DEVELOPMENT OF DENGUE KNOWLEDGE BASE- AN INTEGRATIVE BIOLOGICAL WEB RESOURCE cum DATABASE ON DENGUE VIRAL DISEASE

A Major Project Dissertation submitted

in partial fulfilment of the requirement for the degree of

Master of Technology
In
Bioinformatics

Submitted by

Neha
(2K11/BIO/12)
Delhi Technological University, Delhi, India

Under the supervision of

Dr. B. D. Malhotra
Department of Biotechnology

Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, New Bawana Road, Delhi-110042,
INDIA

CERTIFICATE

This is to certify that the M. Tech. dissertation entitled **"DEVELOPMENT OF DENGUE KNOWLEDGE BASE- AN INTEGRATIVE BIOLOGICAL WEB RESOURCE cum DATABASE ON DENGUE VIRAL DISEASE"**, submitted by **NEHA (2K11/BIO/12)** in partial fulfilment of the requirement for the award of the degree of Master of Engineering, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi), is an authentic record of the candidate's own work carried out by her under my guidance.

The information and data enclosed in this thesis is original and has not been submitted elsewhere for honouring of any other degree.

Date: June, 2013

Dr. B. D. Malhotra
(Project Mentor)
Department of Bio-Technology
Delhi Technological University
(Formerly Delhi College of Engineering, University of Delhi)

ACKNOWLEDGEMENT

"…..the beauty of destination is half veiled and the fragrance of success is half dull, until the traces of all those enlightening the path are left to fly with wind spreading word of thankfulness……"

I hereby take this opportunity to express my deep sense of gratitude to my mentor, **Dr. B.D. Malhotra (**Professor, Department of Biotechnology), who has ungrudgingly and painstakingly spared both time and energy for me and under whose unfettered guidance I have been able to complete my project.

I express my gratitude to **Prof. S. Maji** (Head of Department), Department of Biotechnology, and all the faculty members for their support.

I am obliged to the faculty and staff members of Indian Biosciences Research Institute(IBRI), Noida, for the valuable information provided by them in their respective fields. I am grateful to **Dr. Kumud Sarin**, Technical Head, IBRI and **Mr. Mithilesh Mishra,** IT Head, IBRI, Noida for guiding me at every stage of the project.

Last but not the least, my heartfelt thanks to the Almighty, my family, colleagues and friends for their belief and love.

NEHA
(2K11/BIO/12)

DECLARATION

I hereby declare that the dissertation entitled, "**Development of Dengue Knowledge Base- An Integrative Biological Web Resource cum Database On Dengue Viral Disease"** submitted in the partial fulfilment of the **Master of Technology** Degree in **Bioinformatics** is a record work done by me, during the period January 2013-June 2013 and has not formed the basis for the award of any degree or other similar titles under any University in India or Abroad.

NEHA
(2K11/BIO/12)

CONTENTS

LIST OF FIGURES……

LIST OF TABLES.....

LIST OF ABBREVIATIONS....

DF	Dengue fever
DHF	Dengue hemorrhagic fever
DSS	Dengue shock syndrome
NS1	Non structural protein 1
NS2a	Non structural protein 2a
NS2b	Non structural protein 2b
NS3	Non structural protein 3
NS4a	Non structural protein 4a
NS4b	Non structural protein 4b
NS5	Non structural protein 5
DENV	Dengue Virus
DENV-1	Dengue serotype 1
DENV-2	Dengue serotype 2
DENV-3	Dengue serotype 3
DENV-4	Dengue serotype 4
NCBI	National Center for Biotechnology Information
CDD	Conserved Domain Database
SQL	Structured Query Language
DKB	Dengue Knowledge Base

ABSTRACT

Development of Dengue Knowledge Base
A web resource cum database on Dengue viral disease

Delhi Technological University, Delhi, India

ABSTRACT

With the advent of technology an exponential growth in biological information has been observed. Bioinformatics has played a significant role in integrating, storing and analyzing the biological information. In the era of modern science when scientific research became highly interdisciplinary and collaborative, the researchers seek for Internet resources that would allow them to find and useful information in fast and efficient manner. And databases lie at the core of data management and offer scientists the opportunity to access information efficiently.

Dengue Knowledge Base is a web resource cum database on Dengue Viral Disease that provides a comprehensive and unified access to information regarding various aspects of the Dengue virus and the disease in a user-friendly way. It is developed using SQL (database development technology) and Dot net framework (web resource development technology).

Dengue infection is a burgeoning menace that ails tropical and sub-tropical countries and India is one of the top rankers in the list! The global prevalence of Dengue has grown dramatically over the last decades in tropical and sub-tropical regions around the world.

Features:

- To make the access more user-friendly and comprehensive, the Search section provides the integrative links to navigate through the entire database driven website.
- The Advanced Query section allows to access the quick and relevant information on the protein's attributes as molar extinction co-efficient, isoelectric point, charged , acidic, basic, polar and hydrophobic amino acid profile.

Moreover, it is a repository of relevant information (taking into account all four serotypes of virus) as:

- Virus morphology, severity, pathogenesis, disease manifestations and its control and management.

- Genome wise gene prediction in all reading frames.

- Protein Sequence Computational Analysis Data including :

- Primary structural details(amino acid composition, molar extinction co-efficient, isoelectric point, charged , acidic, basic, polar and hydrophobic amino acid profile etc.).
- Secondary structural details(Residue wise probability of alpha, beta and turns regions, Surface Probability, Antigenic Index, Flexible Regions, Alpha and Beta, Amphipathic Regions etc.), Titration Curves
- Conserved regions
- Epitopes

- Literature reviewed host-virus protein- protein interactions
- Crystallographic protein structures deposited in PDB
- Drug targets.
- Plants metabolites acting as anti viral agents.

All the data are linked with Genbank (NCBI) files, the Protein Data Bank (PDB), the Kyoto Encyclopaedia of Genes and Genomes (KEGG), the primary literature, such as publications found in PubMed or other sequence or structure servers for references.

In a nutshell, useful information on Dengue Viral disease can be retrieved from the Dengue Knowledge Base and its development provides an opportunity to have insights into various aspects of Dengue viral disease on one platform in an organized and user-friendly manner. Till now, there does not exist any database or web resource that unifies such vast information on the same.

INTRODUCTION

OBJECTIVE : Development of Dengue Knowledge Base-A web resource cum database on Dengue viral disease

With a vision to develop a database driven web-resource on Dengue viral disease that includes literature reviewed data on various aspects of dengue viral disease as well as computational analysis of DENV(taking into account all serotypes) genome and protein sequences, DENGUE KNOWLEDGE BASE has been developed.

"Computational biology is part of a larger revolution that will affect how all of science is conducted. This larger revolution is being driven by the generation and use of information in all forms and in enormous quantities and requires the development of intelligent systems for gathering, storing and accessing information."-

In the last two decades as a consequence of development of rapid DNA sequencing techniques, the concomitant progress in computer based technologies has enabled molecular biologists in general and bioinformaticians in particular to cope with this information deluge. The emergence of bioinformatics is a phenomenon that promises to revolutionize people's lives through breakthroughs such as facilitating the creation of new drugs to treat diseases. The field of bioinformatics represents the convergence of biological data and computer technology that is necessary for the acquisition, management, and analysis of large-scale biological data. Most definitions of bioinformatics are similar to that offered: "Bioinformatics is the use of computational techniques for the consolidation and analysis of experimental data in biology." (Moody and Glyn ,2004)

TOPIC OF DATABASE DEVELOPMENT-DENGUE VIRAL DISEASE

The Dengue virus is a member of the virus family *Flaviviridae*. The virus is transmitted to humans by the mosquitoes *Aedes Aegypti* and *Aedes Albopictus*. The incidence of dengue fever (DF), dengue hemorrhagic fever (DHF), and dengue shock syndrome (DSS) are rapidly increasing, and more than 2.5 billion people live in regions endemic for the disease. Presently approximately 50-100 million cases of DF occur yearly with more than 500,000 resulting in severe and potentially fatal forms of the disease (DHF & DSS) (Halstead *et al*, 2007).
Several factors contribute to the threat posed by dengue. Most significant are lack of cross-reactive immunity for the four DEN serotypes (DEN-1, DEN-2, DEN-3, and DEN-4), hyper endemic circulation of the four different serotypes in the same geographical area, frequent worldwide travel, high population density, and lack of effective mosquito control programs. Other socioeconomic factors only amplify the challenge of dengue control(Harris *et al*, 2000).

Dengue virus is now believed to be the most common arthropod-borne disease in the world. DHF (Dengue hemorrhagic fever) and DSS (Dengue shock syndrome) are severe forms of dengue

virus infection leading to morbidity and mortality. Dengue virus are enveloped and contain a single, positive sense RNA genome of about 11kb that encodes a large polyprotiens precursor. Co- and post- transcriptional processing gives rise to 3 tructural and 7 non structural proteins, encoded by genes in order (from 5' to 3') .

THE DENGUE KNOWLEDGE BASE:

The Dengue Knowledge Base is a biological data consortium on Dengue virus and based on the collection of concepts related to it as its morphology, pathogenesis, manifestations, control and management, treatment etc. As well as, a detailed insight into its proteins is encompassed.

The primary literature, such as publications found in PubMed and published by researchers and clinicians in the field of biology, biotechnology, biochemistry, bioinformatics, proteomics, genomics etc forms the base of the Dengue Knowledge Base. There is a large volume of related articles published daily from which relevant text has been captured effectively allowing the user to have a unified access.The Genome sequences and the structural and non-structural proteins sequences of all serotypes of Dengue (retrieved from NCBI) were analyzed for their properties mentioned aforesaid. It aims to provide a structured collection of information about Dengue virus as :

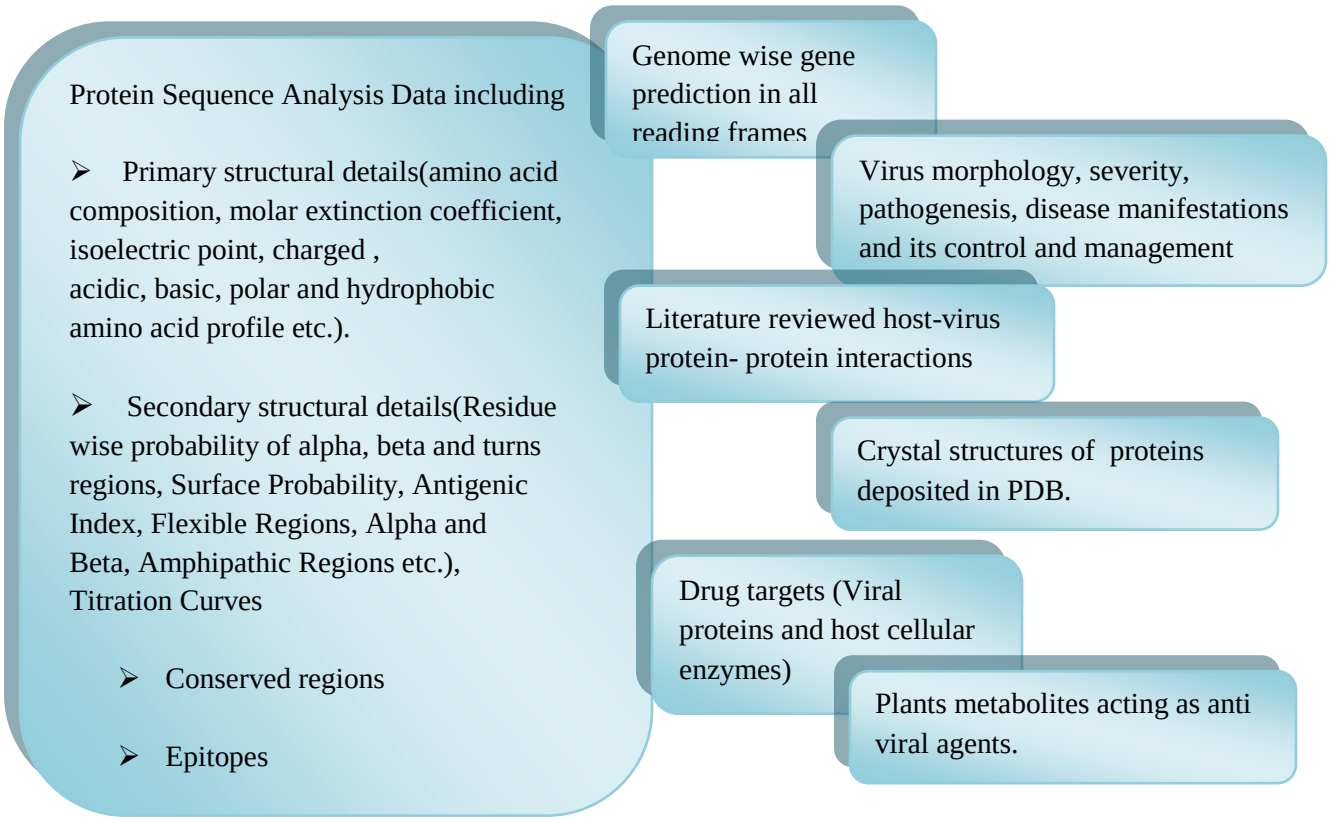


Fig. 1 A peep-in view of features of Dengue Knowledge Base.

ORGANIZATION

The organization of the Dengue Knowledge Base can be summed up as:

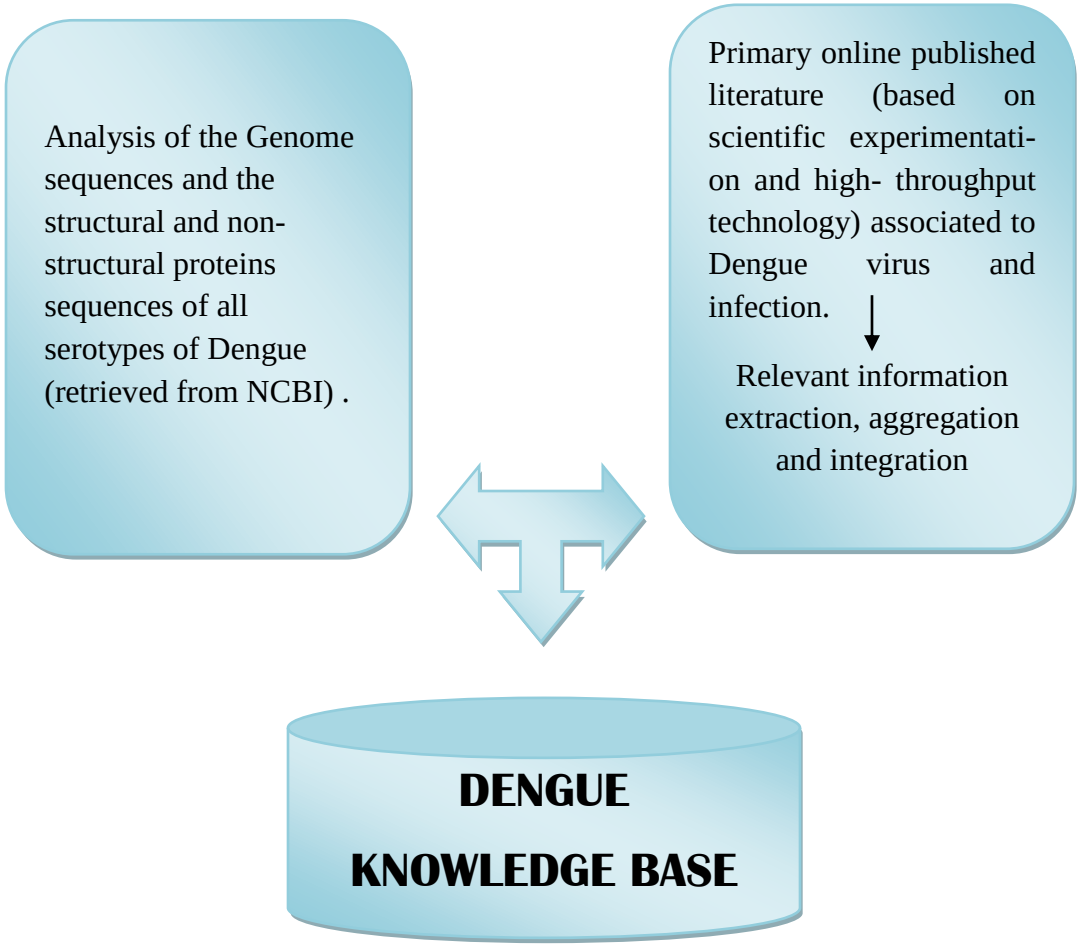


Fig.2 Structure of Dengue Knowledge Base

F EATURES

A snapshot of what areas related to Dengue has been covered in the Dengue Knowledge Base is summed up as under:

S.No.	Feature in Database	Description
1.	Dengue-paedia	Can be termed as encyclopaedia of Dengue virus and the infection caused by it. It covers topics as: • Virus Morphology(Virus genome structure, Replication of virus) • Virus Pathogenesis(Immune response generated by humans upon infection) • Disease Manifestation (DHF,DF) • Organ pathology • Control and management of disease • Epidemiology of disease
2.	Host-Virus Interaction	It encompasses the known interactions between host(human as well as mosquito) and the virus at genetic level. 105 human-virus and 65 mosquito-virus interaction has been illustrated. The gene information can be retrieved from the Genbank(link provided)
3.	Herbal-Shot	This feature puts forward all the literature reviewed plant metabolites that are known to be anti-Dengue. Plants acting as viral-inhibitors as well as larvacidal

		are described. More information about the compound can be retrieved from PubChem(link provided). Plant information can also be retrieved from the Plant database (link provided).
4.	Predicted Genes	The genome sequences of all serotypes of virus(DENV-1, DENV-2, DENV-3 and DENV-4) have been analyzed to predict genes in all six reading frames and also their corresponding protein sequences.
5.	Protein Analysis	The Protein sequences of all serotypes of virus(DENV-1, DENV-2, DENV-3 and DENV-4) have been analyzed for: • Primary structural details(amino acid composition, molar extinction coefficient, iso-electric point, charged ,acidic, basic, polar and hydrophobic amino acid profile etc.). • Secondary structural details(Residue wise probability of alpha, beta and turns regions, Surface Probability, Antigenic Index, Flexible Regions, Alpha and Beta, Amphipathic Regions etc.), • Titration Curves, • Conserved regions, • Epitopes
6.	Protein Structure PDB links.	The Crystal structures of proteins can be accessed from here.
7.	Titration curve	Titration curve of each protein can be accessed through this section.

7.	Search and Advanced search	These sections provide an interface from where the user can get options to retrieve desired information and navigate through the entire database driven web-resource.

COMPUTATIONAL RESOURCES USED:

The data collection, extracting information and analysis and interpretation employed a number of computational tools. They are summed as under:

Table.2 List of computational resources used in present work

	COMPUTATIONAL RESOURCES	PURPOSE OF USE IN PRESENT WORK
1.	NCBI (Gene, Genome, Protein)	• To retrieve Dengue virus genome and protein sequences.
2	(NCBI)PubMed	• To access research papers to extract relevant information on various aspects of Dengue virus and infection as well as other associated information
3	(NCBI) CDD (Conserved Domain Database by NCBI)	• To predict conserved domains in of proteins encoded by Dengue virus genome.
5.	*Chemgenome 3.0*	• To produce and interpret structural annotations for the viral genome of *Dengue virus*.
6.	Protean (DNASTAR's Lasergene Core Suite version 10.1)	• To predict and annotate the structural character of proteins encoded by Dengue virus genome.
7.	Protein Data Bank	• To access the known crystallographic structures of proteins encoded by Dengue virus genome.
8.	Immune Epitope Database and Analysis Resource	• To predict B-cell and T-cell Epitopes.
9.	PubChem	• To access chemical properties and structure of compounds showing anti-Dengue activity.
10.	Microsoft Visual Studio 2010	• To develop web resource.
11.	Microsoft SQL 2005	• To develop database.
12.	Tropicos	• To access information on plants acting as anti-Dengue agents.

REVIEW OF LITERATURE

chapter- 1

1.1 DATABASES:

The history of computing in biology goes back to the 1920s when scientists were already thinking of establishing biological laws solely from data analysis by induction. However, only the development of powerful computers and the availability of experimental data that can be readily treated by computation (for example, DNA or amino acid sequences and three-dimensional structures of proteins) launched bioinformatics as an independent field. Today, practical applications of bioinformatics are readily available through the World Wide Web, and are widely used in biological and medical research (Waterman and Michael S.,1995). Three key areas are

- The organization of knowledge in databases,
- Sequence analysis and
- Structural bioinformatics.

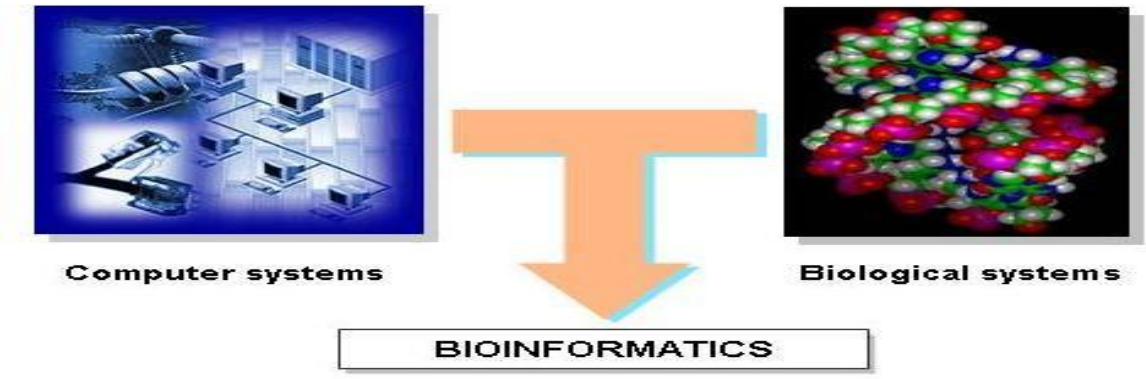


Fig.3 Bioinformatics schematics

1.2 BIOLOGICAL DATABASES

Biological databases are libraries of life sciences information, collected from scientific experiments, published literature, high-throughput experiment technology, and computational analyses. Biological knowledge is distributed amongst many different general and specialized databases. This sometimes makes it difficult to ensure the consistency of information. Integrative bioinformatics is one field attempting to tackle this problem by providing unified access. Biological database design, development, and long-term management is a core area of the discipline of bioinformatics (Johnson *et al*, 2004). Data contents include gene sequences, textual descriptions, attributes and ontology classifications, citations, and tabular data. These are often described as semi-structured data, and can be represented as tables, key delimited records, and XML structures. Biological raw data are stored in public databanks (such as Genbank or EMBL for primary DNA sequences). The data can be submitted and accessed via the World Wide Web. Protein sequence databanks like trEMBL provide the most likely translation of all coding

sequences in the EMBL databank. Sequence data are prominent, but also other data are stored, e. g. yeast two-hybrid screens, expression arrays, systematic gene-knock-out experiments, and metabolic pathways.

