Quasispecies of hepatitis C virus genotype 1 and treatment outcome with Peginterferon and Ribavirin

Ana Carolina Gomes Jardim, Lilian Hiromi Tomonari Yamasaki, Artur Trancoso Lopo de Queiróz, Cíntia Bittar, João Renato Rebello Pinho, Claudia Márcia Aparecida Carareto, Paula Rahal, Isabel Maria Vicente Guedes de Carvalho Mello

A R T I C L E  I N F O

Article history:
Received 14 December 2007
Received in revised form 4 November 2008
Accepted 5 November 2008
Available online xxx

Keywords:
Hepatitis C virus
Genotype 1
Quasispecies
Nonstructural 5A region
Pegylated interferon
Ribavirin
Sequence analysis

A B S T R A C T

The majority of patients with chronic hepatitis C fail to respond to antiviral therapy. The genetic basis of this resistance is unknown. The quasispecies nature of HCV may have an important implication concerning viral persistence and response to therapy. The HCV nonstructural 5A (NS5A) protein has been controversially implicated in the inherent resistance of HCV to interferon (IFN) antiviral therapy. To evaluate whether the NS5A quasispecies pre-treatment composition of HCV 1a/1b is related to responsiveness to combined pegylated interferon (PEG-IFN) and Ribavirin therapy, detailed analyses of the complete NS5A were performed. Fifteen full-length NS5A clones were sequenced from 11 pretreatment samples of patients infected with genotype 1 HCV (3 virological sustained responders, 4 non-responders, and 4 end-of-treatment responders). Our study could not show a significant correlation between the mean number of mutations in HCV NS5A before treatment and treatment outcome, and the phylogenetic construction of complete NS5A sequences obtained from all patients failed to show any clustering associated with a specific response pattern. No single amino acid position was associated with different responses to therapy in any of the NS5A regions analyzed, and mutations were clustered downstream the ISDR, primarily in the V3 region. We observed that the CRS and NLS regions of the NS5A protein were conflicting for some variables analyzed, although no significant differences were found. If these two regions can have antagonistic functions, it seems viable that they present different mutation profiles when compared with treatment response. The patient sample that presented the lowest genetic distance values also presented the smallest number of variants, and the most heterogeneous pattern was seen in the end-of-treatment patients. These results suggest that a detailed molecular analysis of the NS5A region on a larger sample size may be necessary for understanding its role in the therapy outcome of HCV 1a/1b infection.

1. Introduction

The hepatitis C virus (HCV) is one of the most frequent and important causes of chronic viral hepatitis and approximately 130 million people are infected with this virus worldwide (GBD, 2004). Chronic HCV infection was estimated to be involved in about 250,000–350,000 deaths per year, essentially related to severe liver cirrhosis, end-stage liver disease and hepatocellular carcinoma (Chevaliez and Pawlotsky, 2007).

HCV, the only Hepacivirus of the family Flaviviridae, is an enveloped virus with a positive, single-stranded RNA genome of approximately 9.6 kb, encoding a single polyprotein of about 3000 amino acids that is co- and post-translationally cleaved by viral and cellular proteases into structural and nonstructural proteins (Forns and Bukh, 1999).

Based on the nucleotide sequence diversity of the viral genome, HCV is classified into six genotypes, HCV-1 to -6, with each
genotype further subdivided into subtypes, such as HCV-1a and 1b (Simmonds et al., 1993, 2005). Furthermore, in infected individuals, HCV circulates as a population of many different but closely related viral variants, referred to as "quasispecies" (Martell et al., 1992; Domingo et al., 2006). New variants are continuously generated during viral replication, as a result of errors made by the viral RNA-dependent RNA polymerase, that lacks proofreading activity (Pawlotsky, 2003a).

Consensus meetings have proposed the combination of interferon (IFN) or pegylated IFN with Ribavirin (RBV) as treatment of choice for chronic hepatitis C. Nevertheless, even with this treatment regimen, in patients infected with genotypes 1a/1b, sustained virological response rates still hover at approximately 50% (Manns et al., 2001; Fried et al., 2002).

Various explanations have been forwarded for the differences in IFN-based therapy outcomes. A number of host parameters, disease characteristics, and virus-related factors play a role in the likelihood of permanent viral clearance after therapy (Pawlotsky, 2003b, 2005). Specific HCV proteins probably play a role in the treatment efficacy, as suggested by the fact that patients infected with different genotypes respond differently to therapy (Manns et al., 2001; Fried et al., 2002; Hadziyannis et al., 2004).

The nonstructural 5A (NS5A) protein is one of the nonstructural HCV proteins reported to be implicated in interferon resistance (Le Guillou-Guillemette et al., 2007). Many interactions between NS5A and cellular signaling pathways have been reported, and the interacting sequence has been identified for some of them (Macdonald and Harris, 2004).

Gale et al. (1997) identified in the NS5A C-terminal codons 2209 and 2274 a protein kinase R (PKR)-binding domain. They demonstrated in vitro that the NS5A protein contains a PKR-binding domain that is able to inhibit PKR activation. Therefore, the NS5A protein has been suggested to interfere with the interferon-induced intracellular antiviral pathway, leading to resistance to the interferon-based therapy.

