Concordance of Hepatitis C Virus Subtyping by Non-structural 5A and Non-structural 5B Sequencing

Michael O. Baclig[1,2], Veronica F. Chan[1], John Donnie A. Ramos[1,3], Juliet Gopez-Cervantes[4] and Filipinas F. Natividad[2]

[1] *Graduate School, University of Santo Tomas, Philippines*
[2] *Research and Biotechnology Division, St. Luke's Medical Center, Philippines*
[3] *Research Center for Natural Sciences, University of Santo Tomas, Philippines*
[4] *Center for Liver Diseases, St. Luke's Medical Center, Philippines*

ABSTRACT

The non-structural 5B (NS5B) gene is the target region to identify hepatitis C virus (HCV) subtypes. However, it is not always possible to amplify this region because of inherently high sequence variability. Nucleotide sequences of the non-structural 5A (NS5A) and NS5B genes and its concordance were determined from patients infected with HCV genotype 1 (HCV-1). Among the 30 HCV-1 samples, 7 (23%) were identified as subtype 1a and 23 (77%) were identified as 1b by NS5A sequencing. Sequence analysis of the NS5B showed that 13 (43%) were identified as 1a and 17 (57%) were identified as 1b. Out of the 13 samples identified as 1a by NS5B, 6 (46%) were correctly identified by NS5A. Of the 17 samples identified as 1b by NS5B, 16 (94%) were correctly identified by NS5A. The presence of glutamic acid (E) or aspartic acid (D) at position 2225 in the NS5A differentiates 1a from 1b subtypes, respectively. This study showed that the NS5A sequencing can identify HCV-1a and 1b subtypes with predictive values of 86% and 70% of cases, respectively. The overall concordance with NS5B was 73%. NS5B sequence analysis remains to be the reference method to identify HCV-1 subtypes. NS5A sequencing may be used to complement NS5B sequencing in case the NS5B gene cannot be successfully amplified.

Key Words: Hepatitis C virus, subtyping, non-structural 5A, non-structural 5B, DNA sequencing

Introduction

Hepatitis C virus (HCV) infection is a global health problem with an estimated 170 million chronic carriers worldwide. Six major genotypes and more than 90 subtypes have been identified worldwide. The HCV genotypes vary in different geographical regions. Genotype 1 is distributed worldwide and is predominant in Europe and the United States. Genotype 2 represents 10% to 30% of HCV genotypes and is particularly common in Japan and Italy. Genotype 3 is prevalent among intravenous drug users in Europe and the United States. Genotypes 4 and 5 are essentially restricted to the Middle East and South Africa, respectively. Genotype 6 seems to be confined in Southeast Asia particularly in Hong Kong and Vietnam.[1,2,3,4,5] In the Philippines, a vast majority of HCV samples were found to be either subtype 1a or 1b.[6,7,8,9] HCV creates a huge disease burden, since it accounts for 20% to 30% of cases of acute hepatitis, 70% to 80% of cases of chronic hepatitis, 40% of cases of end-stage liver disease, 50% to 76% of cases of hepatocellular carcinoma (HCC) and 30% to 40% of liver transplants.[10,11,12]

HCV is a positive strand RNA virus that has been classified in the genus *Hepacivirus* of the family *Flaviviridae*. The positive sense RNA genome is translated to produce a polyprotein which is processed to generate the mature structural and non-structural proteins.[13,14] Analysis of the nucleotide sequence homology of various hepatitis C virus genomic regions such as the 5'untranslated region (5'UTR) and regions coding for the envelope (E1), core, and non-structural 5B (NS5B) led to the identification of six (6) major genotypes, formerly denoted as clades, and numerous closely related subtypes within the genotypes.[15,16] The NS5B region sequence analysis is considered the reference method and has been recommended for HCV genotyping and subtyping in consensus proposals.[17,18] However, it has been shown that it is not always possible to amplify this region because of primer-target mismatch within the highly variable NS5B sequence. In addition, failure in sequencing the NS5B gene has been reported by some investigators despite successive attempts to amplify cDNA with the NS5B primers.[19,20] The present study evaluated the accuracy of subtyping of HCV-1 samples by partial NS5A sequencing as compared with the partial NS5B sequence.

