IJPSR (2011), Vol. 2, Issue 7 **(Research Article)**

INTERNATIONAL JOURNAL
OF
PHARMACEUTICAL SCIENCES
AND
RESEARCH

Received on 11 April, 2011; received in revised form 14 May, 2011; accepted 26 June, 2011

FUNCTIONAL GENOMICS OF *MYCOBACTERIUM TUBERCULOSIS*

Anitha Jabamalai Raj [1], Sangilimuthu Algar Yadav [2], Velliyur Kanniappan Gopalakrishnan [3], Akila Selvaraj [4] and Sudarsanam Dorairaj*[5]

Research and Development Centre, Bharathiar University [1], Coimbatore- 641 046, Tamilnadu, India
Department of Biotechnology, Karpagam University [2], Coimbatore- 641021, Tamilnadu, India
Department of Bioinformatics, Karpagam University [1, 3, 4], Coimbatore- 641021, Tamilnadu, India
Department of Advanced Zoology, Biotechnology and Bioinformatics Unit, Loyola College [5], Chennai- 600034, Tamil Nadu, India

Keywords:
Mycobacterium tuberculosis,
Hypothetical Proteins,
Sequence Similarity,
Bioinformatics tools,
Drug target

Correspondence to Author:

Dr. D. Sudarsanam

Department Advanced Zoology and
Biotechnology, Medical lab Technology &
Biomedical Instrumentation ScienceLoyola
College, Chennai-600034, India

ABSTRACT

Mycobacterium tuberculosis is a deadly infectious disease, there is rising death of humans every year because of this disease, availability of genome sequences of *Mycobacterium tuberculosis* has provided tremendous amount of information that can be useful in drug target and new vaccine development. Sequence similarity provides accurate annotation for genes in newly sequenced genomes. In this present work about 107 hypothetical proteins of *Mycobacetrium tuberculosis* were taken and its functions were predicted using bioinformatics tools BLAST, BLOCKS, COGs, InterProScan and PFP. From our analysis of 107 hypothetical proteins only two shows 100% functions these proteins may serve as target for few antibiotics.

INTRODUCTION: The proteomics of *Mycobacterium tuberculosis* was intensively studied to expand our knowledge base and identify proteins or enzymes that would serve as targets for few antibiotics, many number of proteins present without any function. Tuberculosis is a common and often deadly infectious disease caused by mycobacteria, mainly *Mycobacterium tuberculosis* [1].

Mycobacterium tuberculosis is reputed to have the highest annual global mortality among all of the pathogens. The rise in tuberculosis incidence over the last two decades is partly due to tuberculosis deaths in HIV-infected patients and partly due to the emergence of multidrug resistant strains of the bacteria [2]. This rapid increase in the disease has led to potential finding arrangements aimed at a large effort towards stopping this disease before it becomes a global epidemic. Due to its slow growth and high virulence it is extremely difficult to work with the tuberculosis bacterium. However, rapidly evolving Mycobacterium genomics with complete genome sequences known along with powerful bioinformatics approaches. The proteomics of *Mycobacterium tuberculosis* will be intensively studied to expand our knowledge base and identify proteins or enzymes that would serve as targets for few antibiotics [3]; *Mycobacterium tuberculosis* has totally 4189 number of proteins.

In these proteins 2038 hypothetical proteins were present out of the 107 hypothetical proteins were taken for this work and its functions was predicted using bioinformatics tools such as BLAST, BLOCKS, COGs [4], InterProScan [5] and PROTEIN FUNCTION PREDICTION (PFP) [6].

The main aim of this present work was to predict the function of selected hypothetical proteins.

MATERIALS AND METHODS: The complete genome for *Mycobacterium tuberculosis* was sequenced at the Institute of Genome Research (TIGR) and complete genome sequence of *Mycobacterium tuberculosis* was downloaded from ExPASY [7] databases and NCBI database [8]. In the genome sequence of *Mycobacterium tuberculosis* 2038 hypothetical proteins were present. Out of that only 107 hypothetical proteins of genome sequence was downloaded from the site [9] the function for these hypothetical proteins were predicted using bioinformatics tools such as BLAST [10] (11), BLOCKS [11], COGs [12], InterProScan [13] and PROTEIN FUNCTION PREDICTION (PFP) [14] . The confidence level can be measured on the basis of above tools.

