Rev. Inst. Med. trop. S. Paulo
46(3):145-152, May-June, 2004

MOLECULAR ANALYSIS OF THE DENGUE VIRUS TYPE 1 AND 2 IN BRAZIL BASED ON SEQUENCES OF THE GENOMIC ENVELOPE-NONSTRUCTURAL PROTEIN 1 JUNCTION REGION

Cecília Luiza S. SANTOS(1), Maria Anice M. SALLUM(2), Peter G. FOSTER(3) & Iray Maria ROCCO(1)

SUMMARY

The genomic sequences of the Envelope-Non-Structural protein 1 junction region (E/NS1) of 84 DEN-1 and 22 DEN-2 isolates from Brazil were determined. Most of these strains were isolated in the period from 1995 to 2001 in endemic and regions of recent dengue transmission in São Paulo State. Sequence data for DEN-1 and DEN-2 utilized in phylogenetic and split decomposition analyses also include sequences deposited in GenBank from different regions of Brazil and of the world. Phylogenetic analyses were done using both maximum likelihood and Bayesian approaches. Results for both DEN-1 and DEN-2 data are ambiguous, and support for most tree bipartitions are generally poor, suggesting that E/NS1 region does not contain enough information for recovering phylogenetic relationships among DEN-1 and DEN-2 sequences used in this study. The network graph generated in the split decomposition analysis of DEN-1 does not show evidence of grouping sequences according to country, region and clades. While the network for DEN-2 also shows ambiguities among DEN-2 sequences, it suggests that Brazilian sequences may belong to distinct subtypes of genotype III.

KEYWORDS: Dengue1; Dengue2; E/NS1; Brazil; Phylogeny.

INTRODUCTION

Dengue is an acute febrile illness that threatens 2.5 – 3 billion people living in tropical and subtropical regions of Africa, Asia and the Americas[12]. The etiological agent is a single-stranded, positive-sense RNA virus, which comprises four genetically and antigenically distinct serotypes of the dengue virus (DEN-1 to DEN-4) of the family *Flaviviridae*, genus *Flavivirus*. The endemic cycle of DEN viruses involve human hosts and mainly the mosquito vector *Aedes aegypti* in most urban centers of the tropics[10,11].

Infections by any of the four serotypes can result in a mild, self-limited illness, the dengue fever (DF) or a more severe form, the dengue hemorrhagic fever and dengue shock syndrome (DHF/DSS). Infection by one serotype does not protect against infection by a second serotype. In fact, secondary infection by a heterologous dengue serotype is considered the main risk factor associated to the occurrence of DHF/DSS[14]. On the other hand, severe manifestation of the disease observed in primary infection suggests that viral factors might also be involved [9,28,30]. However, the viral factors responsible for emergence of DHF/DSS are not well understood, mainly because of absence of *in vivo* and *in vitro* models to be used in laboratory studies.

In Brazil, dengue is of great concern. About 2,000,000 cases have been registered since its dramatic resurgence as epidemic of dengue fever in 1986, in an outbreak of DF caused by DEN-1 in Rio de Janeiro. Dengue type 2 virus was isolated in this state in 1990 during an epidemic with several cases of DHF/DSS[7]. Serotype DEN-3, detected for the first time in 1999, was isolated from a patient who arrived from Nicaragua[26]. Autochthonous infections caused by DEN-3 were registered in Rio de Janeiro in January 2001[22]. After that it spread throughout Brazil, co-circulating with DEN-1 and DEN-2 in several communities, increasing the risk of emergence of DHF/DSS.

In São Paulo State, DEN-1 illness was first detected in 1987 in a focal outbreak in the municipalities of Guararapes and Araçatuba. In the summer of 1990-1991, a large epidemic started in Ribeirão Preto. It expanded rapidly to other regions and since then periodic dengue epidemics associated with periods of hyperendemicity have been occuring in São Paulo State. DEN-2 virus was identified in 1996 and autochthonous cases of DEN-3 were notified in 2002 (Centro de Vigilância Epidemiológica da Secretaria de Estado da Saúde de São Paulo).

Aedes aegypti was eliminated from Brazil after a successful yellow fever mosquito vector control in the beginning of the 20th century. However, this species was reintroduced in 1967 and spread throughout the country. Nowadays *Ae. aegypti* infests more than two thirds of the municipalities in São Paulo State. Being the wealthiest state in Brazil, São Paulo is a commercial center, and has a lot of traffic in goods and people. It is well known that human travelers infected and incubating DEN virus can act as a reservoir and thus introduce new serotypes or variants in a country or region. Consequently, molecular epidemiologic

This work was supported by Secretaria de Estado da Saúde, São Paulo and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), grant no. 99/00558-2.

(1) Serviço de Virologia, Instituto Adolfo Lutz, Av. Dr. Arnaldo 355, 01246-902 São Paulo, Brasil.
(2) Núcleo de Pesquisas Taxonômica e Sistemática em Entomologia Médica, Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo, Av. Dr. Arnaldo 715, 01246-904 São Paulo, Brasil.
(3) Department of Zoology, The Natural History Museum, Cromwell Road, SW75BD, London, UK.
Correspondence to: Cecília Luiza Simões dos Santos, Instituto Adolfo Lutz, Av. Dr Arnaldo 355, 01246-902 São Paulo, SP, Brasil. Tel. 55.11. 30682903. E.mail: simoes.santos@uol.com.br and csantos@ial.sp.gov.br

studies are now stressed in order to monitor eventual appearance of genetic changes in dengue viruses, to identify which genotypes are circulating in an area, and to monitor introduction of new genotypes.

Genetic diversity of DEN-1 and DEN-2 viruses were addressed, based mainly on comparative sequence analyses that led to the recognition of several genotypes within each serotype. For DEN-1, three to five genotypes were proposed[4,8,24] and for DEN-2 five genotypes were suggested[19,24,25]. In Brazil, the genetic diversity of DEN-1 has been poorly explored and consequently few strains have been characterized at the molecular level. In the present study, the genomic sequences of the Envelope-Nonstructural protein 1 (E/NS1) regions of 84 DEN-1 and 22 DEN-2 isolates from Brazil were determined. Most of these strains were isolated in the period from 1995 to 2001 in endemic and regions of recent dengue transmission in São Paulo State. In addition, representative sequences of genotypes I to V of DEN-1, and genotypes II and III of DEN-2 were included for comparison with Brazilian isolates. These sequences include both previously published sequences[21,24,25] and sequences newly determined for this study. The main objectives of this study are: (1) to characterize the genotypes of DEN-1 and DEN-2 serotypes that are circulating in the State of São Paulo, (2) to identify potential origins of isolates, and (3) to examine phylogenetic relationships between sequences.

