# Phylogenetic subtyping of hepatitis C virus 5' UTR isolated in Egypt and the effect of 2 transitions in subdomain IIId on the apical loop structure

*Amal Mahmoud and Medhat H. Hashem

Genetic Engineering & Biotechnology Research Institute, Menoufia University, Sadat City Branch, Menoufia, Egypt
*amalmh3@yahoo.com

**Abstract:** HCV 5′-UTR of 3 non-responding isolates (Sohag 1, 2 and 3), collected from Sohag-Egypt was compared with the 6 genotypes collected from GenBank and HCV database sequences included a responder isolate from Egypt. Multiple alignment comparison showed that TTGGGT sequence located in the IIId subdomain loop was conserved in all genotypes. Phylogenetic tree revealed that isolate Sohag 1 was clustered with the responder isolate within subtype 4a. Sohag 2 was clustered with isolates of subtypes g and o, however, isolate 3 was grouped with isolates from subtype 4a and q. Two transitions $T_{175}\rightarrow C$ and $C_{183}\rightarrow T$ were detected in the stem of IIId subdomain and were highly represented in genotype 2 followed by genotype 6: 1.0, 95.5, 2.0, 14.5, 0.0 and 33.5% for genotypes 1, 2, 3, 4, 5 and 6, respectively. The predicted secondary structure of HCV 5′-UTR showed that the formation of UUGGGU loop was affected by the 2 transitions $T_{175}\rightarrow C$ and $C_{183}\rightarrow T$. In the absence of these 2 transitions, the apical loop was replaced by double helix structure in most predicted folding. Taken together, the % of the 2 transitions in different genotypes and the modification in the apical loop structure could affect the response to therapeutic treatment.

[Amal Mahmoud and Medhat H. Hashem. **Phylogenetic subtyping of hepatitis C virus 5' UTR isolated in Egypt and the effect of 2 transitions in subdomain IIId on the apical loop structure**. Life Science Journal. 2012;9(1):903-909] (ISSN:1097-8135). http://www.lifesciencesite.com.133

**Keywords:** Hepatitis C virus; 5` UTR; IIId subdomain; Secondary structure

## 1. Introduction

Egypt has the highest prevalence of Hepatitis C virus (HCV) worldwide, where it infects about 15% of the general population (Egyptian Ministry of Health, 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, 2000).

HCV is an RNA virus and a member of the Hepacivirus genus classified into the Flaviviridae family. HCV presents high mutation rates and because of that it has been evolved to different genotypes based on nucleotide sequence heterogeneity and classified in six major genotypes and more than 80 subtypes (Mizokami *et al*., 1996; Robertson *et al*., 1998 and Simmonds *et al.,* 2005).

The 5′-UTR of HCV is an essential component of the internal ribosome entry site (IRES) that regulates Cap-independent translation of HCV (Wang *et al.*, 1994). Three domains: I, II, and III, are inside this region. Mutations in IIId domain disrupt IRES-mediated translation initiation and also affect the RNA structure, demonstrating the importance of correct RNA folding to IRES function (Jubin *et al*., 2000). Hazari *et al*., 2005 suggested that the antiviral action of IFN-2b blocks IRES-mediated translation and this effect is the same among HCVs of other genotypes.

In this study, a comparison between our non-responder isolates with a responder one and a classification at the subtype level in the 5`-UTR of HCV was done. We analyzed the effect of 2 transitions exist in IIId subdomain of the IRES on its secondary structure.

## 2. Material and Methods

### 2.1. Samples

In a previous study, (Hemeida *et al*., 2011), We collected 92 HCV positive isolates from Center of Cardiac and Digestive System, Sohag, Egypt. The confirmed patients were received 12 vials of PEG-IFNα-2a for 12 weeks (180IU/ml weekly) plus ribavirin (1000 mg for ≤ 75 Kg or 1200 mg for > 75Kg- Roche) and follow up by RT-PCR. The results showed that 67 patients (72.8%) responded to the treatment, while 25 patients (27.2%) were non responders. We selected 3 non responder isolates (Sohag 1, 2 and 3), then were sequenced and analyzed (Hemeida *et al*., 2011).

