Infection, Genetics and Evolution 12 (2012) 309–314

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

Comparative evolutionary epidemiology of dengue virus serotypes

Raquel L. Costa [a,1], Carolina M. Voloch [b,1], Carlos G. Schrago [b,*]

[a] Laboratório Nacional de Computação Científica, Petrópolis, Brazil
[b] Departamento de Genética, Universidade Federal do Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 18 August 2011
Received in revised form 29 November 2011
Accepted 12 December 2011
Available online 27 December 2011

Keywords:
Dengue
Bayesian skyline plot
Timescale
Positive selection

ABSTRACT

Evolutionary studies on dengue virus have frequently focused on intra-serotype diversity or on specific epidemics. In this study, we compiled a comprehensive data set of the envelope gene of dengue virus serotypes and conducted an extensive comparative study of evolutionary molecular epidemiology. We found that substitution rates are homogeneous among dengue serotypes, although their population dynamics have differed over the past few years as inferred by Bayesian coalescent methods. On a global scale, DENV-2 is the serotype with the highest effective population size. The genealogies also showed geographical structure within the serotypes. Finally, we also explored the causes of dengue virus serotype diversification by investigating the plausibility that it was driven by adaptive changes. Our results suggest that the envelope gene is under significant purifying selection and the hypothesis that dengue virus serotype diversification was the result of stochastic events cannot be ruled out.

1. Introduction

Dengue is an arthropod-borne disease that affects 50–100 million people per year in tropical and subtropical regions, resulting in death rates between 0.03% and 1.4% (WHO, 2005). Dengue virus (DENV) is a positive-sense single-stranded RNA (ssRNA+) *Flavivirus* transmitted by mosquitos of the genus *Aedes*. Dengue virus exists as four antigenically different serotypes (DENV-1–4), with uncertain evolutionary origins (Holmes and Burch, 2000). Within the serotypes, several authors have also characterized genotypes, based on genetic diversity and geographical distribution (Araujo et al., 2009; Klungthong et al., 2008; Kukreti et al., 2009; Rico-Hesse, 1990; Vasilakis et al., 2007).

Evolutionary studies of DENV are abundant. However, works have focused on the inference of substitution rates and the intra-serotype phylogeny (Goncalvez et al., 2002; Lanciotti et al., 1997, 1994; Twiddy et al., 2002a). It has been shown that evolutionary lineages within serotypes generally present characteristic geographical structure that reflects the spatial dynamics of epidemics (Holmes, 2008; Zhang et al., 2005). Dengue virus substitution rates, mainly measured from the envelope protein gene (E), have been estimated using several methodological approaches and are comparable to the evolutionary rates of other RNA viruses (Twiddy et al., 2002b; Wang et al., 2000).

Epidemiological surveys of dengue were conducted throughout the last century and, until recently, these were the only data available for obtaining historical information about the spread of the dengue virus (Gubler, 1998). Such surveys are essential for designing effective public health policies. However, the genetic history of the virus is as important as traditional epidemiological surveillance for understanding the full picture of epidemics (Bennett et al., 2010). This is because genetic material harbors information that can be used to reconstruct the history of epidemics in time and space. Moreover, recent coalescence techniques have also permitted the inference of population parameters such as growth rate and the effective number of infections (Drummond et al., 2003).

Several works have applied these methods to tackle DENV evolution (Bennett et al., 2010; Mondini et al., 2009; Villabona-Arenas and Zanotto, 2011). However, these studies focused on specific serotypes and an extensive comparative analysis is still lacking. A comparative analysis of dengue serotypes is important because there is increasing evidence that their epidemic potentials are different and that episodes of secondary infection with different serotypes frequently result in higher morbidity rates caused by the severe forms of the infection (Guzman et al., 2007).

Moreover, the causes of dengue serotype diversification are still unclear (Twiddy et al., 2002a). It has been suggested that the four serotypes originated from independent zoonotic passage to humans (Vasilakis et al., 2011). If this was so, we would expect that the DENV lineages have adapted to the new host and, consequently, there would be signatures of positive selection along the DENV genome. However, studies involving the estimation of d_N/d_S ratios on the envelope gene revealed extensive purifying selection (Holmes, 2003; Klungthong et al., 2004; Zhang et al., 2006).

* Corresponding author. Address: Departamento de Genética, A2-092, Inst. Biologia, CCS, Universidade Federal do Rio de Janeiro, Ilha do Fundão, Rio de Janeiro, RJ 21941-617, Brazil. Tel.: +55 21 4063 8278; fax: +55 21 2562 6397.
E-mail address: guerra@biologia.ufrj.br (C.G. Schrago).
[1] These authors contributed equally to this work.

1567-1348
doi:10.1016/j.meegid.2011.12.011

The present study aimed to perform a comprehensive comparative analysis of the history of dengue virus serotypes as inferred from a large number of publicly available sequences from the last few decades. By using coalescent-based approaches in a Bayesian framework, we assessed the demographic dynamics of epidemics throughout the last few centuries and estimated evolutionary parameters relevant for understanding the evolutionary aspects of DENV serotype emergence. We have also investigated the occurrence of adaptive molecular evolution on the emergence of DENV serotypes as well as the heterogeneity of selective forces along the envelope gene.