MATERIALS AND METHODS

Virus: All new DEN-1 and DEN-2 strains used in this study were isolated from the original patient serum between 1995-2001. Sera were collected from endemic areas in Araçatuba, São José do Rio Preto, Ribeirão Preto, and Barretos, and from regions of recent dengue transmission in Santos, Campinas, São Paulo, and Barueri. We also included DEN-1 isolates, which were obtained from reference center collection at the "Instituto Adolfo Lutz", São Paulo. These samples were obtained from patients during the first DEN-1 epidemic that occurred in Ribeirão Preto in 1990 and from patients who were infected in the states of Mato Grosso do Sul and Alagoas in 1991. All dengue virus isolates were from DF patients except for DEN-2 strain 190947, which was obtained from a DHF patient. All nucleotide sequences of E/NS1 gene junction region of DEN-1 and DEN-2 serotypes generated for this study are deposited in GenBank under accession numbers AY306015 to AY306098 (DEN-1), and AY306099 to AY306120 (DEN-2).

Viral isolation and serotype identification: Dengue viruses were isolated by inoculation of 20 µl serum aliquot/tube in *Aedes albopictus* clone C6/36 cell culture. Serotype identification was achieved by indirect fluorescent antibody test using serotype-specific monoclonal antibodies according to GUBLER *et al*.[13].

RNA extraction: Viral RNA was extracted from the culture supernatants of the first passage to C6/36 infected cells according to the procedure described by CHOMCZYNSKI & SACCHI[3].

Amplification and sequencing of the E/NS1 gene junction: A region encompassing 240 nucleotides of the E/NS1 gene junction was amplified by RT-PCR with primers described by RICO-HESSE[24]. A single reaction tube procedure was performed with the "SuperScript™ One-Step RT-PCR with Platinium® *Taq* System" (Invitrogen/Life Technologies, CA, USA). The RT-PCR products were sequenced directly using the "ABI Prism® Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit" according to the manufacturer's protocol. Sequences were determined in an ABI sequencer model 377 (PE Applied Biosystems, Foster City, CA, USA).

Sequence alignment: Nucleotide sequences of the E/NS1 gene junction of DEN-1 and DEN-2 viruses were aligned separately using the multiple sequence alignment method as implemented in CLUSTALX[34].

Phylogenetic analysis: Isolates that exhibited identical sequences were excluded from all phylogenetic analyses. Serum strains, locations, years of isolation, code and GenBank accession numbers of the sequences utilized in the phylogenetic and split decomposition analyses are listed in Table 1 (DEN-1) and Table 2 (DEN-2). Maximum likelihood (ML) analyses were performed using PAUP 4.0b10[31]. The best model in PAUP was chosen using ModelTest 3.06[23]. This program uses both a hierarchical likelihood ratio test (LRT) and the Akaike Information Criterion (AIC) to choose among available models; when the LRT and AIC disagreed the simpler model was chosen. After choosing a model in this way, the ModelTest tree was re-evaluated with a codon based site-specific (SS) model. Under the adopted model and using a NJ tree as the starting tree for branch-swapping, five iterative rounds of ML analysis were performed using the less intensive (NNI) to those using more intensive (TBR) branch-swapping. The most likely tree identified during each of these rounds was used as the starting tree for the next search, both for calculation of updated parameter values and for the initiation of branch-swapping. Branch-swapping were respectively, Nearest Neighbor Interchange (NNI), Subtree Pruning Regrafting (SPR), and SPR, Tree Bisection Reconnection (TBR) and TBR. Bootstrapping used 100 pseudoreplicates, each using a search as described above except that the branch-swapping regime was shortened to NNI, a single SPR, and a single TBR. When using the SS model, the program P4 (available from PGF) was used to bootstrap the data, followed by use of PAUP to analyze the pseudoreplicates.

Bayesian inference of phylogeny was carried out using the computer program MrBayes 2.01[16]. Since the kinds of models available in this program have limited overlap with those in PAUP*, Bayesian analysis was carried out using GTR[27] plus site-specific model. Program default values for prior probabilities were used. The Markov Chain Monte Carlo (MCMC) was allowed to run 5,000,000 generations for DEN-1, and 3,000,000 generations for DEN-2 both sampled every 100 generations after a burn-in of 50,000 generations. Conflicting phylogenetic signals were examined using the Split Decomposition method[2], which is developed in the program SplitsTree[17]. The maximum likelihood distance matrix between sequences was obtained using the TrNef + Γ model[33] for DEN-1 and GTR + Γ model[27] for DEN-2 data.

RESULTS

Nucleotide and amino acid sequence analysis: Nucleotide sequences of the E/NS1 genome junction of 84 DEN-1 strains (77 samples isolated in São Paulo State between 1995-2001, four samples isolated in Ribeirão Preto, State of São Paulo in 1990, one from Mato Grosso do Sul and two from Alagoas in 1991) and 22 DEN-2 strains (isolated in São Paulo State between 1996-2001) were used to assess the degree of genetic diversity among Brazilian sequences.

