Hepatitis C Virus Genotyping in Developing Countries: Results from Cuba, India and Turkey

Juan Roca, Masood Ahmad, Subrat Kumar Panda, Guillermo Padron, Shahid Jameel

1Division of Diagnostics and Immunotechnology, Center for Genetic Engineering and Biotechnology, PO Box 6162, Havana 10600, Cuba. Fax: (53-7) 21 8008; E-mail: diag@cigb.edu.cu
2Virology Group, International Center for Genetic Engineering and Biotechnology, PO Box 10504, Aruna Azaf Ali Marg, New Delhi 110067, India.
3Department of Pathology, All India Institute of Medical Sciences, New Delhi, India.

ABSTRACT

Infection by hepatitis C virus (HCV) results in liver disease with a high rate of associated chronicity. Significant genetic variation is seen among HCV isolates based on nucleotide sequence homologies, which has allowed grouping into a number of genotypes. However, most of the information about HCV genotypes is based on studies from developed countries with less information available from developing countries. Here we present the results of HCV genotype determinations in 128 sera from 74 patients in three countries: Cuba, India and Turkey. An established PCR-based genotyping method was optimized to accurately detect multiple genotypes in a given sample. Type II (1b) HCV, which correlates with more aggressive forms of the disease and lower response rates to interferon was most commonly found in the patients from the countries studied. While the majority of patients (58.1%) were infected with a single genotype, dual infections were also found frequently (28.4%). This report discusses the utility of a genotyping assay and the prevalence of genotypes.

Keywords: HCV genotypes, hepatitis C virus, PCR

Biotecnologia Aplicada 1999;16:88-92

RESUMEN

La infección por el virus de la hepatitis C (VHC) causa una enfermedad hepática caracterizada por una alta tasa de cronicidad. Se han observado considerables variaciones genéticas entre los diferentes aislamientos del VHC, basadas en la homología de sus secuencias nucleotídicas, lo que ha permitido su agrupamiento en genotipos. No obstante, la mayor parte de la información sobre los genotipos del VHC se basa en estudios realizados en países desarrollados, siendo menor la información procedente de países en desarrollo. En este trabajo se presentan los resultados de la determinación de genotipos de pacientes de Cuba, la India y Turquía. Se estudiaron 128 muestras de suero de 74 pacientes mediante el empleo de un método ya establecido, basado en reacción en cadena de la polimerasa, optimizada para detectar con certeza múltiples genotipos en una muestra. El genotipo del VHC más frecuentemente encontrado en los tres países fue el de tipo II (1b), que ha sido asociado con una enfermedad más agresiva y con menores tasas de respuesta al interferón. Aunque en la mayoría de los pacientes se encontró solo un genotipo (58.1%), la infección doble se observó también con frecuencia (28.4%). Se discute la utilidad de los ensayos de genotipaje y la prevalencia de los genotipos.

Palabras clave: genotipos del VHC, RCP, virus de la hepatitis C

Introduction

Hepatitis C virus (HCV) has been identified as the main cause of post-transfusion non-A, non-B (NANB) hepatitis and produces more than half of the sporadic NANB hepatitis cases [1, 2]. HCV is an RNA virus similar to flaviviruses and pestiviruses [1], and has a high spontaneous mutation rate with an estimated frequency of $10^{-4}$ mutations/nucleotide/year [3]. As a result, HCV exists as a heterogeneous group of viruses showing approximately a 70% of overall homology [4].

Four HCV genotypes (1-IV) were initially recognized, which differed from each other in $21\%$ of the approximately 9,500 nucleotides of the genome [5]. Based on these criteria, a fifth genotype (V) was later recognized [6]. According to Chan et al. [7] and Simmonds et al. [8], major HCV isolates are classified on a phylogenetic basis into types 1, 2 and 3, each one further divided into subtypes. In this classification, subtypes 1a and 1b correspond to genotypes I and II, respectively, 2a and 2b to III and IV, and 3a and 3b to V and VI. Other typing methods have also been proposed based on the nonstructural region NS5 [9] and the putative envelope glycoprotein E1 [10]. A universal nomenclature has recently been proposed [4] to clarify some of the confusions arising from the use of multiple genotyping systems.

Many investigators have tried interferon-alpha 2b (IFNα) for the treatment of patients with chronic hepatitis C. Any decision to administer expensive therapies like IFNα to patients from poor countries should be taken after a careful consideration of their effectiveness. Since the infecting genotype of HCV appears to be a good primary predictor of the response to interferon therapy [11, 12], it makes sense to evaluate the HCV genotypes prevalent in these countries. Furthermore, because of the different epidemiological distribution of genotypes [10] and their association with different clinical outcomes [13, 14],

© 1999 Elslos Scientiae
this will also serve as a predictive indicator of HCV disease in these regions.

