PHYLOGENETIC AND BIOGEOGRAPHICAL RELATIONSHIP OF DENV-3 IN LATIN AMERICA

CHRISTIAN JULIÁN VILLABONA ARENAS

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE BIOLOGIA
2008

PHYLOGENETIC AND BIOGEOGRAPHICAL RELATIONSHIP OF DENV-3 IN LATIN AMERICA

CHRISTIAN JULIÁN VILLABONA ARENAS

Trabajo de investigación para optar el titulo de Biólogo

DIRECTOR:

Daniel Rafael Miranda Esquivel

PhD Ciencias Naturales

CODIRECTOR:

Raquel Elvira Ocazionez Jiménez

PhD Imunologia Básica e Aplicada

UNIVERSIDAD INDUSTRIAL DE SANTANDER

FACULTAD DE CIENCIAS

ESCUELA DE BIOLOGIA

2008

ACKNOWLEDGEMENTS

We thank to B. Parra (from Universidad del Valle, Colombia) and V. Vordam (from CDC, Atlanta) for providing isolates for this study, E. A. Gould and E. C. Holmes for their contributions to manuscript, and F. J. Diaz (from Universidad de Antioquia, Colombia) for his technical recommendations. This study was supported by Universidad Industrial de Santander, Colombia (grant # 2007-5636) and was partially supported by Secretaria de Salud of Santander State, Colombia (grant # 2007-068-000-0070)

CONTENT

LIST OF FIGURES

Pag.

LIST OF TABLES

Título: RELACIÓN FILOGENÉTICA Y BIOGEOGRÁFICA DEL DENV-3 EN LATINO AMÉRICA *

Autor: Christian Julián Villabona Arenas**

Palabras clave: Virus del dengue*, Latinoamérica, filogenia, Biogeografía.*

El virus del Dengue tipo 3 (DENV-3) fue reportado en las Américas en 1994 y desde entonces ha expandido su rango geográfico hacia la mayor parte de los países latinoamericanos. En Colombia, la introducción del DENV-3 fue por primera vez documentada en el 2001 en Santander, un departamento cercano a la frontera con Venezuela. Para determinar la relación filogenética y biogeográfica del DENV-3 en Latinoamérica, se secuenció el gen de la envoltura viral de 21 aislados colombianos muestreado en un periodo de siete años, y junto con los aislados disponibles en las bases de datos se llevaron acaba análisis filogenéticos usando parsimonia y Máxima Verosimilitud (ML). Todas las optimizaciones con ML y los diferentes costos con parsimonia recuperaron en mismo cladograma general pero con distintos niveles de resolución. En todos los análisis los aislados latinoamericanos agruparon con la cepa de Sri Lanka, perteneciendo al genotipo III. Todos los análisis mostraron que el DENV-3 circulando en Latinoamérica desde 1994 puede ser agrupado en 5 grupos donde las asociaciones filogenéticas están dadas con la proximidad geográfica. Un análisis biogeográfico usando un análisis de dispersión-vicarianza mostro tres distintos escenarios que representan las áreas ancestrales más plausible para el genotipo. El área común a los tres escenarios fue Centroamérica. Finalmente, teniendo en cuenta la distribución temporal y las asociaciones filogenéticas, se hipotetiza que existen al menos dos rutas de entrada del genotipo hacia Colombia, a través de Venezuela y Ecuador-Perú.

*Trabajo de Investigación

**Facultad de Ciencias, Escuela de Biología, Daniel Rafael Miranda Esquivel – Raquel Elvira Ocaziones Jiménez

Title: PHYLOGENETIC AND BIOGEOGRAPHICAL RELATIONSHIP OF DENV-3 IN LATIN AMERICA*

Author: Christian Julián Villabona Arenas**

Key words: Dengue virus*, Latin America, Phylogeny, Biogeography.*

Dengue virus type 3 (DENV-3) was reported in the Americas in Panama in 1994 and has spread subsequently to most Latin American countries. In Colombia, the introduction of DENV-3 was first reported in 2001 in Santander, a state near to Venezuela. To determine the phylogenetic and biogeographical relationship of DENV-3 in Latin America, we sequenced the complete envelope gene sequence of 21 Colombian viruses sampled over a seven year period, and together with available Latin American sequences deposited in Genbank we performed detailed phylogenetic analyses using both maximum likelihood (ML) and parsimony methods. All the optimizations with ML and the different cost schemes with parsimony recovered the same general cladogram with different levels of resolution. In all the analyses the American isolates grouped with a Sri-Lankan strain, belonging to subtype III. All analyses also showed that DENV-3 circulating in Latin America since 1994 could be grouped in five main clades where the phylogenetic associations of the strains correlated with their geographic proximity. A biogeographical analysis using the dispersion-vicariance optimal distribution showed three different areas of the Americas represented the most likely ancestral sources for the DENV-3. One of these areas was Central America. Finally, taking into account the temporal distribution and the phylogenetic associations, we hypothesize that DENV-3 was introduced into Colombia through Venezuela and through Ecuador or Peru.

*Trabajo de Investigación

**Facultad de Ciencias, Escuela de Biología, Daniel Rafael Miranda Esquivel – Raquel Elvira Ocaziones Jiménez

1. INTRODUCTION

Dengue viruses belong to the genus *Flavivirus*, family *Flaviviridae* and comprise four antigenically distinct virus serotypes designated DENV-1, 2, 3 and 4. They are enveloped viruses with a single stranded ~11 kb positive-sense RNA genome that encodes three structural proteins and seven non-structural proteins. The four serotypes are responsible for dengue hemorrhage fever/dengue shock syndrome (DHF/DSS). These are one of the most important infectious diseases affecting tropical urban areas (Gubler, 1998, 2002, 2006). Each year there is an estimated number of 50-100 million dengue infections globally, with 500000 cases of hospitalized DHF and 20000 – 25000 deaths (WHO, 1997, 2000). Phylogenetic studies of the Dengue viruses has shown variation within each serotype organized as discrete clusters on trees, named subtypes or genotypes (Rico-Hesse, 1990; Lanciotti *et al*., 1994, 1997; Twiddy *et al*., 2003). The emergence of epidemics has been characterized by the co-occurrence of multiple DENV serotypes in the same locality and a spatial pattern within each of the four serotypes where their genotypes often have a disjunct geographical distribution (Holmes, 2004).

