SciForum
Mol2Net

Phylogenetic and Genetic Analysis of Envelope Gene of the Prevalent Dengue Serotypes in India in Recent Years

Sumanta Das*, Centre for Interdisciplinary Research and Education, Jodhpur

Park, Kolkata 700068, India, Email: sumantadey13@gmail.com; anandy43@yahoo.com

Ashesh Nandy, Centre for Interdisciplinary Research and Education, Jodhpur

Published: 4 December 2015

Abstract: A fresh wave of dengue infection, particularly dengue serotypes 1 and 3, have been observed all across India in recent times and has led to several fatalities. Since the surface situated envelope protein of the dengue virion is responsible for virus entry into the host cell, we have laid special emphasis on its characterization and analyses of the envelope gene with an aim to eventually develop inhibitors of the dengue virus. There are four serotypes of the dengue virus of which types 1 and 3 are the most widely prevalent in India. 2D graphical representations of the envelope gene from various countries show that the gene from an Indian dengue type 1 virus bears a strong resemblance to the genes from Asia, whereas in the case of dengue type 3, the Indian strain representation shows simlarity to strains from North America. Phylogenetic trees using alignment procedures also bear this out, implying an inherent cross-national spread of the dengue virus. Moreover, hydropathy analysis shows that amino acid compositional changes are tending to increase hydrophobic residues in the dengue type 3 viruses leading to morphological changes that may explain, in part, the higher pathogenicity of the dengue virus in India in recent times. These exercises serve to show the urgency of comprehensive genetic surveillance of the dengue virus to anticipate further damaging changes in the viral sequence.

Keywords: Dengue envelope gene, 2D graphical representation, Phylogenetic analysis, transition/transversion ratio, hydropathy plot, amino acid composition changes, viral pathogenicity.

Introduction

The dengue virus (DENV) is estimated to infect over 50-100 million people across 60 countries annually [1], with fatalities of around 50 to 100 thousand per year [2, 3]. This virus, genus *Flavivirus*, family Falviviridae, consists of four antigenically distinct serotypes (DENV 1 to 4) [4, 5]. The genome comprises a single strand of RNA encoding three structural proteins - the capsid (C), premembrane/membrane (PrM/M) and envelope (E), and seven non-structural (NS)

proteins - NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 [6], arranged sequentially as shown in a 2D graphical representation in Fig.1. Of these, the surface situated envelope glycoprotein is usually targeted by vaccines and drugs for inhibition [1, 7]. Moreover it is found that dengue virus infectivity depends on envelope protein binding to target cell [8, 9, 10]; it is also interesting to note that even a single nucleotide change in the envelope protein gene of dengue DENV2 can affect its neurovirulence in mice [11], and could be important in the case of human patients also.

Because of the wide prevalence of dengue infections in India, we have done a bioinformatics study of the dengue focusing on the outer capsid envelope gene which is responsible for viral endocytosis. We report here results of the study for a selection of dengue viruses of DENV3 and DENV1 types from various countries of the world along with the same sera prevalent in India over a period of several years. In India these two serotypes are mostly observed [12], but in recent times cases of DENV3 appear to be much more prevalent. In this report we characterize and draw phylogeny of envelope gene sequences of DENV3 and DENV1 from India and other countries for comparative study along with 2D graphical representation of the gene for visualization of the base distributions and structure. We also report here the transition/transversion ratio, amino acid composition as well as hydropathy index of the envelope protein to get an indication of the protein morphological differences between the serotypes DENV3 and DENV1, with emphasis on the changes in DENV3 envelope protein in relation to DENV1.

Result and Discussions

Figs. 2 and 3 show the phylogenetic relationship between the envelope gene sequences of DENV3 and DENV1 respectively from various countries in recent times. The results show that DENV3 envelope gene from India is closely related to American strains, whereas for DENV1, the Indian gene is more closely related to Chinese and other Asian strains, which we had also reported in our previous paper [14].

Figs. 4 and 5 show the comparative analysis of the envelope gene by 2D graphical representation. These representations clearly show that the envelope gene of DENV 3 strains from India (e.g., Locus ID: JQ686083) are closely related to American strains (e.g., EU596494) whereas in case of DENV 1 strains from India (e.g., JF967939) are related to Asian strains (e.g., China KC006933).