Table 1
DEN-1 virus strains used in the phylogenetic and split decomposition analysis

Strain	Code	Year of isolation	Location	GenBank acc.#
SPH 158985(A)*	SP95A	1995	Araçatuba, São Paulo / Brazil	AY306015
SPH 184232(B)	SP99B	1999	Santos, São Paulo / Brazil	AY306066
SPH 184477(C)	SP99C	1999	São José do Rio Preto, São Paulo / Brazil	AY306035
SPH 189720(D)	SP00D	2000	Ribeirão Preto, São Paulo / Brazil	AY306058
SPH 194573(E)	SP01E	2001	Barretos, São Paulo / Brazil	AY306075
SPH 182491(F)	SP98F	1998	Ribeirão Preto, São Paulo / Brazil	AY306054
SPH 197221(G)	SP01G	2001	Barueri, São Paulo / Brazil	AY306093
SPH 198979(H)	SP01H	2001	São Paulo, São Paulo / Brazil	AY306088
SPH 197357(I)	SP01I	2001	Barueri, São Paulo /Brazil	AY306097
SPH 159181(J)	SP95J	1995	Ribeirão Preto, São Paulo / Brazil	AY306049
SPH 167911(K)	SP97K	1997	Ribeirão Preto, São Paulo / Brazil	AY306053
SPH 191091(L)	SP00L	2000	Campinas, São Paulo / Brazil	AY306074
SPH 184069(M)	SP98M	1998	São José do Rio Preto, São Paulo / Brazil	AY306033
SPH 167148(N)	SP97N	1997	São José do Rio Preto, São Paulo / Brazil	AY306032
SPH 190477(O)	SP00O	2000	Campinas, São Paulo / Brazil	AY306073
SPH 202241(P)	SP01P	2001	São Paulo, São Paulo / Brazil	AY306091
SPH 194579(Q)	SP01Q	2001	Barretos, São Paulo / Brazil	AY306077
SPH 197218(R)	SP01R	2001	Barueri, São Paulo / Brazil	AY306092
SPH 117252(S)	SP90S	1990	Ribeirão Preto, São Paulo / Brazil	AY306040
SPH 194757	Brd1SP	2001	Barretos, São Paulo / Brazil	AF520798
Mochizuki	Japan43	1943	Nagasaki / Japan	M32929
IBH28326	Nige68	1968	Nigeria	M32927
691475	SriL69	1969	Sri Lanka	M32913
228682	Phil74	1974	Manila / Philippines	M32919
228686	Burm76	1976	Burma	M32920
IBH13689	Nige78	1978	Nigeria	M32928
DAK29177	Sene79	1979	Bandia / Senegal	M32909
1351	Colo82	1982	Colombia	M32900
1378	Mexi83	1983	Mexico	M32901
ArA15120	Ivor85	1985	Ivory Coast	M32922
CEA147	Braz86	1986	Ceará / Brazil	M32923
391094	Colo87	1987	Guaviare / Colombia	M32911
766602	Taiw87	1987	Kaohsiung / Taiwan	M32915
28973	Braz88	1988	Brazil	M32908
36589	Ango88	1988	Angola	M32912

* Representative sequence variants detected in this study are shown between brackets

The nucleotide alignments of E/NS1 sequences of DEN-1 and DEN-2 isolates indicate the presence of 19 DEN-1 and seven DEN-2 sequence variants, which were arbitrarily designated by letters: A to S for DEN-1 (Table 1), and A to G for DEN-2 (Table 2). These variants are not specific for a particular geographic region (Table 3). In addition, they are characterized by base substitutions that occur mainly in the third codon position and yield silent mutations; some of them are unique to an individual strain, whereas other base substitutions are shared by several dengue isolates. Comparison of the deduced amino acid sequences encoded in the E/NS1 region of DEN-1 and DEN-2 strains from Brazil included in this study reveals that amino acid sequence similarities range from 97.5 to 100.0% for dengue virus serotype 1 and 90.0 to 100.0% for dengue virus serotype 2 (MEGALIGN, DNASTAR, Inc.).

The E/NS1 alignment for 35 sequences of DEN-1 utilized in both phylogenetic and split decomposition analyses consists of 240 positions of which 65 are variable and 28 are parsimony-informative. For 29 sequences of DEN-2, the alignment generated 240 sites of which 59 are variable and 28 are parsimony-informative. Uncorrected ("p") distance between DEN-1 sequences ranged from 0% between Brd1SP and SP01P strains to 11.25% between Taiw87 and SriL69, and Japan43 and SriL69; for DEN-2 sequences uncorrected ("p") distances ranged from 0% between ES95 and SP97A, BR90_1 and BR90_2, and RJ90_3 and RJ90_4 strains to 10.41% between Thai64 and RJ98 strains. Nucleotide frequencies of E/NS1 region for DEN-1 data are: 29.01% of A, 19.92% of C, 26.04% of G and 25.01% of T, and DEN-2 are: 30.27% of A, 19.28% of C, 25.67% of G and 24.77% of T. The standard χ^2 test for base homogeneity implemented in PAUP 4.0b10[31] was unable to reject homogeneity of base frequencies among either DEN-1 or DEN-2 sequences, either using all sites or just using variable sites (p = 1.000000).

Phylogenetic and Split Decomposition analyses

Dengue 1 E/NS1 gene region: ModelTest 3.06[23] was used to choose a model. This program uses both a hierarchical likelihood ratio test (LRT) and the AIC to choose among available models in PAUP. The LRT

Table 2
DEN-2 virus strains used in the phylogenetic and split decomposition analysis

Strain	Code	Year of isolation	Location	GenBank acc.#
SPH 167231 (A)*	SP97A	1997	São José do Rio Preto, São Paulo / Brazil	AY306099
SPH 173072 (B)	SP98B	1998	Ribeirão Preto, São Paulo / Brazil	AY306107
SPH 194766 (C)	SP01C	2001	Barretos, São Paulo / Brazil	AY306115
SPH 196706 (D)	SP01D	2001	Santos, São Paulo / Brazil	AY306111
SPH 183966 (E)	SP98E	1998	Santos, São Paulo / Brazil	AY306109
SPH 190947 (F)	SP00F	2000	Campinas, São Paulo / Brazil	AY306114
SPH 182452 (G)	SP98G	1998	Ribeirão Preto, São Paulo / Brazil	AY306108
SPH 194757	Brd2SP	2001	Barretos, São Paulo / Brazil	AF520799
39056BR90	BR90_1	1990	Rio de Janeiro / Brazil	U91859
40247BR90	BR90_2	1990	Rio de Janeiro / Brazil	U91861
BR39122RJ90	RJ90_3	1990	Rio de Janeiro / Brazil	AF529064
BR39325RJ90	RJ90_4	1990	Rio de Janeiro / Brazil	AF529065
BR51502BA95	BA95_1	1995	Bahia / Brazil	AF529066
BR52582ES95	ES95	1995	Espírito Santo / Brazil	AF529068
BR51504BA95	BA95_2	1995	Bahia / Brazil	AF529067
BR62515RJ98	RJ98	1998	Rio de Janeiro / Brazil	AF529072
BR61654RN98	RN98_1	1998	Rio Grande do Norte / Brazil	AF529071
BR64020RN98	RN98_2	1998	Rio Grande do Norte / Brazil	AF529073
BR64905RJ99	RJ99	1999	Rio de Janeiro / Brazil	AF529075
BR66985RJ00	RJ00	2000	Rio de Janeiro / Brazil	AF529078
BR66703ES00	ES00_1	2000	Espírito Santo / Brazil	AF529076
BR66718ES00	ES00_2	2000	Espírito Santo / Brazil	AF529077
16681	Thai64	1964	Thailand	M32941
8110827	Jamaica81	1981	Jamaica	M32950
516	Thai83	1983	Thailand	M32947
766635	Taiw87	1987	Kaohsiung / Taiwan	M32949
57S	Viet87	1987	Saigon / Vietnam	M32948
028	Phil88	1988	Manila / Philippines	M32932
K0074	Thai94	1994	Thailand	U87349