2. Materials and methods

2.1. Data set

We have downloaded all envelope gene sequences, which codes for the envelope protein, available in GenBank that contained information on the collection date, geographical location and that comprised the whole gene sequence. Then, a data set was composed of 596 sequences sampled to obtain the maximum geographical and chronological diversity (Table 1). This was achieved by excluding sequences collected in the same year from the same geographical location. Recombining sequences were excluded from the data set after inspection in the RDP3 program (Martin et al., 2010), using the RDP method (Martin and Rybicki, 2000). This sampling was performed to permit the implementation of sophisticated evolutionary analyses, which are unfeasible to be run on large amount of sequences. The data set covers a temporal range of 64 years (1944–2008) from 71 countries and consists of the largest DENV data set compiled to date. Sequences were aligned in ClustalW and inspected in MEGA 4.1. The final alignment included 1485 nucleotides. A list of access numbers is provided as supplementary information. The alignment used is available from the authors upon request. Analyses were conducted on each serotype independently and on the full data set.

2.2. Evolutionary rates and genealogical analysis

The comparative analysis of the DENV serotypes consisted of inferring the genealogy of the serotypes and the historical dynamics of the epidemics. We also investigated the action of differential selective forces along the envelope gene. Nucleotide substitution model choice was performed in HyPhy (Pond et al., 2005) and the GTR + G was selected since it significantly increased the likelihood of the data. All coalescent-based analyses were conducted in BEAST 1.6.1 (Drummond and Rambaut, 2007), which implements a Bayesian inference of evolutionary parameters via the Markov chain Monte Carlo (MCMC) method. The chronology of the dengue virus diversification was estimated using the relaxed molecular clock, in which the priors for the evolution of rates among branches were assumed to follow an uncorrelated lognormal distribution. The Bayesian skyline was the tree topology coalescent prior applied, since there is no indication whether the effective population size of DENV serotypes follow simple demographic models (i.e., constant, exponential or logistic growth). During MCMC, the parameters were visited for 50,000,000 generations and sampled every thousand cycles, resulting in 50,000 topologies and associated parameters. Of these, 10% were discarded as burn-in. Convergence of the MCMC run was checked in TRACER by inspecting the trace plots and calculating the effective sample sizes. To reconstruct the dengue virus demographic history, we employed the Bayesian skyline plot (BSP) model (Drummond et al., 2003, 2005), which generates piecewise constant population size trajectories. In order to test the robustness of the estimates of effective population sizes of serotypes to the number of sequences analyzed, we have composed an additional data set in which DENVs-1, 2 and 3 samples were reduced to 90 randomly chosen sequences. Note that this is also the number of sequences present in the smallest data set (DENV-4).

Table 1
Characteristics of DENV data set used in this study.

Serotype	Number of sequences	Number of countries	Time intervals
DENV-1	139	39	1944–2008
DENV-2	206	47	1944–2008
DENV-3	161	37	1956–2008
DENV-4	90	28	1956–2008

2.3. Differential selection along codon sites

Analysis of differential selective pressures along codon sites of the envelope gene was conducted in the CODEML program of the PAML 4.4 package (Yang, 2007) by calculating codon-specific d_N/d_S values (the ω parameter). Alignments of each DENV serotype were tested for positive selection independently, and jointly, using the M1a–M2a, the M7–M8 and the M8–M8a models (Wong et al., 2004). A qualitative measure of the distribution of ω values along codon sites in DENV serotypes was obtained by plotting the beta distribution with the parameters inferred by the M7 model (Yang et al., 2000). A comparison of the strength of selection along the envelope gene among the serotypes was conducted using the approach of Choisy et al. (2004), who used a paired Wilcoxon rank sum test to verify whether or not the weighted ω for each codon, estimated under the discrete M3 model (Yang et al., 2000), significantly differed between serotypes. Finally, the SLAC and the FEL algorithms (Pond and Frost, 2005) of the HyPhy package (Pond et al., 2005) were also used to access codons under positive selection in DENV serotypes independently and on the joint data set.

2.4. Differential selection on internal branches of DENV phylogeny

In order to investigate the action of adaptive molecular evolution on the diversification of dengue serotypes, we allowed internal branches of the phylogeny containing all DENV serotypes (Fig. 1) to evolve under different d_N/d_S ratios using the approach described in Yang (1998), implemented in the PAML 4.4 package. We have also run the IFEL test of HyPhy (Pond et al., 2006) and the newly proposed branch-site test of episodic diversifying selection (Kosakovsky Pond et al., 2011), available at www.datamonkey.org, which permits the inference of lineage-specific events of positive selection. These analyses were conducted to verify whether the hypothesis of diversification lead by genetic drift (founder effects) could be rejected.

2.5. Analysis of geographic correlation and ancestral states

In order to understand the spatial dynamics of DENV evolution comparatively, we have verified the association between phylogeny and geography for each serotype independently. This test was conducted in Befi-BaTS, v0.1.1 (Parker et al., 2008), which implements several tests of phylogeny–trait correlation considering phylogenetic error. We have inferred the parsimony score (PS), the association index (AI) and the monophyletic clade size statistics to test phylogeny–geography correlation. As geography traits, we have considered the continent where the samples were obtained. The reconstruction of the ancestral geographical states was implemented in Mesquite 2.7 (Maddison and Maddison, 2009) using the parsimony algorithm.

