ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 8, Issue, 09, pp.37617-37634, September, 2016

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

RESEARCH ARTICLE

IN-SILICO IDENTIFICATION OF PUTATIVE DRUG TARGETS OF *CHLAMYDIA PSITTACI* 6BC BY SUBTRACTIVE GENOME APPROACH

***Khalida Naveed, Ahsan Jabbar and Syeda Tatheer Fatima**

Baqai Institute of Information Technology, Baqai Medical University, Karachi, Pakistan

ARTICLE INFO

Article History:
Received 20th June, 2016
Received in revised form
02nd July, 2016
Accepted 11th August, 2016
Published online 20th September, 2016

Key words:

Chlamydia psittaci, Subtractive genome, Drugability, Structural homology modeling, Molecular docking.

ABSTRACT

Chlamydia psittaci is a pathogenic and bio-warfare agent that causes infections in multiple hosts including birds and mammals. In human, it mostly causes severe lung diseases that are associated with high motility rate. This pathogen also causes a major setback to the world economy. Therefore, there is an urgent need to develop antimicrobial agent that can cure infections caused by these agents. In the present study, we have predicted number the possible putative drug targets against it with the aid of subtractive genome approach along with their subcellular localization, drugable potential and involvement in essential metabolic pathway. Furthermore, we have also predicted the 3D molecular structure using structural homology methodology with of one of the identified drug target that affect the isoprenoid biosynthesis in C. psittaci (a crucial pathway needed by the pathogen to survive). After molecular modeling, we also performed molecular docking to predict the possible binding mechanism of the selected protein with its ligand. This study can be utilized to identify potent inhibitor of 4-hydroxy-3-methylbut-2-enyl diphosphate reductase protein by virtual screening of number of chemical entities against the modeled protein.

Citation: Khalida Naveed, Ahsan Jabbar and Syeda Tatheer Fatima, 2016. "*In-silico* identification of putative drug targets of *Chlamydia psittaci* 6bc by subtractive genome approach", *International Journal of Current Research*, 8, (09), 37617-37634.

INTRODUCTION

The *chlamydiaceae*is a family of bacteria that comprises of nine different anddistinctive species, namely *C. trachomatis*, *C. suis*, *C. muridarum*, *C. psittaci*, *C. pneumoniae*, *C. abortus*, *C. felis, C. pecorum* and *C. caviae*. Among them, *Chlamydophilapsittaci* is a pathogenic, zoonotic, gram negative, obligate intracellular bacteria that mostlycause infections in parrot (Psittaciforme) but can also infect several other birds' species and a wide variety of mammalian organisms(Hotzel *et al.*, 2004; Kaleta & Taday, 2003; Read *et al.*, 2013). *C.psittaci* is a principalrisk for the poultry, farmhouse and research industries linked with its great setback to the world economy and specifically United States (Gaede *et al.*, 2008; Miller *et al.*, 1987; Smith *et al.*, 2011). It causes diverse range of infections including respiratory disorders, arthritis, enteritis and abortion in their primary host and can transmit these infections to mammals including humans (Branley *et al.*, 2008; Harkinezhad *et al.*, 2009; Hughes *et al.*, 1997; Lee *et al.*, 2014; Smith *et al.*, 2005). In humans, it mostly associated with severe pulmonary disorder that can often involve multiple organs and has notably high morbidity and death toll(Smith *et al.*, 2011). According the Center for Disease Control and Prevention, *C. psittaci* is categorized in B group of bio warfare agent due to its potential pathogensis in human (Meselson *et al.*, 2002) The exact pathogenesis of *C. psittaci* is not well explored. But it is well known that it has a unique developmental process can be divided into bi-episodic cycles i.e., extracellular elementary bodies (EBs), and intracellular reticula bodies (RBs). EBs is a metabolically sedentary infectious phase while RBs is non-infectious phase(AbdelRahman & Belland, 2005). Genetic engineering methodologies cannot authentically reliable over *chlamydiaceae* family of bacteria, due to their intracellular mode of replication. Consequently, investigation of the precise relations and assistances of genes in the involvement pathogenesis, virulence and replication are limited. The juxtapositions among the numerous genotypes and between species of *chlamydiaceae* are restricted regardless of some current developments in the characterization of species and the usage of surrogate organisms (Peters *et al.*, 2007). However, some members of the chlamydiaceae family have been sequenced and studied for their host and tissue preferences and virulence mechanisms. Till to date, only 16 have been completely characterized (Read *et al.*, 2000; Read *et al.*, 2013;

Corresponding author: Khalida Naveed,
Baqai Institute of Information Technology, Baqai Medical University, Karachi, Pakistan.

Read *et al.*, 2003; Stephens *et al.*, 1998; Thomson *et al.*, 2005; Van Lent *et al.*, 2012; Voigt *et al.*, 2012). The recent advancement in sequencing technologies has augmented the rate of deciphering the genetically and functional annotation existing in the genomes of pathogenic microbes. The information acquired *via* sequencing projects aided the characterization of gene that encrypt for virulence factors. The genome of disease causing bacteria mostly consists of hypothetical genes whose functions are not known. Therefore, this class of genes are of substantial attraction for medicinal research as they can lead to the development of novel and new putative targets that can be used for the expansion of therapeutics. Using computational and bioinformatics based tools; these hypothetical genes can be identified and targeted for the development of putative potential drugs and vaccines. Among these, comparative and subtractive genomics approaches are commonly applied tactics in the current era of drug discovery (Shahbaaz *et al.*, 2016).

In the present study, we have applied a newer approach of subtractive genomics in order to isolate crucial and vital genes of the selected pathogen, these are non- homologous to the human genome and can aid in the process of drug development (Shoukat *et al.*, 2012).Numerous literature supports successful drug targets prediction with the application of subtractive genome approach (Abadio *et al.*, 2011; Allsop *et al.*, 1995; Anishetty *et al.*, 2005; Sakharkar *et al.*, 2004). After identifying the putative targets, we have selected 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, an iron protein as a drug target for three dimensional structure determination using Swiss model(Biasini *et al.*, 2014).This enzyme is a part of isoprenoid biosynthesis of *C.psittaci*, a crucial biosynthetic pathway for many pathogenic bacteria including *C. psittaci*(Heuston *et al.*, 2012). The ligand interaction was studied through Dock 6. This would give a better understanding of 4-hydroxy-3-methylbut-2-enyl diphosphate reductase as drug target and can therefore eventually targeted for the development of putative drug against the *C. psittaci.*

MATERIALS AND METHODS

The 2.2.26 version of standalone BLAST+ from NCBI (Altschul *et al.*, 1990) (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/) was downloaded and installed over Linux workstation with Intel Xeon quad core processor. The flow chat of current scheme is showed in Fig.1.

Retrieval of complete proteome

The UniProtKB (Universal Protein Resource) provides a stable, complete, easily available, fundamental resource on protein sequences and functional annotation and can be accessed from http://www.uniprot.org. This database is friendly to user and comprised of manually curated protein sequences aided with computational study, sequence archival and retrieval (Consortium, 2008).The NCBI (National Center for Biotechnology Information) is a collection of large amount of online resources for biological information and data, majorly comprised of data bases of the nucleic acid sequence, PubMed references and abstracts for published natural sciences journals. This database also provides Identical Protein Report that exhibits the accessions of entire other protein sequence archives that are identical to a particular protein complimented with their relative CDS (coding DNA sequence) in Nucleotide for the respective protein(NCBI, 2015). All the resources of this database can be opened through http://www.ncbi.nlm.nih.gov.

Identification of non-homologous proteins against human proteome

The entire proteome of *C. psittaci* and *H. sapiens* were retrieved from UniProtKB and NCBI respectively on October, 20 2015. BLASTp of selected organism was run against human proteins having E-value of 10^{-3}, in order to find out the non-homologous sequences to human proteome (Haag *et al.*, 2011; Kerfeld & Scott, 2011). The non-homologous sequences having no hits against the human host were considered for the further analysis while the homologous sequences were excluded from the investigation

Determination of non-homologous essential genes in *C. psittaci*

The genes that are crucial for the survival of the micro-organism are termed as essential genes as they are vital for their cellular activities. The DEG (Database of essential gene) version 6.8 was downloaded from the DEG website (http://www.essentialgene.org/). This database consists of essential genes and their proteins(Zhang & Lin, 2009; Zhang *et al.*, 2004). BLASTp (expectation value of 10^{-5}) of the obtained non-homologous proteins were subjected against DEG database

Drugability prediction of non-homologous hypothetical proteins

In order to evaluate the drugable nature of the identified proteins, the BLASTp of these proteins with in-built parameters having expectation value of 10^{-3}, was performed against the Drug Bank database (Knox *et al.*, 2011). The database has proteins targets in reference to drug IDs approved by U S Food and Drug Administration authority (FDA).

Evaluation of Subcellular localization

In order to find out the location of identified non-homologous essential proteins, the PSORTb version 3.0 was used. This computational tool is renowned bioinformatics tool used to assess the locality of unknown proteins(Nancy *et al.*, 2010). The Sub Cellular Localization BLAST (SCL BLAST) is used by PSORTb, which consecutively run BLASTp of non-homologous essential proteins with the proteins having defined subcellular location. It predicts ranges of protein sites including cytosol, plasma membrane, and cell wall, extracellular and even anonymous too.

Classification of non-homologous hypothetical proteins into functional families

Support Vector Machine of Proteins (SVM-PROT) is a web based server is utilized to classify selected non-homologous

hypothetical proteins into the protein families, according to their function(Cai *et al.*, 2003). This server guess the functional classes from the primary structure of protein into classes like enzymes, receptors, transporters, channels, DNA-binding proteins and RNA-binding proteins etc.

Analyses of KEGG metabolic pathway

KEGG (Kyoto Encyclopedia of Genes and Genomes) database that offers the information for the understanding of extremely multifaceted functions exists in a living organism (Kanehisa *et al.*, 2011). The most common server KAAS (KEGG Automated Annotation Server) was used to predict the metabolic pathway. This server execute a BLASTp similarity quest of identified non-homologous essential proteins against the updated KEGG database(Moriya *et al.*, 2007). Along with the metabolic pathways, the KAAS outcomes also have other supplementary information like alternative pathway, Enzyme Commission (EC) numbers and KO list assignment. Theses predicted metabolic pathways can serve as the potential therapeutic targets against the evaluated proteins.

Homology modeling and model evaluation

The three dimensional structure of protein can also be determined by structure homology with the proteins having known 3D structure using various bioinformatics based tools. The template protein for homology modeling was retrieved through BLASTp of selected 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (WP_006343407.1) was performed against updated Protein Databank (PDB)(Bernstein *et al.*, 1978). Then, the 3D structure of the selected protein was predicted with a web based tool Swiss Model (Arnold *et al.*, 2006). After modeling the structure, the quality of the modeled protein was assessed by Ramachandran plot *via* online software PROCHECK (Thapa *et al.*, 2013).