### 2.2. Sequence analysis

DNA sequencing of three random serum samples from HCV non-responder patients (sohag 1, 2 and 3) was carried out as previously described (Hemeida *et al*., 2011). BLAST program (NCBI) was

used to identify the similarity between isolates. Secondary structure of 5′-UTR of HCV was deduced using mFOLD software. Version 3.2 program(Zucker,1989) (http://mfold.bioinfo.rpi.edu/). Our sequences were submitted to GenBank database. [GenBank: JQ228803–JQ228805 for isolates Sohag 2, 3 and 1, respectively]. Sequences of HCV different genotypes were retrieved from GenBank and HCV database. An isolate responded to the pegylated IFN alpha-2a plus ribavirin treatment from Egypt was used in this study for comparison (Zekri *et al*. 2007). Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 4 (Tamura *et al*., 2007). Standard error estimates are shown and were obtained by a bootstrap procedure (500 replicates).

.

## 3. Results

### 3.1. Multiple sequence alignment of HCV 5′-UTR

Our non-responder isolates (Sohag isolates) showed identity of 94-98% with those of GenBank databases. The alignment of the responder isolate against Sohag isolates (192nt) resulted in 8 variable sites (highlighted positions is in Fig. 1). The responder isolate was characterized by one insertion ($T_9$) and 2 deletions ($C_{36}$ and $A_{117}$).


Figure 1. Nucleotide sequence alignment comparison of the HCV 5′-UTR region between responder and non-responders isolates (Sohag-1, 2 and 3).

Multiple alignment comparison with the 6 known genotypes retrieved from GenBank database is shown in Fig. 2. Hyper variable region was located from nt 85 to 161 in the 5'UTR of different HCV isolates. Located in the IRES IIId subdomain loop, TTGGGT sequence at position 177-182 nt, this sequence was conserved in all genotypes (highlighted and boxed in Fig. 2). Two transitions $T_{175}$→C and $C_{183}$→T exist in the stem of IIId subdomain were highly represented in genotype 2 and 6. Using HCV database, we retrieved 200 sequences for each genotype to analyze the percentage of the $T_{175}$→C and $C_{183}$→T transitions in each genotype and it was as follow: 1.0, 95.5, 2.0, 14.5, 0.0 and 33.5% in genotypes 1, 2, 3, 4, 5 and 6, respectively. Also, the 2 transitions were detected in isolate Sohag 3 but not in Sohag 1, 2 or the responder isolate. Multiple alignment comparison against hepatocellular carcinoma (HCC), responder and non-responder isolates, collected from HCV database was done (Fig. 3). We noticed that the occurrence of the 2 transitions was not related to HCC or response to therapeutic treatment (Fig. 3).

### 3.2. Phylogenetic subtyping of HCV 5′-UTR of Sohag isolates

According to the phylogenetic analysis, genotypes 1, 2, 3, 4, 5 and 6 were separated in 6 clusters (Fig. 4a). The Egyptian isolates (responder and non-responders) were clustered into genotype 4. Phylogentic relationship between our isolates and genotype 4 is shown in Fig. 4b. Sohag 1 and the responder isolates were clustered with subtype 4a. However, isolate Sohag 2 was clustered with isolates FJ462432 (4g), AB548316.1 (4o) and FJ462440 (4o). Isolate Sohag 3 was clustered with isolates FJ462434 (4q) and AB550017.1 (4a).

### 3.3. Secondary structure of HCV 5′-UTR

To investigate the effect of the 2 transitions $T_{175}$→C and $C_{183}$→T, located within IIId subdomain stem, the secondary structure was analyzed using mFOLD where the window parameter controls how many foldings will be automatically computed and how different they will be from one another. The secondary structure of the 5′-UTR III domain resulted in 4 stem loop subdomains IIIa-d (Fig. 5A). The absence of the 2 transitions $T_{175}$→C and $C_{183}$→T in IIId subdomain stem, affected the formation of the UUGGGU apical loop, it was replaced by double helix structure in most predicted folding (Fig. 5B and C, respectively).