Dengue virus circulating in the Americas was reported for the first time in 1953 and was classified as DENV-2 (Anderson *et al*., 1956). DENV-3 was introduced in 1963 and the last isolations took place in the 1970s in Puerto Rico and Colombia (Pinheiro & Corber, 1997). After 17 years of its disappearance from the region, DENV-3 re-appeared in 1994 in Nicaragua and Panama and later in Mexico, other Central American and the Caribbean countries (CDC, 1995; Guzmán *et al*., 1997; PAHO, 1995-2000; Usuku *et al*., 2001; Peyrefitte *et al*., 2003). The serotype was isolated in South America in 2000 in Brazil (Nogueira *et al*., 2001), Ecuador, Peru (Kochel *et al*., 2008), and Venezuela (Uzcategui *et al*., 2003); and subsequently in

nearby countries (Aquino *et al*., 2006; Aquino *et al*., 2008; Barrero & Mistchenko, 2008; Kochel *et al*., 2008).

DENV-3 has been classified into four (Lanciotti *et al*., 1994) or five (Wittke *et al*., 2002) genotypes. Previous phylogenetic studies showed that the DENV-3 from Latin America circulating during the 1960s and 1970s was a genotype IV virus, while viruses isolated since 1994 have been grouped into genotype III (Wittke *et al*. 2002; Uzcategui *et al*., 2003; Messer, *et al*., 2003). Although DENV-3 genotype III have been the genotype circulating in the Americas since its first detection, the DENV-3 genotype I was isolated in 2008 in Brazil (Barcelos Figueiredo *et al*., 2008).

In Colombia, the introduction of DENV-3 genotype III was reported for the first time in 2001 in Santander, a state near to the Venezuelan border (Ocazionez *et al*., 2006). The serotype was detected in subsequent years in the Northern, Southern, and Southeastern regions of the country, becoming the predominant serotype and causing DF outbreaks in numerous localities of the country (Mendez & Bernal, 2002; Ospina, 2004). In Santander, the introduction and subsequent predominance of the serotype between 2001 and 2004 coincided with an extensive epidemic when at least 7106 dengue laboratory-confirmed cases were registered (Ocazionez *et al*., 2006). In the current study we use phylogenetic methods to determine the genotype of Colombian viruses isolated between 2001 and 2007 and to reconstruct the relationships of DENV-3 in the Americas. In addition, we used biogeographical methods to reconstruct the history of the distribution of DENV-3 in America since its introduction in 1994.

2. METHODS

2.1 VIRUS DATA

Colombian viruses from Santander and Norte de Santander were isolated in the Centro de Investigaciones en Enfermedades Tropicales (CINTROP) of the Universidad Industrial de Santander, Bucaramanga, Colombia. Colombian viruses from Valle del Cauca were kindly provided by B. Parra (Universidad del Valle, Cali, Colombia). Isolates from Honduras and Sri Lanka were donated by the Center for Disease Control and Prevention (CDC, Fort Collins, Colorado). Virus isolation was undertaken in both institutions, CINTROP and Universidad del Valle, by culturing in C6/36 cells. Colombian samples consisted of 21 DENV-3 isolates (19 from Santander and Norte de Santander in the Northeastern region, and 2 from Cali in the Southwestern region) from outbreaks and epidemics occurring between 2001 and 2007.

2.2 RNA EXTRACTION, RT-PCR AND SEQUENCING

Viral RNA was extracted from a 140 uL aliquot of supernatant from virus-infected cultures using the QIAamp Viral RNA mini kit (Qiagen) according to the manufacturer's instructions. The viral cDNA was synthesized using 10 uL of RNA containing 50 ng of random hexamer oligonucleotides and each dNTP to a final

concentration of 0.5 mM. The mixture was incubated at 70 °C for 5 min and chilled on ice for 5 min. 1 uL of 0.1 M dithiothreitol (DTT), 4 uL of 5X first-strand buffer and 1 uL of Superscript III Reverse Transcriptase (200 units/uL) (Invitrogen) were added. The final volume was adjusted with RNase/DNase free water to 20 uL. Reverse transcriptions were conducted at 25 °C for 10 min, at 55 °C for 1 h and finally at 75 °C for 15 min. The primers previously published by Aquino *et al*. (2006) were used for amplification and/or sequencing. The region of the genome encoding the envelope gene was amplified using a reaction mixture containing 2 uL of the cDNA, 25 uL of the GoTaq Green Master Mix (Promega) and each primer to a final concentration of 0,5 uM. The final volume was adjusted with RNase/DNase free water to 50 uL. The amplification was carried out as follows: 95 ºC for 2 min, 39 cycles at 95 °C for 10 s, 55 °C for 30 s, and 72 °C for 150 s, followed by a final incubation at 72 °C for 5 min. The products of PCR were purified using the Wizard PCR Preps DNA Purification System (Promega) and sequenced in a commercial manufacturer (Macrogen). Sequence assembly was performed with the Lasergene package (version 7.0; DNASTAR).

2.3 PHYLOGENETIC ANALYSIS

Parsimony optimizations were conducted under different cost schemes (Indel:substitution cost relations of 1:1 and 2:1, and transversion:transition cost relations of 1:1, 2:1 and 4:1) using the dynamic homology approach (Wheeler 1996; Wheeler 2001; Wheeler *et al*., 2006) as implemented in POY version 4.0 build 2885 software (Varón *et al*., 2008). The analysis were done using tree bisection-reconnection (TBR) and subtree pruning-regrafting (SPR) swapping on wagner trees generated from 100 random taxon addition sequences, with one tree held in memory per replication. Those cost schemes were used to evaluate the sensitivity of the results to the variation of the cost values in terms of taxa grouping.

