OPEN ACCESS

Saudi Journal of Medicine
Abbreviated Key Title: Saudi J Med
ISSN 2518-3389 (Print) | ISSN 2518-3397 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

In silico Evaluation of SNPs and Molecular Modelling Study in Dengue

Lingaraj Murmu[1], Krishna Kumar Das[1], Santosh Kumar Behera[2*]

[1]BIF Centre, Department of Bioinformatics, Centre for Post Graduate Studies, Orissa University of Agriculture and Technology, Bhubaneswar-751003, India
[2]National Institute of Pharmaceutical Education and Research, Ahmedabad, Gujarat, India

DOI: 10.36348/sjm.2023.v08i03.004 | **Received:** 13.01.2023 | **Accepted:** 20.02.2023 | **Published:** 09.03.2023

***Corresponding Author:** Santosh Kumar Behera
National Institute of Pharmaceutical Education and Research, Ahmedabad, Gujarat, India

Abstract

Dengue may be a spectrum of disease caused by 5 serotypes of the foremost current arthropod- borne virus touching humans nowadays. The term dengue came into general use solely once 1828. Dengue viruses (DV) belong to family Flaviviridae and there are 5 serotypes of the virus observed as DV-1, DV-2, DV-3, DV-4 and DV5. DV could be a positive-stranded encapsulated polymer virus and consists of 3 structural macromolecule genes that write in code the nucleocapsid or core (C) macromolecule, a membrane-associated (M) macromolecule, associate degree enclosed (E) conjugated protein and 7 non-structural (NS) proteins. It is transmitted mainly by *Aedes aegypti* mosquito and also by *Ae. Albopictus.* The present strategy of bioinformatics analysis is to exploit the current data available both on gene and genome-wide association study (GWAS) meta-analysis of dengue to integrate these at novel levels of understanding of gene network interactions and expression levels. The study revealed 2 studies of dengue from GWAS with a total of 2 unique genes, namely MICB and PLCE1 that were mapped to discrete genomic locations of human genome which represented 15 rsIDs (SNPs) associated with MICB gene and 5 rsIDs (SNPs) associated with PLCE1 gene. The consensus results of the online tools like SIFT ,SNP & GO, PANTHER , Ploy phen2.0 and I- mutant for structure and functional studies depicted rs1051788, rs1051788, rs41293883 and rs45583740 of MICB gene to be deleterious/ diseased and effecting the structure and function of MICB gene which may the reason for occurrence of dengue in human counterpart. The DAVID bioinformatics functional enrichment analysis reported 234 genes and 468 GO terms for biological processes (BP).In this study a total of 307 genes pertaining to dengue and its associated diseases were mined from various databases like GWAS, GEO & Jeans lab disease database. Out of 307 genes only 22 genes were consider for interaction study with different drug molecules as they are treated as the key factors that play a vital role in dengue.

Keywords: *Aedes Aegypti*, MICB, PLCE1, SNPs, GEO, Molecular modelling.

INTRODUCTION

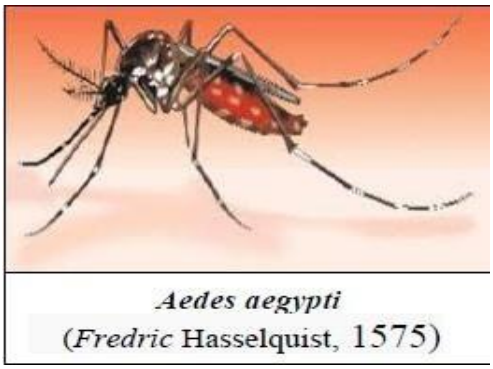
Dengue is the world's most important arboviral disease in terms of number of people affected. Virus is tiny agents that can infect a variety of living organisms, including bacteria plants and animals. Like other viruses, the dengue virus is a microscopic structure that can only replicate inside a host organism. Viral disease known as viral infection which occurs when organism's body invaded by pathogenic viruses, infectious virus particles that consists of an RNA or DNA core with protein shell called as Capsid. Dengue is an acute systemic viral disease that has established itself globally in both endemic and epidemic transmission cycles. The *Aedes aegypti* mosquito can transmit the viruses that cause Dengue fever and it's also called as *Yellow fever*. *Yellow fever* is a viral infection transmitted by bites from infected mosquitoes. Dengue virus (DENV) is the most life threatening disease prevalent arthropod-borne or mosquito borne viral disease in humans, this disease used to be called "break-bone" disease is now as significant problems in many countries with potential fatal complications caused by five dengue virus serotype (DENV -1, DENV-2, DENV-3, DENV-4 and 5).

Dengue viruses are transmitted from person to person mainly by breeding of (*Female Aedes aegypti* mosquito and also by *Ae. Albopictus*) in the domestic environment. Other members of the same genus include *West nail Virus*, *St. Louis encephalitis virus*, *Japanese encephalitis virus*, *tick-borne encephalitis virus*, *Kyasanur forest disease virus*, and *Omsk hemorrhagic*

fever virus. The causative dengue viruses are members of the genus *Flavivirus*, within the family *Flaviviridae*. DENV is a positive-stranded encapsulated RNA virus and is composed of three structural protein genes. The RNA genome of dengue virus is about 0.7 kb and which encode the nucleocapsid or core (C) Protein, a membrane- associated (M) protein, an enveloped(E) glycoprotein, and seven non-structural (NS) proteins (NS1, NS2a, NS2b, NS3,NS4a, NS4b and NS5) are encoded by viral genome. The incidence of dengue is currently increasing dramatically and it is now one of the diseases said to be re-emerging.

Classical dengue fever (DF) disease is a sudden onset of severe headache, chills, pain upon moving the eyes, and low backache and painful aching in the legs and joints severe pain is known as "break-bone fever or bone-crusher disease" occurs during the first hours of illness. In this stage the temperature raises quickly as high as 40º C, with low heart rate and low blood pressure and also some associated symptoms like abdominal pain, nausea, vomiting. Dengue Hemorrhagic Fever (DHF) is the flu-like symptoms caused by same viruses and it's signalized by increased vascular permeability, hypovolaemia and abnormal blood clotting mechanisms. It's very deadly complications with symptoms but after several days patient becomes irritable, restless and sweaty. The DHF continuously present for two to seven days and temperature high as 41º C. Dengue Shock Syndrome (DSS) is a dangerous complication of dengue infection and it is associated with high mortality. In this stage rapid and weak pulse and narrow pulse pressure (<20mm Hg) and hypertension and cold, clammy skin and restlessness, internal bleeding, organ failure, and death may occur. The disease, mostly during and shortly after the rainy season predominantly found in tropical and subtropical areas. The suitable temperature for development of the *Aedes* mosquito ranging from 15 to 30° C, at this temperature face lower mortality rate. Dengue is characterized by a sudden onset of high fever (103°F -106°F).

The primary vector of dengue is mosquito, *Aedes aegypti* and transmission may also cause by *Ae. albopictus*. Through the bites of infected female mosquitoes, the virus is transmitted to humans. Female mosquito feeds on blood because it needs protein for laying eggs. After incubation of virus for about 4 to 10 days, an infected mosquito is capable of transmitting the virus to the host for the rest of its life. The main carriers and multipliers of the virus are infected humans. The mosquito Ae. *aegypti* found in urban habitats and breeds commonly in man-made containers like flower vases, water storage jars, unused toilets bowls and chocked roof gutters.



Aedes aegypti
(*Fredric* Hasselquist, 1575)

Kingdom	: Animalia
Phylum	: Arthropoda
Class	: Insecta
Subclass	: Dipetra
Family	: Culicidae
G enus	: Aedes
Species	: A.aegypti

Aedes aegypti is a mosquito that can spread the dengue fever, Chikungunya and yellow fever viruses, and other diseases. The mosquito is a small, dark mosquito of approximately 4 to 7 millimetres with typical white markings on the legs and a marking of the form of a lyre on the thorax. It dwell in tropical and subtropical regions all over the world, mainly between the latitudes of 35°N and 35°S where the winter temperature is no colder than 10°C. They require a warm climate; they typically do not live at altitudes above 1000 m, where the temperature is colder. These mosquitoes are associated with the living spaces of humans.

MATERIALS AND METHOD

Mining of Genes Associated with Dengue from GWAS Catalog

Disease search for "Dengue" with a p-value threshold of p<10-5was performed to retrieve GWAS studies on Dengue from GWAS Catalog (http://www.genome.gov/gwastudies/ currently https://www.ebi.ac.uk/gwas/).A total of 2 GWAS studies resulted in a total of 2 unique genes mapped to discrete genomic locations of human genome.

