Publications
JOURNAL PUBLICATIONS based on this Thesis


Bilachi S Ravindranath, Venkatappa Krishnamurthy, Venkatarangaiah Krishna "Modelling and Docking of potential drug target nrdB "has been communicated.

Bilachi S Ravindranath, Venkatappa Krishnamurthy, Venkatarangaiah Krishna "Synteny based comparative analysis of Chalmydophila.pneumoniae conserved regions for target discovery" (In preparation)
CONFERENCE PUBLICATIONS

Bilachi S Ravindranath, Venkatappa Krishnamurthy, Venkatarangaiah Krishna
"A Synteny based comparative study genomic study of respiratory pathogens
with reference to Chlamydophila pneumonia . National symposium "Biosatum-
2012" as oral presentation in Dept of Biotechnology, BIET, Davangere, May

Bilachi S Ravindranath, Venkatappa Krishnamurthy, Venkatarangaiah Krishna
"Computational approach for the genome analysis of Chlamydophila.pneumonia
with reference to infective chronic obstructive pulmonary disease (COPD) "as
oral presentation in the AOHK National conference at JJMC Medical college,
In silico synten based comparative genomics approach for identification and characterization of novel therapeutic targets in Chlamydophila pneumoniae

Bilachi S Ravindranath¹*, Venkatappa Krishnamurthy¹, Venkatarangaiah Krishna² & Sunil Kumar C¹

¹Department of Biotechnology, PES Institute of Technology (Recognized research centre of Kuvempu University), BSK III Stage, Bangalore - 560085, India; ²Department of Biotechnology and Bioinformatics, Kuvempu University, Shankarghatta -577451, India; Bilachi S Ravindranath Email: ravindranathbs@gmail.com; Phone:+91 9538277478; *Corresponding author

Received May 02, 2013; Accepted May 07, 2013; Published June 08, 2013

Abstract:
Chlamydophila pneumoniae is one of the most important and well studied gram negative bacterial strain with respect to community acquired pneumonia and other respiratory diseases like Chronic obstructive pulmonary disease (COPD), Chronic asthma, Alzheimer’s disease, Atherosclerosis and Multiscerosis which have a great potential to infect humans and many other mammals. According to WHO prediction, COPD is to become the third leading cause of death by 2030. Unfortunately, the molecular mechanisms leading to chronic infections are poorly understood and the difficulty in culturing C pneumoniae in experimental conditions and lack of entirely satisfactory serological methods for diagnosis is also a hurdle for drug discovery and development. We have performed an insilico synten synten based comparative genomics analysis of C pneumoniae and other eight Chlamydial organisms to know the potential of C pneumoniae which cause COPD but other Chlamydial organisms lack in potential to cause COPD though some are involved in human pathogenesis. We have identified total 354 protein sequences as non-orthologous to other Chlamydial organisms, except hypothetical proteins 70 were found functional out of which 60 are non homologous to Homo sapiens proteome and among them 18 protein sequences are found to be essential for survival of the C pneumoniae based on BLASTP search against DEG database of essential genes. CELLO analysis results showed that about 80% proteins are found to be cytoplasmic. Among which 5 were found as bacterial exotoxins and 2 as bacterial endotoxins, remaining 11 proteins were found to be involved in DNA binding, RNA binding, catalytic activity, ATP binding, oxidoreductase activity, hydrolase activity and proteolysis activity. It is expected that our data will facilitate selection of C pneumoniae proteins for successful entry into drug design pipelines.

Key Words: C pneumoniae, Chlamydial, COPD, comparative analysis.

Background:
The genus Chlamydia was established in 1966, Chlamydiaceae family consists of obligate intracellular gram-negative bacterium, which possesses a distinctive and a composite biphasic developmental cycle [1]. C pneumoniae was separated as a distinct species in 1992 [2]. It is perhaps one of the most successful Chlamydial species, which have established a forte in a range homoeothermic and poikilothermic hosts, including humans, animals, amphibians and reptiles [3]. C pneumoniae is a common cause of upper respiratory tract infections and pneumonia and has been associated with several chronic inflammatory conditions such as atherosclerosis [4] and COPD
It is a very common bacterium worldwide, and almost everyone is infected at some point of their life. In some cases, acute C. pneumoniae infection can become chronic. The molecular mechanism of chronic infections is poorly understood, which has lead to a major setback in combating these pathogens. C. pneumoniae infection has also been reported as a cause of lower respiratory tract infection. The lower tract infection has a direct impact on the pathogenesis, diagnosis and prognosis of COPD in several ways. Several recent group studies suggest that lung growth is impaired by childhood lower respiratory tract infection, in making these individuals more susceptible to develop COPD. Several recent group studies suggest that lung growth is impaired by childhood lower respiratory tract infection, making these individuals more susceptible to develop COPD. This chronic colonization of the lower respiratory tract by bacterial pathogens could induce a chronic inflammatory response with lung damage.

