World Journal of Pharmaceutical Sciences
ISSN (Print): 2321-3310; ISSN (Online): 2321-3086
Available online at: http://www.wjpsonline.org/
Original Article

Genetic Analysis of MTB-H37RV by using *VFDB*-Database

Krishna Kumar Das[1], Sudeepa Rout[2], Yangya Prasad Nath Sharma[2], Lingaraj Jena[3*]

[1,2]Department of Bioinformatics, C.P.G.S, OUAT, Bhubaneswar
[3*]JBTDRC-MGIMS, Sevagram, Wardha, Maharashtra, India

Received: 04-02-2019 / Revised Accepted: 25-03-2019 / Published: 30-03-2019

ABSTRACT

The virulence factor database (VFDB) is an integrated and comprehensive online resource for curating information about virulence factors of bacterial pathogens. Since its inception in 2004, VFDB has been dedicated to providing up-to-date knowledge of VFs from various medically significant bacterial pathogens. **Virulence factors** (VFs) refer to the properties (i.e., gene products) that enable a microorganism to establish itself on or within a host of a particular species and enhance its potential to cause disease. In my study by using this database we got 234 genes are associated with the Bacteria MTB-H37Rv. A total of 139 genes obtained from significantly enriched biological processes are termed as key genes that were carried out for network construction of MTB-H37Rv through STRING database. From STING we got the most interacting 68 genes in the bacteria which are called core genes for the disease TB and from GENE MANIA we get the common 13 genes associated with both the organism (MTB-H37Rv and Homo sapiens) during TB infection. The common genes of both the organisms are NFKB1, NFKB2, NFAT5, NFATC2, NFATC1, NFATC3, MMD, MMD2, ADIPOR2, PAQR3, PAQR5, PAQR6 and PAQR9.The structure and details of these genes are predicted by using bioinformatics tools and databases.

Key words: -MTB-H37Rv; Virulence factors; VFDB; STRING; GENE MANIA

Address for Correspondence: Lingaraj Jena, JBTDRC-MGIMS, Sevagram, Wardha, Maharashtra, India; **Email Id:** lingaraj.jena@gmail.com

How to Cite this Article: Krishna Kumar Das, Sudeepa Rout, Yangya Prasad Nath Sharma, Lingaraj Jena. Genetic Analysis of MTB-H37RV by using *VFDB*-Database. World J Pharm Sci 2019; 7(4): 9-19.

INTRODUCTION

Tuberculosis (TB) is a contagious infectious disease caused in humans mainly by Mycobacterium tuberculosis (MTB). MTB is spread essentially through the air: when an infectious person coughs, sneezes, talks or spits, saliva droplets containing tubercle bacilli are projected into the air and can be inhaled by a nearby person. Indeed, tubercle bacilli enter the human body mainly through the respiratory route after inhalation of these tiny droplets expelled into the air (**Fig. 1**).

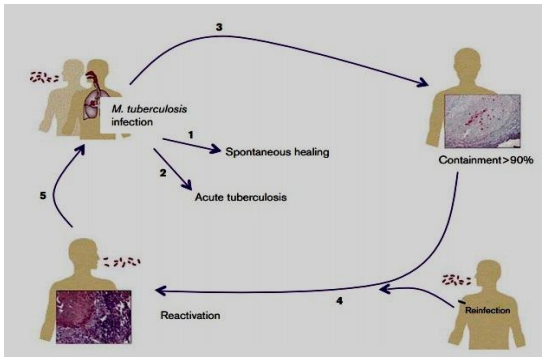

Fig.1:-Cycle based on Kaufmann et al. (2006) and Godreuil et al. (2007b). MTB enters the host by inhalation of aerosols.Different scenarios are possible: (1) immediate elimination of MTB by the pulmonary immune system; (2) infection progressesto active tuberculosis; (3) infection does not progress to active disease and MTB enters a latency phase; (4) after the latencyphase, MTB can become active following endogenous reactivation or a new exogenous infection or both; (5) at this stage,this is MTB dissemination and transmission.[22]

These particles are small enough to be able to reach the lower airways (Dannenberg, 1991). The infection success and the development of the pulmonary form of TB (lungs are the main target of this bacterium) depend on four successive steps: phagocytosis of the bacilli, their intracellular multiplication, the latent contained phase of infection and finally the active lung infection.

M. tuberculosis is an intracellular bacterium whose preferred host cell is the human macrophage. It is a slow-growing and rod-shaped bacterium with a generation time of about 24 hours. It is classified as a Gram-positive bacterium, but has a cell wall with an additional outer layer of unusual lipids, mainly mycolic acid. This feature has been used to identify mycobacteria in the laboratory for over a century by using staining techniques referred to as acid-fast or Ziehl-Neelsen staining, first developed by Franz Ziehl and Friedrich Neelsen[23]. M. tuberculosis arose from M. bovis, which causes tuberculosis in cattle, around the time when man first domesticated cattle roughly 10,000 years ago. However, comparative genome analyses have shown that M. tuberculosis is unlikely to have been derived from its bovine counterpart, and the more likely scenario is that tuberculosis was actually transferred from man to animal [24]. M. tuberculosis can be further subdivided into several strain lineages that are genetically distinct from each other. a striking difference between M. tuberculosis and many other pathogens is that the most evolutionary conserved parts of the genome are in fact epitopes recognized by human T cells, implying that M. tuberculosis actually benefits from recognition by the human immune system[25].

