Virology **301,** 148–156 (2002)
doi:10.1006/viro.2002.1549

Extinction and Rapid Emergence of Strains of Dengue 3 Virus during an Interepidemic Period

V. Wittke,* T. E. Robb,* H. M. Thu,* A. Nisalak,† S. Nimmannitya,† S. Kalayanrooj,† D. W. Vaughn,‡ T. P. Endy,‡ E. C. Holmes,§ and J. G. Aaskov*,[1]

*School of Life Sciences, Queensland University of Technology, Brisbane, Australia; †Queen Sirikit National Institute of Child Health, Bangkok, Thailand; ‡Department of Virology, Armed Forces Research Institute of Medical Science, Bangkok, Thailand; and §Department of Zoology, University of Oxford, Oxford, United Kingdom

Received January 30, 2002; returned to author for revision March 7, 2002; accepted May 9, 2002

Strains of dengue 3 (DEN-3) virus circulating in Thailand prior to 1992 appear to have disappeared from that location and to have been replaced by two new lineages which have evolved locally, rather than being introduced. Similar DEN-3 virus extinctions may have occurred previously in Thailand in 1962 and 1973. Although no causal relationship could be shown, this strain replacement event was accompanied by DEN-3 replacing DEN-2 as the serotype recovered most frequently from patients in Thailand. Although this implies a change in selection pressure, we found no evidence for positive natural selection at the level of either the E protein or the E protein gene. Further, the extinction of the pre-1992 strains and the appearance of the new lineages occurred during an interepidemic period, suggesting that a genetic bottleneck, rather than selection, might have been important in the emergence of these two new strains of virus. The pre-1992 DEN-3 virus lineage could still be found in 1998, to the west, in Myanmar. The ratio of nonsynonymous-to-synonymous nucleotide changes within a DEN-3 virus population from a single patient was less than the ratio among the consensus sequences of DEN-3 viruses from different patients, suggesting that many of the nonsynonymous nucleotide changes which occurred naturally in the E protein were deleterious and removed by purifying selection.

Key Words: dengue virus; evolution; Thailand.

INTRODUCTION

Human infection with dengue (DEN) viruses results in a spectrum of symptoms ranging from inapparent to haemorrhagic fever, hypovolemic shock, and death. It has been estimated that three billion people between the Tropics of Cancer and Capricorn are at risk of infection with dengue viruses (Gubler, 1998) and that between 500,000 and 50,000,000 cases of clinical infection occur annually (Gubler and Meltzer, 1999). There are four dengue virus serotypes (DEN-1 to DEN-4) and extensive genetic variation exists in some genes within each serotype. Phylogenetic analyses suggest that there are regional foci of virus evolution and virgin soil dengue outbreaks following introduction of viruses into susceptible populations (Chungue *et al.,* 1995; Hanna *et al.,* 2001; Lanciotti *et al.,* 1974, 1977; Lewis *et al.,* 1993; Rico-Hesse, 1990; Rico-Hesse *et al.,* 1997, 1998). There are no dengue vaccines available, commercially, after more than fifty years of research activity. However, attenuated vaccines are at various stages of development (Bhamarapravati and Yoksan, 1997; Sardelis *et al.,* 2000) and there is an extensive literature relating to efforts to develop dengue vaccines through recombinant DNA technology (reviewed in Trent *et al.,* 1997).

Although comparisons of consensus nucleotide sequences have provided some measure of genetic diversity of dengue viruses sampled from different localities, there have been fewer measures of genetic variation between the consensus sequences of viruses from the same location, recovered on a regular basis, over a prolonged period. There has been no measure of genetic variation *within* dengue virus populations from individual hosts, particularly for DEN-3. Equally, it is unclear what evolutionary processes shape the genetic diversity of dengue viruses in nature, most notably whether the virus responds to local ecological conditions and host immune pressures by fixing adaptively favourable mutations. Such analyses have been hindered by the chronological and geographical gaps in the virological record from areas where dengue is endemic, due to the failure to isolate dengue virus from patients and mosquitoes routinely.

Determination of the evolutionary processes acting on dengue viruses may have an impact on vaccination strategies. For a vaccine to be effective in all geographical locations, it must protect against infection with all dengue viruses. If virus populations in particular localities are very diverse, there is a high likelihood that they will be able to adapt rapidly to new ecological niches or escape the host immune response and, perhaps, pro-

[1] To whom correspondence and reprint requests should be addressed at School of Life Sciences, Queensland University of Technology, GPO Box 2434, Brisbane, Australia 4001. Fax: 61 7 3864 1534. E-mail: j.aaskov@qut.edu.au.