Table 1 shows the transition/transversion ratio matrix of the dengue envelope genes of DENV3 and DENV1 from India, influenza A surface genes (Hamaglutinin and Neuraminidase) and mammalian beta globin genes. Numbers in bold are the transition frequencies (%), others are transversion frequencies (%). The data above show that while the rate of transition to transversion mutations is about 55: 45 in mammalian and other viral genes, in the case of dengue genes this ratio at 88:12 is significantly different.

Figs. 6 and 7 show the hydropathy index plot of the envelope gene sequences of both the serotypes 3 and 1, respectively, from India after or from the year 2000. The results indicate that hydrophobicity of the envelope protein of the Indian serotype DENV3 shows a slight tendency to increase with time but in case of the Indian serotype DENV1 it shows tendency towards decreasing with time implying morphological changes in the protein structure.

Fig 8 shows the differences in the amino acid composition of DENV1 and DENV3 for each amino acid of Indian strains in the period under study: positive values imply higher frequencies in DENV3. The figure shows changes vary mainly in amino acids like Asparagine, Isoleucine, Tyrosine, Aspartic acids, Arginine, Serine, Threonine, Valine and Glutamic acids while changes are at a minimum in amino acids like Cysteine, Lysine, Methionine, Glutamine and Tryptophan between the two serotypes, DENV3 and DENV1.

We infer these data have the following implications:

> Since, Isoleucine and Valine have large aliphatic hydrophobic side chains, their molecules are rigid and their mutual hydrophobic interactions are important for correct folding of proteins. So changes in the composition of these amino acids, e.g. higher Isoleucine for DENV3, can affect the 3D structure of the envelope protein for both the strains.

> Tyrosine contains a large rigid aromatic group on the side chain and is also one of the biggest amino acids. Moreover like Isoleucine and Valine, Tyrosines are hydrophobic and trend to orient towards the interior of the folded protein molecule. Excess Tyrosine in DENV3 could be making it more hydrophobic.

> Arginine contains a large flexible side chain with a positively-charged end. The flexibility of the chain makes Arginine suitable for binding to molecules with many negative charges on their surfaces. The strong charge makes the amino acid prone to be located on the outer hydrophilic surfaces of the proteins. Since the envelope protein is surfaced exposed, change in the Arginine composition of DENV3 might reduce the binding property of the protein.

> Serine and Threonine with a hydroxyl group are very hydrophilic. Their reduced frequency in DENV3 leads to higher hydrophobicity and consequent morphological change in the DENV3 envelope protein

Table 1 showing transition/transversion matrix

Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution

	A	T	C	G
A	-	*1.22*	*1.09*	**12.05**
T	*1.77*	-	**29.69**	*1.52*
C	*1.77*	**33.01**	-	*1.52*
G	**14.05**	*1.22*	*1.09*	-

DENV3 Envelope genes

Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution

	A	T	C	G
A	-	*1.57*	*1.49*	**11.9**
T	*2.34*	-	**28.85**	*1.96*
C	*2.34*	**30.32**	-	*1.96*
G	**14.2**	*1.57*	*1.49*	-

DENV1 Envelope genes

Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution

	A	T	C	G
A	-	*5.87*	*4.42*	**10.55**
T	*7.95*	-	**11.59**	*5.91*
C	*7.95*	**15.4**	-	*5.91*
G	**14.18**	*5.87*	*4.42*	-

Influenza virus A surface genes (Hemaglutinin and Neuraminidase)

Maximum Likelihood Estimate of Substitution Matrix

	A	T/U	C	G
A	-	*5.76*	*5.63*	**13.67**
T/U	*4.85*	-	**15.10**	*6.68*
C	*4.85*	**15.45**	-	*6.68*
G	**9.94**	*5.76*	*5.63*	-