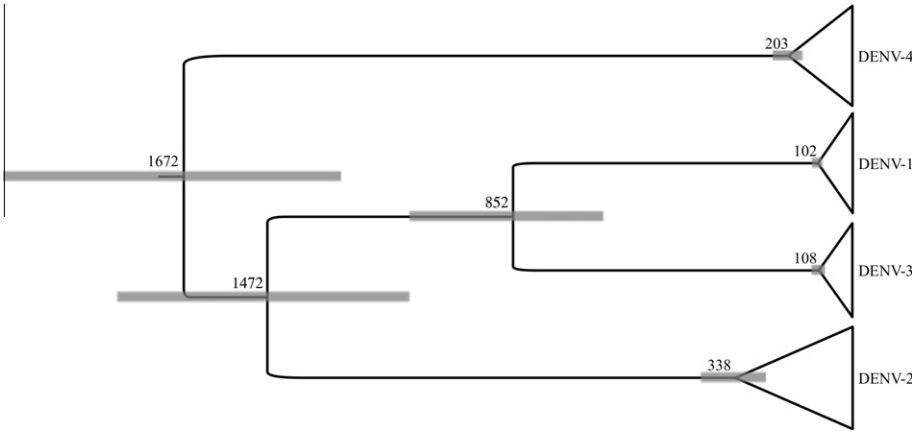

Fig. 1. DENV virus phylogeny, picturing the evolutionary affinities of the serotypes. Numbers on nodes indicate the TMRCA of the clades (years before 2008). Bars depict the 95% HPD intervals of the time estimates. The posterior probabilities of all clades are 100%.

3. Results and discussion

3.1. History and geography of dengue virus

The four serotypes of DENV were clearly arranged in well-defined clades with maximum statistical support (Fig. 1). The time to the most recent common ancestor (TMRCA) of all of the dengue serotypes, and also the time of the split of the DENV-4 lineage, was estimated at 1672 years ago, i.e., in the fourth century (Table 2). The 95% highest probability density (HPD) interval, however, was wide ranging from 2294 to 1158 years ago. It is interesting to note that the earliest records of an illness clinically similar do dengue fever are from the Chin dynasty of China, between the third and fifth centuries (Gubler, 1998), which is close to our mean estimate for the TMRCA of all serotypes. The serotype with the oldest coalescence time was found to be DENV-2, at ca. 340 years ago. Again, our estimates are in agreement with the historical records of the first outbreaks of illnesses that were clinically similar to dengue in the Americas (French West Indies and Panama) at the end of the seventeenth century (Gubler, 1998). Such estimates indicate that the events that gave rise to the serotypes had already occurred before the year 1650, as we inferred.

DENV genealogy revealed significant geographical structure (Fig. 2, Supplementary material). For all serotypes, the PS and AI tests were significant (p-value <0.01). However, the monophyletic clade size statistics was different for each serotype. For DENV-1 and DENV-2, the clade statistics recovered significant trait association for America and Asia. DENV-4 is also structured for all sampled continents (Asia and Americas) except for Oceania, however this serotype did not included any samples from Africa, while DENV-3 diversity is structured for all sampled continents except for Africa. The absence correspondent geographical structuring among serotypes implies differential migration during the evolutionary history of DENV. The geographical association revealed in this study is not associated with chronological sampling bias, i.e., sequences came from a particular outbreak, because they were collected at different time intervals. This indicates that, at a given time point, the probability is high that the ancestral lineage of the sample existed in the same geographic region. However, the global geographical structure within serotypes must be thoroughly tested by increasing the number of samples from Africa and Oceania, continents that are poorly available in public databases. Based on the data we have collected, the dynamics of dengue epidemics is fundamentally restricted by geographical distance at the continental level (Twiddy et al., 2002a).

Except for DENV-3, the ancestral state reconstruction points parsimoniously to an Asian origin for all serotypes. In DENV-3, an American origin is equally parsimonious. Early historical records of dengue epidemics in the Americas, supposedly a result of the slave trade, and the incidence of DENV in African monkeys would lead to an African origin. However, an Asian origin is consistent with the geographical occurrence of the sylvatic cycle of DENVs 1–4, while only sylvatic DENV-2 was reported in Africa (Vasilakis et al., 2011).

3.2. Rates of molecular evolution

The global rate of molecular evolution of the envelope gene of dengue virus was inferred as 7.6×10^{-4} substitutions/site/year (s/s/y) (Table 2). Substitution rates were very homogeneous among the serotypes: DENV-1 and DENV-2 presented an average of 7.5×10^{-4} s/s/y, while for DENV-3 and DENV-4 the values were estimated as 8.2×10^{-4} and 7.8×10^{-4} s/s/y, respectively. The 95% HPD interval of the estimates did not allow the rejection of the hypothesis that substitution rates on the envelope gene are the same among the serotypes. Within the serotypes, homogeneity of evolutionary rates along branches of the tree, which characterizes the molecular clock, could not be observed (Table 2). The 95% HPD interval of the estimate of the coefficient of variation of the serotypes did not contain 0, thus, the strict clock was not appropriate for these data. Although our estimates are in agreement with previous studies (Carrington et al., 2005; Jenkins

Table 2
Time to the most recent common ancestor and evolutionary parameters of DENV.

