JOURNAL OF CLINICAL MICROBIOLOGY, June 2006, p. 2229–2232 Vol. 44, No. 6
0095-1137/06/$08.00+0 doi:10.1128/JCM.02569-05

NOTES

Phylogenetic Analysis of Previously Nontypeable Hepatitis C Virus Isolates from Argentina†

María Inés Gismondi,[1]* Pablo Daniel Becker,[2] Pamela Valva,[1] Carlos Alberto Guzmán,[2] and María Victoria Preciado[1]

Laboratorio de Virología, Hospital de Niños Ricardo Gutiérrez, Gallo 1330, C1425EFD Ciudad de Buenos Aires, Argentina,[1] *and Department of Vaccinology, GBF-German Research Centre for Biotechnology, Mascheroder Weg 1, D-38124 Braunschweig, Germany*[2]

Received 9 December 2005/Returned for modification 6 February 2006/Accepted 21 March 2006

Phylogenetic analysis of hepatitis C virus isolates from Argentina that were previously nontypeable by restriction fragment length polymorphism (RFLP) analysis revealed that they belong to genotype 1a. A substitution at position 107 (G→A), which is the landmark of these strains, was shown to be distributed among isolates worldwide. The RFLP patterns obtained for these isolates should be added to the ones reported for genotype 1 isolates.

Hepatitis C virus (HCV), an RNA virus belonging to the genus *Hepacivirus* in the family *Flaviviridae*, is a major cause of chronic progressive liver disease worldwide that may lead to cirrhosis and/or hepatocellular carcinoma. Given that about 3% of the world's population is chronically infected with HCV and that 3 million to 4 million people are newly infected each year, it has become a major problem in public health (19). In fact, up to 35% of the approximately 170 million people chronically infected with HCV will develop liver cirrhosis, and hepatocellular carcinoma may develop in 2 to 7% of cirrhotic patients (26).

HCV isolates are classified into six major genetic groups referred to as genotypes, which show 30 to 35% dissimilarity at the nucleotide level. Isolates of the same genotype are further classified into subtypes, which show 20 to 25% nucleotide sequence variability (22). The geographical distribution of HCV genotypes has been described, with genotypes 1a, 1b, and 3 being the most widely distributed all over the world (20).

The HCV 5′ untranslated region (5′ UTR), which is 341 bases long and which shows a highly ordered secondary structure, is the most slowly evolving region of the viral genome (22). It is widely used for HCV genotyping, since it contains certain sequence polymorphisms that facilitate HCV genotype and even subtype identification by different methods, including restriction fragment length polymorphism (RFLP) analysis following 5′ UTR amplification by reverse transcription-PCR (6).

We have recently reported on the identification of HCV isolates displaying a new restriction site for RsaI (an enzyme used to determine HCV genotype by 5′ UTR RFLP analysis), which renders these isolates nontypeable by this method (8). The present study was performed to determine whether these strains represent a novel subtype of genotype 1.

A total of eight plasma samples belonging to six children and two adults with chronic HCV infection were analyzed. Except for a mother-infant pair, all patients were unrelated. Total RNA was extracted from plasma, and reverse transcription was performed at 42°C for 60 min with the specific antisense primer. Thereafter, the HCV 5′ UTR (251 bp), core/E1 (810 bp), and NS5B (379 bp) regions were amplified by PCR with a DNA polymerase exhibiting 3′-5′ exonuclease activity. The primers and PCR amplification conditions used in this study are shown in Table 1.

The amplification products were sequenced bidirectionally, and the sequences were aligned by using ClustalX software (24). As for the 5′ UTR, all samples displayed the previously described G→A substitution at position 107 (relative to the sequence reference strain H77; data not shown), which accounts for the novel RFLP pattern obtained when restriction was performed with HaeIII/RsaI (8).

The sequence similarity with other genotype 1 and non-genotype 1 isolates available online (http://www.euhcvdb.ibcp.fr; http://www.hcv.lanl.gov) was ascertained by distance-based methods by the use of the most appropriate model of evolution obtained from each alignment (16) and the neighbor-joining algorithm for tree reconstruction. Molecular phylogeny was estimated from the E1 and NS5B regions by use of a parsimony approach. Statistical support for each node in the trees drawn by both the neighbor-joining and the parsimony algorithms was obtained by performing 1,000 bootstrap replicates of the original nucleotide sequence alignment (7).

