



## **INSILICO COMPARATIVE STUDY OF VIRAL GENOMES OF PARVOVIRIDAE**

**A.Sumathi\* and L.Surendran**

Department of Biotechnology, Hindusthan College of Arts and Science, Coimbatore-641028, India.  
Department of Bioinformatics, K.G College of Arts and Science, KGiSL Campus, Saravanampatti, Coimbatore-641035, India.

### **ABSTRACT**

The Current Research deals with comparative study and functional similarities of viral genomes of the family *parvoviridae*. *Parvoviridae* is a viral family of single stranded DNA viruses with some identical proteins for the infecting the animals. Current study reveals the similarity of genomes and protein sequences of different viral species namely Adeno associated virus, Bovine adeno associated virus, Canine minute virus, Minute virus of mice, Mouse parvo virus, Muscovy duck parvo virus. The above mentioned six species of single stranded DNA animal viruses, those with genomic similarity with each other and also some identical proteins were isolated. The Open Reading Frame (ORF) of these genomes is encoding for the production of some specific proteins such as replication catalytic domain, capsid protein, Non structural protein and Coat protein. From this research, it was found that the similar infectious and replicative proteins present in the several species of viral family *parvoviridae*.

**Key words:** *Parvoviridae*, Capsid Protein, Replication Catalytic Domain.

### **INTRODUCTION**

Viruses are generally subdivided into three major groups based on the kind of host they infect. Bacterial viruses, animal viruses, plant viruses. Also the viruses have been found to, infect insects, birds, reptiles, mammals, and Plants. They are also causative for many human diseases like mumps, measles, rabies and influenza.

The Single stranded dna viruses contain the coat protein called Capsid. It is also comprises separate protein segments known as capsomers. *Parvoviridae* is icosahedral, 18-26nm diameter and consists only protein and DNA. Six species of virus from genus parvo virus, dependo virus and boca virus were taken. In each viral species three Capsid proteins and Replication catalytic domains were encoded, while several viral species had encoded nonstructural and structural

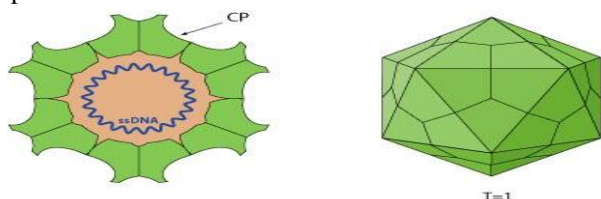
replication catalytic domain. Similarities of proteins were compared among above mentioned viruses. This study serves as baseline information to develop new antiviral drugs in future [1-5].

### **MATERIALS AND METHODS**

The viral species were analyzed using some bioinformatics tools and online reference to retrieve the similarity of proteins and genomes among them. The following bioinformatics tools were used to determine the protein and genomic similarity.

### **NCBI (National Centre for Biotechnology Information)**

National centre for biotechnology information is an online reference among the life science community, which provides information digitalized sequences of genomes and protein ranging from prokaryotes to eukaryotes. By using biotechnology information online reference, the genome sequences were retrieved for the developmental purpose. Viral sequences were retrieved from viral zone of nation centre for biotechnology information online reference.



\*Corresponding Author **A.Sumathi** E mail: [sumathiphd@yahoo.co.in](mailto:sumathiphd@yahoo.co.in)

**Clustal w (EMBL-European Bioinformatics Institute EBI)**

The genome and protein sequences were taken from NCBI online reference. The sequences of different species of viruses were submitted in online sequence alignment tool clustal w. The clustal w, which compares the query, sequences each other given by the user and it will list out the most similar sequences at first according to the score given for similarity of sequence with query. The result revealed the most similar sequences of viral species taken by comparing the sequences of each other.

**Phylo draw (Pusan National University)**

The Phylo draw is a tool that provides origin and relationship among the species. The viruses of our desired species came from particular origin and each species are closely related and also related to other viral species. In order to conclude, whether they are closely related or distantly related, the phylo draw tool is used. The phylogenetic tree pictured the closely and distantly related species of virus.

**Genscan**

The genscan is tool for retrieving the protein sequence from the desired genome. At first the genome sequence were retrieved from the NCBI online reference and submitted in the genscan tool. The result gives the protein sequences of the submitted genome.

**ORF finder (National centre for Biotechnology Information)**

ORF open reading frame of genomes and segment that will be encoded by ribosome during the translation. The genome sequence will transcribe into the protein sequence. The protein may involve in essential activities of virus. ORF of our genome sequences encoding three major kinds of proteins involved in the transcription and translation process. They include Replication catalytic domain play in replication process of the viral cycle. The Capsid protein, which involves in the protection and enclosing the genome of viral genetic material. Coat protein, also plays a vital role in the formation of surface proteins.

