Available online at www.sciencedirect.com

SCIENCE DIRECT®

Virology 329 (2004) 168–179

VIROLOGY

www.elsevier.com/locate/yviro

The molecular epidemiology of dengue virus serotype 4 in Bangkok, Thailand☆

Chonticha Klungthong[a,b,1], Chunlin Zhang[a,1], Mammen P. Mammen Jr.[a], Sukathida Ubol[b], Edward C. Holmes[c,*]

[a]*Department of Virology, U.S. Army Medical Component–Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand*
[b]*Department of Microbiology, Mahidol University, Bangkok, Thailand*
[c]*Department of Zoology, University of Oxford, Oxford OX1 3PS, UK*

Received 28 May 2004; returned to author for revision 8 July 2004; accepted 6 August 2004
Available online 17 September 2004

Abstract

Dengue represents a major public health problem in Thailand, with all four viral serotypes co-circulating. Dengue virus serotype 4 (DENV-4) is the least frequently sampled serotype, although one that is often associated with hemorrhagic fever during secondary infection. To determine the evolutionary forces shaping the genetic diversity of DENV-4, and particularly whether its changing prevalence could be attributed to instances of adaptive evolution in the viral genome, we undertook a large-scale molecular epidemiological analysis of DENV-4 in Bangkok, Thailand, using both E gene and complete coding region sequences. This analysis revealed extensive genetic diversity within a single locality at a single time, including the discovery of a new and divergent genotype of DENV-4, as well as a pattern of continual lineage turnover. We also recorded the highest average rate of evolutionary change for this serotype, at 1.072×10^{-3} nucleotide substitutions per site, per year. However, despite this abundant genetic variation, there was no evidence for adaptive evolution in any gene, codon, or lineage of DENV-4, with the highest rate of nonsynonymous substitution observed in NS2A. Consequently, the rapid turnover of DENV-4 lineages through time is most likely the consequence of a high rate of deleterious mutation in the viral genome coupled to seasonal fluctuations in the size of the vector population.

Keywords: Dengue virus; Molecular epidemiology; Genotype; Substitution rate; Mosquito

Introduction

Dengue is the most prevalent vector-borne viral infection of humans; each year, an estimated 100 million people suffer an acute febrile illness designated dengue fever (DF), while at least 250 000 experience the far more serious condition of dengue hemorrhagic fever (DHF), which is associated with case fatality rates of up to 15% depending on the availability of appropriate clinical management (Gubler, 2002). As well as its public health burden, dengue-associated illness is responsible for considerable monetary losses from lost work hours and tourism, and the virus may be expected to grow in prevalence due to such factors as human population size growth, international transportation systems, and global warming. As dengue vaccine development is actively underway, an understanding of the evolutionary factors that determine viral genetic diversity is of utmost importance.

Dengue viruses are single-stranded positive-sense RNA viruses (genus *Flavivirus*) that are assigned to four antigenically distinct serotypes—denoted DENV-1 to DENV-4. These viruses now co-circulate in many parts of the tropical and subtropical world following the geographical distribu-

☆ The opinions or assertions contained herein are the private ones of the authors and are not to be construed as official or as reflecting the view of the U.S. government.

* Corresponding author. Department of Zoology, University of Oxford, Tinbergen Building, South Parks Road, Oxford OX1 3PS, UK. Fax: +44 1865 281997.
E-mail address: Edward.Holmes@zoo.ox.ac.uk (E.C. Holmes).

[1] These authors have contributed equally to this study.

0042-6822/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2004.08.003

tion of *Aedes* sp. mosquitoes. The viral genome, which is translated as a single polyprotein, encodes three structural proteins (C, M, and E) and seven nonstructural proteins (NS1, NS2a, NS2b, NS3, NS4a, NS4b, and NS5).

In recent years, considerable attention has been directed toward characterizing the extent of genetic diversity in each of the four serotypes of dengue virus, particularly through the sequencing and phylogenetic analyses of complete envelope (E) gene sequences from strains sampled globally (reviewed in Holmes and Twiddy, 2003). In general, these studies reveal that strains of dengue virus from individual serotypes often fall into well-defined phylogenetic groups supported by high bootstrap values, which may be called "genotypes" or "subtypes" (although such definitions are likely to change with increased sampling). A more revealing pattern is that viral genotypes often have differing patterns of geographic distribution, with some restricted to particular geographic regions (notably Southeast Asia), while others are found in more diverse geographical localities, reflecting the extensive migration of both infected hosts and vectors. For example, in DENV-2, where E gene sequence data is particularly abundant, some genotypes have distributions that cover much of the land mass within the tropics of Cancer and Capricorn, so that they can be thought of as having a "cosmopolitan" geographic structure (Twiddy et al., 2002a). In sum, the current genetic structure of dengue virus populations appears to reflect an intricate pattern of gene flow, most likely caused by the movement of infected humans, and population subdivision, as the virus spreads within local communities, highlighting the complex epidemiological history of these viruses in the 20th century (Holmes and Twiddy, 2003). Finally, in the case of DENV-2 and DENV-4, genotypes confined to a sylvatic (or "jungle") cycle have also been described (Wang et al., 2000). Here, dengue viruses circulate exclusively in nonhuman primates and are transmitted by a variety of rural *Aedes* species. It is likely that phylogenetically distinct sylvatic genotypes also occur in DENV-1 and DENV-3.

Thailand is representative of countries that experience annual dengue virus transmission with the co-circulation of all four virus serotypes. Since its first appearance in the late 1950s, DHF case numbers have increased, albeit with changing frequencies of isolation of each serotype and an oscillation between relatively low and high dengue case years (Nisalak et al., 2003), most likely reflecting complex immunodynamics. Although DENV-4 is the least frequently isolated serotype, has not been associated with severe dengue outbreaks, and is responsible for only 2% of dengue cases observed at the Queen Sirikit National Institute of Child Health (QSNICH) in Bangkok (formerly known as Bangkok Children's Hospital) between 1973 and 1999, it was the predominant serotype during 1993 and 1994 (Nisalak et al., 2003). More strikingly, despite the low prevalence of DENV-4, it was responsible for 10% of DHF cases in children; the vast majority of DENV-4 DHF cases (359 of 368) being associated with secondary dengue virus infection (Nisalak et al., 2003). This implies that DENV-4 requires the action of heterologous immune factors (e.g., antibodies vis-a-vis antibody-dependent enhancement) to cause severe manifestations of disease. Thailand also represents an ideal locality to document the molecular epidemiology of dengue virus. Not only is dengue a major public health problem in this country, but associated epidemiological data is available, so that the dynamics of this virus infection can be investigated with relative precision. Such a study recently revealed that DHF epidemics in Thailand occur on a 3-year cycle, spreading in waves from an origin in Bangkok (Cummings et al., 2004).

