MOLECULAR EVOLUTION, DISTRIBUTION AND GENETIC RELATIONSHIP AMONG THE DENGUE 2 VIRUSES ISOLATED FROM DIFFERENT CLINICAL SEVERITY

Basu Dev Pandey, Kouichi Morita, Futoshi Hasebe, Maria del Carmen Parquet and Akira Igarashi

Department of Virology, Institute of Tropical Medicine, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan

Abstract. Twenty-two strains of dengue 2 virus, isolated in China, Latin America, New Guinea and Thailand were subjected to phylogenetic analysis. The UPGMA analysis was carried out on each gene region of dengue virus and demonstrated that outcome from most of the gene regions showed similar results except those from NS4B and 3′UTR with very short nucleotide length. Among ten regions examined, the results from E gene documented the geographical differences of the virus strains most clearly and all the American strains (Mara 4, IQT1797 and S1) were distantly related to the Asian isolates. As for the 16 Thai strains isolated in 1993, they were clustered into three groups and a strain from a DSS patient formed a distinct branch compared to the other two groups. This finding from phylogenetic analysis is consistent with earlier conclusion and support the severity related subtyping of dengue 2 virus based on amino acid changes.

INTRODUCTION

Dengue (DEN) fever results from infection with one of the four serotypes of dengue viruses (DEN-1, DEN-2, DEN-3 and DEN-4). This mosquito-borne virus infection is currently one of the most important flavivirus infections because of its progressively extended geographic distribution and increasing number of cases with severe hemorrhagic manifestations. It is prevalent in the tropical and subtropical areas of Africa, Latin America, Oceania and, most notably, in Southeast Asia. At least 100 million cases of dengue fever (DF) and more than 30,000 deaths due to dengue hemorrhagic fever (DHF) as well as dengue shock syndrome (DSS) occur every year (Gubler, 1995; Monath *et al*, 1996; Knudsen, 1997). Most patients develop an acute febrile illness including undifferentiated classical dengue fever. Others, however, go on to experience two more serious clinical manifestations: DHF and DSS.

In the last 20 years DF, DHF and DSS have emerged as a most important disease in the subtropical and a tropical countries, and a leading cause of hospitalization and death among children in Southeast Asian countries. The major factors in the emergence of DEN as one of the most serious health problems facing the developing world are the ever increasing size, rapid urbanization, increasing frequency of mobility of human populations as well as the lack of an effective vaccine (Monath, 1994: Gubler, 1998).

Multiple serotypes of dengue virus are prevalent in Thailand but DEN- 2 appeared to be more associated with DHF (Burke *et al*, 1989). An outbreak of DHF in Thailand was first recognized in Bangkok in 1958. Since then DEN virus has spread to all parts of country establishing the endemo-epidemic pattern (Nimmanitya, 1987). It is not yet understood whether this observation reflects differences in virus-host interaction or higher virulence of DEN-2 viruses. It was observed that the Southeast Asian genotypes are more commonly associated with causing DHF than the American strains (Leitmeyer *et al*, 1999). Severity related amino acid differences on DEN-2 virus has also been observed which could be correlated with the pathogenesis of DHF (Mangada and Igarashi, 1998). We are able to classify 19 strains of DEN-2 viruses, which were isolated in the same epidemic area during the same epidemic season into 3 subtypes on the basis of nonsynonymous amino acid changes (Pandey and Igarashi, 2000) which correlated with severity of the disease.

Dengue virus genome is a single, positive-stranded RNA, approximately 11 Kb in length with single open reading frame flanked by 5′ and 3′ UTR. The gene order of the dengue polyprotein is

Correspondence: Dr Kouichi Morita, Department of Virology, Institute of Tropical Medicine, Nagasaki University, Sakamoto 1-12-4, Nagasaki 852-8523, Japan.
Tel:+(81-95) 849-7829; Fax: +(81-95) 849-7830; E-mail: moritak@net.nagasaki-u.ac.jp

