MDR	Multiple drug resistance
RND	Resistance nodulation division
MATE	Multidrug and toxic resistance
ABC	ATP binding cassette
EI	Efflux inhibitors
ATP	Adenosine triphosphate
BLAST	Basic alignment search tool

Chapter 1

1.1 Introduction

Mycobacterium tuberculosis was discovered by Dr. Robert Koch in 1882. Tuberculosis has existed throughout history, but the name has changed frequently over time. Tuberculosis (TB) is a bacterial infection spread through inhaling tiny droplets from the coughs or sneezes of an infected person. It is a serious condition but can be cured with proper treatment. TB mainly affects the lungs. However, it can affect any part of the body, including the bones and nervous system. (Dong park *et.al*. 2003)

Typical symptoms of TB (Storla *et.al.* 2008) include:
a persistent cough for more than three weeks that brings up phlegm, which may be bloody

- weight loss
- night sweats
- high temperature (fever)
- tiredness and fatigue
- loss of appetite

Mycobacterium tuberculosis, bacteria that causes *tuberculosis*, has been around for centuries. Recently, fragments of the spinal columns from Egyptian mummies from 2400 B.C.E. were found to have definite signs of the ravages of this terrible disease. Also called consumption, TB was identified as the most widespread disease in ancient Greece, where it was almost always fatal. But it wasn't until centuries later that the first descriptions of the disease began to appear. Starting in the late seventeenth century, physicians began to identify changes in the lungs common in all consumptive, or TB, patients. At the same time, the earliest references to the fact that the disease was infectious began to appear.

In 1720, the English doctor Benjamin Marten was the first to state that TB could be caused by "wonderfully minute living creatures." He went further to say that it was likely that ongoing contact with a consumptive patient could cause a healthy person to get sick. Although Marten's findings didn't help to cure TB, they did help people to better understand the disease.

The sanitarium, which was introduced in the mid-nineteenth century, was the first positive step to contain TB. Hermann Brehmer, a Silesian botany student who had TB, was told by his doctor to find a healthy climate. He moved to the Himalayas and continued his studies. He survived his bout with the illness, and after he received his doctorate, built an institution in Gorbersdorf, where TB patients could come to recuperate. They received good nutrition and were outside in fresh air most of the day. This became the model for the development of sanatoria around the world.

In 1865, French military doctor Jean-Antoine Villemin demonstrated that TB could be passed from people to cattle and from cattle to rabbits. In 1882, Robert Koch discovered a staining technique that allowed him to see the bacteria that cause TB under a microscope.

Until the introduction of surgical techniques to remove infected tissue and the development of x-rays to monitor the disease, doctors didn't have great tools to treat TB. TB patients could be isolated, which helped reduce the spread of the disease, but treating it remained a challenge. (Russell DG. 2012)

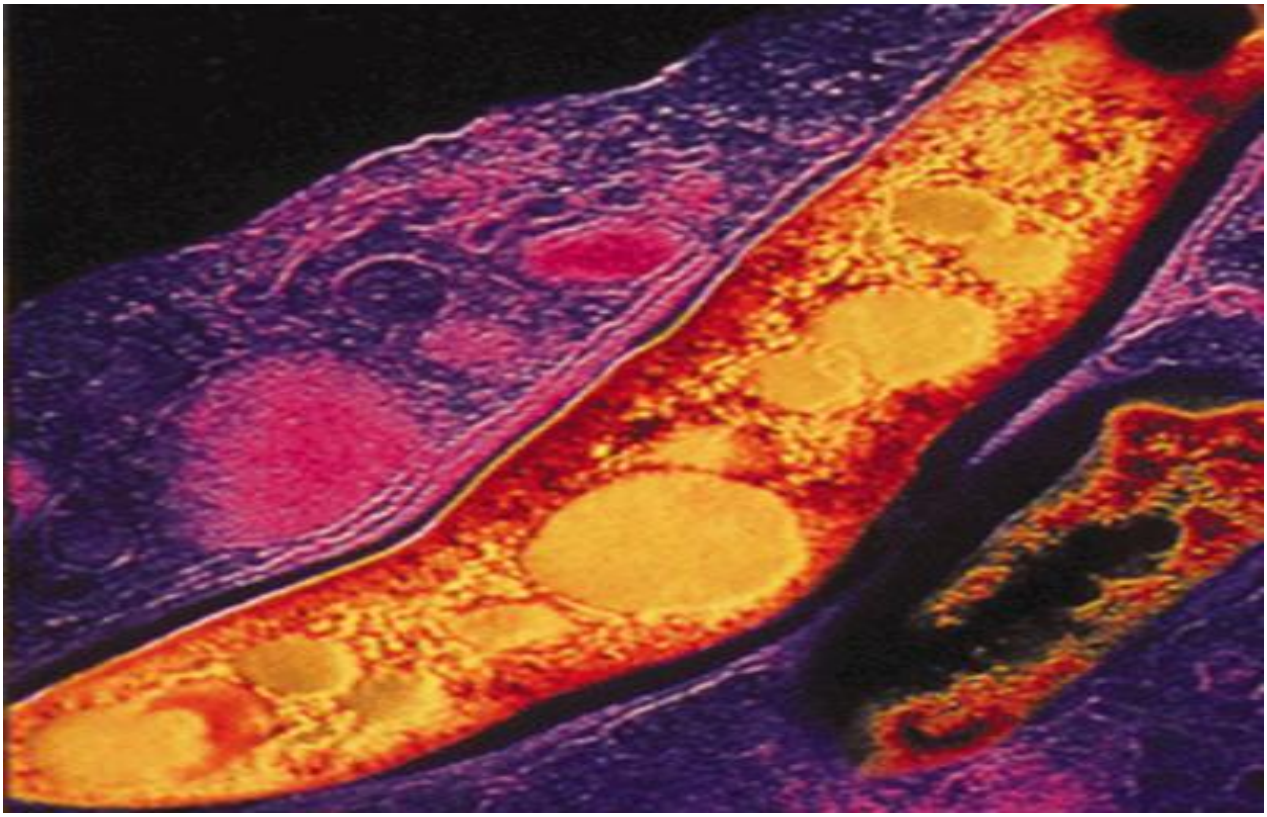


Fig. 1.1: the Mycobacterium tuberculosis strain

1.2 Important statistics (http://www.who.int/mediacentre/factsheets/fs104/en/)

Global:

- Tuberculosis is second only to HIV/AIDS as the greatest killer worldwide due to a single infectious agent.
- In 2012, 8.6 million people fell ill with TB and 1.3 million died from TB.
- In 2012, an estimated 530 000 children became ill with TB and 74 000 HIV-negative children died of TB.

India:

- Tuberculosis is the biggest health issue that lies around India, but what makes is worse is the newly and recently discovered global phenomenon of **TDR-TB - Totally Drug-Resistant Tuberculosis** which is **most dangerous one**.
- An experiment was conducted at **Hinduja Hospital** in **Mumbai** in January, 2012 on *four* patients to test how accurate the "new category" of ***TDR-TB*** is. These patients were given all the **first-line drugs** and **second-line drugs** that usually are prescribed to treat TB, and as a result were resistant to all.

1.3 Genetics of *M. tuberculosis*

The complete genome of *M. tuberculosis* was sequenced in 1998.TB structure consortium has a collection of structures of over 400 proteins from *M.tuberculosis* many review articles and publications have analyzed these structures in the context of functional information.(CW *et.al* 2003)

