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

Ashesh Nandy, Centre for Interdisciplinary Research and Education, Jodhpur Park, Kolkata 700068, India, Email: anandy43@yahoo.com

Mol2Net, 2015, 1 (Section B), slides 1-12,ppt, doi: xxx-xxxx

Abstract

A fresh wave of Dengue infection, particularly Dengue serotype 1and 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 strong likeness 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 morphologtical 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.

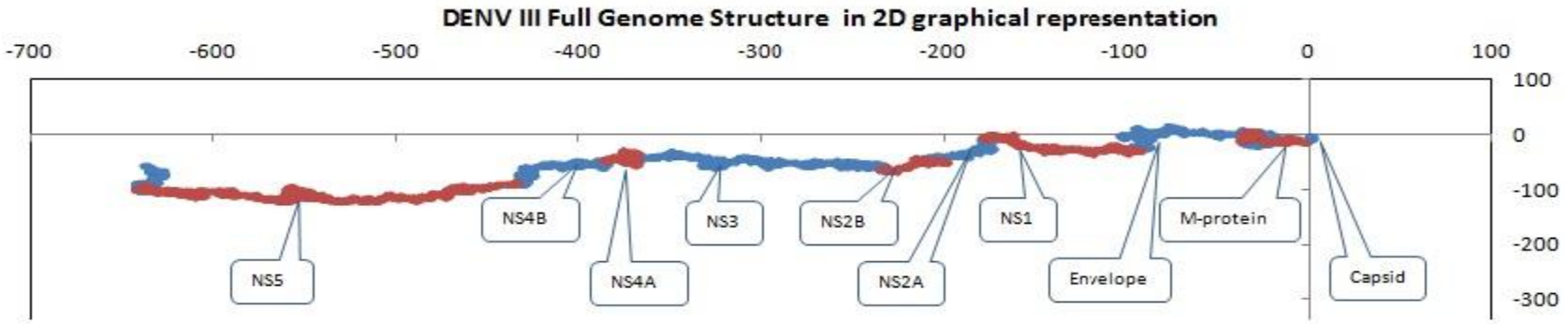
Dengue virus (DENV) in brief

- Infects >50 – 100 million people across 60 countries annually [1]

- Fatalities ~ 50 – 100 thousands per year [2,3]

- DENV (*Flavivirus* sp.Falviviridae) consists of four antigenically distinct serotypes (DENV 1 to 4).

- +ve ssRNA genome, comprises 10 segments: 3 structural proteins - capsid (C), premembrane/membrane (PrM/M), envelope (E); 7 non-structural (NS) proteins, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 as in the following figure [4]:



Methods and Materials

1. Sequence Data downloaded from NCBI GenBank (last accessed Sep 15, 2015).
2. Comparative sequence analysis - visualization by 2D graphical representation method (5).
3. Phylogenetic tree, Transition/Transversion ratios and Amino acid composition determination through MEGA 5.2 (6).
4. Hydropathy Index calculation through ExPasy Server (7).

Phylogenetic Analysis

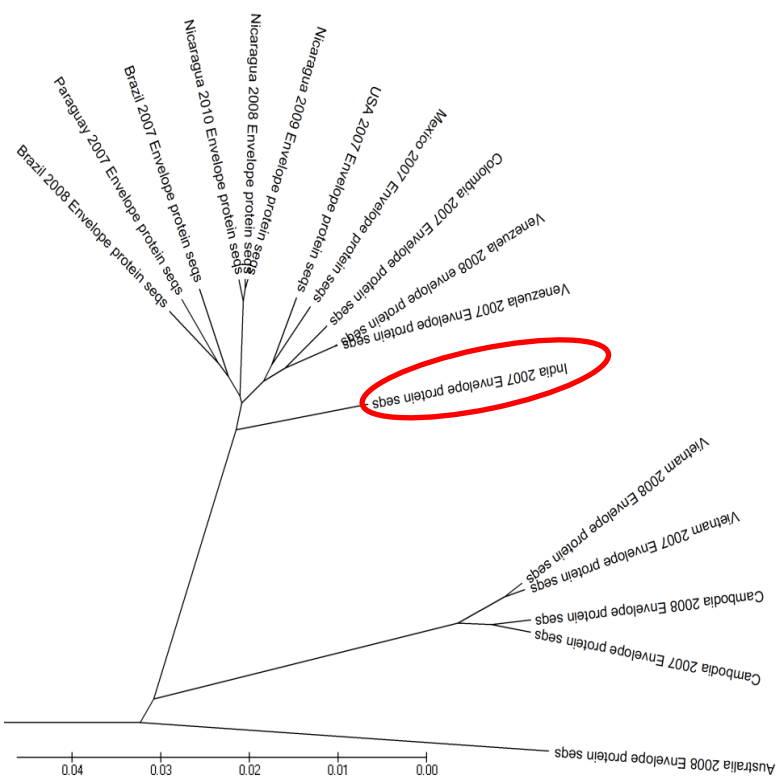


Figure 1 shows phylogenetic relationship between 18 envelope gene sequences of DENV 3 from various countries in recent times

