View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com

SCIENCE DIRECT

Virology 336 (2005) 163 – 172

VIROLOGY

www.elsevier.com/locate/yviro

# Lineage extinction and replacement in dengue type 1 virus populations are due to stochastic events rather than to natural selection

Hlaing Myat Thu[a,b], Kym Lowry[a], Limin Jiang[a], Thaung Hlaing[b], Edward C. Holmes[c], John Aaskov[a,*]

[a]*School of Life Sciences, Queensland University of Technology, 2, George Street, Brisbane 4001, Australia*
[b]*Department of Medical Research, Ziwaka Road, Yangon 11191, Myanmar*
[c]*Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA*

Received 18 January 2005; returned to author for revision 25 February 2005; accepted 19 March 2005
Available online 18 April 2005

**Abstract**

Between 1996 and 1998, two clades (B and C; genotype I) of dengue virus type 1 (DENV-1) appeared in Myanmar (Burma) that were new to that location. Between 1998 and 2000, a third clade (A; genotype III) of DENV-1, which had been circulating at that locality for at least 25 years, became extinct. These changes preceded the largest outbreak of dengue recorded in Myanmar, in 2001, in which more than 95% of viruses recovered from patients were DENV-1, but where the incidence of severe disease was much less than in previous years. Phylogenetic analyses of viral genomes indicated that the two new clades of DENV-1 did not arise from the, now extinct, clade A viruses nor was the extinction of this clade due to differences in the fitness of the viral populations. Since the extinction occurred during an inter-epidemic period, we suggest that it was due to a stochastic event attributable to the low rate of virus transmission in this interval.
© 2005 Elsevier Inc. All rights reserved.

*Keywords:* Dengue viruses; Evolution; Phylogeny; Genetic bottleneck

## Introduction

Dengue is a disease caused by four serotypes of a flavivirus of the same name. Infection with dengue viruses may be inapparent or may result in "influenza-like" symptoms (dengue fever, DF), hemorrhagic fever (dengue hemorrhagic fever, DHF), or hypovolemic shock (dengue shock syndrome, DSS) (World Health Organization., 1997). The severity of symptoms is influenced by the serotype and genotype of the infecting virus (Rico-Hesse et al., 1997; Vaughn et al., 2000) and by previous infections with dengue viruses (Halstead, 1970). It has been estimated that up to 100 million dengue infections occur annually (Monath, 1994) and the case fatality rate can be as high as 44% (Rigau-Perez et al., 1998). Dengue occurs in more than 100 countries and 2.5 billion people are at risk of infection (Gubler, 1998). In urban settings in Asia, the most common mosquito vector for this virus is *Aedes aegypti* (Strickman and Kittayapong, 2003). The competence of Aedes mosquitoes to become infected with dengue viruses and their ability to transmit the virus to a new host are believed to be influenced by its genotype (Bosio et al., 2000; Gubler and Rosen, 1976).

Dengue viruses (DENV-1 to DENV-4) have RNA genomes of approximately 11 kb which encode three structural proteins (core protein [C], membrane-associated protein [M], and envelope protein [E]) and seven non-structural (NS) proteins. The 5′ and 3′ untranslated regions (UTRs) of flaviviruses play an important role in initiating translation, viral replication, and transcription (reviewed in Rice, 1996). The absence of a proof-reading capacity in RNA-dependent RNA polymerases (Steinhauer et al., 1992) is believed to give rise to approximately one nucleotide change in a dengue virus genome during each cycle of

* Corresponding author. Fax: +61 7 3864 1534.
*E-mail address:* j.aaskov@qut.edu.au (J. Aaskov).

0042-6822/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2005.03.018

replication. Although the majority of mutations that arise within each host are likely to be deleterious (Holmes, 2003), such variation may give rise to a pool of diverse viruses able to occupy new ecological niches or to adapt to sudden pressures on the cycle of replication. Indeed, although analyses of the ratio of nonsynonymous to synonymous nucleotide changes per site ($d_N/d_S$) within and between DENV populations suggest they are subject to strong purifying selection (Holmes, 2003; Wittke et al., 2002), with relatively little evidence for adaptive evolution (Bennet et al., 2003; Holmes and Twiddy, 2003; Twiddy et al., 2002), there is growing experimental evidence for differences in fitness among genotypes (Cologna et al., 2005). In three instances where sudden changes in the genotype of dengue viruses at a single locality have been observed (Sittisombut et al., 1997; Thu et al., 2004; Wittke et al., 2002), the changes appeared to be due to stochastic events giving rise to large population bottlenecks. However, these three analyses focused on the envelope (E) protein genes of DENV and so would not have detected changes due to selective pressures acting on other viral proteins or on the 3′ and 5′ UTRs of the genome. Where changes in DENV populations have been attributed to selection pressures, the changes usually have occurred in proteins other than E (Bennet et al., 2003; Holmes and Twiddy, 2003; Twiddy et al., 2002). A lack of longitudinal studies of the evolution of DENV at single locations has also hampered our understanding of how frequently extinctions of DENV genotypes might have occurred or of the evolutionary processes responsible.

The extinction of a lineage of DENV-1 has been documented recently in Myanmar (Thu et al., 2004). In this case, a minority of the sequences of the E genes of DENV-1 from Myanmar were assigned to genotype III and all were from viruses sampled no later than 1998. In contrast, the majority of Myanmar DENV-1 strains fell into genotype I (although at a variety of positions within this clade), including all those sampled post-1998. To determine, in more detail, the evolutionary processes responsible for the lineage extinction that apparently occurred in 1998, we determined the nucleotide sequences of genomes of DENV-1 collected at Yangon, in Myanmar, between 1971 (the earliest isolate available; Ming et al., 1974) and 2002. In particular, we asked whether there was any evidence for natural selection acting on proteins other than E playing a role in lineage extinction. These analyses also included the only examples of DENV-1 recovered from individual mosquitoes.

## Results

Nucleotide sequences were derived from the 5′ UTR to the end of the second cyclization sequence (CS2) in the 3′ UTR (nucleotides 1–10,614) of thirteen viruses (nine from dengue patients and four from the more than 600 mosquitoes assayed). The nucleotide sequences of the remaining 120 nucleotides of 3′ UTR could be obtained for only two viruses (D1.Myanmar.31459/98 and D1.Myanmar.31987/98); one from each of the two clades (B and C) of viruses circulating at that location now (see below). The nucleotide sequences of the open reading frames (ORF) of the viruses varied at 1285 sites resulting in deduced amino acid changes at 175 sites. The nucleotide and amino acid variation (changes per ORF) between the consensus sequences of the ORFs of viruses from the 11 patients (nucleotide changes at 1213 sites resulting in 139 deduced amino acid changes) was similar to that from the four mosquitoes (nucleotide changes at 492 sites resulting in 60 deduced amino acid changes).

