Genetic variability of genotype 1 hepatitis C virus isolates from Tunisian haemophiliacs

Ahlem Djebbi, Olfa Bahri, Houda Langar, Amel Sadaou, Selma Mejri, Henda Triki

Laboratory of Clinical Virology, Institut Pasteur of Tunis, Tunisia

INTRODUCTION

Hepatitis C virus (HCV) infection is frequently observed in haemophiliacs, and many cohort studies have reported chronic liver disease and high prevalence of anti-HCV seropositivity among these patients (Brettler et al., 1990; Morfini et al., 1994). The majority of these infections occurred after the introduction of clotting factors concentrates in the early 1970s and before the development of effective inactivation procedures for these products in the mid 1980s (Kasper and Kipnis, 1972; Hasiba et al., 1977). HCV is a positive-stranded RNA virus with significant genomic variability. Sequencing of viral variants isolated from distinct geographical regions in the world has allowed the classification of the virus into at least 6 genotypes and 70 subtypes (Robertson et al., 1998). A correlation between HCV genotypes and different clinical and epidemiological features was established. Patients infected with genotype 1b are generally less sensitive to antiviral therapy than those infected with the other genotypes (Yoshioka et al., 1992; Davis, 1994; Tsubota et al., 1994). Some genotypes have a worldwide distribution with a prevalence varying from country to country (i.e., genotypes 1 and 2) while others seem to be restricted to distinct geographical areas, such as genotypes 4, 5 and 6 (Zein, 2000).
Previous studies in Tunisia, conducted on both healthy and HCV infected subjects, reported a large predominance of genotype 1b with prevalences ranging from 79 to 90% of circulating genotypes (Djebbi et al., 2003; Ben Moussa et al., 2003; Mejri et al., 2005). Little has been published on HCV infection in haemophiliacs. The present work aimed to study HCV prevalence and genotypes in 95 haemophilic Tunisian patients. Sequencing in the NS5b region of the viral genome was conducted to assess phylogenetic relatedness patients’ isolates among each others and to HCV sequences previously reported from Tunisia and from other parts of the world.

MATERIALS AND METHODS

Patients
The study included 95 Tunisian haemophiliacs (87 with factor VIII deficiency and 8 with factor IX deficiency), all males, followed in the Care Unit of Haematology of Aziza Othmana Hospital of Tunis and originated from 12 different districts of Tunisia. Patients ages ranged from 5 months to 56 years (mean =19 years). They were divided into 3 groups with regard to the years 1985 and 1994, date of adoption of effective inactivation procedures for clotting factors and date of introduction of systematic HCV screening in Tunisian blood donors, respectively. Group 1 (n =45) comprises patients born before 1985, Group 2 (n =34) includes haemophiliacs born between 1985 and 1994, and Group 3 (n = 16) includes patients born after 1994. Most of patients harbour severe haemophilia forms and received cryoprecipitates locally prepared from Tunisian blood donors and commercial concentrates of factor VIII or IX imported from Europe. Samples were collected in 2003, divided into separate aliquots and stored at -80°C.

Diagnosis of HCV infection
Primary screening for HCV infection was performed using the 4th generation serological commercial enzyme immunoassays (Murex anti-HCV version 4.0 from Abbott Murex Biotech). ELISA anti-HCV positive samples were assessed by immunoblot using the Wellcozyme HCV Western blot VK68 (Abbott Murex Biotech) to confirm the ELISA seropositivity. HCV RNA was also assessed by PCR using the commercial Cobas Amplicor HCV test (Hoffman-LaRoche AG-Germany).

Genotyping HCV isolates
Genotypes were assessed using the INNO LiPA HCV version II (Innogenetics-Belgium) hybridization test and partial sequencing in the NS5b genomic region. Viral RNA was first extracted from all sera tested positive by the Cobas amplicor HCV test, using the commercial QIAamp viral RNA Mini Kit (QIAGEN). Complementary DNA (cDNA) was prepared from the extracted RNA by reverse-transcription using random hexamers (Pharmacia) and MMLV reverse transcriptase (Gibco BRL). The INNO LiPA Hybridization test was then performed on the obtained cDNA according to the manufacturer’s instructions; the assay system is based on hybridization of labelled PCR products from the 5’UTR (un-translated region) with oligonucleotide probes immobilized as parallel lines on membranes strips. Partial sequencing in the NS5b region also used the same cDNA. A nested PCR was conducted, using previously published primers (Thiers et al., 1997; Desenclos et al., 2001). Amplicons consistent with the expected size of 364 pb were purified using the QIAquik Gel Extraction kit (QIAGEN) and subjected to direct sequencing on both strands using the cycle sequencing method (Big Dye Terminator, Applied Biosystems), with the same inner primers used for nested PCR, in an automated ABI PRISM 377 DNA Sequencer (Applied Biosystems). The obtained sequences were aligned with the reference HCV sequences representing the different HCV genotypes and subtypes, genotypes were assigned according to the grouping pattern of the isolates with the reference sequences.