1. 100% confidence level was given to those proteins that indicate same functions in all the given five tools.
2. 80% confidence level was given to those proteins that indicate same functions in any of the given four tools other are different.
3. 60% confidence level was given to those proteins that indicate same functions in any of the given three tools other two given different functions.
4. 40% confidence level was given to those proteins that indicate same functions in any of the give two tools other three given different functions.
5. 20% confidence level was given to those proteins that indicates some function in any of the given tools
6. Moreover 0% confidence level was given to those proteins that indicate unknown functions in all the given tools.

RESULTS AND DISCUSSION: *Mycobacterium tuberculosis* has totally 4189 proteins. In this 2038 proteins are hypothetical proteins from which 107 hypothetical proteins of *Mycobacterium tuberculosis* were retrieved for this work from NCBI Database hypothetical proteins were submitted to bioinformatics tools BLAST, InterProScan, BLOCKS, PFP and COGs. The results of different tools were analyzed and the functions observed in each output were noted down and tabulated against the sequence ID.

Table 1 (Functions of Hypothetical Protein) shows functional genomics of *Mycobacterium tuberculosis* using Bioinformatics tools. Protein function can be predicted from sequence, including posttranslational modifications, subcellular localization, and physical/chemical properties [15]. Function prediction of genomic sequences for hypothetical proteins of *Mycobacterium tuberculosis* will help in the identification of new drug targets.

TABLE 1: FUNCTIONAL GENOMICS OF *MYCOBACTERIUM TUBERCULOSIS*

GENE ID	PROTEIN ID	BLAST	INTERPRO	BLOCKS	PROTEIN FUNCTION PREDICTION	COG	CONFIDENCE LEVEL
15839425	P71705	FUNCTION UNPREDICTED	PROTEIN OF UNKOWN FUNCTION	PROTEIN OF UNKNOWN FUNCTION	HYDROLASE'S ACTIVITY	NO RELATED COG	20%
15839426	P64677	FUNCTION UNPREDICTED	NO HITS	ClpX C4-TYPE ZINC FINGER	DNA -BINDING	NO RELATED COG	20%
15839428	Q7DAJ3	FUNCTION UNPREDICTED	HEMOPEXIN	Na+/Pi-COTRANSPOTER	FAD -BINDING	NO RELATED COG	20%
15839434	P64679	FUNCTION UNPREDICTED	NO HITS	L6-MEMBRANE	CALICUM-TRANSPORTING ATPASE ACTIVITY	NO RELATED COG	20%
15839436	O53604	FUNCTION UNPREDICTED	NO HITS	RNA POLYMERASE	ACETATE-COA LIGASE ACTIVITY	NO RELATED COG	20%
15839437	O53605	APPR-1-P PROCESSING	APPR-1-P-PROCESSING	PROTEIN OF UNKNOWN FUNCTION	PROTEIN FUNCTION UNKNOWN	UN CHARACTERIZED ACR	40%
15839438	Q8VKS4	FUNCTION UNPREDICTED	UNINTEGRATED	DRF AUTOREGLUATOR	GTPASE ACTIVATOR ACTIVITY	NO RELATED COG	20%
15839439	Q8VKS3	FUNCTION UNPREDICTED	NO HITS	PROTEIN OF UNKONOWN FUNCTION	TUMOR NECROSIS FACTOR RECEPTOR BINDING	NO HITS	20%
15839442	Q8VKS2	FUNCTION UNPREDICTED	UNINTEGRATED	VACULOR ATPASE	PROTEIN FUNCTION UNKNOWN	NO RELATED COG	20%
15839443	O53609	FUNCTION	PROTEIN OF UNKOWN	PROTEIN OF UNKNOWN	PROTEIN FUNCTION	UNCHARACTERIZED ACR	0%

