International Journal of Research Instinct
(Peer-reviewed Journal)

ISSN ONLINE 2348-2095
Research Article

COMPARITIVE *IN-SILICO* SUBTRACTIVE GENOMIC STUDY FOR THE IDENTIFICATION OF POTENTIAL DRUG TARGETS FROM VARIOUS METABOLIC PATHWAYS IN *STAPHYLOCOCCUS AUREUS* NCTC 8325

[1]Sherlin Rosita . A and [2*]Kirubakaran S.A.
[1]Research scholar and [2*]Assistant professor,
PG and Research Department of Biotechnology,
Srimad Andavan Arts and Science College (Autonomous), Trichy-5.
E-mail: *envirokiruba@gmail.com*

ABSTRACT

Large genomic sequencing projects of pathogens as well as human genome leads to immense genomic and proteomic data which could be beneficial for the novel target identifications in pathogens. Subtractive genomic approach is one of the recently adopted methodology in which the subtraction of sequence between the host and pathogen proteome which provides information for a set of proteins that are likely to be essential to the pathogen but absent in the host. With this aim, we performed an *in silico* comparative analysis of metabolic pathways of the host *Homo sapiens* and the pathogen *Staphylococcus aureus* strain NCTC 8325. We identified several proteins in the metabolic pathways that are unique to the pathogen. Subsequently proteins of pathogen which are unique were compared with host proteins. Our study revealed 196 genes are non homologous to human genome, and by screening these genes using Database of Essential Genes resulted in the identification of 129 genes as essential genes of the bacteria out of which 122 genes have no human homolog. We further assessed the drugability potential of the targets by comparing them with Drug Bank Database.The study was successful in listing out potential drug targets from the *Staphylococcus aureus* strain NCTC 8325 proteome involved in vital aspects of the pathogen's metabolism. This systematic evaluation of metabolic pathways of host and pathogen through reliable and conventional bioinformatic methods can be extended to other pathogens of clinical interest.

KEYWORDS

Comparative metabolic pathways, *Staphylococcus aureus*, homolog, Subtractive Genomics, druggability, KEGG.

INTRODUCTION

Staphylococcus aureus (*S. aureus*) is a gram positive bacteria and it commonly cause skin and soft tissue infections worldwide and has also been implicated in conditions such as furuncle/ carbuncles, aellulities, septic thrombophlebitis, bacteraemia, epidural abscess, osteomyelitis and prosthetic joint infection[1]. *S.aureus* strain NCTC 8325 is generally regarded as the protypical strain for all genetic manipulation. It produces numerous toxins including super antigens that cause unique disease entities such as toxic-shock syndrome and staphylococcal scarlet fever[2]

Since existing antibiotics are ineffective against S.*aureus* strain NCTC 8325, the discovery of novel antibiotics is of prime need. Analysis of complete genomes allows us to compile a list of potential gene products and identify the functions present in the host and absent in the pathogen[3]. This reduces the problem of searching for potential drug targets from a large list to selecting from a chosen few. As most currently known, antibacterials are essentially inhibitors of certain bacterial enzymes, all enzymes specific to bacteria can be considered as potential drug targets[4,5]. The Subtractive genomic approach is one of the recently adopted methodology in which the subtraction of sequence between the host and pathogen proteome provides information for a set of proteins that are likely to be essential to the pathogen but absent in the host[6,7,8]

In this study, we have adopted a strategy of comparative metabolic pathway analysis with the help of KEGG database. The enzymes in the pathways of S. *aureus* strain NCTC 8325, which do not show similarity to any protein from the host, represent attractive potential drug targets. The elimination of pseudo drug targets is essential since the cost involved in the investigation of drug targets is prohibitive. The aim of the current study is to propose those targets which are essential for the survival of the pathogen and at the same time should not have homology with the human host. It was expected that our results would facilitate the selection of *S.aureus* NCTC 8325 drug targets which could enter into a successful drug design pipeline.

MATERIALS AND METHODS

The methodology comprised of the following steps:

Comparative metabolic pathway analysis

The complete metabolic pathway information of the host *Homo sapiens* and the pathogen *S. aureus* NCTC 8325 was retrieved from the pathway database of Kyoto Encyclopedia of Genes and Genomes (KEGG).The three letter code specific for each organisms i.e, *Homo sapiens* (hsa) and *S.aureus* (sao) were placed in the genome comparison and combination box for comparison. Each pathway name and their proteins were enlisted and manually compared with human host. As a result the pathways lacking in *Homo sapiens* but existing in *S. aureus* NCTC 8325 were indicated as

unique to *S. aureus* NCTC 8325 (24 unique metabolic pathways).The remaining pathways were enlisted as the common with the host. The sequences of the respective protein involved in the unique pathways were then indentified. The protein sequences which are unique for *S. aureus* NCTC 8325 were searched for its sequence homology with human proteome using BLASTp program available at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) bit score cut off < 100 and minimum expectation value (E-value) cut off E^{-10} were taken to identify homology exhibiting significant differences with their human counterpart. Protein sequences less than 100 amino acids in length were unlikely to represent essential to *S. aureus* NCTC 8325 hence such sequences were excluded from analysis.