The nucleotide sequences of the ORFs of the fifteen Myanmar DENV-1 fell into three clades (denoted A, B, and C; Fig. 1). The earliest examples of DENV-1 from Myanmar (D1.Myanmar.40553/71 and D1.Myanmar.40568/76) and two more recent isolates (D1.Myanmar.23819/96 and D1.Myanmar.32514/98) fell in genotype III, with the sequences from the viruses from 1996 and 1998 forming the most closely related pair. For ease of comparison in this paper, these viruses have been designated as “clade A”. The remaining viruses segregated into two distinct clades within genotype I. “Clade B” contained two of the four virus isolates recovered from *A. aegypti* mosquitoes (D1.Myanmar.194/01, D1.Myanmar.206/01) and was related, although distantly, to isolates sampled in Djibouti in 1998 and in China in 1980. The remaining viruses, including two further isolates from mosquitoes (D1. Myanmar.059/01 and D1.Myanmar.305/01), formed a separate “clade C” which was related to an isolate recovered in Cambodia in 1998 and more distantly to a strain sampled in Singapore in 1990. All clades are supported by high bootstrap values. Similar phylogenetic analyses employing the nucleotide sequences of the individual DENV-1 genes of these viruses revealed similar phylogenetic relationships to those for the entire ORF (data not shown). There were no significant changes in the positions of the Myanmar strains in the phylogenetic trees derived with the nucleotide sequences of individual genes, indicating an absence of recombination. Finally, no attempt was made to identify nucleotide or amino acid changes associated with disease of varying severity because almost all low-passage, contemporary DENV-1 (clades B and C) from Myanmar were obtained from patients with mild disease (DF, DHF I or II).

Using a newly developed Bayesian MCMC method, we estimated the overall rate of nucleotide substitution in the ORF of DENV-1 assigned to genotypes I and III. These rates were found to be almost identical, at 0.599 and 0.583 $\times 10^{-3}$ substitutions per site, per year, respectively, and with relatively little statistical uncertainty as reflected in high probability densities [HPD] of 0.482–0.726 $\times 10^{-3}$ and 0.489–0.677 $\times 10^{-3}$, respectively. These substitution rates fall within the range seen in other RNA viruses (Jenkins et al., 2002).

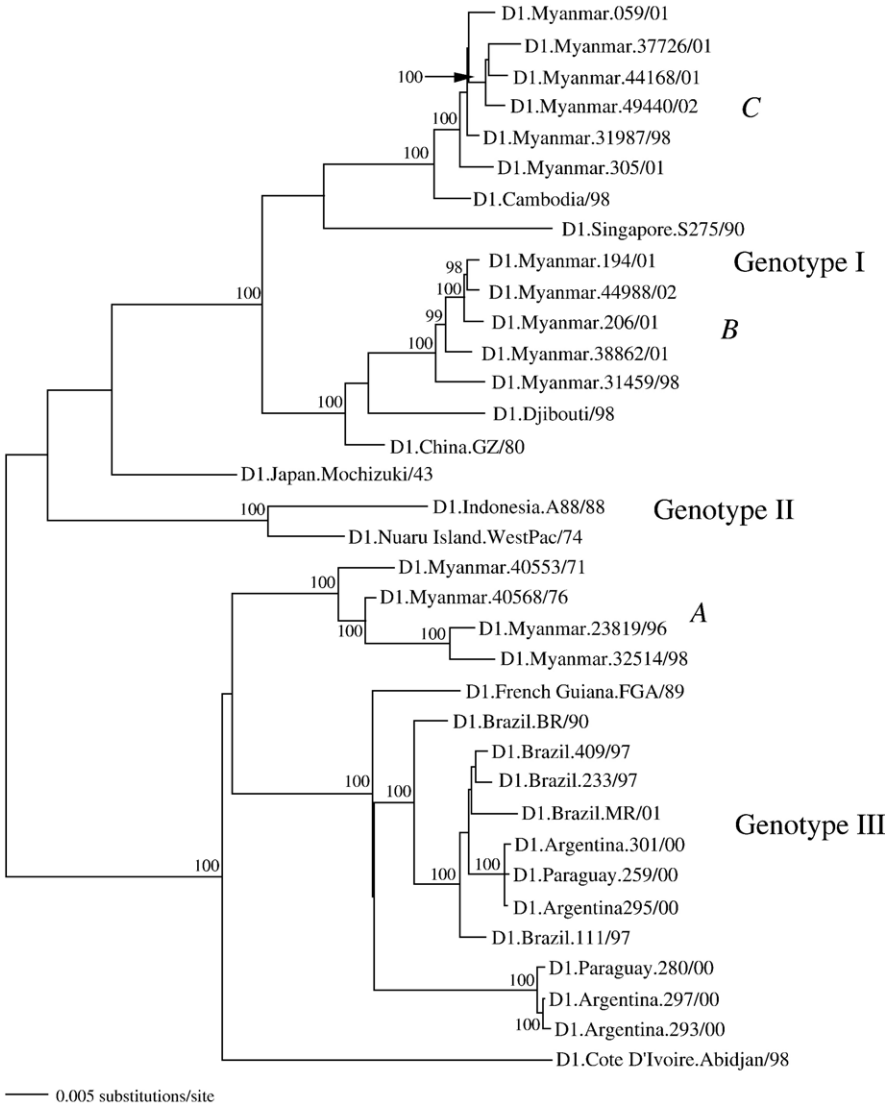


Fig. 1. Maximum likelihood phylogenetic tree of 35 complete genomes (ORF regions) of DENV-1. The three genotypes of DENV-1 are indicated, as are the three clades of Myanmar viruses (A, B, and C) within these genotypes. Bootstrap support values of >95% are shown. All horizontal branch lengths are drawn to scale and the tree is mid-point rooted for clarity only.