GEO Micro Array

Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) data GSE51808 was utilized to find out differentially expressed genes associated with Dengue virus. GSE51808 dataset contain 56 samples. The dataset was divided into three groups, i.e. Convalenscent Vs Healthy Control Dengue,

Hermorrhagic Fever Vs Healthy Control, and Dengue Fever Vs Healthy Control. Geo2R application of GEO was utilized to find differentially expressed genes.

Jeans Lab Disease Database

Jeans lab disease database integrates the results from text mining with manually curated disease– gene associations, cancer mutation data, and genome-wide association studies from existing databases. Information about genes with are associated with dengue disease are retrieved from Jeans lab disease database.

Functional Annotation and GO Association of Dengue Genes by DAVID

The functional annotation of a total genes that have been mined from GWAS cataloge, GEO database and Jeans lab disease database, was performed through Gene Ontology (GO) analysis which describes the functions along the three categories viz., molecular functions (MF), biological processes (BP) and the cellular components (CC). The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used for GO term annotation (i.e., the common vocabulary for the functional description of genes and gene products) annotation.

Analysis of Natural Variant

Natural variants are otherwise called as Single Nucleotide Polymorphisms (SNPs) that is a variation in a single nucleotide that occurs at a specific position in the genome, where each variation is present to some appreciable degree within a population. Information about the natural variants of the genes was retrieved from Uniprot database. The rsIDs and its respective mutated/substitued position were being identified and pursued for analysis of the respective mutated position using SIFT ,SNP & GO, PANTHER , Ploy phen2.0 and I-mutant to find out the effect of mutation on the protein function and structure.

HOPE Server

HOPE is an easy-to-use web service that analyses the structural effects of a point mutation in a protein sequence. In this server, protein sequence and the position of the mutation of genes are submitted. This server collect and combine available information from a series of web services and databases and produces a report, complete with results, figures and animations about the effect of mutation on the structure and function of the protein.

Generation of Gene Network and its Interactions using STRING Database

Gene networks present a graphical view at the level of gene activities and genetic functions and help us to understand complex interactions in a meaningful manner. The STRING database aims to provide such a global perspective for as many organisms as feasible. Known and predicted associations are scored and integrated, resulting in comprehensive protein networks covering >1100 organisms. The key and hub genes have been identified based on the interaction found in protein-protein interaction network.

Gene-Disease Association Study through WebGestalt

WebGestalt (WEB-based Gene Set Analysis Toolkit) , one of the first software applications that integrate functional enrichment analysis and information visualization for the management, information retrieval, organization, visualization and statistical analysis of large sets of genes. In addition to significant data expansion, WebGestalt has also improved user friendliness and added new visualization features that help users better understand the enrichment results. WebGestalt was used for gene–phenotype association, gene–disease association and Drug association analysis.

Retrieval of Drugs and Proteins (Corresponding Targets)

The Structure Data Format (SDF) 3D structure of the reported drugs was retrieved from the NCBI PubChem database along with its PubChem ID, Molecular weight and Molecular formula. The compounds were converted into pdb format structure using the PyMol (academic version) tool, Discovery Studio v4.1 visulizer tools and online SMILES translator web server.

UNIPROT

The Universal Protein Resource (UniProt) is a comprehensive resource for protein sequence and annotation data. The corresponding protein sequences encoded by these genes were retrieved from UniProtKB database. The structures of proteins which are not present in the database, derived by modelling the structure using MODELLER tool.

Model Refinement, Evaluation and Structure Validation

The structure validations of proteins were validated using various web servers like WhatIF, used to refine the structure. Quality of the generated model was evaluated with Procheck by Ramachandran plot analysis. Stereochemical quality and accuracy of the selected models was further improves by subjecting to energy minimization with GROMOS96 43b1 parameter set, implementation of swiss pdb viewer. Validation of generated model was further performed by Varify3D and ERRAT program. ProsA was used for analysis of Z score and energy plot. The three dimension structures of the modeled proteins were analyzed using deep view swiss pdb viewer.

Prediction of Binding Site

Structural and active site studies prediction of the proteins were done by using CASTP (Computed Atlas of Surface Topography of Proteins).

Docking Approach

AutoDock 4.2 was used for docking studies which is widely distributed public domain molecular docking software. The docking analysis was carried out for the reported drugs (can be said as ligands) with their corresponding targets (proteins) using AutoDock4.2 tool. The interactions of ligand and proteins were studied using Visulizer tools and softwares. The various bonding interactions of ligand and proteins were explored using the above tools.

RESULTS AND DISCUSSION

Genome Wide Association Studies

The GWAS reported two studies of dengue with a total of two unique genes and the results are represent in Table 1, namely MICB and PLCE1 that were mapped to discrete genomic locations of human genome.

Table 1: GWAS Disease gene result

Sl No.	Reported Genes	RsID
1	MICB	rs3132468
2	PLCE1	rs3765524

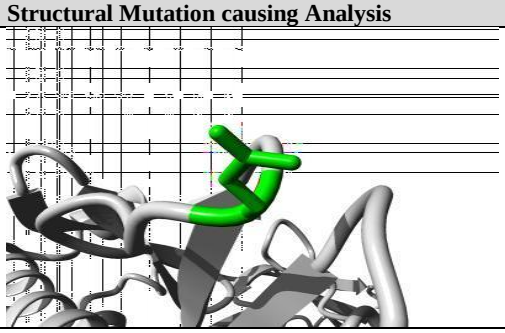
MICB and PLCE1 Natural Variants Analysis

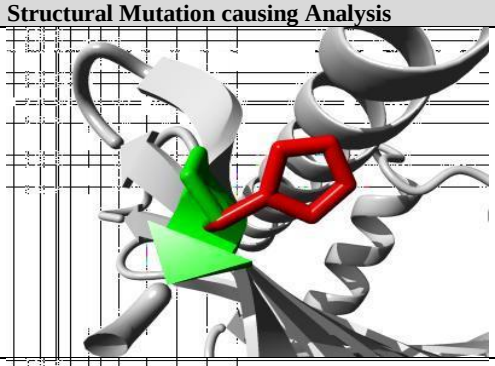
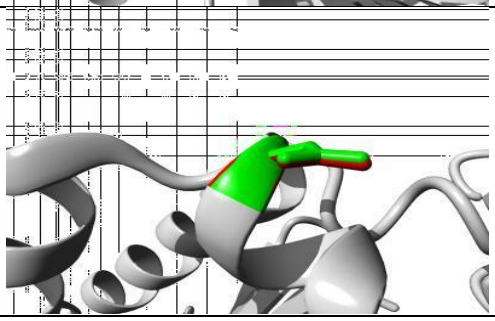
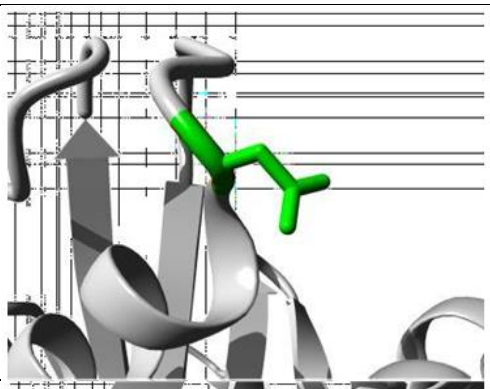
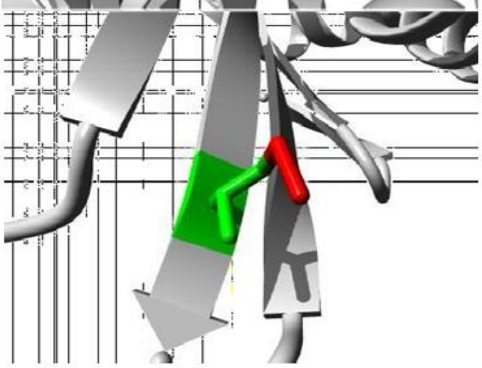
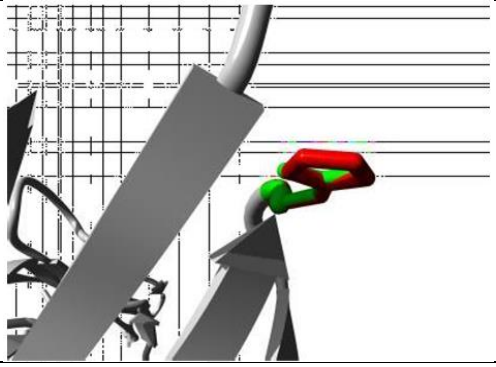
The rsIDs of the above said genes were retrieved from Uniprot database. The results represented 15 rsIDs (SNPs) associated with MICB gene and 5 rsIDs (SNPs) associated with PLCE1 gene. The effect of mutation/substitution on the using various protein function and structure was analyzed online tools like SIFT, SNP & GO, PANTHER, Ployphen2.0 and I-mutant to find out the consensus results. The consensus result depicted rs1051788, rs1051788, rs41293883 and rs45583740 of MICB gene to be deleterious/ diseased and effecting the structure and function of MICB gene that may result in the dengue. The result was represented in Table 2.