Subtractive Genomic Approach

Non human homolog proteins were then searched against DEG (http://tubic.tju.edu.cn/deg/).A non-human homologue (not present in the host but present in the pathogen) is considered as a good target against the pathogen[9]. The amino acid sequences of unique enzymes were subjected to similarity search against the Database of Essential genes (DEG) (http://tubic.tju.edu.cn/deg/) with an E-value cut off 10^{-10} and a minimum bit score of 100 to find the essentiality of the unique enzymes which results in the identification of potential drug targets.

RESULTS AND DISCUSSION

Modern tools of computational biology greatly enhance the speed and reliability of antimicrobial discovery[10].With an objective of identifying proteins potentially useful as drug targets, we have relied on the use of proteomes and a subtractive genomic approach for the identification of putative drug targets. A subtractive genomics approach utilizes the whole proteome of host and pathogen to identify proteins exclusively present in the pathogen by deducing the homologous proteins[11,12,13]. The essential proteins of S.*aureus* strain NCTC 8325 were identified using subtractive genomics approach and further analyzed for identification of putative potential drug targets. The results are given in table 1

Table 1: Summary of subtractive analytical result of S.*aureus* NCTC 832

S.No	Name of the step	Numbers
1	Complete pathways of the pathogen from KEGG (sao) S.aureus NCTC 8325	101
2	Complete pathways of the human from KEGG (hsa) *Homosapiens*	291
3	Comparison between sao and hsa Common metabolic pathways Unique metabolic pathways	77 24
4	Protein sequence of unique metabolic pathways	196

5	Removal of redundant KEGG IDs (manually)	160
6	BLASTp of unique metabolic protein sequences against prokaryotic protein database of DEG	129
7	CD-hit of unique essential protein sequence against human host proteome	122
8	BLASTp of unique essential non-homolog protein sequence against DBD	77

Comparative Metabolic pathway analysis

The basic aim of this step was to perform a comparison between the complete metabolic pathways of S.*aureus* NCTC 8325 and its human host which eventually resulted in listing the unique as well as the common metabolic pathways. Subsequent retrieval of the protein sequences of only the unique metabolic pathways would lead to prioritized potential drug targets. KEGG database was used to perform the genome-wide metabolic pathway analysis. The complete metabolic pathway information of the pathogen S.*aureus* NCTC 8325(sao) and the host *Homo sapiens*(has) were retrieved from the KEGG database. KEGG contains 101 and 291 metabolic pathways of sao and has respectively. Each pathway name and their proteins were enlisted and manually compared with the human host. As a result, the pathways lacking in human but existing in *S.aureus* strain NCTC 8325 (24 unique metabolic pathways) and the remaining pathways were enlisted as common with the host. A complete list of unique as well as common metabolic pathways along with their KEGG pathway IDs and corresponding number of proteins are provided in Table 2 and 3.

Table : 2 Common metabolic pathways of *S.aureus* NCTC 8325 against the human host based upon KEGG annotations

S.NO	KEGG pathway IDs	Common metabolic pathways	No. of. Proteins
1	Sao00010	Glycolysis/Gluconeogenesis	37
2	Sao00020	Citrate cycle(TCA cycle)	22
3	Sao00030	Pentose phosphate pathway	20
4	Sao00040	Pentose and glucuronate interconversion	10
5	Sao00051	Fructose and mannose metabolism	16
6	Sao00052	Galactose metabolism	20
7	Sao00053	Ascorbate and aldarate metabolism	6

8	Sao00061	Fattyacid biosynthesis	14
9	Sao00071	Fatty acid degradation	11
10	Sao00072	Synthesis and degradation of Ketone bodies	3
11	Sao00130	Ubiqinone and other terpenoid quinone biosynthesis	8
12	Sao00190	Oxidative phophorylation	22
13	Sao00230	Purine metabolism	54
14	Sao00240	Pyrimidine metabolism	46
15	Sao00250	Alanine, aspartate and glutamate metabolism	18
16	Sao00260	Glycine, serine and threonine metabolism	31
17	Sao00270	Cysteine and methionine metabolism	21
18	Sao00280	Valine, leucine and isoleucine degradation	13
19	Sao00290	Valine, leucine and isoleucine biosynthesis	13
20	Sao00300	Lysine biosynthesis	13
21	Sao00310	Lysine degradation	11
22	Sao00330	Arginine and proline metabolism	27
23	Sao00340	Histidine metabolism	15
24	Sao00350	Tyrosine metabolism	5
25	Sao00360	Phenylalanine metabolism	3
26	Sao00380	Tryptophan metabolism	9
27	Sao00400	Phenylalanine, tyrosine and tryptophan biosynthesis	19
28	Sao00410	β-Alanine metabolism	6
29	Sao00430	Taurine and hypotaurine metabolism	6
30	Sao00450	Selenocompond metabolism	8
31	Sao00460	Cyanoaminoacid metabolism	3
32	Sao00471	D-Glutamine and D-Glutamate metabolism	3
33	Sao00472	D-Arginine and D-Ornithine metabolism	1
34	Sao00480	Glutathione metabolism	7
35	Sao00500	Starch and sucrose metabolism	14