1.3 FUNCTIONALITY OF DATABASES:

The stored data need to be accessed in a meaningful way, and often contents of several databanks or databases have to be accessed simultaneously and correlated with each other. Special languages have been developed to facilitate this task (such as the Sequence Retrieval System (SRS) and the Entrez system). An unsolved problem is the optimal design of inter-operating database systems. Databases provide additional functionality such as access to sequence homology searches and links to other databases and analysis results. For example, SWISSPROT contains verified protein sequences and more annotations describing the function of a protein. Protein 3D structures are stored in specific databases (for example, the Protein Data Bank, now primarily curated and developed by the Research Collaborators for Structural Bioinformatics). Organism specific databases have been developed (such as ACEDB, the A C. Elegans DataBase for the C. elegans genome, FLYBASE for D. melanogaster etc). A major problem are errors in databanks and databases (mostly errors in annotation), in particular since errors propagate easily through links.

Also databases of scientific literature (such as PUBMED, MEDLINE) provide additional functionality, e.g. they can search for similar articles based on word-usage analysis. Text recognition systems are being developed that extract automatically knowledge about protein function from the abstracts of scientific articles, notably on protein-protein interactions.

Information from various sources explicitly presented in the form of biological databases is generally maintained by many institutions. They vary widely in their content, format and access method (Kihler2004). Biological data are being deposited in a database and the need for data analysis has made molecular biology databases vital tools for research. Although several databases are reported in literature, many researchers are not aware of their availability and these databases were not effectively utilized.

Interfaces to Molecular Biology databases aim at overcoming the following obstacles: Limited data awareness, complex data retrieval, limited data analysis tools availability, limited literature reference availability.

Fig.4 Schematics of Biological databases: Biological databases often have web interfaces, which allow users to send queries to the databases

Chapter-2

2.1 DENGUE VIRUS

Dengue virus is the most widespread vector borne viral disease of humans. DHF (Dengue hemorrhagic fever) and DSS (Dengue shock syndrome) are severe forms of dengue virus infection leading to morbidity and mortality (Sittisombut *et al*, 1997).

Four dengue virus serotype exists, and DHF/DSS occur almost in patients who are re-infected with a different virus serotype. Globally it is estimated that approximate 50 to 100 million dengue virus infections occur annually (Freedman DO *et al*, 2006). Among these there are 200,000 to 500,000 cases of potential life threatening DHF & DSS, characterized by thrombocytopenia and increased vascular permeability. The death rate associated with more severe forms DHF/DSS is approx 5% predominantly in children under the age of 15, and now in adults also.

2.2 DENV CLASSIFICATION

DENV belongs to the family *Flaviviridae*, genus *Flavivirus*, and is transmitted through the mosquito vector *Aedesaegypti*. Flaviviruses are enveloped, positive-stranded RNA viruses. DENV shares the *Flaviviridae* classification with eight other viruses with similarities in both structural and non-structural (NS) proteins. Other family members include: Aroa virus, Japanese encephalitis, Kolobera virus, West Nile virus, and Yellow fever virus. DENV also belongs to a larger heterogeneous group of viruses transmitted by insect vectors, called arboviruses. For these viruses, the transmission to vertebrate hosts is dependent on arthropod vectors, which for DENV is *Aedes aegypti*. Currently about two thirds of the world population live in areas with *Aedes aegypti* (Pinheiro and Corber, 1997).

2.3 EPIDEMIOLOGY OF DISEASE

The first cases of Dengue Fever (DF) were recorded In 1779 in Batavia, Indonesia, and Catro. For the past 200 years, pandemics have been recorded. In tropical and subtropical climates at 10 to 30 year intervals. In 1944, Albert Sabin successfully isolated the virus that causes DF and found that belonged to the Flaviviridae virus family.

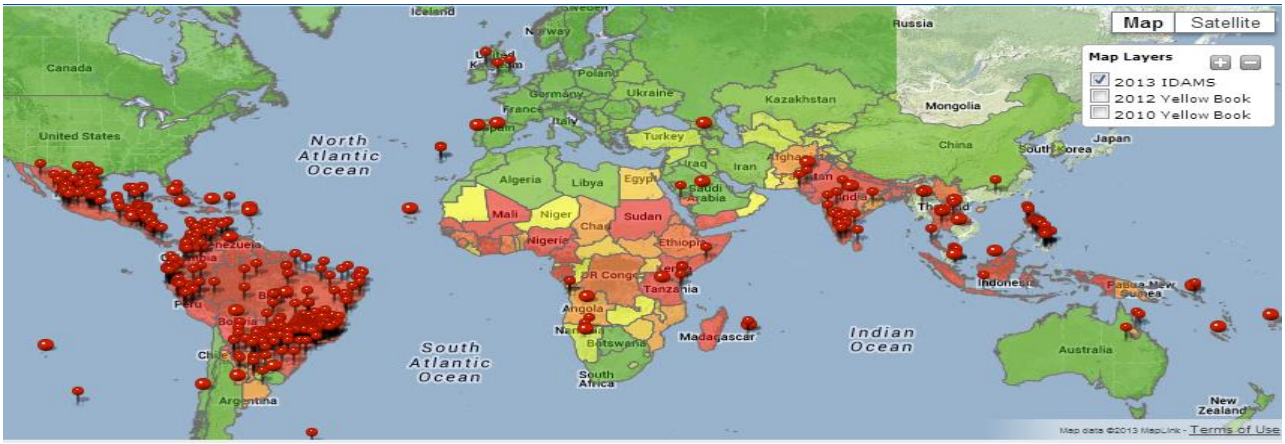


Fig.5 Dengue affected regions (courtesy: CDC, Health map)

Dengue virus infection is the most common arthropod-borne disease worldwide with an increasing incidence in the tropical regions of Asia, Africa, and Central and South America. There are four serotypes of the virus (Service, 1992).
The global prevalence of dengue has grown dramatically in recent decades. The disease is now endemic in more than 100 countries in Africa, the Americas, the Eastern Mediterranean, South-east Asia and the Western Pacific. South-East Asia and the Western Pacific are most seriously affected. Some 2,500 million people - two fifths of the world's population - are now at risk from dengue. It is estimated that there may be 50 million cases of dengue infection worldwide every year (Guha *et al* 2006).

Ethnicity is nonspecific, but the disease's distribution is geographically determined. Fewer cases have been reported in the black population than in other races

Significant outbreaks of dengue fever tend to occur every five or six years. The cyclicity in numbers of dengue cases is thought to be the result of seasonal cycles interacting with a short-lived cross-immunity for all four strains, in people who have had dengue (Weaver SC and Reisen WK, 2010). When the cross-immunity wears off, the population is then more susceptible to transmission whenever the next seasonal peak occurs. Thus in the longer term of several years, there tend to remain large numbers of susceptible people in the population despite previous outbreaks because there are four different strains of the dengue virus and because of new susceptible individuals entering the target population, either through childbirth or immigration.

On the basis of disease occurrence and severity, countries have been categorised into:

- CATEGORY-A (Indonesia, Myanmar and Thailand)
- CATEGORY-B (India, Bangladesh, Maldives and Sri-Lanka)
- CATEGORY-C (Bhutan and Nepal)
- CTEGORY-D (Korea)

2.4 DENGUE IN INDIA

DF has been recognized for many years m India, since the outbreak of dengue occurred in 1912 In Kolkata (Kennedy, 1912), the proportion of DHF or DF cases with hemorrhagic manifestations has Increased In the last 5-6 years, all the states of the country have reported outbreaks The onset of the disease occurs immediately after the monsoon season which varies in duration from state to state, between July and November. The attack rates in the outbreaks ranged from 20% to 80% of the population. In affected localities, all age-groups have been affected by DF, and all the four serotypes of the virus have been isolated and are in circulation in the country: more than one serotypes are commonly present during many of the outbreaks in urban areas. The sero-positivity of localities affected by dengue can be quite high ranging from 8% to 91%(Victor *et. al.*. 2002).

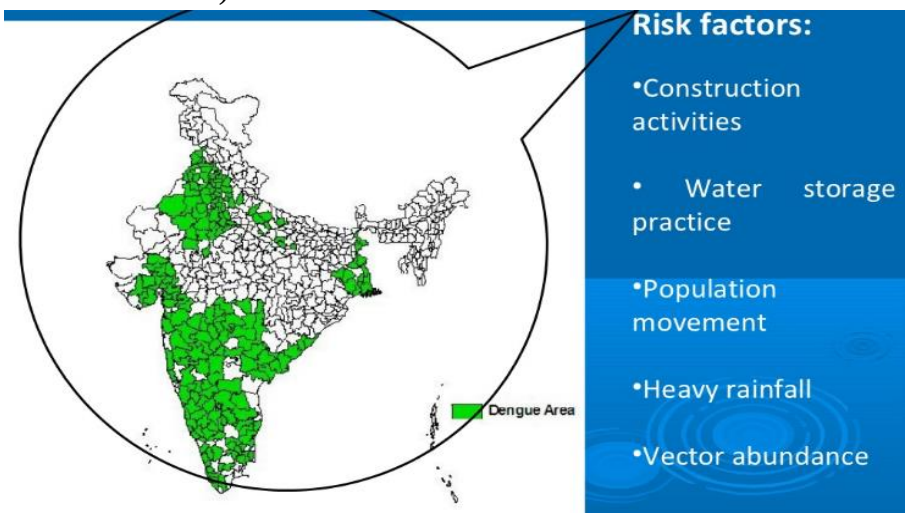


Fig. 6(a)Overview of Dengue effected areas in India and associated risk actors.

The first outbreak of dengue fever in India was recorded in 1812 (Jatanasen S and Thongcharoen P,1993). In spite of preventive measures taken by the respective governments since then, recurrent outbreaks have occurred, and over the last 10 to 15 years DF has been the major cause of hospitalization and mortality after acute respiratory and diarrhoeal infections among children. New Delhi, the capital of India located in the northern region of the country, experienced seven major outbreaks between 1967 and 2003 (Broor *et al*, 1997, Gupta *et al*, 2005). Then in 2006 another major outbreak occurred with more than 11,000 reported cases and 165 reported fatal cases. Data obtained from the World Health Organization (WHO) exhibits the number of DF cases reported in India from 1991 to 2008 as well as the annual reported fatality rate during this period.

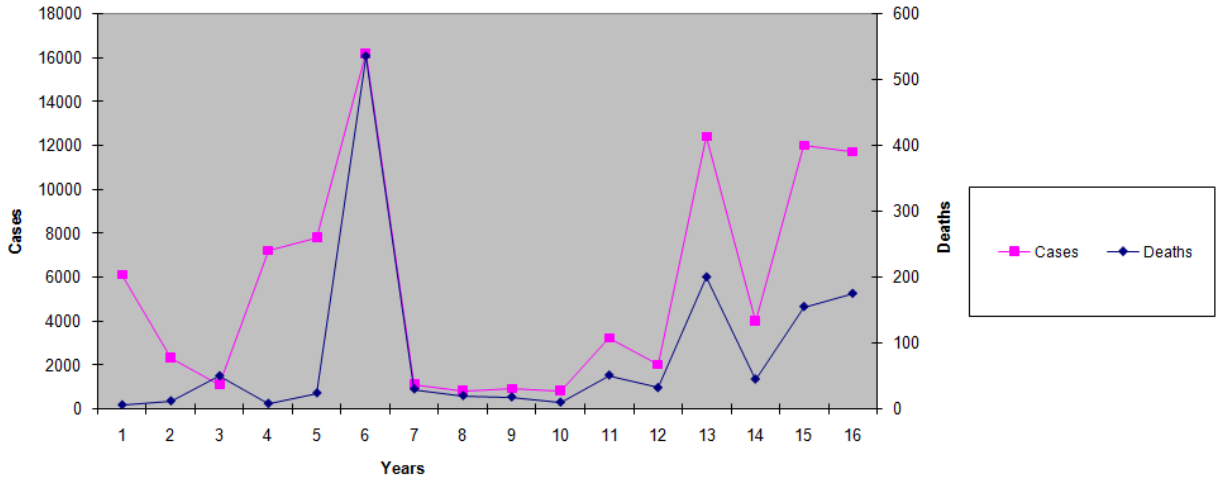


Fig 6(b) DF reported cases and deaths in India from 1991 till 2010.

2.5 *A. aegypti* VECTOR FOR DISEASE

The primitive enzootic transmission cycle of dengue viruses involves canopy dwelling Aedes mosquitoes and lower primates in the rain forests of Asia and Africa. Current evidence suggests that these viruses do not regularly move out of the forest to urban areas (Rico-Hesse, 1990). An epidemic transmission cycle may occur in villages or islands, where the human population is small (Kittigui *et al*, 2007). Introduced viruses quickly infect the majority of susceptible individuals in these areas, and increasing herd immunity causes the virus to disappear from the population. A number of Aedes spp. may act as a vector in these situations, depending on the geographic area, including A, aegypti and *A*, albopictus.

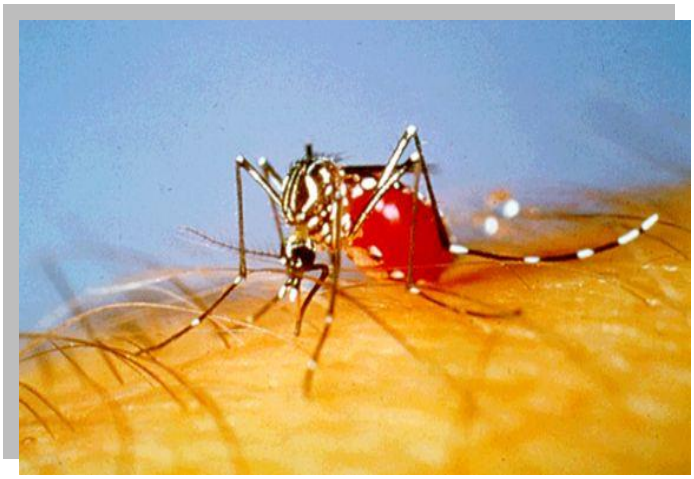


Fig.7 Aedes Aegypti mosquito- Vector for Dengue viral manifestation

The viruses are maintained in an A. Aegypti-human-A. Aegvpti cycle with periodic epidemics. Often, multiple virus serotypes co circulate in the same city (hyperendemicity) (Service, 1992). The vectors, *A*, aegypti and A. Albopictns, are widespread in India and their local densities can

be quite high; however, the role of A. Albopictus has not yet been established (Victor *et al*., 2002). Most of the DF and DHF outbreaks have occurred in localities where the larval house index was more than 20%. There is no regular vector surveillance and control programme in India.

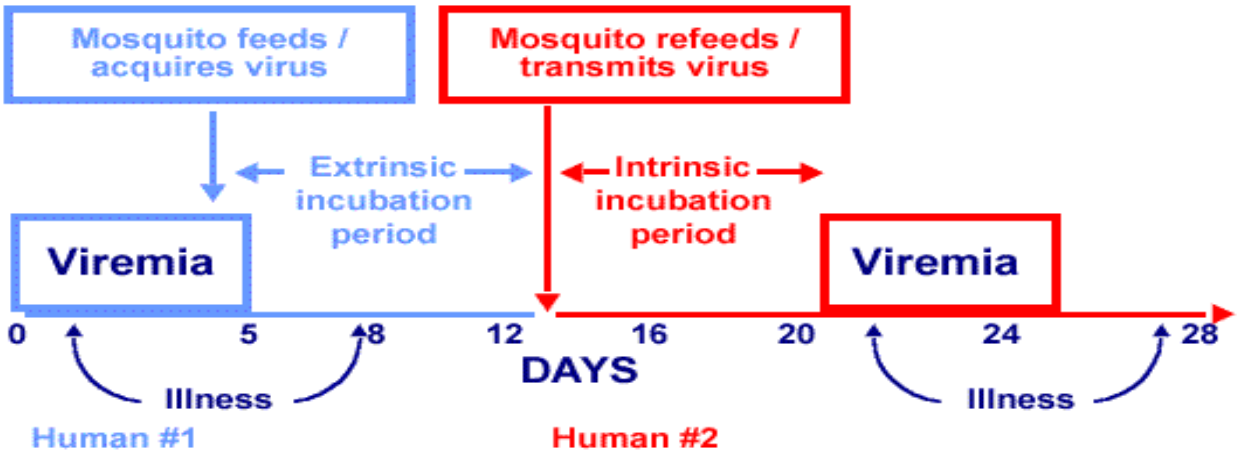


Fig. 8 Transmission of disease.

2.6 DENV GENE EXPRESSION:

The virion RNA is infectious and serves as both the genome and the viral messenger RNA. The whole genome is translated into a polyprotein, which is processed co- and post-translationally by host and viral proteases. Dengue virus are enveloped and contain a single, positive sense RNA genome of about 11kb that encodes a large polyprotiens precursor. Co- and post- transcriptional processing gives rise to 3 structural and 7 non structural proteins, encoded by genes in order (from 5' to 3')

Structural proteins:
C (capsid),
preM (premembrane),
E(envelope)

Non-structural proteins:
NS1,NS2a, NS2b, NS3,
NS4b and NS5

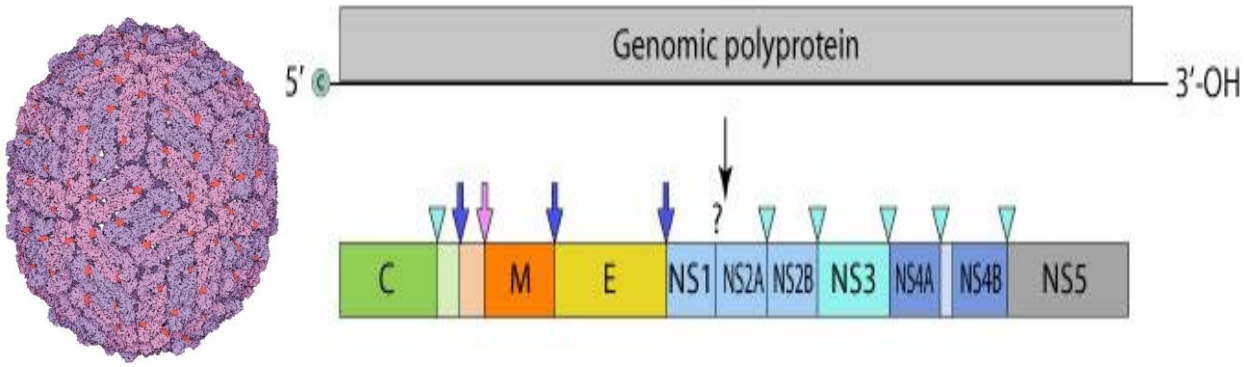


Fig 9 DENV genomic polyprotein(courtesy:PDB)

2.7 PATHOGENESIS OF DENGUE DISEASE:

1. Upon inoculation of DENV into the dermis, Langerhans cells and keratinocytes will primarily be infected.
2. The virus subsequently spreads via the blood (primary viremia) and infects tissue macrophages in several organs, especially the macrophages in the spleen. The replication efficiency of DENV in DC, monocytes, and macrophages, as well as its tropism for and replication efficiency in EC, bone marrow stromal cells, and liver cells, collectively determine the viral load measured in blood.
3. This viral load represents an important risk factor for development of severe disease. Essentially, infection of macrophages, hepatocytes, and EC influences the hemostatic and the immune responses to DENV.
4. Infected cells die predominantly through apoptosis and to a lesser extent through necrosis.
5. Necrosis results in release of toxic products, which activate the coagulation and fibrinolytic systems. Depending on the extent of infection of bone marrow stromal cells and the levels of IL-6, IL-8, IL-10, and IL-18, hemopoiesis is suppressed, resulting in decreased blood thrombogenicity.
6. Platelets interact closely with EC, and a normal number of functioning platelets is necessary to maintain vascular stability.
7. A high viral load in blood and possibly viral tropism for EC, severe thrombocytopenia, and platelet dysfunction may result in increased capillary fragility, clinically manifested as petechiae, easy bruising, and gastrointestinal mucosal bleeding, which is characteristic of DHF.
8. At the same time, infection stimulates development of specific antibody and cellular immune responses to DENV. When IgM antibodies that crossreact with EC, platelets, and plasmin are produced, the loop that results in increased vascular permeability and coagulopathy is amplified.
9. In addition, enhancing IgG antibodies bind heterologous virus during secondary infection and enhance infection of APCs, thereby contributing to the increased viral load that is seen during secondary viremia in some patients.
10. Furthermore, a high viral load overstimulates both low- and high-avidity cross-reactive T cells. In the context of certain HLA haplotypes, cross-reactive T cells delay virus clearance, while producing high levels of proinflammatory cytokines and other mediators.
11. Ultimately, these high levels of soluble factors, many of which still remain to be identified, induce changes in EC leading to the coagulopathy and plasma leakage characteristic of DSS.
    (Whitehorn J *et al*, 2011**)**

2.8 CLINICAL MANIFESTATIONS OF DENGUE VIRAL DISAESE:

There are actually four dengue clinical syndromes: (WHO, 2007)

1. Undifferentiated fever	Classic dengue fever	Dengue hemorrhagic fever or DHF	Dengue shock syndrome or DSS

Clinical Case Definition for Dengue Fever

Classical Dengue fever or Break bone fever is an acute febrile viral disease frequently presenting with headaches, bone or joint pain, muscular pains, rash and leucopenia. (Guilarde *et al*, 2008)

Clinical Case Definition for Dengue Hemorrhagic Fever

4 Necessary Criteria:

1. Fever, or recent history of acute fever
2. Hemorrhagic manifestations
- Low platelet count (100,000/mm3 or less)
4. Objective evidence of 'leaky capillaries'

Clinical Case Definition for Dengue Shock Syndrome

- 4 criteria for DHF + Evidence of circulatory failure manifested indirectly by all of the following:
- Rapid and weak pulse, Narrow pulse pressure (≤ 20 mm Hg) or hypotension for age, Cold, clammy skin and altered mental status(Noisakran S *et al*,2008).
- Frank shock is direct evidence of circulatory failure

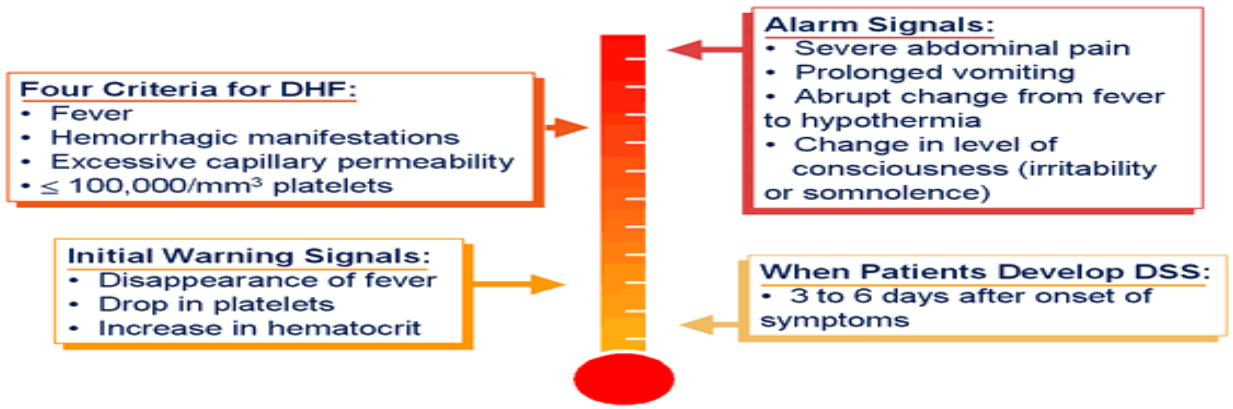


Fig.10 Warning signs and signals of DENV infection

2.9 DIAGONOSIS:

Diagnosis of dengue falls into two stages:

- stage I, fever and viraemia accompanied by NSI antigens in blood; and
- stage II, the early post-febrile period lasting a few weeks when IgM and IgG antibodies are in higher levels (Halstead. 2007).

