BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 9(12)

Comparison of full-length genomics sequences between dengue virus serotype 3, parental strain, and its derivatives, and B-cell epitopes prediction from envelope region

Siriwattana Thaisonthi[1], Jundee Rabablert[1] & Sutee Yoksan[2]*

[1]Department of Biology, Faculty of Science, Silpakorn University, Nakhon Pathom, Thailand; [2]Center for Vaccine Development, Institute of Molecular Biosciences, Mahidol University, Nakhon Pathom, Thailand; Sutee Yoksan – Email: sutee.yok@mahidol.ac; *Corresponding author

Received March 29, 2013; Revised June 4, 2013; Accepted April 05, 2013; Published July 12, 2013

Abstract:
Biological markers are normally used to evaluate the candidate of live-attenuated dengue vaccines. D3V 16562 Vero 23 and D3V 16562 Vero 33 which were derivatives of D3V 16562, parental strain, showed the similar biological data. We used molecular techniques and computational tools to evaluate these derivatives. The nucleotide and amino acid sequences of the derivatives were compared to their parent. The secondary structures of untranslated regions and B-cell epitopes were predicted. The results showed that nucleotide substitutions mostly occurred in NS5 and NS5 of V2 was unusual because of amino acid change at 3349 (tryptophan →stop codon). The nucleotide substitutions in 5'UTR, prM, E, NS1, NS2A, NS3, and 3'UTR were 4, 1, 2, 2, 1, 3, and 2, respectively. The secondary structure of 5'UTR of V2 was different from P and V1. The secondary structure of 3'UTR of V2 was similar to P and certainly distinct from V1. Furthermore, B-cell epitopes prediction revealed that there were 21 epitopes of envelope and the interesting epitope was at position 297-309 because it was in domain III in which the neutralizing antibody is induced. For this study, the attenuation of derivatives was caused by the nucleotide substitutions in 5'UTR, 3'UTR, and NS5 regions. The genotypic data and B-cell epitope make the derivatives attractive for the chimeric and peptide DENV vaccine development.

Key words: Dengue virus, Live-attenuated dengue vaccine, dengue epitope, 5'-3'UTR secondary structure.

Background:
Dengue viruses (DENV) cause a major viral mosquito-borne human infection which are the member of *flavivirus* genus and there are 4 distinct serotypes (DENV1-4) **[1].** DENV has spread widely in tropical and subtropical regions due to recent changes in human ecology and travellers to areas where DENV is endemic are the potential source of the spread **[2].** About 50 million, dengue infections have been occurred annually and 2.5 billion people live in dengue endemic countries **[3].** DENV can be transmitted to human by the bite of the infected vector. *Aedes aegypti* and *Aedes albopictus* are the vectors of dengue which found in tropical areas. Primary infection provides homotypic immunity and probably lifelong but not heterologous immunity. The infection results in a spectrum of clinical illness ranging from asymptomatic to severe symptom including dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) which results from secondary infection **[4].**

The genome of DENV is a single-stranded positive-sense RNA of approximately 11 kb and contains a single open reading frame that is expressed as a large polyprotein. The RNA genome is capped at the 5' end and lack a 3' terminal poly (A) tail. The gene organization is 5'UTR-capsid (C)-premembrane / membrane (prM/M)-envelope (E)-nonstructural protein 1 (NS1)-NS2A-MS2B-NS3-NS4A-NS4B-NS5-3'UTR. This polypeptide is co-translationally and post-translationally

processed by viral and cellular protease into 3 structural proteins (C, M and E) and 7 nonstructural proteins (NS1-NS5) **[5].**

The envelope protein is responsible for several activities, including dengue binding to the host cell receptors and entry into the target cell. Hence, this protein affects host range, cellular tropism and, in part, the virulence of virus **[6].** Moreover, the protective and neutralizing antibody can be induced by envelope protein **[7].** It is now concerned that it should be the target for dengue vaccine development. Although, the structure-function relationships of the dengue virus glycoprotein E was illustrated but the location of well-defined B-cell for glycoprotein E are still unknown. This study was performed to find the regions associated with the attenuation of DENV3, to find B-cell epitopes related to neutralizing antibody inducement and we expected that these data could be used in chimeric or peptide vaccine development.

