Multitypic Hepatitis C Virus Infection Identified by Real-Time Nucleotide Sequencing of Minority Genotypes

Andrew J. Buckton,1* Siew-Lin Ngui,1 Catherine Arnold,1 Katie Boast,2 Joanne Kovacs,2 Paul E. Klapper,3 Bharat Patel,4 Imam Ibrahim,5 Savita Rangarajan,2 Mary E. Ramsay,1 and Chong-Gee Teo1

Centre for Infections, Health Protection Agency, Colindale Avenue, London, United Kingdom; Haemophilia Reference Centre, St. Thomas’ Hospital, London, United Kingdom; West Yorkshire Health Protection Unit, Bridle Path, York Road, Leeds, United Kingdom; HPA Collaborating Laboratory, Department of Microbiology, North Middlesex University Hospital, London, United Kingdom; and Microbiology Department, Ealing Hospital, Uxbridge Road, Southall, Middlesex, United Kingdom

Received 9 August 2005/Returned for modification 4 November 2005/Accepted 12 June 2006

The prevalence of concurrent multitypic hepatitis C virus (HCV) infection is uncertain. A sensitive and specific approach to identifying minority HCV genotypes in blood is presented. Following serum extraction and reverse transcription PCR to amplify cDNA originating from the viral 5’ noncoding region, the amplified product mixture was treated with genotype-specific restriction endonuclease to digest the dominant genotype. Residual amplicons were subjected to PCR cloning and then to real-time DNA sequencing using a Pyrosequencer to identify the remaining genotypes. Dilution experiments showed that minority genotypes may be detected when they represent 1:10,000 of the total population and in serum specimens with viral loads as low as 1,000 IU/ml. Of 37 patients with bleeding disorders and 44 injecting drug users, infection by more than one HCV genotype was found in 7 (19%) and 4 (9%) patients, respectively. The low rate of detection in people at high risk of repeated HCV infection suggests that multitypic HCV carriage is uncommon.

Hepatitis C virus (HCV) exhibits substantial genetic variation (4). Six distinct viral genotypes and a variable number of subtypes have been described (3). Genotypes tend to segregate geographically. Thus, in England and Wales, genotype 1 (subtypes 1a and 1b) and genotype 3 (subtype 3a) are the most prevalent, while genotypes 4, 5, and 6 are not (8). As patients infected with different genotypes respond differently to antiviral drugs, identification of the infecting HCV genotype has become important to guide the choice of the mode and duration of antiviral therapy (24, 25).

Multitypic HCV infection has been reported in individuals at risk of multiple or repeated HCV exposures, e.g., patients with bleeding disorders given inadequately sterilized clotting factor concentrates (11, 16, 21) and injection drug users (IDUs) (1). Infection of a host with multiple HCV genotypes potentially complicates pre-antiviral therapy assessments. It has been proposed as a possible mechanism for nonresponse to alpha interferon-based antiviral therapy, as selection of minority genotype populations more refractory to antiviral treatment can emerge during treatment (18). Moreover, the precision of studies on the transmission of HCV also depends on whether or not minority genotypes can be tracked along with the dominant ones.

Genotyping has been carried out by identifying type-specific nucleotide sequence polymorphisms in products derived by consensus PCR sequencing of HCV cDNA (5, 19). This approach has limited value when applied to detect minority genotypes, as it lacks sensitivity. Furthermore, the predominant viral strain tends to be preferentially amplified (9). Thus, studies using these approaches to evaluate the prevalence of multitypic HCV infection likely lead to underestimates. Other approaches have been adopted to study multitypic HCV infection specifically (5, 9, 10, 18, 20–22). Those involving PCR cloning and sequencing are potentially the most effective to identify minority populations, but the impracticality and expense of sequencing large numbers of clones restrict their wider applicability.

