Bioinformatic Analysis of Hepatitis C Virus Internal Ribosomal Entry Site Secondary Structure Related To Response to Interferon Combined Therapy in Egyptian Patients

(Received: 08. 09.2019; Accepted: 23.09.2019)

Saad I. M. [1], Attaby F. A.[2], El-Gelil F. A. [2], Salwa, Sabet[3], Saber M.A.[1]
[1]Department of Biochemistry and Molecular Biology, Theodore Bilharz Research Institute, Giza, Egypt.
[2]Department of Chemistry, Faculty of Science, Cairo University, Cairo, Egypt.
3Department of Zoology, Faculty of Science, Cairo University, Cairo, Egypt.

ABSTRACT

Hepatitis C virus (HCV) infection is the major cause of chronic hepatitis. The HCV genotype 4a is predominant in Egyptian patients that are not responding to the combined pegylated interferon-alpha /ribavirin therapy. The HCV IRES is a well-defined structure of about 341 nucleotides in its 5′-untranslated region (UTR). This study conducted on 46 Quantification of HCV-RNA in serum has been carried out at the beginning of treatment (W0) and at week12 (W12), patients 26 responders and 20 non-responders to study HCV IRES sequence and secondary structure for genotype 4a and find the correlation to paginated interferon\ribavirin Responding Treatment in Egyptian patients. There was no statistically significant difference between responders and non-responders regarding gender (P=0.1), age (P=0.3), AFP (P=0.4) and Albumin (P=0.1) while There was a significant increase for AST/ALT (P=0.002, P=0.001), respectively. HCV IRES RT-PCR and sequence analysis including genotyping, MSA and Sequence variation showed that 98% genotype 4a and there was no significant difference in sequence between responders and non-responders patients in addition to the prediction of HCV IRES secondary structure MFE. Centroid revealed that there was no effect on IRES secondary structure and it is conserved among all HCV genotypes.

Key words: *HCV, genotype 4a, qPCR, INF/RBV treatment, 5'UTR, IRES secondary structure, MFE, Centroid, mountain plot, Bioinformatics.*

INTRODUCTION

The HCV is a strong pathogenic virus. It is estimated that up to 150-200 million people are infected with HCV, globally ~3% of the world's population (Atsbaha *et al.,* 2016). It is clearly evident that the incidence of HCV is higher among less developed nations reaching as high as 14.7% in Egypt with a significant reduction in the overall prevalence of HCV to 10.0% in 2017(Kandeel *et al*., 2017). HCV can eventually lead to permanent liver damage, cirrhosis and hepatocellular carcinoma (HCC); the third leading cause of cancer-related death worldwide (Axley *et al*., 2018).

The HCV persists as a collection of virus quasispecies, classified into 6 genotypes and more than 24 sub-genotypes have been identified; where genotype 4a is common in North Africa and the Middle East and especially in Egypt more than 92.5% are infected with genotype 4a (Omran *et al.*, 2018).

The Interferon-α (IFN-α) therapy has been playing a central role in anti-HCV strategies pegylated-IFN-α plus RBV combination therapy came to be a standard treatment, which provided an SVR in about 40%-50% of the patients with HCV infections (Asselah *et al.*, 2010). Patients infected with HCV genotype 1 had SVR rates of about 40% and 50% to 80% SVR rates were accomplished in patients infected with genotypes 2, 3, 5, and 6. Lower SVR rates of 40% to 60% were reached in patients with HCV genotype 4a (Hathorn and Elsharkawy, 2016).

At the 5´-untranslated region (UTR) of the HCV RNA genome, the HCV IRES is a well-defined secondary structure of about 341 nucleotides that play important role in HCV life cycle and HCV infection by either affecting the virus replication or by interfering with the responding to the interferon therapy. Mutations within the HCV genome, genotypes, IRES sequence and secondary structure (Ashraf *et al.,* 2016) may provide an explanation for the divergent data related to the response to interferon therapy. The search for IRES element and its secondary structure confirmations have become a rapidly growing research field that can contribute the scientists to understand the HCV replication and the response to combined interferon/ribavirin therapy (Floden *et al.*, 2016).

This study conducted on 46 Quantification of HCV-RNA in serum has been carried out at the beginning of treatment (W0) and at week12 (W12), patients 26 responders and 20 non-responders to study HCV IRES sequence and secondary structure for genotype 4a and find the correlation to paginated interferon\ribavirin Responding Treatment in Egyptian patients.

MATERIALS AND METHODS

This work was have been done on forty-six patients infected with HCV at Theodor Bilharz Research Institute (TBRI) in Biochemistry and Molecular Biology Department, from December 2014 to April 2019. The patients enrolled in this study were treated for 12 weeks by pegylated interferon alpha-2a (Peg-IFN-α) (PEGASYS®; Hoffmann-La Roche, Basel, Switzerland) at a dose of 180 µg/kg once/week combined with ribavirin (1000-1200 mg/kg) (Hu et al, 2019). The viral load was quantitated before (W0) and after treatment (W12) using the quantitative Real-Time polymerase chain reaction (qPCR). The samples were a part of project funded by Science and Technology Development Fund in Egypt (STDF) under ID: 1763 (Project type: TC/2 Healthl2009/Hep). All subjects in this study were approved by the ethics committee of Theodor Bilharz research institute (TBRI) according to "the institutional committee for the protection of human subjects and adopted by the 18th world medical assembly, Helsinki, Finland".

HCV RNA extraction and quantitation

The HCV total RNA extraction was performed using Abbott® mSample preparation system kit, cat.no (02K02-96) (Abbott Molecular, Inc., Des Plaines, IL). The HCV qPCR have been quantified using the Abbott® real-time HCV amplification reagent pack RT-qPCR (cat.no. 1N30) (Abbott Molecular, Inc., Des Plaines, IL) with a detection limit of 12 IU/ml.