Phylogenetic analyses in the NS5b region
The same sequences in the NS5b genomic region were also compared to each other and to the most related HCV published sequences from other countries. Blast and Clustal X programs were used for the selection of sequences from GenBank, phylogenetic analyses were performed with the software program Mega 2.1 and the Phylip program package, using the nucleotide Kimura-2 parameter for the distance estimation and the neighbour-joining method for tree constructions. Robustness of the tree was accessed
with branch supporting-values from bootstrap statistic analyses (1000 replicates). The HC-G9/1c strain (GenBank accession number D14853) was used as an out-group for subtypes 1a and 1b dendrogram constructions.

**Nucleotide sequences accession numbers**

The new sequences described in this report have been submitted to Genbank and can be retrieved under accession numbers AM279711 to AM279721 for Genotype 1a and AM279722 to AM279730 for Genotype 1b. The following sequences from GenBank were used for phylogenetic studies:

**Genotype 1a**: Tunisia: AF462275 to AF462282; USA: 00888-14 (AY683108), 02301-1 (AY682782), 04749-1 (AY682788), C107_91 (AB079693), CF_N55B_83 (AB079698), JL_NSSB_77 (AB079711), 00650-1 (AY682758), PR_N55B_78 (AB079720), 00282-12 (AY683099), 00624-1 (AY682757), 1D4W (AY100184), 01373-1 (AY682765); France: FrSSD27 (AJ291247), 1M1 (AY100156), Mart07 (AY257428), FrSSD54 (AJ291260), (FR.ILH5), (AJ231490); UK: 1GS2 (AY100144), 1GS7 (AY100149), 1UCL3 (AY100177); USA: 7400-93 (AF071979), 694-95 (AF071986), 700-95 (AF071986); France: patient C02 (AJ880706), patient C18 (AJ880720); Scotland: Scotland T1825 (L23440S); Genotype 1b: Tunisia: AF462283 to AF462297 and AY387600 to AY387639; France: patient DEL (X59934), MRS113 (AF516061), P15.1.Rcp (AF515987), P18.Don (AF515993), Mart05 (AY257426); USA: 11605 (AY689289), 04333-1 (AY682785); UK: BD268 (AY003968), AD17 (AY321569); Sweden: 177 (AF388495), 730 (AF388401); Spain: ns5bIC-02 (AY898898), ns5bIC-27 (AY898918); Italy: N19155 (AY944674), I19169 (AY944670); Japan: Uz27h (AB081038), Uz88 (AB081060); Germany: HCV-AD78 (AJ132996); Turkey: HCV-TR1 (AF483269); Russia: HIA331 (DQ001182); China: gz53162 (AY835023).

**Statistical analyses**

Statistical data analyses were performed using the Pearson’s $\chi^2$ test.

**RESULTS**

**Prevalence of HCV infection**

Among the 95 serum samples tested, antibodies to HCV were detected by ELISA in 53 samples. The seropositivity was confirmed by immunoblot and/or PCR in 48 cases: 37 samples were positive by immunoblot and PCR, 6 samples were immunoblot positive with no detectable RNA by PCR and 5 samples had an indeterminate immunoblot with positive PCR test (collected from 2 HIV/HCV and 3 HGV/HCV coinfected patients). The remaining 5 ELISA positive samples were negative with both immunoblot and PCR. Consequently, the overall prevalence of confirmed HCV infection in our haemophiliac patients was 50.5% (48 out of 95) (Table 1). Thirty-four HCV positive patients belonged to Group 1 (born before 1985), 14 to Group 2 (born between 1985 and 1994) and all patients from Group 3 (born after 1994; n = 16) were HCV negative. Accordingly, the prevalence of HCV infection was 75.5% (34 out of 45) in Group 1 and 41.1% (14 out of 34) in Group 2 with a significant decrease from Group 1 to Group 3 (Group 1 vs Group 2: $p=0.00194$, Group 2 vs Group 3: $p=0.00055$) (Table 1).