**P fam (Wellcome Trust Sanger Institute, United Kingdom)**

P fam (Sanger institute, united kingdom) is protein family analysis tool from sanger institute situated at united kingdom. The online protein analysis tool that analyses the protein sequence with each other. The tool will analyze and compares the submitted protein sequences with other protein [6-11].

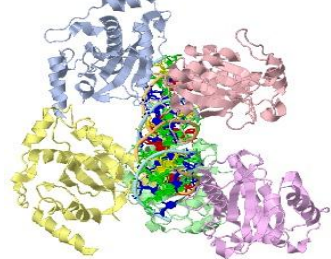
**RESULT AND DISCUSSION**

The result of this project describes the possible homology of genome and proteins of different viral species of the family *parvoviridae*. The genome sequences were retrieved, analyzed, transcribed and declared by using above mentioned bioinformatics and biotechnology

online resources and tools from authorized intuitions around the world.

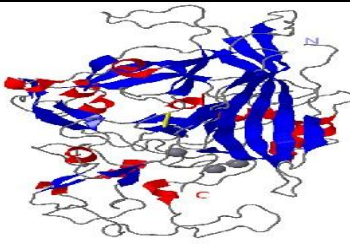
**The Replication Protein Catalytic Domain**

Adeno-associated virus1, Bovine Adeno-Associated Virus, Muscovy duck parvo virus were encoded the same replication catalytic domain and also they have a similar structure with PDB entry 1rz9. The replication catalytic domains of different viral species were isolated from NCBI and analyzed using bioinformatics tools showed that the same replication catalytic domain was encoded by the above viruses. Also the replication catalytic domain involved as a helicase activity in transcription process.

viral species	Replication Catalytic domain PDB Entry 1rz9
A. Adeno-Associated Virus1.	
B. Bovine Adeno-Associated Virus.	
C. Muscovy Duck Parvo Virus	

**Capsid Protein**

The Pfam analysis of Adeno Associated virus 1, Bovine adeno associated virus, Muscovy duck parvo virus depicted the capsid proteins of above viral species similar to like that of the Replication catalytic domain of above viral species. The capsid protein is layer of virus and it is involved in the protection of genetic material against other foreign organisms. The protein was formed by the combination of peptide sequence encoded by the genome as viral protein 1, viral protein 2 and viral protein 3. The combine together to form the capsid protein.

Viral Species	Capsid Protein Pdb Entry 1c8f
A. Adeno Associated Virus.	
B. Bovine Adeno-Associated Virus.	
C. Muscovy Duck Parvo Virus.	

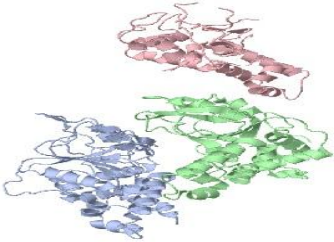
This viral protein 2, together with VP1 forms a capsomer. Both of these proteins are formed from the same transcript using alternative splicing. As a result, VP1 and VP2 differ only in the N-terminal region of VP1. VP2 is involved in packaging the viral DNA.

Although among these proteins, another essential

viral protein 3 plays a vital role in packaging, protecting the genetic material and enzymes. However, viral protein will be active only with combination of viral protein 1 and viral protein 2.

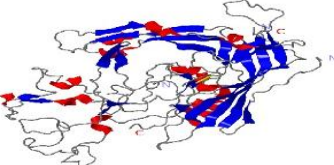
**Nonstructural Protein**

The viral species Adeno associated virus, Canine minute virus, Minute virus of mice and Muscovy duck parvo virus were encoded structural and non structural proteins. There were two ORF isolated from those viruses, that are responsible for structural and non structural proteins. They are identical proteins produced by above mentioned viral species while infecting with host and life cycle. Both of these involved in the replication and transcription activities of above mentioned viral species and the two proteins essential for DNA helicase and ATPase activity during viral cycle [12-14].

Viral Species	Non-Structural Protein PDB Entry 1s9h
A. Adeno-Associated Virus1	
B. Canine Minute Virus	
C. Minute Virus of Mice	
D. Mouse Parvovirus	

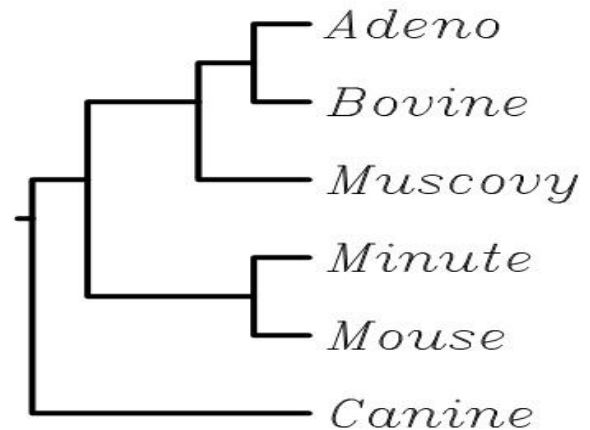
**Coat Protein**

The ORF of viral species encodes protein molecule of coat protein. In the parvo virus the Coat protein is a combination of peptides that is responsible for the formation of coat protein. The viral coat protein works as a protective barrier against other foreign organisms and it is involved during the infection with the host. The viral species Minute virus of mice, Mouse parvovirus1, Muscovy duck parvovirus were encoded identical coat protein during gene expression.