C-PrM(M)-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5. Phylogenetic analysis based on nucleotide sequence data has been used for molecular epidemiology of dengue viruses. Most of the available reports are based on the analysis of partial gene sequences (Rico-Hesse, 1990; Blok *et al*, 1991; Duebel *et al*, 1993). The increasing prevalence of DEN virus infection especially in Latin America seems to occur as a reinvasion of that continent by *Aedes aegypti*. It highlights the mosquito as an emerging principal vector of dengue virus. But where did the virus come from and what factors have facilitated its spread? Molecular phylogenetic analysis could be a useful tool to provide the answers to these questions. In this study, we report the phylogenetic analysis of DEN-2 viruses isolated from different countries and different clinical severities which could give a better understanding of the origin and genetic relationship of DEN-2 virus strains.

MATERIALS AND METHODS

Viruses

Origin, clinical diagnosis, disease severity and genebank accession number of each of the 22 virus strains used in this study are shown in Table 1. Nineteen Thai strains were isolated from patients in Nakhon Phanom Provincial Hospital, Thailand in 1993. Samples from DF cases were collected from the outpatient department, while the other samples were taken from hospitalized patients with DHF/DSS. Clinical diagnosis and disease severity grading were determined according to the criteria of the World Health Organization (WHO, 1986). The nucleotide sequence of American strains (Mara 4, IQT1797, S1) and China 04, 16681 (Kinney, 1997) and NGC strains were obtained from the genebank.

Nucleotide sequencing

Viral RNA was extracted from infected C6/36 cell culture fluid by using Trizol LS (GIBCO BRL, USA). cDNA products were prepared from the RNA using superscript II, RNase H-reverse transcriptase according to the manufacturer's instruction (GIBCO BRL, USA). cDNA was used for PCR amplification and RT-PCR product was purified using Microcon microconcentrators (Amicon, USA) followed by direct sequencing using the sequencer (Applied Biosystem Prism, 310) as described earlier (Pandey and Igarashi, 2000). DNASIS-Mac version 3.6 Software System, Hitachi, 1995, was used for primer selection, homology search and comparison of all obtained sequences. Construction and analysis of the genomic DEN DNA were done using sequence Navigator and Factura software (Perkin-Elmer/Applied Biosystems, USA).

Phylogenetic analysis

The sequence in each genome was used to draw a phylogenetic tree for the comparison among the 22 DEN-2 strains, including Thai, prototype NGC, China 04 and the American strains (Mara 4, IQT1797, S1). Sequences of DEN-1, 3 and 4 viruses were included as outgroups (data not shown).

The PHYLIP package of software programs distributed by Felsenstein (SeqBoot, DNADIST and UPGMA) was used to calculate nucleotide evolutionary distances and to prepare the phylograms. The trees were viewed using the TREEVIEW program as described previously (Felenstein, 1993). Thousand times bootstrap reassembling of the data set was done to ascertain support for major branches of the tree.

RESULTS

Analysis of the structural protein

PrM/M protein: The phylogenetic analysis and nucleotide evolutionary distances within the 19 Thai strains, 16681, DEN-2 NGC, Mara 4, IQT1797 and DEN-2 S1 were analyzed. The American strain Mara 4 was distantly related to the Thai strains with the lowest evolutionary distance of 0.0769 with ThNH28/93 and the highest 0.1009, with 16681. Based on the phylogenetic tree constructed using the PrM gene (Fig 1) all the Thai isolates came from DHF and DF showed a closer relationship with the 16681 strain, belonging to the Southeast Asian genotypes (Rico-Hesse, 1990), whereas the American isolates Mara 4, IQT1797 and DEN-2 S1 were distantly related to the Thai isolates. The subtype III viruses isolated from mild DF cases were classified in a distinct branch and subtype II virus strains isolated from 13 DHF cases and 2 DF cases were grouped into a separate cluster. Subtype I virus isolated from a DSS patient formed a distinct branch from subtype II or III virus.

Envelope (E) protein: E protein is the major protein component of the virion of DEN virus possessing antigenic properties. The phylogram

Table 1
Information on clinical severity, location and year of isolation of the dengue 2 viruses.