The mycobacterial cell wall

Many of the characteristics of M. tuberculosis, as well as many of its virulent traits can be attributed to its unique cell wall composition. The mycobacterial cell wall consists of an outer segment and inner segment. The outer segment is mainly composed of free lipids, while the inner segment consists of peptidoglycan covalently linked to arabinogalactan, which in turn is attached to mycolic acid. This inner segment or cell wall core is referred to as the mycolyl-arabinogalactan-peptidoglycan complex [26]. Throughout the cell wall, both cell wall proteins and lipoglycans can be found.

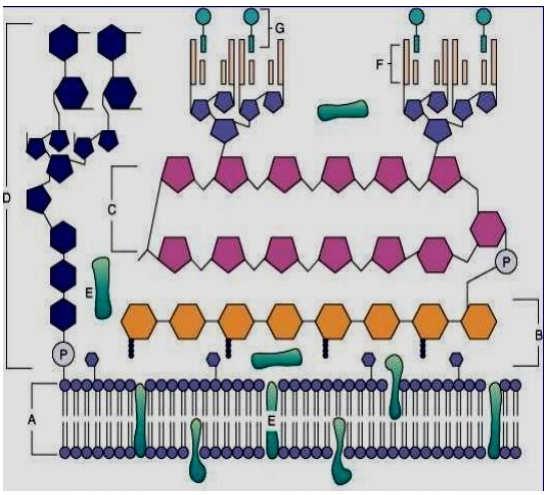

Fig.2:-Mycobacterial cell wall structure. The components include the (A) plasma membrane, (B) peptidoglycans, (C) arabinogalactan, (D) mannose-capped lipoarabinomannan, (E) plasmaassociated and cell wall-associated proteins, (F) mycolic acids, and (G) glycolipid surface molecules associated with the mycolic acids. (Redrawn from Karakousis et al: Cell Microbiol 6:105-116, 2004.)

These lipoglycans include phosphatidylinositol mannosides (PIMs) that upon additional glycosylation steps can form lipomannan (LM) and lipoarabinomannan (LAM). The PIMs, LM and LAM all attach non-covalently to the plasma

membrane by their phosphatidyl-myo-inositol anchor from where they extend outwards [27]. Though not being the only bioactive lipids present in the cell wall, PIM, LM and especially LAM has received a lot of attention due to their ability to modulate the host immune response. LAM can be further modified by addition of either mannose (ManLAM) or phosphoinositol (PILAM) caps, where the former are mainly found in pathogenic mycobacteria such as M. tuberculosis, while the latter are more associated with nonpathogenic mycobacterial species. PILAM as well as the precursor LM can induce production of cytokines (interleukin (IL)-12, IL-8 and tumor necrosis factor (TNF)) as well as to induce apoptosis in macrophages, while ManLAM from pathogenic bacteria fails to do so[28,29]. In fact, ManLAM from M. tuberculosis actively inhibit many important antimicrobial mechanisms in macrophses[30-32].Certain cell wall proteins, such as the 19 kDa lipoprotein, also act in a similar fashion, modulating the immune response of the host[33].

Mycobacterium Tuberculosis H37rv

In 1967, Runyon et al. (1) published species descriptions for Mycobacterium tuberculosis (Zopf) Lehmann and Neumann, M. bovisKarlson and Lessel, and M. microti Reed. In this publication, strain H37Rv [34] was proposed as the neotype of the species M. tuberculosis. Because the publication by Runyon et al. (1) did not meet the requirements for the designation of a neotype[34] and because there was only minimal information regarding the in vitro characteristics, pathogenicity, and maintenance of virulence of strain H37Rv, it was deemed necessary to designate properly the neotype of M. tuberculosis and to describe fully its properties. Furthermore, the use of strain H37Rv as the "standard" for M. tuberculosis in taxonomic studies, drug susceptibility testing, and pathology-immunology experiments has resulted in the world-wide distribution of this strain to the extent that it is maintained by virtually every mycobacterial culture collection in existence. Unfortunately, through lack of knowledge or experience, many cultures of H37Rv no longer are virulent for their common laboratory animal hosts. It is hoped that this publication will clarify the problem of maintenance of virulence, especially as it pertains to this widely used and newly designated neotype.

Strain H37Rv. M. tuberculosis strain H37 was isolated from a human patient in 1905 by E. R. Baldwin, Trudeau Sanatorium, Saranac Lake, N.Y.[34]. The early studies of Steenken et al.[34]revealed that this strain could dissociate into avirulent and virulent colony forms. Strangely enough, the terms first used to describe these variants were "R" and "S," but these letters did not refer to the words rough and smooth commonly used to describe colonial variants; instead the terms were used to describe those variants resistant (R) or susceptible (S) to the environment of the culture medium, even though both colonial variants were rough. If the original designations had persisted, the virulent variant of the neotype strain would be H37S. Fortunately, however, the inapt choice of the letters R and S was realized, and Steenken[35]changed the designation of the virulent variant from S to Rv (for rough virulent). In a subsequent publication [36] the full designation H37Rv was used for the first time, and the virulence of this variant was compared with that of the avirulent colony form (H37Ra).