Viral Species	Parvovirus Coat Protein Pdb id 1p5y
A. Minute Virus Of Mice	
B. Mouse Parvovirus1	
C. Muscovy Duck Parvovirus	

**Phylogenetic Analysis**

The phylogenetic analysis helps in finding out the close relationship among the species. Different types of phylogenetic trees are available. Based on the algorithms used by phylogenetic tool it may vary. The output of Phylogenetic tree displayed the closely related viruses by genomic and protein similarity. From this result each viral species has shared a genome or protein with another species of virus and they produce identical proteins as like neighbour virus.



**CONCLUSION**

This project depicted the similarity of the genome and identical proteins among six viral species. The genome sequences and protein sequences were retrieved from the viral zone (National Centre for Biotechnology Information) and the genome was analyzed in each viral species of the parvoviridae family by using above mentioned bioinformatics tools and online databases. The Insilico identification of the open reading frame in viral genome was done, which plays a vital role in regulation of gene expression during its life cycle. The result revealed the sequential similarity of the each genome, taken as graphical and other formats, such as phylip, nexus. Further the study explained about the domains, which are target identification for the future studies on drug discover. All the similar domains were identified in the six viral species of family *parvoviridae* by pfam online reference (Wellcome Trust Sanger Institute). The above viruses were found to be encoded identical Replication catalytic domain, Capsid protein, Non structural protein and Coat protein with each other viral species.

**REFERENCES**

- Walker SL, Wonderling RS, Owens RA. Mutational analysis of the Adeno associated virus type 2 Rep68 protein helicase motifs. *J Virol*, 1997, 71, 6996-7004.
- Glotzer JB, Saltik M, Chiocca S, Michou AI, Moseley P, Cotton M. Activation of heat shock response by adeno virus is essential for virus replication. *Nature*, 407, 2000, 207-211.
- Weger S, Hammer E, Heilbronn R. Topors, a p53 and topoisomerase I binding protein, interacts with the adeno associated virus(AAV-2)Rep78/68 protein and enhances AAV-2 gene expression. *J Gen Virol*, 83, 2002, 511-516.
- Qiu,j YY, Tullis GE and Pintel DJ. 2006. Parvovirus RNA Processing stratigies, In M.e.Bloom(ed.), Parvo viruses.Hodder Arnold, London, United Kingdom, 2006, 253-273.

5. Samulki RJ, Chang LS and Shenk T. Helper-freestock of recombinant adeno associated viruses: normal integration does not require Viral gene expression. *J.Virol*, 63, 1989, 3822-3828.
6. Brown KE, Anderson SM and Young NS. Erythrocyte Cellular receptor B19 parvovirus. *Science*, 262, 1993, 114-117.
7. Burgert HG, Blusch JH. Immuno modulatory functions encoded by the E3 transcription unit of adenoviruses. *Virus genes*, 21, 2000, 13-25.
8. Luo W and Astell CR. A novel protein encoded by small RNAs of parvovirus B19. *Virology*, 195, 1993, 448-445.
9. Shackelton LA, Holmes EC. The evolution of large DNA viruses: Combining Genomic information of viruses and their hosts. *Trends Microbial*, 12, 2004, 458-465.
10. Gather D. Genome signatures, self organizing map sandhigher order phylogenies: A parametric analysis. *Evol Bioinf*, 3, 2007, 211-236.
11. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: Patterns and determinants. *Nat Rev Genet*, 9, 2008, 267-276.
12. Pruitt KD, Tatusova T, Maglott DR. NCBI refrence sequences(Ref Seq):Acurated Non-redundant sequence data base of genomes,transcript sand proteins. *Nucleic Acids Res*, 35, 2007, D61-65.
13. Novatchkova M, Eisenhaber F. Can molecular mechanisms of Biological processes be extracted from expression profiles? Case study:endothelial contribution to tumor-induced Angiogenesis. *Bioassays*, 23, 2001, 1159-1175.
14. Hermonat PL, Santin AD, Batchu RB, Zhan D. The adeno associated virus Rep78 major regulatory protein binds the cellular TATA-binding protein *in vitro* and *in vivo*. *Virology*, 245, 1998, 120-127.