Methods for genotyping have included genomic amplification and DNA sequencing of certain regions of the HCV genome, polymerase chain reaction (PCR) with genotype-specific primers, restriction fragment length polymorphism (RFLP) analysis of the PCR amplicons, differential hybridization, and serological genotyping [15].

Here we show the HCV genotyping results in sera of patients from three developing countries: Cuba, India and Turkey. We have based the HCV genotyping primarily on sequences within the core gene and therefore we have used the Okamoto nomenclature [4]. This is, however, supplemented with the Simmonds’ [7] nomenclature for the sake of clarity.

Materials and Methods

Patients

The patients studied in this report were identified by screening programs. The Cuban and Indian groups included clinically proven, acute and chronic liver disease patients, with only anti-HCV positive patients in the former. The three groups included patients at high risk for HCV infection. Samples from all patients were screened for HCV viremia by reverse transcription and subsequent PCR amplification (RT-PCR), based on the viral 5’ noncoding region. Only those positive for HCV RNA were then used for genotype determination.

RNA preparation

Viral RNA for RT-PCR was extracted from 50–100 μL of serum by the guanidinium isothiocyanate-acid phenol method [16], and resuspended in 25 μL of diethylpyrocarbonate-treated water.

Amplification of the 5’ noncoding region

For detection of HCV RNA by RT-PCR, a nested amplification was carried out using primers located in the highly conserved 5’ noncoding region of the HCV genome (HCV1-4) as described by Cha et al. [17].

Genotyping of HCV

Genotypes of HCV, designated as 1 (1a), 2 (1b), 3 (2a), 4 (2b) and 5 (3a), were determined by an RT-PCR method, based on the amplification of part of the HCV core gene. The reaction conditions used were exactly as described by Okamoto et al. [18, 19]. Briefly, the first step involved the amplification by RT-PCR of a fragment spanning nucleotides 480–751 (272 bp) with a set of universal primers (primers 256 and 186), deduced from conserved sequences. After 1 h at 42 °C for the RT step, the PCR was carried out for 35 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 1.5 min and extension at 72 °C for 2 min. One-tenth of the reaction product was re-amplified with a universal sense primer (primer 104), and a mixture of five type-specific antisense primers (primers 132, 133, 134, 135 and 339), resulting in amplified fragments of different sizes for the different genotypes. These are: type 1 (1a), 57 bp; II (1b), 144 bp; III (2a), 174 bp; IV (2b), 123 bp and V (3a), 88 bp. The second PCR was carried out for 30 cycles of denaturation at 94 °C for 1 min, annealing at 60 °C for 1 min and extension at 72 °C for 1.5 min.

Amplification, cloning, sequencing and analysis of the core gene

The RNA equivalent to 25 μL of serum was amplified as follows, except for the amplification profile which was 35 cycles of 94 °C for 1 min, 60 °C for 1 min and 72 °C for 2 min, for both PCR. The primers used for amplification were:

outer set:

5’-GAAGGCCTTGTAATGATGCTCCTAAAGGGTGGTCCGTGAGT-3’ (sense)
805’-ATCCCGTTGTACAGTTCAAGGCTGCTCT-3’ (antisense)

inner set:

298’-AGTGGGAGGCCTCGGTAGACCAGCGCATCAGGACAC-3’ (sense)
789’-CTGTGGCATAGTTCAAGGGGTACTCTCCAGACCGGGGACACC-3’ (antisense)

The amplified fragment of approximately 450 bp, encompassing the core gene, was cloned in the PCR-Script vector (Strategene, USA) by standard protocols [20]. Multiple independent clones were sequenced in both directions using the T7 sequencing kit (Pharmacia, USA) and the T3/T7 primers following the instructions of the manufacturer.

Serological tests

Anti-HCV testing was carried out by a second generation EIA (Abbot, USA), following the instructions of the manufacturer.

Results

Optimization of the multiplex PCR for HCV genotyping

On initial analysis of 50 HCV RNA-positive sera, only type II was detected using the published protocol for genotyping PCR [18, 19]. However, when the type II-specific antisense primer was omitted from the second stage PCR, other genotypes were also detected. Sera of a few patients were also found to contain multiple genotypes of HCV. Some of these sera were used to optimize the multiple detection of HCV genotypes.

