MOLECULAR CHARACTERIZATION OF HEPATITIS C VIRUS (HCV) CORE REGION IN HCV-INFECTED THAI BLOOD DONORS

A. THEAMBOONLERS[1], T. CHINCHAI[2], K. BEDI[1], P. JANTARASAMEE[1], M. SRIPONTONG[1], Y. POOVORAWAN[1]*

[1]Viral Hepatitis Research Unit, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand; [2]Inter-Department of Medical Microbiology, Faculty of Graduate School, Chulalongkorn University, Bangkok, Thailand

Received August 16, 2002; accepted October 7, 2002

Summary. – In order to investigate the distribution of Hepatitis C virus (HCV) genotypes in Thailand, we performed phylogenetic analysis based on the virus core region and in this way we identified and reliably distinguished HCV genotypes 1–6 as well as subtypes. Among 100 plasma samples randomly selected from blood donors positive for antibodies to HCV (anti-HCV) 90 (90%) were found positive for HCV RNA and 77 of them were subjected to nucleotide sequencing. The following types and subtypes were identified in this group: 1a in 16 samples (20.8%), 1b in 14 samples (18.2%), 3a in 29 samples (37.7%), 3b in 5 samples (6.5%), and 6a in 13 samples (16.9%). Although this study allowed identification and characterization of HCV among blood donors, more extensive studies are needed to explore the HCV distribution in other population groups and in other geographical regions and to exploit the virus core-based characterization of HCV for evaluation of treatment and clinical outcome and epidemiological purposes.

Key words: HCV; virus core; blood donor; phylogenetic tree; genotype; subtype

Introduction

HCV infection, which can cause chronic liver diseases, cirrhosis and hepatocellular carcinoma, is still a major worldwide problem with approximately 170 million (3%) people already infected and 3–4 million de novo infected each year. (World Health Organization, 1997)

HCV, a positive single-stranded RNA virus of approximately 9.4 kb showing structural similarities with flaviviruses and pestiviruses, and hence closely related to Dengue, Japanese B encephalitis and Yellow fever viruses, has been identified as the major etiological agent of the parenterally transmitted non-A, non-B hepatitis (NANB, Choo *et al.*, 1989).

Upon comparing the sequences of its variants collected from different geographical areas, at least six major genotypes have been identified and classified. Among these, many contain a variety of more closely related but distinctive subtypes of the virus (Simmonds *et al.*, 1993). Attention has been focused on clinically relevant differences in liver disease caused by HCV, which may be attributable to infection with different genotypes, for example, leading to patients' different responses to the interferon therapy (Martinot-Peignoux *et al.*, 1998). Virus classification into genotypes may be achieved by phylogenetic analysis of complete genomic sequences (Okamoto *et al.*, 1992) or genomic regions such as core, E1, NS4 or NS5. Yet, as the respective genotypes are apparently crucial for patient management and as investigating the distribution of these sequences in Southeast Asia is imperative, sequences in the core region have been amplified which led to a reliable distinction of types 1–6 (Mellor *et al.*, 1996). The core region is relatively well conserved with nucleotide sequence similarity ranging from 81 to 88% in isolates of different genotypes (Simmond *et al.*, 1994). Due to the high degree of conservation these regions have been chosen for

*Corresponding author. E-mail: Yong.P@chula.ac.th; fax: +662-2564929.

Abbreviations: HCV = Hepatitis C virus; NANB = non-A, non-B hepatitis; RFLP = restriction fragment length polymorphism; PCR = polymerase chain reaction; RT-PCR = reverse transcription – PCR; IVDU = intravenous drug user; LiPA = line probe assay

the design of generic primer sequences to allow direct polymerase chain reaction (PCR) amplification and sequencing and hence, HCV characterization. Our objective has been to examine the genotype prevalence among blood donors using as a reliable method characterization of the HCV core region and to compare the genotype distribution with that reviewed in previous studies.

Materials and Methods

Population study. Plasma specimens were collected from new blood donors, who had screened positive for HCV antibodies at the National Blood Centre, Thai Red Cross, Bangkok, conducted between June 2001 and January 2002 for the purpose of basic immunology research on HCV infection. The project was approved by the Ethics Committee, Ministry of Public Health and Faculty of Medicine, Chulalongkorn University, Bangkok. The anti-HCV positive plasma samples were collected by code number and kept at the Viral Hepatitis Research Unit, Chulalongkorn University and Hospital, Bangkok, Thailand. All samples were stored at -70°C until tested.