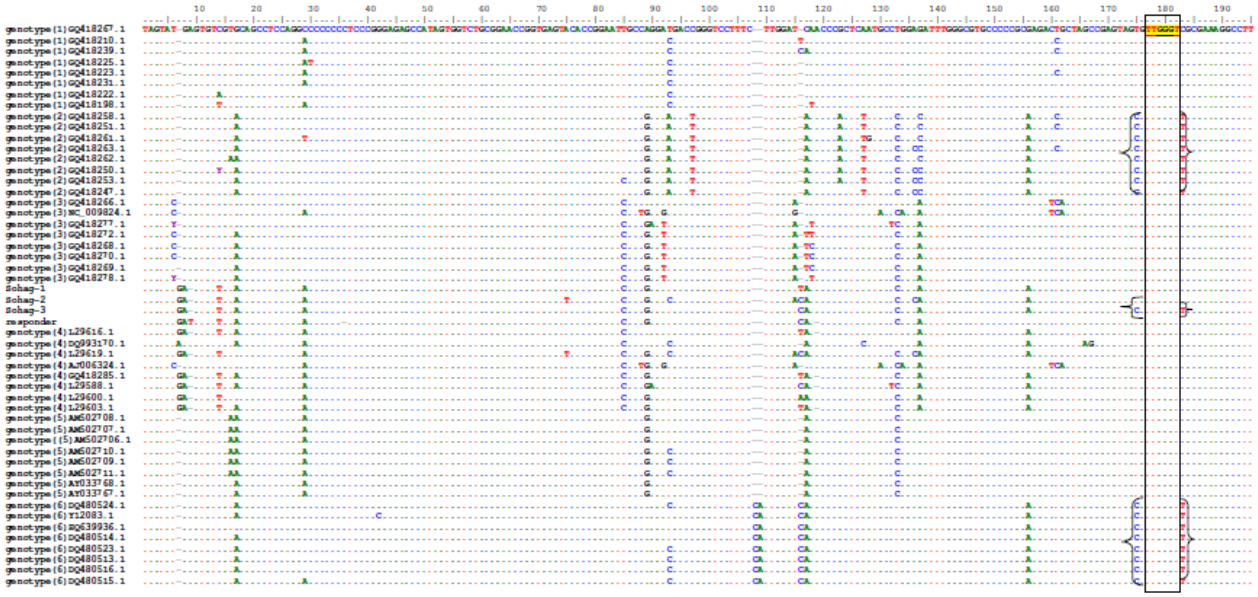

Figure 2. Multiple alignments of the nucleotide sequence of HCV 5′-UTR genotypes. The TTGGGT conserved sequences located in the IRES IIId (177-182nt) are boxed; $T_{175}$→C and $C_{183}$→T substitutions could be related to IIId loop formation are parenthesized.


Figure 3. Multiple alignments of the nucleotide sequence of HCV 5′-UTR using isolate Sohag 3 against hepatocellular carcinoma (HCC), responder (RS) and non-responder (NR) isolates. Arrows mention to the 2 transitions $T_{175}$→C and $C_{183}$→T.

