ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

VIROLOGY

Virology 306 (2003) 126–134

www.elsevier.com/locate/yviro

Molecular evolution and phylogeny of dengue type 4 virus in the Caribbean

Jerome E. Foster,[a,1] Shannon N. Bennett,[b,1] Helen Vaughan,[c] Vance Vorndam,[d] W. Owen McMillan,[b] and Christine V.F. Carrington[a,*]

[a] *Department of Pre-Clinical Sciences, Faculty of Medical Sciences, University of the West Indies, St. Augustine, Trinidad*
[b] *Department of Biology, University of Puerto Rico, Rio Piedras, San Juan, Puerto Rico*
[c] *Caribbean Epidemiology Centre, 16-18 Jamaica Boulevard, Federation Park, Port of Spain, Trinidad*
[d] *Centers for Disease Control and Prevention, San Juan Dengue Branch, San Juan, Puerto Rico*

Received 4 June 2002; returned to author for revision 30 September 2002; accepted 9 October 2002

Abstract

We sequenced the E gene and adjacent prM/M and NS1 junctions (1940 bp) of 48 Dengue-4 (DEN-4) isolates collected between 1981 and 1999 from 8 Caribbean islands and from 7 South and Central American countries. Phylogenetic analysis confirms a single introduction in the early 1980s and a high degree of gene flow resulting in a pattern of evolution defined more by time period than geographic origin, especially within the Caribbean basin. A modern Caribbean clade consisting of four distinct lineages has arisen, comprised of isolates from Caribbean islands and nearby regions of South America. This clade is defined by three amino acid substitutions in the E (aa 163 and 351) and NS1 (aa 52) proteins. These findings highlight the importance of migration and gene flow in dengue viral change and suggest that efforts to understand disease dynamics in the Caribbean basin need to focus at regional, rather than local scales.

Keywords: Dengue; DEN-4; Phylogeny; Evolution; Caribbean; Envelope; NS1

Introduction

Dengue virus is a single-stranded, positive-sense RNA virus belonging to the genus *Flavivirus*, family *Flaviviridae*. There are four antigenically distinct serotypes (DEN-1–4), all of which can cause dengue fever (DF), a relatively mild febrile illness, or the potentially fatal dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS). The four serotypes are closely related, and evidence suggests that they represent at least three independent introductions from sylvatic primates into human populations, with the most recent occurring within the last 100 years (Wang et al., 2000). Considered among the most important re-emergent infectious diseases, dengue virus annually infects more than 50 million people globally, leading to approximately 10,000 DHF/DSS-related infant deaths (WHO, 1996).

The emergence of epidemic dengue in the Americas, as well as worldwide, has been characterized by a rise in hyperendemicity (the co-occurrence of multiple dengue serotypes in the same locality). Before the 1980s, regional outbreaks of dengue in the Americas were caused by single serotypes and were geographically restricted and largely self-limiting (Gubler, 1998). However, the elimination of mosquito control programs and increases in urbanization and global travel have significantly changed the epidemiologic landscape of the disease (Gubler, 1997; Gubler, 1998). Since the early 1980s, dengue virus epidemics in the region have become steadily larger and more frequent with a concomitant rise in the numbers of DHF/DSS cases (WHO, 1996). All four serotypes have now been reported in the Caribbean basin and most epidemics have consisted of a mixture of DEN serotypes (Dietz et al., 1996). This mounting hyperendemicity is believed to be one of the most

* Corresponding author. Fax: +1-868-662-1873.
E-mail address: ccarrington@fms.uwi.tt (C.V.F. Carrington).
[1] Contributed equally to the manuscript.

0042-6822/03/$ – see front matter © 2003 Elsevier Science (USA). All rights reserved.
PII: S0042-6822(02)00033-8

significant factors contributing to increases in disease severity and DHF/DSS (Gubler, 1997). There is strong correlation between disease severity and secondary infection with a heterologous serotype: individuals experiencing a second infection are at least 15 times more likely to contract DHF/DSS (Halstead, 1988; Kliks et al., 1989; Thein et al., 1997). Viral factors such as strain virulence may also play a role in determining disease severity and thus contribute to the changing pattern of the disease (Kliks, 1990; Morens et al., 1991; Rico-Hesse et al., 1997; Vaughn et al., 2000). There is mounting evidence that genetic change can lead to shifts in the epidemic potential or pathogenicity of dengue virus serotypes and their relatives (Gubler et al., 1981; Leitmeyer et al., 1999; Brault et al., 2002; Bennett et al., manuscript submitted).

Here we describe the evolution of DEN-4 across the Caribbean and adjacent regions since its arrival and establishment in 1981. Despite the importance of viral population dynamics in disease severity, few studies have attempted to characterize movement or gene flow among established viral populations. Dengue viral serotypes can clearly spread rapidly and colonize new areas: DEN-4 was introduced into the Americas in 1981 and within three years had spread to most countries in the region (Gubler, 1998). Once a serotype has been established, however, the relative impact of continued gene flow and/or local viral evolution on disease dynamics is unknown. Within a geographic area marked by frequent viral exchange, local evolution and subsequent spread of new variants could prove the most serious threat for severe epidemic disease transmission.