Serological test. Anti-HCV serology was performed using a commercially available third generation ELISA test kit (Abbott Laboratory) according to the manufacturer's instructions.

HCV RNA detection. Anti-HCV-positive plasma samples were randomly selected for detection of HCV RNA by reverse transcription–PCR (RT-PCR) of the virus core region. Total RNA was extracted from plasma samples by a guanidinium method and reversely transcribed into cDNA using the primer 410 targeting the core region. RT-PCR was performed as described earlier (Theamboonlers *et al.,* 1995, 2000).

For amplification, the nested primer pairs 954 (an outer sense primer, 5'-ACTGCCTGATAGGGTGCTTGCGAC-3', nt (-30)–(-51) and 410 (an outer anti-sense primer, 5'-GCCG-ATCTCAT GGGGTATAT-3', nt 350–332), and 953 (an inner sense primer, 5'-AGGTCTCGTAGACCGTGCACCATG-3', nt (-16)–(3) and 951 (an inner anti-sense primer, 5'-TCAT-CGATACCCTTACATG-3', nt 324–306), respectively, were used (Mellor *et al.,* 1995). The amplification protocol described earlier (Theamboonlers *et al.,* 1995) was used; just the annealing temperature was changed to 48°C. The PCR product of 405 bp (nt (-21)–(383)) was analyzed by electrophoresis in a 1.5% agarose gel, staining with ethidium bromide and visualization by UV light.

Sequencing. The target PCR product in the agarose gel was purified for sequencing using the Perfectprep Gel Cleanup Kit (Eppendorf) according to the manufacturer's instructions, and subjected to agarose gel (1.5%) electrophoresis in order to ascertain its purity. The concentration of the purified PCR products was determined in a Shimadzu UV 160 A spectrophotometer (A_{260}= 1.0 for 50 mg of double-stranded DNA per ml). Between 10 and 30 ng of DNA per µl (3–6 µl) was used in cycle sequencing with 8 µl of a dye terminator from a DNA sequencing kit (Big Dye Terminator V.3.0 Cycle Sequencing Ready Reaction (Foster City, CA) and 3.2 pmoles of a specific primer in a final reaction volume of 20 µl in a 9600 Perkin Elmer Cetus thermocycler. This round of amplification was performed according to the manufacturer's instructions using the primer 953 and was confirmed by using the primer 951 to amplify the particular DNA strand for further sequencing. The extension products were subsequently purified by removing the unincorporated dye terminator by ethanol precipitation according to the manufacturer's instructions (ABI Sequencing Kit, ABI) and subjected to sequence analysis by an ABI Prism 310 Genetic Analyzer. The rest of the analysis was performed according to the manufacturer's instructions.

Phylogenetic analysis. In order to determine the genotype of HCV isolates we generated the multiply aligned Clustal X Program Version 1.4. Bootstrap analysis was performed for values representing 1000 replicates using the SEQBOOT Program. Distances between pairs of sequences were estimated using the DNADIST Program of the PHYLIP Package Version 3.5c obtained from Dr. J. Felsenstein, Department of Genetics, University of Washington, Washington. The distances served for phylogenetic clustering by means of the NEIGHBOR and CONSENSE Software from the PHYLIP Package. Equivalent phylogenetic relationships were also found in the maximum likelihood analysis. The TREEVIEW Program Version 1.5 was run for phylogenetic tree construction based on nucleotide sequences of the core region. The genomic sequences of 11 different HCV strains/isolates obtained from the GenBank database served as reference for the phylogenetic tree analysis. The HCV sequences clustering in the same node were considered of the same genotype.

