Virology **303,** 181–191 (2002)
doi:10.1006/viro.2002.1635

Sequence Diversity of the Capsid Gene and the Nonstructural Gene NS2B of Dengue-3 Virus *in Vivo*

Wei-Kung Wang,*[,1] Tzu-Ling Sung,* Chun-Nan Lee,† Tsai-Yu Lin,* and Chwan-Chuen King‡

*Institute of Microbiology, †Graduate Institute of Medical Technology, College of Medicine, and ‡Institute of Epidemiology, College of Public Health, National Taiwan University, Taipei, Taiwan

Received April 23, 2002; returned to author for revision May 22, 2002; accepted June 21, 2002

Previously, we studied the envelope (E) gene of dengue virus and reported that dengue-3 virus is present as a quasispecies. To investigate the extent of intrahost sequence variation of other dengue viral genes, we examined in this study the capsid (C) gene and the nonstructural gene, NS2B, derived directly from plasma dengue viruses from 18 confirmed dengue-3 patients. Using reverse transcription-PCR, multiple clones of a 360-nucleotide region covering the C gene and of a 404-nucleotide region covering the NS2B gene from each patient were completely sequenced and analyzed. Our findings of the intrahost sequence variation of the C and the NS2B genes (mean pairwise p-distance: 0.12 to 1.02%, and 0.16 to 1.20%, respectively) demonstrate the quasispecies structure of dengue virus *in vivo*. A linear relationship was found between the extent of sequence variation of the C and NS2B proteins, suggesting that intrahost sequence variation of dengue-3 virus is likely to reflect genetic drift. The extent of intrahost sequence variation observed is in the same range as that of acute human immunodeficiency virus or hepatitis C virus infection, indicating that the random mutation frequency of dengue virus is similar to that of other RNA viruses *in vivo*. Consistent with a previous report of the E gene, the observations of genome-defective clones in both the C and the NS2B genes (3.9 and 5.0% of the clones, respectively) suggest a higher frequency of defective viruses *in vivo*. These findings would add to our understanding of the evolution of dengue-3 virus.

Key Words: dengue virus; capsid; NS2B; sequence diversity; quasispecies; defective virus.

INTRODUCTION

Among the 80 or so arthropod-borne flaviviruses, epidemics of the four serotypes of dengue virus (DEN-1, DEN-2, DEN-3, and DEN-4) continue to be a major public health problem in tropical and subtropical areas (Gubler, 1998; Innis, 1995; Monath, 1994). It has been estimated that approximately 100 million dengue infections occur annually throughout the world (Halstead, 1988; Monath, 1994; Gubler, 1998). Infection by any of the four dengue viruses results in a spectrum of clinical features ranging from asymptomatic or a mild, self-limited illness, dengue fever (DF), to severe and potentially life-threatening disease, dengue hemorrhage fever/dengue shock syndrome (DHF/DSS) (Gubler, 1998; Innis, 1995; WHO, 1997).

Dengue virus belongs to the genus flavivirus of the family *Flaviviridae*. It contains a positive-sense single-stranded RNA genome of approximately 11.7 kb in length. Flanked by the 5′ and 3′ nontranslated regions, the single open reading frame in the dengue genome encodes a polyprotein precursor which is subsequently cleaved into three structural proteins, the capsid (C), precursor membrane (PrM)/membrane (M), and envelope (E), as well as seven nonstructural proteins, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 (Gubler, 1998; Lindenbach and Rice, 2001).

The C gene is located at the 5′ part of the genome. The nascent polyprotein is cleaved at the C-PrM junction by both cellular signalase and viral protease. Cleavage by the signalase at ER results in a partially processed C (the anchC), which contains the 14 amino acid residues of PrM signal peptide. Cleavage at the characteristic dibasic site by NS2B/NS3 protease generates the mature C (virion C) protein (Lindenbach and Rice, 2001). The mature C protein, about 11 kDa in molecular weight, contains several basic residues at both N- and C-termini and is involved in the packaging of genomic RNA. An internal hydrophobic domain of approximately 22 resides was recently reported to mediate membrane integration and play a role in virion assembly (Markoff *et al.,* 1997; Lindenbach and Rice, 2001). The NS2B gene is located at the middle part of the genome. It encodes the NS2B protein of approximately 27 kDa, which contains a highly charged central domain flanked by two hydrophobic regions. The central hydrophilic domain has been shown to be required for the NS2B/NS3 protease activity probably through interaction with the NS3 protein (Falgout *et al.,* 1993; Droll *et al.,* 2000; Lindenbach and Rice, 2001).

Partly due to the nonproofreading nature of viral RNA polymerase or reverse transcriptase, many RNA viruses

[1] To whom correspondence and reprint requests should be addressed at Institute of Microbiology, College of Medicine, National Taiwan University, No.1 Sec. 1 Jen-Ai Road, Taipei, Taiwan. Fax: 886-2-2391-5293. E-mail: wwang60@yahoo.com.

0042-6822/02 $35.00

display prominent genomic heterogeneity among isolates, not only from different individuals but also within the same individual (Martell *et al.,* 1992; Zhu *et al.,* 1993; Wolinsky *et al.,* 1996). RNA virus is therefore a population made of a complex and dynamic swarm of mutants that have genetically different but related genomes known as quasispecies (Steinhauer and Holland, 1987; Domingo *et al.,* 1988; Holland *et al.,* 1992). Quasispecies are believed to be important for the survival and evolution of RNA viruses, and the pathogenesis of disease as well. Two well-studied examples are the human immunodeficiency virus type 1 (HIV-1) and the hepatitis C virus (HCV) (Zhu *et al.,* 1993; Wolinsky *et al.,* 1996; Farci *et al.,* 2000).

Previously, we studied the E gene, one of the structural genes, of plasma DEN-3 viruses from six patients and reported that DEN-3 virus is a population of closely related genomes: quasispecies *in vivo* (Wang *et al.,* 2002). However, the extent of sequence variation of other dengue viral genes within infected individuals remains unknown. In this study, we investigated intrahost sequence diversity of the C gene, another structural gene, and the NS2B gene, a nonstructural gene, derived from plasma dengue virus through reverse-transcription polymerase chain reaction (RT-PCR) and cloning sequencing. The 18 study participants were dengue patients from a DEN-3 outbreak in southern Taiwan in 1998 (King *et al.,* 2000). Since they had different disease severity (10 DF and 8 DHF), the relationship between the extent of sequence variation and disease severity was also examined. We report here that intrahost sequence diversity of the C gene and the NS2B gene is in agreement with the quasispecies structure of dengue virus *in vivo*. The observations of genome-defective clones in three different dengue viral genes (C, NS2B, and E reported previously) suggest a higher frequency of defective viruses in *vivo* and raise the possibility that defective virus may influence the evolution of dengue virus. The extent of intrahost sequence variation observed is in the same range as that of acute HIV-1 or HCV infection, indicating that the random mutation frequency of DEN-3 virus is similar to that of other RNA viruses. These findings would provide new insights into our understanding of the evolution of DEN-3 virus.

RESULTS

Study participants

The basic demographic and clinical information of the 18 patients, including 10 DF and 8 DHF patients, were summarized in Table 1. There were no significant differences in gender, age, duration of fever, or sampling day between the DF and DHF groups ($P = 0.52$, Fisher's exact test, $P = 0.27$, $P = 0.17$, and $P = 0.46$, Mann–Whitney test, respectively).

