

## Short Communication

# Complete genomes of hepatitis C virus (HCV) subtypes 6c, 6l, 6o, 6p and 6q: completion of a full panel of genomes for HCV genotype 6

Ling Lu,<sup>1</sup> Chunhua Li,<sup>1,2</sup> Yongshui Fu,<sup>3</sup> Fengxiang Gao,<sup>4</sup> Oliver G. Pybus,<sup>5</sup> Kenji Abe,<sup>6</sup> Hiroaki Okamoto,<sup>7</sup> Curt H. Hagedorn<sup>1</sup> and Donald Murphy<sup>8</sup>

### Correspondence

Ling Lu  
llu@kumc.edu

Donald Murphy  
Donald.Murphy@inspq.qc.ca

<sup>1</sup>Division of Gastroenterology/Hepatology, Department of Medicine, University of Kansas Medical Center, 4035 Delp, MS 1023, Kansas City, KS 66160, USA

<sup>2</sup>The First People's Hospital of Yunnan Province, Kunming, China

<sup>3</sup>Gunagzhou Blood Center, Guangzhou City, Guangdong Province, China

<sup>4</sup>Public Health Laboratories, New Hampshire Department of Health and Human Services, Concord, NH, USA

<sup>5</sup>Department of Zoology, Oxford University, Oxford, UK

<sup>6</sup>Department of Pathology, National Institute of Infectious Diseases, Tokyo, Japan

<sup>7</sup>Division of Virology, Department of Infection and Immunity, Jichi Medical University School of Medicine, Tochigi-Ken, Japan

<sup>8</sup>Institut National de Santé Publique du Quebec, Laboratoire de Santé Publique du Quebec, Sainte-Anne-de-Bellevue, Quebec, Canada

Five hepatitis C virus (HCV) complete genome sequences (Th846, 537796, QC227, QC216 and QC99) from a blood donor in Thailand and three Asian immigrants and one Caucasian in North America were determined. Phylogenetically, they represent the first complete genomes for subtypes 6c, 6l, 6o, 6p and 6q, respectively. Similarity analysis showed no evidence of inter- or intrasubtype recombination. Further analysis in conjunction with partial sequences from the Los Alamos HCV database led to the identification of other closely related isolates from south-eastern Asia or immigrants from that region. However, Th846 did not cluster with any reference sequence and is the sole isolate of subtype 6c reported so far. This study completes the full genome sequencing of all 17 assigned HCV genotype 6 subtypes (6a–6q). The utility of this panel of complete sequences for accurate detection and classification of infection, and for estimating the origin of this genotype of HCV, is discussed.

Received 24 December 2006

Accepted 7 February 2007

*Hepatitis C virus* (HCV) is classified in the genus *Hepacivirus* of the family *Flaviviridae*. It has a single-stranded, positive-sense RNA genome of approximately 9600 nt. The genome contains a single open reading frame (ORF) that spans nearly the complete length of the genome. The ORF encodes three structural (core, E1 and E2) and seven non-structural (P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B) proteins (Major & Feinstone, 1997). Phylogenetic analysis has classified HCV into six genotypes and, within each genotype, closely related variants are

grouped into a number of subtypes (Robertson *et al.*, 1998). In a recent HCV nomenclature proposal, up to 71 subtypes have been classified; within genotype 6, 17 subtypes in total (6a–6q) have been assigned (Simmonds *et al.*, 2005). Complete genome sequences represent the 'gold standard' for HCV classification and molecular epidemiology; they are vital for the identification of putative recombinant strains and can be compared and analysed with any set of subgenomic sequences.

Geographically, genotype 6 variants have been identified exclusively in south-eastern Asia or immigrants from this region (Apichartpiyakul, *et al.*, 1994; Bernier *et al.*, 1996; Doi *et al.*, 1996; Mellor *et al.*, 1995, 1996; Noppornpanth *et al.*, 2006a; Shinji *et al.*, 2004; Stuyver *et al.*, 1995; Thaikrua *et al.*, 2004; Theamboonlers *et al.*, 2002). In previous studies, complete genomic sequences representing 12 subtypes within genotype 6 have been isolated from

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this study are EF424625–EF424629.

A table showing pairwise nucleotide similarities of the five HCV complete genomic sequences with other genotype 6 subtypes and a diagram outlining the strategy used to amplify the five complete HCV genomic sequences are available as supplementary material in JGV Online.

samples obtained from this region (Tokita *et al.*, 1998; Simmonds *et al.*, 2005; Li *et al.*, 2006; Lu *et al.*, 2006a, 2007). However, five other officially declared subtypes (6c, 6l, 6o, 6p and 6q) have not yet had their genomes sequenced fully. Here, we report and analyse the complete genome sequences of these five subtypes.