The recovered implicit alignments were used in another parsimony optimization with TNT version 1.1 software (Goloboff *et al*., 2008) with the previous cost schemes and similar procedures to evaluate whether or not both parsimony optimization approaches found the same results. The obtained topologies were compared taking into account the number of clades in common among them with the treeC software (Arias & Miranda-Esquivel, 2006). To assess the nodal support in the obtained parsimony topologies the relative bremer support was used because it exhibits the advantage to take into account the number of extra steps needed to collapse a branch, providing an approximate measure of the amount of favorable/contradictory evidence for each node. Relative bremer support values (Goloboff & Farris, 2001) were calculated with TNT version 1.1 software (Goloboff *et al*., 2008) using suboptimal trees 25 steps longer and holding 25000 trees in memory. To select the outgroup, different data sets that included strains of DENV-3 from different genotypes and/or strains for other serotypes and/or strains of related flaviviruses were used along with an indel:transversion:transition cost matrix of 1:1:1 in a parsimony optimization with POY version 4.0 build 2885 software (Varón *et al*., 2008). The criteria for selecting among the different outgroups were the maintenance of the length of the recovered alignments and the number of clades in common between the obtained tree with respect to the in-group optimized alone, along with the highest localities variety of the included taxa. The initial sequences were aligned with Muscle version 3.7 software (Edgar, 2004) using the default parameters. The costs scheme used to evaluate the sensitivity of the results in the parsimony analysis could not be implemented in Muscle software because the program only generated aligned blocks when a gap opening cost value as small as -1600 for any given transversion:transition cost relation was used. The best-fit model of nucleotide substitution was determined using a hierarchical likelihood ratio test (Posada and Crandall, 2001) as implemented in the Modeltest version 3.7 software (Posada and Crandall, 1998) using the ingroup alone and the ingroup plus the selected outgroup. The obtained alignment from Muscle 3.7 software (Edgar, 2004) was used in a Maximum Likelihood (ML) optimization analyses in phyML version 3.0 software (Guindon and Gascuel, 2003)

along with the selected model. The starting tree was found using the neighbor-joining (NJ) method and a bootstrap analysis using 10000 replicates was conducted to place confidence values on grouping within the trees.

2.4 BIOGEOGRAPHICAL ANALYSIS

A phylogenetic analysis shows the pattern of historical relationships among the taxa under study, Hence, to deduce events such as dispersion or *in situ* evolution from the topology itself, even if the patterns of spatial and temporal sampling are available, is just matter of speculation unless the appropriate methods are applied. To reconstruct the ancestral distributions for the clades found in ML tree a dispersal-vicariance optimization (Ronquist, 1997) was undertaken. The distributions of the groups and their ancestral area were described in terms of a set of unit areas. For the analysis, the taxa in the ML tree were replaced by their distribution. Sister taxa were collapsed if they had the same distribution. The distributional criterion used was the geographical regions where the terminal strains were found: Caribbean (Ca), Central America (Ce), northern South America (No), northeastern South America (Ne), Pacific (Pa), and Southeastern South America (Se). Because there is a tendency for the ancestral distribution reconstructions to include most or all of the areas occupied by the terminals, constraints on the maximum number of unit areas allowed were imposed to find the best candidate(s) in terms of area. Hence, the number of unit areas allowed in the ancestral distributions was set to the minimal possible value in DIVA version 1.2 software (Ronquist, 1996), a value of two.

3. RESULTS AND DISCUSSION

The in-group included twenty-three new DENV-3 envelope sequences: twenty-one from three states in Colombia (Norte de Santander and Santander from the northeastern region and Valle del Cauca from the southwestern region) over a period of seven years (2001–2007) and one from Honduras-1995 (GenBank accession numbers FJ189449-FJ189469, FJ204475-FJ204476) (Table 1). These sequences were used with the available South American isolates deposited in GenBank (see Supplementary Table A). The selected outgroup was composed of the strains Philippines-1956, Puerto Rico-1977, Sri-Lanka-1990 (obtained in the present study), Fiji-1992, and Thailand-1997 which represented all the different DENV-3 genotypes. The final data set comprised 119 sequences of 1479 nucleotides each.

All the different cost schemes used in both optimizations with parsimony recovered the same general cladogram with a different resolution level where the most representative groups in terms of country grouping were conserved. The implied alignments for each cost scheme suggest a different level and number of gaps. The obtained alignment with Muscle software (Edgar, 2004) was of 1479 nt long which had the same length as the DENV-3 envelope gene suggesting no gaps in the aligned blocks.

The general cladogram representing the relationships among DENV-3 is shown in Figure 1. This topology showed that the DENV-3 that have been circulating in Latin America since 1994 could be grouped in five main clades: Clade I that grouped viruses from Center America and Mexico, Clade II that grouped the remaining viruses from Mexico, clade III with some Colombian strains (strains from Norte de

Santander and Santander) and Venezuela, clade IV with viruses from Cuba, Ecuador, Peru and Colombia (strains from Valle del Cauca and one strain from Santander as clade II) and clade V that grouped viruses from Argentina, Cuba, Bolivia, Brazil, Martinique and Paraguay. There were no groupings by country, and in some cases strains from the same locality appeared in separate clades, eg the Colombian and Cuban viruses.

The Tamura and Nei + Γ distribution model (TrN + Γ model) (Tamura & Nei, 1993) was the best fit to the data with an α (shape parameter) value of 0.26. Under this model each possible substitution had its own probability and it used a continuous distribution to estimate variables rates at sites indicating that all sites in the sequences did not evolve according to the same substitution rate matrix. The α value indicated that most sites were invariable but a few had very high rates of substitution. There were no differences in the model selected when the ingroup was used alone and only the base frequencies and Γ changed slightly (α value of 0.29). The groups recovered in the parsimony analyses were found in the ML reconstruction, except for clade V that only was recovered under the latter analysis (Figure 1) and for clade I which appeared as the sister group of the grouping of clade II plus clade III with a bootstrap value of 78% (Figure 2). Clades I and the node that groups clade II and III had a bootstrap value of 45% and the remaining main clades had values above 55%. The p-value interpreted as the probability observed whether or not the DENV-3 taxa were evolving according to a molecular clock was 0.00000 and therefore the molecular clock hypothesis was rejected. Given this, we could not relate the divergence time among taxa to the number of molecular differences measured among them. Branch lengths which represent the number of substitutions per site showed that the differences among Latin American strains and the remainder are huge. Remarkably, the Argentinean isolates had much longer branch lengths in contrast with the other Latin American viruses. The assignment of DENV strains to a genotype depends on their association with other strains in a formal phylogenetic analysis, and because all the Latin American sequences isolated since 1994 appeared as a monophyletic group with the virus

strains Sri Lanka-1990, the data presented here stated that they all belong to genotype III (Lanciotti *et al*., 1994; Uzcategui, *et al*., 2003).