TABLE 1

Demographic and Clinical Information of Study Participants

Patient[a]	Disease[b]	Age (years)	Gender	Fever duration (days)	Sampling day[c]
ID3	DF	24	M	2	d2
ID4	DF	36	M	5	d4
ID5	DF	54	M	2	d2
ID6	DF	51	F	4	d4
ID7	DF	68	M	4	d3
ID8	DF	36	F	9	d7
ID9	DF	49	F	5	d4
ID15	DF	36	F	3	d8
ID16	DF	26	M	4	d5
ID17	DF	28	M	6	d7
ID18	DHF	57	M	3	d3
ID19	DHF	67	M	6	d3
ID20	DHF	57	F	7	d4
ID21	DHF	23	F	5	d4
ID22	DHF	67	F	5	d8
ID23	DHF	27	M	7	d8
ID24	DHF	63	M	4	d5
ID25	DHF	38	F	6	d8

[a] ID, identification.

[b] DF, dengue fever; and DHF, dengue hemorrhagic fever, according to WHO definition (WHO, 1997).

[c] Sampling day 1 (d1) is the first day of fever.

Nucleotide sequence diversity of the C gene

Ten clones containing the PCR products of the C gene derived from plasma sample from each of the 18 patients were completely sequenced, aligned, and analyzed in the 318-bp region (excluding the sequences of primers). The results were summarized in Table 2. There were more nonsilent substitutions than silent substitutions for most samples analyzed. To examine the extent of sequence variation, we first determined the mean diversity, which is the number of substitutions divided by total nucleotides sequenced (Zhu *et al.,* 1993; Wang *et al.,* 2002). It ranges from 0.06 to 0.57%, indicating that the extent of sequence diversity varies among different patients. Within each patient, some clones had more substitutions than others. We thus employed another method, pairwise comparison of each nucleotide sequence from the same sample, to assess the extent of sequence variation. The mean pairwise p-distance of the 18 patients ranges from 0.12 to 1.02% (Table 2). The extent of sequence variation determined by the mean diversity correlated well with that determined by the mean pairwise p-distance (simple linear regression, coefficient of correlation, $r = 0.99$), suggesting that a similar extent of sequence variation was seen in most clones analyzed. There is no correlation between the mean pairwise p-distance and gender, age, duration of fever, or sampling day ($P = 0.90$, Mann–Whitney test, $P = 0.67$, $P = 0.41$, $P = 0.29$, simple linear regression, respectively). The mean pairwise p-distance of the DF group (0.12 to 0.85%)

TABLE 2

Sequence Diversity of the Capsid Gene in Dengue Patients

Group/ patient[a]	No. of clones	Nucleotide sequence					Amino acid sequence		
		Substitutions[b]		Mean[c] diversity	p-distance[d]		Mean[e] diversity	p-distance[f]	
		Nonsilent	Silent		Mean	(Range)		Mean	(Range)
DF									
ID3	10	8	2	0.31%	0.61%	(0–1.57%)	0.75%	1.45%	(0–3.77%)
ID4	10	7	2	0.28%	0.56%	(0–1.26%)	0.66%	1.32%	(0–2.83%)
ID5	10	9	5	0.44%	0.85%	(0–2.52%)	0.85%	1.66%	(0–5.66%)
ID6	10	1	1	0.06%	0.12%	(0–0.63%)	0.09%	0.19%	(0–0.94%)
ID7	10	6	0	0.19%	0.36%	(0–1.26%)	0.57%	1.09%	(0–3.77%)
ID8	10	6	0	0.19%	0.38%	(0–0.94%)	0.57%	1.13%	(0–2.83%)
ID9	10	7	1	0.25%	0.49%	(0–1.26%)	0.66%	1.28%	(0–3.77%)
ID15	10[g,i]	10	3	0.41%	0.71%	(0–1.89%)	0.94%	1.54%	(0–2.86%)
ID16	10[g]	5	3	0.25%	0.44%	(0–1.27%)	0.47%	0.76%	(0–2.86%)
ID17	10	3	0	0.09%	0.19%	(0–0.63%)	0.28%	0.56%	(0–1.89%)
DHF									
ID18	10	4	2	0.19%	0.38%	(0–0.63%)	0.38%	0.75%	(0–1.89%)
ID19	10[g]	8	2	0.31%	0.57%	(0–1.59%)	0.75%	1.32%	(0–4.76%)
ID20	10	2	2	0.13%	0.25%	(0–1.26%)	0.19%	0.38%	(0–1.89%)
ID21	10	3	1	0.13%	0.25%	(0–0.63%)	0.28%	0.56%	(0–1.89%)
ID22	10	13	3	0.50%	0.92%	(0–2.52%)	1.23%	2.20%	(0–5.66%)
ID23	10	6	2	0.25%	0.50%	(0–1.57%)	0.57%	1.13%	(0–4.72%)
ID24	10	4	3	0.22%	0.44%	(0–1.26%)	0.38%	0.75%	(0–1.89%)
ID25	10[g,h]	13	5	0.57%	1.02%	(0–1.59%)	1.23%	2.14%	(0–3.77%)

[a] Disease group includes dengue fever (DF) and dengue hemorrhagic fever (DHF). ID: identification.
[b] The nucleotide substitutions consist of nonsilent and silent substitutions.
[c,e] The mean diversity is the number of substitutions divided by the total number of nucleotides or amino acids sequenced.
[d,f] p-distance is calculated by pairwise comparison of nucleotide or amino acid sequences between clones by the program MEGA.
[g] One clone contains a stop codon.
[h] One clone contains a single-nucleotide deletion.
[i] Two clones contain single-nucleotide deletions (one each).

was not significantly different from that of the DHF group (0.25 to 1.02%) ($P = 0.70$, Mann–Whitney test) (Fig. 1A).

Deduced amino acid sequences of the C protein

To investigate the extent of sequence variation at the protein level, the deduced amino acid sequences of 10 clones from each patient were aligned and analyzed. As summarized in Table 2, the mean diversity of amino acid ranges from 0.09 to 1.23%, and the mean pairwise p-distance ranges from 0.19 to 2.20%. Overall, the results are in agreement with those at the nucleotide level in that there is a good correlation between the mean p-distance of amino acid and that of nucleotide (simple linear regression, $r = 0.96$).

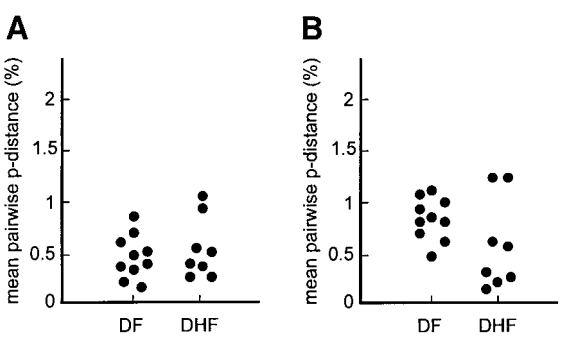

FIG. 1. The relationship between disease severity and intrahost sequence diversity of the C gene (A) and the NS2B gene (B). The extent of sequence variation is shown by the mean pairwise p-distance of nucleotide. DF, dengue fever; DHF, dengue hemorrhagic fever.