*Representative sequence variants detected in this study are shown between brackets

Table 3
DEN-1 and DEN-2 sequence variants recovered from autochthones cases of dengue viruses infection detected in the geographic regions of the study

Locality	Serotype	
	DEN-1	DEN-2
	Variant	Variant
Araçatuba/São Paulo State	A; C	not present
Barueri/São Paulo State	A; G; I; R	not present
Barretos/São Paulo State	E; P; Q	C
Campinas/São Paulo State	A; L; O	A; F
Ribeirão Preto/São Paulo State	A; D; F; J; K; S	B; G
Santos/São Paulo State	A; B; C	D; E
São José do Rio Preto/São Paulo State	A; C; D; M; N	A; B
São Paulo / São Paulo State	A; B; H; P	not present
Mato Grosso do Sul State	A	not studied
Alagoas State	A	not studied

indicated the TrNef + Γ model[33], while the AIC indicated the SYM + Γ model[40]; the TrNef + Γ model was chosen, it being the simpler of the two. Having chosen TrNef as the rate matrix, we looked at whether a codon based site specific (SS) among site rate variation could model the data better than the gamma model indicated by ModelTest. The NJ tree used by ModelTest was evaluated with the TrNef + SS model and gave an increase in log likelihood of 23.7 relative to the TrNef + Γ model, at cost of 1 parameter, and so the TrNef + SS model was chosen as the best-fit model for DEN-1 data.

Maximum likelihood analysis was carried out under TrNef + Γ and TrNef + SS models. A single ML topology was generated under TrNef + Γ model with a log likelihood of - 897.38601. Similarly, a single most likely tree with a log likelihood of - 873.75807 was generated under TrNef + SS model. These two topologies are identical and define a major clade, which includes six Brazilian sequences (Braz86, Braz88, Brd1SP, SP01H, SP01P and SP01R), five strains from Africa, five strains from Asia and three strains from Latin America. Within this major clade, Braz86 and Braz88 formed a subgroup with Taiw87, Phil74 and Japan43, whereas the position of Brd1SP, SP01H and SP01P was unresolved (Fig. 1). Phylogenetic relationships among the remaining DEN-1 sequences from the São Paulo State are unresolved (Fig. 1). ML bootstrap analysis under either TrNef + Γ (not shown) or TrNef + SS models provides generally very poor support for relationships among DEN-1 sequences, with only six clades achieving ≥ 50% bootstrap proportions (Fig. 1). Bayesian 50% majority rule consensus tree (not shown) is similar to ML topology generated under both TrNef + Γ and TrNef + SS models except for the placement of Braz88, which was recovered in a basal position within the clade consisting of Burm76, Colo87, Braz86, Taiw87, Phil74 and Japan43.

In ML tree, Braz88 was placed within the clade consisting of Braz86, Taiw87, Phil74 and Japan43. In addition, SirL69 was recovered as sister to Ivor85, whereas Sene79 appeared as outgroup to (Ivor85, SriL69), and SP01H strain shared a sister-group relationship with Mexi83. Posterior probability for relationships among DEN-1 sequences are generally higher than ML bootstrap values, however, relationships among most sequences is poorly supported (Fig. 1). The network diagram generated with SplitTest[17] using the maximum likelihood distance matrix recovered under the TrNef + Γ model shows that all the sequences branched off separately from Nige68 except for (Taiw87, Phil74, Japan73), (SP99C, SP00O), (Sene79, SirL69) and (Brd1SP, SP01P), which cluster together having arisen from Nige68 (Fig. 2). This network does not show any evidence of grouping sequences according to country, region and date. Excluding Burm76, Ango88, Ivor85, Sene79, Nige68, SriL69, Nige78, Taiw87, Phil74 and Japan43 from the split decomposition analysis, it was possible to visualize two well-separated groups (not shown). The first group consists of six Brazilian sequences and three Caribbean strains of the genotype I. Although the relation among these sequences has some degree of ambiguity, it is reasonable to consider that they arose from Mexi83 sequence. In addition, the second group consists of several sequences isolated from the State of São Paulo, which may belong to a single outbreak.