HCV IRES amplification

The HCV IRES RT-PCR amplification was performed in total volume of 50 µL containing 10 µL of 5x green GoTaq® Reaction Buffer (1.5mM $MgCl_2$ included), 200 µM dNTPs, 1.5 unit Taq DNA polymerase (Gotaq ® DNA, Cat No. M3005, Promega, Inc., USA), 50 units of Moloney Murine Leukemia Virus (MMLV) reverse transcriptase, (Cat No. M1701, Promega, Inc. USA), 30 µL HCV RNA (Warkad *et al.,* 2018) and 20 pmol of each of denovo specific primers which is designed specifically to amplify HCV-4a IRES by our research team and submitted in NCBI database under (PopSet: 1333190374, with the accession no Genbank KY981528-KY981556), IRES

Forward primer: 5' TTGGGGGCGACACTCCAC 3' and IRES Reverse primer: 5' CTTTGAGGTTTAGGAATTCGTGCTC 3'.

The IRES RT-PCR reaction was performed in a T100^{TM} thermal cycler (Bio-Rad® PCR systems, Inc. USA). The thermal cycling condition was as follows : The RNA was denatured by heating at 70^{o}C for 20 min to overcome highly RNA structure prior to RT-PCR, followed by 42°C for 40 min, 95°C for 5 min, and 30 cycles of 95°C for 30 min, 55°C for 30 sec, 72°C for 30 sec, and final extension at 72°C for 10 minutes.

The final IRES PCR products obtained were subjected to 3% agarose gel electrophoresis (Green and Sambrook 2019).A clear sharp band was observed at 360bp in both responders and Non responder's samples and purified using Genedirex®Gel Extraction kit (Cat No. NA006-0100, Genedirex® Inc., Taiwan).

The DNA sequencing of HCV IRES region

The sequences of the responders and non-responder's IRES amplified fragments were determined by direct sequencing in two directions (Sanger *et al.,* 1977). Formation of contig sequence and trimming the low quality bases at the end of two-directional sequencing samples have been done using DNA Chromatogram Explorer Lite, Heracle BioSoft(2013).

Table (1): Characteristics and biochemical data of patients.

Patients characteristics	**R (n= 26)**	**NR (n= 20)**	P. value
Age (yrs.)	38.6 ± 5.7	37.1 ± 3.83	0.3
Gender (F/M)	12/14	14/6	0.1
ALT (IU/L)	36.96±1.30	45.48±1.16	0.001*
AST (IU/L)	33±0.87	44.83±1.12	0.002*
Albumin (g/dL)	3.74 ± 0.36	3.8 ± 0.29	0.1
AFP (ng/mL)	2.4±2.23	2.83±1.70	0.2
HCV PCR (W0) (IU/ml)	4.41 ± 4.63	4.42±4.37	
HCV PCR (W12) (IU/ml)	Negative	3.33±3.27	0.07

Data are expressed as mean ± SD,R= Responders, NR = Non-Responders, n= number of patients, F: female, M: male, ALT: alanine aminotransferase, AST: aspartate aminotransferase and AFP: alpha-fetoprotein.

Bioinformatics analysis of HCV IRES sequences

Sequence analysis has been done using different Softwares and online servers to understand its features, function, structure, and evolution. The bioinformatics module was performed as following: HCV genotyping, homology and identity against the NCBI database using blastn limited to secondary non-redundant (nr) database and to the HCV organism (taxoid #:11103). Multiple/Pairwise Sequence Alignment has been done using Clustal Omega, kalign; EMBOSS Water *(Li et al.,* 2015) and Bio edit software. The Construction of Rooted circle, Cladogram and Real distance Phylogenetic trees have been done for the HCV IRES samples using Simple Phylogeny by EMBL-EBI Centre *(*De Bruyn *et al.,* 2014).

```
CLUSTAL O(1.2.4) multiple sequence alignment (all )

R19     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R18     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR3     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR15    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R15     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R4      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR1     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR18    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R1      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R8      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R25     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR5     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R7      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR19    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR6     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR2     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR4     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R22     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R14     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R16     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR11    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR20    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R10     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R11     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R24     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R2      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R13     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R5      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR8     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R3      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R17     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R21     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR17    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R23     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR10    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R26     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR12    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R12     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R20     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R6      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
R9      TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR9     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR16    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR14    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR7     TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
NR13    TTGGGGGCGACACTCCACCATGAACCGCTCCCCTGTGAGGAACTACTGTCTTCACGCAGA    60
        ************************************************************
```

Fig.(1): Nucleotide sequence comparison of the HCV IRES for R(26 isolate) and NR(20 isolates) where The conserved sequences are shown as dots in (A) MSA of samples using CLUSTAL Omega, high of blocks with a percentage identity value was equal to 97.52% in (B) Bioedit®

Prediction and identification of HCV IRES secondary structure

The minimum free energy (MFE) optimal secondary structure prediction and centroid structure model have been predicted using to annotate and understanding the relation between IRES sequence and structure corresponding to its biological function in translation initiation of HCV polyprotein and response to combined INF/RBV treatment (Fricke *et al.,* 2015).

The Prediction of HCV IRES secondary structure module was performed as following: detecting the presence of potential IRES functional elements in the sequences have been done using IRESPred© (http://196.1.114.46:1800/IRESPred/home.html). HCV IRES MFE and Centroid optimal secondary structure was predicted by single stranded RNA sequence traces using RNAfold, (http://rna.tbi.univie.ac.at//cgibin/RNAWebSuite/RNAfold.cgi).

Dot-Bracket secondary structure, Mountain plot representations and Thermodynamic Ensemble of the structures for IRES Sequences have been predicted using RNAeval (http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/RNAeval.cgi).TheDot plot Sequence Base pairing probability has been plotted and visualized using mfold (v3.6 UNAFold), (http://unafold.rna.albany.edu/?q=unafold-man-pages).