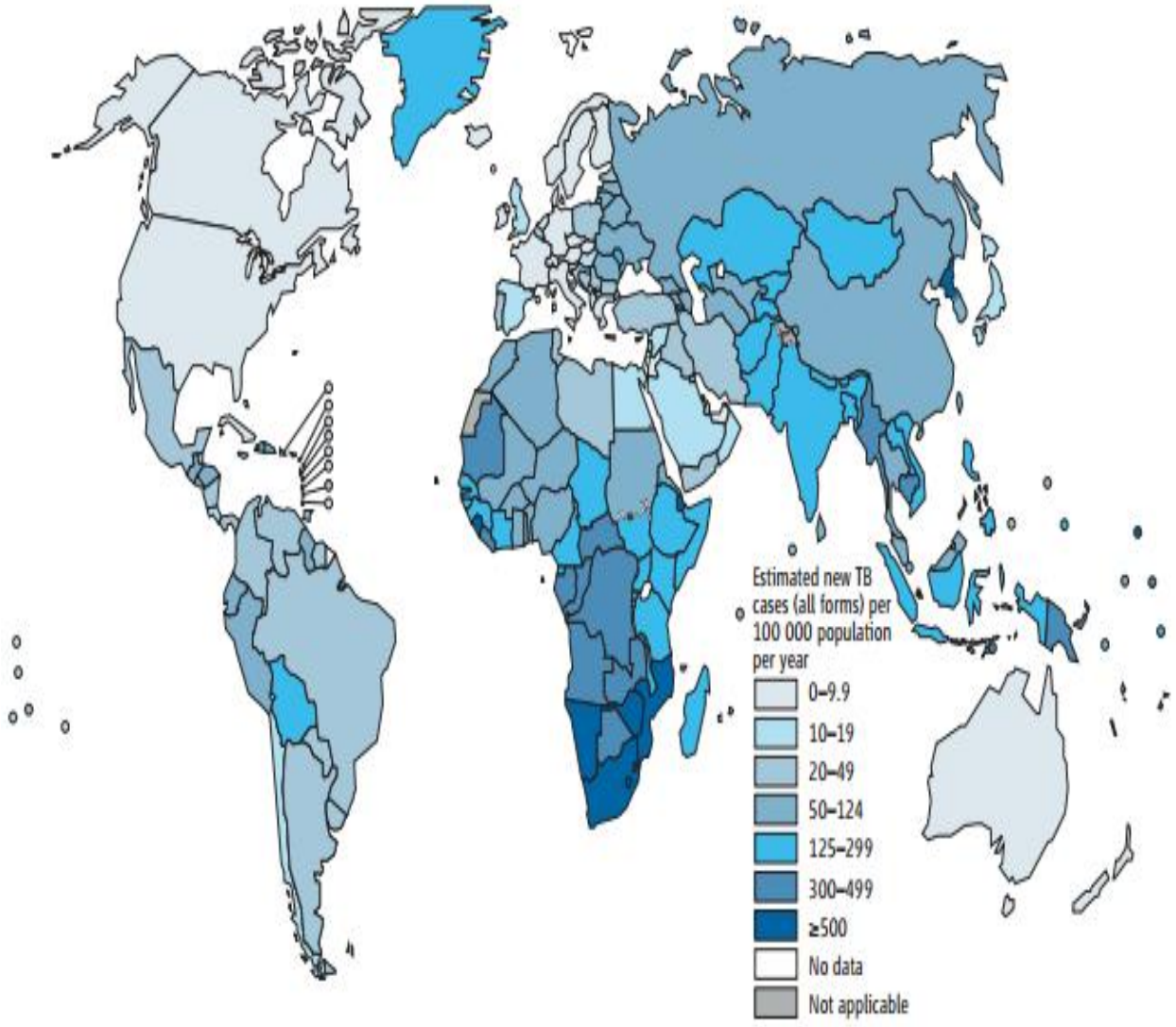
1.4 How does TB spread

TB is spread through the air from one person to another. The TB bacteria are put into the air when a person with active TB disease of the lungs or throat coughs, sneezes, speaks, or sings. People nearby may breathe in these bacteria and become infected.

1.5 TV and HIV

For many people, initial HIV symptoms are not clearly visible. But when they go for TB diagnosis, they realize that they are also suffering from HIV which probably increases susceptibility to infection with M. Tuberculosis. HIV increases the risk of progression of M. Tuberculosis infection to TB disease. This risk increases with increasing immunosuppressant. HIV infection also interferes with TB diagnosis. HIV increases not only the risk but also the rate of progression of recent or latent M. Tuberculosis infection to disease.

1.6 WorldScenario



(http://www.stoptb.org/assets/images/countries/GTBCR2013_incidence)

Figure 1.2-Estimated TB incidence rates,2012

1.7 Therapy

The standard "short" course treatment for TB is ioniazid rifampcin (also known as rifampin in the United States), pyrazinamide, and ethambutol for two months, then isoniazid and rifampicin alone for a further four months. The patient is considered cured at six months (although there is still a relapse rate of 2 to 3%). For latent tuberculosis, the standard treatment is six to nine months of isoniazid alone.

First line

All first-line anti-tuberculosis drug names have a standard three-letter and a single-letter abbreviation:

- pyrazinamide is pza or z
- ethambutol is emb or e
- rifampicin is rmp or r
- isoniazid is inh or h

US commonly use abbreviations and names that are not internationally recognized: rifampicin is called rifampin and abbreviated rif;

Most regimens have an initial high-intensity phase, followed by a continuation phase (also called a consolidation phase or eradication phase): the high-intensity phase is given first, then the continuation phase, the two phases divided by a slash.

Means isoniazid, rifampicin, ethambutol, pyrazinamide daily for two months, followed by four months of isoniazid and rifampicin given three times a week.

Second line

There are six classes of second-line drugs used for the treatment of TB. A drug may be classed as second-line instead of first-line for one of three possible reasons: it may be less effective than the first-line drugs (e.g. *p*-amino salicylic acid) or it may have toxic side-effects (e.g. cycloserine) or it may be unavailable in many developing countries (e.g. Fluoroquinolones)

- fluoroquinolones: e.g. ciprofloxacin , levofloxacin, moxifloxacin
- amino glycosides: e.g. amikacin , kanamycin
- cycloserine (the only antibiotic in its class);
- polypeptides: e.g. capreomycin, viomycin, enviomycin
- *p*-amino salicylic acid
- thioamides: e.g. Ethionamide, prothionamide

1.8 Bacterial efflux pump

Efflux pumps are proteinous transporters localized in the cytoplasm membrane of all kinds of cells. They are active transporters meaning that they require a source of chemical energy to perform their function. Bacterial efflux transporters are classified into five major super families, based on the amino acid sequence and the energy source used to export their substrates. (Marquez B. *et.al.* 2005) [5]

1. The major facilitator super family (MFS)
2. The ATP -binding cassette super family (ABC)
3. The small multidrug resistance family (SMR)
4. The resistance-nodulation-cell division super family (RND)
5. The Multi antimicrobial extrusion protein family (MATE).

1.9 The importance of efflux pumps in bacterial antibiotic resistance:

Efflux pumps are transport proteins involved in the extrusion of toxic substrates (including virtually all classes of clinically relevant antibiotics) from within cells into the external environment. These proteins are found in both Gram-positive and -negative bacteria as well as in eukaryotic organisms.(1)Pumps may be specific for one substrate or may transport a range of structurally dissimilar compounds (including antibiotics of multiple classes); such pumps can be associated with multiple drug resistance (MDR). In the prokaryotic kingdom there are five major families of efflux transporter (2) MF, MATE, RND, SMR and ABC .All these systems utilize the proton motive force as an energy source, (3) apart from the ABC family, which utilizes ATP hydrolysis to drive the export of substrates. Recent advances in DNA technology and the advent of the genomic era have led to the identification of numerous new members of the above families, and the ubiquitous nature of efflux pumps is remarkable. Transporters that efflux multiple substrates, including antibiotics, have not evolved in response to the stresses of the antibiotic era. All bacterial genomes studied contain several different efflux pumps; this indicates their ancestral origins. It has been estimated that ~5–10% of all bacterial genes are involved in transport and a large proportion of these encode efflux pump. (weber *et.al.*2003) [6]

1.10 Bacterial mechanisms of antibiotic resistance:

Several mechanisms have evolved in bacteria which confer them with antibiotic resistance. These mechanisms can chemically modify the antibiotic, render it inactive through physical removal from the cell, or modify target site so that it is not recognized by the antibiotic.

The most common mode is enzymatic inactivation of the antibiotic. An existing cellular enzyme is modified to react with the antibiotic in such a way that it no longer affects the microorganism. An alternative strategy utilized by many bacteria is the alteration of the antibiotic target site. These and other mechanisms are shown in the figure and accompanying table below. (Kenneth Todar, Online textbook of bacteriology: Bacterial Resistance to Antibiotics)[7]

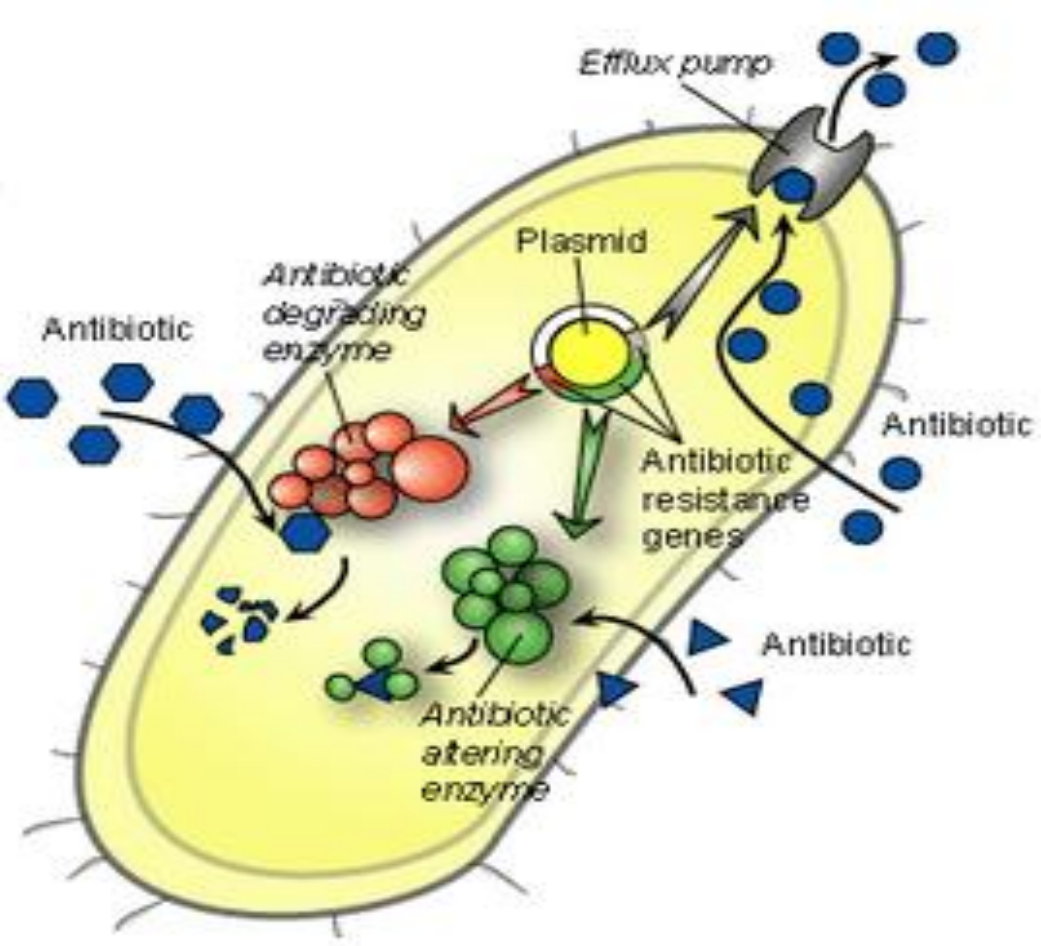


Figure-1.3 Mechanism of antibiotic resistance

Inherent resistance Bacteria may be inherently resistant to an antibiotic. For example, an organism lacks a transport system for an antibiotic or an organism lacks the target of the antibiotic molecule or as in the case of Gram-negative bacteria, the cell wall is covered with an outer membrane that establishes a permeability barrier against the antibiotic.