36	Sao00520	Aminosugar and nucleotide sugar metabolism	29
37	Sao00561	Glycerolipid metabolism	15
38	Sao00562	Inositol phosphate metabolism	3
39	Sao00564	Glycerophospholipid metabolism	15
40	Sao00590	Arachidonic acid metabolism	2
41	Sao00592	Alpha- Linolenic acid	1
42	Sao00620	Pyruvate metabolism	38
43	Sao00630	Glyoxylate and dicarboxylate metabolism	13
44	Sao00640	Proponate metabolism	18
45	Sao00650	Butanoate metabolism	16
46	Sao00670	One carbon pool by folate	9
47	Sao00730	Thiamine metabolism	7
48	Sao00740	Riboflavin metabolism	6
49	Sao00750	Vitamin B6 metabolism	4
50	Sao00760	Nicotinate and nicotinamide metabolism	7
51	Sao00770	Pantothenate and COA biosynthesis	16
52	Sao00780	Biotin metabolism	11
53	Sao00785	Lipoic acid metabolism	4
54	Sao00790	Folate biosynthesis	18
55	Sao00860	Porphyrin and chlorophyll metabolism	14
56	Sao00900	Terpenoid backbone biosynthesis	14
57	Sao00910	Nitrogen metabolism	13
58	Sao00920	Sulfur metabolism	9
59	Sao00970	Aminoacyl tRNA biosynthesis	85
60	Sao01040	Biosynthesis of unsaturated fatty acid	2
61	Sao01200	Carbon metabolism	-
62	Sao01212	2-Oxocarboxylic acid	-
63	Sao01220	Degradation of aromatic acid	-
64	Sao01230	Biosynthesis of aminoacid	-
65	Sao02010	ABC transporters	97

66	Sao03010	Ribosome	71
67	Sao03018	RNA degradation	14
68	Sao03020	RNA polymerase	5
69	Sao03030	DNA replication	16
70	Sao03060	Protein export	16
71	Sao03410	Base excision repair	10
72	Sao03420	Nucleotide excision repair	8
73	Sao03430	Mismatch excision repair	17
74	Sao03440	Homologous recombination	20
75	Sao04122	Sulfur relay system	8
76	Sao05150	Staphylococcus aureus infection	38

Table :3 Unique metabolic pathways of S.*aureus* NCTC 8325 against the human host based upon KEGG annotations

S.NO	KEGG pathway Ids	Unique metabolic pathways	No of proteins
1	Sao00121	Secondary bile acid biosynthesis	1
2	Sao00281	Geraniol degradation	2
3	Sao00362	Benzoate degradation	5
4	Sao00401	Novobiocin biosynthesis	3
5	Sao00473	D-Alanine metabolism	6
6	Sao00521	Streptomycin biosynthesis	3
7	Sao00550	Peptidoglycan biosynthesis	20
8	Sao00625	Chloroalkane and Chloroalkene degradation	5
9	Sao00626	Napthalene degradation	2
10	Sao00627	Aminobenzoate degradation	3
11	Sao00642	Ethylbenzene degradation	2
12	Sao00660	C5-branched dibasic acid metabolism	10

13	Sao00680	Methane metabolism	19
14	Sao00900	Terpenoid backbone biosynthesis	14
15	Sao00903	Limonene and pinene degradation	4
16	Sao00906	Carotenoid biosynthesis	5
17	Sao01110	Biosynthesis of secondary metabolites	-
18	Sao01120	Microbial metabolism in diverse environments	-
19	Sao01501	β-Lactam resistance	-
20	Sao01502	Vancomycin resistance	-
21	Sao02020	Two-component system	52
22	Sao02060	Phosphotransferase system	24
23	Sao03070	Bacterial secretion system	13
24	Sao05100	Bacterial invasion of epithelial cell	2