0042-6822/02 $35.00

TABLE 1

Nucleotides at 18 Hypervariable Sites in the E Gene of Thai DEN-3 Viruses

	Nucleotides at the positions indicated																	
Virus	952	982	1024	1030	1304	1328	1357	1447	1448	1459	1498	1543	1642	1966	2056	2197	2275	2370
D87.007	A	G	U	C	C	C	U	U	A	A	U	A	G	U	U	U	U	C
D87.1384	A	G	U	C	U	C	U	U	A	A	U	A	G	U	U	A	U	C
D87.561	A	G	U	C	U	C	U	U	A	A	U	A	G	C	U	A	U	C
D88.086	A	G	U	C	U	C	U	C	A	A	U	A	G	U	U	U	U	C
D89.273	A	G	U	C	U	C	U	C	A	A	C	A	G	U	U	A	U	C
D91.538	A	G	U	C	C	C	U	U	A	A	G	A	G	U	U	U	U	C
D91.393	A	G	U	C	C	C	U	C	A	A	U	A	G	U	U	C	U	C
D92.431	A	G	U	C	C	C	U	C	A	A	U	A	G	U	U	C	U	C
D92.423	A	A	U	C	C	U	C	U	G	G	C	G	A	C	C	U	C	U
D93.044	A	A	U	C	C	U	C	U	G	G	C	G	A	C	C	U	C	U
D93.674	G	G	C	U	C	C	C	U	A	G	U	G	A	C	C	U	C	U
D94.283	A	A	C	C	C	U	C	U	G	G	C	G	A	C	C	U	C	U
D94.122	G	G	C	U	C	C	C	U	A	G	U	G	A	C	C	U	C	U
D95.0014	A	A	U	C	C	U	C	U	G	G	C	G	A	C	C	U	C	U
D95.0400	G	G	C	U	C	C	C	U	A	G	U	G	G	C	C	U	C	U
D96.330	A	A	U	C	C	U	C	U	G	G	C	G	A	C	C	U	C	U
D96.313	G	G	C	U	C	C	C	C	A	G	U	G	A	C	C	U	C	U
D97.0144	A	A	U	C	C	U	C	U	G	G	C	G	A	C	C	U	C	U
D97.0106	G	G	C	U	C	C	C	C	A	G	U	G	A	C	C	U	C	U
D97.0291	G	G	C	U	C	C	C	U	A	G	U	G	A	C	C	U	C	U
	Deduced amino acid changes																	
					E124 Pro Ser	E132 His Tyr			E172 Ile Val									E479 Val Ala

duce viruses with new phenotypic properties. One way to examine this possibility is to study changes that are occurring in dengue virus populations in areas where there are high rates of infection and where the majority of the human population becomes immune early in life; that is, in localities where viruses are exposed to continuous immunological pressure. Thailand is one such country. Almost every year from 1973 until 1989, DEN-2 virus was recovered most frequently from patients admitted to the Queen Sirikit National Institute of Child Health (formally the Bangkok Children's Hospital) in Bangkok, Thailand. Since 1995, DEN-3 has replaced DEN-2 as the most frequently isolated virus and in the largest dengue epidemic in Thai history, in 1997–1998, DEN-3 was the most frequently isolated dengue virus serotype (A. Nisalak, unpublished observations).

Most epitopes involved in the serological neutralisation of dengue virus infection are located on the envelope (E) protein (Roehrig, 1997). The E protein may also contain determinants responsible for attachment of virus to host cells and for fusion of virions with host cell membranes (Crill and Roehrig, 2001). As such, it could be the target of nonimmunological selective pressures as the virus moves between nonimmune human hosts and mosquito vectors or the target of immunological pressure where hosts have preexisting antidengue virus antibody.

This study addressed the questions of how much and how quickly the E protein (gene) of dengue 3 viruses changed in areas of high transmission and high herd immunity. State-of-the-art techniques of evolutionary analysis were used to reconstruct the molecular epidemiological history of DEN-3 in Thailand and to determine the selection pressures acting on the virus.

RESULTS

Consensus nucleotide sequences were derived for the E protein gene of 21 DEN-3 viruses recovered from patients in Thailand infected between 1987 and 1997—two years in which large epidemics of dengue occurred, and from one recent (1998) DEN-3 isolate from Myanmar by sequencing cDNA derived by RT-PCR from viral RNA. Excluding virus D88.303 (for reasons outlined below), nucleotide variation occurred at 96 sites in the 1479 nucleotide E protein genes of these viruses with 13 changes resulting in amino acid changes. Only four of the amino acid changes were nonconservative (E124 Ser-Pro, E132 Tyr-His, E404 Ala-Thr, E414 Gly-Glu). At 45 of these 96 sites, only one nucleotide differed from the consensus sequence of all the isolates; at 28 sites, two

TABLE 2

Nucleotide and Deduced Amino Acid Variation between 20 E Protein Gene Clones from a Population of DEN-3 Virus (D92.431)

Position	Nucleotide change	Amino acid change	Clone with this change	Other examples from Genbank
1051	T-C		62	21, e.g., L11422[a]
1069	T-C		66	8, e.g., L11422
1090	A-G		65	
1123	A-G		05, 09	7, e.g., L11424
1154	T-C	E74 Cys-Arg	65	
1171	A-G		05, 09, 42	2, e.g., L11433
1188	A-G	E85 Glu-Gly	48	
1444	A-G		48	
1547	A-G	E205 Met-Val	09	
1563	G-A	E210 Trp-Stop	10	
1589	A-G	E219 Thr-Ala	58	
1652	A-G	E240 Asn-Asp	65	
1764	T-C	E277 Phe-Ser	55	
1801	A-G		66	
1915	A-T		54	
1924	T-C		42	2, e.g., L11423
2006	A-G	E358 Lys-Glu	62	
2062	A-G	E376 Ile-Met	57	
2094	T-C		37	
2095	T-C	E387 Ile-Thr	54	15, e.g., L11422
2221	A-G		62	
2271	T-C	E446 Phe-Ser	41	
2314	C-T		09, 42	7, e.g., L11422
2367	T-C	E478 Ile-Thr	38	
2370	C-T	E479 Ala-Val	55	12, e.g., L11422

[a] Genbank accession number.

nucleotides differed from the consensus sequence and at 5 sites, three nucleotides differed. Four or more nucleotide changes occurred at each of 18 of the 96 sites mentioned above (Table 1). Four of these sites contained three different nucleotides. The 18 sites at which four or more nucleotides varied from virus population consensus sequence were not randomly distributed throughout the E gene. Four were found between nucleotides 952 and 1030, six between nucleotides 1304 and 1459, and three between nucleotides 1966 and 2197. The nucleotide changes at 4 of these 18 sites results in amino acid changes (E124 Ser-Pro, E132 Tyr-His, E172 Ile-Val, E479 Val-Ala).