The alignments of the deduced amino acid sequences of 10 clones from each of the three patients, ID7, ID8, and ID25, are shown in Fig. 2. Within the 106 amino acid region analyzed, there were six amino acid substitutions for ID7 (Fig. 2A). The amino acid substitutions included both conservative (such as phenylalanine to leucine) and drastic changes (such as lysine to isoleucine) (Fig. 2A). For ID8 and ID25, there were 6 and 12 amino acid substitutions, respectively (Figs. 2B and 2C). In agreement with our previous observation of genome-defective clones in the E gene, defective clones were found in the C gene of ID25. One in-frame stop codon was found at amino acid residue 55 of clone 2B, and a single nucleotide deletion was noted at the first base of amino acid

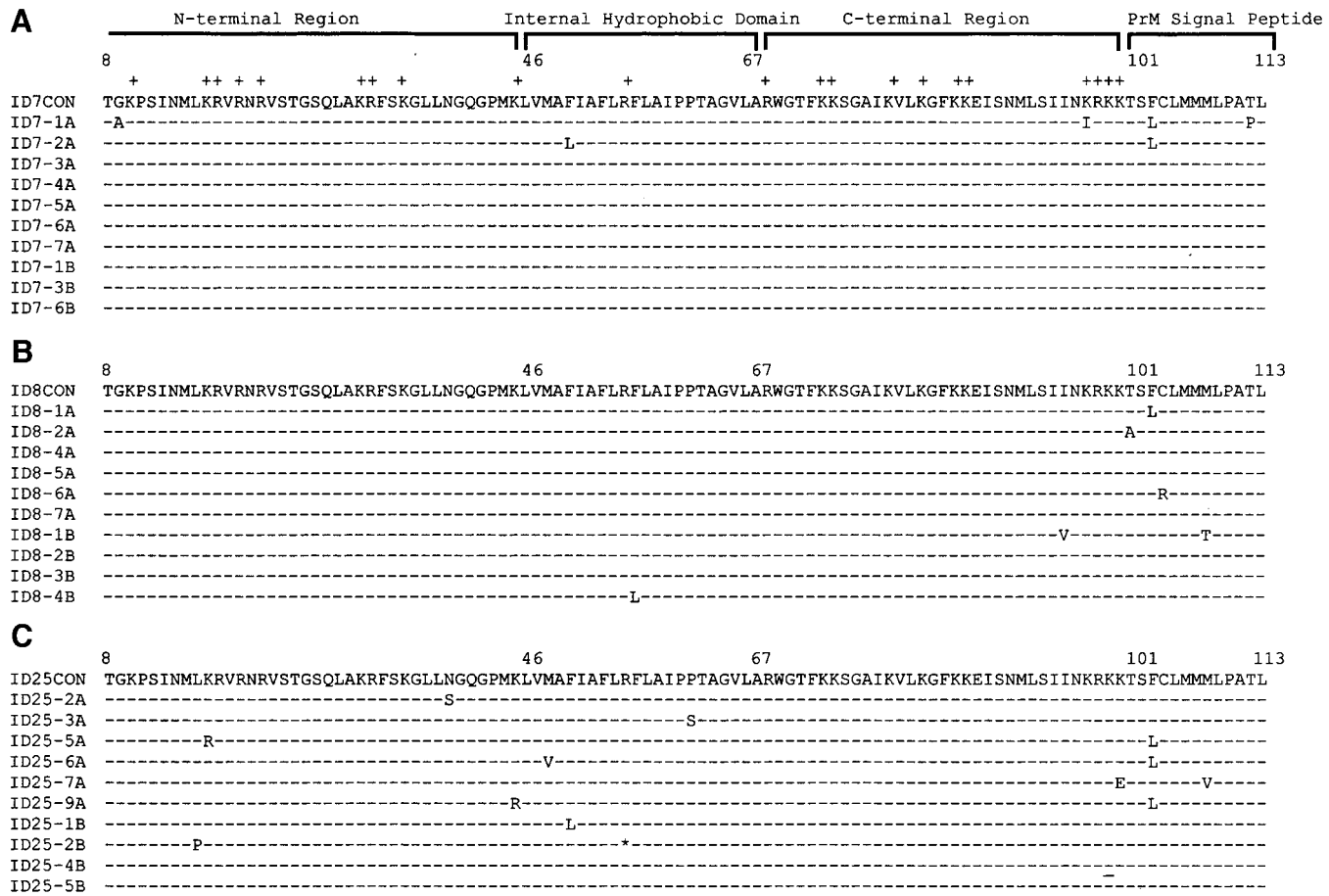

FIG. 2. Alignments of deduced amino acid sequences of the mature C protein and PrM signal peptide of multiple clones from three patients, ID7 (A), ID8 (B), and ID25 (C). The positions of amino acid residues and the corresponding regions are shown on top. Plus signs indicate positively charged residues. A consensus sequence (CON) was generated for each sample. Dashes indicate sequence identity; asterisk indicates in-frame stop codon, and underline indicates deletion at that position. Individual clone number is shown at the left following patient ID, with the letters A and B indicating two separate PCRs.

residue 99 of clone 4B, which resulted in a frameshift and a premature stop several residues downstream (Fig. 2C). The frequency of defective virus containing either stop codons or deletions was summarized by each patient in Table 2. Overall, there were seven defective clones among the total of 180 clones of the C gene sequenced, corresponding to a frequency of defective virus of 3.9%. There is no difference in the frequency of defective virus between the DF and DHF groups (4.0% vs 3.8%, $P = 0.76$, proportion test).

Distribution of amino acid variation in the C protein

The 106 amino acid region examined in this study covers the N-terminal region, the internal hydrophobic domain, and the C-terminal region of the mature C protein, and the PrM signal peptide (Fig. 2) (Markoff *et al.,* 1997; Lindenbach and Rice, 2001). To investigate the distribution of amino acid substitutions in different regions, the frequency of amino acid substitutions in each region, corrected for the length of the region, was calculated for each clone and analyzed. The overall distribution of amino acid substitutions in different regions for total 180 clones examined was summarized in Table 3. As exemplified by the three patients in Fig. 2, there were more amino acid substitutions in the PrM signal peptide than in each of the three C regions ($P < 0.001$, $P = 0.002$, and $P < 0.001$, respectively, chi-square test). Within the mature C protein, there was no difference in the frequency of amino acid substitutions between the internal hydrophobic domain and the N-terminal region or the C-terminal region (Table 3) ($P = 0.73$ and $P = 0.96$, respectively, chi-square test).

Among the 180 clones examined, there were 29 amino acid substitutions in the PrM signal peptide. Signal peptides are known to have a common structural motif (von Heijne, 1990; Lee *et al.,* 2000). Analysis of the predicted secondary structure of the PrM signal peptide revealed that the majority of amino acid substitutions in this region did not change its secondary structure drastically (Fig. 3B). The internal hydrophobic domain, which has been

TABLE 3

Summary of the Distribution of Amino Acid Substitutions in Different Regions of the Capsid Protein

Region[a]	N-terminal region	Internal hydrophobic domain	C-terminal region	PrM signal peptide
Length[b]	38	22	33	13
No. of amino acid substitutions/total no. of amino acids[c]	33/6840	21/3960	32/5940	29/2340
Relative frequency of substitutions in each region[c]	0.48%	0.53%	0.54%	1.24%

[a] The mature capsid protein contains the internal hydrophobic domain, the N-terminal, and C-terminal regions.

[b] The length is the number of amino acid residues in each region.

[c] Total number of amino acid substitutions of the 180 clones studied in each region is divided by the total number of amino acids sequenced in each region to determine the relative frequency of substitutions in each region.

shown to be involved in membrane integration, is structurally conserved among flaviviruses (Markoff *et al.,* 1997). In the case of DEN-3 virus, it is composed of a segment of nine hydrophobic residues, an arginine residue in the middle (residue 55), and a segment that is less hydrophobic, and contains two proline and one glycine residues (Fig. 3A). Of the 15 amino acid substitutions involving the hydrophobic residues in this domain, 13 were conservative changes including 9 substitutions to hydrophobic residues and 4 to threonine residues (Fig. 3A).