"*In situ*" evolution of DENV-3 strains following its introduction into Latin America has been reported before (Uzcategui *et al*., 2003; Rodriguez-Roche, *et al*., 2005). This conclusion was derived from a phylogenetic tree given that the strains appeared as a unique clade. Nonetheless, the fact that the isolates from the geographical regions emerge from a common node on the tree, only suggests a common origin and a subsequent diversification. Hence, those clades rather than proving "*in situ*" evolution, denote that there were no direct introductions of DENV-3 virus from Asia to each and every Latin American country, but there was a descendant of the virus that was gradually dispersing and diverging through the continent once it was first introduced into Panama.

The dispersion-vicariance optimal distribution (Ronquist, 1997) showed as the most probable ancestral areas for all the DENV-3 in the Americas, three different possibilities: Caribbean and Central America, Central America and Pacific, and Central America and Southeastern South America. Consequently, and according to the isolation dates, Central America represents the introduction point of DENV-3 in America. For the grouping of clades I, II and III the most likely ancestral area was Central America. Taking into account the temporal distribution and the phylogenetic grouping, the data hypothesized the introduction into Venezuela in 2000 (Uzcategui *et al*., 2003) from Central America-Mexico, spreading then westwards into Colombia during 2001 (Ocazionez *et al*., 2006). Our analysis identified the Caribbean-Southeastern South America as one of the likely ancestral areas for the node that groups clade IV and V indicating a transmission route from this region through South America as mentioned before (Aquino *et al*., 2006). Given that Cuba-2000 and Cuba-2001/2002 appeared in separate clades previous papers concluded that DENV-3 was introduced at least twice into Cuba from the Latin American region (Rodriguez-Roche *et al*., 2005). For clade IV, which included some Colombian and Cuba-2001/2002 strains the most likely ancestral area was

the Pacific suggesting the introduction of viruses into this country from Ecuador-Peru where the serotype was detected prior to the former ones. The optimization showed for the strain CO07a, which was isolated in the northeastern region of Colombia but appeared within the Pacific Clade, that it was probably imported to Santander from Valle del Cauca and did not represent the reintroduction of another strain from another country. For clade V the most likely ancestral areas were Caribbean and Southern South America. The Brazilian isolates predated the isolates from Argentina and Paraguay and in all cases they were associated in the same clades in the tree. Therefore, the most likely source for Paraguayan and Argentinean isolates was most likely to be Brazil. Remarkably, in this clade the optimization suggested in more than one clade that the Caribbean region could be the most likely source of the Brazilian isolates which indicated that more than one introduction in different points of time has occurred.

Our study showed that DENV-3 viruses in Latin America formed distinct clades within the genotype III and that the geographic proximity was related with the phylogenetic associations of the strains. Furthermore, this pattern suggests the movement of dengue viruses over long distances where the national border does not represent a geographic barrier. The intensive trade between nearby countries may explain this, indicating that to understand the virus dynamics we need to focus on the regional context in which the viral populations exist rather than in a country center perspective given that monophyly in terms of countries was not found.

DENV-3 outbreaks have been reported in Central America since 1994 and then the epidemic spread to Mexico and Caribbean Islands and finally southwards into South America (Messer *et al*., 2003). Our data suggested not only introductions into South America from both the Caribbean Islands and Central America but movement of viruses to the Caribbean from the countries in the Pacific region as well. Eight years after its introduction in Panama, DENV-3 had spread into Colombia. Nonetheless, in accordance with the phylogenetic tree it was not introduced from this country but through Venezuela and Ecuador or Peru. The

finding of two different strains of DENV-3 during 2007 in Santander-Colombia could be an indicator of the active circulation of Dengue virus populations through the country.

4. FIGURES AND TABLES

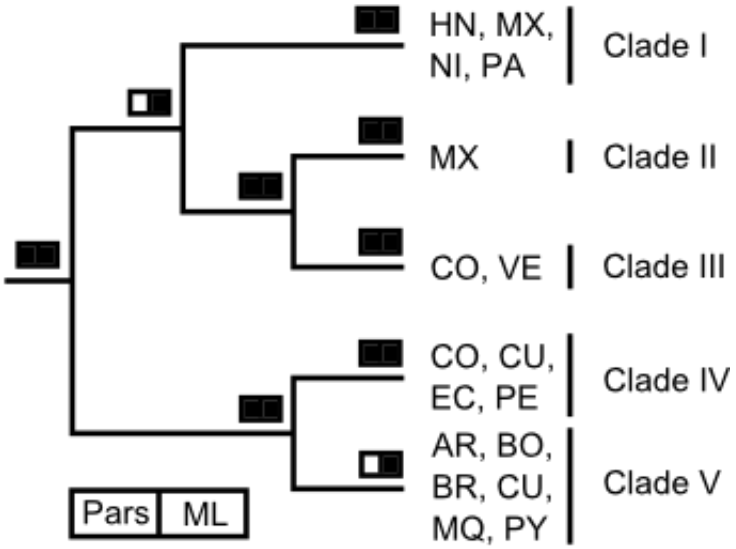

Fig. 1. General cladogram showing the groups found using the Maximum Likelihood and Parsimony optimizations.