To determine the extent of sequence diversity within the DEN-3 consensus sequences, cDNA used for the determination of the nucleotide sequence of virus D92.431 was cloned into pGEM-T vectors and used to transform *Escherichia coli.* The E protein gene inserts in plasmids from 20 colonies of transformed bacteria were sequenced. The consensus sequence derived from the 20 clones was the same as that derived by sequencing the cDNA used in the cloning step. The E protein gene in plasmids from five bacterial colonies had the same nucleotide sequence as that derived from the original cDNA and the remainder varied from the consensus sequence at 25 sites (Table 2). The sequence of six clones varied from the consensus nucleotide sequences at one site, five at two sites, three at three sites, and one at four sites. The "variant" nucleotides found at 8 of these 25 sites had been reported previously for other DEN-3 viruses (Table 2). Twelve of the nucleotide changes resulted in amino acid changes and one produced a stop codon. The amino changes appeared to cluster in a nonrandom manner, i.e., E74,85, E205,210,210,240, E358,376,387, and E478,479. One of these cloned E protein genes had nucleotides which varied from that of the consensus sequence at positions 1764 (T-C, E277 Phe-Ser) and 2370 (C-T, E479 Ala-Val). The consensus sequence of all Thai DEN-3 isolates recovered after 1992 had T at position 2370 in contrast to the C at this site in the pre-1992 samples (see below).

To determine the evolutionary history of the Thai dengue viruses, the nucleotide sequence of their E protein genes were compared with those from all previously published E gene sequences of DEN-3 virus available in GenBank (excluding putative recombinants; Worobey *et al.,* 1999). Phylogenetic analyses of these 50 nucleotide sequences separated the viruses into five main groups, which may be considered genotypes (Fig. 1). These genotypes are similar to those identified by Lanciotti *et al.* (1994), although our phylogenetic analysis shows that genotype I of Lanciotti *et al.* (1994) encompasses two distinct groups of viruses that we have classified separately by constructing a new "genotype V." As in previous studies of DEN-3 diversity, the 1963 Puerto Rico strain

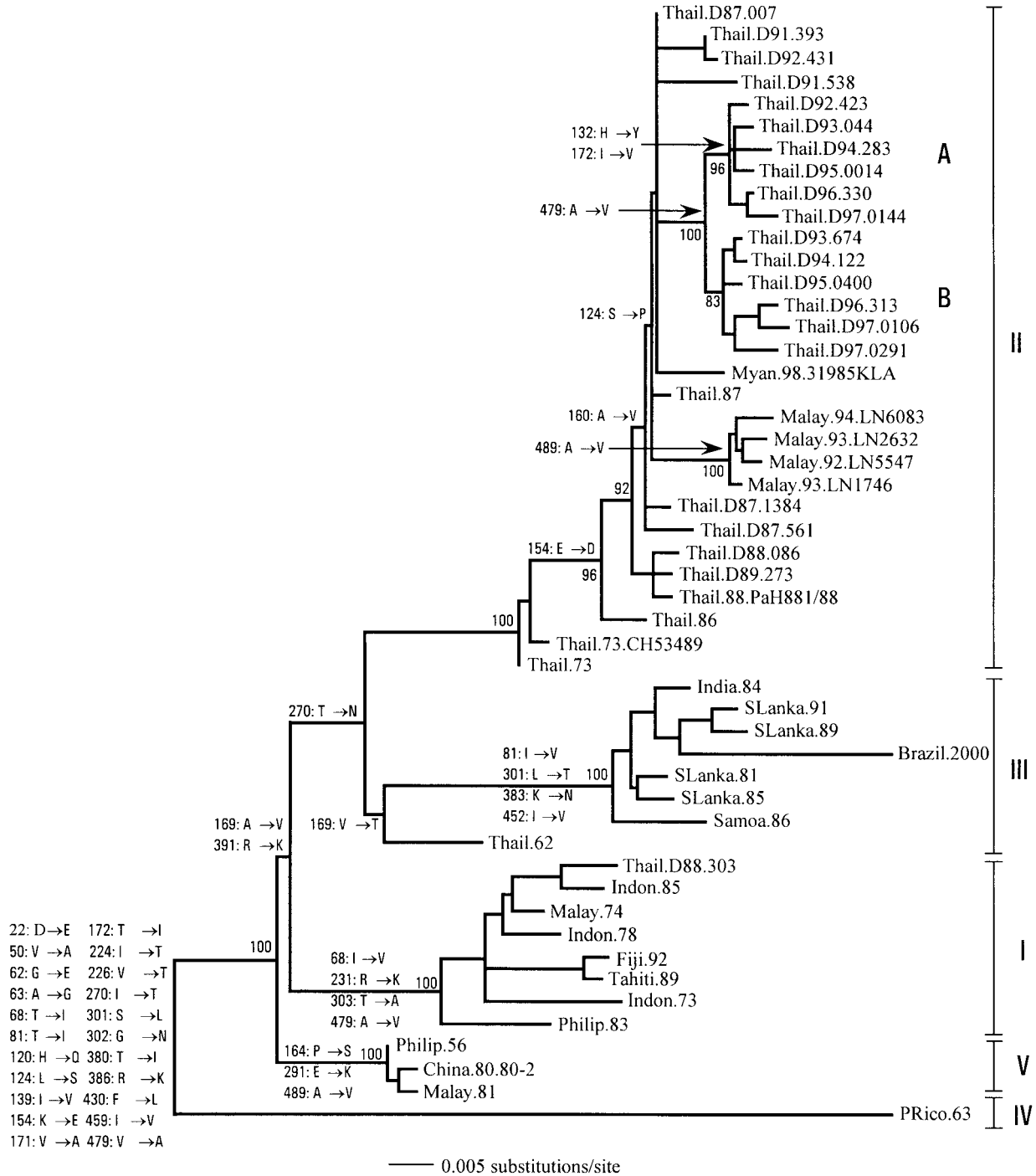

FIG. 1. Maximum likelihood phylogenetic tree of E gene sequences from a global sample of 50 DEN-3 viruses. The genotypes are labelled according to the scheme of Lanciotti *et al.* (1994), with genotype V newly identified here, and the amino acid changes distinguishing each group of viruses are marked on the tree. The tree is mid-point rooted for purposes of clarity only and all horizontal branch lengths are drawn to scale. Bootstrap support values are shown for key nodes only.