Although dengue virus is comprised of four distinct serotypes, relatively few comprehensive studies of DENV-4 genetic diversity have been undertaken, particularly in Southeast Asia. DENV-4 is of particular interest in that it is the first dengue serotype to diverge in phylogenetic analyses of the genus *Flavivirus* (Gaunt et al., 2001; Kuno et al., 1998; Jenkins et al., 2001), so that it may have developed unique phenotypic features, and is generally found at low frequency in Southeast Asia. Early molecular epidemiological studies confirmed that DENV-4 could be divided into a series of discrete genotypes (Chungue et al., 1995; Lanciotti et al., 1997), although some intragenotype recombination has also been observed (AbuBakar et al., 2002; Worobey et al., 1999). To date, three major genotypes of DENV-4 virus have been described (Lanciotti et al., 1997); one found in Southeast Asia (genotype I), a second described in Southeast Asia and the Americas (genotype II), and a third found exclusively as sylvatic strains in Malaysia. More recently, extensive molecular epidemiological studies have documented the spread of DENV-4 in the Caribbean, most notably Puerto Rico, where genotype II seems to have become established since the early 1980s after having been imported from Southeast Asia (Bennett et al., 2003; Foster et al., 2003). In particular, a large-scale analysis of DENV-4 genetic diversity in Puerto Rico produced phylogenetic trees with a strong temporal structure, such that the position of lineages generally corresponded to their time of sampling; for example, the oldest viral strains, sampled in 1981–1982, fell nearest the root of the tree, while the most recently sampled strains fell furthest from the root (Bennett et al., 2003). Moreover, the major epidemic of DF associated with DENV-4 that affected Puerto Rico in 1998 was linked to a high ratio of nonsynonymous (d_N) to synonymous (d_S) substitutions per site in the NS2A protein, corresponding to three amino acid changes (at residues 14, 54, and 101). In contrast, the strongly antigenic E protein remained highly conserved in sequence over the 17-year sampling period. This was interpreted to mean that a selectively favored strain of DENV-4 was responsible for the surge of dengue in 1998, and the overall temporal structure of the tree, but that this selection pressure did not involve mechanisms to escape from neutralizing antibodies. Although the selection pressures acting on NS2A remain unclear, it is noteworthy that

the escape from T-cell-mediated immunity involving epitopes in the nonstructural genes has been documented, and may play a key role in dengue pathogenesis (Mongkolsapaya et al., 2003).

In this paper, we present the results of an extensive molecular epidemiological analysis of DENV-4 in Bangkok, Thailand, to determine the structure of viral genetic diversity in this locality and the evolutionary processes responsible for this structure. In particular, we wished to determine whether the change in DENV-4 prevalence through time could be associated with specific genetic changes in the viral genome; that is, to what extent does adaptive evolution in the viral genome influence long-term epidemiological patterns? To this end, we have sequenced the E gene sequences from 53 dengue viral isolates recovered during a 27-year period in Thailand and further determined the complete coding region of the viral genome from six of these strains.

Results

Prevalence of DENV-4

The changing prevalence of DENV-4 in Thailand is shown in Fig. 1. This is characterized by distinct outbreaks of 3–5-year duration interspersed between periods of relatively little disease/virus isolation. As has been observed in other studies of viral genetic diversity (Shurtleff et al., 2001; Uzcategui et al., 2003), there was no obvious association between disease status and the phylogenetic position of their associated DENV-4 strains; strains associated with DHF fell indiscriminately on the phylogenetic tree and did not tend to group together. Hence, DHF is not associated with a specific DENV-4 E gene sequence. As described below, the most important factor shaping the DENV-4 tree was the time of sampling.

Molecular epidemiology of DENV-4 viruses

Our phylogenetic analysis of 109 E gene sequences shows that DENV-4 viruses sampled globally can be placed into four distinct and well-supported groups, which can be thought of as genotypes, one of which is only found in nonhuman primates in Malaysia (Fig. 2). The majority of the Thai DENV-4 strains sampled here fell into the previously described genotypes I and II; 47 fell into genotype I that has been widely reported in Asiatic regions, including Thailand, while a single strain (ThD4-0734/00) fell into a subclade of genotype II. The latter appears to have an Asian origin and was then exported to Latin America, The Caribbean (and also the South Pacific), where it has evolved in situ for almost 25 years.

The most striking observation, however, was that five of the Thai viruses obtained here fell into a phylogenetic group that has not been described previously, is distinct from the other Thai viruses, and seemingly the first human lineage of DENV-4 virus to diverge. This is denoted “genotype III” in Fig. 2. Not only is this new genotype phylogenetically distinct (although its position relative to the other two human genotypes in the E gene could not be assigned with certainty in the bootstrap analysis), but it is comprised of Thai DENV-4 strains sampled relatively recently, from 1997 to 2001. This suggests two possibilities: strains of this new and divergent DENV-4 genotype existed before 1997 but had not been sampled previously, or the genotype came into existence recently following a recombination event between diverse DENV-4 viruses.

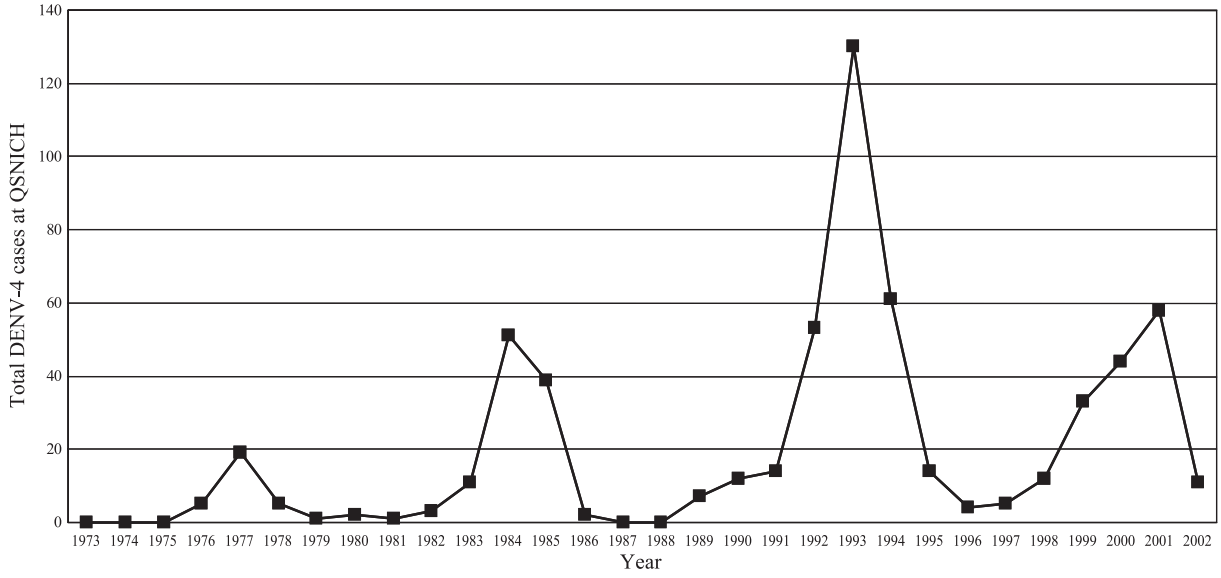

Fig. 1. The total cases of DENV-4 recorded in children at the QSNICH, Bangkok, Thailand during the period 1973 to 2002.