Analysis of HCV IRES secondary structure Multiple / Pairwise Alignment have been performed by The BEAGLE® and LocARNA® between two or more HCV IRES secondary structures. The method exploits RNA secondary structure substitution matrix and computes multiple alignments of HCV IRES based on their sequence and structure similarity (Ayres *et al.,* 2019).

Statistical Analysis

Data were analyzed using the SPSS; comparison between Responders and Non responder groups was made by the Chi-square or Fisher exact test and the Students t-test. A probability value ($P \leq 0.05$) was considered statistically significant.

RESULTS

Characteristics of patients

This study has been conducted on forty-six patients infected with HCV- 4a. Among them, 26 responders and 20 non-responders; the majority of patients were in the age group of twenty-five to fifty-five years. Physical and clinical characteristics of patients are given in Table (1). Our results showed that no statistically significant difference between R and NR regarding age (p=0.3), gender (p=0.1), AFP (p=0.4), Albumin (p=0.1) Also, there is a significant increase among AST/ALT (p=0.002, p=0.001) respectively, after 12 weeks of treatment with combined PEG-IFNα2a/RBV.

HCV viral load qReal-Time quantification

The viral load was quantitated before (W0) and after treatment (W12) which reveals that the HCV patients were divided into two main groups depending on their responses to the treatment, 26 responders and 20 non-responders. There was no significant difference have been observed (P=0.07) as shown in Table (1), indicating that patients were selected randomly.

Table (2): The HCV IRES predicted secondary structure folding parameters and their characteristics for both groups R and NR, respectively.

	MFE - kcal/mol		Thermodynamic Ensemble - kcal/mol		Frequency %		Ensemble Diversity		Centroid - kcal/mol	
sample no	R N=26	NR N=20	R N=26	NR N=20	R N=26	NR N=20	R N=26	NR N=20	R N=26	NR N=20
1	127.30	127.30	131.32	131.40	0.15	0.13	30.3	30.73	126.50	126.50
2	123.00	123.90	128.10	129.76	0.03	0.01	92.46	95.80	73.60	89.89
3	127.50	127.60	131.75	132.58	0.10	0.03	51.17	92.61	119.1	82.60
4	121.90	123.30	126.00	128.46	0.13	0.02	66.91	62.09	120.00	114.40
5	126.1	128.20	131.56	133.09	0.01	0.04	52.63	61.98	120.8	117.90
6	123.50	122.30	129.08	128.29	0.01	0.01	88.09	54.23	80.40	114.30
7	128.00	124.20	133.33	129.91	0.02	0.01	67.87	94.61	95.50	80.10
8	127.20	125.90	131.16	130.23	0.16	0.09	30.04	77.03	126.40	103.10
9	124.80	128.90	130.24	132.48	0.01	0.30	75.65	37.76	92.90	126.20
10	118.80	125.50	124.03	130.13	0.02	0.05	45.78	92.06	116.10	87.40
11	119.60	125.10	125.90	130.79	0.00	0.01	47.01	87.51	110.00	89.00
12	124.80	128.90	129.71	133.19	0.03	0.10	96.75	66.05	72.80	100.00
13	125.90	124.20	130.51	129.91	0.06	0.01	36.90	94.48	122.30	80.10
14	123.90	125.20	128.88	130.05	0.03	0.04	93.02	93.57	75.00	82.60
15	120.40	132.10	125.33	135.25	0.03	0.61	62.79	20.95	108.50	131.70
16	126.30	125.70	130.95	131.11	0.05	0.02	76.22	88.97	104.60	67.50
17	126.70	124.50	131.23	130.92	0.06	0.00	39.00	87.07	125.00	76.30
18	129.20	128.70	133.31	132.69	0.13	0.15	45.75	46.81	121.00	127.90
19	130.20	116.20	134.81	122.47	0.06	0.00	63.34	70.07	121.80	102.50
20	124.80	125.50	129.71	130.61	0.03	0.03	97.44	84.32	72.80	92.00
21	118.50		124.47		0.01		93.97		75.30	
22	129.50		133.80		0.09		72.03		119.00	
23	123.80		128.61		0.04		58.34		105.20	
24	115.70		122.13		0.00		107.09		58.70	
25	127.00		132.34		0.02		63.66		111.50	
26	121.70		127.45		0.01		59.16		103.80	
Mean ± SD	124.5±3.7	125.7±3.2	129.5±3.3	130.7±2.6	3 (1 - 6.75)	3 (1 – 9.75)	65.9±22.3	71.9±23.5	103.0±20.9	99.6±19.7
P.value	0.2		0.1		0.9		0.3		0.5	

All parameters are represented as Mean ± SD, except frequency is represented as median (Interquartile range) (25%-75%)
P-value <0.01 is significant, P-value <0.001 is highly significant.