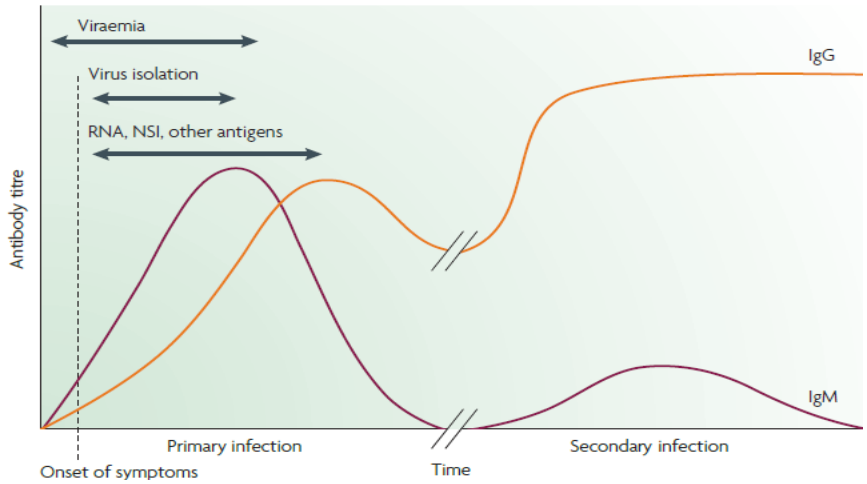


Fig.11 Dengue diagonostics

During primary infection, viraemia more or less coincides with fever. However, during a secondary infection; the duration of viraemia can he 2 or 3 days. Whereas presence of NS I antigens in blood lasts somewhat longer. Serological diagnosis will not be positive until defervescence. In individuals with DHF and DSS. vascular permeability is recognized usually at defervescence. at which time the IgM-capture serological test should he positive but tests to detect virions, dengue RNA. or dengue proteins could be negative(Singh et al., 2006). Commercial dengue serological tests in most countries are not subjected to quality control (Blacksell *et al*, 2006).

3.2.10 THE HUNT FOR DENGUE CURE

Although there are no ideal vaccines or therapy for the prevention and treatment of DHF, the understanding of the life cycle of dengue virus has made great progress over the past few years, and all the life cycle stages can represent potential targets for antiviral drug discovery (Rui-feng Qi *et al*, 2009).

High-throughput screening method have been used to identify mosquito and human proteins that physically interact with dengue proteins and are conserved across serotypes,moving us closer to a complete host – dengue protein interactome. These data should be useful for understanding the

interplay between dengue and its hosts and may provide candidates for drug targets and vector control strategies.(Mairiang D *et al,* 2013)

A dengue vaccine has proven difficult to develop, in part because there are four major subtypes of dengue virus, each with slightly different viral proteins. Many researchers currently believe that the deadly dengue hemorrhagic disease is caused when a person is infected with one subtype, and then infected later by a second subtype (Halstead, 1993). The antibodies, and immunity, gained from the first infection appear to assist with the infection by the second subtype, instead of providing a general immunity to all subtypes. This means that an effective vaccine will have to stimulate protective antibodies against all four types at once, a feat that has not yet been achieved.

A primary infection with any one of the four serotypes by means of a mosquito bite leads to lifelong immunity to that serotype. Secondary infection requires productive infection by a different DENV serotype, heterologous from the primary infecting serotype, and leads to 90% of all severe cases, increasing the risk of DHF by 15-80 fold. Thus, one antibody type only provides partial short-term protection against other DENV serotypes (Sabin, 1952). A challenge for any vaccine under development is protection against DHF. Furthermore, immunizing against a single serotype, or incomplete vaccination against any single serotype, may lead to increased risk of a move severe infection (Rothman, 2004).

A large number of diverse dengue vaccine candidates are in preclinical development. These preclinical candidates ensure a continued influx of innovation into the vaccine pipeline, which is critical for maximizing the chances of success for dengue vaccine development. One possible scenario is that a first generation of licensed dengue vaccines will arise from candidates currently in clinical development, while some of the current preclinical stage candidates could become next generation dengue vaccines licensed at a later stage. Next generation dengue vaccines could serve to complement existing vaccines.(Julia, 2011)

Careful planning and a coordinated approach to safety assessment are recommended to ensure adequate long-term evaluation of dengue vaccines that will support their introduction and continued use.(Bentsi-Enchill AD et al, 2013).

chapter-3

3.1 WEB RESOURCE DEVELOPMENT TECHNOLOGY:VISUAL STUDIO 2010

Microsoft Visual Studio is an integrated development environment (IDE) from Microsoft.
It is used to develop console and graphical user interface applications along with Windows Forms or WPF applications, web sites, web applications, and web services in both native code together with managed code for all platforms supported by Microsoft Windows, Windows Mobile, Windows CE, .NET Framework, .NET Compact Framework and Microsoft Silverlight.

Visual Studio supports different programming languages by means of language services, which allow the code editor and debugger to support (to varying degrees) nearly any programming language, provided a language-specific service exists. Built-in languages include C/C++ (via Visual C++), VB.NET(via Visual Basic .NET), C# (via Visual C#), and F# (as of Visual Studio 2010). Support for other languages such as M, Python, and Ruby among others is available via language services installed separately. It also supports XML/XSLT, HTML/XHTML, JavaScript and CSS. Individual language-specific versions of Visual Studio also exist which provide more limited language services to the user: Microsoft Visual Basic, Visual J#, Visual C#, and Visual C++.

FEATURES:

Some of the numerous features of visual studio 2010 are:

Code editor:

- Supports syntax highlighting and code completion using IntelliSense
- As code is being written, Visual Studio compiles it in the background in order to provide feedback about syntax and compilation errors, which are flagged with a red wavy underline. Warnings are marked with a green underline

Debugger:

- A computer program used to test and debug other programs

Designer:

- Visual Studio includes a host of visual designers to aid in the development of applications

- Visual Studio includes a web-site editor and designer that allows web pages to be authored by dragging and dropping widgets.

Solution Explorer:

- The *Solution Explorer* is used to manage and browse the files in a solution

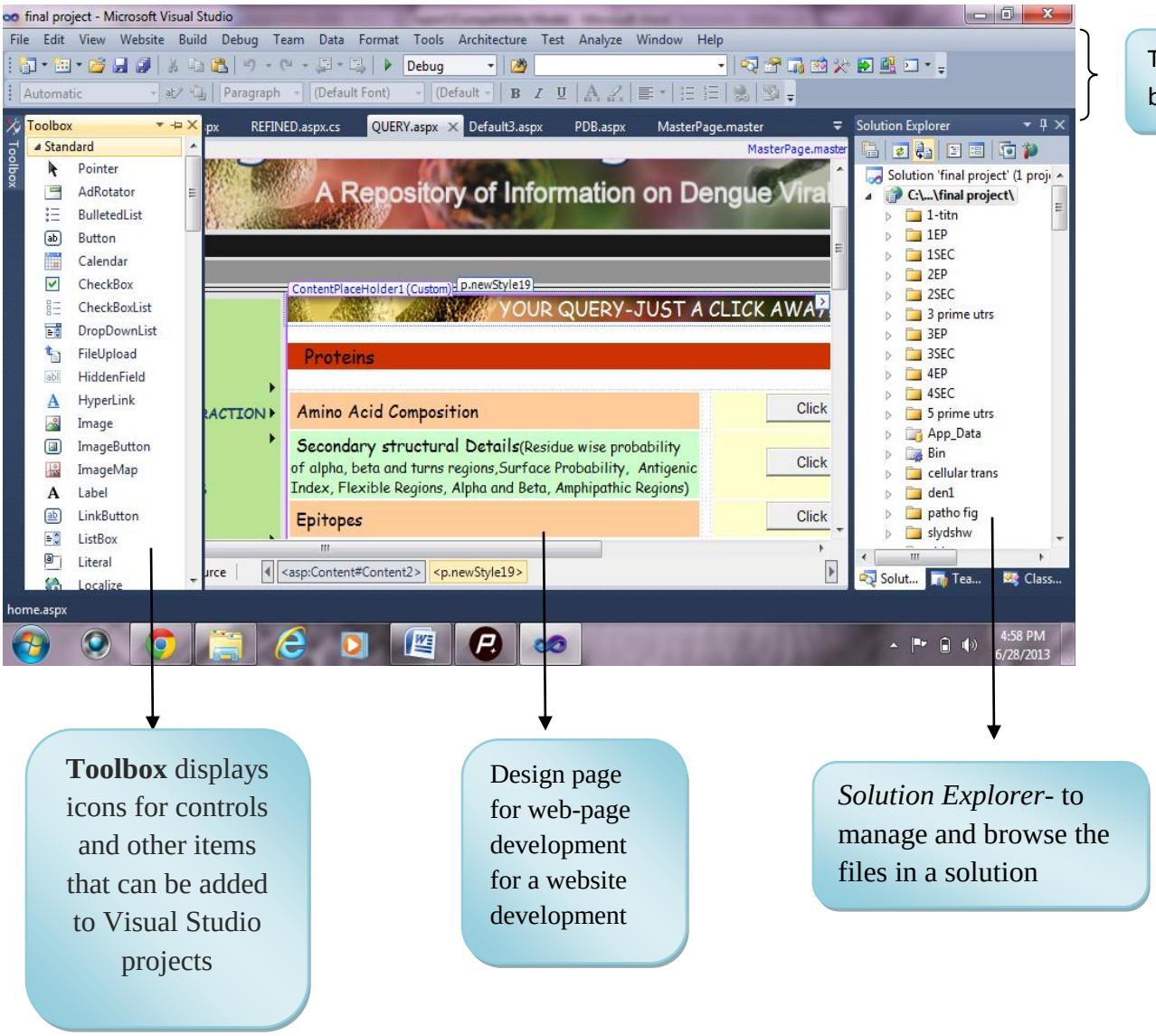


Fig. 12 Designer page in Visual studio 2010

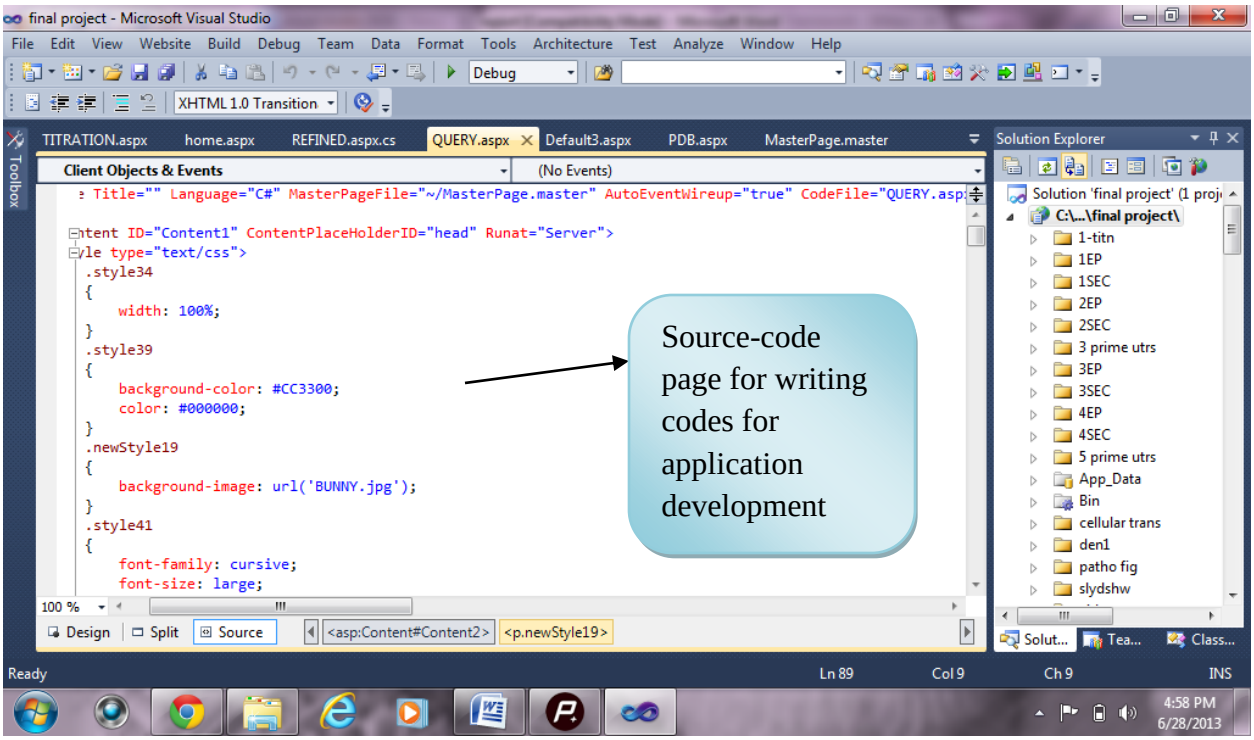


Fig.13 Source code page in Visual studio 2010

3.2 DATABASE DEVELOPMENT TECHNOLOGY:SQL

A database is a computer system involved in data storage and retrieval, most often focused on the relational model. The relational model is based on tables for data storage and relationships between those tables. SQL (structured query language) is the primary language used to communicate with relational databases.

SQL (Structured Query Language) is a special-purpose programming language designed for managing data held in a relational database management system (RDBMS).

SQL became a standard of the American National Standards Institute (ANSI) in 1986, and of the International Organization for Standards (ISO) in 1987.

SQL:

- Allow users to access data in relational database management systems.
- Allow users to describe the data.
- Allow users to define the data in database and manipulate that data.
- Allow to embed within other languages using SQL modules, libraries & pre-compilers.
- Allow users to create and drop databases and tables.
- Allow users to create view, stored procedure, functions in a database.
- Allow users to set permissions on tables, procedures, and views

METHODOLOGY

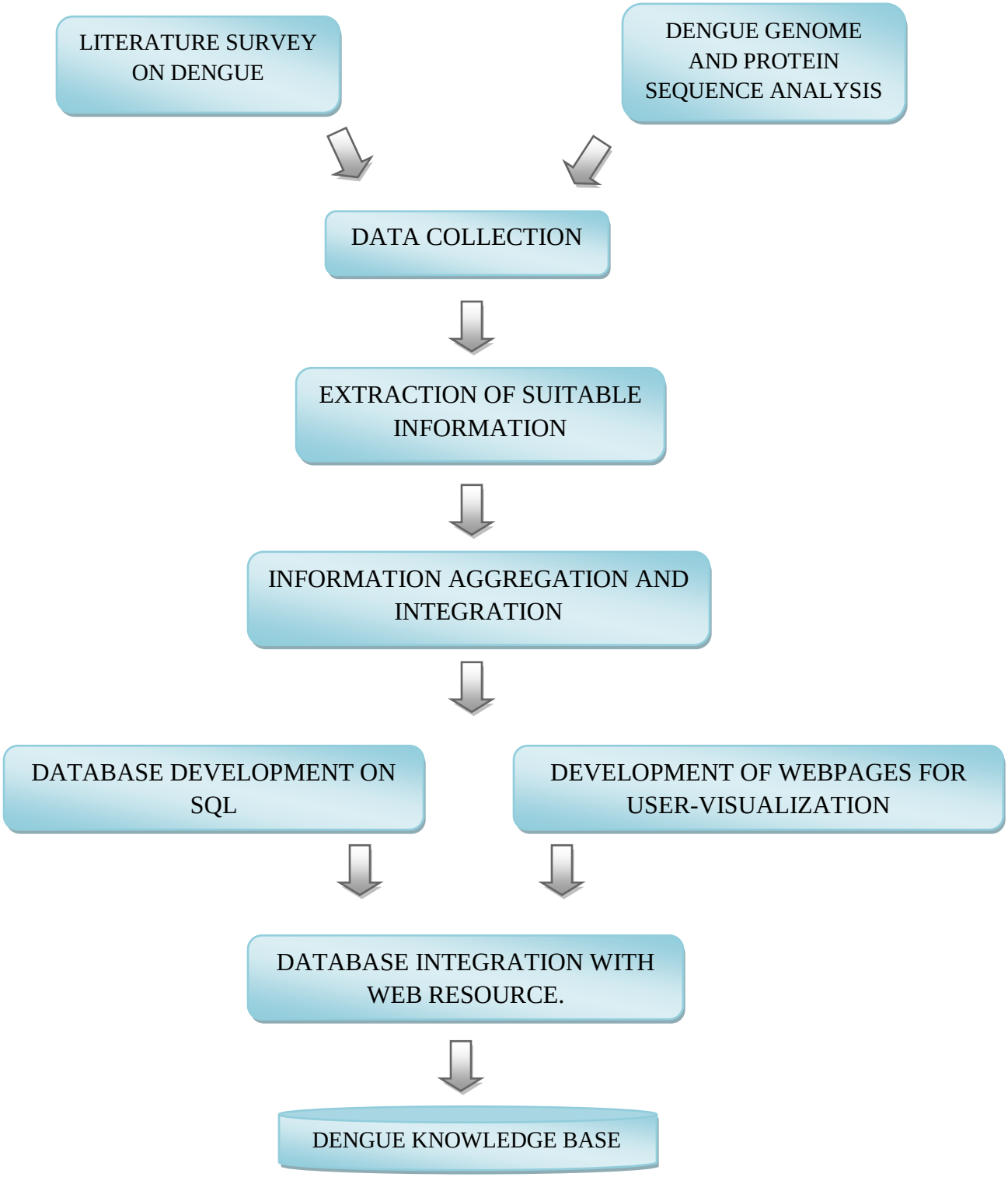


Fig. 14 Methodology adopted in the development of the Dengue Knowledge Base

STEP-1 DATA COLLECTION

The Dengue Knowledge base is a library of information on Dengue, collected from scientific experiments, published literature, high-throughput experiment technology, and computational analyses.

Two types of data were collected for the purpose of its development:

1. Literature reviewed data on Dengue
2. Data obtained after Genome sequence analysis
3. Data obtained after Protein sequence analysis

1.1 COLLECTION OF LITERATURE REVIEWED DATA ON DENGUE:

- More than 50 years of research on dengue has resulted in a host of literature. This plethora of literature was reviewed through non-sequence based information resource, PubMed.
- PubMed is a free database accessing primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics.
- Hundreds of research papers were reviewed to extract relevant information on various aspects of Dengue virus and infection as well as other associated information as:

1. Virus Morphology(Virus genome structure, Replication of virus)
2. Virus Pathogenesis(Immune response generated by humans upon infection)
3. Disease Manifestation (DHF,DF)
4. Organ pathology
5. Control and management of disease
6. Epidemiology of disease
7. Host-virus interactions
8. Plants metabolites showing anti-Dengue activities

- While representing the data in the database, every research finding has been linked to its corresponding PMID so that user can access the whole record.

[A PMID (PubMed identifier or PubMed unique identifier) is a unique number assigned to each PubMed record.]

- References section in the database includes the references of all the accessed research papers.

1.2 COLLECTION OF GENOME STRUCTURAL ANNOTATION DATA

STEPS INVOLVED IN GENOME STRUCTURAL ANNOTATION: TO IDENTIFY GENES AND PREDICT THEIR CORRESPONDING PROTEINS:

In the present study, *Chemgenome 3.0*, (the SCFBio tool) has been used to produce and interpret structural annotations for the viral genome of *Dengue virus*.

Whole genome sequences of all serotypes of Dengue virus(DENV-1, DENV-2, DENV-3, DENV-4) were retrieved from NCBI.

S.No	Dengue virus type	NCBI RefSeq
1.	DENV-1	NC_001477.1
2.	DENV-2	NC_001474.2
3.	DENV-3	NC_001475.2
4.	DENV-4	NC_002640.1

[Table 3. List of NCBI RefSeq Dengue Genome sequences]

Gene are then predicted in all six reading frames (5'-3')and(3'-5') which are then translated into protein sequences using Chemgenome 3.0

Tabular summaries of the obtained data were organized for inclusion in the database.

Fig.15 Steps involved in Dengue Genome structural annotation

1.3 COLLECTION OF PROTEIN SEQUENCE COMPUTATIONAL ANALYSIS DATA

Protein sequence computational analysis was done to obtain information on the following areas:

1) Prediction and annotation of the structural character

I. Primary structural details (Amino acid composition, molar extinction co-efficient, isoelectric point, charged , acidic, basic, polar and hydrophobic amino acid profile etc.).

II. Secondary structural details (Residue wise probability of alpha, beta and turns regions, Surface Probability, Antigenic Index, Flexible Regions, Alpha and Beta, Amphipathic Regions etc.) and Titration Curves

2) T-cell and B-cell Epitope Prediction

3) Conserved Domain Prediction

4) Obtaining Titration curve for each protein

PROTEIN SEQUENCE RETREIVAL :

Co- and post- transcriptional processing of virus polyprotein gives rise to 3 structural and 7 non structural proteins, encoded by genes in order (from 5' to 3').