The 40 first amino acids of the PKR-binding domain have been designated as interferon sensitivity-determining region (ISDR). In 1995, Enomoto et al. described a correlation between the number of mutations in ISDR and response to IFN therapy in genotype 1b-infected patients (Enomoto et al., 1995). These results were afterwards confirmed by other Japanese studies (Chayama et al., 1997; Arase et al., 1999; Nakano et al., 1999), but not by most West European (Duverlie et al., 1998; Berg et al., 2000; Sarrazin et al., 2000) and American studies (Chung et al., 1999; Nousbaum et al., 2000; Murphy et al., 2002), which found that most of the ISDR sequences harbored an intermediate resistance profile. A recent meta-analysis focusing on the number of mutations within NS5A ISDR confirmed the predictive usefulness of ISDR, but a geographic difference between European and Japanese patients was highlighted (Pascu et al., 2004).

Inchauspe et al. (1991) identified another domain localized in the C-terminal region of NS5A (codons 2356–2379), designated as the V3 domain. Some studies demonstrated a relationship between mutations in V3 and response to interferon therapy (Nousbaum et al., 2000; Hofmann et al., 2005).

Although full-length NS5A protein appears to be located exclusively in the cytoplasm, N-terminal deleted forms containing the nuclear localization signal (NLS) at aa positions 343–362 have been found in the nucleus, suggesting that such forms could move into the nucleus and act as potent transcriptional activators (Ide et al., 1996; Pawlotsky and Germanidis, 1998). Other experiments have indicated the presence of a 27-aa N-terminal cytoplasmatic retention signal (CRS) capable of sequestering NS5A in the cytoplasm with dominant activity over the NLS (Satoh et al., 2000). Similar results have demonstrated not only a function of N-terminal CRS, but also of C-terminal CRS (aa 353–447) and of a separate region (aa 27–38) to mask NLS activity (Song et al., 2000).

The nonstructural 5A region has become a hotspot to study, as it has been correlated to response to therapy and to a possible prognostic role. However, there is limited information about the genetic profile of the entire NS5A and different responses to IFN therapy. The previous studies have been based mainly on direct sequencing or analyses of a limited region of the NS5A gene, which does not provide precise information about viral quasispecies of NS5A. In the present study, we examined mutations in the entire NS5A region, including its PKR-binding domain, ISDR, V3 and the less studied regions NLS and CRS, of the HCV genotypes 1a and 1b from patients submitted to pegylated IFN plus Ribavirin therapy. The mutation profile of NS5A was analyzed, and HCV quasispecies of twelve patients were selected for analysis, based on their response to the treatment.

2. Methods

2.1. Patients

Twelve naïve patients infected chronically with HCV RNA genotypes 1a or 1b from the Hepatology Department of the São José do Rio Preto University Hospital were enrolled. Patients with other concomitant liver diseases [hepatitis B virus (HBV) or other hepatotropic virus infections, alcohol abuse, autoimmune hepatitis and hereditary liver diseases] were excluded. The Ethics Committee of the University approved the study, and written informed consent was obtained from all patients. Plasma samples were collected before starting treatment. During a 48-week treatment, the patients received once a week PEG-IFN-α-2b (according to body weight) subcutaneously and daily RBV, taken orally at a dose of 600–1200 mg (according to body weight).

The patients were classified into three groups, according to their response to therapy: four patients (1–4) showed sustained virological response (SVR); four patients (5–8) were non-responders (NR), and four patients (9–12) were end-of-treatment virological responders (ETR) (Table 1). SVR was defined as absence of HCV RNA in plasma by qualitative PCR 6 months after the end of the therapy. NR was defined as continued presence of HCV RNA in plasma during treatment and six months after. Patients who were HCV RNA-negative at the end of the therapy but experienced a rebound in HCV viremia afterwards were classified as ETR.

Demographic data, HCV genotypes and treatment responses of the studied patients, as well as the number of clones studied and the sequence fragment length for each patient, are shown in Table 1.

2.2. RNA extraction, RT-PCR and NS5A amplification

Total RNA was extracted from 140 μl plasma samples, using the commercially available QIAamp Viral RNA Kit (Qiagen, Uniscience, SP, Brazil). RNA was reverse-transcribed into cDNA using a High-Capacity cDNA Archive kit (Applied Biosystems, Foster City, CA, USA) and random primers. The mixture was incubated at 37 °C for 2 h. For amplification of the entire NS5A region of the HCV genome, a nested polymerase chain reaction (PCR) was performed. The first-round primer set consisted of degenerated forward primer NS5B151abS (5’-CTSCYGCCATCCTCCTC-3’, nucleotide [nt] position 5988–6004) and degenerated antisense primer NS5B215abA (5’-TMYACTCTTGAACC-3’, nt 7816–7799). The second-round primer set consisted of forward primer NS4B604abS (5’-GTGATGCGATGAACC-3’, nt 6087–6094) and degenerated antisense primer NS5B160abA (5’-AKGTACYYTCTTTCCG-3’, nt 7781–7744). For genotype 1b samples, which could not be amplified with the described primers, another approach was used.
with forward primer NS4B514bS (5′-TCCTCCATCTCTCC-3′, nt 5988–6004) and reverse primer NS5B215bA (5′-TCATCTCTTGTGAC-3′, nt 7816–7799) in the first round, and antisense primer NS5B160bA (5′-AGGTACCCCTCTGCC-3′, nt 7761–7744) in the second round. The standard sequence accessed in GenBank for which the primer positions are shown above is H77 (accession NC_004102.1).

