

Molecular epidemiology of Malaysian dengue 2 viruses isolated over twenty-five years (1968-1993)

M.-Y. Fong ^{(1) (*)}, C.-L. Koh ⁽²⁾ and S.-K. Lam ⁽³⁾

⁽¹⁾ *Department of Parasitology,*

⁽²⁾ *Institute of Biological Sciences and*

⁽³⁾ *Department of Medical Microbiology, University of Malaya, 50603 Kuala Lumpur (Malaysia)*

SUMMARY

The limited sequencing approach was used to study the molecular epidemiology of 24 Malaysian dengue 2 viruses which were isolated between 1968 and 1993. The sequences of a 240-nucleotide-long region across the envelope/non-structural 1 protein (E/NS1) gene junction of the isolates were determined and analysed. Alignment and comparison of the nucleotide and deduced amino acid sequences of the isolates revealed that nucleotide changes occurred mostly at the third position of a particular codon and were of the transition (A↔G, C↔U) type. Five nucleotide changes resulted in amino acid substitutions. Pairwise comparisons of the nucleotide sequences gave divergence values ranging from 0 to 9.2%. At the amino acid level, the divergence ranged between 0 and 3.8%. Based on the 6% divergence as the cut-off point for genotypic classification, the isolates were grouped into two genotypes, I and II. Comparison of the nucleotide sequences of the Malaysian dengue isolates with those of the dengue viruses of other regions of the world revealed that members of genotypes I and II were closely related to viruses from the Indian Ocean and Western Pacific regions, respectively.

Key-words: Dengue 2 virus; Molecular epidemiology, Nucleotide sequence divergence, Dendrogram, Genotypes, Malaysia.

INTRODUCTION

Dengue viruses are members of the family *Flaviviridae* (Westaway *et al.*, 1985) and can be grouped into four serotypes (1, 2, 3 and 4) which are distinguished by serological techniques. Dengue viruses are transmitted by *Aedes* mosquitoes, of which *A. aegypti* is the most efficient. The viruses are a major public health concern as they are the aetiological agents of dengue fever and the more severe dengue haemorrhagic fever and den-

gue shock syndrome. Dengue is documented in virtually every part of the tropical and subtropical regions of the world. Dengue is endemic in Malaysia and dengue fever was first reported in the country in 1902 (Skae, 1902). The first case of dengue haemorrhagic fever in Malaysia was reported in 1962. (Rudnick *et al.*, 1965). Serious dengue outbreaks have occurred in Malaysia in four-year cycles between 1974 and 1990 (i.e., 1974, 1978, 1982, 1986 and 1990). However, this cycle is no longer apparent as cases of dengue

Submitted April 6, 1998, accepted September 16, 1998.

(*) Corresponding author.

remained high since 1990. Dengue 2 and 3 viruses are currently the predominant serotypes in Malaysia.

The dengue virus genome consists of a single-stranded positive sense RNA of 11 kb with the following gene order: 5'-C-prM(M)-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5-3' (Mackow *et al.*, 1987; Deubel *et al.*, 1988; Hahn *et al.*, 1988; Osatomi and Sumiyoshi, 1990; Fu *et al.*, 1992). The RNA is directly translated into a long polypeptide which is eventually cleaved and modified to produce three structural and seven non-structural proteins (Rice *et al.*, 1985). Analysis of nucleotide sequences of dengue viruses can reveal genetic variation among strains within the same serotype, and geographical movement of strains. Rico-Hesse (1990) determined the molecular evolution and distribution of dengue 1 and 2 viruses by performing quantitative comparisons of a 240-nucleotide-long stretch across the E/NS1 gene junction. This stretch accounts for only 2.2% of the dengue virus genome. Consequently, the limited sequence approach has been used in numerous studies on genetic relatedness and epidemiology of dengue viruses (Chungue *et al.*, 1993; Deubel *et al.*, 1993; Lanciotti *et al.*, 1994; Chungue *et al.*, 1995).

