Molecular epidemiology of the hepatitis C virus in Brazil

Solange Busek¹ and Guilherme Oliveira¹²

¹Laboratório de Parasitologia Celular e Molecular, Centro de Pesquisas René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, MG, Brazil
²Programa de Pós-Graduação e Pesquisa, Santa Casa de Misericórdia de Belo Horizonte, Belo Horizonte, MG, Brazil
Corresponding author: G. Oliveira
E-mail: oliveira@cpqrr.fiocruz.br

Received November 27, 2002
Published March 31, 2003

ABSTRACT. Hepatitis C virus (HCV) is a major cause of liver disease throughout the world. The genome of this virus consists of approximately 10,000 bp and codes for 10 mature polypeptides. Genome sequence comparison has revealed the existence of six major genotypes and a large number of subtypes. The genotypes can be distinguished by whole genome or genome fragment sequencing, genotype specific amplification of a genomic region or PCR amplification, followed by hybridization or restriction digestion, among other methods. There is a markedly heterogeneous geographical distribution of the HCV genotypes in the world. Different genotypes have been linked to distinct clinical outcomes and to differences in the susceptibility of the virus to interferon treatment. Several studies have been conducted to determine the distribution of HCV genotypes among different groups of individuals in Brazil. Most of these studies indicate a higher prevalence of genotype 1, followed by genotypes 3 and 2. Differences in genotypes can affect serological detection as well as the clinical outcome of the disease and sensibility to interferon treatment. Further studies need to be conducted to determine the degree of differentiation of circulating HCV genotypes in different patient groups in Brazil.

Key words: Hepatitis C virus, Genotype, Molecular epidemiology
INTRODUCTION

Hepatitis C virus (HCV), a member of the Flaviviridae, is the leading cause of chronic liver disease worldwide. It is estimated that about 170 million people are chronically infected with the HCV (Boyer and Marcellin, 2000). The HCV genome is a positive-stranded RNA molecule of approximately 10,000 nucleotides, which contains a single uninterrupted open-reading frame that encodes a protein of about 3,000 amino acids (Choo et al., 1989, 1990). The structural region consists of core and envelope (E1 and E2) proteins, and a 3′ region codes the nonstructural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B), which after processing constitute 10 mature viral proteins (Figure 1).

![Figure 1](image-url). Genetic organization of hepatitis C virus. C: core, E: envelope, NS: nonstructural, NCR: noncoding region.

Sequence analysis of the HCV genome has revealed considerable variation (Simmonds et al., 1993). Some regions of the genome, such as E1/E2, are highly variable, while others, such as the 5′ noncoding region (NCR), contain invariant domains that are important for the initiation and control of polyprotein translation (Bukh et al., 1995; Duffy et al., 2002). The observed sequence variants of the complete genome or of genomic fragments have been used to classify HCV into six major groups, or genotypes (Simmonds et al., 1993). Other authors have classified HCV into more groups. Bukh and collaborators (1995) described nine genotypes and at least 30 subgroups. However, genotypes 1 through 6 are the most frequently found, and some genotypes contain several more closely related subtypes (McOmish et al., 1994; Forns and Bukh, 1999). Significant sequence differences have been found among these genotypes. For example, the amino acid sequences of the E1 protein of different HCV isolates varies by as much as 51% (Bukh et al., 1993).

The nomenclature proposed by Chan et al. (1992) reflects a two-tiered hierarchical genotypic division. The HCV types are designated by Arabic numerals in the order of discovery (genotypes 1, 2, 3, etc.) and subtypes are designated using these numerals, followed by lower case letters, also in order of discovery (subtypes 1a, 1b, 2a, 2b, etc.).