Acquired resistance. Several mechanisms are developed by bacteria in order to acquire resistance to antibiotics. All require either the modification of existing genetic material or the acquisition of new genetic material from another source.

1.11 Inhibitors of Efflux pump mechanism:

Induction of genes encoding efflux pumps influence general pathway to drug resistance which ultimately leads to high-level of chromosomal-mutation-related resistance in mycobacteria .Nevertheless, 20-30% of INH (isoniazid) resistant *Mycobacterium tuberculosis* isolates do not have mutations in any of the genes implicated with INH resistance.

The efflux inhibitors (EIs) particularly verapamil decreased resistance in the INH induced resistant-strains and also promoted a reversal of resistance in some of the strains tested .Although Ser531Leu and His526Asp mutations in RpoB gene result in Rifampin (RIF) drug resistance, but expression analysis has indicated that Rv2936 and Rv0783 may be responsible for the RIF drug resistance . The induction of efflux pumps affect multiple pathways and facilitate sequential acquisition of mutations that leads to the development of MDR strains to ethambutol monotherapy where there is isoniazid resistance.

Verapamil's R isomer and its metabolite norverapamil have substantially less calcium channel blocking activity, yet were similarly active as verapamil at inhibiting macrophage-induced drug tolerance. Our finding that verapamil inhibits intracellular *M. tuberculosis* growth and tolerance suggests its potential for treatment shortening. Norverapamil, R-verapamil and potentially other derivatives present attractive alternatives that may have improved tolerability.(Adams KN1 *et.al.*2014) [8]

Chapter 2

METHODS

2.1 Earlier Efflux pump protein:

There were 21 efflux pump protein in the literature. Then the extraction of the sequence (fasta) of all these efflux pumps protein. And then to do blast of these efflux protein with the protein from *M.tuberculosis* H37Rv. (Viveiros *et.al.* 2012)[9]

2.2 Addition of Efflux pump protein:

Stand alone blast for these 21 literature obtained efflux pump protein with the *M.tuberculosis* h37rv. And then check the simmilarity and characterization of then and maximum match should be extracted by code.

Now there are 10 additional efflux pump protein.Then add up these protein with 21 literature efflux protein.Now at last there are 31 efflux pump protein

2.3 Pathogenic bacteria:

Uni Drug Target is a computational tool which help to identify unique drug target in pathogenic bacteria. In this pathogenic or non pathogenic organism with particular strain are there.

So from this differntiate the more pathogenic, pathogenic, non-pathogenic and baneficial organism.(Chanumolu et.al.2012)[10]

More pathogenic-M..leprae, P.aeruginosa, V.cholerae, H.pylori, S.pneumoniae, M.tuberculosis.

Pathogenic-M.bovis, B.subtilis, M.avium.

Non-Pathogenic-M.smegmatis, E.coli.

Beneficial-L.acidophilus, S.salivarius, B.coagulans.

2.4 Genome sequence:

Extraction of genomic sequence of all these organism from NCBI.

[ftp://ftp.ncbi.nlm.nih.gov/genomes/]

Chapter 3

Results and Discussion:

3.1 Final Efflux pump protein:

Table 1.1 Total efflux pump protein

EFFLUX PUMP PROTEIN	LENGTH	RV	EFFLUX PUMP PROTEIN	LENGTH	RV
M.tb(gi|15607816)	964	Rv0676c	M.tb(gi|444895257)	865	Rv1747
M.tb(gi|15610079)	920	Rv2942	M.tb(gi|15607335)	1194	Rv0194
M.tb(gi|15609983)	530	Rv2846c	M.tb(gi|15608358)	314	Rv1218c
M.tb(gi|57117052)	107	Rv3065	M.tb(gi|15607923)	540	Rv0783c
M.tb(gi|15607483)	640	Rv0342	M.tb(gi|397674751)	439	RV2836c
M.tb(gi|15607482)	479	Rv0341	M.tb(gi|15610714)	413	Rv3578
M.tb(gi|15607484)	493	Rv0343	M.tb(gi|15610816)	386	Rv3680
M.tb(gi|15608073)	276	Rv0933	M.tb(gi|15610815)	340	Rv3679
M.tb(gi|15610073)	331	Rv2936	M.tb(gi|444895701)	379	Rv2184c
M.tb(gi|15610074)	289	Rv2937	M.tb(gi|444894758)	419	Rv1258c
M.tb(gi|15610075)	276	Rv2938	M.tb(gi|397674552)	498	RV2643
M.tb(gi|15608398)	419	Rv1258c	M.tb(gi|15610126)	233	Rv2989
M.tb(gi|15609014)	687	Rv1877	M.tb(gi|15609973)	439	Rv2836c
M.tb(gi|15608772)	471	Rv1634	M.tb(gi|15609824)	237	Rv2687c
M.tb(gi|15609470)	537	Rv2333c	M.tb(gi|15609825)	301	Rv2688c
M.tb(gi|15609823)	252	Rv2686c			

3.2 Characteristics of protein:

- These protein are involved in the many of the physiological process like to maintain the ph of homeostatis ,cell wall division and secretion of intracelluar metabolites.
- Based on bioenergetic criteria, these efflux pump can be considerd as primary and secondary transporters.
- Primarry transporter are energized by the hydrolysis of ATP and it make the ATP binding cassette i.e ABC family.
- Secondary transporters harness energy stored in an electrochemical gradient to the surface of the cell, also known as proton motive force.
- Secondary transporters are classified into four families
    1. The major facilitator super family.
    2. The multidrug and toxic compound extrusion family.
    3. The small multidrug resistance family.
    4. The resistance nodulation division family

- These efflux pump make contribution for the development of drug resistance in many ways like a natural decreased susceptibility to one or more drugs increased expression of genes that code for efflux pumps may be the first step in the development of clinically-relevant drug resistance.

- The diversity of compounds that can be extruded by efflux pumps allows them to confer a low-level multidrug-resistant phenotype.

- The decrease in the intracellular concentration of given antimicrobial by the activity of efflux pumps make it allows the bacteria to survive for a greater length of time, until chromosomal mutation arises and conferring high-level resistance to that particular drug.

- There is a way to prevent these events from occurring could be the inhibition of these efflux systems.

- If efflux pumps play a role in the stabilization of mutants, these events should appear with decreased frequency in the presence of an efflux pump inhibitor.

- For that reason, it is necessary to understand the molecular and functional mechanisms behind efflux-mediated resistance in *M. tuberculosis* and how this knowledge can be used to prevent their consequences.

3.3 Functional Information:

Table 1.2 Function of the protein

Protein Name	RV	Function
Mmpl5	0676c	It is thought to be involved in fatty acid transport.
efpA	2846c	It is thought to be involved intransport of undetermined substrate(pssibly drug) across the membrane.
		so responsible for the translocation of the substrate across the membrane.
emrE	3065	Involved in transport of multidrugs (tetraphenylphosphonium, erythromycin, ethidium bromide, acriflavine, safranin O
		,pyorin etc) across the membrane (export): multidrugs resistance by an export mechanism.
iniB	341	it is related to the rifampin-dependent strains , so it can be regarded as an indicator for .
		rifampin-dependent mycobacterium tuberculosis
iniA	342	iniA gene is essential for activity of an efflux pump that confers drug tolerance to both isoniazid and ethambutol.
iniC	343	it is related to the rifampin-dependent strains , so it can be regarded as an indicator for .
		rifampin-dependent mycobacterium tuberculosis
PstB	933	Involved in active transport of inorganic phosphate across the membrane (import); responsible for energy coupling to the transport
		system.This is one of the proteins required for binding-protein-mediated phosphate transport. Have ATP-binding
		ability and ATPase activity.
DrrA	2936	Probably involved in active transport of antibiotic and phthiocerol dimycocerosate (dim) across the membrane (export).
Rv1634	1634	Thought to be involved in transport of drug across the membrane (export). Drug resistance by an export mechanism.
		(confers resistance to toxic compounds by removing them for the cells).
stp	2333c	Involved in transport of drug across the membrane (export)
Rv2686c	2686c	Thought to be involved in active transport of unidentified antibiotic across the membrane (export): antibiotic resistance by an export
		mechanism
Rv2687c	2687c	Thought to be involved in active transport of unidentified antibiotic across the membrane (export): antibiotic resistance by an export
		mechanism
Rv2688c	2688c	Thought to be involved in active transport of unidentified antibiotic across the membrane (export): antibiotic resistance by an export
		mechanism. Responsible for energy coupling to the transport system.