In this paper, we used the approach of Rico-Hesse (1990) to study the genetic relatedness and molecular epidemiology of Malaysian dengue 2 viruses which were isolated over twenty-five years (1968-1993). The results obtained gave a better understanding of the origin and movement of dengue 2 strains in Malaysia and its neighbouring regions.

MATERIALS AND METHODS

Dengue 2 virus isolates

Twenty-four dengue 2 virus isolates from human patients were selected (table I). These viruses were isolated between 1968 and 1993. All isolates except

Table I. Malaysian dengue 2 virus isolates used in the molecular epidemiology study.

Isolate	Year isolated	Passage level
P7-845	1968	P3
P9-573	1969	P2
P70-104	1970	P2
P73-718	1973	P2
P73-987	1974	P3
P75-559	1976	P3
P78-182	1978	P3
P78-612	1979	P4
V32491	1980	P1
V34380	1982	P1
V48171	1983	P3
V54451	1986	P3
V60350	1987	P1
M2	1987	P4
V62779	1988	P1
V83440	1989	P1
V89821 (Kuala Selangor)	1989	P1
V92436	1990	P1
JB349 (Johor Bahru)	1990	P1
KS30 (Sarawak)	1990	P1
TG2 (Tanjung Karang)	1990	P1
V94035	1991	P1
V05375	1992	P1
V03857	1993	P1

All isolates, unless indicated in brackets, were from confirmed dengue cases in Klang Valley in the state of Selangor. Kuala Selangor and Tanjung Karang are two rural districts, situated about 55-60 km north east of Klang Valley. Johor Bahru is a town at the southern tip of the Malaysia Peninsula. Sarawak is a Malaysian state on the island of Borneo (fig. 1).

Isolates of the "P" series (1968-1979) were isolated by the Arbovirus Diagnostic and Research Group of the University of California International Centre for Medical Research (UCICMR). The group was funded by the Hooper Foundation, USA and was based at the Department of Medical Microbiology, University of Malaya between 1960 and 1980. From 1980 onwards, virus isolation was carried out by the WHO Collaborating Centre for Arbovirus Reference and Research (DF/DHF), in the Department of Medical Microbiology, University of Malaya.

V89821, TG2, JB349 and KS30 were from dengue cases in the Klang Valley in Peninsular Malaysia. V89821 and TG2 were isolated from patients following local dengue outbreaks in two neighbouring rural districts about 55 km northeast of Klang Valley. JB349 was isolated from a patient in Johor

PCR = polymerase chain reaction.
RT-PCR = reverse transcription/PCR.

Taq = *Thermus aquaticus*.

Bahru, a town situated at the southern tip of the Malaysian peninsula. KS30 was an isolate from Sarawak, a Malaysian state on the island of Borneo.

Virus propagation and viral RNA extraction

All virus isolates were propagated in the C6/36 mosquito cell line. Essentially, a monolayer of C6/36 cells grown in a 150-cm² tissue culture flask was inoculated with 0.5 ml of patient's serum or infected cell culture supernatant and kept at 37°C for 1 h to allow virus adsorption. After adsorption, the infected cell culture was incubated at 28°C. Virus growth was monitored by assaying the supernatant for haemagglutinating activity. Once the haemagglutinating titre reached 1:16 or higher, the supernatant was removed. Viral RNA was extracted from the infected cells using the method of Kautner and Lam (1992).

