Bioinformation
www.bioinformation.net

open access
Hypothesis

Computational genome analyses of metabolic enzymes in *Mycobacterium leprae* for drug target identification

Anusuya Shanmugam[1], Jeyakumar Natarajan[2]*

[1]Department of Bioinformatics, VMKV Engineering College, Vinayaka Missions University, Salem; [2]Department of Bioinformatics, Bharathiar University, Coimbatore; Jeyakumar Natarajan - Email: n.jeyakumar@yahoo.co.in; *corresponding author.

Received November 04, 2009; accepted March 06, 2010; published March 31, 2010

Abstract:
Leprosy is an infectious disease caused by *Mycobacterium leprae*. *M. leprae* has undergone a major reductive evolution leaving a minimal set of functional genes for survival. It remains non-cultivable. As *M. leprae* develops resistance against most of the drugs, novel drug targets are required in order to design new drugs. As most of the essential genes mediate several biosynthetic and metabolic pathways, the pathway predictions can predict essential genes. We used comparative genome analysis of metabolic enzymes in *M. leprae* and *H. sapiens* using KEGG pathway database and identified 179 non-homologues enzymes. On further comparison of these 179 non-homologous enzymes to the list of minimal set of 48 essential genes required for cell-wall biosynthesis of *M. leprae* reveals eight common enzymes. Interestingly, six of these eight common enzymes map to that of peptidoglycan biosynthesis and they all belong to Mur enzymes. The machinery for peptidoglycan biosynthesis is a rich source of crucial targets for antibacterial chemotherapy and thus targeting these enzymes is a step towards facilitating the search for new antibiotics.

Keywords: Comparative genomics, Mur enzymes, *M. leprae*, Leprosy.

Background:
The availability of the complete genome sequences of several pathogenic bacteria and the completion of the human genome project has revolutionized the field of drug-discovery against threatening human pathogens **[1]**. Novel drug targets are required in order to design new drugs against antibiotic sensitive pathogens. In general, a target should provide adequate selectivity yielding a drug which is specific or highly selective against the pathogen with respect to the human host. Moreover, the target should be essential for growth and viability of the pathogen at least under the condition of infection **[2]**. The search for potential drug targets has increasingly relied on genomic approaches. The entire approach is built on the assumption that the potential target must play an essential role in the pathogen's survival and constitute a critical component in its metabolic pathway. At the same time, this target should not have any well-conserved homolog in the human host. This would preclude the possibilities of unacceptable cross-reactivity that might prove detrimental to the host **[3]**. Leprosy is caused by *Mycobacterium leprae*, which primarily affects the skin, mucous membranes and peripheral nerves causing deformities. Leprosy remains a major global health problem, especially in the developing world. For over a century leprosy has presented major challenges in the fields of microbiology, pathology, immunology, and genetics; it continues to do so today. Computer analysis demonstrated that only half of the sequence contains protein-coding genes. The other half contains pseudo genes and non-coding sequences. These findings indicate that *M. leprae* has undergone a major reductive evolution leaving a minimal set of functional genes for survival [**4**]. Study of the coding region of the sequence provides evidence accounting for the particular pathogenic properties of *M. leprae* which is an obligate intracellular parasite. *Mycobacterium leprae* remains non-cultivable **[5].** *M. leprae* had mutational changes in some of their genes like gyrA, rpoB, and folP which developed resistance against drugs like newer quinolones, refampicin and dapsone [**6**]. Resistant strains of *M. leprae* appeared due to mutations in the macrolide target, the ribosome [**7**]. These findings suggest the emergence of multi-drug resistant *M. leprae*. Hence the mycobacterial cell wall with its specific composition and structure is considered to be a major factor in promoting the natural resistance of mycobacteria to various antibiotics. Early detection of *Mycobacterium leprae* infection is considered an important component of strategies aiming at reducing transmission of infection, but currently available diagnostic tools often lack sufficient sensitivity and specificity to reach this goal **[8]**. In clinical studies, notable progress has been made concerning the immunology and immunopathology of leprosy, the genetics of human resistance, mechanisms of nerve injury, and chemotherapy. In nearly all of these areas, however, leprosy remains poorly understood compared to other major bacterial diseases **[9].** Here, we present a computational approach to identify the genes essential to *M. leprae* using comparative pathway analysis followed by mapping of non-homologues genes with list of minimal set of essential genes required for cell-wall biosynthesis of *M. leprae.* In addition, our approach successfully identified a unique group of common enzymes as promising protein targets for new antibiotic development and further characterization in the laboratory.