We report the development of an approach to identify minority HCV genotypes circulating in people at risk of multiple or repeated exposures to HCV, namely, patients with bleeding disorders and IDUs. This approach involves the eradication of amplicons derived from the cDNA of the dominant genotype by genotype-specific restriction endonuclease digestion followed by real-time DNA sequencing to characterize PCR clones generated from residual, undigested amplicons.

**MATERIALS AND METHODS**

**Strategy to identify minority HCV genotypes.** The strategy to identify minority HCV genotypes is summarized in Fig. 1.

**Patient samples.** A panel of archived sera collected between May 2001 and July 2004 from 37 HCV-seropositive patients with bleeding disorders attending the Hemophilia Reference Center, St. Thomas’ Hospital, London, United Kingdom, was assembled. The bleeding disorders and number of patients were the following: hemophilia A, 26 (70%); hemophilia B, 7 (19%); von Willebrand’s disease, 3 (8%); and afibrinogenaemia, 1 (3%). Bleeding severity was scored as severe in 21 (57%) patients, moderate in 2 (5%), and mild in 10 (27%); the severity was unknown for 4 (11%). All patients were naive to anti-HCV chemotherapy, and their infecting HCV had earlier been genotyped by a protocol based on restriction-fragment length polymorphism (RFLP) analysis (15). Only individuals identified in this assay to be infected with genotypes 1a, 1b, or 3a and 5a were included in the study; those infected with other genotypes were not. Written informed consent was obtained from the patients agreeing to participate in the study. In addition to type of bleeding disorder and disease severity, the human
immunodeficiency virus (HIV) coinfection status was recorded for each patient. Another panel of study samples, collected between March 2002 and December 2003, was derived from unlinked anonymous sera from 44 HCV-seropositive patients identified as IDUs undergoing diagnostic hepatitis C testing from centers participating in the ongoing Hepatitis C Denominator Study of England and Wales. The IDUs were newly identified as HCV infected and had not been given anti-HCV chemotherapy. Only individuals infected with genotype 1a, 1b, or 3a by biochemical competence were included in the study. The two HCV-carrying sera used for sensitivity studies had their viral loads quantified using the COBAS AmpliPrep/COBAS Taqman HCV assay (Roche). The input volume of serum was 1 ml. This approach has a dynamic range between 15 IU/ml and 10^7 IU/ml.

**Cloning and transformant screening.** The digested PCR products were subjected to a reaction to ligate to the TOPO-TA cloning vector (Invitrogen), according to the manufacturer’s instructions, and were transfected into chemically competent TOP10 Escherichia coli. Transformants were selected by overnight incubation on nutrient agar plates containing 50 µg/ml of ampicillin. Using DNA from whole cells as template, typically 15 colonies were screened using PCR amplification with 5'-biotinylated antisense oligonucleotide primer 299, whose sequence is described above, but with an initial incubation of 10 min at 80°C.

**Pyrosequencing and analysis.** Pyrosequencing was performed using a PSQ 96 sample preparation kit and a PSQ MA analyzer (Biotage, Uppsala, Sweden). The technique enables rapid detection of single-nucleotide polymorphisms and the reading of small lengths of nucleotides from the PCR product mixture by a process known as “sequencing by synthesis,” based on the quantification of inorganic pyrophosphate released following nucleotide incorporation into a growing DNA chain. The system uses coupled enzymatic reactions to monitor pyrophosphate released during nucleotide incorporation. Free pyrophosphate activates the enzyme cascade, resulting in light emission. Light signals proportional to the amount of pyrophosphate released are obtained by the concerted action of ATP sulfurylase and luciferase. During each minute-long cycle, unreacted nucleotides are degraded by apyrase, allowing iterative addition of nucleotides. The light signatures produced by the hydrolysis of luciferin are then recorded and plotted into a simple graphic form, known as a pyrogram, from which sequence data are obtained. A 42-nucleotide fragment of the viral 5’NCR, from positions −140 to −98 in the HCV genome (positions follow the system of Choo et al. [4]), constituted the fragment of analysis, as it encompasses genotype-specific base sequence polymorphisms. Oligonucleotide F-new (5’-CGGGTCC TTTTTGGG-3’) was annealed to 20 µl of the digested PCR product by heating for 2 min at 80°C and allowing it to cool to room temperature. The Pyrosequencer operator system enables the selection of a specific nucleotide dispensation order, adding defined nucleotides at each cycle which correspond to the region of interest. In order to determine this nucleotide dispensation order, an alignment of all the published full-length sequences representing the six genotypes (MegAlign 5.03; DNAs star Inc., Madison, WI) was assembled. Nucleotide sequencing was characterized and given genotype assignments following alignments to reference sequences and BLAST analysis (2).

