SHORT COMMUNICATION

Molecular epidemiology and evolutionary analysis of dengue virus type 2, circulating in Delhi, India

Pankaj Sharma¹ · Veena Mittal¹ · Mala Chhabra¹ · Roop Kumari² · Priyanka Singh³ · Srinivas Venkatesh^{1,2,3}

Received: 12 July 2016 / Accepted: 26 August 2016 / Published online: 14 September 2016
© Indian Virological Society 2016

Abstract Dengue virus type 2 (DENV-2) has been associated with severe dengue outbreaks in many countries including India. Its predominance was recorded nearly after a decade in the capital city, Delhi in 2013. The present study characterizes DENV-2 circulated during 2013–2014. Analysis based on envelope (E) gene showed the presence of two clades (I and II) of DENV-2, within the Cosmopolitan genotype. Analysis of time of most recent common ancestor revealed the existence of clade I for more than a decade (95 % HPD 13–16 years) however, clade II showed comparatively recent emergence (95 % HPD 5–13 years). Presence of different clades is of high significance as this may result in increased virus transmission and major outbreaks. Further, the presence of a unique amino acid substitution, Q325H was also observed in an isolate; 14/D2/Del/2013 (KT717981). This substitution falls in immune epitope (epitope id: 150268) and may have important role in host immune response.

Keywords Dengue · Molecular epidemiology · Envelope · Evolutionary analysis

Electronic supplementary material The online version of this article (doi:10.1007/s13337-016-0346-4) contains supplementary material, which is available to authorized users.

✉ Veena Mittal
veena_m12@yahoo.com

¹ Division of Zoonosis, National Centre for Disease Control, 22-Sham-Nath Marg, Delhi 110054, India

² Centre for Medical Entomology and Vector Management, National Centre for Disease Control, Delhi 110054, India

³ Division of Biotechnology, National Centre for Disease Control, Delhi 110054, India

Dengue fever is a mosquito borne infection caused by four antigenically-related serotypes of dengue virus (DENV-1, 2, 3 and 4). During the past decades dengue fever (DF) emerged as a major public health problem of global scale. Nearly 40 % of the world population is living at the risk of dengue infection. According to World Health Organization (available at http://www.searo.who.int/entity/vector_borne_tropical_diseases/data/data_factsheet/en/) it is the major cause of illness and deaths in tropical and subtropical regions, where *Aedes* mosquito population is high during rainy season. Dengue virus (DENV) belongs to the family *Flaviviridae* (genus; *Flavivirus*). Its genome consists of single stranded, positive sense RNA with approximately 11,000 nucleotides. A single open reading frame consists of three structural (capsid; C, premembrane/membrane; prM/M and envelope; E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) flanked by non-translated regions (5' UTR and 3' UTR) [1].

Dengue is endemic in several parts of India. Many outbreaks have been reported from different parts of the country. During last two decades, Delhi, the national capital has emerged as one of the “badly hit” city by dengue. The city witnessed its first major dengue hemorrhagic fever (DHF) outbreak in 1996 [2]. Dengue virus type 2 (DENV-2) was confirmed as main etiological agent for this outbreak. During another outbreak in 2003, co-circulation of DENV-2 and DENV-3 was observed [3]. However, in the following years a downward shift in predominance of DENV-2 was recorded. The outbreaks occurred in 2006 and 2010 were predominated by DENV-3 and DENV-1 respectively [4, 5]. Nearly after a decade, DENV-2 re-emerged as predominant serotype during the outbreak in 2013 [6]. The emergence of DENV-2 is an alarming signal for public health, as this serotype has been associated with major DF and DHF outbreaks [2, 6–8]. Based on genomic

variations DENV-2 has been divided into 6 distinct genotypes; Sylvatic, Asian (further classified into Asian I and Asian II), American, Asian/American and Cosmopolitan [9]. Genotypes don't only differ in their geographical distribution, but also in their virulence and potential to cause large epidemics. [8, 10] Various studies have also suggested the role of small genomic changes, leading to the lineages with greater fitness, epidemic potential and possibly the greater virulence of strains [11]. The present study was carried out to understand the molecular epidemiology and evolutionary dynamics of DENV-2, circulated in Delhi during the post-monsoon period of 2013 and 2014.