Protein sequences of all serotypes of Dengue virus(DENV-1, DENV-2, DENV-3, DENV-4) were retrieved from NCBI.

	Protein name	DENV-1	DENV-2	DENV-3	DENV-4
1	Capsid protein C	NP_059433.1	NP_056776.2	YP_001621843.1	NP_073286.1
2	Membrane glycoprotein	NP_722459.2	NP_739592.2	YP_001531167.1	NP_740316.1
3	Envelope	NP_722460.2	NP_739583.2	YP_001531168.2	NP_740317.1

4	Nonstructural protein NS1	NP_722461.1	NP_739584.2	YP_001531169.2	NP_740318.1
5	Nonstructural protein NS2a	NP_733808.1	NP_739585.2	YP_001531170.2	NP_740319.1
6	Nonstructural protein NS2b	NP_733809.1	NP_739586.2	YP_001531171.3	NP_740320.1
7	Nonstructural protein NS3	NP_722463.1	NP_739587.2	YP_001531172.2	NP_740321.1
8	Nonstructural protein NS4a	NP_733810.1	NP_739588.2	YP_001531173.2	NP_740322.1
9	Nonstructural protein NS4b	NP_733811.1	NP_739589.2	YP_001531175.2	NP_740324.1
10	Nonstructural protein NS5	NP_722465.1	NP_739590.2	YP_001531176.2	NP_740325.1

[Table 4: list of retrieved Protein sequences (accession no.)]

A. STEPS INVOLVED IN PROTEIN STRUCTURAL ANNOTATION

Computational resource used: PROTEAN from DNASTAR's Lasergene Core Suite version 10.1

Prediction and annotation of the structural character of all above mentioned protein sequences was done to obtain:

- Primary structural details (Amino acid composition, molar extinction co-efficient, isoelectric point, charged , acidic, basic, polar and hydrophobic amino acid profile etc.).
- Secondary structural details (Residue wise probability of alpha, beta and turns regions, Surface Probability, Antigenic Index, Flexible Regions, Alpha and Beta, Amphipathic Regions etc.) and Titration Curves,

The graphical display and tabular summaries of sequence composition and other data were organized for inclusion in the database.

Fig. 16 Steps involved in Dengue Protein structural annotation

B. STEPS INVOLVED IN PREDICTION OF PROTEIN CONSERVED DOMAINS

Computational resource used: CDD (Conserved Domain Database by NCBI)

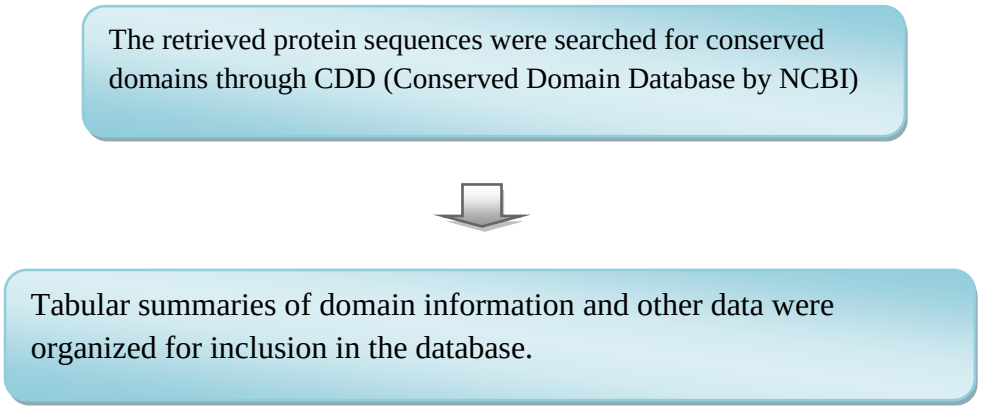


Fig.17 Steps involved in prediction of DENV protein conserved regions

C. STEPS INVOLVED IN PROTEIN T-CELL AND B-CELL EPITOPE PREDICTION

Computational resource used: Immune Epitope Database and Analysis resource(IEDB).

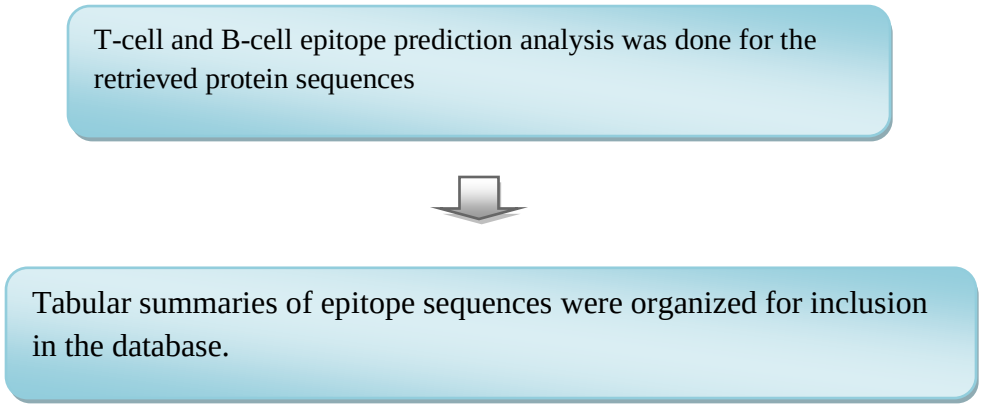


Fig.18 Steps involved in epitope prediction for DENV proteins

D. STEPS INVOLVED TO OBTAIN PROTEIN TITRATION CURVE

Computational resource used: PROTEAN from DNASTAR's Lasergene Core Suite version 10.1

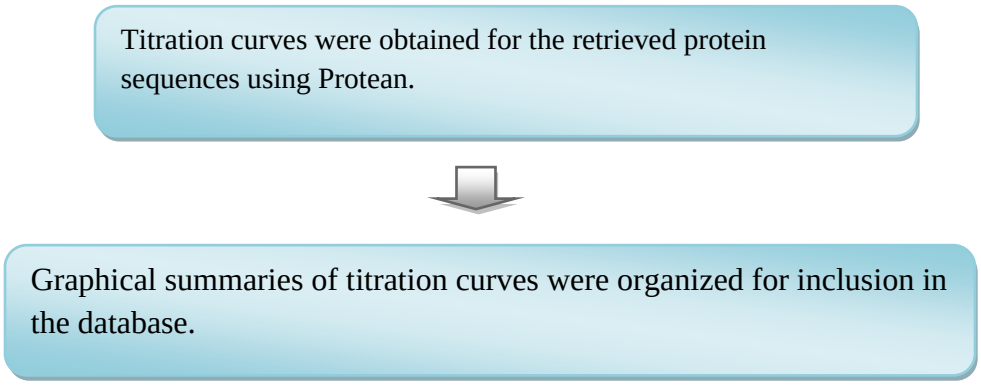


Fig.19 Steps to obtain Titration curves for DENV proteins

STEP-2 DATABASE DEVELOPMENT ON SQL(Backend)

The data was managed in tables using SQL as follows:

Fig. 20 Use of SQL for database development.

STEP-3 DEVELOPMENT OF WEBPAGES FOR USER VISUALIZATION(Front-End)

TOOL USED: MICROSOFT VISUAL STUDIO 2010

Microsoft Visual Studio 2010 offers many new features for development of ASP.NET web applications.

Using .NET Framework 4 tools and technologies in Microsoft Visual Studio 2010, all the above mentioned data (i.e. data collected in Step-1) were represented on Webpages. Around 300 webpages were developed.

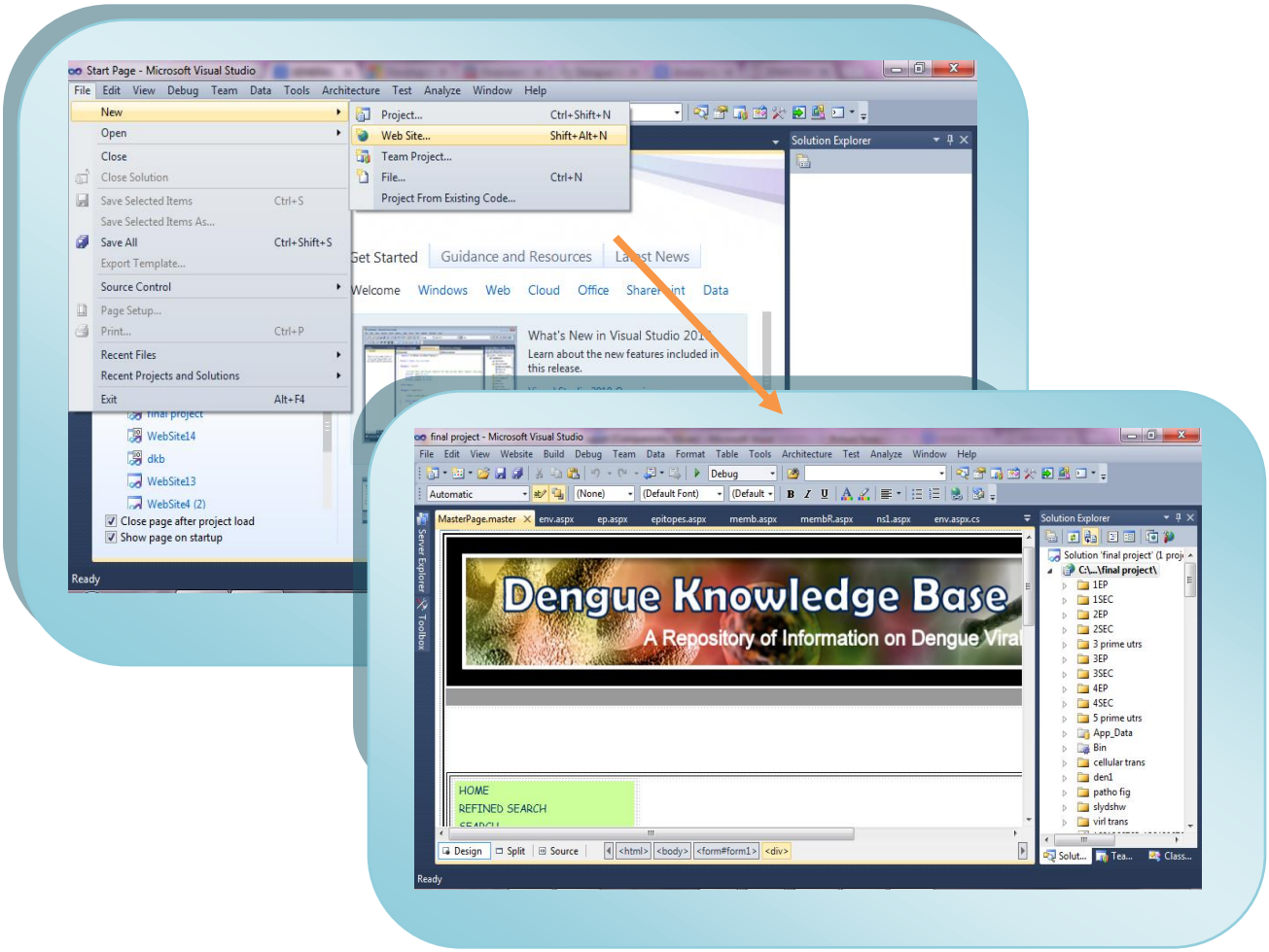
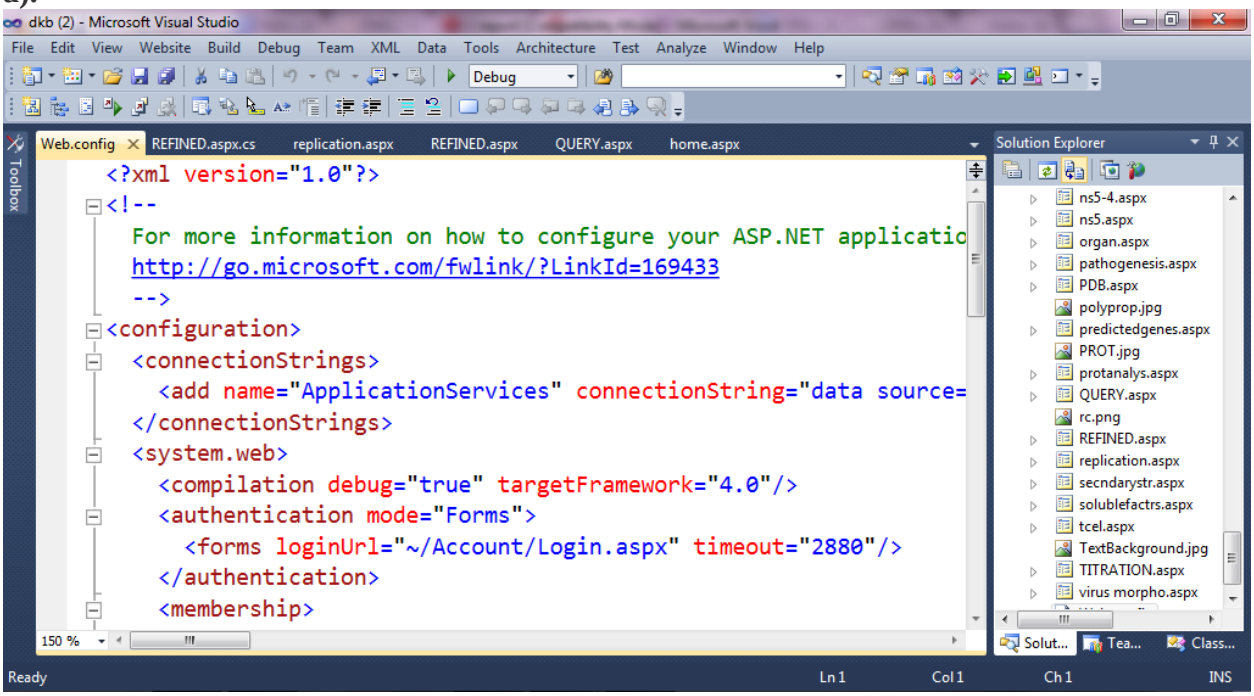


Fig. 21 Use of visual studio 2010 for website development.

STEP-4 DATABASE AND WEBPAGES INTEGRATION

For interfacing the database(backend) built on SQL with the web(front-end), ASP.NET technology was used.

a).



b).

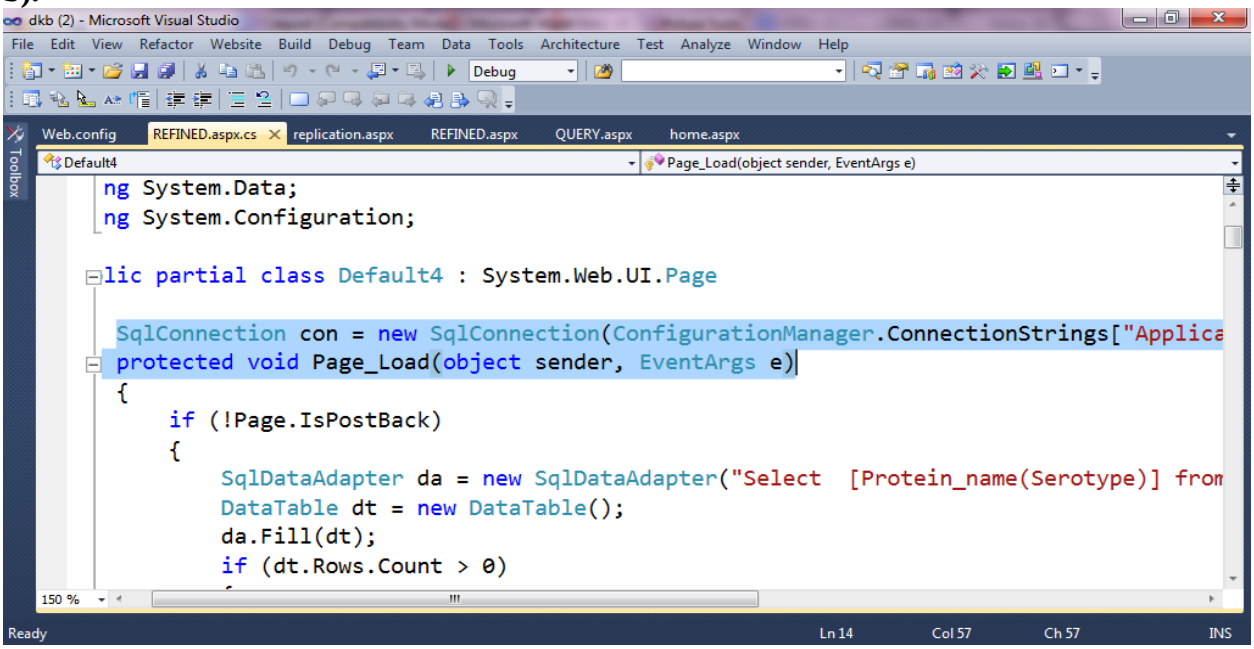


Fig. 22 Snapshots of codes for integrating database into web resource.

RESULTS

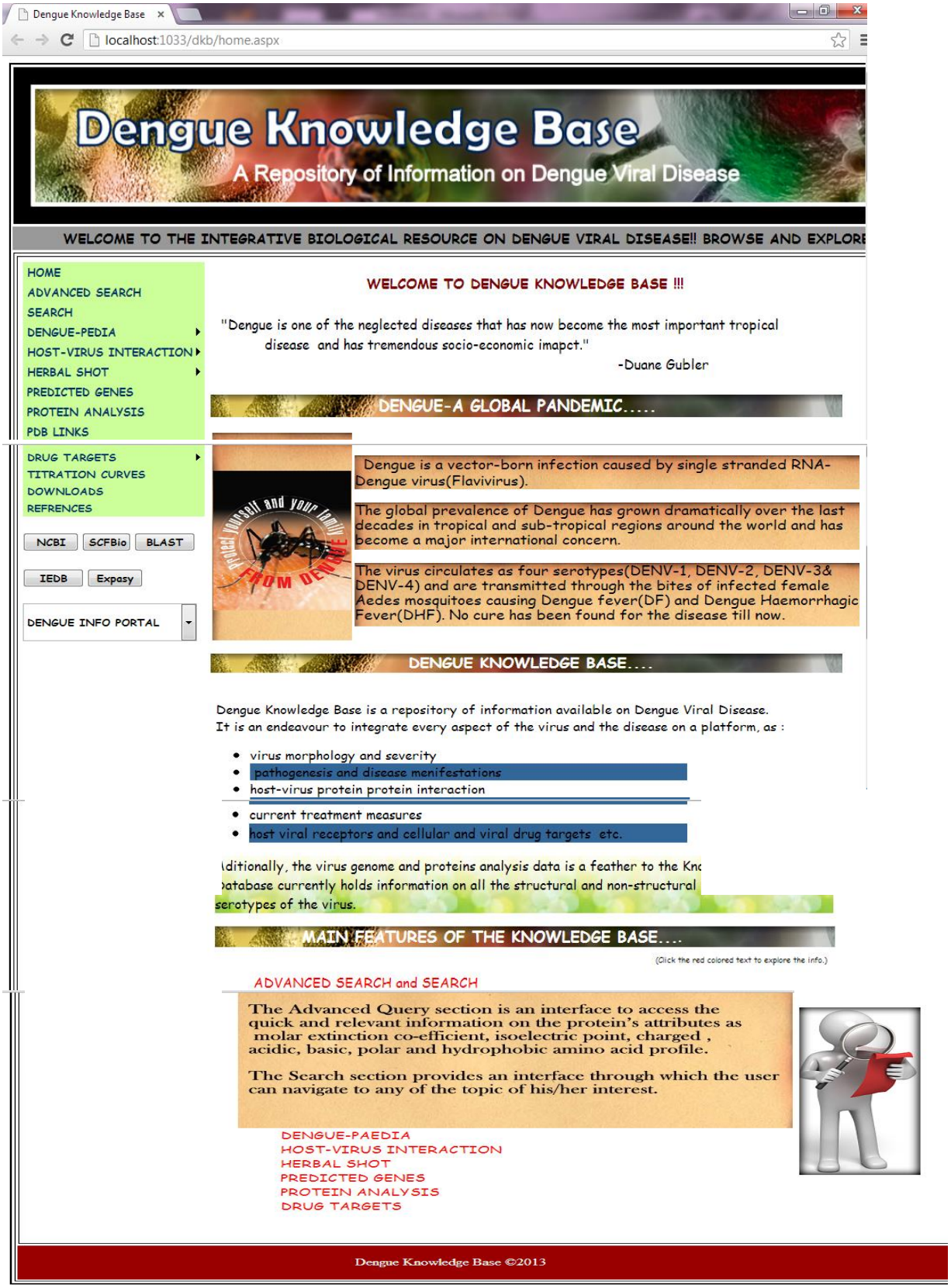


Fig 23. Home Page of Dengue Knowledge Base

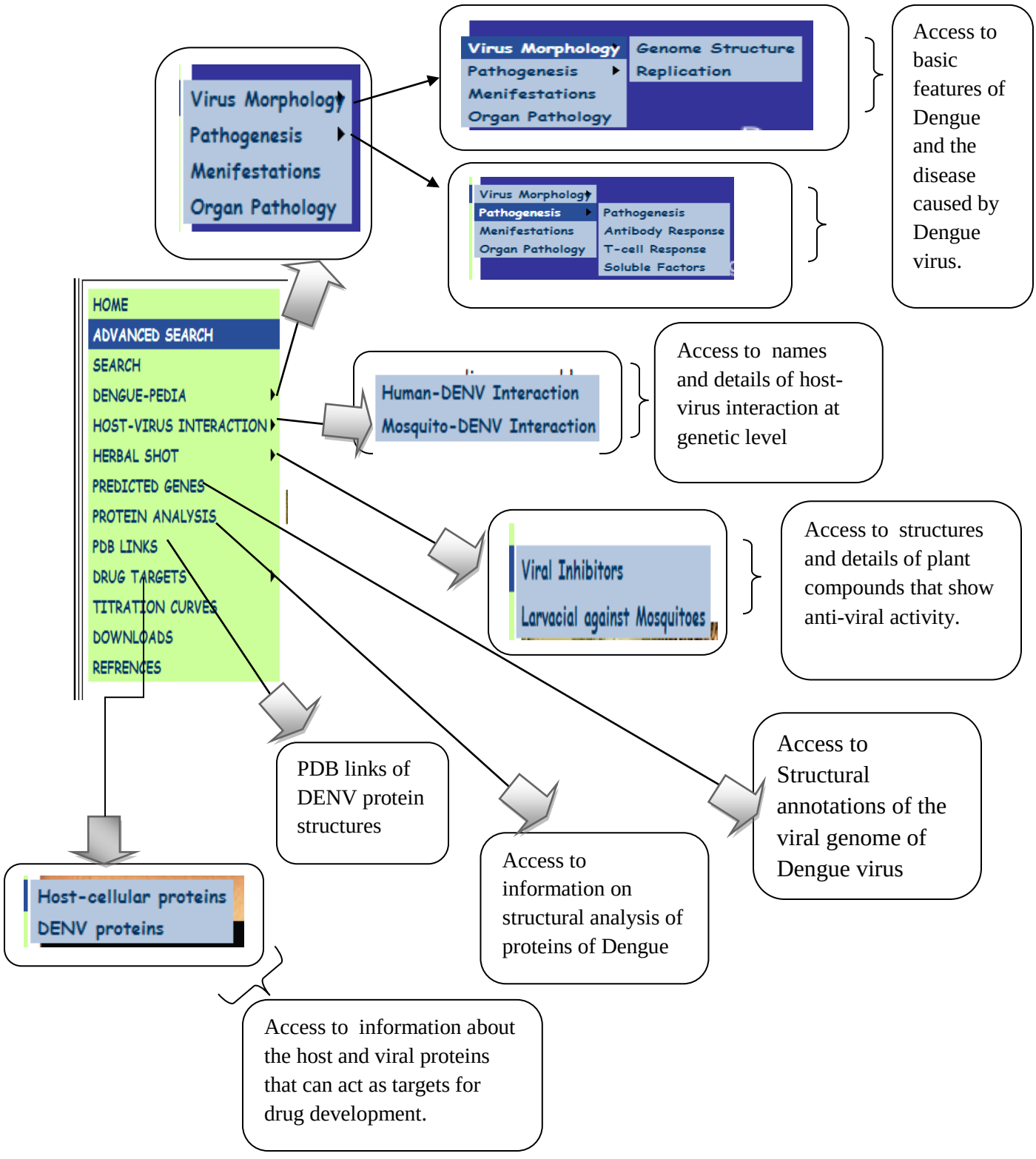


Fig.24 Overview of menu and its subdivisions to navigate through the database

ADVANCED SEARCH:

- The Advanced Query section is an interface to access the quick and relevant information on the protein's attributes as molar extinction co-efficient, isoelectric point, charged, acidic, basic, polar and hydrophobic amino acid profile.