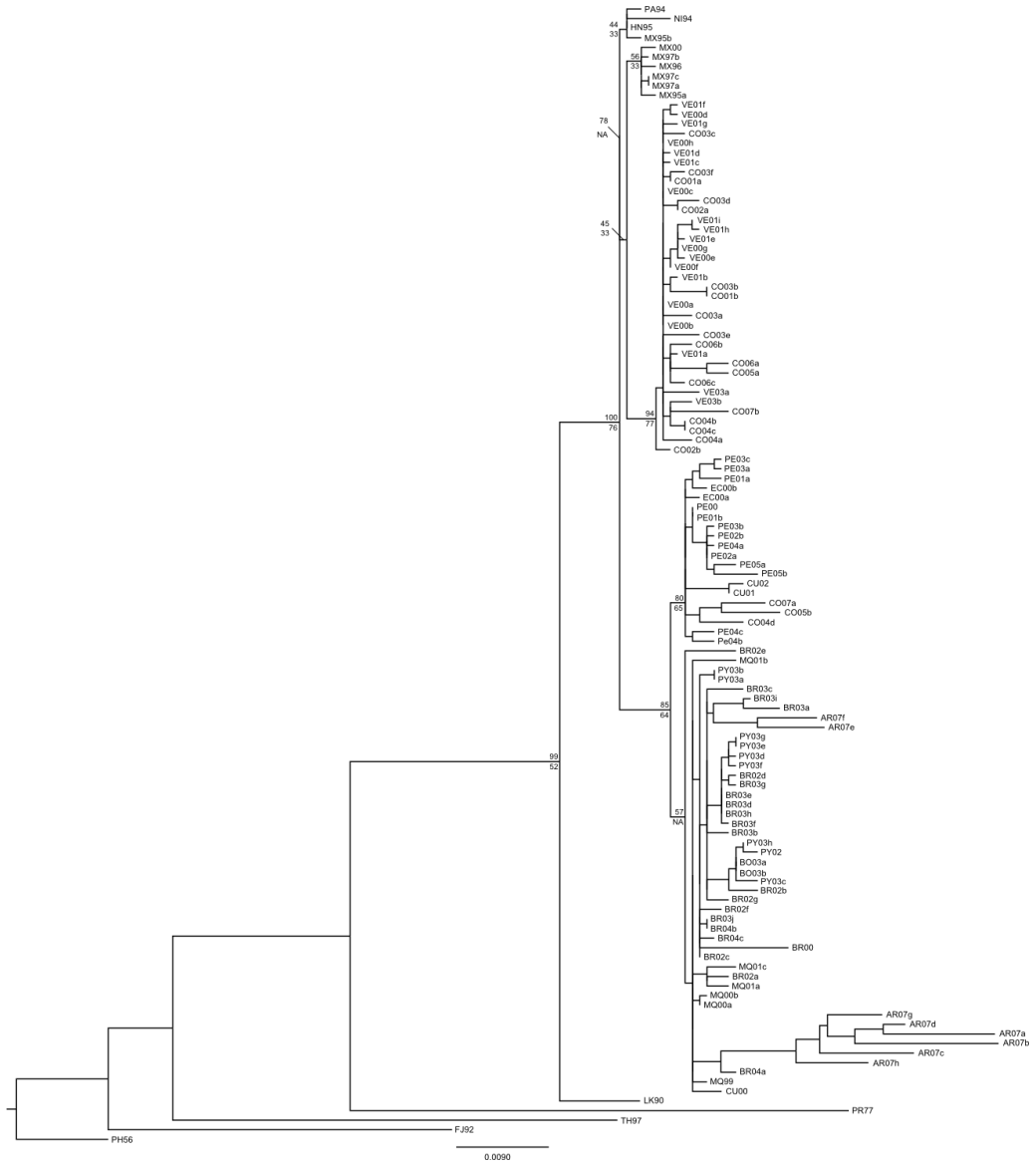

Fig. 2. Maximum likelihood tree showing the relationships among the envelope genes of 119 strains of DENV-3. Upper numbers correspond to percentage bootstrap values based on 10000 pseudo-replicates and lower numbers correspond to relative Bremer support values for the parsimony reconstruction with the cost scheme 8:8:1. Conventions as in Table 1.

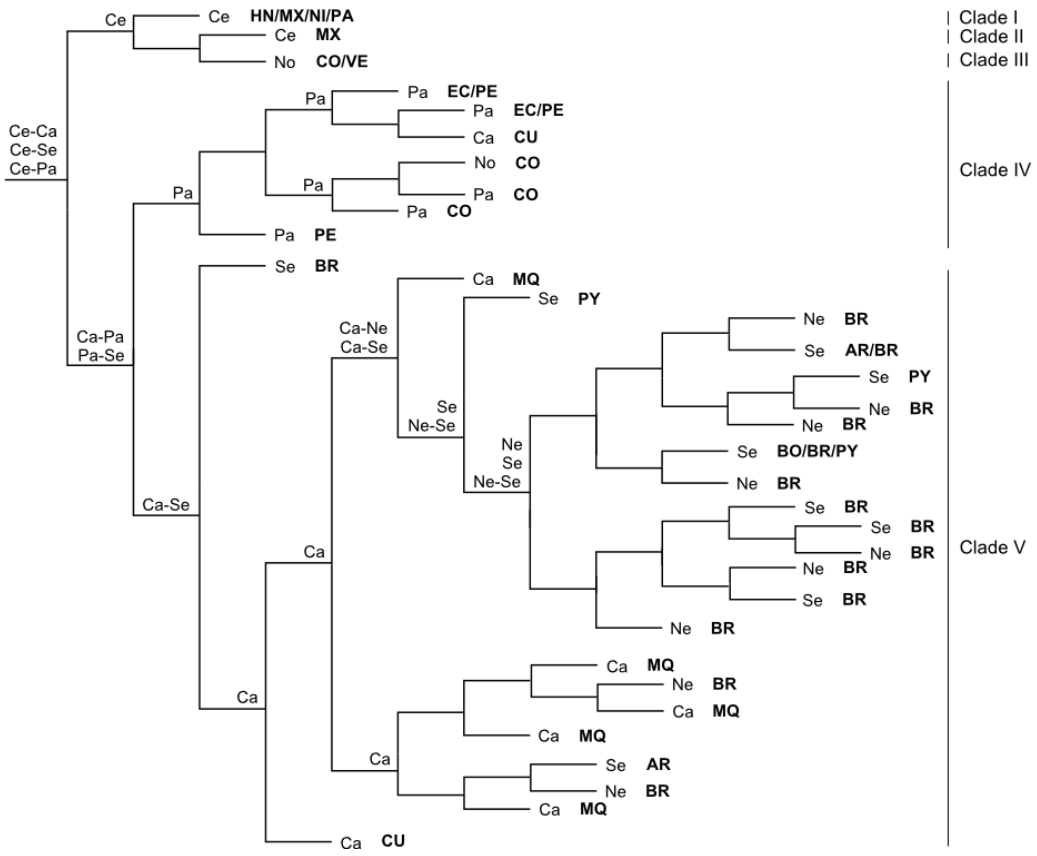

Fig. 3. Reconstruction of the ancestral geographic ranges in the DENV-3 based on dispersal–vicariance analysis (Ronquist, 1996). Labels for areas codes as in methods and for countries (in bold) ISO codes.