First, we analyzed the effect of varying the annealing temperature in the second stage PCR (data not shown). None of the temperatures tested (50, 55, 60 and 65 °C) allowed the detection of types other than type II (1b) without nonspecific amplifications, and 60 °C provided the best result for type II (1b) detection. Other types could be detected unequivocally only when the type II-specific antisense primer was omitted. Two separate second stage amplification reactions, with and without the type II-specific primer were therefore carried out to determine the complete spectrum of HCV types. Figure 1 shows the genotyping results of one Cuban patient (C/E7) and two interferon-treated Indian patients (1/5 and 1/23). Patient C/E7 was positive for types II (1b) and IV (2b). Patients 1/5 and 1/23 were infected with multiple genotypes before interferon therapy. Six months after withdrawal of interferon therapy these patients showed the persistence of genotypes 1 (1/5) and II (1/5 and 1/23).

Sera from two patients (C8 and D8) were then used to optimize the primer concentrations (Figure 2). The RT and the first stage PCR were carried out with 0.5 mM of each of the universal sense and
antisense primers. For the second stage PCR, the universal sense primer was used at 1 mM. All type-specific antisense primers were used at 0.5 mM, except type II (1b) primer for which the concentration was varied. The selected annealing temperature was 60 °C. In patient D8, infected with HCV I (1a) and II (1b), both types were detected with 0.05 mM of type II primer, but only type II (1b) was seen when the concentration of this primer was increased to 0.1 mM. In patient C8, infected with HCV I (1a), II (1b) and V (3a), type V became undetectable even at 0.02 mM of type II primer, and type I (1a) became undetectable at 0.05 mM of this primer.

The original protocol was therefore modified for the analysis of the HCV genotypes reported here. The first stage RT-PCR was carried out as described with 0.5 mM of each of the universal sense and antisense primers. The second stage PCR was then carried out separately for type II (1b) and the other types. For type II (1b), 0.5 mM of each of the universal sense and type II-specific antisense primers were used. For the other types, 1 mM of the universal sense primer and 0.5 mM of each of the type-specific antisense primers were used.

HCV RNA and the anti-HCV response in patients

A total of 128 sera from three groups of patients was screened for the presence of HCV RNA. The results are summarized in Table 1. On ELISA screening, 27 (i.e. 61.4%) of the 44 Turkish hemophiliacs were anti-HCV positive, although a same individual was not always positive for HCV RNA and anti-HCV antibodies. Of the 36 hemophiliacs positive for either marker, 18 (i.e. 50%) had both markers, while 9 (i.e. 25%) were positive for either HCV RNA alone or anti-HCV antibodies alone.

Prevalence of HCV genotypes

Sera positive for HCV RNA based on RT-PCR with the 5'-noncoding region primers were then used to determine the infecting HCV genotype. Of the 74 patients studied, 7 (9.5%) were repeatedly positive for HCV RNA, but negative for the known genotypes of HCV based on the modified multiplex PCR method used here. These included 3 Cuban and 4 Turkish patients.

Table 2 shows the distribution of HCV genotypes in the three groups. While infection with a single HCV genotype was more frequent (41 of 74 cases, 58.1%), a large number of dual infections was also present (21 of 74 cases, 28.4%). Infection with more than two HCV genotypes was rare (3 of 74 cases, 4%). Among the three groups, genotype II (1b) was the most common, being present in 69 (93.2%) of the 74 patients. Of the other genotypes, type I (1a) was found in 7 (9.5%), type III (2a) in 1 (1.3%), type IV (2b) in 10 (13.5%), and type V (3a) in 16 (21.6%) cases. Among the patients infected with more than one HCV genotype, type II (1b) was always present, followed by type V (3a) (14 of 24 cases, 58.3%), type I (1a) (7 of 24 cases, 29.1%), type IV (2b) (6 of 24 cases, 25%) and type III (2a) (1 of 24 cases, 4.2%).

Figure 1. RT-PCR-based genotyping of three patients (C/E7, I/5 and I/23). For the second stage PCR two separate reactions were carried out, one with type II (1a) and the other with the rest of the primers (B). Patient I/5 and I/23 received IFN therapy. Sera from pre-treatment (I) and continuing treatment (II) or post-withdrawal (III) times were analyzed. MW represents the 50 bp molecular weight marker (Amersham, UK) with band sizes (from top to bottom) of 700, 600, 500, 400, 300, 200, 100 and 50 bp. C+: positive control.

A breakdown of the three geographic groups revealed similar genotype distribution patterns: type II (1b) showed the highest prevalence, types I (1a) and IV (2b) moderate prevalence, and type III (2a) was detected in only one case. The prevalence of type V (3a), however, showed significant differences among the three groups tested. This genotype was present in 12 of the 34 (35.3%) Cuban cases and 2 of the 13 (15.4%) Indian cases, but was absent from the Turkish cases. Interestingly, infection with type V (3a) was always found as a co-infection with type II (1b). Type III (2a) was detected in an Indian patient who also showed the presence of types I (1a), II (1b) and IV (2b).