To provisionally designate an HCV strain as a new subtype, a recently published proposal for HCV nomenclature recommends amplification and sequencing of certain fragments of the E1 and the NS5B regions of HCV (21). In the present study, core/E1 and/or partial NS5B region amplification was

* Corresponding author. Mailing address: Laboratorio de Virología, Hospital de Niños Ricardo Gutiérrez, Gallo 1330, C1425EFD Ciudad de Buenos Aires, Argentina. Phone: 5411-4964-3118. Fax: 5411-4962-6770. E-mail: migismondi@yahoo.com.ar.

† Supplemental material for this article may be found at http://jcm.asm.org/.

TABLE 1. PCR amplification conditions and primers

HCV region	PCR round	Sequences (positions) of primers used[a]	Amplification conditions	No. of samples amplified/total no. of samples
5′ UTR	First	5′-GTGAGGAACTACTGTCTTCACGCAG-3′ (47 to 71); 5′-TG CTCATGGTGCACGGTCTACGAGA-3′ (348 to 324)	5 min at 94°C; 30 cycles of 15 s at 94°C, 30 s at 58°C, and 30 s at 68°C	8/8
	Second	5′-TTCACGCAGAAAGCGTCTAG-3′ (63 to 82); 5′-CAC TCTCGAGCACCCTATCAGGCAGT-3′ (313 to 288)	The same as for the first round	
Core/E1	First	5′-GCGTGAACTATGCAACAGGGAA-3′ (823 to 844); 5′-GGT GTTGAGGCTATCATTGCARTT-3′ (1646 to 1623)	5 min at 94°C; 35 cycles of 15 s at 94°C, 30 s at 52°C, and 1 min at 68°C	5/8
	Second	5′-TCCCGGTTGCTCTTTCTCTATCTT-3′ (848 to 871); the reverse primer was the same as that for the first round	The same as for the first round	
NS5B		5′-ATGACACCCGCTGCTTTGAC-3′ (8257 to 8276); 5′-CTR GTCATAGCCTCCGTGAA-3′ (8635 to 8616)	5 min at 94°C and 35 cycles of 15 s at 94°C, 30 s at 53°C, and 40 s at 68°C	5/8

[a] Nucleotide numbering is relative to that of reference sequence H77.

achieved for six of the eight plasma samples (Table 1). The results of analysis of the partial E1 region is shown in Fig. 1A. As expected, all nontypeable isolates grouped together, indicating that they are closely related. Furthermore, Argentine nontypeable isolates and prototypic strains of genotype 1a formed a monophyletic group, showing that the former isolates may be classified as genotype 1a. Of note, one of the Argentine strains of genotype 1a (as determined by RFLP analysis) isolated in our laboratory also grouped with the nontypeable isolates, thus confirming that the method used was appropriate for genotyping.

Figure 1B shows the phylogeny obtained with the partial NS5B region for five plasma samples. The tree topology proved to be similar to the one obtained for the E1 region. The same results were reached when analysis was performed by distance-based methods (data not shown). Thus, we conclude that the previously nontypeable isolates should provisionally be classified as HCV genotype 1a.

The sequence alignments of the 5′ UTR, 5′ UTR-E1, and 5′ UTR-NS5B regions belonging to isolates with complete genome sequences available and of the 5′ UTR-E1 and 5′ UTR-NS5B regions from partially sequenced genomes are included in the supplemental material.

In order to evaluate the global distribution of the previously described substitution at position 107, 5′ UTR sequences from our nontypeable isolates were aligned with other genotype 1a sequences that have been isolated worldwide and that were retrieved from HCV databases (4, 10). Remarkably, the G→A substitution was present in samples from North and South America, Europe, and Asia (Table 2). According to their sequence

