Biochemical and Biophysical Research Communications 436 (2013) 478–485

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Complete genome sequencing and evolutionary analysis of Indian isolates of Dengue virus type 2

Paban Kumar Dash *, Shashi Sharma, Manisha Soni, Ankita Agarwal, Manmohan Parida, P.V.Lakshmana Rao

Division of Virology, Defence R&D Establishment (DRDE), Jhansi Road, Gwalior, MP 474002, India

ARTICLE INFO

Article history:
Received 13 May 2013
Available online 10 June 2013

Keywords:
DENV-2
Complete genome
Phylogeny
Selection pressure
Molecular epidemiology

ABSTRACT

Dengue is the most important arboviral infection of global public health significance. It is now endemic in most parts of the South East Asia including India. Though Dengue virus type 2 (DENV-2) is predominantly associated with major outbreaks in India, complete genome information of Indian DENV-2 is not available. In this study, the full-length genome of five DENV-2 isolates (four from 2001 to 2011 and one from 1960), from different parts of India was determined. The complete genome of the Indian DENV-2 was found to be 10,670 bases long with an open reading frame coding for 3391 amino acids. The recent Indian DENV-2 (2001–2011) revealed a nucleotide sequence identity of around 90% and 97% with an older Indian DENV-2 (1960) and closely related Sri Lankan and Chinese DENV-2 respectively. Presence of unique amino acid residues and non-conservative substitutions in critical amino acid residues of major structural and non-structural proteins was observed in recent Indian DENV-2. Selection pressure analysis revealed positive selection in few amino acid sites of the genes encoding for structural and non-structural proteins. The molecular phylogenetic analysis based on comparison of both complete coding region and envelope protein gene with globally diverse DENV-2 viruses classified the recent Indian isolates into a unique South Asian clade within Cosmopolitan genotype. A shift of genotype from American to Cosmopolitan in 1970s characterized the evolution of DENV-2 in India. Present study is the first report on complete genome characterization of emerging DENV-2 isolates from India and highlights the circulation of a unique clade in South Asia.

1. Introduction

Dengue virus (DENV) is a rapidly-emerging mosquito-borne virus which infects an estimated 50–100 million people annually. DENV cause a wide spectrum of symptoms, from mild disease (dengue fever) to a severe hemorrhagic form (dengue hemorrhagic fever (DHF)/dengue shock syndrome (DSS)) [1]. DENV comprises of four closely related but antigenically and genetically distinct serotypes (DENV-1 through 4), belonging to genus *Flavivirus* of the family *Flaviviridae* [2]. Further, DENV-2 has been classified into six genotypes, including a sylvatic genotype [3,4]. DENV is an enveloped, single-stranded positive-sense RNA virus. The genomic RNA is approximately 11 kb in length and encodes three structural proteins: C (core), M (membrane), and E (envelope) and seven non-structural (NS) proteins: NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5. The coding region is flanked by non-coding regions at both 5′ and 3′ end [5].

Infection with any DENV serotype generally leads to a mild, self limiting febrile illness (dengue fever, DF). Few cases of DF leads to life threatening DHF and DSS. The pathogenesis of DHF and DSS is poorly understood. Several hypotheses, like antibody dependent enhancement (ADE), involvement of a virulent viral genotype, and host factors have been suggested as possible mechanism of pathogenesis [6,7].

In the past decade, number of cases have increased at an alarming rate and dengue emerged as a major public health problem in India. A recent modeling study estimated that India accounts for a third of global dengue infections [8]. DENV-2 was the first DENV serotype isolated in India in 1956 and since then all the four serotypes of DEN viruses have been isolated from various outbreaks [9]. The first major DHF outbreak in India was reported in Delhi in 1996 and DENV-2 was identified as its etiology [10]. Since then, persistence of DENV-2 as a major etiology of dengue outbreaks in different parts of India has been well documented [9,11,12]. The endurance and reemergence of DENV-2 was recently reported in Puerto Rico in spite of large scale changes in the circulating serotypes [13].

Though DENV were isolated in India since 1956, information regarding the genetic makeup of these viruses is still very limited.

* Corresponding author. Fax: +91 751 2341148.
E-mail address: pabandash@rediffmail.com (P.K. Dash).

0006-291X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2013.05.130

The evolutionary analysis based on C-prM gene junction and envelope gene of DENV-2 led to identification of two genotypes (American and cosmopolitan) in India [9–11]. Complete genome of Indian DENV-3 and DENV-1 was reported recently [14,15] but complete genome information of none of the other Indian DENV serotypes is available. The lack of complete genome information remained as a major bottleneck in understanding the circulating DENV-2 in India. Keeping in view of the above facts, the present study was undertaken to decipher the complete genome sequence information of four recent DENV-2 (2001–2011) isolates and one older DENV-2 from 1960, isolated from a viremic patient in India. Extensive phylogenetic analysis was also carried out to understand the genetic relationship of Indian DENV with other global isolates of last five decades.

2. Materials and methods

2.1. Dengue virus type 2

Four DENV-2 isolates from serum samples of viremic human patients collected from different parts of India during 2001–2011 were employed in this study for the complete genome sequencing. These viruses were designated as GWL18 (Madhya Pradesh-2001), RR44 (Delhi-2009), 1392 (Andhra Pradesh-2010) and Od2112 (Odisha-2011). GWL18 and 1392 were isolated from DF patients, whereas, RR44 and Od2112 were isolated from DHF patients, as per WHO classification. In addition, one older Indian DENV-2 strain obtained from National Institute of Virology (NIV), Pune, India and designated as P23085 (Tamil Nadu-1960) at passage level 47 in suckling mice, was included. The four recent DENV-2 isolates in this study were at passage level 2 in C6/36 cells.

