Phylogenetic and Genotyping of Hepatitis C Virus in Egypt


Department of Biomedical Technology, National Research Center, Cairo Egypt.

Abstract: Objective: Genotyping is considered an important tool for epidemiological and clinical studies. The biological differences between genotypes make genotyping important for decision-making regarding disease management and therapeutic intervention. Methods: This study has established a simple and accurate in house system for HCV core gene typing. The data were confirmed by sequencing and phylogenetic analysis. The method used, originally developed by another laboratory, is advantageous over the other methods by efficiently typing most of the subtypes in two multiplex PCR reactions. Besides, it could identify mixed infections. Results and conclusions: The method also showed a sensitivity of (100%) in detecting type 4, the most prevalent genotype in Egypt, as genotype-4 was detected in this study in 283 cases out of 300 (94.3%). Sequencing data from the core gene have been utilized for phylogenetic analyses and demonstrated genetic relevance to HCV types 1 and 2 , prevalent in (Europe and USA) but not to genotypes of other populations.

Key words: HCV, Genotyping, PCR, Multiplex-PCR, Sequencing analysis, Phylogenetic analysis

INTRODUCTION

Hepatitis C virus (HCV) is a major health problem affecting 170 million people worldwide. The seroprevalence rate is about 1% in western countries and North America, 3-4% in some Mediterranean and Asian countries and up to 10-20% in parts of central Africa and Egypt (WHO, 2000; Abdel-Hamid, M., M. El-Daly, 2007). The infection with the HCV is the leading cause of chronic hepatitis worldwide, progressing to liver cirrhosis in approximately 20% of patients after 10 years and to hepatocellular carcinoma (HCC) in a subset of them with a yearly incidence of 3% (Zein, N., 2000; Moradpour, D., A. Cerny 2001). HCV is a positive strand RNA virus of approximately 9.6 Kb in length. Its genome is composed of a 5’ non-coding region (5’NCR), a long open reading frame (ORF) encoding a polyprotein precursor of about 3,000 amino acids and 3’NCR. The 5’NCR functions as internal ribosomal entry site (IRES) essential for cap-independent translation of the viral RNA (Bartenschlager, R., M. Frese 2004). The core region has numerous functional activities. These include its role in encapsidation of viral RNA, a regulatory role on cellular and viral promoters, interactions with a number of cellular proteins, a modulator role in cell death under certain conditions, involvement in cell growth promotion and immortalization, induction of HCC in transgenic mice and a possible immuno-regulatory role (Ray, R.B. and R. Ray, 2001). HCV displays significant genetic heterogeneity as a result of accumulation of mutations during replication. The genetic heterogeneity is not uniform across the genome, the most highly conserved regions of the genome are parts of the 5’NCR and the terminal 3’NCR followed by the core region. In contrast, the most heterogeneous portions of the genome are the genes encoding the envelope proteins (E1 and E2). Accumulation of nucleotide substitution in the HCV genome results in diversification and evolution into different genotypes, subtypes and quasispecies. No fewer than 6 genotypes and more than 50 subtypes have been detected (Farci, P. and R. Purcell, 2000; Kato, N., 2001). Each of the six main genotypes of HCV is equally divergent from one another and varies by as much as 35% of nucleic acid content, while subtypes within a typical genotype differing from each other by 20-23%. Within the infected host the viral pool comprises several different but closely related sequences called quasispecies, these may show up to 10% diversity (Rapicetta, M., C. Argentini, 1998; Farci, P., A. Shimoda 2000). There is great evidence that the quasispecies nature of HCV provides a large reservoir of biologically different viral variants that may have important clinical implications for viral persistence by immune escape.
There is increasing evidence that patients infected with different HCV genotypes have different clinical profiles, severity of liver disease and response to alpha-interferon therapy. Hence, a convenient and reliable HCV genotyping system is essential for large-scale epidemiological and clinical studies. In this study, the major goal was to establish a simple, accurate and a reliable genotyping system of HCV for use in Egyptian patients. The method was utilized for comparing genotypes in sera versus peripheral blood mononuclear cells (PBMC) within the same patients. Besides core gene sequences were determined in 10 Egyptian quasispecies and were used for the construction of a phylogenetic tree correlating the genetic divergence from the other globally known genotypes.

**Subjects and Methods:**

This study was conducted on 300 Egyptian patients infected with Hepatitis C virus (HCV). Patients were subdivided into two groups. Serum samples were obtained from 200 patients chronically infected with HCV (151 males (75.5%) and 49 females (24.5%)). PBMC samples obtained from 100 patients. All patients had chronic active hepatitis as confirmed by ultrasonographic examinations, were seropositive for anti-HCV antibodies and were also positive by PCR testing for HCV RNA.