The selection pressures acting on each gene of DENV-1, as measured by average $d_N/d_S$ values among all 35 genomes, are given in Table 1. The C and NS2A genes were characterized by the highest $d_N/d_S$ values (and NS2A exhibited the most nonsynonymous diversity), while the M, NS3, and NS4B genes were the most constrained. However, there were no significant differences ($P \geq 0.05$; $\chi^2$ test) between the $d_N/d_S$ ratios for any of the individual genes. Interestingly, although the E gene is the main target for neutralizing antibodies (reviewed in Roehrig, 1997), this gene was characterized by an intermediate level of non-synonymous variation and in all cases the low $d_N/d_S$ ratios indicating purifying selection were the main evolutionary process acting on DENV-1. Similar results were obtained in the codon and lineage-specific analyses of $d_N/d_S$ (results not shown, available from authors). Under these analyses, evidence for weak ($d_N/d_S = 1.986$), positive selection was found only at amino acid positions 127, 135, and 669 in the NS5 protein ($P \leq 0.002$) and none of these changes distinguished all extinct from all contemporary clades of

Table 1
Ratio of nonsynonymous to synonymous nucleotide changes per site ($d_N/d_S$) in the genes of DENV-1

| Protein | No. amino acids | $d_N$ | $d_S$ | $d_N/d_S$ |
|---|---|---|---|---|
| Capsid | 114 | 0.012 | 0.142 | 0.085 |
| Membrane | 166 | 0.007 | 0.296 | 0.024 |
| Envelope | 495 | 0.010 | 0.266 | 0.038 |
| NS1 | 352 | 0.013 | 0.292 | 0.045 |
| NS2A | 218 | 0.020 | 0.321 | 0.062 |
| NS2B | 130 | 0.008 | 0.312 | 0.026 |
| NS3 | 619 | 0.006 | 0.254 | 0.024 |
| NS4A | 150 | 0.013 | 0.284 | 0.046 |
| NS4B | 249 | 0.007 | 0.279 | 0.025 |
| NS5 | 899 | 0.008 | 0.267 | 0.030 |

DENV-1, although a Tyr at position 127 did distinguish the clade B sequences. Among the Myanmar viruses, 69 of the nonsynonymous nucleotide changes resulted in clade-specific amino acid changes. Eighteen of the 40 amino acid changes that distinguished clade B and C viruses from the earlier, extinct, clade A viruses were non-conservative (data available from authors).

Nucleotide changes were observed at five sites in the 5′ UTR of these viruses; positions 23, 24, 38, 49, and 65. A C at position 49 distinguished the extinct clade (A) from present viruses (clades B and C), but the change had no effect on the predicted secondary structure of the 5′ UTR RNA (data not shown). A single virus isolate from human sera (D1.Myanmar.49440/02) had changes at nucleotide positions 23 (A → C) and 24 (C → A), but these did not result in significant changes in either the predicted secondary structure of the 5′ UTR or in ΔG values (−17.8 to −18.4 kcal/mol). A single nucleotide change (A → G) at position 65 in the 5′ UTR of a virus isolate from a mosquito (D1.Myanmar.059/01) had no significant effect on ΔG values (changed from −17.8 to −17.1 kcal/mol), but it introduced a second stem-loop into the 5′ end of the UTR (data not shown). The putative genome cyclization sequence UCAAUAUG, between nucleotides 38 and 46 of the C gene (Hahn et al., 1987), was conserved in all viruses.

Nucleotide changes were detected at 34 sites in the 3′ UTR between the end of the ORF and CS2 (nucleotides 10,274–10,614) (data available from authors). There was significantly more nucleotide variation (twenty-four variable sites; $P < 0.05$, $\chi^2$ test) in the first 100 nucleotides than in any other region of the 3′ UTR. The nucleotide sequences of the 3′ UTR of all clade A viruses were the same in this region except at sites 10,355 and 10,476. The change at nt 10,355 distinguished the 1971 and 1976 viruses from the more recent (1996 and 1998) isolates in this, now extinct, clade. The nucleotide sequences of the 3′ UTR of all clade C viruses also resembled each other in this region except D1.Myanmar.37226/01, which had an A at position 10,305 rather than G, and D1.Myanmar.059/01 which had a U at position 10,434 rather than a G. These changes resulted in neither a change in predicted secondary structure of the 3′ UTR nor in a significant change in ΔG values. In contrast, there was extensive variation in this region of the 3′ UTR of clade B viruses including a nine nucleotide deletion in the gene of D1.Myanmar.31459/98. Nucleotide changes at eight sites distinguished clades B and C from clade A viruses. If the nucleotide changes which distinguished the clade B and C viruses from the extinct viruses (nt. 10,283 C-U; 10,285 C-U; 10,288 G-A; 10,317 A-G; 10,335 U-A; 10,440 G-A; 10,466 C-U; 10,537 G-A) were made to the nucleotide sequence of the clade A virus D1.Myanmar.40553.71, the lowest energy secondary RNA structures were different (Figs. 2A and B), but the ΔG values for the optimum secondary structure of the RNAs were similar (−78.7 kcal/mol for the extinct viruses and −75.6 kcal/mol for the structure with the changes characteristic of the clade B and C viruses).

## Discussion

Phylogenetic analyses of the sequences of DENV-1 genomes confirmed that three distinct clades (from two genotypes) have circulated in Myanmar since the early 1970s, when dengue hemorrhagic fever was recognized first in that country (Ming et al., 1974), and that one of these clades (clade A, genotype III) became extinct after 1998 to be replaced by viruses from two other clades (B and C) assigned to genotype I. In both cases, the replacement genotype I clades appeared to have origins in Asia. In particular, the “new” clade B viruses were related to a strain of DENV-1 isolated in Guangzhou, in southern China, in 1980. Given the extensive commercial links between China and Myanmar, it is possible that this clade was introduced into Myanmar from China, although the virological record to confirm this does not exist. Similarly, the “new” clade C viruses were related a strain of DENV-1 recovered in Cambodia in 1998, the year that this clade of virus was detected first in Myanmar. While Myanmar has less direct contact with Cambodia than with China, there are extensive

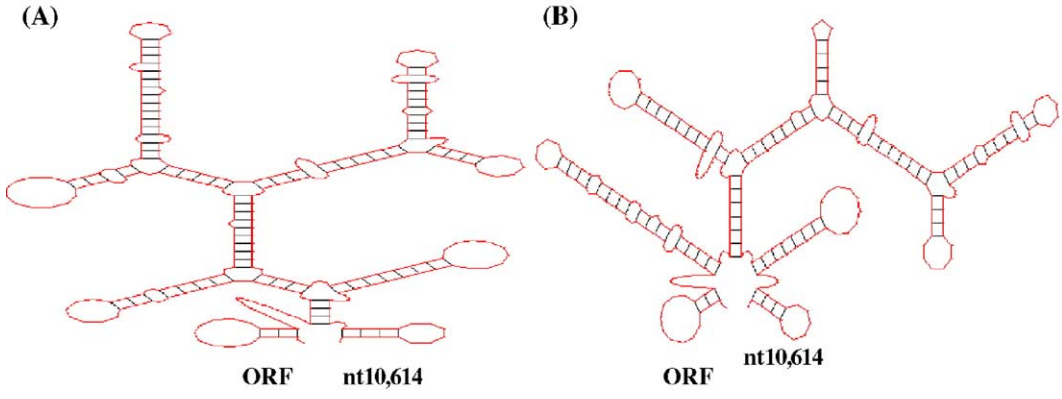


Fig. 2. The predicted secondary structure of (A) the 3′ region of D1.Myanmar.40553/71 and (B) the same region with the nucleotide changes characteristic of the post-1998 strains (clades B and C) of DENV-1 at positions 10,284, 10,286, 10,289, 10,318, 10,336, 10,441, 10,467, and 10,538.