Identification of essential and non-host unique proteins

Essential genes/gene products are required by pathogenic organism for survival. Identifying unique essential proteins in microorganism is pivotal in the development of novel drugs[14.] According to Haag *et al*[15]and Butt *et al*[16] , the identification of potential drug target is to identify genes essential for the survival of pathogen and to identify genes absent in the human host is by using DEG. In the present study it was observed that the protein sequences present in the unique metabolic pathways of the pathogens are196. The resulting 196 sequences were then subjected to Blastp using DEG. From that only 129 essential protein sequences were identified. These identified sequences are viable for pathogen since they are involved in maintaining the significant metabolic pathways during the pathogenic life cycle. Recent antimicrobial discoveries are mostly dependent upon the inhibition of the metabolic mechanism of the bacterial pathogens and therefore, such protein sequences could be considered as possible therapeutic targets[17,18,19.] With this aspect the identified 129 proteins which were retrieved were further analyzed through subtractive genome approach. Firstly, these proteins were run in CD-hit tool using 60% identity as threshold value which resulted in 122 non-paralog proteins. Among these 77 proteins were sorted out having more than 100 amino acids in length. These 77 proteins were then subjected to BLASTp of unique

essential non-homolog protein sequence against drug bank database. The comparison with DBD resulted in the identification of FDA approved drugs or drug like compounds for which the experimental evidence of binding with proteins similar to the *S. aureus* NCTC 8325 were identified are presented in Table.4.

Table:4 Non-Homologous essential proteins searched within Drug Bank Database consisted of FDA approved drug targets.

S.NO	KEGG ID	Unique essential non host drug target proteins	DBD target ID and name	DBD IDs
1	Sao00336	AcetylCOA acetyltransferase	S-Hydroxy cysteine Coenzyme A S-Acetyl cysteine S-Butyryl cysteine Pantothenyl Amino ethanol acetate pivalic acid Acetoacetyl Coenzyme Pantothenyl aminoethanol-11-pivalic acid (3R)-3-hydroxy-2,2-dimethyl-4-oxo-4-({3-oxo-3-[(2-sulfanylethyl) amino]propyl}amino}butyl2-2-dimethyl proponate	DB01915 DB01992 DB02039 DB02160 DB03045 DB03059 DB08328 DB08408
2	Sao02202	Acetolactate synthase large sub unit	Thiamin diphosphate 2-[2-{2-hydroxy-ethoxy}-ethoxy]-ethanol. 2{(9as)-9a-[(1s)-1-hydroxyethyl]-2,7-Dimethyl-9a,10-Dihydro-5h-pyrimido[4,5-D][1,3]Thiazolo[3,2-a]Pyrimidin-8-41} Ethyl trihydrogen diphosphate	DB01987 DB02327 DB03361
3	Sao01216	SuccinylCOA synthetase subunit beta	Succinic acid	DB00139
4	Sao01218	SuccinylCOA isomerase subunit alpha	Succinic acid	DB00139
5	Sao02287	Isopropyl malate isomerase large subunit	Aconitate ion	DB04351
6	Sao02286	3-Isopropyl malate dehydrogenase	3-Isopropylmalic acid	DB04279
7	Sao00608	Alcohol dehydrogenase	Trifluoroethanol	DB03226

8	Sao00113	Bifunctional acetaldehyde–COA	1,2-Propanediol	DB02059
9	Sao00132	Aldehyde dehydrogenase	NADH	DB00157
10	Sao01400	Alanine racemase	N(5'-phosphopyridoxyl)-D-Alanine Pyridixamine-5'phosphate PmP- Hydroxyisoxazole Pyridoxamine-5-phosphate-hydroxyisooxazole {1-[13-hydroxy-methyl-5-phosphonoxy-methyl-pyridin4-yimethyl}-amino}-ethyl] phosphonic acid Pyridoxyl-N,O-cyclo serylamide-5-monophosphate Propanoic acid Lysine NZ-carboxylic acid Pyridoxyl –Alanine-5-phosphate	DB01993 DB02142 DB03097 DB03327 DB03579 DB03766 DB03801 DB04467
11	Sao02318	D-Alanyl-alanine synthetaseA	3-Chloro-2,2-dimethyl-N-{4-(Trifluoromethyl)Phenyl}Propanamide	DB07805
12	Sao01867	D-Alanine aminotransferase	Pyridoxamine-5'-phosphate	DB02142
13	Sao00869	D-Alanine –poly (phosphoribitol) ligase subunit	Adenosine monophosphate Adenosine triphosphate	DB00131 DB00171
14	Sao00195	AcetylCOA acyl transferase	3,6,9,12,15-Pentaoxatricosan	DB08249
15	Sao01364	Prephenate dehydrogenase	Hydroxy phenyl propionic acid 3-(4-hydroxy-phenyl) Pyruvic acid	DB03897 DB07718
16	Sao03012	Histidinol phosphate aminotransferase	Pyridoxyl –Glutamic acid-5' monophosphate Pyridoxamine-5'-Phosphate Phosphoric acid Mono-[2-amino-3-(3h-imidazol-4-yl)-propyl]Ester	DB01813 DB02142 DB03997
17	Sao02337	UDP-N-acetylglucosamine 1-carboxyvinyl transferase	{S}-2-{Methyl-[2-(Naphthalene-2-sulfonylamino)-5-(Naphthalene-2-sulfonyloxy)-Benzoyl]-Amino}-Succinic acidAminomethylcyclohexane Cyclohexyl ammonium ion L-Iso-Aspartate 3'-1-Carboxy-1-phosphonoxy-Ethoxy-Uridine-Diphosphate-N-Acetylglucosamine 1-Anilino-8-Naphthalene sulfonate	DB01879 DB02435 DB02995 DB03089 DB04174 DB04474