Maintenance of H37Rv. For the past 3 years, strain H37Rv has been maintained at the Trudeau Institute as a frozen suspension (-70 C) in the phosphatebuffered (0.067 M, pH 6.8) gelatin (1%). Prior to this time it was maintained by routine passage of pellicle growth at 10 to 14 day intervals on the surface of Proskauer and Beck (PB) medium.[37]

MATERIALS AND METHODS

The virulence factor database (VFDB) is an integrated and comprehensive online resource for curating information about virulence factors of bacterial pathogens.

Prediction of genes associated with MTB H37Rv: A bacterial pathogen is usually defined as any bacterium that has the capacity to cause disease. Its ability to cause disease is called pathogenicity. Virulence provides a quantitative measure of the pathogenicity or the likelihood of causing disease. Virulence factors refer to the properties (i.e., gene products) that enable a microorganism to establish itself on or within a host of a particular species and enhance its potential to cause disease. Virulence factors include bacterial toxins, cell surface proteins that mediate bacterial attachment, cell surface carbohydrates and proteins that protect a bacterium, and hydrolytic enzymes that may contribute to the pathogenicity of the bacterium.

The virulence factor database (VFDB) is an integrated and comprehensive online resource for curating information about virulence factors of bacterial pathogens. Since its inception in 2004, VFDB has been dedicated to providing up-to-date knowledge of VFs from various medically significant bacterial pathogens. The motivation for constructing VFDB was twofold:

- First, to provide in-depth coverage major virulence factors of the best-characterized bacterial pathogens, with the structure features, functions and mechanisms used by these pathogens to allow

them to conquer new niches and to circumvent host defense mechanisms, andcause disease. Second, to provide current knowledge of the wide variety of mechanisms used by bacterial pathogens for researchers to elucidate pathogenic mechanisms in bacterial diseases that are not yet well characterized and to develop new rational approaches to the treatment and prevention of infectious diseases[2-5].

Annotation of genes according to biological processes: The DAVID Gene Functional Classification Tool [DAVID Gene Functional Classification Tool [http://david.abcc.ncifcrf.gov/gene2gene.jsp] and DAVID Functional Annotation Clustering Tool [DAVID Functional Annotation Clustering [http://david.abcc.ncifcrf.gov/summary.jsp] are two new components integrated in DAVID Bioinformatics Resources [DAVID Home Page [http://david.abcc.ncifcrf.gov]. Therefore, to illustrate the key scientific concepts, we describe only the major procedures of the DAVID Gene Functional Classification Tool. These procedures consist of three major steps: measurement of functional relationship of gene pairs, DAVID agglomeration method to partition genes into functional gene groups, and visualization of results in text and graphic modes [6].

Networking of protein: The STRING database has been designed with the goal to assemble, evaluate and disseminate protein–protein association information, in a user-friendly and comprehensive manner. As interactions between proteins represent such a crucial component for modern biology, STRING is by far not the only online resource dedicated to this topic. Apart from the primary databases that hold the experimental data in this field [7-11] and hand-curated databases serving expert annotations[12,13], a number of resources take a meta-analysis approach, similar to STRING. These include GeneMANIA[14], ConsensusPathDB[15], I2D[16], VisANT[17] and, more recently, hPRINT[18], HitPredict[19], IMID [20] and IMP[21]. Within this wide variety of online resources and databases dedicated to interactions, STRING specializes in three ways: **(i)** it provides uniquely comprehensive coverage, with >1000 organisms, 5 million proteins and >200 million interactions stored; **(ii)** it is one of very few sites to hold experimental, predicted and transferred interactions, together with interactions obtained through text mining; and **(iii)** it includes a wealth of accessory information, such as protein domains and protein structures, improving its day-to-day value for users.

UniProt-(http://www.uniprot.org/)
The Universal Protein Resource (UniProt) is a comprehensive resource for protein sequence and annotation data. The corresponding protein sequences encoded by these genes were retrieved from UniProtKB database.[38,39]

Retrieval of proteins: The structures of the corresponding proteins of reported genes were retrieved from PDB Protein Data Bank (PDB). The compounds were converted and visualised into pdb format structure using the PyMol[40] (academic version) tool, Discovery Studio v4.1 visualize tools[41] as per our requirement.

RESULTS AND DISCUSSION

The virulence factor database (VFDB) is an integrated and comprehensive online resource for curating information about virulence factors of bacterial pathogens. Since its inception in 2004, VFDB has been dedicated to providing up-to-date knowledge of VFs from various medically significant bacterial pathogens. From this database we got 234 genes are associated with the Bacteria MTB-H37Rv. These are listed in the **Table.1.**

Table.1:- The genes associated with the bacteria MTB-H37Rv from VFDB

GLNA1	MMAA4	STF0	MCE2C
LEUD	CMAA2	PAPA2	MCE2D
LYSA	MYMA	PAPA1	MCE2E
PROC	LIPR	PKS2	MCE2F
PURC	SADH	MMPL8	MCE3A
TRPD	ADHD	LPQY	MCE3B
NARX	CHP	SUGA	MCE3C
NARG	TGS4	SUGB	MCE3D
NARH	FADD13	SUGC	MCE3E
NARJ	FADD26	CHP1	MCE3F
NARI	FADD28	SAP	MCE4A
NARK2	PPSA	ICL2	MCE4B
NUOG	PPSB	KASB	MCE4C
CYP125	PPSC	ICL	MCE4D