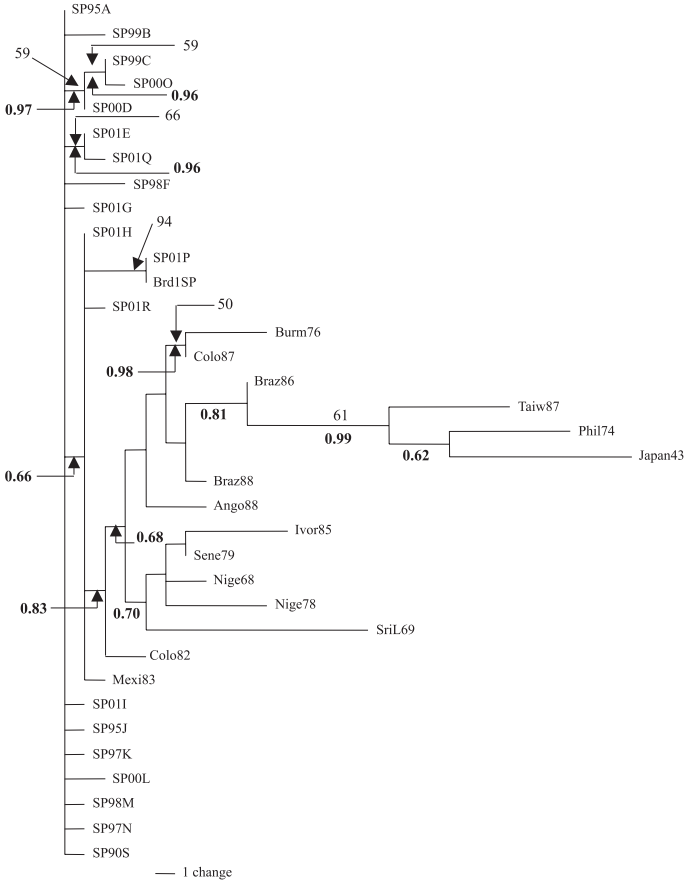

Fig. 1 - The single tree identified by maximum likelihood analysis of the E/NS1 junction region for DEN-1 data under the TrNef + SS model of nucleotide substitution. Numbers above branches indicate ML bootstrap proportions obtained under TrNef + SS model, and numbers below branches indicate posterior probabilities obtained under the GTR + SS model.

Dengue 2 E/NS1 gene region: The NJ tree was evaluated in the program ModelTest 3.06[23]. The likelihood ratio test found the K80 + Γ model[18] and the AIC found the GTR + Γ model[27] to be the best-fit models for DEN-2 data. The simpler of the two, the K80 + Γ model, was chosen. Again, SS among site rate variation was tested by comparing the K80 + Γ to the K80 + SS models. The SS model gave an increase in log likelihood of 13.5 with one additional parameter, indicating that the SS model has a better fit to the data.

Maximum likelihood analysis was performed under K80 + Γ and K80 + SS. A single most likely tree with a log likelihood = -786.88522 was generated under the K80 + Γ model. An identical ML topology was found under the K80 + SS model with log likelihood = -773.37693 (Fig. 3). The majority of DEN-2 sequences were recovered within a main clade, consisting of 16 Brazilian sequences, six Asian sequences and Jamaica81 sequence. Three subgroups were recovered within this clade: a basal subgroup consisting of RN98_1, which shares a sister-group relationship with RJ98; a clade leading to six Asian strains, Jamaica81, and four Brazilian strains; and a clade consisting of 10 Brazilian strains, which includes isolates from the states of São Paulo, Espírito Santo, Rio de Janeiro, Rio Grande do Norte and Bahia. In addition, two Brazilian sequences designated as BR90_2 and BR90_1 share a sister-group relationship with Jamaica81. Six sequences (four from São Paulo, one each from Bahia and Espírito Santo) clustered outside the large clade. Maximum likelihood bootstrap support for the majority of relationships among DEN-2 E/NS1 sequences is very poor with only eight groups achieving > 50% bootstrap proportion. The grouping leading to Asian strains except Viet87 is moderately well supported (82% bootstrap value) (Fig. 3). Similarly, the sister-group relationships between Brazilian strains (BR90_2 and BR90_1) and Jamaica81 are supported by 82% bootstrap

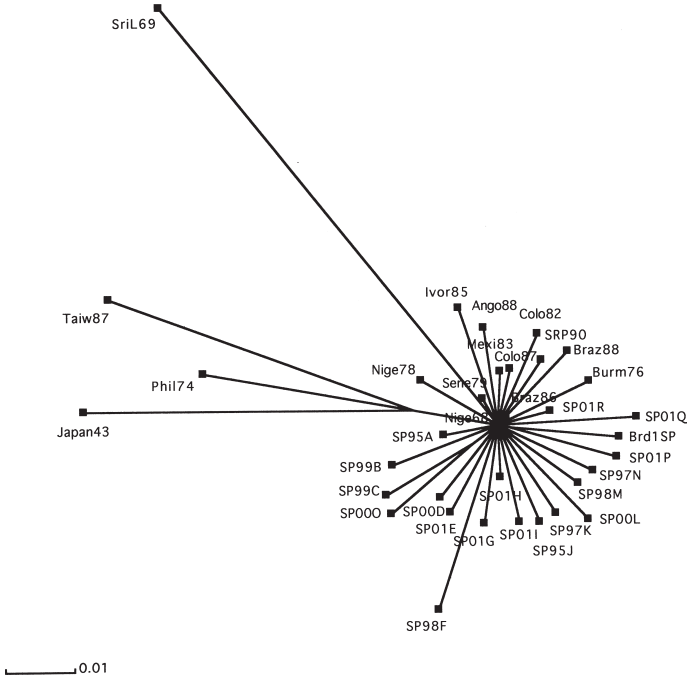

Fig. 2 - A graphical representation of the split-decomposable part of the evolutionary distance between DEN-1 sequence data generated using the distance matrix obtained under the TrNef + Γ model (drawn to scale; force triangle inequalities).

proportion. The posterior probabilities for most relationships among DEN-2 strains are higher than the corresponding ML bootstrap values. In fact, the posterior probabilities for 14 groups are higher than 0.65 (Fig. 3). Bayesian analyses recovered two major clusters: a cluster consisting of several Brazilian sequences, which were isolated from the State of São Paulo, Espírito Santo, Rio de Janeiro, Bahia and Rio Grande do Norte, and a cluster leading to four Brazilian strains, two isolated from Rio de Janeiro and both BR90_1 and BR90_2, and sequences representing genotypes II and III. The posterior probability for the split leading to the former cluster is 0.69 and for the latter cluster is 0.81. Also, posterior probability for the group leading to Jamaica81 and both BR90_1 and BR90_2 sequences is 1.0 (Fig. 3). The split decomposition analysis reveals that conflicting relationships exist among DEN-2 sequences (Fig. 4). It also shows three main clusters, however, none of them represent previously defined genotypes of the virus. The first cluster consists of sequences of five Asian countries; however relation between these sequences is ambiguous. The second cluster consists of two Brazilian sequences (BR90_1 and BR90_2), Jamaica81 and Viet87, suggesting that they are of the same virus genotype. A third group is composed of sequences, which have been circulating in several regions in Brazil. The network shows ambiguities among these sequences but suggests that they belong to different subtypes of genotype III (Fig. 4). Excluding Phil88, Thai94, Thai83, Thai64 and Taiw87 from the split decomposition analysis, four subgroups were identified, a subgroup composed of Jamaica81, Viet87 and four Brazilian strains, which can be assigned to genotype III (not shown). The remaining Brazilian sequences cluster into three subgroups, which probably belong to different subtypes of genotype III and may be of different origin. The network graph shows ambiguities between all four subgroups (not shown).