Molecular docking

As modeled protein has not any ligand, we use the ligand of the template protein as the ligand for our targeted protein. The molecular docking was executed between the modeled structure of 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (WP_006343407.1) enzyme as a receptor and ligand of template protein with the aid of DOCK 6 program. This was achieved by superimposing the ligand on to a negative image of active site of modeled protein (Dolinsky *et al.*, 2007; Fu *et al.*, 2014; Lang *et al.*, 2009).

RESULTS AND DISCUSSION

The aim of the present study was to discover new and potent drug targets against *Chlymadia psittaci* that can be used as potential therapeutics.

We had implied the subtractive genomic approach for the determination of unique putative drug targets against *C.psittaci* (Amineni *et al.*, 2010; Dutta *et al.*, 2006; Georrge & Umrania, 2011; Reddy *et al.*, 2010; Sharma *et al.*, 2008; Shoukat *et al.*, 2012). The work flow of the present study is presented in Fig. 1 and Table 1 showed the precised results of each phase.

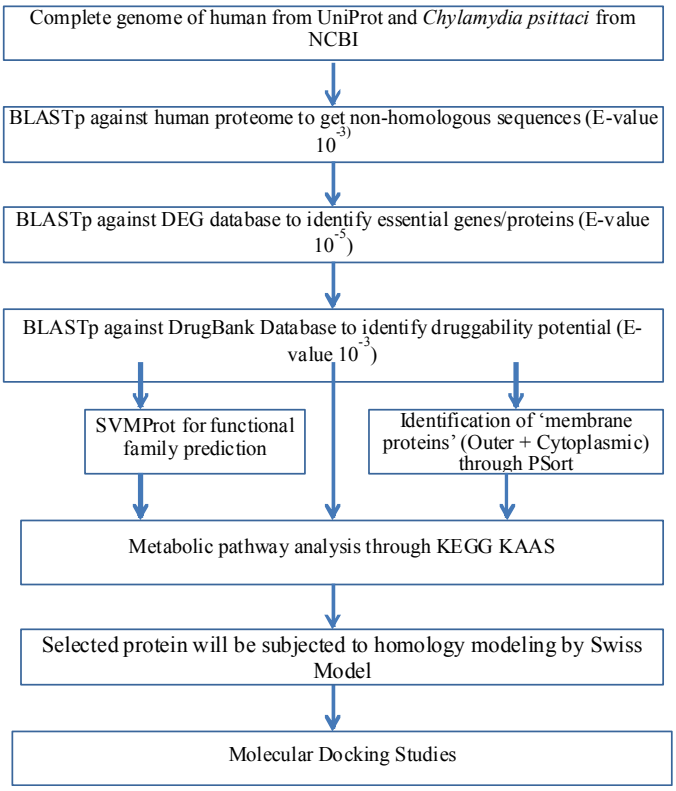

Figure 1. Scheme of workflow

Table 1. Stepwise summary of results

S. No.	Steps	6BC
1	Total numbers of proteins	979
2	Number of proteins against *H. sapiens* using BLASTp (E-value 10^{-3})	660
3	Essential proteins in DEG (E-value 10^{-5})	255
4	Essential drug targets proteins (E-value 10^{-3})	76
5	Number of essential cell membrane proteins (PSORT)	76
6	Number of non-homologous essential proteins (SVM)	76
7	Essential proteins involved in metabolic pathway (KEEG)	62
8	Selected protein will be subjected to homology modeling by Swiss-model	1
9	Molecular docking using DOCK6	1

Identification of non-homologous essential proteins against human proteome

The whole proteome of *C.psittaci* strain 6BC consisting of 979 number of proteins was repossessed from NCBI (National Center for Biotechnology Information) available at http://www.ncbi.nlm.nih.gov. These proteins were subjected to blastp (E-value 10^{-3})against human host proteome (retrieved from UniProtKB), in order to eliminate the common in both organisms (Gasteiger *et al.*, 2003). Out of 979 proteins, *C.psittaci* have 660 non-homologous proteins. Moreover, we determined whether these non-homologous proteins are essential for the survival and pathogenesis of the *C.psittaci* with the aid of BLASTp having cut off E-value $10^{-5}$of against the Essential Gene (DEG) database (Zhang & Lin, 2009; Zhang *et al.*, 2004). This step has further shortlisted the proteins up to 255 in numbers, that could be explore as the drug targets in order to cure *C.psittaci* infections if their drugability potential is well understood.

Table 2. List of identified potential drug targets

S. No.	Gene ID	Protein Name and ID	Drugbank ID
1	WP_006342686.1	P80373 30S ribosomal protein S4	(DB08185)
2	WP_006342696.1	O66496 2-dehydro-3-deoxyphosphooctonate aldolase	(DB01709; DB01819; B02053; DB02433; DB02992; DB03248; DB03745; DB03937)
3	WP_006342724.1	P44469 Penicillin-binding protein 2	(DB00303; DB00671)
4	WP_006342755.1	P44868 tRNA (cytidine(34)-2'-O)-methyltransferase	(DB01752)
5	WP_006342760.1	P03007 DNA polymerase III subunit epsilon	(DB01643)
6	WP_006342763.1	P47205 UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine Deacetylase	(DB07861)
7	WP_006342764.1	O25928 3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	(DB07445)
8	WP_006342765.1	O25927 Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	(DB01694; DB08558)
9	WP_006342777.1	Q5SHP7 30S ribosomal protein S17	(DB08185)
10	WP_006342805.1	Q96MA6 Adenylate kinase 8	(DB01717)
11	WP_006342820.1	P33590 Nickel-binding periplasmic protein	(DB03374)
12	WP_006342832.1	P41789 Nitrogen regulation protein NR(I)	(DB01857)
13	WP_006342846.1	P33038 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	(DB01879; DB02435; B02995; DB03089; DB04174; DB04474)
14	WP_006342849.1	P0A6I0 Cytidylate kinase	(DB02456; DB02883; B03403; DB04555)
15	WP_006342881.1	P39594 Thiamine-phosphate synthase	(DB01788; DB02254; B02885; DB03145; DB03416; DB07782)
		P25053 Regulatory protein TenI	(DB03570)
16	WP_006342891.1	P0AFU8 Riboflavin synthase	(DB00140)
17	WP_006343008.1	P08877 Phosphocarrier protein HPr	(DB01899)
		P07515 Phosphocarrier protein HPr	(DB04522)
18	WP_006343009.1	Q6FEW8 Phosphoenolpyruvate-protein phosphotransferase	(DB08357)
		P22983 Pyruvate, phosphate dikinase	(DB02522)
19	WP_006343049.1	P22188 UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate Ligase	(DB02314; DB03590; B03801)
20	WP_006343057.1	P03007 DNA polymerase III subunit epsilon	(DB01643)
21	WP_006343078.1	P28248 Deoxycytidine triphosphate deaminase	(DB02333; DB03258)
22	WP_006343084.1	P0AGE0 Single-stranded DNA-binding protein	(DB04243)
23	WP_006343107.1	Q5SKW1 RNA polymerase sigma factor	(DB08226; DB08266)
24	WP_006343119.1	P45568 1-deoxy-D-xylulose 5-phosphate reductoisomerase	(DB02496; DB02948; B03649; DB04272)
25	WP_006343123.1	P0A988 DNA polymerase III subunit beta	(DB06998)
26	WP_006343147.1	P96618 Holo-[acyl-carrier-protein] synthase	(DB01992; DB04447)
27	WP_006343157.1	O25927 Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	(DB01694; DB08558)
28	WP_006343161.1	P0A6R0 3-oxoacyl-[acyl-carrier-protein] synthase 3	(DB01034; DB01992; B02039; DB02316; DB03661; DB04524)
29	WP_006343185.1	P0A721 Thymidylate kinase	(DB03280)
30	WP_006343186.1	P06710 DNA polymerase III subunit tau	(DB02930)
31	WP_006343194.1	P69772 Probable aromatic acid decarboxylase	(DB03247)
32	WP_006343261.1	Q9KPI8 5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	(DB07463)
33	WP_006343263.1	P42216 3-deoxy-manno-octulosonate cytidylyltransferase	(DB02344; DB02431; DB03403; DB04482; DB04555)
34	WP_006343266.1	P06202 Periplasmic oligopeptide-binding protein	(DB07365)
35	WP_006343268.1	P06202 Periplasmic oligopeptide-binding protein	(DB07365)
36	WP_006343275.1	P00512 6-phosphofructokinase	(DB02726; DB04493)
37	WP_006343277.1	P00512 6-phosphofructokinase	(DB02726; DB04493)
38	WP_006343322.1	P43912 tRNA (guanine-N(1)-)-methyltransferase	(DB01752)
39	WP_006343376.1	P04036 4-hydroxy-tetrahydrodipicolinate reductase	(DB03969; DB04267)