FADE28	PPSD	LIPF	MCE4E
FADE29	PPSE	SAPM	MCE4F
CAEA	PAPA5	PAND	CTPV
ERP	MAS	PANC	CTPC
PAPA3	MMPL7	PLCA	IRTA
MMPL10	DDRA	PLCB	IRTB
MMPS4	DDRB	PLCC	MMPL3
RMLA	DRRC	PLCD	MMPL11
GTF1	TESA	MCE1A	IDER
GTF2	LPPX	MCE1B	MGTC
MPS1	PKS15	MCE1C	MBTH
FADE5	PKS1	MCE1D	MBTG
FAD23	FADD22	MCE1E	MBTF
PE	FADD29	MCE1F	MBTE
HBHA	KEFB	MCE2A	MBTD
LPRG	PCAA	MCE2B	MBTC
MBTB	FBPB	ESXD	SIGE
MBTA	FBPC	PPE69	SIGF
MBTJ	EIS	PE36	SIGH
MBTI	PKNG	ECCC2	SIGL
FADD33	SECA2	ECCB2	SIGM
FADE14	ESXA	ECCA3	WHIB3
MBTK	ESXB	ECCB3	LPQH
NDK	ECCA1	ECCC3	HSPX
PE_PGRS30	ECCB1	PE5	FBPA
PTPA	ECCCA1	PPE4	ECCA5
PAFA	ECCCB1	ESXH	ECCE5
MPA	PE35	ESPG3	CYP143
ZMP1	ECCD1	ECCD3	PPE25
RELA	ESPC	MYCP3	PPE26
DEVR/DOSR	ESPA	ECCE3	PPE27
DEVS	ESPB	ESXG	PE19
MOSR	ESPK	ESXT	AHPC
MPRA	ECCE1	ESXU	KATG
MPRB	MYCP1	ECCC4	SODC
PHOP	ESPD	CCCD4	SODA
PHOR	PPE68	MYCP4	OMPA
PRRA	ESPI	ECCB4	ESPH
PRRB	ESPJ	ECCB5	SENX3
REGX3	ESPL	PPE41	SIGA/RPOV
ESPE	ESPR	ESXM	SIGD
ESPF	ECCA2	ECCCB5	MYCP2
ESPG1	ECCE2	ECCCA5	ECCD2
PE18	ECCD5		ESPG2
MYCP5	ESXN		ESXC

The DAVID bioinformatics functional enrichment analysis reported on the essential role of biological processes, 139 genes and 40 GO terms obtained from significantly enriched biological processes were termed as key genes that were used for network construction of gene-gene/ protein-protein interaction of the bacteria MTB-H37Rv. These genes played a vital role for causing the Tuberculosis disease. The results were represented in **Table.2**.

Table.2:- Genes associated with biological processes (BP) of MTB-H37Rv

PKNG	SUGC	SUGB	ECCD5	MMPL8
CTPV	ESPC	MCE1A	ECCB5	FBPC
SIGL	ESPB	SUGA	ECCD3	MMPL10
PAFA	ESPI	FADE29	ECCB3	SAPM
MBTB	ESXT	ECCD1	ECCC3	KATG
SIGE	SENX3	CYP125	ECCB1	SODC
HBHA	SECA2	MCE1F	EIS	ECCD2
LPRG	ESPJ	MCE1D	HSPX	NARG
SIGD	MMPL11	ECCA1	FBPA	AHPC
PANC	ESPR	MCE1C	MCE2F	PKS15
NUOG	MMPS4	MCE1B	DRRC	MBTF
PAND	REGX3	TESA	ADHD	MBTE
PLCA	IRTB	FADD28	LIPR	MBTD
PLCB	IRTA	PAPA5	FADD13	MBTC
PLCC	CAEA	FADD26	FBPB	NARK2
MPA	PTPA	PPSE	TGS4	NARX
LPQY	MPRA	PPSD	LIPF	NARJ
PHOP	MPRB	PPSC	MBTI	MBTJ
ECCCA1	ESXG	PKS1	FADE5	MMAA4
PPE41	PCAA	PPSA	PKS2	KASB
FADE28	ESXH	PPSB	IDER	PRRA
MYCP1	ESXB	FADD22	MBTG	PAPA3
RELA	GLNA1	FADD29	MBTK	PRRB
LPQH	WHIB3	CMAA2	MBTA	DEVS
MMPL7	ESXA	MMPL3	ECCA3	PAPA2
LPPX	SODA	ECCE1	SIGM	PAPA1
ECCCB1	ESPA	ECCE3	SIGF	ESPH
ESPD	ESPE	ECCE5	SIGH	

A total of 139 genes obtained from significantly enriched biological processes are termed as key genes that were carried out for network construction of MTB-H37Rv through STRING database. The result of the string was represented in **Figure.3**. The MTB-H37Rv network of STRING database reported the genes namely MBTH, MBTG, MBTF, MBTE, MBTD, MBTC, MBTB, MBTA, MBTJ, MBTI, MBTK, PPSA, PPSB, PPSC, PPSD, PPSE, TESA, PAPA1,PAPA2, PAPA3, PAPA5, FADD13, FADD22, FADD26, FADD28, FADD29, FADE28, FADE29, FADE5, PKS15, PKS1, CYP125, IRTA, IRTB, ADHD, KASB, DRRC, FBPA, FBPB, FBPC, DEVS, PHOP, PRRA, PKNG, MPRB, CTPV, REGX3, SENX3, MPRA, MCE1C, MCE1B, MCE1A, MCE1D, MCE1F, MCE2F, ECCC3, ECCB3, ECCA3, ECCB5, MYCP1, NARX, NARJ, NARG, NARK2, SNM1, SNM2, MYCP and ESPE at the core region of the network. These 68 genes may be said to play a key in MTB-H37Rv as well as can be differentially expressed in MTB-H37Rv bacteria.