**The Patients Were Subjected to the Following:**

1. **Full History Taking, Clinical and Laboratory Examinations:**
2. **HCV Antibody Detection Using ELISA Technique:**
   (Third generation enzyme immuno-assay, Dia Sorion-Italy)
3. **RNA Extraction:**

   Using the method of Chomczynski and Sacchi (Chomczynski, P., Sacchi N., 1987). In this method total RNA was isolated by a single-step extraction using an acid guanidinium thiocyanate-phenol-chloroform mixture.

4. **HCV-RNA Amplification by Reverse Transcription-PCR (RT-PCR):**

   This was performed using primer sequences derived from the highly conserved non-coding region of HCV genome as reported earlier (EL-Awady, M., S. Ismail, 1999; EL-Awady, M., M. Abdel Rahman, 2003). The RT-PCR was performed in a 50 ul volume single-step reaction using the Ready-To-Go RT-PCR Beads (Pharmacia Amersham Biotech, USA), 10uM of each of the down-stream primer 1CH and the up-stream primer 2CH and the reverse primer P2. The Thermal cycling protocol was 30 minutes at 42°C for reverse transcription followed by 5 minutes at 95°C then 30 cycles of 1 minute at 94°C, 1 minute at 55°C and 1 minute at 72°C. The nested PCR amplification was performed in 50 ul reaction mixture, containing: 0.2 mM dNTPs, 10 uM of reverse nested primer D2 and the forward nested primer F2, 2 units of Taq DNA polymerase (Promega, USA). Then 10 ul of the nested PCR product were electrophoresed on a 2% agarose gel, stained with ethidium bromide and evaluated under U.V light.

   Primer sequences were as follows:

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1CH</td>
<td>5’ GGT GCA CGG TCT ACG AGA CCT C 3’</td>
</tr>
<tr>
<td>2CH</td>
<td>5’ AAC TAC TGT CTT CAC GCA GAA 3’</td>
</tr>
<tr>
<td>P2</td>
<td>5’ TGC TCA TGG TGC ACG GTC TA 3’</td>
</tr>
<tr>
<td>D2</td>
<td>5’ ACT CGG CTA GCA GTC TCG CG 3’</td>
</tr>
<tr>
<td>F2</td>
<td>5’ GTG CAG CCT CCA GGA CCC 3’</td>
</tr>
</tbody>
</table>

5. **HCV Genotyping:**

   (5:A) Nested PCR Amplification of HCV Core Gene Using Genotype Specific Primers:

   Genotyping was performed by utilization of the method reported previously (Ohno, T., M. Mizokami, 1997). Denatured RNA samples were reverse transcribed into cDNA, using: 20 units of reverse transcriptase (Promega, USA), 2.5 uM of AC primer (anti-sense primer of the core region) and 0.2 mM dNTPs. Two rounds of PCR amplification were carried out on cDNA as follows: The first round of amplification utilizes 2.5 uM of primers SC1 and AC1 (which are the sense and the anti-sense primers for the core region respectively), 0.2 mM dNTPs and 4 units of Taq DNA polymerase (Promega, USA). Thermal cycling of first PCR was as follows: 20 cycles of (94°C for 1 minute, 45°C for 1 minute and 72°C for 1 minute) followed by 20 cycles of (94°C for 1 minute, 60°C for 1 minute and 72°C for 1 minute). Two second-round PCR amplifications were performed for each sample.
One with primer mixture (1) containing the S7, S2a, G1b, G2a, G2b and G3b with 0.2 mM dNTPs and 4 units of Taq DNA polymerase. The other second PCR with primer mixture (2) containing the S7, G1a, G3a, G4, G5a and G6a with 0.2 mM dNTPs and 4 units of Taq DNA polymerase (Promega, USA). Thermal cycling of second PCR was as follows: 30 cycles of (94°C for 1 minute, 62°C for 45 seconds and 72°C for 1 minute). The detection of genotype-specific products in multiplex (WHO, 2000) and multiplex (Abdel-Hamid, M., M. El-Daly, 2007) was designed so that the differences in the sizes of PCR products could be evaluated on gels easily. Primer sequences were reported in (Ohno, T., M. Mizokami, 1997).