Methodology:

Samples

The parental strain (D3V 16562, represented by P) and its 2 derivatives (D3V 16562 Vero-23 and D3V 16562 Vero-33, represented by V1 and V2, respectively) were obtained from Assoc. Prof. Sutee Yoksan, Center for Vaccine Development, Mahidol University, Thailand.

Primer design

All whole genome DENV3 in Genbank were aligned by CLUSTALW to find conserved region among them. 5 primer pairs were designed which cover the whole genome.

cDNA synthesis

RNA was extracted by using E.A.N.A.™ Viral RNA Kit (Omega, Norcross, GA, USA) according to the manufacturer's protocol and use immediately or store at -80°C until use. cDNA was synthesized by using the SuperScript®VILO™ cDNA Synthesis Kit (Invitrogen Life Technologies,CA, USA) according to the manufacturer's protocol and store at -20°C

Polymerase Chain Reaction

cDNA of all DENV3 was amplified by using Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen Life Technologies,CA, USA) according to the manufacturer's protocol. The designed forward and reverse oligonucleotide primers, **Table 1 (see supplementary material),** were used to amplify 5 overlapping amplicons. The cycling conditions for 35 cycles are listed below:

Initial denaturation	95°C	30 second
Initial denaturation	95°C	30 second
Initial annealing	56°C	30 second
Initial extension	68°C	2.30 minutes
Final extension	68°C	5 minutes

PCR product (about 2kb) were separated in a 1% agarose gel and visualized under UV light with SYBR Gold (Invitrogen Life Technologies,CA, USA)

Transformation and Analyzing Transformants

The reaction was done by using TOPO TA Cloning® Kit for Sequencing (Invitrogen Life Technologies,CA, USA) according to the manufacturer's protocol. Each transformation was spread on prewarmed selective plate and incubated overnight at 37°C. 2-6 colonies were taken and cultured overnight in LB medium containing 50 µg/ml kanamycin. To isolate plasmid, PureLink™ Quick Plasmid Miniprep Kit (Invitrogen Life Technologies,CA, USA) was applied and processed accordingly the manufacturer's protocol. The insertion of PCR product could be checked by the methods below.

Polymerase Chain Reaction

About 5-10 colonies were picked and resuspended invidually in 50 µl of the PCR cocktail which comprised forward and reverse primers of the target and follow the cycle above.

Restriction analysis

Plasmid was cut by using FastDigest® *EcoRI* (Fermentas, Glen Burnie, Maryland, USA) according to the manufacturer's protocol and check the result on 1% agarose gel.

Sequence analysis

Sequence alignment will be performed using CLUSTALW algorithm **[8]** and optimized by visual inspection.

5′ and 3′UTR secondary structure Prediction

We predicted secondary structure of 5′ and 3′ UTR of all DENV3 via entering the sequences into the RNAfold WebServer (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) to assess and anticipate the viral replication and mRNA translation.

B cell epitopes prediction

The IEDB Analysis Resource (http://tools.immuneepitope.org /tools/bcell/iedb_input) shows the peptide sequences of envelope protein which play a role of antibody response inducement. Antigenic peptides are determined using the method of Kolaskar and Tongaonkar **[9].** We enter the amino acid sequence of E of all DENV3 strains.

Results & Discussion:

Identification of nucleotide and amino acid substitutions in heterogeneous regions at full genomic scale of DENV3 and its derivatives

In current, vaccine development is still not succeeded because of some reasonable factors such as lacking of a good animal model and fact that humans and mosquitoes represent the only two natural hosts **[10]** and reliable surrogate markers of immunity **[11].** However, biological markers are normally utilized to evaluate live-attenuated DENV vaccine. These markers are comprised of plaque size, temperature sensitivity, neurovirulence and neutralizing antibody. In this study, we had 2 derivatives which act as the candidate DENV3 vaccines and they showed the similar biological markers. Then, we expected that the full-length genome comparison could aid us in indication of which derivative is the most reliable to vaccine development. This study showed that the full-length RNA genomes were 10,696 nt. The single open reading frame (ORF) was located at 95-10,267 nucleotide position, coding for a polyprotein of 3,390 amino acids. Chao *et al.* **[12]** reported that the transitions were generally higher than transversions but we found that the transversions were higher than the transitions and the replacement rates of derivatives are alike. The transitions and transversions of V1 were 39.4% and 60.6% while