indirect links between Cambodia and Myanmar through China and Thailand. It also is possible that the viruses were introduced into Myanmar from other sources (notably Thailand), but the available virological data are inadequate to identify this source. Although recombinant DENVs composed of "Myanmar" and "Chinese" strains of DENV-2 were isolated from a mosquito collected in Yangon in 2000 (Craig et al., 2003), no evidence of recombination was found in the genomes of the Myanmar DENV-1 sampled here as all genes produced congruent trees. Hence, there is no evidence to suggest that the viruses circulating most recently in Myanmar had arisen from the, now extinct, clade A viruses by recombination.

The most important question addressed by our work is whether the extinction of one clade of DENV-1 and its replacement by two clades new to Yangon in Myanmar was due to direct competition between the three viral populations. Such a scenario is possible as all three clades were co-circulating in Yangon in 1998. During the intense virus transmission accompanying the dengue outbreak in Myanmar in 1998 (Thu et al., 2004), in which similar numbers of DENV-1, 2, and 3 were recovered from patients, the two new DENV-1 genotypes (clades B and C) may have outcompeted the older strains, as well as other DENV serotypes, so that by 2000, the year before the next major dengue outbreak, the two new clades of DENV-1 represented those viruses most frequently recovered from patients (Thu et al., 2004). An alternative explanation was that some stochastic event associated with low rates of virus transmission during the inter-epidemic interval, 1999–2000, resulted in the elimination of the older, clade A, viruses. Similar rapid extinctions and appearances of new genotypes of DENV-2 and 3 have been described in Thailand during inter-epidemic periods (Sittisombut et al., 1997; Wittke et al., 2002).

To test between these competing hypotheses, we undertook an extensive analysis of the selection pressures acting on complete genomes of DENV-1, examining these pressures in genes, along lineages, and at individual codons. An analysis of the open reading frames of 35 genomes of DENV-1, including strains sampled from Myanmar over the period 1971–2002, failed to find evidence of strong positive selection contributing to the extinction of pre-1998 strains of virus or to the appearance of the two new clades of viruses. Indeed, evidence for only weak positive selection in one protein–NS5–was found, and the overriding evolutionary process was that of strong purifying (negative) selection. This notwithstanding, the selected sites in NS5 (NS5-127, 135, and 669) were all located close to two critical functional determinants in that protein. Amino acids NS5-127 and NS5-135 are in motif III of the RNA cap (nucleoside-2′-*O*-)-methyltransferase portion of NS5 and straddle amino acids NS5-131 and 132 which interact with S-adenosyl-L-methionine to aid capping of the viral RNA (Egloff et al., 2000). The second site found to be under selective pressure (NS5-669) was adjacent to a potential nucleoside triphosphate binding motif $^{661}$GDD$^{663}$ in motif VI (Koonin, 1991) of the polymerase component at the C-terminal end of NS5. However, the amino acid residues at these sites were not common to all members of the extinct clade of viruses or to either of the two clades that replaced it.

As well as these, putatively, selected sites, there were also a number of non-conservative amino acid changes that distinguished the extinct from the contemporary clades of virus. Too little is known of structure–function relationships in DENV proteins to assign effects to each of these changes, but for a number, it was possible to speculate on a possible effect of such changes on the fitness of viral populations. The Ser–Gly change at residue 70 of the core (C) protein occurred beside a Trp residue (C-69), which is conserved in all flaviviruses, and at the end of an hydrophobic alpha helix (alpha 3) which may interact with the lipid membrane of these viruses (Ma et al., 2004). Amino acid deletions from the alpha 3 helix region of C of tick-borne encephalitis virus resulted in non-infectious virions (Kofler et al., 2002). The Asn–Asp change at E-37 of the envelope protein occurred in a cross-reactive, immunodominant, serological epitope (Aaskov et al., 1989) and that at E-369 (Thr–Ala) was in a region of Domain III of the E gene which influences cell trophism and is a target for neutralizing antibodies (reviewed in Roehrig, 1997). The change from His to Tyr at NS1-111 occurred at the first position of a linear DENV-1-specific serological epitope $^{111}$HKYSWK$^{116}$ (Yao et al., 1995). The changes in the NS3 protein which distinguished the current viruses from the extinct clade fell in both the N-terminal serine-protease domain (NS1-44, 85, and 112) among the catalytic triad of amino acids (NS1-51, 75, and 135) and adjacent (NS3-293) to the $^{285}$DEAH$^{288}$ motif in the C-terminal helicase region in NS3. DENV-2 with a Val in place of Phe at NS3-283 was temperature sensitive, replicating at 33 °C but not at 37 °C (Matusan et al., 2001).

In contrast, there were no clade-specific nucleotide substitutions or changes in the predicted secondary structure of the 5′ UTR of these viruses that were likely to have contributed to the extinction of clade A viruses or have given a fitness advantage to all clade B and C viruses. While Cahour et al. (1995) found a 6-nt deletion between positions 62 and 68 of 5′ UTR destroyed DENV-4 infectivity, the change observed at position 65 of a single mosquito isolate, D1.Myanmar.059/01, was a single nucleotide transition (A–G) rather than a deletion. The nucleotide sequence of the cyclization region (CS1, Hahn et al., 1987; Khromykh et al., 2001) was conserved in the C gene of all viruses and in the 3′ UTR of both genomes in which this region could be sequenced. The sequences of the CS2 and RCS2 region of the 3′ UTR of all viruses were also conserved. The observation that the nine nucleotide deletion in the 3′ UTR of D1.Myanmar.31459/98 occurred in the same area (nt 10,297–10,305) as a ten nucleotide deletion in the 3′ UTR of DENV-2 (Shurtleff et al., 2001) together with the hypervariability in the sequence of the 100 nucleotides at

the 5′ end of the 3′ UTR of DENV-1 suggests that this region may act as a "spacer" to separate the coding role of the ORF from the replicative role of the UTR. Although the clade-specific nucleotide changes in the 3′ UTR changed the predicted secondary structure of this region, the ΔG values for the structures were so similar that the changes were unlikely to have resulted in significant fitness advantage. These analyses focused on untranslated regions of RNA and lengths of RNA which could be analyzed with some confidence by secondary structure prediction programs like Mfold. Recent studies have suggested that there may be "genome scale secondary structures" (GORS) in viruses such as GB virus-C and hepatitis C that can be destroyed by synonymous nucleotide changes, or changes in UTRs (Simmonds et al., 2004). It is possible that clade-specific, synonymous, nucleotide differences between the clade A and the clades B and C DENV-1 conferred a fitness advantage to the clades B and C viruses, but there is currently no data to support such a contention.