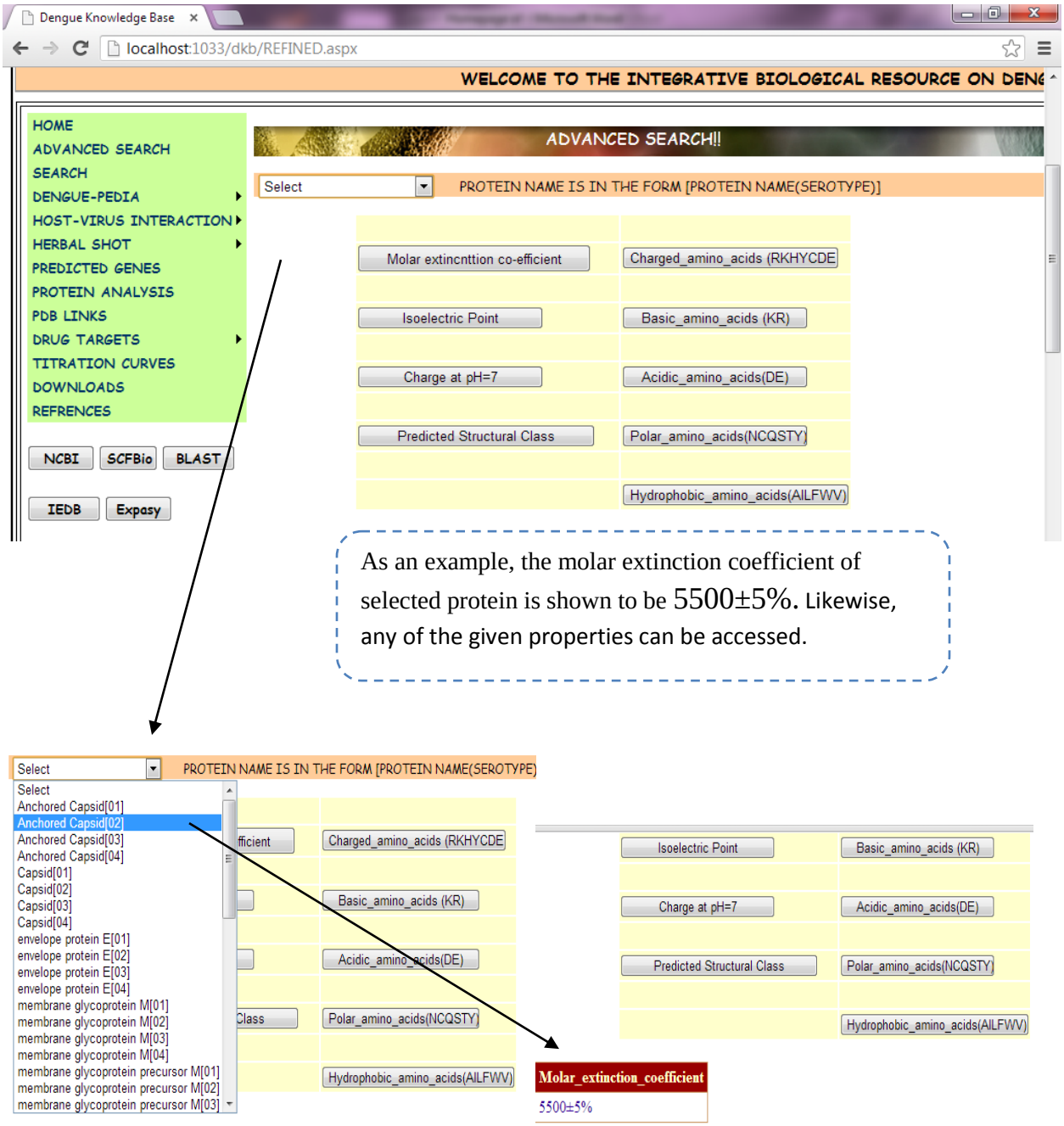


Fig. 25 Web page of DKB showing ADVANCED SEARCH.

SEARCH:

- To make the access more user-friendly and comprehensive, the Search section provides the integrative links to navigate through the entire database.

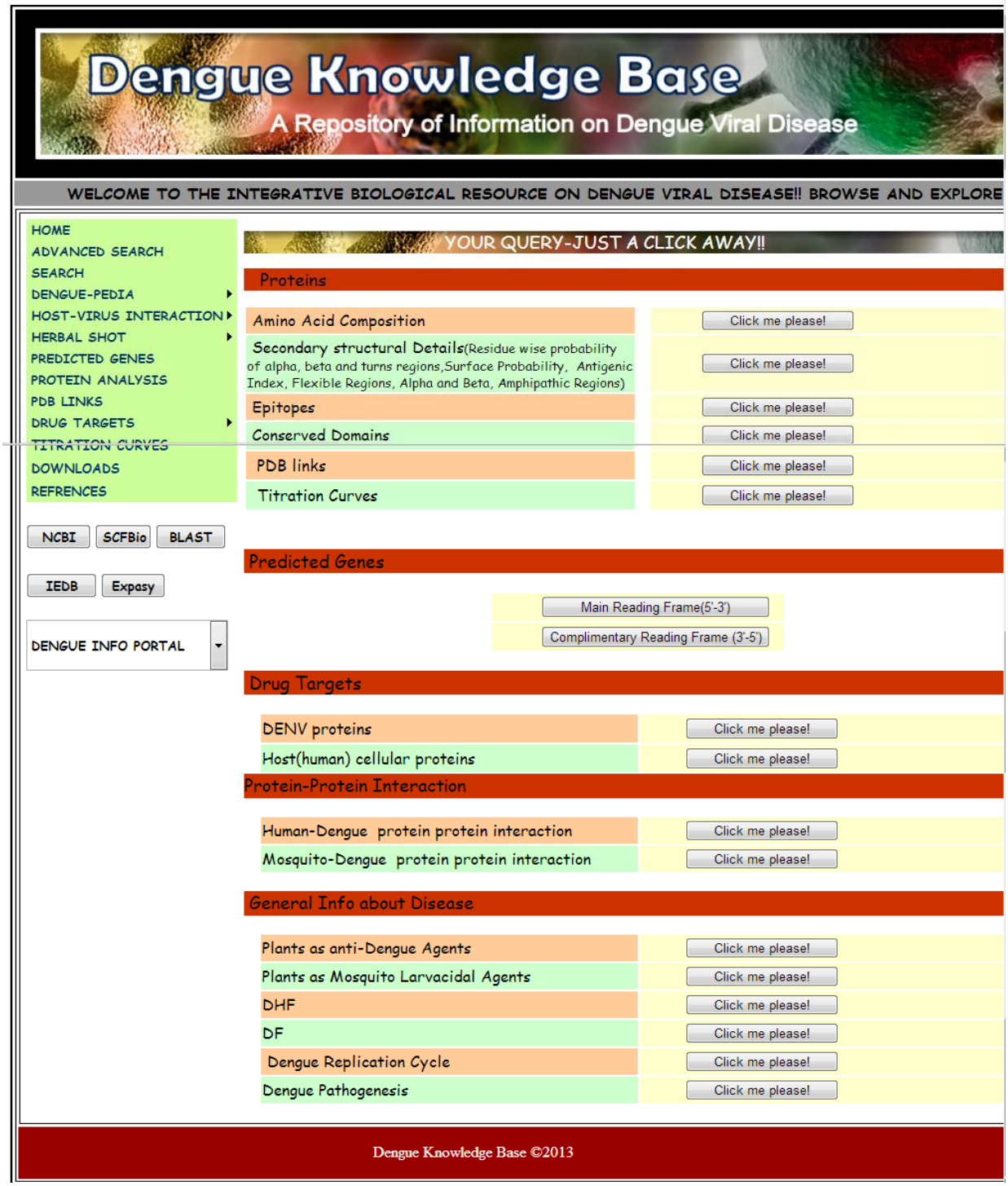


Fig.26 Web page of DKB showing SEARCH options.

DENGUE-PAEDIA:

- The information on virus morphology, replication, pathogenesis, host-infection, manifestations, organ pathology, control and management and treatment that could be accessed through Dengue-paedia (encylopaedia of Dengue) feature of Dengue Knowledge Base.

Navigation to genome structure:

DENGUE-PAEDIA --→ virus morphology --→genome structure

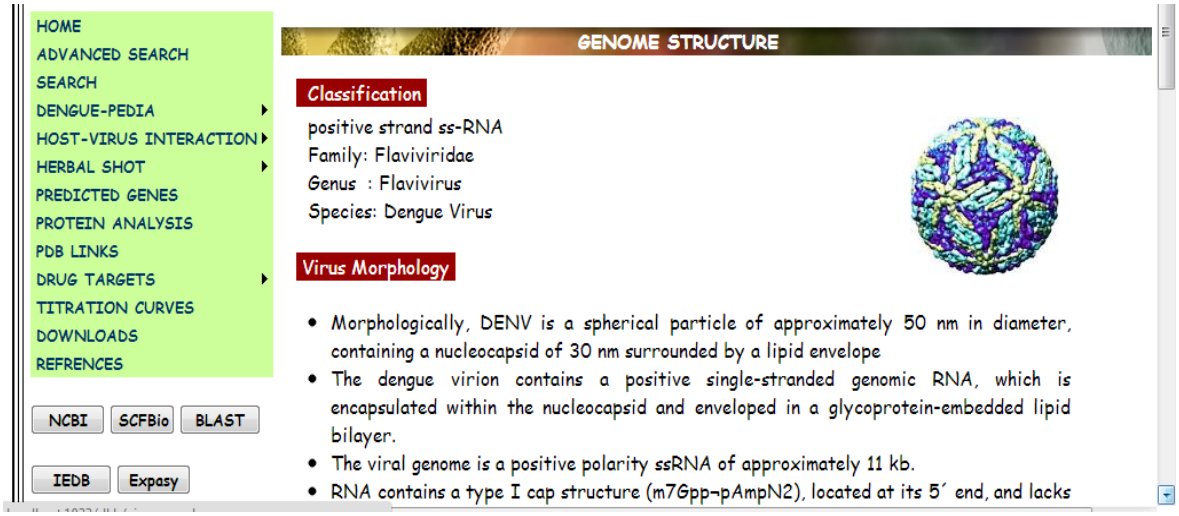


Fig. 27 Web page of DKB showing information on genome structure.

Navigation to virus replication:

DENGUE-PAEDIA --→ virus morphology --→replication

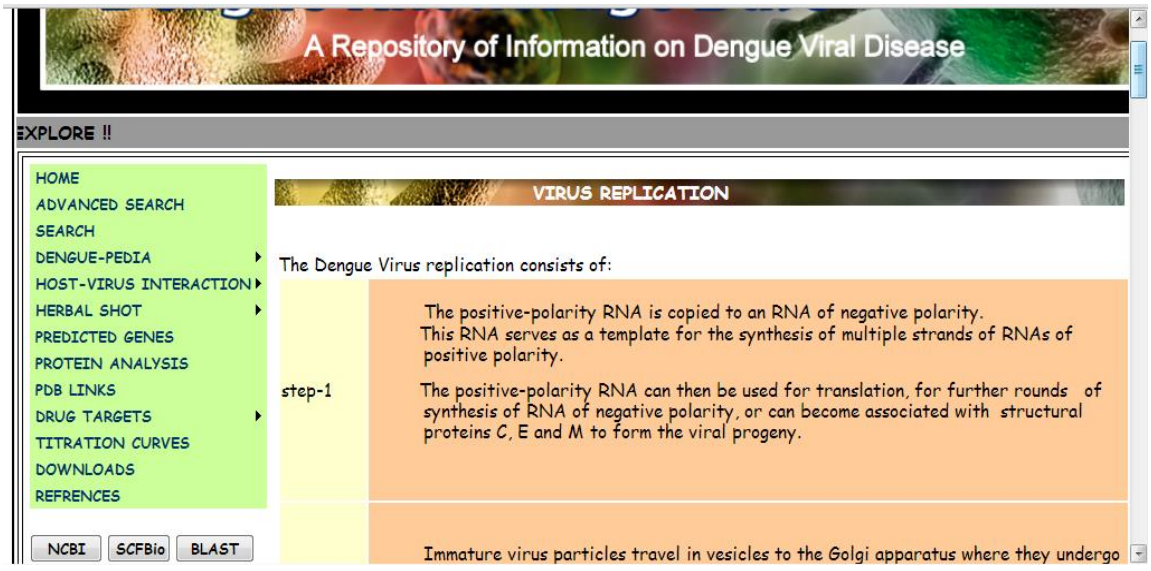


Fig.28 Web page of DKB showing information on genome replication.

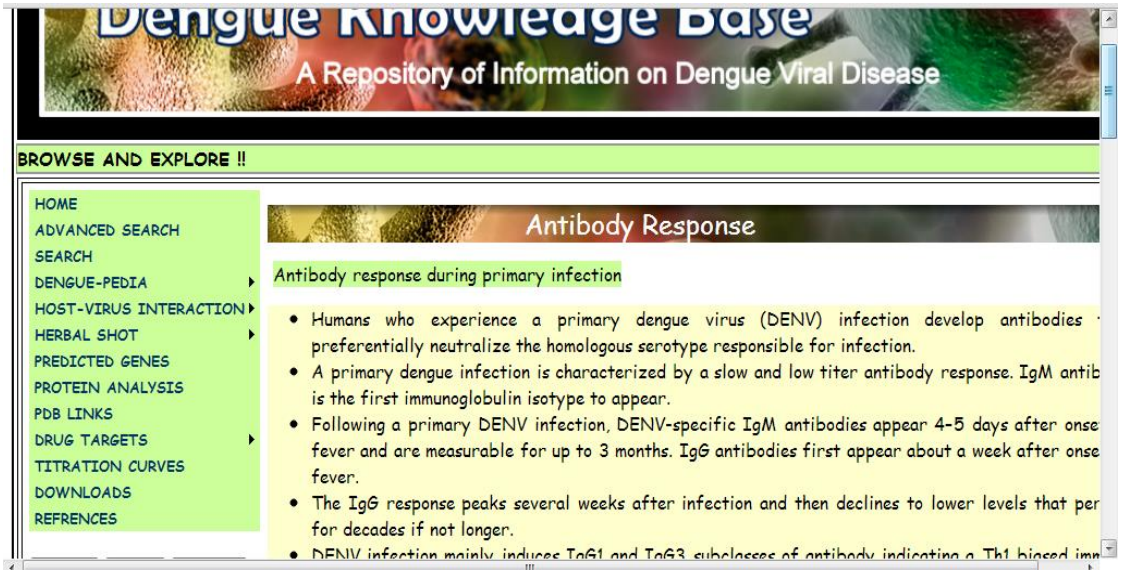
Navigation through pathogenesis:

DENGUE-PAEDIA --→ pathogenesis --→Antibody response

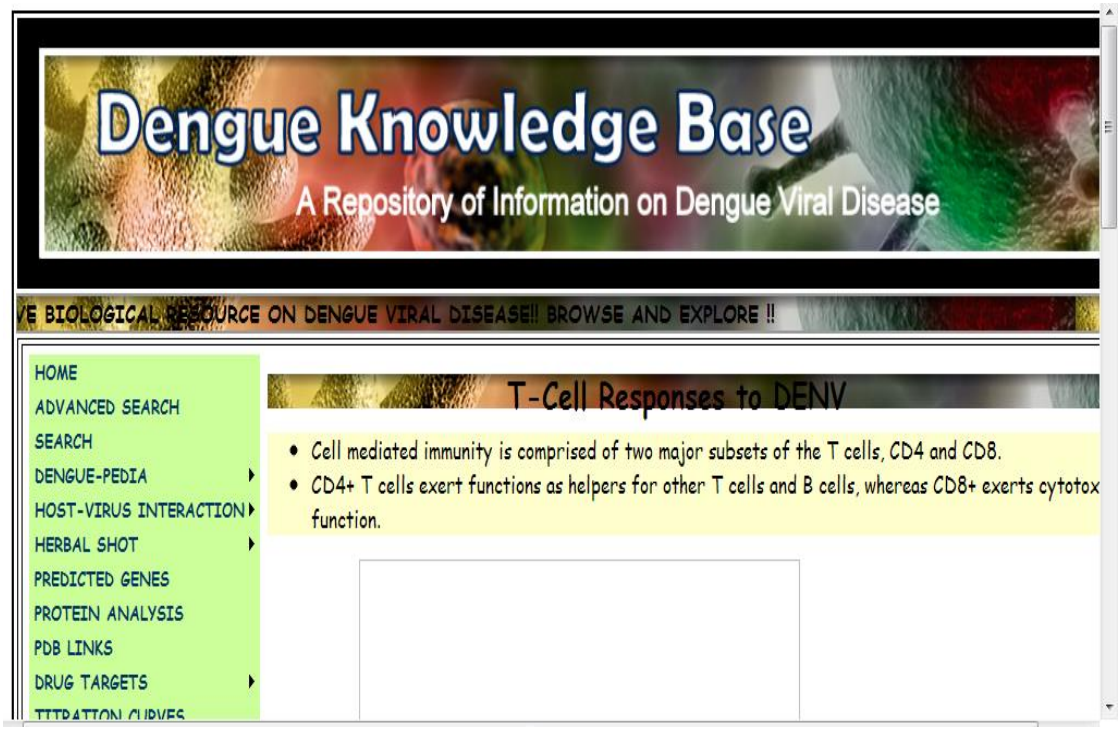
➔ T-cell response

➔ Soluble factors released during pathogenesis

a)



b)



c)

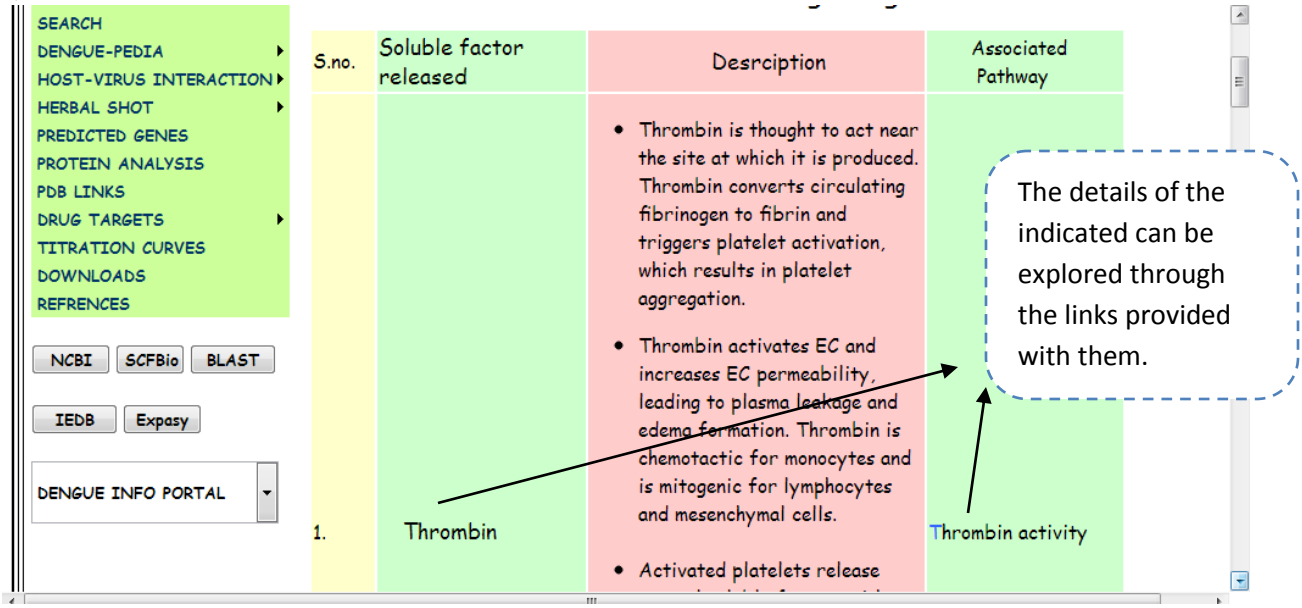


Fig 29 a,b,c. Webpages of DKB showing information on Antibody response, T-cell response and Soluble factors released during pathogenesis respevtively.