cDNA synthesis and RT-PCR

Both the cDNA and RT-PCR processes were carried out according to the methods described by Chan *et al.* (1994). Basically, a small quantity ($\approx 1 \mu\text{g}$) of the extracted viral RNA was used as the template for random priming reverse transcription, and the resulting cDNA served as the template for amplification by PCR. Amplification was carried out in PCR buffer containing the four deoxyribonucleoside 5'-triphosphates (dATP, dCTP, dGTP and dTTP), *Taq* DNA polymerase and a pair of primers: (5'-GACAACAA-TGATGGGAGCGAAGAGAATGGC^{3'}, forward) and (5'-TCAACTCA GTGGGCTGAGGCCGCAGAGATC^{3'}, reverse). The primers annealed at map sites 2145-2174 and 2726-2755 [numbers according to Hahn *et al.* (1988)] on the dengue 2 virus genome and flanked the E/NS1 gene junction. The amplification parameters were as follows: 72°C for 1.5 min; followed by 35 cycles of denaturation at 94°C (1 min), primer annealing at 55°C (1 min), primer extension at 72°C (1 min); and finally 72°C for 12 min. After amplification, a 611-base pair double-stranded DNA fragment which encompassed the 240-nucleotide E/NS1 gene junction region was produced. The PCR product was purified by using silica particles (GeneClean, Bio101, USA).

Sequencing the PCR product

The sequence at the E/NS1 gene junction was determined by directly sequencing the PCR product. The chain termination approach of Sanger *et al.* (1977) was used in the sequencing step. Sequencing reactions were performed by using the "Sequenase Version 2.0 DNA Sequencing Kit" (USB Corp., USA) and were primed by the primers described earlier.

Computer analyses of sequence data

Analyses of nucleotide and deduced amino acid sequences were performed by using the CLUSTAL V programme of Higgins *et al.* (1992). The analyses included pairwise sequence comparisons, alignment of sequences, computation of divergence values and construction of dendrograms.

RESULTS

Figures 1 and 2 show the alignments of nucleotide and deduced amino acid sequences, respectively, of the E/NS1 junction region of the virus isolates. In each alignment, the oldest isolate, P7-845, was taken as the reference strain.

Alignment of the nucleotide sequences shows that changes were distributed throughout the E/NS1 junction region. Taken together, nucleotide changes occurred at 38 positions (15.8%) of the sequenced region. Generally, the changes were at the third position of a particular codon, and were of the transition (A \leftrightarrow G, C \leftrightarrow U) type. Therefore, most of the changes were silent. Nucleotide changes which resulted in amino acid substitutions were located at five positions: 2318, 2320, 2344, 2386 and 2478. The corresponding amino acid substitutions were: 741 (V \rightarrow A), 742 (V \rightarrow I), 750 (S \rightarrow T), 764 (V \rightarrow I) and 794 (I \rightarrow M). Only at position 2478 did the nucleotide change occur at the third position of the codon.

Pairwise comparisons of the nucleotide sequences of the isolates show differences which ranged from 0 to 22. Therefore, maximal divergence was 9.2%, and this is seen in the comparisons of V03857 with V34380, V60350, M2, V62779 and V94035. As expected, amino acid sequence similarity among the isolates was very high; maximal divergence was only 3.8%. More interestingly, in some cases of full homology (i.e. 0% divergence), the isolates involved were from different geographic origins and were temporally separated by more than 20 years. This is observed, for example, in the comparisons of P7-845 and P9-573 with those isolated in 1990, such as V92436, TG2, JB349 and KS30.

Genetic relationships among the isolates were clear when presented in the form of a dendrogram.