While this study has identified a number of DENV-1 clade-specific changes, the study of which may provide insights into the structure and function of DENV, the issue of the relative fitness of dengue virus populations is complex and crucially dependent on the particular ecological environment where the sequence evolution took place. In particular, it is possible that the presence of multiple DENV serotypes will result in complex patterns of cross-immunity which might determine which clades survive and which become extinct (Thu et al., 2004). It is difficult to conceive of laboratory experiments that might provide meaningful insights into the fitness of different dengue viruses in replicative cycles in nature. The primary and/or major sites of replication of dengue viruses in humans have not been defined unambiguously and so it is unclear what cells or cell lines might be appropriate substrates for experiments relevant to the human condition. Furthermore, we do not have unpassaged virus in sufficient quantities for comparative studies and there are no mosquitoes from 1996 to 1998 from Yangon for studies of virus infection and transmission.

Collectively, our data suggest that positive natural selection, at either the amino acid, nucleotide, or RNA secondary structure levels, was not responsible for the extinction of the pre-1998 clade A strains of DENV-1 in Myanmar. Moreover, we found no difference in the overall rates of nucleotide substitution between the genotype I (clades B and C) and genotype III (clade A) viruses, indicating that they were subjected to broadly similar evolutionary processes before and after the clade extinction. As such, this study of DENV-1 genomes provides further evidence that the process of lineage extinction and replacement that is being observed in dengue virus populations may often be due to the seasonal ("dry season" and "monsoon") variations in the size of the mosquito vector populations. More importantly, these dramatic changes in the genetic structure of dengue virus populations have preceded significant changes in patterns of disease on at least two occasions (Thu et al., 2004; Wittke et al., 2002) and indicate that studies of DENV evolution should not ignore the associated ecological background.

## Materials and methods

### *Viruses*

Viruses were recovered from acute phase sera from dengue patients admitted to the Yangon Children's Hospital and from female *A. aegypti* mosquitoes collected from within a radius of 100 m of the homes of patients (Thu et al., 2004). DENV-1 isolates recovered in Rangoon (Yangon) in 1971 and 1976 were kindly provided by the Centres for Disease Control, Atlanta, USA. The nucleotide sequences of genomes from 15 DENV-1 strains from Myanmar determined here were combined with 20 complete DENV-1 genomes taken from GenBank. All virus strains used for this study are listed in Table 2. The following nomenclature has been used to identify isolates—virus serotype.country.strain/year of isolation.

### *Reverse transcription and polymerase chain reaction (RT-PCR)*

RNA was extracted from the virus isolates using QIAamp Viral RNA extraction columns (Qiagen, Australia) as per manufacturer's instructions. The RNA and random hexamer primers (Roche, Germany) were heated at 72 °C for 10 min then cooled on ice for 1 min. Avian myeloblastosis virus (AMV) reverse transcriptase (Roche) and deoxynucleoside triphosphates (Roche) were added and the mixture incubated at 55 °C for 10 min followed by 45 °C for 60 min (Wittke et al., 2002). The resultant cDNA was stored at −20 °C until required.

PCR was performed by adding cDNA to a mixture containing 30 pmol/μl sense and anti-sense primers (Table 3), Expand DNA polymerase (a mixture of Taq and Pwo polymerases; Roche), and deoxynucleotide triphosphates (Roche). The cycling conditions used were 93 °C for 2 min, followed by 40 cycles of 92 °C for 40 s, 56 °C for 40 s with extension at 68 °C and the extension times increasing from 1 min 30 s to 1 min 50 s, to 2 min and to 2 min 20 s at intervals of 10 cycles with a final extension for 10 min at 68 °C. cDNA was recovered following electrophoresis of the PCR product in a 1.5% w/v agarose Tris-acetate/EDTA gel by excising the band of interest and purifying it using a High Pure PCR purification kit (Roche) according to the manufacturer's instructions.

### *Cloning*

Because sufficient cDNA, corresponding to the 3′ UTR, could not be obtained for sequencing, this region of the

Table 2
Strains of DENV-1 used in this study

| Virus strain | Year of isolation | Host, country of origin | Passage history | GenBank accession number |
|---|---|---|---|---|
| D1.Argentina.293/00 | 2000 | Argentina | Unknown | AY206457 |
| D1.Argentina.295/00 | 2000 | Argentina | Unknown | AF514885 |
| D1.Argentina.297/00 | 2000 | Argentina | Unknown | AF514889 |
| D1.Argentina.301/00 | 2000 | Argentina | Unknown | AF514876 |
| D1.Brazil.BR/90 | 1990 | Brazil | Unknown | AF226685 |
| D1.Brazil.111/97 | 1997 | Brazil | Unknown | AF311956 |
| D1.Brazil.233/97 | 1997 | Brazil | Unknown | AF311958 |
| D1.Brazil.409/97 | 1997 | Brazil | Unknown | AF311957 |
| D1.Brazil.MR/01 | 2001 | Brazil | Unknown | AF513110 |
| D1.Cambodia/98 | 1998 | Cambodia | Unknown | AF309641 |
| D1.China.GZ/80 | 1980 | China | Unknown | AF350498 |
| D1.CoteD'Ivorie.Abidjan/98 | 1998 | Ivory Coast | Unknown | AF298807 |
| D1.Djibouti/98 | 1998 | Djibouti | Unknown | AF298808 |
| D1.French Guiana.FGA/89 | 1989 | French Guiana | Unknown | AF226687 |
| D1.Indonesia.A88/88 | 1988 | Indonesia | Unknown | ABO74761 |
| D1.Japan.Mochizuki/43 | 1943 | Japan | Unknown | ABO74760 |
| D1.Nauru Island.WestPac/74 | 1974 | Nauru | Unknown | U88535 |
| D1.Paraguay.259/00 | 2000 | Paraguay | Unknown | AF514883 |
| D1.Paraguay.280/00 | 2000 | Paraguay | Unknown | AF514878 |
| D1.Singapore.S275/90 | 1990 | Singapore | Unknown | M87512 |
| D1.Myanmar.40553/71 | 1971 | Human, Myanmar | 1 passage in C6-36 cells | – |
| D1.Myanmar.40568/76 | 1976 | Human, Myanmar | 1 passage in C6-36 cells | – |
| D1.Myanmar.23819/96 | 1996 | Human, Myanmar | 1 passage in C6-36 cells | AY589692 |
| D1.Myanmar.31459/98 | 1998 | Human, Myanmar | 1 passage in C6-36 cells | AY588272 |
| D1.Myanmar.31987/98 | 1998 | Human, Myanmar | 1 passage in C6-36 cells | AY588273 |
| D1.Myanmar.32514/98 | 1998 | Human, Myanmar | 1 passage in C6-36 cells | AY600860 |
| D1.Myanmar.37726/01 | 2001 | Human, Myanmar | 1 passage in C6-36 cells | AY606062 |
| D1.Myanmar.38862/01 | 2001 | Human, Myanmar | 1 passage in C6-36 cells | AY618210 |
| D1.Myanmar.44168/01 | 2001 | Human, Myanmar | 1 passage in C6-36 cells | AY618211 |
| D1.Myanmar.44988/02 | 2002 | Human, Myanmar | 1 passage in C6-36 cells | – |
| D1.Myanmar.49440/02 | 2002 | Human, Myanmar | 1 passage in C6-36 cells | – |
| D1.Myanmar.059/01 | 2001 | Mosquito, Myanmar | 1 passage in C6-36 cells | – |
| D1.Myanmar.194/01 | 2001 | Mosquito, Myanmar | 1 passage in C6-36 cells | – |
| D1.Myanmar.206/01 | 2001 | Mosquito, Myanmar | 1 passage in C6-36 cells | – |
| D1.Myanmar.305/01 | 2001 | Mosquito, Myanmar | 1 passage in C6-36 cells | – |