Methodology:
Collection of metabolic pathway enzymes of *M. leprae*
Kyoto Encyclopedia of Genes and Genomes (KEGG) **[10]** is a collection of online databases dealing with genomes, enzymatic pathways, and biological chemicals. KEGG maintains five main databases. They are KEGG Atlas, KEGG Pathway, KEGG Genes, KEGG Ligand and KEGG BRITE. First, we collected all the metabolic pathways of *M. leprae* and *H. sapiens* from KEGG pathway database. Each of the pathways of *M. leprae* was compared with all the available pathways of *H. sapiens* to identify whether that particular pathway of *M. leprae* is present in *H. sapiens* or not. The pathways which were present in both *M. leprae* and *H. sapiens* were separated out and were named as shared pathways. The pathways which were present only in *M. leprae* but were not present in *H. sapiens* were grouped together and were called as unique pathways. The gene name and the enzyme commission number (EC) of all the enzymes present in both shared and unique pathways were identified and collected from KEGG Genes database.

Retrieval of protein sequences and BLAST
The protein sequence of all enzymes in both shared and unique pathways of *m. leprae* were retrieved from UNIPROT **[11]** in FASTA format. Each protein sequence was subjected to BLASTP analysis against the *H. sapiens* at an E-value cutoff of 10^{-4} **[12].** BLAST results with no hits with *H. sapiens* were identified as non-homologues enzymes of *M. lepare*.

Identification of essential Enzymes
The minimal set of essential genes required for cell envelope biosynthesis of *m. leprae* was reported previously using comparative genome sequence method by Vissa and Brennan **[13]**. The *M. leprae* enzymes which were non homologous to *H. sapiens* were mapped with the gene list of Vissa & Brennan and the most common *M. leprae* genes were identified and further explored.

Results and discussion:
Metabolic pathway information
In KEGG pathway database we found 99 metabolic pathways for *M. leprae* and 210 metabolic pathways for *H. sapiens*. Out of 99 metabolic pathways for *M. leprae* five pathways are unique to *M. leprae* alone and comprised of 29 enzymes and remaining 94

pathways are present in *H. sapiens* as well and comprised of 731 enzymes. The five unique pathways of *m. leprae*, 29 enzymes specific to these pathways and their corresponding gene id and the EC number were represented **Table 1 (see supplementary material)**.

Prediction of enzymes which were non homologous to human

Removing enzymes from the pathogen that share a similarity with the host protein ensures that the targets have nothing in common with the host proteins and thereby, eliminating undesired host protein-drug interactions. BLASTP similarity search of all these 760 (29 unique + 731 shared) enzymes at an *E*-value cutoff of 10^{-4} resulted 179 non-homologues enzymes of *m. leprae* of which ten enzymes from the unique pathways and the remaining 169 belong to enzymes from shared pathways. All these 179 enzymes with their corresponding gene-id and EC number were represented in **Table 2 (see supplementary material)**.

Comparison of non-homologues enzymes with essential gene set

The 179 (10 + 169) non-homologues enzymes were further compared to the minimal set of 48 essential genes required for cell-wall biosynthesis of *M. leprae* and reported by Vissa and Brennan **[13]**. There are eight enzymes common in both data sets (**Table 3 in supplementary material**). Among the eight common enzymes only one enzyme was found to be present in unique pathway and the remaining seven enzymes were found to be present in shared pathways. All these eight enzymes were categorized as essential enzymes of *m. leprae.*

Role of essential enzymes of *M. leprae*

All the eight essential enzymes were further analyzed for the identification of potential drug targets. One of the eight essential enzymes Alanine racemase (alr) is the enzyme found in D-Alanine metabolism which is a unique pathway of *M. leprae*. It is also found in Alanine and Aspartate metabolism which is a shared pathway of *M. leprae*. Another essential enzyme Putative dTDP-4-dehydrorhamnose 3, 5-epimerase (rmlC) was found to be essential for Nucleotide sugar metabolism which is a shared pathway of *M. leprae*. It was also found to be essential for polyketide sugar unit biosynthesis which is a unique pathway of *M. leprae*. The remaining 6 essential enzymes murC, murD, murE, murF, murG and murY were found to be essential for Peptidoglycan biosynthesis. It is noteworthy that all these 6 enzymes belong to the same family. This particular pathway, peptidoglycan biosynthesis was analyzed for the prediction of drug targets.