This basic knowledge on HCV genomic heterogeneity has clinical implications in prognostication as well as treatment. The prognostic significance lies in the association of specific genotypes and subtypes with severe liver disease such as chronic hepatitis, cirrhosis and HCC. Moreover, some genotypes and subtypes has been shown to be resistant to antiviral treatment, with HCV genotypes 2 and 3 being associated with better response, compared with genotype 1. Marto *et al.*, (2008) found that the identification of the infecting genotype and subtype is an

Corresponding author: Michael O. Baclig, PhD
Research and Biotechnology Division
St. Luke's Medical Center
279 E. Rodriguez Sr. Blvd. Quezon City, Philippines 1102
Telephone: +632 727-5540
Telefax: +632 726-0467
Email: mobaclig@stluke.com.ph

important tool to optimize its treatment type, its duration of therapy and its correct dose.[18]

Methods

Patients

Thirty (30) blood samples from patients infected with HCV-1, previously confirmed by PCR-RFLP and clinically diagnosed with chronic hepatitis C, collected from May 2005 to December 2008 at St. Luke's Medical Center, Philippines were analyzed. There were 17 males and 13 females with ages ranging from 32 to 76 years old. Patients are excluded if they have other causes of liver diseases such as autoimmune hepatitis and history of alcoholism, and if they were found to be reactive to hepatitis B surface antigen (HBsAg) and antibody to hepatitis B core antigen (anti-HBc).

HCV RNA Extraction

The viral RNA from HCV-infected patient plasma was extracted from peripheral blood using the QIAamp® Viral RNA Mini kit from Qiagen according to manufacturer's instructions. Briefly, 140 µl of plasma was digested at 56ºC for 5 minutes with AVL buffer as provided by the kit. Absolute ethanol was added for the purpose of precipitation and the mixture was transferred to the spin column and washed three times using the provided buffer to purify the viral RNA that binds to the column. The viral RNA was then eluted and collected from the column. The nucleic acid concentration and purity was determined by measuring the absorbance at 260/280 nm using the NanoDrop® 1000 Spectrophotometer (Thermo Fisher Scientific).

cDNA Synthesis

Reverse transcription was carried out from 10 µl of viral RNA extract using the SuperScript™ III reverse transcriptase (Invitrogen™).

NS5A Nested PCR Amplification

The NS5A amplification was carried out with 0.25 pmol of outer sense primer 5′-CAGTGCTCACTTCCATGCTCA-3′; and outer antisense primer 5′-ACGGATATTTCCCTCTCATCC-3′, 0.02 U Phusion™ DNA polymerase, 0.2 mM dNTPs, 1X Phusion™ HF buffer, and PCR-grade water with the following thermal profile: initial denaturation at 98°C for 30 seconds followed by 35 cycles at 98°C for 10 seconds, 56°C for 10 seconds, 72°C for 30 seconds, and final extension at 72°C for 10 minutes.[21] The conditions of the second PCR were the same as described above using inner pair of sense and antisense primers 5′-ACCCCTCCCACATTACAGCAG-3′ and 5′-CCGAAGCGGATCGAAAGAGTCCA-3′.

NS5B Nested PCR Amplification

The NS5B amplification was carried out with 0.5 pmol of outer sense primer 5′-TGGGGTTCTCGTATGATACCC-3′ and outer antisense primer 5′-CCTGGTCATAGCCTCCGTGAA-3′, 0.02 U Phusion™ DNA polymerase, 0.2 mM dNTPs, 1X Phusion™ HF buffer, and PCR-grade water with the following thermal profile: initial denaturation at 95°C for 1 minute, followed by 40 cycles at 95°C for 20 seconds, 56°C for 30 seconds, 72°C for 1 minute, and final extension at 72°C for 10 minutes.[22] The conditions of the second PCR were the same as described above using inner pair of sense and antisense primers 5′-GATACCCGCTGCTTTGACTC-3′and 5′-CCTCCGTGAAGGCTCTCAG-3′.

DNA Purification and Nucleotide Sequencing of the NS5A and NS5B

The nested PCR products were purified using the GFX™ PCR DNA and Gel Band Purification kit (GE Healthcare). Amplicons in the NS5A and NS5B regions were sequenced using Big Dye Terminator Cycle Sequencing Ready Reaction Kit and Applied Biosystem 3730xl automated sequencer (Macrogen, Korea). Sequence data were aligned with the consensus sequences of known subtypes using bioinformatics tools such as ChromasPro version 1.34 and BioEdit version 7.0. The nucleotide sequences were compared for identity with sequence from the National Center for Biotechnology Information (NCBI) using the basic local alignment search tool program (BLAST). The GenBank/EMBL/DDBJ accession numbers of HCV sequences used in the analysis were M62321 (HCV-1) for subtype 1a, D90208 (HCV-J) for subtype 1b. Biosafety and quality control measures were observed throughout the conduct of this study.

