Hepatitis C virus genotype distribution in Turkey remains unchanged after a decade: Performance of phylogenetic analysis of the NS5B, E1, and 5’UTR regions in genotyping efficiency

**ABSTRACT**

**Background/Aims:** Hepatitis C virus (HCV) genotyping has a considerable effect on therapy. The aim was to determine the change in prevalence of HCV genotypes in Turkey during the last decade and to compare the performance of DNA sequencing of different targets in the HCV genome (NS5B, E1, and 5’UTR).

**Materials and Methods:** Five hundred HCV RNA-positive patients (226 males, 274 females) were included in the study. The NS5B, E1, and 5’UTR regions of the HCV genome were amplified by polymerase chain reaction (PCR) in patients where possible. Amplified PCR products were sequenced directly, and phylogenetic analysis was performed. A commonly used database, namely www.hcv.lanl.gov, was also used to determine the genotypes.

**Results:** Phylogenetic analysis of the NS5B, E1, and 5’UTR regions showed that 1b was the most frequent genotype, with percentages of 92.5%, 93.5%, and 87.7%, respectively. Genotype 1a was the second most prevalent genotype, with ratios of 6.7%, 5.6%, and 6.6%, whereas genotype 2a was detected in proportions of 0.4%, 0.2%, and 0.8%, respectively. Genotype 5 or 6 was not detected among patients. The phylogenetic analysis showed discordant results with 18 patients’ genotypes for different targets. The phylogenetic analysis showed similar results with the hcv.lanl.gov database for the E1 and NS5B sequences.

**Conclusion:** There has been no change in genotyping profiles of Turkey during the last decade, representing 1b as the most prevalent subtype, followed by 1a. Phylogenetic analysis of HCV indicated high performance compared with the hcv.lanl.gov database when sequences of E1 and NS5B regions were analyzed.

**Keywords:** Genotyping, hepatitis C virus, phylogenetic analysis

**INTRODUCTION**

Hepatitis C virus (HCV) infection is known to be the major contributor of chronic hepatitis and liver cirrhosis, often leading to hepatocellular carcinoma. HCV is an enveloped virus with a positive-sense RNA genome, containing a single, large open reading frame composed of 9600 nucleotides, encoding a polyprotein precursor consisting of about 3000 amino acids (1,2).

Analysis of the HCV genome has revealed both highly conserved and highly variable regions. The highly conserved region is found in the 5’ untranslated region (UTR) (3), which has a complex secondary structure that is essential for replication and translation. In contrast, the envelope (E2) glycoprotein region of HCV is the most hypervariable region within the genome (4). Relatively well-conserved regions of the genome (5’UTR, C, E1, and NS5B) have been studied extensively and used as the basis for classification. The high genetic diversity of HCV is due to high replicative activity and lack of proof-reading activity of RNA-dependent RNA polymerase. In terms of nucleotide variation, HCV is classified into 7 major genotypes and 67 subtypes, showing different distributions by geographical region (5). Variations of the viral genes that encode cellular and humoral immune response proteins play a crucial role in escape from the host immune system (6). Therefore, the determination of HCV genotype has a decisive clinical
effect on duration and dosage of the therapy. It has been re-
ported in several studies that genotype 1b, which was found to
be the most prevalent genotype in Turkey (7-10), shows poor
response to IFN therapy (11-14), even in combination with riba-
virin (14). However, up to 80% of patients infected with geno-
types 2 and 3 can have a sustained virological response with
interferon and ribavirin combination therapy with a shorter
duration of therapy (15). New treatments for both treatment-
naive and treatment-experienced patients with HCV genotype
1 are now licensed in several countries. These current treat-
ments add either telaprevir or boceprevir protease inhibitors
to PEG-IFN-α and ribavirin combinations. Therefore, the deter-
mination of HCV genotype is essential prior to therapy, which
may ensure that genotype 1 patients take advantage of current
optimal treatment with triple combination (16).

The aim of this study is to detect whether there have been al-
terations in HCV genotypes in Turkey during the last decade
and to compare the performance of the phylogenetic analysis
of different target HCV sequences (NS5B, E1, and 5’UTR).

MATERIALS AND METHODS

Study population
A total of 500 HCV RNA-positive patients (226 males, 274 fe-
males), admitted to Ankara University Medical School, Depart-
ment of Gastroenterology, Hepatology Laboratory, were in-
cluded in the study. Mean age of the patients was 57.2±11.1.
Quantitative HCV RNA levels of the patients were between
10^2-10^7 copy/mL. Serum samples were stored at -80°C under
RNase-free conditions until analysis.

