

# Identification of Genetic Variation among Dengue Virus DEN-4 Isolates with Heteroduplex Analysis<sup>+</sup>

by

**Haijun Yao<sup>\*#</sup>, Meiyu Fang<sup>\*\*</sup>, Wenzhong Zhao<sup>\*\*</sup>, Fuqin Duan<sup>\*</sup>, Lihui Lin<sup>\*\*</sup>,  
Cuihua Chen<sup>\*\*</sup> and Huiyu Guo<sup>\*\*\*</sup>**

<sup>\*</sup>Dept. of Molecular Virology, Guangzhou Hygienic & Epidemic Prevention Institute, No.23, 3rd Zhongshan Rd, Guangzhou, 510080, P.R. China.

<sup>\*\*</sup>Guangzhou Military Medical Research Institute, Yantang, Guangzhou 510507, P.R. China.

<sup>\*\*\*</sup>Dept. of Microbiology, Sun Yat-sen University of Medical Sciences, No. 74, 2nd Zhongshan Road, Guangzhou, 510089, P.R. China.

## Abstract

Heteroduplex Analysis (HA) was applied for the identification of genetic variation among five DEN-4 isolates. A 398 bp fragment from NS2a-NS2b region was amplified by reverse transcription/polymerase chain reaction (RT/PCR), and the products were first analysed by HA with different strain as reference strain. In order to confirm the results from HA, each RT/PCR product was also cloned into appropriate vector and were sequenced. HA results showed that the isolates from the DF epidemic in 1990 in southern China shared the same band pattern. The band pattern of the isolate from the DF epidemic in 1978 in southern China was obviously different from the band patterns produced by the 1990 isolates. Sequence results confirmed that the 1990 isolates shared the same sequence, and the sequence of the 1978 isolate was indeed different from the sequence of the 1990 isolates. These results indicated that HA can rapidly identify variations among the dengue viruses, and thus HA could be used as a useful tool in the molecular epidemiological studies of dengue virus.

**Keywords:** Dengue virus, variation analysis, Heteroduplex Analysis (HA).

## Introduction

Dengue virus is a member of *flavivirus*; it is a single-strand positive sense RNA virus; and it has four distinct serotypes, DEN-1, DEN-2, DEN-3, and DEN-4. Its genome is about 11,000 nucleotides in length. The genome

consists of a single open reading frame which encodes a precursor polyprotein. Proteolytic cleavage of the polyprotein results in the formation of the core(C), membrane (M) and envelope (E) proteins, and the nonstructural proteins NS1, NS2a, NS2b, NS3, NS4a, NS4b, and NS5.

+ This study was supported by the China Guangzhou Health Bureau Medical Research Foundation.

# For reprint or correspondence: Dr Haijun Yao, Dept. of Microbiology, Louisiana State University Health Science Center in Shreveport, 1501 Kings Highway, Shreveport, LA 71130, U.S.A. Telephone: (318) 6755759, Fax (318) 6755764, E-mail: yaohaijun@yahoo.com

Dengue viruses are transmitted to humans by the mosquito vectors, like *Aedes aegypti* and *Aedes albopictus*. Approximately 2.5 billion people live in areas at risk for the epidemic transmission of dengue. It can cause dengue fever (DF) and dengue haemorrhagic fever/dengue shock syndrome (DHF/DSS). There are about 100 million cases of DF and 250,000 cases of DHF/DSS occurring annually. In 1978, dengue fever occurred in an epidemic form in southern China after an absence of nearly 40 years. Since then, DF has become a serious health problem in the area. The DF epidemic occurs every 2-3 years in China and the epidemic areas are enlarging, though the endemics are mainly restricted to southern China. However, molecular epidemiological studies of dengue virus in mainland China are limited<sup>(1)</sup>. Therefore, further study of dengue virus' isolates in mainland China is needed to identify genetic characteristics that may influence their epidemiology, virulence and often biological characteristics of the viruses.

Molecular techniques that have been used for the identification of genetic variation among dengue virus strains include oligonucleotides fingerprinting (ONF)<sup>(2)</sup>, primer-extension<sup>(3)</sup>, antigen signature analysis<sup>(4)</sup>, Restriction Fragment Length Polymorphisms (RFLP)<sup>(5)</sup>, Single-Strand Conformation Polymorphism (SSCP)<sup>(6)</sup>, and sequencing<sup>(7)</sup>. Nucleotide sequencing is considered to be the gold standard for the analysis of genetic variation, but because of the high expenses, it is not cost-effective for the study of large numbers of isolates as molecular epidemiological studies generally require analysis of a large number of isolates. Thus, it is difficult to use this technique to

identify and study genetic variation on a large scale.