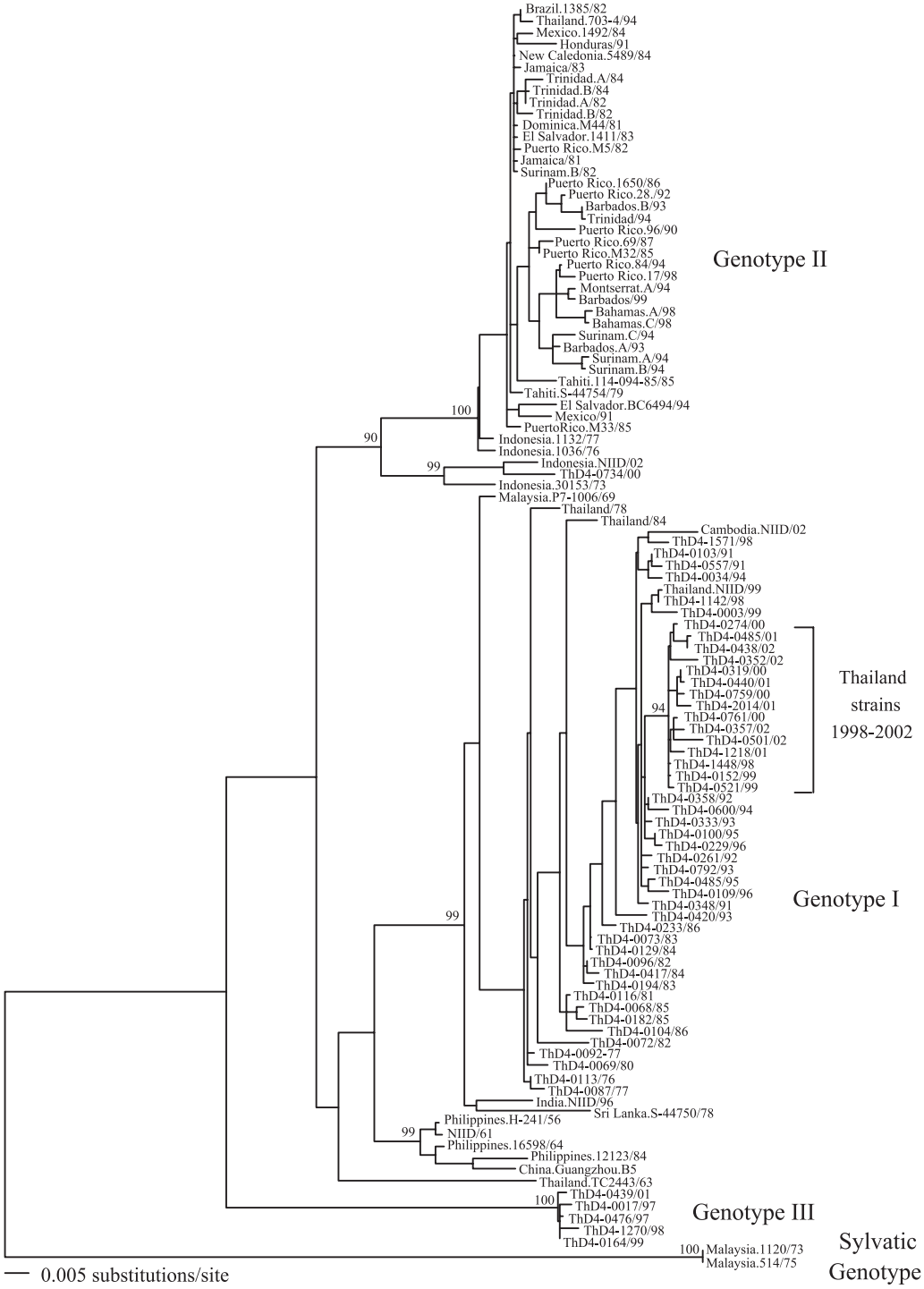

Fig. 2. Maximum likelihood (ML) phylogenetic tree of 109 E sequences of DENV-4 sampled on a global basis. The four viral genotypes, including that newly described here (III), are indicated. Bootstrap support values (where >90%) are shown for a number of key nodes on the tree. The tree is rooted by the two sylvatic (nonhuman primate) strains and all horizontal branch lengths are drawn to scale.

To determine whether the new Thai genotype of DENV-4 is a product of divergent evolution or recombination, we compared complete coding regions from six Thai strains representing the full phylogenetic diversity of DENV-4 in this country, including two from the new genotype III (strains ThD4-0017/97 and ThD4-0476/97), as well as two complete DENV-4 genomes available on GenBank (Dominica/81 and China.Guangzhou.B5). These eight strains cover the full phylogenetic diversity of human DENV-4. Phylogenetic trees were estimated for all 10 genes from these 8 viral strains. In every case, strains ThD4-0017-97 and ThD4-0476/97 fell into a separate and divergent clade,

confirming that they constitute an entirely new genotype of DENV-4 and are not the product of recombination (trees not shown; available upon request). As expected, the divergent phylogenetic position of genotype III was also observed on a phylogenetic tree of the complete coding region from 58 viral isolates (Fig. 3). In this case, the early division of genotype III was also supported by a high number of bootstrap replications, indicating that this genotype indeed has a long independent evolutionary history.

The phylogenetic tree of the DENV-4 E gene produced here is also notable for its strong temporal structure, particularly in the Thai DENV-4 strains assigned to genotype I; viral strains isolated at the earliest time points (such as ThD4-0087/77 sampled in 1977) tended to fall near the root on the phylogenetic tree, while those sampled more recently (such as ThD4-0352/02 from 2002) were located at the most distal tips. This temporal ordering is caused by the continual emergence and extinction of viral lineages; new lineages are regularly produced by mutation, but most go extinct relatively rapidly and few progress to subsequent sampling times. A particularly clear example of this process of strain turnover seems to have taken place during 1998–1999; all Thai strains sampled from genotype I from 2000 to 2002, as well as one strain from 1998 and two from 1999, fell into a distinct and well-supported clade on the phylogenetic tree that is separate from those sampled in earlier years (highlighted in Fig. 2). This suggests that an

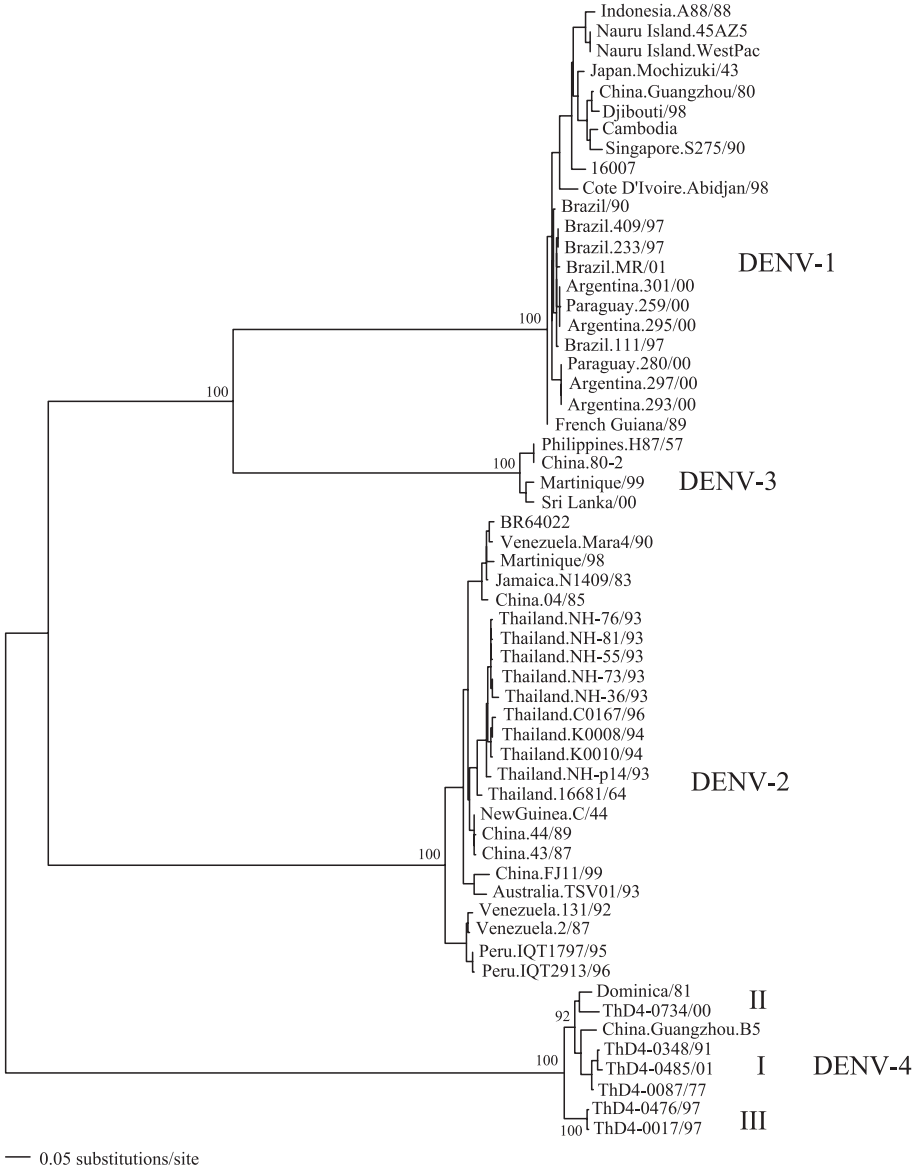

Fig. 3. ML phylogenetic tree of 58 complete coding regions (10185 bp) from dengue viruses, including 8 from DENV-4. The three human genotypes of DENV-4 are indicated. Bootstrap support values (where >90%) are shown for nodes depicting the relationships of the four viral serotypes and new genotype of DENV-4. The tree is mid-point rooted for purposes of clarity only and all horizontal branch lengths are drawn to scale.

extinction–replacement event took place after the sampling in 1999.