A total of 139 genes obtained from VFDB database significantly enriched biological processes are termed as key genes and the genes collected from literature survey were used for network construction for Gene-Gene interaction of MTB-H37Rv were analyzed through Gene Mania database. The result of the network is represented in **Figure.4**. The MTB-H37Rv network of Gene Mania database reported the 30 genes namelyPAQR4, PAQR3, ADIPOR1, PAQR9, MMD2, PAQR6, PAQR5, ADIPOR2, MMD, LIPJ, REL, RELB, NFATC4, NFATC3, NFATC1, NFATC2, NFAT5, LIPA, NFKB1, NFKB2, PTPA, CMPK2, PAQR8, PAQR7, GLI3, CHP1, RELA, ELK3, PROC and LIPF are the interacting genes from the above target genes. But from the network analysis 13 genes are at the core region of the

network, they are NFKB1, NFKB2, NFAT5, NFATC2, NFATC1, NFATC3, MMD, MMD2, ADIPOR2, PAQR3, PAQR5, PAQR6 and PAQR9.

These genes may be said to play a key in MTB-H37Rv bacteria related with Homo sapiens.

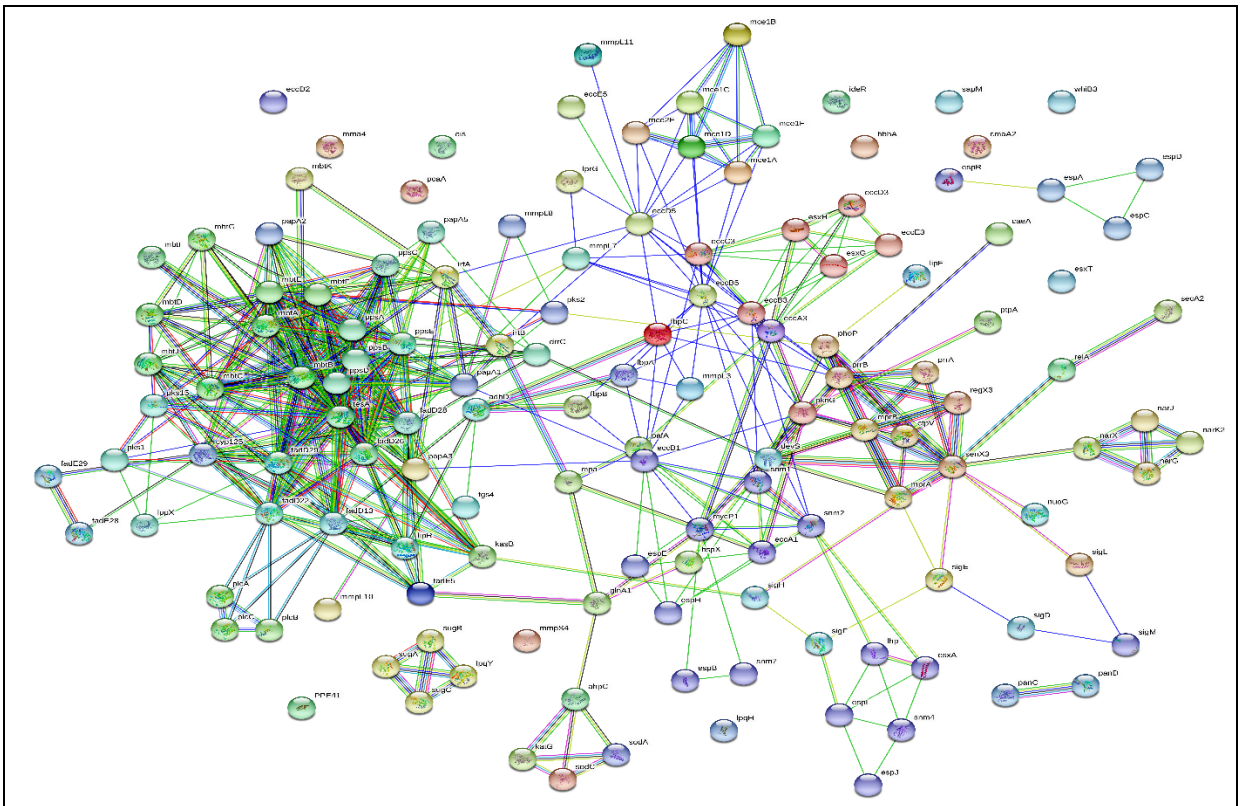

Fig.3: Interaction of genes of MTB-H37Rv in STRING database

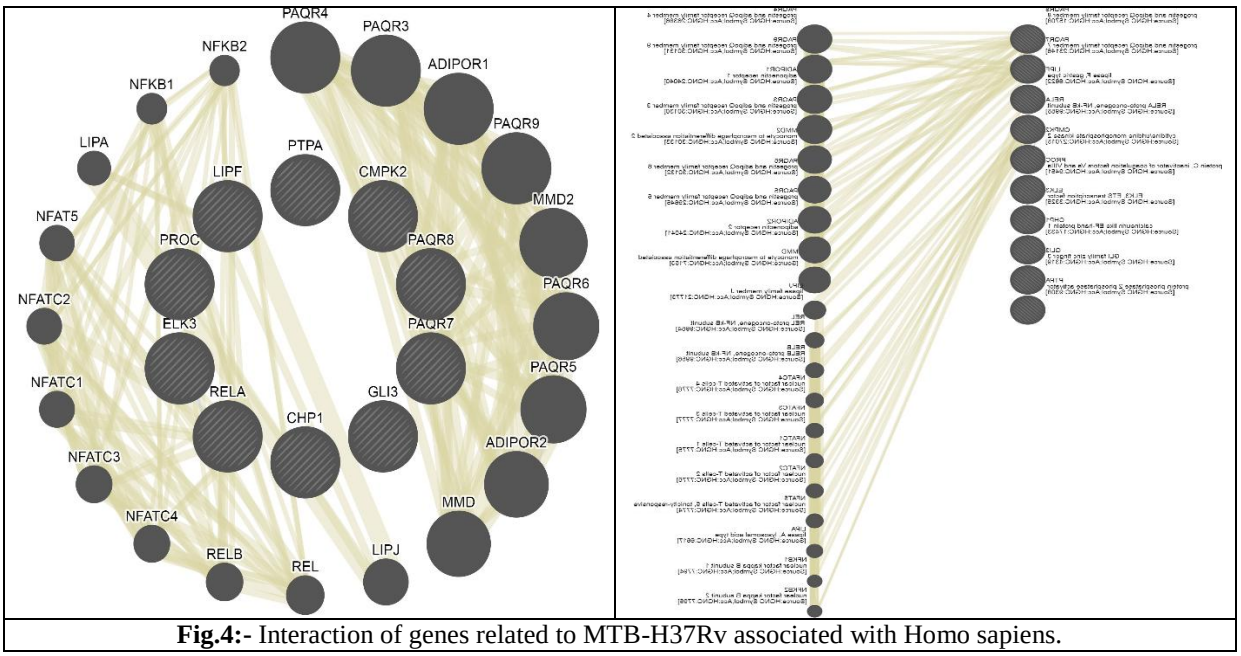

Fig.4:- Interaction of genes related to MTB-H37Rv associated with Homo sapiens.

From this we also get that by comparing the genes related to MTB-H37Rv and Homo sapiens, the above 30 genes are affected by the bacteria in the human body throughout the period of disease caused by this bacteria i.e TB. From the networking we get the above 13 genes play a vital role during the bacterial infection in human. The genes common in both the bacteria and human are the core genes, which associated with both the organism at the stage of infection by the bacteria in human during the disease Tuberculosis and the correlated disease of TB.