Acute phase serum samples, confirmed for the presence of DENV-2 serotype were included in the study. Samples were referred to National Centre for Disease Control for diagnosis of DENV from different geographical location of Delhi, during the post-monsoon period of 2013 and 2014. Presence of DENV was detected in the samples by NS1 antigen detection assay and the serotype of virus was confirmed through sequencing of capsid-premembrane (CprM) gene junction (data not shown here). The approval of the institutional ethical committee was obtained to carry out the present study.

Viral RNA, was extracted from 140 µl of each serum sample by using QIAmp Viral RNA Mini kit (Qiagen, Germany) in accordance to the manufacturer's protocol. For amplification of E gene forward primer; DEN2 FS2 (5'-GCTCTCCGAGAACAGGCC-3') and reverse primer; DEN2 R (5'-CCAGCTGCACAACGCAACCAC-3') were used. [12]. cDNAs were synthesized with 6 µl RNA and E gene specific reverse primer (DEN2 R), using GoScript™ Reverse Transcription System (Promega, USA) as per the manufacturer's protocol. For amplification of target gene GoTaq® Green Master Mix (Promega, USA) was applied, using 1 µl cDNA as a template. Initial denaturation was done at 95 °C for 2 min, followed by 35 cycles of denaturation (95 °C for 30 s), annealing (55 °C for 30 s) and extension (72 °C for 1 min). A final extension was provided at 72 °C for 10 min. Amplified PCR products (~930 bp) were visualized on 1.2 % agarose gel stained with ethidium bromide.

For sequencing PCR products were purified by QIAquick PCR purification kit (Qiagen, Germany). Sequencing reaction was carried out by BigDye Terminator Cycle Sequencing Kit version 3.1 (Applied Biosystems, USA) on an automatic DNA sequencer; ABI 3130xl Genetic Analyzer (Applied Biosystems, USA). Sequences thus obtained, were submitted to gene bank (www.ncbi.nlm.nih.gov) and accession numbers were acquired (Supplementary Table). Prior to this a BLAST search (www.ncbi.nlm.nih.gov/blast/BLASTinfo/information3.html) was carried out to confirm the identity of strains. For sequence comparison, CLUSTAL W based Multiple Sequence

Alignment (MSA) was carried out by using BioEdit v 7.0.5.3. Nucleotide diversity (π) of the aligned sequences was calculated by using DNA SP v 5.10. Sliding window approach (window size, 300 nt; step size, 25 nt) was applied to identify the region of highest nucleotide diversity. To estimate the nucleotide substitution rate and time of most recent common ancestor (tMRCA), BEAST v 1.7.4 was applied. Strict, relaxed uncorrelated lognormal and relaxed uncorrelated exponential clocks with different demographic models (constant size and exponential growth) were employed for this purpose. Six independent MCMC (Markov Chain Monte Carlo) analyses, each for 15,000,000 steps, were performed for each combination of branch rate and demographic model. Convergence of parameters was analyzed on the basis of effective sample size (ESS >200) after a 10 % burn-in using Tracer v 1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>). For model selection, analysis of Bayes factor was done on the basis of relative marginal likelihoods [13]. Maximum clade credibility (MCC) tree was generated by using Tree Annotator (available in BEAST). FigTree v1.4.2. (<http://tree.bio.ed.ac.uk/software/figtree/>) was used to visualize the annotated tree. Ninety-five percent (95 %) highest posterior density (HPD) intervals were obtained to determine degree of uncertainty in the parameter estimates.