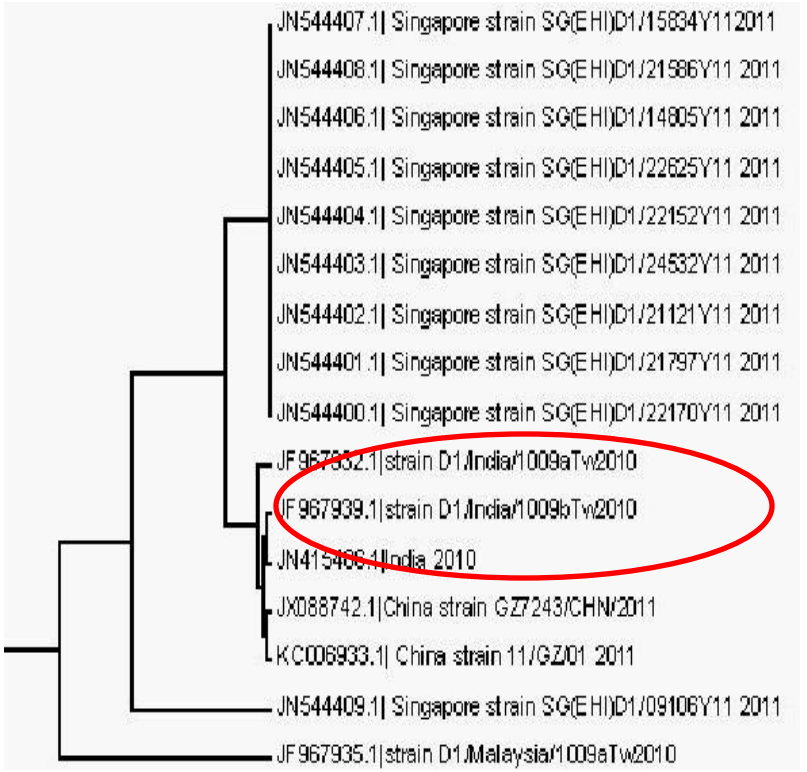
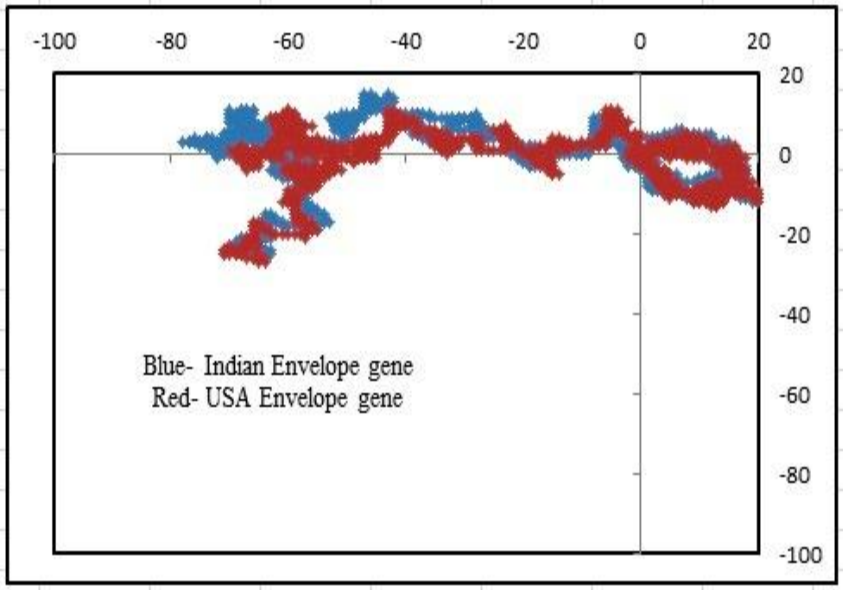