Five samples were used in this study. Th846 was from a commercial blood donor in Thailand and represents a subtype 6c infection (Tokita *et al.*, 1995). Sample 537796 was from an Asian immigrant in the USA and represents subtype 6l. This isolate was identified in the US Third National Health and Nutrition Examination Survey (NHANES III) (Nainan *et al.*, 2006). Three samples (QC227, QC216 and QC99) represent subtypes 6o, 6p and 6q, respectively. QC227 was obtained from a Caucasian individual who was infected following a blood transfusion in Canada in the mid-1980s, whilst QC216 and QC99 were respectively from a Vietnamese and a Cambodian immigrant in Quebec, Canada.

From 100 µl of each of these samples, complete HCV genomes were amplified by using strategies listed in Supplementary Fig. S1 (available in JGV Online). The amplicons were sequenced directly and sequence information was analysed as described elsewhere (Lu *et al.*, 2007). The lengths and genomic organizations of the resultant five genomes are described in Table 1. Pairwise comparison of the five sequences with 14 genotype 6 reference sequences revealed that they have nucleotide similarities of 72.6–81.5% (mean, 75.3%) over the entire genome. The highest sequence similarities were observed between isolates 537796 (6l), VN405 (6k) and km41 (6?). The lowest similarities were found when the five sequences were compared with those of subtypes 6a and 6b. Similarities were the highest in the core (mean, 87.4%) and NS5B (mean, 78.2%) genes and lowest in the P7 (mean, 65.8%) and NS2 (mean, 68.2%) regions (see Supplementary Table S1, available in JGV Online). According to the criterion that HCV subtypes are defined by nucleotide similarities of <75–80% over the entire genome (Simmonds *et al.*, 1994), the similarities calculated here indicate that these five sequences represent five distinct HCV subtypes.

Phylogenetic analysis demonstrated that the five isolates were all classified into genotype 6. These strains therefore

complete a panel of 17 subtypes with determined complete genomes, designated 6a–6q within genotype 6 (Simmonds *et al.*, 2005). Novel isolates km41 and gz52557, for which the entire genome sequences are also available, may represent two novel subtypes (Lu *et al.*, 2006a). In the phylogenetic tree, each isolate is represented by a separate branch, forming a highly diverse genotype 6 clade (Fig. 1a). Within this clade, four subclades can be determined, with each subclade having a bootstrap value of 100%. The first subclade contains subtypes 6a and 6b; the second consists of 6c, 6d, 6e, 6f, 6o, 6p and 6q; the third includes 6g and the unassigned isolate gz52557; and the fourth contains 6h, 6i, 6j, 6k, 6l, 6m, 6n and the unassigned km41. The second subclade can be further split into three pairs (6c/6d, 6e/6q and 6o/6p) and a single branch, 6f. The fourth subclade can be divided into three groups (6h/6i/6j, 6k/km41/6l and 6m/6n), with each group having a bootstrap value of 100%.

A second phylogeny was reconstructed by using the deduced amino acid sequences of the isolates. The phylogeny had an identical genotype 6 topology, indicating the robustness of this branching pattern (Fig. 1b). Nucleotide sequences from ten protein-encoding regions were analysed separately, which also showed that the 19 isolates were consistently distinct (data not shown). To identify potentially recombinant isolates, pairwise similarity scores (i.e. diversity plots) were calculated along the HCV genome between the five isolates and 14 reference genotype 6 sequences. Each genome showed similar patterns of genetic similarity to other isolates along the genome and showed no evidence of recent recombination events (data not shown).

