Am. J. Trop. Med. Hyg., 71(4), 2004, pp. 478–484

COMPARISON OF DENGUE-1 VIRUS ENVELOPE GLYCOPROTEIN GENE SEQUENCES FROM FRENCH POLYNESIA

MANOLA LAILLE AND CLAUDINE ROCHE

Laboratoire de Recherche en Virologie Médicale, Institut Louis Malardé, Papeete, Tahiti, Polynésie Française

Abstract. Dengue (DEN) is the leading arboviral infection of humans, with 100 million cases annually in the tropical areas of the world. The recent severe DEN-1 epidemic in French Polynesia in 2001, with an incidence rate of 16% and more than 45% of the cases with dengue hemorrhagic fever/dengue shock syndrome among 1,400 hospitalized children and eight fatalities, led us to study this new circulating strain. The entire envelope (E) gene of two French Polynesian DEN-1 virus isolates from the two epidemics of 1988–1989 (FP89) and 2001 (FP01) were sequenced and compared with 29 published DEN-1 virus E gene sequences. Phylogenetic relationships showed that the FP89 strain belonged to genotype V and the FP01 strain to genotype IV based on studies on the same region of DEN-1 virus genome (1,485 nucleotides). The recent dengue epidemic in French Polynesia in 2001 was probably due to the introduction of a new DEN-1 virus from Southeast Asia, since the minimum nucleotide divergence was 3.3% with A88, the Indonesian strain isolated in 1988 in Jakarta.

INTRODUCTION

Over the past 20 years, a dramatic increased frequency of epidemic dengue fever (DF) and the emergence of the most severe form, dengue hemorrhagic fever/dengue shock syndrome (DHF/DSS), were reported in tropical regions worldwide.[1–3] Currently, dengue is endemic/epidemic in tropical areas including Asia, Africa, the Americas, and some Pacific Islands. There are 100 million DF cases and hundreds of thousands of DHF/DSS cases per year. Overall, the case-fatality rate of DHF in Asian countries and the Americas is 4–8%.[4]

In French Polynesia, an alarming increase in dengue transmission has been observed since the dengue-2 (DEN-2) outbreak in 1971, in which there were more than 35,000 cases, including the first three deaths recorded.[5] Subsequently, dengue outbreaks involving all four dengue serotypes (DEN-1, DEN-2, DEN-3, and DEN-4) have occurred: DEN-1 in 1975–1976, 1988–1989, and 2001; DEN-4 in 1979; DEN-3 in 1989–1990, and DEN-2 in 1996–1997.[6,7] No co-circulation of the four serotypes has been observed in French Polynesia.

To determine the origin and evolution of dengue viruses and to examine virus diversity, sequence analysis of the premembrane (prM) locus,[8,9] the envelope (E) gene, the E/nonstructural 1 (NS1) junction region, and other genes such as nonstructural 3 (NS3) gene has been undertaken by several groups. With regard to DEN-1 viruses, studies suggest there up to five DEN-1 virus genotypes or groups that cluster geographically.[10–12]

The viral E protein of dengue virus contains the sites required for virus binding and penetration into susceptible cells. Within the E protein, neutralizing epitopes are localized with two potential N-linked glycosylation sites, Asn-67 and Asn-153. Mutations of these N-linked glycosylation sites may affect virus-mediated membrane fusion and neurovirulence.[13,14] Therefore, most sequence analysis of the dengue viruses have been performed on this gene.

For surveillance and epidemiologic studies, we compared the entire DEN-1 virus E gene sequences derived from French Polynesian strains isolated during the 1988–1989 and 2001 epidemics with those described previously and available in the GenBank library (Table 1). The E gene sequence of the Singapore S275/90 strain (SIN90) included in this study was that revised on July 5, 2002 and is also available on the GenBank library. No deletion or extra cysteine residues at amino acid positions 389-390 were observed in this latter strain.

MATERIALS AND METHODS

Virus isolation. The DEN-1 virus strains were first isolated from acute-phase sera by inoculation into *Aedes albopictus* clone C6/36 cell line (American Type Culture Collection, Manassas, VA). After a one-hour adsorption of the inoculum (serum diluted 1:40 in culture medium) onto the cells at 28°C, cell cultures were incubated for seven days at 28°C. Cells were harvested and infection was confirmed by an immunofluorescence assay using serotype-specific monoclonal antibodies as previously described.[15] The C6/36 cells were grown at 28°C to confluency in RPMI 1640 medium (Sigma, St. Louis, MO) supplemented with 10% fetal calf serum (Sigma) and 2.5 μg/mL of fungizone (Sigma).

Extraction of RNA and reverse transcriptase–polymerase chain reaction (RT-PCR). The virus genomic RNA was extracted from either acute-phase sera from a DF patient directly (FP89) or from the culture media virus from infected C6/36 cells collected after one passage (FP01) using the QIAamp viral RNA mini kit (Qiagen Inc., Hilden, Germany) according to manufacturer's instructions. Briefly, a 140 μL aliquot of serum or infected cellular supernatant was mixed with the buffer AVL-carrier RNA, incubated for 10 minutes, and loaded onto the spin column. This was followed by two washes with buffer AW. The final nucleic acid extracts were obtained in a total volume of 60 μL of elution buffer AVE.

