ORIGINAL RESEARCH PAPER

Microbiology

KEY WORDS: in silico analysis, Eggplant mottled dwarf virus, protein interaction

IN SILICO ANALYSIS AND PROTEIN STUDIES IN EGGPLANT MOTTLED DWARF VIRUS

DSVGK Kaladhar Department of Microbiology and Bioinformatics, Bilaspur University, Bilaspur (CG)

ABSTRACT

Proteins are the main components present within the host cell that control many biological funtions of viruses. Eggplant mottled dwarf virus (EMDV) is a member of the genus *Nucleorhabdovirus* causes mottled disease in eggplants (*Solanum melongena*). Gene prediction from GenBank: NC_025389 using FSENESV0 was predicted with seven genes like nucleocapsid protein, X protein, phosphoprotein, putative movement protein, matrix protein, glycoprotein and RNA-directed RNA polymerase protein. The genes of EMDV were shown relationhip with TVCV, MMV, PCMV and PYDV. The results shown only protein 4 and protein 6 are stable. In *Coprinopsis cinerea with* CC1G_00342 - RNA-directed RNA polymerase, CC1G_00194 – glycoprotein and CC1G_08598 - G-X-X-X-Q-X-W domain-containing protein didnt shown any interactions. CC1G_01977 - phosphoprotein phosphatase and CC1G_10289 - nuclear matrix protein NMP200 shown good interactions with mitochondrial and golgi proteins.

INTRODUCTION

Proteins can control several biological systems within a host cell and perform functions of viruses independently, interact with other proteins and genes for transforming into improper biological activities of the host[1]. Proteins are the workhorses of living cells that are facilitating most biological processes like cell growth, gene targetting and expression, morphology, gene repair, motility and intercellular communications[2]. Cells respond to a functional stimuli of external pathogenic species and controls protein expression into a dynamic process. The basic characteristics and interaction of proteins suggests a complexity mechanisms which is difficult to investigate using traditional approaches, especially in understanding protein function in a proper biological network context[3-5]. Some of the critical aspects that are required to understand the functions of a proteins includes: sequence to structure to function, conserved sequences and evolutionary history, post-translational modifications, protein characterisation, protein expression profile, protein-protein interactions and Intracellular localization[6-8].

Eggplant mottled dwarf virus (EMDV) is a plant rhabdovirus which infects eggplants (*Solanum melongena*), was first detected in southern Italy. EMDV natural means of transmission was remained unknown causes and may be transmitted by grafting, inoculation of sap, leafhoppers and through seeds [9-11]. The present studies shows *in silico* analysis like identification, charecterisation, phylogenetic analysis and protein interaction network analysis of Eggplant mottled dwarf virus (EMDV) genome.

METHODOLOGY

The National Center for Biotechnology Information is a database that created advances in science and health by providing access to various biomedical and genomic information (https://www.ncbi.nlm.nih.gov/). Genome Sequence of Eggplant mottled dwarf virus (EMDV) isolate Agapanthus, complete genome (GenBank: NC_025389) was retrieved from the NCBI database. Protein and gene sequence similarity was analyzed by BLAST tools for searching the NCBI non-redundant sequence database. fgenesV0 is used for prediction of genes that uses generic parameters Markov chain-based viral gene prediction (http://linux1.softberry.com/berry.phtml). A phylogenetic tree or evolutionary tree was conducted using MEGA. Protein charecterization using ProtParam was conducted to computed parameters include the molecular weight, theoretical pI, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY). STRING v 10 is a database of known and predicted protein-protein interactions.

RESULTS AND DISCUSSION

The Genome Sequences Eggplant mottled dwarf virus (EMDV) isolate Agapanthus, complete genome (GenBank: NC_025389) was retrieved from the NCBI database. The sequence submitted for gene prediction using FGENESV0 has shown seven genes and the results were produced as shown below

Table 1: Gene Identification in Eggplant mottled dwarf virus (EMDV) genome (NC_025389)