those of V2 were 40.8% and 59.2%. In addition, the most transitions are A → G or G → A with 73.08% of V1 and 62.07% of V2. The substitutions in 5′UTR, prM, E, NS1, NS2A, NS3, NS5, and 3′UTR were 4, 1, 2, 2, 1, 3, 19, and 2, respectively and there was no substitution in NS2B, NS4A, and NS4B. This result accorded with Cáceres *et al*. **[13]** who reported that the substitutions were mostly found in NS5. However, we focused on 5′UTR, E, NS5, and 3′UTR because they concerned with viral replication, viral translation, and neutralizing antibody inducement.

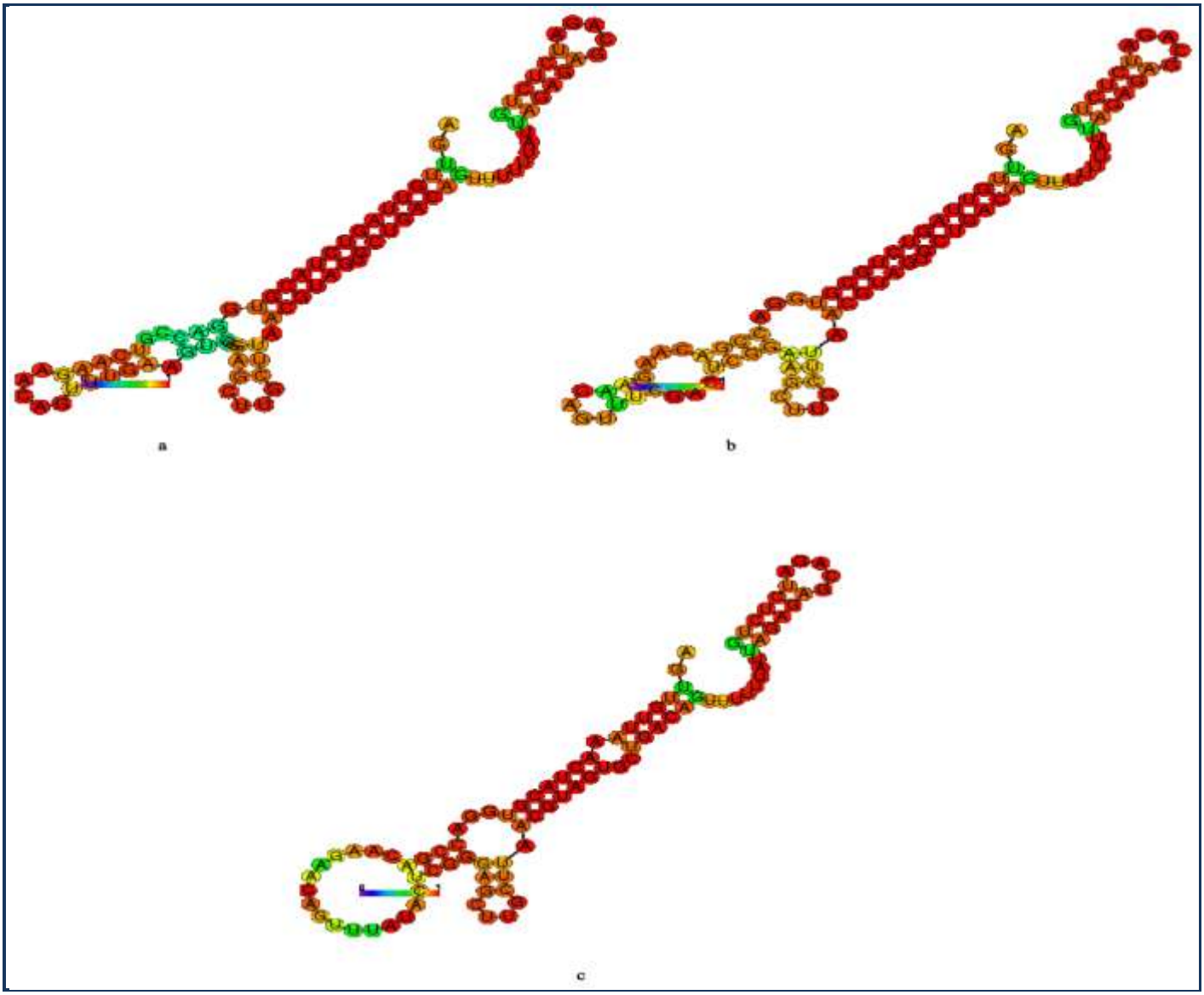

Figure 1: 5′UTR secondary structure prediction of **a)** D3V 16562; **b)** D3V 16562 Vero-23; **c)** D3V 16562 Vero-33.