DENV-1 genome was amplified by RT-PCR using the protocol above and primers shown in Table 3. To enable cloning into the pGEM-TEasy plasmids, ATP extensions were added to the cDNA by incubating it with 10 μM ATP (Roche) and 5 U of Taq polymerase (Roche) at 72 °C for 30 min. The "A-tailed" cDNA was purified using a High Pure gel extraction kit (Roche) according to the manufacturer's instructions, ligated into pGEM-T Easy plasmids (Promega, Australia), and used to transform maximum efficiency DH5-α *E. coli* (Invitrogen, USA) according to the manufacturer's instructions. Plasmids were purified from individual bacterial colonies grown on LB (Luria-Bertani media) supplemented with 100 μg of ampicillin/ml, 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG), and 80 μg/ml of 5-bromo-4-chloro-3-indoyl-β-D-galactosidase/ml using QIAprep Miniprep kits (Qiagen) according to the manufacturer's instructions.

*Nucleotide sequencing and data analysis*

Fifty to 100 ng of purified cDNA or plasmid was sequenced using 1 μl of 3.2 pmol/μl oligonucleotide primers (Table 3) and the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Version 3.1). The product was purified using Dye-EX spin columns (Qiagen) according to the manufacturer's instructions and the samples were sequenced in an Applied Biosystems 3730XL Nucleotide Sequencer at the Australian Genome Research Faculty, University of Queensland.

Consensus nucleotide sequences for each isolate were compiled using SeqMan (Version 4.03 DNASTAR Inc., USA). The nucleotide and derived amino acid sequences of the ORFs of the genomes were aligned using EclustalW software from the Australian National Genome Information Service (ANGIS). This resulted in a final alignment (complete ORF only) of 35 sequences, 10,176 bp in length for sequence analysis. A maximum likelihood (ML) tree for these data was constructed using the GTR + Γ + I model of nucleotide substitution (parameter values available from the authors on request). To assess the robustness of particular clades on the phylogenetic tree, a bootstrap analysis was performed using 1000 replicate neighbor-joining (NJ) trees estimated under the ML substitution matrix described above. All these analyses were conducted using the PAUP*

Table 3
Oligonucleotide primers used for amplification and sequencing of the genomes of DENV-1 strains

| Primer name | Sequence (5′–3′) | Nt. position in the genome[a] | Primer name | Sequence (5′–3′) | Nt. position in the genome* |
|---|---|---|---|---|---|
| D1-1bF[b] | AGTTGTTAGTCTACGTGGACC | 1–21 | D1-5958F | GATCACGCTCATTGGACAGAAG | 5966–5987 |
| D1-202R[c] | GGCCTGAGAGCAATCCTTTTGAG | 181–205 | D1-6073R | TCCCCGTCTATTGCTGCACTC | 6066–6086 |
| D1-764F | CAAATACAAAAGTGGAGACCTGGGC | 791–816 | D1-6177R | GAGTACTGGAAGCCTTCTGAGGC | 6167–6191 |
| D1-789F | CTCTGAGACACCCAGGATTCAC | 780–801 | D1-7026F | CAGGCAGCTATATTGATGGGAC | 7026–7047 |
| D1-832R | TGCTAGAAAAAGGGCTATCACCGTG | 812–836 | D1-7192R | GCAGTCCAGGTCCAATTATG | 7174–7194 |
| D1-1063R | GTCTTCAAGAGTTCAATGTC | 1044–1063 | D1-8023R | GTTGGGTTCGGAGAGGA | 8019–8035 |
| D1-1573F | CGGGGGCTTCAACATCCCAAGAGAC | 1599–1623 | D1-8285F | CACAATGGCTCACAGGAAACC | 8288–8308 |
| D1-1704F | GAGCGACAGAAATCCAAACG | 1706–1725 | D1-8962F | GAAGTCGTGCAATATGGTACATG | 8967–8989 |
| D1-1847R | CAGCCACTTCCTTCTCAAC | 1826–1845 | D1-8990F | CACTCAGTGGAGTGGAAGG | 8998–9016 |
| D1-2267F | GCGGTGTTTCCTGGACCATG | 2277–2296 | D1-9186R | CTGTTATTCTTGTATCCCATCC | 9165–9186 |
| D1-2310F | GACATGGCTAGGATTAAACTC | 2313–2333 | D1-9240F | GCCCTACTGGCCACGTCAGTC | 9244–9264 |
| D1-2467R | GACTTCATTGGTGACAAAAATGCCGC | 2445–2467 | D1-9525R | CAGTCATCTCCGCTGATTGCC | 9505–9525 |
| D1-2554F | GGGAGGAGGGTGTGTGTG | 2557–2574 | D1-9536F | CAGCGGAGATGACTGCGTGGT | 9536–9553 |
| D1-3350R | GATCTGCAGCACCATTCATGG | 3336–3356 | D1-9610R | GAAGGTTCCCATTGTGGTATG | 9411–9611 |
| D1-3789F | CTAGTGGCATCCGTGGAGCTACC | 3785–3797 | D1-9670R | GATTGCGGCATGGCACCAC | 9652–9670 |
| D1-3972R | CCATAGCCATTGTCTTCCATG | 3953–3973 | D1-10115R | CAAGCCTATCAAGGATCCACAC | 10,103–10,124 |
| D1-4610F | GGTCCAAGATGATGGAACCAT | 4617–4638 | D1-10145F | GGCCACCAACATACAAGTGGCC | 10,145–10,166 |
| D1-4663F | CCACACAATGTGGCATGTCACCAG | 4642–4665 | D1-10376R | CACAGGCAGCATCGCTCTTACC | 10,358–10,379 |
| D1-4728F | GACCTGATCTCATATGGAGGAG | 4725–4746 | D1-10653R | GTCTCTCCCAGCGTCAATATGCT | 10,632–10,654 |
| D1-4995F | GCCATAAATCCCGCTAAAGCATC | 4997–5029 | D1-10734R | GACCTGTTGATTCAACAGCAC | 10,713–10,734 |
| D1-4793R | CTGCACTTCTTCTCCCGTGTTC | 4783–4804 | T7† | TAATACGACTCACTATAGGG | pGEM-T plasmid |
| D1-5466R | CTGTGGAAAGGCCTCCACCGATCC | 5451–5474 | SP6‡ | ATTTAGGTGACACTATAGAA | pGEM-T plasmid |
| D1-5728R | GTGACGACATAATCCCAGTCG | 5707–5729 | | | |