<u>Navigation to clinical manifestations:</u>
DENGUE-PAEDIA --→ manifestations

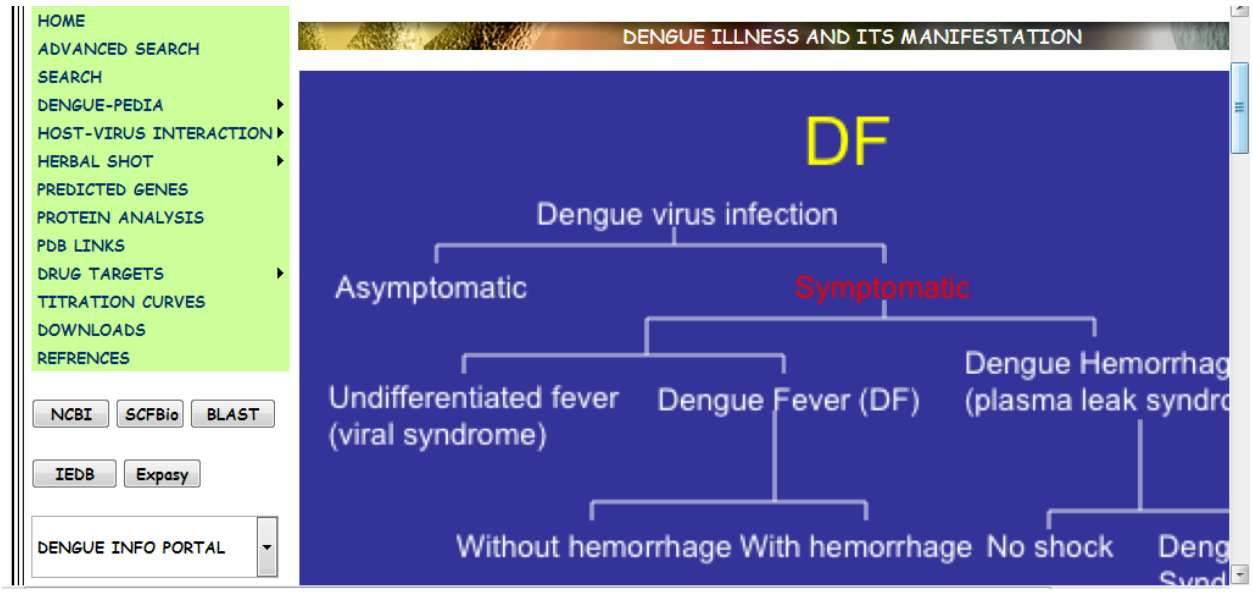


Fig. 30 Webpages showing information on manifestations of dengue illness.

Navigation to organ pathology

DENGUE-PAEDIA --→ organ pathology

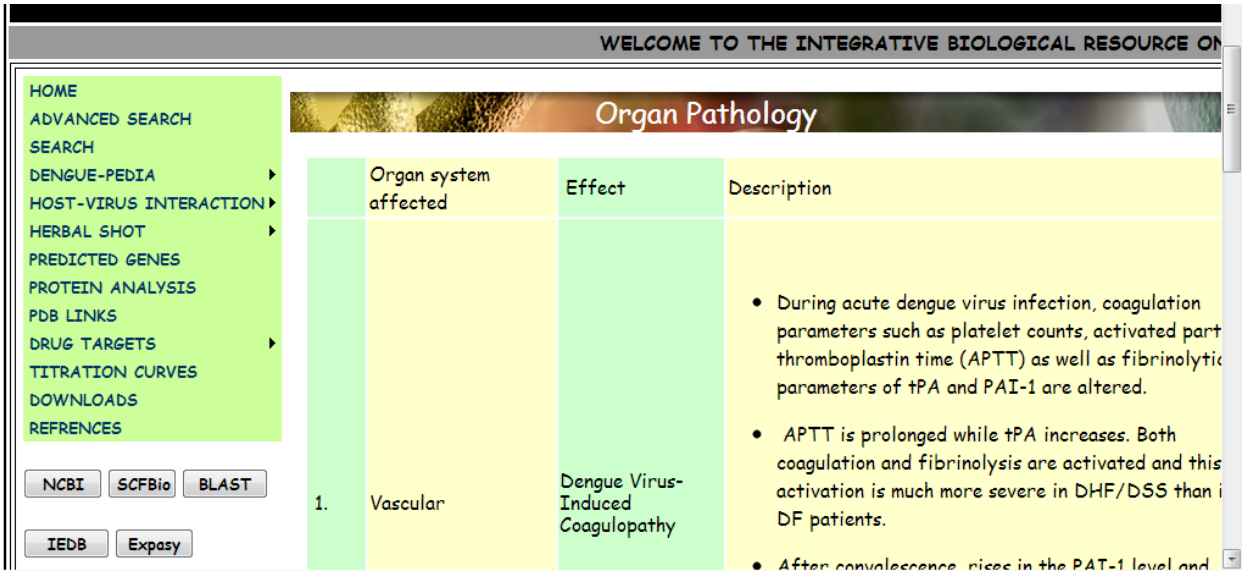


Fig. 31 Webpages showing information on organ pathology during dengue illness.

HOST-VIRUS INTERACTION:

- The host-virus interaction section in the resource emphasized on interactions between the host(including humans and *Aedes* mosquito) and virus at genetic level.

Navigation through host-virus interaction

HOST-VIRUS INTERACTION --→ human-virus interaction

--→mosquito-virus interaction

.

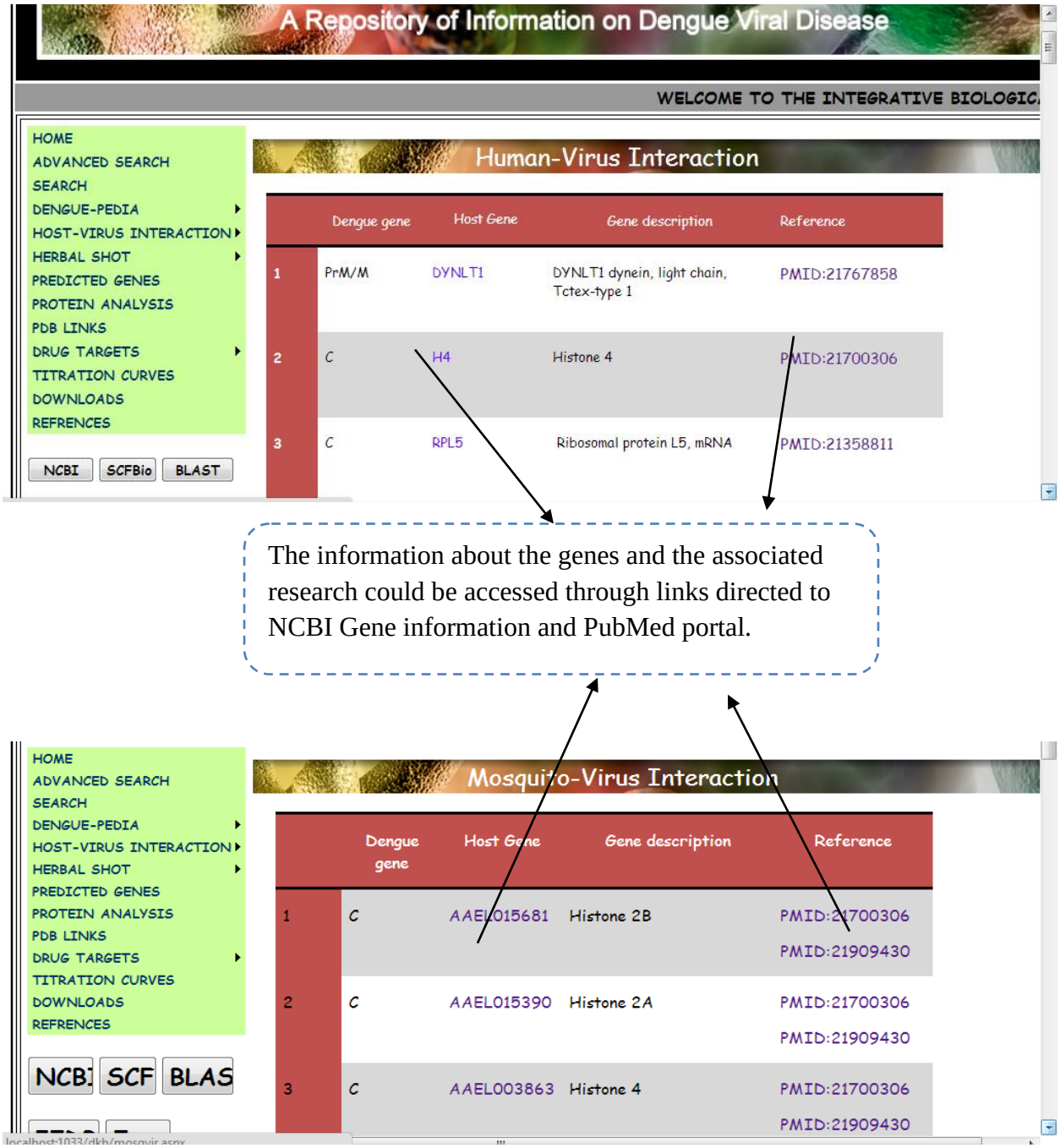


Fig.32 Web page of DKB for HOST-VIRUS INTERACTIONS

HERBAL SHOT:

- The Herbal Shot feature ensembles information on the plants whose compounds show anti-viral or larvacidal activity.

Navigation through herbal shot

HERBAL SHOT --→ viral inhibitors

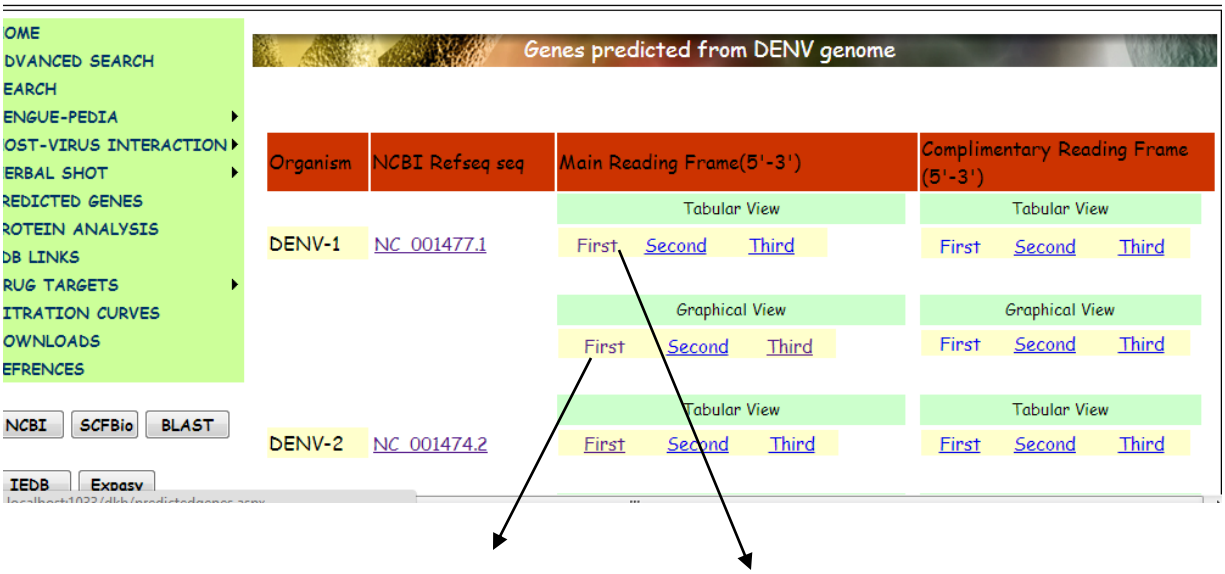
--→larvacidal against mosquitoes.

Fig.33 Web page of DKB for HERBAL-SHOT

PREDICTED GENES:

- The Predicted Genes section puts forward the structural annotations for the viral genome of Dengue virus(as retrieved from NCBI). The prediction of genes was done through computational resource Chemgenome 3.0 in all six reading frames (5'-3')and(3'-5') which are then translated into protein sequences.

Navigation to predicted genes --→PREDICTED GENES



Click on the links provides the graphical and tabular view of genes as predicted by Chemgenome3.0.

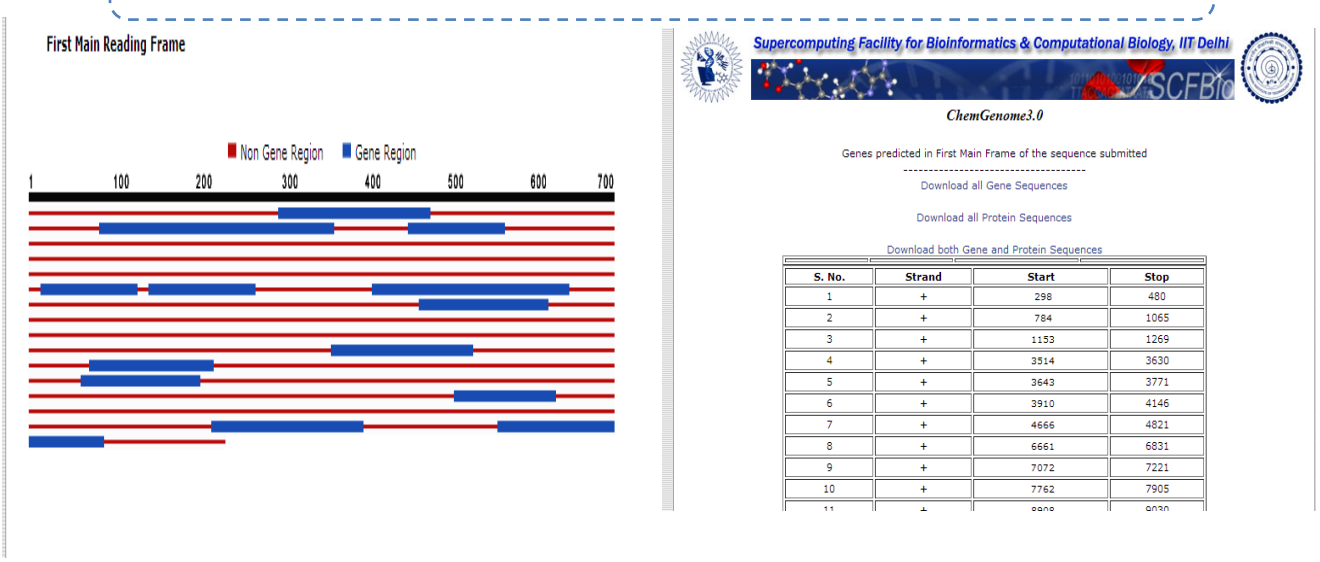


Fig.34 Web page of DKB for PREDICTED GENES

PROTEIN ANALYSIS:

- The protein Analysis section portrays annotation of the structural character three structural and seven non structural proteins(as retrieved from NCBI), encoded by genes in order (from 5' to 3') of all the serotypes of virus.

Navigation to protein analysis -->PROTEIN ANALYSIS

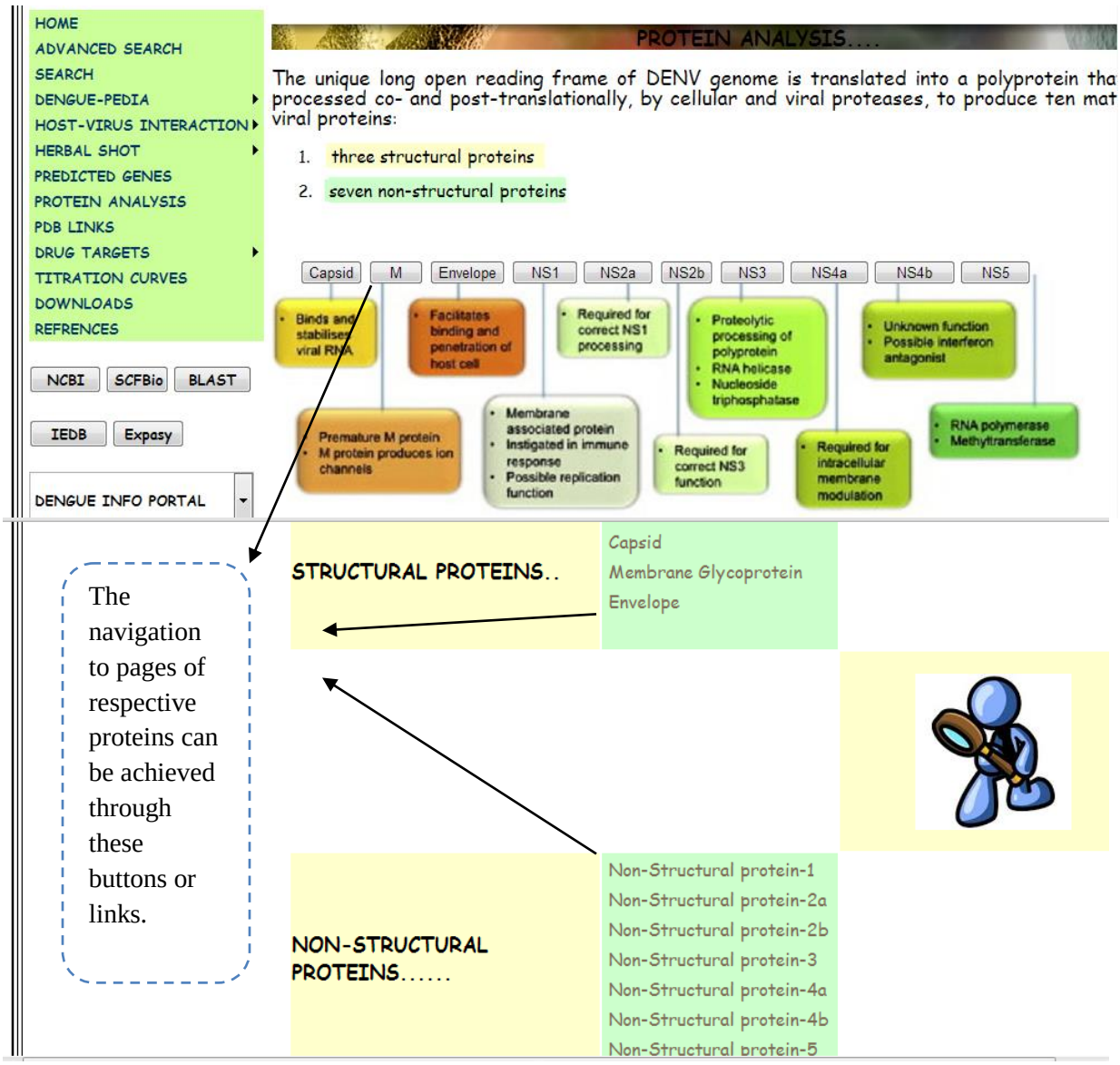


Fig.35 Web page of DKB for PROTEIN ANALYSIS.

A)

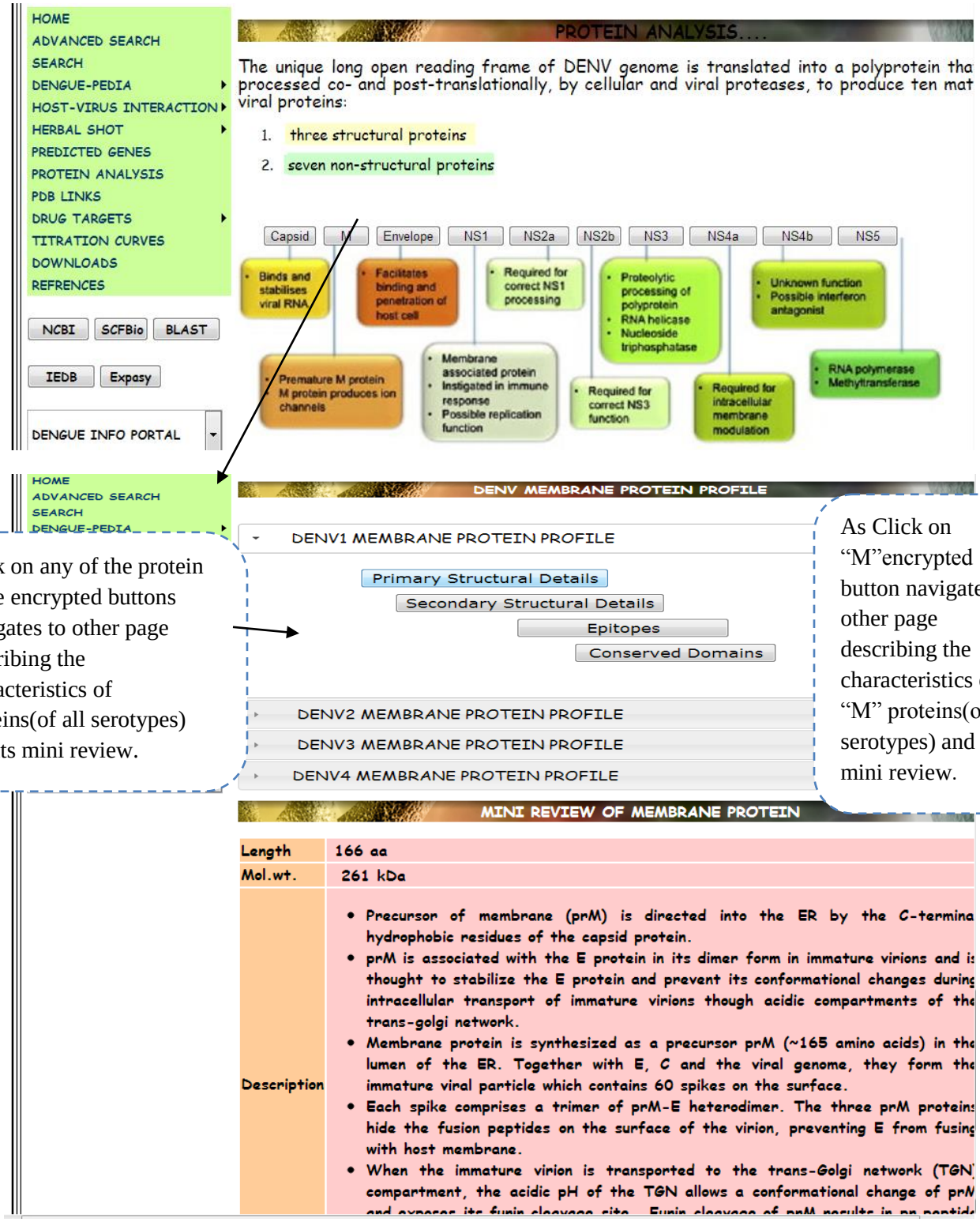
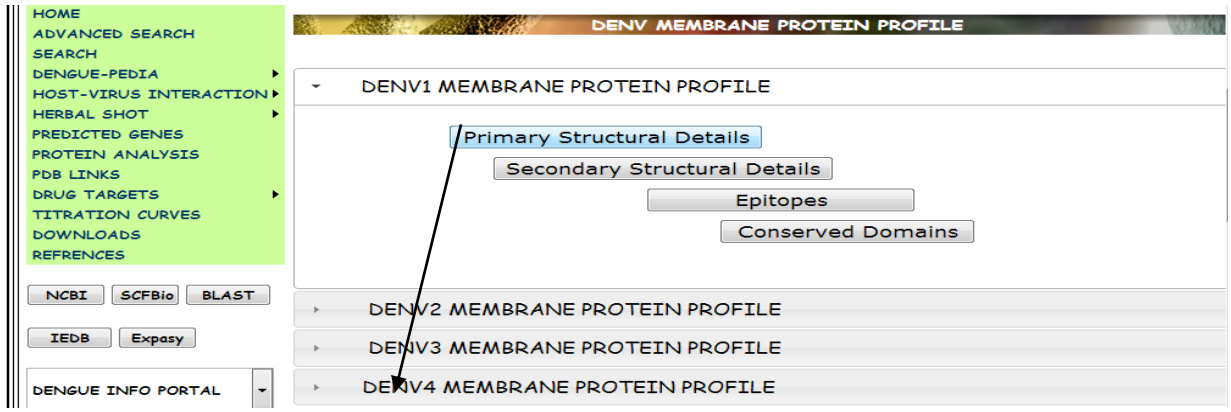