Protein Name	RV	Function
Rv1747	1747	Thought to be involved in active transport of undetermined substrate (possibly lipooligosaccharide) across the membrane.
		Responsible for energy coupling to the transport system and for the translocation of the substrate across the membrane.
Rv0194	194	Thought to be involved in active transport of drugs across the membrane (export): multidrug resistance by an export mechanism
		Responsible for energy coupling to the transport system and for the translocation of the substrate across the membrane.
Rv1218c	1218c	Thought to be involved in active transport of tetronasin across the membrane (export): tetronasin resistance by an export mechanism.
arsB2	3578	Thought to be involved in transport of arsenic across the membrane (export): arsenic resistance by an export mechanism.
		Form the channel of an arsenite pump responsible for the translocation of the substrate across the membrane.
dinF	2836c	Function unknown; induction by DNA damage.
Rv3680	3680	Anion-transporting ATPase; supposedly catalyzes the extrusion of undetermined anions
		[catalytic activity: ATP + H(2)O + undetermined anion(in) = ADP + phosphate + undetermined anion(out)].
Rv3679	3679	Anion-transporting ATPase; supposedly catalyzes the extrusion of undetermined anions
		[catalytic activity: ATP + H(2)O + undetermined anion(in) = ADP + phosphate + undetermined anion(out)].
Rv2184c	2184c	Probably invoved in many.
arsC	2643	Involved in transport of arsenic compounds across the membrane (export): arsenic resistance by an export mechanism.
		Responsible for the translocation of the substrate across the membrane.
Rv1258c	1258c	Thought to be involved in transport of undetermined substrate (possibly macrolide) across the membrane (export).
		Responsible for the translocation of the undetermined substrate across the membrane.
Rv2989	2989	Involved in transcriptional mechanism.
MmpL7	2942	Invoved in the translocation of phthiocerol dimycocerosate
		in the cell wall.

Table 1.3 Function of the protein

Protein Name	RV	Function
DrrB	2937	Probably involved in active transport of antibiotic and phthiocerol dimycocerosate (dim)
		(dim)across the membrane. Probably responsible for the translocation of the substrate across
		the membrane and localization of dim into the cell wall.
DrrC	2938	Probably involved in active transport of antibiotic and phthiocerol dimycocerosate (dim)
		across the membrane. Probably responsible for the translocation of the substrate across
		the membrane and localization of dim into the cell wall.
EmrB	0783c	Translocate that confers resistance to substances of high hydrophobicity. Involved in transport :
		of multidrug across the membrane (export) multidrug resistance by an export mechanism.
		Responsible for the translocation of the substrate across the membrane.
Rv1877	1877	Involved in transcriptional mechanism.

3.4 BLAST against Four Groups:

Blast of these efflux pump protein with four different group separately.

3.4.1 Blast with More pathogenic organism.

Table 1.4 Blast between efflux protein and more pathogenic organism

Efflux Pump Protein	Length	More Pathogenic	Alignment
M.tb(gi|15607816)	964	M.Leprae(YP_002504171.1)	933
M.tb(gi|15610079)	920	M.leprae(YP_002502862.1)	881
M.tb(gi|15609983)	530	M.leprae(YP_002503716.1)	535
		P.aeruginosa(NP_250745.1)	391
M.tb(gi|57117052)	107	M.leprae(YP_002503834.1)	107
M.tb(gi|15607483)	640		
M.tb(gi|15607482)	479		
M.tb(gi|15607484)	493		
M.tb(gi|15608073)	276	P.aeruginosa(NP_254053.1)	263
		S.pneumoniae(YP_817302.1)	243
		V.cholerae(YP_002818996.1)	248
		M.leprae(YP_002504064.1)	254
M.tb(gi|15610073)	331	M.leprae(YP_002504161.1)	331
M.tb(gi|15610074)	289	M.leprae(YP_002504160.1)	289
M.tb(gi|15610075)	276	M.leprae(YP_002504159.1)	276
M.tb(gi|15608398)	419	P.aeruginosa(NP_252155.1)	377
M.tb(gi|15609014)	687	P.aeruginosa(NP_248937.1)	439
M.tb(gi|15608772)	471		
M.tb(gi|15609470)	537	M.leprae(NP_250007.1)	441
M.tb(gi|15609823)	252		
M.tb(gi|15609824)	237		
M.tb(gi|15609825)	301	M.leprae(YP_002503130.1)	304
M.tb(gi|444895257)	865		
M.tb(gi|15607335)	1194	P.aeruginosa(NP_253684.1)	513
M.tb(gi|15608358)	314	P.aeruginosa(NP_252084.1)	297
		P.pneumoniae(YP_816983.1)	292
M.tb(gi|15607923)	540	P.aeruginosa(NP_249149.1)	405
M.tb(gi|397674751)	439	V.cholerae(YP_002818354.1)	454
		P.aeruginosa(NP_253508.1)	445
		H.pylori(YP_001910084.1)	432
M.tb(gi|15610714)	413	P.aeruginosa(NP_250968.1)	417
M.tb(gi|15610816)	386	M.leprae(YP_002504131.1)	373
M.tb(gi|15610815)	340	M.leprae(YP_002504130.1)	341
M.tb(gi|444895701)	379	M.leprae(YP_002503305.1)	378
M.tb(gi|444894758)	419	P.aeruginosa(NP_252155.1)	377
M.tb(gi|397674552)	498		
M.tb(gi|15610126)	233	P.aeruginosa(NP_252198.1)	224
M.tb(gi|15609973)	439	P.aeruginosa(NP_253508.1)	445
		v.cholerae(YP_002818354.1)	440

3.4.2 BLAST with pathogenic organism:

Table 1.5 BLAST between efflux protein and pathogenic organism

EFFLUX PUMP PROTEIN	LENGTH	PATHOGENIC	ALIGNMENT
M.tb(gi|15607816)	964	M.bovis(NP_854353.1)	964
		M.avium(YP_881702.1)	948
M.tb(gi|15610079)	920	M.bovis(NP_856612.1)	920
		M.avium(YP_884395.1)	863
M.tb(gi|15609983)	530	M.bovis(NP_856516.1)	530
		B.subtilis(YP_007664427.1)	308
M.tb(gi|57117052)	107	M.bovis(NP_856737.1)	107
		M.avium(YP_883109.1)	107
		B.subtilis(YP_007662538.1)	104
M.tb(gi|15607483)	640	M.bovis(NP_854013.1)	640
M.tb(gi|15607482)	479	M.bovis(NP_854012.1)	479
M.tb(gi|15607484)	493	M.bovis(NP_854014.1)	493
M.tb(gi|15608073)	276	M.bovis(NP_854615.1)	213
		B.subtilis(YP_007663157.1)	243
		M.avium(YP_880040.1)	237
M.tb(gi|15610073)	331	M.bovis(NP_856606.1)	331
		M.avium(YP_882432.1)	309
		B.subtilis(YP_007661599.1)	305
M.tb(gi|15610074)	289	M.bovis(NP_856607.1)	289
		M.avium(YP_882433.1)	243
M.tb(gi|15610075)	276	M.bovis(NP_856608.1)	276
		M.avium(YP_881223.1)	250
M.tb(gi|15608398)	419	M.bovis(NP_854942.1)	419
		M.avium(YP_880648.1)	402
		B.subtilis(YP_007661611.1)	414
M.tb(gi|15609014)	687	M.bovis(NP_855561.1)	404
		M.avium(YP_882019.1)	671
		B.subtilis(YP_007663928.1)	538
M.tb(gi|15608772)	471	M.bovis(NP_855313.1)	471
		M.avium(YP_882324.1)	469
M.tb(gi|15609470)	537	M.bovis(NP_856010.1)	508
		M.avium(YP_880630.1)	397
M.tb(gi|15609823)	252	M.bovis(NP_856351.1)	252
M.tb(gi|15609824)	237	M.bovis(NP_856352.1)	237

Table 1.6 BLAST between efflux protein and pathogenic organism

EFFLUX PUMP PROTEIN	LENGTH	PATHOGENIC	ALIGNMENT
M.tb(gi|15609825)	301	M.bovis(NP_856353.1)	301
		B.subtilis(YP_007661684.1)	194
M.tb(gi|444895257)	865	M.bovis(NP_855428.1)	865
		M.avium(YP_882196.1)	862
M.tb(gi|15607335)	1194	M.bovis(NP_853865.1)	1194
M.tb(gi|15608358)	314	M.bovis(NP_854904.1)	311
		M.avium(YP_880606.1)	296
		B.subtilis(YP_007661765.1)	292
M.tb(gi|15607923)	540	M.bovis(NP_854463.1)	540
		M.avium(YP_880005.1)	535
M.tb(gi|397674751)	439	M.bovis(NP_856506.1)	439
M.tb(gi|15610714)	413	M.bovis(NP_857248.1)	413
		M.avium(YP_879858.1)	398
M.tb(gi|15610816)	386	M.bovis(NP_857344.1)	386
		M.avium(YP_879730.1)	366
M.tb(gi|15610815)	340	M.bovis(NP_857343.1)	340
		M.avium(YP_879731.1)	323
M.tb(gi|444895701)	379	M.bovis(NP_855855.1)	379
		M.avium(YP_881514.1)	382
M.tb(gi|444894758)	419	M.bovis(NP_854942.1)	419
		M.avium(YP_880648.1)	402
M.tb(gi|397674552)	498	M.bovis(NP_856322.1)	498
		M.avium(YP_880736.1)	361
		B.subtilis(YP_007664752.1)	337
M.tb(gi|15610126)	233	M.bovis(NP_856658.1)	233
		M.avium(YP_883006.1)	220
M.tb(gi|15609973)	439	M.bovis(NP_856506.1)	439