Oligonucleotide primers (both sense and antisense) used in the amplification and sequencing protocols were designed with the OMIGA 1.1 software (Oxford Molecular Ltd., Oxford, United Kingdom). For the entire E gene, three primer pairs were used as follows: D1/732 (5′-GAAACAAGAACCGAAACG-3′) and D1/1509 (5′-CCATCTCATTAAAGTCCAGC-3′), D1/1399 (5′-CAATTGCAACCATAACACC-3′) and D1/2106 (5′-TGCTTCCTTTCTTGAACC-3′), and D1/1966 (5′-GATTGATAACAGCCAATCC-3′) and D1/2685 (5′-CAACAACATCTCCTACAACC-3′). The nucleic acid sequences of these three E selected regions of DEN-1 virus were amplified using the OneStep RT-PCR Kit (Qiagen Inc.). Briefly, 10 μL of extracted RNA were added to 40 μL of an RT-PCR mixture as recommended by the manufacturer. Reverse transcription

TABLE 1
Dengue-1 virus isolates used for genome sequence analysis*

Geographic origin	Isolation year	Host	Strain	Code	GenBank accession no.	Reference
Angola	1988	H	RIO H 36589	ANG88	AF425610	12
Argentina	2000	H	297arg00	ARG297	AF514889	30
Argentina	2000	H	301arg00	ARG301	AF514876	30
Aruba	1985	H	st495-1	ARU85	AF425609	12
Australia	1983	H	AUS HAT17	AUS83	AF425612	12
Brazil	1990	H	BR/90	BR90	AF226685	31
Brazil	1997	H	BR/97-111	BR971	AF311956	31
Brazil	1997	H	BR/97-233	BR972	AF311958	31
Brazil	1997	H	BR/97-409	BR974	AF311957	31
Brazil	2001	H	BR/01-MR	BR01	AF513110	32
Cambodia	1998	H	CAMB/98/658	CAM98	AF309641	24
Colombia	1985	H	INS 347869	COL85	AF425616	12
Colombia	1996	H	INS 371869	COL96	AF425617	12
Costa Rica	1993	H	Cesara 1	CR93	AY153755	
Djibouti	1998	H	DJIB/98/606	DJI98	AF298808	24
French Guiana	1989	H	FGA/89	FGA89	AF226687	31
French Polynesia	1989	H	FP/89/5103	FP89	AY630408	
French Polynesia	2001	H	FP/01/192206	FP01	AY630407	
Indonesia	1988	H	A88	A88	ABO74761	19
Jamaica	1977	H	PRS 288690	JAM77	AF425621	12
Malaysia	1972	M	P72-1244	MAL72	AF425622	12
Nauru	1974	H	WPAC74	NAU74	U88535	32
Nigeria	1968	H	IBH 28328	NIG68	AF425625	12
Paraguay	2000	H	280par00	PAR00	AF514878	30
Peru	1991	H	DEI 0151	PER91	AF425626	12
Philippines	1984	H	836-1	PHI84	D00503	33
Singapore	1990	H	S275/90	SIN90	M87512	34
Thailand	1954	H	TH-SMAN	THA54	D10513	35
Thailand	1980	H	AHF82-80	THA80	D00502	33
Trinidad	1986	H	CAREC 86471	TRI86	AF425639	12
Venezuela	1995	H	5345	VEN95	AF425635	12

* H = human; M = monkey.

TABLE 2
Nucleotide and amino acid pairwise distances between dengue-1 virus strains used in this study for the envelope glycoprotein*

	FP89	FGA89	ARU85	JAM77	CR93	ANG88	BR01	ARG301	BR972	BR974	BR971	VEN95	BR90	COL85	TRI86
FP89		0.6	0.4	0.6	1.0	1.0	1.0	1.2	1.0	0.8	0.8	0.8	0.8	1.2	0.6
FGA88	0.8		0.6	0.8	1.2	1.2	1.2	1.4	0.8	1.0	1.0	1.0	1.0	1.4	0.8
ARU85	0.8	1.0		0.4	0.8	0.8	0.8	1.0	0.8	0.6	0.6	0.6	0.6	1.0	0.4
JAM77	1.7	1.8	0.9		0.4	0.4	0.4	0.6	0.4	0.2	0.2	0.2	0.2	0.6	0.0
CR93	2.3	2.3	1.4	1.1		0.8	0.8	1.0	0.8	0.6	0.6	0.6	0.6	1.0	0.4
ANG88	2.3	2.4	1.5	1.0	1.4		0.4	0.6	0.8	0.6	0.6	0.6	0.6	1.0	0.4
BR01	3.1	3.1	2.3	1.8	2.2	2.1		0.2	0.4	0.2	0.2	0.2	0.2	1.0	0.4
ARG301	3.3	3.3	2.5	2.1	2.6	2.5	0.7		0.6	0.4	0.4	0.4	0.4	1.2	0.6
BR972	3.1	2.8	2.4	1.9	2.3	2.4	0.5	1.0		0.2	0.2	0.2	0.2	1.0	0.4
BR974	2.9	2.8	2.3	1.8	2.3	2.3	0.5	1.0	0.2		0.0	0.0	0.0	0.8	0.2
BR971	3.2	3.0	2.5	2.0	2.4	2.5	0.7	1.2	0.5	0.4		0.0	0.0	0.8	0.2
VEN95	2.5	2.6	1.7	1.2	1.9	1.7	1.4	1.8	1.5	1.4	1.6		0.0	0.8	0.2
BR90	2.6	2.5	1.7	1.2	1.9	1.8	1.5	1.8	1.6	1.6	1.7	1.1		0.8	0.2
COL85	2.3	2.4	1.5	1.0	1.7	1.6	2.1	2.3	2.3	2.2	2.3	1.6	1.4		0.6
TRI86	1.9	2.0	1.1	0.6	1.3	1.2	1.8	2.2	2.0	1.9	2.1	1.3	1.3	0.8	
COL96	3.1	3.2	2.4	1.9	2.6	2.6	3.0	3.6	3.3	3.3	3.4	2.6	2.5	2.3	1.9
PAR00	3.8	3.8	3.0	2.5	3.0	2.8	3.5	3.7	3.6	3.5	3.6	3.0	2.9	2.4	2.4
ARG297	3.8	3.8	3.0	2.5	3.0	2.8	3.5	3.7	3.6	3.5	3.5	3.0	2.9	2.4	2.4
PER91	3.0	3.0	2.1	1.6	2.2	2.0	2.6	3.1	2.9	2.8	2.9	2.3	2.2	1.9	1.6
NIG68	4.8	5.0	3.9	3.7	4.0	4.2	4.6	4.8	4.7	4.6	4.8	4.2	4.4	3.9	3.6
FP01	10.1	10.0	9.5	9.4	9.4	9.5	9.7	10.0	10.1	10.0	10.2	9.7	9.5	9.6	9.2
A88	9.9	9.9	9.3	9.1	9.3	9.3	9.4	9.4	9.6	9.6	9.8	9.4	9.1	9.4	9.0
AUS83	9.4	9.3	8.6	8.5	8.5	8.5	8.9	9.0	9.3	9.2	9.4	9.2	8.8	8.7	8.3
PHI84	9.4	9.2	8.6	8.3	8.4	8.4	8.9	9.1	9.2	9.1	9.4	8.9	8.6	8.5	8.1
NAU74	9.0	8.9	8.3	7.9	8.1	8.1	8.6	8.8	8.9	8.8	9.0	8.6	8.3	8.2	7.8
SIN90	9.0	8.7	8.3	7.9	8.1	8.1	8.6	8.8	8.9	8.8	9.0	8.6	8.3	8.2	7.8
THA80	8.6	8.5	7.9	7.6	7.9	7.9	8.3	8.6	8.6	8.6	8.7	8.2	7.8	7.9	7.5
THA54	8.7	8.7	8.0	7.3	7.8	8.2	8.0	8.4	8.5	8.4	8.7	8.0	7.9	7.7	7.4
MAL72	8.2	8.3	7.6	7.4	7.6	7.3	8.0	8.1	8.4	8.3	8.5	8.0	7.3	7.4	7.2
CAM98	9.1	9.0	8.3	8.0	8.3	8.2	8.9	8.9	9.0	8.9	9.0	8.7	8.6	8.5	8.2
DJI98	9.1	9.0	8.8	8.6	8.3	8.7	8.9	9.1	9.2	9.1	9.2	9.1	9.1	8.9	8.8