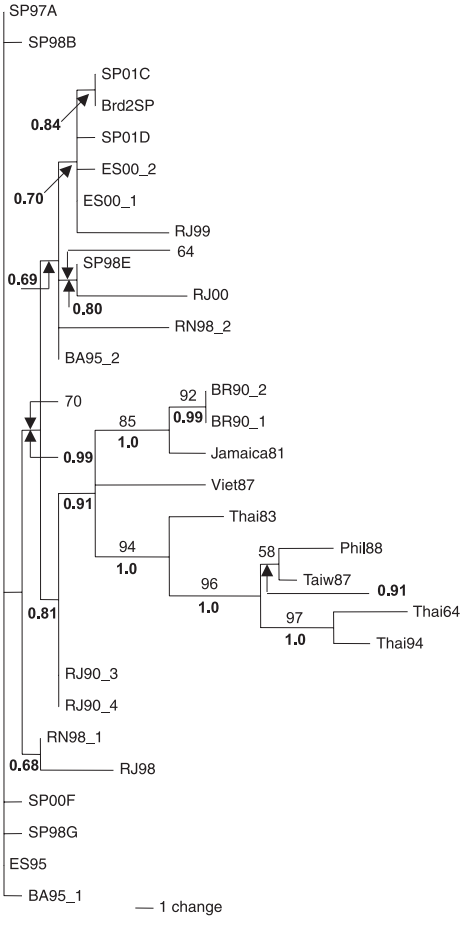

Fig. 3 - The single tree identified by maximum likelihood analysis of the E/NS1 junction region for DEN-2 data under the K80 + SS model of nucleotide substitution. Numbers above branches indicate ML bootstrap proportions obtained under the K80 + SS model, and numbers below branches indicate posterior probabilities obtained under the GTR + SS model.

DISCUSSION

According to the Brazilian Ministry of Health data, almost 70% of notified cases of dengue infection are concentrated in urban areas of municipalities with more than 50,000 inhabitants, which are undergoing to economic development. Commercial exchange among urban areas is considered to be responsible for *Ae. aegypti* dispersion and for the spread of dengue virus infection throughout Brazilian municipalities.

Dengue epidemics in the State of São Paulo, which is the most economically developed state in Brazil, and considered the main point for converging and dispersing national and international goods and people, have increased dramatically in recent years. Consequently, monitoring genetic changes in circulating or introduced viruses may be

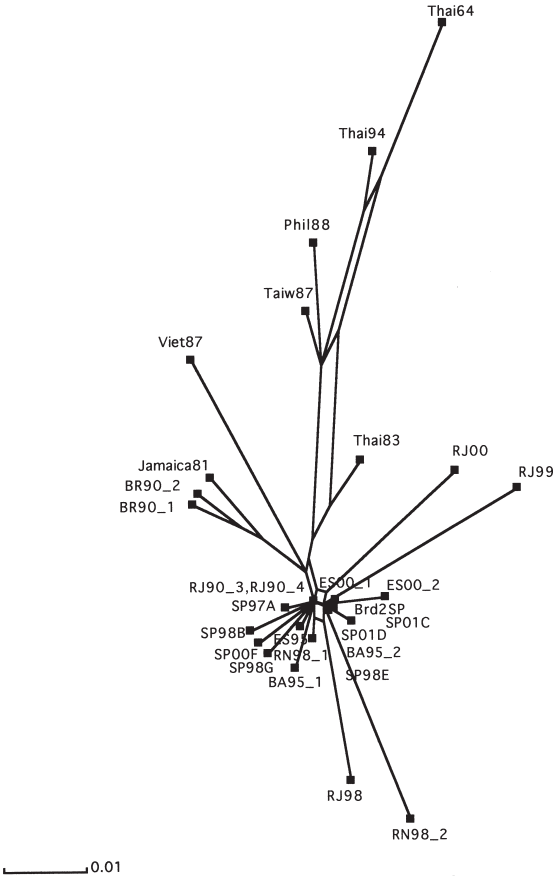

Fig. 4 - A graphical representation of the split-decomposable part of the evolutionary distance between DEN-2 sequence data generated using the distance matrix obtained under the GTR + Γ model (drawn to scale; force triangle inequalities).

useful to understand emergence of hyperendemicity associated with epidemics in an area[12]. In the present study, genetic diversity of several DEN-1 and DEN-2 isolates from autochthonous cases in Brazil was analyzed. The analysis involved the sequencing of a fragment of 240 base pairs of the E/NS1 gene junction of dengue virus. Although representing about two percent of the total dengue virus genome, E/NS1 region was used to generate evolutionary information, which may be useful for molecular epidemiologic studies[24, 25]. Nineteen DEN-1 and seven DEN-2 sequence variants were identified among 84 DEN-1 and 22 DEN-2 isolates. There were no significant differences in the distribution of these variants in either endemic areas or recent regions of dengue virus transmission. Variant A of DEN-1, for example, comprises the most fully dispersed one, which were isolated in Araçatuba (1995, 1996), São José do Rio Preto (1995, 1996, 1999, 2000), Ribeirão Preto (1990, 1995, 1996, 1999, 2000), Santos (1997, 1998, 2001), Campinas (1998, 2000), Barueri (2001) and during the first epidemic in São Paulo City in 2001. Furthermore, pairwise comparisons of variant SP95A and the sequence of a strain isolated in the State of Rio de Janeiro in 1990 considered the reference for DEN-1 epidemic in South America[5,38] reveal 100% nucleotide identity. Sequences variants SP01P (DEN-1) and SP01C (DEN-2) are identical to Brd1SP and Brd2SP isolated from a patient with concurrent dengue infection[29].