Continue……

40	WP_006343381.1	P12996 Biotin synthase	(DB03754; DB03775)
41	WP_006343383.1	P13000 ATP-dependent dethiobiotin synthetase BioD 1	(DB01715; DB02927; B02941; DB03624; DB03775)
42	WP_006343386.1	P0A6D3 3-phosphoshikimate 1-carboxyvinyltransferase	(DB01942; DB03116; B04328; DB04539)
43	WP_006343388.1	P56122 Chorismate synthase	(DB03247)
44	WP_006343389.1	Q6GGU4 3-dehydroquinate synthase	(DB02592)
45	WP_006343390.1	P15770 Shikimate dehydrogenase	(DB03461; DB04447)
46	WP_006343407.1	P62623 4-hydroxy-3-methylbut-2-enyl diphosphate reductase	(DB01785; DB04714)
47	WP_006343436.1	Q9X286 N utilization substance protein B homolog	(DB04272)
48	WP_006343437.1	P61432 UDP-N-acetylenolpyruvoylglucosamine reductase	(DB03147)
49	WP_006343469.1	P49228 50S ribosomal protein L32	(DB01361)
50	WP_006343476.1	Q83LD8 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	(DB03687; DB04395)
51	WP_006343479.1	Q5SLP8 30S ribosomal protein S6	(DB08185)
52	WP_006343506.1	Q59771 L-phenylalanine dehydrogenase	(DB02494; DB03884)
53	WP_006343524.1	P14900 UDP-N-acetylmuramoylalanine--D-glutamate ligase	(DB01673; DB02314; B03801; DB08105; DB08106; DB08107; DB08108; DB08112)
54	WP_006343526.1	Q8DNV6 UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	(DB06970)
55	WP_006343531.1	P0AE12 AMP nucleosidase	(DB03464)
56	WP_006343542.1	P07771 Benzoate 1,2-dioxygenase electron transfer component	(DB03147)
57	WP_006343552.1	O66529 6,7-dimethyl-8-ribityllumazine synthase	(DB02214; DB04128; DB04262)
58	WP_006343559.1	P06709 Bifunctional protein BirA	(DB04651)
59	WP_006343561.1	P45124 Ribosomal small subunit pseudouridine synthase A	(DB01955)
60	WP_006343606.1	P37093 Type II secretion system protein E	(DB04395)
61	WP_006343637.1	P0A910 Outer membrane protein A	(DB04233)
62	WP_006343654.1	P75430 Probable 5-formyltetrahydrofolate cyclo-ligase	(DB02800)
63	WP_006343659.1	Q81VW8 Dihydropteroate synthase	(DB03592; DB03705; DB04047; DB04196)
64	WP_006343660.1	P56740 Dihydroneopterin aldolase	(DB01778; DB01906; B02119; DB02489; DB03231; DB03571; DB04168; DB04400; DB04425; DB06906)
65	WP_006343661.1	Q18BX5 RNA polymerase sigma factor	(DB08874)
66	WP_006343668.1	P0ACC7 Bifunctional protein GlmU	(DB01992; DB03397; B03814)
67	WP_006343678.1	P03692 DNA primase/helicase	(DB02452; DB03222)
68	WP_013462583.1	P08506 D-alanyl-D-alanine carboxypeptidase DacC	(DB00274; DB00303; B00430; DB01329; DB01331)
69	WP_013462587.1	P61177 50S ribosomal protein L22	(DB00199; DB00207; B01369)
70	WP_013462588.1	P0A7Z4 DNA-directed RNA polymerase subunit alpha	(DB00615)
71	WP_013462596.1	Q9X180 Multi-sensor signal transduction histidine kinase	(DB02731)
72	WP_013462606.1	Q8EBR3 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	(DB07780)
73	WP_013462639.1	Q51504 Penicillin-binding protein 3	(DB01147; DB01413; B09050)
74	WP_013462674.1	P06202 Periplasmic oligopeptide-binding protein	(DB07365)
75	WP_013462741.1	P17443 UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	(DB02196)
76	WP_028444368.1	Q9WYH8 Phospho-2-dehydro-3-deoxyheptonate aldolase	(DB01819; DB03937)

Druggability nature of the shortlisted non- homologous essential proteins

This study was additionally augmented *via* eliminating out the non-homologous essential gene on the basis of their drugability nature. This was achieved by screening these proteins on the basis of BLASTp (10^{-3}) with the available FDA approved drug targets deposited in Drug bank database. This lead to the identification of 76 proteins that are non-homologous, crucial for the survival of *C. psittaci* and have drugable potential. The summarized results of these drugable protein with their drug bank ID is presented in Table 2, while the detailed results can be viewed in the supplementary table (S1).

Sub-cellular localization of drugable proteins

It is necessary to find out the sub-cellular compartment of the identified drugable proteins where theses drug targets exist in order to design the action of an appropriate selective drug in their specific sub-cellular site. The PSORT analysis of non-homologous drugable proteins was performed to locate their sub-cellular compartments (Nancy *et al.*, 2010). The results showed that most of the drugable proteins were found in the cytoplasm (i.e., 75%) and few in cytoplasmic membrane. The remaining 12% of the drugable proteins were those whom locations were not predicted by PSORT. The graphical presentation of PSORT results are shown in Fig.2.

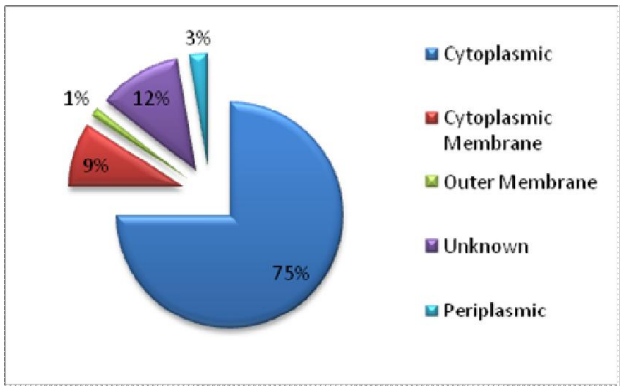

Figure 2. Subcellular localization of drug targets

Classification of non-homologous hypothetical proteins into functional families

Functional classification of the drugable proteins was performed by SVMprot method. The precision of SVMprot results is indicated in term of percent P-value. We have predicted the functional classes of 76 hypothetic drugable proteins and their results are shown in Fig. 3. Most of the drugable proteins are found to be associated with enzyme transferases, lipid binding, Zinc binding, transmembrane and DNA binding functional classes and can be targeted as a potential drug targets. The entire lists of functional classes presented in supplementary data (Tables S2).

Analyses of KEGG metabolic pathway

The drugable proteins were subjected to KEGG Automated Annotation Server (KAAS), in order to find out the protein involved inessential metabolic pathways. These results are shown in Fig. 4 and their detailed results were presented in the supplementary information (Table S3). Concisely, a total of 76 identified drugable proteins was subjected to KEGG and the metabolic pathways of 62 proteins were predicted through the database. Among the predicted drugable proteins, 22% were involved in amino acid metabolism, 18% in nucleotide, 18% in cofactor and vitamins and 13% glycan biosynthesis and metabolism etc.

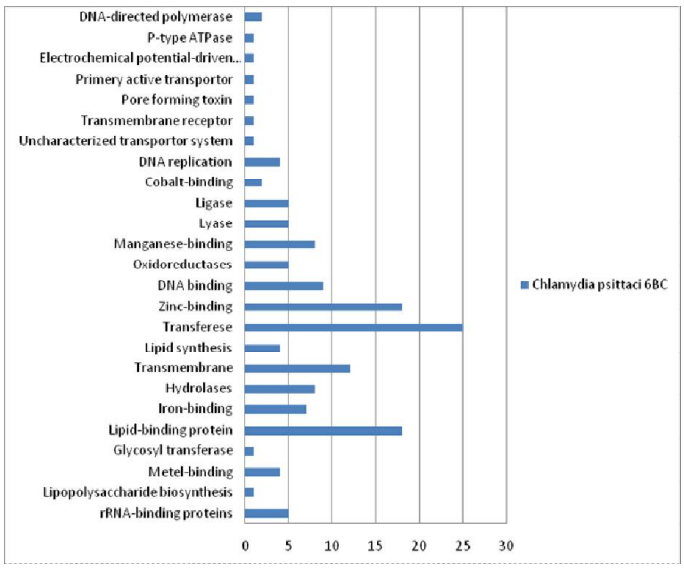

Figure 3. Predicted functional classes of hypothetical drugable proteins

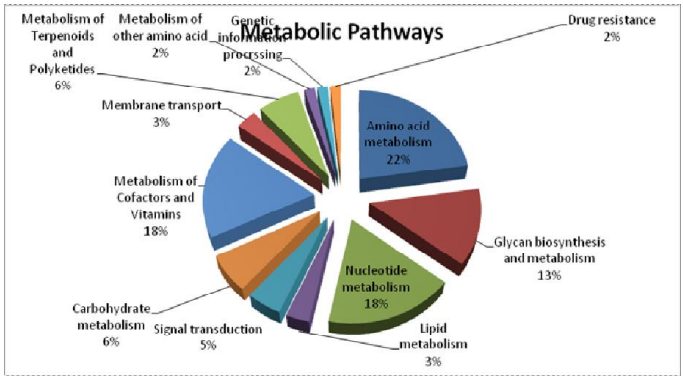

Figure 4. Categorization of drugable proteins according to different metabolic pathways with the aid of KASS

Homology modeling of 4-hydroxy-3-methylbut-2-enyl diphosphate reductase

The information regarding drug design, function and active residue of a protein is embedded in its three dimensional structure, which is experimentally determined by X-ray crystallography or NMR spectroscopy. As these methodologies are time consuming and expensive as compared to bioinformatics based computational methods(Jaroszewski, 2009; Kopp & Schwede, 2004). The prediction of three dimensional structure (Fig 5a) was done by Swiss model using chain A of E-1-hydroxy-2-methyl-but-2-enyl-4-diphosphate reductase from *Plasmodium falciparum* (PDB: 4N7B) as a

template with 52% sequence identity with the query sequence (WP_006343407.1). The quality of the predicted model in term of stereo-chemical properties was assessed by Ramachandran plot from PROCHECK server showed 92.7% of most favorable region, 6.5% in additional allowed region, 0.4% in generously and disallowed region respectively (Fig 5b).

Molecular docking of4-hydroxy-3-methylbut-2-enyl diphosphate reductase

The ligand of the template protein (PDB: 4N7B) was docked with the predicted modeled enzyme (WP_006343407.1) with the aid of DOCK 6 program. We have predicted the major and predominant binding mode of the modeled protein with the ligand shown in Fig 6 along with the scoring function of docking shown in Table 3.