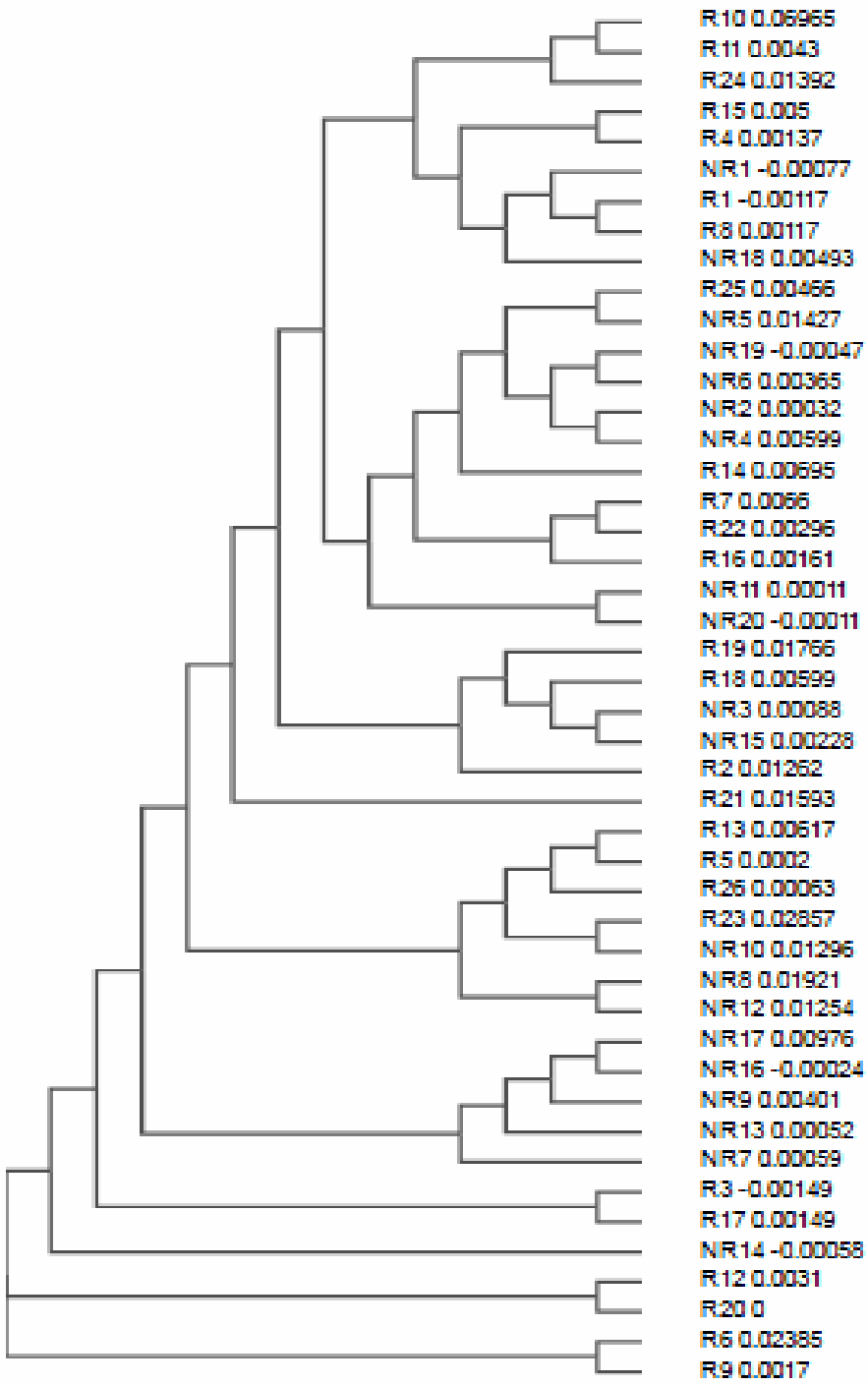

Fig.(2): Cladogram Phylogenetic tree using Neighbor-joining without distance correction between HCV IRES isolates. The tree shows the phylogenetic correlation of HCV IRES genotype 4a sequences which share the common ancestor and branch off the same clade with a maximum distance between isolate was equal to 0.009; no clustering have been observed.

DNA sequencing of HCV IRES region

The HCV IRES PCR product of 360bp was sequenced by the automated direct sequencing in two directions. The HCV IRES sequence for R and NR reported in this study has been deposited in the GenBank nucleotide sequence databases with (PopSet: 1333190374, with the accession no Genbank KY981528-KY981556).

Sequence Homology Analysis and genotyping

Homology sequence identity and HCV Genotyping against the NCBI database shows identities equal to 98% HCV-4a.

MSA and phylogenetic tree analysis

Multiple Sequence Alignment (MSA) of samples using Clustal Omega and Bioedit in which, homology, mutation and the evolutionary relationships between the sequences have been annotated. The conserved (consensus) sequences are shown as high of blocks with a percentage identity value was equal to 97.52% (Fig.1). The cladogram phylogenetic tree shows the phylogenetic correlation of HCV IRES genotype 4a sequences, revealed that no clustering have been observed as shown in Fig. (2).

HCV IRES secondary structure prediction

MFE structure drawing and Centroid structure predicted shows the same pattern for R and NR isolate encoding positional entropy as color hue. Where the positional entropy for each position with Scale of (0 - 2.4), (0-2.6), respectively, revealing the core of IRES element UUGGGU in the (IIId apical loop) is shown and persists in both isolates with well-defined and indeed predicted correctly. Statistically there is no significant difference regarding their MFE (P=0.2), Thermodynamic Ensemble (P=0.1), Frequency (P=0.9), Ensemble Diversity (P=0.3) and Centroid (P=0.5) as shown in Table (2).

The Dot Plot and the mountain plot of base pairs enclosing a sequence position (height) versus the position. Shows three curves, mountain plots derived from the MFE structure and the fold pairing probabilities and the centroid structure in addition to a positional entropy curve. In general the closer the curves, the better-defined structure. The HCV IRES for R and NR isolates shows a similar Dot and mountain plot curves as shown in Fig (3).

HCV IRES secondary structures Pairwise alignments between two sets (RNA1, RNA2) for the R and NR samples, respectively using BEAGLE®, where the Z-score for all the alignments was greater than 5.78 reveals that there was no significant difference between the R and NR HCV IRES secondary structures has been observed (Fig. 4).

DISCUSSION

Egypt represents the world's highest prevalence of HCV infection (Abdel-Ghaffar *et al.,* 2015) reaching as high as 14.7% in Egypt with a significant reduction to 10.0% (Kandeel *et al.,* 2017) .HCV Genotype-4a of the virus is the common strain in Egypt(El-Tahan *et al.,* 2018). It responds poorly to the PEG-IFN-α2a/RBV combination therapy and had SVR rates of about 40% (Papastergiou and Karatapanis 2015,(Asselah *et al.,* 2010. and Hathorn and Elsharkawy, 2016).