We used a phylogenetic approach to examine region-wide patterns of viral change in DEN-4 over the period of its establishment in the area. The region of the genome chosen for analysis included the entire envelope (E) gene and flanking regions. Because of its role in host cell binding, entry, and elicitation of immune response (Roehrig et al., 1998), the E gene is the one most commonly surveyed in dengue molecular epidemiologic studies (Lanciotti et al., 1997; Rico-Hesse et al., 1997; Leitmeyer et al., 1999; Wang et al., 2000). We investigated whether this serotype's evolution has been a regional phenomenon, marked by frequent genetic exchange among localities, or whether it is instead best characterized as a collection of discrete, independently evolving populations undergoing little genetic exchange. This type of information will be critical for identifying geographic areas of high viral genetic diversity and regional patterns of gene flow. Ultimately, by understanding the links between viral evolution, gene flow and selection, it should be possible to identity those viral lineages that are most likely to spread and seed new epidemic cycles.

Results

The 48 DEN-4 virus sequences analyzed represent fifteen countries from the Greater and Lesser Antilles and from South and Central America (Fig. 1). Sequences were more than 94% similar and fell within the viral lineage designated as "genotype II" by Lanciotti et al. (1997; Fig. 2). A total of 349 substitutions have occurred in the envelope gene (or 0.018 substitutions per site) since DEN-4's introduction into the Caribbean basin, most of which (80%) are silent mutations. The maximum divergence within this Caribbean subset is still relatively low compared to the degree of divergence that separates the two DEN-4 subtypes (0.018 vs 0.06 substitutions per site). Nonetheless, as few as one or two point mutations can have profound epidemiologic effects (Brault et al., 2002; Leitmeyer et al., 1999).

The initial introduction of DEN-4 into the Caribbean, represented by the 1981 isolate from Dominica, underwent little divergence until 1985: numerous Caribbean isolates from the early 1980s, up to 1984, along with 1982 isolates from northern South America (Suriname) fall together with the introduction isolate at the base of our phylogram (Fig. 2, expanded tree). We designated this group the "1981 introduction group" (delineated by green bar labeled "A" in Fig. 2, expanded phylogeny). From 1985 onward, this introduction group has evolved into a well-supported modern Caribbean basin lineage or clade (labeled "B" in Fig. 2), at the base of which fall divergent lineages from greater South and Central America, including isolates from Ecuador (1994), Honduras (1991), El Salvador (1993) and Mexico (1991, 1995).

The modern Caribbean basin clade as well as three of the Central American isolates (El Salvador [1993] and Mexico [1991, 1995]) are distinct from the introduction group by a single amino acid change from methionine to threonine in the E gene (amino acid position 163), a highly non-conservative substitution from a hydrophobic to a polar amino acid (Fig. 2, indicated by red bar on internal branch). The modern Caribbean basin clade ("B") itself includes only isolates collected after 1984, from the Caribbean, and nearby coastal South America (Venezuela, Suriname), with the exception of a single Costa Rican isolate from 1996. This well-supported lineage is distinctive based on two conservative amino acid substitutions (Fig. 2, black bars), from isoleucine to valine (amino acid position 351, E gene) and from lysine to arginine (amino acid position 52, NS1 gene), in addition to four silent base substitutions (not shown on figure). Within this Caribbean basin lineage, isolates are grouped more by year of isolation than by geographic origin.

The modern Caribbean basin clade ("B") includes four major and distinct lineages, none of which are defined by amino acid changes across the region examined. They are instead defined by anywhere from two to six silent base substitutions and are well supported by bootstrap analysis. The first lineage is comprised of Puerto Rican isolates from 1985 to 1987 (solid yellow bar at the base of lineage "B," Fig. 2). No isolates from other geographic areas fell within this group, but this observation may reflect the paucity of samples from other regions in the mid to late 1980s. An-