2.2. Extraction of viral RNA

Viral RNA was extracted from 140 μl of infected culture supernatant using QIAamp viral RNA mini kit (Qiagen, Germany) in accordance with the manufacturer's instructions. Finally, RNA was eluted in 50 μl of nuclease free water and stored at −80 °C until use.

2.3. Reverse transcription-polymerase chain reaction (RT-PCR)

A total of 25 overlapping amplicons spanning the complete genomic region were amplified using 50 primers (data not shown). The RT-PCR was carried out in two steps. First, complementary DNA (cDNA) was synthesized in a 10 μl reaction volume with RT mix comprising of 5X-RT buffer, dNTPs, RNasin® ribonuclease inhibitor and Moloney murine leukemia virus reverse transcriptase (MMLV-RT) (Promega, USA) with antisense primer (Operon, Germany). RT mix was incubated at 37 °C for 1 h followed by 99 °C for 5 min (for inactivation of MMLV-RT). Amplification of cDNA was carried out in 50 μl reaction volume with PCR mix containing *Pfu* DNA Pol 10× buffer with 20 mM $MgSO_4$, dNTPs (10 mM each), *Pfu* DNA polymerase (Promega, USA), using respective sense primer in a thermal cycler (BioRad, USA). The thermal profile of the PCR reaction was – initial denaturation at 95 °C for 2 min, followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 54–59 °C for 1 min, extension at 72 °C for 2 min and final extension at 72 °C for 10 min. The PCR products were gel purified from 1.2% agarose gel using the QIAquick gel extraction kit (Qiagen, Germany) and used as template in sequencing reaction.

2.4. Sequencing reaction

Double pass sequencing was carried out employing Big dye terminator cycle sequencing ready reaction kit (Applied Biosystems, USA). Briefly, each sequencing reaction was carried out in a final volume of 20 μl by mixing the Big Dye terminator mix containing the thermostable Ampli*Taq* DNA polymerase, dNTPs and four dye-labeled dideoxy nucleotide terminators (ddNTPs), 25 ng of purified PCR product, and 3.2 pmol of respective primer following standard cycle sequencing profile [14]. The reaction mixture was column purified and loaded on to the ABI 3130 automated DNA sequencer (Applied Biosystems, USA).

2.5. Sequence and phylogenetic analysis

The nucleotide sequences were aligned and analyzed using the SeqScape (Applied Biosystems, USA) and *EditSeq* and *MegAlign* modules of Lasergene 5 software package (DNASTAR Inc, USA). Multiple sequence alignments were carried out employing MUSCLE and Phylogenetic analysis was conducted using MEGA version 5.03 [16]. Tamura Nei model of nucleotide substitution with gamma-distribution rates among invariant sites (with 5 categories) (termed as TN93 + G + I model), found to be the best model in MEGA, was used to construct the Maximum-Likelihood tree. The tree topologies were evaluated using 1000 replicates of the data set. Two different phylogenetic trees were drawn based on the nucleotide sequence of complete coding region (10173 nucleotides) and complete envelope gene (1485 nucleotides) of representative DENV-2 of diverse geographical origins (Table S1). Further, the phylogenetic grouping was also confirmed using Bayesian analysis. The deduced amino acid was determined from the nucleotide sequence using the *EditSeq* module of Lasergene 5.

2.6. Selection pressure analysis

To assess the selection pressure acting on individual codons of the ORF, a dataset comprising of all genotypes DENV-2 ($n = 146$) was prepared (Table S2). This dataset was prepared following elimination of closely related sequences (>99.9% nucleotide identity) and sequences having ambiguous characters. The analysis was carried out using HyPhy open-source software package available under the Datamonkey web-server (http://www.datamonkey.org/) [17]. The ratio of non-synonymous (dN) to synonymous (dS) substitutions per site (dN/dS or ω) were estimated using four different approaches including single-likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), mixed effects model of evolution (MEME) and fast unbiased Bayesian approximation (FUBAR). Best nucleotide substitutions model as determined through the available tool in Datamonkey server was adopted in the analysis. GARD recombination analysis of the dataset was carried out before final analysis.

3. Results

The RT-PCR amplification resulted in generation of 25 specific amplicons, which were gel purified and subjected to sequencing. The editing and alignment of sequences of overlapping fragments led to deciphering a sequence of 10,670 nucleotides in length. These five complete genome sequences along with sequences of 36 other representative DENV-2 of diverse geographical origins were retrieved from GenBank (Table S1) and were aligned to the homologous genomic sequences of the prototype strain of DENV-2 (NGC) (Nucleotide sequence alignment not shown). Comparison of sequences revealed base substitutions scattered throughout the entire stretch of the alignment. No nucleotide deletions or insertions were detected in the protein coding region. However, several nucleotide deletions were recorded at the 3′NTR in all the genotype V viruses, including an eight nucleotide stretch (10276-GCAAA-ACT-10283). Surprisingly, this deletion was not noticed in

P23085 (a genotype V virus), sequenced in this study. In this stretch, there is an insertion of 'AAAGGAAAA' in P23085, compared to all other genotype V viruses analyzed in this study. Similarly, there is a unique 'C' insertion at position 10300, compared to other genotype V viruses. All the four recent Indian DENV-2 isolates sequenced in this study revealed 96.4–98.2% nucleotide sequence identity among them. Compared to a closely related DENV-2 (V2416), these viruses revealed an identity of 96.3–97.2% (Table 1). When these viruses were compared with the older Indian DEN-2 isolate (P23085), there was a sequence identity of 90.3–90.6%. Compared to the prototype DENV-2 (NGC), the recent Indian DENV revealed a sequence identity of 93.1–93.6% (Table S3). Overall, the nucleotide sequence identity among all the DENV-2 viruses was 89.9–100%. The genome was found to be AU rich (53.68%).