3.4.3 BLAST with non-pathogenic organism:

Table 1.7 BLAST between efflux protein and non- pathogenic organism

EFFLUX PUMP PROTEIN	LENGTH	NON-PATHOGENIC	ALIGNMENT
M.tb(gi|15607816)	964	M.smegmatis(YP_006568174.1)	950
M.tb(gi|15610079)	920		
M.tb(gi|15609983)	530	M.smegmatis(YP_006567320.1)	531
M.tb(gi|57117052)	107	M.smegmatis(YP_006568342.1)	104
M.tb(gi|15607483)	640	M.smegmatis(YP_006565453.1)	629
M.tb(gi|15607482)	479		
M.tb(gi|15607484)	493	M.smegmatis(YP_006565456.1)	496
M.tb(gi|15608073)	276	M.smegmatis(YP_006570356.1)	251
M.tb(gi|15610073)	331	M.smegmatis(YP_006571058.1)	214
M.tb(gi|15610074)	289		
M.tb(gi|15610075)	276		
M.tb(gi|15608398)	419	M.smegmatis(YP_006569644.1)	286
M.tb(gi|15609014)	687	M.smegmatis(YP_006568240.1)	666
M.tb(gi|15608772)	471	M.smegmatis(YP_006568479.1)	419
M.tb(gi|15609470)	537	M.smegmatis(YP_006570255.1)	419
M.tb(gi|15609823)	252	M.smegmatis(YP_006566236.1)	241
M.tb(gi|15609824)	237	M.smegmatis(YP_006566235.1)	237
M.tb(gi|15609825)	301	M.smegmatis(YP_006566234.1)	286
M.tb(gi|444895257)	865	M.smegmatis(YP_006566371.1)	904
M.tb(gi|15607335)	1194	M.smegmatis(YP_006570245.1)	575
M.tb(gi|15608358)	314	M.smegmatis(YP_006569688.1)	290
M.tb(gi|15607923)	540		
M.tb(gi|397674751)	439	M.smegmatis(YP_006567331.1)	433
M.tb(gi|15610714)	413	M.smegmatis(YP_006570640.1)	402
M.tb(gi|15610816)	386	M.smegmatis(YP_006570761.1)	380
M.tb(gi|15610815)	340	M.smegmatis(YP_006570760.1)	344
M.tb(gi|444895701)	379	M.smegmatis(YP_006568902.1)	381
M.tb(gi|444894758)	419	M.smegmatis(YP_006569644.1)	286
M.tb(gi|397674552)	498	M.smegmatis(YP_006565913.1)	341
M.tb(gi|15610126)	233	M.smegmatis(YP_006567089.1)	233
M.tb(gi|15609973)	439	M.smegmatis(YP_006567331.1)	433

3.4.4 BLAST with beneficial organism:

Table 1.8 BLAST between efflux protein and beneficial organism

EFFLUX PUMP PROTEIN	LENGTH	BENEFICIAL	ALIGNMENT
M.tb(gi|15607816)	964		
M.tb(gi|15610079)	920		
M.tb(gi|15609983)	530		
M.tb(gi|57117052)	107		
M.tb(gi|15607483)	640		
M.tb(gi|15607482)	479		
M.tb(gi|15607484)	493		
M.tb(gi|15608073)	276	B.coagulans(YP_004858474.1)	243
		S.salivarius(YP_006068268.1)	249
M.tb(gi|15610073)	331	S.salivarius(YP_006069053.1)	312
		L.acidophilus(YP_193170.1)	314
		B.coagulans(YP_004859822.1)	223
M.tb(gi|15610074)	289		
M.tb(gi|15610075)	276		
M.tb(gi|15608398)	419		
M.tb(gi|15609014)	687	B.coagulans(YP_004858537.1)	406
M.tb(gi|15608772)	471		
M.tb(gi|15609470)	537		
M.tb(gi|15609823)	252		
M.tb(gi|15609824)	237		
M.tb(gi|15609825)	301	S.salivarius(YP_006068055.1)	287
		B.coagulans(YP_004859822.1)	210
M.tb(gi|444895257)	865		
M.tb(gi|15607335)	1194		
M.tb(gi|15608358)	314	B.coagulans(YP_004859571.1)	299
M.tb(gi|15607923)	540		
M.tb(gi|397674751)	439		
M.tb(gi|15610714)	413		
M.tb(gi|15610816)	386		
M.tb(gi|15610815)	340		
M.tb(gi|444895701)	379		
M.tb(gi|444894758)	419		
M.tb(gi|397674552)	498	B.coagulans(YP_004859964.1)	343
M.tb(gi|15610126)	233		
M.tb(gi|15609973)	439		

3.5 Efflux protein only in more pathogenic and pathogenic organism:

There are 7 proteins those are exclusively present in more pathogenic and pathogenic organism.

Because these proteins are present in more pathogenic and pathogenic organism so these proteins may more involved in pathogenesis.

Table 1.9 Efflux proteins present in more pathogenic and pathogenic organism

Efflux Pump Protein	Length	RV	More Pathogenic	Alignment	Pathogenic	Alignment
M.tb(gi|15610079)	920	2942	M.leprae(YP_002502862.1)	881	M.bovis(NP_856612.1)	920
					M.avium(YP_884395.1)	863
M.tb(gi|15607482)	479	0341			M.bovis(NP_854012.1)	479
M.tb(gi|15610074)	289	2937	M.leprae(YP_002504160.1)	289	M.bovis(NP_856607.1)	289
					M.avium(YP_882433.1)	243
M.tb(gi|15610075)	276	2938	M.leprae(YP_002504159.1)	276	M.bovis(NP_856608.1)	276
					M.avium(YP_881223.1)	250
M.tb(gi|15608398)	419	1258c	P.aeruginosa(NP_252155.1)	377	M.bovis(NP_854942.1)	419
					M.avium(YP_880648.1)	402
M.tb(gi|15607335)	1194	0194	P.aeruginosa(NP_253684.1)	513	M.bovis(NP_853865.1)	1194
M.tb(gi|15607923)	540	0783cv	P.aeruginosa(NP_249149.1)	405	M.bovis(NP_854463.1)	540
					M.avium(YP_880005.1)	535

3.6 Function of these proteins:

Table 1.10 Function of the proteins

Protein Name	Rv	Function
MmpL7	2942	Invoved in the translocation of phthiocerol dimycocerosate in the cell wall.
DrrB	2937	Probably involved in active transport of antibiotic and phthiocerol dimycocerosate (dim) (dim)across the membrane. Probably responsible for the translocation of the substrate across the membrane and localization of dim into the cell wall.
iniB	341	rv0341 is related to the rifampin-dependent strains , so it can be regarded as an indicator for . rifampin-dependent mycobacterium tuberculosis
DrrC	2938	Probably involved in active transport of antibiotic and phthiocerol dimycocerosate (dim) across the membrane. Probably responsible for the translocation of the substrate across the membrane and localization of dim into the cell wall.
Probable conserved integral membrane transport protein	1258c	Thought to be involved in transport of undetermined substrate (possibly macrolide) across the membrane (export). Responsible for the translocation of the undetermined substrate across the membrane.
Probable transmembrane multidrug efflux pump	194	Thought to be involved in active transport of drugs across the membrane (export): multidrug resistance by an export mechanism. Responsible for energy coupling to the transport system and for the translocation of the substrate across the membrane.
EmrB	0783c	Translocate that confers resistance to substances of high hydrophobicity. Involved in transport : of multidrug across the membrane (export) multidrug resistance by an export mechanism. Responsible for the translocation of the substrate across the membrane.

3.7 BLAST with Mycobacterium and Bacteria:

7 efflux pump protein those were present in more pathogenic and pathogenic organism is used to do blast with mycobacterium. So After that blast can give a more result interested for knowing the more organisms of bacteria which are involved in highly pathogenicity. So comparisons are as followings:

3.7.1 Rv0194:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv0194	M.africanum	100%	0	99%
	M.bovis	100%	0	99%
	M.orygis	100%	0	99%
	M.canettii	100%	0	97%
	M.marinum	99%	0	97%
	M.ulcerans	94%	0	76%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv0194	Streptomyces purpureus	97%	0	54%

3.7.2 Rv0341:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv0341	M.canettii	100%	0	95%
	M.bovis	100%	0	100%
	M.marinum	91%	2.00E-41	50%

3.7.3 Rv783c:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv783c	M.africanum	100%	0	99%
	M.canettii	100%	0	99%
	M.marinum	94%	0	62%
	M.ulcerans	98%	0	67%
	M.avium	98%	0	67%
	M.intracellulare	98%	0	67%
	M.parascrofulaceum	98%	0	67%
	M.kansasii	93%	0	67%
	M.liflandi	94%	0	62%
	M.colombiense	95%	0	63%

3.7.4 Rv1258c:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv1258c	M.africanum	100%	0	99%
	M.canettii	100%	0	99%
	M.marinum	98%	0	77%
	M.avium	98%	0	72%
	M.parascrofulaceum	98%	0	73%
	M.kansasii	97%	0	80%
	M.liflandi	98%	0	76%
	M.smegmatis	96%	0	71%
	M.vanbaalenii	95%	0	70%
	M.chubuense	96%	0	69%
	M.mageritense	96%	0	70%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv1258c	Salinispora pacifica	90%	2.00E-99	51%

3.7.5 Rv2937:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2937	M.ulcerans	99%	3.00E-123	66%
	M.leprae	99%	3.00E-115	64%
	M.canettii	100%	0	99%

3.7.6 Rv2938:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2938	M.marinum	100%	6.00E-144	72%
	M.kansasii	98%	6.00E-168	87%
	M.smegmatis	93%	5.00E-81	53%
	M.ulcerans	100%	8.00E-144	72%
	M.leprae	100%	4.00E-154	79%
	M.canettii	100%	0	99%

3.7.7 Rv2942:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2942	M.canettii	100%	0	99%
	M.bovis	100%	0	99%
	M.kansasii	90%	0	62%
	M.marinum	96%	0	59%
	M.liflandi	96%	0	59%
	M.ulcerans	96%	0	59%
	M.laprae	95%	0	69%

So, after making comparisons as done with these above 7 proteins with mycobacterium. The outcome is like there are many of the mycobacteriums organisms come in comparison with these above 7 proteins.