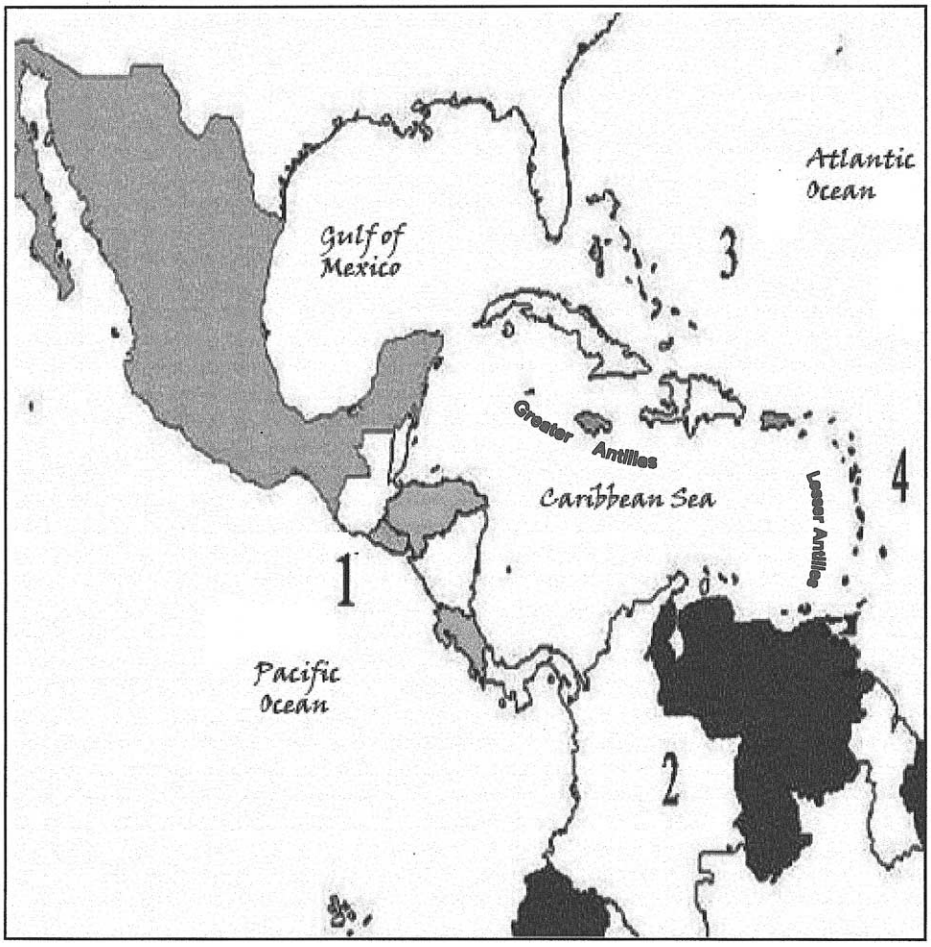

Fig. 1. Map of the Americas showing countries represented in this study. Regions within the Americas, i.e., Central America, South America, the Greater Antilles and the Lesser Antilles, are labeled 1–4 respectively. Countries represented in this study are indicated by shading: Costa Rica, El Salvador, Honduras and Mexico (Region 1); Ecuador, Suriname and Venezuela (Region 2); The Bahamas, Jamaica and Puerto Rico (Region 3); Barbados, Dominica, Martinique, Montserrat and Trinidad (Region 4).

other lineage within the modern Caribbean basin clade is comprised of Puerto Rican isolates from 1986 (upper yellow bar within "B," Fig. 2) and 1990 to 1994, as well as a 1993 Barbadian isolate and a Trinidad isolate from 1994 (Fig. 2, solid light-blue bars and lower solid dark-blue bar within "B"). The third and fourth lineages within the modern Caribbean basin clade are associated, forming a group of mid to late 1990s Caribbean isolates. Within this group, the third lineage is composed of isolates collected between 1993 and 1996 in the Lesser Antilles (Barbados and Martinique), Suriname and Costa Rica (Fig. 2, cross-hatched bar within "B" and solid dark-blue bar immediately below). Finally, the fourth lineage, best described as a late 1990s Caribbean basin group, is geographically diverse within the Caribbean basin and is composed of late 1990s isolates from Puerto Rico (1998), the Bahamas (1998), and the Lesser Antilles (1999; Fig. 2, solid red bar), along with a few earlier isolates, i.e., Montserrat isolates from 1994, and a Venezuelan isolate from 1995 (Fig. 2, solid dark-blue bar immediately below red area). This lineage is supported 98% of the time according to bootstrap analysis (Fig. 2).

There are three instances where isolates from the same time period at a given locality fall into different lineages. A 1986 isolate from Puerto Rico (PR 1986 115) falls in to the outside of the 1985–87 Puerto Rican lineage. Another Puerto Rican isolate, from 1985 (PR 1985 M33), lies entirely outside of lineage "B." There are also two 1993 isolates from Barbados that appear in different lineages: one in the early 1990s lineage (BDS 1993B) and the other in the mid-1990s lineage (BDS 1993A). There was no evidence for past recombination events in our dataset, at least across the gene region examined.

Discussion

This is the first large-scale regional analysis of the dynamics of DEN-4 populations. Our samples were drawn from multiple locations over the period that marked the rise of endemic DEN-4 in the Americas. Phylogenetic evidence is consistent with a single introduction and rapid colonization of the Americas by DEN-4 "genotype II" during the

