

<http://www.pjbs.org>

PJBS

ISSN 1028-8880

**Pakistan
Journal of Biological Sciences**

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Geminivirus Database (GVDB): First Database of Family Geminiviridae and its genera Begomovirus

¹Rajneesh Prajapat, ¹Avinash Marwal, ²Zuber Shaikh and ¹Rajarshi Kumar Gaur

¹Department of Science, Faculty of Arts, Science and Commerce,

Mody Institute of Technology and Science, Lakshmangarh, Sikar-332311, Rajasthan, India

²Department of Web Application Development, Aroit Infoservices Pvt. Ltd., Jaipur, Rajasthan, India

Abstract: *Geminivirus* Database (GVDB) is an online interactive database of Geminiviridae family. GVDB comprises of partial and complete nucleotide sequences along with duly annotated expressed genes of isolated *Begomovirus* species. The *in silico* homology modeling, docking and recombination results obtained for different begomoviral sequences are also mentioned. This database is endowed with comprehensive information about *Geminivirus* members which grounds infection in various plants species in India assorting from crops, ornamentals plants and common weeds. The home page of this database offers various links associated with current research projects and also the publications related to molecular and *in silico* study of *Begomovirus* infection. The main feature of GVDB includes flexible database designs based on platform of PHP allows easy retrieval of the information. The database is made available at www.wikigeminivirus.org.

Key words: GVDB, *Begomovirus*, weeds, PHP, *in silico* study

INTRODUCTION

Agricultural crops are threatened to get vandalized in a short time frame by recent explosion of insect pests posing a serious threat to the overall production yield. Apart from this, the effect of other microscopic elements like viruses cannot be negated which are causing an immense annual loss to the agriculture sector. The list of viruses damaging these crops is immeasurable. In recent years, the epidemics of Geminiviridae range observed from 30 to 100% of the crops (Rojas, 2004). The members of Geminiviridae are transmitted from one plant to other by the white fly (*Bemisia tabaci*), leaf hopper and tree hopper (Marwal *et al.*, 2013).

Geminiviruses are major plant pathogens in tropical and subtropical countries (Boulton, 2003; Mansoor *et al.*, 2003) affecting a wide variety of crops with a catastrophic effect on productivity. The family Geminiviridae is one of the largest plant virus family and its members have a circular, single-stranded DNA (ssDNA) genome of approximately 2.7-5.2 kb. Based on their genome arrangement and biological properties, the family Geminiviridae is classified into four genera, i.e., *Mastrevirus*, *Curtovirus*, *Topocuvirus* and *Begomovirus* (Stanley *et al.*, 2005; Medina-Ramos *et al.*, 2008).

Begomovirus is one of the biggest genera of the Geminiviridae family. It comprises of around 200 species existing worldwide, principally infecting the

dicotyledonous plant species (Sunter and Bisaro, 1991). Its carrier *Bemisia tabaci* prevails more in the tropical and subtropical regions. It has a bipartite genome i.e., DNA-A and DNA-B with a size of 2.7 to 3 kb. An additional circular DNA reported as β satellite also contributes to *Begomovirus* virulation (Briddon *et al.*, 2002).

Although the researchers have spawned vast information about *Geminivirus* and its genera in the past years there has been always a need for a separate database for the plant virus family. Taking this into account and after serious thought process, GVDB was designed to compile and manage the research outcome information (Fig. 1) especially with respect to *Begomovirus* infection.

The GVDB [www.wikigeminivirus.org] provides an easy access to exhaustive study based on the sets of divergent genes and proteins of *Geminivirus* strain. *In silico* sequence analysis begin with the *Begomovirus* isolates sequences submission in GenBank-NCBI [<http://www.ncbi.nlm.nih.gov>] through BankIt. Also the homologous sequences were retrieved from GenBank-NCBI/DDBJ/EMBL for understand of structure functional relationships. The studies are facilitated in developing strategies for protein engineering. The database enables to comprehend protein sequence and its structure in a much better way and the function of proteins can also be easily implicated.