		UNPREDICTED	FUNCTION	FUNCTION	UNKNOWN		
15839444	O53610	FUNCTION UNPREDICTED	UNINTEGRATED	PiLT PROTEIN N-TERMINAL	OXIDO REDUCTASE ACTIVITY	NO RELATED COG	20%
15839455	O53621	FUNCTION UNPREDICTED	UNINTEGRATED	RETICULON	TRANSCRIPTION ACTIVATOR ACTIVITY	NO RELATED COG	20%
15839458	Q8VKS0	FUNCTION UNPREDICTED	NO HITS	PROTEIN OF UNKONOWN FUNCTION	VIRION BINDING	NO RELATED COG	20%
15839459	Q8VKR9	FUNCTION UNPREDICTED	UNINTEGRATED	PROTEIN OF UNKONOWN FUNCTION	ATP BINDING	NO RELATED COG	20%
15839460	O53624	RIBOSOME ASSOCIATED PROTEIN	NO HITS	PHENOL HYDROLASE'S	CHROMATIN BINDING	NO RELATED COG	20%
15839461	Q7DAH5	FUNCTION UNPREDICTED	FMN-BINDING SPLIT BARREL	PROTEIN OF UNKNOWN FUNCTION	TRIOSE PHOSPHATE ISOMERASE ACTIVITY	NO RELATED COG	20%
15839468	P0A5C3	DIMETHYLADENOSINE TRANSFERASE	UNINTEGRATED	STREPTOMYCES CYCLASE/DEHYDRASE	ZINC- ION BINDING	NO RELATED COG	20%
15839470	P64683	FUNCTION UNPREDICTED	UNINTEGRATED	PROTEIN OF UNKNOWN FUNCTION	OLIGOPEPTIDE TRANSPORTER ACTIVITY	UNCHARACTERIZED MEMBRANE PROTEIN	20%
15839471	Q8VKR8	FUNCTION UNPREDICTED	UNINTEGRATED	RIBOSOMAL PROTEIN	PROTEIN BINDING	NO RELATED COG	20%
15839474	Q10890	FUNCTION UNPREDICTED	UNINTEGRATED	GLUCONATE TRANSPORTER	SULFATE PORTER ACTIVITY	NO RELATED COG	20%
15839475	Q8VKR7	FUNCTION UNPREDICTED	HNH NUCLEASE	PROTEIN OF UNKNOWN FUNCTION	NUCLEIC ACID BINDING	NO RELATED COG	20%
15839478	P64685	FUNCTION UNPREDICTED	NO HITS	CYSTIC FIBROSIS TRANSMEMBRANE	ATP BINDING	NO RELATED COG	20%
15839482	P64689	PREDICTED MEMBRANE PROTEIN	UNINTEGRATED	NA+ DEPENDENT NUCLEOSIDE TRANSPORT	CALICUM-SODIUM ANTIPORTER ACTIVITY	NO RELATED COG	20%
15839486	P64691	FUNCTION UNPREDICTED	COBALAMIN (VITAMIN B12) BIOSYNTHESIS COBW, C-TERMINAL	COBALAMINE SYNTHESIS	DNA DIRECTED RNA POLYMERASE	PUTATIVE GTPASE	40%
15839489	Q7DAG8	FUNCTION UNPREDICTED	NO HITS	PROTEIN OF UNKNOWN FUNCTION	PROTEIN TYROSINE KINASE	NO RELATED COG	20%
15839781	P95206	FUNCTION UNPREDICTED	UNINTEGRATED	SAPOSIN TYPE A	PHOSPHOTRANSFERASE ACTIVITY	NO RELATED COG	20%