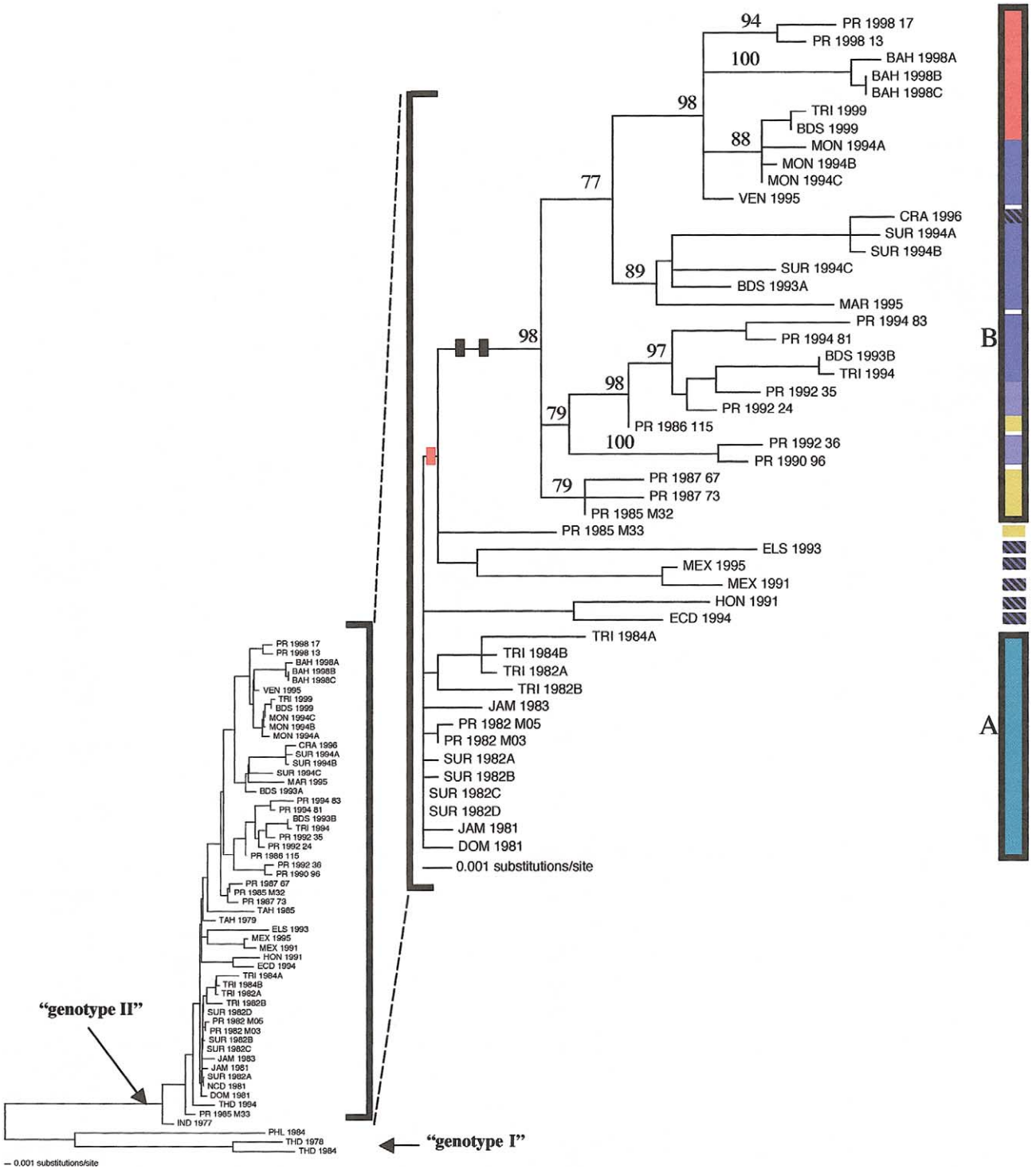

Fig. 2. Maximum likelihood tree of DEN-4 based on nucleotide sequences of E gene and adjacent regions of membrane and NS1 gene (1940 bp) of 48 isolates from around the Caribbean, Central and South America (1981–1999). The tree shown is expanded from a neighbor-joining tree representing the clade's position relative to Asian isolates of both the "genotype I" and "genotype II" DEN-4 subtypes (PHL (Philippines) 1984, THD (Thailand) 1978 and THD 1984 represent "genotype I"; IND (Indonesia) 1977, THD 1994 and NCD (New Caledonia) 1981 represent "genotype II" sample sequences acquired from GenBank). Beside nodes of interest in the expanded American tree, we provide bootstrap support values representing the percentage of times the given node was obtained, out of 1000 replicate neighbour-joining trees reconstructed under the ML substitution model (Posada and Crandall, 1998). The names of isolates refer to country of origin and year of isolation. In cases where there is more than one isolate from a given country and year, a unique isolate number (and/or letter) is also indicated. Countries represented are Bahamas (BAH), Barbados (BDS), Costa Rica (CRA), Dominica (DOM), Ecuador (ECD), El Salvador (ELS), Honduras (HON), Jamaica (JAM), Puerto Rico (PR), Montserrat (MON), Martinique (MAR), Mexico (MEX), Suriname (SUR), Trinidad and Tobago (TRI), and Venezuela (VEN). Scale shown represents number of nucleotide substitutions. Red and black bars on internal branches indicate instances of non-synonymous base changes. Letters "A" and "B" indicate the "1981 introduction group" and "modern Caribbean clade" respectively, delineated by right-marginal black boxes. Isolates from different time periods are delineated by different colors: 1981 to 1984 (green bar); 1985–1987 (yellow bar); 1990-1992 (light-blue bar); 1993–1996 (dark-blue bar); 1998–1999 (red bar). Isolates from the Caribbean basin have a solidly colored bar, whereas those from greater South and Central America have cross-hatched bars. Clades are separated by white spaces in the color bar.

early 1980s. Early 1980s samples collected in the Americas were similar to isolates collected in South East Asia and the Pacific during that same time period (Fig. 2, unexpanded tree). Since its introduction, several lineages have emerged independently from the base of the phylogenetic tree suggesting some degree of independent divergence of DEN-4 viral populations across the Americas (Fig. 2). Nonetheless, the most striking aspect of our phylogeny was the clear coupling of different regional populations through time, which suggests a high degree of continual gene flow between populations. This pattern was particularly true within the Caribbean basin, the region on which we focused our sampling efforts. Indeed, most isolates collected after 1987 within the Caribbean basin, including coastal South America, fell within a single clade defined by three amino-acid substitutions.

This major clade, the modern Caribbean basin clade, consisted of a number of distinct lineages that were generally widespread and often temporally clustered. For example, Lesser Antillean isolates from the early 1990s (BDS 1993B and TRI 1994) are clustered together but nested within a larger group of Greater Antillean (Puerto Rican) isolates from the same time period (Fig. 2; lower dark-blue bar). Isolates from the same areas, but in 1999, fell within another highly distinctive clade that again was grouped with Greater Antillean isolates from the same time period (Fig. 2; red bar). In view of the absence of any major Caribbean basin lineages defined solely on the basis of geographic location, ongoing genetic exchange is both common and important in DEN-4 evolution in the Caribbean. The simultaneous circulation of isolates from different lineages in the same locality (as seen in Puerto Rico, mid-1980s, and Barbados, 1993) is evidence that significant variation co-exists within a population, regardless of whether it derives from mutation or gene flow. That these lineages or variants have different long-term fates suggests that certain variants are more successful than others, some seeding later outbreaks while others became extinct. Analysis of a larger number of isolates would be necessary to confirm this.