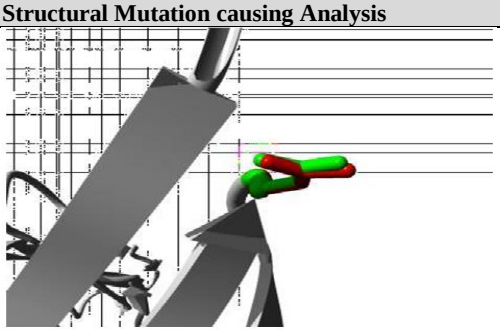
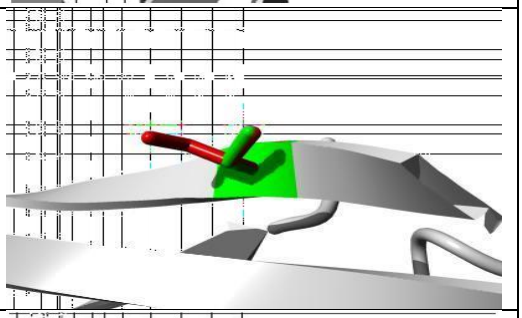
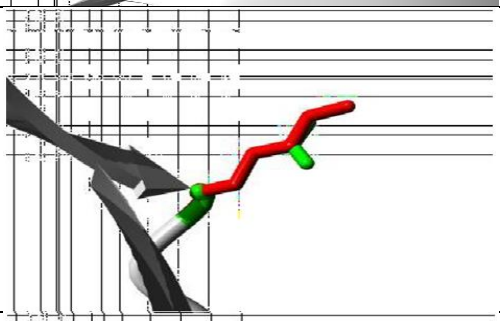
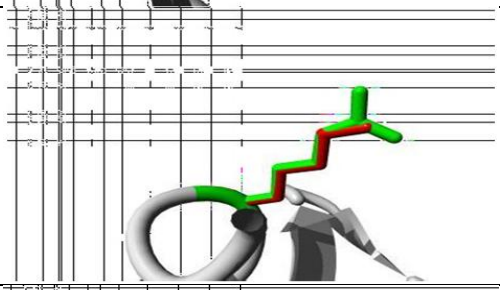
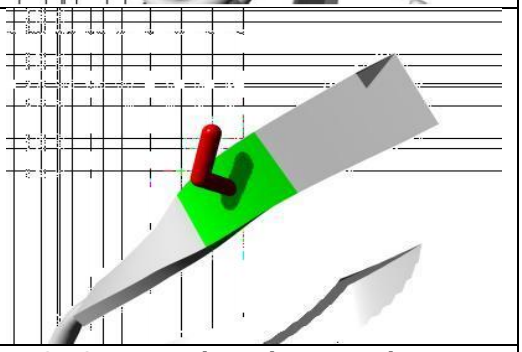
Table 2: Effect of mutation on *MICB* and *PLCE1* gene

Sl. No.	*MICB* rsID	Substitution	SIFT	PANTHER	SNPs & GO	Polyphen-2	I-Mutate2.0
1	rs45578846	E39G	Deleterious	Deleterious	Neutral	Deleterious	Decrease
2	rs45583740	P68H	Deleterious	Deleterious	Disease	Deleterious	Decrease
3	rs3131639	D75N	Deleterious	Deleterious	Neutral	Deleterious	Decrease
4	rs1065075	K80E	Deleterious	Deleterious	Neutral	Deleterious	Decrease
5	rs45486091	D88G	Deleterious	Deleterious	Neutral	Deleterious	Decrease
6	rs45502297	D105G	Deleterious	Deleterious	Neutral	Deleterious	Decrease
7	rs3134900	I121M	Deleterious	Deleterious	Neutral	Deleterious	Decrease
8	rs1051788	D136H	Deleterious	Deleterious	Disease	Deleterious	Decrease
9	rs1051788	D136N	Deleterious	Deleterious	Disease	Deleterious	Decrease
10	rs41293883	T212I	Deleterious	Deleterious	Disease	Deleterious	Decrease
11	rs45624537	E215K	Deleterious	Deleterious	Neutral	Deleterious	Decrease
12	rs45587032	R279K	Deleterious	Deleterious	Neutral	Deleterious	Decrease
13	rs41273040	G291S	Deleterious	Deleterious	Neutral	Deleterious	Decrease
14	rs45470602	V300A	Deleterious	Deleterious	Neutral	Deleterious	Decrease
15	rs1065076	A383T	Deleterious	Deleterious	Neutral	Deleterious	Decrease
	PLCE1 rsID						
1	rs17417407	R548L	Deleterious	Deleterious	Disease	Deleterious	Decrease
2	rs2274224	R1575P	Deleterious	Deleterious	Neutral	Deleterious	Decrease
3	rs3765524	T1777I	Deleterious	Deleterious	Neutral	Deleterious	Increase
4	rs2274223	H1927R	Deleterious	Deleterious	Neutral	Deleterious	Increase
5	rs17508082	S469T	Deleterious	Deleterious	Neutral	Deleterious	Increase
6	rs17417407	R548L	Deleterious	Deleterious	Disease	Deleterious	Decrease

Table 3: Effect of mutation/substitution on *MICB* and *PLCE1* gene obtained from HOPE server

Sl No	MICB gene Substi- tution Position	HOPE Server Results	Structural Mutation causing Analysis
1	E39G	Inheritable Diseases	

Sl No	MICB gene Substi- tution Position	HOPE Server Results	Structural Mutation causing Analysis
2	P68H	Inheritable Diseases	
3	D75N	Inheritable Diseases	
4	K80E	Inheritable Diseases	
5	D88G	Inheritable Diseases	
6	D105G	Inheritable Diseases	
7	I121M	Inheritable Diseases	
8	D136H	Inheritable Diseases	

Sl No	MICB gene Substi- tution Position	HOPE Server Results	Structural Mutation causing Analysis
9	D136N	Inheritable Diseases	
10	T212I	Inheritable Diseases	
11	E215K	Inheritable Diseases	
12	R279K	Inheritable Diseases	
13	G291S	Inheritable Diseases	
14	V300A	Inheritable Diseases	No 3D-Structure and Template is Found
15	A383T	Inheritable Diseases	No 3D-Structure and Template is Found
PLCE1gene Substitution Position		**HOPE Server Results**	**Structural Mutation causing Analysis**
16	S469T	Inheritable Diseases	No 3D-Structure and Template is Found
17	R548L	Inheritable Diseases	No 3D-Structure and Template is Found
18	S1484L	Inheritable Diseases	No 3D Structure and Template is Found
19	R1575P	Inheritable Diseases	No 3D Structure and Template is Found
20	T1777I	Inheritable Diseases	No 3D Structure and Template is Found
21	H1927R	Inheritable Diseases	No 3D Structure and Template is Found

GEO Analysis

The GEO database represented a total of 307 genes after removal of duplicity which seems to be differentially expressed during Dengue infections like convalenscent, hermorrhagic fever, dengue fever & dengue shock syndrome. Result from Jeans lab disease database suggests association of 134 genes with dengue infection.

DAVID Gene Ontology

The DAVID bioinformatics functional enrichment analysis reported 234 genes and 468 GO terms for biological processes (BP). Based on the essential role of biological processes, 234 genes and 468 GO terms obtained from significantly enriched biological processes are termed as key genes that were used for network construction of Dengue.

Protein-Protein Interaction using STRING

Protein-protein interaction of genes from Convalescent vs Healthy control dengue sample (93 genes).