18	Sao01856	UDP-N-acetyl muramate-L-alanine ligase	Uridine -5'-diphosphate-N-acetyl muramoyl-L-alanine Adenosine-5'-[Beta,Gamma-methylene]triphosphate Phosphoamino phosphonic acid-Adenylate ester	DB01673 DB03909 DB04395
19	Sao01147	UDP-N-acetyl muramoyl-L-alanyl-D-glutamate synthetase	Uridine-5'-diphosphate-N-Acetyl muramoyl-L-Alanine Uridine-5'-diphosphate-N-Acetyl muramoyl-L-Alanine-D-Glutamate Lysine NZ carboxylic acid N-{(6-Butoxy Naphthalen-2-YL) Sulfonyl}-L-Glutamic acid N-{(6-Butoxy Naphthalen-2-YL) Sulfonyl}-D-Glutamic acid N-{(6-Pentyloxy Naphthalen-2-YL) Sulfonyl}-D-Glutamic acid N-[{6-[{4-Cyanobenzyl}oxy]Naphthalen-2-YL}sulfonyl]-D-glutamic acid N-[{6-[{4-Cyano-2-fluorobenzyl}oxy]Naphthalen-2-YL}sulfonyl]-D-glutamic acid	DB01673 DB02314 DB03801 DB08105 DB08106 DB08107 DB08108 DB08112
20	Sao00752	UDP-N-acetylenol pyruvoyl glucosamine reductase	Flavin adenine dinucleotide	DB03147
21	Sao02317	UDP-N-acetyl muramoyl alanyl –D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanyl ligase	2-Chloro-N-{3-cyano-5,6-Dihydro-4H-cyclopenta[B]Thiophen-2-Yc}-5-Dimethylsulfamoyl-Benzamide	DB06970
22	Sao01424	Undecaprenyl diphospho-muramoyl pentapeptide beta-N-acetylglucosaminyl transferase	Uridine -Diphosphate –N-Acetylgalactosamine	DB02196
23	Sao01467	Penicillin binding protein 2	Lauryl Dimethylamine-N-oxide	DB04147
24	Sao00646	Penicillin binding protein 4	Cetmetazole Ertapenem Cefpiramide Cefoperazone Cefoxitin	DB00274 DB00303 DB00430 DB01329 DB01331

25	Sao00954	UDP-N-acetyl muramoyl alanyl-D-glutamate-L-lysine ligase	Uridine-5'-diphosphate-N-acetylmuramoyl-L-Alanine-D-Glutamate 2,6 Diaminopimelic acid LysineNZ-carboxylic acid	DB02314 DB03590 DB03801
26	Sao01840	Transglycosylase domain containing protein	Piperacillin Ampicillin Cefalotin Dicloxacillin Cefotaxime Cephalexin Nafcillin Oxacillin Hetacillin Cefadroxil	DB00319 DB00415 DB00456 DB00485 DB00493 DB00567 DB00607 DB00713 DB00739 DB01140
27	Sao02012	Glycosyl transferase	Cefmetazole Ertapenem Cefpiramid Ceftazidine Cefazolin Cefanicid Cefoperazone Cefoxitin Ceftizoxime Cefradine	DB00274 DB00303 DB00430 DB00438 DB01327 DB01328 DB01329 DB01331 DB01332 DB01333
28	Sao01145	Penicillin binding protein 1	Cefprozil Ceftobiprole	DB01150 DB04918
29	Sao01652	Penicillin binding protein 3	Piperacillin Ampicillin Cefalotin Dicloxacillin Cefotaxime Cephalexin Nafcillin Oxacillin Hetacillin Cefditoren	DB00319 DB00415 DB00456 DB00485 DB00493 DB00567 DB00607 DB00713 DB00739 DB01066
30	Sao00247	Choloyl glycine hydrolase	Penicillin V Dithiane Diol Cysteine sulfonic acid	DB00417 DB01822 DB03661
31	Sao02793	Phosphoglucomutase	Alpha-D-Glucose-6-Phosphate Alpha-D-Glucose-1-Phosphate D-Mannose 1 phosphate Alpha-D-Mannose-6-Phosphate Phosphonoserine	DB02007 DB02843 DB02867 DB02900 DB04522