Strain	Diagnosis	Sex	Age (year)	Year of Isolation	Location	GenBank accession no
ThNH7/93	DSS	F	12	1993	Thailand	AF022434
ThNH28/93	DHF (2)	M	10	1993	Thailand	AF022435
ThNH52/93	DHF(1)	M	7	1993	Thailand	AF022436
ThNHp11/93	DF	M	14	1993	Thailand	AF022437
ThNHp12/93	DF	F	11	1993	Thailand	AF022438
ThNHp14/93	DF	M	11	1993	Thailand	AF022439
ThNHp16/93	DF	F	12	1993	Thailand	AF022440
ThNHp36/93	DF	F	9	1993	Thailand	AF022441
ThNH29/93	DHF(2)	M	11	1993	Thailand	AF169678
ThNH36/93	DHF(2)	M	13	1993	Thailand	AF169679
ThNH45/93	DHF(2)	F	9	1993	Thailand	AF169680
ThNH54/93	DHF(1)	M	13	1993	Thailand	AF169682
ThNH55/93	DHF(1)	M	7	1993	Thailand	AF169681
ThNH62/93	DHF(2)	M	13	1993	Thailand	AF169683
ThNH63/93	DHF(1)	F	11	1993	Thailand	AF169684
ThNH69/93	DHF(1)	M	9	1993	Thailand	AF169685
ThNH73/93	DHF(1)	F	8	1993	Thailand	AF169686
ThNH76/93	DHF(1)	F	9	1993	Thailand	AF169687
ThNH81/93	DHF(1)	F	8	1993	Thailand	AF169688
16681	DHF	*	*	1964	Thailand	U87411
China 04	DHF	*	*	1985	China	AF119661
D2/NGC	DF	*	*	1944	New Guinea	M32941
Mara 4	DHF	*	*	1990	Venezuela	AF100149
IQT 1797	DF	*	*	1995	Peru	AF100147
PR 159S1	DF	*	*	1969	Puerto Rica	M32953

*Information not available

Table 2
Amino acid subtyping on the Thai dengue 2 strains[a].

	Position of amino acid replacement in the dengue 2 virus genome										
	PreM		NS1		NS2A				NS3		NS5
Subtype	16	81	278	281	41	136	139	215	13	118	337
I	I	T	G	D	V	I	D	N	I	T	T
II	I	T	D	D	V	I	N	N	M	T	T
III	R	A	D	E	M	M	N	S	M	A	M

[a]Pandey and Igarashi (2000)

showed that ThNHp11, p12 and p14 (subtype III) are closely related to the ThNH7 (Fig 1). All isolates of subtype II virus were clustered together in one group, almost similar to the results obtained in the PrM region, the American isolates were distantly related to the Thai strains. The results of this region most clearly represent the geographical differences of the virus strains. The homologies within the Thai Nakhon Phanom strains were very high, ranging from 98.7-100% (data not shown). However divergence between the Chinese and other Thai strains were about 7%. On the other hand the American strains differ by 11% from the Thai strains.