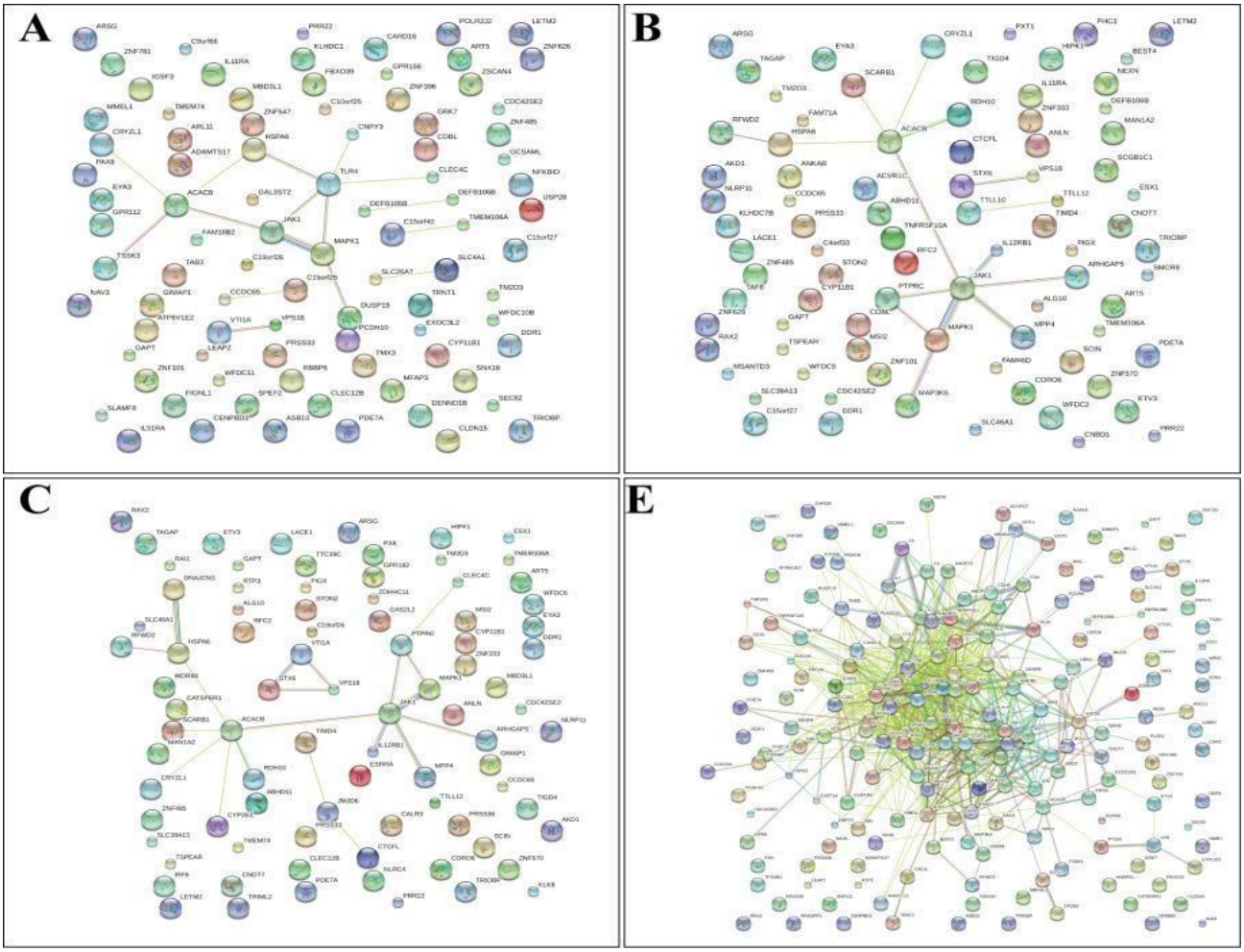


Fig. 1: Network construction of dengue analyzed through STRING database (A) Convalenscent Vs Healthy Control Dengue (B) Hermorrhagic Fever Vs Healthy Control (C) Dengue Fever Vs Healthy Control (D) Dengue Shock Syndrome

Hermorrhagic fever vs Healthy control sample (91 genes), Dengue fever vs Healthy control dengue (93 genes) and Dengue shock syndrome samples (134 genes) were generated using STRING database.

From the protein-protein interaction study 27 genes, namely TNF, ALB, F2, CD209, CD40LG, CD83, CEBPB, DDR1, CCL2, F3, ICAM1, IFNAR2, IFNG, IL10, IL4, IL6, IL8, IRF9, VCAM1, KAT2B, MAPK1, SYT1, PTPRC, PLAT, SCARB1, JMJD6, FGA show strong interaction among each other.

Web Gestalt represented various graphs from the input of 216 BP genes like BP-BP graph, BP- CC graph and BP-MF graph. The charts are represented in (Figure 2). All the three graphs represented differential states of expression of genes in various processes, locations and functionalities. The PHEWAS (Phenome wide association studies) was carried out in Web Gestalt by considering PHEWAS BP-BH (Top 10) and BP-Bonferoni of 234 genes. Both the graphs represented variations among them this may be due to difference in statistical terms (Figure 3).

The Drug association analysis of WebGestalt has reported 21 drugs interacted with 27 genes and its corresponding proteins. The results of WebGestalt pertaining drugs against dengue and its corresponding genes/proteins were cross checked by literature survey, substantially presented in Table 4.

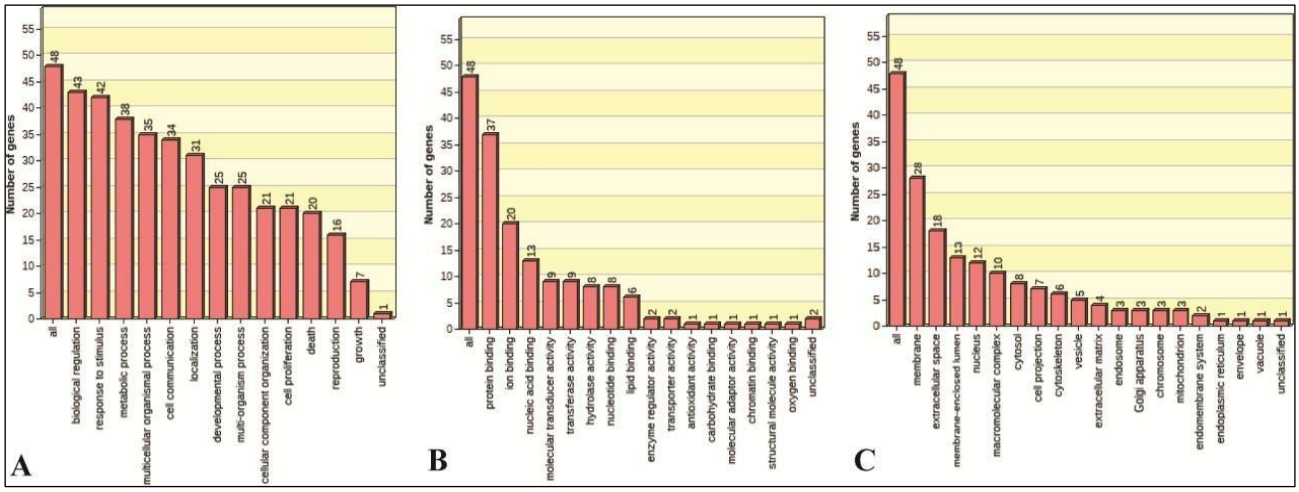


Fig. 2: (A) WebGestalt BAR Chat of BP-BP Graph of 234 Genes (B) WebGestalt BAR Chat of BP-MF Graph of 234 Genes

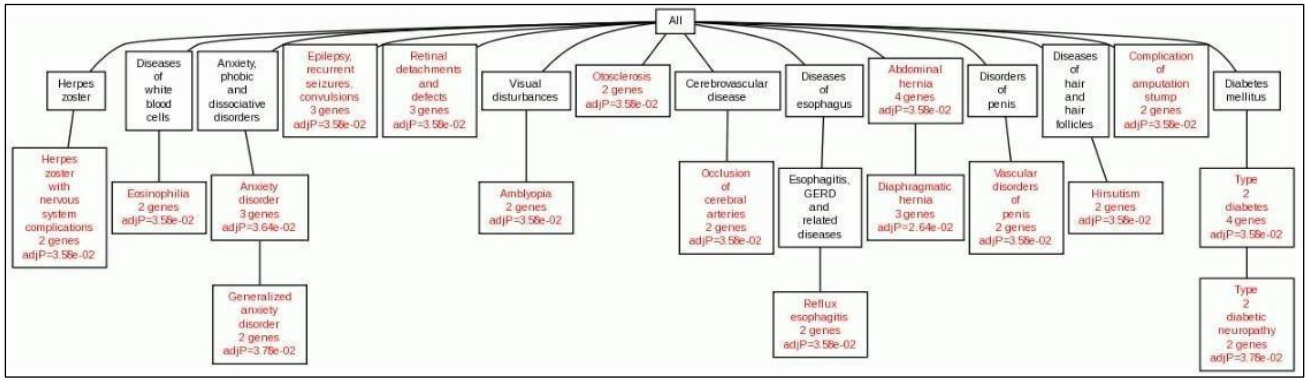


Fig. 3: WebGestalt PHEWAS BP-BH (Top-10) of 234 genes

Out of 27 reported genes, the structures of the 24 genes and its corresponding proteins were retrieved from Protein data bank. Whereas the remaining 3 namely CD83, IRF9 and SCARB1 were modelled using Modeller9.15.The results are represented in Table 4, which reports the Uniprot Id , PDB Id, Full name and short name of the 27 targets along with the region considered for *in silico* docking studies.