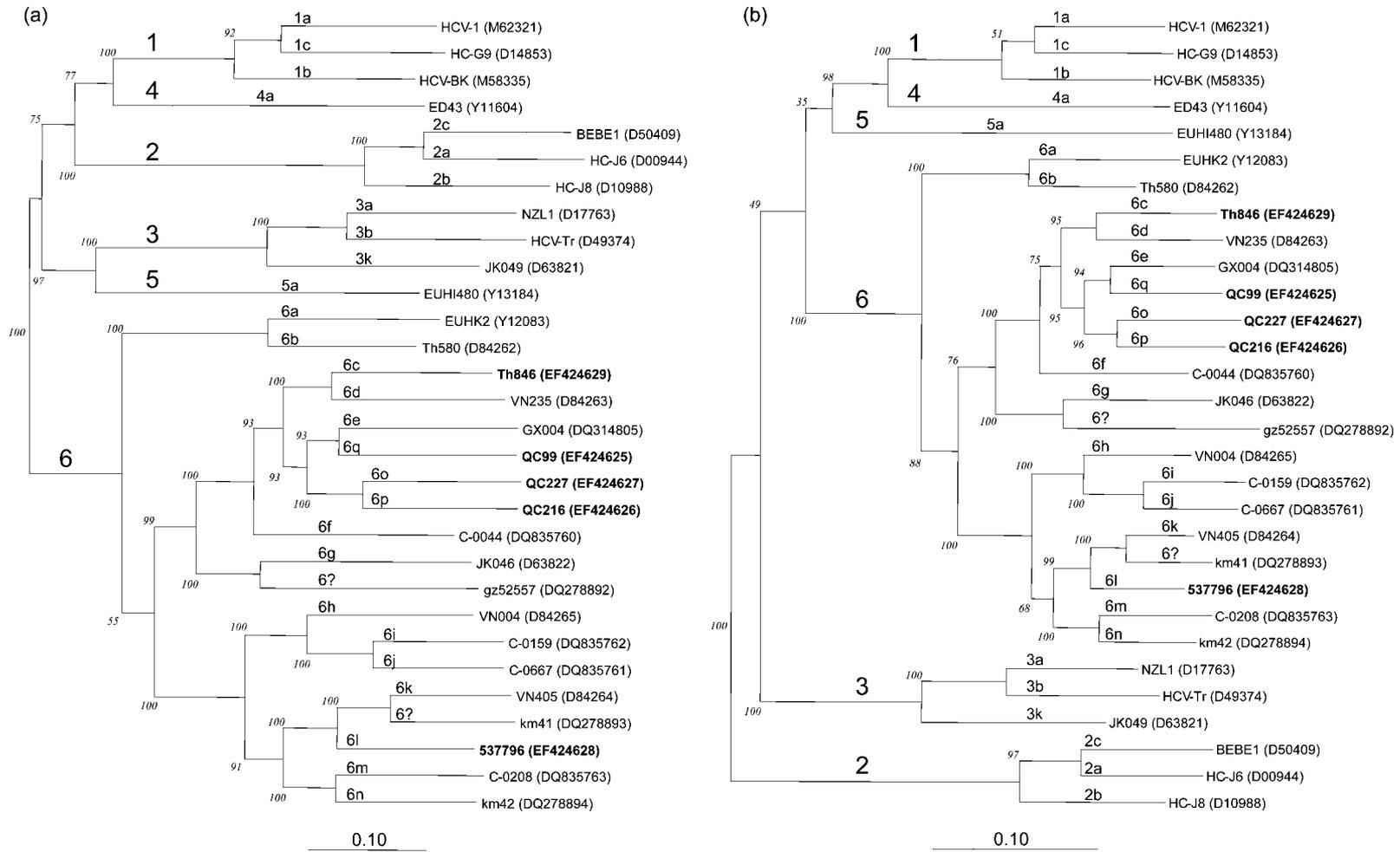
To relate subtypes 6c, 6l, 6o, 6p and 6q to HCV variants characterized by partial sequences, three short, subgenomic regions of the five new isolates were analysed alongside other sequences retrieved from GenBank. The subgenomic segments were (i) 350 nt of the core gene, (ii) 424 nt of the E1 gene, and (iii) 329 nt of the NS5B gene, corresponding to positions 340–689, 867–1290 and 8295–8623 of the H77 genome (GenBank accession no. NC\_004102), respectively. Phylogenetic trees estimated from these three segments