Sequence diversity of the NS2B gene

To further investigate the extent of sequence variation in the nonstructural genes, 10 clones containing the PCR products of the NS2B gene, a nonstructural gene, derived from each patient were completely sequenced, aligned, and analyzed in the 366-bp region. The results are summarized in Table 4. The mean diversity and the mean pairwise p-distance of nucleotide range from 0.08 to 0.60% and from 0.16 to 1.20%, respectively (Table 4). As with the analysis in the C gene, the mean diversity

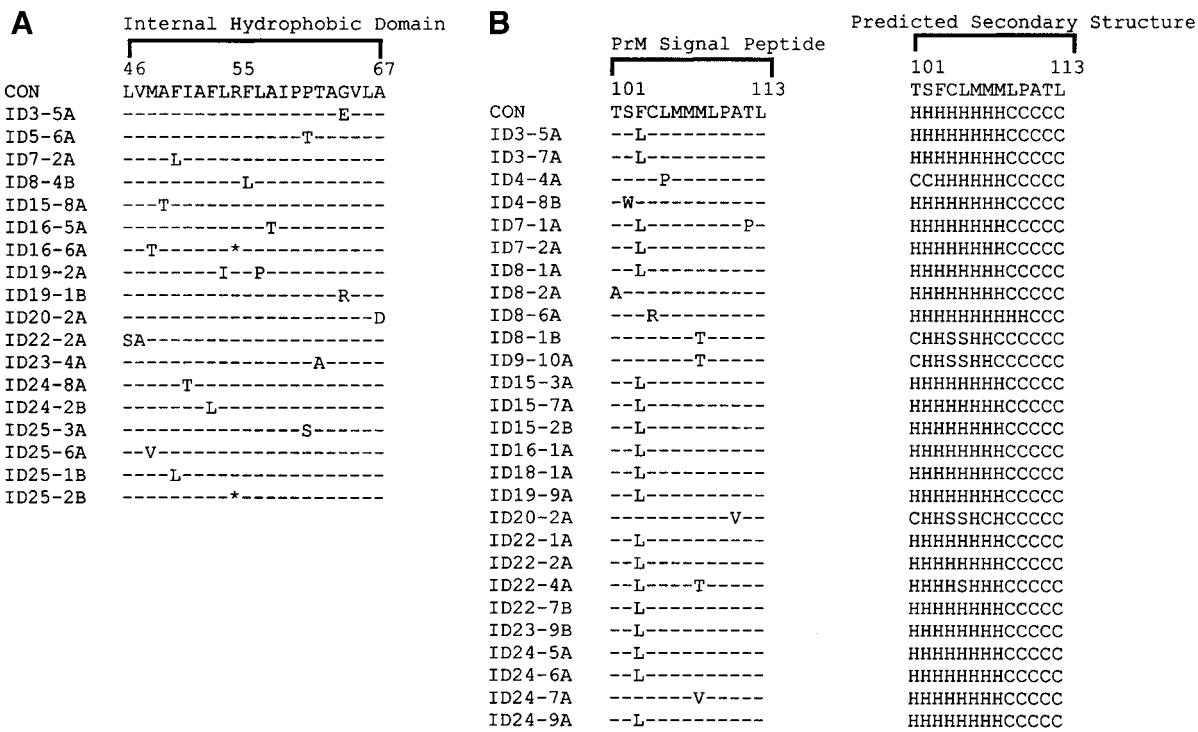

FIG. 3. Amino acid sequence variation in the internal hydrophobic domain of the mature C protein (A) and the PrM signal peptide (B). The positions of amino acid residues are shown on top. Based on the 180 clones examined, consensus sequences (CON) were generated for the internal hydrophobic domain and the PrM signal peptide, respectively. Sequences of each variant clone are shown. Dashes indicate sequence identity, and asterisks indicate in-frame stop codons. Individual clone number is shown at the left following patient ID, with the letters A and B indicating two separate PCRs. For the PrM signal peptide, the predicted secondary structure of each sequence is shown at the right. C, coils; H, helix; S, strand.

TABLE 4

Sequence Diversity of the NS2B Gene in Dengue Patients

Group/ patient[a]	No. of clones	Nucleotide sequence					Amino acid sequence		
		Substitutions[b]		Mean[c] diversity	p-distance[d]		Mean[e] diversity	p-distance[f]	
		Nonsilent	Silent		Mean	(Range)		Mean	(Range)
DF									
ID3	10[g,i]	16	4	0.55%	1.02%	(0–2.19%)	1.31%	2.41%	(0–5.74%)
ID4	10	11	3	0.38%	0.77%	(0.27–1.64%)	0.90%	1.80%	(0.82–4.10%)
ID5	10	9	3	0.33%	0.66%	(0–1.64%)	0.74%	1.48%	(0–4.10%)
ID6	10[h]	9	2	0.30%	0.49%	(0–1.64%)	0.74%	1.15%	(0–4.10%)
ID7	10[i]	6	5	0.30%	0.60%	(0–1.09%)	0.49%	0.99%	(0–2.46%)
ID8	10[g]	11	6	0.46%	0.86%	(0–1.91%)	0.90%	1.59%	(0–4.10%)
ID9	10	13	5	0.49%	0.98%	(0–2.73%)	1.07%	2.11%	(0–6.56%)
ID15	10[j]	9	5	0.38%	0.77%	(0–1.64%)	0.74%	1.48%	(0–3.28%)
ID16	10[j]	16	4	0.55%	1.07%	(0.27–2.20%)	1.31%	2.41%	(0–5.79%)
ID17	10	11	4	0.41%	0.81%	(0–1.91%)	0.90%	1.79%	(0–4.10%)
DHF									
ID18	10	6	0	0.16%	0.33%	(0–0.82%)	0.49%	0.98%	(0–2.46%)
ID19	10	1	2	0.08%	0.16%	(0–0.55%)	0.08%	0.16%	(0–0.82%)
ID20	10	3	2	0.14%	0.27%	(0–0.82%)	0.25%	0.49%	(0–2.46%)
ID21	10	4	0	0.11%	0.22%	(0–0.55%)	0.33%	0.66%	(0–1.64%)
ID22	10	18	4	0.60%	1.20%	(0–2.73%)	1.48%	2.95%	(0–6.56%)
ID23	10	8	2	0.27%	0.55%	(0–1.37%)	0.66%	1.48%	(0–3.28%)
ID24	10[j]	5	6	0.30%	0.60%	(0–1.37%)	0.41%	0.82%	(0–3.28%)
ID25	10	16	6	0.60%	1.20%	(0–2.46%)	1.31%	2.62%	(0–6.56%)

[a] Disease group includes dengue fever (DF) and dengue hemorrhagic fever (DHF). ID, identification.
[b] The nucleotide substitutions consist of nonsilent and silent substitutions.
[c,e] The mean diversity is the number of substitutions divided by the total number of nucleotides or amino acids sequenced.
[d,f] p-distance is calculated by pairwise comparison of nucleotide or amino acid sequences between clones by the program MEGA.
[g] One clone contains a stop codon.
[h] Two clones contain stop codons (one each).
[i] One clone contains a single-nucleotide deletion.
[j] One clone contains a three-nucleotide deletion.

correlates well with the mean pairwise p-distance in the NS2B gene (simple linear regression, $r = 0.99$). The mean pairwise p-distance does not correlate with gender, age, duration of fever, or sampling day ($P = 0.57$, Mann–Whitney test, $P = 0.29$, $P = 0.80$, $P = 0.07$, simple linear regression, respectively). There is no difference in the mean pairwise p-distance between the DF group (0.49 to 1.07%) and the DHF group (0.16 to 1.20%) ($P = 0.12$, Mann–Whitney test) (Fig. 1B).