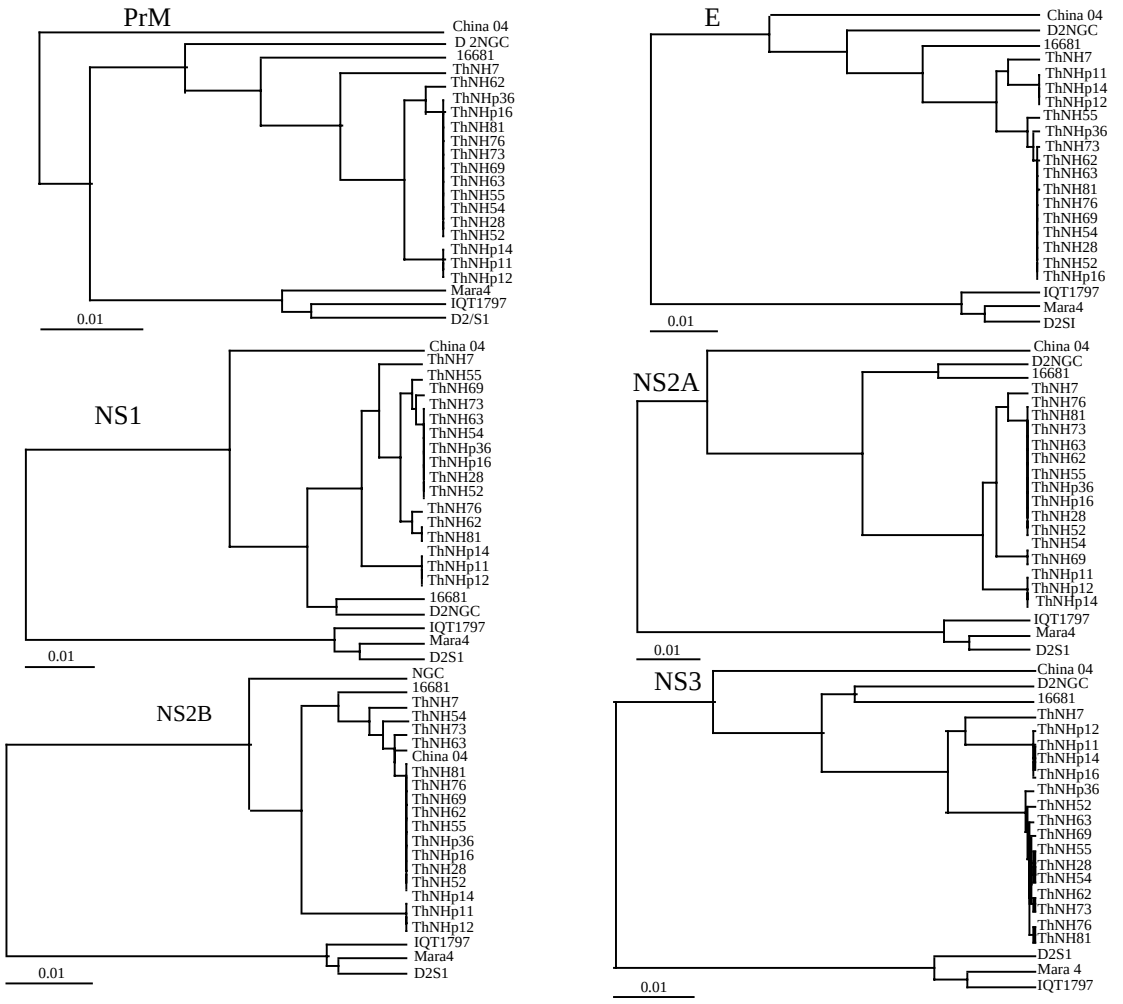

Fig 1–Phylogenetic tree generated using UPGMA utilizing PrM, E, NS1, NS2A, NS2B and NS3 sequences of Thai, China, prototype NGC and the three American strains.

Analysis of the nonstructural proteins

Nonstructural protein 1 (NS1): The phylogenetic tree using the nucleotide sequence of the NS1 region showed similar results to those obtained for the PrM and E region (Fig 1). The American isolates clustered together and were distantly related to all the Thai, China 04 and NGC strains. The IQT1797 strain was distantly related to the other strains. The subtype III viruses were clustered in one group whereas the ThNH7 was in a different branch, all strains from DHF cases were closely related to each other.

Nonstructural protein 2A and 2B (NS2A and NS2B): All the three American strains were distantly related to all of the Thai strains when analyzing NS2A (Fig 1). NGC and 16681 were clustered together in one distinct branch with an evolutionary distance of 0.0280. The subtype III viruses were more closely related to NGC than to ThNH7. The Thai strains from DHF and DF cases were clustered in one group. These findings were consistent with NS2B except that the Chinese strains seems to be closer to the Thai DHF strains in the NS2B region.

