Geographic Distribution of Hepatitis C Virus Genotype 6 Subtypes in Thailand

Srunthron Akkarathamrongsin,1 Kesmanee Praianantathavorn,1 Nisachol Hacharoen,1 Apiradee Theamboonlers,1 Pisit Tangkijvanich,2 Yasuhito Tanaka,3 Masashi Mizokami,3,4 and Yong Poovorawan1*

1Faculty of Medicine, Center of Excellence in Clinical Virology, Chulalongkorn University, Bangkok, Thailand
2Faculty of Medicine, Department of Biochemistry, Chulalongkorn University, Bangkok, Thailand
3Department of Clinical Molecular Informatice Medicine, Nagoya City University Graduate School of Medical Sciences, Kavaasumi, Mizuho, Nagoya, Japan
4Research Center for Hepatitis and Immunology, International Medical Center of Japan Kounodai Hospital, Kounodai, Ichikawa, Japan

The nucleotide sequence of hepatitis C virus (HCV) genotype 6 found mostly in south China and south-east Asia, displays profound genetic diversity. The aim of this study to determine the genetic variability of HCV genotype 6 (HCV-6) in Thailand and locate the subtype distribution of genotype 6 in various geographic areas. Four hundred nineteen anti-HCV positive serum samples were collected from patients residing in the central part of the country. HCV RNA positive samples based on reverse transcriptase-polymerase chain reaction (RT-PCR) of the 5'UTR were amplified with primers specific for the core and NS5B regions. Nucleotide sequences of both regions were analyzed for the genotype by phylogenetic analysis. To determine geographic distribution of HCV-6 subtypes, a search of the international database on subtype distribution in the respective countries was conducted. Among 375 HCV RNA positive samples, 71 had HCV-6 based on phylogenetic analysis of partial core and NS5B regions. The subtype distribution in order of predominance was 6f (56%), 6n (22%), 6i (11%), 6j (10%), and 6e (1%). Among the 13 countries with different subtypes of HCV-6, most sequences have been reported from Vietnam. Subtype 6f was found exclusively in Thailand where five distinct HCV-6 subtypes are circulating. HCV-6, which is endemic in south China and south-east Asia, displays profound genetic diversity and may have evolved over a considerable period of time.

INTRODUCTION

Hepatitis C virus infection is a worldwide public health problem, with an estimated 170 million chronic infections globally. Although acute presentation of HCV infection is generally mild and asymptomatic, about 80% of infected individuals develop chronic infection and progress to cirrhosis and hepatocellular carcinoma [Hoofnagle, 2002]. HCV is a single-stranded positive-sense RNA virus of the family Flaviviridae. The viral genome is approximately 9.6 kb in length, flanked by 5' and 3'UTR, and encodes a polyprotein precursor of about 3,000 amino acids. The precursor is cleaved into at least 10 different proteins comprising the structural proteins, core, E1, E2, and p7 as well as the non-structural proteins, NS2, NS3, NS4A, NS4B, NS5A, and NS5B [Choo et al., 1991; Moradpour et al., 2007].

HCV has been classified into six major genotypes and numerous subtypes [Simmonds et al., 1993, 2005]. A new HCV genotype 7a isolate has also been described [Murphy et al., 2007]. Direct sequencing of the NS5B, the core, and the envelope regions has proven a reliable method for classification of HCV genotypes [Simmonds et al., 2005]. Some genotypes (genotypes 1, 2, and 3) are distributed globally, while others are endemic in different geographically restricted areas [Simmonds, 2004]. In south-east Asia, genotype 3a is the most common genotype followed by 1b and 6 variants [Mellor et al., 1996; Kanistanon et al., 1997; Sunanchaikarn et al., 2007].
HCV genotype 6 (HCV-6) distributed primarily in south China and south-east Asia displays pronounced genetic diversity. Novel subtypes of genotype 6 have been discovered continuously and now comprise 22 subtypes, 6a–6v [Lu et al., 2008; Nopporpanth et al., 2008; Xia et al., 2008a,b; Wang et al., 2009]. These may represent the evolution of this genotype for a considerable time span that exceeds probably 1,000 years [Pybus et al., 2009]. Accumulation of nucleotide mutations may drive continual development of HCV genetic diversity. This genotype variability appears to occur at a higher frequency in Vietnam and Myanmar than in Thailand [Lwin et al., 2007; Sunanchaikarn et al., 2007].