Table 4: Protein and its UniProt Id & PDB ID

SL No	Target	UniProt Id	PDB Id	Short Name of Target	Full name of Target	Region
1	TNF	P01375	4TSV	TNF-a	or necrosis factor	84-233
2	IFNG	P01579	1FYH	IFN-gamma	Interferon gamma	28-156
3	PTPRC	PO8575	5FN7	L-CA	Receptor-type tyrosine-protein phosphatase C	223-392
4	CD40LG	P29965	1ALY	CD40-L	CD40 ligand	116-261
5	CD209	Q9NNX6	2XR6	Dendritic cell-specific ICAM-3-grabbing non-integrin 1	CD209 antigen	250-404
6	IL8	P10145	5D14	L-CA	Interleukin-8	30-99
7	IL4	P05112	2D48	IL-4	Interleukin-4	25-153
8	IL10	P22301	2ILK	IL-10	Interleukin-10	20-178
9	ICAM1	P05362	1IAM	ICAM-1	cellular adhesion molecule 1	28-212
10	CD83	Q01151	NA			
11	F2	P00734	5AFY	Coagulation factor II	Prothrombin	364-621
12	FGA	P02671	1BBR	Fibrinopeptide A	Fibrinogen alpha chain	26-35
13	PLAT	P00750	1RTF	t-PA	Tissue-type plasminogen activator	311-562
14	F3	P13726	4YLQ	TF	Tissue factor	33-251
15	CCL2	P13500	1DOK	MCAF	C-C motif chemokine 2	24-99

SL No	Target	UniProt Id	PDB Id	Short Name of Target	Full name of Target	Region
16	VCAM1	P19320	1VCA	V-CAM 1	Vascular cell adhesion protein 1	25-226
17	SCARB1	Q8WTV0	NA			
18	JMJD6	Q6NYC1	3IDB	Protein PTDSR	Bifunctional arginine demethylase and lysyl-hydroxylase JMJD6	1-334
19	SYT1	P21579	3F04	Syt I	Synaptotagmin-1	141-266
20	IL6	P05231	1ALU	IL-6	Interleukin-6	
21	DDR1	Q08345	4BKJ	Epithelial discoidin domain receptor 1	Epithelial discoidin domain-containing receptor 1	601-913
22	MAPK1	P28482	4ZZN	MAP kinase 1	Mitogen-activated protein kinase 1	11-360
23	IRF9	Q00978	NA			
24	CEBPB	P17676	1GU4	C/EBP beta	CCAAT/enhancer- binding protein beta	259-336
25	KAT2B	Q92831	5FE6	Histone acetylase PCAF	Histone acetyltransferase KAT2B	715-831
26	ALB	P02768	1N5U	NA	Serum albumin	25-609
27	IFNAR2	P48551	3S9D	IFN-alpha/beta receptor 2	Interferon alpha/beta receptor 2	37-232

Table 5: The structure validation scores of the modelled structures

Server		CD83	IRF9	SCARB1
PROCHECK	Most favored regions (%)	75.6	90.3	87.0
	Additional allowed regions (%)	21.7	9.7	12.7
	Generously allowed regions (%)	2.7	0.0	0.3
	Disallowed regions (%)	0.0	0.0	0.0
	Overall G-factor	0.02	0.22	0.12
Verify 3D	Averaged 3D-1D Score > 0.2 (%)	19.51	32.06	67.39
ERRAT	Overall quality	10.204	17.021	42.920

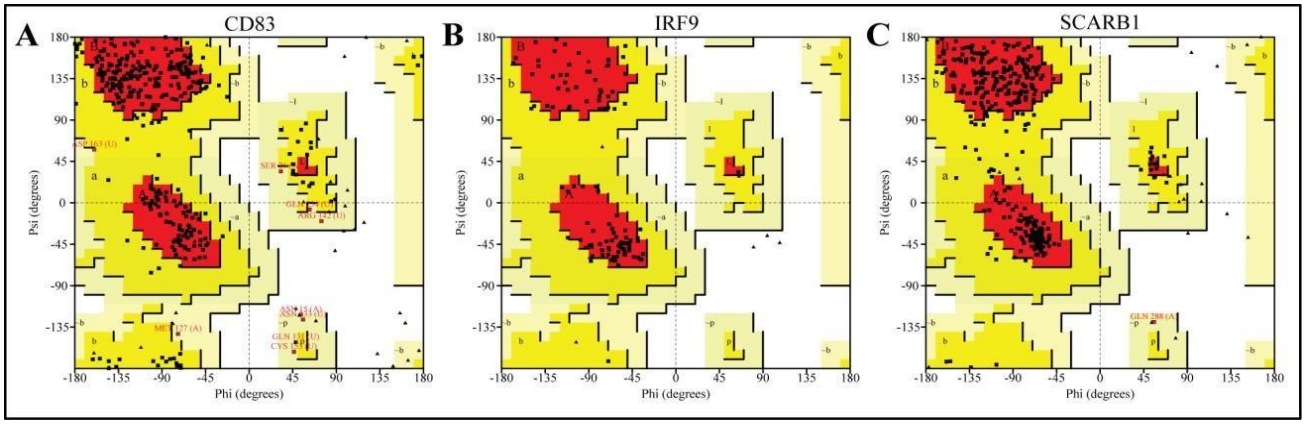


Fig. 4: Ramachandran PLOT of (A) CD83, (B) IRF9, (C) SCARB1

Prediction of binding site

CASTp server was used to identify the active site binding of the 27 proteins. Out of 27, only 22 proteins shows Active/Binding site. The result is represented in Table 6. Five structures, namely FGA, ALB, CD83, IRF9 and SCARB1 whose active sites were unpredictable need further simulation using GROMAC software.

Table 6: Result of CASTp for 22 proteins

SL No	Target Name	Binding Site Region
1	CD40LG	SER128,GLU129,ALA130,SER131,THR134,THR135,LEU138,GLN139,TRP140,PRO244,SER245 ,VAL247,HIS249,THR254
2	CD209	SER307,SER308,ASN311,ARG312,PHE313,PHE359,SER360,GL Y361
3	ICAM1	LEU91,ALA92,LEU94,THR105,LEU106,ARG107,TYR180
4	IFNG	PHE29,GLN48,ILE49,SER51,PHE52,PHE54,LYS55,PHE57,LYS58,LYS74,ARG89,PHE92,GLU9 3,LEU95,THR96,ASN97,TYR98,VAL100,GLN248,SER251,PHE252,PHE254,LYS255,PHE257,L YS258,LYS261,VAL270,GLU271,LYS274,ASN278,PHE282,ARG289,PHE292,GLU293,LEU295, THR296,ASN297,TYR298,VAL300
5	IL4	ILE32,PHE33,ALA35,LYS37,ASN38,THR39,THR44,ARG47,AR G115,ILE119,LYS123
6	IL8	LYS1,GLU2,LEU3,ARG4,CYS5,GLN6,ARG24,ILE26,HIS31,CY S32,ASN34,GLU36
7	IL10	LEU23,LEU26,ARG27,PHE30,VAL33,LYS34,PHE37,GLN38,ASP41,LEU47,LEU48,LS49,LEU5 2,LEU53,PHE56,LEU65,MET68,ILE69,PHE71,TYR72,VAL76,MET77,ALA80,VAL91,LEU94,L E U98,LEU101,ARG102,LEU105