### 2.3. Cloning and sequencing

The PCR products of approximately 1.7 kb were purified and ligated into the pCR-XL-TOPO-vector using the TOPO XL PCR cloning Kit (InvitrogenTM Life Technologies, Carlsbad, CA, USA). Fifteen transformants were randomly chosen for further studies, and plasmid DNA was isolated after an overnight 3.0 Milk broth culture using the QIAprep Miniprep kit (Qiagen). Recombinant pCR-XL-TOPO-NS5A clones were sequenced by dyeoxy terminator automated sequencing (ABI Prism Ready Reaction Mix; Applied Biosystems, Foster City, CA, USA) using an ABI Prism 377 sequencer, according to the manufacturers’ instructions (Applied Biosystems Inc., Foster City, CA, USA). Six sequencing reactions were performed for each clone, using flanking primers M13 Forward and M13 Reverse (Invitrogen TM Life Technologies, Carlsbad, CA, USA) and four internal forward primers covering the entire NS5A sequence, being NS5A1a267 (5′-CATTAACCCACAC-3′, nt 7351–7369), NS5A1a618 (5′-CTCCCTATACACAGGAG-3′, nt 6875–6893), NS5A1a823 (5′-GAGAACAAAGTGGTGATTC-3′, nt 7080–7098) and NS5A1a1094 (5′-AATCCACCAGTATCTGTC-3′, nt 7351–7369) for genotype 1a, and NS5A1b267 (5′-CATCAACGATCCACAC-3′, nt 6524–6541), NS5A1b622 (5′-CATACCATACAGCAGAG-3′, nt 6879–6896), NS5A1b872 (5′-AGGATGAGGAGGAATATCT-3′, nt 7129–7147) and NS5A1b1232 (5′-AGTCCATTCTCTCATGC-3′, nt 7489–7506) for genotype 1b. The standard sequence accessed in GenBank for which the primer positions are shown above is H77 (accession NC_004102.1).

### 2.4. Genetic and phylogenetic analysis of NS5A

All sequences obtained were submitted to Phred-Phrap-Consed programs (Ewing and Green, 1998; Ewing et al., 1998; Gordon et al., 1998). These programs analyzed the quality of the sequences and aligned them in one contig representing the entire NS5A sequence. Altogether, 163 full-length NS5A sequences were generated in this study. These sequences resulted from 15 clones of 9 patients, and 14 clones from patients 3 and 5. For patient 1, no good-quality N-terminal sequences of NS5A were generated, resulting in 15 sequences with 633 nucleotides, as shown in Table 1. The nucleotide sequence contigs (1344 nucleotides for genotype 1a and 1341 for genotype 1b) were aligned with the Clustal X program (version 1.81) (Thompson et al., 1997) and amino acid sequences were obtained. Primer sequences were removed from all sequences by using the BioEdit program (version 7.0.5.3) (Hall, 1999). Two nonsense mutations were detected at the NS5A codon 9 in two clones, one from patient 4 and the other from patient 9 (Table 1 and Fig. 1). The nucleotide and amino acid substitution numbers for each isolate and the genetic distance (p-distance) between pairs of sequences were calculated using the MEGA V4.0 software (Tamura et al., 2007). dN and dS, where dN is the frequency of non-synonymous substitutions per non-synonymous site and dS is the frequency of synonymous substitutions per synonymous site, were calculated by pairwise comparison of every single sequence using the SNAP—Synonymous Non-synonymous Analysis Program (Korber, 2000), that was also applied to calculate the dN/dS ratio. The obtained results were analyzed by two different approaches, between samples (bs) and within samples (ws). The between samples analyzes consisted of calculating the means of genetic distances, dS, dN and dN/dS ratios among all the clones of the response group, to obtain the group value. The within-sample analysis consisted of calculating the means of genetic distances, dS, dN and dN/dS ratios among the clones of one patient, and then calculate the mean of the four values obtained for each patient of the response group, to obtain the group value. The genetic diversity (Pi) analysis was performed using the software DnaSP version 4.5 (Rozas et al., 2003). To construct the phylogenetic tree of the NS5A variants obtained from the eleven patients, we used the PAUP* version 4 program (Swofford, 2002). Neighbor-joining phylogenetic tree was constructed with the HKY model of substitution, as determined by hierarchical likelihood ratio test score criteria in Modeltest 3.06 (Posada and Crandall, 1998). Base frequency, gamma distribution, and transition/transversion ratio were determined (from the data) by Modeltest 3.06. A thousand replicates were used to test reliability of the tree topology, and bootstrap values > 70 were considered significant (McCormack and Clewley, 2002). Genetic and phylogenetic analyses were performed using as reference the sequences NC004102 for genotype 1a and D50481 for genotype 1b, obtained from the GenBank.

### 2.5. Analysis of quasispecies diversity

For analyzing the quasispecies variability, software LOCQSPEC 1.0 was designed for our team (Marucci et al., in press). It is a tool