Selection pressures in DENV-4

To determine the nature of the selection pressures acting on DENV-4, we undertook an extensive analysis of the numbers of synonymous (d_S) and nonsynonymous (d_N) changes per site, at the gene, codon, and lineage levels. This analysis confirmed recent reports that the predominant evolutionary force acting on this virus is relatively strong purifying selection (Twiddy et al., 2002b; Woelk and Holmes, 2002). In all viral genes analyzed, d_S greatly exceeded d_N, indicating that purifying selection is strong (Table 1). In terms of increasing selective constraint (decreasing d_N/d_S), the order of genes was as follows: NS2A, C, NS1, NS2B, E, NS4A, NS5, M, NS3, NS4B, with a 6.8-fold variation in d_N/d_S across genes (ranging from 0.013 to 0.088).

A similar picture of strong purifying selection was observed in the codon and lineage-specific analysis of d_N/d_S. At the codon level, there was no evidence for positive selection at any individual codon in the sequence alignment (the neutral M7 model could not be rejected by the selection M8 model, $P = 0.225$), and although there was evidence for lineage-specific variation in d_N/d_S (M0 was significantly rejected by FR, $P = 0.012$), no individual lineage was distinguished by an unusually high d_N/d_S, as expected if positive selection had been in operation. Indeed, under M0, the mean d_N/d_S across the tree was 0.055, again indicating that purifying selection dominates DENV-4 evolution.

Rates of evolution in DENV-4

The strong temporal structure of DENV-4 in Thailand also allowed us to estimate the rates of molecular evolution (nucleotide substitution) of this virus. Using a Bayesian MCMC method, we estimated that the overall substitution rate of DENV-4 in 47 Thai (Bangkok) isolates spanning the years 1976 to 2002 and assigned to genotype I was on average 1.072×10^{-3} substitutions/site/year (highest posterior density [HPD] of 0.841×10^{-3}–1.311×10^{-3}). This falls within the range seen in other RNA viruses (Jenkins et al., 2002).

Table 1
Selection pressures in DENV-4 on a gene-by-gene basis

Protein	Length (amino acids)	d_N	d_S	d_N/d_S
Capsid	113	0.013	0.229	0.057
Membrane	166	0.008	0.400	0.020
Envelope	495	0.011	0.335	0.033
NS1	352	0.015	0.335	0.045
NS2A	218	0.031	0.354	0.088
NS2B	130	0.009	0.245	0.037
NS3	618	0.005	0.342	0.015
NS4A	150	0.013	0.436	0.030
NS4B	245	0.005	0.372	0.013
NS5	900	0.008	0.337	0.024

Discussion

The molecular epidemiology of DENV-4

We have performed the most extensive molecular epidemiological study of DENV-4 to date. Strikingly, three different viral genotypes were found in a single location—Bangkok, Thailand—at equivalent sampling times. Although viruses assigned to genotype I were the most common form isolated, it is clear that Bangkok acts as a major reservoir for dengue virus diversity. This may also explain why Bangkok has been identified as the epicenter for DHF in Thailand (Cummings et al., 2004). Of the three genotypes present in Bangkok, one (genotype III) is newly identified, and is also the most divergent human strain of DENV-4 identified to date. It is likely that a more extensive sampling procedure would reveal more genetic diversity within DENV-4, including other members of genotype III and perhaps novel genotypes.

Although the geographic range of current DENV-4 sampling is limited, it is notable that all the human genotypes of DENV-4 were present in Southeast Asia, particularly Thailand, and that the sylvatic genotype has been found in monkeys in Malaysia. At face value, this suggests that the ultimate origin of DENV-4 might lie in Southeast Asia. Such an interpretation is supported by the observation that even in the cosmopolitan genotype II, which contains isolates from a wide range of geographical locations, those from Asia (including ThD4-0734/00) fall at the deepest locations on the molecular tree. Hence, it appears that Southeast Asia serves as the source population for dengue virus strains that have subsequently spread throughout the tropical world since World War II. Confirmation of this hypothesis will require analysis of a broader geographical sampling of DENV-4 strains, particularly from Africa where sampling has traditionally been sparse.

A greater challenge is to explain the strong temporal structure of the DENV-4 E gene phylogenies, particularly over an approximately 30-year sampling period in Thailand. At one level, the high rate of lineage extinction observed, producing the distinctive ladder-like phylogeny, could be taken as evidence for continual immune-driven positive selection as similarly shaped phylogenies have been observed for the hemagluttinin (HA) gene of human influenza A virus sampled at the interhost level, and the envelope gene of HIV-1 serially sampled within individual hosts (reviewed in Grenfell et al., 2004). In these cases, the short side branches represent lineages that are eventually removed from the population by cross-protective immune responses, while the main trunk of the tree (i.e., that

connects viruses from different sampling times) traces the antigentically favored strains through time. However, we found no evidence for positive selection on any branch or codon in the E gene, in contrast to the powerful selection observed in both interhost influenza A (Bush et al., 1999) and intrahost HIV (Williamson, 2003). Indeed, the very low d_N/d_S values observed in DENV-4 reveal a strong force of purifying selection. Such powerful purifying selection has also been noted by the dramatic reduction in d_N/d_S within infected populations compared to that observed within infected individuals; this strongly suggests that the vast majority of amino acid changes that arise within infected hosts are deleterious in the long term and eventually are removed from the population by purifying selection (Holmes, 2003; Wittke et al., 2002). A more likely explanation for the ladder-like DENV-4 phylogeny and the high rate of lineage extinction is therefore a regular series of population bottlenecks, probably caused by seasonal fluctuations in the size of the mosquito population, coupled with a high rate of deleterious mutation. Similar instances of lineage replacement, thought to be the outcome of population bottlenecks, have previously been observed in both DENV-2 and DENV-3 in Thailand (Sittisombut et al., 1997; Wittke et al., 2002). In sum, we suggest that the dynamics of mosquito transmission may mean that population bottlenecks and rapid lineage turnover are a common observation in dengue virus evolution. As such, the pattern of viral evolution observed is more likely to be a consequence of large-scale epidemiological factors than the cause of changing patterns of prevalence.

The molecular evolution of DENV-4

Our relatively large DENV-4 data set enabled us to undertake an extensive study of the molecular evolution of this virus. First, by using the detailed temporal sampling of DENV-4 strains, we were able to obtain a precise estimate of the rate of molecular evolution (nucleotide substitution) in this virus. The mean rate estimated—1.072×10^{-3} substitutions/site/year—is noteworthy in that it is one of the highest rates estimated for dengue virus to date, particularly for DENV-4 (Twiddy et al., 2003). More specifically, the rate is approximately twice that (0.542×10^{-3} substitutions/site/year) estimated by Twiddy et al. (2003) on the same genotype using similar methods although with a much smaller data set, over 35% greater than DENV-4 rate (0.77×10^{-3} substitutions/site/year) obtained in the large comparative study of Jenkins et al. (2002), and more than 20% higher than the overall DENV-4 rate (0.83×10^{-3} subs/site/year) estimated by Lanciotti et al. (1997) using a simpler linear regression method. Whether the relatively high substitution rate estimated here reflects the high turnover of DENV-4 in Thailand, or is simply a reflection of more extensive sampling, is unclear. However, because we have identified a signal of strong purifying selection in DENV-4 coupled with a process of lineage extinction, all the substitution rates estimated for dengue virus to date are possibly slight overestimates because deleterious mutations that have yet to be selectively purged from the population are still included in the sample.