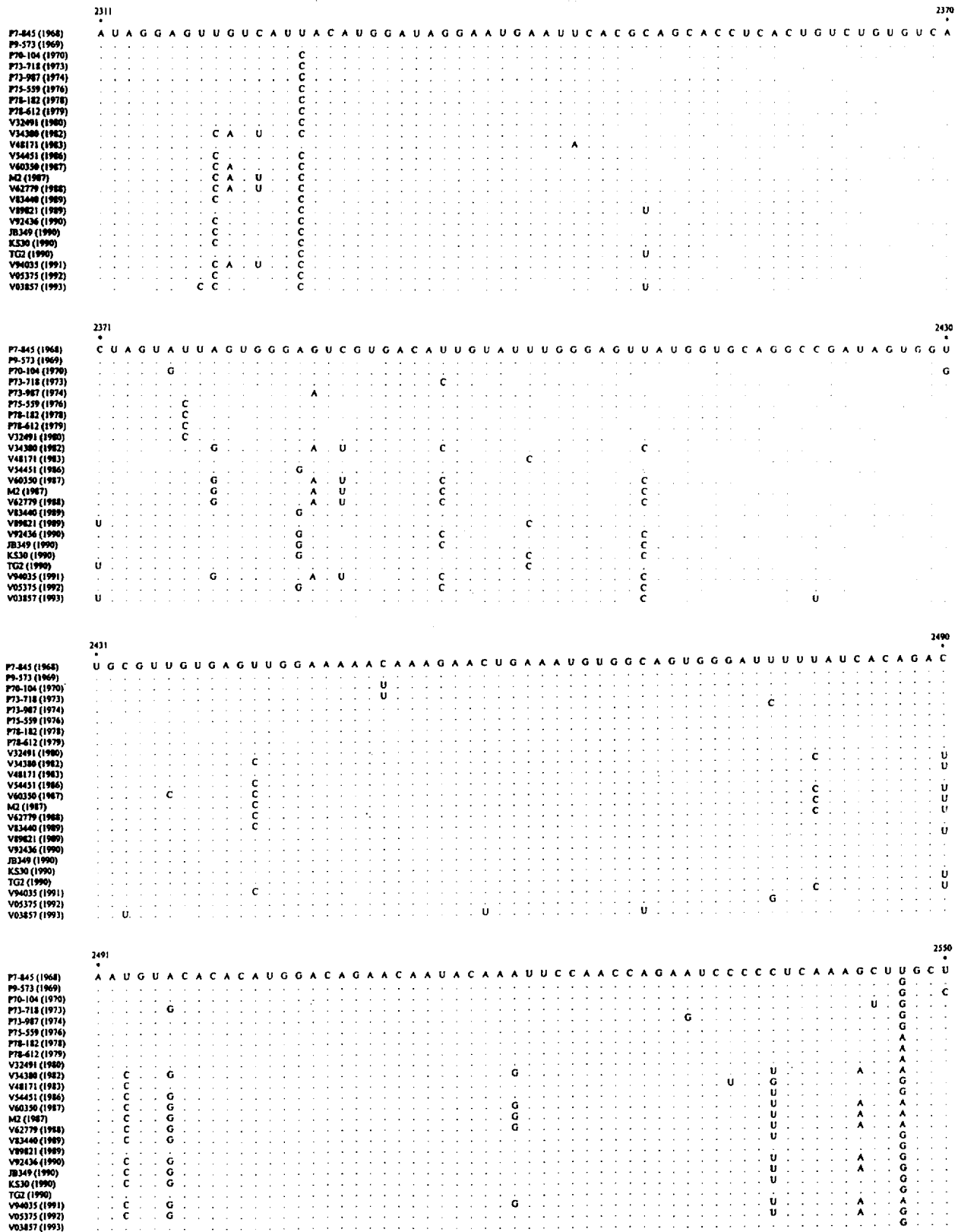


Fig. 1. Alignment of nucleotide sequences at the E/NS1 gene junction of 24 Malaysian dengue 2 virus isolates.

Dots indicate nucleotides identical to those of the reference strain, P7-845 [positions are numbered according to Hahn *et al.* (1988)].