Heteroduplex Analysis (HA) has been recently used to identify genetic variation in HIV-1<sup>(8)</sup>, influenza virus<sup>(9)</sup>, varicella-zoster virus<sup>(10)</sup>, measles virus<sup>(11)</sup>, and Norwalk-like virus<sup>(12)</sup>. HA has proved to be a sensitive method for the detection of genetic variation in viruses. HA is based on the electrophoretic mobility to its corresponding homoduplex with no mismatch<sup>(13)</sup>. This method was originally developed to detect single-base substitution in PCR products. The present study attempts the use of HA to analyse some DEN-4 isolates collected in southern China. The isolates were first amplified by RT/PCR and detected by HA. Finally, the results were confirmed by sequencing each of the strains.

## Materials and methods

Virus isolates and cell line four DV4 Chinese isolates from the 1990 epidemic in Guangdong and one DV4 Chinese isolate from the 1978 epidemic in Guangdong were chosen for analysis (Table). These isolates were from sera of DF patients using routine methods in C6/36 cell line. Their serotype were determined by Indirect Immuno Fluorescence Assay (IFA). The three isolates from the 1990 epidemic were from DF patients in the early, peak and later part of the epidemic. The GDA63 strain was isolated from *Aedes albopictus* in the early stage of the 1990 epidemic. GD7856B2 was isolated from a DF patient in the 1987 epidemic. DV4 H-241 is the prototype of DV4, which was isolated from a DF patient in the Philippines in 1956.

**Table:** Description of DV4 virus isolates compared by sequence analysis

Strains	Receiving date	Geographical origin	Source	GeneBank Accession No.
GDA63	9/15, 1990	China	<i>Aedes albopictus</i>	Y19171
GD9006A1	9/13, 1990	China	DF patient	Y19172
GD9033A1	9/30, 1990	China	DF patient	Y19173
GD9049A2	10/23, 1990	China	DF patient	Y19174
GD7856B2	1978	China	DF patient	Y19175
DV4 H-241	1956	Philippines	DF patient	Y19176

### **RNA extraction and RT/PCR amplification of viral RNA**

Viral RNA was extracted directly from virus stock as the method described by Yao HJ et al.<sup>(1)</sup>. The primers used in RT and PCR are as follows. Upstream primer: D4S, 5'-CCATTATGGCTGTGTTGTTT-3', 3973nt---3992nt; downstream primer: D4C, 5'-TTCATCCTGCTTCACTTCT-3', 4370nt---4352nt. RT was performed in 20µl of Tris-HCl 50mM, pH8.3, KCl 50mM, MgCl<sub>2</sub> 8mM, dithiothreitol 10mM, 0.5 mM dATP, 0.5 mM dCTP, 0.5mM dGTP, 0.5 mM dTTP, 40 units RNase inhibitor and two units of AMV reverse transcriptase (Promega), 100 ng down stream primer, and 100ng RNA. The reaction was incubated at 42°C for one hour and then at 99°C for 5 minutes to inactivate the AMV reverse transcriptase and degrade the RNA template. Stored at -20°C for further use.

The PCR amplification was done in 50µl reaction volume; it contained 10 mM Tris-HCl, pH 8.4, 50mM KCl, 2 mM MgCl<sub>2</sub>, 10µg gelatin, 0.25mM each dNTP (dATP, dCTP, dGTP, and dTTP), 0.3µg of upstream and downstream primers, 5µl RT reaction

product, preheated at 94°C for 5 min and then added 2.5 units of Pfu DNA Polymerase (Gibco). Amplification was done with the following parameters: denaturation at 94°C 30 s, annealing at 55°C 45 s, and extension at 72°C 45 s, 30 cycles. After the last cycle, samples were maintained at 72°C for five minutes. Take 10 µl of the reaction mixture for electrophoresis in 1% Agarose gel in 1 X TAE containing ethidium bromide (0.5µg/ml).

### **Cloning of the PCR product**

Specific PCR products were purified with GeneClean II kit (HB101, Inc). Then, this product was inserted into PCR 4Blunt-TOPO vector (Invitrogen) according to the protocols in the kit. Briefly, the product was ligated with the pCR 4Blunt-TOPO vector at appropriate ratio according to kit's manual. Take 2 µl ligation mixture for transformation. It was transformed into DH5α electro-competent cell with electrotransfection method (25 µF, 2.5 kV, 200 Ω). Then, plated onto LB plate containing Ampicillin (100 µg/ml).