was performed at 50°C for 30 minutes, followed directly by 30 cycles of amplification at 94°C for 30 seconds, 54°C for one minute, and 72°C for one minute, and a final extension at 72°C for 10 minutes. The RT-PCR was performed in single-tube (0.2 μL) in a thermocycler (Model Mastercycler® Gradient 5331; Eppendorf, Hamburg, Germany). Electrophoresis of the amplified DNA products was then performed on a 1.5% (w/v) agarose gel and staining with ethidium bromide (0.5 μg/mL). The expected band was sliced from the gel, and purified using the QIAquick gel extraction kit (Qiagen Inc).

Sequencing of DNA. Sequencing reactions were performed using the BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer Corp.–Applied Biosystems Inc., Warrington, United Kingdom). Unincorporated dideoxy terminators were removed using the DyeEx Spin columns (Qiagen Inc). The products were analyzed using an automated ABI Prism 310 genetic analyzer (Applied Biosystems, Foster City, CA). Both strands of the cDNA PCR products were sequenced.

Phylogenetic analysis. The overlapping nucleic acid sequences obtained from individual sequencing reactions were combined for analysis and edited using the AutoAssembler 2.1 software (Perkin-Elmer Corp.–Applied Biosystems Inc.). Multiple alignment of sequences from French Polynesian isolates and the published E DEN-1 virus sequences accessed from the GENBANK library (Table 1) was performed using the CLUSTAL W 1.81[16] algorithm with default parameters. A phylogenetic tree was constructed using the neighbor joining method[17] using maximum likelihood distance parameters (Kimura two-parameter formula) using the Molecular Evolutionary Genetics Analysis software package.[18] Bootstrap analysis with 1,000 replicates was used to determine the robustness of the tree and the evolutionary relationship of the viruses.

RESULTS

Six overlapping E segment PCR products were amplified and sequenced in both directions, covering a continuous region of the entire E gene (1,485 nucleotides) of two French Polynesian DEN-1 virus isolates (FP89 and FP01). Nucleotide sequence identity among all the available sequenced DEN-1 viruses ranged from 89.8% to 99.8%.

Sequence comparisons of the entire E gene of the 31 DEN-1 virus isolates showed a maximum divergence of 10.2% between BR971 and FP01, which translates into a 1.6% divergence at the protein level. The minimum nucleotide divergence of 0.2% was observed between two Brazilian 1997 strains, BR972 and BR974, with 0.2% amino acid divergence (Table 2). It is interesting to note that the FP01 strain is the most divergent DEN-1 virus isolate at the gene level (9.2–10.2%), while the Australian strain AUS83 is the most divergent at the protein level (1.8–3.6%).

Alignment of the E sequences of the two French Polynesian DEN-1 virus strains showed considerable variability. The nucleotide and amino acid divergence rates between DEN-1 virus strains in E gene are as follows: FP89 versus FP01, 10.1% and 2.2%; FP89 versus NAU74 (a prototype strain), 9.0% and 2.6%; and FP01 versus NAU74, 3.8% and 1.2%, respectively.