formed a distinct outlier. Differing numbers of amino acid replacements have occurred on the lineages leading to the five genotypes (Fig. 1). The distinct 1963 Puerto Rico strain is separated from other DEN-3 viruses by 22 amino acid changes dispersed across the E protein. In contrast, genotype I is distinguished by four amino acid replacements, genotype V by three such replacements, while only a single change separates all members of genotype III, although this also rises to five if the divergent "Thail.62" strain is excluded. Finally, it is noticeable that not a single amino acid change has occurred on the lineage leading to genotype II.

The viruses sampled in this study fell into a number of locations on the global phylogeny. Virus D88.303, isolated in Thailand in 1988, clustered with viruses from Indonesia, Malaysia, the Philippines, and the South Pacific in genotype I. The remaining Thai isolates fell into genotype II, therein clustering with a large group of viruses including some previously sampled from Thailand as well as four viruses from Malaysia (strains LN1746, LN2632, LN5547, LN6083; unpublished, direct submission to GenBank). There also was an important temporal division among the genotype II Thai viruses. Specifically, strain D92.423 and all others sampled after 1992 form a phylogenetic group that is distinct (100% bootstrap support) from all earlier isolates and a second 1992 strain, D92.431. Consequently, it appears that the post-1992 strains have replaced the pre-1992 strains, with isolate D92.423 the first representative of the new viruses. A DEN-3 isolate made from a patient in Myanmar in 1998 (3198KLA) also clustered with the pre-1992 Thai viruses. Finally, the post-1992 groups of viruses can be further subdivided into two cocirculating lineages (designated "A" and "B" on Fig. 1), each with strong bootstrap support.

The replacement of the pre-1992 viruses that occurred in Thailand could have been due to natural selection—so that the later viruses had some phenotypic characteristic that allowed them to out-compete the preexisting viruses—or be a chance occurrence, perhaps caused by a population bottleneck. To choose between these hypotheses, the selection pressures acting on these sequences were examined in more detail. First, the ratio of nonsynonymous (d_N) to synonymous (d_S) changes (ratio $d_N/d_S = \omega$) was analysed using a recently developed maximum likelihood approach which has been found to provide a more accurate measure of selective pressure than earlier pairwise methods (Yang and Bielawski, 2000). Application of this method to the DEN-3 data provided a picture of relatively strong selective constraints and no good evidence for positive selection. Specifically, none of the models of codon evolution that allow for positive selection were favoured over neutral alternatives. Under the best supported model of codon evolution (M3), the majority (83%) of nucleotide sites was very highly constrained ($d_N/d_S = 0.0001$), a smaller proportion (15.6%) was moderately constrained ($d_N/d_S = 0.261$), and a single site, codon 380, fell into the ambiguous zone between neutral evolution and weak positive selection ($d_N/d_S = 1.346$). (Full results available from the authors on request.) Furthermore, the branch leading to the post-1992 Thai viruses is characterised by eight nucleotide substitutions (nucleotide positions 1357, 1459, 1543, 1642, 1966, 2056, 2275, 2370), only one of which resulted in the change of an amino acid (E479, A → V).

The possibility of selection acting at synonymous codon positions was considered by examining patterns of codon usage bias and RNA secondary structure. Codon usage by the pre-1992 viruses and the two more recent groups of Thai DEN-3 viruses were similar so it is unlikely that this represents a selected difference between them (results not shown, available on request). The possibility of selection acting on RNA secondary structure stems from the apparently nonrandom distribution of the 16 hypervariable sites. The free energy of predicted RNA secondary structures in the three regions of the E gene where hypervariable sites clustered were compared using representative members of each of three groups of Thai DEN-3 viruses—pre-1992, and the two lineages of post-1992 viruses. Viruses selected varied only at one or more of the hypervariable sites in the region under study. The free energy of the predicted RNA secondary structure of the first region (nucleotides 935–1175) varied from −51.4 to −54.4 kcal/mol between the three groups, that of the second region (nucleotides 1356–1575) from −37.5 to −42.4 kcal/mol, and for the third region (nucleotides 2115–2401) from 46.7 to −48.2 kcal/mol. Hence, there is little evidence that the post-1992 Thai viruses gained a selective advantage through changes in RNA secondary structure. Finally, no evidence was found that the post-1992 viruses had arisen by recombination between DEN-3 viruses circulating before that date in either Thailand or elsewhere, as there was no evidence strains moved position across the tree as expected under recombination.

DISCUSSION

The data presented here suggest that strains of DEN-3 virus circulating in Thailand up to 1992 have disappeared from that country but are still circulating in at least one neighbouring locality (Myanmar, represented by strain 31985KLA sampled in 1998). Since 1992, two new lineages of DEN-3 viruses have appeared from a single common ancestor and are cocirculating. There also appeared to have been an incursion into Thailand by an "Indonesian/Malaysian-like" strain of DEN-3 in 1988 (represented by strain D88.303) but this failed to become established as none of the later Thai samples grouped with this isolate. It is unclear whether this was due to competition with more fit strains of DEN-3 virus circulating in Thailand, to regional differences in mosquito vector competence for each of these viruses (Gubler *et al.,* 1979), or because the scope of the virus sampling was too limited to detect "rare" populations of virus. Finally, the distinct temporal nature of the DEN-3 phylogeny—with progressively more recent viruses tending to fall toward the tips of the tree—suggests that there might be regular extinctions of strains of DEN-3 virus in Thailand (e.g., the 1962, 1973, and the 1986–1992 strains) and in other localities.