**Table 1.** Genomic organization of the five HCV isolates

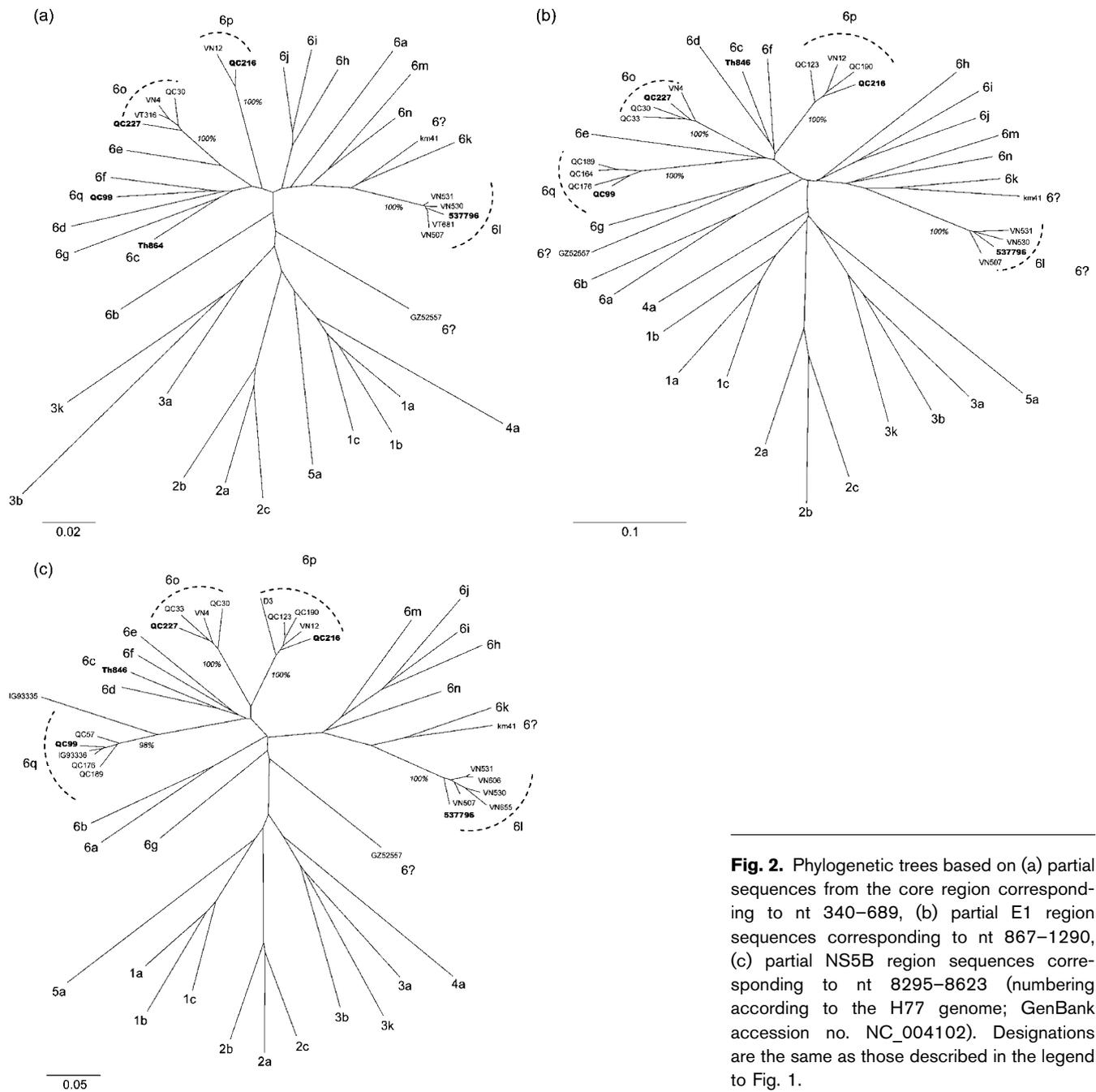
Bold type indicates regions varying in length between isolates.

Isolate	Amplicons (n)*	Length (nt)														
		Genome	ORF	5' NCR	Core	E1	E2	P7	NS2	NS3	NS4A	NS4B	NS5A	NS5B	3' NCR	Poly(U)
Th846/6c	25	9459	9048	339	573	576	<b>1089</b>	189	651	1893	162	783	<b>1356</b>	1776	72	24
537796/6l	22	9453	9051	338	573	576	<b>1095</b>	189	651	1893	162	783	<b>1353</b>	1776	64	28
QC227/6o	34	9450	9051	338	573	576	<b>1092</b>	189	651	1893	162	783	<b>1356</b>	1776	61	28
QC216/6p	31	9453	9051	339	573	576	<b>1092</b>	189	651	1893	162	783	<b>1356</b>	1776	63	31
QC99/6q	19	9463	9051	339	573	576	<b>1092</b>	189	651	1893	162	783	<b>1353</b>	1776	76	28

\*Number of overlapping fragments covering the complete genomic sequence.



**Fig. 1.** Phylogenetic tree based on (a) complete nucleotide sequences and (b) deduced amino acid sequences. The six HCV genotypes are indicated by numbers 1–6; subtypes are designated 1a–6q and 6? (subtype unassigned). Reference HCV sequences are indicated by an isolate name followed by a GenBank accession number in parentheses. The five HCV variants sequenced completely in this study are shown in bold. Bootstrap analysis values are shown in italics. Bars, 0.10 nucleotide substitutions per site.