A total of 23 samples, were amplified and sequenced for E gene. Among the samples, mean age 26.11 years (SD $\pm$ 13.02) and male to female ratio of 1.3:1 was observed. The common clinical symptoms; fever, headache, body-ache, joint-pain and nausea were observed in all the patients. The period of the acute phase of illness ranged from 1 to 5 days (mean; 2.60, SD: $\pm$ 1.23). To understand the evolutionary dynamics of circulating virus, sequences were aligned with the prototype strain of DENV-2; NGC44 (AF038403) and the other geographically diverse isolates (Supplementary Table). Nucleotide alignment (nt 1531–2379) revealed no base insertion or deletion. However, substitutions for many bases were observed. Majority of substitutions were synonymous in nature except a few, which were nonsynonymous and resulted into change in amino acid. Nucleotide diversity (π) of 0.047 was observed for the aligned region. Sliding window method showed highest nucleotide diversity (π = 0.057) for the region from nt 1831–2130. In MCC tree all isolates sequenced in the study, grouped into two clades (I and II) under the Cosmopolitan genotype of DENV-2 (Fig. 1). Isolates of clade I were clustered with Indian isolates, OD2112 (JQ955624) and GWL228 (DQ448237) in close proximity to isolates from Sri Lanka (GQ252677 and AB194883) and Pakistan (KF041237 and JX042511). Isolates of clade I (KT717968-85 and KT717989-90) revealed the mean sequence identity of 98.94 % (98.46–99.41 %) and 99.41 % (98.93–99.88 %) to the isolates OD2112 and

Table 1 Parameter estimates for different clock models

Clock- (demographic model)	Posterior (L–U) ^a	Marginal likelihood (±SE)	Mean rate × 10 ^{−4} (L–U)
Strict- (constant size)	−4797.42 (−4817.06, −4780.40)	−4409.70 (±0.25)	5.81 (4.83, 6.83)
Strict- (exponential growth)	−4793.08 (−4811.07, −4774.32)	−4409.74 (±0.22)	5.87 (4.89, 6.87)
Uncorrelated lognormal relaxed- (constant size)	−4735.61 (−4761.41, −4710.28)	−4366.82 (±0.28)	6.38 (5.04, 7.78)
Uncorrelated lognormal relaxed- (exponential growth)	−4729.84 (−4757.62, −4704.35)	−4366.66 (±0.27)	6.45 (5.07, 7.85)
Uncorrelated exponential relaxed-(constant size)	−4715.67 (−4742.15, −4691.60)	−4354.68 (±0.26)	6.70 (4.92, 8.40)
Uncorrelated exponential relaxed-(exponential growth)	−4709.28 (−4735.51, −4684.01)	−4355.28 (±0.28)	6.68 (4.88, 8.49)

^a Highest probability density (95 % HPD)—L (lower) and (U) upper bound

The parameter estimates for best fit model are indicated in bold

GWL228 respectively. Three isolates; 3/D2/Del/2014 (KT717986), 4/D2/Del/2014 (KT717987) and 5/D2/Del/2014 (KT717988) clustered with Indian isolate, 1392 (JX475906) under clade II and exhibited mean sequence identity of 99.48 % (99.17–99.64 %) with the same. Mean divergence between clade I and II was nearly 3.08 %. Sequence comparison of all Delhi isolates (KT717968–90) with Seychelles isolate (L10048) revealed nearly 3.86 % mean divergence. In molecular clock analysis, among all the models implemented, highest marginal likelihood was obtained for relaxed uncorrelated exponential clock (constant population size) (Table 1). The model was favored over the other models in Bayes factor analysis. Under this model, estimates of nucleotide substitution rates were found 6.70×10^{-4} substitutions/site/year (95 % HPD 4.92×10^{-4} – 8.40×10^{-4}). Estimate of tMRCA for all the genotypes (node A) was about 131 years (95 % HPD 87–215) with respect to the recent isolate of 2014, which correspond to the year 1883 (95 % HPD, years 1799–1927). Isolates of Asian-I, Asian, II and American/Asian genotypes shared common ancestors (node B) nearly 87 years (95 % HPD 73–114 years) ago. Common ancestor of all the isolates of Cosmopolitan genotype (node C) was about 61 years (95 % HPD 49–83 years) ago. Isolates of clade I (node D) shared common ancestor nearly 14 years ago (95 % HPD 13–16 years) corresponding to the year 2000 (95 % HPD, years 1998–2001). However, the common ancestor of clade II (node E) was observed nearly 8 years ago (95 % HPD 5–13 years) in 2006 (95 % HPD, years 2001–2009). Isolates of American genotype showed common ancestors (node F) 69 years ago (95 % HPD 60–87 years) corresponding to year 1945 (95 % HPD, years 1927–1954).

