Effectiveness of Mutated Nucleotides in Apical Region

Alaa Eldin Abdallah Yassin[1], Amr Badr2, Ahmed Farouk Al-Sadek[1]

[1] Central Lab for Agricultural Experts Systems, Ministry of Agriculture and Land Reclamation
Giza, Egypt

[2] Department of Computer Science, Faculty of Computers and Information, Cairo University
Cairo, Egypt

Abstract

The biochemical studies have proven that; some domains in the genome have importance more than other regions, because it contains the function parts of those regions. Apical domain in internal ribosome entry site (IRES) of foot and mouth disease virus (FMDV) is one of these domains. There are some nucleotides almost located in conserved motifs of these domains; effect on the structure of these domains which may change its function. Some researches interested in studying the effects which happen in the structure and function of the sequence when some mutation occurred, where these mutations sometimes have a big impact up to producing a new serotype of the viruses as a result of these mutations like SAT1, SAT 2 and SAT3 serotypes. In this paper we produced an observation to the changes of Apical domain in IRES of FMDV on 282 Apical sequences belonging to four different serotypes collected from NCBI portal, and we studied the stable locations and the variable parts in the sequence, and also the paper explained the effectiveness of these mutations in the structure. We introduce this effort as a step to help biologists in knowing the unstable parts in the importance regions which may help in founding the suitable drug or expect the new serotypes of this dangerous virus.

Keywords: IRES; Virus mutation; Apical region; Nucleotides substitution.

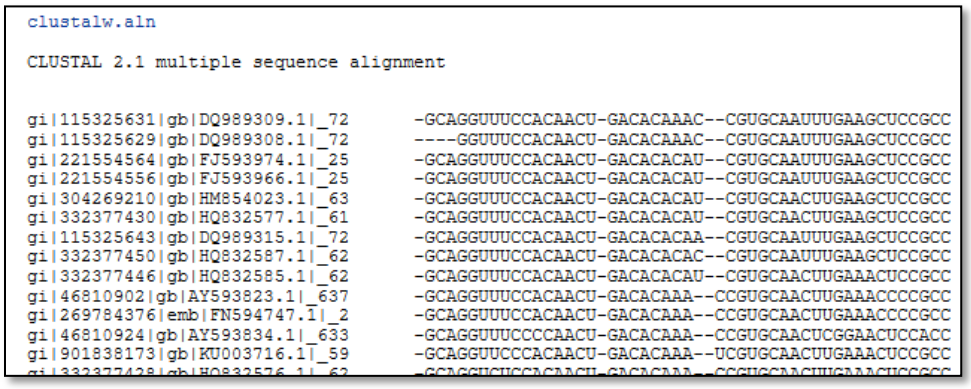
1. Introduction

Replication process in viruses don't depend only on coded regions, but also untranslated regions like Internal ribosome entry site (IRES) which play an important role in replication process [1,2,3]; because it acts as the key of this process. IRES region is a long RNA region exceed ~450 nucleotides and folded in five or four domains depending on its family. IRES in picornavirus family folded in 5 domains, the most important is the third one because its include the function domain of IRES at all, which called Apical region [1,3,4]. Apical region in foot and mouth disease (FMDV) is about 116 nucleotides and contains some conserved motifs which affect directly on its structure [5]. This paper interested in the changes or mutations which happened in Apical region to help biologists in their studies on these mutations which may causes a new serotypes of this virus. We observed the changes on created data set consisting of 282 records; represent 282 different Apical sequences belonging to 4 serotypes which are collected from NCBI portal [6]. The importance of this paper came from that; a single nucleotide substitution in the internal ribosome entry site of FMDV sometimes cause changes in structure at all and leads enhanced cap-independent translation in vivo [12], so some researches focus on this changes only, to study what happen when nucleotide change to another like C to U[5].