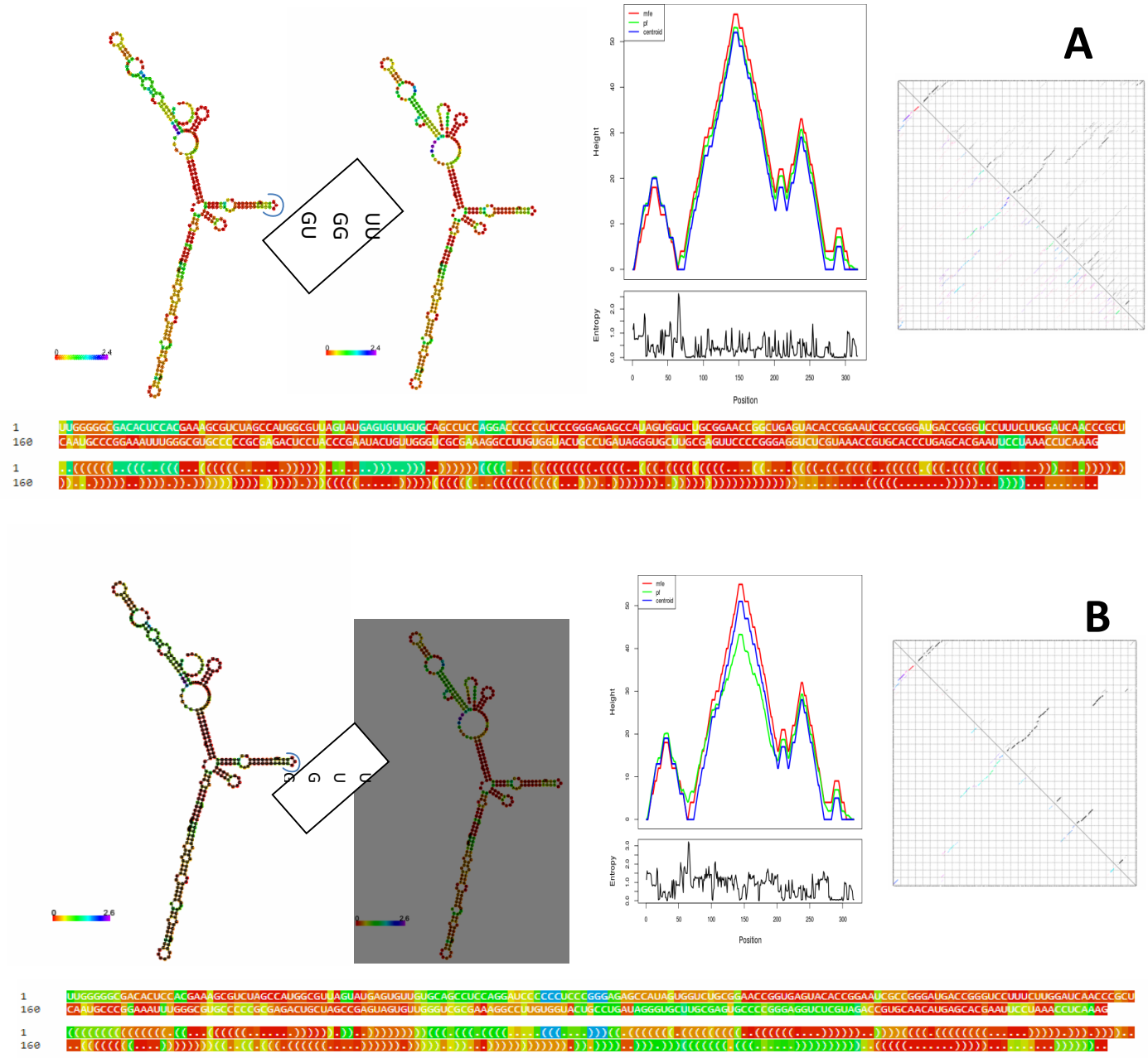

Fig.(3): A-Responder, B-Non responder, MFE and Centroid structure predicted encoding positional entropy as color hue, red(low entropy, well-defined) via green to blue and violet (high entropy, ill-defined).core of IRES element UUGGGU in the (IIId apical loop) is shown and persists with well-defined(red).The mountain plotting, sequence position (height) versus the position, Terminal loops (peaks), nested helices (stacking) and multiple-loops (plateau).three curves, MFE structure mountain plots (red) and fold pairing probabilities (green) and centroid structure (blue), positional entropy curve (black). Dot Plot Computed base pair probability matrices and align the sequences based on their full structures.

Toggle Color Download Results

	RNA 1 (Id)	RNA 2 (Id)	Seq Id (%)	Str Id (%)	Str Sim(%)	p-value	z-score
	R1	NR1	100	100	100	0.0001	11.234
	R8	NR1	99.37	99.68	99.68	0.0001	11.141
	R4	NR1	99.68	88.96	98.74	0.0001	10.756
	R14	NR2	98.11	100	100	0.0001	10.629
	R9	NR2	99.68	98.73	100	0.0001	10.445
	R5	NR6	99.68	91.75	99.68	0.0001	10.276
	R13	NR6	99.68	85.03	96.82	0.0001	9.608
	R20	NR7	99.68	81.27	83.81	0.0001	9.194
	R12	NR7	99.37	81.27	83.81	0.0001	9.181
	R3	NR7	99.68	77.14	83.17	0.0001	8.927
	R20	NR10	97.78	78.41	82.86	0.0001	8.822
	R12	NR10	97.46	78.41	82.86	0.0001	8.81

Toggle Color Download Results

	RNA 1 (Id)	RNA 2 (Id)	Seq Id (%)	Str Id (%)	Str Sim(%)	p-value	z-score
	R1	NR1	100	100	100	0.0001	11.234