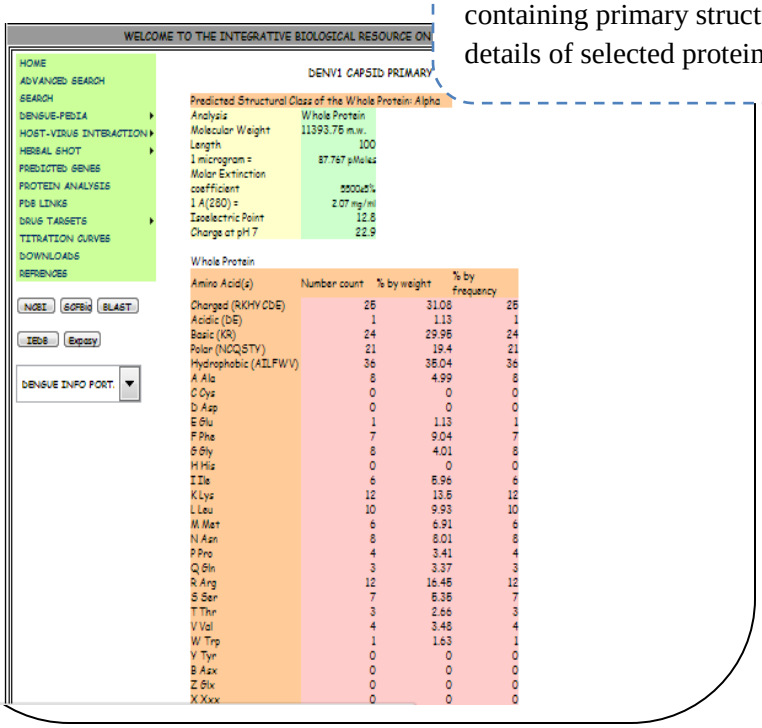
Fig.36 Web page of DKB for PROTEIN ANALYSIS.

B.

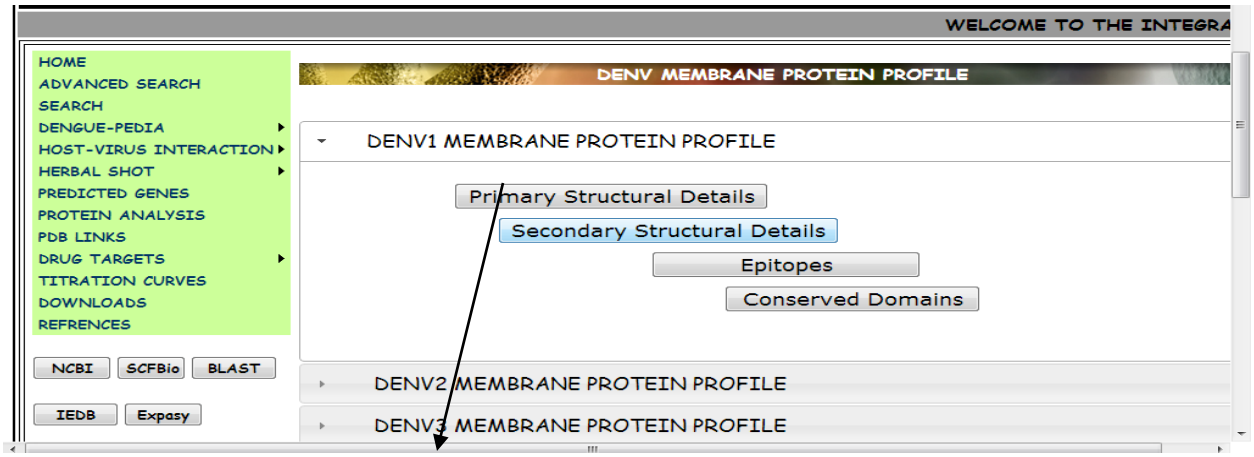


Further click on "primary str. details", redirects to the page containing links to primary structural details of each protein.

Further click on links given on "primary str. details" page, redirects to the page containing primary structural details of selected protein.

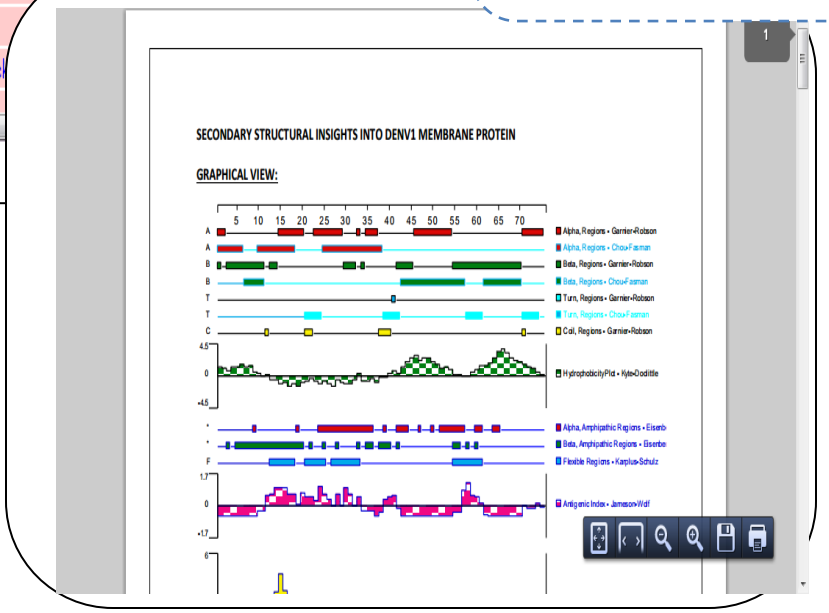


C.

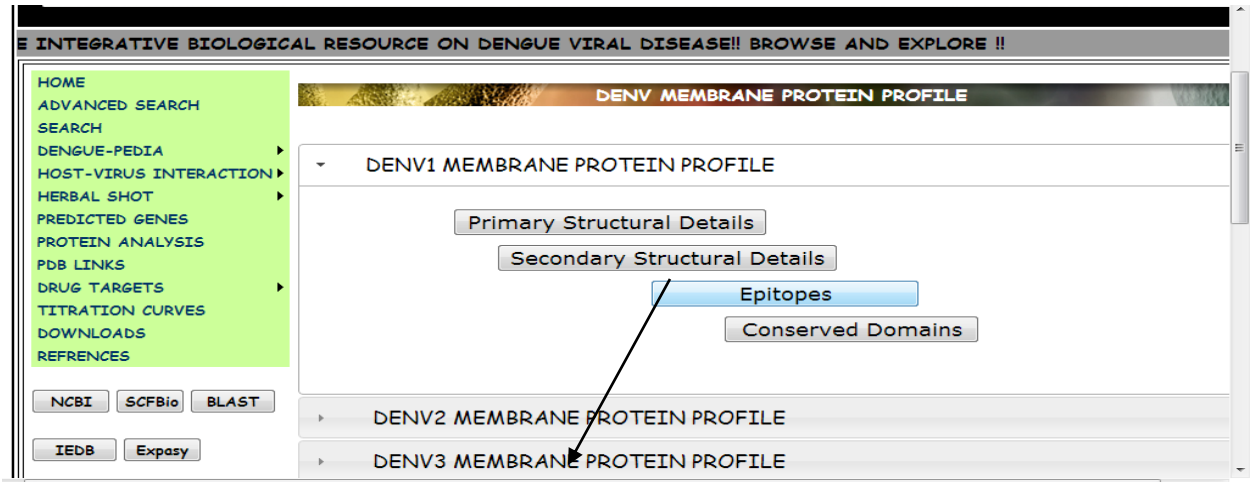


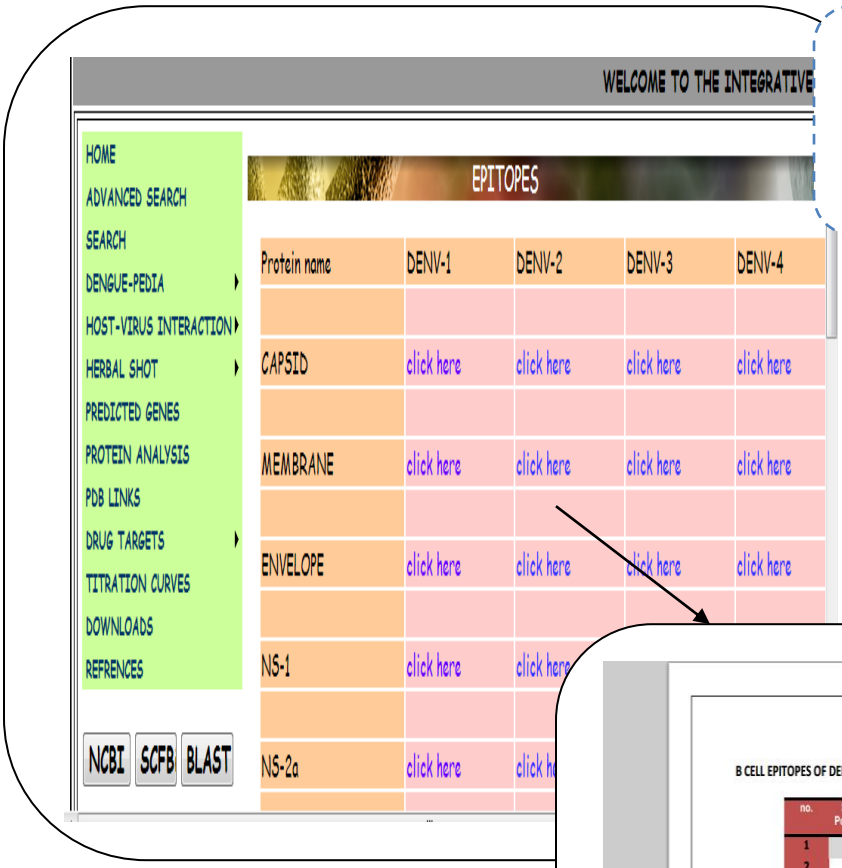
Further click on "secondary str. details", redirects to the page containing links to secondary structural details of each protein.

Further click on links given on "secondary str. details" page, redirects to the pdf page containing secondary structural details of selected protein.



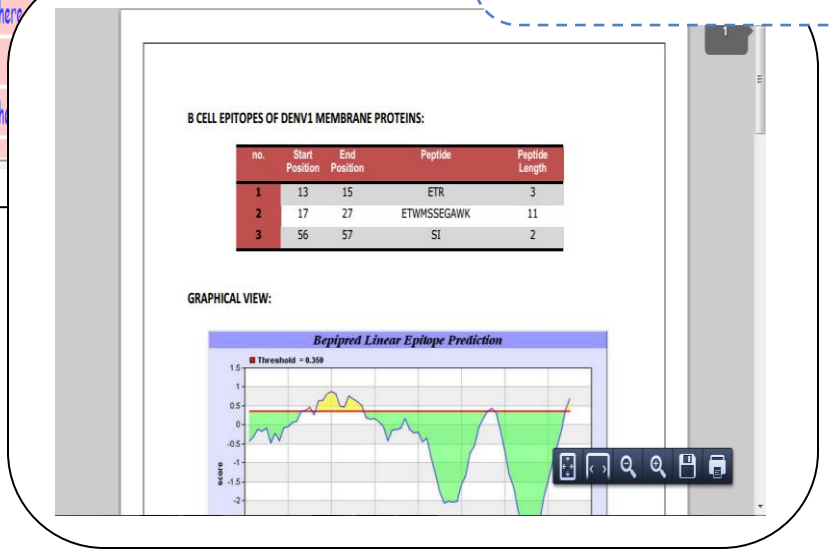
D.



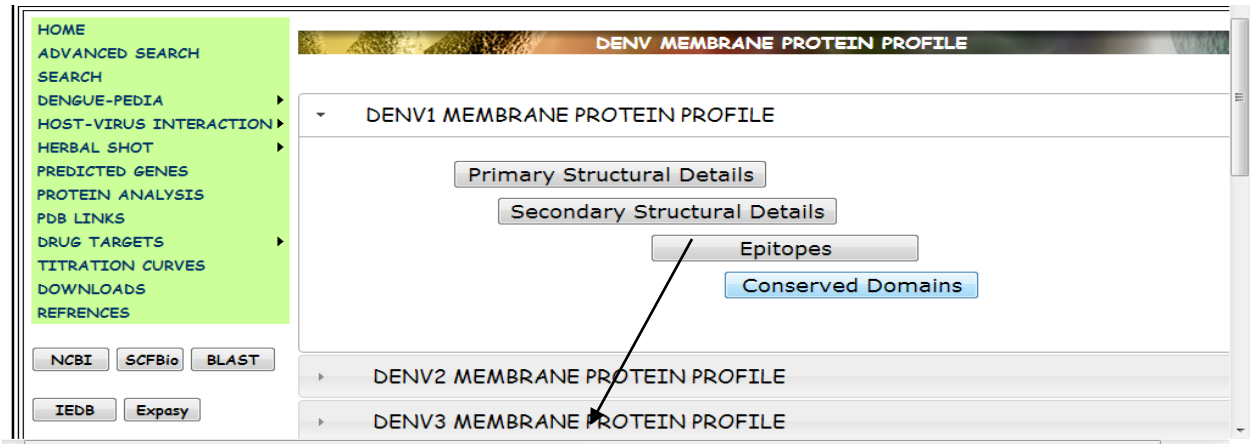


Further click on "epitopes. details", redirects to the page containing links to epitope details of each protein.

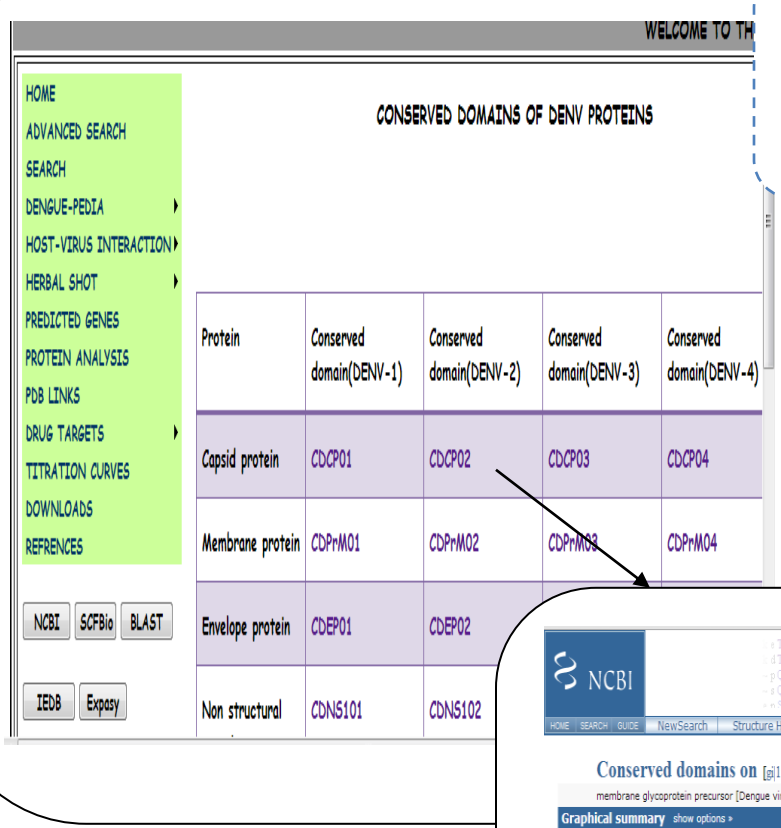
Further click on links given on "epitopes details" page, redirects to the pdf page containing epitope regions of selected protein.



E.



Further click on "conserved domains", redirects to the page containing links to conserved domains of each protein.



Further click on links given on "conserved domains" page, redirects to the CDD tool containing conserved domains of selected protein.

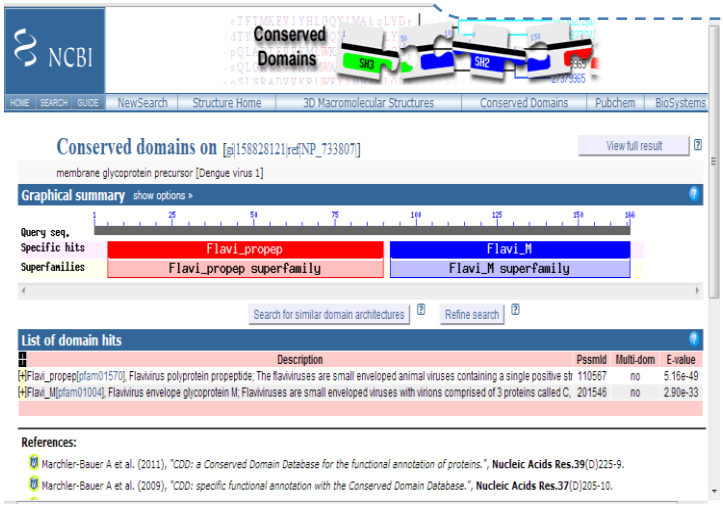


Fig.37 A, B, C, D, E – Insights into webpages of DKB exhibiting Protein Analysis details

DRUG TARGETS:

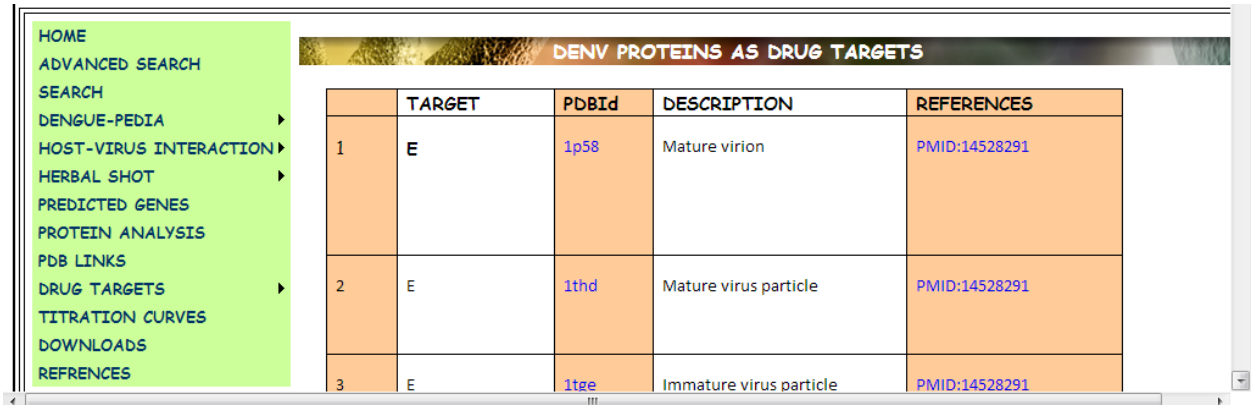
The Drug Target feature includes the information on viral and host(human) cellular parts that can act as targets for drug development.

Navigation to drug targets

DRUG TARGETS --→ viral proteins

--→human cellular proteins.

A)



B)

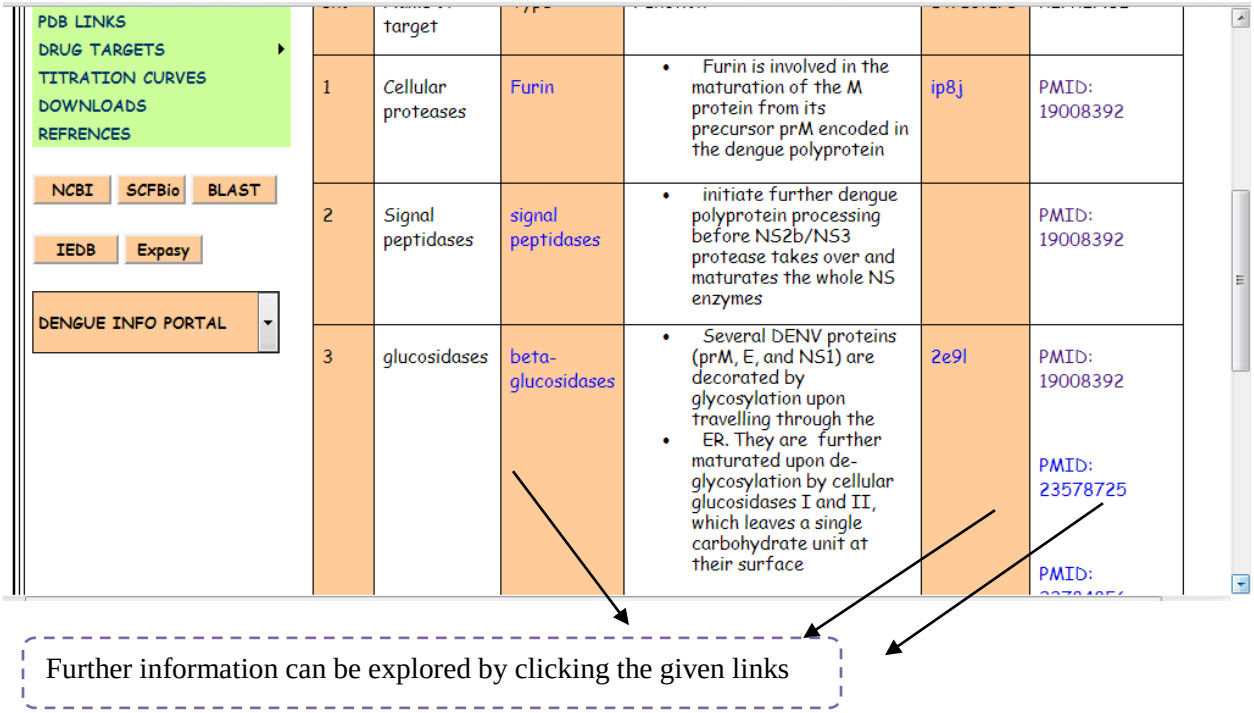


Fig. 38 webpage of DKB for Drug Targets.