Reliable methods for determining the genotype of the HCV isolates are essential for diagnostic and epidemiological studies. A definitive determination of genotypes and subtypes can be obtained by sequence analysis of PCR-amplified genomic fragments. For example, by sequencing a fragment of 100 nucleotides within the E1 region it was possible to distinguish isolates representing genotypes 1 to 6, and 12 subtypes (Bukh et al., 1993). Other typing methods that are not based on DNA sequencing have been developed. Although faster and less expensive, these methods usually address only a limited number of the existing genotypes because they depend on distinguishing genotypes through a few specific nucleotide changes. Some methodologies are based on PCR amplification with universal primers followed by a second round of amplification with type-specific primers directed, for example, towards the core or the NS5B region (Okamoto et al., 1992, 1993; Chayama et al., 1993; Widell et al., 1994; Zusinaite et al., 2000). Other methods involve hybridization with type-specific probes deduced from the 5′ NCR (van Doorn et al., 1994), core (Viazov et al., 1994), E1 (Ravaggi et al., 1994), NS3 (Qu
et al., 2002) or NS5B regions (Takada et al., 1993). Finally, there are techniques based on
digestion with different restriction enzymes of the amplified genomic regions (RFLP), such as
the 5' NCR (Chan et al., 1992) or NS5B gene (Nakao et al., 1991), which generate banding
patterns that are genotype specific.

Recombinant proteins cloned from the prototype virus and synthetic peptides based on
the viral sequence have been used to detect HCV antibodies (Kuo et al., 1989; Hosein et al.,
1991). However, many samples still remain seronegative or indeterminate with the use of
commercial serological tests (Esteban et al., 1990). One of the possible reasons is the
immunological window that occurs after infection and before the generation of specific
antibodies (Grant et al., 2002). The degree of sequence variability found among HCV genotypes
would be expected to profoundly affect the antigenicity of many of the HCV protein epitopes,
as demonstrated by variations in the NSSA protein (Dou et al., 2002). It is possible that some of
these false-negative serological results are a consequence of infection by extreme sequence
variants of HCV that elicit an antibody response with limited or no cross-reactivity with the
peptide antigens used in serological assays. In fact, HCV variants that have markedly different
sequences compared to the original prototype HCV have been found (Chan et al., 1992).

The existence of different hepatitis C viral genotypes has been correlated with distinct
clinical disease syndromes associated with infection. The variations in clinical manifestations
may reflect underlying differences in the pathogenicity of the different types of virus (Chan et
al., 1992). HCV infection has also been associated with other diseases, including some autoimmune
diseases (Lunel, 1994), type II cryoglobulinemia (Agnello et al., 1992; Thiel et al., 2002) and
lymphoma (Hermine et al., 2002). Investigations of the molecular mechanisms involved in the
development of pathologies have greatly profited from the development of self-replicating RNA
culture systems (Lohmann et al., 2001) and from the use of transgenic mice. One such example
is the involvement of the core protein in the development of steatosis, which was determined by
targeting the activity of the microsomal triglyceride transfer protein and changes in VLDL
assembly and secretion (Perlemuter et al., 2002).

The quasispecies character has been reported to affect the host immune system escape
mechanism (Kato et al., 1993; Taniguchi et al., 1993; Bukh et al., 1995; Sakai et al., 1999) and
also to be a predictive factor for sensitivity to interferon therapy (Moribe et al., 1995; Hopf et
al., 1996; Sakai et al., 1998), in which case genotype 1 viruses are less likely to respond to
treatment than individuals infected with genotype 2 or 3 viruses (Boyer and Marcellin, 2000).
The molecular mechanism involved in interferon resistance may involve the interaction of the
interferon sensitivity-determining region (ISDR), of the NS5A protein, with PKR protein
kinase (Gale et al., 1998). Sequence variations in the ISDR region have been correlated with
susceptibility to treatment (Brecha, 1999); however, the genetic profile of the infected human
population may also play a role (Nakano et al., 1999).

Several authors have reported that the distribution of the different genotypes of the
HCV vary geographically. Genotypes 1, 2 and 3 are predominantly found in Europe, Japan and
the United States (Bukh et al., 1995), genotype 4 in Central and North Africa and the Middle
East (McOmish et al., 1994), genotype 5 in South Africa (Smuts and Kannemeyer, 1995),
genotype 6 in Hong Kong (McOmish et al., 1994), and the usually rare genotypes 7, 8 and 9, in
Vietnam (Takada et al., 1993).