A second important insight into the molecular evolution of DENV-4 comes from our gene-specific analysis of the selection pressures acting on the complete coding region of the virus genome. In particular, we found that some of the nonstructural genes, most notably NS2A, exhibited higher levels of nonsynonymous variation than the structural and antigenic E gene, which is often thought to harbor the greatest amount of genetic diversity because of antibody-mediated selection. As noted above, the maximum likelihood analysis of selection pressures in the E gene also revealed a process of strong purifying selection. The basis of the (relatively) high level of genetic variation in NS2A is unclear and requires further explanation. Although there was insufficient data to undertake a codon-specific analysis of selection pressures, nucleotide substitutions in this gene were previously cited as the virological cause of a recent epidemic of DENV-4 in Puerto Rico (Bennett et al., 2003) and another nonstructural gene, NS3, has been found to harbor T-cell epitopes (Mongkolsapaya et al., 2003). As such, our results suggest that more attention be directed toward the nonstructural genes as a source of adaptively relevant genetic variation, particularly NS2A.

Materials and methods

Specimen data

DENV-4 was isolated from 53 children (mean age = 8.7 years) hospitalized at QSNICH over a 27-year period spanning 1976 to 2002 (Table 2). Of these, 21 were associated with DF and 32 with DHF. Grading of these diagnostic specimens was conducted by QSNICH physicians using WHO classification guidelines. Further testing at AFRIMS characterized the specimens according to their dengue virus serotype and nature of the dengue virus infection. Primary versus secondary DENV-4 infection was determined solely by hemagglutination assay inhibition (HAI) before 1990 and by MAC-ELISA supported by HAI subsequently. Viral collections before 1980 had been isolated after two to three serial passages in LLC-MK2 cells, whereas subsequently they were isolated after amplification in *Toxorhynchites splendens* mosquitoes followed by one to three passages in C6/36 cells.

RNA extraction and reverse transcriptase-polymerase chain reaction (RT-PCR) amplification

Virus RNA was extracted from cell culture supernatant using Trizol (GIBCO BRL) reagent. One milliliter of Trizol was mixed with 100 μl of each DENV-4 virus and 200 μl of RNase-free chloroform (Sigma) was added to each sample

Table 2
Description of DENV-4 isolates sequenced in this study

Virus[a]	Disease	Infection	Sex[b]	Age (Years)	Virus[a]	Disease	Infection	Sex[b]	Age (Years)
ThD4-0113/76	DHF III	Secondary	F	6.00	ThD4-0109/96	DHF IV	Secondary	F	8.92
ThD4-0087/77	DHF III	Secondary	M	11.00	ThD4-0229/96	DF	Secondary	F	11.25
ThD4-0092/77	DF	Secondary	F	8.00	ThD4-0017/97	DF	Secondary	M	10.42
ThD4-0696/80	DHF III	Unknown	M	12.00	ThD4-0476/97	DHF II	Secondary	F	7.83
ThD4-0116/81	DHF III	Secondary	F	5.00	ThD4-1142/98	DHF I	Secondary	M	12.75
ThD4-0072/82	DF	Secondary	M	10.00	ThD4-1270/98	DHF III	Secondary	M	8.75
ThD4-0096/82	DHF III	Secondary	M	3.00	ThD4-1448/98	DHF II	Secondary	F	10.92
ThD4-0073/83	DHF III	Secondary	M	10.00	ThD4-1571/98	DF	Secondary	F	3.33
ThD4-0194/83	DF	Secondary	M	3.00	ThD4-0003/99	DHF I	Secondary	F	10.25
ThD4-0129/84	DHF III	Secondary	F	11.00	ThD4-0152/99	DHF II	Secondary	F	10.33
ThD4-0417/84	DF	Secondary	M	12.00	ThD4-0164/99	DF	Unknown	F	3.92
ThD4-0068/85	DHF IV	Secondary	F	5.00	ThD4-0521/99	DHF III	Secondary	F	11.00
ThD4-0182/85	DF	Secondary	M	10.00	ThD4-0274/00	DHF IV	Secondary	F	2.33
ThD4-0104/86	DF	Secondary	F	6.00	ThD4-0319/00	DHF II	Secondary	F	11.08
ThD4-0233/86	DHF III	Secondary	F	15.00	ThD4-0734/00	DF	Secondary	M	14.92
ThD4-0103/91	DHF I	Primary	F	0.58	ThD4-0759/00	DF	Secondary	F	4.50
ThD4-0348/91	DF	Secondary	M	10.00	ThD4-0761/00	DF	Secondary	F	10.58
ThD4-0557/91	DHF II	Secondary	F	6.00	ThD4-0439/01	DHF II	Secondary	F	12.17
ThD4-0261/92	DHF III	Secondary	M	8.00	ThD4-0440/01	DHF I	Secondary	F	14.08
ThD4-0358/92	DF	Secondary	M	8.00	ThD4-0485/01	DHF I	Secondary	M	5.5
ThD4-0333/93	DF	Secondary	M	12.75	ThD4-1218/01	DF	Secondary	M	12.58
ThD4-0420/93	DHF III	Secondary	F	11.00	ThD4-2014/01	DF	Secondary	F	1.92
ThD4-0792/93	DHF[c]	Primary	M	0.50	ThD4-0352/02	DF	Secondary	M	8.75
ThD4-0034/94	DF	Unknown	F	11.67	ThD4-0357/02	DHF II	Secondary	M	8.42
ThD4-0600/94	DHF[c]	Secondary	F	10.17	ThD4-0438/02	DHF I	Secondary	F	3.33
ThD4-0100/95	DHF I	Secondary	M	13.08	ThD4-0501/02	DHF III	Secondary	F	9.92
ThD4-0485/95	DF	Secondary	M	11.42					

[a] The middle numeral is the sample number while the last two numerals indicate the year of isolation. The highlighted (gray color) samples were those selected for complete genome sequencing.
[b] F is female and M is male.
[c] Unclassified grade.

mixture followed by vigorous vortexing. The mixture was centrifuged at 12 000 × *g* at 10 °C for 15 min, and supernatant was taken to mix with 500 µl of isopropanol and then kept at room temperature for 5 min before centrifugation at 12 000 × *g* at 20 °C for 10 min. Each pellet was washed with 1 ml of 75% ethanol and centrifuged at 12 000 × *g* at 20 °C for 2 min. The RNA pellet was dried at room temperature for 30 min, and then resuspended in 25 µl of RNase-free water (Sigma). The RNA solution was used for RT-PCR. Oligonucleotide primers for RT-PCR amplification and sequencing were designed by using Primer Design program (http://www.hgmp.mrc.ac.uk/GenomeWeb/nuc-primer.html). The sequences and positions of primers for DENV-4 genome are shown in Table 3. All specimens used for sequencing were pretested by using a Lanciotti RT-PCR and nested PCR to confirm correct viruses for the DENV-4 serotype and to adjust virus concentration. The RT-PCR was performed according to the protocol of Lanciotti et al. (1992).