TITRATION CURVES:

- The Titration curve section provides the titration curve of each DENV protein.

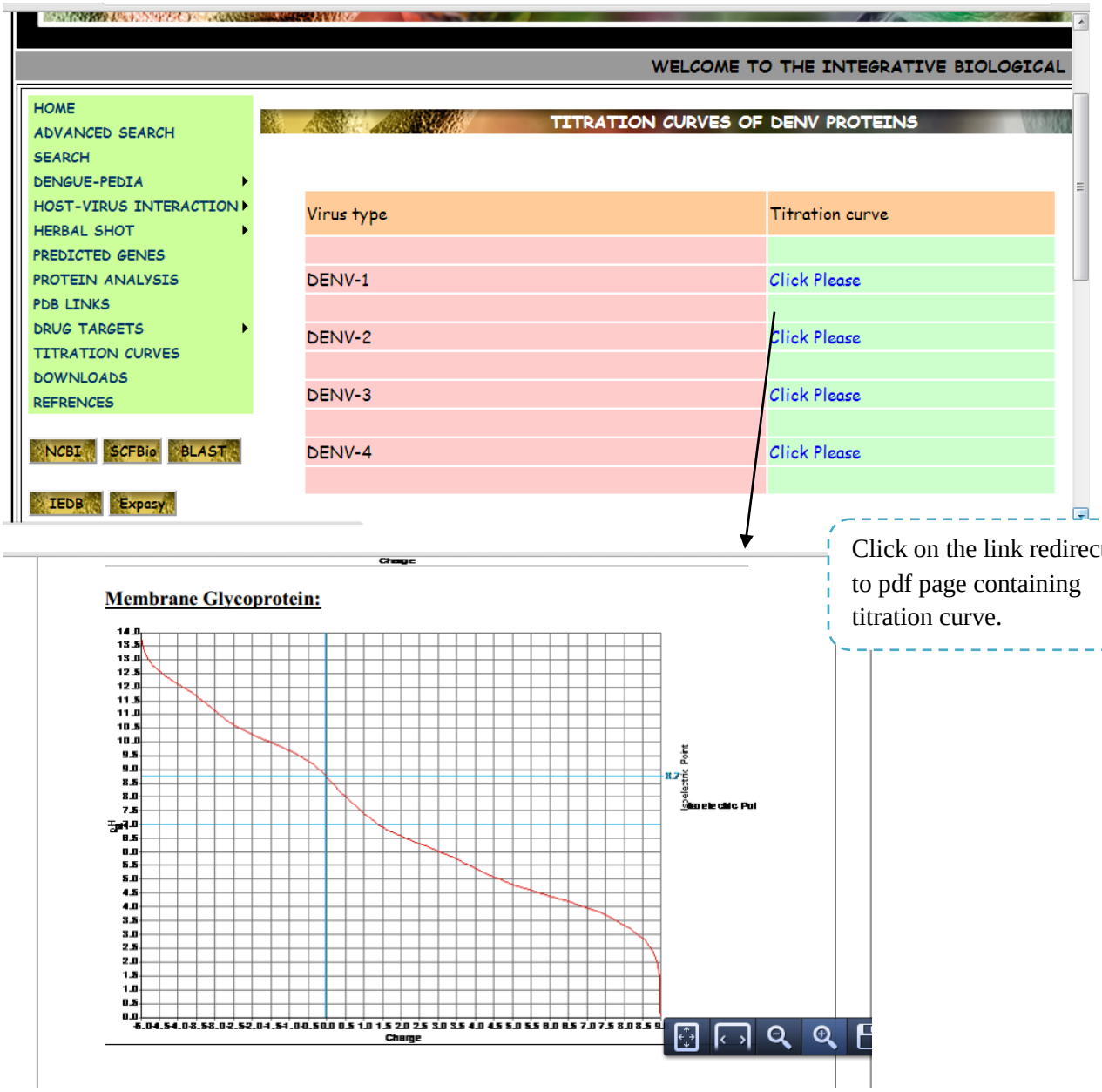


Fig.39 webpage of DKB for titration curve

PDB LINKS

- This feature assembles the known-protein crystal structures deposited in PDB and provides a link to access the respective structure.

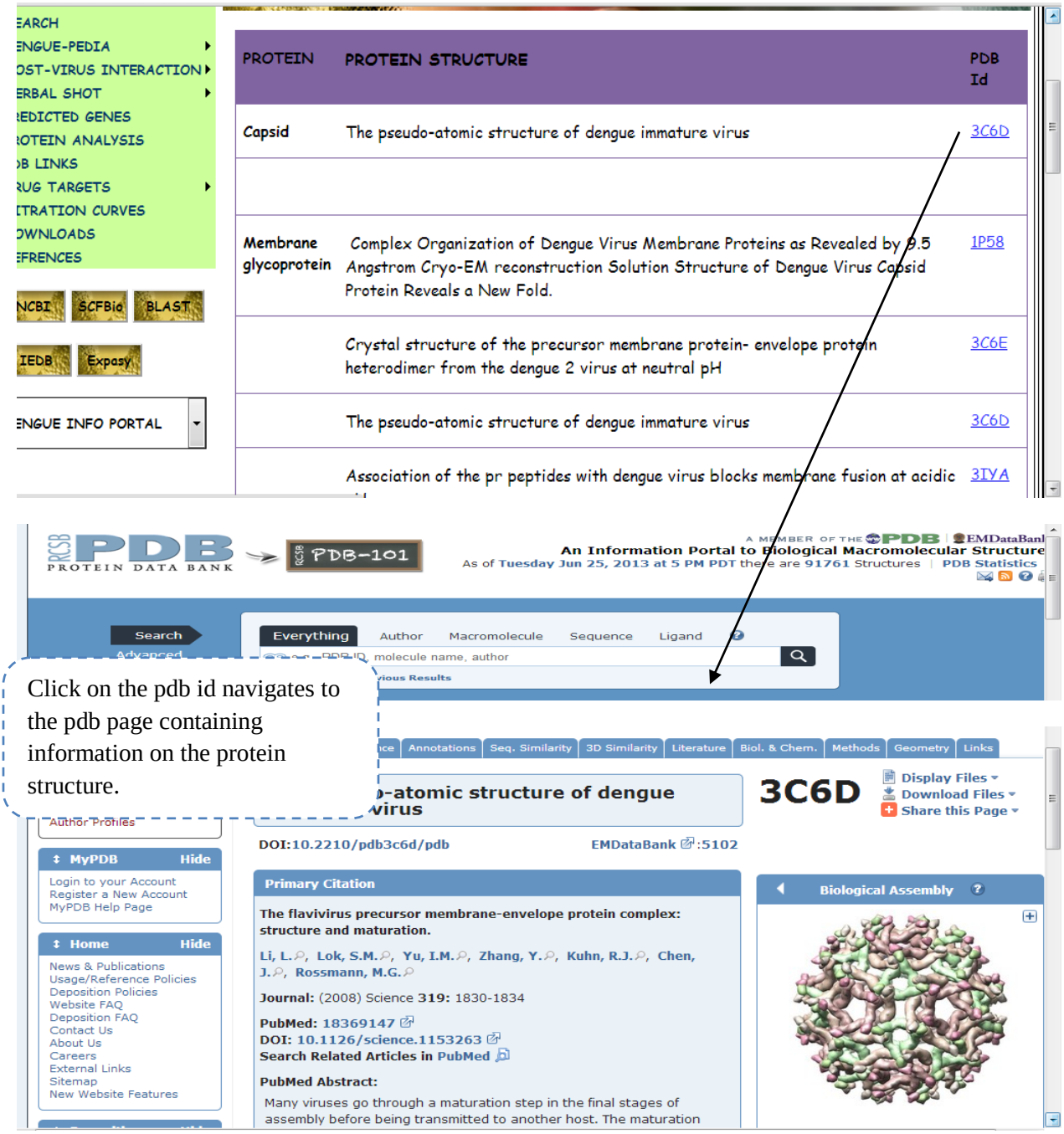


Fig.40 webpage of DKB for PDB links

DISCUSSION

A challenge for the computational biology field arising from the new opportunities is to expand training and educational opportunities at the interface of biology with computer and information science, bioengineering, and the physical sciences. Computational science is now poised to be a partner in the armament of biological tools, maintaining an essential triangle of theory, computation, and experimentation in scientific areas.

There is a plethora of online published research papers presenting scientific experiments, high-throughput experiment technology and computational analyses on Dengue virus and the viral disease manifestations as DF and DHF. But very less web resources have been developed exclusively for Dengue virus that consolidates the available information in a user friendly manner on one platform.

There exist few databases that provide information regarding the virus, but comparatively more information is provided in Dengue Knowledge Base where curation of most features of the virus and the infection is done to produce searchable and comprehensive review.

Dengue Knowledge Base is such an endeavour to integrate various facets of online available information on Dengue viral disease. It is developed to provide a comprehensive web-based biological resource to the scientific community and act as a repository of information on Dengue viral disease. The intended audience(user) for the database are researchers, academicians and students.

And to an extent covers most of the features of dengue virus and provides an insight into its genetic and protein profile. The current version of Dengue Knowledge Base

- Supports literature reviewed information on virus morphology, replication, pathogenesis, host-infection, manifestations, organ pathology, control and management,treatment, interactions between the host(including humans and *Aedes* mosquito) and virus at genetic level, natural compounds(derived from plants) that show anti-viral or larvacidal activity, vaccine initiatives taken so far and viral and host cellular targets for drug development.

- Besides, it holds information on structural annotations for the viral genome of Dengue virus(4 genome sequences) and protein sequences(40 proteins) that was done using computational resources.

The information retrieved from the database could be useful for comparative analysis studies, drug and vaccine development and structural and functional analysis studies.

However, there is always a room for improvement. The database is expected to continue to grow and improve by allowing curation of more viral genes and proteins that are submitted online. Additional research is also needed, as there are a number of drug and vaccine trials in progress. A continuous updating of information is required.

CONCLUSION AND FUTURE PROSPECTIVE

This project was undertaken to create a database for Dengue viral disease to provide a comprehensive web-based biological resource to the scientific community. The database is named "**Dengue Knowledge Base- A repository of information on Dengue viral disease".**

The SQL(Database development technology) and Visual studio 2010(Website development technology) have been used to

- Integrate the collected information.
- Construct a user-friendly web interface for Dengue Knowledge Base.

The virus is transmitted to humans by the mosquitoes Aedes aegyptii and Ae. albopictus. The incidence of dengue fever (DF), dengue hemorrhagic fever (DHF), and dengue shock syndrome (DSS) are rapidly increasing, and more than 2.5 billion people live in regions endemic for the disease. Presently approximately 50-100 million cases of DF occur yearly with more than 500,000 resulting in severe and potentially fatal forms of the disease (DHF & DSS). Several factors contribute to the threat posed by dengue. Most significant are lack of cross-reactive immunity for the four DEN serotypes (DEN-1, DEN-2, DEN-3, and DEN-4), hyperendemic circulation of the four different serotypes in the same geographical area, frequent worldwide travel, high population density, and lack of effective mosquito control programs. Other socioeconomic factors only amplify the challenge of dengue control.

From the present work following conclusions can be drawn:

- There exist few databases that provide information regarding the virus, but comparatively more information is provided in Dengue Knowledge Base where curation of most features of the virus and the infection is done to produce searchable and comprehensive review.
- Literature reviewed data and the data obtained after analysis of Dengue genome and proteins(taking into account all the serotypes) have been coupled to produce the repository of information.
- The plethora of literature was reviewed through non-sequence based information resource, PubMed, to access and assemble the information on virus morphology, replication, pathogenesis, host-infection, manifestations, organ pathology, control and management and treatment that could be accessed through Dengue-pedia (encylopaedia of Dengue) feature of Dengue Knowledge Base.
- The host-virus interaction section in the resource emphasized on interactions between the host(including humans and *Aedes* mosquito) and virus at genetic level. The information about the genes and the associated research could be accessed through links directed to NCBI Gene information and PubMed portal.
- The Herbal Shot feature ensemble information on the plants whose compounds show anti-viral or larvacidal activity. The structure of compound deposited in PubChem ,the

information of plants from plant database Tropicos and research work from PubMed could be accessed through links provided to their portal.

- The Predicted Genes section put forward the structural annotations for the viral genome of Dengue virus(as retrieved from NCBI). The prediction of genes was done through computational resource Chemgenome 3.0 in all six reading frames (5'-3')and(3'-5') which are then translated into protein sequence.
- The protein Analysis section portrayed annotation of the structural character three structural and seven non structural proteins(as retrieved from NCBI), encoded by genes in order (from 5' to 3') of all the serotypes of virus. The protein sequence analysis was accomplished using computational resources Protean (Dnastar lasergene coresuite version10.1), Immune Epitope analysis resource and Database(IEDB) and NCBI Conseved Domain Database(CDD). The analysis included:

    - Primary structural details (Amino acid composition(number count, % by weight, % by frequency), molar extinction co-efficient, isoelectric point, charged , acidic, basic, polar and hydrophobic amino acid profile etc.).
    - Secondary structural details (Residue wise probability of alpha, beta and turns regions, Surface Probability, Antigenic Index, Flexible Regions, Alpha and Beta, Amphipathic Regions etc).
    - T-cell and B-cell Epitope.
    - Conserved Domains.
    - Titration curve for each protein.

- The Drug Target feature has included the information on viral and host(human) cellular parts that can act as targets for drug development.
- The Titration curve section has provided the titration curve of each DENV protein.
- The PDB links feature has provided access to all the known crystal structures of protein deposited at PDB.
- The Downloads section contains the genome, gene and protein sequences used in the analysis part.
- To make the access more user-friendly and comprehensive, the Search section has provided the integrative links to navigate through the entire database.
- The Advanced Query section is an interface to access the quick and relevant information on the protein's attributes as molar extinction co-efficient, isoelectric point, charged , acidic, basic, polar and hydrophobic amino acid profile.

The provision of links to online bioinformatics resources like NCBI, Expasy, SCFBio (supercomputing facility) etc allows user to associate more to the search.

The information on Dengue Virus is now just a click away with Dengue Knowledge base!

FUTURE PROSPECTIVES:

However, there is always a room for improvement. The database is expected to continue to grow and improve by allowing curation of more viral genes and proteins that are submitted online. Additional research is also needed, as there are a number of drug and vaccine trials in progress. A continuous updating of information will be required.

REFERENCES

- Bentsi-Enchill AD, Schmitz J, Edelman R, Durbin A, Roehrig JT, Smith PG, Hombach J, Farrar J.(2013).Long-term safety assessment of live attenuated tetravalent dengue vaccines: Deliberations from a WHO technical consultation. Vaccine, 31(23):2603-9.
- Blacksell SD, Newton PN, Bell D, Kelley J, Mammen MP Jr, Vaughn DW, Huthlekanun V, Sungkakum A, Nisalak A, Day NP, (2006). The comparative accuracy of 8 commercial rapid immunochromatographic assays for the diagnosis of acute dengue virus infection. Clin Infect DIS *42*, 1127-34.
- Broor S, Dar L, Sengupta S, Chakaraborty M, Wali JP, Biswas A, Kabra (1997) Recent Dengue Epidemic in Delhi, India. In: Saluzzo JE and Dodet B, editors. Factors in the emergence of arboviruses disease. Paris: Elsevier 123-127.
- Freedman DO, Weld LH, Kozarsky PE, Fisk T, Robins R, von Sonnenburg F(2006) Spectrum of disease and relation to place of exposure among ill returned travelers. J Med. 12;354(2):119–30.
- Garg P, Nagpal J, Khalrnar P, Seneviratne SL, (2008) Economic burden of dengue infections In India. Trans R Soc Tropical Med Hyg , *102*, 570-77.
- Gubler DJ, (1997) Dengue and dengue hemorrhagic fever, history and resurgence as a global public health problem, Wallingford. UK CAB International. 1-12.
- Gubler DJ(2002). Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st Century. Trends Microbiol. 10, 100–103
- Guha-Sapir, D., and B. Schimmer. (2005). Dengue fever: new paradigms for a changing epidemiology. Emerging. Themes Epidemiology. 2:1.
- Guilarde, A. O., M. D. Turchi, J. B. Siqueira, Jr., V. C. Feres, B. Rocha, J. E. Levi, V. A. Souza, L. S. Boas, C. S. Pannuti, and C. M. Martelli. (2008). Dengue and dengue hemorrhagic fever among adults: clinical outcomes related to viremia, serotypes, and antibody response. J. Infect. Dis. 97:817–824.
- Gupta E, Dar L, Narang P, Srivastava VK, Broor S (2005) Serodiagnosis of dengue during an outbreak at a tertiary care hospital in Delhi. Indian J Med Res 121**:** 36-38.
- Halstead, S. B. (2007). Dengue. Lancet. *370*:1644–1652.
- Harris, E., E. Videa, L. Perez, E. Sandoval, Y. Tellez, M. L. Perez, R.Cuadra, J. Rocha, W. Idiaquez, R. E. Alonso, M. A. Delgado, L. A. Campo, F. Acevedo, A. Gonzalez, J. J.

Amador, and A. Balmaseda. (2000). Clinical, epidemiologic, and virologic features of dengue in the 1998 epidemic in Nicaragua. Am. J. Trop. Med. Hyg. 63:5–11.

- Jatanasen S and Thongcharoen P (1993) Dengue hemorrhagic fever in South East-Asian countries. Monograph on dengue/dengue haemorrhagic fever. NewDelhi: WHO 23-30.
- Johnson, George; Tai Te Wua (2000). "Biological Database and its applications: 30 years after the first variability plot". Nucleic Acids Res 28 (1): 214–218.
- Julia Schmitz, John Roehrig, Alan Barrett, Joachim Hombach.(2011).Next generation dengue vaccines: A review of candidates in preclinical development.Vaccine *29*, 7276–728.
- Kennedy RS, (1912) Some notes on an epidemic of Dengue-form fever amongst Indian troops, Calcutta Indian Med Gar, XVII 42.
- Kittigui, P. Pitakarnjanakul, D. Sujirarat, and K. Siripanichgon. (2007).The differences of clinical manifestations and laboratory findings in children and adults with dengue virus infection. J. Clin. Virol. *39*, 76–81.
- Kliks SC, Nisalak A, Brandt WE, Wahl L, Burke DS (1989).: Antibody-dependent enhancement of dengue virus growth in human monocytes as a risk factor for dengue hemorrhagic fever. *Am. . Trop. Med. Hyg. 40*, 444–451.
- Mairiang D, Zhang H, Sodja A, Murali T, Suriyaphol P, et al. (2013) Identification of New Protein Interactions between Dengue Fever Virus and Its Hosts, Human and Mosquito. PLoS ONE 8(1).
- Moody, Glyn (2004). Digital Code of Life: How Bioinformatics is Revolutionizing Science, Medicine, and Business. ISBN 978-0-471-32788-2.
- Murphy BR, Whitehead SS(2011). Immune response to dengue virus and prospects for a vaccine. Annu Rev Immunol,*29*,587–619.
- Noisakran S & Pemg GC, (2008) Alternate hypothesis on the pathogenesis of dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) in dengue virus infection. Exp 5101 Med ,*233*, 401-8.
- Rico-Hesse, R, (1990). Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. Virology , *174*: 479-93.

- Rothman AL, and Ennis FA (2004) Immunopathogenesis of dengue hemorrhagic fever. Virology *257*, 1–6.
- Rui-feng Qi, Ling Zhang and Cheng-wu Chi (2009)Biological characteristics of dengue virus and potential targets for drug design. Acta Biochim Biophys Sin :91-101.
- Sabin (1952) *In vivo* enhancement of dengue virus infection in rhesus monkeys by passively transferred antibody. J Infect Dis, *9*,527-533.
- Service, MW, (1992) Importance of ecology In Aedes aegypti . J Trop Med Pubhc Health, *23*, 68 1-90.
- Singh K, Lale A, Eong 001 E, Chlu LL, Chow VT, Tambyah P, Koay ES, (2006). A prospective clinical study on the use of reverse transcription-polymerase chain reaction for the early diagnosis of Dengue fever J Moi Dragn ,*8*, 613-6.
- Sittisombut, N., Sistayanarain, A., Cardosa, M. J., Salminen, M.,Damrongdachakul, S., alayanarooj, S., Rojanasuphot, S.,Supawadee, J. & Maneekarn, N. (1997). Possible occurrence of a genetic bottleneck in dengue serotype 2 viruses between the 1980 and 1987 epidemic seasons in Bangkok, Thailand. Am J Trop Med Hyg *57*,100–108.
- Victor TI, Malatlu M, Gurusamy D, Desai A, Raw V, Narayanasamy G, Anuradha L, Rant C, Knshnamurthy P, (2002) Dengue fever outbreaks In two vlllages of Dharmapun dlstnct In Tam11 Nadu Indran J Med Res , 116 133-9.
- Waterman, Michael S. (2009). Introduction to Computational Biology: Sequences, Maps and Genomes. CRC Press.
- Weaver SC, Reisen WK(2010). Present and future arboviral threats. *Antiviral Res*. *85*, 328–345.
- Whitehorn J, Simmons CP (2011) The pathogenesis of dengue. Vaccine,*29(42):*7221-8.
- World Health Organization, (2012) Dengue hemorrhagic fever diagnosis, treatment, Prevention. 1-84

Websites used:

1. http://www.ncbi.nlm.nih.gov
2. http://www.scfbio-iitd.res.in
3. http://www.ncbi.nlm.nih.gov/pubmed

4. http://www.ncbi.nlm.nih.gov/gene
5. http://www.ncbi.nlm.nih.gov/protein
6. http://www.ncbi.nlm.nih.gov/cdd
7. http://www.iedb.org/