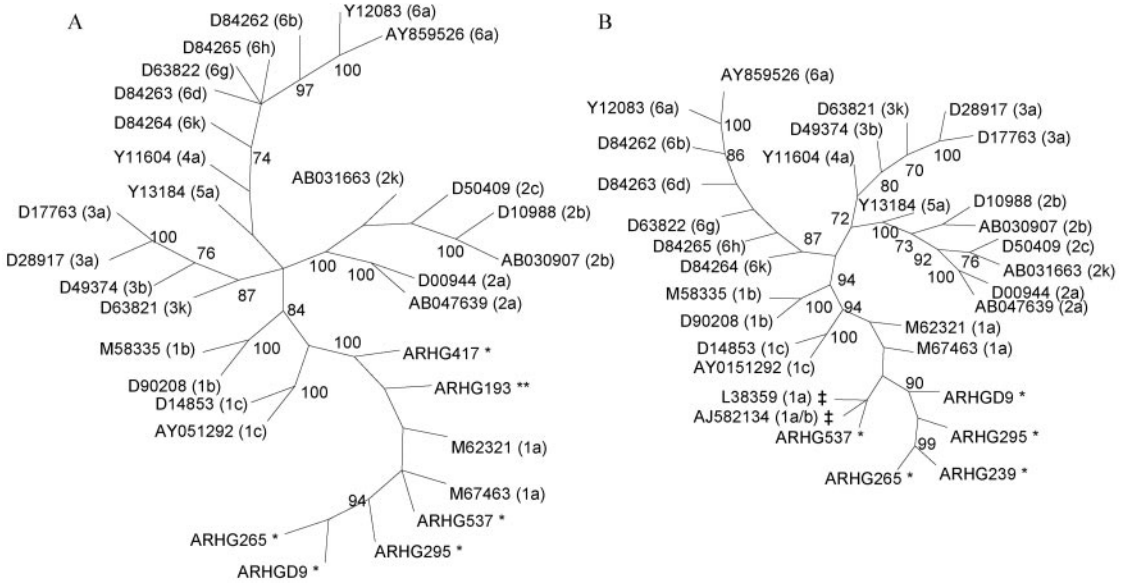

FIG. 1. Phylogenetic analysis of nontypeable isolates, other South American isolates, and prototypic HCV strains. Isolates are named according to their GenBank accession numbers. HCV genotype and subtype are indicated in parentheses. Nontypeable isolates are labeled with an asterisk. Bootstrap values greater than 70% after 1,000 resamplings are indicated at the nodes. (A) Phylogeny estimated by the use of partial E1 sequences (nucleotides 878 to 1292 of reference isolate H77). The strict consensus tree of the eight most parsimonious trees is represented. **, Argentine isolate of genotype 1a, as determined by RFLP analysis of the 5′ UTR. (B) Phylogeny estimated by the use of partial NS5B sequences (nucleotides 8276 to 8615 of reference isolate H77). The strict consensus tree of the two most parsimonious trees is shown. ‡, samples from Brazil (L38359) and Peru (AJ582134) retrieved from the HCV database (10).

TABLE 2. HCV isolates from all over the world showing a G→A nucleotide substitution at position 107

Isolate name	GenBank accession no.	Country of origin	Reference or source
ARHG265	DQ313453	Argentina	Present study
ARHG295	DQ313454	Argentina	Present study
ARHG537	DQ313455	Argentina	Present study
ARHGD9	DQ313456	Argentina	Present study
ARHG417	DQ313457	Argentina	Present study
ARHG205	AY376834	Argentina	8
ARHG74	AY376837	Argentina	8
ARHG239	AY376836	Argentina	8
QC40	U05029	Canada	12
186	DQ087249	China	Unpublished
14224	AY188158	Denmark	5
16878	AY188159	Denmark	5
16168	AY188160	Denmark	5
PM	AY188161	Denmark	5
AL	AY188162	Denmark	5
CS	AY188163	Denmark	5
2296	AY188164	Denmark	5
FR117	U51750	France	14
S1	AF245282	Germany	18
NIV-33	AF158606	India	Unpublished
NIV-13	AF158612	India	Unpublished
K-0001-1A	D29817	Japan	Unpublished
NL57	X58939	The Netherlands	9
VL1_1	M94499	Spain	11
SOT7	AF217297	United Kingdom	Unpublished
XF223	AF511949	United Kingdom	Unpublished
34510	AY766578	United Kingdom	Unpublished
41907	AY766618	United Kingdom	Unpublished
44617	AY766631	United Kingdom	Unpublished
Glasgow	AY885238	United Kingdom	15
AG094	Z14099	United States	2
US10	L38352	United States	23
INC9-48	AY446047	United States	1
PD102	AY695436	United States	Unpublished
PD101	AY695437	United States	Unpublished
Montevideo 1	AJ012831	Uruguay	3
Montevideo 2	AJ012832	Uruguay	3