Analysis of HCV from non-genotyped patients

In the screening used here, HCV in 7 patients could not be assigned to any genotype. Two of these patients, T/38 and T/40, belonging to the Turkish group, were selected for further analysis by sequencing >90% of the core region. As an internal control, two patients with previously assigned HCV genotypes, C/E14 with types I (1a) and II (1b) and C/F2 with types II (1b) and V (3a), were also submitted to the same analysis.

Of the previously nongenotyped cases, the HCV core region sequence from patient T/38 revealed a maximum homology (97.5%) to the corresponding region of type I (1a). Similarly, the sequence of patient T/40 showed a maximum homology (93.6%) to

Table 1. HCV positivity in the three groups.

<table>
<thead>
<tr>
<th></th>
<th>Cuban</th>
<th>Indian</th>
<th>Turkish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total samples tested</td>
<td>33 (42)</td>
<td>32 (42)</td>
<td>44 (44)</td>
</tr>
<tr>
<td>HCV RNA positive</td>
<td>30 (37)</td>
<td>11 (15)</td>
<td>26 (28)</td>
</tr>
<tr>
<td>Positive for genotypes</td>
<td>27 (34)</td>
<td>11 (15)</td>
<td>24 (24)</td>
</tr>
<tr>
<td>Negative for genotypes</td>
<td>03 (03)</td>
<td>00 (00)</td>
<td>04 (04)</td>
</tr>
</tbody>
</table>

*Numbers represent patient and sera.
type II (1b). The core region from patients C/E14 and C/F2, typed earlier, showed a maximum homology of about 95% to type II (1b) sequences. This was in agreement with the genotyping PCR results.

**Discussion**

In this study we have examined the distribution of HCV genotypes in three developing countries from different continents. For genotyping, we followed the method and classification scheme of Okamoto et al. [6, 18, 19]. The PCR-based genotyping protocol [18, 19] required modifications to detect the full repertoire of HCV genotypes present in a given serum. The most important modification was to amplify type II (1b) separately from all of the other genotypes. When amplified together, type II (1b) was able to conceal the detection of all of the other types present. There are two major reasons for this. First, it has been documented that serum concentrations of HCV type II (1b) RNA are generally higher than those of other HCV types [11, 12]. Since genotyping by PCR utilizes a common sense primer, a competition effect will take place. Second, analysis of the genotyping primers (Oligo 4.0 software, USA) revealed that, compared to the other primers, the type II-specific antisense primer had a lower propensity to form stem-loop structures and primer dimers. In other words, the type II primer was the most efficient of all, and in a multiplex situation should compete favorably for the preferential detection of HCV type II (1b). These proposals are supported by our experimental data (Figure 2) indicating that a 10 to 25-fold lower concentration of the type II primer was capable of completely concealing the detection of other HCV types in this multiplex PCR assay.

The rates of HCV positivity by PCR (5' noncoding region) varied quite significantly in the three geographic groups tested (Table 1). However, since the selection criteria were not uniform, this cannot be taken as a measure of generalized HCV positivity in the population. Evaluation of anti-HCV antibodies and HCV RNA by PCR in one same group (Turkish) showed equivalent rates of positivity, averaging 60%. This emphasizes the need of using both testing procedures for complete screening of HCV positivity in patients with liver disease.

On genotyping HCV in sera from the three groups, type II (1b) was found to be the most prevalent, being present in almost 95% of the cases. According to another study [10], this is true worldwide as well, though not at such high rate. HCV type II (1b) is also the dominant genotype found in the Japanese population, both in apparently healthy blood donors as well as in patients with NANB hepatitis [18]. In our geographic groups, type I (1a), the dominant genotype in America and Western Europe [10], showed a low prevalence. Type III (1b), another variant detected frequently in Japan, was almost absent in our geographic groups. However, nearly 35% of the cases studied here showed the presence of type IV (2b) and V (3a). In other studies [7, 9, 10], type IV (2b) was not found to be restricted to any particular geographic region, and type V (3a), originally detected in Thailand [6], was subsequently found in cases from diverse geographic regions including New Zealand, USA, Western Europe, Italy and Japan [10]. In the geographic groups studied here, type V (3a) was frequently detected in Cuba (38.2%) and India (20%), but not detected in Turkey.