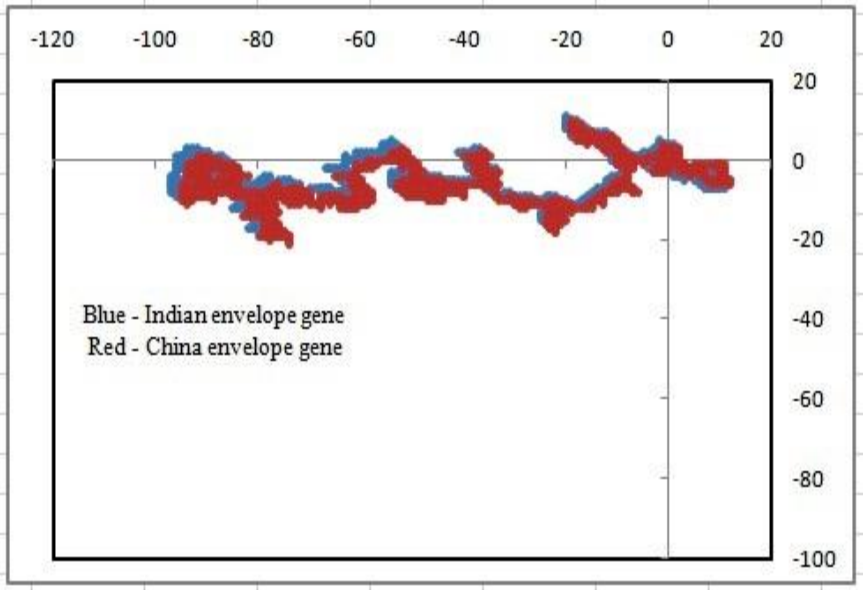
Figure 2. shows phylogenetic relationship between envelope gene sequences of DENV 1 from Asian countries in recent times

Fig 1 shows that DENV 3 Envelope gene for India is closely related to American strains, whereas for DENV 1, fig 2, the Indian gene is more closely related to Chinese and other Asian strains (8).

Genome comparative analysis by 2D graphical representation method.



2D graphical representations of DENV3 envelope gene sequences from India and USA



2D graphical representations of DENV1 envelope gene sequences from India and China.

Graphical representations above clearly shows that the envelope gene of DENV 3 strains from India (Locus ID: JQ686083) are closely related to American strains (EU596494) whereas in case of DENV 1 strains from India (JF967939) are related to Asian strains (China KC006933).

Transition/Transversion Matrix

Caption Expert

Mammalian Globin Genes

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*	-

Influenza A surface 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*	-

M5: MCL Transition/Transversion Bias

Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution

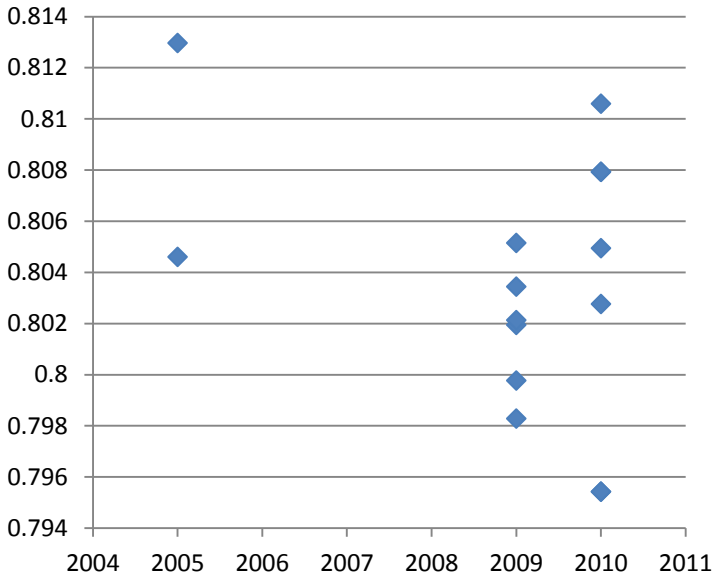
	A	T	C	G
A	-	*1.22*	*1.09*	**12.05**
T	*1.77*	-	**29.69**	*1.52*
C	*1.77*	**33.01**	-	*1.52*
G	**14.05**	*1.22*	*1.09*	DENV 3

Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution

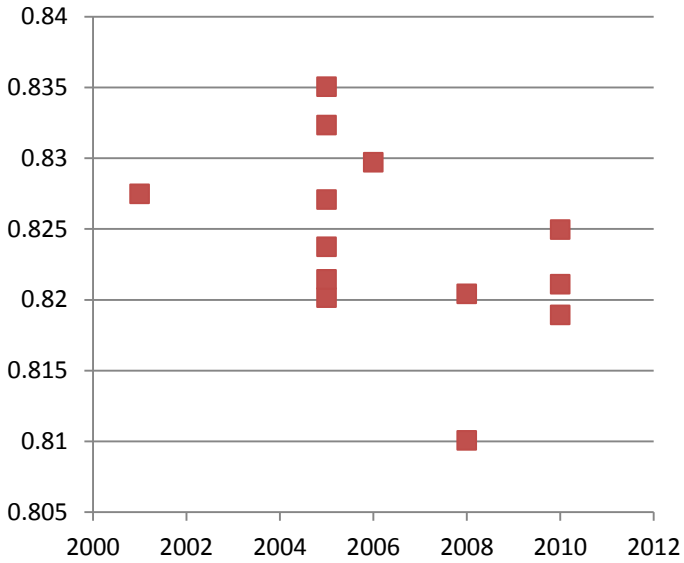
	A	T	C	G
A	-	*1.57*	*1.49*	**11.9**
T	*2.34*	-	**28.85**	*1.96*
C	*2.34*	**30.32**	-	*1.96*
G	**14.2**	*1.57*	*1.49*	DENV 1

Numbers in bold are the transition frequencies (in %), others are transversion frequencies (%). The data above show that while the rate of transition to transversion mutations is about 55: 45 in mammalian and influenza genes, in the case of dengue envelope genes this ratio at 88:12 is significantly different.