Mammalian Beta globin genes

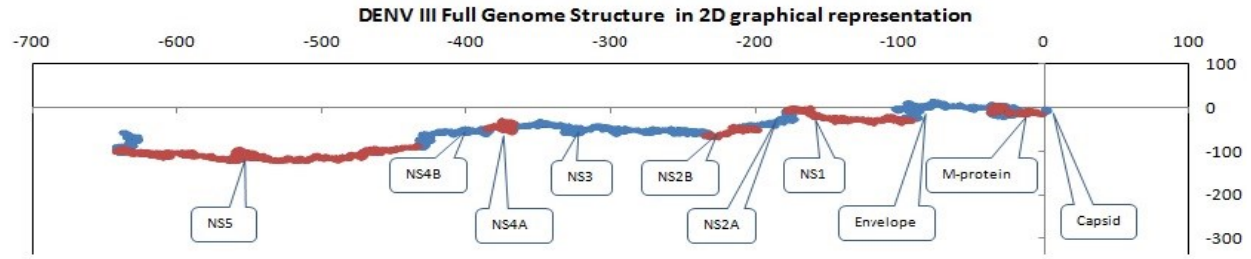

Figure 1: Genome structure of Indian DENV3 in 2D graphical representation

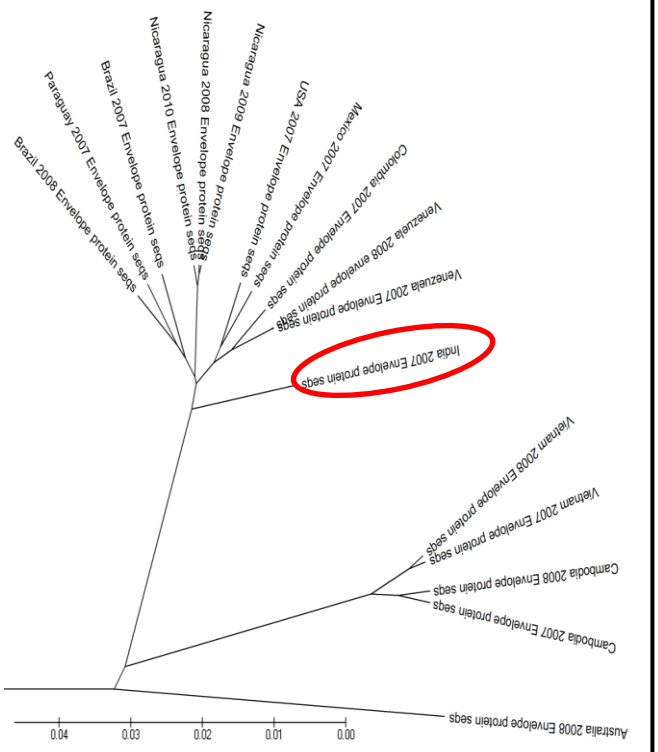

Figure 2: Phylogenetic relationship between 18 envelope gene sequences of DENV 3 from various

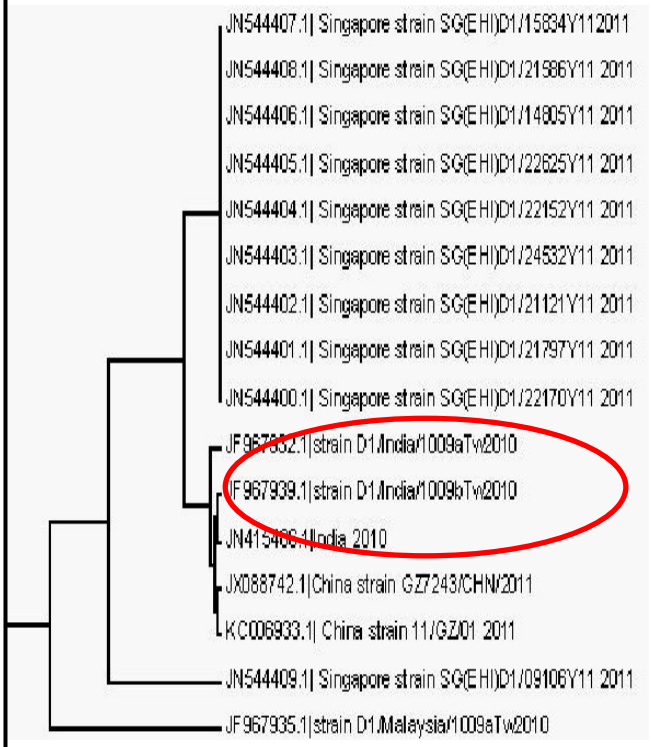

Figure 3: Phylogenetic relationship between envelope gene sequences of DENV 1 from Asian countries in recent

