Biomedical Data Mining

Extended Abstract

Conserved epitopes of DENV structural and non-structural proteins for exploring universal vaccine targets

Mansi Verma

Abstract

Dengue is a serious arising arthropod-borne viral illness happening all around the world. Around two-fifths of the total populace or up to 3.9 billion individuals are at a danger of dengue contamination. Disease instigates a deep rooted defensive resistance to the homologous serotype however gives just fractional and transient insurance against resulting contamination against other serotypes. Accordingly, there is a requirement for an antibody which is fit for giving a deep rooted assurance against all the serotypes of dengue infection. In our investigation, similar genomics of Dengue virus (DENV) was led to investigate possible contender for novel immunization targets.

From our investigation we effectively discovered 100% moderated epitopes in Envelope protein (RCPTQGE); NS3 (SAAQRRGR, PGTSGSPI); NS4A (QRTPQDNQL); NS4B (LQAKATREAQKRA) and NS5 proteins (QRGSGQV) in all DENV serotypes. Some serotype explicit monitored themes were additionally found in NS1, NS5, Capsid, PrM and Envelope proteins. Utilizing near genomics and insusceptible informatics approach, we could discover rationed epitopes which can be investigated as peptide antibody possibility to battle dengue around the world. Serotype-explicit epitopes can likewise be misused for fast diagnostics. Every one of the ten proteins are investigated to locate the saved epitopes in DENV serotypes, subsequently making it the most widely examined viral genome up until now.

Keywords: DENV, Serotype, Epitope, Comparative genomics

Introduction

Dengue virus or as commonly called DENV is a single stranded RNA virus that infects approximately 390 million people each year, putting more than two-fifth of the world's population under the threat of this efficacious virus. The dengue fever has, thus become one of the most widespread disease. The virus belongs to the family *Flaviviridae* and genus *Flavivirus* . DENV is an arbovirus, having two known mosquito vectors *Aedes aegypti* and *Aedes albopictus* . The positive stranded RNA genome of dengue virus is of 10.7 Kb size and composed of three structural proteins (Envelope, Capsid, Membrane) and seven non-structural proteins (NS1, NS2A, NS2B,NS3, NS4A, NS4B, NS5). There are atleast four serotypes and they show 65% similarity in the genome structure.

The dengue infection is caused by one of the four serotypes of DENV that are spread by *Aedes* mosquito. During primary infection, the body develops immune responses in the form of antibodies against the particular serotype attacked. But the main complexity of DENV arises during the secondary infection with another serotype, leading to serious version of dengue infection like Dengue Haemorrhagic fever (DHF) and Dengue Shock Syndrome (DSS). This is caused due to the antibodies produced during primary attack which complicate the secondary DENV infection by a phenomenon known as Antibody Dependent Enhancement (ADE). During ADE, there is a cross reaction between the antibodies of the primary infection and virus of secondary infection such that there is an increased infection in macrophages and monocytes. These challenges bring the importance

Mansi Verma
Sri Venkateswara College, Delhi University, India, E-mail: mansiverma20@gmail.com

of an archetypal dengue vaccine which can provide life time immunity against all the serotypes.

Currently, the vaccine candidates that are under various stages of clinical trial are the live attenuated viruses, chimeric vaccine, recombinant vaccine with adjuvants, reverse vaccinology, purified and inactivated virions, subunit proteins and plasmid DNA. Among these, live attenuated DENGVAXIA or CYD-TDV, a tetravalent chimeric dengue vaccine, developed by Sanofi Pasteur in December 2015, is the first licensed vaccine in some Asian and Latin American countries. These clinical manifestations caused by the vaccine are ascribed to inefficiency of the vaccine in producing competent T- cells that protect against DENV disease. Moreover, the vaccine does not encode any non- structural proteins which are required by the virus to evade immune response of the host. All these studies imply that a vaccine that is tetravalent and simultaneously prevents antibody- dependent enhancement (ADE) needs to be designed urgently. These concerns led to the need for a relatively new technique of vaccine development i.e. Epitope or synthetic peptide based vaccines. As DENV has both structural and non-structural proteins for its viral activity, conserved epitopes may prove to be useful in designing synthetic peptide based vaccine. This can be easily initiated in today's time, as there is no dearth of information about genome sequences in the databases.

Materials and Methods

A sum total of 23,622 partial sequences of structural and non-structural proteins of dengue virus were retrieved from NCBI. After retrieving, the sequences were aligned using multiple sequence alignment program CLUSTAL_X. These aligned sequence files were then used for further analysis.

Sequences were used for detecting the species-specific signature sequences or **motifs**. Motifs were obtained for each serotype of DENV individually using an online tool multiple em for motif elicitation or MEME *Suite* . Motifs common for all serotypes were also obtained using the same method. In order to get a maximum number of motifs, the default setting was adjusted from 3 motifs to 10 motifs.

The motifs were then analyzed for the presence of B cell epitopes. The linear B cell epitopes were found using BCPRED and BEPIPRED tools of immune epitope database with default settings. Bepipred used Hidden Markov model for the prediction of B cell epitopes.