RNA extraction and RT PCR
RNA was extracted from 200 uL of serum sample by the High
Pure Viral RNA kit (Roche Diagnostics, Germany) according to
the manufacturer’s protocol. The RNA was finally eluted in 50
uL of ribonuclease-free water and stored at -80°C until further
analysis. The RNA pellet was reverse-transcribed to comple-
mentary DNA (cDNA) using proper primers for the NS5B re-
gion, E1 region, and 5’UTR region; dNTP; and 10 units of AMV
reverse-transcriptase in a final volume of 20 uL. This mixture
was incubated in a thermal cycler (Eppendorf, Hamburg, Ger-
many) at 42°C for 1 hour.

Amplification of HCV NS5B, E1, and 5’UTR regions
The polymerase chain reaction (PCR) primers used for ampli-
fication of the NS5B (17), E1 (18), and 5’UTR (3,19) regions are
given in Table 1. For the amplification of NS5B, semi-nested
PCR was used, and for the amplification of the E1 and 5’UTR
regions, nested PCR was used. The PCR conditions were as
follows: initial denaturation at 95°C for 5 min followed by 30
cycles of denaturation at 95°C for 30 seconds (sec), anneal-
ing at 56°C for NS5B, 55°C for both E1 and 5’UTR for 1 minute
(min), and extension at 72°C for 1 min, followed by a final ex-
tension step at 72°C for 10 min. For the second rounds of the
PCR, the same cycle programs were performed using internal
sense and antisense primers of the proper region (Table 1). All
PCR reactions were performed in an Eppendorf Master-Cycler
Personal thermocycler (Eppendorf, Hamburg, Germany). Amp-
licons were run on 2% agarose gels stained with ethidium
bromide, and DNA was purified by using a silica-based meth-

Table 1. Primers used for PCR amplification

<table>
<thead>
<tr>
<th>Regions</th>
<th>Primer Sequences</th>
<th>Ref. No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS5B region</td>
<td>External</td>
<td>NS5B B1 5’TAT GAY ACC CGY TGC TTT GAC 3’</td>
</tr>
<tr>
<td></td>
<td>External</td>
<td>NS5B B2 5’GAG GAG CAA GAT GTT ATC AGC TC 3’</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>NS5B B1 5’TAT GAY ACC CGY TGC TTT GAC 3’</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>NS5B B3 5’GAA TAC CTG GTT ATC ATC GCA GCC 3’</td>
</tr>
<tr>
<td>E1 region</td>
<td>External</td>
<td>DM 110 5’GTR GGN GAC CAR TTC ATC ATC A 3’</td>
</tr>
<tr>
<td></td>
<td>External</td>
<td>DM 111 5’GCA ACA GGG AAY YTD CYY GGT TGC TC 3’</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>DM 108 5’TTC ATC ATC ATR TCC CAN GCC AT 3’</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>DM 109 5’AAY YTD CCC GGT TGC TCT TTY TCT AT 3’</td>
</tr>
<tr>
<td>5’UTR</td>
<td>External</td>
<td>HCV 11 5’ATA CTC GAG GTC GAC CAG CTA CGA GAC CT 3’</td>
</tr>
<tr>
<td></td>
<td>External</td>
<td>HCV 12 5’CAC TCT CGA GCA CCC TAT CAG GCA GT 3’</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>HCV 13 5’CTG TGA GGA ACT ACT GTG TT 3’</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>HCV 14 5’TTC ACG CAG AAA GGG TCT AG 3’</td>
</tr>
</tbody>
</table>

5’UTR: 5’ untranslated region.

Sequencing and phylogenetic analysis
Purified DNA fragments were directly sequenced by an ABI
3100 Genetic Analyzer with the Big Dye Terminator V3.1 Se-
quencing kit (Applied Biosciences, US). Phylogenetic analysis
was performed using the neighbor-joining method in Molecu-
lar Evolutionary Genetics Analysis (MEGA) 5.2 software. The lo-
cus and accession numbers of the reference sequences used

Kabakçı Alagöz et al. Phylogenetic analysis of HCV in Turkey Turk J Gastroenterol 2014; 25: 405-10
are given in Table 2. Phylogenetic comparison was done by distance matrix analysis using the Kimura two-parameter (K2P) distance method followed by ClustalW neighbor-joining. The database www.hcv.lanl.gov was used for the determination of the genotypes.

Recombination analysis
The sequence of potential recombinant strains were aligned and analyzed with references sequences HPCPLYPRE and HPCCGAA for subtype 1a, HPCHUMR for subtype 1b, HPCPOLP for subtype 2a, and HPCK3A for subtype 3a using the Recombination Analysis Tool (RAT) (20).

RESULTS
In this study, HCV genotyping was carried out in 500 patients (274 females, 226 males) by using NS5B, E1, and 5’UTR sequences. In addition to the phylogenetic analysis performed by the MEGA software tool, genotypes were also determined by using the most common database, www.hcv.lanl.gov. The results are summarized in Table 3.