The pattern of DEN-4 evolution in the Caribbean basin is perhaps not surprising given the enormous amount of social and commercial activity within this region. Both human and mosquito hosts can be transported freely throughout the region via ships and commercial airliners that move frequently between islands. Indeed, DEN-4 spread throughout the Americas within the two years following its initial diagnosis in Dominica in 1981. Rapid initial spread of DEN-4 reflects to some extent the absence of a previous infection history in the region. Dengue is an acute viral illness and exposure to a serotype appears to provide lifelong immunity to that serotype (Gubler, 1997). Prior to the 1980s, DEN-4 was absent from the region; all individuals were thus susceptible to the disease and viral spread would be unimpeded by host immunity. Subsequent to the initial epidemic, local host immunity could have acted as a barrier to continued regional gene flow. However, our data suggest that regional gene flow remains an important source of variation within populations of DEN-4 in the region, effectively unifying most of the Caribbean basin into a single evolutionary unit.

Although our sampling design does not allow us to identify specific transmission routes within the region, the Lesser Antilles island chain provides an obvious bridge between mainland South America and the Greater Antilles and Bahamas. Trinidad, the southernmost island in the Lesser Antilles, is just 11 km off the coast of Venezuela and there is frequent traffic between the two both by air and by sea. The divergent evolution of isolates from Central America and Ecuador may reflect weaker geographic and economic links. Analysis of DEN-4 isolates from different sublocalities within these mainland countries may help to determine whether the divergent evolution is correlated with geographic or economic factors. Despite the observed divergence, similar branch lengths for given years of collection indicate that evolution has been occurring at a relatively equal rate throughout the Americas. The presence of a single Central American isolate within the modern Caribbean basin clade may indicate an instance of genetic exchange between the Caribbean islands and the mainland, possibly the result of vertebrate host movement in the past. We have no recent collections to determine the result of that introduction. Analysis of a larger number of Central American isolates would be required to assess the degree of gene flow between these regions and its importance in dengue evolution.

The history of local evolution coupled with contemporaneous gene flow that we observe among DEN-4 populations has important epidemiologic ramifications. First, these data clearly suggest that eradication and control efforts should focus on a regional rather than local scale. Traditional control methods (e.g., education and vector control) applied at a single location are unlikely to succeed over the long term simply because the probability of recolonization is very high. Second, high levels of gene flow effectively expand susceptible host population size. This will allow for the maintenance of continuous DEN-4 transmission cycles throughout the region despite local or seasonal extinctions (Anderson and May, 1979). Third, and more insidiously, regional population dynamics marked by local evolution and gene flow would facilitate the spread of better-adapted DEN-4 variants. For dengue, which cycles between humans and mosquitoes, selection could act to increase viral replication rate and/or alter antigenic properties in either host. Both modes will likely have strong impact on viral transmission rates and/or disease severity.

Relative to other viruses DEN-4 has an intermediate rate of evolution of 4.0×10^{-4} to 1.0×10^{-3} substitutions per site per year (Bennett et al., manuscript submitted; Jenkins et al., 2002). There is a growing body of evidence that suggests that viral genotype is important to virulence and viral transmission potential and that a small number of amino acid changes can have profound epidemiologic sig-

nificance (Gubler et al., 1981; Gubler, 1997). For example, Leitmeyer et al., (1999) suggest that the presence of a single amino acid substitution in the envelope glycoprotein, the dominant virus antigen, responsible for virus attachment, fusion, and assembly (Gubler et al., 1981; Gubler, 1997; Chen et al., 1996; Roehrig et al., 1998), is one of the primary determinants of DHF in patients infected with DEN-2.

Despite a history marked by viral gene flow and lineage spread, there is no direct evidence in these data that adaptive evolution is promoting the success and spread of newly evolved lineages in the Caribbean. Of the three amino acid changes that defined the modern Caribbean basin lineage, two are highly conservative, and the third, although resulting in a hydrophobic to polar amino acid shift, does not fall into a known antigenic domain of the envelope glycoprotein. Furthermore, subsequent evolution within the region has been dominated by silent nucleotide substitutions. None of the major lineages within the basin was defined by amino acid changes within envelope and adjacent gene regions. Nonetheless, natural selection in other viral gene regions (undetectable in the present analysis) may be driving the pattern of evolutionary change that we observe. The most recent DEN-4 epidemic in Puerto Rico (1998) was dominated by viruses bearing mutations within a non-structural gene region under natural selection (Bennett et al., manuscript submitted). Isolates from this 1998 epidemic fall within a lineage that was detected as early as 1994 and is presently widespread in the Caribbean basin (Bahamas, Barbados, Montserrat, Suriname, Trinidad, and Venezuela; Fig. 2). The Bahamas also experienced a DEN-4 epidemic in 1998 (CAREC, 1999a; CAREC 1999b).