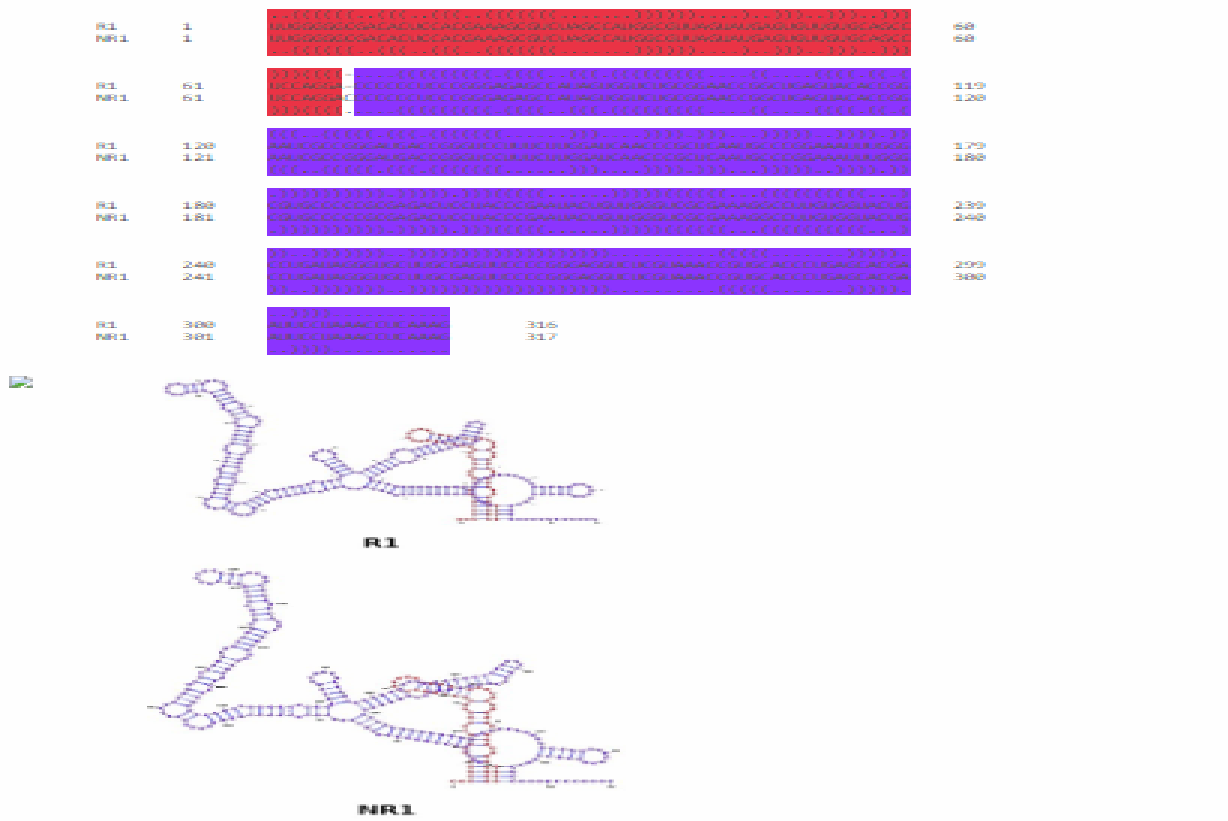

Fig.(4): Pairwise alignments of HCV IRES secondary structures have been done between two sets (RNA1, RNA2) for the R and NR samples respectively samples by using BEAGLE®, where the Z-score for all the alignments was greater than 5.78 which reveals that there were no significantly difference between the R and NR to the HCV IRES secondary structures .

The viral RNA genome is of importance for response to interferon-based treatment. Mutations within the HCV genome, genotypes, IRES sequence and secondary structure (Ashraf *et al.,* 2016) provide an explanation for the divergent data related to the response to interferon therapy. Bioinformatics analysis and computational techniques applied (Barria et al., 2009) helps in revealing and identification of the HCV IRES sequence and secondary structure to find the correlation between IRES and respond to combined pegylated interferon\Ribavirin therapy.

In this study, all the 46 patients infected with HCV were treated by pegylated interferon alpha-2a in combination with ribavirin for 12 weeks, the qPCR reveals that the HCV patients were 26 responders and 20 non-responders. Statistically, there was no significant difference (p=0.07) indicating that the patients were selected randomly (Sharvadze *et al.,* 2009) , regarding age (P=0.3), gender (P=0.1), AFP (P=0.4) and Albumin (P=0.1) Also, regarding AST/ALT, There is a significant increase (P=0.002, P=0.001) respectively. This explains that there was no clear evidence between responding to the treatment and clinical picture of the patient (Hetta *et al.,* 2015). In our study, the MSA revealed that there was sequence variation in bases with no statically difference at the nucleotide sequence level that can be figured out which shows a high level of conserved regions upon all samples with percentage identity (97.52%). These results were consistent with previous studies .It was observed, as mentioned previously, that the sequence variability of IRES does not appear to correlate with any difference in serum HCV-RNA concentration.

This study reports, that IIId domain sequence is conserved where the core the GGG triplet sequence at the positions of 253-255nt at the apical loop are well conserved among HCV IRES R and NR samples and persist constantly in the predicted IRES secondary structure . These finding are similar to Shimoike *et al.,* (2006) and agree with Jubin *et al.,* (2000),who reported that UUGGGU were absolutely conserved across HCV genotypes ,Another study by El Awady *et al.* (2009) showed that the mutation at the core GGG triplet sequence have a dramatical inhibition effect on HCV translation efficacy our findings shows that there was no correlation between HCV IRES translation efficacy and INF treatment response, these results are in agree Saiz *et al.,* (1999) while another study by Yasmeen *et al.,* (2006) showed a correlation at which the inhibition of translation efficacy was grater with the responder patients but the reported difference did not reach statical significance.