Deduced amino acid sequence revealed that most amino acid changes were shared by all isolates of clade I and II except the I322 V substitution which was present in isolates of clade I only. The amino acid substitutions, V252A, Q325H and E360G were also observed in three isolates of clade I; 2/D2/Del/2013 (KT717969), 14/D2/Del/2013

(KT717981) and 16/D2/Del/2013 (KT717983), respectively.

Among all the 4 serotypes DENV-2 is genetically most diverse with six distinct genotypes. During the outbreak of DHF in 1996 in Delhi, predominance of Cosmopolitan genotype (genotype IV) was reported. Vigilance over the years recorded its continued circulation in the following years [2, 8]. However, with emergence of DENV-3 in 2003, a downward shift in predominance of DENV-2 occurred, and its spread narrowed to sporadic cases [3–5, 14]. After being persistent at low key for nearly a decade, it emerged as predominant serotype in 2013 outbreak [6]. This study characterizes circulating DENV-2 on the basis of E gene region. The E protein, a major antigenic determinant, consists of three domains (I, II and III) and plays important role in virulence, receptor binding and fusion during the entry of virus particle [15]. Partial nucleotide sequence of E gene is widely used to infer the phylogeny of DENV [2, 6].

MCC tree analysis carried out for E gene region (nt 1531–2379) revealed that most isolates clustered within the same group (clade I) and showed mean sequence identity of 99.41 % to the isolate GWL228 (DQ4448237), reported from northern India in 2001. Three isolates formed a separate group (clade II) with Indian isolate 1392 (JX475906) in close proximity to Sri Lankan isolate V2416 (FJ882602). Presence of different clades of DENV-2 indicates the genomic variations occurring in Cosmopolitan genotype. This is further supported by increased mean divergence (3.86 %) of currently circulating Delhi isolates (KT717968–90), to the Seychelles strain (L10048). As in DENV-2 isolates from 1996 outbreak, 3.5 % mean divergence to this strain (L10048) was recorded [2].

Deduced amino acid sequence revealed no specific amino acid difference between clade I and II except the I322V substitution. This substitution was absent in clade II isolates. In a previous study by Afreen et al. [16] the association of this substitution with a unique lineage of DENV-2 has been identified. Other substitutions; V252A

Fig. 1 Maximum clade credibility tree (MCC) of E gene (nt 1531–2379) of DENV-2 with the constant size demographic model under the relaxed uncorrelated exponential clock. The tMRCA

estimates of the key nodes are marked (filled circle) and labeled (A–F). Isolates sequenced in this study are marked (filled triangle) and shown in **bold**

and E360G, recorded in our study have been observed in isolates reported earlier. However, a unique conservative substitution Q325H (replacement of glutamine by histidine) was also observed in an isolate; 14/D2/Del/2013. Amino acid histidine has property to undergo protonation and get charged upon exposure to acidic environment. The protonation of histidine residues trigger structural changes at low pH which plays crucial role in membrane fusion during the entry of virus particle [17]. Various histidine residues (His244, His261, and His317) in E protein of DENV-2 have been reported to undergo conformational changes during conversion to post-fusion structure [18]. The presence of histidine residues (His325) may have

important role in conformational rearrangements of E protein. Since this substitution also falls in immune epitope (epitope id; 150268) and has not been reported in any other previously reported isolate it is highly imperative to understand its role in host immune response, virus entry and membrane fusion.