32	Sao01055	Inositol monophosphate family protein	Lithium L-Myo-inositol-1-phosphate	DB01356 DB03542
33	Sao01205	Signal recognition particle docking protein Ftsy	Citric acid	DB04272
34	Sao01207	Signal recognition particle protein	Formic acid Guanosine-5'-diphosphate	DB01942 DB04315
35	Sao02877	Squalene Synthase	Flavin adenine dinucleotide	DB03147
36	Sao02871	Hypothetical protein	1-{Issopropyl thio}-Beta-Galactopyranside Coenzyme A 1-O-[P-Nitrophenyl]-Beta-D-Galactopyranside	DB01862 DB01992 DB02632
37	Sao00142	Formate dehydrogenase	D-Tartaric acid	DB01694
38	Sao02582	Formate dehydrogenase subunit alpha	Selenocyteine	DB02345
39	Sao02354	Serine hydroxyl methyl transferase	5-Formyl-5,6,7,8-Tetrahydrofolate	DB03256
40	Sao00799	Phospho pyruvate hydratase	Glycerol	DB04077
41	Sao02366	Fructose-biphosphate aldolase	Phosphoglycohydroxamic acid	DB03026
42	Sao01807	6-Phosphofructokinase	2-Phosphoglycolic acid Fructose-6-phosphate	DB02726 DB04493
43	Sao00574	Phosphotransacetylase	Acetyl phosphate	DB02897
44	Sao01846	AcetylCOA synthetase	Coenzyme A Adenosine-5'-propylphosphate	DB01992 DB03230
45	Sao00359	Phospho glycerate mutase family protein	Phosphoaminophosphonic acid-Adenylate ester Fructose-6-phosphate	DB04395 DB04493
46	Sao00798	Phosphoglyceromutase	2-Phosphoglyceric acid 3-Phosphoglyceric acid	DB01709 DB04510
47	Sao01833	D-3-Phosphoglycerate dehydrogenase	D-Tartaric acid	DB01694
48	Sao01029	Phosphoenol pyruvate protein phosphotransferase	1-Ethoxy-2-{2-Ethoxy Ethoxy} Ethane	DB08357
49	Sao00220	2-C-methyl-D-erythritol -4-phosphate cytidylyl transferase	Cytidine-5'-triphosphate	DB02431

50	Sao00466	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	4-Diphosphocytidyl-1-2-C-methyl-D-Erythritol Phosphoaminophosphonic acid-adenylate ester	DB03687 DB04395
51	Sao02860	Hydroxy methyl glutaryl-COA synthase	S-Acetyl cysteine Acetoacetyl-coenzyme A (S)-Hmg-COA	DB02039 DB03059 DB03169
52	Sao02859	Hydroxy methyl glutaryl-COA reductase	Coenzyme A (S)-Hmg-COA (R)-Mevalonate (3r,5r)-7-{(1r,2r,6s,8r,8as)-2,6-Dimethyl-8{[(2r)-2-methyl butanoyl]oxy}-1,2,6,7,8,8a- Hexahydronaphthalen-1-Yl}-3,5-Dihydroxy-heptanoic acid	DB01992 DB03169 DB03518 DB03785
53	Sao00577	Mevalonate kinase	Farnesyl Thiopyrophosphate	DB04695
54	Sao02623	Isopentyl pyrophosphate isomerase	Riboflavin monophosphate	DB03247
55	Sao01618	Geranyl transtransferase	Dimethyl allyl-S-thioiodiphosphate Isopentyl pyrophosphate Diphosphate	DB02270 DB02508 DB04160
56	Sao01486	Heptaprenyl diphosphate synthase component II	Isopentyl pyrophosphate	DB02508
57	Sao01237	Undecaprenyl pyrophosphate synthase	Farnesyl thiopyrophosphate Isopentyl pyrophosphate (1-Hydroxy-1-phosphono-2-[1,1',3',1'']Terphenyl-3-Yl-ethyl)-phosphonic acid (1-Hydroxy-1-phosphono-2-[1,1',4',1'']Terphenyl-3-Yl-ethyl)-phosphonic acid [2-{3-Dibenzofuran-4-YL-phenyl}-1-hydroxy-1-phosphono-ethyl]-phosphonic acid [1-hydroxy-2-{1,1',3',1''-Terphenyl-3-Yloxy}Ethane-1,1-Diyl]Bis(phosphonic acid) Farensyl diphosphate	DB04695 DB04714 DB07404 DB07409 DB07410 DB07426 DB07780