---

Table 1

Demographic data, HCV genotype, treatment response, number of clones studied and sequence fragment length of the studied patients.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Gender</th>
<th>Age</th>
<th>HCV genotype</th>
<th>Treatment response*6 months after end of therapy</th>
<th>No. of clones</th>
<th>No. of NS5A nucleotides</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>36</td>
<td>1a</td>
<td>SVR</td>
<td>15</td>
<td>633a</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>44</td>
<td>1b</td>
<td>SVR</td>
<td>15</td>
<td>1341</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>40</td>
<td>1a</td>
<td>SVR</td>
<td>14</td>
<td>1344</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>35</td>
<td>1a</td>
<td>SVR</td>
<td>15</td>
<td>1344</td>
</tr>
<tr>
<td>5</td>
<td>Male</td>
<td>41</td>
<td>1a</td>
<td>NR</td>
<td>14</td>
<td>1344</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>27</td>
<td>1b</td>
<td>NR</td>
<td>15</td>
<td>1341</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>52</td>
<td>1b</td>
<td>NR</td>
<td>15</td>
<td>1341</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>39</td>
<td>1a</td>
<td>ETR</td>
<td>15</td>
<td>1344</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>27</td>
<td>1a</td>
<td>ETR</td>
<td>15</td>
<td>1341</td>
</tr>
<tr>
<td>10</td>
<td>Male</td>
<td>56</td>
<td>1b</td>
<td>ETR</td>
<td>15</td>
<td>1344</td>
</tr>
<tr>
<td>11</td>
<td>Male</td>
<td>51</td>
<td>1b</td>
<td>ETR</td>
<td>15</td>
<td>1344</td>
</tr>
<tr>
<td>12</td>
<td>Female</td>
<td>25</td>
<td>a</td>
<td>ETR</td>
<td>15</td>
<td>1344</td>
</tr>
</tbody>
</table>

*a SVR, sustained virological response; NR, non-response; ETR, end-of-treatment response.

b Sequences without NS5A N-terminal region.

c One clone with a nonsense substitution.

---

Fig. 1. Schematic representation of the deduced amino acid sequence of the NS5A protein. Upper rows: sequences of HCV-1a isolates and the 1a reference sequence. Lower rows: sequences of HCV-1b isolates and the 1b reference sequence. The blue lines represent sustained virological responders, the pink lines non-responders, and the green lines end-of-treatment responders. The gray line represents the non-sequenced region from patient 1. The vertical lines represent amino acid changes. The CRS, PKR-bd, ISDR and NLS regions and the V3 domain are shown in boxes. Reference sequence on top for NS5A1a is NC004102 and for NS5A1b is D50481. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

built in C++ programming language that allows comparing a set of sequences, finding similar or different sequences as results. The 165 full-length NSSA sequences were submitted to the program, and the contigs that presented the same nucleotide or amino acid sequences were grouped together. Each group corresponds to a set of identical sequences which can be found more than once. The construction of each group started with an identity verification based on parity among sequences which have been grouped. The first sequence of the set was directly associated to the first group. For the other sequences, their identity was checked with regard to the entire set of existing groups. There was an optimization in this process, to stop the verification whenever a mismatch was found. Otherwise there was identity, the sequence was associated to the group, and the next sequence started to be verified. If all existing groups were verified and no identity was found, a new group was formed and the new sequence was associated to it. Thus, it was possible to determine the number of different and of identical quasispecies in the same sample.

2.6. Statistical analysis

Analysis of variance (ANOVA) was used to compare the variables means, and comparisons between groups were made by Fisher’s test for categorical variables and Tukey’s test for quantitative variables. The genetic distance and nucleotide diversity between groups were analyzed through the non-parametric Kruskal-Wallis test. The values of quantitative variables were expressed as means ± SD. Values of \( P < 0.05 \) were considered significant.

2.7. Nucleotide sequence accession numbers

The nucleotide sequence data reported here have been submitted to the GenBank nucleotide sequence database with accession numbers from EU309511 to EU309673, as follows: Patient 2 (EU309511 to EU309525), Patient 3 (EU309586 to EU309599), Patient 4 (EU309600 to EU309614), Patient 5 (EU309615 to EU309628), Patient 6 (EU309526 to EU309540), Patient 7 (EU309541 to EU309555), Patient 8 (EU309556 to EU309570), Patient 9 (EU309629 to EU309643), Patient 10 (EU309644 to EU309658), Patient 11 (EU309571 to EU309585) and Patient 12 (EU309659 to EU309673).

3. Results

3.1. Patient characteristics

The patients’ characteristics are summarized in Table 1. The non-response patients showed a higher mean age than the patients with other kinds of response to treatment (NR 42.5 vs SVR 38.75 vs ETR 36.75), the population was predominantly male (NR 75% vs SVR 25% vs ETR 25%), and patients infected with HCV genotype 1b were prevalent (NR 75% 1b, 25% 1a vs SVR 25% 1b, 75% 1a vs ETR 25% 1b, 75% 1a). However, there were no statistically significant differences between the groups for these characteristics.

3.2. Within-sample NSSA mutation profiles

A total of 165 pre-treatment full-length HCV NSSA sequence samples from 11 patients chronically infected with HCV 1a or 1b were investigated. The pattern of response to treatment (SVR, ETR or NR) was compared to pre-treatment quasispecies sequence features, as shown in Table 2 (sequences were compared with reference sequences for genotypes 1a or 1b). The differences between response groups were not significant when all variables were analyzed.

Table 2

HCV quasispecies in pre-treatment isolates from sustained treatment responders (STR), non-responders (NR) and end-of-treatment responders (ETR).