Results and Discussion

In 2001 there were 66,340 new blood donors recorded at the National Blood Centre of the Thai Red Cross Society. Six hundred and forty-nine (0.98%) plasma samples were positive for anti-HCV (National Blood Centre, Thai Red Cross, 2002, personal communication). Between June, 2001 and January, 2002, 100 anti-HCV positive samples were randomly selected for detection of HCV RNA and 90 (90%) of them were HCV RNA-positive. Of those, 77 samples were selected for direct sequencing. The obtained sequences were submitted to the GenBank which made them accessible under Acc. Nos. AF 484969-74, AF 484976-83, AF 484985-90, AF 525883-905, AF 525907-08, AF 528323-26, AY 089747-51, AY 089754-63, AY 089765-66 and AY 089772-82. The obtained HCV core region sequences were aligned with those of the isolates of known genotype and subjected to phylogenetic analysis (Fig. 1).

The largest survey conducted to identify HCV genotypes in 1995 (Davidson *et al.,* 1995) has revealed that throughout Europe genotypes 1, 2 and 3 were predominant among blood donors. In North America, the genotype 1a was most frequent in blood donors in Pittsburgh; and varying proportions of type 1a to 1b were found in NANB patients elsewhere in the USA (Lau *et al.,* 1994). For the Far-East and South-East Asia major intra-region differences in genotype distribution were apparent in Japan and Taiwan, the genotypes 1b, 2a and 2b were the only variants detected, and in Hongkong