**Fig. 2.** Phylogenetic trees based on (a) partial sequences from the core region corresponding to nt 340–689, (b) partial E1 region sequences corresponding to nt 867–1290, (c) partial NS5B region sequences corresponding to nt 8295–8623 (numbering according to the H77 genome; GenBank accession no. NC\_004102). Designations are the same as those described in the legend to Fig. 1.

demonstrated that Th846 represents the sole published representative of subtype 6c, whereas 537796, QC227, QC216 and QC99 clustered with six, four, four and five other sequences, respectively (Fig. 2). Among these other sequences, six 6l (VN507, VN530, VN531, VN606, VN655 and VT681), two 6o (VN4 and VT316) and two 6p (D3 and VN12) sequences had origins in Vietnam (Noppornpanth *et al.*, 2006b; Stuyver *et al.*, 1995; Tokita *et al.*, 1994). Two 6o (QC30 and QC33), two 6p (QC123 and QC190) and four 6q (QC57, QC164, QC176 and QC189) sequences were isolated in Vietnamese or Cambodian immigrants in Quebec, Canada (Bernier *et al.*, 1996; Murphy *et al.*, 2003). Another 6q sequence (IG93336) was recorded from Laos,

but detailed information on this isolate has not been published. Bootstrap values for these clusters varied from 98 to 100 %.

Several studies have suggested that unique and highly divergent variants of HCV genotype 6 exist in south-eastern Asia. In one study, two partial HCV sequences (1582 and 1142 nt long; GenBank accession numbers D37843 and D37857, respectively) were determined for strain Th846, which was isolated from a commercial blood donor in Thailand. Analysis of these two fragments initially led to the classification of Th846 as a novel genotype 7d virus (Tokita *et al.*, 1995). After reclassification of

genotypes 7, 8 and 9 as subtypes within genotype 6, this isolate (Th846) was assigned to subtype 6c (Simmonds *et al.*, 1996). The latter designation has now been used in the consensus proposal for HCV nomenclature (Simmonds *et al.*, 2005). Although several studies of HCV infection among blood donors in Thailand and other south-eastern Asian countries have been performed, no other 6c isolate has been reported to date (Thaikruea *et al.*, 2004), suggesting that Th846 represents a rare HCV variant. In this regard, Th846 is comparable to gz52557, which is also not related closely to other genotype 6 variants (Lu *et al.*, 2006a). As the prevalence and epidemiological features of these strains are not well understood, sequencing their entire genomes may provide important new insights regarding the evolution and spread of genotype 6 variants. In addition, characterization of the full range of global HCV diversity is essential for the development and implementation of automated virus-genotyping tools (de Oliveira *et al.*, 2005).

In the US NHANES III study, partial sequences of the core, E1 and NS5B regions were determined for HCV isolate 537796, obtained from an Asian immigrant (Nainan *et al.*, 2006). Initial phylogenetic analysis classified this isolate as subtype 6l. Here, we obtained the entire genome sequence for this isolate. Although isolate 537796 is distinct from subtype 6k and the unassigned km41, it is related to these strains (genome similarities of 81.0 and 81.5%, respectively). These similarities are comparable to those between subtypes 6i and 6j, and between subtypes 6m and 6n (genome similarities of 81.3–82.7%). Furthermore, subtypes 6k, 6l and the unassigned km41 form a clade of closely related subtypes (Fig. 1), suggesting that this genotype 6 lineage has been present and circulating in Vietnam for a considerable period of time. In support of this, a search of the Los Alamos HCV database identified two subtype 6k (VN405 and D33) and six subtype 6l (VN507, VN530, VN531, VN606, VN655 and VT681) isolates. All 6k and 6l isolates were sampled exclusively from Ho Chi Minh City, Vietnam (Tokita *et al.*, 1994), except for VN405, which was from Hanoi. Although isolates km41 and km45 were isolated in China, they were from the Kunming region, bordering Vietnam. The isolation of strain 537796 from an Asian immigrant suggests that it might also originate from Vietnam, but this information was not reported in the NHANES III study (Nainan *et al.*, 2006).

Previously, partial sequences have been determined for a number of HCV variants from patients in Quebec, Canada (Murphy *et al.*, 2003). Among these, variants QC227, QC216 and QC99 were assigned to recently recognized subtypes 6o, 6p and 6q, respectively (Simmonds *et al.*, 2005), and were selected for complete sequencing here. VN4 was the first isolate of subtype 6o (obtained in 1995 from a Vietnamese patient), although it was initially designated subtype '7c' (Stuyver *et al.*, 1995). QC30 was the second isolate of subtype 6o detected from a Vietnamese immigrant in Quebec, Canada (Bernier *et al.*, 1996). QC33

and QC227 were later isolated from another Vietnamese immigrant and a Caucasian in Quebec, respectively, and both were classified as subtype 6o (Murphy *et al.*, 2003). Interestingly, the Caucasian had acquired HCV infection in Quebec after a blood transfusion in the mid-1980s. This would be the first report of post-transfusional hepatitis C caused by a genotype 6 isolate in North America. VT316 is another subtype 6o strain isolated from Vietnam (<http://hcv.lanl.gov/content/hcv-db/index>).