5′ and 3′ UTR secondary structure prediction

The untranslated regions of the genome play roles in the regulation of translation and genome replication **[14, 15].** The secondary structure of 5′UTR influences the translation of the genome and serves as a site of initiation for positive-strand synthesis during RNA replication which 5′UTR is in the negative strand. The 3′UTR enhances translation of mRNA and can interact with the viral replicase proteins NS3 and NS5 **[16].** In this study, 5′UTR secondary structure of P and V1 were similar but V2 was distinct **(Figure 1)** because of nucleotide substitution. Sirigulpanit *et al.,* **[17]** reported that mutation in 5′UTR caused the partial attenuation of DENV2. We could say that the translation and replication of V2 was not as good as P due to unusual structure of V2. Furthermore, it could be one of the reasonable factors for the attenuation of V2. However, this incident was opposite 3′UTR that is to say the 3′UTR secondary structure of P and V2 were similar but V1 was different **(Figure 2).** Blaney JE Jr *et al.,* **[18]** reported that deletions in 3′UTR caused the attenuation of DENV3. It could be said that the unusual structure of V1 could cause lower replication than P because of low interaction efficacy of 3′UTR and the viral replicase proteins and this could cause the attenuation of V1.

Envelope protein and B-cell epitopes prediction

The changes in envelope could affect immunogenicity or cell entry **[14].** If there were a lot of changes in envelope of derivative, it might hard entry the cell. The question is if only envelope is interested for vaccine development, which derivative is suit for the aim. To answer this question, nucleotide and amino acid sequences of envelopes were

examined. In this study, the substitutions in envelope of V1 and V2 were 10 and 4, respectively. It could be said that the cell entry ability of V2 was more similar to P than V1. Then, we could assume that V2 was better for live-attenuated vaccine development.