Table 2
Continued

COL96	PAR00	ARG297	PER91	NIG68	FP01	A88	AUS83	PHI84	NAU74	SIN90	THA80	THA54	MAL72	CAM98	DJI98
1.2	1.4	1.4	1.0	1.6	2.2	2.4	3.4	2.8	2.6	2.6	2.8	2.4	2.2	3.0	2.8
1.4	1.6	1.6	1.2	1.8	2.4	2.6	3.6	3.0	2.8	2.8	3.0	2.6	2.4	3.2	3.0
1.0	1.2	1.2	0.8	1.4	2.0	2.2	3.2	2.6	2.4	2.4	2.6	2.2	2.0	2.8	2.6
0.6	0.8	0.8	0.4	1.0	1.6	1.8	2.8	2.2	2.0	2.0	2.2	1.8	1.6	2.4	2.2
1.0	1.2	1.2	0.8	1.4	2.0	2.2	3.2	2.6	2.4	2.4	2.6	2.2	2.0	2.8	2.6
1.0	1.2	1.2	0.8	1.4	2.0	2.2	3.2	2.6	2.4	2.4	2.6	2.2	2.0	2.8	2.6
1.0	1.2	1.2	0.8	1.4	1.8	1.8	3.0	2.6	2.4	2.4	2.4	2.2	2.0	2.8	2.6
1.2	1.4	1.4	1.0	1.6	2.0	2.0	3.2	2.8	2.6	2.6	2.6	2.4	2.2	3.0	2.8
1.0	1.2	1.2	0.8	1.4	1.8	1.8	3.0	2.6	2.4	2.4	2.4	2.2	2.0	2.8	2.6
0.8	1.0	1.0	0.6	1.2	1.6	1.6	2.8	2.4	2.2	2.2	2.2	2.0	1.8	2.6	2.4
0.8	1.0	1.0	0.6	1.2	1.6	1.6	2.8	2.4	2.2	2.2	2.2	2.0	1.8	2.6	2.4
0.8	1.0	1.0	0.6	1.2	1.6	1.6	2.8	2.4	2.2	2.2	2.2	2.0	1.8	2.6	2.4
0.8	1.0	1.0	0.6	1.2	1.6	1.6	2.8	2.4	2.2	2.2	2.2	2.0	1.8	2.6	2.4
1.2	1.4	1.4	1.0	1.6	2.2	2.4	3.4	2.8	2.6	2.6	2.8	2.0	1.8	3.0	2.8
0.6	0.8	0.8	0.4	1.0	1.6	1.8	2.8	2.2	2.0	2.0	2.2	1.8	1.6	2.4	2.2
	1.0	1.0	1.0	1.6	2.2	2.4	3.4	2.8	2.6	2.6	2.8	2.4	2.2	3.0	2.8
3.3		0.0	0.4	1.6	1.6	1.8	2.8	2.2	2.0	2.0	2.6	2.6	2.4	2.8	2.6
3.3	0.4		0.4	1.6	1.6	1.8	2.8	2.2	2.0	2.0	2.6	2.6	2.4	2.8	2.6
2.8	1.2	1.2		1.2	1.4	1.6	2.6	2.0	1.8	1.8	2.4	2.2	2.0	2.4	2.2
5.0	5.3	5.3	4.6		2.4	2.6	3.6	3.0	2.8	2.8	3.0	2.8	2.6	3.4	3.2
9.6	9.7	10.1	9.7	9.6		0.8	1.8	1.4	1.2	1.2	1.6	2.6	2.0	2.4	2.2
9.8	9.4	9.7	9.4	8.7	3.3		2.0	1.6	1.4	1.4	1.8	2.8	2.0	2.6	2.4
9.3	8.9	9.3	8.9	8.9	4.6	4.1		2.6	2.0	2.0	2.6	3.4	2.8	3.0	3.0
9.0	8.8	9.1	8.7	8.2	4.2	4.0	3.5		0.6	1.0	1.2	3.0	2.6	3.0	2.4
8.7	8.5	8.8	8.4	7.9	3.8	3.6	2.9	0.5		0.4	1.4	2.6	2.0	2.4	2.2
8.7	8.5	8.8	8.4	7.9	3.6	3.6	2.9	0.8	0.3		1.4	2.6	2.0	2.4	2.2
8.3	8.5	8.8	8.3	7.8	5.4	5.2	4.5	2.1	2.3	2.3		3.2	2.6	3.4	3.2
8.2	8.5	8.8	8.1	7.3	7.9	8.6	7.4	6.8	6.6	6.6	7.2		2.2	3.0	2.6
8.2	8.4	8.7	8.1	6.9	8.3	8.2	7.4	7.1	6.8	6.8	7.3	6.5		2.4	2.2
9.2	8.0	8.2	8.2	8.6	9.0	9.2	8.1	7.9	7.6	7.6	8.4	7.8	8.0		1.4
9.1	8.5	8.7	8.8	8.9	9.5	9.2	8.5	8.2	8.1	8.1	8.7	7.9	8.3	4.0	

* Nucleotide pairwise distances (%) are shown below the diagonal and amino acid pairwise distances (%) are shown above the diagonal.

According to the classification of Rico-Hesse[10] in which maximum genetic distance within a given genotypic group or subgroup was <6%, five genotypes were observed from the complete sequence of the E gene of 31 DEN-1 virus strains involved worldwide. Phylogenetic analysis of the E gene sequences using the neighbor-joining method (Figure 1) showed these five clusters and confirmed the results of previous studies.[12] These include cluster I, the first dengue-1 virus strains from Japan, Hawaii, Asia, and Djibouti; cluster II, Thai strains; cluster III, the Malaysian sylvatic strain; cluster IV, strains from southeast Asia, the south Pacific, and Australia; and cluster V, strains from the Americas, Africa, and southeast Asia.

The two French Polynesian strains (FP89 and FP01) are assigned to group V and IV, respectively (Figure 1). Within a genotype, nucleotide divergence among viruses was 4.0% for genotype I, ranging from 0.3% to 5.4% and 0.2% to 5.3% for genotypes IV and V, respectively. Within genotype V, FGA89 and ARU85 are the most closely related strains to the FP89 isolate, with 0.8% divergence at the nucleotide level and 0.6% and 0.4% divergence at amino acid level, respectively.