In the absence of animal models and *in vitro* correlates of the disease, detailed examination of the pattern of genetic change provides a powerful strategy toward understanding the relationship between dengue viral genotype and changing disease dynamics. The pattern of DEN-4 evolution in the Americas shows a complex history of rapid colonization, local evolution, and contemporary gene flow. Although there is some evidence for extended periods of independent evolution within the Americas, since DEN-4's arrival, newly evolved genotypes have regularly spread throughout the region. The Caribbean basin, in particular, is defined by a history of high gene flow irrespective of geopolitical boundaries. These types of population dynamics will have significant impacts on efforts to control and monitor the disease. DEN-4 is not the only serotype to demonstrate the ability to span regional boundaries: the relatively recent (1994) re-emergence of DEN-3 in the Americas provides further evidence of the ability of the dengue virus to spread rapidly (CDC, 1995; CAREC, 1999c; CAREC, 2000). A similar history of local evolution and migration will likely define the population history of other dengue virus serotypes. These data argue that efforts to understand disease dynamics need to focus not only on local scales, but also on the regional context in which these local populations exist.

Materials and methods

Viral archive, RNA extraction and RT-PCR

Virus isolates were obtained from the collections of the Caribbean Epidemiology Centre (CAREC), Trinidad, and the Centers for Disease Control and Prevention (CDC), Dengue Branch, San Juan, Puerto Rico. At both CAREC and CDC, viruses were isolated by culturing in C6/36 mosquito cells. Sera are regularly submitted to both CAREC and CDC for laboratory testing by medical institutions throughout the Caribbean and South and Central America, and hence usually represent symptomatic infections. Since these two institutions represent principal sources of dengue identification and serotyping in the Caribbean basin, viruses deposited there provide the best available representation of the symptomatic host population. We selected viruses from the collections to represent the multiple localities and years in which DEN-4 has been recorded since its most recent introduction to the Americas. Table 1 shows the year, country of origin, and number of isolates included in this study.

Our samples consisted of 48 American DEN-4 isolates (34 from Caribbean islands, 9 from South America, and 5 from Central America) from outbreaks and epidemics occurring between 1981 and 1999. We extracted RNA from samples using QIAamp Viral RNA Mini kits (Qiagen GmbH, Germany) according to the manufacturer's instructions. For each isolate, we amplified a 1940 bp region of the genome encoding the envelope (E) gene and adjacent 3′ region of the membrane gene (225 bp) and 5′ portion of the NS1 gene (353 bases) as two overlapping fragments (917 bp and 1054 bp) in separate single-tube reverse transcriptase–polymerase chain reactions (RT-PCRs). RT-PCR amplification primer sequences are provided in Table 2. Where necessary, RT-PCR products were purified prior to sequencing using Qiagen PCR purification kits (Qiagen GmbH, Germany), as outlined by the manufacturer.

RT-PCR fragment sequencing

Cycle sequencing reactions were performed using *Taq* DyeDeoxy Terminator Cycle Sequencing Kits (Applied Biosystems, Foster City, CA), according to the manufacturer's instructions, with sequencing primers described in Table 2. We cleaned products by standard precipitation before sequencing on an ABI PRISM 377 automated DNA sequencer (Applied Biosystems, Foster City, CA). We aligned sequences against a DEN-4 reference sequence (Zhao et al., 1986; GenBank number M14931) and combined fragments for a given isolate using Sequencher 3.1.1 software (Gene Codes, MI) to obtain a continuous nucleotide sequence for each sample. All sequences used in this study can be accessed in GenBank (Accession numbers, in order of appearance in Table 1: AY152364–66 (Bahamas); AY152368, AY152375–76 (Barbados); AY152104 (Costa Rica); AY152360 (Dominica); AY152292 (Ecuador); AY152300 (El Salvador); AY152379 (Honduras);

Table 1
Location, year of isolation, and number of DEN-4 isolates included in this study

Country of isolation	Date of isolation															
	1981	1982	1983	1984	1985	1986	1987	1990	1991	1992	1993	1994	1995	1996	1998	1999
Bahamas															3	
Barbados											2					1
Costa Rica														1[a]		
Dominica	1															
Ecuador												1[a]				
El Salvador											1[a]					
Honduras									1							
Jamaica	1		1													
Martinique													1[a]			
Mexico									1				1[a]			
Monserrat												3				
Puerto Rico		2[a]			2[a]	1[a]	2[a]	1[a]		3[a]		2[a]			2[a]	
Suriname		4										3				
Trinidad		2		2								1				1
Venezuela													1[a]			

[a] Samples from Bennett et al., manuscript submitted.

AY152384, AY152389 (Jamaica); AY152100 (Martinique); AY152378, AY152304 (Mexico); AY152369–71 (Monserrat); AY152056, AY152068, AY152112, AY152144, AY152148, AY152188, AY152208, AY152224, AY152236, AY152268, AY152336, AY152344, AY152855-57 (Puerto Rico); AY152372–74, AY152385–88 (Suriname); AY152367, AY152377, AY152380-83 (Trinidad); AY152092 (Venezuela)).

Phylogenetic analysis

Sequences were aligned together in MEGALIGN v 3.1.7 (Lasergene), using the Clustal algorithm (Higgins and Sharp, 1988; Higgins et al., 1996). A phylogenetic tree consisting of 56 isolates was generated using PAUP 4.0b8 (Swofford, 1998). We based tree calculations on models of sequence evolution whose parameters were estimated by maximum likelihood, using Modeltest 3.06 PPC (Posada and Crandall, 1998). The resultant phylogenetic tree was rooted with the 1981 isolate from Dominica, believed to represent the first introduction of the serotype into the Americas. We assessed the reliability of the tree topology by bootstrapping the data set 1000 times to generate neighbor-joining trees. Internal nodes of particular interest are labeled with percent times obtained, if greater than 50 (Fig. 2).