More difficult to discern are the evolutionary causes of the extinctions and replacements documented, in particular, whether they are due to underlying differences in

fitness or to the vagaries of population bottlenecks. It is intriguing that by 1993 the pre-1992 strains of DEN-3 disappeared from the sites studied in this project, and by 1995 (2–3 years after the putative genetic bottleneck) DEN-3 had become the most frequently isolated dengue virus serotype in Thailand (A. Nisalak, personal communication). This suggests that the strain replacement we document was associated with the emergence of DEN-3 as the dominant dengue serotype, compatible with a selectively driven event. However, other than on the branch leading to the Puerto Rico 1963 strain, the E gene of DEN-3 is relatively conserved at the amino acid level. However, four amino acid changes occurred more frequently within the genotype II Thai strains—E124 S-P, E132 H-Y, E172 I-V, E479 V-A (Table 1). Those at E124 (S-P) and E132 (H-Y) were nonconservative and there are several reports of amino acid changes in the E protein of dengue viruses in the vicinity of amino acids E124 and E132 that influence neurovirulence for mice (possibly as a result of changes in cell trophism; Bray *et al.,* 1998; Gualano *et al.,* 1998) and cell fusion (Guirakhoo *et al.,* 1993) and this region contains significant linear and, possibly, conformational serological epitopes (Aaskov *et al.,* 1989; Innis *et al.,* 1989; Lok *et al.,* 2001). There also is evidence of linear and conformational epitopes in the region E160–172 of dengue viruses (Aaskov *et al.,* 1989; Serafin and Aaskov, 2001), where one of the conservative amino acid changes in the DEN-3 virus E protein was observed (E172 I-V). It is more difficult to identify a selective advantage associated with a conservative amino acid change in the second, putative, transmembrane region of the DEN-3 E protein (E479 A-V), although this is the only amino acid difference between the pre- and post-1992 viruses. The change to a more hydrophobic amino acid (V) might aid insertion, or retention, of this region in the lipid envelope of the virion or alternatively it might influence the efficiency of the posttranslational cleavage of E-NS1. These changes were all quite distinct from the E390, domain III, amino acid changes that has been suggested to influence virulence of DEN-2 virus in humans (Leitmeyer *et al.,* 1999) and which was also found to be associated with variation in neurovirulence of DEN-2 virus in mice (Sanchez and Ruiz, 1996).

Despite these putative functional differences, statistical analyses of the selection pressures acting on amino acid sites in the E gene provided no compelling evidence for adaptive evolution (positive selection) on those lineages where strain replacement was most apparent. Although this rules out natural selection on the envelope glycoprotein as the underlying cause of strain replacement, it is possible that adaptively useful mutations have been fixed in other genes or noncoding regions, but this could only be resolved with sequencing the complete genome. Furthermore, it is possible that natural selection was acting on RNA secondary structure or on codon usage. There is at least one example of the tissue-specific availability of t-RNA influencing the replicative capacity of a virus (Zhou *et al.,* 1999), although this has not been documented in RNA viruses. However, there was no evidence of a change in codon usage with evolution of the Thai DEN-3 strains. Also, while accepting the limitations of RNA secondary structure prediction programs such as MFold (Zuker, 1989), there was no evidence that regions of the E protein gene undergoing evolutionary change in these viruses were evolving to more stable forms of RNA secondary structure.

On the other hand, there are good reasons why stochastic processes may have a major effect on the genetic diversity of DEN-3 virus in Thailand. Mosquito transmission of virus to a human host requires as little as 1–3 μl of infectious saliva, providing an ideal opportunity for a genetic bottleneck in a virus population. On the other hand, both theoretical and laboratory studies suggest that a genetic bottleneck is most likely to give rise to a virus population that is less fit than the one from which it was derived (Escarmis *et al.,* 1996; Novella *et al.,* 1999). Furthermore, 1992 was an interepidemic year when the number of dengue cases in Thailand was relatively low, suggesting that virus transmission would also have been low (41,125 cases in 1992 compared with 174,285 in 1987 and 101,689 in 1997; A. Nisalak, unpublished observations). In contrast, in Myanmar, where the pre-1992 DEN-3 genotype was still circulating in 1998, DEN-1, 2, and 3 viruses were being recovered from patients in approximately similar numbers (H. M. Thu, unpublished observations). Moreover, there is a previous report suggesting that a genetic bottleneck occurred between 1980 and 1987 in Thailand, which may have been responsible for a change in DEN-2 virus populations over that time interval (Sittisombut *et al.,* 1997). In this study, which covered similar areas over an equivalent time interval, DEN-2 populations were found to be stable and homogeneous (Rico-Hesse *et al.,* 1998) at a time when DEN-3 viruses were undergoing rapid change.

Finally, the variation between the nucleotide sequences of 20 E protein genes within a single virus population represented by a single consensus sequence (D92.431) was, as expected, significantly less than the variation seen between different virus consensus sequences. Furthermore, the number of synonymous (12) and nonsynonymous (13) nucleotide changes within a single DEN-3 virus consensus sequence was almost equal (Table 2), whereas between different viral consensus sequences, the number of synonymous changes greatly exceeded the number of nonsynonymous changes 83:13. This difference in d_N/d_S ratios indicates that the E gene sequences contained within a single consensus sequence are a measure of the total mutant spectrum, which would predict more even numbers of synonymous and nonsynonymous changes, and that many of the nonsynonymous changes are deleterious and are subsequently removed by purifying selection.