Moreover, within genotype IV, DEN-1 virus strain FP01 showed the highest (5.4%) nucleotide divergence rate when compared with the THA80 strain (AHF82-80 strain) isolated in Thailand in 1980 and the lowest (3.3%) nucleotide divergence rate when compared with the A88 strain, an Indonesian strain isolated in 1988 from a DHF patient.[19] It is interesting to note the lowest divergence (3.6%) at nucleotide level of this strain with the revised S275/90 strain.

The deduced amino acid sequences of the entire E gene were aligned and compared. The E proteins have retained an amino acid sequence similarity greater than 96% over the studied 46-year period. These data suggest that domains responsible for predicted flavivirus protein architecture and/or biologic function are strictly conserved. More than 80% of the nucleotide mutations were silent and among the other mutations, most are transitions.

Within the E glycoprotein, 59 of 495 amino acid changes were observed when the worldwide DEN-1 virus strains were compared. The amino acid sequence changes are summarized in Table 3.

Analysis of the aligned amino acid sequences of the two French Polynesian strains (FP89 and FP01) showed 11 changes (Table 4). The five amino acid substitutions at positions 37 (domain I), 339, 352, 379 (domain III), and 439 (trans-membrane region) are non-conservative. Domain III (amino acids 302-404) is thought to play an important role in cell attachment and tropism. The impact of these changes on biologic properties and disease outcome is not known. The six

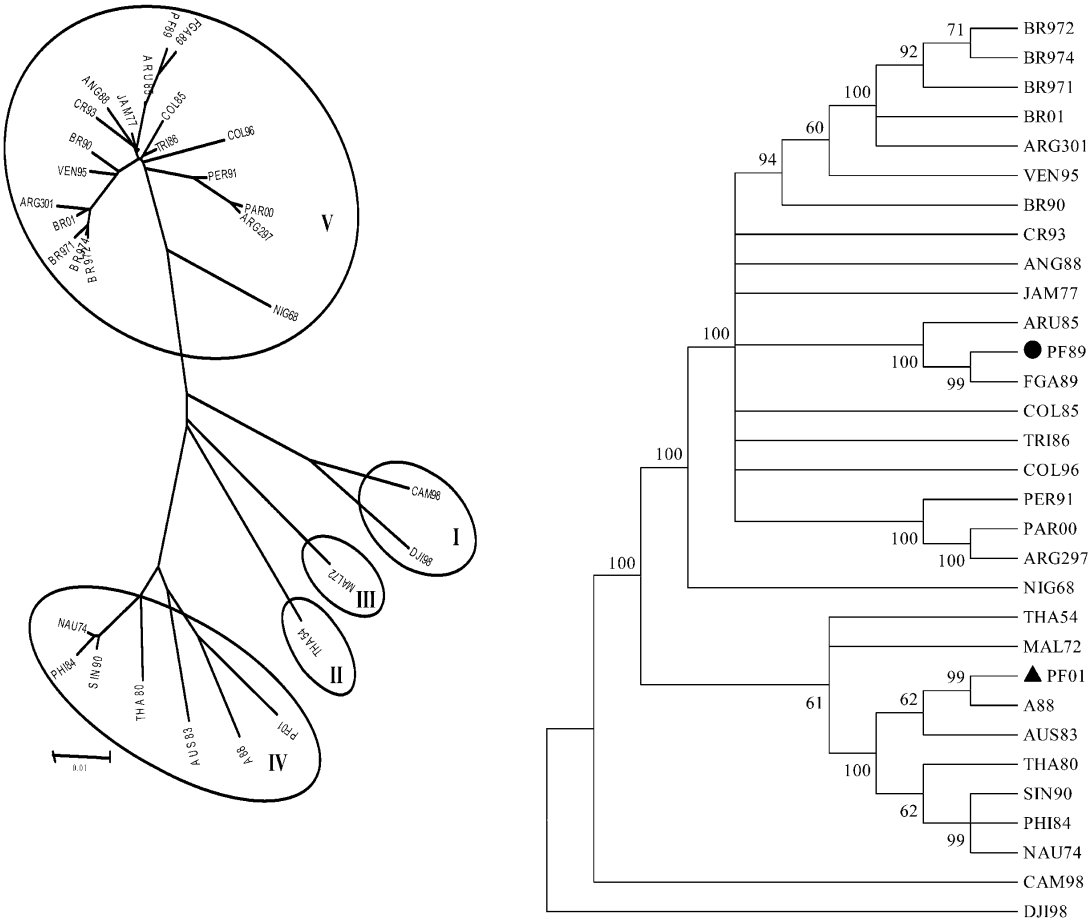

FIGURE 1. Phylogenetic tree derived from envelope protein gene nucleotide sequence of 31 strains of dengue-1 (DEN-1) viruses generated by neighbor-joining analysis. Twenty-nine nucleic acid sequences of DEN-1 virus strains involved worldwide were obtained from GenBank. The numbers displayed next to the nodes correspond to the bootstrap values (1,000 replicates) supporting that genotype. Boostrap values ≥70 corresponded to a probability of 95% or higher that the corresponding cluster was correctly identified.[22]