The estimate of nucleotide substitution rate for all the genotypes was 6.70×10^{-4} substitutions/site/year, which was comparable to the earlier estimates [12, 19]. The emergence of DENV-2 genotype was observed nearly 131 years ago (95 % HPD 87–215 years) which was also in agreement with the earlier estimates [19]. Clustering of majority of sequences under clade I confirmed

predominance of this clade. Mean sequence identity of 99.41 % (98.93–99.88 %) to GWL 228 (DQ448237) isolate, indicated the existence of this clade for more than a decade, as also suggested by tMRCA which was nearly 14 years ago (95 % HPD 13–16 years). However clade II isolates (node E) showed comparatively recent emergence as indicated by tMRCA nearly 8 years ago (95 % HPD 5–13 years) corresponding to year 2006 (95 % HPD, years 2001–2009). Emergence of this clade may be a public health concern as it may become the predominant clade in coming years. However further studies to understand the distribution, predominance and fitness of clade will be necessary to derive such conclusion.

In our previous study, presence of different clades of DENV-2 on the basis of CprM gene region has been discussed [20]. In another study by Afreen et al. [16], emergence of different lineages of DENV-2 in Delhi has also been described on the basis of 457 bp (nt 1801–2257) of E gene region. Even in the present study which was carried out on 849 bp (nt 1531–2379) of E gene region, similar findings were confirmed. The study also revealed region (nt 1831–2130) of high nucleotide diversity. This region corresponds to domain III of E protein. The substitutions; I322V, Q325H and E360G also fall in the same region. The region may be important to understand the evolutionary dynamics of DENV.

Continuous existence of cosmopolitan genotype of DENV-2 in the capital, speculate its fitness over the other genotypes. Accumulation of genomic changes over the time period may have been providing it better selective advantage.

Acknowledgments The authors sincerely acknowledge the partial financial support from the Indian Council of Medical Research (ICMR) during the course of this study.