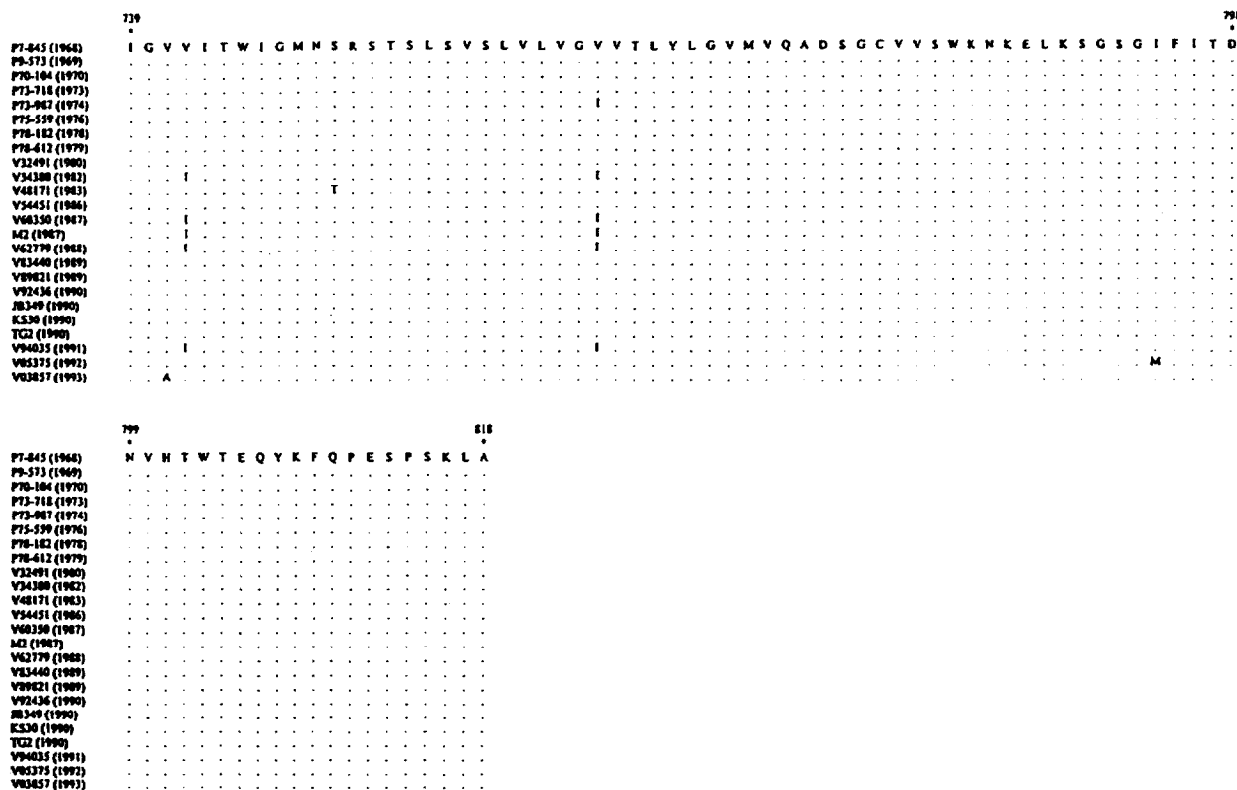


Fig. 2. Alignment of deduced amino acid sequences at the E/NS1 gene junction.

Dots indicate amino acids identical to those of the reference strain, P7-845 [amino acid positions are numbered according to Hahn *et al.* (1988)].

In our study, the dendrogram was constructed with the aid of the CLUSTAL V programme (Higgins *et al.*, 1992). All pairwise comparisons and the respective divergence values were computed and, in this way, eliminated any bias introduced by comparison to only a single reference strain. However, it should be pointed out that the dendrogram is without time coordinates and therefore does not reveal the evolution rates of the isolates.

Figure 3 shows the dendrogram based on nucleotide sequence divergence among the isolates. Based on 6% divergence as a cut-off point for the definition of a dengue virus genotype within a serotype (Rico-Hesse, 1990), it was possible to classify the isolates into two main genotypes.

The first genotype (I) contained three subgroups. The first subgroup consisted mainly of

isolates of older origin (1968-1990); the second and third subgroups comprised more recent isolates (1986-1992 and 1993, respectively).

The second genotype (II) contained members isolated between 1982 and 1991. With the exception of V60350, all the isolates were identical (0% divergence) to each other. Therefore, no distinct or significant subgroups were noted within genotype II.