Serotype	TMRCA (years ago)	Substitution rate ($\times 10^{-4}$ s/s/y)	Coefficient of variation
DENV-1	102.9 (88.6–119.2)	7.5 (6.6–8.4)	0.3 (0.2–0.5)
DENV-2	338.1 (227.1–462.9)	7.5 (6.0–8.6)	0.4 (0.3–0.5)
DENV-3	108.9 (87.5–132.6)	8.2 (7.2–9.1)	0.4 (0.2–0.5)
DENV-4	203.4 (116.4–295.7)	7.8 (6.3–9.5)	0.4 (0.3–0.6)
Joint	1672.1 (1158.4–2294.7)	7.6 (6.6–8.7)	0.4 (0.3–0.4)

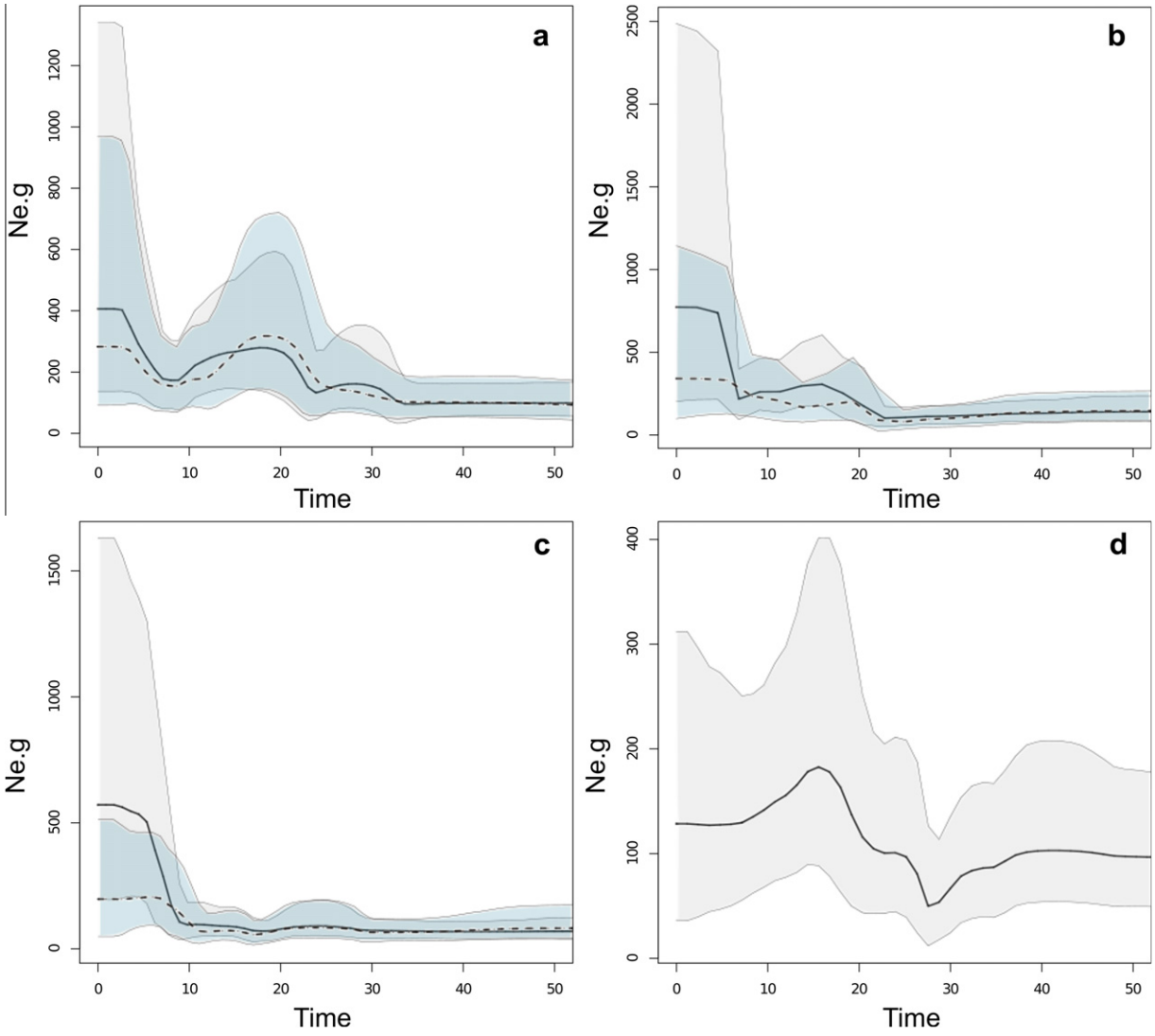

Fig. 2. Bayesian skyline plot of the past population dynamics of the four serotypes of dengue virus. Solid black line: full data set, dashed black line: reduced data set. Grey areas correspond to the 95% HPD intervals of the full data, blue areas correspond to the 95% HPD intervals of the reduced data. The effective population size (*y*-axis) is measured by the product of the effective population size (N_e) and virus generation time (g).

et al., 2002; Lanciotti et al., 1997; Zanotto et al., 1996). Twiddy et al. (2003) reported slower substitution rates for DENV-1 (4.6×10^{-4} s/s/y) and Klungthong et al. (2004) inferred a faster rate for DENV-4 (10.7×10^{-4} s/s/y). These differences are, however, possibly not significant, since these studies used different methods and different sequence samplings. For instance, Klungthong et al. (2004) surveyed only DENV-4 genotype I sequences from Thailand.