Sequence data were also used to examine phylogenetic relationships among sequences of dengue virus type 1 and type 2, isolated mainly from the State of São Paulo but also from different regions of Brazil and of the world. Although we have employed several sequences, which were also used in RICO-HESSE's[24] study, phylogenetic relationships generated from the analyses performed for this study differ from those proposed genetic groups. Differences in the results may be partially attributable to the different methods of analysis employed in both studies, and also because DEN-1 and DEN-2 sequence data are not identical. GONÇALVEZ *et al.*[8] examined genetic diversity and phylogenetic relationships among 44 strains of DEN-1 from different regions of the world. They suggested the existence of five genotypes, with a single genotype circulating in the Americas associated with both DF and DHF/DSS. Similarly, AVILÉS *et al*.[1] indicated that a single genotype of DEN-1 circulated in Argentina and Paraguay during epidemics in 2000. The results of all the analyses carried out for the present study indicate that different subtypes of a single genotype of DEN-1 virus are co-circulating in Brazil. Also, the network diagram generated when Burm76, Ango88, Ivor85, Sene69, Nige68, SriL69, Nige78, Taiw87, Phil74 and Japan43 were excluded, suggest that the genotype that is circulating in Brazil is closely related to Colo87, Colo82 and Mexi83 of genotype I of RICO-HESSE[24]. Additionally, based on the results of all the analyses performed for the present study, it is reasonable to suppose that divergence among DEN-1 serotypes isolated from different regions of the world is not strong, corroborating the results of GONÇALVEZ *et al.*[8] who estimated that divergence among the DEN-1 epidemic genotypes occurred approximately 100 years ago.

Previous studies have shown that DEN-2 strains, which are circulating in Brazil belong to genotype III within RICO-HESSE's classification[21,25]. In the present study, we have confirmed that BR90_1 and BR90_2 strains belong to genotype III. In addition, the results of both phylogenetic and split decomposition analyses indicate that the remaining strains isolated in several regions in Brazil are closely related and belong to distinct subtypes of the genotype III. However, relationship among them is ambiguous.

Finally, the results of the ML and Bayesian analysis are consistent with the conclusion that the E/NS1 gene region consisting of 240 base pairs does not contain enough information for recovering phylogenetic relationships within DEN-1 and DEN-2 sequences used for the present study. In addition, the results suggest that divergences among DEN-1 and DEN-2 sequences are very low and may have occurred too recently [8,15,37] to be tracked by information in the E/NS1 data. The low genetic distance separating all the sequences seems to support this conclusion. In addition, it is also important to take into consideration that the maximum likelihood and Bayesian methods for phylogenetic inference assume that evolution is hierarchical. However, in virus genome, recombination, reassortment and horizontal transfer occur frequently. Regarding to dengue virus, recombination event is stressed since there has been increasing evidence for its occurrence in natural virus population[35,36,39]. Consequently, all standard methods for phylogenetic reconstruction, which assumes that evolution is a tree-like process, may recover incorrect topologies and poor resolution when used to analyze virus genome sequence data[6,20].

RESUMO

Análise molecular dos vírus dengue tipo 1 e 2 no Brasil, baseada nas seqüências da região da junção dos genes do envelope e da proteína não estrutural 1

Foram determinadas as seqüências nucleotídicas da junção dos genes do envelope e da proteína não estrutural 1 (E/NS1) de 84 cepas de DEN-1 e 22 cepas de DEN-2 do Brasil. A maioria dessas cepas foi isolada no período de 1995-2001, em regiões endêmicas e de transmissão recente no Estado de São Paulo. Seqüências da junção E/NS1 de DEN-1 e DEN-2 de outras regiões geográficas brasileiras e mundiais, obtidas do GenBank, foram também utilizadas neste estudo. As análises foram efetuadas utilizando-se as técnicas de Verossimilhança Máxima e Bayesiana de inferência filogenética. Os resultados das análises das seqüências de DEN-1 e DEN-2 são ambíguos e o suporte para a maioria dos grupos é baixo, sugerindo que a região E/NS1 não é filogeneticamente informativa. O gráfico gerado na análise de decomposição dos grupos de DEN-1 não mostrou evidências de agrupamento das seqüências de acordo com os países, as regiões ou clados. No entanto, para DEN-2 evidenciou a existência de ambigüidades entre as seqüências, sugerindo que as brasileiras pertencem a subtipos distintos do genotipo III.

ACKNOWLEDGMENTS

Secretaria de Estado da Saúde do Estado de São Paulo and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Grant 1999/00558-2 for financial support. To M.A.A. Bastos for laboratory assistance, and Tiyo Sakurai and Gizelda Katz for suggestions and providing helpful comments to improve the paper. MAMS was supported by grant from FAPESP, São Paulo, Brasil. (Grant nº 99/10517-1).

REFERENCES

1. AVILÉS, G.; ROWE, J.; MEISSNER, J. *et al.* - Phylogenetic relationships of dengue-1 viruses from Argentina and Paraguay. **Arch. Virol., 147:** 2075-2087, 2002.

2. BANDELT, H.-J. & DRESS, A. W. M**. -** Split decomposition: a new and useful approach to phylogenetic analysis of distance data. **Molec. Phylogenet. Evol., 1:** 242-252, 1992.

3. CHOMCZYNSKI, P. & SACCHI, N. - Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. **Analyt. Biochem., 162:** 156-159, 1987.

4. CHUNGUE, E.; CASSAR, O.; DROUET, M.T. *et al*. - Molecular epidemiology of dengue –1 and dengue-4 viruses. **J. gen. Virol., 76:** 1877-1884, 1995.

5. DESPRÈZ, P.; FRENKIEL, M.P. & DEUBEL, V. - Differences between cell membrane fusion activities of two dengue type 1 isolates reflect modifications of viral structure. **Virology, 196:** 209-219, 1993.

6. DOPAZO, J.; DRESS, A. & HAESELER, A. VON - Split decomposition: a technique to analyze viral evolution. **Proc. nat. Acad. Sci. (Wash.), 90:** 10320-10324, 1993.

7. FIGUEIREDO, L.T.M. **-** History, present and future of dengue fever in Brazil. In : TRAVASSOS DA ROSA, A.P.A.; VASCONCELOS, P.F.C. & TRAVASSOS DA ROSA, J.F.S., ed. **An overview of arbovirology in Brazil and neighbouring countries**. Belém, Instituto Evandro Chagas, 1998. p. 155-163.

8. GONÇALVEZ, A.P.; ESCALANTE, A.A.; PUJOL, F.H. *et al.* - Diversity and evolution of the envelope gene of dengue virus type 1. **Virology, 303:** 110-119, 2002.