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

2. Methodology

The proposed pipe line pass over several steps or phases to produce our observations. **First:** create the data set by enter the Apical reference sequence: Foot-and-mouth disease virus (FMDV) strain C, isolate c-s8c1 as an input to NCBI portal and use BLAST tool to found other similar Apical regions in the different serotypes, the result of this step is: data set consisting of 282 Apical sequences belonging to 4 different serotypes, see for example "*75Apical_A.txt*" file at supplementary material. **Second**: Make multiple sequence alignments of Apical sequences of each serotype in individual using CLUSTALW program [7], the output of this step is a file contains the result of alignment process, see "*clustalw_Apical_A_Sero.aln*" file at supplementary material and figure 1. **Third**: Detect nucleotides frequency at the poll of collected sequences using "web logo" program [8]; by using output file gained from CLUSTALW as an input to this process, because "web logo" needs a sequences with the same length to make its task, see the frequency of the nucleotides of " serotype A" at figure 3, where the size of the nucleotide at its position represents its frequency and appearance in the sequences at all. **Finally**: Observation phase: type each mutated nucleotide in the sequence in a table each one in individual row, and type the mutation occurred to this nucleotide in new column in the same row, see table 1 which explain that if the nucleotide have two probability its presented in two columns; First frequency column which contains the nucleotide appear with the greater value or frequency in this position, second frequency column which contains the nucleotide appear with the later value or frequency in the same position, and so on, for example in table 1, column 1 represent the nucleotide position in the sequence, on positions 22 have 3 different nucleotides A,G,U ordered by their frequency. Then we visualize the effectiveness of each change on the folded shape of the sequence using prediction and visualization programs and tools like MFOLD [9]. The flow chart of the work activity is shown at figure 2.



```
clustalw.aln

CLUSTAL 2.1 multiple sequence alignment


gi|115325631|gb|DQ989309.1|_72      -GCAGGUUUCCACAACU-GACACAAAC--CGUGCAAUUUGAAGCUCCGCC
gi|115325629|gb|DQ989308.1|_72      ----GGUUUCCACAACU-GACACAAAC--CGUGCAAUUUGAAGCUCCGCC
gi|221554564|gb|FJ593974.1|_25      -GCAGGUUUCCACAACU-GACACACAU--CGUGCAAUUUGAAGCUCCGCC
gi|221554556|gb|FJ593966.1|_25      -GCAGGUUUCCACAACU-GACACACAU--CGUGCAAUUUGAAGCUCCGCC
gi|304269210|gb|HM854023.1|_63      -GCAGGUUUCCACAACU-GACACACAU--CGUGCAACUUGAAGCUCCGCC
gi|332377430|gb|HQ832577.1|_61      -GCAGGUUUCCACAACU-GACACACAU--CGUGCAACUUGAAGCUCCGCC
gi|115325643|gb|DQ989315.1|_72      -GCAGGUUUCCACAACU-GACACACAA--CGUGCAAUUUGAAGCUCCGCC
gi|332377450|gb|HQ832587.1|_62      -GCAGGUUUCCACAACU-GACACACAC--CGUGCAAUUUGAAGCUCCGCC
gi|332377446|gb|HQ832585.1|_62      -GCAGGUUUCCACAACU-GACACACAU--CGUGCAACUUGAAACUCCGCC
gi|46810902|gb|AY593823.1|_637      -GCAGGUUUCCACAACU-GACACAAA--CCGUGCAACUUGAAACCCCGCC
gi|269784376|emb|FN594747.1|_2      -GCAGGUUUCCACAACU-GACACAAA--CCGUGCAACUUGAAACCCCGCC
gi|46810924|gb|AY593834.1|_633      -GCAGGUUUCCCCAACU-GACACAAA--CCGUGCAACUCGGAACUCCACC
gi|901838173|gb|KU003716.1|_59      -GCAGGUUCCCACAACU-GACACAAA--UCGUGCAACUUGAAACUCCGCC
gi|332377428|gb|HQ832576.1|_62      -GCAGGUCUCCACAACU-GACACAAA--CCGUGCAACUUGAAACUCCGCC
```

Figure 1: Result of multiple sequence alignments using "CLUSTALW" program.

3. Apicals of individuals serotypes:

This section is the core of the proposed work where we conducted all process on each serotype of collected data set individually; in the later subsections we study the 4 serotypes A, O, C and Asia1 by taking the great frequency nucleotides that appears in web logo, for example in figure 3 we have in position 5 two nucleotides with different frequency C and U, first we consider C as the nucleotide of the sequence in position 5 (section 3.1.1) , and in the next analysis we consider U nucleotide as the next frequency one (section 3.1.2) , sometime more than two nucleotides appear in the same position like position 22 we have A,G,U so we will consider A at first iteration, G at second iteration and U at the third iteration. In the following subsections we shown the mutated nucleotides at each serotype.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

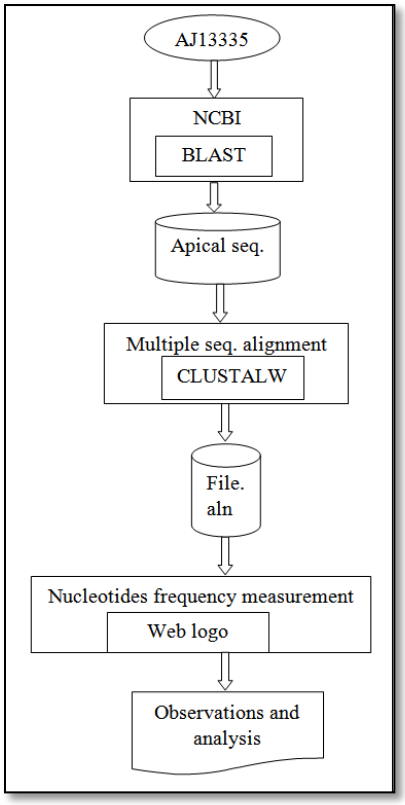


Figure 2: activity flow chart.

3.1. Apical of serotypes A:

The similarity process conducted on BLAST tool produced 75 Apical sequences belonging to sero type A of FMDV, we entered these sequences in multiple sequence alignment tool and then made frequency measurement for the nucleotides and then observe as we explain in the flow chat; figure 2.