The deduced amino acid sequences of 10 clones from each patient were aligned and analyzed. As summarized in Table 4, the mean diversity of amino acids ranges from 0.08 to 1.48%, and the mean pairwise p-distance ranges from 0.16 to 2.95%. The mean p-distance of amino acid correlates with that of nucleotide (simple linear regression, $r = 0.96$). Overall, the extent of intrahost sequence variation of the NS2B was in the same range as that of the C at the levels of both nucleotide and amino acid (Fig. 1, Tables 2 and 4). Genome-defective clones containing either stop codons or deletions were also found in the NS2B gene (Table 4). There were 9 defective clones among the total of 180 clones sequenced, corresponding to a frequency of defective virus of 5.0%. There is no difference in the frequency of defective virus between the DF and DHF groups ($P = 0.09$, proportion test).

Relationship between intrahost sequence diversity of the C and the NS2B genes

Since intrahost sequence diversity of the C protein and of the NS2B protein varies among different patients, we next examined the relationship between the extent of sequence variation of the C protein and that of the NS2B protein. As shown in Fig. 4, there was a trend of increase in the mean pairwise p-distance of amino acid of the C protein as that of the NS2B protein increase (simple linear regression, $r = 0.60$, $P = 0.009$) (Fig. 4). Similarly, a linear relationship was also observed when comparing the mean pairwise p-distance of nucleotides between these two genes (simple linear regression, $P = 0.017$).

DISCUSSION

To our knowledge, this is the first study examining the extent of sequence variation of two dengue viral genes

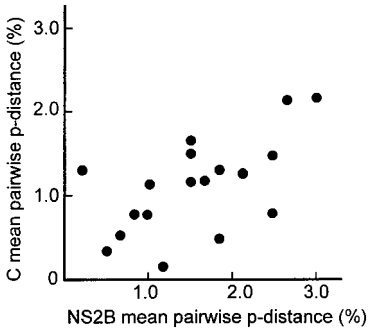

FIG. 4. The relationship between intrahost sequence diversity of the C protein and that of the NS2B protein. The extent of sequence variation is shown by the mean pairwise p-distance of amino acid.

within infected individuals. The approach of RT-PCR of dengue viral RNA derived directly from plasma avoids the potential selection due to *in vitro* passage of virus. Information derived from this type of analysis would have implications to our understanding of the quasispecies of dengue virus *in vivo* and disease pathogenesis, the structure/function of the dengue viral proteins examined, as well as evolution of dengue virus.

The quasispecies structure of dengue virus *in vivo* and disease severity

We used the approach of RT-PCR and molecular cloning, which has been shown to be a simple and valuable method for characterization of mutant spectra of virus quasispecies (Arias *et al.,* 2000). In agreement with our previous report of the E gene, analysis of the C and the NS2B genes in this study demonstrates the quasispecies structure of DEN-3 virus *in vivo*. The sequence variation observed in our study is unlikely to be due to *in vitro* artifacts, though it cannot be excluded that a minority of the mutations might have been introduced by RT or *Taq* polymerase. This is because the error frequency of the *Taq* polymerase after 60 cycles of PCR amplification (2.3 to 5.5×10^{-4}) (Meyerhans *et al.,* 1989; Martell *et al.,* 1992; Smith *et al.,* 1997), plus that of the RT are lower than the mean diversity of the C gene (0.06 to 0.57%) and that of the NS2B gene (0.08 to 0.60%) observed in our samples, as has been addressed previously (Wang *et al.,* 2002). We also did a control experiment in which a DEN-3 C clone with known sequence was serially diluted and subjected to *in vitro* transcription, RT, and PCR amplification under identical conditions. The product derived from the most diluted template concentration was cloned and sequenced. Among the eight clones sequenced, there was one substitution of 2544 bases sequenced, corresponding to an error frequency of 0.04%. This was lower than the mean diversity observed in our sample.

There is no difference in the extent of sequence variation of the C or the NS2B gene between DF and DHF patients (Fig. 1). Comparing the deduced amino acid sequences in the C and the NS2B proteins between the DF and DHF groups, there is no particular amino acid substitution associated with DHF (data not shown). This is in agreement with two recent studies of full-length genome of DEN-2 virus, in which no specific amino acid change consistently correlated with disease outcome was found (Mangada and Igarashi, 1998; Leitmeyer *et al.,* 1999) and this suggests that the C protein or the NS2B protein does not contain signature sequence for DHF. When compared with the DEN-3 sequences available in GenBank, the consensus nucleotide sequence of the NS2B gene of our samples revealed a homology of 95.6% to that of the prototype DEN-3 strain, H87, which was originated from Philippines in 1956 (Osatomi and Sumiyoshi, 1990). Similarly, the consensus nucleotide sequence of the C gene had a homology of 96.5% to that of the H87 strain. Interestingly, a higher degree of homology, 98.8%, was found between the C gene of our samples and that of another DEN-3 strain, CH53489, which was isolated in Bangkok in 1973. This finding suggests that the DEN-3 virus of the 1998 outbreak in Taiwan is closer to the DEN-3 virus from Thailand.

Sequence variation of the C and the NS2B proteins

Comparing the relative frequency of amino acid substitutions in the 106 amino acid region examined, more amino acid substitutions were found in the PrM signal peptide than in each of the three regions of the mature C protein (Table 3). The majority of amino acid substitutions in the PrM signal peptide did not drastically change the predicted secondary structure (Fig. 3B). These findings are consistent with the characteristics of signal peptides, namely, high variability with a common structural motif (von Heijne, 1990, Lee *et al.,* 2000). Within the mature C protein, the internal hydrophobic domain was not more conserved than the N-terminal or the C-terminal region (Table 3). Nonetheless, most of the amino acid substitutions involving the hydrophobic residues of the internal hydrophobic domain were conservative changes (Fig. 3A), indicating that these hydrophobic residues are important.

There are 9 and 11 positively charged residues at the N-terminal and C-terminal region of the C protein, respectively (Fig. 2). Studies of the eukaryotic membrane protein indicated that positively charged residues amino-terminal to a transmembrane domain constitute a cytoplasmic retention signal (Hartmann *et al.,* 1989; Parks and Lamb, 1991). Mutational study of the C protein of DEN-4 virus has shown recently that the net positive charge in the N-terminal region determines the membrane orientation of the C protein (Markoff *et al.,* 1997). A closer examination of the amino acid substitutions revealed that 5 of the 8 substitutions involving the positively charged residues in the N-terminal region were conservative change (to positively charged residues),

whereas only 3 of the 13 substitutions of such in the C-terminal region were conservative (data not shown). These findings suggest that the positively charged residues at the N-terminal region may have important roles.

The 122 amino acid region of the NS2B protein examined in this study contains a central hydrophilic domain flanked by the N-terminal and the C-terminal hydrophobic regions. The central hydrophilic domain has been shown to be required for the NS2B/NS3 protease activity (Falgout *et al.,* 1993; Droll *et al.,* 2000; Lindenbach and Rice, 2001). The frequency of amino acid substitutions in each region, corrected for the length of the region, was also examined for each clone and analyzed (data not shown). There was no difference in the frequency of amino acid substitution between the central hydrophilic domain and the N-terminal or the C-terminal hydrophilic region ($P = 0.49$ and $P = 0.16$, respectively, chi-square test).