Table 1. Dengue virus type 3 sequences sequenced in this study

Geographical origin	Year	Strain	Identifier	GenBank accession number
Colombia, Santander	2001	COD3_01072	CO01a	FJ189450
Colombia, Santander	2001	COD3_01330	CO01b	FJ189451
Colombia, Santander	2002	COD3_02200	CO02a	FJ204475
Colombia, Santander	2002	COD3_02211	CO02b	FJ189452
Colombia, Santander	2003	COD3_LV016	CO03a	FJ189454
Colombia, Santander	2003	COD3_LV038	CO03b	FJ189455
Colombia, Santander	2003	COD3_LV057	CO03c	FJ189456
Colombia, Santander	2003	COD3_LV058	CO03d	FJ189457
Colombia, Santander	2003	COD3_G-048	CO03e	FJ189453
Colombia, Santander	2003	COD3_LV073	CO03f	FJ189458
Colombia, Santander	2004	COD3_LV357	CO04a	FJ189459
Colombia, Santander	2004	COD3_LV433	CO04b	FJ189461
Colombia, Santander	2004	COD3_LV428	CO04c	FJ189460
Colombia, Valle del Cauca	2004	COD3_C88	CO04d	FJ189467
Colombia, Norte de Santander	2005	COD3_OC092	CO05a	FJ189462
Colombia, Valle del Cauca	2005	COD3_C108	CO05b	FJ189468
Colombia, Norte de Santander	2006	COD3_OC283	CO06a	FJ204476
Colombia, Norte de Santander	2006	COD3_OC384	CO06b	FJ189463
Colombia, Norte de	2006	COD3_OC400	CO06c	FJ189464

Santander				
Colombia, Santander	2007	COD3_07181	CO07a	FJ189465
Colombia, Santander	2007	COD3_07302	CO07b	FJ189466
Honduras	1995	HN179	HN95	FJ189469
Sri Lanka	1990	SK698	LK90	FJ189449

Table 2. Dengue virus type 3 sequences from GENBANK used in this study

Geographical origin	Year	Strain	Identifier	GenBank accession number
Argentina, Buenos Aires	2007	ARG6475-07	AR07a	EU052792
Argentina, Buenos Aires	2007	ARG6541-07	AR07b	EU052793
Argentina, Buenos Aires	2007	ARG6645-07	AR07c	EU052794
Argentina, Buenos Aires	2007	ARG6694-07	AR07d	EU052795
Argentina, Buenos Aires	2007	ARG6733-07	AR07e	EU052796
Argentina, Buenos Aires	2007	ARG6768-07	AR07f	EU052797
Argentina, Buenos Aires	2007	ARG11586-07	AR07g	EU052798
Argentina, Buenos Aires	2007	ARG11595-07	AR07h	EU052799
Bolivia, Santa Cruz,	2003	FS B413/2003	BO03a	DQ177886
Bolivia, Santa Cruz,	2003	FS B439/2003	BO03b	DQ177887
Brazil, Rio de Janeiro	2000	68784	BR00	AY038605
Brazil, Boa Vista	2002	D3BR/BV4/02	BR02a	DQ118865
Brazil,	2002	D3BR/CU6/0	BR02b	DQ118866

Cuiaba		2		
Brazil, Manaus	2002	D3BR/MA1/02	BR02c	DQ118869
Brazil, Porto Velho	2002	D3BR/PV5/02	BR02d	DQ118875
Brazil, Rio de Janeiro	2002	BR74886/02	BR02e	AY679147
Brazil, Sao Geraldo do Araguaia	2002	D3BR/SG2/02	BR02f	DQ118880
Brazil, Sao Luis	2002	D3BR/SL3/02	BR02g	DQ118881
Brazil, Goiania	2003	D3BR/GO5/03	BR03a	DQ118867
Brazil, Iguape	2003	D3BR/IG10/03	BR03b	DQ118868
Brazil, Marituba	2003	D3BR/MR9/03	BR03c	DQ118870
Brazil, Porto Velho	2003	D3BR/PV1/03	BR03d	DQ118871
Brazil, Porto Velho	2003	D3BR/PV2/03	BR03e	DQ118872
Brazil, Porto Velho	2003	D3BR/PV3/03	BR03f	DQ118873
Brazil, Porto Velho	2003	D3BR/PV4/03	BR03g	DQ118874
Brazil, Porto Velho	2003	D3BR/PV6/03	BR03h	DQ118876
Brazil, Ribeirao	2003	D3BR/RP1/03	BR03i	DQ118877

Preto				
Brazil, Ribeirao Preto	2003	D3BR/RP2/03	BR03j	DQ118879
Brazil, Braganca	2004	D3BR/BR8/04	BR04a	DQ118864
Brazil, Paranapebas	2004	D3BR/PP15/04	BR04b	DQ118878
Brazil, Santarem	2004	D3BR/ST14/04	BR04c	DQ118882
Cuba	2000	Cuba116/00	CU00	AY702032
Cuba	2001	Cuba580/01	CU01	AY702030
Cuba	2002	Cuba21/02	CU02	AY702031
Ecuador, Cañar	2002	OBS8857/2000	EC02	DQ177899
Ecuador, Guayas	2002	OBS8852/2000	EC02	DQ177898
Fiji	1992	29472	FJ92	L11422
Martinique	1999	D3/H/IMTSSA-MART/1999/1243	MQ99	AY099337
Martinique	2000	D3/H/IMTSSA-MART/2000/1567	MQ00a	AY099338
Martinique	2000	D3/H/IMTSSA-MART/2000/1706	MQ00b	AY099339

Martinique	2001	D3/H/IMTSSA-MART/2001/2012	MQ01a	AY099340
Martinique	2001	D3/H/IMTSSA-MART/2001/2023	MQ01b	AY099341
Martinique	2001	D3/H/IMTSSA-MART/2001/2336	MQ01c	AY099342
Mexico	1995	MEX6097	MX95a	AY146763
Mexico, Yucatan	1995	4841/YUCATAN-MX/95	MX95b	DQ341202
Mexico, Yucatan	1996	6584/YUCATAN-MX/96	MX96	DQ341203
Mexico, Yucatan	1997	6883/YUCATAN-MX/97	MX97a	DQ341204
Mexico, Quintana Roo	1997	6889/QUINTANA ROO-MX/97	MX97b	DQ341205
Mexico, Quintana Roo	1997	6896/QUINTANA ROO-MX/97	MX97c	DQ341206
Mexico, Oaxaca	2000	OAXACA-MX/00	MX00	DQ341207
Nicaragua	1994	Nicaragua24/94	NI94	AY702033