3.1. Amino acid sequence diversity

To determine the degree of relatedness of these viruses at protein level, the deduced amino acid sequences corresponding to complete ORF of all these DENV-2 were aligned. On the pair-wise comparison, the four recent Indian DENV-2 revealed 99.1–99.5% sequence identity among them. The recent Indian DENV-2 revealed around 99%, 97% and 98% amino acid sequence identity with the V2416, P23085 and NGC respectively (Tables 1 and S3). The sequence alignment revealed amino acid replacement throughout the aligned region. The amino acid substitutions among the recent Indian DENV-2 sequenced in this study along with closely related DENV-2 from Asian countries was shown in Table 2. A number of substitutions involving hydrophilic to hydrophobic amino acid and neutral to charged residues were identified. Further the substitutions compared to older Indian DENV-2 (P23085) and prototype NGC was provided in Table S4. A number of unique residues (GWL18: 92M, 430T, 898I; 1392: 102V, 163V, 1160L, 1313R, 1661R; RR44: 660V, 2266L, 2666D, 2927N, 3135H; Od2112: 903S, 916T, 963V, 2010E, 2258F, 2569M, 2950M) were identified in recent Indian DENV-2.

3.2. Phylogenetic analysis

The phylogenetic tree based on complete coding region (ORF) revealed five different genotypic groups (Fig. 1). All the four recent Indian DENV-2 isolates cluster into one group along with three Sri Lankan (V2416, V2421 and V2422) and a Chinese isolate (FJ11). These four viruses formed a close branch with five other Asian viruses; three from Indonesia (BA05i, 98900665 and 1023DN) and one each from Singapore (4155) and Australia (TSV01). All these isolates together belong to cosmopolitan genotype (IV). The older Indian isolate (P23085) cluster into American genotype (V) along with isolates from Tonga (74), Fiji (UH21), Venezuela (Ven2), Peru (IQT1797), and Colombia (V3358).The prototype DENV-2 (NGC) was grouped into Asian 2 genotype (I) along with a Colombian and a Chinese isolate. Four viruses from Thailand were grouped into a close branch and belonged to Asian 1 genotype (II) along with an isolate from Vietnam. The American/Asian genotype (III) was represented by viruses from Dominican Republic, Brazil, Cuba, Martinique, Colombia, Puerto Rico, Vietnam, Venezuela, China and Cambodia. The Bayesian phylogenetic analysis also revealed similar topology (Fig. S1).

The tree based on complete envelope gene was drawn by including a total of one hundred representative strains of DENV-2 from geographically diverse areas (Fig. 2). The tree revealed that the four recent Indian DENV-2 isolates sequenced in this study were grouped into cosmopolitan genotype (IV) along with a large number of isolates from India (since 1974 onwards), Sri Lanka, Bhutan, Bangladesh, China, Malaysia, Saudi Arabia, Uganda, Indonesia, Thailand, Singapore, Australia, Somalia, Senegal and Mexico. Another DENV-2 (P23085, sequenced in this study) was grouped into American genotype (V) along with other Indian DENV-2 (1956–1980) and isolates mostly from Americas (Puerto Rico, Tonga, Fiji, Colombia, Mexico, Venezuela and Peru). The Asian 2 genotype (I) was represented by the prototype strain (NGC) along with isolates from China, Taiwan, Vietnam, Colombia and Mexico. The Asian 1 genotype (II) was represented by the virus isolates exclusively from South East Asian countries viz., Malaysia, Myanmar, Vietnam, Thailand and China. The American/Asian genotype (III) comprised of viruses from both Americas and South East Asia. Further examination revealed two distinct clades, one comprising viruses from Americas (Brazil, Jamaica, Cuba, Martinique, Puerto Rico, Venezuela, Colombia, Dominican Republic) and other from Asia (Thailand, China, Cambodia, Myanmar and Vietnam). Four viruses of sylvatic origin from Malaysia, Senegal, Ivory Coast and Guinea were clustered into a separate basal group in the phylogenetic tree.

3.3. Selection pressure analysis

The site specific selection pressure in DENV-2 was analyzed using SLAC, FEL, MEME and FUBAR. In case of SLAC, FEL and MEME, sites with p-value <0.1; and in FUBAR, posterior probability <0.9, were considered significantly positive in the present study. The analysis revealed that the codons in DENV-2 were under strong negative selection. A total of 32 sites were identified to be under positive pressure by at least three methods (Table 3). Out of these, six sites (capsid-102; NS3-14, 186, 549; NS4b-19; NS5-401) with significant evidence of positive selection by at least two different approaches were identified. Out of all the approaches, SLAC could identify four sites (capsid-102; NS3-14; NS4b-19; NS5-401) to be under positive selections, which were also identified by more powerful methods like FEL and MEME.