15839784	P95210	FUNCTION UNPREDICTED	UNINTEGRATED	PROTEIN OF UNKNOWN FUNCTION	G-PROTEIN COUPLED RECEPTOR ACTIVITY	NO RELATED COG	20%
15839788	P0A5K0	FUNCTION UNPREDICTED	MEMBRANE PROTEIN	MYCOBACTERIUM MEMBRANE	ATPASE ACTIVITY	NO RELATED COG	40%
15839789	Q8VKL6	FUNCTION UNPREDICTED	NO HITS	PROTEIN OF UNKNOWN FUNCTION	PROTEIN UNKNOWN FUNCTION	NO HITS	0%
15839798	Q7D9V2	FUNCTION UNPREDICTED	UNINTEGRATED	BACTERIAL CONJUGATION Trbl-LIKE PROTEIN	MOTOR ACTIVITY	NO RELATED COG	20%
15839804	Q8VKL5	FUNCTION UNPREDICTED	UNINTEGRATED	MoeZ/MoeB	LIGASE ACTIVITY	NO RELATED COG	20%
15839806	P96265	LIPOPROTEIN PEPTIDASE	METALLOPEPTIDASE	L-SERYL-tRNA SELENIUM TRANSFERASE	NEUROTENSIN RECEPTOR ACTIVITY	PREDICTED METALLOPROTEASE	20%
15839807	P96266	FUNCTION UNPREDICTED	UNINTEGRATED	ALPHA-L-ARABINOFURONSIDASE B	ZINC- ION BINDING	NO RELATED COG	40%
15839808	P96267	FUNCTION UNPREDICTED	ESTERASE/LIPASE/THIOESTREASE	2-NITROPROPANE DIOXYGENASE, NPD	PROTEIN BINDING	PREDICTED HYDROLASE'S	20%
15839811	Q8VKL4	FUNCTION UNPREDICTED	NO HITS	RIBOSOMAL PROTEIN S5	FUNCTION UNKNOWN	NO RELATED COG	20%
15839813	Q8VKL3	FUNCTION UNPREDICTED	UNINTEGRATED	HERPESVIRUS	ATP BINDING	NO RELATED COG	20%
15839815	Q7D9U2	FUNCTION UNPREDICTED	UNINTEGRATED	G2 NIDOGEN AND FIBULIN G2F	EXTRACELLULAR MATRIX CONFERRING ELASTICITY	NO RELATED COG	20%
15839817	Q7D9U1	FUNCTION UNPREDICTED	HELIX -TURN- HELIX FIS TYPE	GOLGI PHOSPHOPROTEIN	ARGININE -Trna LIGASE	NO RELATED COG	20%
15839820	P64693	CARBOXYLATE AMINE LIGASE	GULUTAMATE CYSTEINE LIGASE	GLUTAMATE-CYSTEINE LIGASE ACTIVITY	GLUTAMATE-CYSEINE LIGASE ACTIVITY	UNCHARACTERIZED BCR	20%
15839821	Q7D9T9	PEPTIDASE	PEPTIDASE	ELK POTASSIUM CHANNEL	ATP BINDING	LON PROTEASE	60%
15839822	Q8VKL2	FUNCTION UNPREDICTED	UNINTEGRATED	PROTEIN OF UNKNOWN FUNCTION	PROTEIN FUNCTION UNKNOWN	NO HITS	40%
15839831	Q7D9T5	FUNCTION UNPREDICTED	UNINTEGRATED	RIBOSOMAL PROTEIN L25	ISOMERASE ACTIVITY	NO RELATED COG	20%
15839832	O53729	FUNCTION	UNINTEGRATED	3-OXO-5-ALPHA-STEROID 4-	SPERMIDINE SYNTHASE	NO RELATED COG	20%