Results

Nucleotide sequencing of the NS5A revealed that 7 (23%) and 23 (77%) of the 30 samples were identified as subtypes 1a and 1b, respectively. In contrast, sequence analysis of the NS5B showed that 13 (43%) were identified as 1a and the remaining 17 (57%) were identified as 1b (Table 1). Out of the 13 samples identified as 1a by NS5B, 6 (46%) were correctly identified by NS5A. Of the 17 samples identified as 1b by NS5B, 16 (94%) were correctly identified by NS5A. The predictive value of NS5A to subtype 1a was 6/7 (86%). For subtype 1b, the predictive value of NS5A was 16/23 (70%). The overall concordance between NS5A and NS5B was 73% (Table 2).

Table 1. Genomic identification of HCV-1 subtypes by partial NS5A and NS5B sequencing.

	SEQUENCING	
	NS5A	NS5B
1a	7 (23%)	13 (43%)
1b	23 (77%)	17 (57%)
TOTAL	30	30

Table 2. Concordance of HCV-1 subtyping by NS5A and NS5B sequencing.

		NS5B Subtyping		
		1a	1b	Total
NS5A sequencing	1a	6	1*	7
	1b	7*	16	23
	Total	13	17	30

The NS5A sequencing is horizontally indicated and the NS5B is presented vertically.

* Inconsistent subtype assignment.

In this study, it was shown that HCV-1a viruses always presented with one substitution, glutamic acid (E) at position 2225, as compared to HCV-1b reference sequence (Table 3). While HCV-1b viruses always presented with one substitution, aspartic acid (D) at position 2225, with respect to the HCV-1a reference sequence (Figures 1 and 2).

Table 3. Identification of key amino acids in the NS5A gene.

	D2225	E2225	Total
1a	0 (0%)	7 (100%)	7
1b	23 (100%)	0 (0%)	23
Total	23	7	30

Discussion

Accurate identification of HCV genotypes and subtypes has important clinical and epidemiological implications.[23] Until now, the non-structural 5B (NS5B) is the preferred region for both genotyping and subtyping, and has been recommended in consensus proposals. However, it has been shown that it is not always possible to amplify this region because of lack of conservation in the primer-binding sites. Our study evaluated the accuracy of subtyping of HCV-1 samples by partial NS5A sequencing compared with the partial NS5B sequence.

The present study showed that the NS5A sequencing can identify subtypes 1a and 1b with predictive values of 86% and 70% of cases, respectively. Overall, among the 30 HCV-1 samples, 22 (73%) were concordantly subtyped by NS5A sequencing (Tables 1 and 2). The non-structural 5 gene was chosen because this region is sufficiently variable, and could be readily amplified from plasma of HCV-infected individuals and because there is already considerable quantity of published sequence data in this region. Nucleotide sequencing of a subgenomic region is preferred to identify HCV subtypes because full length sequence analysis is time-consuming, expensive and impractical in a clinical setting. This is based on the assumption that a particular subgenomic region is representative of the entire viral genome.[16] In 2007, Ross *et al*., reported that this supposition is still valid because both inter-genotypic and inter-subtype recombinations of HCV are rare events.[23] In addition, it has been suggested that with the exception of the 5'untranslated region (5'UTR), any region of the genome can be used as the basis for virus subtype identification, provided sequences from major databases such as the GenBank, European Molecular Biology Laboratory (EMBL), and DNA Databank of Japan (DDBJ) are available.[24]

Our study demonstrated that glutamic acid (E) at position 2225 was present in 7/7 (100%) subtype 1a viruses, while aspartic acid (D) at position 2225 was present in 23/23 (100%) subtype 1b viruses, a change that can be considered subtype specific (Table 3). Results suggest that glutamic acid (E) at position 2225 can be used as a marker to reliably identify HCV-1a subtypes. On the other hand, aspartic acid (D) at position 2225 in the NS5A region of the HCV genome can be used as a marker to reliably identify HCV-1b subtypes. Recently, it has been reported that alignments obtained with sequences from databases confirmed that glutamic acid (E) at position 2225 are widely distributed in HCV-1a subtype, while aspartic acid

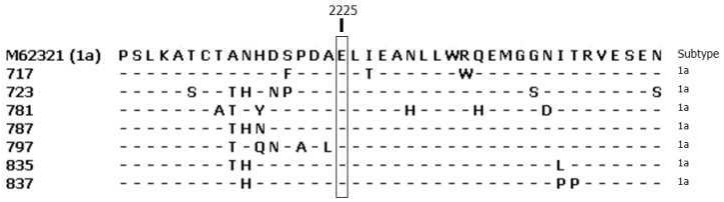

Figure 1. Deduced amino acid alignment of NS5A of 7 HCV-1a confirmed samples. At the top, HCV-1a reference sequence (GenBank accession number M62321) is given. At the left, the sample identification numbers are shown. The key amino acid associated with 1a is boxed.