3.3. Population dynamics

While the dengue serotypes were found to be similar in measures of substitution patterns over time, the population dynamics of these lineages differed in both the complete and reduced data sets (Fig. 2). The phylodynamics of the serotypes are best recorded over the last 50 years. In the complete data set, DENV-1 serotype presented three periods in which the effective population size, as measured by the product of the effective population size and generation time, increased – the beginning of the 1970s, the 1980s and by the year 2000. The phylodynamic pattern inferred for DENV-2 exhibited periods of a rapid increase in the effective population size (the early 1980s and early 2000s). An abrupt increase in the effective population size in the early 2000s also characterized the evolution of DENV-3. Finally, DENV-4 displayed an interesting pattern. The effective population size of this serotype expanded during the 1980s. By the mid-1990s, however, its population size had decreased. The DENV-4 serotype was the only one that presented such a tendency. For all of the other serotypes, the population size increased dramatically since the early 2000s. Such increase is probably associated with increased of human population movement as well as uncontrolled growth of mosquito population. Also, commercial routes, which might spread mosquito eggs in diapause, may contribute to the recent increase in DENV effective population size worldwide.

DENV-2 is the serotype with the highest effective population size, followed by DENV-3, DENV-1 and DENV-4. Therefore, on a global scale, DENV-2 is the epidemiologically most relevant serotype. The significant increase of the population size of DENV serotypes 1–3 is corroborated by the data from the dengue incidence from public health agencies (e.g., http://www.who.int/mediacentre/factsheets/fs117/en). Such correlation corroborates the recent findings of Bennett et al. (2010), who showed that the cyclic pattern of DENV-4 epidemic in Puerto Rico is correctly inferred by the BSP. Interestingly, Villabona-Arenas and Zanotto (2011) also conducted the BSP analysis on global sequences of DENV-4 and

found that the genotypes I and II have opposite patterns of effective population growth over the last decate (GI is increasing, while GII is decreasing). Our analysis, thus, inferred DENV-4 growth as nearly constant over the same period. Therefore, a detailed analysis of the genotypes within DENV serotypes will lead to a finer picture of dengue evolution.

The historical dynamics of the effective population size of DENV serotypes should be interpreted with caution. When the sample size was reduced to 90 sequences, the phylodynamics inferred for DENVs 1, 2 and 3 differed. Although there still exists a cyclic pattern of population expansion and contraction during the last 50 years, the accelerated growth rates during the last decade of DENV-2 and DENV-3 were not recovered in the reduced data set (Fig. 2b). This indicates that the robustness of the BSP analysis to virus sequence sampling should be further explored.

3.4. Selection analysis and the causes of dengue diversification

The average d_N/d_S values over all codons of the envelope gene were similar in all of the serotypes and all estimates were significantly lower than one, varying from 0.06 to 0.07, which implies the action of negative selection, as previously documented (Holmes, 2003). The parameters of the beta distribution, used to model the probability of ω values along the sequence sites (M7 model), resulted in distributions of similar shapes between DENV-1 and DENV-2 and between DENV-3 and DENV-4 (Fig. 3). The DENV-3 and 4 serotypes presented a group of sites with a higher probability of being assigned to ω values close to one, indicating the relaxation of negative selection. However the highest density of the distribution was found when ω assumed values <1.

In nearly all DENV serotypes, the M1a–M2a, M7–M8 and M8–M8a tests of positive selection on codon sites failed to reject the null models in which the $\omega > 1$ category is absent (Table 3). In fact, the log-likelihoods of the models M1a and M2a were identical in all of the comparisons performed. A similar pattern was found in M8–M8a comparisons. However, the M7–M8 model comparisons in DENV-3 and DENV-4 indicated the existence of positive selection in these lineages, and the M8–M8a comparison could not reject the null hypothesis of absence of positive selection in DENV-4. The SLAC test of HyPhy also failed to estimate codon sites under positive selection in serotypes and in the joint data set (Table 3). The FEL algorithm inferred sites with $\omega > 1$ in DENV-3, DENV-4 and the joint data set.

Therefore, the null hypothesis of absence of positively selected codon sites in DENV serotypes could not be rejected by all tests implemented. Thus, there is no undisputed evidence of adaptive molecular evolution occurring along codon sites of the envelope gene. The tests that rejected positive selection, the M1a–M2a model comparison and the SLAC analysis, are subjected to lower type I error rates (Pond and Frost, 2005). Thus, we prefer to assume that the evidence of positive selection within DENV lineages is scant because it is not detectable by all methodologies available. Moreover, the Wilcoxon test failed to reject the null hypothesis that the strength of selection was homogeneous along the envelope gene among the serotypes, which further corroborates that codon sites of the envelope gene are, in general, under strong purifying selection.