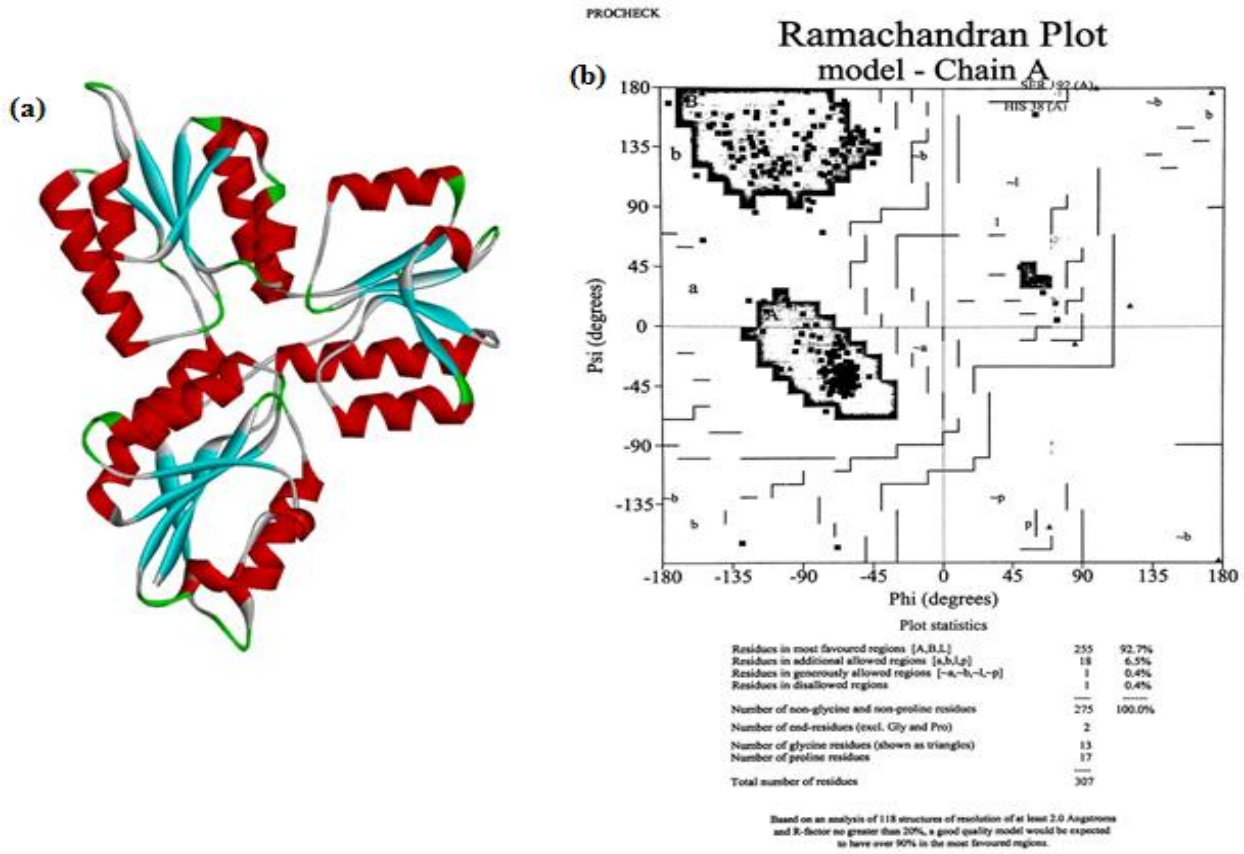

Figure 5. (a-b): (a) Predicted 3D structure of modeled protein (WP_006343407.1) using Swiss model; (b) Ramachandran plot of modeled protein (WP_006343407.1) using PROCHECK showing amino acid residue in different regions

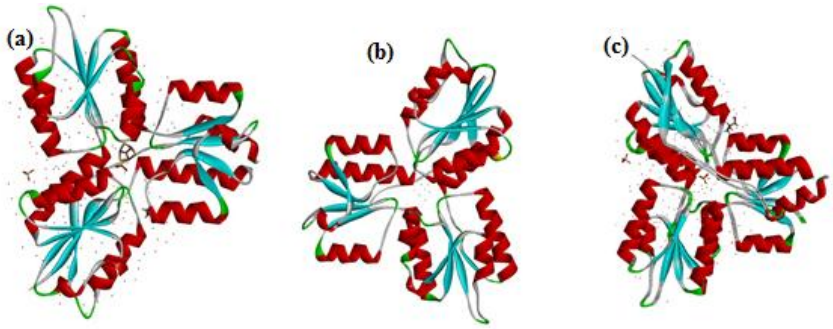

Figure 6(a-c): (a)Molecular docking of template protein (PDB: 4N7B) with its ligand; (b)Molecular docking of modeled protein (WP_006343407.1) with the template ligand; (c)Superimposed image of both template and modeled protein with the ligand

Table 3. Scoring functions with score of docked modeled protein

Scoring functions	Scores
Grid Score	-23.029579
Grid_vdw	8.395700
Grid_es	-31.425278
Internal energy	0.000000

Supplementary tables:

Table S4: List of drugable protein

S. No.	Gene ID	Protein Name and ID	Drugbank ID
1	WP_006342686.1	P80373 30S ribosomal protein S4	(DB08185)
		P0A7V8 30S ribosomal protein S4	(DB00254; DB00256; DB00453; DB00595; DB00618; DB01017)
2	WP_006342696.1	O66496 2-dehydro-3-deoxyphosphooctonate aldolase	(DB01709; DB01819; DB02053; DB02433; DB02992; DB03248; DB03745; DB03937)
		P0A716 2-dehydro-3-deoxyphosphooctonate aldolase	(DB01819; DB02433; DB03113; DB03936)
		Q9WYH8 Phospho-2-dehydro-3-deoxyheptonate aldolase	(DB01819; DB03937)
3	WP_006342724.1	P44469 Penicillin-binding protein 2	(DB00303; DB00671)
		P0AD65 Penicillin-binding protein 2	(DB00303; DB00438; DB00948; DB01163; DB01327; DB01328; DB01329; DB01413; DB01415; DB01598)
		Q9X6V3 Penicillin-binding protein 2	(DB00438; DB01413)
		Q70KI2 Penicillin-binding protein 2	(DB00923; DB01326)
		P0AD69 Peptidoglycan synthase FtsI	(DB05659)
		P0AD68 Peptidoglycan synthase FtsI	(DB00267; DB00274; DB00303; DB00430; DB00438; DB01327; DB01328; DB01329; DB01331; DB01332; DB01413; DB01415; DB01416; DB04918)
		B8DCL9 Penicillin-binding protein 3	(DB00485)
		P0A3M6 Penicillin-binding protein 2B	(DB00319; DB00415; DB00456; DB00485; DB00493; DB00567; DB00607; DB00713; DB00739; DB01066; DB01140; DB01163; DB01212; DB01331; DB01603; DB03313; DB08795)
		P0A3M5 Penicillin-binding protein 2B	(DB01147; DB01150; DB05659)
		P07944 Beta-lactam-inducible penicillin-binding Protein	(DB02443; DB02968; DB04041)
		Q51504 Penicillin-binding protein 3	(DB01147; DB01413; DB09050)
		Q60FT7 PBP3	(DB00535)
		P45059 Peptidoglycan synthase FtsI	(DB00303)
		P08149 Penicillin-binding protein 2	(DB00535)
		Q65HB5 Penicillin-binding protein 2A	(DB00689)
		P42971 Penicillin-binding protein 3	(DB00355; DB00493; DB01598; DB04570)
		Q2FGH1 Penicillin-binding protein 3	(DB01053)
		P59676 Penicillin-binding protein 2X	(DB03190)
		P14677 Penicillin-binding protein 2x	(DB01150; DB04918)
4	WP_006342755.1	P44868 tRNA (cytidine(34)-2'-O)-methyltransferase	(DB01752)
5	WP_006342760.1	P03007 DNA polymerase III subunit epsilon	(DB01643)
6	WP_006342763.1	P47205 UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine Deacetylase	(DB07861)
		O67648 UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase	(DB01991; DB04257; DB07355; DB07536; DB08231)
7	WP_006342764.1	O25928 3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	(DB07445)
		Q5G940 3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	(DB04216; DB06949; DB06950; DB06978; DB07044; DB07097; DB07098; DB07352; DB07715; DB08517)
		P0A6Q3 3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase	(DB03813)
		P0A6Q5 3-hydroxydecanoyl-[acyl-carrier-protein] Dehydratase	(DB03813)
8	WP_006342765.1	O25927 Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	(DB01694; DB08558)
		P07464 Galactoside O-acetyltransferase	(DB01862; DB01992; DB02632)
9	WP_006342777.1	Q5SHP7 30S ribosomal protein S17	(DB08185)
		P62658 30S ribosomal protein S17	(DB08185)
10	WP_006342805.1	Q96MA6 Adenylate kinase 8	(DB01717)
11	WP_006342820.1	P33590 Nickel-binding periplasmic protein	(DB03374)
		P06202 Periplasmic oligopeptide-binding protein	(DB07365)
		Q9X0V0 ABC-type transporter, periplasmic subunit	(DB01942)
12	WP_006342832.1	P41789 Nitrogen regulation protein NR(I)	(DB01857)
		P13632 C4-dicarboxylate transport transcriptional regulatory protein DctD	(DB04077)
		O32393 Adenylate cyclase	(DB02355; DB02596; DB07706)
		P0AE67 Chemotaxis protein CheY	(DB02461; DB03487; DB04156)
		Q9A5I5 Response regulator PleD	(DB01972)
13	WP_006342846.1	P33038 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	(DB01879; DB02435; DB02995; DB03089; DB04174; DB04474)
		P0A751 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	(DB03397)
		P0A749 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	(DB00828)
14	WP_006342849.1	P0A6I0 Cytidylate kinase	(DB02456; DB02883; DB03403; DB04555)
15	WP_006342881.1	P39594 Thiamine-phosphate synthase	(DB01788; DB02254; DB02885; DB03145; DB03416; DB07782)
		P25053 Regulatory protein TenI	(DB03570)