Five subtype 6p isolates have been characterized so far, all with origins from south-eastern Asia. VN12 was the first isolate of subtype 6p discovered in a Vietnamese patient (Stuyver *et al.*, 1995), followed by QC123, QC190 (identified among Cambodian immigrants in Quebec) and QC216 (from a Vietnamese immigrant; Murphy *et al.*, 2003). Recently, isolate D3 was obtained from a blood donor in Ho Chi Minh City. It represents a recombinant between subtypes 2i and 6p (Noppornpanth *et al.*, 2006a). A recombination event could reflect a relatively high degree of prevalence or co-infection.

In addition to QC99, which was sequenced fully in this study, five subtype 6q sequences are available from GenBank: QC57, QC164, QC176, QC189 and IG93336, all isolated from Vietnamese or Cambodian immigrants in Quebec (Murphy *et al.*, 2003) or recorded to be from Cambodia (IG93336; <http://hcv.lanl.gov/content/hcv-db/index>). Isolate IG93335 from Laos is recorded as subtype 6q (<http://hcv.lanl.gov/content/hcv-db/index>) but, in our phylogenetic analysis, it appears to be a distinct, novel subtype (IG93335 has similarities of 82.4–85.3% with the six subtype 6q sequences). This range of nucleotide similarity is comparable to those observed for 6i versus 6j (82.9%), 6c versus 6d (84%) and 6k versus km41 (85.1%) when the same NS5B regions were compared pairwise.

Almost all subtypes of HCV genotype 6 (except for subtype 6a) are characteristic of endemic transmission. In other words, isolates of these subtypes appear to be relatively rare, sampled from a restricted geographical region (or from immigrants from that area) and highly genetically diverged from each other. The low frequency of detection of these strains could be due to biased or incomplete sampling or low prevalence. The latter could be a result of either inefficient, non-parenteral transmission or the presence of background immunity among local populations because of a long-term endemic infection. In contrast, only a single case of person-to-person transmission of these strains was confirmed outside south-eastern Asia (the Caucasian QC227, infected in Canada). It suggests a long-term co-circulation and evolution of genotype 6 variants in south-eastern Asia, with little exchange of virus lineages with outside populations. Collectively, these features are consistent with a zoonotic origin for HCV genotype 6, which in some ways resembles the pattern observed for genotypes III and IV of hepatitis E virus (Lu *et al.*, 2006b).

The accelerating rate of globalization and economic development in south-eastern Asia will lead to increased

urbanization, commercialization, transportation and tourism. More immigrants and visitors are expected to travel between south-eastern Asia and other parts of the world. This could enhance the spread of genotype 6 variants to new regions. Modern, efficient transmission routes, such as blood transfusion, injection drug use and unsafe medical practice, are thought to have been crucial for the world-wide spread of epidemic HCV strains during the past 50–70 years (Pybus *et al.*, 2001). These transmission routes, especially injection drug use, may have increased the likelihood of HCV co-infection and recombination, such as that observed recently between subtypes 1b and 2k in St Petersburg, Russia (Kalinina *et al.*, 2002, 2004). Whilst a similar recombination event was found between subtypes 6p and 2i in Vietnam, its prevalence and possible impact remain to be addressed (Noppornpanth *et al.*, 2006a).

To date, a high prevalence has not been recorded for any HCV genotype 6 subtype, except for subtype 6a. Although lesser symptoms and better responses to interferon therapy have been reported for genotype 6 infections (Dev *et al.*, 2002; Doi *et al.*, 1996), this does not necessarily indicate that such strains are less pathogenic, as they might have caused less recognizable, atypical hepatitis in many cases. In addition, current HCV detection systems were established to detect the highly prevalent worldwide epidemic strains of HCV, and may be failing to detect genetically diverse genotype 6 strains correctly. Furthermore, these variants display 5' non-coding region (NCR) sequences that are similar to those of genotype 1, whilst showing greater genetic variability in other regions (Noppornpanth *et al.*, 2006a). Thus, these variants would be identified as being of genotype 1 and consequently missed by 5' NCR-based genotyping assays. The completion of full genomic sequences for a panel of 17 assigned genotype 6 subtypes reported here will hopefully provide useful sequence information to enable more effective diagnostic testing to be developed.

This study will form the basis for a more precise estimation of the origin of HCV and the history of HCV evolution. Previously, the origin of HCV was estimated to be 500–2000 years ago by using a constant-rate molecular-clock model and based on analysis of short segments of the E1 and NS5B regions (Pybus *et al.* 2001; Simmonds, 2004; Smith *et al.*, 1997). The estimation of the origin of HCV should be improved in the future by combining the complete HCV genomic sequences now available with a newly developed and accurate 'relaxed-clock' model of evolution (Drummond *et al.*, 2006).