Amplification of the NS5B region was successful in 494 of 500 samples. Phylogenetic analysis showed that 490 patients (99.2%) were infected with genotype 1. Among them, subtype 1b was found to be the most prevalent genotype and was determined in 457 (92.5%) patients. HCV subtype 1a was the second most prevalent genotype, with 33 (6.7%) infected patients, followed by HCV subtype 2a in 2 (0.4%) patients. Genotype 4 was detected in 2 patients, with subtypes 4a in 1 patient and 4d in the other.

According to the analysis of E1 sequences, the results were similar for HCV subtype 1b (93.3%) and 1a (5.6%). These genotypes were followed by subtype 4d (0.4%) and 3a (0.4%). Only 1 patient was found to be infected with subtype 2a (0.2%). Of the 500 samples, 496 could be amplified when primers specific for the E1 region were used.

Results of the phylogenetic analysis of 5’UTR sequences were also in parallel with those of the E1 and NS5B sequences for HCV subtype 1b (87.7%) and 1a (6.6%). Four patients were determined as subtype 2a. Genotype 3 was found in 2 (0.4%) patients. Two samples (0.2%) were determined as genotype 4. In 20 samples, the genotype could not be determined. Moreover, all but 1 of the samples could be amplified when 5’UTR-specific primers were used.

Table 3. Comparison of genotyping results obtained with analyses of NS5B, E1, and 5’UTR regions

<table>
<thead>
<tr>
<th>NS5B Region</th>
<th>E1 Region</th>
<th>5’UTR Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>hcv.lanl.gov</td>
<td>Phylogenetic analysis</td>
<td>hcv.lanl.gov</td>
</tr>
<tr>
<td>1a</td>
<td>34 (6.9%)</td>
<td>27 (5.4%)</td>
</tr>
<tr>
<td>1b</td>
<td>457 (92.5%)</td>
<td>464 (93.5%)</td>
</tr>
<tr>
<td>2a</td>
<td>1 (0.2%)</td>
<td>1 (0.2%)</td>
</tr>
<tr>
<td>3a</td>
<td>2 (0.4%)</td>
<td>2 (0.4%)</td>
</tr>
<tr>
<td>3b</td>
<td>1 (0.2%)</td>
<td>1 (0.2%)</td>
</tr>
<tr>
<td>4a</td>
<td>1 (0.2%)</td>
<td>1 (0.2%)</td>
</tr>
<tr>
<td>4d</td>
<td>1 (0.2%)</td>
<td>2 (0.4%)</td>
</tr>
<tr>
<td>?</td>
<td>1 (0.2%)</td>
<td>4 (0.8%)</td>
</tr>
<tr>
<td>-</td>
<td>6 (1.2%)</td>
<td>4 (0.8%)</td>
</tr>
</tbody>
</table>

?: Can not be determined
(-): No amplification
5’UTR: 5’ untranslated region
In this study, phylogenetic analysis of the three different target regions (NS5B, E1, 5’UTR) showed discordant results for 18 samples when the non-amplified samples were excluded (Table 4). Out of 6 non-amplified samples in the NS5B region, 2 of them were determined as genotype 3, showing a difference in subtypes (3a with E1 sequences and 3b with 5’UTR sequences), and 1 of them was found to be subtype 1b in both the E1 and 5’UTR regions. The other 3 samples could not be amplified with E1-specific primers, either, although their subtype was detected as 2a in the 5’UTR region. Besides these 3, 1 other non-amplified sample in region E1 was found to be subtype 2a both in the NS5B and 5’UTR regions.

The genotypes of 20 (4%) samples could not be identified by phylogenetic analysis of 5’UTR sequences, although the genotypes of all samples were clearly determined when the phylogenetic analysis was performed using sequences of NS5B and E1.

The presence of putative recombinant strains was also analyzed using the RAT application, and no recombination was found.

Subtypes were also determined using the www.hcv.lanl.gov database (Table 2). Phylogenetic analysis of the 5’UTR sequences determined genotype 1 in 94.3% samples, although this number was approximately 98.5% when the database was used for analysis. Comparison of the results of the phylogenetic analysis and hcv.lanl.gov database showed very similar distributions for both the E1 and NS5B regions (Table 3). None of the samples was determined as genotype 5 or 6 by using the phylogenetic analysis or HCV database, www.hcv.lanl.gov.

DISCUSSION

The results of the current study indicate that: (i) the prevalence of different HCV genotypes in Turkey has not changed during the last decade; (ii) the phylogenetic analyses of NS5B and E1 sequences of HCV showed better consistency; and (iii) the hcv.lanl.gov database performed better with the sequences of both the E1 and NS5B regions.