Table 3
Tests of positive selection in DENV virus serotypes and in the joint data set. $2\Delta\ln L$, the likelihood ratio test statistic. PSS, positively selected site.

	M1a–M2a	M7–M8	M8–M8a	SLAC	FEL
DENV-1	$2\Delta\ln L = 0.0$	$2\Delta\ln L = 1.6$	$2\Delta\ln L = 0.0$	PSS = 0	PSS = 0
DENV-2	$2\Delta\ln L = 0.0$	$2\Delta\ln L = 5.1$	$2\Delta\ln L = 0.0$	PSS = 0	PSS = 0
DENV-3	$2\Delta\ln L = 0.0$	$2\Delta\ln L = 21.1^{***}$	$2\Delta\ln L = 0.1$	PSS = 0	PSS = 2 (132, 380)
DENV-4	$2\Delta\ln L = 0.0$	$2\Delta\ln L = 14.8^{***}$	$2\Delta\ln L = 2.9^{*}$	PSS = 0	PSS = 1 (108)
Joint	$2\Delta\ln L = 0.0$	$2\Delta\ln L = 0.2$	$2\Delta\ln L = 0.0$	PSS = 0	PSS = 1 (228)

* $0.01 < p\text{-value} < 0.05$.
*** $p\text{-value} < 0.001$.

With respect to branch-specific d_N/d_S analyses, we also found no evidence of adaptive molecular evolution acting on the basal diversification of dengue virus serotypes. In CODEML, although the null hypothesis of a single d_N/d_S for all branches of the dengue tree was rejected in favor of the alternative hypothesis that the basal branches had distinct d_N/d_S values, the ratio inferred for the basal branches was very small (0.003); it was actually smaller than the value estimated for the inside-serotype branches (0.054). This further indicates that global positive selection did not act significantly during divergence of the serotypes as revealed by the E gene. In HyPhy, the IFEL analysis could not differentiate d_N/d_S values along the internal and terminal branches and the branch-site test of episodic diversifying selection also did not inferred any branch under $d_N/d_S > 1$. Therefore, our results indicate that the emergence of DENV serotypes was probably led by stochastic events (founder effects) followed by periods of geographical isolation. Such a scenario was also proposed to explain the evolution of HIV-1 M subtypes (Rambaut et al., 2004).

Studies of adaptive molecular evolution on DENV are not abundant. Previous analyses that found evidence of positive selection on the envelope gene were conducted mainly on DENV-2 (Bennett et al., 2006; Twiddy et al., 2002a,b), and they indicated a few sites under positive selection. Here, no codon site was inferred with $P > 95\%$ at the $\omega > 1$ category in DENV-2. Actually, the M1a model could not be rejected for all serotypes. One explanation for such discrepancy is that the augmented sample size, including sequences from all over the world and from different time periods, reduced the rate of false positives, which are known to be an issue with maximum likelihood tests of selection (Suzuki and Nei, 2004). Another hypothesis is that the global sample masked intra-genotype events of adaptive evolution, especially in the passage from the sylvatic to the human cycle. However, as already reported by Twiddy et al., 2002a, at least for DENV-2, there is no evidence of adaptive evolution in the E gene during this event.

In conclusion, the present study has conducted a broad comparative investigation of the evolutionary epidemiology of dengue serotypes by means of the analysis of the envelope gene. We have shown that evolutionary rates are statistically identical in all serotypes and that the E gene codon sites are under purifying selection. No evidence of adaptive molecular evolution was found on the basal branches of the DENV phylogeny, which suggests

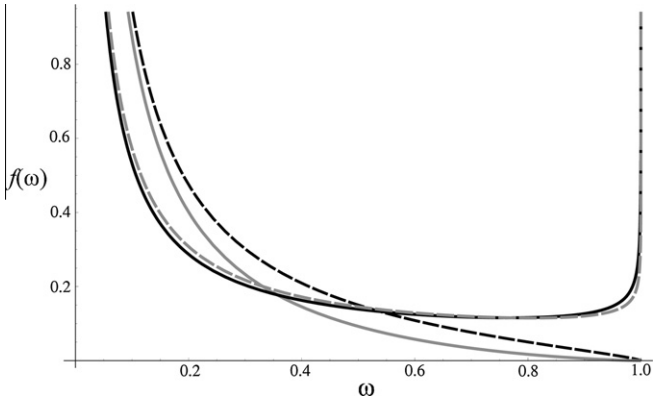

Fig. 3. Probability distribution of the ω values for codon sites along the envelope gene, as modeled by the Beta distribution (M7). Dashed black line: DENV-1; solid grey line: DENV-2; solid black line: DENV-3 and dashed grey line: DENV-4.

that early serotype diversification can be the result of founder effects. DENV genetic diversity presents geographic structure on a global scale. Such picture indicates that DENV spatial dynamics is limited. Finally, the effective population sizes of DENV-1, 2 and 3 have increased since the early 2000s. DENV-2 is the serotype with the highest global growth rate over the last decade, while DENV-4 has kept the effective population size nearly constant during the same period. However, effective sample size inference should be interpreted with caution, since we have demonstrated that the analysis is not robust to the sampling of virus sequences.