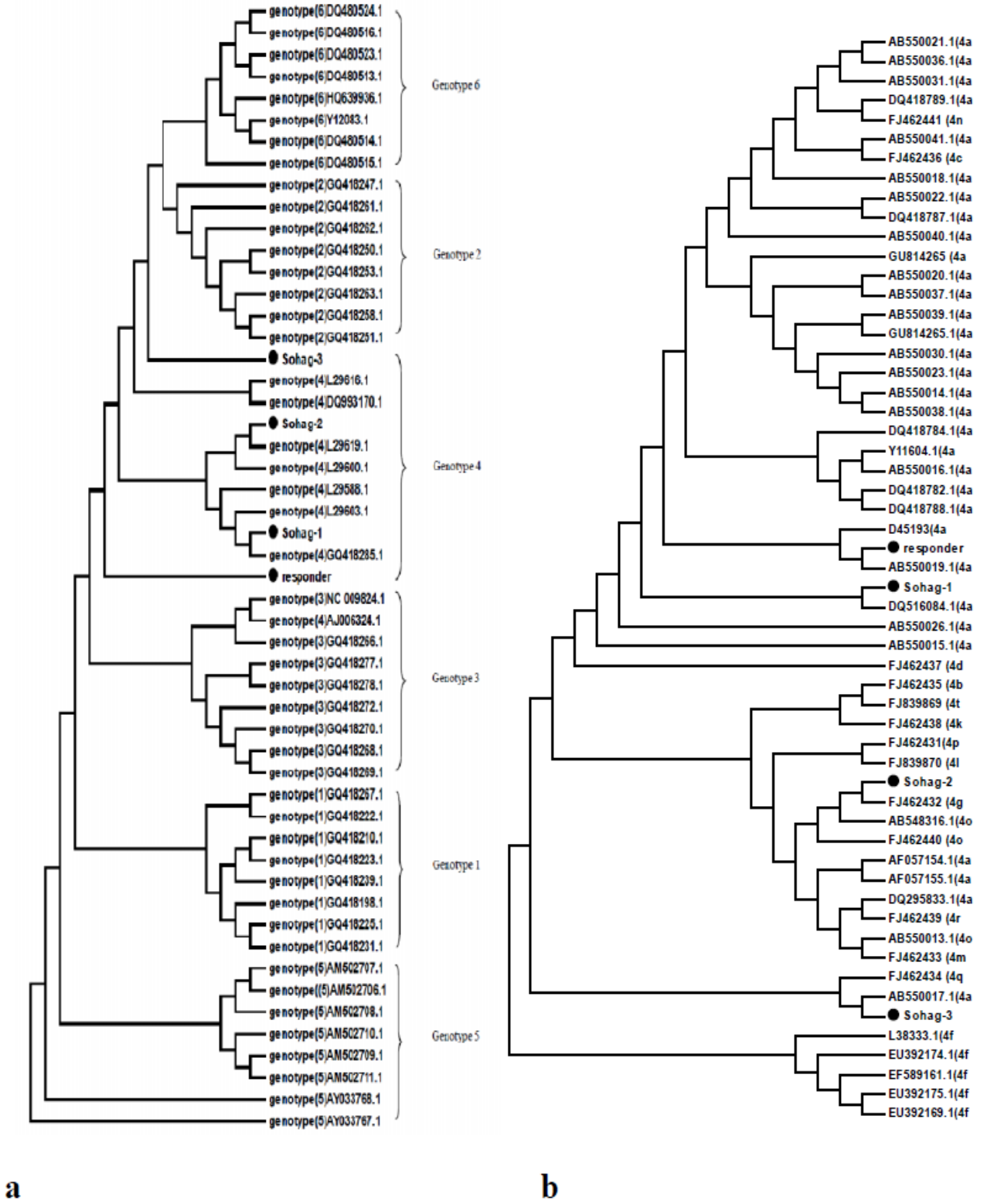


Figure 4. Rooted neighbor-joining tree of HCV 5′-UTR using genotypes 1-6 (a) and subtypes of genotype 4 (b). Bootstrapping of 1000 replicates was carried out.