By applying the gene symbol on Uniprot, we can get the details about the genes. Here we put the 13 genes in UniprotKB to get the structure and the details of the genes. The details of the genes are represented in the **Table.3.** The structure according to the details are visualized by software PyMol and shown in the **Figure.5**.

Table.3:- Details of genes associated with MTB-H37Rv and Homo sapiens.

SL No	UNIPROT ID	GENE NAME	PROTEIN FULL NAME	PDB ID	REGION	CHAIN NAME	RESOLUTION
1	P19838	NFKB1	Nuclear factor NF-kappa-B p105 subunit	1SVC	2-365	P	2.60 Å
2	Q00653	NFKB2	Nuclear factor NF-kappa-B p100 subunit	1A3Q	37-327	A/B	2.10 Å
3	O94916	NFAT5	Nuclear factor of activated T-cells 5	1IMH	264-544	C/D	2.86 Å
4	Q13469	NFATC2	Nuclear factor of activated T-cells, cytoplasmic 2	1P7H	393-678	L/M/N/O	2.60 Å
5	O95644	NFATC1	Nuclear factor of activated T-cells, cytoplasmic 1	5SVE	384-400	C	2.60 Å
6	Q12968	NFATC3	Nuclear factor of activated T-cells, cytoplasmic	2XRW	141-154	B	1.33 Å
7	Q15546	MMD	Monocyte to macrophage differentiation factor	NO	NO	NO	NO
8	Q8IY49	MMD2	Monocyte to macrophage differentiation factor 2	NO	NO	NO	NO
9	Q86V24	ADIPOR2	Adiponectin receptor protein 2	3WXW	100-386	A	2.40 Å
10	Q6TCH7	PAQR3	Progestin and adipoQ receptor family member 3	NO	NO	NO	NO
11	Q9NXK6	PAQR5	Membrane progestin receptor gamma	NO	NO	NO	NO
12	Q6TCH4	PAQR6	Membrane progestin receptor delta	NO	NO	NO	NO
13	Q6ZVX9	PAQR9	Membrane progestin receptor epsilon	NO	NO	NO	NO