So by doing comparison there were many of the organism those were in blast result. So by taking some criteria like query coverage as up to 70% and identity up to 50%. There was some of the organism those were best fit with the criteria. So these organisms can be useful for study regarding the pathogenecity of the organism.

3.8 Efflux proteins common to all organisms:

There are 5 proteins which are common to more pathogenic, pathogenic, non-pathogenic and beneficial organism. So these are the proteins which can be seen in many of the organism like the organism who are more involved in pathogenecity and some organism those are somewhere less pathogenic and some were in non pathogenic and also involved in the beneficial organism.

Table 1.8 proteins common to all groups

Efflux Pump Protein	Length	RV
M.tb(gi|15608073)	276	0933
M.tb(gi|15610073)	331	2936
M.tb(gi|15609014)	687	1877
M.tb(gi|15609825)	301	2688c
M.tb(gi|15608358)	314	1218c

3.9 BLAST of these protein to mycobacterium, all and bacteria excluding mycobacterium.

The basis for doing blast of these 5 common protein with mycobacterium and bacteria and with all organism is to know that or for study the many of the organism those can come somewhere related for these proteins.

So doing blast with mycobacterium is to find that what are different mycobacterium organism those can be considered as or to check the number of organism those are best with the 5 common protein.

And to make comparison with bacteria it will give the different bacterial organism for study regarding the pathogenecity.

Rv0933:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv0933	M.colombiense	97%	3.00E-160	83%
	M.intrcellulare	87%	3.00E-144	82%
	M.yongonense	87%	6.00E-144	82%
	M.canettii	100%	0	99%
	M.avium	89%	1.00E-141	78%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv0933	M.canettii	100%	0	99%
	Thermonaerobacter Pseudethanolicus	87%	5.00E-87	51%
	Acholeplasma brassicae	89%	7.00E-87	52%
	Desulfotomaculum carboxydivorans	87%	4.00E-85	52%
	Sphingomonas sanxanigenens	93%	2.00E-84	50%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv0933	Thermoanaerobacter pseudethanolicus	87%	2.00E-87	51%
	Acholeplasma brassicae	89%	4.00E-87	52%
	Desulfotomaculum carboxydivorans	87%	2.00E-85	52%
	Sphingomonas sanxanigenens	93%	1.00E-84	50%
	Clostridium colicanis	87%	1.00E-84	50%
	Morella thermoacetica	90%	6.00E-84	50%
	Faecalibacterium prausnitzii	91%	1.00E-83	52%

So here blast of Rv0933 with mycobacterium has given some of the organism those were taken upon some criteria like query coverage of up to 70% and identity up to 50%. So this mycobacterium organism can be considered as involved in pathogenecity.

And as blast with bacteria also given some of the bacterial organism taken by same criteria as above best for study.

And also with all organisms it has also given some of the organism best considered for pathogenecity like study.

Rv1877:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv1877	M.canettii	99%	0	100%
	M.kansasii	99%	0	72%
	M.intrcellulare	97%	0	63%
	M.yongonense	97%	0	63%
	M.avium	97%	0	63%
	M.parascrofolaceum	94%	0	61%
	M.smegmatis	96%	0	60%
	M.colombiense	95%	0	60%
	M.megritense	95%	0	60%
	M.vaccae	98%	0	57%
	M.vanbaaleni	97%	0	57%
	M.rhodesiae	98%	0	57%
	M.cosmeticum	95%	0	60%
	M.vulneris	95%	0	57%
	M.fortuitum	93%	0	56%
	M.abscessus	95%	0	53%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv1877	M.africanum	100%	0	99%
	Streptomyces	91%	0	51%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv1877	Streptomyces	98%	0	49%

Here the blast of Rv1877 has given the name of many mycobacterium organisms. But when did blast with bacterial and all organism then only few of organism were fit with criteria.

Rv1218c:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv1218c	M.bovis	100%	0	99%
	M.canettii	100%	0	99%
	M.orygis	100%	0	99%
	M.marinum	99%	0	83%
	M.colombiense	92%	0	89%
	M.lifandi	99%	0	82%
	M.kansasii	98%	0	82%
	M.avium	95%	0	87%
	M.parascrofolaceum	97%	1.00E-180	84%
	M.xenopi	95%	4.00E-178	84%
	M.cosmeticum	93%	3.00E-167	79%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv1218c	M.bovis	100%	0	99%
	Nocardia nova	96%	6.00E-154	74%
	Rhodococcus wratislaviensis	92%	2.00E-150	73%
	Gordonia kroppenstedtii	92%	1.00E-146	76%
	Ktedonobacter racemifer	92%	7.00E-139	66%
	Actinokineospora enzanensis	92%	1.00E-138	67%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv1218c	Ktedonobacter racemifer	93%	7.00E-140	67%
	Actinokineospora enzanensis	92%	5.00E-139	67%
	Kutzneria albida	92%	3.00E-137	67%
	Streptomyces avermitilis	94%	3.00E-137	67%
	Amycolatopsis balhimycina	94%	9.00E-136	67%
	Saccharothrix espanaensif	94%	2.00E-135	66%
	Micromonospora aurantiaca	95%	1.00E-134	67%
	Stackebrandita nassauensis	92%	1.00E-133	66%
	Cellulomonas flavigena	93%	2.00E-128	63%

Here also blast of Rv1218c with mycobacterium, bacteria and with all organisms has given many of the organisms those were can be considered as best for studying the pathogenecity of the organism.

Rv2688c:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2688c	M.bovis	100%	0	99%
	M.canettii	100%	0	99%
	M.kansasii	97%	1.00E-171	82%
	M.avium	99%	2.00E-179	82%
	M.parascrofolaceum	97%	1.00E-180	84%
	M.xenopi	95%	4.00E-178	84%
	M.cosmeticum	96%	4.00E-171	82%
	M.yongonense	99%	1.00E-178	81%
	M.intracellulare	94%	9.00E-177	85%
	M.vulneris	95%	3.00E-173	83%
	M.megritense	99%	3.00E-172	78%
	M.fortuitum	96%	1.00E-170	80%
	M.marinum	95%	1.00E-169	80%
	M.smegmatis	95%	1.00E-160	81%
	M.abscessus	95%	4.00E-160	77%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2688c	M.bovis	100%	0	99%
	Bacillus licheniformis	94%	1.00E-113	55%
	oceanobacillus picturae	94%	1.00E-112	55%
	Halococcus salifodinae	94%	2.00E-112	58%
	Methanocella arvoryzae	94%	2.00E-111	56%
	Paenibacillus sabinae	94%	6.00E-110	54%
	Chloroflexus aurantiacus	94%	7.00E-109	56%
	Roseiflexus castenholzii	94%	8.00E-108	58%
	Cohnella laeviribosi	94%	9.00E-103	53%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2688c	Amycolatopsis nigrescens	94%	6.00E-149	73%
	Nocardiopsis potens	94%	2.00E-135	68%
	Stackebrandtia nassauensis	95%	3.00E-134	68%
	Bacillus licheniformis	94%	6.00E-114	65%
	Paenibacillus lactis	94%	7.00E-112	56%
	Chloroflexus aurantiacus	94%	4.00E-109	56%
	Roseiflexus castenholzii	94%	5.00E-108	58%
	Hyphomonas jannaschiana	95%	2.00E-106	56%
	Oceanobacillus kimchii	94%	2.00E-105	50%
	Bacillus maethanolicus	94%	3.00E-104	51%
	Paenibacillus senegalensis	94%	4.00E-103	53%
	Cohnella laeviribosi	94%	5.00E-103	53%
	Paenibacillus dendrotiformis	94%	6.00E-103	55%

So here also the comparison of Rv2688c with mycobacterium and bacteria and with all organisms gives the some different organism and some of the same by which it can be kept in the study of pathogencity of the organism.

Rv2936:

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2936	M.bovis	100%	0	99%
	M.canettii	100%	0	99%
	M.kansasii	100%	0	86%
	M.leprae	100%	0	85%
	M.marinum	99%	0	83%
	M.lifandi	99%	0	83%
	M.ulcerans	99%	0	83%
	M.smegmatis	93%	3.00E-158	72%
	M.vanbaalenii	95%	2.00E-151	68%
	M.chubuense	95%	5.00E-150	67%
	M.gilvum	95%	5.00E-149	97%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2936	M.canettii	100%	0	99%
	Nocardia niigatensis	93%	6.00E-142	65%
	Smaragdicoccus niigatensis	93%	6.00E-142	65%
	Gordonia polyisoprenivorans	93%	4.00E-136	65%
	Rhodococcus triatomae	96%	3.00E-127	58%

Protein Name	Organism Name	Q.Coverage	E.Value	Identity
Rv2936	Synechococcus	94%	5.00E-98	52%
	Actinomadura flavelba	89%	6.00E-96	54%
	Nocardiopsis dassonvillei	95%	1.00E-92	50%
	Streptomyces globisporus	94%	2.00E-92	52%
	Amycolatopsis mediterranei	96%	2.00E-91	50%

So at last the blast of Rv2936 with mycobacterium, bacteria and with all organisms also given the result with the name of many organisms there. But with as criteria set gives some organism those are best fit are taken here. And can be considered as involved in pathogenecity a function similarity with Rv2936 pump. So from here it can be make out the organism compared with this particular efflux pump are best involved in pathogenecity.