### Identification of correct clones

Colonies on plates were picked and identified with enzyme digestion (with Eco. RI) and PCR. Heteroduplex Analysis (HA) Mix 5µl specific PCR product with an equal amount of reference isolate's PCR product, mixed well, denatured at 95°C for five minutes, then renatured at 55°C for 15 minutes. For polyacrylamide gel electrophoresis (PAGE), the renatured samples were loaded along with 2µl of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, and 30% glycerol) into a 1.5mm thick, 20cm x 20cm square, 6% nondenaturing polyacrylamide gels containing 1M urea and 1% glycerol in 1 x TBE, 160V for 3 hours. After electrophoresis, the gels were stained with 0.5µg/ml of ethidium bromide in 1xTBE for 40 minutes at 4°C. DNA bands were observed under ultraviolet light.

### DNA sequencing

The nucleic acid sequencing was performed in automatic ABI PRISM 377 DNA sequencer with Bigdye Terminator Cycle Sequencing Ready Reaction Kit (PERKIN ELMER). Protocols used were depicted in the kit's manual.

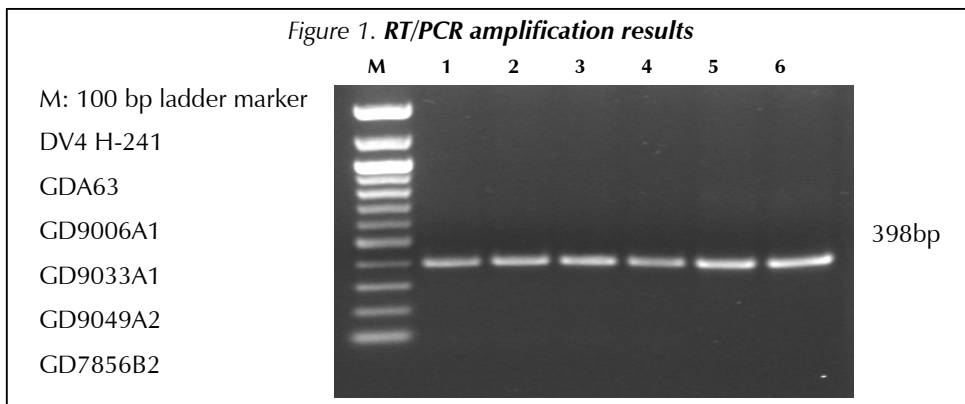
### DNA sequence analysis

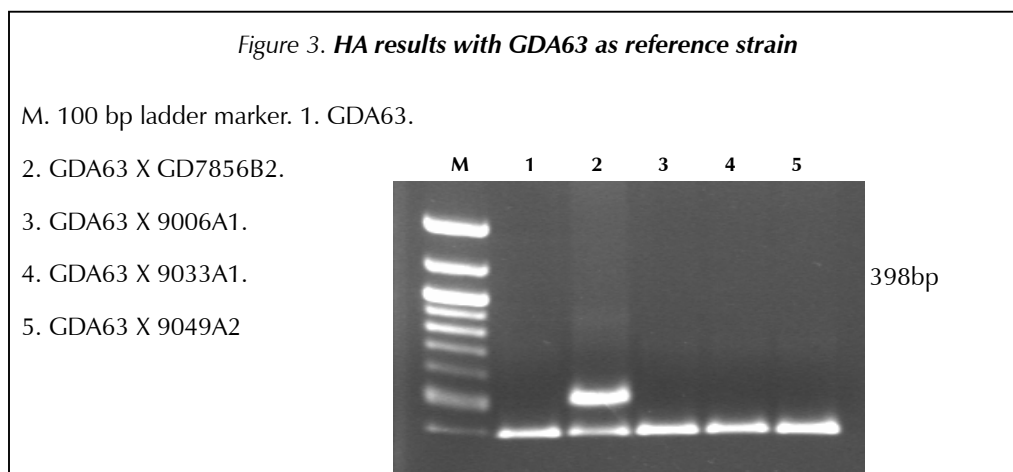
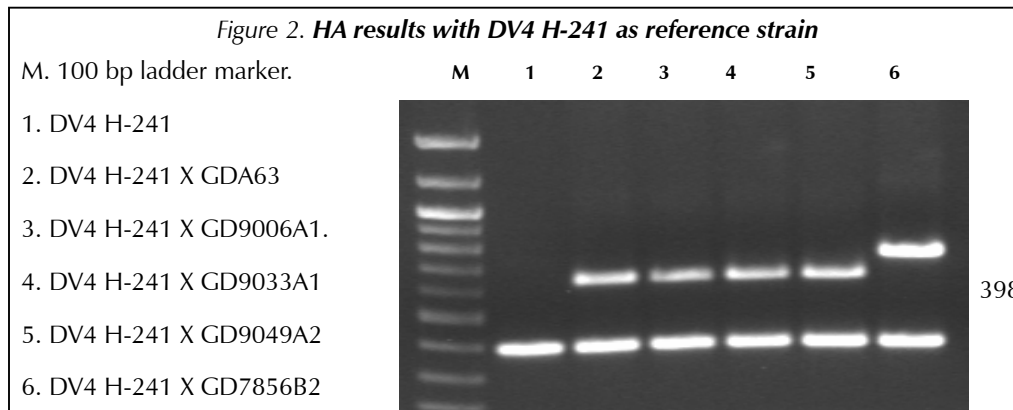
Clustal V programmes (from Baylor College of Medicine) were used to align the above sequences.