		UNPREDICTED		DEHYDROGENASE	ACTIVITY		
15839834	Q7D9T2	TRANSMEMBRANE PROTEIN	3-OXO5-ALPHA DEHYDROGENASE	LYSINE EXPORTER PROTEIN	S-ISOPRENYLCYSTEINE O-METYL TRANSFERASE	NO RELATED COG	20%
15839836	Q7D9T0	PERIDCTED NAD/FAD BINDING PROTEIN	PROTEIN OF UNKOWN FUNCTION	PROTEIN OF UNKONOWN FUNCTION	ENDONUCLEASE ACTIVITY	UNCAHARACTERIZED ACR	20%
15839837	O53734	PREDICTED NAD/FAD BINDING PROTEIN	MONOMETHYLAMINE METHYL TRNASFERASE	ADRENODOXIN REDUCTASE FAMILY	OXIDO REDUCTASE ACTIVITY	PREDICTED NAD/FAD BINDING PROTEIN	40%
15839839	P0A5K2	FUNCTION UNPREDICTED	MYCOBACTERIUM MEMBRANE	MYCOBACTERIUM MEMBRANE	OXIDO REDUCTASE ACTIVITY	NO RELATED COG	40%
15839840	Q7D9S8	MONOMETHYLAMINE TRANSFERASE	BACTERIAL REGULATORY PROTEIN Ter	BACTERIAL REGULATOR PROTEIN TetR,	TRANSCRIPTION FACTOR ACTIVITY	TRANSCRIPTIONL REGULATORY	60%
15839842	Q8VKL1	FUNCTION UNPREDICTED	NO HITS	PURINE AND PHOSPHORYLASE	PROTEIN FUNCTION UNKNOWN	NO RELATED COG	20%
15839843	Q7D9S6	FUNCTION UNPREDICTED	UNINTEGRATED	FLAGIN HOOK IN	HYDROLASE'S ACTIVITY	NO RELATED COG	20%
15839845	Q8VKL0	FUNCTION UNPREDICTED	UNINTEGRATED	AROMATIC AMINO ACID PERMEASE	PROTEIN TYROSINE PHOSPHOTASE ACTIVITY	NO RELATED COG	20%
15839848	O53744	FUNCTION UNPREDICTED	PROTEIN OF UNKOWN FUNCTION	PROTEIN OF UNKNOWN FUNCTION	ENDOPEPTIDASE INHIBITOR ACTIVITY	UNCAHARACTERIZED ACR	20%
15839849	Q7D9S2	FUNCTION UNPREDICTED	UNINTEGRATED	SUGAR TRANSPORTER SIGNATURE	SODIUM SYMPORTER ACTIVITY	NO RELATED COG	20%
15839850	Q7D9S1	FUNCTION UNPREDICTED	UNINTEGRATED	NADH-UBIQUINONE OXIDOREDUCTASE	MALATE SYNTHASE ACTIVITY	NO RELATED COG	20%
15839852	O53748	FUNCTION UNPREDICTED	UNINTEGRATED	MIP FAMILY	PRENYL TRANSFERASE ACTIVITY	NO RELATED COG	20%
15839853	O53749	FUNCTION UNPREDICTED	CARBOXYMUCONOLACTONE DECARBOXYLASE	CARBOXYMUCONOLACTON E DECARBOXYLASE	AMINO PEPTIDASE ACTIVITY	UNCAHARACTERIZED ACR	40%
15839855	O53751	ACYL THIOESTERASE	ACYL ACP THIOESTERASE	FUNCTION UNKNOWN	ACYL CARRIER ACTIVITY	PREDICTED THIOESTERASE	80%
15839859	Q8VKK9	FUNCTION UNPREDICTED	1,4-DIHYDROXY 2NAPHTHOATEOCTOPRENYL TRANSFERASE	6-PHOSPHOGULCONATE DEHYDROGENASE	TRANSFERASE ACTIVITY	1, 4DIHYDROXY2-NAPHTHOATE OCTOPRENYLTRANSFERASE	40%