Panama	1994	PANAMA/94	PA94	DQ341209
Paraguay, Asunción	2002	D3PY/AS12/02	PY02	DQ118884
Paraguay, Asunción	2003	D3PY/AS10/03	PY03a	DQ118883
Paraguay, Asunción	2003	D3PY/AS9/03	PY03b	DQ118885
Paraguay, Fernando de la Mora	2003	D3PY/FM11/03	PY03c	DQ118886
Paraguay, Pedro Juan Caballero	2003	D3PY/PJ4/03	PY03d	DQ118887
Paraguay, Pedro Juan Caballero	2003	D3PY/PJ5/03	PY03e	DQ118888
Paraguay, Pedro Juan Caballero	2003	D3PY/PJ6/03	PY03f	DQ118889
Paraguay, Pedro Juan Caballero	2003	D3PY/PJ7/03	PY03g	DQ118890
Paraguay, Yaguaron	2003	D3PY/YA2/03	PY03h	DQ118891
Peru, Tumbes	2000	OBT412/Tumbes-2000	PE00	DQ177903
Peru, Piura	2001	FSP581/Piura-2001	PE01a	DQ177890
Peru,	2001	OBT1467/Tu	PE01b	DQ177900

Tumbes		mbes-2001		
Peru, Loreto	2002	FSL706/Loreto-2002	PE02a	DQ177889
Peru, Iquitos	2002	IQD1728/Iquitos-2002	PE02b	DQ177895
Peru, Tumbes	2003	FST145/Tumbes-2003	PE03a	DQ177891
Peru, Iquitos	2003	IQD5132/Iquitos-2003	PE03b	DQ177896
Peru, Piura	2003	OBT2812/Piura-2003	PE03c	DQ177901
Peru, Yurimaguas	2004	FSL1212/Yurimaguas-2004	PE04a	DQ177888
Peru, Tumbes	2004	FST289/Tumbes-2004	PE04b	DQ177892
Peru, Tumbes	2004	FST 346/Tumbes-2004	PE04c	DQ177894
Peru, Iquitos	2005	MFI624/Iquitos-2005	PE05a	DQ177897
Peru, Comas	2005	OBT4024/Lima-Comas-2005	PE05b	DQ177902
Philippines	1956	H87	PH56	L11423
Puerto Rico	1977	1340	PR77	L11434
Thailand	1997	ThD3_1309_97	TH97	AY676405
Venezuela, Aragua	2000	LARD5990	VE00a	AY146764
Venezuela,	2000	LARD6007	VE00b	AY146765

Aragua				
Venezuela, Aragua	2000	LARD6218	VE00c	AY146766
Venezuela, Aragua	2000	LARD6315	VE00d	AY146767
Venezuela, Aragua	2000	LARD6318	VE00e	AY146768
Venezuela, Aragua	2000	LARD6397	VE00f	AY146769
Venezuela	2000	LARD6411	VE00g	AY146770
Venezuela, Aragua	2000	LARD6456	VE00h	AY146771
Venezuela, Aragua	2001	C02-003/Maracay 2001	VE01a	DQ367720
Venezuela, Aragua	2001	C09-006/Maracay 2001	VE01b	DQ371245
Venezuela, Aragua	2001	LARD6666	VE01c	AY146772
Venezuela, Aragua	2001	LARD6667	VE01d	AY146773
Venezuela, Aragua	2001	LARD6668	VE01e	AY146774
Venezuela, Aragua	2001	LARD6722	VE01f	AY146775
Venezuela, Aragua	2001	LARD7110	VE01g	AY146776
Venezuela, Aragua	2001	LARD7812	VE01h	AY146777

Venezuela, Aragua	2001	LARD7984	VE01i	AY146778
Venezuela, Aragua	2003	C23-009/Maracay 2003	VE03a	DQ367721
Venezuela, Aragua	2003	C29-008/Maracay 2003	VE03b	DQ367722

5. REFERENCES

Aquino, J. D. J. D., Tang, W. F., Ishii, R., Ono, T., Eshita, Y., Aono, H. & Makino, Y. (2008). Molecular epidemiology of dengue virus serotypes 2 and 3 in Paraguay during 2001–2006: The association of viral clade introductions with shifting serotype dominance. *Virus Res,* doi:10.1016/j.virusres.2008.07.011.

Aquino, V. H., Anatriello, E., Gonçalves, P. F., Da Silva, E. V., Vasconcelos, P. F. C., Vieira, D. S., Batista, W. C., Bobadilla, M. L., Vazquez, C. & other authors (2006). Molecular epidemiology of dengue type 3 in Brazil and Paraguay, 2002-2004. *Am J Trop Med Hyg* **75**, 710-715.

Anderson, C. R., Downs, W. G., & Hill, A. E. (1956). Isolation of dengue virus from a human being in Trinidad. *Science* **124**, 224-225.

Arias, J. S. & Miranda-Esquivel, D. R. (2006). Tree C. Laboratorio de Sistemática y Biogeografia. http://tux.uis.edu.co/labsist/intro.html.

Barcelos Figueiredo, L., Batista Cecílio, A., Portela Ferreira, G., Paiva Drumond, B., Germano de Oliveira, J., Bonjardim, C. A., Peregrino Ferreira, P. C. & Kroon, E. G. (2008). Dengue virus 3 genotype 1 associated with dengue fever and dengue hemorrhagic fever, Brazil. *Emerg Infect Dis* **14**, 314-316.

Barrero, P. R. & Mistchenko, A. S. (2008). Genetic analysis of Dengue virus type 3 isolated in Buenos Aires, Argentina. *Virus Res* **135**, 83–88.

Méndez, J. & Bernal. M. P. (2002). Serotipos de dengue aislados por semana epidemiológica dentro de la vigilancia de enfermedades febriles. Bogotá:

Laboratorio de Virología, Instituto Nacional de Salud. http://www.ins.gov.co/pdf_investiga/v_dengue.PDF

Boschell, J., Groot, H., Gacharna, M., Márquez, G., González, M., Gaitán, M. O., Berlie, C. H. & Martínez, M. (1986). Dengue en Colombia. *Biomedica* **6**, 101-102.