The evolutionary phylogenetic tree shows that the 46 isolates share the common ancestor and branch off the same clade, no clustering have been observed; those of results are similar to (Ashraf, Chakravarti et al. 2016). IRES MFE and centroid secondary structures prediction shows there was no statistically difference (P=0.2, P=0.5) respectively, Thermodynamic Ensemble) P=0.1), Frequency (P=0.9), Ensemble Diversity (P=0.3). these results reveals statistically no correlation to response to treatment (Hermann, 2016) as there was no sequence substitutions. An earlier study by Araujo *et al.* (2008) agree with these results. There was no correlation between the HCV IRES secondary structure and the genotype of HCV as the IRES Conesus secondary structure was highly conserved among genotypes and no correlation to response to combined interferon \ Ribavirin treatment.

To our knowledge, this study is the first in finding HCV IRES genotype 4a regarding

the MFE structure, centroid structure and mountain plots as well as ensemble diversity, thermodynamic ensemble and its correlation to INF/RBV combination treatment response.

CONCLUSION

The studied HCV IRES sequence traces and secondary structures were conserved in responders compared to non-responders in HCV-4a Egyptian patients and there was no correlation to the combined pegylated interferon alpha /ribavirin therapy. These results suggest that HCV IRES sequence and secondary structure prediction cannot be used as single or combined biomarkers to predict the responding of to the combined pegylated interferon alpha /ribavirin therapy in HCV-4 Egyptian patients.

REFERENCES

Abdel-Ghaffar, T. Y., Sira M. M.and El Naghi S. (2015). "Hepatitis C genotype 4: The past, present, and future." World J. Hepatol .,7(28): 2792-2810.

Araujo, F. M., Sonoda I. V. RodriguesN. B., Teixeira R. Redondo R. A. and Oliveira G. C. (2008). "Genetic variability in the 5' UTR and NS5A regions of hepatitis C virus RNA isolated from non-responding and responding patients with chronic HCV genotype 1 infection." Mem Inst Oswaldo Cruz 103(6): 611-614.

Ashraf, A., Chakravarti A. Roy P. Kar P. and Siddiqui O. (2016). "Frequency of nucleotide sequence variations in the internal ribosome entry site region of hepatitis C virus RNA isolated from responding and non-responding patients with hepatitis C virus genotype 3 infection." Virusdisease 27(3): 251-259.

Asselah, T., Estrabaud E. Bieche I. Lapalus M. De Muynck S. Vidaud M. Saadoun, D. Soumelis V. and Marcellin P. (2010). "Hepatitis C: viral and host factors associated with non-response to pegylated interferon plus ribavirin." Liver Int., 30(9): 1259-1269.

Atsbaha, A. H., Asmelash Dejen T. Belodu R., Getachew, K., Saravanan M. and Wasihun A. G. (2016). "Sero-prevalence and associated risk factors for hepatitis C virus infection among voluntary counseling testing and anti retroviral treatment clinic attendants in Adwa hospital, northern Ethiopia." BMC Res Notes 9: 121.

Axley, P., Ahmed Z. Ravi and S. Singal A. K. (2018). "Hepatitis C Virus and Hepatocellular Carcinoma: A Narrative Review." J. Clin. Transl. Hepatol. 6(1): 79-84.

Ayres, D. L., Cummings M. P., Baele G. Darling A. E. Lewis, P. O., Swofford, D. L., Huelsenbeck, J. P., Lemey, P., Rambaut A. and Suchard M. A. (2019). "BEAGLE 3: Improved Performance, Scaling, and Usability for a High-Performance Computing Library for Statistical Phylogenetics." Syst Biol.

Barria, M. I., Gonzalez A. Vera-Otarola J. Leon U., Vollrath V., Marsac D., Monasterio O., Perez-Acle T., Soza A. and Lopez-Lastra M. (2009). "Analysis of natural variants of the hepatitis C virus internal ribosome entry site reveals that primary sequence plays a key role in cap-independent translation." Nucleic Acids Res., 37(3): 957-971.

De Bruyn, A., Martin D. P. and Lefeuvre P. (2014). "Phylogenetic reconstruction methods: an overview." Methods Mol. Biol., 1115: 257-277.

El-Tahan, R. R., Ghoneim A. M. and Zaghloul H. (2018). "5' UTR and NS5B-based genotyping of hepatitis C virus in patients from Damietta governorate, Egypt." J. Adv. Res. 10: 39-47.

El Awady, M. K., Azzazy H. M., FahmyA. M., Shawky S. M., Badreldin, N. G., Yossef, S. S., Omran, M. H., Zekri A .R. and Goueli S. A. (2009). "Positional effect of mutations in 5'UTR of hepatitis C virus 4a on patients' response to therapy." World J. Gastroenterol., 15(12): 1480-1486.

Floden, E. W., Khawaja, A. Vopalensky V. and Pospisek M. (2016). "HCVIVdb: The hepatitis-C IRES variation database." BMC Microbiol., 16(1): 187.

Fricke, M., Dunnes N., Zayas M., Bartenschlager R., Niepmann M. and Marz M. (2015). "Conserved RNA secondary structures and long-range interactions in hepatitis C viruses." RNA 21(7): 1219-1232.

Green, M. R. and Sambrook J. (2019). "Agarose Gel Electrophoresis." Cold Spring Harb Protoc., 2019(1): pdb prot100404.

Hathorn, E. and Elsharkawy A. M. (2016). "Management of hepatitis C genotype 4 in the directly acting antivirals era." BMJ Open Gastroenterol., 3(1): e000112.

Hermann, T. (2016). "Small molecules targeting viral RNA." Wiley Interdiscip Rev RNA 7(6): 726-743.

Hetta, H. F., Mekky M. A., Khalil N. K., Mohamed W. A., El-Feky M. A., Ahmed, S. H., Daef, E. A., Nassar, M. I., Medhat, A., Sherman K. E. and Shata M. T. (2015). "Association of colonic regulatory T cells with hepatitis C virus pathogenesis and liver pathology." J. Gastroenterol. Hepatol., 30(10): 1543-1551.

Hu, J. H., Chang M. L., Huang T. J., Yeh C. T., Chiu W. N. Chiang M. S. and Chen M. Y. (2019). "Comparison of Compliance and Efficacy of Pegylated Interferon alpha-2a and alpha-2b in Adults with Chronic Hepatitis C." J Interferon Cytokine Res., 39(4): 205-213.