TABLE 3

Summary of amino acid differences among all dengue-1 virus genotypes*

E gene position	NAU74 strain	Genotype I	Genotype II	Genotype III	Genotype IV	Genotype V
6	I	·	·	V	·	V†
7	G	·	·	·	C†	·
8	N	S	·	·	I†	·
10	D	·	·	·	H†	·
37	D	·	N	N	·	N†
70	T	·	·	·	·	P†
77	Q	·	·	·	·	H†
88	A	·	·	·	T‡	S†
96	T	·	·	·	·	V†
114	I	·	·	·	·	**L**
120	K	·	·	E	·	·
122	V	·	A	·	·	·
132	Y	·	·	·	·	H†
155	T	**S**	·	·	·	·
157	E	·	·	G	·	·
161	T	·	·	·	I‡	**I**
171	S	T	·	·	·	·
173	I	·	·	·	V†	·
180	A	·	·	·	·	T†
195	E	·	·	·	R†	·
196	M	·	·	·	V†	·
202	K	·	·	·	E†	·
203	E	·	·	·	K‡	·
225	S	T	·	·	·	
226	T	·	·	·	·	I†
251	V	·	A	·	·	·
287	L	·	·	·	·	I†
293	T	·	·	·	I†	·
297	M	·	·	·	I†/T†/V†	T†
305	S	·	·	·	P†	·
324	V	I	·	·	·	·
337	F	·	·	·	·	L†
338	S	·	·	·	·	L†
339	S	**T**	T	T	·	A†/I†
345	V	·	·	I	·	A†
351	L	V	·	·	·	·
352	I	·	·	·	V†	·
359	T	·	·	·	I†	·
365	V	·	·	·	·	I†
369	A	·	·	·	·	**T**
379	V	·	·	·	·	I†/M†
380	V	I	·	·	·	·
386	A	·	·	·	·	T†
397	S	T	I		T†	·
399	G	·	·	·	W†	·
402	F					L†
425	I	·	·	·	I	M†
428	V	·	·	·	·	M†
432	V	·	M	·	·	M†
436	I	**V**	V	V	V†	**V**
439	I	·	V	·	·	**V**
442	T	·	·	·	·	A†
461	I	**V**	·	·	·	·
472	S	·	N	·	·	·
473	T	·	·	·	·	A†
478	T	·	M	·	·	·
481	A	·	·	·	·	V†
484	M	**L**	·	·	·	·
494	Q	·	·	·	H†	·

* Amino acid positions were numbered by reference to the sequence of the Nauru Western Pacific strain.

E = envelope; Dots = no change compared with the prototype strain Nau74. Changes for all strains within the genotype are in **bold.**

† Change for some strains within the genotype.

‡ Change for the majority of strains within the genotype.

TABLE 4

Amino acid changes in the envelope (E) protein between the two French Polynesian dengue-1 virus strains compared to the prototype strain Nau74 and to a consensus generated from all strains used for each genotype

E gene position	FP/89	FP/01	Nau74	Consensus
37	N	D	D	N
88	A	T	T	A
114	L	I	I	L
287	I	L	L	L
297	M	I	M	M
339	T	S	S	T
352	I	V	I	I
369	T	A	A	T
379	I	V	V	V
439	V	I	I	V
473	A	T	T	T

other amino acid substitutions at positions 88, 114, 287, 297, 369, and 473 are conservative.

DISCUSSION

The E glycoprotein is the major structural protein exposed on the surface of dengue viruses. It encodes the important biologic functions of viral attachment to a specific cell surface receptor and membrane fusion.[13,14] The E protein is also both a target and a modulator of the host immune response.[20]

Studies of dengue virus molecular evolution and sequence data were performed to determine phylogenetic relationships of the viruses within each serotype. A threshold of 6% divergence is currently used to separate different genotypes within a dengue serotype.[10,21] Several phylogenetic studies analyzing a large number of DEN-1 virus genomic sequences have been published.[10,11] They analyzed either a 240 nucleotide sequence in the E/NS1 junction region of the genome or a 180 nucleotide sequence in the E gene and suggested that there were potentially five or three DEN-1 virus genotypes, respectively. A recent study, based on nucleotide sequences of the entire E gene,[12] showed that DEN-1 viruses can be classified into five genotypes or monophyletic groups.

The present study confirmed the existence of these five genotypes with variations in the classification of the Singapore S275/90 strain. This strain belongs to genotype IV instead of genotype V, as in the last analysis.[12] No deletion at nucleotide positions 1,045–1,047 was observed in the July 5, 2002 revised sequence (1,485 nucleotides instead of 1,482 nucleotides). Moreover, the THA80 (AHF82-80) strain isolated in 1980 in Thailand was assigned to genotype IV in this study, which is different from the analysis of Goncalvez and others,[12] in which Thai80 (genotype I) corresponded to another isolate, the PUO 359 strain, also isolated in 1980 in Thailand. American and African strains were distinct from other isolates. The Malaysian sylvatic monkey isolate represents one distinct genotype, although bootstrap values supporting these groupings were low (Figure 1).

Bootstrap analysis showed that while several of the major nodes of the tree were not well supported (values $\leq$60%), many others were robust (values $\geq$70%).[22] In most phylogenetic analyses, bootstrap values provide highly conservative estimates of the probability of correctly inferring the corresponding clusters or clades.

The phylogram generated by the sequencing of the entire E gene showed that the French Polynesian isolate FP89 be-

longed to subtype V. The high similarity obtained by nucleotide and amino acid sequence between ARU85 and FP89 (99.2% and 99.6%, respectively) confirmed the American origin of DEN-1 viruses isolated in 1989 in French Polynesia, which were previously demonstrated by comparing a 180-basepair fragment in the E gene.[11] For the latest DEN-1 virus involved in the epidemic in 2001 in French Polynesia, it was assigned to genotype IV. Pairwise comparisons of strains isolated in Indonesia in 1988 (A88) and in French Polynesia in 2001 (FP01) showed a nucleotide divergence of 3.3% for the entire E gene. This observation suggests that the recent and severe epidemic in French Polynesia was probably due to the introduction of genotype IV DEN-1 viruses that were previously involved in an outbreak of DHF in Jakarta, Indonesia in 1988,[23] rather than the re-emergence of a previously strain circulating in the area at a lower level between epidemics. It must be noted that the recombinant strain[24] S275/90 (SIN90) is also closely related to the FP01 isolate with a 3.6% nucleotide divergence.