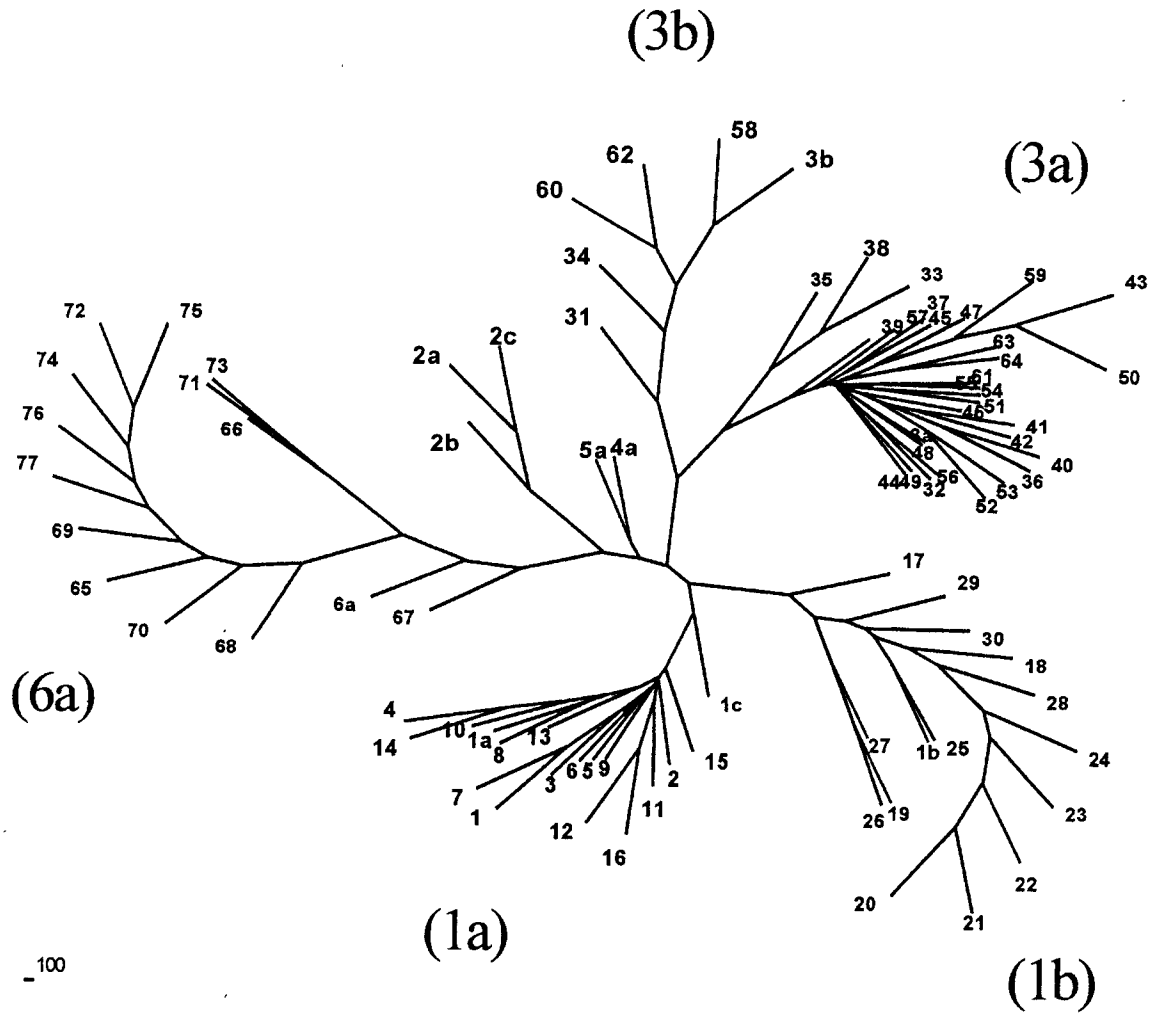

Fig. 1

Phylogenetic analysis of 77 HCV isolates from Thai blood donors and reference HCV isolates on the basis of the sequence of HCV core region

1a, 1b, 3a, 3b and 6a represent genotypes clustering on the tree.

a high frequency of type 6a infection was reported (McOmish *et al.*, 1994). The genotype 3a has been found in Malaysia, Singapore and Thailand (Davidson *et al.*, 1995; Theamboonlers *et al.*, 2000) and the genotype 3 predominated in Bangladesh, Pakistan and India. In South Africa, the genotype 5a was predominant in blood donors. Now it has been established with reasonable certainty that the genotype 4 is predominant in Zaire and elsewhere in Central and North Africa (Bukh *et al.*, 1993; Stuyver *et al.*, 1993; Simmond *et al.*, 1993; McOmish *et al.*, 1994).

The variation in African and South-East Asian regions appears to be different from that in Europe and other Western countries. The genotypes 1a, 1b, and 3a are most frequent, the occurrence of others in these countries is often attributable to recent travelling or immigration; e.g., a Canadian blood donor infected with the genotype 6 was found to be a recent immigrant from Vietnam (Murphy *et al.*, 1994).

Based on our previous study of IVDUs using RT-PCR and RFLP, the HCV genotype distribution consisted of 3a (73.5%), 1b (17. 9%), 6a (3.55%), 3b (1.44%), 1a (1.0%), coinfection 3a and 6a (2.1%) and non-typable (0.6%) (Hansurabhanon *et al.*, 2002). In this study, infected Thai blood donors harbored the genotypes 1a (20.8%), 1b (18.2%), 3a (37.7%), 3b (6.5%), and 6a (16.9%). The results showed that the sporadic genotypes 1a, 1b and 3a were nearly

Table 1. HCV genotypes among different population groups in Thailand as determined by various methods

Reference	Population group	HCV genome region	No.	HCV genotype No. (%)						Genotyping method
				1a	1b	2	3a	3b	6a	Untypable	
Luengrojanakul *et al.*, 1994	Blood donor	Core	8	–	2 (22)	–	4 (44)	–	2 (22)	–	TS primer
	Liver disease	Core	39	6 (15)	12 (31)	2(5)	19 (49)	–	1 (3)	–	TS primer
	Renal failure	Core	43	8 (19)	7 (16)	1(2)	20 (47)	2 (5)	6 (14)	–	TS primer
Apichartpiyakul *et al.*, 1994	Blood donor	NS5	23	3 (13)	2 (9)	–	8 (35)	3 (13)	3 (13)	4 (17)	SA
	IVDU	NS5	17	5 (29)	2 (12)	–	5 (29)	3 (18)	1 (6)	–	SA
	OP/HE[a] smoker	NS5	7	2 (29)	2 (29)	–	1 (14)	–	–	–	SA
Doi *et al.*, 1996	Liver disease	5'-NCR	46	10 (22)	7 (15)	–	13 (28)	1 (2)	1 (2)	7 (15)	SA
Greene *et al.*, 1995	Liver disease	E1, E2, NS4, NS5	22	3 (14)	8 (36)	–	10 (46)	1 (4)		–	SA
Mellor *et al.*, 1996	Blood donor	Core/5'-NCR	70	9 (13)	12 (17)	–	36 (51)	6 (9)	7 (10)	–	RFLP
Davidson *et al.*, 1995	Blood donor	5'NCR	76	10 (13)	24 (32)	–	36 (47)	6 (8)		–	RFLP
Kanistanon *et al.*, 1998	Blood donor	5'-NCR/Core	201	17 (8)	42 (21)	–	83 (41)	7 (4)	34 (17)	18 (8)	RH
Theamboonlers *et al.*, 2000	Blood donor	5'-NCR/Core	41	3 (7)	13 (32)	–	18 (44)	3 (7)	4 (10)	–	RFLP
	Liver disease	5'-NCR/Core	86	12 (14)	33 (38)	–	33 (38)	0 (0)	5 (6)	–	RFLP
Hansurabhanon *et al.*, 2002	IVDU	Core	290	3 (1)	52 (18)	–	213 (73)	4 (1)	10 (3)	1 (0.3)	RFLP
This study	Blood donor	Core	77	16 (21)	14 (18)	–	29 (38)	5 (6)	13 (17)	–	SA