Continue……

16	WP_006342891.1	P0AFU8 Riboflavin synthase	(DB00140)
17	WP_006343008.1	P08877 Phosphocarrier protein HPr	(DB01899)
		P07515 Phosphocarrier protein HPr	(DB04522)
18	WP_006343009.1	Q6FEW8 Phosphoenolpyruvate-protein phosphotransferase	(DB08357)
		P22983 Pyruvate, phosphate dikinase	(DB02522)
19	WP_006343049.1	P22188 UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate Ligase	(DB02314; DB03590; DB03801)
		Q8DNV6 UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine Ligase	(DB06970)
		P45066 UDP-N-acetylmuramate--L-alanine ligase	(DB01673; DB03909; DB04395)
		P08192 Bifunctional protein FolC	(DB01015; DB02437; DB03830)
		P14900 UDP-N-acetylmuramoylalanine--D-glutamate Ligase	(DB01673; DB02314; DB03801; DB08105; DB08106; DB08107; DB08108; DB08112)
20	WP_006343057.1	P03007 DNA polymerase III subunit epsilon	(DB01643)
21	WP_006343078.1	P28248 Deoxycytidine triphosphate deaminase	(DB02333; DB03258)
22	WP_006343084.1	P0AGE0 Single-stranded DNA-binding protein	(DB04243)
23	WP_006343107.1	Q5SKW1 RNA polymerase sigma factor	(DB08226; DB08266)
		Q18BX5 RNA polymerase sigma factor	(DB08874)
24	WP_006343119.1	P45568 1-deoxy-D-xylulose 5-phosphate reductoisomerase	(DB02496; DB02948; DB03649; DB04272)
25	WP_006343123.1	P0A988 DNA polymerase III subunit beta	(DB06998)
26	WP_006343147.1	P96618 Holo-[acyl-carrier-protein] synthase	(DB01992; DB04447)
		P0A2W7 Holo-[acyl-carrier-protein] synthase	(DB01812)
27	WP_006343157.1	O25927 Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	(DB01694; DB08558)
		P0ACC7 Bifunctional protein GlmU	(DB01992; DB03397; DB03814)
28	WP_006343161.1	P0A6R0 3-oxoacyl-[acyl-carrier-protein] synthase 3	(DB01034; DB01992; DB02039; DB02316; DB03661; DB04524)
		Q820T1 3-oxoacyl-[acyl-carrier-protein] synthase 3	(DB07429)
		P0A574 3-oxoacyl-[acyl-carrier-protein] synthase 3	(DB03264; DB07611; DB07650; DB08171; DB08684; DB08712)
		Q9F6D4 3-oxoacyl-[acyl-carrier-protein] synthase 3	(DB01992)
29	WP_006343185.1	P0A721 Thymidylate kinase	(DB03280)
30	WP_006343186.1	P06710 DNA polymerase III subunit tau	(DB02930)
31	WP_006343194.1	P69772 Probable aromatic acid decarboxylase	(DB03247)
32	WP_006343261.1	Q9KPI8 5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	(DB07463)
33	WP_006343263.1	P42216 3-deoxy-manno-octulosonate cytidylyltransferase	(DB02344; DB02431; DB03403; DB04482; DB04555)
		P44490 3-deoxy-manno-octulosonate cytidylyltransferase	(DB04482)
		P0A0Z8 N-acylneuraminate cytidylyltransferase	(DB04555)
		Q8NFW8 N-acylneuraminate cytidylyltransferase	(DB02485)
34	WP_006343266.1	P06202 Periplasmic oligopeptide-binding protein	(DB07365)
		P33590 Nickel-binding periplasmic protein	(DB03374)
		Q5LRQ9 ABC transporter, periplasmic substrate-binding protein	(DB02078)
35	WP_006343268.1	P06202 Periplasmic oligopeptide-binding protein	(DB07365)
		P33590 Nickel-binding periplasmic protein	(DB03374)
36	WP_006343275.1	P00512 6-phosphofructokinase	(DB02726; DB04493)
37	WP_006343277.1	P00512 6-phosphofructokinase	(DB02726; DB04493)
38	WP_006343322.1	P43912 tRNA (guanine-N(1)-)-methyltransferase	(DB01752)
		P0A876 tRNA (guanine-N(1)-)-methyltransferase	(DB01752)
39	WP_006343376.1	P04036 4-hydroxy-tetrahydrodipicolinate reductase	(DB03969; DB04267)
		P72024 4-hydroxy-tetrahydrodipicolinate reductase	(DB04267)
40	WP_006343381.1	P12996 Biotin synthase	(DB03754; DB03775)
41	WP_006343383.1	P13000 ATP-dependent dethiobiotin synthetase BioD 1	(DB01715; DB02927; B02941; DB03624; DB03775)
42	WP_006343386.1	P0A6D3 3-phosphoshikimate 1-carboxyvinyltransferase	(DB01942; DB03116; B04328; DB04539)
		Q9S400 3-phosphoshikimate 1-carboxyvinyltransferase	(DB04328; DB04539)
		P0A751 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	(DB03397)
		P0A749 UDP-N-acetylglucosamine 1-carboxyvinyltransferase	(DB00828)
43	WP_006343388.1	P56122 Chorismate synthase	(DB03247)
		P0A2Y7 Chorismate synthase	(DB03247; DB03350)
44	WP_006343389.1	Q6GGU4 3-dehydroquinate synthase	(DB02592)
45	WP_006343390.1	P15770 Shikimate dehydrogenase	(DB03461; DB04447)
		P43876 Shikimate dehydrogenase	(DB02363)
46	WP_006343407.1	P62623 4-hydroxy-3-methylbut-2-enyl diphosphate reductase	(DB01785; DB04714)
47	WP_006343436.1	Q9X286 N utilization substance protein B homolog	(DB04272)
48	WP_006343437.1	P61432 UDP-N-acetylenolpyruvoylglucosamine reductase	(DB03147)
		P08373 UDP-N-acetylenolpyruvoylglucosamine reductase	(DB03147; DB07296)
49	WP_006343469.1	P49228 50S ribosomal protein L32	(DB01361)
50	WP_006343476.1	Q83LD8 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	(DB03687; DB04395)
51	WP_006343479.1	Q5SLP8 30S ribosomal protein S6	(DB08185)
52	WP_006343506.1	Q59771 L-phenylalanine dehydrogenase	(DB02494; DB03884)
		P49448 Glutamate dehydrogenase 2, mitochondrial	(DB00142; DB00157)
53	WP_006343524.1	P14900 UDP-N-acetylmuramoylalanine--D-glutamate ligase	(DB01673; DB02314; B03801; DB08105; DB08106; DB08107; DB08108; DB08112)

Continue……

54	WP_006343526.1	Q8DNV6 UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	(DB06970)
55	WP_006343531.1	P0AE12 AMP nucleosidase	(DB03464)
		P0A1F6 Uridine phosphorylase	(DB04627)
		P12758 Uridine phosphorylase	(DB01629; DB02256; B02681; DB03101; DB04485; DB06872; DB06873; DB07437; DB07439)
		Q8I3X4 Purine nucleotide phosphorylase, putative	(DB03881)
56	WP_006343542.1	P07771 Benzoate 1,2-dioxygenase electron transfer component	(DB03147)
		P39662 Flavohemoprotein	(DB03147; DB03979)
		P24232 Flavohemoprotein	(DB03147)
		P58558 Ferredoxin--NADP reductase	(DB03147; DB03461)
		P21890 Ferredoxin--NADP reductase	(DB03147; DB03461)
		P14779 Bifunctional P-450/NADPH-P450 reductase	(DB03440; DB04257; DB08086)
57	WP_006343552.1	O66529 6,7-dimethyl-8-ribityllumazine synthase	(DB02214; DB04128; DB04262)
		P11998 6,7-dimethyl-8-ribityllumazine synthase	(DB04162)
		P66034 6,7-dimethyl-8-ribityllumazine synthase	(DB01692; DB02135; B02184; DB02290; DB02693; DB02711; DB03022; DB03812; DB03973; DB08016)
		P61713 6,7-dimethyl-8-ribityllumazine synthase 2	(DB04162)
58	WP_006343559.1	P06709 Bifunctional protein BirA	(DB04651)
59	WP_006343561.1	P45124 Ribosomal small subunit pseudouridine synthase A	(DB01955)
		P0AA43 Ribosomal small subunit pseudouridine synthase A	(DB03419; DB03685)
60	WP_006343606.1	P37093 Type II secretion system protein E	(DB04395)
		Q7BK04 Cag alpha	(DB02930)
61	WP_006343637.1	P0A910 Outer membrane protein A	(DB04233)
62	WP_006343654.1	P75430 Probable 5-formyltetrahydrofolate cyclo-ligase	(DB02800)
63	WP_006343659.1	Q81VW8 Dihydropteroate synthase	(DB03592; DB03705; DB04047; DB04196)
		P0C0X1 Dihydropteroate synthase 1	(DB00250)
		P0AC13 Dihydropteroate synthase	(DB00259; DB00263; B00576; DB00634; DB01015; DB01298; DB01581; DB01582; DB06729)
		P53848 Folic acid synthesis protein FOL1	(DB00634)
		P0A578 Dihydropteroate synthase 1	(DB03592)
		Q49184 Dihydropteroate synthase type-1	(DB00891)
		P0C002 Dihydropteroate synthase type-1	(DB00634)
		Q27738 Dihydropteroate synthetase	(DB00359; DB00664; B01145; DB01299; DB06147; DB08798)
		P0C0X2 Inactive dihydropteroate synthase 2	(DB00250)
		P26281 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine yrophosphokinase	(DB02119; DB02596; B03197; DB04047; DB04158; DB04610)
		P43777 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine yrophosphokinase	(DB02278)
		P64143 2-amino-4-hydroxy-6-hydroxymethyldihydropteridine yrophosphokinase	(DB00233)
64	WP_006343660.1	P56740 Dihydroneopterin aldolase	(DB01778; DB01906; B02119; DB02489; DB03231; DB03571; DB04168; DB04400; DB04425; DB06906)
65	WP_006343661.1	Q18BX5 RNA polymerase sigma factor	(DB08874)
		Q5SKW1 RNA polymerase sigma factor	(DB08226; DB08266)
66	WP_006343668.1	P0ACC7 Bifunctional protein GlmU	(DB01992; DB03397; B03814)
		Q97R46 Bifunctional protein GlmU	(DB03397)
		P43889 Bifunctional protein GlmU	(DB08344)
67	WP_006343678.1	P03692 DNA primase/helicase	(DB02452; DB03222)
68	WP_013462583.1	P08506 D-alanyl-D-alanine carboxypeptidase DacC	(DB00274; DB00303; B00430; DB01329; DB01331)
		B8DD61 D-alanyl-D-alanine carboxypeptidase DacA (DD-peptidase) (DD-carboxypeptidase) (CPase) (Penicillin-binding protein5) (PBP-5)	(DB00485)
		P0AEB3 D-alanyl-D-alanine carboxypeptidase DacA	(DB01147; DB04647)
		P0AEB2 D-alanyl-D-alanine carboxypeptidase DacA	(DB00274; DB01329; B01331; DB04647)
		P72161 D-alanyl-D-alanine endopeptidase	(DB00438)
		Q75Y35 Penicillin-binding protein 3	(DB00319; DB00415; B00438; DB00447; DB00456; DB00485; DB00567; DB00607; DB00713; DB00739; DB00833; DB00948; DB01000; DB01140; DB01163; DB01330; DB01331; DB01603; DB03313; DB08795)
		P0AFI5 D-alanyl-D-alanine endopeptidase	(DB01331)
69	WP_013462587.1	P61177 50S ribosomal protein L22	(DB00199; DB00207; B01369)
70	WP_013462588.1	P0A7Z4 DNA-directed RNA polymerase subunit alpha	(DB00615)
		Q5SHR6 DNA-directed RNA polymerase subunit alpha	(DB08266)
		Q9Z9H6 DNA-directed RNA polymerase subunit alpha	(DB08226)
71	WP_013462596.1	Q9X180 Multi-sensor signal transduction histidine kinase	(DB02731)
		P23222 Sensor protein FixL	(DB02671)
		Q9X2W8 PPH	(DB04066)
		O32393 Adenylate cyclase	(DB02355; DB02596; B07706)
		P0AEJ4 Osmolarity sensor protein EnvZ	(DB04395)

Continue.....