Hydropathy Index



Hydropathy plot of envelope gene sequences of DENV 3 strains from India



Hydropathy plot of envelope gene sequences of DENV 1 strains from India

Hydrophobicity of envelope protein of the Indian serotype (DENV 3) shows a slight tendency to increase with time but in case of the Indian serotype (DENV 1) shows tendency towards decreasing with time implying morphological changes in the protein structure.

Amino acid compositional difference between DENV1 and DENV3

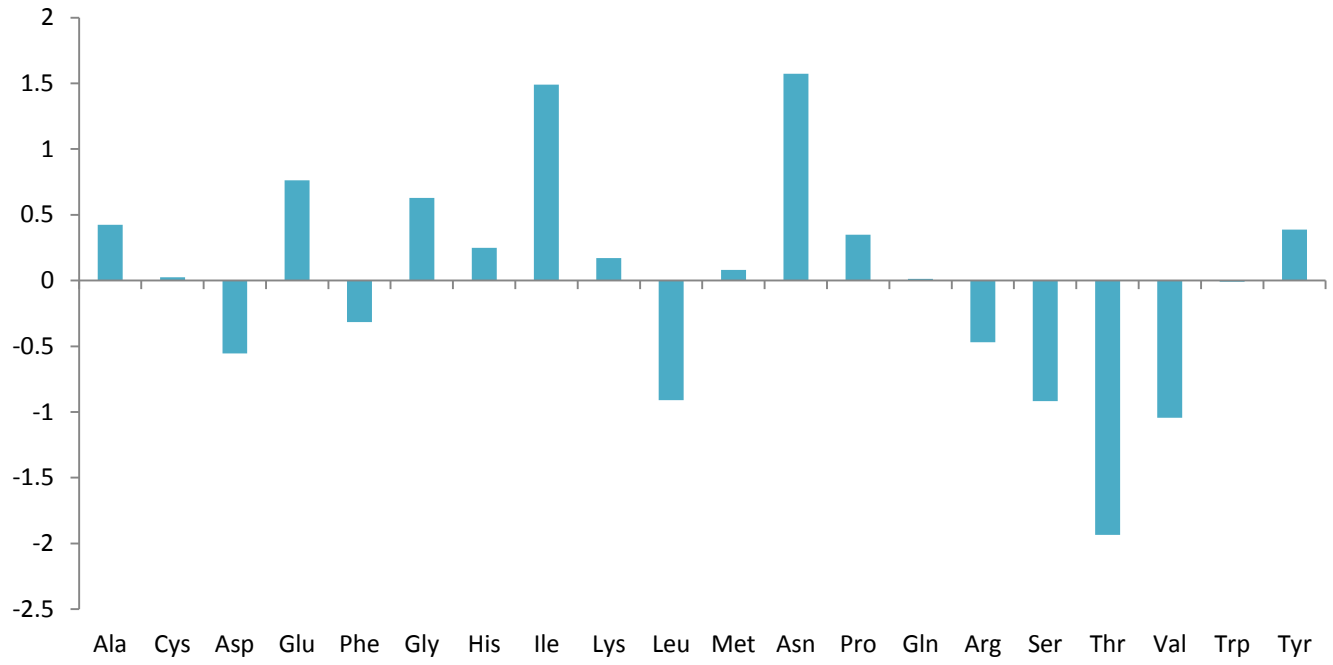


Chart of the differences in amino acid composition of DENV1 and DENV3 for each amino acid. Positive values imply higher frequencies in DENV3.

The chart above 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.

Causes of morphological changes in Envelope protein structure of DENV3 as compared to DENV1

> 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 contain 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.

Conclusions

- From phylogenetic as well as through 2D graphical represntation point of view it is evident that Indian DENV 3 strains are closely related to American strains, whereas Indian DENV 1 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.

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. 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.
5. Nandy, A. A new graphical representation and analysis of DNA sequence structure: I. Methodology and Application to Globin Genes, Current Sci. 66(4), 309-314 (1994).
6. www.megasoftware.net
7. www.**expasy**.org
8. 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)