Nonstructural protein 3 (NS3): NS3 protein is important for viral replication as it possesses the protease and helicase motifs (Arias *et al*, 1993). We found a strain specific amino acid replacement

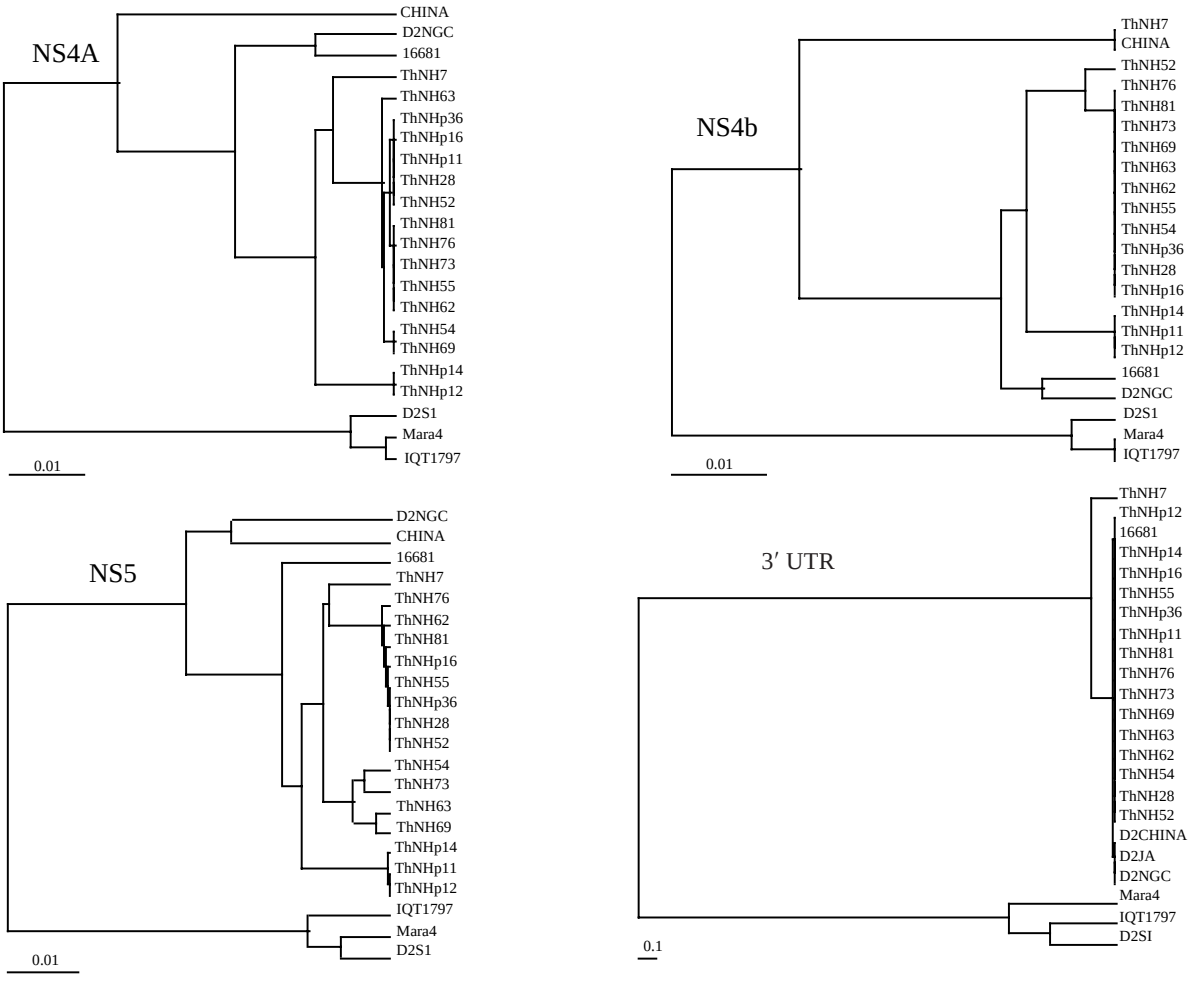

Fig 2–Phylogenetic tree generated using UPGMA utilizing NS4A, NS4B, NS5 and 3′ UTR sequences.

in position 13 (M↔I) of the NS3 region in subtype I virus (Pandey and Igarashi, 2000). The phylogenetic tree (Fig 1) showed similarities with the results observed for the E region but ThNHp16 was included into the group of ThNHp11, p12, and p14. The subtype I virus (ThNH7) was closer to subtype III viruses than the rest of the strains. The American strains were distantly related to all the Asian strains with an evolutionary distances of 0.0885.