[a] Numbering from Ishak et al. (2001) for D1.Japan.Mochizuki/43.
[b] Sense primer.
[c] Anti-sense primer.

package (Swofford, 2003). Secondary structures in viral RNA were predicted using programs Mfold and Plotfold (Squiggles) from ANGIS.

To determine the nature of the selection pressures acting on the complete genomes of DENV-1, the ratio of nonsynonymous ($d_N$) to synonymous ($d_S$) nucleotide changes per site was calculated, with $d_N > d_S$ indicative of positive selection. First, the mean $d_N/d_S$ ratio for each gene in 35 complete DENV-1 genomes was measured using the pairwise method of Nei and Gojobori (1986) as implemented in the MEGA2 package (Kumar et al., 2001). Next, to obtain both codon and lineage-specific measures of selection pressures on each gene (including a test for positive selection), the maximum likelihood method in the PAML sequence analysis package (program CODEML) was employed (Yang, 1999; Yang et al., 2000). This involved the comparison of four models of codon evolution which differed in how $d_N$ and $d_S$ vary among codons or lineages. To analyze selection pressures at individual codons, the M7 and M8 models were compared; the former allows individual codons to take on one of 10 categories of $d_N/d_S$, all estimated from the data, but where no category has $d_N/d_S > 1.0$ so that only neutral evolution is allowed. In contrast, M8 allows positive selection by including an 11th category of codons at which $d_N/d_S$ can exceed 1.0. Evidence for positive selection is obtained if M8 rejects M7 under a likelihood ratio test and at least one category of codons in M8 has a $d_N/d_S$ ratio > 1. To analyze selection pressures along each lineage of the DENV-1 complete genome phylogeny, we compared the M0 model, in which each lineage has an identical $d_N/d_S$ ratio, to the FR ("free ratio") model, in which lineages are allowed to take on individual values of $d_N/d_S$. For M0 and FR, $d_N/d_S$ was estimated from the data and their support was assessed using a likelihood ratio test. Finally, the overall rate of nucleotide substitution of complete genomes from genotypes I and III of DENV-1 was estimated. This analysis was undertaken using the Bayesian Markov Chain Monte Carlo (MCMC) method available in the BEAST package (http://www.evolve.zoo.ox.ac.uk/beast/) and was based on the comparison of branch lengths from viruses sampled at different times. For genotype I, 16 sequences were available (including D1.Japan.Mochizuki/43) and the effective sample size was 880.986. In the case of genotype III, 17 sequences were available and the effective sample size was 512.814. Insufficient data were available to conduct similar analyses on genotype II viruses. As with the ML trees inferred above, the GTR + Γ + I model of nucleotide substitution was employed and a constant viral population size was assumed. In both cases, a burn-in of 200,000 and a final chain length of 2 million were employed.

## Acknowledgments

This study was supported by the Wellcome Trust, the World Health Organization, and the Lee Foundation. The

authors wish to thank the staff of Yangon Children's Hospital and of the Virology and Medical Entomology Research Divisions, Department of Medical Research (Lower Myanmar), for their assistance in collecting material used in this study.

## References


Aaskov, J.G., Geysen, H.M., Mason, T.J., 1989. Serologically defined linear epitopes in the envelope protein of dengue 2 (Jamaica strain 1409). Arch. Virol. 105, 209–221.

Bennet, S.N., Holmes, E.C., Chirivella, M., Rodriguez, D.M., Beltran, M., Vordam, V., Gubler, D.J., McMillan, W.O., 2003. Selection-driven evolution of emergent dengue virus. Mol. Biol. Evol. 20, 1650–1658.

Bosio, C.F., Fulton, R.E., Salasek, M.L., Beaty, B.J., Black, W.C., 2000. Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito *Aedes aegypti*. Genetics 156, 687–698.

Cahour, A., Pletnev, A., Vazeille-Falcoz, M., Rosen, L., Lai, C.-J., 1995. Growth restricted dengue virus mutants containing deletions in the 5′ non-coding region of the RNA genome. Virology 207, 68–76.

Cologna, R., Armstrong, P.M., Rico-Hesse, R., 2005. Selection for virulent dengue viruses occurs in humans and mosquitoes. J. Virol. 79, 853–859.

Craig, S., Thu, H.M., Lowry, K., Wang, X.F., Holmes, E.C., Aaskov, J., 2003. Diverse dengue type 2 virus populations contain recombinant and both parental viruses in a single mosquito host. J. Virol. 77, 4463–4467.

Egloff, M.-P., Benarroch, D., Selisko, B., Ronette, J.-L., Canard, B., 2000. An RNA cap (nucleoside-2′-*O*-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. EMBO J. 21, 2757–2768.

Gubler, D.J., 1998. Dengue and dengue haemorrhagic fever. Clin. Microbiol. Rev. 11, 480–496.

Gubler, D.J., Rosen, L., 1976. Variation among geographic strains of *Aedes albopictus* in susceptibility to infection with dengue viruses. Am. J. Trop. Med. Hyg. 25, 318–325.

Hahn, C.S., Hahn, Y.S., Rice, C.M., Lee, E., Dalgarno, L., Strauss, E.G., Strauss, J.H., 1987. Conserved elements in the 3′ untranslated regions of flavivirus RNAs and potential cyclization sequences. J. Mol. Biol. 198, 33–41.