58	Sao01799	Histidine kinase	Adenosine-5'-RP-Alpha-Thio-Triphosphate Alpha,Beta-Methylene adenosine-5'-triphosphate 2,3,17Beta-Trihydroxy-1,3,5{10}-Estratriene	DB02355 DB02596 DB07706
59	Sao01800	Alkaline phosphate synthesis transcriptional regulatory protein	Guanosine-5'-monophosphate	DB01972
60	Sao01389	Phosphate ABC transporter substrate binding protein	N-[2-(1-maleimidyl)Ethyl]-7-Diethyl aminocoumarin-3-carboxamide Dihydrogen phosphate ion	DB02799 DB04066
61	Sao02314	Sensor protein Kdp D	Para-coumaric acid	DB04066
62	Sao02315	DNA-binding response regulator	Guanosine-5'-monophosphate	DB01972
63	Sao02311	Potassium-transporting ATPase subunit B	Phosphoaminophosphonic acid-Adenylate ester	DB04395
64	Sao01585	Respiratory response protein Srr B	Adenosine-5'-RP-Alpha-Thio-Triphosphate Alpha,Beta-Methylene adenosine-5'-triphosphate 2,3,17 Beta-Trihydroxy-1,3,5(10)-Estratriene	DB02355 DB02596 DB07706
65	Sao01586	DNA-binding response regulator	Adenosine-5'-RP-Alpha-Thio-Triphosphate Alpha,Beta-Methylene adenosine-5'-triphosphate 2,3,17 Beta-Trihydroxy-1,3,5(10)-Estratriene	DB02355 DB02596 DB07706
66	Sao00021	Sensory box histidine kinase Vick	Adenosine-5'-RP-Alpha-Thio-Triphosphate Alpha,Beta-Methylene adenosine-5'-triphosphate 2,3,17 Beta-Trihydroxy-1,3,5(10)-Estratriene	DB02355 DB02596 DB07706
67	Sao00020	Two component response regulator	Guanosine-5'-monophosphate	DB01972
68	Sao00714	Sensor histidine kinase Sae S	Adenosine-5'-RP-Alpha-Thio-Triphosphate Alpha,Beta-Methylene adenosine-5'-triphosphate 2,3,17 Beta-Trihydroxy-1,3,5(10)-Estratriene	DB02355 DB02596 DB07706

69	Sao00715	Response regulator	Guanosine-5'-monophosphate	DB01972
70	Sao00666	Hypothetical protein	Adenosine-5'-RP-Alpha-Thio-Triphosphate Alpha,Beta-Methylene adenosine-5'-triphosphate 2,3,17 Beta-Trihydroxy-1,3,5(10)-Estratriene	DB02355 DB02596 DB07706
71	Sao00665	Hypothetical protein	Guanosine-5'-monophosphate	DB01972
72	Sao03036	ABC transporter ATP binding protein	Adenosine triphosphate	DB00171
73	Sao01810	NAD-dependent malic enzyme	NADH Tartronate	DB00157 DB03680
74	Sao02098	DNA-binding response regulator Vra R	Benzoic acid	DB03793
75	Sao02675	Hypothetical protein	Adenosine-5'-RP-Alpha-Thio-Triphosphate Alpha,Beta-Methylene adenosine-5'-triphosphate 2,3,17 Beta-Trihydroxy-1,3,5(10)-Estratriene	DB02355 DB02596 DB07706
76	Sao01287	Glutamine synthetase	2-Amino-4-(Hydroxymethyl-phosphinyl)Butanoic acid	DB02663
77	Sao00184	Response regulator receiver domain containing protein	Phosphoaspartate	DB01857

These 77 protein drug targets listed in Table.4 has to be investigated further for its potentiality. To establish that it can indeed be a drug target, more insight into the activity and the essential nature of the target in the viability of the pathogen in the host should be gathered from literature where possible.

CONCLUSION

In this study, we have performed a comparative metabolic pathway analysis of the host *H. sapiens* and the pathogen *M. tuberculosis*. Though sequence similarity greater than 25% implies homology, we have adopted a stringent measure of listing out only those enzymes, which have no similarity (or negligible similarity above the e-value threshold of 0.005) to the host proteins as potential targets. The approach was successful in listing out many targets from the *M. tuberculosis* proteome, which are involved in vital aspects of the pathogen's metabolism, persistence, virulence and cell wall biosynthesis.With an objective of identifying proteins potentially useful as drug targets, we have relied on the use of proteomes and a subtractive genomic approach for the identification of putative drug targets. A subtractive genomics approach utilizes the whole proteome of host and pathogen to identify proteins exclusively present in the pathogen by deducing the homologous proteins. Therefore from the present *in silico* study with the use of subtractive genomic approach we have identified 22 unique metabolic pathways and 129 unique essential proteins as drug targets Drug bank database. Unique pathways are usually ideal for drug targets to avert toxicity and cross reactions as they are not present in the host. This approach is cost effective approach for screening of drug and vaccine targets for any given pathogen and thereby reducing the laborious, capital intensive and rigorous process of screening a whole organism for drug targets on whole genome scale in the laboratory.