Peptidoglycan biosynthesis and Mur enzymes

M. leprae posses a multilayered cell envelope which basically consisted of, from inner to outer layer, a plasma membrane (PM), a peptidoglycan layer (PG), an electron translucent layer (ETL), and an irregular electron dense outer layer (OL) **[14].** This bacterial cell envelope provides strength and rigidity to counteract internal osmotic pressure, and protection against the environment. The peptidoglycan layer gives the cell wall its strength, and helps to maintain the overall shape of the cell. The basic peptidoglycan structure of both Gram-positive and Gram-negative bacteria is comprised of a sheet of glycan chains connected by short cross-linking polypeptides. Biosynthesis of peptidoglycan is a multi-step process comprising three main stages: (1) Formation of UDP-N-acetylmuramic acid (UDPMurNAc) from N-acetylglucosamine (GlcNAc). (2) Addition of a short polypeptide chain to the UDPMurNAc. (3) Addition of a second N-acetylglucosamine (GlcNAc) to the disaccharide-pentapeptide building block and transport of this unit through the cytoplasmic membrane and incorporation into the growing peptidoglycan layer.

The second step of the petidoglycan biosynthesis was carried out by four of the mur ligase enzymes MurC, MurD, MurE and MurF. These four Mur ligases are responsible for the successive additions of L-alanine, D-glutamate, meso-diaminopimelate or L-lysine, and D-alanyl-D-alanine to UDP-N-acetylmuramic acid. The final step in the formation of peptidoglycan was carried out by murG enzyme (N-acetyl glucosaminyl transferase). This enzyme is peripherally associated with the inner face of the cytoplasmic membrane. Therefore, the peptidoglycan subunit is completely assembled before it traverses the cytoplasmic membrane. Phospho-N-acetylmuramoyl-pentapeptide-transferase (mraY) is an important enzyme in murein synthesis. It is responsible for the formation of the first lipid intermediate of the cell wall peptidoglycan synthesis **[15]**. As the layer of the bacterial cell wall that confers strength is the peptidoglycan meshwork if we target murC, murD, murE and murF which catalyze the addition of a short polypeptide chain to the UDP-N-acetylmuramic acid (UDPMurNAc), we can easily prevent the synthesis of bacterial cell wall. Thus, these are excellent candidates for further exploration.

Conclusion:

The availability of full genome sequences and computer-aided analysis to identify probable antimicrobial drug targets has become a new trend in pharmacogenomics. The use of a comprehensive set of unique pathways and enzymes present in these pathways of *M. leprae* to identify new drug targets were documented in this study. We have found peptidoglycan biosynthetic pathway and the six mur enzymes (murC, murD, murE, murF, murG and murY) involved in this pathway to be used as potential drug targets. Protein structure and inhibitors of these important enzymes are not currently available. Further analysis on the structural studies on these mur enzymes is believed to provide valuable insights towards the design of an inhibitor specific to the peptidoglycan biosynthesis of *M. leprae* for the treatment of leprosy. The availability of the newer anti-leprotic drugs in the future would definitely support our present findings such that there would be a possibility of mur enzymes which were proposed by us for targeting *M. leprae*.

Acknowledgment:

We would like to acknowledge Dr. R. Ravishankar and Dr. S. Sivamani for critical reading of the manuscript and for helpful discussions.

References:

[1] L Miesel *et al.*, *Nat. Rev. Genet.* (2003) **4**: 442 [PMID:12776214]
[2] MK Sakharkar *et al.*, *In Silico Biol.* (2004) **4:** 0028 [PMID:15724285]
[3] B Frtiz, GA Raczniak, *BioDrugs.* (2002) **16:** 331 [PMID:12408737]
[4] N Honore, *Med Trop (Mars)* (2002) **62:** 473 [PMID:12616936]
[5] DM Scollard *et al., Clin Microbiol Rev* (2006) **19:** 338 [PMID:16614253]
[6] S Maeda, *Antimicrob Agents Chemother.* (2001) **45:** 3635 [PMID:11709358]
[7] F Doucet-Populaire *et al*., *Curr Drug Targets infect Disord* 2002 Dec; 2(**4**): 355-70 [PMID:12570741]
[8] A Geluk *et al*., *Infect Immun*. 2005 Sep; 73(**9**): 5636-44 [PMID:16113281]
[9] DM Scollard *et al., Clin Microbiol Rev*. 2006 Apr; 19(**2**): 338-81 [PMID:16614253]
[10] http:// www.genome.jp/kegg.html
[11] http:// www.uniprot.org/
[12] Altschul *et al.*, *Nucleic Acids Res*. 1997 Sep, 1:25(**17**): 3389-3402 [PMID:9254694]
[13] VD Varalakshmi, PJ Brennan, *Genome Biology* (2001) 2(8):Reviews1023 [PMID: 11532219]
[14] AI Takade *et al*., *Microbiol Immun*. (2003) 47(**4**): 265-70 [PMID:12801063]
[15] http://www.ebi.ac.uk/interpro