## Acknowledgements

This work was supported by NIH grants 5 P20 RR016443-04 and CA063640.

## References

- Apichartpiyakul, C., Chittivudikarn, C., Miyajima, H., Homma, M. & Hotta, H. (1994). Analysis of hepatitis C virus isolates among healthy blood donors and drug addicts in Chiang Mai, Thailand. *J Clin Microbiol* **32**, 2276–2279.
- Bernier, L., Willems, B., Delage, G. & Murphy, D. G. (1996). Identification of numerous hepatitis C virus genotypes in Montreal, Canada. *J Clin Microbiol* **34**, 2815–2818.
- de Oliveira, T., Deforche, K., Cassol, S., Salminen, M., Paraskevis, D., Seebregts, C., Snoeck, J., van Rensburg, E. J., Wensing, A. M. & other authors (2005). An automated genotyping system for analysis of HIV-1 and other microbial sequences. *Bioinformatics* **21**, 3797–3800.
- Dev, A. T., McCaw, R., Sundararajan, V., Bowden, S. & Sievert, W. (2002). Southeast Asian patients with chronic hepatitis C: the impact of novel genotypes and race on treatment outcome. *Hepatology* **36**, 1259–1265.
- Doi, H., Apichartpiyakul, C., Ohba, K. I., Mizokami, M. & Hotta, H. (1996). Hepatitis C virus (HCV) subtype prevalence in Chiang Mai, Thailand, and identification of novel subtypes of HCV major type 6. *J Clin Microbiol* **34**, 569–574.
- Drummond, A. J., Ho, S. Y., Phillips, M. J. & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biol* **4**, e88.
- Kalinina, O., Norder, H., Mukomolov, S. & Magnius, L. O. (2002). A natural intergenotypic recombinant of hepatitis C virus identified in St Petersburg. *J Virol* **76**, 4034–4043.
- Kalinina, O., Norder, H. & Magnius, L. O. (2004). Full-length open reading frame of a recombinant hepatitis C virus strain from St Petersburg: proposed mechanism for its formation. *J Gen Virol* **85**, 1853–1857.
- Li, C., Fu, Y., Lu, L., Ji, W., Yu, J., Hagedorn, C. H. & Zhang, L. (2006). Complete genomic sequences for hepatitis C virus subtypes 6e and 6g isolated from Chinese patients with injection drug use and HIV-1 co-infection. *J Med Virol* **78**, 1061–1069.
- Lu, L., Nakano, T., Li, C., Fu, Y., Miller, S., Kuiken, C., Robertson, B. H. & Hagedorn, C. H. (2006a). Hepatitis C virus complete genome sequences identified from China representing subtypes 6k and 6n and a novel, as yet unassigned subtype within genotype 6. *J Gen Virol* **87**, 629–634.
- Lu, L., Li, C. & Hagedorn, C. H. (2006b). Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. *Rev Med Virol* **16**, 5–36.
- Lu, L., Li, C., Fu, Y., Thaikruea, L., Thongswat, S., Maneekarn, N., Apichartpiyakul, C., Hotta, H., Okamoto, H. & other authors (2007). Complete genomes for hepatitis C virus subtypes 6f, 6i, 6j and 6m: viral genetic diversity among Thai blood donors and infected spouses. *J Gen Virol* **88**, 1505–1518.
- Major, M. E. & Feinstone, S. M. (1997). The molecular virology of hepatitis C. *Hepatology* **25**, 1527–1538.
- Mellor, J., Holmes, E. C., Jarvis, L. M., Yap, P. L. & Simmonds, P. (1995). Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification. The International HCV Collaborative Study Group. *J Gen Virol* **76**, 2493–2507.
- Mellor, J., Walsh, E. A., Prescott, L. E., Jarvis, L. M., Davidson, F., Yap, P. L. & Simmonds, P. (1996). Survey of type 6 group variants of hepatitis C virus in Southeast Asia by using a core-based genotyping assay. *J Clin Microbiol* **34**, 417–423.
- Murphy, D., Willems, B., Deschênes, M., Hilzenrat, N. & Mousseau, R. (2003). Use of nucleotide sequence analysis of the NS5B region for routine genotyping of hepatitis C virus: identification of numerous variants not classifiable into the known HCV genotypes. *Hepatology* **38**, 423A.
- Nainan, O. V., Alter, M. J., Kruszon-Moran, D., Gao, F. X., Xia, G., McQuillan, G. & Margolis, H. S. (2006). Hepatitis C virus genotypes

and viral concentrations in participants of a general population survey in the United States. *Gastroenterology* **131**, 478–484.