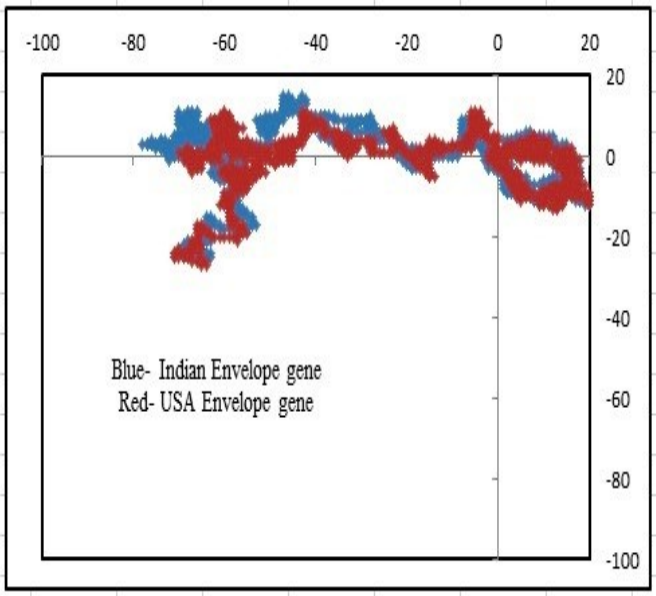

Figure 4:2D graphical representations of DENV3 envelope gene sequences from India (Locus ID: JQ686083) and USA (EU596494)

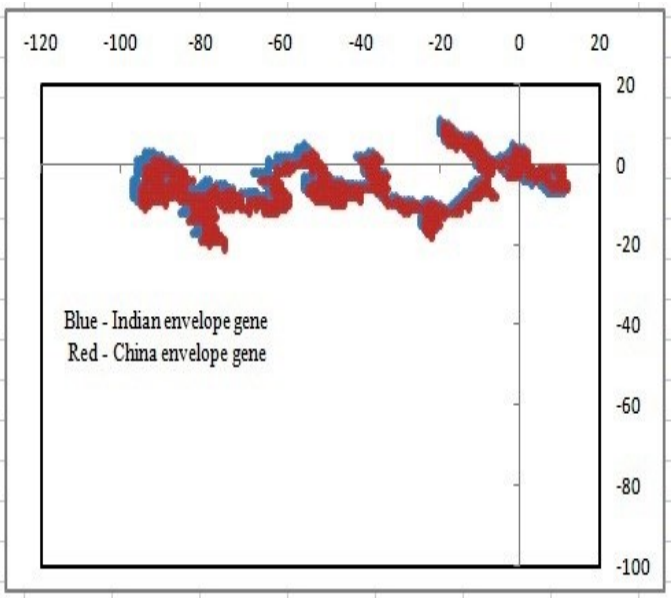

Figure 5:2D graphical representations of DENV1 envelope gene sequences from India (JF967939) and China (KC006933).

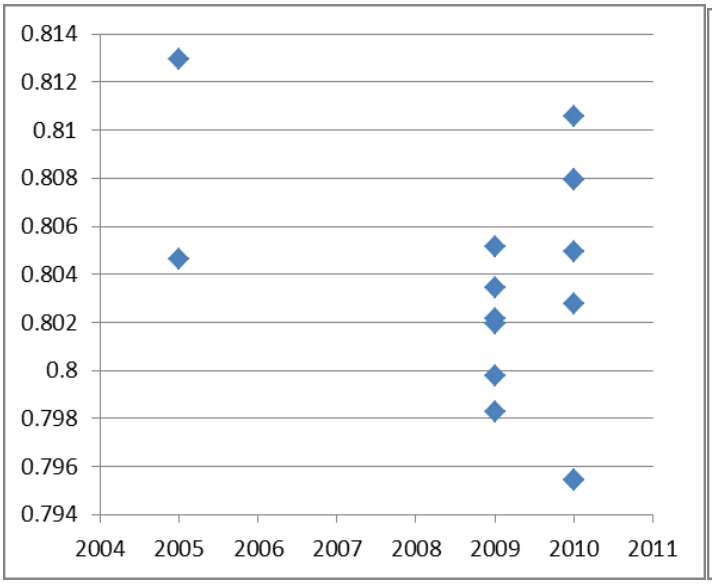

Figure 6: Hydropathy plot of envelope gene sequences of DENV 3 strains from India