SL No	Target Name	Binding Site Region
8	PTPRC	CYS94,THR95,GLN96,LYS99,THR103,ILE104,CYS105,LYS138, LEU139,GLU140
9	TNF	GLY66,GLY68,CYS69,PRO70,THR72,HIS73,VAL74,PRO100,C YS101,TRP114,TYR141
10	CCL2	SER27,TYR28,ARG29,ILE42,LYS44,GLU50
11	F3	SER1,GLY2,THR4,ASN5,THR6,VAL7,VAL30,ASN31,GLN32
12	VCAM1	VAL118,TYR119,PRO120,PHE121,ASP122,ALA144,ASP145,AR G146,SER148,GLU150
13	F2	LEU41,CYS42,HIS57,CYS58,TYR60,LEU99,ASP189,ALA190,CYS191,GLU192,GLY193,SER195,VAL213,SER214,TRP215,GLY216,GLY219,CYS220,GLY226,PHE227,TYR228
14	JMJD6	ALA16,ARG17,GLU19,LYS100,CYS101,GLY102,GLU103,SER128,PRO129,LEU130,TYR131,PHE133,SER135,TRP174,VAL176,GLY183,THR184,GLY185,ILE186,HIS187,ASP189,PRO190,ALA195,TRP196,ASN197,LYS204,TRP206,LYS219,VAL220,GLN229,ASP230,GLU231,TRP235, VAL267,HIS273,VAL275,ASN277,AL A283,THR285,ASN287
15	SYT1	VAL183,PHE184,LEU185,ASP188,LYS190,LYS191,LYS192,PHE193,PHE212,LYS213,VAL214
16	PLAT	PRO124,LEU128,GLN129,LEU130,PRO131,ASP132,LEU162,TYR163,PRO164,SER165,CYS168 ,VAL176,THR177,ASP178,MET180,LEU181,CYS182,VAL210,LYS230,VAL231,THR232
17	IL6	LEU33,ILE36,SER37,ARG40,THR43,CYS44,CYS50,GLU51,HIS164,LEU167,ARG168,LYS171
18	DDR1	LEU616,GLY617,GLU618,GLY619,GLN620,PHE621,GLY622,GLU623,VAL624,ALA653,VAL654,LYS655,ILE656,LEU657,ALA665,ASP668,PHE669,LYS671,GLU672,ILE675,MET676,LEU679,,ILE684,ILE685,LEU687,MET699,THR701,ASP702,TYR703,MET704,GLY707,ASP708,GLN711,LEU757,PHE762,VAL763,HIS764,ARG765,ASP766,LEU773,ILE782,ALA783,ASP784,PHE785,GLY786,MET787,SER788,ARG789,TYR796,ALA803,VAL804,LE U805,PRO806,PHE820
19	MAPK1	ILE29,GLY30,GLU31,GLY32,ALA33,TYR34,GLY35,MET36,VAL37,ALA50,LYS52,LYS53,ARG65,GLU69,LEU73,ILE82,GLN103,ASP104,LEU105,MET106,GLU107,THR108,ASP109,LYS112,ASP147,LYS149,SER151,ASN152,LEU154,CYS164,ASP165
20	CEBPB	LYS315,GLU318,ARG322
21	KAT2B	TRP746,PRO747,PHE748,GLU750,PRO751,VAL752,GLU756,ALA757,TYR760,TYR761,PRO767,MET768,ASP769,VAL795,ASN798,CYS799,TYR802,ASN803,TYR809
22	IFNAR2	SER8,LEU9,GLY10,SER11,ARG12,ARG13,LEU15,PHE27,LEU30,ARG33,ASP35,PHE36,GLY37,PHE38,GLN40,LEU42,PHE43,THR44,ILE45,SER47,LYS48,PRO49,GLU50,GLU51,THR52,ILE53,VAL54,LYS56,ILE60,ASP70,GLU71,ARG73,SER74,THR75,HIS76,GLU77,TYR79,VAL82,GLU84,GLU87,GLN90,GLN91,SER94,LEU95,SER96,CYS98,TRP100,ALA102,ILE103,ASN119,HIS120,PHE127,GLN136,PHE137,ASP138,LEU139,SER140,LEU141,GLU146,ARG149,SER152,LEU153,HIS154,LYS155,ASN156,GLU157,MET162,PHE166,THR167,TYR168,ILE169,ILE170,ASP171,L YS172,GLU186,HIS187

Protein-ligand Complex and Interaction studies

Protein-ligand Interactions were carried out using only 22 genes whose active sites were predicted using Castp.AutoDock 4.2 (autodock.scripps.edu/) that was used for docking studies revealed docking score with energy minimization values, Binding energy, and Ligand Efficiency, Inhibition Constant and Electrostatic energy for 21 ligands/drugs-22 potential targets interactions are represented at Table 7.

The result obtained from docking studies revealed that Cyclosporine and Heparin may act as potential drug for treatment of Dengue disease. The molecular docking studies have reported the drug-target interactions of IL10- cyclosporine was -12.24, F2-heparin was -19.41, F3- heparin was -14.41, PLAT-heparin was -16.37. F2-heparin has been identified with the highest docking scores with energy minimization and interactions of F3-Phosphatidylserine was -2.91 which is identified as the lowest docking scores with energy minimization out of 22 ligand- protein interactions.

Table 7: Docking analysis of 22 protein and 21 drugs using Autodock 4.2 software

SL No	Target	Drug	Binding Energy	Ligand Efficiency	Hydrogen Bonding	Hydrophobic	Electrostatic
1	CD40LG	Immunoglobulin	-3.37	-0.48	SER131,PRO244,VAL247	NA	NA
2	CD209	Immunoglobulin	-3.31	-0.47	ARG312,GLY361,ARG312,ARG312,ASN311,ASN311,SER308	NA	NA
3	ICAM1	Immunoglobulin	-2.77	-0.4	LEU91,SER177,ALA178,LEU91,GLU90	NA	NA
4	IFNG	Immunoglobulin	-3.62	-0.52	NA	NA	NA
5	IL4	Immunoglobulin	-3.33	-0.48	LYS102,ARG47,PRO100,GLU103, ASN38	NA	NA
6	IL8	Immunoglobulin	-3.7	-0.53	ARG24,GLN6,GLU46	NA	NA
7	IL10	Immunoglobulin	-5.02	-0.72	LYS57,GLU50,GLU54	NA	NA
8	PTPRC	Immunoglobulin	-3.17	-0.45	LYS138,GLU140	NA	NA
9	TNF	Immunoglobulin	-4.55	-0.65	LYS112,CYS69,PRO100,GLU116, TRP114	TRP114	NA

SL No	Target	Drug	Binding Energy	Ligand Efficiency	Hydrogen Bonding	Hydrophobic	Electrostatic
10	CCL2	Lovastatin	-4.42	-0.15	ARG29,THR10	CYS52	NA
11	F3	Lovastatin	-5.84	-0.2	ASN31,ASN5,	VAL7,VAL30	NA
12	ICAM1	Lovastatin	-7.06	-0.24	LEU94,LEU91	TRP97,TYR180	NA
13	VCAM1	Lovastatin	-6.59	-0.23	ILE88,TYR119,GLU66,PRO120, TYR119	ALA144	NA
14	F2	Phosphati dylserine	-2.19	-0.14	LYS60F,SER195,LEU41,LEU40	NA	NA
15	F3	Phosphati dylserine	-0.56	-0.04	SER1,ASN31,THR4,GLY2,ASN5	NA	NA
16	JMJD6	Phosphati dylserine	-2.07	-0.13	ASN197,THR285,THR285,ASP189	NA	TRP174
17	PTPRC	Phosphati dylserine	-3.36	-0.21	LYS138,THR103,GLU140	NA	NA
18	SYT1	Phosphati dylserine	-0.61	-0.04	LYS213	NA	NA
19	IFNG	Nitric oxide	-4.31	-1.08	PHE54	NA	NA
20	PLAT	Nitric oxide	-2.84	-0.71	TYR163,LEU181	LYS230	LYS230
21	TNF	Nitric oxide	-4.25	-1.06	HIS73,VAL74,CYS101,VAL74	HIS73	HIS73
22	VCAM1	Nitric oxide	-2.99	-0.75	PHE121,VAL118,SER148	NA	NA
23	IL6	Pentoxify lline	-4.73	-0.24	GLN175,ARG40,LYS171	ARG40,ARG168, LYS17, ARG40, LYS171	NA
24	IL8	Pentoxify lline	-4.64	-0.23	LEU3,ARG24,THR72	ILE26	NA
25	TNF	Pentoxify lline	-6.41	-0.32	GLN67,THR72,TRP114,TY R141	PRO70,CYS69, CYS101,HIS73,	NA
26	DDR1	Collagenase	-4.55	-0.18	VAL763,ARG765,ASP784, HIS764	PHE820,ARG765,	ASP784
						LEU805,PRO806,MET810,ARG 765,LEU805,PHE820	
27	IL8	Collagenase	-5.58	-0.21	ARG4,CYS5,LEU3	ILE8	NA
28	MAPK1	Collagena se	-3.21	-0.12	ALA33,TYR34,GLY32	ILE29,LEU105, MET106,LEU 154	NA
29	TNF	Collagena se	-6.15	-0.24	GLN67,ARG138,TYR141,GLY68, GLY66	PRO70,VAL74,CYS69,CYS101, HIS73	NA
30	IL8	Genisterin	-5.99	-0.3	ASN34,ILE37,LEU49,	CYS7,PRO51	GLU36
31	MAPK1	Genisterin	-7.55	-0.38	NA	NA	NA
32	IFNG	Cyclospo	-8.89	-0.1	ASN297,GLU293,ASP290,T	NA	ASP290,GLU
		rine			HR27		293
33	IL10	Cyclospo rine	-12.24	-0.14	NA	MET77,LEU94, LEU52,PHE30, PHE56	ASP41,TYR72
34	TNF	Cyclospo rine	-3.67	-0.04	NA	LEU75	NA
35	CEBPB	Dexamet hasone	-3.63	-0.13	THR326,GLU323,	ARG322	NA
36	IL6	Dexamet hasone	-4.92	-0.18	LYS171,GLN175,ASP34	NA	NA
37	IL8	Dexamet hasone	-5.18	-0.19	GLN6,GLU27	HIS31	NA
38	CD209	Ciproflox acin	-6.31	-0.26	ARG312,GLY361,ARG309,ASN362,SER308	ARG309	NA
39	KAT2B	Ciprofloxacin	-8	-0.33	ARG754,THR755,LYS753,TYR761	LYS753,TYR761, PRO751,ARG754	NA
40	F2	Aminoca proic acid	-6.08	-0.68	SER195,CYS191,ASP189,GLY219	NA	NA
41	PLAT	Aminoca proic acid	-3.68	-0.41	LYS230,MET180,SER165	NA	NA
42	CCL2	Prednisone	-6.65	-0.26	TYR13,GLU50,CYS11	CYS52	NA
43	IL10	Prednison e	-7.23	-0.28	TYR72,ASP41	VAL33,LEU94,PHE30,PHE37,TYR72	NA
44	TNF	Prednisone	-6.19	-0.24	LYS65,TYR141	PRO70,TYR141	NA
45	CCL2	Morphine	-6.12	-0.29	TYR13,CYS11,GLU50	CYS52	NA
46	IL6	Morphine	-4.38	-0.21	NA	NA	NA
47	F2	Warfarin	-9.68	-0.42	GLY193,ASP194,SER195,SER214	TRP60D,GLU192, CYS191,GLY216,	NA
						TRP215,ALA190	
48	IL6	Suramin	-6.33	-0.07	GLN111,TYR31,GLY35,SER37, ASP34	NA	ASP34,TYR31
49	F2	Heparin	-19.41	-0.54	ARG221,GLY226,ASP189,ASP222,TYR184,ASP221,GLY184,ARG187,GLY216, TYR225	NA	NA
50	F3	Heparin	-14.11	-0.39	ALA80,VAL30,GLN32,ASN5,VAL30,ALA80, THR6,PRO79	NA	NA