Comparison of intrahost sequence variation of different genes

The extent of sequence variation observed in this study was in the same range as what has been reported for acute infection of HIV-1 and HCV (Martell *et al.,* 1992; Zhu *et al.,* 1993; Farci *et al.,* 2000). This finding suggests that the random mutation rate of dengue virus is similar to that of other RNA viruses *in vivo* (Holland *et al.,* 1982; Lanciotti *et al.,* 1994). A study of sequence variation of HIV-1 after sexual transmission revealed that the nucleotide mean diversity of the E gene (gp120) was 0.24% and that of the gag gene (p17) was 0.50% (Zhu *et al.,* 1993). The observation that the mean diversity of gp120 was lower than that of p17 has been interpreted as a strong selection for specific gp120 at the time of transmission or from exposure to seroconversion (Zhu *et al.,* 1993; Zhang *et al.,* 1993). Our findings that the mean diversity of the C gene (0.06 to 0.57%) was similar to that of the E gene (0.10 to 0.84%) reported previously would suggest that there is no strong selection on the E gene during the acute stage of dengue virus infection (Wang *et al.,* 2002). Consistent with this interpretation was that the ratio of synonymous (dS) nucleotide substitutions per site to nonsynonymous (dN) nucleotide substitutions per site, dS/dN, was generally higher than 1 for both E and C genes, suggesting a lack of strong positive selection in most cases (data not shown).

Pairwise comparisons of each of the 11 proteins from different flaviviruses have shown that the relative phylogenetic distances of the C, NS2A, and NS2B were similar and higher than those of the NS1 and NS3 (Blok *et al.,* 1992). This suggests different mutation rates or different plasticity of various flaviviral genes during evolution. Our findings that intrahost sequence diversity of the C and NS2B proteins was in the same range is consistent with the relative phylogenetic distance of the C and NS2B proteins between different flaviviruses (Blok *et al.,* 1992). Recently, an analysis of the relative divergence of different viral proteins from isolates of the same RNA virus including HIV, HCV, enteroviruses, and flaviviruses revealed a linear relationship (Sala and Wain-Hobson, 2000). The smoothness of the protein sequence diversification over a wide variety of hosts or niches led to the conclusion that the majority of amino acid changes likely reflect genetic drift. Our findings of the linear relationship between the extent of sequence variation of the C protein and that of the NS2B protein (Fig. 4), together with the observation that the dS/dN ratio, was generally higher than 1, suggest that intrahost sequence variation of DEN-3 virus is likely to reflect genetic drift.

Genome-defective virus

Genome-defective clones containing either stop codons or deletions were found in 3.9 and 5.0% of the clones of the C gene and the NS2B gene, respectively. This is similar to the frequency of defective clones of the E gene, 5.8%, reported previously (Wang *et al.,* 2002). The 318-bp C gene and the 366-bp NS2B gene plus the 393-bp E gene examined previously correspond to 10.1% of the dengue virus genome. Although variation observed in a 10.1% fraction may not represent the entire genome, our findings of defective clones in three different viral genes would suggest that the frequency of defective virus might be higher *in vivo*. The frequency of defective clones observed in our study was lower that that of HIV-1, in which it was 10 to 15% for the tat gene and 15% for the gag and E genes (Meyerhans *et al.,* 1986; Goodenow *et al.,* 1989). Of note was that the ratio of genome copy number to infectious unit was reported to be 10^2 to10^5 for dengue virus, whereas it was 10^4 to 10^7 for HIV-1 (Piatak *et al.,* 1993; Wang *et al.,* 2000; Sudiro *et al.,* 2001).

There is a wealth of evidence indicating that defective virus may interfere with viral replication *in vitro*, modulate immune response and pathogenesis *in vivo*, or lead to establishment of persistent infection (Barrett and Dimmock, 1986; Domingo *et al.,* 1988; Huang, 1988). Defective interfering (DI) particles have been reported in flaviviruses *in vitro* (Lancaster *et al.,* 1998; Lindenbach and Rice, 2001). A recent study revealed that DI particles can affect the fitness of eastern equine encephalitis virus (EEEV), an arthropod-borne virus, and suggested that defective virus may influence the evolution of arbovirus (Weaver *et al.,* 1999). Genome replication interfered by defective viruses would result in a slow rate of genetic change. In addition, DI particles have been shown to associate with high-multiplicity passage *in vitro*, which was found to have slow rate of genetic change in EEEV (Weaver *et al.,* 1999).

Study of the structural proteins of different flaviviruses has revealed a high degree of conservation among mem-

bers of the same serological subgroup, including different serotypes of dengue viruses (Mandl *et al.,* 1988). In the case of DEN-3 viruses, the amino acid homology of the PrM/E proteins was found to be greater than 95% over a 36-year period (Lanciotti *et al.,* 1994). Such conservation has been attributed to structural or functional constraints imposed on dengue virus and other arthropod-borne viruses that replicate in both vertebrae and arthropod hosts (Mandl *et al.,* 1988; Weaver *et al.,* 1991; Lanciotti *et al.,* 1994). Our study of the intrahost sequence diversity of two DEN-3 viral genes suggests that the relatively low rates of evolution of DEN-3 virus is unlikely due to innate properties of the genome replication such as high fidelity of viral polymerase, since the extent of intrahost sequence variation observed was similar to that of two other RNA viruses (Zhu *et al.,* 1993; Lanciotti *et al.,* 1994; Farci *et al.,* 2001). Moreover, genome-defective clones were found in three different dengue viral genes in this study. Although it remains to be determined whether defective virus is also present in mosquito, another nature host in dengue virus life cycle, our findings of a significant proportion of defective clones in human host raise the possibility that defective virus may also contribute to the evolutionary conservation of DEN-3 virus.

MATERIALS AND METHODS

Study participants

The diagnoses of DF and DHF followed the WHO clinical definition (WHO, 1997). Detection of dengue genomic sequences in plasma by a previously described RT-PCR assay was the laboratory criteria of confirmation for all cases (Lanciotti *et al.,* 1992; Harris *et al.,* 1998). Eighteen confirmed dengue patients from three hospitals (Chi-Mei Foundation Medical Center, Kuo General Hospital, and Sin-Lau Christian Hospital) during an outbreak in southern Taiwan in 1998 were included in this study (King *et al.,* 2000). The day of onset of fever (oral temperature $\geq$ 38°C) is defined as Day 1 of illness (d1). Acute blood samples were collected in EDTA-containing tubes between Day 2 and Day 8 of illness. Plasma was prepared within 6 h of collection and stored at −80°C until use (Wang *et al.,* 2000). The patients were closely observed during hospitalization and monitored with routine laboratory tests. The serotype of the 18 patients was found to be DEN-3 using the RT-PCR assay, which can distinguish the four serotypes by the size of the products (Lanciotti *et al.,* 1992; Harris *et al.,* 1998).

Isolation of viral RNA, RT, and PCR

Dengue viral RNA was isolated directly from plasma using the QIAamp viral RNA mini kit (Qiagen, Germany) (Wang *et al.,* 2000). The RNA eluate was subjected to RT using the cDNA synthesis kit and the Superscript II RT

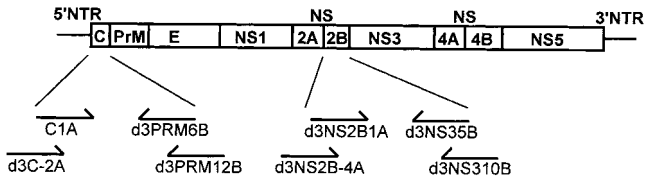

FIG. 5 Schematic diagram of the dengue virus genome and the C and NS2B genes examined in this study. The relative positions of primers used in PCR are shown. NTR, nontranslated region; C, capsid; PrM, precursor membrane; NS, nonstructural.