15839860	O53756	FUNCTION UNPREDICTED	TRANSMEMBRANE REGION	TatD DEOXYRIBONUCLEASE	NEUROTENSIN RECEPTOR ACTIVITY	NO RELATED COG	20%
15839862	Q7D9R2	CONSERVED TRANSMEMBRANE	PROTEIN OF UNKOWN FUNCTION	CLASS I CYTOCHROME C	RNA POLYMERASE ACTIVITY	UNCHARACTERIZED MEMBRANE PROTEIN	20%
15839864	Q8VKK8	FUNCTION UNPREDICTED	NO HITS	ANTISTASIN	BINDING	NO HITS	20%
15839865	P0A5P6	FUNCTION UNPREDICTED	GTP BINDING SIGNAL RECOGNITION SRP54	ACID SHOCK	ACETYL-COA CARBOXYLASE ACTIVITY	NO RELATED COG	20%
15839866	Q8VKK7	FUNCTION UNPREDICTED	NO HITS	FUNCTION UNKNOWN	NUCLEIC ACID BINDING	NO HITS	20%
15839867	P64697	FUNCTION UNPREDICTED	TRANSMEMBRANE REGION	BACTERIAL CONJUGATION TrbI-LIKE PRO	METHYL TRANSFERASE ACTIVITY	NO RELATED COG	20%
15839869	P64699	FUNCTION UNPREDICTED	TRANSMEMBRANE REGION	CATION DEPENDENT MANNOSE 6-PHOSPHATE	PROTEIN FUNCTION UNKNOWN	NO RELATED COG	20%
15839871	P64701	FUNCTION UNPREDICTED	NO HITS	DNA-DIRECTED DNA POLYMERASE	RNA POLYMERASE ACTIVITY	NO RELATED COG	40%
15839873	P64703	ErfK/YbiS/YcfS/YnhG	ErfK/YbiS/YcfS/YnhG	ErfK/YbiS/YcfS/YnhG	PROTEIN FUNCTION UNKNOWN	NO RELATED COG	60%
15839877	P64709	FUNCTION UNPREDICTED	NO HITS	RIBOSOMAL PROTEIN L5 SIGNATURE	PROTEIN FUNCTION UNKNOWN	NO RELATED COG	20%
15839878	Q8VKK6	FUNCTION UNPREDICTED	NO HITS	FUNCTION UNKNOWN	PROTEIN FUNCTION UNKNOWN	NO HITS	20%
15839884	Q11157	FUNCTION UNPREDICTED	NO HITS	ANTRAX RECEPTOR C-TERMINAL	FERRIC IRON BINDING	NO RELATED COG	20%
15839885	Q11158	FUNCTION UNPREDICTED	NO HITS	GALACTOSE-3-O-SULFOTRANSFERASE	PROEIN FUNCTION UNKNOWN	NO RELATED COG	20%
15839887	P64713	FUNCTION UNPREDICTED	SIGNAL PEPTIDE	KINESIN ASSOCIATED	NUTRIENT RESERVOIR ACTIVITY	NO RELATED COG	20%
15839888	P65786	EXOPOLYPHOSPHATASE	PPX/GPPA PHOSPHATASE	PPX/GPPA PHOSPHOTASE	HYDROLASE'S ACTIVITTY	EXOPOLYPHOSPHATASE	80%
15839889	P64715	FUNCTION UNPREDICTED	UNINTEGRATED	FUNCTION UNKNOWN	HYDROGEN TRANSPORTING ATPASE	NO RELATED COG	20%
15839890	P64717	SUGAR PHOSPHATASE ISOMERASE	XYLOSE ISOMERASE LIKE TIM BARREL	RETROVIRAL M-DOMAIN	FLAP ENDONUCLEASE	SUGAR PHOSPHATE ISOMERASE/EPIMERASE	60%