<table>
<thead>
<tr>
<th>Patients</th>
<th>HCV infection nb (%)</th>
<th>Genotype 1a prevalence nb (%)</th>
<th>Genotype 1b prevalence nb (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1: born before 1985 (n = 45)</td>
<td>34$^b$ (75.5%)</td>
<td>18$^c$ (40%)</td>
<td>10$^d$ (22%)</td>
</tr>
<tr>
<td>Group 2: born between 1985 and 1994 (n = 34)</td>
<td>14$^b$ (41.1%)</td>
<td>2$^c$ (5.9%)</td>
<td>7$^d$ (20.6%)</td>
</tr>
<tr>
<td>Group 3: born after 1994 (n = 16)</td>
<td>0$^b$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total: (n = 95)</td>
<td>48 (50.5%)</td>
<td>20 (52.6%)</td>
<td>17 (44.7%)</td>
</tr>
</tbody>
</table>

$^a$Genotype determined by either INNO LiPA or partial sequencing; $^b$Group 1/Group 2 (p = 0.00194); Group 2/Group 3 (p = 0.00055); $^c$Group 1/Group 2 (p = 0.0014); $^d$Group 1/Group 2 (p = 0.86).
Genotypes
All samples with detectable HCV RNA by the Amplicor HCV PCR assay (n = 42) were assessed for HCV genotyping by sequencing in the NS5b and using the INNO LiPA commercial test which targets the 5'UTR of the HCV genome; the results are shown in Table 2.

Four samples did not give suitable amplification in the two targeted genomic regions, 12 supplemental samples did not amplify in the NS5b region and were analysed by the INNO LiPA test only, 26 samples could be analysed by both tests. Seventeen samples showed identical results in the NS5b and by INNO LiPA (Table 2, rows 1 and 4). For 5 samples, the INNO LiPA test identified the genotype but not the subtype which could be determined by sequence analysis in the NS5b region (Table 2, rows 3, 6, 8 and 9).

Four samples showed discordant results and included single (n=2) and multiple (n=2) genotypes (Table 2, rows 7, 12 and 14). INNO LiPA results were retained for these samples; probable recombination following the 5'NC region was admitted for the two first samples and mixed-genotype infections were assumed for the two others. Mixed infections were detected with the INNO LiPA in 15.7% of cases (6 out of 38), most of them (5 patients) were from Group 1. Globally, out of the 38 tested samples, genotypes 1a and 1b were the most frequent with 52.6% (20 out of 38) and 44.7% (17 out of 38) of prevalence respectively; including single and mixed-genotype infections. Subtype 1a was more frequent among patients from Group 1: 40% (18/45) and 5.9% (2/34) in Group 1 and Group 2 respectively (p=0.0014) (Table 1).

In contrast, the prevalence of subtype 1b was statistically similar between the two groups: 22% (10/45) and 20.6% (7/34) in Group 1 and Group 2 respectively (p=0.86).

Phylogenetic analysis in the NS5b region
The 14 sequences of genotype 1a and the 9 of genotype 1b obtained from the patients in the NS5b genomic region were compared to other sequences from the same subtypes published in GenBank. For each sequence, the 40 most related HCV sequences from other countries were

<table>
<thead>
<tr>
<th>Lines</th>
<th>INNO LiPA</th>
<th>Partial sequencing in NS5b region</th>
<th>N° of patients</th>
<th>Final genotyping result</th>
<th>N° of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1a</td>
<td>1a</td>
<td>11</td>
<td>1a</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>1a</td>
<td>NA</td>
<td>3</td>
<td>1a</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1a</td>
<td>1</td>
<td>1a</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>1b</td>
<td>1b</td>
<td>6</td>
<td>1b</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>1b</td>
<td>NA</td>
<td>4</td>
<td>1b</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1b</td>
<td>2</td>
<td>1b</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>1b</td>
<td>1a</td>
<td>2</td>
<td>1b</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>2c</td>
<td>1</td>
<td>2c</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>3a</td>
<td>1</td>
<td>3a</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>NA</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>1a + 1b</td>
<td>NA</td>
<td>2</td>
<td>1a + 1b</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>1a + 2</td>
<td>2c</td>
<td>1</td>
<td>1a + 2c</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1a + 2</td>
<td>NA</td>
<td>1</td>
<td>1a + 2</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>1a + 2 + 4</td>
<td>1b</td>
<td>1</td>
<td>1a + 2 + 4</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>1b + 3a</td>
<td>NA</td>
<td>1</td>
<td>1b + 3a</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>NA</td>
<td>NA</td>
<td>4</td>
<td>-</td>
<td>4</td>
</tr>
</tbody>
</table>