3.10 Result and Discussion:

- There were 21 efflux pump protein present in the literature. After this the extraction of sequence (fasta) of all these 21 efflux pumps protein.
- Then blast of these efflux pump protein with the additional efflux pump protein from M. tuberculosis H37Rv. And then there were thousands of tuberculosis h37rv protein.
- Considering some criteria over there like query coverage around 70% and 50% of identity.
- So by this method there were 10 additional efflux proteins were there. So at last there were 31 efflux pump protein for study.
- After this there were grouping of organisms like more pathogenic, pathogenic, non pathogenic and beneficial organisms.
- And then the blast of these above 31 efflux pump protein with these grouping organism separately.
- There were many of the organisms by this BLAST.
- But considering or setting criteria like query coverage of 70% and identity of 50% given some organism those were taken for further analysis.
- So following table shows the different groups of organism along with 31 efflux pump proteins.

Efflux Pump Protein	Length	Highly Pathogenic	Alignment	MEDIUM PATHOGENIC	ALIGNMENT	NON-PATHOGENIC	ALIGNMENT	BENEFICIAL	ALIGNMENT
M.tb(gil15610079)	920	M.leprae(YP_002502862.1)	881	M.bovis(NP_856612.1)	920				
				M.avium(YP_884395.1)	863				
M.tb(gil15607482)	479			M.bovis(NP_854012.1)	479				
M.tb(gil15610074)	289	M.leprae(YP_002504160.1)	289	M.bovis(NP_856607.1)	289				
				M.avium(YP_882433.1)	243				
M.tb(gil15610075)	276	M.leprae(YP_002504159.1)	276	M.bovis(NP_856608.1)	276				
				M.avium(YP_881223.1)	250				
M.tb(gil15608398)	419	P.aeruginosa(NP_252155.1)	377	M.bovis(NP_854942.1)	419				
				M.avium(YP_880648.1)	402				
				B.subtilis(YP_007661611.1)	414				
M.tb(gil15607335)	1194	P.aeruginosa(NP_253684.1)	513	M.bovis(NP_853865.1)	1194				
M.tb(gil15607923)	540	P.aeruginosa(NP_249149.1)	405	M.bovis(NP_854463.1)	540				
				M.avium(YP_880005.1)	535				
M.tb(gil15608073)	276	P.aeruginosa(NP_254053.1)	263	M.bovis(NP_854615.1)	213	M.smegmatis(YP_006570356.1)	251	B.coagulans(YP_004858474.1)	243
		S.pneumoniae(YP_817302.1)	243	B.subtilis(YP_007663157.1)	243			S.salivarius(YP_006068268.1)	249
M.tb(gil15610073)	331	M.leprae(YP_002504161.1)	331	M.bovis(NP_856606.1)	331	M.smegmatis(YP_006571058.1)	214	S.salivarius(YP_006069053.1)	312
				M.avium(YP_882432.1)	309			L.acidophilus(YP_193170.1)	314
M.tb(gil15609825)	301	M.leprae(YP_002503130.1)	304	M.bovis(NP_856353.1)	301	M.smegmatis(YP_006566234.1)	286	S.salivarius(YP_006068055.1)	287
				B.subtilis(YP_007661684.1)	194				
M.tb(gil15608358)	314	P.aeruginosa(NP_252084.1)	297	M.bovis(NP_854904.1)	311	M.smegmatis(YP_006569688.1)	290	B.coagulans(YP_004859571.1)	299
		P.pneumoniae(YP_816983.1)	292	M.avium(YP_880606.1)	296				
M.tb(gil15609014)	687	P.aeruginosa(NP_248937.1)	439	M.bovis(NP_855561.1)	404	M.smegmatis(YP_006568240.1)	666	B.coagulans(YP_004858537.1)	406
				M.avium(YP_882019.1)	671				
				B.subtilis(YP_007663328.1)	538				
M.tb(gil397674552)	498			M.bovis(NP_856322.1)	498	M.smegmatis(YP_006569644.1)	286	B.coagulans(YP_004859964.1)	343
M.tb(gil15608772)	471			M.bovis(NP_855313.1)	471	M.smegmatis(YP_006568479.1)	419		
				M.avium(YP_882324.1)	469				
M.tb(gil15609470)	537	M.leprae(NP_250007.1)	441	M.bovis(NP_856010.1)	508	M.smegmatis(YP_006570255.1)	419		
				M.avium(YP_880630.1)	397				
M.tb(gil15609823)	252			M.bovis(NP_856351.1)	252	M.smegmatis(YP_006566236.1)	241		
M.tb(gil15609824)	237			M.bovis(NP_856352.1)	237	M.smegmatis(YP_006566235.1)	237		
M.tb(gil444895257)	865			M.bovis(NP_855428.1)	865	M.smegmatis(YP_006566371.1)	904		
				M.avium(YP_882196.1)	862				
M.tb(gil397674751)	439	V.cholerae(YP_002818354.1)	454	M.bovis(NP_856506.1)	439	M.smegmatis(YP_006567331.1)	433		
		P.aeruginosa(NP_253508.1)	445						
		H.pylori(YP_001910084.1)	432						
M.tb(gil15610714)	413	P.aeruginosa(NP_250968.1)	417	M.bovis(NP_857248.1)	413	M.smegmatis(YP_006570640.1)	402		
				M.avium(YP_879858.1)	398				
M.tb(gil15610816)	386	M.leprae(YP_002504131.1)	373	M.bovis(NP_857344.1)	386	M.smegmatis(YP_006570761.1)	380		
				M.avium(YP_879730.1)	366				
M.tb(gil15610815)	340	M.leprae(YP_002504130.1)	341	M.bovis(NP_857343.1)	340	M.smegmatis(YP_006570760.1)	344		
				M.avium(YP_879731.1)	323				
M.tb(gil444895701)	379	M.leprae(YP_002503305.1)	378	M.bovis(NP_855855.1)	379	M.smegmatis(YP_006568902.1)	381		
				M.avium(YP_881514.1)	382				
M.tb(gil444894758)	419	P.aeruginosa(NP_252155.1)	377	M.bovis(NP_854942.1)	419	M.smegmatis(YP_006569644.1)	286		
				M.avium(YP_880648.1)	402				
				M.avium(YP_880736.1)	361				
M.tb(gil15610126)	233	P.aeruginosa(NP_252198.1)	224	M.bovis(NP_856658.1)	233	M.smegmatis(YP_006567089.1)	233		
				M.avium(YP_883006.1)	220				
M.tb(gil15609973)	439	P.aeruginosa(NP_253508.1)	445	M.bovis(NP_856506.1)	439	M.smegmatis(YP_006567331.1)	433		
		v.cholerae(YP_002818354.1)	440						
M.tb(gil15608398)	350	P.aeruginosa(NP_253312.1)	352	M.bovis(NP_854942.1)	350	M.smegmatis(YP_006569644.1)	204		
				M.avium(YP_880648.1)	340				
M.tb(gil15607816)	964	M.Leprae(YP_002504171.1)	933	M.bovis(NP_854353.1)	964	M.smegmatis(YP_006568174.1)	950		
				M.avium(YP_881702.1)	948				
M.tb(gil15609983)	530	M.leprae(YP_002503716.1)	535	M.bovis(NP_856516.1)	530	M.smegmatis(YP_006567320.1)	531		
		P.aeruginosa(NP_250745.1)	391						
M.tb(gil57117052)	107	M.leprae(YP_002503834.1)	107	M.bovis(NP_856737.1)	107	M.smegmatis(YP_006568342.1)	104		
				M.avium(YP_883109.1)	107				
M.tb(gil15607483)	640			M.bovis(NP_854013.1)	640	M.smegmatis(YP_006565453.1)	629		
M.tb(gil15607484)	493			M.bovis(NP_854014.1)	493	M.smegmatis(YP_006565456.1)	496		

- And then the efflux pump proteins those were classified as different.
- Like some of the proteins those were only present in more pathogenic organisms and pathogenic organisms.
- Those were 7 proteins found as such. So then there BLAST with mycobacterium and bacteria excluding bacteria has done by considering some criteria like query coverage of 70% and 50% of identity.
- So after this much of process there were some organism those were found in studies those can be considered or involved in more pathogenecity.
- These organism along with table are discussed in above pages.
- Now there were some of the proteins those were present in all of the grouping organisms i.e. they are common efflux pump protein.
- Then BLAST of these proteins with mycobacterium and bacteria excluding mycobacterium and with all organism has done whether considering the query coverage of 70% and 50% of identity
- So then also got some of the organism can be considered involved in or survival of the organism or more involved in pathogenecity.