DISCUSSION

The study by Rico-Hesse (1990) demonstrated that limited sequence analysis on the viral RNA is sufficient for determining the epidemiology and classification of dengue virus strains. This approach was further improved by performing

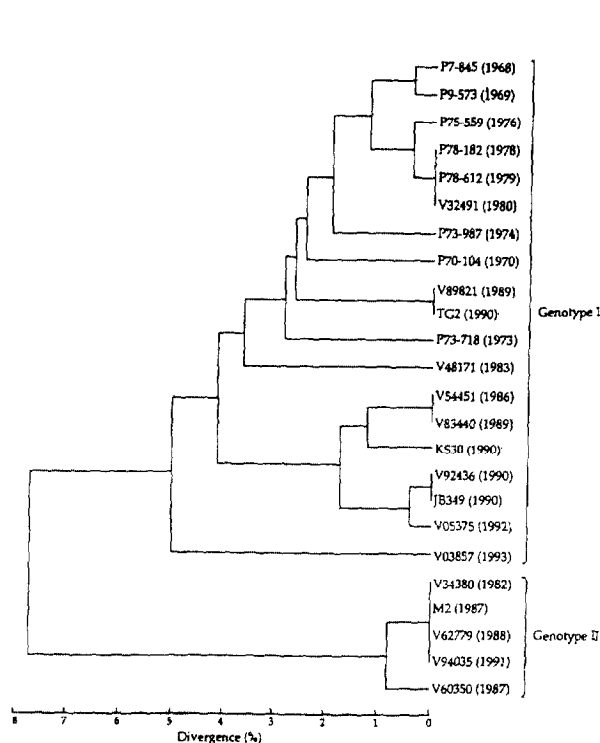


Fig. 3. Dendrogram based on nucleotide sequence divergence (%) among Malaysian dengue 2 virus isolates.

direct sequencing on RT-PCR products. Apart from saving time, this approach enables characterization of the strains without extensive virus propagation and RNA purification. In addition, direct sequencing of the dominant population of RT-PCR products reduces misincorporation errors by the *Taq* DNA polymerase (Lewis *et al.*, 1992). Hence, in this study, direct sequencing of a RT-PCR product containing the E/NS1 gene junction region of dengue 2 virus isolates was carried out to rapidly generate sequence data to be interpreted epidemiologically.

Dengue has long been endemic in Malaysia, and dengue 2 and 3 viruses are currently the prevalent serotypes. Because of the endemicity of the disease and the proximity of Malaysia with countries having high levels of dengue activity, it would not be unexpected that the Malaysian dengue virus isolates show complex or multitypic groupings.

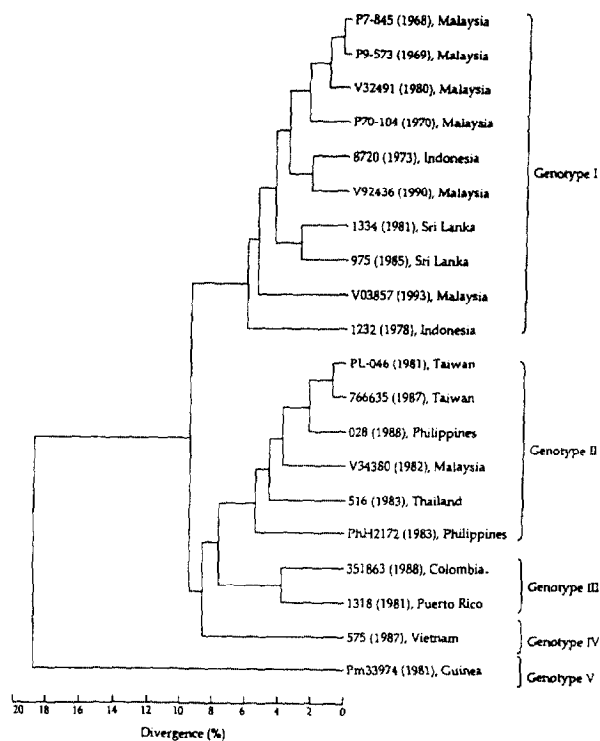


Fig. 4. Dendrogram based on nucleotide sequence divergence (%) among dengue 2 viruses from Malaysia and other regions of the world.