4. Discussion

The complete genome of the Indian DENV-2 sequenced in this study was found to be 10,670 bases long with an ORF coding for 3391 amino acids. The genetic analysis revealed the uniqueness of the currently circulating DENV-2 in India, as these are found to be more than 3% divergent from closely related DENV-2 from

Table 1
Percentage sequence identity of the complete genome of Indian DENV-2, as compared to closely related DENV-2 isolates.

V2416 (SRIL-96)	V2421 (SRIL-03)	V2422 (SRIL-04)	FJ11 (CHIN-99)	GWL18 (INDI-01)	RR44 (INDI-08)	1392 (INDI-09)	Od2112 (INDI-11)	
***	97	97	97.1	97.2	96.8	96.9	96.3	**V2416 (SRIL-96)**
99.2	***	99.9	97	98.6	97.5	96.5	97.7	**V2421 (SRIL-03)**
99.2	100	***	96.9	98.6	97.5	96.4	97.7	**V2422 (SRIL-04)**
99	99.1	99.1	***	97.5	97	96.5	96.7	**FJ11 (CHIN-99)**
99.3	99.7	99.6	99.1	***	98.1	97	98.2	**GWL18 (INDI-01)**
99.3	99.5	99.5	99.2	99.5	***	96.7	97.1	**RR44 (INDI-08)**
99.5	99.2	99.2	99.1	99.3	99.3	***	96.4	**1392 (INDI-09)**
99	99.4	99.3	98.8	99.4	99.3	99.1	***	**Od2112 (INDI-11)**

Note: The upper-right and lower-left matrix corresponds to nucleotide sequences and amino acid sequences respectively.

Table 2
Description of unique amino acid substitutions in recent Indian DENV-2 compared to closely related DENV-2 isolates.

Sl. No.	AA[a] position (ORF)	AA[a] position (Protein)	V2416 (SRIL-96)	V2421 (SRIL-03)	FJ11 (CHIN-99)	GWL18 (INDI-01)	RR44 (INDI-08)	1392 (INDI-09)	Od2112 (INDI-11)
Capsid									
1	**71**	**71**	M	T	T	T	T	T	T
2	81	80	I	V	V	I	V	V	V
3	92	92	L	L	L	M	L	L	L
4	102	102	A	A	A	A	A	V	A
5	104	104	M	V	M	V	M	M	V
prM									
6	163	49	I	I	I	I	I	V	I
7	**248**	**134**	T	T	T	T	T	T	A
8	**262**	**148**	Y	Y	Y	Y	H	Y	H
Envelope									
9	421	141	I	V	V	V	V	I	V
10	**430**	**150**	A	A	A	T	A	A	A
11	442	162	I	I	I	I	I	V	I
12	602	322	I	V	I	V	I	I	V
13	660	380	I	I	I	I	V	I	I
14	**758**	**478**	A	S	S	S	S	S	S
15	764	484	V	V	V	V	I	V	V
NS1									
16	898	123	M	M	M	I	M	M	M
17	**903**	**128**	P	P	P	P	P	P	S
18	**916**	**141**	A	A	A	A	A	A	T
19	963	188	I	I	I	I	I	I	V
20	1101	326	D	E	E	E	E	E	E
NS2a									
21	**1129**	**2**	Y	H	H	H	H	H	H
22	1139	12	V	V	V	V	V	V	I
23	1160	33	V	V	A	A	A	L	V
24	1184	57	V	M	M	M	M	M	M
25	1260	133	I	I	I	I	I	I	V
26	**1313**	**186**	Q	Q	Q	Q	Q	R	Q
NS3									
27	**1643**	**168**	T	T	T	T	T	A	S
28	1661	186	R	K	K	K	K	R	K
29	2010	535	D	D	D	D	D	D	E
30	2042	567	I	I	I	I	I	V	I
NS4b									
31	2256	13	L	F	F	F	F	F	F
32	2258	15	L	L	L	L	L	L	F
33	2266	23	P	P	P	P	L	P	P
NS5									
34	**2510**	**19**	A	S	A	S	S	A	S
35	2521	30	K	R	R	R	R	K	R
36	2569	78	V	V	V	V	V	V	M
37	**2666**	**175**	N	N	N	N	D	N	N
38	**2687**	**196**	A	T	A	T	A	A	T
39	2927	436	D	D	D	D	N	D	D
40	**2953**	**462**	L	I	T	I	I	L	M
41	3012	521	D	E	E	E	E	D	E
42	**3118**	**627**	E	G	G	G	G	G	G
43	**3122**	**631**	G	G	G	G	G	S	S
44	3128	637	A	A	A	A	A	A	T
45	3132	641	I	V	I	V	T	I	V
46	3135	644	Q	Q	Q	Q	H	Q	Q
47	**3136**	**645**	N	D	D	D	D	D	D
48	3178	687	V	V	V	I	V	V	V
49	3291	800	R	R	R	R	R	K	R
50	3367	876	N	S	S	S	S	N	S

Note: The amino acid site involving hydrophobic to hydrophilic amino acid substitutions or vice versa are written in bold font.
The amino acid site involving substitutions of neutral to charged residues or vice versa are written in bold and italics font.
The amino acid positions are in respect to GenBank Accession No FJ882602 and NC_001474.2.
[a] AA: amino acid.

neighboring countries. The newer Indian DENV-2 isolates were also found 9% divergent from the older Indian strain (P23085), which is primarily attributed to genotype difference of these viruses. The higher passage history of the P23085 in suckling mice might have also contributed to these observations.