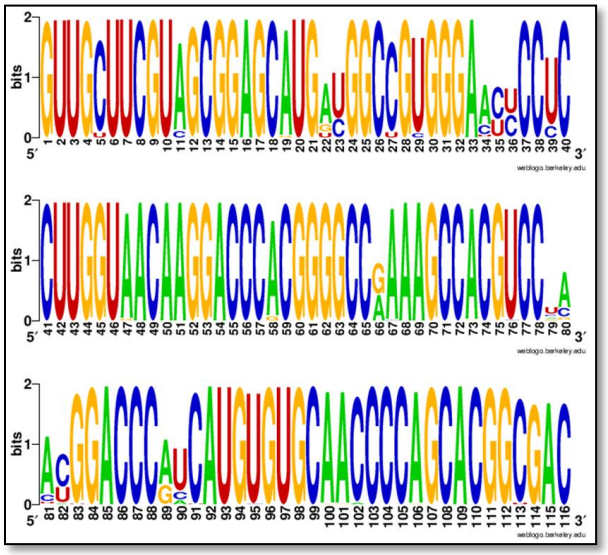


Figure 3: weblogo result: detect the frequency of each nucleotide in apical sequences of serotype A.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

3.1.1. Great probability sequence of serotype A

In this subsection we shown the mutated nucleotides in the sequence; we remark them by bold and underline as fowling sequence, also we type all sequence in table and identify the mutated nucleotide by oval shape abounded it. Figure 4 illustrated the folded shape of the Apical sequence at the best energy value and in the worst energy value, when we consider the great frequency value as the mutated nucleotide, and figure 5 when we consider the second frequency nucleotide.

GUUG**C**UUCGU**AG**CGGAGC**A**UGA**U**GGC**C**G**U**GGGA**ACU**CC**U**CCUUGGU**A**ACAAGGAC
CC**A**CGGGGCC**G**AAAGCCACG**U**CC**UAACG**GACCC**AUC**AUGUGUGCAA**C**CCCAGCACG
G**C**GAC

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
G	U	U	G	**C**	U	U	C	G	U	**A**	**G**	C	G	G	A	G	C	**A**	U
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
G	**A**	**U**	G	G	C	**C**	G	**U**	G	G	G	A	**A**	**C**	**U**	C	C	**U**	C
41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
C	U	U	G	G	U	**A**	A	C	A	A	G	G	A	C	C	C	**A**	C	G
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
G	G	G	C	C	**G**	A	A	A	G	C	C	A	C	G	**U**	C	C	**U**	**A**
81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
A	**C**	**G**	G	A	C	C	C	**A**	**U**	**C**	A	U	G	U	G	U	G	C	A
101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116				
A	**C**	C	C	C	A	G	C	A	C	G	G	**C**	G	A	C				

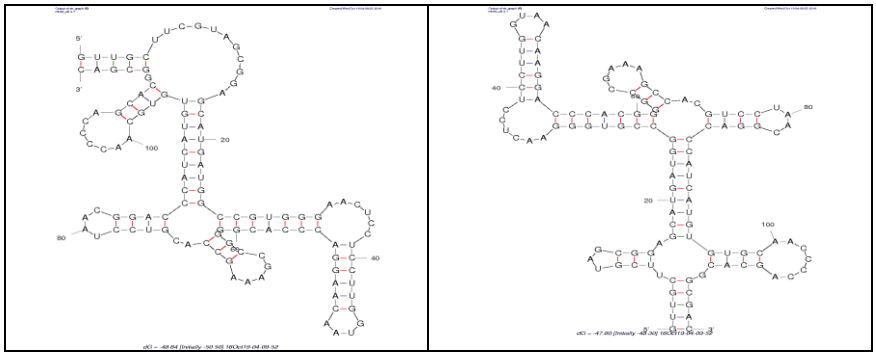


Figure 4 : the second structure for the Apical of serotypeA with the great frequency; the left one in the best energy structure and the right one is the worst energy structure

3.1.2. Second probability sequence of serotype A

GUUG**U**UUCGU**CA**CGGAGC**G**UG**GC**GGC**U**G**C**GGGA**CUC**CC**C**CCUUGGU**G**ACAAGGAC
CC**G**CGGGGCC**A**AAAGCCACG**C**CC**CCCUU**GACCC**GCU**AUGUGUGCAA**A**CCCAGCACG
G**U**GAC

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
G	U	U	G	**U**	U	U	C	G	U	**C**	**A**	C	G	G	A	G	C	**G**	U
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
G	**G**	**C**	G	G	C	**U**	G	**C**	G	G	G	A	**C**	**U**	**C**	C	C	**C**	C
41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
C	U	U	G	G	U	**G**	A	C	A	A	G	G	A	C	C	C	**G**	C	G
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
G	G	G	C	C	**A**	A	A	A	G	C	C	A	C	G	**C**	C	C	**C**	**C**
81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
C	**U**	**U**	G	A	C	C	C	**G**	**C**	**U**	A	U	G	U	G	U	G	C	A
101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116				
A	**A**	C	C	C	A	G	C	A	C	G	G	**U**	G	A	C				