TABLE 3

Patients, Location, and Passage History of DEN-3 Viruses Used in this Study

Virus[a]	Disease[b]	Infection	Passage[c]	Location	Virus	Disease	Infection	Passage	Location
D87.007	DF	Primary	Ts-1, C6-36-4	Bangkok	D93.674	DF	Secondary	Ts-1, C6-36-2	Bangkok
D87.1384	DHF III	Secondary	Ts-1, C6-36-2	Bangkok	D94.122	DHF III	Secondary	Ts-1, C6-36-2	Surin
D87.561	DHF II	Secondary	Ts-1, C6-36-2	Patumthani	D94.283	DHF II	Secondary	Ts-1, C6-36-2	Samutprakan
D88.086	DHF III	Secondary	Ts-1, C6-36-2	Bangkok	D95.0014	DHF II	Primary	Ts-1, C6-36-2	Samutprakan
D88.303	DHF III	Secondary	Ts-1, C6-36-2	Patumthani	D95.0400	DHF III	Primary	Ts-1, C6-36-2	Bangkok
D89.273	DHF III	Secondary	Ts-1, C6-36-2	Bangkok	D96.313	DF	Primary	Ts-1, C6-36-2	Bangkok
D91.393	DHF I	Primary	Ts-1, C6-36-2	Bangkok	D96.330	DHF II	Secondary	C6-36-4	Surin
D91.538	DHF II	Secondary	Ts-1, C6-36-3	Bangkok	D97.0106	DHF III	Secondary	C6-36-4	Pratumthani
D92.423	DHF II	Secondary	Ts-1, C6-36-2	Bangkok	D97.0144	DHF II	Secondary	C6-36-4	Bangkok
D92.431	DF	Primary	Ts-1, C6-36-2	Bangkok	D97.0291	Unknown	Secondary	C6-36-4	Samutpraken
D93.044	DHF II	Primary	Ts-1, C6-36-2	Nonthaburi	31985KLA	DHF II	Secondary	C6-36-1	Yangon, Myanmar

[a] The first two numerals indicate the year of isolation and those after the point, the sample number. Sample 31985KLA was isolated in 1998.
[b] DF, dengue fever; DHF, dengue haemorrhagic fever (I, II, II grade I, II, or III).
[c] Ts1, passaged once in *T. splendens* mosquitoes. C6-36 -2, 3, 4, passaged 2, 3, or 4 times in C6-36 mosquito cells.

The virus containing a genome with a stop codon (Table 2, clone 10, nt 1563) is unlikely to be viable and may represent a defective interfering particle. While it is possible that some of the changes observed could have been due to errors introduced by *Taq* polymerase and not removed by *Pwo* during PCR, the error rate should have been the same for the cDNA used to produce the clones and for the cDNA used to determine the nucleotide sequences of the viral genome and so might not be expected to alter these conclusions.

In sum, this study provides an example of how quickly dengue virus gene sequences, and perhaps phenotypes, can change in local populations. It also is salutary to observe, in a region where all for dengue serotypes are circulating, that rapid change can occur in one serotype at times of relative stasis in a second and serves to highlight, yet again, the differences between the disease patterns of each of the dengue virus serotypes (Vaughn *et al.,* 2000).

MATERIALS AND METHODS

Virus

Viruses, their origins, and passage histories are shown in Table 3. Most Thai viruses were isolated in *Toxorhynchites splendens* mosquitoes. All viruses were passaged at least once in C6-36 *Aedes albopictus* mosquito cells before sequencing.

Nucleotide sequencing

Nucleotides have been numbered from the 5′ end of the DEN-3 genome according to the scheme of Osatomi and Sumiyoshi (1990). Deduced amino acid sequences have been numbered from the amino-terminal of the E protein.

RNA was recovered, using viral RNA extraction columns (Qiagen, Australia) according to the manufacturer's instructions, from virus in supernatant from cultures of C6-36 cells infected 7 days previously. The RNA was transcribed to cDNA using random hexamer oligonucleotides (Roche, Germany) and avian myeloblastosis virus (AMV) reverse transcriptase (Roche) and incubations of 55°C for 10 min followed by 45°C for 60 min. The cDNA was amplified using sense (P722) and antisense (CP 2550) oligonucleotide primers (Lanciotti *et al.,* 1994) and "Expand polymerase" (a mixture of *Taq* and *Pwo* polymerases; Roche) with cycling times as follows

94°C	2 min	1 cycle
94°C	30 s	
45°C	30 s	35 cycles
68°C	90 s	
68°C	5 min	1 cycle

The PCR product was analysed on 1.5% w/v agarose/Tris acetate–EDTA (TAE) gels and the cDNA band of interest excised and purified using silica "glass milk" according to the manufacturer's instructions (Bresatech, Australia).

This cDNA was sequenced (see below) or in the case of some cDNA from virus D92.431, "tailed" with ATP to allow cloning into pGEM-T plasmids (Promega). Briefly, the cDNA was incubated with 0.1 μmol dATP and 1 U *Taq* polymerase (Perkin Elmer) for 1 h at 72°C to add ATP at the 3′ end of each strand of DNA and then purified using a High Pure PCR Product Purification Kit (Roche). Eight microliters of the A-cDNA was ligated into 50 ng plasmid by incubating the mixture at 4°C overnight in the presence of T4 ligase (Promega). Competent *E. coli* (DH5 alpha; Life Technologies, Australia) were transformed with 4 μl of the ligation mix and then plated on Lauria–Bertani agar supplemented with 50 μg/ml ampicillin, 0.5 mM ispropyl-β-D-thiogalactopyranoside, and 80 μg/ml

5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside and incubated at 37°C for 24 h. Discrete white colonies from the plates were inoculated into 2 ml SOC broth and incubated for 24 h at 37°C with shaking at 225 rpm. Plasmids were recovered from each broth culture using a Plasmid Purification Kit according to the manufacturer's instructions (Qiagen).

cDNA and plasmid inserts were sequenced using the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer) according to the manufacturer's instructions and the products were anlaysed on an Applied Biosystems Nucleotide Sequencer. The oligonucleotide primers of Lanciotti *et al.* (1994) were used to sequence both strands of the ds-cDNA. Where sequence ambiguity was detected, cDNA was resequenced or fresh cDNA prepared and sequenced.