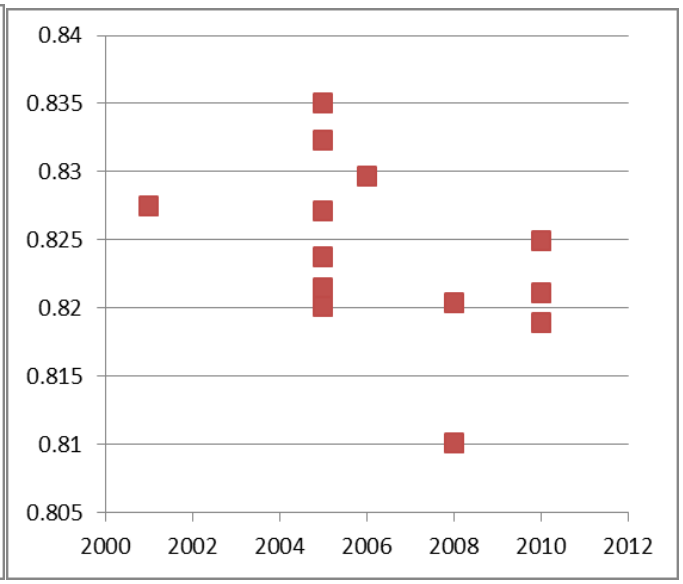

Figure 7: Hydropathy plot of envelope gene sequences of DENV 1 strains from India

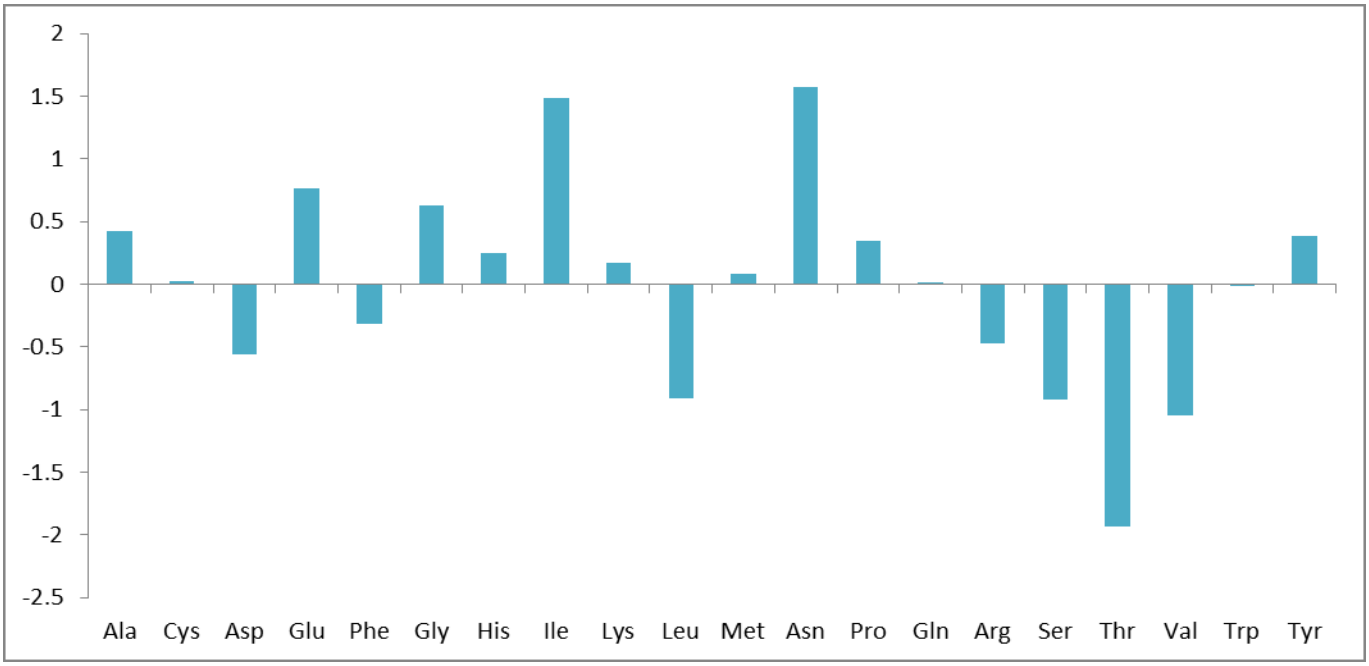

Figure 8: Chart of the differences in amino acid composition of DENV1 and DENV3 for each amino acid. Positive values imply higher frequencies in DENV3.