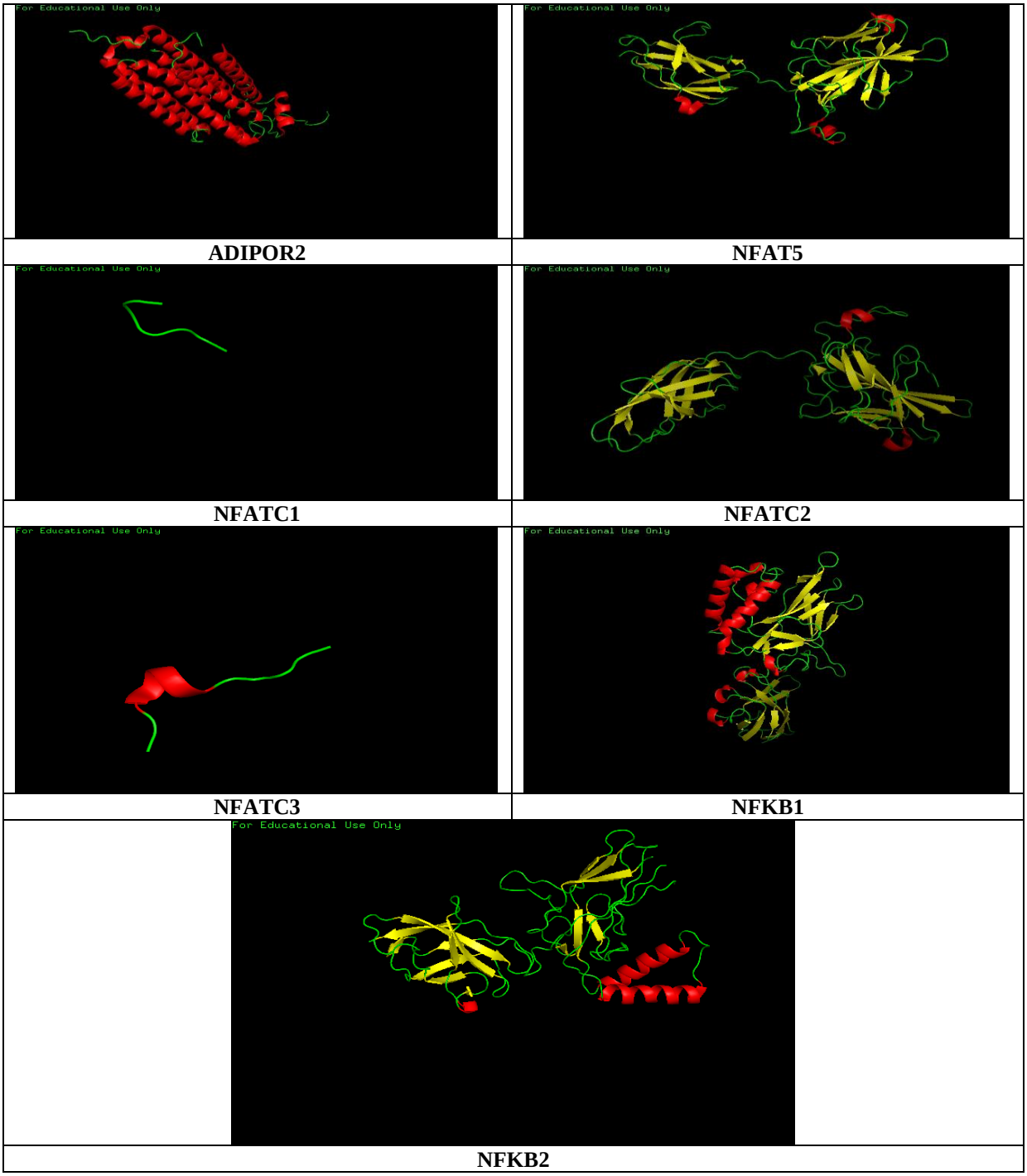

Fig.5:- The structure of the common associated Genes

CONCLUSION

In this study we predict the genes associated with the bacteria MTB-H37Rv, which is the main causative organism of the disease Tuberculosis (TB) in human. Here we get the associated genes of the bacteria from VFDB database and by purifying the data through the bioinformatics genetic database and tools; we get the most associative genes through network analysis by the STRING and GENE MANIA database. Both the database are for gene-gene interaction analysis. But from STING we got the most interacting 68 genes in the bacteria which are called core genes for the disease TB and from GENE MANIA we get the common 13 genes associated with both the organism during TB infection.

Future Aspects

You can use the genetic data given above for future use. In the above the core genes (13) of both the organisms predicted from VFDB are given with their details. By using these data you can get all the information about the genes and use for your future requirements