Sequencing and analysis

DENV-4 genomic RNA was converted to cDNA by using random hexamer oligonucleotides with SuperScript First-Strand Synthesis System (Invitrogen) according to the manufacturer's instructions. The DNA fragments of the E gene for all 53 DENV-4 and eight overlapping DNA fragments that covered the entire DENV-4 genome for each of isolates ThD4-0087/77, ThD4-0348/91, ThD4-0017/97, ThD4-0476/97, ThD4-0734/00, and ThD4-0485/01 (Fig. 4) were amplified by PCR with *Taq* DNA polymerase and eight pairs of primers described in Fig. 4 and Table 3. The PCR-amplified DNA fragments were purified using Qiagen PCR purification kits (Qiagen, Germany) according to the manufacturer's instructions. Purified DNA fragments were used for sequencing. For an unknown reason, only forward direction sequencing was available for regions nt 9708–9943 and nt 10178–10653 of sample ThD4–0476/97 and nt 9679–9943 and nt 10451–10649 of sample ThD4–0485/01. However, in all cases, the chromatography peak for each base was unambiguous.

Cycle sequencing reactions were performed using DYEnamic ET Dye Terminator Sequencing Kit (Amersham Pharmacia Biotech), according to the manufacturer's instructions, with sequencing primers described in Table 3. The sequencing products were cleaned by standard precipitation before sequencing on a MegaBASE 500 automated DNA sequencer (Amersham Pharmacia Biotech). Over-

Table 3
DENV-4 sequencing primers used in this study

Name[a]	Sequence (5′→3′)	Position[b]	Name[a]	Sequence (5′→3′)	Position[b]
5′D4F1[c]	AGTTGTTAGTCTGTGTGGACC	1–21	3′D4R235	ATCCGTAAGGGTCCTTTCC	217–235
5′D4F136	TCAATATGCTGAAACGCGCGAGAAACCG	136–163	3′D4R646	TTGCACCAACAGTCAATGTCTTCAGGTTC	618–646
5′D4F717[c]	GTAGCTTTAACACCACATTC	717–736	3′D4R836[c]	TCCTGGGTTTCTGAGTATCCA	816–836
5′D4F1162	CAACGCAAGGAGAGCCTTA	1162–1180	3′D4R1285	CATGTCACAACTCCTCCTT	1267–1285
5′D4F1631	GAATTACAAAGAGAGAATGGT	1631–1651	3′D4R1847	TGAACACATCGTGTATGAC	1829–1847
5′D4F2064	AGCTACATAGTGATAGGTGT	2064–2083	3′D4R2163	CACCTCTGTATGTGGACTC	2145–2163
5′D4F2373[c]	TGCATAGCTGTTGGAGGAAT	2373–2392	3′D4R2519[c]	GTACTGTTCTGTCCAAGTGT	2500–2519
5′D4F2824	TAATAGACGGACCAGACAC	2824–2842	3′D4R2772	TCTTCCATGAATATTTCAGA	2753–2772
5′D4F3244	AGATAGACTTTGGAGAATGC	3244–3263	3′D4R2773	GTCTTCCATGAATATTTCA	2755–2773
5′D4F3543[c]	TGCTTGAGGAGAAGAGTCAC	3543–3562	3′D4R3191	GTAATTGTGCTGTGAAAAAG	3172–3191
5′D4F3564	AGGAAACACATGATATT	3564–3580	3′D4R3270	TTCCGGGGCATTCTCCAA	3253–3270
5′D4F3992	GTTTGTGGTCACACTCATT	3992–4010	3′D4R3642[c]	GTAGTAAGTCCATCCATGTG	3623–3642
5′D4F4439	ACTGATAACAGTGTCAGGTC	4439–4458	3′D4R4111	AGAGTCATTAGGTACACTGG	4092–4111
5′D4F4883	TGGAGAAATTGGAGCAGTA	4883–4901	3′D4R4128	TTGAGGCTCCTTTCATGAG	4110–4128
5′D4F5298	CTCATGTGTCATGCAACCT	5298–5316	3′D4R4519	TGTGTT(T,C)TCACTTGCCA	4503–4519
5′D4F5313[c]	ACCTTCACAACAAGACTTTT	5313–5332	3′D4R4528	CCTGATCTTTGTGTTTTCAC	4509–4528
5′D4F5788	TGATAGACC(C,T)CAGGAGATG	5788–5806	3′D4R4972	TAGAGTCCGATGACTTTTCC	4953–4972
5′D4F5862	ATTCCAGTGACTCCAGCA A	5862–5880	3′D4R5005	TAATCACCTGATTTGGT	4989–5005
5′D4F6343	TGAAGGATTTCAAGGAGT	6343–6360	3′D4R5383[c]	TGTGCTTCATCCATCACTAT	5364–5383
5′D4F6441	AAGCT(C,A)GCCCTTGACAA	6441–6457	3′D4R5717	TTTCGTTTTTGGATACTCT	5699–5717
5′D4F6783	TACGTCATATTGACCATTCT	6783–6802	3′D4R5758	TCAGATATGTCTGT(G)GT	5742–5758
5′D4F6796	CCATTCTCAC(T,C)ATTATTG	6796–6813	3′D4R6160	AGCCACACCGGAAGGTCT	6143–6160
5′D4F6975[c]	AGACACACCATAGAAAACAC	6975–6994	3′D4R6180	CAGAAGCTACCTTATAGCTC	6161–6180
5′D4F7401	ACTTTGGCCACAGGACCA	7401–7418	3′D4R6676	ATTTCTGC(C,T)ACCCAGAG	6660–6676
5′D4F7414	GACCAATCTTGACCTTGT	7414–7431	3′D4R7127[c]	GTTCACTTGAGAATAGCATC	7108–7127
5′D4F7882	GTCCAGGACATGAAGAAC	7882–7899	3′D4R7470	CTATGGTCGTGTTCCAAA	7453–7470
5′D4F8350[c]	GAAGTGTCTCCACTGAAAC	8350–8368	3′D4R7905	GAATCGGTTCTTCATGTC	7888–7905
5′D4F8645	CAGAACACCACAACCAAA	8645–8662	3′D4R8326	TCCTTCTCATA(A,G)GTGGG	8310–8326
5′D4F8682	ATGACCAC(G,A)ACAGCCAA	8682–8698	3′D4R8331	CTACGTCCTTCTCATAAGTG	8312–8331
5′D4F9069	TCATGGAGTGGA(G,C)TGGAA	9069–9086	3′D4R8768[c]	TTTTGAGATGAACTCTTCC	8750–8768
5′D4F9070	CATGGAGTGGAGTGGAAG	9070–9087	3′D4R9302	GACTTTCACCACTTTGTTT	9284–9302
5′D4F9481	TGAAAGAAAGAGT(T,C)GAGA	9481–9498	3′D4R9375	CAAC(C,T)TGTCCACTACCT	9359–9375
5′D4F9485[c]	AGAAAGAGTTGAGAAATGG	9485–9503	3′D4R9750	TACA(T,C)GGAACAACCAGTGA	9732–9750
5′D4F9944	AACATGGTCAATCCACGC	9944–9961	3′D4R9764[c]	TTCATCCTGGTTTCTACAT	9746–9764
5′D4F9984	GAAGA(T,C)ATGCTCAAAGTG	9984–10001	3′D4R10180	TTCCTGAC(C,T)TGGGTTATG	10163–10180
5′D4F10411	ATATTGGACTAGCGGTTAGA	10411–10430	3′D4R10228	CTT(C,T)TCATGACTGGCATGT	10210–10228
			3′D4R10649[c]	AGAACCTGTTGGATCAAC	10632–10649

[a] Primer names with F indicate forward direction and R indicate reverse direction.
[b] Based on the nucleotide numbering of DENV-4 strain 814669 (GenBank Accession No. AF326573).
[c] These primers were used for both amplification and sequencing.