a = opium/heroin; RH = reverse hybridization (line probe assay); SA = sequence analysis; TS primer = type-specific primer.

equally distributed in the blood donor group. Among IVDUs the genotype 3a (73.5%) was most frequent in contrary to the blood donor group, which might be associated with sharing needles/syringes and/or drug paraphernalia among IVDUs. The genotype patterns among different population groups in Thailand are shown in Table 1.

The purpose of the present study was to investigate the HCV genotype distribution among Thai blood donors confirmed anti-HCV-positive by a third generation ELISA and HCV RNA-positive by RT-PCR. There are numerous methods available to determine HCV genotypes such as RFLP, INNO LiPA and nucleotide sequencing of different regions of core, NS5, and 5'-NCR regions. As for the distribution of HCV genotypes in south-east Asia, most data have been obtained by RFLP (Davidson *et al.*, 1995), INNO LiPA (Stuyver *et al.*, 1993), and assays using genotype-specific primers (Okamoto *et al.*, 1993). These analyses could identify only 50% of all the 6 known genotype/subtype variants. However, the INNO LiPA for HCV (Innogenetics, Ghent, Belgium) could not distinguish all virus subtypes or some novel genotypes discovered in Thailand and Vietnam, the 5'-NCR sequences of which were identical to type 1 viruses so that all 6a genotype samples obtained from sequencing were 1b by the INNO LiPA for HCV (Chinchai *et al.*, unpublished data).

Developing new genotype assays based on the sequences amplified from the more variable virus core region will help investigating HCV genotype distribution in south-east Asia. The most reliable method might be the direct nucleotide sequencing. Hence, this study aimed at determination of the prevalence of HCV infection as well as the HCV genotype distribution among blood donors in Thailand in order to devise preventive measures with special consideration as to the type and severity of the disease. The main problem with any HCV genotype/subtype assay is the need of its continuous modifications as new HCV genotypes evolve.

In conclusion, our data showed the prevalence of certain HCV genotypes among blood donors in Thailand by using a reliable molecular method characterizing the HCV core region of the genome. These data will facilitate diagnosis, development, treatment, and outcome and epidemiological study.

Acknowledgement. This study was supported by a grant from the Centre of Excellence Research Fund, Chulalongkorn University, the Thailand Research Fund, Senior Research Scholar and the HECSA project from European Commission. The authors also thank the National Blood Centre, the Thai Red Cross for providing the specimens, Ms. P. Hirsch and Venerable Dr. M. Bhikkhu of Wat Rajaorasaram, Bangkok for reviewing the manuscript.

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.

Bukh J, Miller RH, Kew MC, Purcell RH (1993): Hepatitis C virus RNA in southern African blacks with hepatocellular carcinoma. *Proc. Natl. Acad. Sci. USA* **90**, 1848–1851.

Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M (1989): Isolation of a cDNA derived from a blood borne non-A, non-B hepatitis genome. *Science* **244**, 659–362.

Davidson F, Simmonds P, Ferguson JC, Jarvis LM, Dow BC, Follett EA, Seed CR, Krusius T, Lin C, Medgyesi GA *et al.* (1995): Survey of major genotypes and subtypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. *J. Gen. Virol.* **76**, 1197–1204.

Doi H, Apichartpiyakul C, Ohba K, 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.