Nonstructural protein NS4A, NS4B and NS5: The American isolates were clustered in one group in NS4A and NS4B and far from all the Asian strains, similar to the rest of the genome. The NGC strain is closely related to the three strains ThNHp11, p12 and p14 of subtype III virus in the NS4B region than the others (Fig 2). The lowest evolutionary distance between the NGC and subtype III virus is 0.0304 in NS4B region whereas the evolutionary distance between ThNH7 and NGC is 0.0653 (data not shown).

NS5 is the major nonstructural protein important for viral replication, in association with NS3. A phylogram constructed using the majority of the other genes of viral proteins gave almost identical result with the phylogram constructed using NS5 gene (Fig 2). Viruses isolated from DF and belonging to subtype III were clustered in one group whereas all DHF strains and 2 DF strains

were clustered in another group. On the other hand the virus isolated from DSS showed a separate branch similar to that observed in other genes.

Analysis on 5′ and 3′ UTR: The 5′ UTR is responsible for the initiation of translation while the 3′ UTR region of the viral genome is essential for viral replication and growth: it serves as a signal for the initiation of minus-strand synthesis (Hahn *et al*, 1988; Men *et al*, 1996). All the isolates possess identical sequences with that of the prototype NGC strain in 5′ UTR. However, subtype I virus possesses one nucleotide insertion at position 335 of the 3′ UTR giving it a total genome length of 10,724 nucleotides compared to 10,723 nucleotides of the other Thai isolates. On the other hand, all the American strains have an 8 nucleotide deletion at the beginning of the 3′ UTR (Leitmeyer *et al*, 1999). Two definite clusters were observed in the phylogram in 3′ UTR. The American strains clustered in one group and the rest of the strains in another (Fig 2).

DISCUSSION

Molecular epidemiology is one of the emerging sciences and an useful tool to study the origin, genetic relationship and classification of the viruses. Nucleotide sequence analysis of dengue viruses can reveal genetic variation among strains within the same serotype, and geographical movement of strains. Several studies have been carried out on the molecular evolution of DEN viruses by performing quantitative comparisons of nucleotides based on small fragments of the genome (Rico-Hesse, 1990; Blok *et al*, 1991; Duebel *et al*, 1993). Most of the previous studies on genetic relationship and epidemiology of dengue viruses were based on short fragment sequences.

DEN-2 is currently the predominant serotype associated with DHF in Thailand. Some of the genotypes have wide distributions ranging from Asia, across the Pacific to Latin America (Monath, 1996). Such a geographical range implies that dengue virus frequently moves between populations because of human activities. It leads to epidemics perhaps due to the importation of new strains into populations with no existing previous immunity. As there are no animal models for DEN to test the hypothesis for the severity of the disease, comparing nucleotide and deduced amino acid sequences from patients exhibiting different clinical manifestations could be one approach to study the pathogenesis of DHF (Mangada and Igarashi 1998; Pandey and Igarashi, 2000). The studies of Rico-Hesse (1990) and Leitmeyer *et al* (1999) on DEN-2 viruses indicated that Southeast Asian genotype is associated with severe clinical manifestations, while the Caribbean genotype is associated with milder diseases.

The result obtained in this study of DEN-2 virus provides important information on the genetic relationships and origin of the virus in Asia and America. Here, we could infer possible groupings according to the clinical severity in most of the regions of the DEN genome. Subtype III viruses isolated from mild DF cases were clustered in a distinct branch whereas subtype II virus strains isolated from DHF and DF cases were grouped into a separate cluster. Subtype I virus isolated from a DSS patient formed a distinct branch compared to subtype II or III virus (Figs 1 and 2). The American isolates Mara 4, IQT1797 and S1 were distantly related with Asian isolates. This result support our previous finding of severity related molecular differences (Pandey and Igarsahi, 2000).