<table>
<thead>
<tr>
<th>Patients and groups</th>
<th>Nucleotide mutations</th>
<th>Amino acid mutations</th>
<th>Genetic diversity (Pi)(^{b})</th>
<th>Genetic distance(^{w})</th>
<th>Genetic distance(^{b})</th>
<th>ds(^{w})</th>
<th>ds(^{b})</th>
<th>dn(^{w})</th>
<th>dn(^{b})</th>
<th>dn/ds(^{w})</th>
<th>dn/ds(^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVR</td>
<td>111.2</td>
<td>30.9</td>
<td>0.0053</td>
<td>0.0050</td>
<td>0.1340</td>
<td>0.0161</td>
<td>0.3508</td>
<td>0.0018</td>
<td>0.0639</td>
<td>0.0064</td>
<td>0.0185</td>
</tr>
<tr>
<td>NR</td>
<td>108.1</td>
<td>27.6</td>
<td>0.0089</td>
<td>0.0088</td>
<td>0.1230</td>
<td>0.0268</td>
<td>0.3426</td>
<td>0.0030</td>
<td>0.0513</td>
<td>0.0035</td>
<td>0.0131</td>
</tr>
<tr>
<td>ETR</td>
<td>101.8</td>
<td>28.1</td>
<td>0.0130</td>
<td>0.0130</td>
<td>0.1170</td>
<td>0.0416</td>
<td>0.3183</td>
<td>0.0039</td>
<td>0.0522</td>
<td>0.0279</td>
<td>0.0671</td>
</tr>
<tr>
<td>CBF</td>
<td>5.4</td>
<td>1.0</td>
<td>0.0011</td>
<td>0.0013</td>
<td>0.1030</td>
<td>0.0026</td>
<td>0.3062</td>
<td>0.0000</td>
<td>0.0504</td>
<td>0.0010</td>
<td>0.0033</td>
</tr>
<tr>
<td>NR</td>
<td>8.4</td>
<td>2.1</td>
<td>0.0038</td>
<td>0.0038</td>
<td>0.0920</td>
<td>0.0108</td>
<td>0.2855</td>
<td>0.0001</td>
<td>0.0419</td>
<td>0.0011</td>
<td>0.0030</td>
</tr>
<tr>
<td>ETR</td>
<td>4.1</td>
<td>1.2</td>
<td>0.0104</td>
<td>0.0105</td>
<td>0.0920</td>
<td>0.0458</td>
<td>0.3129</td>
<td>0.0002</td>
<td>0.0366</td>
<td>0.0011</td>
<td>0.0017</td>
</tr>
<tr>
<td>PVR binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVR</td>
<td>13.8</td>
<td>2.7</td>
<td>0.0027</td>
<td>0.0027</td>
<td>0.1200</td>
<td>0.0085</td>
<td>0.4322</td>
<td>0.0003</td>
<td>0.0717</td>
<td>0.0011</td>
<td>0.0035</td>
</tr>
<tr>
<td>NR</td>
<td>12.5</td>
<td>2.6</td>
<td>0.0076</td>
<td>0.0078</td>
<td>0.1110</td>
<td>0.0205</td>
<td>0.4720</td>
<td>0.0016</td>
<td>0.0527</td>
<td>0.0013</td>
<td>0.0049</td>
</tr>
<tr>
<td>ETR</td>
<td>12.8</td>
<td>2.8</td>
<td>0.0099</td>
<td>0.0098</td>
<td>0.0930</td>
<td>0.0354</td>
<td>0.3154</td>
<td>0.0019</td>
<td>0.0267</td>
<td>0.0014</td>
<td>0.0022</td>
</tr>
<tr>
<td>ISDR</td>
<td>7.1</td>
<td>1.4</td>
<td>0.0029</td>
<td>0.0030</td>
<td>0.1170</td>
<td>0.0078</td>
<td>0.3543</td>
<td>0.0002</td>
<td>0.0452</td>
<td>0.0011</td>
<td>0.0035</td>
</tr>
<tr>
<td>SVR</td>
<td>7.0</td>
<td>0.7</td>
<td>0.0079</td>
<td>0.0078</td>
<td>0.1120</td>
<td>0.0308</td>
<td>0.3871</td>
<td>0.0008</td>
<td>0.0285</td>
<td>0.0011</td>
<td>0.0028</td>
</tr>
<tr>
<td>NR</td>
<td>6.0</td>
<td>0.7</td>
<td>0.0052</td>
<td>0.0050</td>
<td>0.0840</td>
<td>0.0141</td>
<td>0.2816</td>
<td>0.0005</td>
<td>0.0256</td>
<td>0.0011</td>
<td>0.0017</td>
</tr>
<tr>
<td>NLS</td>
<td>3.0</td>
<td>0.6</td>
<td>0.0146</td>
<td>0.0145</td>
<td>0.1290</td>
<td>0.0265</td>
<td>0.2813</td>
<td>0.0033</td>
<td>0.0622</td>
<td>0.0012</td>
<td>0.0032</td>
</tr>
<tr>
<td>NR</td>
<td>1.8</td>
<td>0.3</td>
<td>0.0046</td>
<td>0.0045</td>
<td>0.1150</td>
<td>0.0153</td>
<td>0.2842</td>
<td>0.0000</td>
<td>0.0365</td>
<td>0.0010</td>
<td>0.0027</td>
</tr>
<tr>
<td>ETR</td>
<td>2.8</td>
<td>0.0</td>
<td>0.0141</td>
<td>0.0143</td>
<td>0.1230</td>
<td>0.0388</td>
<td>0.3496</td>
<td>0.0019</td>
<td>0.0229</td>
<td>0.0011</td>
<td>0.0017</td>
</tr>
<tr>
<td>V3</td>
<td>11.5</td>
<td>5.6</td>
<td>0.0073</td>
<td>0.0075</td>
<td>0.2510</td>
<td>0.0156</td>
<td>0.2673</td>
<td>0.0011</td>
<td>0.2443</td>
<td>0.0013</td>
<td>0.0035</td>
</tr>
<tr>
<td>NR</td>
<td>10.4</td>
<td>4.6</td>
<td>0.0108</td>
<td>0.0108</td>
<td>0.2430</td>
<td>0.0121</td>
<td>0.3045</td>
<td>0.0045</td>
<td>0.2203</td>
<td>0.0015</td>
<td>0.0060</td>
</tr>
<tr>
<td>ETR</td>
<td>10.1</td>
<td>5.7</td>
<td>0.0165</td>
<td>0.0165</td>
<td>0.2630</td>
<td>0.0350</td>
<td>0.2983</td>
<td>0.0074</td>
<td>0.2490</td>
<td>0.0015</td>
<td>0.0069</td>
</tr>
</tbody>
</table>

* The partial sequences obtained from patient 1 was considered.