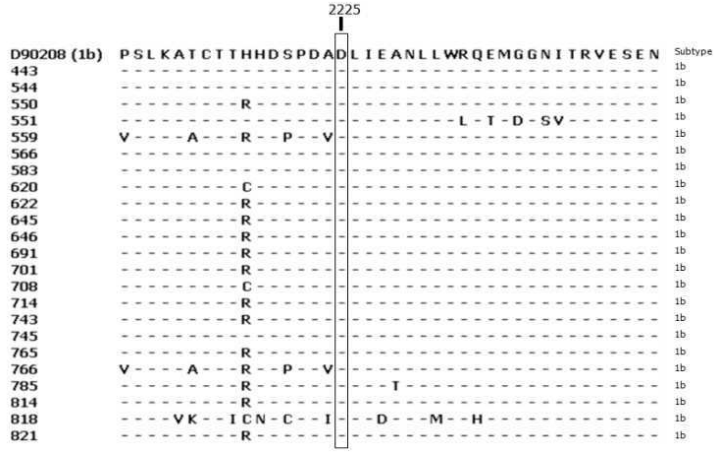

Figure 2. Deduced amino acid alignment of NS5A of 23 HCV-1b confirmed samples. At the top, HCV-1b reference sequence (GenBank accession number D90208) is given. At the left, the sample identification numbers are shown. The key amino acid associated with 1b is boxed.

(D) at position 2225 are widely distributed in HCV-1b subtype and thus, can be considered as 1a and 1b markers. This study corroborates with the findings of Torres-Punte *et al*., (2008) showing that glutamic acid (E) and aspartic acid (D) are distributed worldwide among HCV-1a and 1b viruses and thus, could be considered subtype-specific.[25] In addition, results suggest that sequence analysis of variable regions of the HCV genome, specifically the NS5A, spanning from nucleotides 6954 to 7073 may be used as an alternative target region for identification of HCV-1 subtypes.

Recently, it has been reported that there is a slight difference in treatment outcomes between HCV-1a and 1b-infected patients.[12] HCV-1b for example, is associated with less favorable prognosis following antiviral treatment.[26] In addition, based on a local study by Barzaga *et al*., (1994), certain subtypes appear to be more responsive to interferon therapy. Thus, accurate typing is an indispensable tool for tailoring antiviral treatment as well as in epidemiological investigations.[27]

Overall, these findings have major implications for epidemiologic HCV investigations as to its origin and spread of infection, its distribution, its routes of transmission including outbreak studies and novel transmission risks, its association with certain risk groups and for viral evolutionary studies.

It has been shown that patients infected with subtype 1b have been found to have more advanced liver disease, which may suggest increased pathogenicity. Thus, accurate subtyping may later on help in decision making for clinical management of HCV infection. In addition, molecular characterization of HCV-1 subtypes is likely to facilitate and contribute to the development of an effective vaccine against infection with HCV.

Conclusion

NS5A sequencing can identify HCV-1a and 1b subtypes with predictive values of 86% and 70% of cases, respectively. The overall concordance with NS5B was 73%. NS5B sequence analysis remains to be the reference method to identify HCV-1 subtypes. NS5A sequencing may be used to complement NS5B sequencing in case the NS5B gene cannot be successfully amplified.

Acknowledgment

This study was supported by a research grant to the St. Luke's Liver Diseases Study Group (Project 06-013) through the Research and Biotechnology Division, St. Luke's Medical Center, Philippines.