diversity, HCV types and subtypes may be categorized as "endemic" or "epidemic" (20, 22). Genotype 1a belongs to the "epidemic" group of strains, which have shown increasing diversification in the recent past and which are considered to be the greatest threat to public health in the near future (17). Indeed, Colina et al. (3) and Vega et al. (25) have shown the increasing diversification of HCV isolates in South America, whereas others suggested the migration of genotype 1a strains from the United States to Brazil (13). Hence, the wide distribution of the substitution mentioned above is a further indicator of HCV diversification.

The intragenotype variability of unrelated HCV isolates may be explained by neutral theory, in which neutral sequence changes that occur in coding or noncoding regions become fixed by chance (20). The previously detected G→A substitution may therefore represent one such event of neutral evolution within the HCV 5′ UTR. However, further structural biology studies need to be carried out in order to examine the biological relevance of such mutation.

Finally, genotype determination is indicated before the onset of antiviral treatment, since genotypes 1 and 4 have been reported to be more resistant to alpha interferon, and thus, patients infected with these genotypes require more prolonged therapy (20). Our results reveal that genotype 1a isolates from Argentina, North America, Europe, and Asia may be misclassified by standard genotyping methods. Therefore, we propose that the restriction patterns observed for these previously nontypeable isolates (8) be considered in addition to the ones described by Davidson et al. (6) for HCV genotype 1.

Nucleotide sequence accession numbers. The GenBank accession numbers for the sequences reported in this work are DQ313453 to DQ313468.

M.V.P. is a member of the National Research Council (CONICET) Research Career Program, M.I.G. was supported by a fellowship from CONICET, and P.V. has a fellowship from the National Ministry of Health. This study was supported by grants from the National Agency for Scientific and Technological Promotion (PICT no. 25344) and CONICET (PIP no. 5359).

We are indebted to Karina Watzke for her outstanding technical help.