(5:B) Reverse Hybridization Analysis of 5’ Non-coding Region by Using the INNO-LIPA HCV-II Assay as Confirmatory Test (Innogenetics-Belgium):

This assay allows an easy and fast determination of 6 HCV genotypes and their subtypes. The assay is based on variations found in the 5’ untranslated region (5’NCR) of the different HCV genotypes (Stuyver, L., R. Rossau, 1993). This assay includes nested-PCR amplification of 5’NCR with specific primers followed by reverse hybridization assay with type-specific probes.

(5c) Sequencing Analysis:

The reference standard and most definitive method for HCV genotyping is sequencing of a specific PCR-amplified portion of the HCV genome obtained from the patient, followed by the phylogenetic analysis. In this study, we planned to drive a genetic correlation between quasispecies of genotype 4 and the comparable sequences derived from the same regions of other genotypes for this region. The 99 bp amplified product from genotype 4 in Egyptian patients was copared with the same regions within other genotypes/subtypes. The purified PCR product was cloned into the pGEM-T vector (Promega, Madison WI). At least 3 clones from each patient Thus confirmatory test for HCV genotype was done by cloning the purified PCR product into the pGEM-T vector (Promega, USA) and the samples were utilized for sequence analysis using the Big Dye Terminator Cycle Sequencing and in the Automated SequencerABI Prism 310 Genetic Analyzer(Applied biosystems,CA).

(6) Phylogenetic Analysis:

Phylogenetic trees were generated by use the PHYLIP suit of programs, version (3.572) (Felsenstein, J., 1989). Genetic distance matrices were calculated with the DNADIST program. The distance matrix was then used to generate a tree by use of the neighbor-joining algorithm as implemented in NEIGHBOR. In a similar manner, bootstrap confidence values were calculated after randomly permuting the sequence alignment 100 times with the SEQBOOT program. Permuted trees were generated by use of the NEIGHBOR program with random addition and a consensus topology was derived by use of CONSENSE program.

RESULTS AND DISCUSSION

Results:

This study was conducted on 300 Egyptian patients chronically infected with Hepatitis C Virus (HCV). Included serum samples obtained from 200 patients and PBMC samples from 100 patients chronically infected with HCV.

I-Clinical Data:

1-Age of Patients: The ages of the 300 patients ranged from 8-65 years old with a mean ±S.D (43.5 ±5.6) years old. 2-Sex Distribution: The samples studied included 236 males (78.7%) and 64 females (21.3%) out of the total 300 patients studied. 3-Liver Enzymes: Elevations in liver enzymes (ALT & AST) were observed in 53% of the 300 cases studied. 4-HCV Transmission: Out of the 300 patients, 54 patients (18%) had histories of blood transfusions, 108 patients (36%) had histories of different surgeries and dental procedures. 24 patients (8%) had no identifiable risk factors for acquiring HCV infection, 114 patients (38%) with history of Schistosomiasis infection and treatment with non-disposable syringes, and 9 patients(3%)had been co-infected with HBV (Figure 1).

II: 1 Assessment of HCV Genotyping Methods:

Using 30 infected samples, the method described by Ohno et al (1997) was compared with a previously reported method (Okamoto et al, 1993) and a commercial genotyping method (Inno-Lipa, Innogenetics, Belgium). Ohno’s method matches the genotypes obtained by Inno-Lipa’s method in 29 of 30 cases (96.7%). On the other hand, both methods have no match with Okamoto’s method. The latter was originally designed to type HCV genotypes 1a, 1b, 2a, 2b and 3a but not type 4 which explains its failure to type the Egyptian
strains.

Fig. 1: Risk Factors for HCV transmissions in Egyptian patients.

II: 2 HCV genotyping in Egyptian patients:

This study was conducted on 300 Egyptian patients chronically infected with HCV. The results of this study demonstrated that among the 300 Egyptian patients, HCV genotype 4 (both single and mixed infection) was detectable in 283 patients (94.3%) with a 95% confidence interval of lower value (91.7%) and higher value (96.9%). This reveals that the most prevalent genotype among these patients is type 4. Either types 1a or 2a were detected in only 1% among the studied samples of HCV patients. It is interesting that neither types 3 and 5 were detected in the studied samples. Figure (2) shows the percent of each HCV genotype in the 300 studied samples. Figure (3) shows the amplicons of two patients of different genotypes.

Fig. 2: Percent of HCV genotypes in 300 chronic hepatitis patients from Egypt.