**Determination of assay sensitivity using PCR amplicon titrations.** A 265-nucleotide PCR fragment of the 5’NCR (generated using oligonucleotide primers 57 and 321) was cloned from a serum sample known to be carrying genotype...
1a and another sample carrying genotype 3a HCV, both of which had been referred for genotyping. Individual clones were screened for inserts, and plasmid DNA was isolated from overnight cultures using the Miniprep spin kit (QIAGEN). From each clonal template further amplicons were generated using primers 126 and 299, and the resulting DNA content was spectrophotometrically quantified at 260 nm. Amplified product mixtures produced from the same template were pooled and resuspended to a concentration of 1 ng/μl in nuclease-free water (representing a weakly positive PCR amplification). A panel consisting of the type 1-amplified products diluted in graduation with the type 3-amplified products were prepared in the following series: 1:2, 1:10, 1:100, 1:1,000, 1:10,000, and 1:100,000 (each having a total volume of 50 μl). A similar panel comprising the type 3-amplified products prepared with graduating dilutions of the type 1-amplified products were also prepared. Each dilution mixture was treated with the appropriate restriction endonuclease to digest the dominant genotype, and the undigested product was subjected to the same cloning and pyrosequencing procedure as that described above.

**RESULTS**

**Genotype identification by pyrosequencing.** Representative pyrograms generated from study specimens that represent HCV genotypes 1 to 5 are shown in Fig. 2.

**Assay sensitivity.** Studies using the two dilution panels showed that, with the total amplicon concentration set at 50 μg/μl, the assay was capable of detecting minority genotypes at a ratio of 1:100,000. Each dilution mixture was exposed to the minority genotype detection procedure without replicate testing.

**DNA sequencing.** For each study sample from which a minority genotype was identified by pyrosequencing, the result was confirmed by nucleotide sequence analysis of PCR amplicons originating from one representative clone using the CEQ 2000 dye terminator cycle sequencing kit (Beckman-Coulter, Fullerton, CA).

**FIG. 2.** Characteristic pyrogram traces generated from 5′NCR of HCV belonging to different genotypes. The y axis shows signal intensity in arbitrary light units, and the x axis shows temporal dispensation of nucleotides in minutes. E, dispensation of enzyme mix. S, dispensation of lumogenic substrate.
ng/reaction, the minority genotype could be identified even when constituting 1:10,000 of the total population. This represents the lowest dilution at which at least one randomly selected clone contains an insert originating from the minority type. This level of sensitivity was achieved whether the genotype 1-amplified product mixture was diluted in the genotype 3-amplified product mixture or vice versa. A 10-fold dilution series of a mixture of two sera, with the serum carrying genotype 3a diluted in the serum carrying genotype 1a, was prepared. Fifteen randomly selected clones were analyzed at each dilution. From all test specimens representing an HCV genome concentration of 1,000 IU/ml or greater, one or more colonies could be found to contain a PCR insert derived from genotype 3a.

**Comparison of multitypic infection between the two study groups.** A total of 81 specimens were tested, 37 from patients with bleeding disorders and 44 from IDUs. A total of 29 patient specimens failed to amplify using the single-round amplification strategy and therefore required reamplification by nested PCR. The infecting genotype as determined by RFLP analysis of the 5’NCR showed a predominance of type 1 over type 3 infection in the bleeding disorders group (23/37; 62%) compared to the IDU group (17/44; 39%) \( (P = 0.003; \text{Fisher’s exact test}) \). Of the total number of specimens tested, 11 (14%) were found to carry more than one genotype. The difference in proportion of multitypic HCV carriage did not differ significantly between the samples from the bleeding disorders group (7/37; 19%) and the IDU group (4/44; 9%).