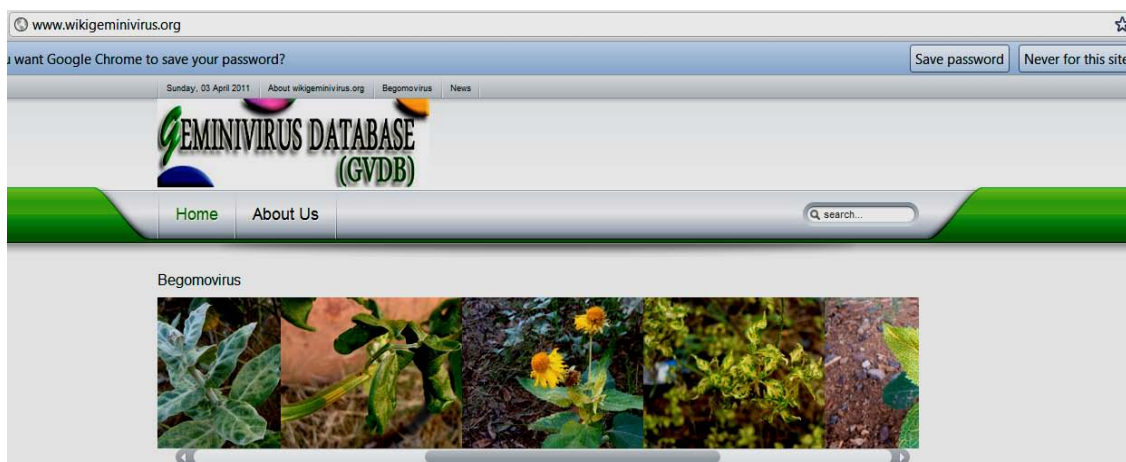


Fig. 1: Home page view of *Geminivirus* database (GVDB, www.wikigeminivirus.org)

MATERIALS AND METHODS

A detailed survey was carried out in various regions of Rajasthan, Haryana and Punjab, India, for acquiring essential information for GVDB. The plants samples were collected based on symptoms of typical *Begomovirus* infection. The various symptoms identified include vein yellowing, leaf curling and shortening, chlorosis and stunting of plant. CTAB (Cetyl Trimethyl Ammonium Bromide) method (Doyle and Doyle, 1987) was used for DNA extraction. Polymerase Chain Reaction (PCR) was then carried out by using a pair of universal coat protein primer (Hallan, 1998). After sequencing, nucleotide sequences were submitted to NCBI through BankIt (Accession of submitted sequences are JQ693143, JQ693145, JQ693149, JQ407224, JQ693150, JQ693148, JQ693146, JQ693144, JQ693142, JQ693140, JQ693138, JQ693136, JQ693151, JQ693147, JQ693141, JQ693139, JQ693137, JN998453, JN998451, JN998449, JN998447, JN998445, JN998443, JN998441, JN998452, JN998450, JN998448, JN998446, JN998444, JN998442, HQ876467, JN009667, JN009665, JN009666, JN009664, JF968443, JF968444, NC_015631, JN000703, JN000701, JN000702, JN000700, HQ631431, HQ631429, HQ631430 etc.).

Thereafter, homology search was carried out using BLAST (Basic Local Alignment Search Tool) (Altschul *et al.*, 1997) and homologous sequence information was used for phylogenetic tree construction in MEGA4 (Tamura *et al.*, 2007). The 3D-Jigsaw server was used to build three-dimensional models for proteins (ods) based on homologues of known structure. The UCLA-DOE server (<http://nihserver.mbi.ucla.edu>) was used for model validation through Verify-3D and

PROCHECK tools. The Verify-3D (Heinrichs, 2008) and PROCHECK (Laskowski *et al.*, 1993) outcomes were displayed in the form of profile search and Ramachandran plots (Prajapat *et al.*, 2010). The protein-protein docking was done by using Hex 6.3 (Prajapat *et al.*, 2011) to predict the interaction of proteins with each other.