C. pneumoniae, usually regarded as a human respiratory pathogen, has been demonstrated to cause intracellular infections of the upper and lower respiratory tract respiratory tissue worldwide. C. pneumoniae infections are often importunate, and an acute infection may sometimes turn chronic. Acute C. pneumoniae infection can cause bronchitis, emphysema and pneumonia, in addition, more serious diseases such as atherosclerosis and stroke. Myocarditis, Alzheimer's disease and multiple sclerosis. COPD have been associated with chronic C. pneumoniae infection. Chronic infection with C. pneumoniae is being seriously investigated as a cause of several systemic diseases, studies reveal the elevated incidence of C. pneumoniae infection in COPD.

These chronic infections of respiratory tissue could contribute to the pathogenesis of COPD by altering the host response which has proven to be extremely difficult to diagnose and impossible to treat with current antibiotics. Thus, development of safe and effective vaccines represents a cost-effective approach that would have a greater impact on the high prevalence of Chlamydia infections and the prevention of severe long-term sequelae. Thus, new antichlamydial drug targets are urgently needed to be identified. The first identified case of C. pneumoniae infection was reported in Taiwan. It was identified as TWAR organism, originally derived from the names of the two isolates - Taiwan (TW-183) and an acute respiratory isolate designated as AR-39. By the availability of genome sequence of C. pneumoniae AR-39, we considered AR-39 as a reference species as all the C. pneumoniae strains share ~99.5% genome in common but comparably AR-39 has more number of coding genes and gene products. Availability of the whole genome sequence which was first sequenced in 1999 and has been deposited in Genbank database has paved a way to this research.

By the application of newer molecular and genomics research techniques/tools with the modern approach like Synteny Based Comparative Genomics (SBCG) is helping us to find precisely the important genes which are conserved and play a role of bacterial infection in COPD which can be identified as potential drug targets.

Figure 1: Screen shot showing the Synetview tool view with selected organisms for comparison, and grey colored synteny blocks indicates non-orthologous regions predicted.
Methodology:

Reference organism selection and genome retrieval
C. pneumonia AR39, the reference species was selected based on literature survey, database availability of complete genomic information, moderate genome size in comparative to other strains of C. pneumonia pathogens (share ~99.5% identical genes). The genome sequence of the C. pneumonia was retrieved from Syntebase, which is a prokaryotic genome database, and holds a relational database as the raw data retrieved from public genome databanks (GenBank/EMBL/DDJB). The raw data are organized into two tables, genome table contain the information about the prokaryotic genomes compared and the gene/protein table contains many features like amino acid sequences, length, species name, location of gene, etc.

Genome comparison of reference organism with other selected Chlamydial species
Syntebase 1.0 is an inbuilt tool available in Syntebase at http://www.syntebase.unine.fr/, which was designed to visualize conserved syntenic blocks in a large number of prokaryotic genomes after choosing C. pneumonia AR39 as a reference organism. It was employed to identify the non-orthologous syntenic blocks in all the Chlamydial pathogens considered in this study. In the present study Web service mode was used to retrieve and analyze the data from Syntebase. The reference species was selected by using “Reference” tab and the “Compared” tab was used for selecting and adding the compared species, includes 8 species of Chlamydia viz., Chlamydia trachomatis, Chlamydia psittaci, Chlamydia muridarum, Chlamydia pecorum, Chlamydophila caviae, Chlamydophila abortus, Chlamydophila felis, Candidatus Parachlamydia acanthamoebae (Figure 1).