(Life Technologies, Rockville, MD) as described previously (Wang *et al.,* 2002). Based on the DEN-3 sequences available in GenBank, outer (d3C-2A, d3PRM12B) and inner (C1A, d3PRM6B) primers were designed to amplify a 360-nucleotide region covering the C gene (Fig. 5). The sequences of the primers are as follows: d3C-2A, 5′-TCTCTGATGAACAACCAACGG-3′ (corresponding to genome positions 89 to 109 of the DEN-3 H87 strain) (Osatomi and Sumiyoshi, 1990); d3PRM12B, 5′-GCGCG-GCTCTCCATCTCGTG-3′ (positions 469 to 450 of H87 strain); C1A, 5′-ATGAACAACCAACGGAAAAAG-3′ (positions 95 to 115 of H87 strain); d3PRM6B, 5′-TCGTGAAGT-TAAGTGGAAAGC-3′ (positions 454 to 434 of H87 strain). For the NS2B gene, outer (d3NS2B-4A, d3NS310B) and inner (d3NS2B1A, d3NS35B) primers were designed to amplify a 404-nucleotide region covering the NS2B gene (Fig. 5). The sequences of the primers are as follows: d3NS2B-4A, 5′-CTCAAAAGGAGAAGCTGGCC-3′ (corresponding to genome positions 4111 to 4130 of the H87 strain) (Osatomi and Sumiyoshi, 1990); d3NS310B, 5′-GGGGGCTGGGTACGTCCC-3′ (positions 4543 to 4526 of H87 strain); d3NS2B1A, 5′-AGCTGGCCACTGAATGAGG-3′ (positions 4123 to 4141 of H87 strain); d3NS35B, 5′-CATAGGACGCCGGATCTTTG-3′ (positions 4526 to 4507 of H87 strain).

An aliquot of cDNA was subjected to the first- and the second-round PCR, using the super *Taq* polymerase (HT Biotechnology, Cambridge, U.K.) plus the outer and inner primers, respectively. PCR was performed in a separate room from that used for RNA isolation, and precautions for PCR were followed to avoid contamination (Kwok and Higuchi, 1989). The PCR conditions for the C gene were 95°C for 5 min, followed by 30 cycles of 95°C 1 min, 58°C 1 min and 72°C 1 min, and then 72°C 5 min. The PCR conditions for the NS2B gene were 95°C for 5 min, followed by 30 cycles of 95°C 1 min, 60°C 1 min and 72°C 1 min, and then 72°C 5 min.

Cloning and sequencing

Each PCR product was cloned to the T/A cloning vector, pCRII-TOPO, which was transformed to TOP10 competent cells (Invitrogen, San Diego, CA). To explore the extent of sequence diversity and to avoid the bias due to preferential amplification of certain templates in a

single PCR, 10 clones derived from two separate PCR were picked up and completely sequenced, using the BigDye terminator cycle sequencing kit and the ABI 377 automated sequencer (Applied Biosystems, Foster City, CA).

Sequence analysis

Ten C clones and ten NS2B clones from each plasma sample were aligned in the 318- and 366-bp regions, respectively (excluding the sequences of primers), using the program Dnaman Version 4.15 (Lynnon Biosoft, Canada). Consensus sequences of the C gene and NS2B gene were thus generated. The mean diversity of nucleotide and amino acid is the number of substitutions divided by total nucleotides and total amino acids sequenced, respectively (Zhu *et al.,* 1993; Wang *et al.,* 2002). Pairwise comparisons of both nucleotide and amino acid sequences between clones were performed using the program MEGA version 1.02 (Molecular Evolutionary Genetics Analysis, Pennsylvania State University, PA) to determine the mean and range of proportion of difference (p-distance). The number of synonymous nucleotide substitutions per site (dS) and the number of nonsynonymous nucleotide substitutions per site (dN) for each sample were calculated using the program MEGA based on the method of Nei and Gobojori (1986). The ratio of dS to dN was also determined. The predicted secondary structure of the PrM signal peptide was analyzed using the program Dnaman Version 4.15 (Lynnon Biosoft).

Statistical test

A nonparametric statistical method, the Mann–Whitney test in the software SPSS base 10.0 (SPSS Inc., Chicago, IL), was used to compare age, fever duration, sampling time, and the extent of sequence variation (mean pairwise p-distance of nucleotide and amino acid) between groups. Fisher's exact test was used to compare the difference in gender between groups. Proportion test was used to determine the difference in the frequency of defective clones between groups (software SPSS base 10.0, SPSS Inc.). Regression analysis was performed to examine the correlation between the mean pairwise p-distance and the mean diversity, age, duration of fever, or sampling day (software SPSS base 10.0, SPSS Inc.). Chi-square test was performed to compare the numbers of amino acid substitutions (corrected by the length) in different regions of the C or NS2B protein among the 180 clones sequenced (software SPSS base 10.0, SPSS Inc.).

ACKNOWLEDGMENTS

We thank Dr. Chien-Ming Li at the Chi-Mei Foundation Medical Center, Dr. Shih-Chung Lin at the Kuo General Hospital, and Dr. Shih-Ting Ho at the Sin-Lau Christian Hospital for kindly providing clinical samples, and Day-Yu Chao, Yu-Chen Tsai, and Su-Ru Lin for technical assistance. The sequences have been submitted to the GenBank, and the Accession Nos. are AF495889–AF495906 and AF497721–AF497738. This work was supported by the National Health Research Institute (NHRI-CN-CL8903P), and in part by the National Science Council (NSC91-2320-B-002-191), Taiwan, Republic of China.

REFERENCES

Arias, A., Lazaro E., Escarmis C., and Domingo E. (2000). Molecular intermediates of fitness gain of an RNA virus: Characterization of a mutant spectrum by biological and molecular cloning. *J. Gen. Virol.* **82,** 1049–1060.

Barrett, A. D. T., and Dimmock, N. J. (1986). Defective interfering viruses and infections of animals. *Curr. Top. Microbiol. Immunol.* **128,** 55–84.

Blok, J., McWilliam, S. M., Butler, H. C., Gibbs, A. J., Weiller, G., Herring, B. L., Hemsley, A. C., Aaskov, J. G., Yoksan, S., and Bhamarapravat, N. (1992). Comparison of a dengue-2 virus and its candidate vaccine derivative: Sequence relationships with the flaviviruses and other viruses. *Virology* **187,** 573–590.

Domingo, E. J., Holland, J. J., and Ahlquist, P. (1988). "RNA Genetics" Vol. 3, CRC Press, Boca Raton, FL.

Droll, D. A., Murthy, K. H. M., and Chambers, T. J. (2000). Yellow fever virus NS2B-NS3 protease: Charged-to-alanine mutagenesis and deletion analysis define regions important for protease complex formation and function. *Virology* **275,** 335–347.

Falgout, B., Miller, R. H., and Lai, C. J. (1993). Deletion analysis of dengue virus type 4 nonstructural protein NS2B: Identification of a domain required for NS2B-NS3 protease activity. *J. Virol.* **67,** 2034–2042.