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

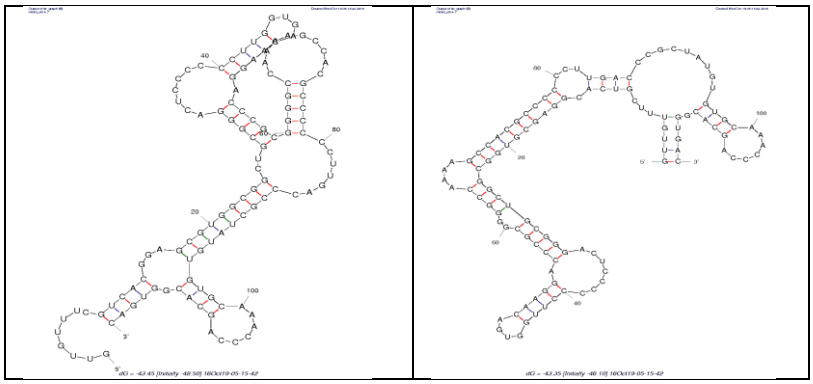


Figure 5 : The second structure for the Apical of serotype A with the second frequency; the left one in the best energy structure and the right one is the worst energy structure

3.1.3. Serotype A: observation and analysis

As shown in table 1 we count the mutated nucleotides as follow:

- Number of stable nucleotides = 90 nucleotides (~ 78 %).
- Number of nucleotides with dual probability = 26 nucleotides (~22%).
- Number of nucleotides with three probability = 7 nucleotides (~.06%).
- Number of nucleotides with four probability (very mutated) = only one nucleotide.

Table 1: All mutated nucleotides in serotype "A"

#	First frequency	Second frequency	Third frequency	Fourth frequency
5	C	U		
11	A	C		
12	G	A		
19	A	G		
22	A	G	U	
23	U	C		
27	C	U		
29	U	C		
34	A	C	U	
35	C	U	-	
36	U	C	-	
39	U	C	-	
47	A	G		
58	A	G		
66	G	A		
76	U	C		
79	U	C	A	G
80	A	C	G	
81	A	C	U	
82	C	U	A	
83	G	U		
89	A	G		
90	U	C	A	
91	C	U		
102	C	A		
113	C	U		

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

3.2. Apical of serotypes O:

The similarity process conducted on BLAST produced 165 Apical sequences belonging to sero type O of FMDV, we entered these sequences in multiple sequence alignment tool and then made frequency measurement for the nucleotides and then observe as we explain in the flow chat.

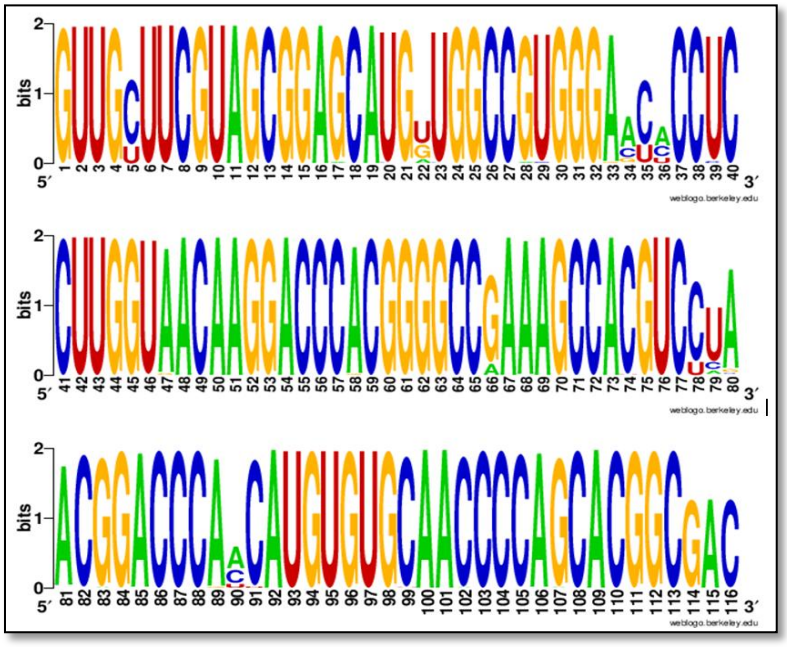


Figure 6: weblogo result: detect the frequency of each nucleotide in apical sequences of serotype O.