Total 42

NA: Not Applicable.
first selected from GenBank using the blastn program. All sequences were then pooled, repeated ones were removed and only those with the highest affiliations to the Tunisian strains were maintained for the phylogenetic tree construction. Figure 1 and Figure 2 show the results obtained for subtypes 1a and 1b, respectively. In addition to the sequences from other countries (31 of genotype 1a and 21 of genotype 1b), the trees also included previously published Tunisian sequences obtained from non-haemophiliac individuals (5 of subtype 1a, 52 of subtype 1b). The phylogenetic tree of subtype 1a (Figure 1) arranged into two distinct clades with high bootstrap resampling (98%).

**FIGURE 1** - Phylogenetic analysis in the NS5b region of sequences of subtype 1a. All Tunisian sequences are written in bold; the asterisk indicates those obtained from haemophiliac patients. Two clades can be differentiated: Clade 1aH for sequences obtained from haemophiliacs and Clade 1aNH for sequences obtained from non-haemophiliacs. All sequences from Tunisian haemophiliac patients grouped in the Clade 1aH which also included 30 non-Tunisian sequences, 24 of them are from USA. Numbers in the branches show the number of occurrences of the repetitive groups to the right out of 1,000 bootstrap samples.

**FIGURE 2** - Phylogenetic analysis in the NS5b region of sequences of subtype 1b. All Tunisian sequences are written in bold; the asterisk indicates those obtained from haemophiliac patients. There is no segregation between haemophiliac and non-haemophiliac strains; sequences from Tunisian haemophiliacs are randomly distributed throughout the tree among other Tunisian and non-Tunisian sequences. Numbers in the branches show the number of occurrences of the repetitive groups to the right out of 1000 bootstrap samples.
The clade designated “1aH” included all the sequences from Tunisian haemophilic patients (in bold and marked with an asterisk) and 15 non Tunisian sequences, most of them (13 out of 15) from the USA. Some isolates from haemophiliacs were related to each other as they formed small clusters of 3 to 4 sequences (Hem5/TUN04, Hem7/TUN04, Hem8/TUN04), (TN20, TN34 and Hem1/TUN04) and (TN25, TN28, Hem4/TUN04 and Hem6/TUN04).

Of the 14 sequences from Tunisian haemophiliacs, 10 were isolated from patients born before 1985 (Hem1/TUN04, Hem3/TUN04, Hem5/TUN04, Hem7/TUN04, Hem8/TUN04, TN16, TN22, TN25, TN28 and TN34), the remaining sequences were isolated from patients born after this date (Hem2/TUN04, Hem4/TUN04, Hem6/TUN04 and TN20).

The sequences from non haemophilic Tunisian patients grouped outside Clade 1aH, the most related sequences are from European countries: France, United Kingdom, Spain and Sweden. The phylogenetic tree of subtype 1b (Figure 2) was markedly different from the one of genotype 1a. No segregation between the sequences from haemophiliacs and non-haemophiliacs was observed. The sequences from haemophiliacs (in bold and marked with an asterisk) were randomly distributed throughout the tree among other Tunisian and non-Tunisian sequences; the most related sequences were from Tunisia in 5 cases (TN24, TN23, TN38, TN33, gv66/1TUN02) and from European countries in the 4 remaining cases (France, United Kingdom, Sweden and Germany).

DISCUSSION

The present study investigated the prevalence of HCV infection and the genetic characteristics of the infecting viruses in 95 Tunisian haemophilic patients. All patients have repetitively received cryoprecipitates locally prepared from Tunisian blood donors and imported commercial concentrates of factor VIII or IX. HCV infection was detected in 50.5% of cases. Studies conducted in the late 1980s and early 1990s reported higher prevalences of HCV infection in haemophiliacs, frequently exceeding 80% and reaching 95% in some reports (Ludlam et al., 1989; Makris et al., 1990; Blanchette et al., 1991).