Materials and Methods

Sequence data of the envelope genes from across the world including India of the dengue virus serotype 3 and 1 (DENV3 and DENV1) were downloaded from the NCBI website (last accessed Sep15, 2015); the number of dengue gene and genomic sequences from India were comparatively less than from other countries like the USA, China, Australia, etc., but we downloaded as many complete gene sequences as were available. Plots of the gene sequences were made in a 2D graphical representation scheme with ACGT axis system for visualization of the sequence structure and analyzed for numerical characterisation. For the graphical representation the sequences are plotted as a walk in a 2D grid taking one step in the negative x-direction for an adenine, in the positive y-direction for a cytosine, the positive x-direction for a guanine and in the negative y-direction for a thymine/uracil. This yields (x,y) values for each base in a sequence and successive plots of the points generate a curve in the 2D graph [13]. We also used an alignment based procedure, the software MEGA 5.2, to draw phylogenetic trees showing the evolutionary relationship between the sequences. The transition/transversion ratio along with the amino acid composition of the envelope gene sequences were determined using the same software. The hydropathy index was determined using ExPasy server.

Conclusion

From phylogenetic as well as through 2D graphical representation point of view it is evident that Indian DENV3 strains are closely related to American strains, whereas Indian DENV1 strains are similar to Asian strains.

From the genetic point of view, we hypothesize from the hydropathy index and amino acid differences that morphological changes are occurring in the envelope gene structure in recent times. Such changes could be leading to enhanced viral pathogenecity and might explain part of the high incidence of dengue cases being observed now.

The lack of adequate representations of sequences of the dengue genes and genomes from India makes it difficult to define any trends with good statistics. However, the analyses we have done so far does indicate the propensity of morphological changes in the dengue envelope gene which could lead to higher pathogenicity. These exercises serve to show the urgency of comprehensive genetic surveillance of the dengue virus to anticipate further damaging changes in the viral sequence.

Acknowledgments

The authors would like to acknowledge with thanks the opportunity provided by the organizers of the MOL2NET-1 conference to present our findings to a general audience.

Author Contributions

S Dey worked on the problem and wrote the paper, which A Nandy critically reviewed and edited.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Whitehead S S et al, Prospects for a dengue virus vaccine, Nature Rev Microbiol 5 (2007) 518-528.
2. Guzman M G et al, Do escape mutants explain rapid increases in dengue case – fatality rates within epidemics?, Lancet 355 (2000) 1902-1903.
3. Fatima Z et al, Serotype and genotype analysis of dengue virus by sequencing followed by phylogenetic analysis using samples from three mini-outbreaks – 2007-2009 in Pakistan, BMC Microbiology (2011) 11:200.
4. Mukhopadhyay S, Kuhn RJ, Rossmann MG (2005) A structural perspective of the Flavivirus life cycle. Nat Rev Microbiol 3: 13-22
5. Heinz FX, Stiasny K (2012) Flaviviruses and falvivirus vaccines. Vaccine 30: 4301-4306.
6. .Rice, C.M., Lenches, E.M., Eddy, S.R., Shin, S.J., Sheets, R.L., Strauss, J.H., Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. Science 229 (1985), 726– 733.
7. Gubler DJ (1998) Dengue and Dengue Hemorrhagic Fever, Clin Microbiol Rev 11(3): 480-496.
8. Chen Y, Maguire T, Hileman RE, Fromm JR., Esko JD, Linhardt RJ, Marks RM (1997) Dengue virus infectivity depends on envelope protein binding to target cell heparin sulfate. Nat Med 3: 866-871.

9. Leitmeyer KC, Vaughin DV, Watts DM, Salas R, Chacon IVD, Ramos C, Rico-Hesse R (1999) Dengue Virus Structural Differences That Correlate with Pathogenesis. J Viral 73: 4738-4747.
10. Chen Y, Maguire T, Marks RM (1996) Demonstration of binding of dengue virus envelope protein to target cells. J Virol 70: 12, 8765-8772.
11. Sanehez IJ, Ruiz BH (1996) A single nucleotide change in the E protein gene of dengue virus 2 Mexican strain affects neurovirolence in mice. J Gen Virol 77: 2541-2545.
12. Anoop M, Mathew A J. Jayakumar B, Issac A, Nair S, Abraham R, Anupriya MG, Sreekumar E (2012) Complete genome sequencing and evolutionary analysis of dengue virus serotype 1 isolates from an outbreak in Kerala, South India. Virus genes 45: 1-13.
13. Nandy, A. A new graphical representation and analysis of DNA sequence structure: Methodology and Application to Globin Genes, Current Sci. 66(4), 309-314 (1994).
14. Dey, S.; Nandy, A.; Nandy, P.; Das, S. Diversity and evolution of the envelope gene of dengue virus type 1 circulating in India in recent times. Int. J Bioinfor Res and Appl. (in press)