Edited by P. Kangueane
Citation: Shanmugam & Natarajan, Bioinformation 4(9): 392-395 (2010)

Supplementary material:

Table 1: The five unique pathways and 29 enzymes specific to these pathways of *m. leprae*

S. No	Name of the Pathway	Enzyme	EC #
1. D-Alanine metabolism			
1	D-Alanyl-alanine synthetase A	Ddl	6.3.2.4
2	Alanine racemase	Alr	5.1.1.1
2. Ethylbenzene degradation			
3	Acetyl-CoA acyltransferase	fadA	2.3.1.16
4	Mycolytransferase	fbpA	2.3.1.-
3. Benzoate degradation via CoA ligation			
5	Monooxygenase		1.14.13.-
6	Hypothetical protein		2.7.1.-
7	Amidase	amiE	3.5.1.4
8	Antigen 85-A	fbpA	2.3.1.-
9	Putative acyl-CoA synthetase	fadD32	6.2.1.-
10	Enoyl-CoA hydratase	echA1	4.2.1.17
11	Acetyl-CoA acetyltransferase	fadA4	2.3.1.9
12	Succinate dehydrogenase iron-sulfur subunit	sdhB	1.3.99.1
4. 1,1,1-Trichloro-2,2-bis(4-chlorophenyl)ethane (DDT) degradation			
13	Trans-acting enoyl reductase		1.3.1.-
14	Putative oxidoreductase		1.2.-.-
5. Two component system			
15	Acetyl-CoA acetyltransferase	A to B	2.3.1.9
16	cytochrome c oxidase subunit XV assembly protein (A)	CtaA	
17	Glutamine synthetase	GlnA	6.3.1.2
18	Putative two-component response regulator	MprA	
19	Putative two-component system sensor histidine kinase	MprB	2.7.13.3
20	Putative two-component response regulator	MtrA	
21	Putative two-component system sensor histidine kinase	MtrB	2.7.13.3
22	Putative serine protease	PepD	3.4.21.-
23	Two-component response regulator	PrrA	
24	Sensor histidine kinase	PrrB	2.7.13.3
25	Tryptophan synthase subunit alpha	TrpA	4.2.1.20
26	Tryptophan synthase subunit beta	TrpB	4.2.1.20
27	Indole-3-glycerol-phosphate synthase	TrpC	4.1.1.48
28	Anthranilate phosphoribosyl transferase	TrpD	2.4.2.18
29	Anthranilate synthase component I	TrpE	4.1.3.27