By contrast, the rates reported in more recent studies were usually lower, between 27 and 71% (Beltran et al., 2005; Vinelli and Lorenzana., 2005; Silva et al., 2005). These study cohorts may include more patients born after the introduction of virus-inactivation methods of clotting factor concentrates in 1985 and of screening blood donors for anti-HCV in the early 1990s. In our study, we found a significant (p=0.00194) decrease of HCV prevalence between patients born before 1985 (75.5%) and those born after this date (41.1% in patients born between 1985 and 1994); no HCV positives were found among patients born after 1994, when systematic anti-HCV screening of blood donors was introduced in Tunisia. The absence of HCV infection among the 14 patients born after 1994 investigated herein indicates a markedly reduced risk of infection although a residual risk certainly remains.

The genotype of the HCV isolates infecting our haemophiliac patients was determined; subtype 1a was the most prevalent followed by subtype 1b, detected in 52.6% and 44.7% of patients, respectively. These features are different from those previously described in the general population and other patients in Tunisia.

Previous reports from Tunisia found a large predominance of subtype 1b which counted for 79 to 88% of cases and a very low prevalence of the other HCV genotypes and subtypes, subtype 1a was identified in only 1.5 to 5% of isolates (Djebbi et al., 2003; Ben Moussa et al., 2003; Mejri et al., 2005). Subtype 1a was more frequent among haemophiliacs born before 1985 than among those born after, 40% and 5.9%, respectively (p=0.0014). These results suggest a transmission via factor VIII and IX concentrates imported to Tunisia from Europe and usually manufactured from US blood donors among whom genotype 1a is most frequently detected (Preston et al., 1995; Mahaney et al., 1994).

The phylogenetic analysis of NS5b sequences shown in Figure 1 is consistent with this hypothesis; it shows a clear divergence between subtype 1a sequences infecting haemophiliacs from those obtained from non haemophiliacs as well as their affiliation to US isolates. Sequences of subtype 1a from the same cluster (Hem4/TUN04, Hem6/TUN04) were, nevertheless, found in some patients born after 1985, which alerts to a probable persistent infectivity of clotting factors despite
inactivation measures. Virus transmission between haemophiliacs in a nosocomial setting could also be the reason for such sequence clustering. In contrast, the prevalence of subtype 1b was statistically similar in haemophiliac patients of Group 1 (born before 1985) and Group 2 (born after 1985): 22% and 20.6%, respectively (p=0.86). This favours an infection through cryoprecipitates prepared from Tunisian blood donors. This hypothesis is further supported by the close relationship of subtype 1b sequences obtained from Tunisian haemophiliacs and non-haemophiliacs patients, as shown in Figure 2. Nucleotide sequence similarity of the Tunisian 1b sequences with European isolates was also observed and could be explained by the close geographical relationship and important socio-cultural interactions between Tunisia and European countries, especially France, in addition to the high number of Tunisian patients who receive medical care in Europe. It was previously reported that the worldwide dissemination of subtype 1b was provided by widescale medical intervention in the past (Smith et al., 1994).

In conclusion, this study reports the co-dominance of subtypes 1a and 1b in Tunisian haemophiliacs and suggests different transmission routes for the two variants: through imported clotting blood factors for subtype 1a and through locally-prepared cryoprecipitates for subtype 1b. A haemophiliac cohort study demonstrated that from country to country, HCV genotypes differ markedly in patients treated with locally prepared blood products, since they match the locally circulating genotypes. However, in haemophiliacs treated with commercial clotting factors, genotype 1, and especially subtype 1a, was frequently detected, independent of the patient’s country of origin (Preston et al., 1995). Our study also shows a significantly reduced risk of transmission of HCV infection in these patients with the introduction of effective virus inactivation methods of clotting factor concentrates in 1985 and of systematic screening for anti-HCV of Tunisian blood donors in 1995.

ACKNOWLEDGEMENTS
The authors are grateful to Emna Gouider, Aicha Hafsia and Balkis Meddeb (Laboratoire d’Hématologie - Hôpital Aziza Othmana de Tunis) and Raouf Hafsia (Service d’Hématologie Clinique de l’Hôpital Aziza Othmana de Tunis), for samples coordination and data collection. This work was supported by the European Community (Project: Viral Hepatitis Surveillance in Mediterranean Countries - Contract: ICA3-CT-2000-30011) and the Tunisian Secretariat for Scientific Research and Technology SERST (LR: Hépatites et maladies virales épidémiques).

REFERENCES