Farci, P., Shimoda, A., Coiana, A., Diaz, G., Peddis, G., Melpolder, J. C., Strazzera, A., Chien, D. Y., Munoz, S. J., Balestrieri, A., Purcell, R. H., and Alter, H. J. (2000). The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science* **288,** 339–344.

Goodenow, M., Huet, T., Saurin, W., Kwok, S., Sninsky, J., and Wai-Hobson, S. (1989). HIV-1 isolates are rapidly evolving quasispecies: Evidence for viral mixtures and preferred nucleotide substitutions. *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* **2,** 344–352.

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

Halstead, S. B. (1988). Pathogenesis of dengue: Challenges to molecular biology. *Science* **239,** 476–481.

Harris, E., Smith, G. L., Selle, J., Kramer, L. D., Valle, S., Sandoval, E., and Balmaseda, A. (l998). Typing of dengue viruses in clinical specimens, and mosquitoes by singe-tube multiplex reverse transcriptase PCR. *J. Clin. Microbiol.* **36,** 2634–2639.

Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc. Natl. Acad. Sci. USA* **86,** 5786–5790.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and VandePol. S. (1982). Rapid evolution of RNA genome. *Science* **215,** 1577–1585.

Holland, J. J., dela Torre, J. C., and Steinhauer, D. A. (1992). RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.* **176,** 1–20.

Huang, A. S. (1988). Modulation of viral disease processes by defective interfering particles, *In* "RNA Genetics" (E. Domingo, J. J. Holland, and P. Ahlquist, Ed.), Vol. 3, pp. 195–208. CRC Press, Boca Raton, FL.

Innis, B. L. (1995). Dengue and dengue hemorrhagic fever. *In* "Exotic Viral Infections—1995" (J. S. Porterfield, Ed.), pp. 103–146. Chapman & Hall, London, U.K.

King, C. C., Wu, Y. C., Chao, D. Y., Lin, T. H., Chow, L., Wang, H. T., Ku, C. C., Kao, C. L., Hwang, K. P., Lam, S. K., and Gubler, D. J. (2000). Major epidemics of dengue in Taiwan in 1991–2000: Related to intensive virus activities in Asia. *Dengue Bull.* **24,** 1–10, Geneva, WHO.

Kwok, S., and Higuchi, R. (1989). Avoiding false positive with PCR. *Nature* **339,** 237–238.

Lancaster, M. U., Hodgetts, S. I., Mackenzie, J. S., and Urosevic, N. (1998). Characterization of defective viral RNA produced during persistent infection of Vero cells with Murray Valley encephalitis virus. *J. Virol.* **72,** 2474–2482.

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

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

Lee, E., Stocks C. E., Amberg, S. M., Rice, C. M., and Lobigs, M. (2000). Mutagenesis of the signal sequence of yellow fever virus prM protein: Enhancement of signalase cleavage in vitro is lethal for virus production. *J. Virol.* **74**, 24–32.

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

Lindenbach, B. D., and Rice, C. M. (2001). Flaviviridae: The viruses and their replication. *In* "Fields Virology" (D. M. Knipe, P. M. Howley, *et al.* Eds.), 4th ed., pp. 991–1041. Lippincott William & Wilkins, Philadelphia, PA.

Mandl, C. W., Heinz, F. X., and Kunz, C. (1988). Sequence of the structural proteins of tick-borne encephalitis virus (western subtype) and comparative analysis with other flaviviruses. *Virology* **166,** 197–205.

Mangada, M. N., and Igarashi, A. (1998). Molecular and in vitro analysis of eight dengue type 2 viruses isolated from patients exhibiting different disease severities. *Virology* **244,** 458–466.

Markoff, L., Falgout, B., and Chang, A. (1997). A conserved internal hydrophobic domain mediates the stable membrane integration of the dengue virus capsid protein. *Virology* **233,** 105–117.

Martell, M., Esteban, J. I., Ouer, J., Genesca, J., Weiner, A., Esteban, R. Guardia, J., and Gomez, J. (1992). Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: Quasispecies nature of HCV genome distribution. *J. Virol.* **66,** 3225–3229.

Meyerhans, A., Cheynier, R., Albert, J., Martina, S., Kwok, S., Sninsky, J., Morfeldt-Manson, L., Asjo, B., and Wain-Hobson, S. (1989). Temporal fluctuation in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **58,** 901–910.

Monath, T. P. (1994). Dengue, the risk to developed and developing countries. *Proc. Natl. Acad. Sci. USA* **91,** 2395–2400.

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

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

Parks, G. D., and Lamb, R. A. (1991). Topology of eukaryotic type II membrane proteins: Importance of N-terminal positively charged residues flanking the hydrophobic domain. *Cell* **64,** 777–787.

Piatak, M., Saag, M. S., Yang, L. C., Clark, S. J., Kappes, J. C., Luk, K. C., Hahn, B. H., Shaw, G. M., Lifson, J. D. (1993). High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. *Science* **259,** 1749–1754.

Sala, M., and Wain-Hobson, S. (2000). Are RNA viruses adapting or merely changing. *J. Mol. Evol.* **51,** 12–20.

Smith, D. B., McAllister, J., Casino, C., and Simmonds, P. (1997). Virus "quasispecies": Making a mountain out of a molehill. *J. Gen. Virol.* **78,** 1511–1519.

Steinhauer, D. A., and Holland, J. J. (1987). Rapid evolution of RNA viruses. *Ann. Rev. Microbiol.* **41,** 409–433.

Sudiro, T. M., Zivny, J., Ishiko, H., Green, S., Vaughn, D. W., Kalayanarooj, S., Nisalak, A., Norman, J. E., Ennis, F. A., and Rothman, A. L. (2001). Analysis of plasma viral RNA levels during acute dengue virus infection using quantitative competitive reverse transcription-polymerase chain reaction. *J. Med. Virol.* **60,** 29–34.

von Heijne, C. (1990). The signal peptide. *J. Membr. Biol.* **115,** 195–201.

Wang, W. K., Lee, C. N., Kao, C. L., Lin, Y. L., and King, C. C. (2000). Quantitative competitive reverse transcription-PCR for quantification of dengue virus RNA. *J. Clin. Microbiol.* **38,** 3306–3310.

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

Weaver, S. C., Scott, T. W., and Rico-Hesse, R. (1991). Molecular evolution of eastern equine encephalomyelitis virus in north America. *Virology* **182,** 774–784.

Weaver, S. C., Brault, A. C., Kang, W., and Holland, J. J. (1999). Genetic and fitness changes accompanying adaptation of an arbovirus to vertebrate and invertebrate cells. *J. Virol.* **73,** 4316–4326.

Wolinsky, S. M., Korber, B. T. M., Neumann, A. U., Daniels, M., Kunstman, K. J., Whestsell, A. J., Furtado, M. R., Cao, Y., Ho, D. D., Safrit, J. T., and Koup, R. A. (1996). Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science* **272,** 537–542.

World Health Organization. (1997). "Dengue Hemorrhagic Fever, Diagnosis, Treatment, and Control" 2nd ed., World Health Organization, Geneva, Switzerland.

Zhang, L. Q., MacKenzie, P., Cleland, A., Holmes, E. C., Leigh Brown, A. J., and Simmonds, P. (1993). Selection for specific sequences in the external envelope protein of human immunodeficiency virus type 1 upon primary infection. *J. Virol.* **67,** 3345–3356.

Zhu, T., Mo, H., Wang, N., Nam, D. S., Cao, Y., Koup, R. A., and Ho, D. D. (1993). Genotypic and phenotypic chracterization of HIV-1 in patients with primary infection. *Science* **261,** 1179–1181.