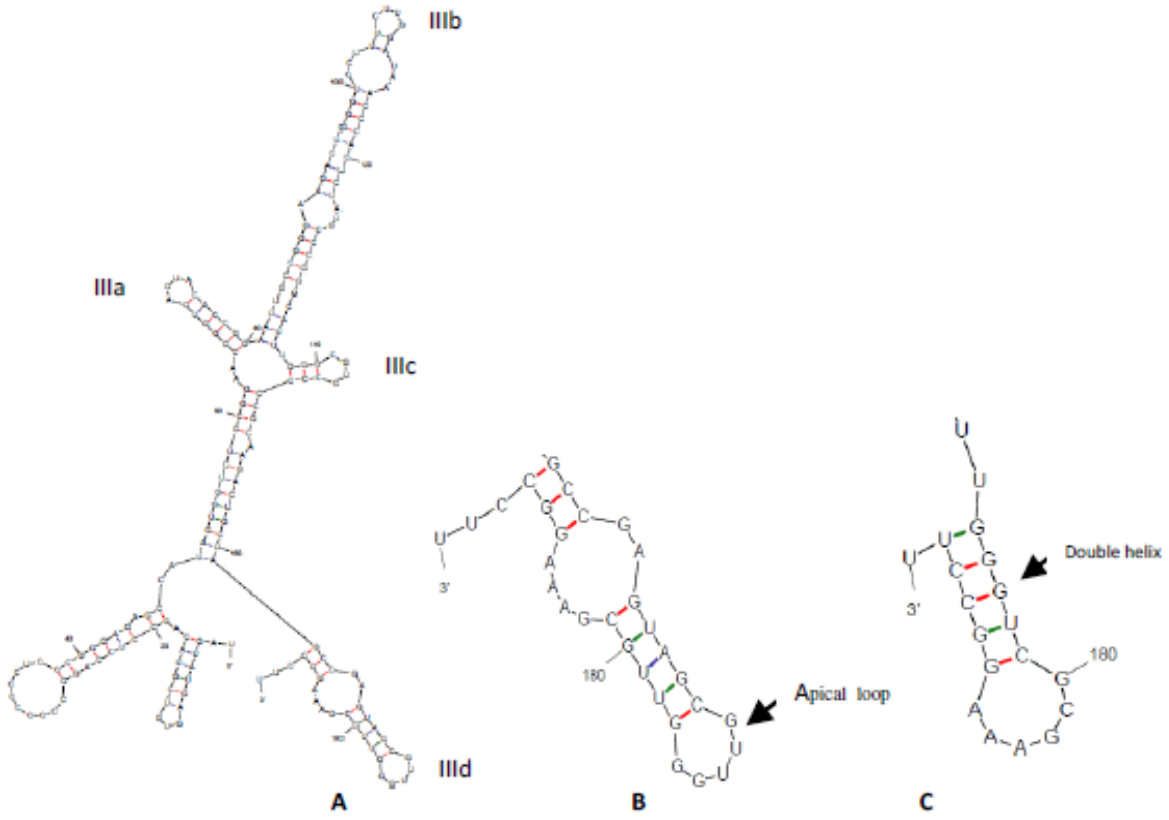


Figure 5. RNA secondary structures of the IRES IIId subdomain of HCV5′ UTR isolates, predicted by mFOLD version 3.2 program. (A) Secondary structure of the entire IRES III domain resulted in 4 stem loop subdomains (IIIa-d). The absence of the 2 transitions $T_{175}$→C and $C_{183}$→T in IIId subdomain stem, affected the formation of the UUGGGU apical loop, it was replaced by double helix structure in most predicted folding (B and C, respectively).

**4. Discussion**

We have previously reported the isolation, amplification and sequence analysis of HCV non-responder isolates (Sohag 1, 2 and 3) from Egypt and they were grouped with genotype 4 (Heimeda *et al*., 2010).

In the present study, sequence of 5′-UTR of the three non-responder isolates (sohag 1, 2 and 3) and one responder isolate reported by Zekri *et al*. (2007) were analyzed. Alignment comparison between the responder and non-responder isolates resulted in one insertion and 2 deletions in the responder isolate. The substitution $G_{154}$→A resulted in our non-responder isolates similar to results reported by Zekri *et al*. (2007). Anila *et al.* (2009) showed that the nucleotide substitutions within the HCV 5`UTR may influence the viral translation and its sensitivity to the antiviral action of interferon.

Extensive studies indicated the importance of HCV genotyping and subtyping in interferon treatment and progression of chronic liver disease (Dammacco *et al.*, 2000; Farci and Purcell, 2000 and Anila *et al*., 2009). Our non-responder and the responder isolates were clustered within genotype 4. Sohag 1 and the responder isolates were clustered with subtype 4a. Sohag 2 was clustered with isolates of subtypes g and o, however, isolate 3 was grouped with isolates from subtype 4a and q. Regrouping of some isolates was previously noticed with isolates genotyped as type 1a or 1b and were found to be wrongly subtyped (Stuyver *et al*., 1995).