Nucleotide sequences were aligned; codon usage was quantitated, and RNA secondary structure predictions were performed with software (ClustalW, codonfrequency, Mfold, respectively) available online from the Australian National Genomic Information Service (www.angis.org.au).

Evolutionary analysis

The nucleotide sequences generated in this study were combined with those of all other DEN-3 E protein gene sequences available on GenBank (Table 4), excluding those previously shown to have evidence of recombination (Worobey *et al.,* 1999). This resulted in a total data set of 50 sequences.

Phylogenetic trees were estimated using the maximum likelihood (ML) method available in the PAUP* phylogeny reconstruction package (Swofford, 2000). The general time-reversible (GTR) model of nucleotide substitution was employed, with the rate of each substitution (A-C, A-G, etc.), the base frequencies, and the rates of nucleotide substitution at each codon position that gave the highest likelihood estimated from the data. Parameter values are available from the authors on request. Successive rounds of branch-swapping were used to find the most likely tree under this model. The robustness of each node on the tree was assessed using the bootstrap resampling method, with all 1000 replicates estimated using the neighbor-joining procedure, but with the input genetic distances produced under the ML substitution model.

The ML method of Yang *et al.* (2000) (implemented in the PAML package; Yang, 1997) was used to examine selection pressures in DEN-3. This approach is based on an analysis of the numbers of nonsynonymous (d_N) and synonymous (d_S) substitutions per site for each codon, using various models of codon evolution. The model with the highest likelihood best explains the data. If this model shows $d_N/d_S < 1$, then the predominant evolutionary process is purifying selection, whereas $d_N/d_S = 1$ reflects neutral evolution and $d_N/d_S > 1$ indicates positive selection. Bayesian methods can also be used to calculate the probability that a particular codon falls into a specific site class, such as those thought to be positively selected. The simplest codon model, M0, estimates a single d_N/d_S for all sites. In contrast, M1 can be thought of as a neutral model as it divides codons into two categories, one for those sites that are invariant ($d_N/d_S = 0$), and second category, representing neutral codon sites, where $d_N/d_S = 1$. M2 can account for positive selection because a third category of codons is incorporated at which d_N/d_S (estimated from the data) can be >1. In M3, the d_N/d_S ratio is estimated from the data for three site classes whose size is also determined from the data. This model also allows for positive selection because the d_N/d_S for any codon class can be >1. Both M7 and M8 use a discrete beta distribution (with 10 categories) to model d_N/d_S ratios among codons but M8 differs from M7 in that it estimates an 11th category of codon sites at which d_N/d_S can be >1. Because some models are nested within each other—both M0 and M1

TABLE 4

DEN 3 Virus Nucleotide Sequences from GenBank Used in Phylogenetic Analyses

Virus	GenBank accession number	Comment*
Fiji.92	L11422	Fiji, 1992
India.84	L11424	India, 1984
Indon.73	L11425	Indonesia, 1973
Indon.78	L11426	Indonesia, 1978
Indon.85	L11428	Indonesia, 1985
Malay.74	L11429	Malaysia, 1974
Malay.81	L11427	Malaysia, 1981[a]
Philip.56	L11423	Philippines, 1956
Philip.83	L11432	Philippines, 1983
PRico.63	L11433	Puerto Rico, 1963
Samoa.86	L11435	Samoa, 1986[a]
SLanka.81	L11431	Sri Lanka, 1981
SLanka.85	L11436	Sri Lanka, 1985
SLanka.89	L11437	Sri Lanka, 1989
SLanka.91	L11438	Sri Lanka, 1991
Tahiti.89	L11619	Tahiti, 1989
Thail.62	L11440	Thailand, 1962
Thail.73	L11620	Thailand, 1973
Thail.86	L11441	Thailand, 1986
Thail.87	L11442	Thailand, 1987
Brazil2000	AY038605	Brazil, 2000[a]
CH53489	M86733	Thailand, 1973[a]
PaH881/88	AF349753	Thailand, 1988[a]
80-2	AF317645	China, 1980
LN1746	AF147458	Malaysia, 1993
LN2632	AF147459	Malaysia, 1993
LN5547	AF147457	Malaysia, 1992
LN6083	AF147460	Malaysia, 1994

[a] There is no information about the sources of these viruses accompanying their GenBank entry. All other viruses in the table were recovered from humans.

are nested with M2 and M3, and M7 is nested with M8—these can be compared directly using standard likelihood ratio test statistics.

ACKNOWLEDGMENTS

This study was supported by grants from the Lee Foundation, the Wellcome Trust, and The Royal Society.

REFERENCES

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

Bhamarapravati, N., and Yoksan, S. (1997). Live attenuated dengue vaccine. *In* "Dengue and Dengue Haemorrhagic Fever." pp. 367–378. CAB International, New York.

Bray, M., Men, R., Tokimatsu, I., and Lai, C.-J. (1998). Genetic determinants responsible for acquisition of dengue type 2 virus mouse neurovirulence. *J. Virol.* **72,** 1647–1651.

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

Crill, W. D., and Roehrig, J. T. (2001). Monoclonal antibodiesthat bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells. *J. Virol.* **75,** 7769–7773.

Duarte, E., Novella, I. S., Ledesma, S., Clarke, S., Moya, A., Elena, S. F., Domingo, E., and Holland, J. J. (1994). Subclonal components of consensus fitness in an RNA virus clone. *J. Virol.* **68,** 4295–4301.