In Thailand, approximately 2.15% of the general population has been infected chronically with HCV. The common genotypes are 3a, 1b, and 6 variants [Sunanchaikarn et al., 2007]. Epidemiological data on the subtypes of HCV-6 are still limited. To determine the genetic variability of these variants, Thai patients infected with HCV have been recruited and investigated emphasis on the diversity of genotype 6. The geographic distribution of the predominant subtypes in south China and south-east Asia has been described in the cohort study.

**MATERIALS AND METHODS**

The study protocol was approved by the Ethics Committee of the Faculty of Medicine, Chulalongkorn University. The anti-HCV positive samples were chosen from blood specimens obtained for screening or investigation and treatment. The specimens were treated as anonymous with a coding number. In addition, all specimens were used exclusively for academic research.

**Sample Collection**

Serum samples were collected from Thai individuals resident mainly in the central part of the country. Each specimen was subjected to enzyme-linked immunosorbent assay (ELISA) for detection of anti-HCV using Abbott, North Chicago, IL) according to the manufacturer’s instructions. HCV antibody positive samples were chosen for further testing. All specimens were stored at –70°C until used.

**RT-PCR Amplification**

Total RNA was extracted from 419 anti-HCV positive sera by the guanidinium method [Theamboonlers et al., 2002] and reverse transcribed into cDNA using random primers (Promega, Madison, WI) and M-MLV reverse transcriptase (RT; Promega). For HCV RNA detection, the DNA fragment of the 5′UTR was amplified. Viral RNA positive samples were selected for amplification of core and NS5B regions. Nested polymerase chain reaction (PCR) amplification of the 5′UTR and the core region was performed as described previously [Mellor et al., 1995; Theamboonlers et al., 2002; Sunanchaikarn et al., 2007]. The NS5B region was amplified by nested PCR resulting in a 471-bp fragment, with outer primers, NS5B_F1 (CAATWSMACBACCATCATGCG, positions 7999–8020), NS5B_R1 (CCAGGARTTRCTG-GAGGTGT, positions 8805–8825); and inner primers, NS5B_F2 (GATGGGHHBSKCMATYGGATCC, positions 8159–8181), NS5B_R2 (CATAGCNTCCGTGANGCTC, positions 8611–8630) (nucleotide numbering is according to reference strain H77, GenBank accession number AF00906). First round PCR was performed by mixing 3 µl of cDNA to a final volume of 25 µl PCR reaction mixture containing 5 pmol each of NS5BF1 and NS5BR1 primers, 200 µM dNTP, 1.5 mM Mg2+, and 1.25 U of Taq DNA polymerase. Samples were amplified under the following conditions: 3 min at 94°C for initial denaturation followed by 40 cycles at 94°C for 1 min, 49°C for 1 min, and 72°C for 1.30 min, and concluded by a final extension step at 72°C for 7 min. For the second round, 0.5 µl of first round PCR product were added to a reaction mixture identical to the first round, except for the inner primer set, NS5BF2 and NS5BR2 (5 pmol each). The amplification reaction was performed under identical conditions to the first round, except for increasing the melting temperature for primer annealing to 52°C. The PCR products were analyzed by electrophoresis in 2% agarose gel stained with ethidium bromide and subsequently visualized under UV light.

**Sequencing**

After gel purification (Perfectprep Gel Cleanup Kit, Eppendorf, Hamburg, Germany), the amplicons from both core and NS5B were subjected to sequencing as described elsewhere [Sunanchaikarn et al., 2007]. The sequences were edited and assembled using Chromas LITE (v.2.01) (www.technelysium.com.au), BioEdit (v.5.0.9) (Ibis Therapeutics, Carlsbad, CA), and SeqMan (DNASTAR, Madison, WI). Sequence similarities between the sequences generated in this study were examined by the BLASTN program (http://www.ncbi.nlm.nih.gov).

**Genotyping and Phylogenetic Analysis**

The edited sequences were analyzed for HCV genotypes using the Viral Genotyping Tool (http://www.ncbi.nlm.nih.gov) and samples designated to genotype 6 were subjected to further analysis. Multiple sequence alignments were performed with ClustalX (v.1.83). Phylogenetic trees of HCV-6 based on both core and NS5B sequences were constructed. Neighbor-joining trees were constructed with the 6-Parameter method and bootstrapped 1,000 times to confirm the reliability of the phylogenetic tree (Shin-I et al., 2008). The reference sequences of HCV-6 variants were retrieved from GenBank, EMBL, and DDBJ as follows: genotype 1a (M62321), 2a (D00944), 3a (D17763), 4a (Y11604), 5a (Y13184), 6a (EU246930, Y12083), 6b (D37841, D84262), 6c (EF424629), 6d (D84263), 6e (DQ814805), 6f (D37846, D37858, D37859, D37860, D37861, D38078, D38079, DQ835760, DQ835764), 6g...
Phylogenetic Analysis of Core and NS5B Regions