** SVR, sustained virological response; NR, non-response; ETR, end-of-treatment response.

\(^{b}\) All differences were significant and \(^{w}\) within-sample; \(^{b}\) between-samples.

\(^{c}\) Significant differences between groups \( p < 0.05 \).

\(^{d}\) When this group was compared with other groups \( p < 0.05 \).

\(^{e}\) When ETR and SVR were compared \( p < 0.05 \).

\(^{f}\) When ETR and NR were compared \( p < 0.05 \).
Although significant differences were not observed, the number of nucleotide and amino acid changes observed was usually higher in HCV isolates from SVR than in those from NR and ETR in the whole NS5A (111.2 vs 108.1 vs 101.8 for nt and 30.9 vs 27.6 vs 28.1 for aa) and in the NLS domains (3.0 vs 1.8 vs 2.8 for nt and 0.6 vs 0.3 vs 0.0 for aa). The same profile was observed in the mean number of nucleotide substitutions in the PKR-binding (13.8 vs 12.5 vs 12.0) and the V3 (11.5 vs 10.4 vs 10.1) domains. However, in the V3 domain the amino acid mean was higher in the ETR (5.7) and SVR (5.6) than in the NR (4.6) group. Conversely to the other domains, CRS had more changes in NR than in the other groups (8.4 in NR vs 5.4 in SVR vs 4.1 in ETR) (Table 2).

One approach to quantify genetic variability between clinical isolates is to determine the mean genetic distance. In Table 2, the mean genetic distance of HCV-1a and HCV-1b pre-treatment isolates for full-length NS5A and other regions analyzed in this study are plotted. Although the within-samples differences were not statistically significant, the mean distance obtained for SVR was lower than those of NR and ETR for the entire NS5A (0.0050 vs 0.0088 vs 0.0130), and of regions CRS (0.0013 vs 0.0038 vs 0.0105), PKR-bd (0.0027 vs 0.0078 vs 0.0098), ISDR (0.0030 vs 0.0078 vs 0.0050) and the V3 domain (0.0075 vs 0.0108 vs 0.0165). As opposed to this finding, NLS showed a higher value for SVR than for NR and ETR (0.0145 vs 0.0045 vs 0.0143). Fig. 2 shows the mean genetic distances corresponding to eleven patients of this study. The lowest values were observed for NR in patients 6 and 7. The genetic diversity (Pi) was also estimated and the population genetic diversity values followed the genetic distance tendencies (Table 2).

We calculated the frequency of synonymous substitutions per synonymous site (dS), the frequency of non-synonymous substitutions per non-synonymous site (dN) and the dN/dS ratio. The dS and dN values were lower for SVR than for NR and ETR in the complete NS5A (0.0161 and 0.0018 vs 0.0268 and 0.0030 vs 0.0416 and 0.0039), in CRS (0.0026 and 0.0000 vs 0.0108 and 0.0001 vs 0.0458 and 0.0002), in PKR-bd (0.0085 and 0.0003 vs 0.0265 and 0.0016 vs 0.0354 and 0.0019) and in ISDR (0.0078 and 0.0002 vs 0.0308 and 0.0008 vs 0.0141 and 0.0005). The exception was the NLS region, that presented lower dS and dN values for NR than for SVR and ETR (0.0153 and 0.0000 vs 0.0265 and 0.0033 vs 0.0388 and 0.0019) (Table 2). The dN/dS ratio for NS5A or other regions were not significantly correlated with therapy outcome.

### 3.3. Between-sample NS5A mutation profiles

Conversely to the within-sample data obtained, the between-sample genetic distance presented higher values in SVR than in NR and ETR for all regions analyzed (NS5A 0.1340 vs 0.1230 vs 0.1170;...
CRS 0.1030 vs 0.0920 vs 0.0920; PKRbd 0.1340 vs 0.1110 vs 0.0930; ISDR 0.1170 vs 0.1120 vs 0.0840 and NLS 0.1290 vs 0.1150 vs 0.1230), except the V3 domain in which ETR showed the highest values (SVR 0.2510 vs NR 0.2430 vs ETR 0.2630) (Table 2). For all the analyzed regions the differences between the response groups were statistically significant (p < 0.05).

Regarding the synonymous and non-synonymous substitution frequency values, dN was higher for ETR in the V3 domain (SVR 0.2443 vs NR 0.2203 vs ETR 0.2490) and in the other regions higher for SVR than for NR and ETR (NS5A 0.0639 vs 0.0513 vs 0.0522; CRS 0.0504 vs 0.0419 vs 0.0366; PKRbd 0.0717 vs 0.0527 vs 0.0267; ISDR 0.0452 vs 0.0265 vs 0.0256 and NLS 0.0622 vs 0.365 vs 0.0365 vs 0.0299). All the differences presented were statistically significant according to the data shown in Table 2.