Efflux Pump Protein	Length	Highly Pathogenic	Alignment	MEDIUM PATHOGENIC	ALIGNMEN	NON-PATHOGENIC	ALIGNMENT	BENEFICIAL	ALIGNMENT
M.tb(gi|15610079)	920	M.leprae(YP_002932602.1)	881	M.bovis(NP_856912.1)	920				
				M.avium(YP_884395.1)	893				
M.tb(gi|15607402)	479			M.bovis(NP_854012.1)	479				
M.tb(gi|15610074)	289	M.leprae(YP_002934160.1)	289	M.bovis(NP_856907.1)	289				
				M.avium(YP_882433.1)	243				
M.tb(gi|15610070)	276	M.leprae(YP_002934159.1)	276	M.bovis(NP_856906.1)	276				
				M.avium(YP_881223.1)	250				
M.tb(gi|15609398)	419	P.aeruginosa(NP_252155.1)	377	M.bovis(NP_854642.1)	419				
				M.avium(YP_880648.1)	402				
				B.subtilis(YP_007801611.1)	414				
M.tb(gi|15607326)	1184	P.aeruginosa(NP_253684.1)	513	M.bovis(NP_853865.1)	1184				
M.tb(gi|15607923)	540	P.aeruginosa(NP_249149.1)	405	M.bovis(NP_854463.1)	540				
				M.avium(YP_880005.1)	535				
M.tb(gi|15609373)	276	P.aeruginosa(NP_254053.1)	263	M.bovis(NP_854615.1)	213	M.smegmatis(YP_006570356.1)	251	B.coagulans(YP_004858474.1)	243
		S.pneumoniae(YP_817302.1)	243	B.subtilis(YP_007802157.1)	243			S.salivarius(YP_000068268.1)	249
M.tb(gi|15610073)	331	M.leprae(YP_002934161.1)	331	M.bovis(NP_856906.1)	331	M.smegmatis(YP_006571058.1)	214	S.salivarius(YP_000068053.1)	312
				M.avium(YP_882432.1)	309			L.acidophilus(YP_193170.1)	314
M.tb(gi|15609825)	301	M.leprae(YP_002933130.1)	304	M.bovis(NP_856263.1)	301	M.smegmatis(YP_006569234.1)	286	S.salivarius(YP_000068055.1)	287
				B.subtilis(YP_007801604.1)	194				
M.tb(gi|15609259)	314	P.aeruginosa(NP_252084.1)	297	M.bovis(NP_854904.1)	311	M.smegmatis(YP_006569688.1)	290	B.coagulans(YP_004858671.1)	299
		P.pneumoniae(YP_818983.1)	292	M.avium(YP_880806.1)	296				
M.tb(gi|15609014)	687	P.aeruginosa(NP_249807.1)	439	M.bovis(NP_855561.1)	404	M.smegmatis(YP_006569240.1)	666	B.coagulans(YP_004858537.1)	406
				M.avium(YP_882019.1)	671				
				B.subtilis(YP_007803608.1)	538				
M.tb(gi|397674552)	496			M.bovis(NP_856322.1)	496	M.smegmatis(YP_006569644.1)	286	B.coagulans(YP_004859964.1)	343
M.tb(gi|15609772)	471			M.bovis(NP_855313.1)	471	M.smegmatis(YP_006569479.1)	419		
				M.avium(YP_882324.1)	469				
M.tb(gi|15609470)	537	M.leprae(NP_250007.1)	441	M.bovis(NP_856010.1)	508	M.smegmatis(YP_006570255.1)	419		
				M.avium(YP_880620.1)	397				
M.tb(gi|15609823)	252			M.bovis(NP_856261.1)	252	M.smegmatis(YP_006569236.1)	241		
M.tb(gi|15609824)	237			M.bovis(NP_856262.1)	237	M.smegmatis(YP_006569235.1)	237		
M.tb(gi|444896257)	865			M.bovis(NP_855408.1)	865	M.smegmatis(YP_006569371.1)	904		
				M.avium(YP_882196.1)	862				
M.tb(gi|397674751)	439	V.cholerae(YP_002818354.1)	454	M.bovis(NP_856936.1)	439	M.smegmatis(YP_006567331.1)	433		
		P.aeruginosa(NP_253500.1)	445						
		H.pylori(YP_001963084.1)	432						
M.tb(gi|15610714)	413	P.aeruginosa(NP_250960.1)	417	M.bovis(NP_857348.1)	413	M.smegmatis(YP_006570640.1)	402		
				M.avium(YP_879858.1)	398				
M.tb(gi|15610810)	386	M.leprae(YP_002934131.1)	373	M.bovis(NP_857344.1)	386	M.smegmatis(YP_006570761.1)	380		
				M.avium(YP_879730.1)	366				
M.tb(gi|15610815)	340	M.leprae(YP_002934130.1)	341	M.bovis(NP_857343.1)	340	M.smegmatis(YP_006570760.1)	344		
				M.avium(YP_879731.1)	323				
M.tb(gi|444896701)	379	M.leprae(YP_002933005.1)	378	M.bovis(NP_855855.1)	379	M.smegmatis(YP_006569902.1)	381		
				M.avium(YP_881514.1)	382				
M.tb(gi|444894758)	419	P.aeruginosa(NP_252155.1)	377	M.bovis(NP_854642.1)	419	M.smegmatis(YP_006569644.1)	286		
				M.avium(YP_880648.1)	402				
				M.avium(YP_880706.1)	361				
M.tb(gi|15610126)	233	P.aeruginosa(NP_252190.1)	224	M.bovis(NP_856658.1)	233	M.smegmatis(YP_006567089.1)	233		
				M.avium(YP_883006.1)	220				
M.tb(gi|15609972)	439	P.aeruginosa(NP_253500.1)	445	M.bovis(NP_856936.1)	439	M.smegmatis(YP_006567331.1)	433		
		v.cholerae(YP_002818354.1)	440						
M.tb(gi|15609398)	350	P.aeruginosa(NP_253312.1)	352	M.bovis(NP_854642.1)	350	M.smegmatis(YP_006569644.1)	204		
				M.avium(YP_880648.1)	340				
M.tb(gi|15607016)	964	M.Leprae(YP_002934171.1)	933	M.bovis(NP_854263.1)	964	M.smegmatis(YP_006569174.1)	950		
				M.avium(YP_881702.1)	948				
M.tb(gi|15609963)	530	M.leprae(YP_002933716.1)	535	M.bovis(NP_856316.1)	530	M.smegmatis(YP_006567320.1)	531		
		P.aeruginosa(NP_250745.1)	391						
M.tb(gi|57117052)	107	M.leprae(YP_002932934.1)	107	M.bovis(NP_856737.1)	107	M.smegmatis(YP_006569342.1)	104		
				M.avium(YP_883109.1)	107				
M.tb(gi|15607403)	640			M.bovis(NP_854013.1)	640	M.smegmatis(YP_006569453.1)	629		
M.tb(gi|15607404)	493			M.bovis(NP_854014.1)	493	M.smegmatis(YP_006569456.1)	496		

The above table is for classifying the efflux protein differently.

- The purple colour shows the efflux pump proteins those are only present in more pathogenic organism or pathogenic organism.
- The red colour shows the efflux proteins those were present in all groups i.e. common efflux pump proteins.
- The green colour proteins present in all groups except the more pathogenic.
- The other proteins those were shown in black colour are present in all groups except the beneficial organisms.

The efflux pump proteins those are in more pathogenic or pathogenic organism are – Mmpl7, iniB, drrB, drrC, Rv1258c, Rv0194 and emrB. So these are the efflux proteins those are in more pathogenic or pathogenic organism. These proteins can be considered more involved in pathogenecity of the organism and can be considered in help for the survival of the organism in vivo.

And some of the proteins like those were present in all groups i.e. in more pathogenic, pathogenic, non-pathogenic or in beneficial organism. Those efflux pump proteins are common to all so when these proteins compared with the mycobacterium, bacteria excluding mycobacterium and with all organism gives some more organisms for further study like those can be considered or may be helpful in the survival of the organism in the vivo.

So this is what here in studies. So the efflux proteins those were classified accordingly with respect to different groups given more organism for which it is considered that it can be best for survival of the organism.

Refrences:

[1] Adams KN1, Szumowski JD, Ramakrishnan L., Verapamil, and Its Metabolite Norverapamil, Inhibit Macrophage-induced, Bacterial Efflux Pump-mediated Tolerance to Multiple Anti-tubercular Drugs., J Infect Dis. (2014) ,doi: 10.1093/infdis/jiu095

[2] Chanumolu SK, Rout C, Chauhan RS (2012) UniDrug-Target: A Computational Tool to Identify Unique Drug Targets in Pathogenic Bacteria. PLoS ONE 7(3): e32833. doi:10.1371/journal.pone.0032833

[3]Dag Gundersen Storla, Solomon Yimer and Gunnar Aksel Bjune. ,A systematic review of delay in the diagnosis and treatment of tuberculosis., BMC Public Health 2008, 8:15 doi:10.1186/1471-2458-8-15

[4]Goulding cw, perry lj, anderson d, sawaya mr, cascio d, apostol mi, chan s, parseghian a, wang ss, wu y, cassano v, gill hs, eisenberg d gstructural genomics of Mycobacterium tuberculosis: a preliminary report of progress at ucla h Biophys Chem.?2003 Sep; 105(2-3):361-70.

[5] Heui-Dong Park1, Kristi M. Guinn1, Maria I. Harrell1, Reiling Liao1, Martin I. Voskuil2, Martin Tompa3, Gary K. Schoolnik2 andDavid R. Sherman1,* dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis Article first published online: 9 APR 2003 DOI: 10.1046/j.1365-2958.2003.03474.x

[6] Kenneth Todar, Online textbook of bacteriology: Bacterial Resistance to Antibiotics (page 3). http://textbookofbacteriology.net/resantimicrobial_3.html.

[7]Marquez B. Bacterial efflux systems and efflux pumps inhibitors. Biochimie. 2005 Dec;87(12):1137-47.

[8] M. A. Webber and L. J. V. Piddock*, The importance of efflux pumps in bacterial antibiotic resistance., Journal of Antimicrobial Chemotherapy (2003) 51, 9–11, DOI: 10.1093/jac/dkg050

[9]Russell dg. "Mycobacterium tuberculosis: here today, and here tomorrow "mol cell biol. 2001 aug;2(8):569-77

[10]Viveiros M1, Martins M, Rodrigues L, Machado D, Couto I, Ainsa J, Amaral L. Inhibitors of mycobacterial efflux pumps as potential boosters for anti-tubercular drugs. Expert Rev Anti Infect Ther. 2012 Sep;10(9):983-98. doi: 10.1586/eri.12.89.