Compared to closely related DENV-2 isolates from other countries, the recent Indian DENV-2 revealed many unique residues in both structural and non structural proteins. The isolates from DHF patients (RR44 and Od2112) also shared amino acid substitutions. One of the important non-conservative amino acid substitution include Y148H in the prM protein, which is located within a predicted B cell epitope, suggesting its role in protective immunity. Maximum numbers of unique amino acid residues were identified in Od2112, which was the most recent isolate in this study. This

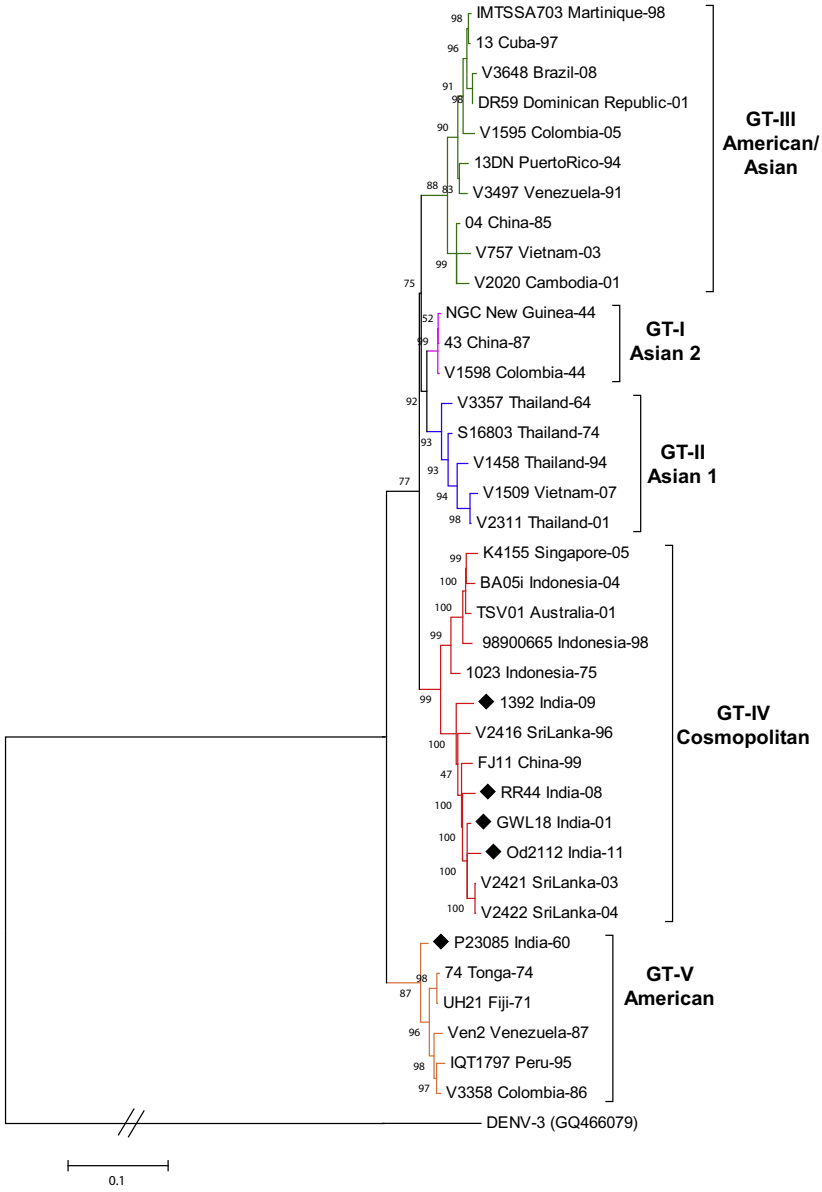

Fig. 1. Phylogenetic tree among Dengue-2 viruses generated by Maximum-Likelihood method based on the complete nucleotide sequence ORF region (10173 nucleotides). Each strain is abbreviated with isolate ID followed by country of origin and last two digits of the year of isolation. The DENV-2 sequenced in this study are marked with solid diamond (♦). Bootstrap values are indicated at the major branch points. An Indian DENV-3 (GenBank Acc. No. GQ466079) was used to root the tree.

was isolated from an eastern Indian state (Odisha), which witnessed an explosive Dengue outbreak in 2011 for the first time in its history. Three out of these seven unique residues were located in NS1. Substitution to a polar residue (A141T) in NS1 was found to be located within a Dengue complex specific B-cell epitope [18]. Other unique substitutions were recorded within helicase domain of NS3 (D535E) and methyl transferase domain of NS5 (V78M). These genetic changes might be linked to the high transmission rate and needs to be further confirmed.

The site-specific selection pressures on DENV-2 revealed strong purifying selection. This is primarily attributed to evolutionary constrains on arboviruses owing to their lifecycle involving two taxonomically divergent hosts viz., human and mosquito [4]. Majority of the positive sites were concentrated on envelope, NS3 and NS5. Many of these sites were identified for the first time in this study and their functional importance need to be addressed in future. Three positively selected sites in NS3 lies in the subdomain II of helicase (aa 180–618), which is involved in RNA unwinding and initiating the RNA replication [19]. Out of the nine positively selected sites in NS5, two (638 and 644) falls within a critical region of RdRP domain of flaviviruses (aa 624–647), which binds to IFNAR2 complex, there by suppressing host innate immune response [20]. The identification of two positively selected sites in NS4B is also interesting, owing to its role as an interferon-signaling inhibitor. Earlier, 13 positive sites were recorded in NS4B of sylvatic DENV-2 [21]. As expected, the sites under positive selection in E protein maps to immunogenic domains II and III, which were responsible for membrane fusion and attachment to cellular receptors.