The possibility that distinct HCV genotypes could affect diagnostic assays, the natural
history of the disease and the effectiveness of treatment, as well as knowledge about the distinct

geographical distribution of the genotypes worldwide, has prompted us and other investigators to investigate the HCV genotype distribution in Brazil. Nevertheless, there are few published studies. Investigations made in the States of São Paulo and Rio de Janeiro have demonstrated the predominance of HCV genotypes 1, 3 and 2 (Bassit et al., 1994; Stuyver et al., 1995). HCV genotype 1 is the most common among hemodialysis and hemophiliac patients in the State of Minas Gerais. However, the second most frequent genotype observed among hemophiliac patients was genotype 3, and for hemodialysis patients it was genotype 2 (Oliveira et al., 1999a; Busek et al., 2002). This prevalence is not different from those previously reported from this same state and from northern Brazil (States of Acre, Pará, Amazonas, Bahia, Rio de Janeiro and São Paulo) with a predominance of genotype 1, followed by genotype 3 (Martins et al., 1998; Oliveira et al., 1999b; Pereira et al., 2002). The same observation was made by other authors who used the line probe assay directed towards the 5' NCR or genotype specific primers towards the core region (Martins et al., 1998; Parana et al., 2000; Silva et al., 2000). Studies of the E1 region sequence in a few isolates from Rio de Janeiro identified subtypes 1a and 1b (Peig Ginabreda et al., 1997). A higher prevalence of genotype 1b was observed in another study (Pereira et al., 2002). In contrast, genotype 3 was the most prevalent, followed by genotypes 1 and 2, in infected individuals in southern Brazil (Krug et al., 1996). A high prevalence of genotype 3 was also observed in the State of Ceará; however, this may be due to nosocomial transmission as the 10 patients tested were from a hemodialysis clinic (Oliveira et al., 1999b). Isolated reports of unusual genotypes have also been published (Bassit et al., 1999; Levi et al., 2002). Summing the results, the general distribution of HCV genotypes in the anti-HCV-positive Brazilian population (Figure 2) is similar to that found in western Europe and the United States, with a generally higher frequency of genotypes 1 and 3 (Schreier et al., 1996).

---

Figure 2. HCV genotype distribution in the different regions of Brazil. In the northern, northeastern and southeastern regions genotype 1 is the most frequent, followed by genotypes 3 and 2. In the southern regions genotype 3 is the most frequent, followed by genotype 1. An isolated higher prevalence of genotype 3 was observed in the State of Ceará in the northeastern region.
CONCLUSIONS

Although further studies of the distribution of the HCV genotypes in Brazil need to be carried out, it is also important to correlate the genotypes with the clinical outcome of the disease. We have not, so far, detected any such correlation in several distinct Brazilian HCV-infected groups of individuals (Oliveira et al., 1999a; Carino et al., 2002; Busek et al., 2002). This may be affected by how the population was selected in relation to the geographical area, clinical follow-up or treatment regimen (Brechot, 1999). The investigation of the prevalence of HCV genotypes could also be used for investigations of nosocomial transmission (Busek et al., 2002). In addition, there are indications suggesting that there are genotypes circulating in the Brazilian population with significant genomic differences from HCV studied in other parts of the world (Oliveira et al., 1999a). We are currently investigating the possible consequences of this genomic variability for serological diagnosis. In conclusion, the interaction between the HCV types and the human genome, taking into consideration the peculiarities of the Brazilian population, needs to be studied at greater depth in order to determine possible consequences for dissemination, susceptibility to interferon treatment, clinical evolution and diagnosis of HCV infection.

ACKNOWLEDGMENTS

Research supported by FIOCRUZ and FAPEMIG.

REFERENCES


cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 244: 359-362.


Grant, P.R., Sims, C.M. and Tedder, R.S. (2002). Quantification of HCV RNA levels and detection of core protein in donations before seroconversion. Transfusion 42: 1032-1036.