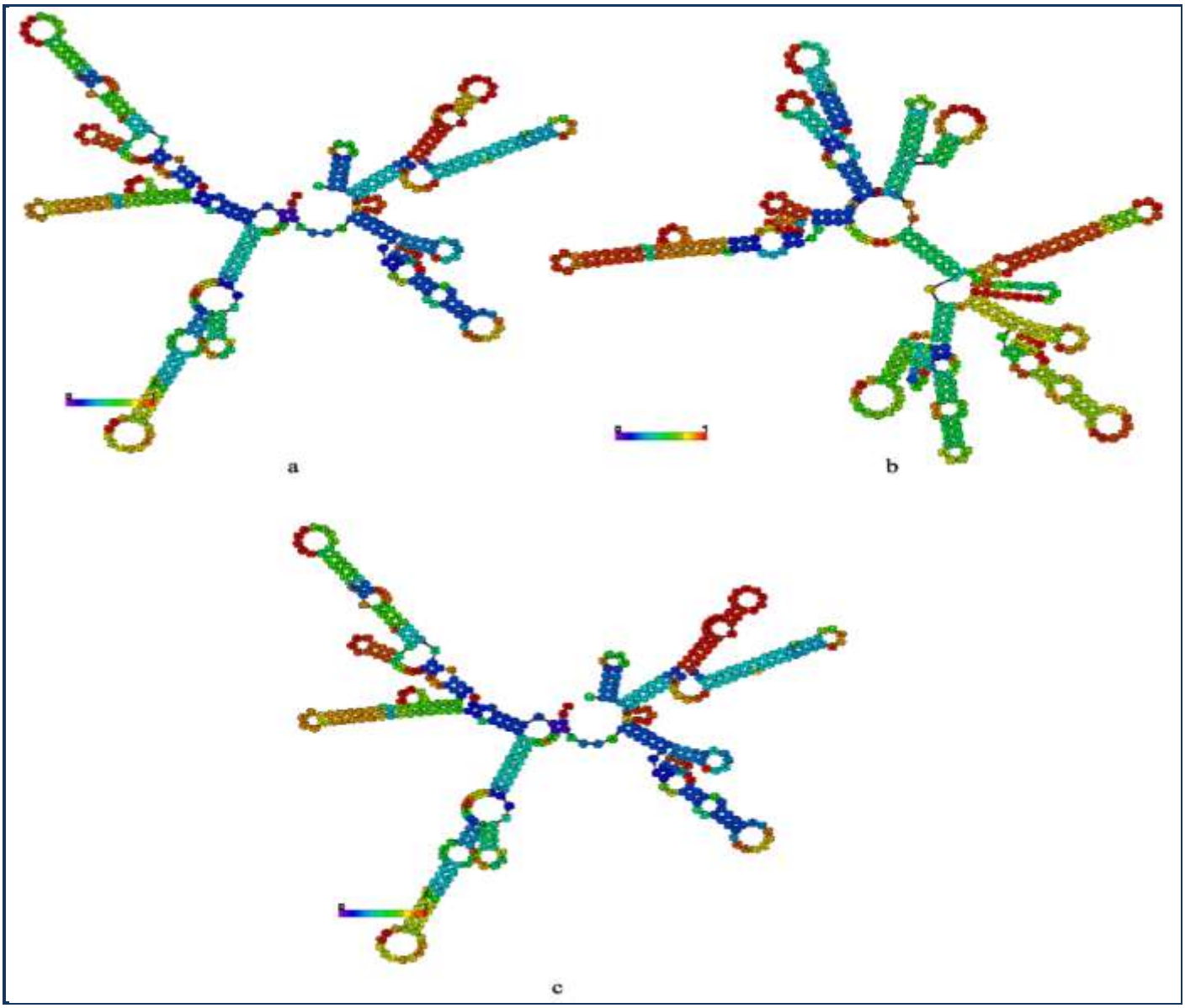

Figure 2: 3′UTR secondary structure prediction of **a)** D3V 16562; **b)** D3V 16562 Vero-23; **c)** D3V 16562 Vero-33.

The envelope is capable of inducing a protective immune response that is neutralizing antibody **[15, 19].** To define epitopes which were able to induce neutralizing antibody, the envelopes were analyzed by programs. Ilyas *et al*., [20] showed the result of 9 predicted B-cell epitopes of DENV3 envelope while Zhong *et al*., **[19]** showed 20 predicted B-cell epitopes which was similar to this study. We found 21 epitopes, **Table 2 (see supplementary material),** 17 epitopes of viruses shared the same sequences. Most monoclonal antibodies that neutralize virus infectivity do so, at least in part, by the blocking of virus adsorption. However, monoclonal antibodies specific for domain III were the strongest blockers of virus adsorption **[20-22].** In this study, there was the peptide sequence at position 297-309 of derivatives was in domain III (immunoglobulin-like domain). We believed that this peptide sequence could induce DENV3 neutralizing antibody and be likely to use for peptide vaccine development.

NS5

The most substitutions located at the third nucleotide of codon (data not shown), however, data in **Table 3 (see supplementary material),** showed the substitutions of derivatives which located at either the first or second nucleotide of codon. In NS5, G→A at 9346 and A→C at 9415 of V1 whereas G→C at 9346 and A→G at 9415 of V2. Because of these replacements, amino acids at 3116 and 3139 were definitely different among them. In addition, G→A at 10046 of V2 caused amino acid change (Tryptophan →stop codon) at 3349. Amino acid change at 3349 resulted in unusual NS5 protein of V2. Takahashi *et al*., **[23]** reported that NS5 is central to the function of the DENV replication and substitution in NS5 could reduce DENV2

replication, hence, the attenuation of V2 could be concerned owing to the reduction of enzyme activity which resulted from unusual NS5 protein of V2.