Table 2: The 179 non-homologous enzymes *m. leprae*

Gene	EC #	Gene	EC #	Gene	EC #	Gene	EC #	Gene	EC #
ppgK	2.7.1.63	Adi	4.1.1.19	murG	2.4.1.227		1.2.1.-	SecA,azi,div	
glpX	3.1.3.11	murI	5.1.1.3	mraY	2.7.8.13		3.3.2.8	SecD	
Fba	4.1.2.13	panD	4.1.1.11	uppP	3.6.1.27		3.1.1.45	SecE	
Fba	1.1.1.42	Asd	1.2.1.11	thiE	2.5.1.3	rpoA		SecF	
rpiB	5.3.1.6	thrB	2.7.1.39	thiL	2.7.4.16	Omega	2.7.7.6	SecG	
otsA	2.4.1.15	Ask	2.7.2.4	ThiC		rplO		SPase I	3.4.21.89
otsB2	3.1.3.12	sdaA	4.3.1.17	ThiG		rplR		SPase II	3.4.23.36
	3.2.1.21	metE	2.1.1.14	ribG	1.1.1.193	rpiU		TatA	
murA	2.5.1.7	ilvC	1.1.1.86	cobT	2.4.2.21	rplV		TatB	
murB	1.1.1.158	ilvD	4.2.1.9	ribC	2.5.1.9	rplY		TatC	
rmlC	5.1.3.13	Hom	1.1.1.3	ribA	3.5.4.25	rpmC		YajC	
glcB	2.3.3.9	dapD	2.3.1.117	ribG	3.5.4.26	rpmD		YidC	
ppdK	2.7.9.1	dapF	5.1.1.7	RIBB		rpmG		dnaE – α subunit	2.7.7.7
Ppc	4.1.1.31	murF	6.3.2.10	YaaD	4.-.-.-	rpmE		dnaN – β subunit	2.7.7.7
gltB	1.4.1.13	murE	6.3.2.13	YaaE	2.6.-.-	rpmF		Subunit – δ	2.7.7.7
cynT	4.2.1.1	hisD	1.1.1.23	nadD	2.7.7.18	rpmH		DnaB	3.6.1.-
cysE	2.3.1.30	hisG	2.4.2.17	nadA		rpmI		DnaG	2.7.7.-
metX	2.3.1.31	hisI	3.5.4.19	COAX	2.7.1.33	rpmJ		rnhB	3.1.26.4
Dxr	1.1.1.267	hisE	3.6.1.31	acpS	2.7.8.7	Rrs		ssb	
pgsA	2.7.8.5	hisB	4.2.1.19	panC	6.3.2.1	Rrf		Nfo	3.1.21.2
pssA	2.7.8.8	hisA	5.3.1.16	panB	2.1.2.11	rpsC		Tag	
Cdh	3.6.1.26	HisF	4.1.3.-	bioB	2.8.1.6	rpsD		UVRC	
Psd	4.1.1.65	HisH	2.4.2.-	folB	4.1.2.25	rpsJ,nusE		xseA	3.1.11.6
dnaN	2.7.7.7	aroE	1.1.1.25	Folk	2.7.6.3	rpsN		Dpo III	2.7.7.7
Dgt	3.1.5.1	ML2472	1.3.1.12	folP	2.5.1.15	rpsO		PriA	3.6.1.-
thyX	2.1.1.148	trpD	2.4.2.18	hemD	2.1.1.107	rpsQ		RecF	
Cmk	2.7.4.14	aroA	2.5.1.19	hemD	4.2.1.75	rpsT		RecO	
pyrH	2.7.4.22	aroG	2.5.1.54	COX15		OppB		RecR	
pyrF	4.1.1.23	aroD	4.2.1.10	MENC	4.2.1.113	OppC		RuvA	
murC	6.3.2.8	pheA	4.2.1.51	MenD	2.5.1.64	ProX		RuvC	3.1.22.4
murD	6.3.2.9	aroB	4.2.3.4	argC	1.2.1.38	PstA		argB	2.7.2.8

CbiQ		hisA	5.3.1.24	argJ	2.3.1.1	pstC		MenD	4.1.1.71
FhuC	3.6.3.34	OppA		argJ	2.3.1.35	RfbA		YrbD	
FhuD		FtsX		YrbE		YadH		PrrA***	
ddl	6.3.2.4	prrB	2.7.13.3	ctaA		trpC	4.1.1.48		2.7.1.-
Alr	5.1.1.1	trpA	4.2.1.20	mprB	2.7.13.3		1.2.-.-		

Table 3: Essential enzymes of *M. leprae*

S. No	Enzyme name	Name of the Enzyme	Name of the pathway
1	**murC**	UDP-N-acetylmuramate—L-alanine ligase	D-Glutamine & D-Glutamate metabolism. Peptidoglycan biosynthesis
2	**murD**	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase	D-Glutamine & D-Glutamate metabolism. Peptidoglycan biosynthesis
3	**murE**	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	Peptidoglycan biosynthesis Lysine biosynthesis
4	**murF**	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanyl ligase (D-alanine:D-alanine-adding enzyme)	Peptidoglycan biosynthesis Lysine biosynthesis
5	**murG**	N-acetylglucosaminyl transferase	Peptidoglycan biosynthesis
6	**mraY**	Phospho-N-acetylmuramoyl-pentapeptide-transferase	Peptidoglycan biosynthesis
7	**rmlC**	Putative dTDP-4-dehydrorhamnose 3,5-epimerase	Nucleotide sugars metabolism. Polyketide sugar unit biosynthesis
8	**alr**	Alanine racemase	D-Alanine metabolism. Alanine & Aspartate metabolism