Fig. 3: HCV genotyping in Egyptian Patients by Ohno’s Method:

Amplified PCR products were visualized on 2% agarose gel stained with ethidium bromide. M represents PhiX 174 DNA-HaeIII digest marker. Lanes (1) and (2) represent the amplified PCR products of patient (X11) in multiplex (WHO, 2000) and multiplex (Abdel-Hamid, M., M. El-Daly, 2007) respectively. Lane (1) shows 2 bands in multiplex (WHO, 2000), a band of 337 bp that is indicative of type 2b and the other band of 190
bp that sometimes appear with type 4 and lane (2) shows one band of 99 bp in multiplex (Abdel-Hamid, M., M. El-Daly, 2007) that is indicative of type 4. Lanes (3) and (4) represent the amplified PCR products of patient (X12): in multiplex (WHO, 2000) and multiplex (Abdel-Hamid, M., M. El-Daly, 2007) respectively. Lane (3) shows no bands in multiplex one, while lane (4) shows a band of 99 bp in multiplex two that is indicative of type 4. Conclusion: Patient (X11) was typed as having mixed infection (4+2b) in serum and Patient (X12) was typed as having type 4 only in PBMC.

II Sequencing & phylogenetic Analysis:
To further analyze different quasispecies within a single type 4 infection in our population, the amplified 99 bp fragment of the type 4 HCV RNA was sequenced in 5 patients and aligned against type 4a in the gene bank. Furthermore, sequencing analyses of the local isolates was performed on a larger portion (355bp) of the core gene. phylogenetic tree (unrooted) was generated using the sequence data from 10 Egyptian clones and the reference sequences of HCV prototype (HCp) and complete HCV genome (HCg) and different genotypes of HCV (1a, 1b, 2a, 2b, 3a, 3b, 4a, 4c, 4f, 5a and 6a) from the Gene bank. As shown in figure (4), local strains MO1-10 have closer sequences to G4a, G4c, G4e, G1b and HCV genome (HCg) than G2a, G2b, G5a, G1a and HCV prototype (HCp), whereas Asian strains G3a, G3b and G6a have the most distant sequence from local strains.

Fig. 4: Phylogenetic tree (unrooted) of the 10 Egyptian clones using reference sequences of HCV prototype (HCp) and complete HCV genome (HCg) and different genotypes of HCV (1a, 1b, 2a, 2b, 3a, 3b, 4a, 4c, 4f, 5a and 6a) from Gene bank. In this figure it is shown that the core gene sequences of our strains had mainly close genetic relation to genotype 4 and their subtypes followed by genotypes 1 and 2.

Discussion:
Although patients’ selection was fairly random in the present study, HCV chronic hepatitis had higher male preponderance over females (3.7:1; P< 0.001), thus confirming earlier reports (Kamal, M., Y. Ghaffer, 1992; Archer, G., M. Buring, 1992; Tanaka, E., K. Kiyosawa,1992; Lau, J., G. Davis, 1993; Waked, I., S. Saleh1995; Frank, C., M. Mohamed, 2000). Such observation is partly related to social risk factors for HCV transmission as drugs, schistosomiasis and occupational exposure. However, the role of androgens in this sexual dimorphism must not be excluded. The mean age of the studied population was 43.5 ± 5.6 years old, which agrees with previous studies (El-Sayed, H., S. Abaza, 1997; Ministry of Health and Population, 1999; Abdel-Aziz, F., M. Habib, 2000). In Egypt, mass campaigns to treat schistosomiasis with injection therapy during 1960 to 1987 might have been responsible for the sustained transmission of HCV during this time and might be related to the mean age of the studied group. In fact 38% of the studied patients had history of schistosomiasis infection, supporting earlier studies (Darwish, M., R. Faris, 1996; Tibbs, C.J., 1997; Helal, T., M. Danial, 1998) and suggesting that history of parental antischistosomal therapy (PAT) played a major role in the spread of HCV throughout Egypt (EL- Awady, M.K., S.S. Youssef,2006). The role of blood transfusion as a risk factor in our population has been documented in 18% of patients supporting earlier reports (Mohamed, M., M. Rakha, 1996).
Higher seroprevalence rates of anti HCV, in specific groups of patients receiving repeated blood transfusion; 70.4% in hemodialysis patients and 75.6% in thalassaemics (El-Gohary, A., A.M. Hosny, 1999) strengthen the belief that blood transfusion is a major route in HCV epidemics in Egypt. The possibility of other routes of infection including previous surgeries and dental procedures were considerably documented (36%). This also suggests that the explorations of contact-associated transmissions, such as shared razors and toothbrushes as well as exposure to skin lesion exude are significant risk factors (El-Gohary, A., A.M. Hosny, 1999; Helal, T., M. Danial, 1998). This calls for investigating the role of HCV in interfamilial transmission. The remaining 5% of the studied cases had no identifiable factors for HCV transmission. This is apparently lower than unidentified routes in other populations (Ackerman, Z., O. Paltil, 1998; Drucker, J., 1997). Clearly, there is much to be learned regarding the mechanisms of HCV transmission.