References

1. Simmonds P, Bukh J, Combet C, et al. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatology. 2005;42:962-973.
2. Bracho M, Carillo-Cruz F, Ortega E, Moya A, Gonzales-Candelas F. A new subtype of hepatitis C virus genotype 1: Complete genome and phylogenetic relationships of an equatorial Guinea isolate. J Gen Virol. 2006;87:1697-1702.
3. Madhavi C, Thippavuzzula R, Rao R, et al. Genotyping of hepatitis C virus in infected patients from South India. Infect Genet Evol. 2007;7:724-730.
4. Spada E, Abbate I, Sicurezza E, et al. Molecular epidemiology of a hepatitis C virus outbreak in a hemodialysis unit in Italy. J Med Virol. 2008;80:261-267.
5. Utama A, Budiato B, Monasari D, et al. Hepatitis C virus genotype in blood donors and associated liver disease in Indonesia. Intervirology. 2008;51:410-416.
6. Katayama Y, Barzaga N, Alipio A, et al. Genotype analysis of hepatitis C virus among blood donors and inmates in Metro Manila, the Philippines. Microbiol Immunol. 1996;40:525-529.
7. Hotta H, Kemapunmanus M, Apichartpiyakul C, Soetjipto, Handajani R, Barzaga N. Differential distribution of hepatitis C virus subtypes in Asia: Comparative study among Thailand, Indonesia, the Philippines and Japan. Southeast Asian J Trop Med Public Health. 1997;28:23-31.
8. Kageyama S, Agdamag D, Alesna E, et al. Tracking the entry routes of hepatitis C virus as a surrogate of HIV in an HIV-low prevalence country, the Philippines. J Med Virol. 2009;81:1157-1162.
9. Maramag F, Rivera M, Predicala R, Baclig M, Matias R, Cervantes J. Hepatitis C genotypes among Filipinos. Phil J Gastroenterol. 2006;2:30-32.
10. Lindenbach B, Rice C. Unraveling hepatitis C virus replication from genome to function. Nature. 2005;436:933-938.
11. Wohnsland A, Hofmann W, Sarrazin C. Viral determinants of resistance to treatment in patients with hepatitis C. Clin Microbiol Rev. 2007;20:23-38.
12. Verbeeck J, Stanley M, Shieh J, et al. Evaluation of Versant hepatitis C virus genotype assay (LiPA) 2.0. J Clin Microbiol. 2008;46:1901-1906.
13. Cohen J. The scientific challenge of hepatitis C. Science. 1999;285:26-30.
14. Simmonds P. Genetic diversity and evolution of hepatitis C virus - 15 years on. J Gen Virol. 2004;85:3173-3188.
15. Robertson B, Myers G, Howard C, et al. Classification, nomenclature and database development for hepatitis C virus and related viruses: Proposals for standardization. Arch Virol. 1998;143:2493-2503.
16. Nolte F, Green A, Fiebelkorn K, et al. Clinical evaluation of two methods for genotyping Hepatitis C virus based on analysis of the 5′ non-coding region. J Clin Microbiol. 2003;41:1558-1564.
17. Nadarajah R, Khan Y, Miller S, Brooks G. Evaluation of a new generation line probe assay that detects 5′untranslated and core regions to genotype and subtype hepatitis C virus. Am J Clin Pathol. 2007;128:300-304.
18. Marto E, Gonzales V, Buckton A, et al. Evaluation of a new assay for hepatitis C virus genotyping targeting both 5′NC and NS5B genomic regions in comparison with reverse hybridization and sequencing methods. J Clin Microbiol. 2008;46:192-197.
19. Tamalet C, Colson P, Tissot-Dupont H, et al. Genomic and phylogenetic analysis of hepatitis C virus isolates: A survey of 535 strains circulating in Southern France. J Med Virol. 2003;71:391-398.
20. Verma V, Chakravarti A. Comparison of 5′ non coding core with 5′ non coding regions of HCV by RT-PCR: Importance and clinical implications. Curr Microbiol. 2008;57:206-11.
21. Zeumen S, Lee J, Roth W. Mutations in the non-structural 5A gene of European Hepatitis C isolates and response to interferon-α. Hepatology. 1997;25:740-744.
22. Chen Z, Weck K. Hepatitis C virus genotyping: Interrogation of the 5′untranslated region cannot accurately distinguish genotypes 1a and 1b. J Clin Microbiol. 2002; 40:3127-3134.
23. Ross R, Viazov S, Roggendorf M. Genotyping of hepatitis C virus isolates by a new line probe assay using sequence information from both the 5′untranslated region and the core regions. J Virol Methods. 2007;143:153-160.
24. Simmonds P, Smith D, McOmish F, et al. Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS5 regions. J Gen Virol. 1994;75:1053-1061.
25. Torres-Punte M, Cuevas J, Jimenez-Hernandez N, et al. Hepatitis C virus and the controversial role of the interferon sensitivity-determining region in the response to interferon treatment. J Med Virol. 2008;80:247-253.
26. Lauer G, Walker B. Hepatitis C virus infection. N Engl J Med. 2001;345:41-52.
27. Barzaga N, Hotta H, Cabanban A, et al. Serological and molecular studies on hepatitis C virus infection in selected areas in Metro Manila. Supplement to JAMA Southeast Asia. 1994; 10:497-500.