Halstead, S.B., 1970. Observations related to the pathogenesis of dengue haemorrhagic fever: IV. Hypothesis and discussion. Yale J. Biol. Med. 42, 350–362.

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

Holmes, E.C., Twiddy, S.S., 2003. The origin, emergence and evolutionary genetics of dengue virus. Infect. Genet. Evol. 3, 19–28.

Ishak, H., Takegami, T., Kamimura, K., Funada, H., 2001. Comparative sequences of two type 1 dengue virus strains possessing different growth characteristics in vitro. Microbiol. Immunol. 45, 327–331.

Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C., 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J. Mol. Evol. 54, 152–161.

Khromykh, A.A., Meka, H., Guyatt, K.J., Westaway, E.G., 2001. Essential role of cyclization sequences in flavivirus RNA replication. J. Virol. 75, 6719–6728.

Kofler, R.M., Heinz, F.X., Mandl, C.W., 2002. Capsid protein C of tick-borne encephalitis virus tolerates large internal deletions and is a favorable target for attenuation of virulence. J. Virol. 76, 3534–3543.

Koonin, E.V., 1991. The phylogeny of RNA-dependent RNA polymerases of positive strand RNA viruses. J. Gen. Virol. 72, 2197–2206.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17, 1244–1245.

Ma, L., Jones, C.T., Groesch, T.D., Kuhn, R.J., Post, C.B., 2004. Solution structure of dengue virus capsid protein reveals another fold. Proc. Natl. Acad. Sci. U.S.A. 101, 3414–3419.

Matusan, A.E., Pryor, M.J., Davidson, A.D., Wright, P.J., 2001. Mutagenesis of the dengue virus type 2 NS3 protein within and outside helicase motifs: effects on enzyme activity and virus replication. J. Virol. 75, 9633–9643.

Ming, C.K., Thein, S., Thaung, U., Tin, U., Myint, K.S., Swe, T., Halstead, S.B., Diwan, A.R., 1974. Clinical and laboratory studies on haemorrhagic fever in Burma, 1970–72. Bull. World Health Organ. 51, 227–235.

Monath, T.P., 1994. Dengue: the risk to developed and developing countries. Proc. Nat. Acad. Sci. U.S.A. 91, 2395–2400.

Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.

Rice, C.M., 1996. Flaviviridae: the viruses and their replication. In: Fields, B.N., Knipe, D.N., Howley, P.M. (Eds.), Fields Virology, 3rd ed. Lippincott-Raven, Philadelphia, pp. 931–959.

Rico-Hesse, R., Harrison, L.M.M., Sala, R.A., Tovar, D., Nisalak, A., Ramos, C., Boshell, J., Mesa, M.T.R., Nogueira, R.M.R., Rosa, A.T.D., 1997. Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. Virology 230, 244–251.

Rigau-Perez, J.G., Clark, G.G., Gubler, D.J., Reiter, P., Sanders, E.J., Vorndam, A.V., 1998. Dengue and dengue haemorrhagic fever. Lancet 352, 971–977.

Roehrig, J.T., 1997. Immunochemistry of dengue viruses. In: Gubler, D.J., Kuno, G.K. (Eds.), Dengue and Dengue Haemorrhagic Fever. CAB International, New York, pp. 199–220.

Shurtleff, A.C., Beasley, D.W.C., Chen, J.J.Y., Ni, H., Suderman, M.T., Wang, H., Xu, R., Wang, E., Weaver, S.C., Watts, D.M., Russell, K.L., Barrett, A.D.T., 2001. Genetic variation in the 3′ non-coding region of dengue viruses. Virology 281, 75–87.

Simmonds, P., Tuplin, A., Evans, D.J., 2004. Detection of genome-scale ordered RNA structures (GORS) in genomes of positive-stranded RNA viruses: implications for virus evolution and host persistence. RNA 10, 1337–1351.

Sittisombut, N., Sistayanarain, A., Cardose, M.J., Salminen, M., Damrongdachakul, S., Kalayanrooj, S., Rojanasupot, S., Supawadee, J., Maneekarn, N., 1997. Possible occurrence of a genetic bottleneck in dengue serotype 2 viruses between the 1980 and 1987 epidemic seasons in Bangkok, Thailand. Am. J. Trop. Med. Hyg. 57, 100–108.

Steinhauer, D.A., Domingo, E., Holland, J.J., 1992. Lack of evidence for proofreading mechanisms associated with RNA virus polymerases. Gene 122, 281–288.

Strickman, D., Kittayapong, P., 2003. Dengue and its vector in Thailand: calculated transmission risk from total pupal counts of *Aedes aegypti* and association of wing length measurements with aspects of the larval habitats. Am. J. Trop. Med. Hyg. 68, 209–217.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis using Parsimony (* and other methods), Version 4. Sinauer Associates, Sunderland, MA.

Thu, H.M., Lowry, K., Myint, T.T., Shwe, T.N., Han, A.M., Khin, K.K., Thant, K.Z., Thein, S., Aaskov, J.G., 2004. Myanmar dengue outbreak associated with displacement of serotypes 2, 3 and 4 by dengue 1. Emerg. Infect. Dis. 10, 593–597.

Twiddy, S.S., Woelk, C.H., Holmes, E.H., 2002. Phylogenetic evidence for adaptive evolution of dengue viruses in nature. J. Gen. Virol. 83, 1679–1689.

Vaughn, D.W., Green, S., Kalayanarooj, S., Innis, B.L., Nimmannitya, S., Suntayakorn, S., Endy, T.P., Raengsakulrach, B., Rothman, A.L., Ennis, F.A., Nisalak, A., 2000. Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. J. Infect. Dis. 81, 2–9.

Wittke, V., Robb, T.E., Thu, H.M., Nisalak, A., Nimmannitya, S., Kalayanrooj, S., Vaughn, D.W., Endy, T.P., Holmes, E.C., Aaskov,

J.G., 2002. Extinction and rapid emergence of strains of dengue 3 virus during an interepidemic period. Virology 301, 148–156.
World Health Organization, 1997. Dengue Haemorrhagic Fever: Diagnosis, Treatment and Control, 2nd ed. World Health Organization, Geneva.
Yang, Z., 1999. PAML: a program package for phylogenetic analysis by maximum likelihood. CABIOS 13, 555–556.
Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M.K., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.
Yao, Z.J., Kao, M.C., Loh, K.C., Chung, M.C., 1995. A serotype-specific epitope of dengue virus identified by phage displayed random peptide library. FEMS Microbiol. Lett. 127, 93–98.