Thus as the viral genotype has been the subject of a large debate on its role in disease progression and response to therapy, it seems prudent that it should be ascertained before it can be related to any clinical findings (Le Guillou- Guillemette, H., 2007). Our main goal in the current study was to establish a reliable and economic HCV genotyping method for local patients and also for HCV patients in other populations. The comparison between different genotyping procedures revealed that the method described by Ohno’s et al (1997) is the simplest, most accurate and least expensive. Most importantly Ohno’s method was appropriate for typing the prevalent viral strains in our patient population. The present results demonstrated that HCV genotype 4 (single and mixed infections) was detected in 283 out of 300 patients (94.3%), a finding that supports earlier studies (Dusheiko, G., H. Schmilovitz; 1994; Abdel-Aziz, I.A., 1995; Chamberlain, R., N. Adams, 1997; Abdel-Karim, A., N. Zein, 1998; Blatt, L., M. Mutchnick, 2000; Ray, S., R. Arthur, 2000; Morice, Y., D. Roulot, 2001). The present data show that the contribution of type 4 into the Egyptian HCV pool is not exclusive but other genetically related genotypes do exist, thus confirming previous studies (Genovese, D., S. Dettori, 2005; Cabot, B., M. Martell, 2001). HCV like other viruses commonly exist in the host as a heterogeneous population of different but closely related genomes (Le Guillou- Guillemette, H., 2007; Simmonds, P., 2000). Sixteen patients out of 300 (5.33%) in this study had mixed-genotype infection with other types. Genetic diversification within single genotype infection was also observed, this finding increases the complexity of the mechanism of HCV genome. Phylogenetic analyses of genotype 4 HCV isolates within the core gene sequences showed that our strains had close genetic relation to the western genotypes 1 and 2 (mainly 1b and 2b). On the other hand, these core sequences are rather genetically distant from the Asian quasispecies G3a, G3b and G6a. Taking together the results of the phylogenetic tree and the high percentage of type 1 and 2 in the mixed infection group (13) out of 16 (80%), these results suggest that the evolution of these non 4 genotypes occurred in the same patient via multiple mutational events of type 4 into distinct but closely related genotypes 1 and 2. The mechanisms of occurrence of mixed HCV infection are not yet clear. However, it seems that both repeated intraindividual mutational events and multiple transfusions of mixed viral population co-exist in the local HCV strains described in this study as well as in other populations (Cannon, N.A., M.J. Donlin, 2008). Accurate detection of mixed-quasispecies infection provides a tool that can be used to address whether multiple episodes of reinfection with different quasispecies would increase the likelihood of acquiring highly pathogenic variants, resulting in more severe hepatitis when compared with that from a single-quasispecies infection (Donlin, M.J., N.A. Cannon, 2007).

Core gene sequence from local clones exhibit homology with reported type 4 HCV strains including 4a, 4c, 4e and 4f. These clones are, therefore, considered as different quasispecies of type 4 as they have displayed diversity up to 10% from different HCV subtypes of genotype 4. Two of these clones (MO3 and MO8) have shown nucleotide frame shift, which ended in a termination stop codon, indicating that these two quasispecies would not be involved in HCV replication. This agrees with the evidence that the quasispecies nature of HCV provides a large reservoir of biologically different viral variants that may have important clinical implications for viral persistence by immune escape mechanism (Morice, Y., D. Roulot 2001; Mao, Q., S. Ray, 2001). In conclusion, HCV genotyping and analysis of the contributing pool of quasispecies and cell tropism may be indicated for patient evaluation and development of effective therapeutic modalities. In any case, assessment of efficacy of HCV vaccines will have to incorporate analysis of genotype differences and extensive research on quasispecies of HCV.

REFERENCES

Ohno, T., M. Mizokami, R. Wu, M.G. Saleh, K. Ohba, et al, 1997. New hepatitis C virus (HCV) genotyping system that allows for identification of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a and 6a. Journal of Clinical Microbiology, 35: 201-207.
Ministry of Health and Population (MOHP), Egypt, Central Department of Preventive Affairs, 1999. The national workshop for the preparation of practical guidelines for prevention and control of viral hepatitis in