Gene No	Start End	Identification
1	282 - 1712	nucleocapsid protein [Eggplant mottled dwarf virus]
2	1618 - 2118	X protein [Eggplant mottled dwarf virus]
3	2191 - 3075	phosphoprotein [Eggplant mottled dwarf virus]
4	3158 - 4117	putative movement protein [Eggplant mottled dwarf virus]
5	4213 - 4968	matrix protein [Eggplant mottled dwarf virus]
6	5185 - 7032	glycoprotein [Eggplant mottled dwarf virus]
7	7123 - 12963	RNA-directed RNA polymerase [Eggplant mottled dwarf virus]

The results were shown that Gene1 predicted as nucleocapsid protein, gene 2 as X protein, gene 3 as phosphoprotein, gene 4 as putative movement protein, gene 5 as matrix protein, gene 6 as glycoprotein and gene 7 as RNA-directed RNA polymerase [Eggplant mottled dwarf virus]. All the genes when submitted to swissmodel was shown no suitable templates that were found in the structure database.

A phylogenetic tree constructed with all 7 proteins except protein 2(divergent). The result was shown in Fig 1.

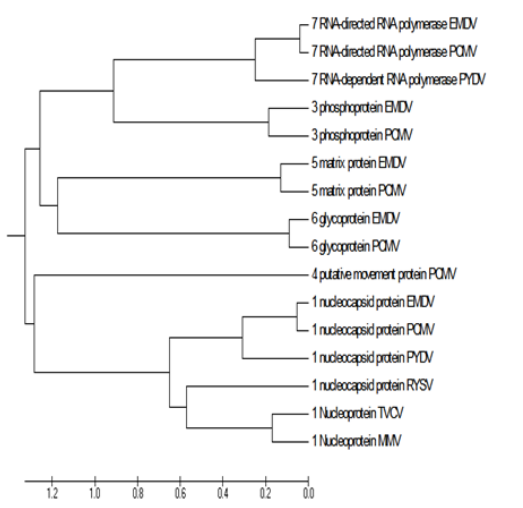

Fig 1: Phylogenetic tree using MEGA for Eggplant mottled dwarf virus

Table 2: PROTPARAM RESULTS for predicted EMDV genome

Gene	Number of amino acids	Molecular weight	Theoretical pI	Extinction coefficients	Estimated half-life	Instability index	Aliphatic index	Grand average of hydropathicity (GRAVY):
1	476	52010.58	6.73	43445	30 hours (mammalian reticulocytes, in vitro) >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	44.62 unstable.	83.89	-0.212
2	166	19089.31	4.18	28920	30 hours (mammalian reticulocytes, in vitro) >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	63.59 unstable.	69.82	-0.455
3	294	32727.68	6.86	24870	30 hours (mammalian reticulocytes, in vitro) >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	51.89 unstable.	74.63	-0.655
4	319	36153.09	8.49	27640	30 hours (mammalian reticulocytes, in vitro) >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	39.65 stable	81.19	-0.146
5	251	27872.92	9.38	13410	30 hours (mammalian reticulocytes, in vitro) >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	42.25 unstable.	90.12	-0.320
6	615	68870.09	4.83	134035	30 hours (mammalian reticulocytes, in vitro) >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	37.34 stable	84.39	-0.084
7	1946	221709.62	6.50	224485	30 hours (mammalian reticulocytes, in vitro) >20 hours (yeast, in vivo). >10 hours (Escherichia coli, in vivo).	41.69 unstable.	92.09	-0.206

The results from table 2 has shown that protein 7 having more number of amino acids and molecular weight. The results shown only protein 4 and protein 6 are stable. Figure 2 shows EMDV proteins that didnt shown any interactions with Coprinopsis cinerea with CC1G_00342 - RNA-directed RNA polymerase, CC1G_00194 – glycoprotein and CC1G_08598 - G-X-X-X-Q-X-W domain-containing protein. CC1G_01977 - phosphoprotein phosphatase and CC1G_10289 - nuclear matrix protein NMP200 of EMDV has shown good interactions with mitochondrial and golgi proteins.