Table 2: All mutated nucleotides in serotype "O"

#	First frequency	Second frequency	Third frequency	Fourth frequency
5	C	U		
17	G	A		
22	U	G	A	
28	G	A		
29	U	C		
33	A	G		
34	A	C	G	
35	C	U	-	
36	A	C	U	
39	U	C	-	
47	A	G		
58	A	G		
66	G	A		
78	C	U		
79	U	C	A	-
80	A	G	C	
83	G	A		
89	A	G		
90	A	C	U	
91	C	U		
102	C	A		
107	G	A		

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

3.2.1. Great probability sequence of serotype O

GUUG**C**UUCGUAGCGGA**G**CAUG**UU**GGCC**GU**GGG**AACA**CC**U**CCUUGGU**A**ACAAGGAC
CC**A**CGGGGCC**G**AAAGCCACGUC**CUA**A**CG**GACCC**AAC**AUGUGUGCAA**C**CCCA**G**CAC
GGCGAC

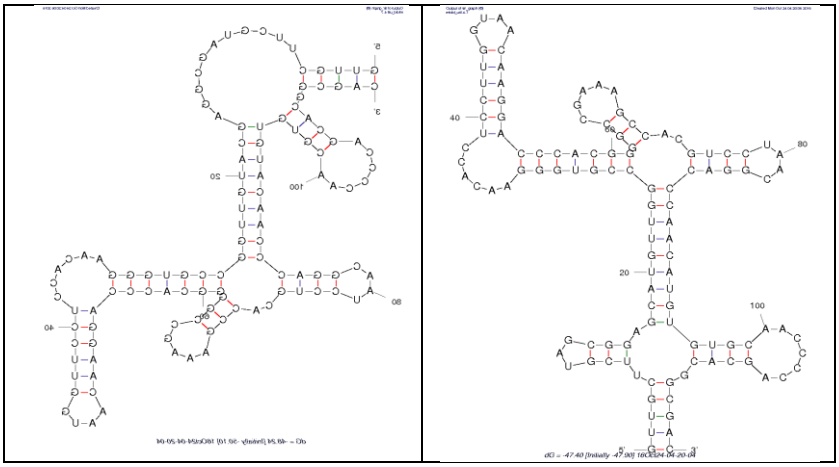


Figure 7 : the second structure for the Apical of serotype O with the great frequency; the left one in the best energy structure and the right one is the worst energy structure.

3.2.2. second probability sequence of serotype O

GUUG**U**UUCGUAGCGGA**A**CAUG**G**UGGCC**AC**GGG**GCUC**CC**C**CCUUGGU**G**ACAAGGAC
CC**G**CGGGGCC**A**AAAGCCACGUC**UCG**A**CA**GACCC**GCU**AUGUGUGCAA**A**CCCA**A**CAC
GGCGAC

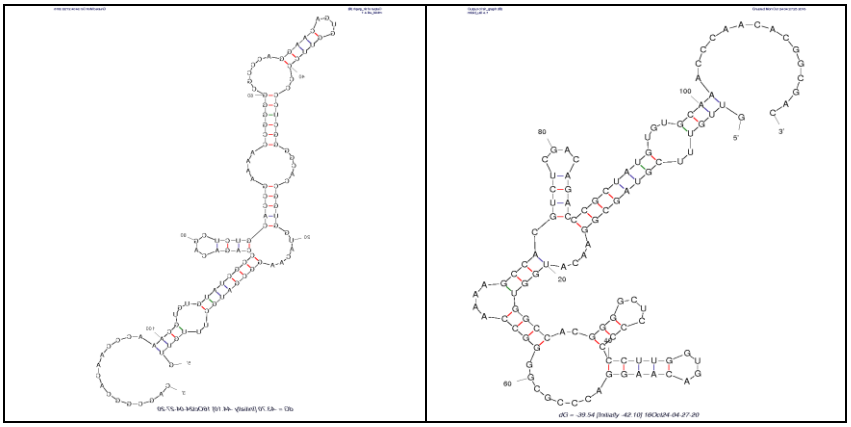


Figure 8 : the second structure for the Apical of serotype O with the second frequency; the left one in the best energy structure and the right one is the worst energy structure

3.2.3. Serotype O: observation and analysis

- Number of stable nucleotides = 94 nucleotides (~ 81 %).
- Number of nucleotides with dual probability = 22 nucleotides (~19%).
- Number of nucleotides with three probability = 6 nucleotides (~.05%).
- Number of nucleotides with four probability (very mutated) = zero nucleotide.

3.3. Apical of serotypes C:

The similarity process conducted on BLAST tool produced 22 Apical sequences belonging to sero type C of FMDV, we entered these sequences in multiple sequence alignment tool and then made frequency measurement for the nucleotides and then observe as we explain in the flow chat.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