Despite statistically differences were found for dS values, no pattern was observed in any region analyzed (Table 2). Also, although significant differences regarding the response to therapy for the complete NS5A (p < 0.05) and regions PKR-bd (p < 0.05), ISDR (p < 0.05), NLS (p < 0.05) and for the ETR group in the CRS region (p < 0.05) were found, no significant correlation between the between-samples dN/dS ratio and therapy outcome could be shown (Table 2).

3.4. Mutation spots, phylogenetic topology and quasispecies variability

A schematic representation of the deduced amino acid changes that were detected in the different analyzed samples is shown in Fig. 1. Mutations were scattered over the full-length NS5A region, and no specific amino acid or motif related to a particular response to therapy was observed. Most of the mutations were concentrated downstream the ISDR, primarily in the V3 region.

Phylogenetic trees were constructed from 163 full-length NS5A sequences as described in Section 2. This analysis did not show any clustering of nucleotide sequences of viral isolates according to the type of treatment response. For most patients, the different HCV variants clustered monophyletically and grouped correctly to the corresponding genotypes 1a or 1b. Some of the patients seem to be epidemiologically closely related, since one patient does not cluster monophyletically, and for a few other patients their monophyletic clustering is not significantly supported (Fig. 3).

A bioinformatics tool made it possible to perform a complex analysis of quasispecies diversity. We employed the LOCQSPEC 1.0 software, designed for our team, to evaluate the number of NS5A quasispecies present in each sample. Fig. 4 depicts the LOCQSPEC 1.0 analyses for all patients, grouped according to treatment response. High variability of quasispecies was observed for the majority of the patients. Considering entire NS5A nucleotide sequences, the ETR group showed to be highly variable between the analyzed clones, 93.3–100% of variability for SVR, and 46.6–100% for NR. With regard to the amino acid sequences, the variability observed was 80–93% for ETR, 46.6–100% for SVR, and 20–93% for NR. The lowest number of quasispecies was observed in the samples from patients 6 and 7 for nucleotide sequences, and from patients 4, 6 and 7 for amino acid sequences. Considering the amino acid sequences, samples 4, 6 and 7 presented a predominant variant that represented 53%, 66% and 80% of the quasispecies population, respectively (Fig. 4).

4. Discussion

Whether antiviral resistance influences therapeutic outcomes in chronic hepatitis C remains an important and controversial issue. The molecular mechanisms underlying treatment failure in chronic infection are still unknown. Viral and host factors have been described as predictors of response to therapy. HCV genotype 1 was found to present a lower rate of sustained virological response to combined IFN and Ribavirin therapy (Poynard et al., 2000), and no significant difference in the virological response rates was reported between its subtypes (Wohnsland et al., 2007). Patients with the same genotype and a similar viral load present varying kinds of response (Medeiros-Filho et al., 2006). This suggests that there must be other factors at least as important as those described so far. In this respect, quasispecies complexity plays an important role in IFN resistance mechanisms (Salmeron et al., 2006).

Genetic variability has been studied in many regions of the HCV genome, mainly in regions E2 (Farn et al., 2000; Polyak et al., 2000; Ueda et al., 2004) and NS5A (Duverlie et al., 1998; Pavlotsky et al., 1998; Sarrazin et al., 2002; Fan et al., 2005). Studies of HCV quasispecies and their relation with the outcome of antiviral therapy have been based mostly on the interferon sensitivity-determining region (ISDR), and controversial results have been obtained (Chayama et al., 1997; Arase et al., 1999; Nakano et al., 1999; Berg et al., 2000; Sarrazin et al., 2000; Ukai et al., 2006; Murayama et al., 2007). However, few reports have focused the complete NS5A (Nousbaum et al., 2000; Fan et al., 2005; Puig-Basagoiti et al., 2005; Veillon et al., 2007).

Furthermore, the studies addressing NS5A and therapy outcome have not classified the end-of-treatment patients as a different response group. Yang et al. (2003) analyzed mutations in the NS5A PKR domain and E2 and compared the mutational profiles of responder and non-responder patients (which include complete non-responders and relapsers). In a study by Sarrazin et al. (2002), SVR, NR and ETR were defined as kinds of therapy response groups. Nevertheless, for mutational analyses, SVR and ETR were considered virological responders. Nousbaum et al. (2000) defined sustained virological response as absence of HCV...
RNA at the end of therapy and 6 months after the end of the therapy. However, all comparisons were based on the end-of-treatment response, grouping SVR and ETR together. El-Shamy et al. (2007) analyzed NS5A and therapy outcome based on the response at the end of the treatment, while other studies analyzed only the SVR and NR groups (Pawlotsky et al., 1998; Puig-Basagoiti et al., 2005; Veillon et al., 2007).

In the present study, we considered it important to analyze the entire NS5A and to classify the end-of-treatment responders as a different group for all analyses performed. Closer investigation of the pre-treatment amino acid sequences of the NS5A and CRS, ISDR, PKR-bd, NLS and V3 domains may be relevant to understand the variability of this region and the possible relations with responsiveness to treatment. In addition, there are no Brazilian studies on HCV quasispecies, and geographic factors may account for different sensitivities of HCV infection towards antiviral therapy (Pascu et al., 2004; Murayama et al., 2007).