Figure 4 shows a dendrogram constructed based on the nucleotide sequences across the E/NS1 gene junction of dengue 2 viruses from Malaysia and other countries in the world [published sequences from Rico-Hesse (1990)]. The Malaysian isolates chosen were representative of each genotype, covering the beginning (1968-1970), middle (1980-1982) and final (1990-1993) periods of the 25 years studied. On the basis of 6% divergence as a cut-off point for genotypic classification, five genotypes could be identified. The groups were generally distributed according to their respective geographic origin: the Western Pacific region (including Thailand), Vietnam, the Indian Ocean region (including Indonesia), Africa and the Americas. The five genotypes obtained in this study were similar to those obtained by Rico-Hesse (1990).

Members of genotype I from Malaysia were grouped together with isolates from Sri Lanka and Indonesia. The Malaysian genotype II iso-

lates, on the other hand, were grouped with isolates from Thailand, the Philippines and Taiwan. The existence in Malaysia of two dengue 2 virus genotypes which were related to those originating from the Indian Ocean and Western Pacific regions, respectively, can easily be explained: Malaysia is sandwiched between the two vast regions, and the frequency and volume of travellers between Malaysia and the two regions are very high. Dengue virus genotypes which evolve from the two regions could therefore reach Malaysia in a relatively short period.

The results obtained in this study have provided valuable information on the genetic relationships and origins of some Malaysian dengue 2 viruses. However, the results could not be used to infer possible groupings which were based on disease severity or virus virulence due to lack of information on disease severity that was produced, particularly by the earlier isolates (1968-1979). Thus, a study to compare limited genome sequences of isolates from clearly defined cases of dengue fever, dengue haemorrhagic fever and dengue shock syndrome should be carried out in the future. Similar studies should also be carried out on dengue 3 viruses because this serotype is actively transmitted in this country. It will be interesting to determine whether Malaysian dengue 3 viruses show geographical distribution and origins similar to those of dengue 2 viruses.

Acknowledgements

We thank Dr. Vincent Deubel (Institut Pasteur, Paris) and Dr. Ingrid Kautner (formerly from the Department of Medical Microbiology, University of Malaya) for their invaluable assistance in the use of the CLUSTAL V programme.

This project was supported by the Ministry of Science, Technology and Environment, Malaysia, and University of Malaya, Kuala Lumpur.

Épidémiologie moléculaire des virus de la dengue 2 de Malaisie isolés de 1968 à 1993

Une approche séquentielle a été utilisée pour l'étude de 24 souches de dengue 2 isolées en Malaisie entre 1968 et 1993. Nous avons déterminé et analysé

les séquences d'une région de 240 nucléotides concernant la jonction génique E/NS1 (enveloppe/protéine 1 non structurale) des souches isolées: les modifications nucléotidiques apparaissent surtout en troisième position d'un codon particulier et sont du type (A↔G, C↔U) de transition. Cinq modifications nucléotidiques se caractérisent par des substitutions d'acides aminés. Des comparaisons par paires de séquences nucléotidiques révèlent des divergences allant de 0 à 9,2%; au niveau des acides aminés les divergences vont de 0 à 2,8%; sur la base de 6% de divergences, on peut grouper les souches en deux génotypes, I et II. La comparaison des séquences nucléotidiques des souches de dengue isolées en Malaisie et des souches originaires d'autres pays montre que les virus ayant le génotype I ou II sont très apparentés à ceux des régions de l'Océan Indien et du Pacifique ouest, respectivement.

Mots-clés: Virus de la dengue 2; Épidémiologie moléculaire, Divergence des séquences nucléotidiques, Dendrogramme, Génotypes, Malaisie.