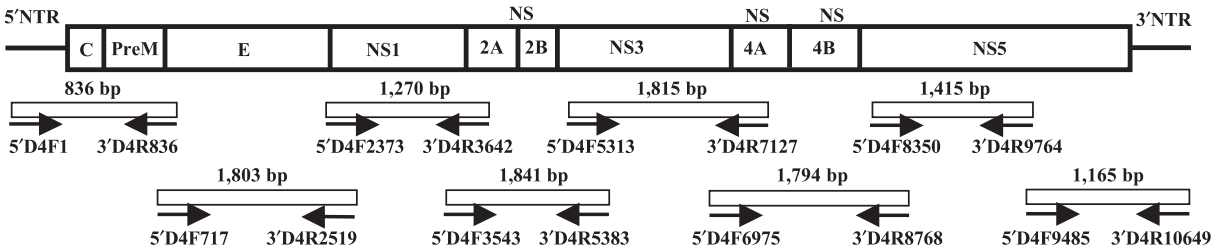

Fig. 4. Schematic diagram of the DENV-4 genome comprising eight overlapping DNA fragments. The relative positions of primers used in PCR to generate these DNA fragments are shown.

lapping nucleic acid sequences were combined for analysis and edited with the aid of the Sequencher software (Gene Code Corp., USA). All sequences produced here have been submitted to GenBank (accession numbers AY618935–AY618993). The other dengue virus isolates used in this analysis are listed in Table 4.

Phylogenetic analysis

A total of 53 DENV-4 E gene sequences from DF and DHF patients in Thailand were obtained. These sequences were combined with 56 DENV-4 E gene sequences taken from GenBank representing a wide range in geographic

Table 4
Sequences from other dengue virus isolates used in this analysis

Country of origin	Year[a]	Accession	Country of origin	Year[a]	Accession	Country of origin	Year[a]	Accession
Bahamas.A/98[b]	1998	AY152364	Puerto Rico.96/90[b]	1990	AY152855	China.44/89[c]	1989	AF204177
Bahamas.C/98[b]	1998	AY152366	Puerto Rico.84/94[b]	1994	AY152084	China.FJ11/99[c]	1999	AF359579
Barbados.A/93[b]	1993	AY152375	Puerto Rico.17/98[b]	1998	AY152056	China.Guangzhou/80[c]	1980	AF350498
Barbados.B/93[b]	1993	AY152376	Sri Lanka.S-44750/78[b]	1978	U18437	Cote D'Ivoire.Abidjan/98[c]	1998	AF298807
Barbados/99[b]	1999	AY152368	Surinam.B/82[b]	1982	AY152386	Djibouti/98[c]	1998	AF298808
Brazil.1385/82[b]	1982	U18425	Surinam.A/94[b]	1994	AY152372	Dominica/81[c]	1981	AF326573
Cambodia.NIID/02[b]	2002	AB111089	Surinam.B/94[b]	1994	AY152373	French Guiana/89[c]	1989	AF226687
China.Guangzhou.B5[b]	NA[d]	AF289029	Surinam.C/94[b]	1994	AY152374	Philippines.H87/57[c]	1957[d]	M93130
Dominica.M44/81[b]	1981	AY152360	Tahiti.S-44754/79[b]	1979	U18438	Jamaica.N1409/83[c]	1983	M20558
El Salvador.1411/83[b]	1983	U18426	Tahiti.114-094-85/85[b]	1985	U18439	Martinique/98[c]	1998	AF208496
ElSalvador.BC6494/94[b]	1994	U18427	Thailand.TC2443/63[b]	1963	U18440	Martinique/99[c]	1999	AY099337
Honduras/91[b]	1991	AY152379	Thailand/78[b]	1978	U18441	Japan.Mochizuki/43[c]	1943	AB074760
India.NIID/96[b]	1996	AB111086	Thailand/84[b]	1984	U18422	Nauru Island.A5AZ5[c]	NA[d]	U88536
Indonesia.30153/73[b]	1973	U18428	Thailand.703-4/94[b]	1994	AF231726	Nauru Island.WestPac[c]	NA[d]	U88535
Indonesia.1036/76[b]	1976	U18429	Thailand.NIID/99[b]	1999	AB111087	New Guinea.C/44[c]	1944	AF038403
Indonesia.1132/77[b]	1977	U18430	Trinidad.A/82[b]	1982	AY152382	Paraguay.259/00[c]	2000	AF514883
Indonesia.NIID/02[b]	2002	AB111088	Trinidad.B/82[b]	1982	AY152383	Paraguay.280/00[c]	2000	AF514878
Jamaica/81[b]	1981	AY152389	Trinidad.A/84[b]	1984	AY152380	Peru.IQT2913/96[c]	1996	AF100468
Jamaica/83[b]	1983	AY152384	Trinidad.B/84[b]	1984	AY152381	Peru.IQT1797/95[c]	1995	AF100467
Malaysia.P7-1006/73[b]	1973	AF231722	Trinidad/94[b]	1994	AY152377	BR64022[c,e]	NA[d]	AF489932
Malaysia.1120/73[b]	1973	AF231724	16607[c]	NA[d]	AF180818	Singapore.S275/90[c]	1990	M87512
Malaysia.514/75[b]	1975	AF231723	Indonesia.A88/88[c]	1988	AB074761	Sri Lanka/00[c]	2000	AY099336
Mexico/91[b]	1991	AY152378	Argentina.293/00[c]	2000	AY206457	Thailand.NH-36/93[c]	1993	AF169679
Mexico.1492/84[b]	1984	U18431	Argentina.295/00[c]	2000	AF514885	Thailand.NH-55/93[c]	1993	AF169681
Montserrat.A/94[b]	1994	AY152369	Argentina.297/00[c]	2000	AF514889	Thailand.NH-73/93[c]	1993	AF169686
NewCaledonia.5489/84[b]	1984	U18432	Argentina.301/00[c]	2000	AF514876	Thailand.NH-76/93[c]	1993	AF169687
NIID/61[b,e]	1961	AB111090	Australia.TSV01/93[c]	1993	AY037116	Thailand.NH-81/93[c]	1993	AF169688
Philippines.H-241/56[b]	1956	U18433	Brazil/90[c]	1990	AF226685	Thailand.NH-p14/93[c]	1993	AF022439
Philippines.16589/64[b]	1964	U18434	Brazil.111/97[c]	1997	AF311956	Thailand.K0008/94[c]	1994	AF100459
Philippines. 12123/84[b]	1984	U18435	Brazil.233/97[c]	1997	AF311958	Thailand.K0010/94[c]	1994	AF100460
Puerto Rico.1650/86[b]	1986	U18436	Brazil.409/97[c]	1997	AF311957	Thailand.C0167/96[c]	1996	AF100464
Puerto Rico.28/92[b]	1992	AY152196	Brazil.MR/01[c]	2001	AF513110	Thailand.16681/64[c]	1964	U87411
Puerto Rico.M5/82[b]	1982	AY152336	Cambodia[c]	NA[d]	AF309641	Venezuela.131/92[c]	1992	AF100469
Puerto Rico.M32/85[b]	1985	AY152856	China.04/85[c]	1985	AF119661	Venezuela.2/87[c]	1987	AF100465
Puerto Rico.M33/85[b]	1985	AY152857	China.80-2[c]	NA	AF317645	Venezuela.Mara4/90[c]	1990	AF100466
Puerto Rico.69/87[b]	1987	AY152252	China.43/87[c]	1987	AF204178			

[a] Year of sampling.
[b] E gene sequence.
[c] Complete coding region sequence.
[d] Year of sampling not available.
[e] Geographical origin not available.

localities, including earlier sampled viruses from Thailand and two sylvatic strains isolated from monkeys in Malaysia. This produced a total data set of 109 E gene sequences, 1485-bp long, on which phylogenetic analysis could be conducted. Known recombinant sequences were excluded from the analysis. For six isolates of DENV-4 from Thailand, the complete coding region of the viral genome was also available for analysis. These sequences were combined with 52 complete coding region sequences taken from GenBank representing all four serotypes of dengue virus, including two other DENV-4 isolates. This produced a total data set of 58 complete coding region sequences, 10 185-bp long. Both sequence alignments are available from the others by request.