Conclusion:

We compared the full-length genome of viruses to find the regions which could cause the attenuation of virus. The main cause of the attenuation of V1 resulted from nucleotide substitutions in 3'UTR and NS5 whereas the attenuation of V2 resulted from nucleotide substitutions in 5'UTR and NS5, especially NS5 of V2 which was shorter than P and V1. Additionally, secondary structures of 5' and 3'UTR implied the low efficacy of the replication and translation which related to virus attenuation. The peptide sequence at position 297-309 was epitope in domain III of envelope which could induce neutralizing antibody. We expected that the sequences of 5'UTR, 3'UTR, and NS5 of V1 and V2 could be used for chimeric DENV vaccine development and the peptide sequence at 297-309 could be used for peptide vaccine development.

Acknowledgment:

This research was supported by Center for Vaccine Development, Mahidol University and partial funding from Faculty of Science, Silpakorn University, Grant No. RGP 54-11.

References:

[1] Christenbury JG *et al. J Virol Methods*. 2010 **169**: 202 [PMID: 20600330]
[2] Vasilakis N *et al. J Virol*. 2007 **81**: 9591 [PMID: 17553878]
[3] http://whqlibdoc.who.int/publications/2009/9789241547871_eng.pdf
[4] Leardkamolkarn V *et al. J Biomed Biotechnol*. 2010 **2010**: 1 [PMID: 20339476]
[5] Blaney JE *et al. J Virol*. 2001 **75**: 9731 [PMID: 11559806]
[6] Kuhn RJ *et al. Cell*. 2002 **108**: 717 [PMID: 11893341]
[7] Gubler DJ, *Clin Microbiol Rev*. 1998 **11**: 480 [PMID: 9665979]
[8] Thompson JD *et al. Nucleic Acids Res*. 1994 **22**: 4673 [PMID: 7984417]
[9] Kolaskar AS & Tongaonkar PC, *FEBS Lett*. 1990 **276**: 172 [PMID: 1702393]
[10] Lima DM *et al. Vaccine*. 2011 **29**: 831 [PMID: 21115054]
[11] Wilder-Smith A *et al. Curr Infect Dis Rep*. 2010 **12**: 157 [PMID: 21308524]
[12] Chao DY *et al. Virol J*. 2005 **2:** 72 [PMID: 16120221]
[13] Cáceres C *et al. Arch Virol*. 2008 **153**: 1967 [PMID: 18815724]
[14] Dos Santos FB *et al. Mem Inst Oswaldo Cruz*. 2002 **97**: 991 [PMID: 12471426]
[15] Murphy BR & Whitehead SS, *Annu Rev Immunol*. 2011 **29**: 587 [PMID: 21219187]
[16] Lindenbach BD *et al. Fields Virology*. 1101
[17] Sirigulpanit *et al. JGV*. 2007 **88**: 1748 [PMID: 17485535]
[18] Blaney JE Jr *et al. Vaccine*. 2008 **26**: 817 [PMID: 18191005]
[19] Zhong H *et al. JBiSE*. 2009 **2**: 123
[20] Ilyas M *et al. African Journal of Biotechnology*. 2011 **10**: 3528
[21] Crill WD & Roehrig JT, *J Virol*. 2001 **75**: 7769 [PMID: 11462053]
[22] Matsui K *et al. J Gen Virol*. 2010 **91**: 2249 [PMID: 20444995]
[23] Takahashi H *et al. Antiviral Res*. 2012 **96:** 305 [PMID: 23072882]

Edited by P Kangueane
Citation: **Thaisonthi *et al.*** Bioinformation 9(12): 622-628 (2013)

BIOINFORMATION

open access

Supplementary material:

Table 1: Names and sequences of primers

Primers name	Sequence
F1SYF3	5'-AGT TGT TAG TCT RCG TGG ACC G-3'
F1SYR3	5'-GTG GCC TCR AAC ATY TTC CC-3'
F2SYF3	5'-CCA CRG AGG ATG GRC AAG G-3'
F2SYR3	5'-GCC ACT ART GGT CCA GCC AT-3'
F3SYF3	5'-GCA TGA GGA AAA CAG ATT GGC-3'
F3SYR3	5'-GGA TCT GAR TAA GTG CGR GC-3'
F4SYF3	5'-ATA CCA GCT CTC TTT GAA CCA G-3'
F4SYR3	5'-CCA TGT TGG GYG TTT CTG G-3'
F5SYF3	5'-GGA ATG CTT GTG AGA AAY CCA C-3'
F5SYR3	5'-AGA ACC TGT TGA TTC AAC AGC AC-3'

Table 2: Nucleotide and amino acid substitutions in untranslated, structural protein, non-structural protein region of DENV3.