The “Start retrieval process” tab was used for genome comparison and identification orthologous syntenic blocks. The program compares each pair of proteomes based on the criteria like E-values <10^-8, the alignment extended for at least 80% of the length of the shorter matching protein [10].

Prediction of non-orthologous (uncommon) genes/proteins
The tool displays the strict conservation of gene order, in comparing to a reference species with compared species. Each gene present in the genome of the reference species is depicted by a rectangle with a color code. The blue, yellow and grey colored rectangles represents positive strand, negative strand and genes without orthologs belonging to a syntenic block in another species. Based on the color code the non-orthologous gene was selected and used for further analysis. Gene information panel located on the left hand side of the tool was used for retrieving the sequence related information like its gi number - Identifier, the name of its species - Species, Gene name, function, its location-Chromosome and sequence information. The sequence information was used to retrieve the protein sequence data and used for further study (Figure 1).

Insilico identification of non-orthologous essential genes and their localization
The predicted non-orthologous protein sequence of the C. pneumonia was subjected to BLASTP at http://blast.ncbi.nlm.nih.gov/ [11] specifically against DEG 7.0 (Database of Essential Genes) at http://tubic.tju.edu.cn/deg/ [12] for screening of essential genes. The DEG database, compiles literature and sequences of experimentally verified essential genes and proteins from Gram-positive and Gram-negative bacteria. The cut off value for database search used was E-value <1E^-10, bit score >100 and percentage of identity >35% at amino acid level were considered.

Insilico prediction of non-host genes and prediction of toxigenesis
The BLASTP analysis was carried out between host and pathogen, for the identification of non-host proteins from C. pneumonia. The database search was made against protein database of H. sapiens (taxid: 9606), through an option available under the BLASTP parameters. The protein hits showing E-value <1E^-10 and sequence identity <35% were selected. Localization of the predicted non-host, non-orthologous, and essential genes was searched using CELLO at http://cello.life.nctu.edu.tw/ [13]. The bacterial toxin producing genes have been identified and classified into exotoxins and endotoxins using BTEXpred tool at http://www.imtech.res.in/raghava/btxpred/ [14], further the proteins were subjected to Interpro server 41.0 at http://www.ebi.ac.uk/interpro/ [15] to extrapolate the domains, signatures, functions and pathways in which they are involved.

Testing the druggability and prediction of therapeutic targets
The druggability of the predicted non-orthologous essential genes involved in production of toxin in pathogenic systems was analyzed using Drug Bank database at http://www.drugbank.ca/ [16] by using Pharma search. Proteins were subjected to BLAST against KEGG database http://www.genome.jp/kegg/ [17] to know in which pathway they are involved. It was found that all the putative therapeutic targets predicted were present in the unique metabolic pathways of C. pneumonia (Unpublished data).

![Figure 2: Showing classification of synteny blocks into putative drug targets.](image)

Results & Discussion:

Genome comparison and Non-Orthologous gene prediction
C. pneumoniae was the reference organism which encloses 1128 synteny blocks and it was compared to identify organism specific synteny blocks as they lack the ability to cause COPD in human beings. Out of these, 354 non-orthologous synteny blocks were identified and protein sequences were retrieved (Table 1). Out of the 354 synteny blocks hypothetical proteins were filtered to avoid the noise.
Non-Orthologous Essential gene prediction and localization
Protein coding sequences were searched for identifying non-
homologous proteins, against human proteome using BLASTP
and 60 proteins were identified as non-homologues to the
human proteome and were subjected to Protein BLAST against
DEG database with E-value cutoff score of 10^-5 and 35%
identity 18 proteins comprising 1% of the total number of
protein coding sequences in C pneumoniae AR39 were found to
be essential and non-host (Figure 2).

Non-Host gene prediction and Bacterial Toxins prediction
Bacterial protein toxins are the most potent poisons known and
may show activity at very high dilutions where toxins are the
major determinant of virulence, usually virulent strains of
bacterium produce range of toxins which is not observed in
non-virulent strains. The protein toxins are soluble proteins
secreted by living bacteria which are essential proteins for their
survival, attack and defensive mechanism. We have been
successful in identifying these toxins in C pneumoniae which can
be classified into exotoxins and endotoxins where the latter is
specific to Gram negative bacteria’s like Chlamydia.