REFERENCES

[1] Bishop EJ and Howden BP. Treatment of *staphylococcus aureus* infections: new issues, 5emerging therapies and future direction. *Expert Opinion on Emerging Drugs*, 2007; 12(1): 1- 22.

[2] Gillapsy A.F, Worrell V, Orvis J, Roe B.A, Dyer D.W, Landolo J.J, The *Staphylococcus aureus* NCTC 8325 .*Genome*, 2006; 381- 412.

[3] Uddin R, Saeed K, Khan W .Metabolic Pathway Analysis Approach: Identification of Novel Therapeutic Target against Methicillin Resistant *Staphylococcus aureus*, Gene, 2015; 48:58-63.

[4] Johnsen PJ, Townsend JP, Bohn T, Simonsen GS, Sundsfjord A, Nielsen KM. Factor affecting the reversal of antimicrobial drug resistance. *Lancet Infectious Diseases,* 2009; 9(6): 357-364.

[5] Michael, Y.G., Eugene, V.K. Searching for drug targets in microbial genomes. Current *Opinion on Biotechnology.*1999; 10 (6) :571–578.

[6] Rajendra H.M and Amol S.W. Subtractive Genomics Approach to identify Potential Therapeutic Targets in *Leishmania Donovani*. *International Journal of Research in Pharmaceutical and Biomedical Sciences*, 2010; 1: 1-6.

[7] Rahman Q Sarangi AN, Aggarwal R, Trivedi N. Subtractive genomics approach for in silico identification and characterization of novel drug targets in *Neisseria meningitides* serogroup B. *Journal of Computer Science and Systems Biology*, 2009; 2: 255- 258.

[8] Munikumar M, Priyadarshini V, Pradhan V, Amine. *In Silico* Identification of common Putative Drug Targets among the Pathogens of Bacterial Meningitis. *Biochemistry Analytical Biochemistry,* 2012; 1:8.

[9] Sakharkar K. R., Sakharkar M. K. and Chow V. T. K. A novel genomics approach for the identification of drug targets in pathogens, with special reference to *Pseudomonas aeruginosa*. *In Silico Biology*,2004; 4: 355-60.

[10] Sharma V, Gupta P, Dixit A. *In silico* identification of putative drug targets from different metabolic pathways of *Aeromonas hydrophila.In silico Biology*, 2008; 8: 331- 338.

[11] Vetrivel U, Subramanian G, Dorairaj S. A Novel *in silico* approach to identify potential therapeutic target in human bacterial pathogens. *Human Genome Organization*,2011; 5: 25- 34

[12] Rathi B, Aditya N, Sarangi AN, Trivedi N. Genome subtraction for novel target definition in *Salmonella typhi*. *Bioinformation,* 2009; 4: 143-150.

[13] Dutta A, Singh SK, Ghosh P, Mukherjee R, Mitter S, Bandyopadhyay D. *In silico* identification of potential therapeutic targets in the human pathogen *Helicobacter pylori*. *In Silico Biology*, 2006; 6: 43- 47

[14] Ononamadu C.J, Umeoguaju U.F, Owolarafe T.J, Udedi S.c, Barau *et.,al. In Silico* Identification of Putative drug targets in Methicillin resistant *Staphylococcus aureus*: a Subtractive Genomic Approach. *International Journal of Computational Bioinformatics and In Silico Modeling*, 2015; 4: 585-591.

[15] Haag N.L, Velk K.K, Wu C.*In silico* identification of drug targets in methicillin/multidrug-resistant *Staphylococcus aureus*. *International Journal of Advanced Life Sciences*,2011; 4(2): 21- 32.

[16] Butt A.M, Nasrullah I, Tahir S, Tong Y.Comparative genomics analysis of *Mycobacterium ulcerans* for the identification of putative essential genes and therapeutic candidates. *PloS One,* 2012; 7: 30-80.

[17] Lemaitre B, Girardin S. Translation inhibition and metabolic stress pathways in the host response to bacterial pathogens. Nature Review Microbiology, 2013; 11:365–369.

[18] Barh D and Misra A. *In silico* identification of membrane associated candidate drug targets in *Neisseria gonorrhoeae*. *International Journal of Integrative Biology;*2009; 6(2): 65- 67.

[19] Anishetty S, Pulimi M, Pennathur G. Potential drug targets in *Mycobacterium tuberculosis* through metabolic pathway analysis. *Computational Biology and Chemistry*,2005;29(5): 368- 378.