### Results

Amplification by RT/PCR A 398 bp PCR product can be obtained in all chosen isolates, as shown in Figure 1.

Analysis of the cDNA by HA. First we compared the local isolates with DV4 prototype (DV4 H-241), as shown in Figure 2. The isolates from the 1990 epidemic shared the same band pattern; the band pattern produced by GD7856B2 was obviously different from that of the isolates from the 1990 epidemic. To further identify the differences between the 1990 isolates and the 1978 isolate, HA analyses of these isolates were done further with GDA63 as reference strain, as shown in Figure 3. Because the band patterns among the isolates from 1990 were the same, they shared the same sequence. However, the band pattern produced by GD7856B2 was different from the patterns produced by the isolates from the 1990 epidemic.





HA has been used recently to identify the genetic variation in some RNA viruses like GBV, HIV-1, Influenza virus, etc. It is an effective tool in the molecular epidemiological studies of these viruses. HA is very rapid, simple to perform, and sensitive. Its size limit is up to 800 bp as compared with SSCP<sup>(14)</sup>. In this study, we tested the potential of HA to identify genetic variation among DEN-4 isolates collected at different times in southern China.

The HA band patterns of the NS2a-NS2b cDNA from the four isolates of the 1990 epidemic showed the same shift pattern; however, the pattern produced by isolate GD7856B2 from the 1978 epidemic was different (Figure 2), and the band patterns obtained with HA were highly reproducible (data not shown). To further check the difference among the isolates from the 1978 and 1990 epidemics, we used GDA63 as the reference strain (Figure 3). The band pattern showed that the

GD7856B2 sequence seemed different from the sequences of the strains from 1990.

The sequence analysis of these strains showed that there were indeed many nucleotides substitutions between the strains from the 1978 and 1990 epidemics. Sequence results also confirmed that the four isolates from 1990 shared the same sequence, and this explained why they had the same band pattern (Figures 2, 3) in HA. Another finding was that they will be seriously retarded if there are more mismatches between the heteroduplex (see Figure 2, 3).

Because GD9006A1, GD9033A1, GD9049A2 and GDA63 shared the same sequence (at least in the sequenced region), this revealed that the epidemic in southern China in 1990 was caused by the same DV4 strain. GDA63 was a strain isolated from *Aedes albopictus* in the epidemic area. Through sequence comparison, we found that it shared the same sequence with the three strains from the patients, which could be an indication of the fact that the virus was transmitted to humans from mosquitoes. Although we did not sequence the whole sequence of the four strains from 1990, it seemed that the dengue viruses were stable in this epidemic. Singh UB et al.<sup>(15)</sup> also sequenced 455 bp E-NS1 region in nine samples of DEN-2 virus isolated in India in 1996, and he found that there were some mutations among these 9 isolates in this region, but were mainly in the third position of a codon. Epidemic strains are relatively stable in one epidemic.

In dengue virus variation studies, the latest method is SSCP<sup>(1,6)</sup>. However, SSCP size limit is about 400 bp; for longer

fragment, it will fail to detect the mutations<sup>(16)</sup>. What is more, the detection rate for G↔C transition is also lower. Conditions for SSCP must be empirically determined because the conformations of the single strand cannot be predicted. Compared with SSCP, HA has a higher size limit (up to 800bp), and it is reproducible. Unlike SSCP, HA does not rely on the secondary structure formation of ssDNA. It relies on heteroduplex formation, thus, the sample's treatment (denature and renature) is simpler than SSCP. One need not worry about quick renaturing of the ssDNA, which sometimes occurs in SSCP.