For both the E gene and complete coding region sequences, maximum likelihood (ML) phylogenetic trees were estimated using the GTR + Γ + I model of nucleotide substitution, with the GTR substitution matrix, the base composition, the gamma distribution of among-site rate variation (Γ) and the proportion of invariant sites (I) all estimated from the data. All parameter values are available from the authors on request. To assess the robustness of particular phylogenetic groupings, a bootstrap resampling analysis was performed using 1000 replicate neighbor-joining (NJ) trees under the ML substitution model described above. All phylogenetic analyses were performed using the PAUP* package (Swofford, 2002).

Selection pressures

To determine the nature of the selection pressures acting on DENV-4, and particularly the incidence of positive selection, an analysis was made of the ratio of non-synonymous (d_N) to synonymous (d_S) nucleotide changes per site, with $d_N > d_S$ indicative of positive selection. This analysis was conducted on (i) a sample of 60 E gene sequences representing the full global diversity of DENV-4, and (ii) the eight DENV-4 viruses for which sequences of the complete coding region were available so that selection pressures could be assessed on a gene-by-gene basis. Two approaches were used to measure d_N and d_S. First, the mean d_N/d_S ratio for each DENV-4 gene was measured using the pairwise method of Nei and Gojobori (1986) as implemented in the MEGA2 sequence analysis package (Kumar et al., 2001). Second, to obtain a codon-specific measure of selection pressures acting on the 60 E gene sequences (including a test for positive selection), the maximum likelihood method in the PAML sequence analysis package (program CODEML) was employed (Yang, 1997; Yang et al., 2000). This involved the statistical comparison of four models of codon evolution that differ in how d_N and d_S vary among codons or lineages. To analyze selection pressures at specific codons, we compared the M7 and M8 models; the former specifies that individual codons can take 1 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 the model only allows neutral evolution, while M8 also allows positive selection by adding an 11th category of codons at which d_N/d_S can exceed 1.0. Significant evidence for positive selection is obtained if M8 significantly 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-4 phylogeny, we compared (again using a likelihood ratio test) model M0, in which each lineage has the same d_N/d_S ratio, with FR (the "free ratio" model) in which lineages are allowed to take on individual values of d_N/d_S.

Rates of molecular evolution

To estimate the overall rate of nucleotide substitution of DENV-4 in Thailand, we compared the differences in branch lengths according to sampling time for 47 of the Thai viruses sampled here, which fell into genotype I of the virus. 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/) using a burn-in of 300 000 and a final chain length of 3 million, giving an effective sample size of 709 098. As with the maximum likelihood trees estimated above, the GTR + Γ + I model of nucleotide substitution was employed.

Acknowledgments

This research was supported by the U.S. Military Infectious Disease Research Program, Ft. Detrick, Maryland, USA. We thank Dr. Suchitra Nimmannitya and Dr. Siripen Kalayanarooj for the clinical grading of all specimens. We also thank the nursing staff of QSNICH and AFRIMS and the previous Chiefs of the Department of Virology at AFRIMS, particularly Dr. David Vaughn and Dr. Timothy Endy, for managing the service specimen collection. Dr. Alexei Drummond (University of Oxford) provided assistance with BEAST.

References

AbuBakar, S., Wong, P.-F., Chan, Y.-F., 2002. Emergence of dengue virus type 4 genotype IIA in Malaysia. J. Gen. Virol. 83, 2437–2442.

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

Bush, R.M., Bender, C.A., Subbarao, K., Cox, N.J., Fitch, W.M., 1999. Predicting the evolution of human influenza A. Science 286, 1921–1925.

Chungue, E., Cassar, O., Drouet, M.T., Guzman, M.G., Laille, M., Rosen, L., Deubel, V., 1995. Molecular epidemiology of dengue-1 and dengue-4 viruses. J. Gen. Virol. 76, 1877–1884.

Cummings, D.A.T., Irizarry, R.A., Huang, N.E., Endy, T.P., Nisalak, A., Ungchusak, K., Burke, D.S., 2004. Travelling waves in the occurrence of dengue haemorrhagic fever in Thailand. Nature 427, 344–347.

Foster, J.E., Bennett, S.N., Vaughan, H., Vorndam, V., McMillan, W.O., Carrington, C.V.F., 2003. Molecular evolution and phylogeny of dengue type 4 virus in the Caribbean. Virology 306, 126–134.

Gaunt, M.W., Sall, A.A., de Lamballerie, X., Falconar, A.K.I., Dzhivanian, T.I., Gould, E.A., 2001. Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. J. Gen. Virol. 82, 1867–1876.

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L.N., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303, 327–332.

Gubler, D.J., 2002. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. Trends Microbiol. 10, 100–103.

Holmes, E.C., 2003. Patterns of intra- and inter-host 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.

Jenkins, G.M., Pagel, M., Gould, E.A., de Zanotto, P.M.A., Holmes, E.C., 2001. Evolution of base composition and codon usage bias in the genus *Flavivirus*. J. Mol. Evol. 52, 383–390.

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.

Kuno, G., Chang, G.-J.J., Tsuchiya, K.R., Karabatsos, N., Cropp, C.B., 1998. Phylogeny of the genus *Flavivirus*. J. Virol. 72, 73–83.

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

Lanciotti, R.S., Calisher, C.H., Gubler, D.J., Chang, G.J., Vorndam, A.V., 1992. Rapid detection and typing of dengue viruses from clinical samples by using reverse transcriptase-polymerase chain reaction. J. Clin. Microbiol. 30, 545–551.

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

Mongkolsapaya, J., Dejnirattisai, W., Xu, X.-N., Vasanawathana, S., Tangthawornchaikul, N., Chairunsri, A., Sawasdivorn, S., Duangchinda, T., Dong, T., Rowland-Jones, S., Yenchitsomanus, P.-T., McMichael, A., Malasit, P., Screaton, G., 2003. Original antigenic sin and apoptosis in the pathogenesis of dengue hemorrhagic fever. Nat. Med. 9, 921–927.

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

Nisalak, A., Endy, T.P., Nimmannitya, S., Kalayanarooj, S., Thisayakorn, U., Scott, R.M., Burke, D.S., Hoke, C.H., Innis, B.L., Vaughn, D.W., 2003. Serotype-specific dengue virus circulation and dengue disease in Bangkok, Thailand from 1973–1999. Am. J. Trop. Med. Hyg. 68, 191–202.

Shurtleff, A.C., Beasley, D.W.C., Chen, J.J.Y., Ni, H., Suderman, M.T., Wang, H., Xu, R., Wand, 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.

Sittisombut, N., Sistayanarain, A., Cardosa, M.J., Salminen, M., Damrongdachakul, S., Kalayanarooj, S., Rojanasuphot, 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. Hyg. Med. 57, 100–108.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates. Sunderland, MA.

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

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

Uzcategui, N.Y., Comach, G., Camacho, D., Salcedo, M., Cabello de Quintana, M., Jimenez, M., Sierra, G., Cuello de Uzcategui, R., James, W.S., Turner, S., Holmes, E.C., Gould, E.A., 2003. The molecular epidemiology of dengue virus type 3 in Venezuela. J. Gen. Virol. 84, 1569–1575.

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

Williamson, S., 2003. Adaptation in the *env* gene of HIV-1 and evolutionary theories of disease progression. Mol. Biol. Evol. 20, 1318–1325.

Wittke, V., Robb, T.E., Thu, H.M., 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.

Woelk, C.H., Holmes, E.C., 2002. Reduced positive selection in vector-borne RNA viruses. Mol. Biol. Evol. 19, 2333–2336.

Worobey, M., Rambaut, A., Holmes, E.C., 1999. Widespread intra-serotype recombination in natural populations of dengue virus. Proc. Natl. Acad. Sci. U.S.A. 96, 7352–7357.

Yang, Z., 1997. 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.