Recombination between viruses, both within and between serotypes, has been demonstrated in dengue (Holmes et al., 1999; Worobey et al., 1999) and can lead to erroneous and conflicting phylogenetic relationships. We examined our dataset for recombinants by comparing neighbor-joining trees based on 500 continuous bases at a time for incongruities. We then used LARD (Holmes et al., 1999), which generates maximum likelihood scores for each set of isolate

Table 2
List of primers used to amplify and sequence envelope gene and adjacent regions of membrane and NSI genes of DEN-4

Primer[a]	Sequence	Function	Fragment	Genome position
U486	5′-CACGTATAAATGCCCCCTACTGGTC	Amplification	EnvA	486
L1786	5′-GCTGTGTTTCTGCCATCTCTTTGTC	Amplification	EnvA	1786
U580	5′-ACCCAGAGCGGAGAACGGAGACGAG	Sequencing	EnvA	580
L803	5′-GGGGCGACCAGCATCATTAGGACAA	Sequencing	EnvA	803
U967	5′-GAACTGACTAAGACAACAGCCAAGG	Sequencing	EnvA	967
L1136	5′-AACAAGCCACAGCCATTGCCCCACC	Sequencing	EnvA	1136
U1363	5′-CCGGACTATGGAGAACTAACACTCG	Sequencing	EnvA	1363
L1658	5′-TGTCCTGCAAACATGTGATTTCCAT	Sequencing	EnvA	1658
U1568	5′-GCAATGGTTTTTGAATCTGCCTCTT	Amplification	EnvB	1568
L2679	5′-CCTTCACATCCCCAGCCACTACAGT	Amplification	EnvB	2679
U1602	5′-GCAGGAGCAGACACATCAGAGGTTC	Sequencing	EnvB	1602
U2114	5′-GAAAGGGAGTTCCATTGGCAAGATG	Sequencing	EnvB	2114
L1985	5′-CAAAGGGGTGGATGAGATGATACGC	Sequencing	EnvB	1985
L2519	5′-TCTCGCTGGGGACTCTGGTTGAAAT	Sequencing	EnvB	2519

[a] Designed by Bennett et al., manuscript submitted.

triplet (potential recombinant and proposed parentals) relationships, with and without a supposed crossover event. This identifies the most likely position for a potential breakpoint. The significance of an improvement in likelihood score by imposing a crossover event was assessed in two ways: statistically, by likelihood ratio test against a chi-square distribution for the appropriate degrees of freedom (Holmes et al., 1999), and biologically, by examining the overall phylogeny with and without a breakpoint inferred for the given isolate. In this way we determined whether the conflicting phylogenetic signal under suspicion was significant relative to the overall pattern of evolution, e.g., whether conflicting signals involved isolates from divergent lineages and thus could impact the overall phylogeny, or merely involved tip isolates within a given lineage.

In reporting our results we use several terms that appear in both evolutionary biology and virology literature, but are not always used in the same sense. For our purposes, we use the term "genotype" in its traditional genetic sense, to refer to the unique sequence of bases possessed by a single isolate (we also may refer to such an individual as a "variant"). Note that each isolate is actually a sample of multiple and potentially variable virus particles within the host (Wang et al., 2002) and presumably represents the most common genotype present. Rare isolates with confirmed sequence polymorphism(s), indicating the presence of multiple and equally common genotypes/virions within a single sample, were excluded from analyses. In contrast to our usage of the term "genotype," virologic nomenclature often uses the term to refer to a class of similar (but not identical) isolates. Instead, we use the terms "lineage" and "clade" to describe groups of similar, and presumably related individuals that share a common ancestor (monophyletic), when that relationship is supported at least 75% of the time by bootstrap analysis. We refer to groups assigned this or a greater level of support as "well-supported."

Finally, in our phylogenetic analysis we consider genetic changes at both the base pair and amino acid level, since the former does not always result in the latter. Based on our phylogenetic results, we discuss potential sources and patterns of genetic variation, using the terms *gene flow* and *genetic exchange* synonymously to refer to the movement of viral genes and thus the movement of dengue viruses between populations (i.e., populations of viruses in their populations of hosts). Because viruses are obligate parasites, the movement of virus particles between populations is also representative of the movement of hosts, either human host or mosquito vector (unintentionally assisted by human means).

Acknowledgments

The authors thank staff at the Caribbean Epidemiology Centre (CAREC) for providing Caribbean isolates for our study, especially A. George and V. Morris-Glasgow. We are also grateful to lab personnel at the Dept. of Biology, University of Puerto Rico and CDC Dengue Branch for their technical assistance and use of their sequence database, particularly D. Rodriguez, M. Chirivella and M. Beltran. We are also grateful to P. Umaharan, D. Gubler and G. Clark for critical reading of the manuscript. This work was funded by a UNDP/World Bank/WHO TDR training grant (to J.E.F.), grants from the University of the West Indies, St. Augustine Campus Research and Publication Fund and the Caribbean Health Research Council (to C.V.F.C.), the National Institute of Health (to W.O.M.) and the University of Puerto Rico's RCMI program.

References

Anderson, R.M., May, R.M., 1979. Population biology of infectious diseases: Part I. Nature 280, 361–367.

Brault, A.C., Powers, A.M., Holmes, E.C., Woelk, C.H., Weaver, S.C., 2002. Positively charged amino acid substitutions in the e2 envelope glycoprotein are associated with the emergence of Venezuelan equine encephalitis virus. J. Virol. 76, 1718–30.