Domain name www.wikigeminivirus.org (domain id: D161538660-LROR) was registered at servers195.ehostpros.com and was developed by using online applications of Joomla as its open source of Content of Management System (CMS). The CMS is used to build web sites and other powerful online applications. CMS is software which keeps track of every piece of content on web site and the content may be simple text, photos, documents etc. It does not require technical skill and knowledge to manage and the users can easily manage their own sites themselves. Results of study were then designed on MySQL 4.1 (www.mysql.com) by using PHP 5.2 (Hypertext Preprocessor, www.php.net) with Linux as the computing platform.

RESULTS AND DISCUSSION

In recent years the Geminiviridae family has been discussed in various international forums and also it falls in the category of one of the most researched plant virus family. *Bemisia tabaci* is commonly found in tropical, sub-tropical and temperate regions, therefore climatic conditions of India are quite favorable for the *Geminivirus* (e.g., *Begomovirus*) infection to proliferate in different plant species. The preliminary survey done in Rajasthan, Punjab and Haryana, India revealed various plants

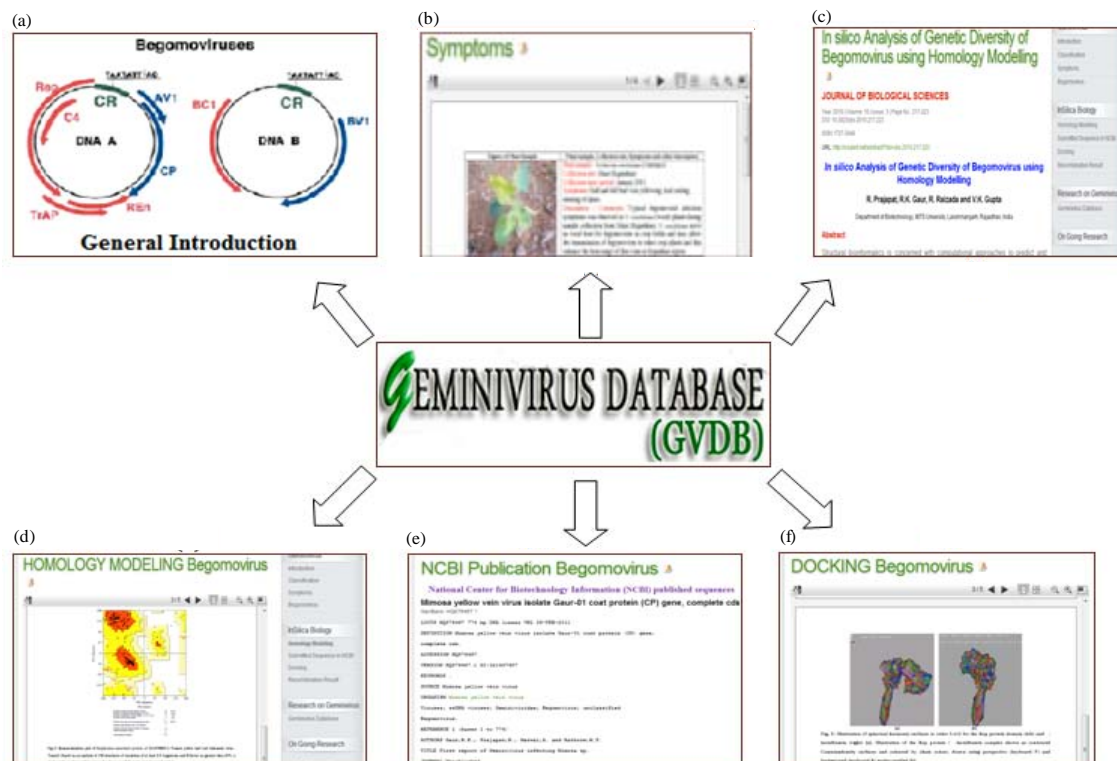


Fig. 2(a-f): The architecture of *Geminivirus* database (GVDB), The GVDB database manages biological and *in silico* data including (a) General introduction on *Geminivirus* and its genera, (b) Virus symptoms information, (c) Research on *Geminivirus*, (d) Homology modeling results, (e) NCBI published sequences and (f) Docking results and other information

species infected with *Begomovirus*. The *Geminivirus* Database (GVDB) has the following modules and sub-modules (Fig. 2) in its architecture:

Module *Geminivirus*: This module of GVDB database includes the four sub-modules:

- **Introduction:** It provides information about *Geminivirus* and its importance in agriculture biotechnology
- **Classification:** It provides information about the classification of Geminiviridae into four genera *Mastrevirus*, *Curtovirus*, *Topocovirus* and *Begomovirus* including their disease carriers in various plant species
- **Symptoms:** It provides information about typical begomoviral infection symptoms in different plant species which were collected from Rajasthan, Haryana and Punjab, India during 2009-2011
- ***Begomovirus*:** Provides general information about *Begomovirus* characteristics (Fig. 2)

Module *in silico* biology: This module of GVDB database included four sub-modules (Fig. 2):

- Homology modeling
- NCBI submitted sequences
- Docking
- Recombination results

However, a solved structure for coat proteins of NCBI reported isolates sequences were not available at the Protein Data Bank (PDB). Therefore, homology modeling model provide information about *Begomovirus* coat proteins by using the x-ray or NMR structure of the best homologous according to the target as templates with 3D-JIGSAW (Marwal *et al.*, 2012). The models were validated using protein structure validated tools such as PROCHECK (<http://www.ebi.ac.uk/thomson-srv/software/PROCHECK/>) and Verify-3D (http://nihserver.mbi.ucla.edu/Verify_3D/) for reliability. Thus, improve chances for anti-begomoviral agent producing capable of recognizing *Begomovirus* coat proteins.

Docking can be used in several ways like to study the mechanism of an enzymatic reaction, to identify possible binding modes for a ligand and to screen a database (Kahraman *et al.*, 2007). Hex 6.3 assigns multiple local coordinate systems to the larger molecule (receptor) and docks the ligand around each local coordinate frame on the receptor. Docking sub-module deals with protein-protein docking of *Begomovirus* proteins (Fig. 2) as receptor and β -lactoglobulin [2Q2M] as ligand.

Recombination is a major mechanism in virus evolution, allowing viruses to evolve more quickly by providing immediate direct access to many more areas of a sequence space than are accessible by mutation alone. Recombination analysis of *Begomovirus* isolates were performed by using Recombination Detection Program version 3.0 [RDP v.3.44] (Martin and Rybicki, 2000) on windows operating system (Ahmad *et al.*, 2011) which detects and analyzes recombination points in a set of aligned DNA sequences.

Module research on *Geminivirus*: This module provides information about geminiviruses ongoing/completed research projects in different research institutes of India. This section basically focuses on the control of *Begomovirus* infection through RNAi technology.

Module publication list: This module provides the list and links of research publications relevant to the projects that are running at Department of Science, Mody Institute of Technology and Science, Lakshmanagarh, Rajasthan, India. This database also provides the facility for the registered users to submission of research articles (Fig. 2).

The objective of the *Geminivirus* Database (GVDB) is to design and provide tool for genetic and *in silico* analysis of *Geminivirus* and specifically for its genera *Begomovirus*. It provides a small platform for the current status and research on *Geminivirus* infection in different economically important crops and weeds across India (Prajapat *et al.*, 2013). Information provided in this database is beneficial to the researchers on *Geminivirus* species.

GVDB database possess various modules that provide general information about *Geminivirus* and its genera (classification, symptoms etc.). *In silico* characterization of *Begomovirus* sequences shows the results in the form of Ramachandran plate, profile plots, protein-protein docking and recombination results.

CONCLUSION

Geminivirus infection cause major yield loss worldwide range in between 80 to 100%. The GVDB molecular, homology modeling, protein-protein docking

and recombination results can be used for screening of inhibitor molecules against the *Geminivirus* proteins and may be further applied for antiviral agent designing. According to best of our knowledge the *Geminivirus* Database (GVDB) is the world first database specifically designed for family Geminiviridae and its genera *Begomovirus*.

ACKNOWLEDGMENTS

The authors would like to acknowledge a vote of thanks to Department of Biotechnology (DBT project No. BT/PR13129/GBD/27/197/2009) and Department of Science and Technology (DST project No. SR/FT/LS-o42/2009), India for their financial support.