REFERENCES

1. Runyon, E. H., R. Bonicke, et al. Mycobacterium tuberculosis, M. bovis and M. microti species descriptions, Zentralbl. Bakteriol. Parasitenk. Infektionskr. Hyg. Abt. Orig. 1967.2O4:4054 13.
2. Chen LH, et al.VFDB : hierarchical and refined dataset for big data analysis-10 years on. Nucleic Acids Res.44 (Database issue), 2016.:D694-D697.
3. Chen LH, et al. VFDB update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res. 40(Database issue), 2012.:D641-D645.
4. Yang J, et al. VFDB release: an enhanced web-based resource for comparative pathogenomics. Nucleic Acids Res. 36(Database issue), 2008:D539-D542.
5. Chen LH, et al. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res.33 (Database issue), 2005:D325-D328.
6. Da Wei Huang, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists, Genome Biology 2007, 8:R183 (doi:10.1186/gb-2007-8-9-r183)
7. Stark,C. et al. The BioGRID interaction database: 2011 update. Nucleic Acids Res., 39, D698–D704.
8. Kerrien,S., et al. The IntAct molecular interaction database in 2012. Nucleic Acids Res., 40, D841–D846.
9. Salwinski,L. et al. The database of interacting proteins: 2004 update. Nucleic Acids Res., 32, D449–D451.
10. Licata,L., et al. (2012) MINT, the molecular interaction database: 2012 update. Nucleic Acids Res., 40, D857–D861.
11. Goll,J., et al. MPIDB: the microbial protein interaction database. Bioinformatics, (2008) 24, 1743–1744.
12. Goel,R., et al. Human protein reference database and human proteinpedia as resources for phosphoproteome analysis. Mol. Biosyst., 2012, 8, 453–463.
13. Croft,D., et al. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res., 2011, 39, D691–D697.
14. Warde-Farley,D., et al.The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res., 38, 2010, W214–W220.
15. Kamburov,A., et al.ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res., 2011, 39, D712–D717.
16. Niu,Y., Otasek,D. and Jurisica,I. Evaluation of linguistic features useful in extraction of interactions from PubMed; application to annotating known, high-throughput and predicted interactions in I2D. Bioinformatics, 2010, 26, 111–119.
17. Hu,Z., et al. VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology. Nucleic Acids Res., 2009, 37, W115–W121.
18. Elefsinioti,A., et al. Large-scale de novo prediction of physical protein-protein association. Mol. Cell. Proteomics, 2011, 10, M111 010629.
19. Patil,A., Nakai,K. and Nakamura,H. HitPredict: a database of quality assessed protein-protein interactions in nine species. Nucleic Acids Res., 2011, 39, D744–D749.
20. Balaji,S., et al. IMID: integrated molecular interaction database. Bioinformatics, 2012, 28, 747–749.
21. Wong,A.K., et al. IMP: a multi-species functional genomics portal for integration, visualization and prediction of protein functions and networks. Nucleic Acids Res., 2012, 40, W484–W490.
22. Anne-Laure Ban˜ uls, et al. Mycobacterium tuberculosis: ecology and evolution of a human bacterium, Journal of Medical Microbiology (2015), 64, 1261–1269, DOI 10.1099/jmm.0.000171.
23. Lamanna C. The Nature of the Acid-fast Stain. J Bacteriol. 1946;52(1):99-103.
24. Smith NH, et al. misconceptions: the origin and evolution of Mycobacterium tuberculosis. Nat Rev Microbiol. 2009;7(7):537-44.
25. Comas I, et al. Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. Nat Genet. 2010;42(6):498-503.
26. Brennan PJ. Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis. Tuberculosis (Edinb). 2003;83(1-3):91-7.
27. Besra GS, et al. Biosynthesis of mycobacterial lipoarabinomannan. J Biol Chem. 1997;272(29):18460-6.
28. Dao DN, et al. Mycobacterium tuberculosis lipomannan induces apoptosis and interleukin-12 production in macrophages. Infect Immun. 2004;72(4):2067-74.
29. Vignal C, et al. Lipomannans, but not lipoarabinomannans, purified from Mycobacterium chelonae and Mycobacterium kansasii induce TNF-alpha and IL-8 secretion by a CD14-toll-like receptor 2-dependent mechanism. J Immunol. 2003;171(4):2014-23.

30. Fratti RA, et al. Mycobacterium tuberculosis glycosylated phosphatidylinositol causes phagosome maturation arrest. ProcNatlAcadSci U S A. 2003;100(9):5437-42.
31. Rojas M, et al. Differential induction of apoptosis by virulent Mycobacterium tuberculosis in resistant and susceptible murine macrophages: role of nitric oxide and mycobacterial products. J Immunol. 1997;159(3):1352-61.
32. Sibley LD, et al. Mycobacterial lipoarabinomannan inhibits gamma interferon-mediated activation of macrophages. Infect Immun. 1988;56(5):1232-6.
33. Wilkinson KA, et al. Genetic determination of the effect of post-translational modification on the innate immune response to the 19 kDa lipoprotein of Mycobacterium tuberculosis. BMC Microbiol. 2009;9(93).
34. Steenken, W., et al. Biological studies of the tubercle bacillus. 111. Dissociation and pathogenicity of the R. and S. variants of the human tubercle bacillus(H3 I). 1934. J. Exp. Med. 60:515-540.
35. Steenken, W., Jr. Lysis of tubercle bacilli in vitro. Proc. SOC. Exp. Biol. Med. 1935, 33:253-255.
36. Oatway, W. H., Jr., and W. Steenken, Jr. The pathogenesis and fate of tubercle produced by dissociated variants of tubercle bacilli. J. Infec. 1936, Dis. 59:306-325.
37. GEORGE P. et al. Designation of Strain H37Rv as the Neotype of Mycobacterium tuberculosis, Vol. 22, No. 2.
38. The UniProtConsortium,UniProt: the universal protein knowledgebase, Nucleic Acids Res. 45: D158-D169 (2017).
39. Poux S; et al., On expert curation and sustainability: UniProtKB/Swiss-Prot as a case study bioRxiv (2017).
40. PyMOL Molecular Graphics System". *SourceForge*, "APBS", "PyMOL v2.0 Release Notes", Unofficial Windows Binaries for Python Extension Packages. https://en.wikipedia.org/wiki/PyMOL.
41. San Diego ,DassaultSystèmes BIOVIA, BIOVIA Workbook; BIOVIA Pipeline Pilot, Release 2017: DassaultSystèmes.http://accelrys.com/about/citations-references/.