The immunogenicity of each epitope was checked using Kolaskar and Tangaonkar antigenicity method with a default threshold value 0.9. Hydrophilicity of the antigenic epitopes, required to check the accessibility were found using Parker Hydrophilicity method at a threshold value of 3.448. Epitopes were checked for their surface accessibility using Emini surface accessibility method with a threshold value of 1.00. Flexibility and Beta turns were checked using Karplus and Schulz Flexibility and Chou and Fasman Beta-turn methods respectively, with a threshold of 1.00 for both. Conservancy of epitopes was checked using Epitope Conservancy Analysis tool.

Models of the 3D structures of DENV proteins were downloaded from RCSB PDB server for mapping the epitopes. Then the location of predicted epitopes in the 3-D model was found using CHIMERA visualization tool. For DENV proteins where no 3D strcture was available, I-TASSER server was used to predict the 3-D structures. The predicted models were saved in pdb format files, which were later used to generate Ramachandran plot using PDBsum –PROCHECK software. This was done to verify the models generated by I-TASSER.

Results

In the present study, comparative genomics was performed on DENV sequences to find novel vaccine targets. For this purpose, all the ten proteins (3 structural and 7 non-structural) of DENV-1, DENV-2, DENV-3, DENV-4 were analyzed individually by retrieving their partial protein sequences and motif analysis was performed by using MEME SUITE. These conserved motif sequences were then used for B-cell epitope analysis. For the matter of space and clarity, we are reporting the figures for one of the conserved epitopes in the main text, while the others are presented in the supplementary materials, Supplementary data).

In total 13 epitopes were found which were further checked for antigenicity, hydrophilicity and surface accessibility. The B-cell epitopes on native proteins are generally composed of hydrophillic amino acids on the protein surface that are topographically accessible to membrane-bound or free antibody. These epitopes tend to be located in flexible regions of an immunogen and display site mobility of epitopes which maximizes complementarity with the antibody's binding site, permitting an antibody to bind with an epitope that it might bind ineffectively if it were rigid. Surface accessibility and hydrophilicity of these predicted epitopes was therefore determined. The flexibility of these epitopes and the presence of beta turn in their structure were also evaluated. Predicted epitopes of Envelope, NS1, NS3, NS4A, NS4B and NS5 proteins were found out to be antigenic as their score was found out to be higher than the threshold value whereas, few antigens were found to be non-antigenic. All the predicted epitopes were found to be flexible in nature.

Discussion

Dengue is one of the most rapidly spreading mosquito-borne viral disease and has emerged as one of the biggest threats to public health. There are four serotypes of DENV which show 65% similarity at genomic level and share same epidemics. Yet the difference in their interactions with antibodies is enough to make the development of a common vaccine, a mammoth task. The urgent need of dengue vaccine development is presented by the alarming rise in the dengue endemics. Although various vaccines are undergoing clinical trials worldwide but presently, there is no specific dengue therapeutics or vaccine available which provides overall protection against this viral infection. Recently, the only licensed DENV vaccine Dengvaxia, has been reported as a failure as it lacks DENV non-structural protein antigens, and is found to be incompetent to raise antibodies against NS1.

In the present study, Comparative Genomics and immunoinformatics was used as a tool to explore the potential candidates of dengue virus to find novel drug and vaccine targets. The advantage of development of an epitope based vaccine over other vaccines is its ability to induce specific immune response without undesirable effects. Also, both time and expenditure needed to screen a large number of epitopes can be saved using such computational approaches. Similar studies have been conducted on Human *Coronavirus* , *Saint Louis encephalitis virus* , Rotaviruses, H1N1 influenza A virus strains and Zika virus, however, the analysis has been conducted for only some selective proteins. Therefore our study is the first one to deal with all the proteins of any viral genome for designing an efficient vaccine.

Conclusion

Dengue Virus (DENV) has emerged as a potential threat to human health worldwide. The therapeutics against DENV are either not available in major parts of the world or if available in some endemic countries, are found to be inefficient. In this study we used computational approaches to find novel vaccine targets. It focuses to explore B-cell epitopes for all serotypes for each protein of dengue virus. We found six highly conserved epitopes in all serotypes of DENV [Envelope protein (**RCPTQGE**); NS3 (**SAAQRRGR, PGTSGSPI**); NS4A (**QRTPQDNQL**); NS4B

Mansi Verma
Sri Venkateswara College, Delhi University, India, E-mail: mansiverma20@gmail.com

(**LQAKATREAQKRA**) and NS5 proteins (**QRGSGQV**)]. Thus, our results suggest any of these proteins can be targeted to stimulate a specific immune response against all serotypes of DENV. These predicted epitopes would be the candidate target for the universal multi-subunit vaccine.

Nevertheless, further studies are needed to confirm the utility of these epitopes.

MV and RL conceptualized the work; MV, SB, KK, NM, SS,SG curated data; formal analysis was done by MV, PSD; MV, PSD were involved in fund acquisition and project administration; MV supervised the work; MV. SB, KK were involved in writing original draft; Review and editing was done by MV, PSD, RL.

This work is partly presented at 9th International Conference on Bioinformatics & System Biology Singapore City, Singapore March 20-21, 2019