REFERENCES

- Ahmad, J., B.R. Singh, A.A. Al-Khedhairi, S. Alarifi, J.A. Khan and J. Musarrat, 2011. Characterization of Sunn hemp based on *in silico* structural and functional analysis of recombinant coat protein. *Afr. J. Biotechnol.*, 10: 2600-2610.
- Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.*, 25: 3389-3402.
- Boulton, M.I., 2003. Geminiviruses: Major threats to world agriculture. *Ann. Applied Biol.*, 142: 143-143.
- Bridson, R.W., S.E. Bull, S. Mansoor, I. Amin and P.G. Markham, 2002. Universal primers for the PCR-mediated amplification of DNA β : A molecule associated with some monopartite begomoviruses. *Mol. Biotechnol.*, 20: 315-318.
- Doyle, J.J. and J.L. Doyle, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.*, 19: 11-15.
- Hallan, V., 1998. Genomic organization of *Geminivirus* causing leaf curl in tomato (*Lycopersicon esculentum*). Ph.D Thesis. University of Lucknow, Lucknow, India.
- Heinrichs, A., 2008. Proteomics: Solving a 3D jigsaw puzzle. *Nat. Rev. Mol. Cell Biol.*, 9: 3-3.
- Kahraman, A., R.J. Morris, R.A. Laskowski and J.M. Thornton, 2007. Shape variation in protein binding pockets and their ligands. *J. Mol. Biol.*, 368: 283-301.
- Laskowski, R.A., M.W. MacArthur, D.S. Moss and J.M. Thornton, 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Applied Cryst.*, 26: 283-291.

- Mansoor, S., R.W. Briddon, Y. Zafar and J. Stanley, 2003. *Geminivirus* disease complexes: An emerging threat. Trends Plant Sci., 8: 128-134.
- Martin, D. and E. Rybicki, 2000. RDP: Detection of recombination amongst aligned sequences. Bioinformatics, 16: 562-563.
- Marwal, A., A.K. Sahu, R. Prajapat, D.K. Choudhary and R.K. Gaur, 2013. Molecular and Recombinational characterization of begomovirus infecting an ornamental plant *Alternanthera sessilis*: A new host of tomato leaf curl kerala virus reported in India. Sci. Int., 1: 51-56.
- Marwal, A., R. Prajapat, A. Sahu and R.K. Gaur, 2012. *In silico* recombination analysis: A study for *Geminivirus* host mobility. Asian J. Biol. Sci., 5: 1-8.
- Medina-Ramos, G., R. De La Torre-Almaraz, R. Bujanos-Muniz, R.G. Guevara-Gonzalez and N. Tierranegra-Garcia *et al.*, 2008. Co-transmission of *Pepper huastecoyellow vein virus* and *Pepper golden mosaic virus* in chili pepper by *Bemisia tabaci* (Genn.). J. Entomol., 5: 176-184.
- Prajapat, R., A. Marwal, A.K. Sahu and R.K. Gaur, 2013. First report of begomovirus infecting *Sonchus asper* in India. Sci. Int., 1: 108-110.
- Prajapat, R., R.K. Gaur and A. Marwal, 2011. Homology modeling and docking studies between AC1 rep protein of *Begomovirus* and whey α -lactalbumin. Asian J. Biol. Sci., 4: 352-361.
- Prajapat, R., R.K. Gaur, R. Raizada and V.K. Gupta, 2010. *In silico* analysis of genetic diversity of begomovirus using homology modelling. J. Biol. Sci., 10: 217-223.
- Rojas, A., 2004. A complex of begomoviruses affecting tomato crops in Nicaragua. Ph.D. Thesis, Swedish University of Agricultural Sciences, Uppsala.
- Stanley, J., D.M. Bisaro, R.W. Briddon, J.K. Brown and C.M. Fauquet *et al.*, 2005. Family Geminiviridae. In Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses.
- Sunter, G. and D.M. Bisaro, 1991. Transactivation in a *Geminivirus*: AL2 gene product is needed for coat protein expression. Virology, 180: 416-419.
- Tamura, K., J. Dudley, M. Nei and S. Kumar, 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol., 24: 1596-1599.