Escarmis, C., Davila, M., Charpentier, N., Bracho, A., Moya, A., and Domingo, E. (1996). Genetic lesions associated with Muller's Ratchet in an RNA virus. *J. Mol. Biol.* **264,** 255–267.

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

Gubler, D. J., and Meltzer, M. (1999). Impact of dengue/dengue haemorrhagic fever on the developing world. *Adv. Virus Res.* **53,** 35–70.

Gubler, D. J., Nalim, S., Tan, R., Saipan, H., and Sulianti-Saroso, J. (1979). Variation ni susceptibility to oral infection with dengue viruses amng geographic strains of *Aedes aegypti. Am. J. Trop. Med. Hyg.* **28,** 1045–1052.

Guirakhoo, F., Hunt, A. R., Lewis, J. G., and Roehrig, J. T. (1993). Selection and partial characterisation of dengue 2 virus mutants that induce fuson at elevated pH. *Virol.* **194,** 219–223.

Gulano, R. C., Pryor, M. J., Cauchi, M. R., Wright, P. J., and Davidson, A. D. (1998). Identification of a major determinant of mouse neurovirulence of dengue type 2 virus using stably cloned genomic length cDNA. *J. Gen. Virol.* **79,** 437–446.

Hanna, J. N., Ritchie, S. A., Phillips D. A., Serafin, I. L., Hills, S. L., van den Hurk, A. F., Pyke, A. T., McBride, W. J. H., Amadio, M. G., and Spark, R. L. (2001). An epidemic of dengue 3 in far north Queensland, 1997–1999. *Med. J. Aust.* **174,** 178–182.

Innis, B. L., Thirawuth, V., and Hemachuda, C. (1989). Identification of continuous epitopes of the envelope glycoprotein of dengue type 2 virus. *Am. J. Trop. Med. Hyg.* **40,** 676–687.

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

Lanciotti, R. S., Lewis, J. G., Gubler, D. J., and Trent, D. W. (1994). Molecular evolution and epidemiology of dengue 3 viruses. *J. Gen. Virol.* **75,** 65–75.

Lewis, J. A., Chang, G.-J., Lanciotti, R. S., Kinney, R. M., Mayer, L. W., and Trent, D. W. (1993). Phylogenetic relationships of dengue 2 viruses. *Virology* **197,** 216–224.

Leitmeyer, K. C., Vaughn, D. W., Watts, D. M., Salas, R., Villalobos de Chacon, I., Ramos, C., and Rico-Hesse, R. (1999). Dengue virus structural differences that correlate with pathogenesis. *J. Virol.* **73,** 4738–4747.

Lok, S. M., Ng, M. L., and Aaskov, J. G. (2001). Amino acid and phenotypic changes in dengue 2 virus associated with escape from neutralisation by IgM antibody. *J. Med. Virol.* **65,** 315–323.

Novella, I. S., Quer, J., Dominingo, E., and Holland, J. J. (1999). Exponential fitness gains of RNA virus populations are limited by bottleneck effects. *J. Virol.* **73,** 1668–1671.

Osatomi, K., and Sumiyoshi, H. (1990). Complete nucleotide sequence of dengue type 3 virus genome RNA. *Virology* **176,** 643–647.

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

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

Rico-Hesse, R., Harrison, L. M., Nisalak, A., Vaughn, D. W., Kalayanarooj, S., Green, S., Rothman, A. L., and Ennis, F. A. (1998). Molecular evolution of dengue type 2 virus in Thailand. *Am. J. Trop. Med. Hyg.* **58,** 96–101.

Roehrig, J. T. (1997). Immunochemistry of dengue viruses. *In* "Dengue and Dengue Haemorrhagic Fever." pp. 199–220. CAB International, New York.

Sanchez, I. J., and Ruiz, B. H. (1996). A single nucleotide change in the E protein of dengue virus 2 Mexican strain affects neurovirulence in mice. *J. Gen. Virol.* **77,** 2541–2545.

Sardelis, M. R., Edelman, R., Klein, T. A., Innis, B. L., Putnak, J. R., Jones, J. W., and Turell, M. J. (2000). Limited potential for transmission of live dengue virus vaccine candidates by *Aedes aegypti* and *Aedes albopictus. Am. J. Trop. Med. Hyg.* **62,** 698–701.

Serafin, I. L., and Aaskov, J. G. (2001). Identification of epitopes on the envelope protein of dengue 2 and dengue 3 viruses using monoclonal antibodies. *Arch. Virol.* **146,** 2469–2479.

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

Swofford, D. L. (2000). Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, MA.

Trent, D. W., Kinney, R. M., and Huang, C. Y.-H. (1997). Recombinant dengue vaccines. *In* "Dengue and Dengue Haemorrhagic Fever." pp. 379–404. CAB International, New York.

Vaughn, D. W., Green, S., Kalayanarooj, S., Innis, B., Nimmannitya, S., Suntayakorn, S., Endy, T., Raengsakulrach, B., Rothman, A. L., Ennis, F. A., and Nisalak, A. (2000). Dengue viraemia titer, antibody response pattern and virus serotype correlate with disease severity. *J. Infect. Dis.* **181,** 2–9.

Worobey, M., Rambaut, A., and Holmes, E. C. (1999). Widespread intraserotype recombination in natural populations of dengue virus. *Proc. Natl. Acad. Sci. USA* **96,** 7352–7357.

Yang, Z. (1997). PAML: A program package for phylogenetic analysis by maximum likelihood. *CABIOS* **13,** 555–556.

Yang, Z., and Bielawski, J. P. (2000). Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15,** 496–502.

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

Zhou, J., Liu, W. J., Peng, S. W., Sun, X. Y., and Frazer, I. (1999). Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. *J. Virol.* **73,** 4972–4982.

Zuker, M. (1989). On finding all suboptimal foldings of an RNA molecule. *Science* **244,** 48–52.