We analyzed the variation of NS5A quasispecies before treatment in patients who presented different kinds of response to combined PEGIFN and Ribavirin therapy. In accordance with other studies, no single amino acid position or motif was associated with different responses to therapy in any of the NS5A regions analyzed, and mutations were clustered downstream the ISDR, primarily in the V3 region (Nousbaum et al., 2000; Sarrazin et al., 2002; Pellerin et al., 2004; Hofmann et al., 2005; Puig-Basagoiti et al., 2005; Kmiecik et al., 2006). A recent study showed a significant correlation between IFN effectiveness and a high number of mutations in the V3 domain, presenting a significant difference between SVR and NR (Veillon et al., 2007). In another study focusing the NS5A region, especially ISDR, PKR-bd and the V3 domain in patients submitted to combined IFN and Ribavirin therapy, the authors also found a greater accumulation of substitutions in the V3 domain than in other domains of NS5A (Puig-Basagoiti et al., 2005). Considering our results, these studies reinforce the hypothesis that this domain has a greater capacity to accumulate mutations (Inchauspe et al., 1991; Nousbaum et al., 2000; Puig-Basagoiti et al., 2005; Veillon et al., 2007). Although the biological function of the V3 domain is not yet completely clear, mutations in this domain may affect the interaction between NS5A and one or several IFN-induced antiviral effectors (Veillon et al., 2007).

In this study, no significant differences could be found among the response groups regarding the mutation rate of the different regions analyzed. Nonetheless, nucleotide and amino acid mutation rates were lower in the non-responders than in the responders and the end-of-treatment responders for almost all regions analyzed. Similar results were reported by other authors (Nousbaum et al., 2000; Sarrazin et al., 2002; Puig-Basagoiti et al., 2005). In our study, the CRS region showed contrasting data when compared to the other regions. Therefore, the mutation profile of these regions and the correlation between the presence of mutations and different kinds of therapy response were evaluated. We observed that the data obtained for the CRS and NLS regions of the NS5A protein were conflicting for some of the variables analyzed. If these two regions can have antagonistic functions, it seems viable that they may present a distinct mutation profile for different treatment responses. Mutations in this region can produce nuclear NS5A forms which can act as transcriptional activator factors, interfere with antiviral effects, and favor the persistence of the infection (Ide et al., 1996; Pawlotsky and Germanidis, 1999). Studies based on mutations in regions CRS and NLS are almost inexistent and more mutation analyses in these regions are necessary.

As also found in previous studies (Pawlotsky et al., 1998; Fan et al., 2005), our results show a usually higher proportion of synonymous substitutions per synonymous site than of non-synonymous substitutions per non-synonymous site. This finding suggests that quasispecies mutations in the region of interest were due mainly to random genetic drift. The dN/dS ratio showed significant differences in the entire NS5A and in the other domains. These results appear to be more interesting as part of a longitudinal study, reinforcing the need to compare the pre-treatment with the during- and after-treatment mutation profiles. More studies addressing HCV quasispecies in the Brazilian population are necessary.

In line with other studies (Farcì et al., 2002; Zekri et al., 2007), our study could not provide statistical evidence that the number of viral strains and the genetic diversity before treatment are correlated with treatment outcome. Consistent with these observations, the phylogenetic construction of complete NS5A sequences obtained from all patients failed to show any clustering associated with a specific pattern of response. In contrast, within the limitations of our data set, analysis of the quasispecies composition may yield important prognostic information in addition to the evolutionary studies of the quasispecies and therapy response.

When we compared the mean genetic distance obtained for each patient sample with the result of strain analyses, we observed that the same patients who showed lower genetic distance values also presented the smaller numbers of strains composing the sample. It seems consistent, given that the mean genetic distances of the same quasispecies were compared. The most heterogeneous pattern was seen in the end-of-treatment patients. These patients presented a high degree of viral diversity, with no single strain representing more than 20% of the total population, which is in accordance with the data of Farcì et al. (2002).

The role of the NS5A region variability in antiviral resistance has not been determined yet, and the results obtained by different authors have been contradictory. Japanese studies described a correlation between the number of mutations in ISDR and response to IFN therapy in genotype 1-infected patients (Chayama et al., 1997; Arase et al., 1999; Nakano et al., 1999), but their results were in disagreement with most West European (Duverlie et al., 1998; Berg et al., 2000; Sarrazin et al., 2000) and American studies (Chung et al., 1999; Nousbaum et al., 2000; Murphy et al., 2002). Murayama et al. (2007) found a clear correlation between ISDR mutation number and response to IFN therapy in the J-type (predominant in Japan) but not in the W-type (distributed worldwide) HCV. It has been suggested that the usefulness of ISDR for predicting the response to therapy is dependent on the geographic area. Additionally, the quasispecies studies have used different parameters to analyze their data, and these factors appear to influence the results obtained. The patients included in these studies were submitted to different treatment conditions, and the kind of response was defined in different manners. Moreover, most studies analyzing the implications of HCV quasispecies in the response to antiviral treatment were conducted in patients treated with IFN monotherapy, with only a few studies on patients submitted to combined PEGIFN and Ribavirin therapy.

In conclusion, our study could not show a significant correlation between NS5A quasispecies of the HCV genotype 1a/1b in pre-treatment isolates and different kinds of response to therapy that may be predictive of the outcome of the combined Peginterferon and Ribavirin treatment. Since our study may be underpowered to see such correlations, further longitudinal studies on a larger sample size, with similar therapeutic approaches such as pegylated and Ribavirin treatment are necessary to improve our understanding of the role of NS5A in the response to IFN. Data on HCV evolution under treatment are very scarce, and we hope that publishing our limited results may contribute to future meta-analysis on the association between HCV quasispecies variation and therapy response.

References