Mutations in IIId domain disrupt IRES-mediated translation initiation and also affect the RNA structure, demonstrating the importance of correct RNA folding to IRES function. (Psaridi *et al*., 1999 and Jubin *et al*., 2000). Sequence comparison showed that apical loop nucleotides (UUGGGU) were absolutely conserved across HCV genotypes (Fig. 2), this was also reported by Jubin *et al*., (2000). This conserved sequence corresponds to position 262-270 of the strain H77, Honda *et al*. (1999). Two transitions $T_{175}$→C and $C_{183}$→T were highly

represented in genotype 2 followed by genotype 6 and were not found in the analyzed isolates of genotype 5. Hazari *et al*., (2005) and Klinck *et al*., (2000) mentioned to the presence of these 2 changes in the primary sequence of IIId shows between the six major HCV genotypes. Dual substitution mutants within the IIId terminal loop demonstrated reductions in activity in the range of 18–42% for the U264:U269 series (Klinck *et al*., 2000). Prediction of secondary structure showed that the 2 transitions affect the formation of the apical loop of IIId subdomain. In the absence of these 2 transitions, the apical loop was replaced by double helix structure in most predicted folding. These 2 transitions could be related to the therapeutic response in genotype 2 and 6. Treatment with pegylated interferon and ribavirin resulted in a significantly higher rate of SVR in patients infected with genotype 2 and 6 than in those infected with the other genotypes (Dev *et al*., 2002, Fung *et al*., 2008, Hui *et al*., 2003 and Phillip *et al*., 2009). Our results showed that the occurrence of the 2 transitions was not related to HCC.

This study revealed that our Egyptian isolates were clustered with subtypes 4a, g, o and q. A correlation between the two transitions $T_{175}$→C and $C_{183}$→T in IIId subdomain, the different genotypes and the secondary structure was detected. The 2 transitions were highly represented in genotype 2 followed by genotype 6 where a significantly high rate of SVR was detected for both genotypes. The formation of UUGGGU loop was highly affected by the presence of the 2 transitions $T_{175}$→C and $C_{183}$→T. Taken together, the % of the 2 transitions in different genotypes and the modification in the apical loop structure could affect the response to therapeutic treatment.

**Acknowledgements:**

We thank Dr. Mahmoud M. El-Hefnawy for his very helpful discussions and suggestions.

**Corresponding Author:**
Dr. Amal Mahmoud
Genetic Engineering & Biotechnology Research Institute, Menoufia University, Sadat City Branch, Menoufia , Egypt.
E-mail: amalmh3@yahoo.com