Acknowledgments

The authors are indebted to the editor and both reviewers of the manuscript. This work was significantly improved by their comments. This study was financially supported by Grants 103.136/2008, 110.838/2010 and 110.028/2011 from Faperj and 308147/2009-0 from CNPq to C.G.S. R.L.C. was supported by a scholarship from CAPES (Brazilian Ministry of Education)

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.meegid.2011.12.011.

References

Araujo, J.M.G., Nogueira, R.M.R., Schatzmayr, H.G., Zanotto, P.M.D., Bello, G., 2009. Phylogeography and evolutionary history of dengue virus type 3. Infection Genetics and Evolution 9, 716–725.

Bennett, S.N., Drummond, A.J., Kapan, D.D., Suchard, M.A., Munoz-Jordan, J.L., Pybus, O.G., Holmes, E.C., Gubler, D.J., 2010. Epidemic dynamics revealed in dengue evolution. Molecular Biology and Evolution 27, 811–818.

Bennett, S.N., Holmes, E.C., Chirivella, M., Rodriguez, D.M., Beltran, M., Vorndam, V., Gubler, D.J., McMillan, W.O., 2006. Molecular evolution of dengue 2 virus in Puerto Rico: positive selection in the viral envelope accompanies clade reintroduction. Journal of General Virology 87, 885–893.

Carrington, C.V.F., Foster, J.E., Pybus, O.G., Bennett, S.N., Holmes, E.C., 2005. Invasion and maintenance of dengue virus type 2 and type 4 in the Americas. Journal of Virology 79, 14680–14687.

Choisy, M., Woelk, C.H., Guegan, J.F., Robertson, D.L., 2004. Comparative study of adaptive molecular evolution in different human immunodeficiency virus groups and subtypes. Journal of Virology 78, 1962–1970.

Drummond, A.J., Pybus, O.G., Rambaut, A., Forsberg, R., Rodrigo, A.G., 2003. Measurably evolving populations. Trends In Ecology and Evolution 18, 481–488.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular Biology and Evolution 22, 1185–1192.

Goncalvez, A.P., Escalante, A.A., Pujol, F.H., Ludert, J.E., Tovar, D., Salas, R.A., Liprandi, F., 2002. Diversity and evolution of the envelope gene of dengue virus type 1. Virology 303, 110–119.

Gubler, D.J., 1998. Dengue and dengue hemorrhagic fever. Clinical Microbiology Reviews 11, 480.

Guzman, M.G., Alvarez, M., Rodriguez-Roche, R., Bernardo, L., Montes, T., Vazquez, S., Morier, L., Alvarez, A., Gould, E.A., Kouri, G., Halstead, S.B., 2007. Neutralizing antibodies after infection with Dengue 1 Virus. Emerging Infectious Diseases 13, 282–286.

Holmes, E.C., 2003. Patterns of intra- and interhost nonsynonymous variation reveal strong purifying selection in dengue virus. Journal of Virology 77, 11296–11298.

Holmes, E.C., 2008. Evolutionary history and phylogeography of human viruses. Annual Review of Microbiology 62, 307–328.

Holmes, E.C., Burch, S.S., 2000. The causes and consequences of genetic variation in dengue virus. Trends in Microbiology 8, 74–77.

Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C., 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. Journal of Molecular Evolution 54, 156–165.

Klungthong, C., Putnak, R., Mammen, M.P., Li, T., Zhang, C., 2008. Molecular genotyping of dengue viruses by phylogenetic analysis of the sequences of individual genes. Journal of Virological Methods 154, 175–181.

Klungthong, C., Zhang, C., Mammen Jr., M.P., Ubol, S., Holmes, E.C., 2004. The molecular epidemiology of dengue virus serotype 4 in Bangkok, Thailand. Virology 329, 168–179.

Kosakovsky Pond, S.L., Murrell, B., Fourment, M., Frost, S.D., Delport, W., Scheffler, K., 2011. A random effects branch-site model for detecting episodic diversifying selection. Molecular Biology and Evolution 28, 3033–3043.

Kukreti, H., Dash, P.K., Parida, M., Chaudhary, A., Saxena, P., Rautela, R.S., Mittal, V., Chhabra, M., Bhattacharya, D., Lal, S., Rao, P.V.L., Rai, A., 2009. Phylogenetic studies reveal existence of multiple lineages of a single genotype of DENV-1 (genotype III) in India during 1956–2007. Virology Journal 6, 1–10.

Lanciotti, R.S., Gubler, D.J., Trent, D.W., 1997. Molecular evolution and phylogeny of dengue-4 viruses. Journal of General Virology 78, 2279–2286.

Lanciotti, R.S., Lewis, J.G., Gubler, D.J., Trent, D.W., 1994. Molecular evolution and epidemiology of dengue-3 viruses. Journal of General Virology 75, 65–75.