72	WP_013462606.1	Q8EBR3 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	(DB07780)
		Q8RQP5 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	(DB03403; DB04555)
		A0R559 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	(DB04714)
		P62617 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	(DB02552)
		P62619 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	(DB01859; DB02552; B03403; DB03687; DB03961; DB04555)
		Q9PM68 Bifunctional enzyme IspD/IspF	(DB02552; DB03403)
73	WP_013462639.1	Q51504 Penicillin-binding protein 3	(DB01147; DB01413; B09050)
		P08149 Penicillin-binding protein 2	(DB00535)
		P0AD69 Peptidoglycan synthase FtsI	(DB05659)
		P0AD68 Peptidoglycan synthase FtsI	(DB00267; DB00274; B00303; DB00430; DB00438; DB01327; DB01328; DB01329; DB01331; DB01332; DB01413; DB01415; DB01416; DB04918)
		P45059 Peptidoglycan synthase FtsI	(DB00303)
		Q60FT7 PBP3	(DB00535)
		Q65HB5 Penicillin-binding protein 2A	(DB00689)
		P14677 Penicillin-binding protein 2x	(DB01150; DB04918)
		P59676 Penicillin-binding protein 2X	(DB03190)
		P44469 Penicillin-binding protein 2	(DB00303; DB00671)
		Q9X6V3 Penicillin-binding protein 2	(DB00438; DB01413)
		P0AD65 Penicillin-binding protein 2	(DB00303; DB00438; B00948; DB01163; DB01327; DB01328; DB01329; DB01413; DB01415; DB01598)
		Q70KI2 Penicillin-binding protein 2	(DB00923; DB01326)
		Q2FGH1 Penicillin-binding protein 3	(DB01053)
		P42971 Penicillin-binding protein 3	(DB00355; DB00493; B01598; DB04570)
		Q7DHH4 MecA	(DB04918)
		Q8DNB6 Penicillin-binding protein 2a	(DB00319; DB00415; B00456; DB00485; DB00493; DB00567; DB00607; DB00713; DB00739; DB01147; DB01163; DB01331; DB01603; DB03313; DB08795)
		P07944 Beta-lactam-inducible penicillin-binding protein	(DB02443; DB02968; B04041)
		B8DCL9 Penicillin-binding protein 3 (Pbp 3) (Pspb20)	(DB00485)
		P0A3M6 Penicillin-binding protein 2B	(DB00319; DB00415; B00456; DB00485; DB00493; DB00567; DB00607; DB00713; DB00739; DB01066; DB01140; DB01163; DB01212; DB01331; DB01603; DB03313; DB08795)
		P0A3M5 Penicillin-binding protein 2B	(DB01147; DB01150; B05659)
74	WP_013462674.1	P06202 Periplasmic oligopeptide-binding protein	(DB07365)
		P33590 Nickel-binding periplasmic protein	(DB03374)
		Q9X0V0 ABC-type transporter, periplasmic subunit	(DB01942)
75	WP_013462741.1	P17443 UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	(DB02196)
76	WP_028444368.1	Q9WYH8 Phospho-2-dehydro-3-deoxyheptonate aldolase	(DB01819; DB03937)
		O66496 2-dehydro-3-deoxyphosphooctonate aldolase	(DB01709; DB01819; B02053; DB02433; DB02992; DB03248; DB03745; DB03937)
		P0A716 2-dehydro-3-deoxyphosphooctonate aldolase	(DB01819; DB02433; B03113; DB03936)

Table S5: List of predicted functional classes of non-homologous hypothetical proteins

S. No.	Gene ID	Prediction	Score	
			R value	P value
1	WP_006342686.1	rRNA-binding proteins	7.6	99.0
2	WP_006342696.1	Lipopolysaccharide biosynthesis	6.0	99.0
3	WP_006342724.1	Metal-binding	3.3	95.7
4	WP_006342755.1	Glycosyltransferases	6.5	99.1
		lipid-binding proteins	4.2	98.0
5	WP_006342760.1	Iron-binding	2.2	86.8
6	WP_006342763.1	Hydrolases	2.8	92.9
		Transmembrane	2.6	91.3
7	WP_006342764.1	Lipid synthesis	1.5	73.8
8	WP_006342765.1	Transferases	3.9	97.5
		Lipid synthesis	3.4	96.1
		Zinc-binding	3.2	95.2
		lipid-binding proteins	2.8	92.9
9	WP_006342777.1	rRNA-binding proteins	4.5	98.3
10	WP_006342805.1	Transferases	3.7	97.0
		Hydrolases	3.2	95.2
		Zinc-binding	3.2	95.2
		DNA-binding	2.7	92.1
11	WP_006342820.1	lipid-binding proteins	5.5	98.9
		Transferases	4.2	98.0
		Iron-binding	3.8	97.3
12	WP_006342832.1	Zinc-binding	5.2	98.8

13	WP_006342846.1	Oxidoreductases	5.8	99.0
		Transferases	7.1	99.0
		Transferases	4.0	97.7
		Manganese-binding	3.9	97.5
14	WP_006342849.1	Transferases	5.7	99.0
15	WP_006342881.1	Lyases	1.4	71.3
16	WP_006342891.1	Transferases	4.9	98.6
17	WP_006343008.1	Hydrolases	1.1	62.2
18	WP_006343009.1	Zinc-binding	4.0	97.7
19	WP_006343049.1	Ligases	7.4	99.0
		Zinc-binding	4.6	98.4
		Lyases	3.2	95.2
		Metal-binding	2.6	91.3
		lipid-binding proteins	2.5	90.3
20	WP_006343057.1	Zinc-binding	2.7	92.1
21	WP_006343078.1	Hydrolases	3.7	97.0
22	WP_006343084.1	DNA replication	3.8	97.3
23	WP_006343107.1	Ligases	2.8	92.9
		DNA-binding	2.5	90.3
24	WP_006343119.1	Oxidoreductases	6.3	99.1
		Manganese-binding	4.5	98.3
		Transferases	3.2	95.2
		Transmembrane	3.1	94.7
		lipid-binding proteins	2.9	93.6
		Cobalt-binding	2.5	90.3
25	WP_006343123.1	Transferases	5.1	98.8
		DNA replication	4.8	98.6
		DNA-binding	3.7	97.0
26	WP_006343147.1	rRNA-binding proteins	2.7	92.1
27	WP_006343157.1	Zinc-binding	10.0	99.2
		Lipid synthesis	4.3	98.1
		lipid-binding proteins	3.2	95.2
		Transferases	2.5	90.3
28	WP_006343161.1	lipid-binding proteins	5.2	98.8
		Lipid synthesis	3.6	96.7
		Transferases	3.4	96.1
29	WP_006343185.1	lipid-binding proteins	2.8	92.9
30	WP_006343186.1	DNA-binding	3.2	95.2
		Zinc-binding	3.2	95.2
31	WP_006343194.1	Transmembrane	1.5	73.8
32	WP_006343261.1	lipid-binding proteins	3.9	97.5
		Transmembrane	2.8	92.9
33	WP_006343263.1	Transferases	2.6	91.3
		Hydrolases	2.6	91.3
		Transmembrane	2.5	90.3
34	WP_006343266.1	Zinc-binding	8.0	99.0
		Incompletely Characterized Transport Systems	2.9	93.6
35	WP_006343268.1	Zinc-binding	3.9	97.5
		Transferases	3.5	96.4
36	WP_006343275.1	Ligases	3.3	95.7
		Transferases	2.7	92.1
37	WP_006343277.1	Transferases	4.0	97.7
		Iron-binding	3.4	96.1
		Metal-binding	2.9	93.6
38	WP_006343322.1	DNA-binding	2.8	92.9
39	WP_006343376.1	Manganese-binding	4.0	97.7
		lipid-binding proteins	3.2	95.2
40	WP_006343381.1	Iron-binding	6.4	99.1
		Zinc-binding	2.8	92.9
41	WP_006343383.1	lipid-binding proteins	2.3	88.1
42	WP_006343386.1	Transferases	5.4	98.9
		lipid-binding proteins	5.4	98.9
		Zinc-binding	3.5	96.4
43	WP_006343388.1	Lyases	3.1	94.7
44	WP_006343389.1	Transmembrane	4.6	98.4
		lipid-binding proteins	3.6	96.7
		Cobalt-binding	3.5	96.4
45	WP_006343390.1	Transferases	4.5	98.3
		DNA replication	2.7	92.1
46	WP_006343407.1	Iron-binding	2.0	83.9
47	WP_006343436.1	transmembrane receptor	1.0	58.6
48	WP_006343437.1	Oxidoreductases	2.2	86.8
49	WP_006343469.1	Pore-forming toxins	1.0	58.6
50	WP_006343476.1	Hydrolases	2.8	92.9
		Oxidoreductases	2.7	92.1
		Hydrolases	2.5	90.3
51	WP_006343479.1	rRNA-binding proteins	4.3	98.1

Continue……

52	WP_006343506.1	Zinc-binding	4.4	98.2
		lipid-binding proteins	3.3	95.7
		Transmembrane	3.0	94.2
		Transferases	2.7	92.1
53	WP_006343524.1	Zinc-binding	6.3	99.1
		Ligases	4.5	98.3
		Lyases	4.3	98.1
		lipid-binding proteins	3.7	97.0
54	WP_006343526.1	Ligases	2.7	92.1
55	WP_006343531.1	Zinc-binding	2.0	83.9
56	WP_006343542.1	Primary Active Transporters	4.4	98.2
		Transmembrane	3.8	97.3
		lipid-binding proteins	3.7	97.0
		Oxidoreductases	3.6	96.7
		Iron-binding	3.0	94.2
		Manganese-binding	2.6	91.3
57	WP_006343552.1	Transferases	2.9	93.6
58	WP_006343559.1	Hydrolases	1.7	78.4
59	WP_006343561.1	Transferases	5.3	98.8
60	WP_006343606.1	Lyases	2.8	92.9
61	WP_006343637.1	Electrochemical Potential-driven transporters	3.0	94.2
62	WP_006343654.1	P-type ATPase	1.0	58.6
63	WP_006343659.1	lipid-binding proteins	4.2	98.0
		Transmembrane	3.2	95.2
		Metal-binding	3.0	94.2
64	WP_006343660.1	Iron-binding	1.1	62.2
65	WP_006343661.1	DNA-directed RNA polymerase	4.0	97.7
		DNA-binding	3.1	94.7
		Zinc-binding	2.5	90.3
66	WP_006343668.1	Manganese-binding	1.3	68.5
67	WP_006343678.1	DNA replication	6.1	99.0
		DNA-binding	4.5	98.3
		Zinc-binding	3.4	96.1
68	WP_013462583.1	Transferases	7.3	99.0
69	WP_013462587.1	rRNA-binding proteins	4.2	98.0
70	WP_013462588.1	DNA-binding	5.3	98.8
		Transferases	4.5	98.3
		DNA-directed RNA polymerase	4.4	98.2
71	WP_013462596.1	Transmembrane	1.9	82.2
72	WP_013462606.1	Magnesium-binding	1.2	65.4
73	WP_013462639.1	DNA-binding	1.8	80.4
74	WP_013462674.1	Transferases	2.1	85.4
75	WP_013462741.1	Transferases	6.7	99.1
		Transmembrane	6.0	99.0
		lipid-binding proteins	3.5	96.4
		Manganese-binding	2.9	93.6
76	WP_028444368.1	lipid-binding proteins	5.0	98.7
		Transmembrane	4.7	98.5
		Zinc-binding	3.5	96.4
		Transferases	2.8	92.9
		Manganese-binding	2.7	92.1