All HCV sequences determined as genotype 6 were subjected to further amplification of core and NS5B regions followed by nucleotide sequencing. The viral genotypes were investigated employing a viral genotyping tool (http://www.ncbi.nlm.nih.gov). Seventy-one samples determined as genotype 6 were subjected subsequently to phylogenetic analysis. The nucleotide sequences determined as genotype 6 were submitted to the GenBank database under designated accession numbers FJ859193, FJ859334.

RESULTS

HCV RNA Detection and Sequencing

Of the 419 anti-HCV positive samples, 375 were positive for viral RNA by RT-PCR amplification of the 5’UTR. These samples were subjected to further amplification of core and NS5B regions followed by nucleotide sequencing. The viral genotypes were investigated employing a viral genotyping tool (http://www.ncbi.nlm.nih.gov). Seventy-one samples determined as genotype 6 were subjected subsequently to phylogenetic analysis. The nucleotide sequences determined in this study were submitted to the GenBank database under designated accession numbers FJ859193–FJ859334.

Phylogenetic Analysis of Core and NS5B Regions

All HCV sequences determined as genotype 6 were confirmed by phylogenetic analysis based on core and NS5B alignments and classified into five specific subtypes. Neighbor-joining trees of the core and NS5B sequences were constructed (Fig. 1) and the sequences clustering closely with the same clade of reference strains were classified as the corresponding genotypes. According to these results, subtypes 6f and 6n are predominant in Thailand, followed by 6i, 6j, and 6e (Table I).

Worldwide Distribution of HCV-6

All HCV-6 strains stored at the GenBank, EMBL, and DDBJ were included in this study. A total of 820 sequences of genotype 6 reported from 13 countries were examined for distribution of subtypes of HCV-6. Most of the reported sequences were from Vietnam (n = 231), followed by China (n = 205) and Thailand (n = 141) (Fig. 2). HCV-6 has not only been isolated in south China and south-east Asia but also in western countries, such as Canada (n = 64), the United States (n = 5), and France (n = 1). According to these databases, Canada shows the highest diversity of genotype 6 with at least 11 subtypes identified and classified as 6a, 6c, 6e, 6f, 6h, 6l, 6o, 6p, 6q, 6r, and 6s.

Recently, a novel subtype, 6v has been identified in Kunming, China [Wang et al., 2009]. HCV-6 exhibits extreme diversity with 22 subtypes identified so far (Fig. 2). Of these 22 subtypes, 6a is found mainly in China and Vietnam. Subtypes 6b and 6c are found predominantly in Thailand. Subtypes 6d and 6t are found only in Vietnam. Subtypes 6e, 6o, and 6p are reported frequently from Cambodia, Thailand, Taiwan, Myanmar, and India stored at the databases have been included. This genotype has been identified in western countries, such as Canada, France, and the United States. Geographic distribution of HCV-6 in different countries was investigated in this study.

DISCUSSION

Phylogenetic analysis based on full-length genome sequences serves as the “gold standard” for HCV genotype or subtype identification. Although this constitutes the most accurate method for virus classification, due to cost and time factors, it is not practical as a routine clinical investigation. Hence, DNA sequencing of subgenomic regions such as the core, the envelope, and NS5B regions is preferred [Simmonds et al., 1993, 2005; Robertson et al., 1998]. Since the 5’UTR is conserved and thus unsuitable for genotype determination, this region has been employed for viral RNA detection. As recombination among HCV strains has emerged [Kalina et al., 2002; Kageyama et al., 2006; Moreau et al., 2006; Noppornpanth et al., 2006; Legrand-Abravanel et al., 2007; Kurbanov et al., 2008], genotyping based on one region may not be effective. To improve accuracy, this study examined the viral genotype by phylogenetic analysis based on the core and NS5B sequences.