	Nucleotide	Nucleotide changes			Amino acid	Amino acid changes		
Region	Position	P	V1	V2	position	P	V1	V2
5′UTR	23	T	A	A				
5′UTR	36	G	C	A				
5′UTR	37	A	G	T				
5′UTR	39	G	C	C				
prM	625	A	T	T	209	I	L	L
E	1744	A	G	G	582	S	G	G
E	2012	A	G	G	671	K	R	R
NS1	2779	A	G	G	927	K	E	E
NS1	3385	C	A	A	1129	Q	K	K
NS2A	3631	T	C	C	1211	F	L	L
NS3	4687	C	A	A	1563	Q	K	K
NS3	4694	T	G	G	1565	V	G	G
NS3	5176	C	A	A	1726	P	T	T
NS5	9214	C	A	A	3072	P	T	T
NS5	9283	C	A	A	3095	P	T	T
NS5	9290	G	C	C	3097	G	A	A
NS5	9310	T	A	A	3104	L	M	M
NS5	9317	A	C	C	3106	D	A	A
NS5	9337	A	G	G				
NS5	9339	A	G	A	3113	K	E	E
NS5	9346	G	A	C				
NS5	9348	A	C	C	3116	G	S	R
NS5	9349	G	C	C				
NS5	9350	T	G	T	3117	V	R	L
NS5	9351	T	T	G				
NS5	9361	G	C	C	3121	A	P	P
NS5	9415	A	C	G	3139	T	P	A
NS5	9421	G	C	C	3141	G	R	R
NS5	9500	A	C	C	3167	D	A	A
NS5	9515	T	A	A	3172	I	N	N
NS5	9530	C	T	T	3177	A	V	V
NS5	10046	G	G	A	3349	W	W	stop
3′UTR	10665	C	A	A				
3′UTR	10666	C	A	A				

Table 3: B cells epitopes prediction of DENV3 envelope protein

No.	Start Position	End Position	Peptide			Peptide Length
			P	V1	V2	
1	18	33	ATWVDVVLEHGGCVTT	ATWVDVLLEHGGCVTT	ATWVDVVLEHGGCVTT	16
2	51	63	TQLATLRKLCIEG	TQLATLRKLCIEG	TQLATLRKLCIEG	13

3	79	84	EAILPE	EAILPE	EAILPE	6
4	88	97	QNYVCKHTYV	QNYVCKHTYV	QNYVCKHTYV	10
5	110	124	KGSLVTCAKFQCLES	KGSLVTCAKFQCLES	KGSLVTCAKFQCLES	15
6	126	134	EGKVVQHEN	EGKVVQHEN	EGKVVQHEN	9
7	136	146	KYTVIITVHTG	KYTVIITVHTG	KYTVIITVHTG	11
8	169	176	AEAILPEY	AEAILPEY	AEAILPEY	8
9	178	184	TLGLECS	TLGLECS	TLGLECS	7
10	211	219	FFDLPLPWT	FFDLPLPWT	FFDLPLPWT	9
11	233	239	ELLVTFK	ELLVTFK	ELLVTFK	7
12	245	253	KQEVVVLGS	KQEVVVLGS	KQEVVVLGS	9
13	277	286	FAGHLKCRLK	FAGHLKCRLK	FAGHLKCRLK	10
14	297	309	YAMCLSSFVLKKE	YAMCLGSFVLKKE	YAMCLGSFVLKKE	13
15	316	323,324	GTILIKVE	GTILIKVEY	GTILIKVE	8,9
16	328	335	DAPCKIPF	DAPCKIPF	DAPCKIPF	8
17	351	358	TANPVVTK	TANPVVTN	TANPVVTK	8
18	374	380	SNIVIGI	SNIVIGI	SNIVIGI	7
19	423	429	VGGVLNS	VGGVLNS	VGGVLNS	7
20	431	448	GKMVHQIFGSAYTALFGG	GKMVHQIFGSAYTALFGG	GKMVHQIFGSAYTALFGG	18
21	456	464	GIGVLLTWI	GIGVLLTWI	GIGVLLTWI	9