**References**

1. Anila Y, Anwar AS, Saeed H, Taranum S, Wasim J, Mats AA (2009). Genetic variations in a well conserved 5'-untranslated region of hepatitis C virus genome isolated in Pakistan. Journal of Virological Methods 160: 38-47.
2. Dammacco FD, Sansonno C, Piccoli V, Racanelli FP, D'Amore Lauletta G (2000). The lymphoid system in hepatitis C infection: Autoimmunity and cryoglobulinemia and overt B cell malignancy. Semin Liv. Dis. 20: 143-157.
3. Dev AT, Mc Caw R, Sundararajan V, Bowden S, Sievert W (2002). Southeast Asian patients with chronic hepatitis C: The impact of novel genotypes and race on treatment outcome. Hepatology 36:1259–65.
4. Egyptian Ministry of Health. Annual Report (2007).
5. Farci P, Purcell RH (2002). Clinical significance of hepatitis C virus genotypes and quasispecies. Semin Liv. Dis. 20: 103-126.
6. Fung J, Lai C-L, Hung I, Young J, Cheng C, Wong D, Yuen M-F (2008). Chronic Hepatitis C Virus Genotype 6 Infection: Response to Pegylated Interferon and Ribavirin. The Journal of Infectious Diseases 198: 808 –1211.
7. Hazari S, Patil A, Joshi V, Sullivan DE, Fermin CD, Garry RF, Elliott RM, Dash S (2005). Alpha interferon inhibits translation mediated by the internal ribosome entry site of six different hepatitis C virus genotypes. Journal of General Virology 86:3047–3053.
8. Hemeida AA, Osman M, El-Shahat M, Hashem M H, Mahmoud A, Dahi H (2011). Genetic variations in a conserved 5'-Untranslated region of hepatitis C virus isolated from Egypt. International J. of Virology 7(3): 91-99.
9. Honda D, Yokochi T, Nakahara T, Raghukumar S, Nakagiri A, Schaumann K, Higashihara T (1999). Molecular phylogeny of labyrinthulids and thraustochytrids based on the sequence of 18S ribosomal RNA gene. J. Eukaryot. Microbiol. 46: 637–647.
10. Hui CK, Yuen MF, Sablon E, Chan AO, Wong BC, Lai CL (2003). Interferon and ribavirin therapy for chronic hepatitis C virus genotype 6: a comparison with genotype 1. J. Infect. Dis. 187(7):1071–4.
11. Jubin R, Vantuno N, Kieft J, Murray M, Doudna J, Lau J, Baroudy B (2000). Hepatitis C Virus Internal Ribosome Entry Site (IRES) Stem Loop IIId Contains a Phylogenetically Conserved GGG Triplet Essential for Translation and IRES Folding. J. of virology p. 10430–10437.
12. Klinck R, Westhof E, Walker S, Afshar M, Collier A, Aboul-ela F (2000). A potential RNA drug target in the hepatitis C virus internal ribosomal entry site. RNA 6:1423–1431.
13. Mizokami M, Gojobori T, Ohba K, Ikeo K, Ge XM, Ohno T, Orito E, Lau JY (1996). Hepatitis C virus types 7, 8 and 9 should be classified as type 6 subtypes. J. Hepatol. 24: 622-624.

14. Phillip SP, Paul JP, Jeffrey SG, Pang PS, Planet PJ, Glenn JS (2009). The Evolution of the Major Hepatitis C Genotypes Correlates with Clinical Response to Interferon Therapy. PLoS ONE 4(8): e6579. e6579 1-12.
15. Psaridi L, Georgopoulou U, Varaklioti A, Mavromara P (1999). Mutational analysis of a conserved tetraloop in the 59 untranslated region of hepatitis C virus identifies a novel RNA element essential for the internal ribosome entry site function. FEBS Lett. 453:49–53.
16. Robertson B, Myers G, Howard C (1998). Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: proposals for standardization. Arch. Virol. 143: 2493-2503.
17. Simmonds P, Bukh J, Combet C, Deleage G, Enomoto N, Feinstone S, Halfon P, Inchauspe G, Kuiken C, Maertens G, Mizokami M, Murphy DG, Okamoto H, Pawlotsky JM, Penin F, Sablon E, Shin T, Stuyver LJ, Thie HJ, Viazov S, Weiner AJ, Widell A (2005). Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatology 42: 962-73.
18. Stuyver L, Wyseur A, van Arnhem W, Lunel F, Laurent-Puig P, Pawlotsky JM, Kleter B, Bassit LN, kengasong J, van Doorn LJ (1995). Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable samples. Virus Res. 38 (2-3):137-57.
19. Tamura K, Dudley J, Nei M Kumar S: MEGA4 (2007). Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24:1596-1599.
20. Wang CY, Sarnow P, Siddiqui A (1994). A conserved helical element is essential for internal initiation of translaion of hepatitis C virus RNA. J. Virol. 68: 7301-7307.
21. Zein N. Clinical significance of hepatitis C virus genotypes (2000). Clinical Microbiology Reviews 13: 223-23.
22. Zekri NA, Alam El-Din MH, Bahnassy AA, Khaled MM, Omar A, Fouad I, El-Hefnewi M, Thakeb F, El-Awady M (2007). Genetic distance and heterogenecity between quasispecies is a critical predictor to IFN response in Egyptian patients with HCV genotype-4. Virology Journal 4(16): 1-12.
23. Zucker M (1989). On finding all suboptimal foldings of an RNA molecule. Science 244: 48–52.

19/2/2012