15839897	Q8VKK4	FUNCTION UNPREDICTED	NO HITS	FUNCTION UNKNOWN	ATP BINDING	NO RELATED COG	20%
15839899	P65376	CONSERVED MEMBRANE PROTEIN	MYCOBACTERIUM MEMBRANE	MYCOBACTERIUM MEMBRANE	CARBOXY-LYASE ACTIVITY	NO RELATED COG	60%
15839901	P64727	FUNCTION UNPREDICTED	GLUTAREDOXIN	FUNCTION UNKNOWN	PROTEIN FUNCTION UNKNOWN	THIODISULPHATE ISOMERASE AND THIOREDOXINS	20%
15839906	Q8VKS4	MONOMETHYLAMINE TRANSFERASE	UNINTEGRATED	ATPASE ACTIVITY	HYDROLASE'S ACTIVITTY	NO RELATED COG	20%
15839907	O33359	FUNCTION UNPREDICTED	UNINTEGRATED	FUNCTION UNKNOWN	PROTEIN FUNCTION UNKNOWN	NO RELATED COG	0%
15839908	Q7D9Q7	FUNCTION UNPREDICTED	HNH ENDONUCLEASE	ANTI-SIGMA FACTOR ANTAGONIST	NUCLEIC ACID BINDING	NO RELATED COG	20%
15839909	Q7D9Q6	FUNCTION UNPREDICTED	SULFATE TRANSPORTER ANTAGONIST	LIPOLYTIC ENZYME	OXALATE TRANSPORTER ACTIVITY	ANTI-ANTI SIGMA FACTOR	20%
15839911	O33363	LIPOLYTIC ENZYME	LIPOLYTIC ENZYME	LIPOLYTIC ENZYME	ATP BINDING	MOLECULAR CHAPERON	60%
15839912	Q8VKK3	FUNCTION UNPREDICTED	UNINTEGRATED	FUNCTION UNKNOWN	ACETYL COA CARBOXYLASE ACTIVITY	NO HITS	20%
15839913	O33364	FUNCTION UNPREDICTED	ESTERASE/LIPASE/THIOESTREASE	ESTERASE	CATALYTIC ACTIVITY	NO RELATED COG	40%
15839915	Q8VKK2	FUNCTION UNPREDICTED	S-ADENOSINE L-METHINONE DEPENDENTMETYL TRANSFERASE	DEOXYNUCLEOSIDE KINASE	ATP BINDING	NO HITS	20%
15839917	Q7D9Q1	FUNCTION UNPREDICTED	MYCOBACTERIUM TUBERCULOSIS PARALOGUS FAMILY	FUNCTION UNKNOWN	HYDROLASE'S ACTIVITY	NO RELATED COG	20%
15839920	O06392	THIODISULPHATE ISOMERASE	THIOREDOXIN RELATED	THIOREDOXIN	PROTEIN FUNCTION UNKNOWN	THIO DISULPHATE ISOMERASE AND THIOREDOXINS	80%
15839922	Q7D9P8	FUNCTION UNPREDICTED	ReSB LIKE	ReSB LIKE	DNA BINDING	ReS PROTEIN REQUIRED FOR CYTOCHROME BIOSYNTHESIS	60%
15839924	Q8VKJ9	FUNCTION UNPREDICTED	UNINTEGRATED	CABYORINIC ACID a,c-DIAMIDE SYNTHASE	ATP BINDING	ATPASE INVOLVED IN CHROMOSOME PARTIONING	20%
15839925	Q8VKJ8	FUNCTION UNPREDICTED	UNINTEGRATED	AFLATOXIN BIOSYNTHESIS REGULATORY PROTEIN	ATP BINDING	NO HITS	20%
15839927	Q8VKJ7	FUNCTION UNPREDICTED	UNINTEGRATED	FUNCTION UNKNOWN	CYSTEINE TYPE PEPTIDASE ACTIVITY	NO RELATED COG	20%
15839933	O06403	FUNCTION UNPREDICTED	UNINTEGRATED	ADENYLATE CYCLASE	FATTY ACID COA LIGASE ACTIVITY	NO RELATED COG	20%
15839934	Q7D9P2	FUNCTION UNPREDICTED	UNINTEGRATED	EXTENSIN PROTEIN LIKE	PROTEIN BINDING	NO RELATED COG	20%
15839936	Q8VKJ5	FUNCTION UNPREDICTED	UNINTEGRATED	3-HYDROXYCOA CARBOXYL DEHYDROGENASE	SODIUM IODIDE SYMPORTER ACTIVITY	UNCHARACTERIZED BCR	20%
15839937	Q8VKJ4	FUNCTION UNPREDICTED	UNINTEGRATED	CARBON STARVATION PROTEIN CstA	DNA HELICASE ACTIVITY	NO RELATED COG	20%
15839939	Q8VKJ3	FUNCTION UNPREDICTED	NO HITS	FUNGAL CHITIN SYNTHASE	1,4 BETA MANNOSIDASE ACTIVITY	NO RELATED COG	20%
15839940	O06410	FUNCTION UNPREDICTED	UNINTEGRATED	NADH-UBIQUINONE OXIDOREDUCTASE	HEME TRANSPORTER ACTIVITY	NO RELATED COG	20%
15839942	O06412	FUNCTION UNPREDICTED	GLYOXALASE	GLYOXALASE	LACTOYL GLUTATHIONE LYASE	LACTOYLGLUTATHIONE LYASE	40%
15839945	Q8VKJ2	FUNCTION UNPREDICTED	NO HITS	FUNCTION UNKNOWN	PROTEIN FUNCTION UNKNOWN	NO RELATED COG	0%
15839946	O06415	PREDICTED NUCLEIC ACID BINDIIG PROTEIN	Pil T PROTEIN N-TERMINAL	FUNCTION UNKNOWN	PROTEIN BINDING	NO RELATED COG	40%
15839947	O06416	FUNCTION UNPREDICTED	RIBBON HELIX-HELIX	OCTOPAMINE RECEPTOR SIGNATURE	ATP BINDING	NO RELATED COG	20%
15839948	Q8VKJ1	FUNCTION UNPREDICTED	UNINTEGRATED	PENCILLIN BINDING PROTEIN TRANSPE	KINASE ACTIVITY	NO RELATED COG	20%
15839950	Q7D9N3	AMINO HYDROLASE'S	METALLO DEPENDENT HYDROLASE'S	AMINOHYDROLASE	HYDROLASE'S ACTIVITTY	PERDICTED METAL DEPENDED HYDROLASE'S	100%
15839954	Q7D9N0	FUNCTION UNPREDICTED	UNINTEGRATED	FecCD TRANSPORT FAMILY	MONOOXYGENASE	NO RELATED COG	20%