Jubin, R., Vantuno N. E., Kieft J. S., Murray M. G., Doudna, J. A., Lau J. Y. and Baroudy B. M. (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. Virol., 74(22): 10430-10437.

Kandeel A., Genedy M., El-Refai S, Funk A. L., Fontanet A. and Talaat M. (2017). "The prevalence of hepatitis C virus infection in Egypt 2015: implications for future policy on prevention and treatment." Liver Int., 37(1): 45-53.

Li, W., Cowley A., Uludag M., Gur T., McWilliam H., Squizzato S., Park Y. M., Buso N. and Lopez R. (2015). "The EMBL-EBI bioinformatics web and programmatic tools framework." Nucleic Acids Res., 43(W1): W580-584.

Omran D., Alboraie M. Zayed R. A., WifiM. N. Naguib, M. Eltabbakh, M., Abdellah M., Sherief A. F., Maklad, S., Eldemellawy H. H., Saad, O. K., Khamiss D. M. and El Kassas M. (2018). "Towards hepatitis C virus elimination: Egyptian experience, achievements and limitations." World J. Gastroenterol., 24(38): 4330-4340.

Papastergiou, V. and Karatapanis S. (2015) . "Current status and emerging challenges in the treatment of hepatitis C virus genotypes 4 to 6." World J. Clin. Cases, 3(3): 210-220.

Saiz, J. C., Lopez de Quinto, S., Ibarrola, N., Lopez-Labrador F. X. Sanchez-Tapias J. M., Rodes J. and Martinez-Salas E. (199 .(Internal initiation of translation efficiency in different hepatitis C genotypes isolated from interferon treated patients." Arch Virol 144(2): 215-229.

Sanger, F., Nicklen S. and Coulson A. R. (1977). "DNA sequencing with chain-terminating inhibitors." Proc. Natl. Acad. Sci. U S A 74(12): 5463-5467.

Sharvadze, L. G., Gogichaishvili S., Sakandelidze T., G., Zhamutashvili M. T. and Chkhartishvili N. I. (2009). "Re-

treatment of patients with hepatitis C who failed to respond (nonresponders) to previous treatment." Georgian Med. News(166): 61-64.

Shimoike T., Koyama C., Murakami K., Suzuki R., Matsuura Y., Miyamura T. and Suzuki T. (2006). "Down-regulation of the internal ribosome entry site (IRES)-mediated translation of the hepatitis C virus: critical role of binding of the stem-loop IIId domain of IRES and the viral core protein." Virology 345(2): 434-445.

Warkad, S. D., Nimse S. B. Song K. S. and Kim T. (2018). "HCV Detection, Discrimination, and Genotyping Technologies." Sensors (Basel), 18(10).

الملخص العربي

استخدام المعلوماتية الحيوية فى التركيب الثانوى لمنطقه الدخول الي الريبوسوم الداخلى فى فيروس الالتهاب الكبدى سي وعلاقته بالاستجابه لعلاج الانترفيرون المدمج فى حالات المرضى المصريين

إسماعيل محمد سعد[١] ،فتحى عبد الجليل[٢] ،سلوى ثابت [٢]،فوزى على عتابى[٣] و محمد على صابر[١]

[١]قسم الكيمياء الحيوية والبيولوجيا الجزيئية – معهد تيودور بلهارس للابحاث – الجيزة – مصر

[٢]قسم الكيمياء – كلية العلوم – جامعة القاهرة –مصر

[٣]قسم علم الحيوان –كلية العلوم –جامعة القاهرة- مصر

عدوى فيروس التهاب الكبد الوبائي هي السبب الرئيسي لالتهاب الكبد المزمن. النمط الوراثي (٤ أ) هو السائد في المرضى المصريين الذين لا يستجيبون لعلاج انترفيرون ألفا / ريبافيرين. إن تسلسل منطقه فيروس التهاب الكبد الوبائي (سي) IRES عبارة عن هيكل محدد جيدًا لحوالي ٣٤١ نيوكليوتيد في منطقته غير المترجمة(٥ ´ UTR). وقد أجريت هذه الدراسة على ٤٦ مريضا مصريا. تم قياس وتقدير كمية فيروس التهاب الكبد الوبائي (سي) في بداية العلاج (الاسبوع الاول) وفي الأسبوع ١٢ من بدايه العلاج باستخدام تقنيه تفاعل البلمرة المتسلسل العكسي بخطوة واحدة و ذلك بعد استخراج الحمض النووي الريبي الفيروسي من مصل المرضى باستخدام طريقة الخرز المغناطيسي. وفقا للقياس الكمي ، تم تقسيم المرضى إلى ٢٠ مستجيبين وتم تقييم ٢٦ غير مستجيبين للعلاج بمضاد للفيروسات ألفا / ريبافيرين مجتمعان و تم دراسه معلومات سريرية مختلفة بما في ذلك العمر والجنس ،و قياس و تقدير مستويات كلا من انزيمات ALT، AST، ALB، AFP فى الدم . تم دراسة جميع تسلسل IRES والهياكل الثانوية في المستجيبين مقارنة مع المرضى المصريين الذين لا يستجيبون للعلاج ولم يكن هناك أي علاقة للعلاج المشترك باستخدام عقار الإنترفيرون ألفا / ريبافيرين. تشير هذه النتائج إلى أنه لا يمكن استخدام تسلسل فيروس التهاب الكبد الوبائي (سي) IRES والتنبؤ بالهيكل الثانوي كمؤشرات حيوية مفردة أو مجتمعة للتنبؤ باستجابة مرضى فيروس التهاب الكبد الوبائي (سي) نمط (٤ أ) الوراثي للعلاج المضاد للفيروسات باستخدام عقار الإنترفيرون ألفا / ريبافيرين المشترك.