Maddison, W.P., Maddison, D.R., 2009. Mesquite: a modular system for evolutionary analysis Version 2.7. Available from: <http://mesquiteproject.org>.

Martin, D., Rybicki, E., 2000. RDP: detection of recombination amongst aligned sequences. Bioinformatics 16, 562–563.

Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuvre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462–2463.

Mondini, A., Bronzoni, R.V.D., Nunes, S.H.P., Neto, F.C., Massad, E., Alonso, W.J., Lazzaro, E.S.M., Ferraz, A.A., Zanotto, P.M.D., Nogueira, M.L., 2009. Spatio-Temporal Tracking and Phylodynamics of an Urban Dengue 3 Outbreak in Sao Paulo, Brazil. Plos Neglected Tropical Diseases 3.

Parker, J., Rambaut, A., Pybus, O.G., 2008. Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty. Infection, Genetics and Evolution 8, 239–246.

Pond, S.L.K., Frost, S.D.W., 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Molecular Biology and Evolution 22, 1208–1222.

Pond, S.L.K., Frost, S.D.W., Grossman, Z., Gravenor, M.B., Richman, D.D., Brown, A.J.L., 2006. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. Plos Computational Biology 2, 530–538.

Pond, S.L.K., Frost, S.D.W., Muse, S.V., 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21, 676–679.

Rambaut, A., Posada, D., Crandall, K.A., Holmes, E.C., 2004. The causes and consequences of HIV evolution. Nature Reviews Genetics 5, 52–61.

Rico-Hesse, R., 1990. Molecular evolution and distribution of dengue viruses type-1 and type-2 in nature. Virology 174, 479–493.

Suzuki, Y., Nei, M., 2004. False-positive selection identified by ML-based methods: examples from the Sig1 gene of the diatom Thalassiosira weissflogii and the tax gene of a human T-cell lymphotropic virus. Molecular Biology and Evolution 21, 914–921.

Twiddy, S.S., Farrar, J.J., Chau, N.V., Wills, B., Gould, E.A., Gritsun, T., Lloyd, G., Holmes, E.C., 2002a. Phylogenetic relationships and differential selection pressures among genotypes of dengue-2 virus. Virology 298, 63–72.

Twiddy, S.S., Holmes, E.C., Rambaut, A., 2003. Inferring the rate and time-scale of dengue virus evolution. Molecular Biology and Evolution 20, 122–129.

Twiddy, S.S., Woelk, C.H., Holmes, E.C., 2002b. Phylogenetic evidence for adaptive evolution of dengue viruses in nature. Journal of General Virology 83, 1679–1689.

Vasilakis, N., Cardosa, J., Hanley, K.A., Holmes, E.C., Weaver, S.C., 2011. Fever from the forest: prospects for the continued emergence of sylvatic dengue virus and its impact on public health. Nature Reviews Microbiology 9, 532–541.

Vasilakis, N., Holmes, E.C., Fokam, E.B., Faye, O., Diallo, M., Sall, A.A., Weaver, S.C., 2007. Evolutionary processes among sylvatic dengue type 2 viruses. Journal of Virology 81, 9591–9595.

Villabona-Arenas, C.J., Zanotto, P.M., 2011. Evolutionary history of Dengue virus type 4: insights into genotype phylodynamics, Infection, genetics and evolution. Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases 11, 878–885.

Wang, E.Y., Ni, H.L., Xu, R.L., Barrett, A.D.T., Watowich, S.J., Gubler, D.J., Weaver, S.C., 2000. Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses. Journal of Virology 74, 3227–3234.

WHO, 2005. Dengue, Dengue Hemorrhagic Fever and Dengue Shock Syndrome in the context of the integrated management of childhood illness. Department of Child and Adolescent Health and Development, Geneva.

Wong, W.S.W., Yang, Z.H., Goldman, N., Nielsen, R., 2004. Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. Genetics 168, 1041–1051.

Yang, Z.H., 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Molecular Biology and Evolution 15, 568–573.

Yang, Z.H., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution 24, 1586–1591.

Yang, Z.H., Nielsen, R., Goldman, N., Pedersen, A.M.K., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.

Zanotto, P.M., Gould, E.A., Gao, G.F., Harvey, P.H., Holmes, E.C., 1996. Population dynamics of flaviviruses revealed by molecular phylogenies. Proceedings of the National Academy of Sciences of the United States of America 93, 548–553.

Zhang, C., Mammen Jr., M.P., Chinnawirotpisan, P., Klungthong, C., Rodpradit, P., Nisalak, A., Vaughn, D.W., Nimmannitya, S., Kalayanarooj, S., Holmes, E.C., 2006. Structure and age of genetic diversity of dengue virus type 2 in Thailand. Journal of General Virology 87, 873–883.

Zhang, C.L., Mammen, M.P., Chinnawirotpisan, P., Klungthong, C., Rodpradit, P., Monkongdee, P., Nimmannitya, S., Kalayanarooj, S., Holmes, E.C., 2005. Clade replacements in dengue virus serotypes 1 and 3 are associated with changing serotype prevalence. Journal of Virology 79, 15123–15130.