E protein is the major target of virus selection in nature. The result obtained here indicates that the E protein most appropriately represented the geographical distribution of DEN-2 viruses. Other regions including PrM, NS1, NS2A, NS2B, NS4B and NS5 could be used for the analysis but E region is the best region to perform the phylogenetic analysis to determine the geographical differences and genetic relationship among the DEN viruses. Furthermore, our results showed that the C, NS3, NS4A and 3′ UTR regions may not be appropriate regions to carry out molecular epidemiological study by constructing phylograms.

ACKNOWLEDGEMENTS

The authors are grateful to Dr GL Enriquez for critical reading of the manuscript. The first author is a recipient of the Monbusho scholarship from the Ministry of Education, Science, Sports and Culture of the Government (MONBUSHO) of Japan for his stay and study in the Institute of Tropical Medicine, Nagasaki University. This study was supported by the Grant-in-Aid for Scientific Research from the MONBUSHO, No. 07457078.

REFERENCES

Arias CF, Preugschat F, Strass JH. Dengue 2 virus NS2B

and NS3 from a stable complex that can cleave NS3 within the helicase domain. *Virology* 1993; 193: 888-99.

Burke DS, Nislak A, Johnson DE, *et al*. A prospective study of dengue infection in Bangkok. *Am J Trop Med Hyg* 1989; 38: 172-80.

Blok J, McWilliam SM, Vitarana, UT. NS1 gene sequences from eight dengue 2 viruses and their evolutionary relationships with other dengue 2 viruses. *Arch virol* 1991; 118: 209-23.

Deubel V, Nogueria RM, Drouet MT, *et al*. Direct sequencing of genomic cDNA fragments amplified by the polymerase chain reaction for molecular epidemiology. *Arch Virol* 1993; 129: 197-210.

Felsenstein J. PHYLIP (Phylogeny Inference Package) Version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, 1993.

Gubler DJ, Clark GG. Dengue/dengue haemorrhagic fever, the emergence of a global health problem. *Emerg Infect Dis* 1995; 1: 55-7.

Hahn YS, Galler R, Hunkapillor T, *et al*. Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. *Virology* 1988; 162: 167-80.

Kinney RM, Butrapet S, Chang GJJ, *et al*. Construction of infectious cDNA clones for dengue 2 virus: strain 16681 and its attenuated vaccine derivative, strain PDK-53. *Virology* 1997; 230: 300-8.

Knudsen AB. Global strategy for the prevention and control of dengue and dengue haemorrhagic fever. In: Saluzzo JF, Doded B, ed. Emerging diseases: Factors in the emergence of arbovirus diseases. Paris: Elsevier, 1997: 129-40.

Leitmeyer KC, Vaughn DW, Watts DM, *et al*. Dengue virus structural differences that correlate with pathogenesis. *J Virol* 1999; 73: 4738-47.

Mangada MNM, Igarashi A. Molecular and *in vitro* analysis of eight dengue type 2 viruses isolated from patients exhibiting different disease severities. *Virology* 1998; 244: 458-66.

Rico-Hesse R. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* 1990; 17: 479-93.

Nimmannitya S. Dengue haemorrhagic fever in Thailand. *Southest Asian J Trop Med Public Health* 1987; 18: 291-4.

Men R, Bray M, Clark D. Dengue type 4 virus mutant containing deletions in the 3' noncoding region of the RNA genome: analysis of the growth restriction in cell culture and altered viremia pattern and immunogenicity in rhesus monkeys. *J Virol* 1996; 70: 3930-7.

Monath TP, Heinz FX. Flaviviruses. In: Fields BN, Knife DM, Howell PM, *et al*, eds. Fields virology, 3rd ed. Philadelphia: Lippincott-Raven, 1996: 961-1034.

Monath TP. Yellow fever and dengue - the interactions of virus, vector and host in the reemergence of epidemic disease. *Semin Virol* 1994; 5: 133-45.

Pandey BD, Igarashi A. Severity related molecular differences among nineteen strains of dengue type 2 viruses. *Microbiol Immunol* 2000; 40: (in press).

World Health Organization, Dengue hemorrhagic fever. Diagnosis, treatment, prevention and control. 1st ed. WHO, 1986.