**Noppornpanth, S., Sablon, E., De Nys, K., Truong, X. L., Brouwer, J., Van Brussel, M., Smits, S. L., Poovorawan, Y., Osterhaus, A. D. & Haagmans, B. L. (2006a).** Genotyping hepatitis C viruses from Southeast Asia by a novel line probe assay that simultaneously detects core and 5' untranslated regions. *J Clin Microbiol* **44**, 3969–3974.

**Noppornpanth, S., Lien, T. X., Poovorawan, Y., Smits, S. L., Osterhaus, A. D. M. E. & Haagmans, B. L. (2006b).** Identification of a naturally occurring recombinant genotype 2/6 hepatitis C virus. *J Virol* **80**, 7569–7577.

**Pybus, O. G., Charleston, M. A., Gupta, S., Rambaut, A., Holmes, E. C. & Harvey, P. H. (2001).** The epidemic behavior of the hepatitis C virus. *Science* **292**, 2323–2325.

**Robertson, B., Myers, G., Howard, C., Brettin, T., Bukh, J., Gaschen, B., Gojbori, T., Maertens, G., Mizokami, M. & other authors (1998).** Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. International Committee on Virus Taxonomy. *Arch Virol* **143**, 2493–2503.

**Shinji, T., Kyaw, Y. Y., Gokan, K., Tanaka, Y., Ochi, K., Kusano, N., Mizushima, T., Fujioka, S., Shiraha, H. & other authors (2004).** Analysis of HCV genotypes from blood donors shows three new HCV type 6 subgroups exist in Myanmar. *Acta Med Okayama* **58**, 135–142.

**Simmonds, P. (2004).** Genetic diversity and evolution of hepatitis C virus – 15 years on. *J Gen Virol* **85**, 3173–3188.

**Simmonds, P., Alberti, A., Alter, H. J., Bonino, F., Bradley, D. W., Brechot, C., Brouwer, J. T., Chan, S. W., Chayama, K. & Chen, D. S. (1994).** A proposed system for the nomenclature of hepatitis C viral genotypes. *Hepatology* **19**, 1321–1324.

**Simmonds, P., Mellor, J., Sakuldamrongpanich, T., Nuchaprayoon, C., Tanprasert, S., Holmes, E. C. & Smith, D. B. (1996).** Evolutionary analysis of variants of hepatitis C virus found in South-East Asia: comparison with classifications based upon sequence similarity. *J Gen Virol* **77**, 3013–3014.

**Simmonds, P., Bukh, J., Combet, C., Deleage, G., Enomoto, N., Feinstone, S., Halfon, P., Inchauspé, G., Kuiken, C. & other authors (2005).** Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology* **42**, 962–973.

**Smith, D. B., Pathirana, S., Davidson, F., Lawlor, E., Power, J., Yap, P. L. & Simmonds, P. (1997).** The origin of hepatitis C virus genotypes. *J Gen Virol* **78**, 321–328.

**Stuyver, L., Wyseur, A., van Arnhem, W., Lunel, F., Laurent-Puig, P., Pawlotsky, J. M., Kleter, B., Bassit, L., Nkengasong, J. & other authors (1995).** Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable samples. *Virus Res* **38**, 137–157.

**Thaikruea, L., Thongsawat, S., Maneekarn, N., Netski, D., Thomas, D. L. & Nelson, K. E. (2004).** Risk factors for hepatitis C virus infection among blood donors in northern Thailand. *Transfusion* **44**, 1433–1440.

**Theamboonlers, A., Chinchai, T., Bedi, K., Jantaramee, P., Sripontong, M. & Poovorawan, Y. (2002).** Molecular characterization of hepatitis C virus (HCV) core region in HCV-infected Thai blood donors. *Acta Virol* **46**, 169–173.

**Tokita, H., Okamoto, H., Tsuda, F., Song, P., Nakata, S., Chosa, T., Iizuka, H., Mishihiro, S., Miyakawa, Y. & Mayumi, M. (1994).** Hepatitis C virus variants from Vietnam are classifiable into the seventh, eighth, and ninth major genetic groups. *Proc Natl Acad Sci U S A* **91**, 11022–11026.

**Tokita, H., Okamoto, H., Luengrojanakul, P., Vareesangthip, K., Chainuvati, T., Iizuka, H., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1995).** Hepatitis C virus variants from Thailand classifiable into five novel genotypes in the sixth (6b), seventh (7c, 7d) and ninth (9b, 9c) major genetic groups. *J Gen Virol* **76**, 2329–2335.

**Tokita, H., Okamoto, H., Iizuka, H., Kishimoto, J., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1998).** The entire nucleotide sequences of three hepatitis C virus isolates in genetic groups 7–9 and comparison with those in the other eight genetic groups. *J Gen Virol* **79**, 1847–1857.