15839957	O06425	FUNCTION UNPREDICTED	UNINTEGRATED	Skb1 METHYLTRANSFERASE	PROTEIN BINDING	NO RELATED COG	20%
15839958	O06426	SAM DEPENDENT METHYL TRANSFERASE	SAM METHYL TRANSFERASE BINDING MOTIF	THIOPURINE S-METHYLTRANSFERASE	SAM METHYL TRANSFERASE	SAM METHYL DEPENDENT TRANSFERASE	100%

The results obtained was 100% confidence level for 2 proteins, with 80% confidence level for 4 proteins and some hypothetical proteins has not been predicted with much confidence level (i.e., Confidence level 60.0%; 40.0%; 20.0% and unknown function). All 107 hypothetical proteins were tabulated against their functions; only two proteins [GENE ID: 15839950 & 15839958] shows 100% results in future these two proteins are taken for modeling which may servers as target for drug designing.

Thus, by this functional genomics approach, the functions of hypothetical proteins of *Mycobacterium tuberculosis* were delineated and their target is identified which paved a way for curing this dreadful disease through docking approach.

CONCLUSION: The function of hypothetical proteins of *Mycobacterium tuberculosis* is predicted using Bioinformatics tools BLAST, BLOCKS, COGs, InterProScan, PROTFUN and PFP (**Table 2**).

TABLE 2: PERCENTAGE OF SIMILARITY

S. No.	Number of proteins in 100 %	Number of proteins in 80 %	Number of proteins in 60 %	Number of proteins in 40 %	Number of proteins in 20 %	Number of proteins in 0%
1	2	3	7	13	78	4

Among 107 hypothetical proteins, only two proteins show 100% confidence level in all the mentioned tools. Three proteins were found to have the same function in any of the four tools. Seven Proteins were found to have the same function in any of the three tools. Thirteen Proteins were found to have the same function in any of the two tools. Seventy eight Proteins were found to have the some function in any of the tools. Three Proteins were found to have no functions. Two proteins those obtained same function such as METHYL TRANSFERASE and HYDROLASE'S. Further these two proteins may serves as target for few antibiotics, which helps in curing this dreadful disease.

REFERENCES:

1. Kumar, Vinay, Abbas, Abul K, Fausto, Nelson, Mitchell and Richard Robbins: Basic Pathology, Saunders Elsevier. 2007; 8, 516-522.
2. Niyaz Ahmed and Seyed E, Hasnain: Genomics of *Mycobacterium tuberculosis*: Old threats & new trends. Indian J Med Res. 2004; 20, 207-212.
3. http://www.pasteur.fr/X-TB/
4. Tatusov R L, Galperin M, Natale D A and Koonin EV: The COG database: a tool for genome-scale analysis of protein functions and evolution. 2000; Nucleic 28(1), 33-36.
5. Mulder N J, Apweiler R, Attwood T K, Bairoch A, Bateman A, Binns D, Bork P, Buillard V, Cerutti L, Copley R, Courcelle E, Das U, Daugherty L, Dibley M, Finn R, Fleischmann W, Gough J, Haft D, Hulo N, Hunter S, Kahn D, Kanapin A, Kejariwal A, Labarga A, Langendijk-Genevaux P S, Lonsdale D, Lopez R, Letunic I, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Nikolskaya A N, Orchard S, Orengo C, Petryszak R, Selengut J D, Sigrist C J, Thomas P D, Valentin F, Wilson D, Wu CH and Yeats C: New developments in the InterPro database. Nucleic Acids Res. 2007; 35, D224-8.
6. Whisstock J C and Lesk A M: Prediction of Protein Function from protein sequence and structure. Q Rev Biophys.2003; 36(3), 307-40.
7. http://expasy.org/
8. http://www.ncbi.nlm.nih.gov
9. http://www.ncbi.nlm.nih.gov/sites/entrez?db=genome&cmd=Retrieve&dopt=Protein%20Table&list_uids=181.
10. http://www.ncbi.nih.gov/BLAST/
11. http://blocks.fhcrc.org/
12. http://www.ncbi.nlm.nih.gov/COG/
13. http://www.ebi.ac.uk/interpro/
14. http://dragon.bio.purdue.edu/pfp
15. Jensen L J, Ussery D W, Brunak S: Functionality of system components: conservation of protein function in protein feature space. Genome Res.2003: 13 (11): 2444-2449.

****