Data were collected from the GenBank, EMBL, and DDBJ nucleotide sequence databases revealed that there are at least seven subtypes of HCV-6 circulating in Thailand (Fig. 2). In the course of this project, only five subtypes, 6e, 6f, 6i, 6j, and 6n could be identified. Subtypes 6b, 6c, and 6m have also been reported from Thailand, but they were not detected in the cohort study. According to previous studies, the target groups were blood donors and intravenous drug users from the northern part of Thailand [Apichartpiyakul et al., 1999; Thaikruuea et al., 2004]. Since the geographic
location of the respective studies may have an influence on the genotype 6 subtype distributions, samples for this project were collected from Thai people residing in the central area. In correlation with the database, the predominant subtype of genotype 6 was 6f, followed by 6n. Information relating to the southern part of the country is still limited and reports on genotype 6 strains have been very rare, in contrast to the data on HCV genotype distribution in the north and center of the country [Sunanchaikarn et al., 2007]. Similarly, genotype 6 has as yet not been reported from Malaysia, the country neighboring Thailand to the south [Mellor et al., 1996]. On the other hand, circulation of genotype 6 has been detected in western countries such as Canada. It appears that distribution of genotype 6 is not restricted to south-east Asia, but tends to spread to other regions of the world. Canada also has the highest diversity of HCV-6. Accumulation of Asian immigrants from endemic areas genotype 6 may have contributed to this phenomenon [Murphy et al., 2007].

This study constitutes the first report on the diversity of HCV-6 and its geographic distribution in Asia. One limitation of this study is that it does not represent the whole population of Thailand. In endemic areas, subtypes of HCV-6 are not distributed evenly, but each

TABLE I. Hepatitis C Virus Genotype 6 Subtypes Determined by Phylogenetic Analysis Based on Core and NS5B Regions

<table>
<thead>
<tr>
<th>Regions</th>
<th>6e (1)</th>
<th>6f (56)</th>
<th>6i (11)</th>
<th>6j (10)</th>
<th>6n (22)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>71</td>
</tr>
<tr>
<td>NS5B</td>
<td>1 (1)</td>
<td>39 (56)</td>
<td>8 (11)</td>
<td>7 (10)</td>
<td>16 (22)</td>
<td>71</td>
</tr>
</tbody>
</table>

subtypes reported from China may have resulted from trans-migration between the countries near the Mae Khong River and The Golden Triangle [Lu et al., 2006; Xia et al., 2008b; Wang et al., 2009]. A high prevalence of intravenous drug abuse has been reported from this area, which may play a role in HCV transmission [Xia et al., 2008a]. The data on HCV-6 in Myanmar, Laos, and Cambodia are still not well advanced.

The transmission process, factors that contribute to the considerable diversity and the manner of the viral spread have been puzzling. The most recent study has reported various subtypes of HCV-6 from Laos [Pybus et al., 2009]. The results of the current study have also confirmed that this genotype is highly variable and endemic in south-east Asia. There is strong evidence that HCV strains from the same country have clustered on the same branch of the phylogenetic tree. Even though HCV has been isolated from other regions, phylogenetic analysis has demonstrated that the subtypes of genotype 6 originated from Asian lineage. Based on the molecular clock and coalescence theory, Pybus et al. [2009] demonstrated that HCV-6 has evolved for more than 1,000 years, and its epidemic proliferation has been established during the early 20th century. This study has investigated genotype 6 from several countries, which may have been subjected to differences in evolutionary rate. Each individual country has experienced a particular historic pattern of transmission, such as civil war, the onset of intravenous drug abuse, the use of unsterile medical service, and healthcare infrastructure, which may be reflected in the different epidemic behavior of HCV infection [Tanaka et al., 2006].

It is concluded that five subtypes of HCV-6 are circulating in Central Thailand while subtype 6f was found exclusively in Thailand. HCV-6 is endemic in south China and south-east Asia displays profound genetic diversity and may have evolved over a considerable period of time, while the mechanism underlying the endemic process and the spread of the virus remain to be elucidated.

ACKNOWLEDGMENTS

We would like to express our gratitude to the Royal Golden Jubilee program of the Thailand Research Fund, Chulalongkorn Hospital, the Commission on Higher Education, the Biomedical Sciences Program of Graduated School of Chulalongkorn University, the Center of Excellence in Clinical Virology, and CU Centenary Academic Development Project for their generous contributions to this research project. We would also like to thank Siriraj, Srinagarind, and Petchaboon Hospitals for providing the specimens, and Ms. P. Hirsch for editing the manuscript.

REFERENCES


![Fig. 2. Geographic distribution of HCV genotype 6 in south China and south-east Asia. Pie charts indicate the proportion of all HCV subtype 6 presents in each country. Pie chart at left bottom shows all HCV subtype 6 of Thailand detected in this study. The number of samples included in this study is indicated. The data have been obtained from GenBank, EMBL, and DDBJ nucleotide sequence databases.](image-url)