Table S6: Predicted metabolic pathway of non-homologous hypothetical proteins

S.No.	Gene id	Pathway	Kegg ID
1	WP_006342686.1	Protein synthesis	chp03010
2	WP_006342696.1	Lipopolysaccharide biosynthesis	Chp00540
		Metabolic pathways	chp01100
3	WP_006342760.1	Purine metabolism	Chp00230
		Pyrimidine metabolism	chp00240
		Metabolic pathways	chp01100
		DNA replication	chp03030
		DNA repair	chp03430
		Homologous recombination	chp03440
4	WP_006342763.1	Lipopolysaccharide biosynthesis	Chp00540
		Metabolic pathways	chp01100
5	WP_006342764.1	Fatty acid biosynthesis	Chp00061
		Biotin metabolism	chp00780
		Metabolic pathways	chp01100
		Fatty acid metabolism	chp01212
6	WP_006342765.1	Lipopolysaccharide biosynthesis	Chp00540
		Metabolic pathways	chp01100
		Cationic antimicrobial peptide (CAMP) resistance	chp01503
7	WP_006342777.1	Protein synthesis	chp03010
8	WP_006342805.1	RNA degradation	chp03018
9	WP_006342832.1	Two-component system	chp02020

10	WP_006342846.1	Amino sugar and nucleotide sugar metabolism	Chp00520
		Peptidoglycan biosynthesis	chp00550
		Metabolic pathways	chp01100
11	WP_006342849.1	Pyrimidine metabolism	Chp00240
		Metabolic pathways	chp01100
12	WP_006342881.1	Thiamine metabolism	chp00730
		Metabolic pathways	chp01100
13	WP_006342891.1	Riboflavin metabolism	chp00740
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
14	WP_006343009.1	Phosphotransferase system (PTS)	chp02060
15	WP_006343049.1	Lysine biosynthesis	chp00300
		Peptidoglycan biosynthesis	chp00550
16	WP_006343057.1	Purine metabolism	chp00230
		Pyrimidine metabolism	chp00240
		Metabolic pathways	chp01100
		DNA replication	chp03030
		Mismatch repair	chp03430
		Homologous recombination	chp03440
17	WP_006343078.1	Pyrimidine metabolism	chp00240
		Metabolic pathways	chp01100
18	WP_006343084.1	Pyrimidine metabolism	chp00240
		Metabolic pathways	chp01100
19	WP_006343107.1	Two-component system	chp02020
20	WP_006343119.1	Terpenoid backbone biosynthesis	chp00900
		Metabolic pathway	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Biosynthesis of antibiotics	chp01130
21	WP_006343123.1	Purine metabolism	chp00230
		Pyrimidine metabolism	chp00240
		Metabolic pathways	chp01100
		DNA replication	chp03030
		Mismatch repair	chp03430
		Homologous recombination	chp03440
22	WP_006343147.1	Pantothenate and CoA biosynthesis	chp00770
23	WP_006343157.1	Lipopolysaccharide biosynthesis	chp00540
		Metabolic pathway	chp01100
24	WP_006343161.1	Fatty acid biosynthesis	chp00061
		Metabolic pathways	chp01100
		Fatty acid metabolism	chp01212
25	WP_006343185.1	Pyrimidine metabolism	chp00240
		Metabolic pathways	chp01100
26	WP_006343186.1	Purine metabolism	chp00230
		Pyrimidine metabolism	chp00240
		Metabolic pathways	chp01100
		DNA replication	chp03030
		Mismatch repair	chp03430
		Homologous recombination	chp03440
27	WP_006343194.1	Ubiquinone and other terpenoid-quinone biosynthesis	chp00130
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
28	WP_006343263.1	Lipopolysaccharide biosynthesis	chp00540
		Metabolic pathways	chp01100
29	WP_006343275.1	Glycolysis / Gluconeogenesis	chp00010
		Pentose phosphate pathway	chp00030
		Fructose and mannose metabolism	chp00051
		Methane metabolism	chp00680
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Microbial metabolism in diverse environments	chp01120
		Biosynthesis of antibiotics	chp01130
		Carbon metabolism	chp01200
		Biosynthesis of amino acids	chp01230
		RNA degradation	chp03018
30	WP_006343277.1	Fructose and mannose metabolism	chp00051
31	WP_006343376.1	Monobactam biosynthesis	chp00261
		Lysine biosynthesis	chp00300
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Microbial metabolism in diverse environments	chp01120
		Biosynthesis of antibiotics	chp01130
		Biosynthesis of amino acids	chp01230
32	WP_006343381.1	Biotin metabolism	chp00780
		Metabolic pathways	chp01100
33	WP_006343383.1	Biotin metabolism	chp00780
		Metabolic pathways	chp01100

Continue……

34	WP_006343386.1	Phenylalanine, tyrosine and tryptophan biosynthesis	chp00400
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Biosynthesis of antibiotics	chp01130
		Biosynthesis of amino acids	chp01230
35	WP_006343388.1	Phenylalanine, tyrosine and tryptophan biosynthesis	chp00400
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Biosynthesis of antibiotics	chp01130
		Biosynthesis of amino acids	chp01230
36	WP_006343389.1	Phenylalanine, tyrosine and tryptophan biosynthesis	chp00400
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Biosynthesis of antibiotics	chp01130
		Biosynthesis of amino acids	chp01230
37	WP_006343390.1	Phenylalanine, tyrosine and tryptophan biosynthesis	chp00400
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Biosynthesis of antibiotics	chp01130
		Biosynthesis of amino acids	chp01230
38	WP_006343407.1	Terpenoid backbone biosynthesis	chp00900
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Biosynthesis of antibiotics	chp01130
39	WP_006343437.1	Amino sugar and nucleotide sugar metabolism	chp00520
		Peptidoglycan biosynthesis	chp00550
		Metabolic pathways	chp01100
40	WP_006343469.1	Protein synthesis	chp03010
41	WP_006343476.1	Terpenoid backbone biosynthesis	chp00900
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Biosynthesis of antibiotics	chp01130
42	WP_006343479.1	Protein synthesis	chp03010
43	WP_006343506.1	Valine, leucine and isoleucine degradation	chp00280
		Valine, leucine and isoleucine biosynthesis	chp00290
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
		Biosynthesis of antibiotics	chp01130
44	WP_006343524.1	D-Glutamine and D-glutamate metabolism	chp00471
		Peptidoglycan biosynthesis	chp00550
		Metabolic pathways	chp01100
45	WP_006343526.1	Lysine biosynthesis	chp00300
		Peptidoglycan biosynthesis	chp00550
		Metabolic pathways	chp01100
		Vancomycin resistance	chp01502
46	WP_006343531.1	Purine metabolism	chp00230
47	WP_006343552.1	Riboflavin metabolism	chp00740
		Metabolic pathways	chp01100
		Biosynthesis of secondary metabolites	chp01110
48	WP_006343559.1	Biotin metabolism	chb00780
		Metabolic pathways	chb01100
49	WP_006343606.1	Bacterial secretion system	chb03070
50	WP_006343654.1	One carbon pool by folate	chb00670
		Metabolic pathways	chb01100
51	WP_006343659.1	Folate biosynthesis	chb00790
		Metabolic pathways	chb01100
52	WP_006343660.1	Folate biosynthesis	chb00790
		Metabolic pathways	chb01100
53	WP_006343678.1	DNA replication	chb03030
54	WP_013462583.1	Peptidoglycan biosynthesis	chb00550
		Metabolic pathways	chb01100
55	WP_013462587.1	Protein synthesis	chb03010
56	WP_013462588.1	Purine metabolism	chb00230
		Pyrimidine metabolism	chb00240
		Metabolic pathways	chb01100
		RNA polymerase	chb03020
57	WP_013462596.1	Two-component system	chb02020
58	WP_013462606.1	Terpenoid backbone biosynthesis	chb00900
		Metabolic pathways	chb01100
		Biosynthesis of secondary metabolites	chb01110
		Biosynthesis of antibiotics	chb01130
59	WP_013462639.1	Peptidoglycan biosynthesis	chb00550
		beta-Lactam resistance	chb01501
60	WP_013462674.1	beta-Lactam resistance	chb01501
		ABC transporters	chb02010
61	WP_013462741.1	Peptidoglycan biosynthesis	chb00550
		Metabolic pathways	chb01100
		Vancomycin resistance	chb01502
62	WP_028444368.1	Phenylalanine, tyrosine and tryptophan biosynthesis	chb00400
		Metabolic pathways	chb01100
		Biosynthesis of secondary metabolites	chb01110
		Biosynthesis of antibiotics	chb01130
		Biosynthesis of amino acids	chb01230

Conclusion

In the current study, we have utilized subtractive genomic approach based on various bioinformatics centered databases and computational tools in order to identify essential non-homologous drugable proteins against *C. psittaci*. We also predicted the subcellular localization, drugable potential and unique essential pathogen's metabolic pathways in which these identified putative targets may involve. Furthermore, we have predicted the molecular modeling and molecular docking of one of the identified drug targets. The selected protein for molecular modeling and docking studies, found to be an iron binding protein and involved in the isoprenoid biosynthesis; a crucial and essential pathway of various pathogenic bacteria. The docking study of selected protein with its ligand can provide a better understanding of active site and its binding affinity. This can further open new paths of proposing new potent inhibitors of the selected targeted protein against *C.psittaci*

REFERENCES

Abadio, A. K. R., Kioshima, E. S., Teixeira, M. M., Martins, N. F., Maigret, B., & Felipe, M. S. S. 2011. Comparative genomics allowed the identification of drug targets against human fungal pathogens. *BMC genomics, 12*(1), 75.

AbdelRahman, Y. M., & Belland, R. J. 2005. The chlamydial developmental cycle. *FEMS microbiology reviews, 29*(5), 949-959.

Allsop, A. E., Brooks, G., Bruton, G., Coulton, S., Edwards, P. D., Hatton, I. K., Kaura, A. C., McLean, S. D., Pearson, N. D., & Smale, T. C. 1995. Penem inhibitors of bacterial signal peptidase. *Bioorganic & Medicinal Chemistry Letters, 5*(5), 443-448.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. 1990. Basic local alignment search tool. *Journal of molecular biology, 215*(3), 403-410.

Amineni, U., Pradhan, D., & Marisetty, H. 2010. In silico identification of common putative drug targets in Leptospira interrogans. *Journal of chemical biology, 3*(4), 165-173.