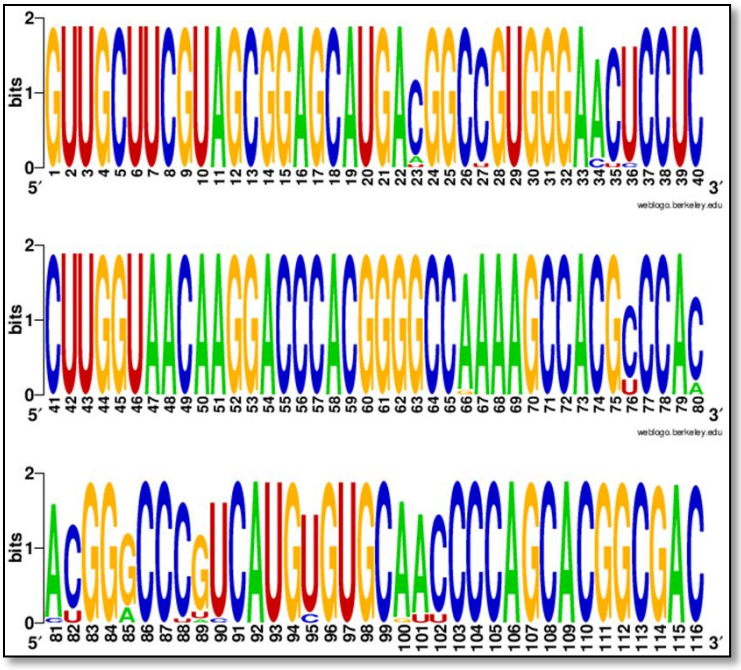


Figure 9: weblogo result: detect the frequency of each nucleotide in apical sequences of serotype C.

3.3.1. Great probability sequence of serotype C

GUUGCUUCGUAGCGGAGCAUGA**C**GGC**C**GUGGGA**ACU**CCUCCUUGGUAACAAGGAC
CCACGGGGCC**A**AAAG**AC**GG**G**CC**CGU**CAU**GU**GUGC**AAAC**AUGUGUGCA**AC**CCCAGC
ACGGCGAC

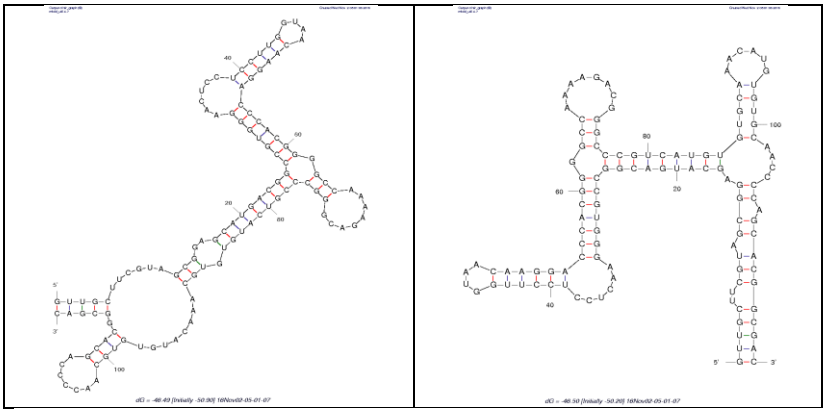


Figure 10 : the second structure for the Apical of serotype C with the great frequency; the left one in the best energy structure and the right one is the worst energy structure.

3.3.2. Second probability sequence of serotype C

GUUGCUUCGUAGCGGAGCAUGA**U**GGC**U**GUGGGA**CUC**CCUCCUUGGUAACAAGGAC
CCACGGGGCC**G**AAAG**CU**GG**A**CC**UUC**CA**U**G**C**GUGC**GGCU**AUGUGUGCA**UU**CCCAGC
ACGGCGAC

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

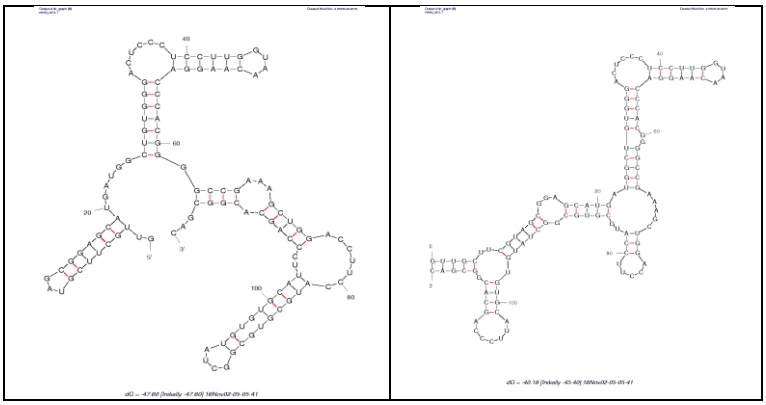


Figure 11 : the second structure for the Apical of serotype C with the second frequency; the left one in the best energy structure and the right one is the worst energy structure

Table 3: All mutated nucleotides in serotype "C"

#	First frequency	Second frequency	Third frequency	Fourth frequency
23	C	A	U	
27	C	U		
34	A	C		
35	C	U	-	
36	U	C		
66	A	G		
78	C	U		
79	U	C	A	-
80	A	G	C	
81	A	C		
82	C	U		
85	G	A		
88	C	U		
89	G	U		
90	U	C		
95	U	C		
100	A	G		
101	A	U		
102	C	U		

3.3.3. Serotype C: observation and analysis

- Number of stable nucleotides = 97 nucleotides (~ 84 %).
- Number of nucleotides with dual probability = 19 nucleotides (~16%).
- Number of nucleotides with three probability = 3 nucleotides (~.03%).
- Number of nucleotides with four probability (very mutated) = zero nucleotide.