The molecular phylogeny based on the complete ORF grouped the recent Indian DENV-2 into genotype IV along with isolates from Asian countries. The genotype V was represented by the older Indian DENV-2 (P23085) along with isolates from Americas. The genotype classification of DENV is based on an arbitrary value of

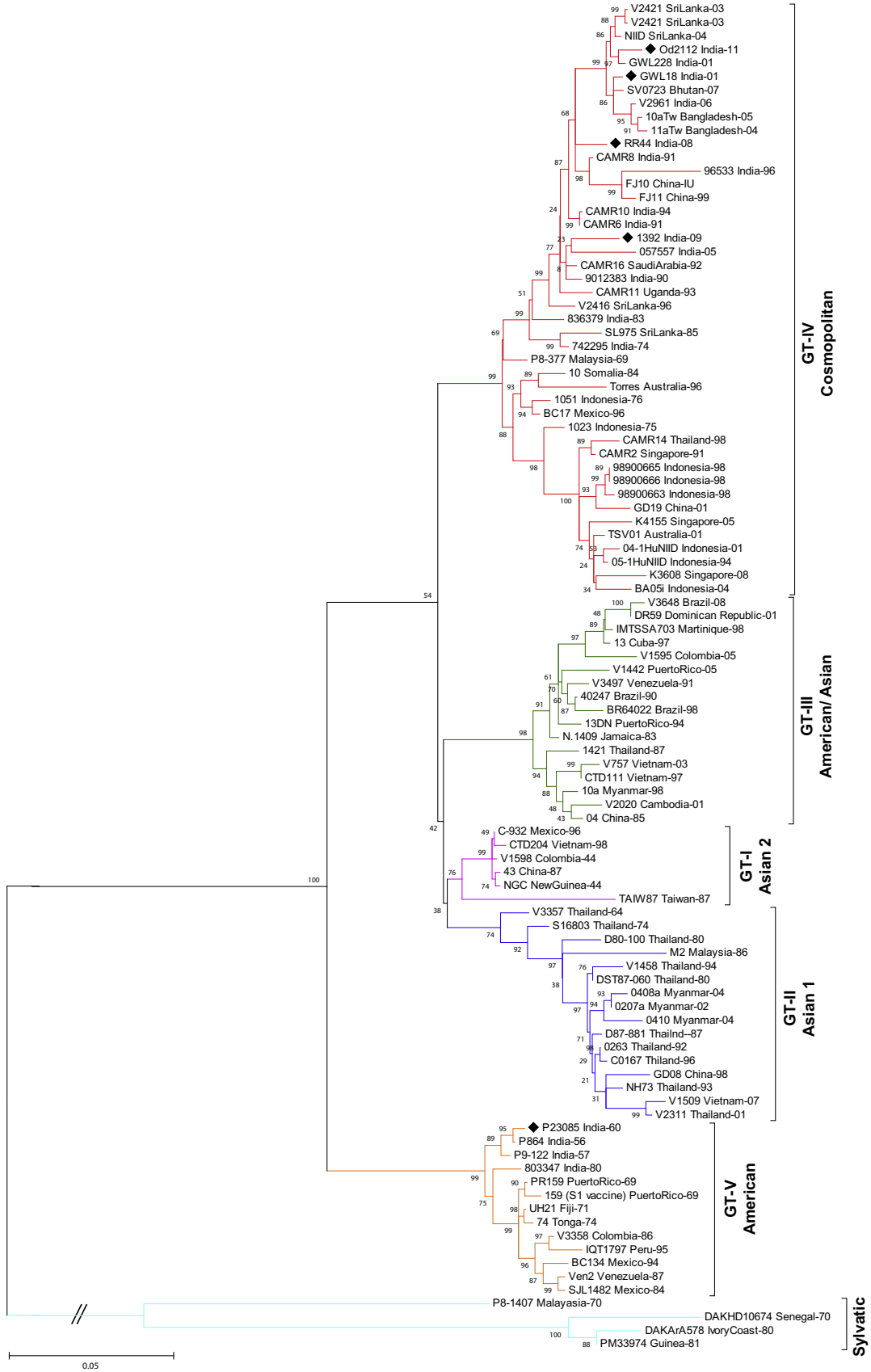

Fig. 2. Phylogenetic tree among Dengue-2 viruses generated by Maximum-Likelihood method based on the complete nucleotide sequence of envelope protein gene (1485 nucleotides). Each strain is abbreviated with isolate ID followed by country of origin and year of isolation. The DENV-2 sequenced in this study are marked with solid diamond (♦). Bootstrap values are indicated at the major branch points.

Table 3
Selection pressure analysis of the open reading frame (ORF) (3391 codons) datasets of all genotypes of DENV-2 (n = 146) using SLAC, FEL, MEME and FUBAR.