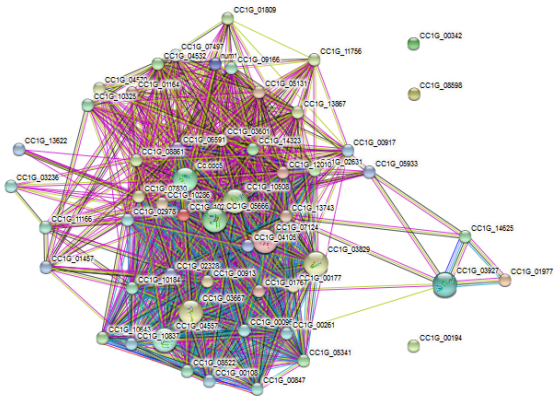

Figure 2: Protein interaction networks for predicted EMDV genome

Genome and the development of the field of proteomics are important in understanding the protein interaction and identification biological networks that shows vital in understanding of proteins function within the living cells[12]. Proteins rarely present alone because their functions tend to be regulated seperately in each machanisms[13]. A large number of protein components are organized and build by protein-protein interactions (PPIs) where many molecular processes takes place within a cell that are carried out by various molecular machines[14]. The present study provided good understanding about EMDV genome using in silico approaches.

CONCLUSION

The genome of EMDV (NC_025389) was predicted to contain 7 genes like nucleocapsid protein, X protein, phosphoprotein, putative movement protein, matrix protein, glycoprotein and RNA-directed RNA polymerase protein. The genomic and proteomic analysis was conducted and shown good results. Further wetlab techniques can provide good understanding in controlling EMDV virus

ACKNOWLEDGEMENT

The author acknowledged to thank management and staff of Bilaspur University in bringing out the research work.

REFERENCES

1. Hartwell, L. H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999). From molecular to modular cell biology. Nature, 402, C47-C52.
2. Jacobs, T. M., B. Williams, T. Williams, X. Xu, A. Eletsky, J. F. Federizon, T. Szyperski, and B. Kuhlman. "Design of structurally distinct proteins using strategies inspired by evolution." Science 352, no. 6286 (2016): 687-690.
3. Brenner, S. (2000). The end of the beginning. Science, 287(5461), 2173-2174.
4. Arkin, M. R., & Wells, J. A. (2004). Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. Nature reviews Drug discovery, 3(4), 301.
5. DSVGK Kaladhar and Shri Kant, In Silico Protein–Protein Interaction Studies Of Sirtuins As Anti-Inflammatory And Anti-Cancer Agents, IJETSR, 4(8), 162-185, 2017
6. Lee, D., Redfern, O., & Orengo, C. (2007). Predicting protein function from sequence and structure. Nature Reviews Molecular Cell Biology, 8(12), 995.
7. Beltrao, Pedro, Peer Bork, Nevan J. Krogan, and Vera van Noort. "Evolution and functional cross–talk of protein post–translational modifications." Molecular systems biology 9, no. 1 (2013): 714.
8. Kaladhar, D. S., Sai, P. C., & Rao, P. V. N. (2012). Towards an understating of signal transduction protein interaction networks. Bioinformation, 8(9), 437.
9. Martelli, G. P., Russo, M., & Rubino, L. (1973). Eggplant mottled dwarf virus. CMI/AAB Descriptions of plant viruses, 115(4).
10. Babaie, G. H., and K. Izadpanah. "Vector transmission of Eggplant mottled dwarf virus in Iran." Journal of Phytopathology 151, no. 11–12 (2003): 679-682.
11. Mavrič, I., M. Tušek Žnidarič, M. Virščeк Marn, P. Dolničar, N. Mehle, D–E. Lesemann, and M. Ravnikar. "First report of Eggplant mottled dwarf virus in potato and tomato in Slovenia." Plant pathology 55, no. 4 (2006): 566-566.
12. Ashburner, Michael, Catherine A. Ball, Judith A. Blake, David Botstein, Heather Butler, J. Michael Cherry, Allan P. Davis et al. "Gene Ontology: tool for the unification of biology." Nature genetics 25, no. 1 (2000): 25.
13. Lindenboim, L., Blacher, E., Borner, C., & Stein, R. (2010). Regulation of stress-induced nuclear protein redistribution: a new function of Bax and Bak uncoupled from Bcl-x L. Cell death and differentiation, 17(2), 346.
14. Yamada, T., & Bork, P. (2009). Evolution of biomolecular networks—lessons from metabolic and protein interactions. Nature Reviews Molecular Cell Biology, 10(11), nrm2787.