HA was proven to be a rapid and reliable tool to screen dengue virus isolates and could be used to identify which isolates should be sequenced for further molecular epidemiological studies. The results in this paper showed that HA was a helpful technique for rapidly identifying genetic variation among dengue viruses. Unlike ONF, primer-extension, antigen signature analysis, RFLP, SSCP and sequencing, HA is easy to perform, and the results can be obtained within one day. It does not need advanced equipment and expensive reagents either. Because HA cannot tell the location and number of nucleotide substitution, it cannot be used for phylogenetic studies and cannot replace sequencing. The major advantage of HA is that it can process a large number of samples in one day, and can also differentiate between the sample and the reference strain. Interesting samples selected by HA can be further studied. HA could be used as a screening tool for dengue virus in molecular epidemiological studies.

## References

1. Yao HJ, Chen HS and Cao H. Variation analysis of dengue virus type 4 Chinese isolates by RT/PCR--SSCP. *Academic Journal of the First Medical University of PLA*, 1997, 17(1): 12.
2. Repik PM, Dalrymple JL, Brandt WE, McCown JM and Russel PK. RNA fingerprinting as a method for distinguishing dengue 1 virus strains. *Am J Trop Med Hyg*, 1983, 32: 577-589.
3. Reco-Hesse R. Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology*, 1990, 174(2): 479-493.
4. Yang PY, Si BY, Yan GZ, Xu PF and Li RX. Antigen signature analysis of dengue-2 viruses strains in Hainan China. *Wei Sheng Wu Xue Bao*, 1989, 29(4): 299-302.
5. Vorndam V, Kuno G and Rosado N. A PCR-restriction enzyme technique for detecting dengue virus subgroups within serotypes. *J Virol Methods*, 1994, 48(2-3): 237-244.
6. Farfan JA, Olson KE, Black IV WC, Goubler DJ and Benty BJ. Rapid characterization of genetic diversity among twelve dengue-2 virus isolates by single-strand conformation polymorphism analysis. *Am J Trop Med Hyg*, 1997, 57(4): 416-422.
7. Chungue E, Cassar O and Drouet MT. Molecular epidemiology of dengue-1 and dengue-4 viruses. *J General Virology*, 1995, 76: 1877-1884.
8. Delwart EL, Shpaer EG, Louwagie J, McCutchan FE, Grez M and Rubsamen-Waigmann H Muliss JI. Genetic relationships determined by a DNA heteroduplex mobility assay: analysis of HIV-1 env genes. *Science*, 1993, 262: 1257-1261.
9. Zou S, Sfansfield C and Bridge J. Identification of New influenza B virus variants by multiplex reverse transcriptase-PCR and the heteroduplex mobility assay. *J Clinical Microbiology*, 1998, 26(6): 1544-1548.
10. Muir WB, Nichols R and Brearer J. Phylogenetic analysis of varicella-zoster virus: evidence of intercontinental spread of genotypes and recombination. *J Virology*, 2002, 76(4): 1971-1979.
11. Kreis S and Whistler T. Rapid identification of measles virus strains by the heteroduplex mobility assay. *Virus research*, 1997, 47: 197-203.
12. Mattick KL, Green J, Punia P, Belda FJ, Gallimore CI and Brown DWG. The heteroduplex mobility assay (HMA) as a pre-sequencing screen for Norwalk-like viruses. *J Virological Methods*, 2000, 87: 161-169.
13. White MB, Carvalho M, Derse D, O'Brien S and Dean M. Detecting single base substitution as heteroduplex polymorphisms. *Genomics*, 1992, 12: 301-306.
14. Xing YD, Wells RL and Elkind MM. Nonradioisotopic PCR heteroduplex analysis: a rapid method of detecting minor gene mutations. *Biotechniques*, 1996, 21: 186-187.
15. Singh UB, Mairtra A and Broor S, et al. Partial nucleotide sequencing and Molecular evolution of epidemic causing dengue 2 strains. *J Infectious Diseases*, 1999, 180:959-65.
16. Yao HJ, Cao H and Guo HY. PCR--SSCP analysis and its progress. *Biotechnology*, 1996, 6(4): 1-8.