3.4. Apical of serotypes Asia1:

The similarity process conducted on BLAST tool produced 20 Apical sequences belonging to sero type Asia1 of FMDV, we entered these sequences in multiple sequence alignment tool and then made frequency measurement for the nucleotides and then observe as we explain in the flow chat.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

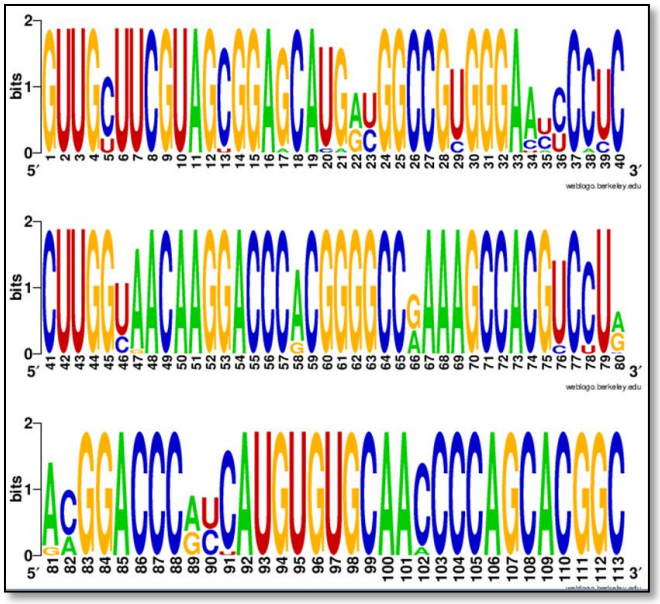


Figure 12: weblogo result: detect the nucleotides frequency of apical sequences of serotype Asia 1.

3.4.1. Great probability sequence of serotype Asia1

GUUG**C**UUCGAG**C**GGA**G**CA**UGAU**GGCCG**U**GGGA**AUC**C**CU**CCUUGG**UA**ACAAGGACC
C**A**CGGGGCC**G**AAAGCCACG**U**C**U**C**U**A**AC**GGACCC**AUC**AUGUGUGCAA**C**CCCAGCACGG
C

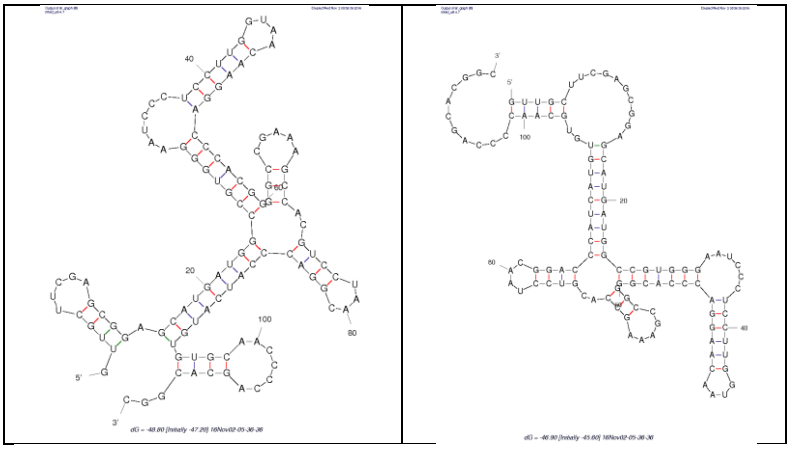


Figure 13 : the second structure for the Apical of serotype Asia1 with the great frequency; the left one in the best energy structure and the right one is the worst energy structure.

3.4.2. Second probability sequence of serotype Asia1

GUUG**U**UUCGAG**U**GGA**A**CA**CAGC**GGCCG**C**GGGA**CCU**C**AC**CCUUGG**CG**ACAAGGACC
C**G**CGGGGCC**A**AAAGCCACG**C**C**UU**GGA**GGACCC**GCU**AUGUGUGCAA**A**CCCAGCACG
GC

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

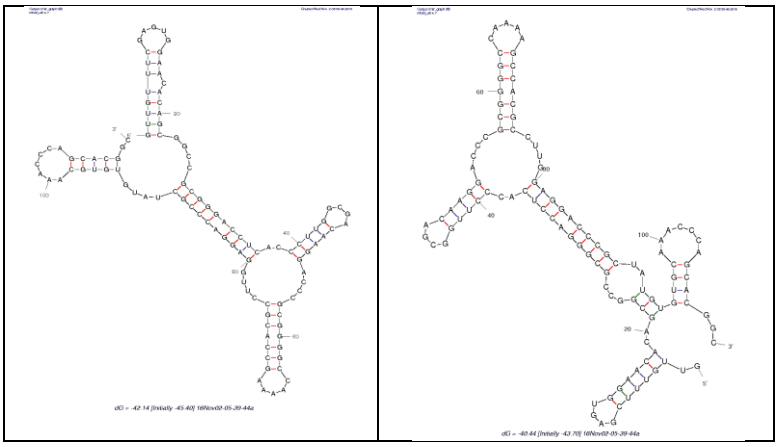


Figure 14 : the second structure for the Apical of serotype Asia1 with the second frequency; the left one in the best energy structure and the right one is the worst energy structure