SL No	Target	Drug	Binding Energy	Ligand Efficiency	Hydrogen Bonding	Hydrophobic	Electrostatic
51	PLAT	Heparin	-16.37	-0.45	LEU130,SER165,TYR163,CYS182	NA	ASP132
52	IFNAR2	Ribavirin	-5.57	-0.33	SER23,GLN136,PHE137,GLU134, PHE21	LEU135,PHE21, LEU18	NA
53	IL10	Ribavirin	-3.5	-0.21	GLN79,GLU75,VAL76	VAL76	NA
54	CD40LG	Aspirin	-5.32	-0.41	SER131,THR135,SER245	NA	NA
55	F2	Aspirin	-6.29	-0.48	SER195,CYS191,GLU192	NA	NA
56	F3	Urokinase	-1.84	-0.03	ASN31,ALA80,ASN82,THR4,GLY 81	NA	NA
57	PLAT	Urokinas e	-7.37	-0.12	ASN233,ASP236,LEU130,ASP132	NA	ASP178,SP13 2,ASP236
58	F2	Sodium lauryl sulphate	-7.02	-0.41	SER195,GLU192,	TYR228,ALA190, VAL213	NA
59	CCL2	Atorvasta tin	-3.45	-0.08	GLU50,TYR28	ARG29,LEU67	NA

The compound (ligand)-target complex was performed for the best binding scores and its interaction studies was visualized in Discovery Studio Visualizer, LigPlot and PyMol visualize depicted in the (Figure 5).

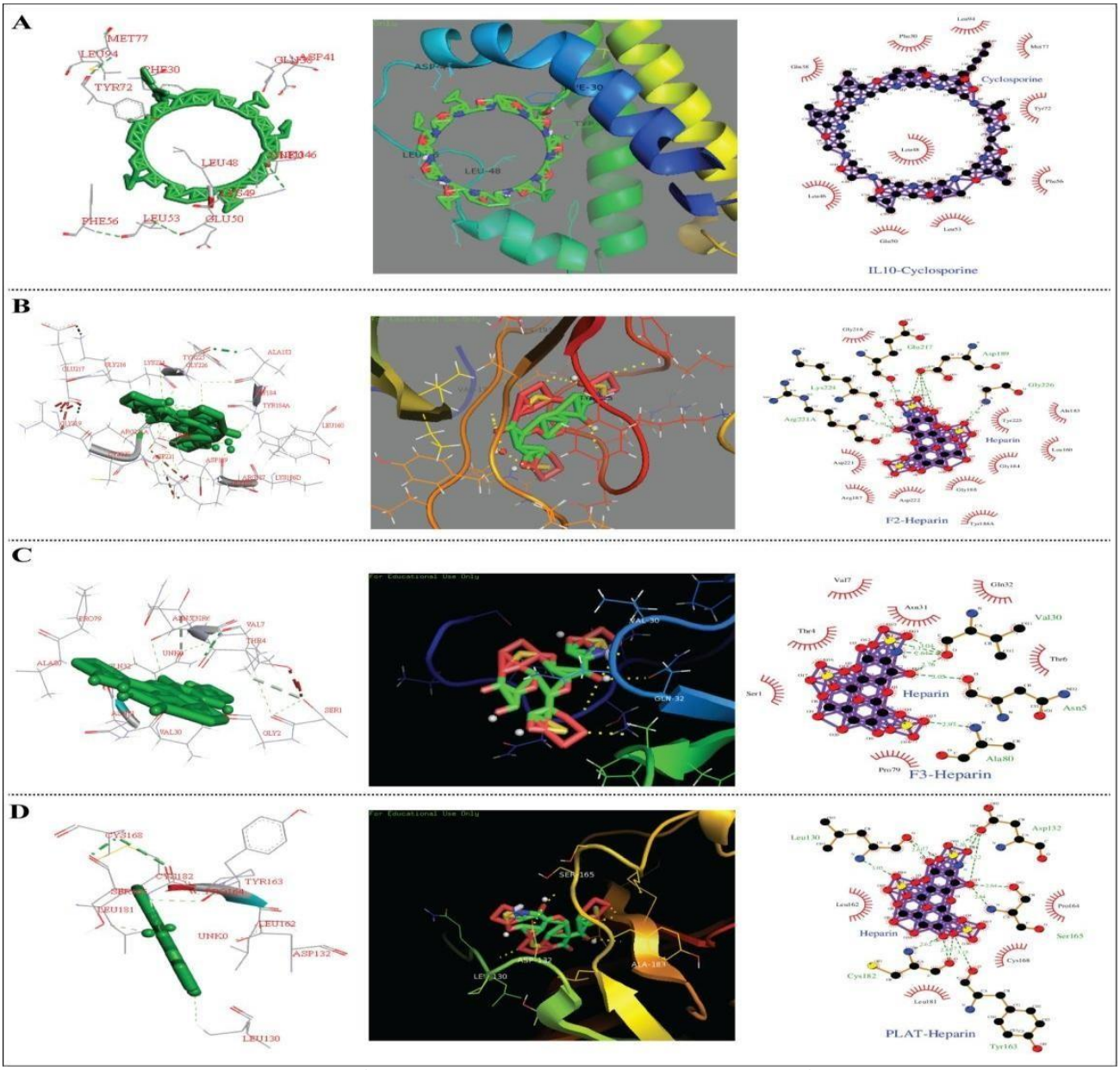


Fig. 5: (A)Molecular docking interaction of IL10-Cyclosporine (B) Molecular docking interaction of F2-Heparin (C) Molecular docking interaction of F3- Heparin (D) Molecular docking interaction of PLAT-Heparin

CONCLUSION

The present investigation was carried out to explore the genes and their interactions pertaining to Dengue which is through in-silico and molecular docking analysis. The present strategy of bioinformatics analysis is to exploit the current data available both on gene and genome-wide association study (GWAS) meta-analysis of dengue to integrate these at novel levels of understanding of gene network interactions and expression levels.

The study revealed 2 studies of dengue from GWAS with a total of 2 unique genes. namely MICB and PLCE1 that were mapped to discrete genomic locations of human genome which represented 15 rsIDs (SNPs) associated with MICB gene and 5 rsIDs (SNPs) associated with PLCE1 gene. The consensus results of the online tools like SIFT ,SNP & GO, PANTHER , Ploy phen2.0 and I-mutant for structure and functional studies depicted rs1051788, rs1051788, rs41293883 and rs45583740 of MICB gene to be deletrious/ diseased and effecting the structure and function of MICB gene which may the reason for occurrence of dengue in human counterpart. The DAVID bioinformatics functional enrichment analysis reported 234 genes and 468 GO terms for biological processes (BP).