Anishetty, S., Pulimi, M., & Pennathur, G. 2005. Potential drug targets in Mycobacterium tuberculosis through metabolic pathway analysis. *Computational biology and chemistry, 29*(5), 368-378.

Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics, 22*(2), 195-201.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. 1978. The Protein Data Bank: a computer-based archival file for macromolecular structures. *Archives of biochemistry and biophysics, 185*(2), 584-591.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., & Bordoli, L. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic acids research*, gku340.

Branley, J., Roy, B., Dwyer, D., & Sorrell, T. 2008. Real-time PCR detection and quantitation of Chlamydophila psittaci in human and avian specimens from a veterinary clinic cluster. *European Journal of Clinical Microbiology & Infectious Diseases, 27*(4), 269-273.

Cai, C., Han, L., Ji, Z. L., Chen, X., & Chen, Y. Z. 2003. SVM-Prot: web-based support vector machine software for functional classification of a protein from its primary sequence. *Nucleic acids research, 31*(13), 3692-3697.

Consortium, U. 2008. The universal protein resource (UniProt). *Nucleic acids research, 36*(suppl 1), D190-D195.

Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., & Baker, N. A. 2007. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic acids research, 35*(suppl 2), W522-W525.

Dutta, A., Singh, S. K., Ghosh, P., Mukherjee, R., Mitter, S., & Bandyopadhyay, D. 2006. In silico identification of potential therapeutic targets in the human pathogen Helicobacter pylori. *In silico biology, 6*(1-2), 43-47.

Fu, L. L., Liu, J., Chen, Y., Wang, F. T., Wen, X., Liu, H. Q., Wang, M. Y., Ouyang, L., Huang, J., & Bao, J. K. 2014. In silico analysis and experimental validation of azelastine hydrochloride (N4) targeting sodium taurocholate co-transporting polypeptide (NTCP) in HBV therapy. *Cell proliferation, 47*(4), 326-335.

Gaede, W., Reckling, K. F., Dresenkamp, B., Kenklies, S., Schubert, E., Noack, U., Irmscher, H. M., Ludwig, C., Hotzel, H., & Sachse, K. 2008. Chlamydophila psittaci infections in humans during an outbreak of psittacosis from poultry in Germany. *Zoonoses and public health, 55*(4), 184-188.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research, 31*(13), 3784-3788.

Georrge, J. J., & Umrania, V. 2011. In silico identification of putative drug targets in Klebsiella pneumonia MGH 78578. *Indian Journal of Biotechnology, 10*(4), 432-439.

Haag, N. L., Velk, K. K., & Wu, C. 2011. *In silico identification of drug targets in methicillin/multidrug-resistant Staphylococcusaureus.* Paper presented at the Proc. Third International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies, BIOTECHNO.

Harkinezhad, T., Verminnen, K., De Buyzere, M., Rietzschel, E., Bekaert, S., & Vanrompay, D. 2009. Prevalence of Chlamydophila psittaci infections in a human population in contact with domestic and companion birds. *Journal of medical microbiology, 58*(9), 1207-1212.

Heuston, S., Begley, M., Gahan, C. G., & Hill, C. 2012. Isoprenoid biosynthesis in bacterial pathogens. *Microbiology, 158*(Pt 6), 1389-1401.

Hotzel, H., Berndt, A., Melzer, F., & Sachse, K. 2004. Occurrence of Chlamydiaceae spp. in a wild boar (Sus scrofa L.) population in Thuringia (Germany). *Veterinary microbiology, 103*(1), 121-126.

Hughes, C., Maharg, P., Rosario, P., Herrell, M., Bratt, D., Salgado, J., & Howard, D. 1997. Possible nosocomial

transmission of psittacosis. *Infection Control, 18*(03), 165-168.

Jaroszewski, L. 2009. Protein structure prediction based on sequence similarity *Biomedical Informatics* (pp. 129-156): Springer.

Kaleta, E., & Taday, E. M. 2003. Avian host range of Chlamydophila spp. based on isolation, antigen detection and serology. *Avian pathology, 32*(5), 435-462.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. 2011. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic acids research*, gkr988.

Kerfeld, C. A., & Scott, K. M. 2011. Using BLAST to teach "E-value-tionary" concepts. *PLoS Biol, 9*(2), e1001014.

Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., & Neveu, V. 2011. DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic acids research, 39*(suppl 1), D1035-D1041.

Kopp, J., & Schwede, T. 2004. Automated protein structure homology modeling: a progress report. *Pharmacogenomics, 5*(4), 405-416.

Lang, P. T., Brozell, S. R., Mukherjee, S., Pettersen, E. F., Meng, E. C., Thomas, V., Rizzo, R. C., Case, D. A., James, T. L., & Kuntz, I. D. 2009. DOCK 6: Combining techniques to model RNA–small molecule complexes. *Rna, 15*(6), 1219-1230.

Lee, M. J., Min, B.-J., Choung, H.-K., Kim, N., Kim, Y. A., & Khwarg, S. I. 2014. Genome-wide DNA methylation profiles according to Chlamydophila psittaci infection and the response to doxycycline treatment in ocular adnexal lymphoma. *Molecular vision, 20*, 1037.

Meselson, M., Guillemin, J., & Hugh-Jones, M. 2002. Public health assessment of potential biological terrorism agents. *Emerging infectious diseases, 8*(2), 225.

Miller, C., Songer, J., & Sullivan, J. 1987. A twenty-five year review of laboratory-acquired human infections at the National Animal Disease Center. *The American Industrial Hygiene Association Journal, 48*(3), 271-275.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic acids research, 35*(suppl 2), W182-W185.

Nancy, Y. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S. C., Ester, M., & Foster, L. J. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics, 26*(13), 1608-1615.

NCBI, R. C. 2015. Database resources of the National Center for Biotechnology Information. *Nucleic acids research, 43*(Database issue), D6.

Peters, J., Wilson, D. P., Myers, G., Timms, P., & Bavoil, P. M. 2007. Type III secretion a la Chlamydia. *Trends in microbiology, 15*(6), 241-251.

Read, T. D., Brunham, R., Shen, C., Gill, S., Heidelberg, J., White, O., Hickey, E., Peterson, J., Utterback, T., & Berry, K. 2000. Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39. *Nucleic acids research, 28*(6), 1397-1406.

Read, T. D., Joseph, S. J., Didelot, X., Liang, B., Patel, L., & Dean, D. 2013. Comparative analysis of Chlamydia psittaci genomes reveals the recent emergence of a pathogenic lineage with a broad host range. *MBio, 4*(2), e00604-00612.

Read, T. D., Myers, G., Brunham, R., Nelson, W., Paulsen, I., Heidelberg, J., Holtzapple, E., Khouri, H., Federova, N., & Carty, H. 2003. Genome sequence of Chlamydophila caviae (Chlamydia psittaci GPIC): examining the role of niche-specific genes in the evolution of the Chlamydiaceae. *Nucleic acids research, 31*(8), 2134-2147.

Reddy, K. G., Rao, N. K., Krishna, P. R. B., & Aravind, S. 2010. In silico identification of potential therapeutic targets in *Clostridium botulinum* by the approach subtractive genomics. *International Journal of Bioinformatics Research, 2*(2).

Sakharkar, K. R., Sakharkar, M. K., & Chow, V. T. 2004. A novel genomics approach for the identification of drug targets in pathogens, with special reference to Pseudomonas aeruginosa. *In Silico Biology, 4*(3), 355-360.

Shahbaaz, M., Bisetty, K., Ahmad, F., & Hassan, I. 2016. Current advances in the identification and characterization of putative drug and vaccine targets in the bacterial genomes. *Current topics in medicinal chemistry, 16*(9), 1040-1069.

Sharma, V., Gupta, P., & Dixit, A. 2008. In silico identification of putative drug targets from different metabolic pathways of Aeromonas hydrophila. *In silico biology, 8*(3-4), 331-338.

Shoukat, K., Rasheed, N., & Sajid, M. 2012. Subtractive genome analysis for in silico identification and characterization of novel drug targets in *C. trachomatis* STRAIN D/UW-3/Cx. *International Journal of Current Research, 4*(7), 017-021.

Smith, K. A., Bradley, K. K., Stobierski, M. G., & Tengelsen, L. A. 2005. Compendium of measures to control Chlamydophila psittaci (formerly Chlamydia psittaci) infection among humans (psittacosis) and pet birds, 2005. *Journal of the American Veterinary Medical Association, 226*(4), 532-539.

Smith, K. A., Campbell, C. T., Murphy, J., Stobierski, M. G., & Tengelsen, L. A. 2011. Compendium of measures to control Chlamydophila psittaci infection among humans (psittacosis) and pet birds (avian chlamydiosis), 2010 National Association of State Public Health Veterinarians (NASPHV). *Journal of Exotic Pet Medicine,* 20(1), 32-45.

Stephens, R. S., Kalman, S., Lammel, C., Fan, J., Marathe, R., Aravind, L., Mitchell, W., Olinger, L., Tatusov, R. L., & Zhao, Q. 1998. Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis. *Science, 282*(5389), 754-759.

Thapa, S., Alum, M. A., Saha, C., Zubaer, A., Akter, A., & Mondal, S. I. 2013. Molecular modeling and functional characterization of a pertinent enzyme in Streptococcus pneumoniae serotype-2: A potential target for the development of novel pneumonia drug. *Indian Journal of Biotechnology,* 12(1), 40-45.

Thomson, N. R., Yeats, C., Bell, K., Holden, M. T., Bentley, S. D., Livingstone, M., Cerdeño-Tárraga, A. M., Harris, B., Doggett, J., & Ormond, D. 2005. The Chlamydophila abortus genome sequence reveals an array of variable proteins that contribute to interspecies variation. *Genome research, 15*(5), 629-640.

Van Lent, S., Piet, J. R., Beeckman, D., van der Ende, A., Van Nieuwerburgh, F., Bavoil, P., Myers, G., Vanrompay, D., & Pannekoek, Y. 2012. Full genome sequences of all nine Chlamydia psittaci genotype reference strains. *Journal of bacteriology, 194*(24), 6930-6931.

Voigt, A., Schofl, G., & Saluz, H. P. 2012. The Chlamydia psittaci genome: a comparative analysis of intracellular pathogens. *PloS one, 7*(4), e35097.

Zhang, R., & Lin, Y. 2009. DEG 5.0, a database of essential genes in both prokaryotes and eukaryotes. *Nucleic acids research, 37*(suppl 1), D455-D458.

Zhang, R., Ou, H. Y., & Zhang, C. T. 2004. DEG: a database of essential genes. *Nucleic acids research, 32*(suppl 1), D271-D272.

****