**Multitypic infection in patients with bleeding disorders.** Patients in the bleeding disorders study group were between the ages of 27 and 71 years (mean, 42 years) at the time of sampling. The dominant HCV genotype as determined by RFLP analysis of the 5’NCR was genotype 1a in 23 (62%) patients, genotype 1b in 8 (22%), genotype 3a in 5 (14%), and genotype 5a in 1 (3%). HIV coinfection had been diagnosed in 10 (27%). Carriage of multitypic HCV was found in the serum of seven patients (19%). The combinations of minority genotypes detected are presented in Table 1. Multitypic infection tended to occur in patients aged between 38 and 53 years rather than in those <38 years or >53 years (maximum likelihood regression analysis against a quadratic effect of age, \( P = 0.038 \)) and in hemophilia A patients than for other forms of bleeding disorders \( (P = 0.08; \text{Fisher’s exact test}) \).

**Multitypic infection in IDUs.** The age range of the IDU study group was from 22 to 66 years (mean, 40 years). Injecting drug use was reported as current in 8 (18%) patients, previous in 19 (43%), and unknown in 17 (39%). The dominant HCV genotype was determined as 1a in 17 (38%) patients, 1b in 7 (16%), and 3a in 20 (54%). Multitypic HCV carriage was identified in 4 (9%) samples, in one current and one former drug user and in two subjects of unknown drug use status. The combinations of minority genotypes detected are presented in Table 2.

**DISCUSSION**

The prevalence of concurrent multitypic HCV infection remains difficult to ascertain. In patients with bleeding disorders receiving blood products from donors who have not been screened for anti-HCV antibody and who accordingly are more prone to multiple HCV infection, the prevalence ranges from 0% to 25% \((10, 16, 20, 22)\). The wide variation reported may be explained by the variety of approaches for detecting mixed-genotype infection. Other than consensus PCR sequencing, approaches used to study the prevalence of multitypic HCV infection include the INNO-LiPA line probe assay \((17)\), type-specific PCR \((7, 20, 22)\), primer-specific mispair extension analysis \((9)\), sequence analysis of PCR clones \((6)\), and serotyping combined with genotyping \((18)\). These differ in terms of sensitivity in the ability to detect minority genotypes and of specificity in the assignment of genotypes.

In the approach described here, which uses real-time nucleotide sequencing to characterize PCR clones generated after amplicons from the dominant HCV genotype carried in serum have been digested by genotype-specific restriction endonuclease action, minority genotypes belonging to types 1 to 5 could be identified. Such an approach confers a number of advantages. First, once the dominant HCV genotype is eliminated, characterization of the PCR clones originating from residual genomes, which represent those of the minority genotype, rapidly leads to assignment of that genotype. Each pyrosequencing reaction takes approximately 40 min to complete, enabling the sequences of a large number of PCR-derived clones to be characterized in a single run without the need for direct nucleotide sequencing. While from each undigested amplicon we usually screen 15 clones to identify minority genotypes, up to 96 clones per pyrosequencing run may be screened. Second, the approach can identify more than one minority genotype at a time. Third, it is exquisitely sensitive compared to other methods used to detect multitypic infection. In a viral population consisting of dominant and minority genotypes, a consensus PCR sequencing approach can detect minority genotypes, if present, at >10% \((14)\), the INNO-LiPA line probe assay can detect minority genotypes, if present, at >3% \((12)\), heteroduplex mobility assay can detect minority genotypes, if present, at >2% \((23)\), and the semiautomated primer-specific mispair extension analysis can detect minority genotypes, if present, at >0.8% \((9)\); these detection levels contrast with the sensitivity levels obtained using the approach described here. We were