Green WK, Cheong MK, Ng V, Yap KW (1995): Prevalence of hepatitis C virus sequence variants in south-east Asia. *J.Gen.Virol.* **76**, 211–215.

Hansurabhanon T, Jirapongsa C, Tunsakun P, Sukbunsung R, Bunyamanee B, Kuirat P, Meedsen S, Waedeng W, Theamboonlers A, Poovorawan Y (2002): Infection with hepatitis C virus among intravenous-drug users: prevalence, genotypes and risk-factor-associated behaviour patterns in Thailand. *Ann. Trop. Med. Parasitol.* (in press).

Kanistanon D, Neelamek M, Dharakul T, Songsivilai S (1998): Genotyping Distribution of Hepatitis C Virus in Different Regions of Thailand. *J.Clin. Microbiol.* 1772–1776.

Lau GK, Fang JW, Wu PC, Davis GL, Lau JY (1994): Detection of hepatitis C virus genome in formalin-fixed paraffin-embedded liver tissue by in situ reverse transcription polymerase chain reaction. *Med. Virol.* **44**, 406–409.

Luengrojanakul P, Vareesangthip K, Chainuvati T, Murata K, Fumio T, Tokita H, Okamoto H, Miyakawa Y, Mayumi M (1994): Hepatitis C virus infection in patients with chronic liver disease or chronic renal failure and blood donors in Thailand. *J. Med. Virol.* **44**, 287–292.

Martinot-Peignoux M, Boyer N, Pouteau M, Castelnau C, Giuily N, Duchatelle V, Auperin A, Degott C, Benhamou JP, Erlinger S, Marcellin P (1998): Predictors of sustained response to alpha interferon therapy in chronic hepatitis C. *J. Hepatol.* **9**, 214–223.

McOmish F, Yap PL, Dow BC, Follett EA, Seed C, Keller AJ, Cobain TJ, Krusius T, Kolho E, Naukkarinen R *et al.* (1994): Geographical distribution of hepatitis C virus genotypes in blood donors: an international collaborative survey. *J. Clin. Microbiol.* **32**, 884–892.

Mellor J, Holmes EC, Jarvis LM, Yap PL, 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 EA, Prescott LE, Jarvis LM, Davidson F, Yap PL, 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, Delage G (1994): Use of the 5' noncoding region for genotyping hepatitis C virus. *J. Infect. Dis.* **169**, 473–475.

Okamoto H, Kurai K, Okada S, Yamamoto K, Lizuka H, Tanaka T, Fukuda S, Tsuda F, Mishiro S (1992): Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. *Virology* **188**, 331–341.

Okamoto H, Tokita H, Sakamoto M, Horikita M, Kojima M, Iizuka H, Mishiro S (1993): Characterization of the genomic sequence of type V (or 3a) hepatitis C virus isolates and PCR primers for specific detection. *J. Gen. Virol.* **74**, 2385–2390.

Simmonds P, Holmes EC, Cha TA, Chan SW, McOmish F, Irvine B, Beall E, Yap PL, Kolberg J, Urdea MS (1993): Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *J. Gen. Virol.* **74**, 2391–2399.

Simmonds P, Smith DB, McOmish F, Yap PL, Kolberg J, Urdca MS, Holmes EC (1994): Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS5 regions. *J. Gen. Virol.* **75**, 1053–1061.

Stuyver L, Rossau R, Wyseur A, Duhamel M, Vanderborght B, Van Heuverswyn H, Maertens G (1993): Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay. *J. Gen. Virol.* **74**, 1093–1102.

Theamboonlers A, Yatsuhashi H, Hamada R, Fujino T, Poovorawan Y, Yano M (1995): Hepatitis C virus antibodies, viral RNA levels and genotypes in Thailand. *Southeast Asian J. Trop. Med. Public Health* **26**, 506–510.

Theamboonlers A, Kaew-in N, Hirsch P, Poovorawan Y (2000): Determination of the genotypes of hepatitis C virus in Thailand, from restriction-fragment length polymorphisms. *Ann. Trop. Med. Parasitol.* **94**, 525–527.

World Health Organization. (1997): Hepatitis C. global prevalence. *Wkly Epidemiol. Rec.* **72**, 65–72.