Several studies have purported to show that dengue virus genotypes differ in their ability to cause severe dengue.[25,26] The two French Polynesian strains isolated in 1989 and 2001 belong to two different genotypes, V and IV, respectively. Epidemiologic differences in disease severity were observed between the DEN-1 epidemics in French Polynesia in 1989 and 2001: in 1989, only DF was reported whereas DHF/DSS occurred at a rate of 2.7 per 1,000 people in 2001. In addition, eight fatal cases were reported among children in the 2001 epidemic. Among the severe hospitalized cases, laboratory tests necessary to determine infection status (primary versus secondary) were performed for 156 cases. A total of 9.6% were primary infections and 73% of these were in individuals less than one year of age (Laille M and others, unpublished data). High-level transmission of this DEN-1 virus was also observed in Hawaii, the Samoan Islands,[27] Fiji, and New Caledonia from 2001 to 2003.(Fijian and New Caledonian Health Authorities, unpublished data). The observed increased severity of DF since 2001 in the Pacific region might be associated with the introduction of this new genotype (IV) from Indonesia, as well as the Southeast Asian genotype DEN-2, which has been reported to potentially cause severe disease.[28]

No relatedness between nucleotide sequence and disease severity could be demonstrated when the French Polynesian E sequences determined from DF cases were compared. Sequence analysis of full-length of French Polynesian virus type 1 strain genomes (from DF and DHF/DSS cases) should be investigated to identify nucleotide and/or amino acid differences responsible for the virulent genotype.

The E protein sequence of geographically diverse DEN-1 virus strains is highly conserved, with an amino acid similarity of at least 96% (Table 2). Conservation of the 12 cysteine residues involved in disulfide bridge formation was observed. The glycine-rich internal element, the flavivirus fusion domain, (amino acids 98-111), is conserved among DEN-1 virus E proteins, as well as the neutralizing epitopes localized in the E protein with two potential N-linked glycosylation sites (Asn-67 and Asn-153).

Nucleotide changes were generally conservative, leading to few amino acid changes (Table 3). The impact of these amino acid changes on biologic properties and pathogenic properties is not known. Nevertheless, in the FP89 and FP01 strains and among the 11 amino acid substitutions, three significant changes were observed within domain III (amino acids 302–404) of the E protein: T 339 S, I 352 V, and I 379 V (Table 4). This domain seems to interact with cellular receptors for virus attachment and entry. Comparison of these three changes with the consensus deduced amino acid sequence generated from all included strains showed that the changes T 339 S and I 352 V may be responsible in part for the observed differences in terms of disease severity and incidence of DHF/DSS. Another observation was that DEN-1 virus FP01 had an N→D non-conservative change at position 37 (domain I) and a V→I significant amino acid substitution at position 439 (trans-membrane region) in the E protein.

Analyzing phylogenetic relationships of dengue virus strains associated with epidemics of DF compared with DHF/DSS, including our French Polynesian strains, may contribute to a better understanding of the pathogenesis of a dengue virus infection, particularly if the amino acid changes showed by sequence analysis may be related to an increase in virulence. Characterization of genetic determinants for virulence would provide further support for the hypothesis that epidemics of DHF/DSS may be caused by the circulation of viral strains with increased virulence[28,29] and expand the range of pathogenic properties for increased transmittability (high viremia).

Received October 28, 2003. Accepted for publication May 28, 2004.

Acknowledgments: We are grateful to Dr. Taiana Darius for help and advice with phylogenetic analysis. We also thank Dr. Allison Imrie for reviewing the manuscript.

Authors' address: Manola Laille and Claudine Roche, Laboratoire de Recherche en Virologie Médicale, Institut Louis Malardé, PO Box 30, 98713 Papeete Tahiti, Polynésie Française, Telephone: 689-41-64-68, Fax: 689-43-15-90, E-mails: mlaille@ilm.pf and croche@ilm.pf.

REFERENCES

1. Monath TP, 1994. Dengue: the risk to developed and developing countries. *Proc Natl Acad Sci USA 91:* 2395–2400.
2. Gubler DJ, 1997. Its history and resurgence as a global public health problem. Gubler DJ, Kuno G, eds. *Dengue and Dengue Hemorrhagic Fever.* London: CAB International, 1–22.
3. McBride WJ, Bielefeldt-Ohmann H, 2000. Dengue viral infections: pathogenesis and epidemiology. *Microbes Infect 2:* 1041–1050.
4. Gubler DJ, Meltzer M, 1999. Impact of dengue/dengue hemorragic fever on the developing world. *Adv Virus Res 53:* 35–70.
5. Moreau JP, Rosen L, Saugrain J, Laigret J, 1973. An epidemic of dengue on Tahiti associated with haemorrhagic manifestations. *Am J Trop Med Hyg 22:* 237–241.
6. Chungue E, Laudon F, Glaziou P, 1993. Dengue and dengue haemorrhagic fever in French Polynesia. Current situation. *Trop Med (Mars): 35:* 209–214.
7. Deparis X, Murgue B, Roche C, Cassar O, Chungue E, 1998. Changing clinical and biological manifestations of dengue during the dengue-2 epidemic in French Polynesia in 1996/97–description and analysis in a prospective study. *Trop Med Int Health 3:* 859–865.
8. Farfan JA, Olson KE, Black IV, Gubler DJ, Beaty BJ, 1997. Rapid characterization of genetic diversity among twelve dengue-2 virus isolates by single-strand conformation polymorphism analysis. *Am J Trop Med Hyg 57:* 416–422.
9. Diaz FJ, Farfan-Ale JA, Olson KE, Loroño-Pino MA, Gubler DJ, Blair CD, Black WC IV, Beaty BJ, 2002. Genetic variation within the premembrane coding region of dengue viruses from the Yucatan peninsula of Mexico. *Am J Trop Med Hyg 67:* 93–101.