AA Position (ORF)	AA Position (Protein)	SLAC		FEL		MEME		FUBAR	
		ω	p-value	ω	p-value	ω	p-value	ω	Post. Pr.
Capsid									
82	82	0.497	0.689	0.161	0.424	>100	**0.032**	−0.377	0.17
102	102	2.982	**0.088**	0.569	0.106	>100	**0.065**	0.163	0.677
113	113	1.327	0.996	0.312	0.902	>100	**0.095**	−1.051	0.148
Envelope									
351	71	1.867	0.265	0.596	0.125	>100	**0.079**	0.116	0.62
400	120	1.491	0.297	0.27	0.261	>100	**0.052**	−0.147	0.353
408	128	0.948	0.522	0.335	0.274	>100	**0.091**	−0.171	0.343
559	279	0.475	0.698	0.08	0.413	>100	**0.091**	−0.257	0.209
639	358	1.48	0.303	0.269	0.263	>100	**0.059**	−0.151	0.349
NS1									
1010	235	0.507	0.663	0.166	0.395	>100	**0.033**	−0.334	0.198
NS2a									
1178	51	0.696	0.541	0.248	0.763	>100	**0.058**	−0.124	0.374
1267	140	0.332	0.999	0.075	0.972	>100	**0.037**	−0.863	0.015
1301	174	2.506	0.191	0.431	0.355	8.492	**0.089**	0.109	0.608
1342	215	2.224	0.252	0.915	**0.069**	2.047	0.444	0.295	0.785
NS2b									
1353	8	0.947	0.477	0.244	0.441	>100	**0.03**	−0.032	0.444
1439	94	1.955	0.271	0.147	0.782	>100	**0.007**	−0.093	0.38
NS3									
1489	14	3.53	**0.058**	1.264	**0.023**	>100	**0.005**	0.478	0.866
1555	80	0.508	0.662	0.166	0.395	>100	**0.069**	−0.334	0.198
1661	186	1.927	0.256	0.867	**0.077**	>100	**0.083**	0.233	0.72
2024	549	2.424	0.176	0.858	**0.076**	>100	**0.088**	0.236	0.725
2033	558	0.497	0.667	0.087	0.521	>100	**0.08**	−0.448	0.105
NS4a									
2157	64	0.497	0.667	0.102	0.342	>100	**0.072**	−0.161	0.302
NS4b									
2262	19	4.291	**0.004**	1.543	**0.011**	>100	**0.004**	1.049	**0.962**
2393	150	0.22	0.933	0.087	0.767	>100	**0.088**	−0.683	0.029
NS5									
2666	175	0.881	0.34	0.504	**0.062**	>100	0.114	0.186	0.717
2891	400	0.535	0.496	0.033	0.952	>100	**0.005**	−0.283	0.221
2892	401	2.437	**0.054**	1.52	**0.027**	>100	**0.043**	0.68	0.888
3087	596	0.92	0.331	0.513	0.166	>100	**0.066**	0.016	0.518
3129	638	1.561	0.131	0.523	0.133	>100	**0.024**	0.108	0.619
3135	644	2.683	0.133	1.624	0.205	3.085	0.15	1.438	**0.901**
3276	785	0.323	0.656	0.08	0.543	>100	**0.032**	−0.469	0.092
3356	865	1.537	0.132	0.481	0.143	>100	**0.031**	0.083	0.594
3367	876	1.177	0.237	0.668	**0.034**	>100	0.19	0.29	0.812

Note: Criteria to consider sites with significant evidence of positive selection: p-value <0.1 in SLAC, FEL and MEME, and posterior probability <0.9 in FUBAR.
Sites that are found positive by at least three methods are included in the list.
Sites with significant evidence of positive selection by at least two and three methods are written in bold and underlined font respectively.
The significant statistical values ω are shown in bold font.
AA: amino acid.
The amino acid positions are in respect to GenBank Accession No NC_001474.2.

less than 6% nucleotide divergence within a selected genomic region due to comparatively lower rate of variation (10^{-4} substitutions/nt/year) [22]. Further, the phylogeny based on the envelope gene, was reported to reveal distinct geographical classification of the dengue viruses [23,24]. Therefore, another phylogenetic analysis was carried out by including 64 additional representative global isolates. This included 13 Indian DENV-2 sequences, so as to cover the representative viruses from the entire duration of DENV-2 circulation in India. This dendrogram provides a comprehensive analysis of phylogeny of Indian isolates in a truly global perspective. This dendrogram revealed six genotypes, one of which included sylvatic strains. The recent Indian viruses belong to genotype IV, along with viruses from several countries of Asia, Australia, Africa and American continent. This genotype is also termed as Cosmopolitan genotype, due to its global spread [23]. Within genotype IV, the recent Indian viruses (sequenced in this study) were found in close cluster with other Indian viruses, isolated during 1974–2006. However, the P23085 (isolated in 1960) belong to genotype V (also known as American genotype), along with Indian viruses, isolated during 1956–1980 and a group of viruses from Americas. The evolutionary trend clearly indicates that the genotype V viruses circulating since mid 1950s were completely replaced by genotype IV viruses in early 1980s. Since then, genotype IV viruses were predominantly circulating and causing outbreaks in many parts of India [10,12]. A critical examination of branching pattern within genotype IV revealed the clustering of Indian viruses into a distinct clade along with isolates from South Asian countries, China and Saudi Arabia. A Malaysian sylvatic DENV-2 (P8-377), isolated from *Aedes albopictus* in 1969 was found to be the ancestor of this clade. Sylvatic viruses, as the ancestor of currently circulating endemic/epidemic DENV have already been well documented [3]. Keeping in mind the efficiency of Cosmopolitan genotype to circulate throughout the world, it seems quite possible that this genotype might have been introduced into India from South East Asia. However, the exact time period and the route of incursion of this virus into India could not be conclusively ascertained due to non-availability of relevant sequences of that period.