In this study a total of 307 genes pertaining to dengue and its associated diseases were mined from various databases like GWAS, GEO & Jeans lab disease database. Out of 307 genes only 22 genes products, namely CD40LG, CD209, ICAM1, IFNG, IL4, IL8, IL10, PTPRC, TNF, CCL2, F3, ICAM1, VCAM1, F2, F3, JMJD6, PTPRC, SYT1, IFNG, PLAT, TNF, VCAM1, IL6, IL8, TNF, DDR1, IL8, MAPK1, TNF, IL8, MAPK1, IFNG, IL10, TNF, CEBPB, IL6, IL8, CD209, KAT2B, F2, PLAT, CCL2, IL10, TNF, CCL2, IL6, F2, IL6, F2, F3, PLAT, IFNAR2, IL10, CD40LG, F2, F3, PLAT, F2, CCL2, were considering for interaction study with different drug molecules. These genes or its corresponding protein could be treated as the key factors that play a vital role in dengue and its associate disease like Convalescent, Hemorrhagic Fever, Dengue Fever and dengue shock syndrome. The results from the above studies depicted the drugs Cyclosporine and Heparin as the potential drug for treatment of Dengue disease. The molecular docking studies have reported the drug-target interactions of IL10-cyclosporine was - 12.24, F2-heparin was -19.41, F3- heparin was -14.41, PLAT-heparin was -16.37. F2-heparin has been identified with the highest docking scores with energy minimization and interactions of F3- Phosphatidylserine was -2.91 which is identified as the lowest docking scores with energy minimization out of 22 ligand-protein interactions. From the above studies we could suggest F2, F3 and IL10 as the potential dengue targets and Cyclosporine and Heparin as potential antidengue drugs for treatment of viral diseases associated with dengue. We would further like to recommend going for *in vivo* and *in vitro* studies of to validate our *in silico* studies.

REFERENCES

- Aguirre, S., Maestre, A. M., Pagni, S., Patel, J. R., Savage, T., Gutman, D., ... & Fernandez-Sesma, A. (2012). DENV inhibits type I IFN production in infected cells by cleaving human STING. *PLoS Pathog*, , 8(10), e1002934.
- Amarasinghe, A., Kuritsky, J. N., Letson, G. W., & Margolis, H. S. (2011). Dengue virus infection in Africa. *Emerging infectious diseases*, *17*(8), 1349-1354.
- Ashour, J., Morrison, J., Laurent-Rolle, M., Belicha-Villanueva, A., Plumlee, C. R., Bernal-Rubio, D., ... & García-Sastre, A. (2010). Mouse STAT2 restricts early dengue virus replication. *Cell host & microbe*, *8*(5), 410-421.
- Beaumier, C. M., Jaiswal, S., West, K. Y., Friberg, H., Mathew, A., & Rothman, A. L. (2010). Differential in vivo clearance and response to secondary heterologous infections by H2b-restricted dengue virus-specific CD8+ T cells. *Viral immunology*, *23*(5), 477-485.
- Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., ... & Hay, S. I. (2013). The global distribution and burden of dengue. *Nature*, *496*(7446), 504-507.
- Brady, O. J., Gething, P. W., Bhatt, S., Messina, J. P., Brownstein, J. S., Hoen, A. G., ... & Hay, S. I. (2012). Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl Trop Dis.*, 6, e1760.
- Deen, J. L., Harris, E., Wills, B., Balmaseda, A., Hammond, S. N., Rocha, C., ... & Farrar, J. J. (2006). The WHO dengue classification and case definitions: time for a reassessment. *The Lancet*, *368*(9530), 170-173.
- Diamond, M. S., & Pierson, T. C. (2015). Molecular insight into dengue virus pathogenesis and its implications for disease control. *Cell*, *162*(3), 488-492.
- Gubler, D. J. (2002). Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends in microbiology*, *10*(2), 100-103.
- Gubler, D. J., & Clark, G. G. (1995). Dengue/dengue hemorrhagic fever: the emergence of a global health problem. *Emerging infectious diseases*, *1*(2), 55-57.
- Guha-Sapir, D., & Schimmer, B. (2005). Dengue fever: new paradigms for a changing epidemiology. *Emerging themes in epidemiology*, *2*(1), 1-10.
- Hammond, S. N., Balmaseda, A., Perez, L., Tellez, Y., Saborío, S. I., Mercado, J. C., ... & Harris, E. (2005). Differences in dengue severity in infants, children, and adults in a 3-year hospital-based study in Nicaragua. *The American journal of*

tropical medicine and hygiene, *73*(6), 1063-1070.

- Johansson, M. A., Dominici, F., & Glass, G. E. (2009). Local and global effects of climate on dengue transmission in Puerto Rico. *PLoS neglected tropical diseases*, *3*(2), e382.
- Khin, M. M., & Than, K. A. (1983). Transovarial transmission of dengue 2 virus by Aedes aegypti in nature. *The American journal of tropical medicine and hygiene*, *32*(3), 590-594.
- Kuhn, R. J., Zhang, W., Rossmann, M. G., Pletnev, S. V., Corver, J., Lenches, E., ... & Strauss, J. H. (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell*, *108*(5), 717-725.
- Lei, Y., Yu, H., Dong, Y., Yang, J., Ye, W., Wang, Y., ... & Zhang, F. (2015). Characterization of N-glycan structures on the surface of mature dengue 2 virus derived from insect cells. *PLoS One*, *10*(7), e0132122.
- Monath, T. P. (1994). Dengue: the risk to developed and developing countries. *Proceedings of the National Academy of Sciences*, *91*(7), 2395-2400.
- Nakhapakorn, K., & Tripathi, N. K. (2005). An information value based analysis of physical and climatic factors affecting dengue fever and dengue haemorrhagic fever incidence. *International journal of health geographics*, *4*, 1-13.
- Pierson, T. C., & Diamond, M. S. (2012). Degrees of maturity: the complex structure and biology of flaviviruses. *Current opinion in virology*, *2*(2), 168-175.
- Plianbangchang, S. Prevention and Control of Dengue and Dengue Haemorrhagic Fever. WHO.
- Reiter, P. (2001). Climate change and mosquito-borne disease. *Environ Health Perspect*, 109, 141–161. *Find this article online*.
- Rodhain, F., & Rosen, L. Mosquito vectors and dengue virus-vector relationships. In: Gubler DJ, Kuno G. Dengue and Dengue Hemorrahgic Fever. CAB International, New York: USA. 45-60.
- Rodhain, F., & Rosen, L. Mosquito vectors and dengue virus-vector relationships. In: Gubler DJ, Kuno G. Dengue and Dengue Hemorrahgic Fever. *CAB International*, New York:USA. 45-60.
- Rouvinski, A., Guardado-Calvo, P., Barba-Spaeth, G., Duquerroy, S., Vaney, M. C., Kikuti, C. M., ... & Rey, F. A. (2015). Recognition determinants of broadly neutralizing human antibodies against dengue viruses. *Nature*, *520*(7545), 109-113.
- Shepard, D. S., Halasa, Y. A., Tyagi, B. K., Adhish, S. V., Nandan, D., Karthiga, K. S., ... & INCLEN Study Group. (2014). Economic and disease burden of dengue illness in India. *The American journal of tropical medicine and hygiene*, *91*(6), 1235-1242.
- Shepard, D. S., Undurraga, E. A., Betancourt-Cravioto, M., Guzman, M. G., Halstead, S. B., Harris, E., ... & Gubler, D. J. (2014). Approaches to refining estimates of global burden and economics of dengue. *PLoS neglected tropical diseases*, *8*(11), e3306. Gupta, N., Srivastava, S., Jain, A., & Chaturvedi, U. C. (2012). Dengue in India. *Indian J Med Res.*, 136(3), 373-90.
- Thai, K. T., Nishiura, H., Hoang, P. L., Tran, N. T. T., Phan, G. T., Le, H. Q., ... & de Vries, P. J. (2011). Age-specificity of clinical dengue during primary and secondary infections. *PLoS neglected tropical diseases*, *5*(6), e1180.
- WHO (2006). Dengue: guidelines for diagnosis, treatment, prevention and control. Geneva, Switzerland.
- WHO (2007a). Scientific Working Group Report on Dengue. Geneva, Switzerland.
- Xiao, P. J., & Samulski, R. J. (2012). Cytoplasmic trafficking, endosomal escape, and perinuclear accumulation of adeno-associated virus type 2 particles are facilitated by microtubule network. *Journal of virology*, *86*(19), 10462-10473.