10. Rico-Hesse R, 1990. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology 174:* 337–343.
11. Chungue E, Cassar O, Drouet MT, Guzman MG, Laille M, Rosen L, Deubel V, 1995. Molecular epidemiology of dengue-1 and dengue-4 viruses. *J Gen Virol 76:* 1877–1884.
12. Goncalvez AP, Escalante AA, Pujol FH, Ludert JE, Tovar D, Salas RA, Liprandi F, 2002. Diversity and evolution of the envelope gene of dengue virus type 1. *Virology 303:* 110–119.
13. Anderson R, King AD, Innis BL, 1992. Correlation of the E protein binding with cell susceptibility to dengue 4 virus infection. *J Gen Virol 73:* 2155–2159.
14. Chen Y, Maguire T, Marks RM, 1996. Demonstration of binding of dengue virus envelope protein to target cells. *J Virol 70:* 8765–8772.
15. Chungue E, Burucoa C, Boutin JP, Philippon G, Laudon F, Plichart R, Barbazan P, Cardines R, Roux J, 1992. Dengue 1 epidemic in French Polynesia, 1988-1989: surveillance and clinical, epidemiological, virological and serological findings in 1752 documented clinical cases. *Trans R Soc Trop Med Hyg 86:* 193–197.
16. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG, 1994. The Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrice choice. *Nucleic Acids Res 22:* 4673–4680.
17. Saitou N, Nei M, 1987. The neighbor joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol 4:* 406–425.
18. Kumar S, Tamura K, Jakobsen IB, Nei M, 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics 17:* 1244–1245.
19. Ishak H, Takegami T, Kamimura K, Funada H, 2001. Comparative sequences of two type 1 dengue virus strains possessing different growth characteristics *in vitro*. *Microbiol Immunol 45:* 327–331.
20. Heinz FX, Roehrig JT, 1990. Flaviviruses. Regenmortel MHV, Neurath AR, eds. *Immunochemistry of Viruses.* Volume II. *The Basis for Serodiagnosis and Vaccines.* Amsterdam: Elsevier Science Publishers BV (Biomedical Division), 289–305.
21. Lewis JA, Chang GJ, Lanciotti RS, Kinney RM, Mayer LW, Trent DW, 1993. Phylogenetic relationships of dengue-2 viruses. *Virology 197:* 216–224.
22. Hillis DM, Bull JJ, 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst Biol 42:* 182–192.
23. Fujita N, Hotta S, Konishi E, Esaki H, Sumarmo, Sujudi, 1997. Dengue hemorrhagic fever in Jakarta, Indonesia in 1988: isolation of dengue virus from patient whole blood using cell cultures. *Am J Trop Med Hyg 56:* 318–321.
24. Tolou HJG, Durand JP, Mercier V, de Pina JJ, de Mico P, Billoir F, Charrel RN, de Lamballerie X, 2001. Evidence for recombination in natural populations of dengue virus type-1 based on the analysis of complete genome sequences. *J Gen Virol 82:* 1283–1290.
25. Rico-Hesse R, Harrison LM, Salas R, Tovar D, Nisalak A, Ramons C, Boshell JR, de Mesa MT, Nogueira RMR, Travassos da Rosa A, 1997. Origins of dengue type 2 viruses associated with increased pathogenicity in the Americas. *Virology 230:* 244–251.
26. Watts DM, Porter KR, Putvatana P, Vasquez B, Calampa C, Hayes CG, Halstead SB, 1999. Failure of secondary infection with American genotype dengue 2 to cause dengue haemorrhagic fever. *Lancet 354:* 1431–1434.
27. Sasaki M, Tammy T, 2002. *Dengue Fever Update. Communicable Disease Report.* Honolulu, HI: Department of Health. Communicable Disease Division.
28. Leitmeyer KC, Vaughn DW, Watts DM, Salas R, Chacon IV, Ramos C, Rico-Hesse R, 1999. Dengue virus structural differences that correlate with pathogenesis. *J Virol 73:* 4738–4747.
29. Pandey BD, Igarashi A, 2000. Severity-related molecular differences among nineteen strains of dengue type 2 viruses. *Microbiol Immunol 44:* 179–188.
30. Aviles G, Rowe J, Meissner J, Manzur Caffarena JC, Enria D, St Jeor S, 2002. Phylogenetic relationships of dengue-1 viruses from Argentina and Paraguay. *Arch Virol 147:* 2075–2087.
31. Duarte dos Santos CN, Rocha CFS, Cordeiro M, Fragoso SP, Rey F, Deubel V, Desprès P, 2002. Genome analysis of dengue type –1 virus isolated between 1990 and 2001 in Brazil reveals a remarkable conservation of the structural proteins but amino acid differences in the non-structural proteins. *Virus Res 90:* 197–205.
32. Puri B, Nelson W, Porter KR, Henchal EA, Hayes CG, 1998. Complete nucleotide sequence analysis of a Western Pacific dengue type-1 virus strain. *Virus Genes 17:* 85–88.
33. Chu MC, O'Rourke EJ, Trent DW, 1989. Genetic relatedness among structural protein genes of dengue 1 virus strains. *J Gen Virol 70:* 1701–1712.
34. Fu J, Tan BH, Yap EH, Chan YC, Tan YH, 1992. Full-length cDNA sequence of dengue type 1 virus (Singapore strain S275/90). *Virology 188:* 953–958.
35. Shiu SY, Jiang WR, Porterfield JS, Gould EA, 1992. Envelope protein sequences of dengue virus isolates TH-36 and TH-Sman, and identification of a type specific genetic marker for dengue and tick-borne flaviviruses. *J Gen Virol 73:* 207–212.