References

- Chambers TJ, Hahn CS, Galler R, Rice CM. Flavivirus genome organization, expression, and replication. *Annu Rev Microbiol.* 1990;44:649–88. doi:[10.1146/annurev.mi.44.100190.003245](https://doi.org/10.1146/annurev.mi.44.100190.003245).
- Singh UB, Maitra A, Broor S, Rai A, Pasha ST, Seth P. Partial nucleotide sequencing and molecular evolution of epidemic causing Dengue 2 strains. *J Infect Dis.* 1999;180(4):959–65. doi:[10.1086/315043](https://doi.org/10.1086/315043).
- Kumar M, Pasha S, Mittal V, Rawat D, Arya SC, Agarwal N, et al. Unusual emergence of Guate98-like molecular subtype of DEN-3 during 2003 dengue outbreak in Delhi. *Dengue Bull.* 2004;28:161–7.
- Singh P, Mittal V, Rizvi MM, Chhabra M, Sharma P, Rawat DS, et al. The first dominant co-circulation of both dengue and chikungunya viruses during the post-monsoon period of 2010 in Delhi, India. *Epidemiol Infect.* 2012;140(7):1337–42. doi:[10.1017/S0950268811001671](https://doi.org/10.1017/S0950268811001671).
- Kukreti H, Chaudhary A, Rautela RS, Anand R, Mittal V, Chhabra M, et al. Emergence of an independent lineage of dengue virus type 1 (DENV-1) and its co-circulation with predominant DENV-3 during the 2006 dengue fever outbreak in Delhi. *Int J Infect Dis.* 2008;12(5):542–9. doi:[10.1016/j.ijid.2008.02.009](https://doi.org/10.1016/j.ijid.2008.02.009).
- Afreen N, Deeba F, Naqvi I, Shareef M, Ahmed A, Broor S, et al. Molecular investigation of 2013 dengue fever outbreak from Delhi, India. *PLoS Curr.* 2013;2013:6.
- Anoop M, Issac A, Mathew T, Philip S, Kareem NA, Unnikrishnan R, et al. Genetic characterization of dengue virus serotypes causing concurrent infection in an outbreak in Ernakulam, Kerala, South India. *Indian J Exp Biol.* 2010;48(8):849–57.
- Dash PK, Parida MM, Saxena P, Kumar M, Rai A, Pasha ST, et al. Emergence and continued circulation of dengue-2 (genotype IV) virus strains in northern India. *J Med Virol.* 2004;74(2):314–22. doi:[10.1002/jmv.20166](https://doi.org/10.1002/jmv.20166).
- Dash PK, Sharma S, Soni M, Agarwal A, Parida M, Rao PV. Complete genome sequencing and evolutionary analysis of Indian isolates of Dengue virus type 2. *Biochem Biophys Res Commun.* 2013;436(3):478–85. doi:[10.1016/j.bbrc.2013.05.130](https://doi.org/10.1016/j.bbrc.2013.05.130).
- Rico-Hesse R, Harrison LM, Salas RA, Tovar D, Nisalak A, Ramos C, et al. Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. *Virology.* 1997;230(2):244–51.
- Khan MA, Ellis EM, Tissera HA, Alvi MY, Rahman FF, Masud F, et al. Emergence and diversification of dengue 2 cosmopolitan genotype in Pakistan, 2011. *PLoS ONE.* 2013;8(3):e56391. doi:[10.1371/journal.pone.0056391](https://doi.org/10.1371/journal.pone.0056391).
- Kumar SR, Patil JA, Cecilia D, Cherian SS, Barde PV, Walimbe AM, et al. Evolution, dispersal and replacement of American genotype dengue type 2 viruses in India (1956–2005): selection pressure and molecular clock analyses. *J Gen Virol.* 2010;91(Pt 3):707–20. doi:[10.1099/vir.0.017954-0](https://doi.org/10.1099/vir.0.017954-0).
- Suchard MA, Weiss RE, Sinsheimer JS. Bayesian selection of continuous-time Markov chain evolutionary models. *Mol Biol Evol.* 2001;18(6):1001–13.
- Chakravarti A, Chauhan MS, Kumar S, Ashraf A. Genotypic characterization of dengue virus strains circulating during 2007–2009 in New Delhi. *Arch Virol.* 2013;158(3):571–81. doi:[10.1007/s00705-012-1522-5](https://doi.org/10.1007/s00705-012-1522-5).
- Stiasny K, Heinz FX. Flavivirus membrane fusion. *J Gen Virol.* 2006;87(Pt 10):2755–66. doi:[10.1099/vir.0.82210-0](https://doi.org/10.1099/vir.0.82210-0).
- Afreen N, Naqvi IH, Broor S, Ahmed A, Kazim SN, Dohare R, et al. Evolutionary analysis of dengue serotype 2 viruses using phylogenetic and Bayesian methods from New Delhi, India. *PLoS Negl Trop Dis.* 2016;10(3):e0004511. doi:[10.1371/journal.pntd.0004511](https://doi.org/10.1371/journal.pntd.0004511).
- Nordlund HR, Hytonen VP, Laitinen OH, Uotila ST, Niskanen EA, Savolainen J, et al. Introduction of histidine residues into avidin subunit interfaces allows pH-dependent regulation of quaternary structure and biotin binding. *FEBS Lett.* 2003;555(3):449–54.
- Kampmann T, Mueller DS, Mark AE, Young PR, Kobe B. The role of histidine residues in low-pH-mediated viral membrane fusion. *Structure.* 2006;14(10):1481–7. doi:[10.1016/j.str.2006.07.011](https://doi.org/10.1016/j.str.2006.07.011).
- Twiddy SS, Farrar JJ, Vinh Chau N, Wills B, Gould EA, Gritsun T, et al. Phylogenetic relationships and differential selection pressures among genotypes of dengue-2 virus. *Virology.* 2002;298(1):63–72.
- Sharma P, Mittal V, Chhabra M, Singh P, Bhattacharya D, Venkatesh S, et al. Continued circulation of DENV-2 (genotype IV) in Delhi, India. *Br Microbiol Res J.* 2016;11(3):1–8.