9. GUBLER, D. J.; REED, D.; ROSEN, L. & HITCHOCK Jr., J.R. - Epidemiologic, clinical, and virologic observations on dengue in the Kingdom of Tonga. **Amer. J. trop. Med. Hyg., 27:** 581-589, 1978.

10. GUBLER, D.J. - The global pandemic of dengue/dengue haemorrhagic fever: current status and prospects for the future. **Ann. Acad. Med. Singapore, 27:** 227-234, 1998a.

11. GUBLER, D.J. - Dengue and dengue hemorrhagic fever. **Clin. Microbiol. Rev., 11:** 480-496, 1998b.

12. GUBLER, D.J**. -** The global emergence/resurgence of arboviral diseases as public health problems. **Arch. med. Res., 33:** 330-342, 2002.

13. GUBLER, D.J.; KUNO, G.; SATHER, G.E.; VELEZ, M. & OLIVER, A. - Mosquito cell cultures and specific monoclonal antibodies in surveillance for dengue viruses. **Amer. J. trop. Med. Hyg., 33:** 158-165, 1984.

14. HALSTEAD, S.B. - Pathogenesis of dengue: challenges to molecular biology. **Science, 239:** 476-481, 1988.

15. HOLMES, E.C. & TWIDDY, S.S. - The origin, emergence and evolutionary genetics of dengue virus. **Infect. Genet. Evol., 3:** 19-28, 2003.

16. HUELSENBECK, J.P. & RONQUIST, F. - MrBayes: Bayesian inference of phylogeny. Rochester, Department of Biology, University of Rochester, 2002. (Distributed by the author).

17. HUSON, D.H. - SplitsTree: analyzing and visualizing evolutionary data. **Bioinformatics, 14:** 68-73, 1998.

18. KIMURA, M. - A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. **J. molec. Evol., 16:** 111-120, 1980.

19. LEWIS, J.G.; CHANG, G.J.; LANCIOTTI, R.S. *et al*. - Phylogenetic relationships of dengue-2 viruses. **Virology, 197:** 216-224, 1993.

20. McCORMACK, G.P. & CLEWLEY, J.P. - The application of molecular phylogenetics to the analysis of viral genome diversity and evolution. **Rev. med. Virol., 12:** 221-238, 2002.

21. MIAGOSTOVICH, M.P.; SEQUEIRA, P.C.; DOS SANTOS, F.B. *et al*. - Molecular typing of dengue virus type 2 in Brazil. **Rev. Inst. Med. trop. S. Paulo, 45:** 17-21, 2003.

22. NOGUEIRA, R.M.R.; MIAGOSTOVICH, M.P.; FILIPPIS, A.M.B. *et al*. - Dengue virus type 3 in Rio de Janeiro, Brazil. **Mem. Inst. Oswaldo Cruz, 96:** 925-926, 2001.

23. POSADA, D. & CRANDALL, K.A. - Modeltest: testing the model of DNA substitution. **Bioinformatics, 14:** 817-818, 1998.

24. RICO-HESSE, R. - Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. **Virology, 174:** 479-493, 1990.

25. RICO-HESSE, R.; HARRISON, L.M.; SALAS, R.A. *et al*. - Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. **Virology, 230:** 244-251, 1997.

26. ROCCO, I.M.; KAVAKAMA, B.B. & SANTOS, C.L.S. **-** First isolation of dengue 3 in Brazil from an imported case. **Rev. Inst. Med. trop. S. Paulo, 43:** 55-57, 2001.

27. RODRIGUEZ, F.; OLIVER, J.L.; MARIN, A. & MADINA, J.R**.** - The general stochastic model of nucleotide substitution. **J. theor. Biol., 142:** 485-501, 1990.

28. ROSEN, L. - The Emperor's New Clothes revisited, or reflections on the pathogenesis of dengue hemorrhagic fever. **Amer. J. trop. Med. Hyg., 26:** 337-343, 1977.

29. SANTOS, C.L.S.; BASTOS, M.A.A.; SALLUM, M.A.M. & ROCCO, I.M. - Molecular characterization of dengue viruses type 1 and 2 isolated from a concurrent human infection. **Rev. Inst. Med. trop. S. Paulo, 45:** 11-16, 2003.

30. SCOTT, R.M. - Shock syndrome in primary dengue infections. **Amer. J. trop. Med. Hyg., 25:** 866-874, 1976.

31. SWOFFORD, D.L. - PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4. Sunderland, Sinauer Associates, 2002.

32. SWOFFORD, D.L.; OLSEN, G.J.; WADELL, P.J. & HILLIS, D.M. - Phylogenetic inference. In: HILLIS, D.M.; MORITZ, C. & MABLE, B.K., ed. **Molecular systematics.** 2. ed. Sunderland, Sinauer Associates, 1996. p. 407-514.

33. TAMURA, K. & NEI, M. - Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. **Molec. Biol. Evol., 10:** 512-526, 1993.

34. THOMPSON, J.D.; GIBSON, T.J.; PLEWNIAK, F. *et al*. - The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality tools. **Nucleic Acids Res., 24:** 4876-4882, 1997.

35**.** TOLOU, H.J.; COUSSINIER-PARIS, P.; DURAND, J.P. *et al*. - Evidence for recombination in natural populations of dengue virus type 1 based on the analysis of complete genome sequences. **J. gen. Virol., 82**: 1283-1290, 2001.

36**.** TWIDDY, S.S. & HOLMES, E.C. - The extent of homologous recombination in members of the genus Flavivirus. **J. gen. Virol., 84**: 429-440, 2003.

37. TWIDDY, S.S.; HOLMES, E.C. & RAMBAUT, A. - Inferring the rate and time-scale of dengue virus evolution. **Mol. Biol. Evol., 20**: 122-129, 2003.

38. WANG, E.; NI, H.; XU, R. *et al*. - Evolutionary relationships of endemic/epidemic and sylvatic dengue viruses. **J. Virol., 74:** 3227-3234, 2000.

39. WOROBEY, M.; RAMBAUT, A. & HOLMES, E.C. - Widespread intra-sorotype recombination in natural populations of dengue virus. **Proc. nat. Acad. Sci. (Wash.), 96:** 7352-7357, 1999.

40. ZHARKIKH, A. - Estimation of evolutionary distances between nucleotide sequences. **J. Molec. Evol., 39:** 315-329, 1994.

Received: 30 September 2003
Accepted: 7 june 2004