This is the first report on complete genome characterization of five dengue virus type 2, isolated from India. This study clearly re-

vealed that a unique clade of Cosmopolitan genotype of DENV-2 are circulating predominantly and causing major outbreaks in most parts of India. Identification of increasing number of amino acid substitutions in Indian DENV-2 over last decade clearly indicates the faster evolving nature of these viruses. The detailed and continuous molecular epidemiological surveillance is warranted to monitor the emergence of newer DENV, which will help to undertake effective control and management strategies at the earliest.

Acknowledgments

The authors are thankful to the Director, DRDE, Gwalior for his keen interest and financial support in this study. The authors are also thankful to the Superintendent, JA Hospital, Gwalior; R&R Hospital, Delhi; NIMS, Hyderabad; and NVBDCP, Odisha for providing the clinical samples. The authors are thankful to Director, NIV, Pune for providing P23085 isolate.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2013.05.130.

References

[1] D.J. Gubler, Dengue and dengue hemorrhagic fever: its history and resurgence as a global public health problem, in: D.J. Gubler, G. Kuno (Eds.), Dengue and Dengue Hemorrhagic Fever, CAB International, London, 1997, pp. 1–22.
[2] E.A. Henchal, J.R. Putnak, The dengue viruses, Clin. Microbiol. Rev. 3 (1990) 376–396.
[3] E. Wang, H. Ni, R. Xu, et al., Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses, J. Virol. 74 (2000) 3227–3234.
[4] S.C. Weaver, N. Vasilakis, Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease, Infect. Genet. Evol. 9 (2009) 523–540.
[5] T.J. Chambers, C.S. Hahn, R. Galler, et al., Flavivirus genome organization, expression and replication, Annu. Rev. Microbiol. 44 (1990) 649–688.
[6] S.B. Halstead, Pathogenesis of dengue: challenges to molecular biology, Science 239 (1988) 476–481.
[7] W.J.H. McBride, H. Bielefeldt-Ohmann, Dengue viral infections; pathogenesis and epidemiology, Microbes Infect. 2 (2000) 1041–1050.
[8] S. Bhatt, P.W. Gething, O.J. Brady, et al., The global distribution and burden of dengue, Nature (2013), http://dx.doi.org/10.1038/nature12060.
[9] S.R. Kumar, J.A. Patil, D. Cecilia, et al., Evolution, dispersal and replacement of American genotype dengue type 2 viruses in India (1956–2005): selection pressure and molecular clock analysesJ, Gen. Virol. 91 (2010) 707–720.
[10] L. Dar, S. Broor, S. Sengupta, et al., The first major outbreak of DHF in Delhi, India, Emerg. Infect. Dis. 5 (1999) 589–590.
[11] P.K. Dash, M.M. Parida, P. Saxena, et al., Emergence and continued circulation of Dengue-2 (Genotype IV) virus strains in northern India, J. Med. Virol. 74 (2004) 314–322.
[12] M. Anoop, A. Issac, T. Mathew, et al., Genetic characterization of dengue virus serotypes causing concurrent infection in an outbreak in Ernakulam, Kerala, South India, Ind. J. Exp. Biol. 48 (2010) 849–857.
[13] K.L. McElroy, G.A. Santiago, N.J. Lennon, et al., Endurance, refuge, and reemergence of Dengue virus type 2, Puerto Rico, 1986–2007, Emerg. Infect. Dis. 17 (2011) 64–71.
[14] S. Sharma, P.K. Dash, S. Agarwal, et al., Comparative complete genome analysis of dengue virus type 3 circulating in India between 2003 and 2008, J. Gen. Virol. 92 (2011) 1595–1600.
[15] M. Anoop, A.J. Mathew, B. Jayakumar, et al., Complete genome sequencing and evolutionary analysis of dengue virus serotype 1 isolates from an outbreak in Kerala, South India, Virus Genes 45 (2012) 1–13.
[16] K. Tamura, D. Peterson, N. Peterson, et al., MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Mol. Biol. Evol. 28 (2011) 2731–2739.
[17] W. Delport, A.F. Poon, S.D. Frost, Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology, Bioinformatics 26 (2010) (2010) 2455–2457.
[18] P. Masrinoul, M.O. Diata, S. Pambudi, et al., Highly conserved region 141–168 of the NS1 protein is a new common epitope region of dengue virus, Jpn. J. Infect. Dis. 64 (2011) 109–115.
[19] A. Sampath, T. Xu, A. Chao, et al., Structure-based mutational analysis of the NS3 helicase from dengue virus, J. Virol. 80 (2006) 6686–6690.
[20] G.S. Park, K.L. Morris, R.G. Hallett, et al., Identification of residues critical for the interferon antagonist function of Langat virus NS5 reveals a role for the RNA-dependent RNA polymerase domain, J. Virol. 81 (2007) 6936–6946.
[21] N. Vasilakis, E.C. Holmes, E.B. Fokam, et al., Evolutionary processes among sylvatic dengue type 2 viruses, J. Virol. 81 (2007) 9591–9595.
[22] R. Rico-Hesse, Molecular evolution and distribution of dengue viruses type 1 and 2 in nature, Virology 174 (1990) 479–493.
[23] S.S. Twiddy, J.J. Farrar, N.V. Chau, et al., Phylogenetic relationships and differential selection pressures among genotypes of dengue-2 virus, Virology 298 (2002) 63–72.
[24] C.V.F. Carrington, J.E. Foster, O.G. Pybus, et al., Invasion and maintenance of Dengue virus type 2 and type 4 in the Americas, J. Virol. 79 (2005) 14680–14687.