REFERENCES

1. **Blackard, J. T., Y. Yang, P. Bordoni, K. E. Sherman, and R. T. Chung.** 2004. Hepatitis C virus (HCV) diversity in HIV-HCV-coinfected subjects initiating highly active antiretroviral therapy. J. Infect. Dis. **189:**1472–1481.
2. **Brown, E. A., H. Zhang, L. H. Ping, and S. M. Lemon.** 1992. Secondary structure of the 5′ nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. Nucleic Acids Res. **20:**5041–5045.
3. **Colina, R., C. Azambuja, R. Uriarte, C. Mogdasy, and J. Cristina.** 1999. Evidence of increasing diversification of hepatitis C viruses. J. Gen. Virol. **80:**1377–1382.
4. **Combet, C., F. Penin, C. Geourjon, and G. Deleage.** 2004. Hepatitis C virus sequences database. Appl. Bioinformatics **3:**237–240. [Online.] http://www.euhcvdb.ibcp.fr.
5. **Corbet, S., J. Bukh, A. Heinsen, and A. Fomsgaard.** 2003. Hepatitis C virus subtyping by a core-envelope 1-based reverse transcriptase PCR assay with sequencing and its use in determining subtype distribution among Danish patients. J. Clin. Microbiol. **41:**1091–1100.
6. **Davidson, F., P. Simmonds, J. Ferguson, L. Jarvis, B. Dow, E. Follet, C. Seed, T. Krusius, C. Lin, G. Medgyesi, H. Kiyokawa, G. Olim, G. Duraisamy, T. Cuypers, A. Saced, D. Teo, J. Conradie, M. Kew, M. Lin, C. Nuchaprayoon, O. Ndimbie, and P. Yap.** 1995. Survey of major genotypes and subtypes of hepatitis C using RFLP of sequences amplified from the 5′ non-coding region. J. Gen. Virol. **76:**1197–1204.
7. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791.
8. **Gismondi, M. I., L. H. Staendner, S. Grinstein, C. A. Guzmán, and M. V. Preciado.** 2004. Hepatitis C virus isolates from Argentina disclose a novel genotype 1-associated restriction pattern. J. Clin. Microbiol. **42:**1298–1301.
9. **Kleter, G. E., L. J. van Doorn, J. T. Brouwer, S. W. Schalm, R. A. Heijtink, and W. G. Quint.** 1994. Sequence analysis of the 5′ untranslated region in isolates of at least four genotypes of hepatitis C virus in The Netherlands. J. Clin. Microbiol. **32:**306–310.
10. **Los Alamos National Laboratory.** 2006. HCV sequence database [Online.] http://www.hcv.lanl.gov.
11. **Martell, M., J. I. Esteban, J. Quer, J. Genesca, A. Weiner, R. Esteban, J. Guardia, and J. Gomez.** 1992. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. J. Virol. **66:**3225–3229.
12. **Murphy, D. G., B. Willems, J. P. Villeneuve, and G. Delage.** 1994. Sequence analysis of the 5′ noncoding region reveals the existence of multiple hepatitis C virus genotypes in Quebec, Canada, p. 296–300. *In* K. Nishioka, H. Suzuki, S. Mishiro, and T. Oda (ed.), Viral hepatitis and liver disease. Springer-Verlag, Tokyo, Japan.
13. **Nakano, T., L. Lu, P. Liu, and O. G. Pybus.** 2004. Viral gene sequences reveal the variable history of hepatitis C virus infection among countries. J. Infect. Dis. **190:**1098–1108.
14. **Nguyen, Q. T., M. Leruez-Ville, F. Ferriere, P. Cohen, D. Roulot-Marullo, T. Coste, P. Deny, and L. Guillevin.** 1998. Hepatitis C virus genotypes implicated in mixed cryoglobulinemia. J. Med. Virol. **54:**20–25.
15. **Patel, J., A. H. Patel, and J. McLauchlan.** 1999. Covalent interactions are not required to permit or stabilize the non-covalent association of hepatitis C virus glycoproteins E1 and E2. J. Gen. Virol. **80:**1681–1690.
16. **Posada, D., and K. A. Crandall.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817–818.
17. **Pybus, O. G., M. A. Charleston, S. Gupta, A. Rambaut, E. C. Holmes, and P. H. Harvey.** 2001. The epidemic behavior of the hepatitis C virus. Science **292:**2323–2325.

18. **Ross, R. S., S. O. Viazov, C. D. Holtzer, A. Beyou, A. Monnet, C. Mazure, and M. Roggendorf.** 2000. Genotyping of hepatitis C virus isolates using CLIP sequencing. J. Clin. Microbiol. **38:**3581–3584.
19. **Seeff, L. B.** 2002. Natural history of chronic hepatitis C. Hepatology **36:**S35–S46.
20. **Simmonds, P.** 2004. Genetic diversity and evolution of hepatitis C virus—15 years on. J. Gen. Virol. **85:**3173–3188.
21. **Simmonds, P., J. Bukh, C. Combet, G. Deleage, N. Enomoto, S. Feinstone, P. Halfon, G. Inchauspe, C. Kuiken, G. Maertens, M. Mizokami, D. G. Murphy, H. Okamoto, J. M. Pawlotsky, F. Penin, E. Sablon, T. Shin-I, L. J. Stuyver, H. J. Thiel, S. Viazov, A. J. Weiner, and A. Widell.** 2005. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatology **42:**962–973.
22. **Stumpf, M. P. H., and O. G. Pybus.** 2002. Genetic diversity and models of viral evolution for the hepatitis C virus. FEMS Microbiol. Lett. **214:**143–152.
23. **Stuyver, L., A. Wyseur, W. Van Arnhem, F. Lunel, P. Laurent-Puig, J. M. Pawlotsky, B. Kleter, L. Bassit, J. Nkengasong, L. J. van Doorn, and G. Maertens.** 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.
24. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24:**4876–4882.
25. **Vega, I., R. Colina, L. García, R. Uriarte, C. Mogdasy, and J. Cristina.** 2001. Diversification of hepatitis C viruses in South America reveals a novel genetic lineage. Arch. Virol. **146:**1623–1629.
26. **World Health Organization.** 2000. Hepatitis C fact sheet no. 164. [Online.] http://www.who.int/mediacentre/factsheets/fs164/en/.