Among the 18 putative uncharacterized essential proteins
predicted toxins were analyzed using the BTxPred tool which
predicts the bacterial toxins based on the protein sequence and
it classifies them into exotoxins and endotoxins, 5 proteins were
identified as bacterial exotoxin proteins and 2 as bacterial
endotoxin proteins which are involved in toxigenesis with the
help of Lipopolysaccharide components attached to produce
cytotoxicity and play an important role in pathogenesis.

Testing Druggability and prediction of Therapeutic targets
We report these bacterial toxins as potential candidate targets.
As no specific drugs are available when checked in the Drug
Bank, so there is an urgent need to develop specific inhibitors of
these targets as these are the primary cause for the pathogenesis
of the C pneumoniae and related organisms which are involved in
causing diseases in humans.We extrapolated the remaining
proteins using Interpro, which helps in analyzing the given
protein’s function and classifies based on their family and
domains involved, and also it provides information on
important sites of protein and integrates the protein signature’s
information from the member databases.

We successfully predicted 1 RNA binding protein, 6 DNA
binding proteins, 4 proteins involved in catalytic activity, 3 ATP
binding proteins, 1 chaperone protein, 1 protein involved in
Oxidoreductase activity,1 protein involved in hydrolyase
activity, 1 protein involved in proteolysis. We recommend 4
proteins namely signal recognition particle protein PtsY and
cysteinyl-tRNA synthetase which are endotoxins involved in
pathogenesis, serine protease and Clp Protease are proteins
which are ubiquitous and serves as virulence factor in causing
the disease.

Conclusion
The SBCG analysis of C pneumoniae, have led to identification of
several proteins in the C pneumoniae genome that can be
targeted for potential drug design and vaccine development. As
various identified drug targets have been reported to play a
vital role in the important metabolic pathways that regulate
bacterial pathogenicity and necessary nutrient uptake, a novel
systematic approach to develop drugs against these targets
would likely be very promising for the treatment of COPD.

The present study has thus led to the identification of several
proteins which are present in unique metabolic pathways of C
pneumoniae (data not shown, unpublished data), and identified
drug targets in the present study are relatively small in number
which can be targeted for effective drug design and vaccine
development against C pneumoniae. It is expected that the drugs
developed against the identified targets will be specific to the
pathogen and less or not toxic for the host. Homology Modeling
of these target proteins will facilitate for identifying crucial sites
which can be targeted for drug designing.

References:
15039303]
10319462]
12086181]
11371760]
19782917]
9259082]
18487846]
10684935]
19087285]
[PMID: 9254694]
18974178]
21059682]
22080510]

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium,
for non-commercial purposes, provided the original author and source are credited
Supplementary material:

Table 1: Showing the reference organism and compared organism’s information.

<table>
<thead>
<tr>
<th>Sl No</th>
<th>Organism name</th>
<th>Accession Id</th>
<th>Number of Nucleotides in bp</th>
<th>GC %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Chlamydophila pneumoniae AR 39</td>
<td>NC_002179.2</td>
<td>1,229,853</td>
<td>40.6</td>
</tr>
<tr>
<td>2.</td>
<td>Chlamydia trachomatis A/HAR-13</td>
<td>NC_007429.1</td>
<td>1044459</td>
<td>41.3</td>
</tr>
<tr>
<td>3.</td>
<td>Chlamydophila felis Fe/C-56</td>
<td>NC_007899.1</td>
<td>1,166,239</td>
<td>39.4</td>
</tr>
<tr>
<td>4.</td>
<td>Chlamydomophila caviae GPIC</td>
<td>NC_003361.3</td>
<td>1,173,390</td>
<td>39.2</td>
</tr>
<tr>
<td>5.</td>
<td>Chlamydomophila abortus S26/3</td>
<td>NC_004552.2</td>
<td>1,144,377</td>
<td>39.9</td>
</tr>
<tr>
<td>6.</td>
<td>Chlamydia muridarum Nigg</td>
<td>NC_002620.2</td>
<td>1,072,950</td>
<td>40.3</td>
</tr>
<tr>
<td>7.</td>
<td>Candidatus acanthamoebae UV-7</td>
<td>NC_015702.1</td>
<td>3,072,383</td>
<td>39</td>
</tr>
</tbody>
</table>