*TABLE 1. Distribution of minority HCV genotypes identified in sera of patients with bleeding disorders*

<table>
<thead>
<tr>
<th>Genotype detected by 5’NCR RFLP</th>
<th>No. of patients with minority genotype(s):</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1a</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1b</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3a</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5a</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

*TABLE 2. Distribution of minority HCV genotypes identified in sera of IDUs*

<table>
<thead>
<tr>
<th>Genotype detected by 5’NCR RFLP</th>
<th>No. of patients with minority genotype(s):</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1a</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1b</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3a</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5a</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
able to detect a minority genotype population of 1:10,000 when the overall viral load of the specimen was greater than 1,000 IU/ml. Some caution may need to be applied to accepting the stated sensitivity limit, as replicate testing of each dilution of the sera was not conducted. Fourth, specificity is assured, as the identity of the minority genotype detected is characterized from the sequence that is directly derived from the genome.

There are limitations to the use of this approach. First, it is qualitative: no information on the viral load of minority genotypes present in the sample can be determined. Second, specimens may carry minority genotypes that are still undetectable. Such a situation can arise in at least three ways. (i) The ability to detect minority genotypes following the digestion of the dominant genotype may relate to the quantity of the dominant genotype in the sample: high quantities may result in incomplete digestion by the given amount of restriction endonuclease added into the digestion reaction. (ii) The likelihood of detecting minority genotypes is directly proportional to the number of PCR clones available after cloning has been applied to the postdigation amplified product mixture. (iii) The restriction endonuclease digestion used to cleave the PCR product originating from the dominant genotype is not completely genotype specific. Thus, the use of MvaI to cleave amplicons from genotype 1 can also cleave minority populations belonging to genotype 6, and the use of HinfI to cleave amplicons from genotype 3 can also cleave certain subtypes from genotype 4. We base this assessment of genotype specificity on the alignment of a small but representative sample of published sequences of genotypes 1 to 5. The possibility that some amplicons will not be cleaved by restriction endonucleases remains. No appropriate restriction endonucleases to specifically cleave amplicons from genotype 2 or 4 are available; for this reason, patients carrying genotype 2 or 4 were excluded from the study. We had intended to study the association between multitypic carriage and the viral load of the samples, but the volume of the residual specimens was too small to permit this. Lastly, the region of sequence used by this technique is insufficient to discriminate between genotypes 1 and 6. For this reason, patients carrying genotype 6 were also excluded from the study.

Despite the application of a very sensitive approach to detect minority HCV genotypes, the prevalence of multitypic infection in the two study groups was found to be low even though they were at high risk of multiple and repeated HCV exposure, particularly among those with bleeding disorders. The proportion was higher among this group but did not differ significantly from that in IDUs. The approach to identify multitypic HCV infection described here potentially permits the complex epidemiology of HCV infection to be unraveled. Its further application allows an evaluation of how the persistence of minority HCV genotypes may be associated with poor response to anti-HCV therapy and of the extent to which infection by minority genotypes affects molecular epidemiologic assessments of HCV transmission events.

ACKNOWLEDGMENTS

This work was supported by a grant funded by the Department of Health, England.

We thank Lisa Brant for providing data relating to specimens originating from IDUs, David Gelb for support with statistical testing, and staff of the HIV/Hepatitis Reference Laboratory for HCV genotyping, William Irving of Queen’s Medical Centre, Nottingham, England, and Clive Taylor of Health Protection Agency North East, Newcastle, England, provided additional specimens.

REFERENCES

  Brechot, and V. Thiers. 1997. Hepatitis C virus genotypes in French haemo-
  of hepatitis C virus is probably a rare event. Infection 28:21–25.
  Simplified hepatitis C virus genotyping by heteroduplex mobility analysis.
  Persing, and the Collaborative Study Group. 1996. Hepatitis C virus geno-
  types in the United States: epidemiology, pathogenicity, and response to