Caribbean Epidemiology Centre (1999a). Communicable Diseases Feedback Report—Reporting Period: Epidemiologic Weeks 13–34, 1998. October 13, 1998. Available at: http://www.carec.paho.org/data/comm-dis/98wks13–34/index.html.

Caribbean Epidemiology Centre (1999b). Communicable Diseases Feedback Report—Reporting Period: Epidemiologic Weeks 35–43, 1998. December 8, 1998. Available at: http://www.carec.paho.org/data/comm-dis/98wks35–43/index.html.

Caribbean Epidemiology Centre (1999c). CAREC Epinote: An Update of Dengue Fever in the Caribbean-1998. 19 April, 1999. Available at: http://www.carec.org/data/dengue/1998update.html.

Caribbean Epidemiology Centre (2000). Dengue Virus Types Identified 1997–2000. 23 November, 2000. Available at: http://www.carec.org/data/dengue/denguevirustypes1997–2000.htm.

Centers for Disease Control, and Prevention, 1995. Dengue Type 3 infection-Nicaragua and Panama, October-November, 1994. Morbid Mortal Weekly Rep. 44, 21–24.

Chen, Y., Maguire, T., Marks, R.M., 1996. Demonstration of binding of dengue virus envelope protein to target cells. J. Virol. 70, 8765–8772.

Dietz, V., Gubler, D.J., Ortiz, S., Kuno, G., Casta-Velez, A., Sather, G.E., Gomez, I., Vergne, E., 1996. The 1986 dengue and dengue hemorrhagic fever epidemic in Puerto Rico: epidemiologic and clinical observations. PR Health Sci. J. 15, 201–210.

Gubler, D.G., Suharyono, W., Lubis, I., Eram, S., Gunarso, S., 1981. Epidemic dengue 3 in Central Java, associated with low viremia in man. Am. J. Trop. Med. Hyg. 30, 1094–1099.

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

Gubler, D.J. (1997). Dengue and dengue hemorrhagic fever: its history and resurgence as a global public health problem, in: Gubler, D.J., Kuno, G. (Eds.), Dengue and Dengue Hemorrhagic Fever, CAB International, pp. 1–22.

Halstead, S.B., 1988. Pathogenesis of dengue: challenges to molecular biology. Science 239, 476–481.

Higgins, D.G., Sharp, P.M., 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73, 237–244.

Higgins, D.G., Thompson, J.D., Gibson, T.J., 1996. Using CLUSTAL for multiple sequence alignments. Methods Enzymol. 266, 383–402.

Holmes, E.C., Worobey, M., Rambaut, A., 1999. Phylogenetic evidence for recombination in dengue virus. Mol. Biol. Evol. 16, 405–409.

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, 156–165.

Kliks, S.C., Nisalak, A., Brandt, W.E., Wahl, L., Burke, D.S., 1989. Antibody-dependent enhancement of dengue virus growth in human monocytes as a risk factor for dengue hemorrhagic fever. Am. J. Trop. Med. Hyg. 40, 444–451.

Kliks, S., 1990. Antibody-enhanced infection of monocytes as the pathogenetic mechanism for severe dengue illness. AIDS Res. Hum. Retroviruses 6, 993–998.

Knudsen, A.B., 1983. *Aedes aegypti* and Dengue in the Caribbean. Mosquito News 43, 269–275.

Lanciotti, R.S., Gubler, D.J., Trent, D.W., 1997. Molecular evolution and phylogeny of dengue-4 viruses. J. Gen. Virol. 78 (7), 2279–2284.

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

Morens, D.M., Marchette, N.J., Chu, M.C., Halstead, S.B., 1991. Growth of dengue type 2 virus isolates in human peripheral blood leukocytes correlates with severe and mild dengue disease. Am. J. Trop. Med. Hyg. 45, 644–651.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rico-Hesse, R., Harrison, L.M., Salas, R.A., Tovar, D., Nisalak, A., Ramos, C., Boshell, J., de Mesa, M.T., Nogueira, R.M., da Rosa, A.T., 1997. Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. Virology 230, 244–51, doi: 10.1006/viro.1997.8504.

Roehrig, J.T., Bolin, R.A., Kelly, R.G., 1998. Monoclonal antibody mapping of the envelope glycoprotein of the dengue 2 virus, Jamaica. Virology 5, 246(2), 317–28, doi: 10.1006/viro.1998.9200.

Swofford, D.L. (1998). PAUP* 4.0-Phylogenetic Analysis Using Parsimony (*and Other Methods). Sunderland, MA: Sinauer.

Thein, S., Aung, M.M., Shwe, T.N., Aye, M., Zaw, A., Aye, K., Aye, K.M., Aaskov, J., 1997. Risk factors in dengue shock syndrome. Am. J. Trop. Med. Hyg. 56, 566–572.

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

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

Wang, W.K., Lin, S.R., Lee, C.M., King, C.C., Chang, S.C., 2002. Dengue type 3 virus in plasma is a population of closely related genomes: quasispecies. J. Virol. 76 (9), 4662–4665.

World Health Organisation. (1996). Dengue and Dengue Haemorrhagic Fever. Fact Sheet, 117, May 1996.

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

Zhao, B., Mackow, E., Buckler-White, A., Markoff, L., Chanock, R.M., Lai, C.J., Makino, Y., 1986. Cloning full-length dengue type 4 viral DNA sequences: analysis of genes coding for structural proteins. Virology 155, 77–88.