References

- Chan, S.Y., Kautner, I.M. & Lam, S.K. (1994), The influence of antibody levels in dengue diagnosis by polymerase chain reaction. *J. Virol. Methods*, 49, 315-322.
- Chungue, E., Deubel, V., Cassar, O., Laille, M. & Martin, P.M.V. (1993), Molecular epidemiology of dengue 3 virus and genetic relatedness among dengue 3 strains from patients with mild or severe form of dengue fever in French Polynesia. *J. Gen. Virol.*, 74, 2765-2770.
- Chungue, E., Cassar, O., Drouet, M.T., Guzman, M.G., Laille, M., Rosen, L. & Deubel, V. (1995), Molecular epidemiology of dengue-1 and dengue-4 viruses. *J. Gen. Virol.*, 76, 1877-1884.
- Deubel, V., Kinney, R.M. & Trent, D.W. (1988), Nucleotide sequence and deduced amino acid sequence of the nonstructural proteins of dengue-2 virus, Jamaican genotype. *Virology*, 165, 234-244.
- Deubel, V., Nogueira, R.M., Drouet, M.T., Zeller, H., Reynes, J.M. & Ha, D.Q. (1993), Direct sequencing of genomic cDNA fragments amplified by the polymerase chain reaction for molecular epidemiology of dengue 2 viruses. *Arch. Virol.*, 129, 197-210.
- Fu, J., Tan, B.H., Yap, E.H., Chan, Y.C. & Tan, Y.H. (1992), Full length cDNA sequence of dengue type-1 virus (Singapore strain 3275/90). *Virology*, 188, 953-958.
- Hahn, Y.S., Galler, R., Hunkapillar, T., Dalrymple, J.M., Strauss, J.H. & Strauss, E.G. (1988), Nucleotide sequence of dengue 2 RNA and comparison of the encoded proteins with those of other flaviviruses. *Virology*, 162, 167-180.
- Higgins, D.G., Bleasby, A.J. & Fuchs, R. (1992), CLUSTAL V: improved software for multiple sequence alignment. *Compu. Appl. Biosci.*, 8, 137-139.

- Kautner, I.M. & Lam, S.K. (1992), Direct nucleotide sequencing using total cellular RNA from dengue virus-infected mosquito cells. *Res. Virol.*, 143, 193-197.
- Lanciotti, R.S., Lewis, J.G., Gubler, D.J. & Trent, D.W. (1994), Molecular evolution and epidemiology of dengue-3 viruses. *J. Gen. Virol.*, 75, 65-75.
- Lewis, J.G., Chang, G.J., Lanciotti, R.S. & Trent, D.W. (1992), Direct sequencing of large flavivirus PCR products for analysis of genome variation and molecular epidemiological investigations. *J. Virol. Methods*, 38, 11-24.
- Mackow, E., Makino, Y., Zhao, B.T., Zhang, Y.M., Markoff, L., Buckler-White, A., Guiler, M., Chanock, R. & Lai, C.J. (1987), The nucleotide sequence of dengue type 4 virus: analysis of the gene coding for non-structural proteins. *Virology*, 159, 217-228.
- Osatomi, K. & Sumiyoshi, H. (1990), Complete nucleotide sequence of dengue type 3 virus genome RNA. *Virology*, 176, 643-847.
- Rice, C.M., Lenches, E.M., Eddy, S.R., Shin, S.J., Sheets, R.L. & Strauss, J.H. (1985), Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution. *Science*, 229, 726-733.
- Rico-Hesse, R. (1990), Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology*, 174, 479-493.
- Rudnick, A., Tan, E.E., Lucas, J.K. & Bin Omar, M. (1965), Mosquito-borne haemorrhagic fever in Malaya. *Brit. Med. J.*, 1, 1269-1272.
- Sanger, F., Nicklen, S. & Coulson, A.R. (1977), DNA sequencing with chain-termination inhibitors. *Proc. Nat. Acad. Sci., USA*, 74, 5463-5467.
- Skae, P.M.T. (1902), Dengue fever in Penang. *Brit. Med. J.*, 2, 1581.
- Westaway, E.G., Brinton, M.A., Gaimamovich, S.Y., Hornizek, M.C., Igarashi, A., Kaariainen, L., Lvov, D.K., Porterfield, J.S., Russell, P.K. & Trent, D.W. (1985), Flaviviridae. *Intervirology*, 24, 183-192.