CDC (1995). Dengue Type 3 infection-Nicaragua and Panama, October-November. *MMWR Morb Mortal Wkly Rep* **44**, 21–24.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797.

Goloboff, P. A. & Farris, J. M. (2001). Methods for quick consensus estimation. *Cladistics* **17**, 26-34.

Goloboff, P., Farris, J. S. & Nixon, K. C. (2008). TNT, a free program for phylogenetic analysis. *Cladistics* **24**, 1–13.

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

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.

Gubler, D. J. (2006). Dengue/dengue haemorrhagic fever: history and current status. In *New treatment strategies for dengue and other flaviviral diseases*, 1st edn, pp. 3-22. Edited by G. Bock & J. Goode. Chichester, United Kingdom: John Wiley & Sons.

Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696-704.

Guzmán, M. G., Vázquez, S., Martínez, E., Álvarez, M., Rodríguez, R., Kourí, G., de los Reyes, J. & Acevedo, F. (1997). Dengue in Nicaragua, 1994: reintroduction of serotype 3 in the Americas. *Pan Am J Public Health* **1**, 193-199.

Holmes, E. C. (2004). The phylogeography of human viruses. *Mol Ecol* **13**, 745-756.

Kochel, T., Aguilar, P., Felices, V., Comach, G., Cruz, C., Alava, A., Vargas, J., Olson, J., & Blair, P. (2008). Molecular epidemiology of dengue virus type 3 in Northern South America: 2000-2005. *Infect Genet Evol* **8**, 682-688.

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

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

Messer, W. B., Gubler, D. J., Harris, E., Sivananthan, K. & de Silva, A. M. (2003). Emergence and global spread of a dengue serotype 3, subtype III virus. *Emerg Infect Dis* **9**, 800-809.

Nogueira, R. M. R., Miagostovich, M. P., de Filippis, A. M. B., Pereira, M. A. S. & Schatzmays, H. G. (2001). Dengue vírus type 3 in Rio de Janeiro, Brazil. *Mem Inst Oswaldo Cruz* **96**, 925-926.

Ocazionez, R. E., Cortés, F. M., Villar, L. A. & Gómez, S. Y. (2006). Temporal distribution of dengue virus serotypes in Colombian endemic area and dengue

incidence. Re-introduction of dengue-3 associated to mild febrile illness and primary infection. *Mem Inst Oswaldo Cruz* **101**, 725-731.

Ospina, M. C. (2004). Vigilancia epidemiológica del dengue en Antioquia. Memorias del 1er. Simposio Nacional de Virología, Medellín 2004. *Iatreia* **17**, 1-9.

PAHO 1995-2000. Number of Reported Cases of Dengue & Dengue Hemorrhagic Fever (DHF), Region of the Americas (by country). http://www.paho.org/English/AD/DPC/CD/dengue.htm.

Peyrefitte, C. N., Couissinier-Paris, P., Mercier-Perennec, V., Bessaud, M., Martial, J., Kenane, N., Durand, J. P. A. & Tolou, H. J. (2003). Genetic characterization of newly reintroduced dengue virus type 3 in Martinique (French Wets Indies). *J Clin Microbiol* **41**, 5195-5198.

Pinheiro, F. & Corber, S. (1997). Global situation of dengue and dengue haemorrhagic fever, and its emergence in the Americas. *World Health Stat Q* **50**, 161-169.

Posada, D. & Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.

Posada, D. & Crandall K. A. (2001). Selecting the best-fit model of nucleotide substitution. *Syst Biol* **50**, 580–601.

Rico-Hesse, R. (1990). Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* **174**, 479-93.

Rodriguez-Roche, R., Alvarez, M., Holmes, E. C., Bernardo, L., Kouri, G., Gould, E. A., Halstead, S. & Guzman, M. G. (2005). Dengue Virus Type 3, Cuba, 2000-2002. *Emerg Infect Dis* **11**, 773-774.

Ronquist, F. (1996). DIVA versions 1.2. Computer program and manual available by anonymous FTP from Uppsala University. http://www.ebc.uu.se/systzoo/research/diva/diva.html.

Ronquist, F. (1997). Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Syst Biol* **46**, 195-203.

Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10**, 512-26.

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.

Usuku, S., Castillo, L., Sugimoto, C., Noguchi, Y., Yogo, Y. & Kobayashi, N. (2001). Phylogenetic analysis of dengue-3 viruses prevalent in Guatemala during 1996-1998. *Arch Virol* **146**, 1381-1390.

Uzcategui, N. Y., Comach, G., Camacho, D., Salcedo, M., Cabello de Quintana, M., Jimenez, M., Sierra, G., Cuello de Uzcategui, R., James, W. S. & other authors (2003). Molecular epidemiology of dengue virus type 3 in Venezuela. *J Gen Virol* **84**, 1569-1575.

Varón, A., Vinh L. S., Bomash I., Wheeler W. C. (2008). POY 4.0.2885. American Museum of Natural History. http://research.amnh.org/scicomp/projects/poy.php.

Wheeler, W. (1996). Optimization alignment: the end of multiple sequence alignment in phylogenetics?. *Cladistics* **12**, 1-9.

Wheeler, W. C. (2001). Homology and the optimization of DNA sequence data. *Cladistics* **17**, 3–11.

Wheeler, W., Aagesen, L., Arango, C. P., Faivovich, J., Grant, T., D'Haese, C., Janies, D., Smith, Wm. L., Varón, A. & Giribet, G. (2006). Character optimization. In *Dynamic Homology and phylogenetic systematics: a unified approach using POY*, 1st edn, pp. 37-62. New York: American Museum of Natural History

Wittke, V., Robb, T. E., Thu, H. M., Nisalak, A., Nimmannitya, S., Kalayanrooj, S., Vaughn, D. W., Endy, T. P., Holmes, E. C. & Aaskov, J. G. (2002). Extinction and rapid emergence of strains of dengue virus during an interepidemic period. *Virology* **301**, 148-156.

WHO (1997). General considerations. In *Dengue Hemorrhagic fever: diagnosis, treatment and control*, 2nd edn, pp. 1-10. Geneva: World Health Organization

WHO (2000). Strengthening implementation of the global strategy for dengue fever/dengue haemorrhagic fever prevention and control. Report of the informal Consultation, 18-20 October 1999. Geneva: World Health Organization