Significant infection by more than one genotype variant of HCV was observed in the studied populations. This is to be expected as high risk patients were selected to determine the complete, rather than a representative spectrum of HCV genotypes present in these countries. The higher rates of multiple genotype infections found, compared to those seen in developed countries, may also be expected as strict screening of the blood supply is not always possible due to the expensive nature of HCV testing. Furthermore, the high population density in developing countries may play a significant role in community-acquired infections.

Of the 74 patients positive for HCV RNA in our study, 7 (i.e. 9.5%) could not be typed by PCR. Two

---

Table 2. Distribution of HCV genotypes in the three groups.

<table>
<thead>
<tr>
<th></th>
<th>Cuban</th>
<th>Indian</th>
<th>Turkish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single genotype cases</td>
<td>16</td>
<td>07</td>
<td>20</td>
</tr>
<tr>
<td>Type I</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Type II</td>
<td>14</td>
<td>07</td>
<td>20</td>
</tr>
<tr>
<td>Type III</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Type IV</td>
<td>03</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Type V</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Dual genotype cases</td>
<td>12</td>
<td>00</td>
<td>04</td>
</tr>
<tr>
<td>Type I+II</td>
<td>02</td>
<td>01</td>
<td>02</td>
</tr>
<tr>
<td>Type I+IV</td>
<td>10</td>
<td>02</td>
<td>02</td>
</tr>
<tr>
<td>Multiple genotypes cases</td>
<td>02</td>
<td>01</td>
<td>-</td>
</tr>
<tr>
<td>Type I+II+V</td>
<td>01</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Type I+IV+V</td>
<td>01</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Type I+II+III+IV</td>
<td>-</td>
<td>01</td>
<td>-</td>
</tr>
</tbody>
</table>

*Number of patients

*One patient showed a genotype switch during IFN therapy

Figure 2. RNA from the sera of two patients (C8 and D8) were amplified as described in Materials and Methods; for the first stage PCR. For the second stage, the primer concentrations were varied. D8, lanes 1-3: 0.05 and 0.1 mmol of the type II (1b)-specific primer, respectively. C8, lanes 1-7: 0.03, 0.05, 0.1, 0.2, 0.3 and 0.5 mmol of the type II (1b)-specific primer, respectively. MW: 50 b.p DNA molecular weight marker (Amersham, UK).
of these, T/38 and T/40, were further studied by sequencing almost the complete core region. Based on the overall sequence homology, these were assigned to type I (1a) and type II (1b), respectively. The reason why these could not be typed by the PCR method could be partially explained by nucleotide sequence alignments in the region of the genotype-specific primers. The core sequence from patient T/40 showed a T→G change at the 3' terminal position in primer 133, the type II (1b)-specific genotyping primer. However, in patient T/38 there was a complete homology between the core sequence generated and the type I (1a)-specific primer. This negative result may be a consequence of a low viremia in this particular patient. It has been our experience that while primers for the 5' noncoding region [17], the complete core region (Panda et al., unpublished) and the type II (1b)-specific [18] primers used in this study are highly efficient, the rest of the type-specific primers [18, 19] are not. This is to be expected since, in general, the core region of HCV is highly conserved between the different types, leaving very little room for designing efficient, yet type-specific primers. In such a situation, if the viremia is low it may not be possible to detect HCV of types other than type II (1b) by the multiplex PCR method. Alternatively, untypeable samples might represent new subtypes, therefore the rest of the patients that could not be typed must be further studied by DNA sequencing.

In conclusion, a quick PCR-based method [18, 19] was optimized to detect the complete spectrum of HCV genotypes in a given patient serum sample. The information obtained from sequencing of the core region suggests that while such a genotyping method may be useful for studying large clinical/population groups, it would still be prudent to sequence interesting and selected cases. Here we present the first report on HCV genotype status in three developing countries: Cuba, India and Turkey, from three different geographic regions. The results show significant rates of infection by more than one HCV variant in these countries. Among the genotype variants of HCV, type II (1b) showed a high prevalence in these countries. Since type II (1b) represents a variant that is more resistant to IFNα and has a higher propensity to cause chronic disease, this study predicts an increase in the HCV carrier pool in these regions in the future.

Acknowledgments

We thank Dr. Kezban Yalcinkaya from the Anadolu University Hospital, Eskisirchir, Turkey; Dr. Enrique Arias from Hermanos Ameijeiras Hospital, Havana, Cuba; Dr. Luis Rivera from Carlos J Finlay Hospital, Havana, Cuba and Dr. SK Sarin from the GB Pant Hospital, New Delhi, India, for sera from HCV patients. This work was supported by internal funds from the ICGEB, Delhi, India.

Received in July, 1998. Accepted for publication in October, 1998.