Previous studies reported from different regions of Turkey showed that genotype 1b (87.2%-97.4%) is the most prevalent subtype in Turkey, followed by genotype 1a (2.6%-11%), genotype 2 (0.9%-3%), genotype 3 (0.6%-2.7%), and genotype 4 (0.6%-1%) (7-10). Similarly, the main consequence of this study is that the predominant genotype is still 1b in Turkey, followed by subtype 1a. The percentages of subtype 1b were found to be 92.5%, 93.3%, and 87.7% when the sequences of NS5B, E1, and 5’UTR were used in the current phylogenetic analysis, respectively. When the results of this phylogenetic analysis were compared with the previous study, which was published in 2004 using a restriction fragment length pattern method, the prevalence of subtype 1a showed a decline from 11% to 5.6%-6.7%, whereas subtype 1b increased from 84% to 87.7%-93.5% (7) (Table 3). In addition, when only the results of the 5’UTR sequences were compared, the percentage of genotype 2 was found to be lower (0.8% vs. 3%) in the current study. The percentages of genotypes 3 and 4 were found to be similar in both studies, with a slight decrease for both genotypes in the present study: genotype 3 (0.4% vs. 1%) and genotype 4 (0.4% vs. 1%).

Many countries, the genotype distribution mostly shows an increase in the percentages of genotypes 3a, 1a, and 4, largely due to the consequences of migration events (21,22) and intravenous drug abuse (23). However, our results showed a relative decrease in the percentages of genotype 1a, genotype 3, and genotype 4 since 2004, although the severity of urban-to-urban migration has continued to increase in Turkey (24). The difference in patient numbers and genotyping methods between studies may also contribute to this diversity.

The PCR reaction was successful with primers targeting the highly conserved 5’UTR region in 499 samples, while no amplification was observed in only 1 sample. This sample was determined as 2a and 1b with the sequences of the NS5B and E1 regions, respectively. When conserved primers of the NS5B and E1 regions were used (17,18), amplification was less successful, with failures in amplification in 6 and 4 samples, respectively. The major factor of these missing samples might be the variability in the primer-binding sequences of the target genome.

In distinct geographical regions, the highly conserved region of the 5’UTR of the HCV genome has been used for effective diagnostic tests and for determination of the prevalence of viral subtypes (25). However, several studies showed that subtyping analysis based on 5’UTR may be limited (18,26,27) due to insufficient sequence variation in these highly conserved regions.

A previous study has indicated that some variants of genotype 6 found in southeast Asia have the same 5’UTR sequences as those of genotypes 1a or 1b (18), which may distort the results of the analysis. In the current study, genotype 5 or 6 was not found in any of the samples, although sequences for the 5’UTR, NS5B, and E1 regions were used, excluding the possibility of inconsistency in genotyping results due to the use of different target sequences of the HCV genome in the analysis. The absence of genotype 5 or 6 is compatible with the results of previous studies from Turkey (7-9), supporting that no considerable change in the distribution of HCV genotype has occurred in Turkey in the last decade.

Hepatitis C virus genotypes of 20 patients could not be determined using phylogenetic analysis of 5’UTR sequences. Analysis of NS5B and E1 sequences of the samples that remained undetermined by analysis of the 5’UTR sequences revealed that they belonged to subtype 1b. This result supports the limited success of subtyping analysis based on 5’UTR. In this respect, when only the results of NS5B and E1 are taken into account, which are the preferred targets for HCV genotyping (5), the current study shows contradictory results for 8 of 500 patients, excluding non-amplified samples (Table 4). The possibility of the
presence of any putative recombinant strains, which may not be common among natural populations of HCV, was eliminated using the RAT application in the current study. One remarkable difference between these regions was seen in the subtype distribution of genotype 1. The highest percentage of subtype 1b was found in the phylogenetic results of the E1 region. This result may be caused by the relatively less conserved feature of the region, which may easily differentiate the genomic similarity between subtypes 1a and 1b. In a study by Qiu et al. (28), the sequences of 1a and 1b shared over 99% similarity in the 5'UTR, 95% in the core, 76% in E1, and 83% in NS5B. In this high variable region, using conserved primers, only 4 of the samples could not be amplified in the current study. This low number of non-amplified samples may indicate the utility of the E1 region for HCV genotyping.

In conclusion, HCV patients from different regions of Turkey participated in the current study, and the results showed no change in genotyping profiles of Turkey during the last decade. Subtype 1b was still the most prevalent, showing no significant decline, followed by 1a. Phylogenetic analysis of HCV indicated high performance with the sequences of the NS5B and E1 regions. Genotype distributions were highly consistent when compared with the hcv.lanl.gov database, particularly using the E1 and NS5B sequences.

Ethics Committee Approval: N/A.

Informed Consent: N/A.

Peer-review: Externally peer-reviewed.


Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES
8. Ozbek TO, Mese S, Atmaca S. Hepatitis C virus genotypes are changing in the southeast of Turkey. Biotechnol Biotech Eq 2009; 23: 1521-3. [CrossRef]