Table 4: All mutated nucleotides in serotype "Asia1"

#	First frequency	Second frequency	Third frequency	Fourth frequency
5	C	U		
13	C	U		
17	G	A		
20	U	C		
21	G	A		
22	A	G		
23	U	C		
29	U	C		
34	A	C	U	
35	U	C	A	
36	C	U		
38	C	A		
46	U	C		
47	A	G		
58	A	G		
66	G	A		
76	U	C		
78	C	U		
80	A	G	C	
81	A	C		
82	C	A		
89	A	G		
90	U	C		
91	C	U		
102	C	A		

3.4.3. Serotype Asia1: observation and analysis

- Number of stable nucleotides = 88 nucleotides (~ 78 %).
- Number of nucleotides with dual probability = 25 nucleotides (~22%).
- Number of nucleotides with three probability = 3 nucleotides (~.03%).
- Number of nucleotides with four probability (very mutated) = zero nucleotide.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

4. Probability of nucleotide mutation in all serotypes

In this section we will explain what nucleotide mutated in what serotype? To know what is the most mutated positions in the Apical region at all; to help biologist to focus their researches in this positions. As shown in table 5; we have 8 nucleotides in the sequence are mutated in all four serotypes (A,O,C and Asia1); those are (34,35,36,66,80,89,90,102), in the fowling subsection we will type our observation on these nucleotides and their effectiveness in the folded shape or second structure of Apical sequence.

Table 5: mutated nucleotides in all serotypes

#	Serotype A	Serotype O	Serotype C	Serotype Asia1
5	√	√		√
11	√			
12	√			
13				√
17		√		√
19	√			
20				√
21				√
22	√	√		√
23	√		√	√
27	√		√	
28		√		
29	√	√		√
33		√		
34	√	√	√	√
35	√	√	√	√
36	√	√	√	√
38				√
39	√	√		
46				√
47	√	√		√
58	√	√		√
66	√	√	√	√
76	√			√
78		√	√	√
79	√	√	√	
80	√	√	√	√
81	√		√	√
82	√		√	√
83	√	√		
85			√	
88			√	
89	√	√	√	√
90	√	√	√	√
91	√	√		√
95			√	
100			√	
101			√	
102	√	√	√	√
107		√		
113	√			

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

4.1. Effectiveness of most mutated nucleotides in the domain structure

Here we discuss the location of the eight most mutated nucleotides (34, 35, 36, 66, 80, 89, 90, 102) in the structure in all serotypes, to know the effectiveness of these mutation in folded shape.

Table 6: location of mutated nucleotides in all serotypes

Nucleotide number	Serotype A	Serotype O	Serotype C	Serotype Asia1
34	Loop	loop	loop	loop
35	Loop	loop	loop	loop
36	Loop	loop	loop	loop
66	Loop	loop	loop	loop
80	Loop	loop	loop	loop
89	Stem	stem	stem	stem
90	Stem	stem	stem	stem
102	Loop	loop	loop	loop

5. Summary and discussion

In the sequences with great or second frequency for the mutated sequences; no change occurs in the second structure of Apical; because the mutated nucleotides are mutated to another nucleotides can pair with the corresponding one according to watson and crick model (A=U, C=G) [10], and wobble stack (G=U) [11] like nucleotide at position number 27 was C and changed to U, but this change don't affects at the structure because nucleotide number 27 is paired with nucleotide number 60 at the folded structure which is G and G can paired with C or U; so the structure still unchanged. Or the mutated nucleotide is located at loop in the folded shape; like nucleotide number 11 was A mutated to C; no changes is happened because nucleotide 11 located at loop.

For the same previous reason the structure of the sequences of the mutated sequence with the sometimes changed, for example nucleotide number 5 is changed from C to U but the paired one not the nucleotide 112 as the reference structure because MFold folded program use dynamic programming algorithm in folding process.

Table 6 explains the location of the most mutated nucleotides at Apical sequence to help biologist in studying the changes in wet lab.

References

[1] Segun Jung and Tamar Schlick, "Candidate RNA structures for domain 3 of the foot-and-mouth-disease virus internal ribosome entry site", Nucleic Acids Research, 2013, Vol. 41, No. 3 1483–1495,doi:10.1093/nar/gks1302.

[2] S Lopez de Quinto, Elafuente and E Martenz-Salas, et al., "IRES interaction with translation initiation factors: Functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII", RNA (2001) 7:1213-1226, DOI: 10+1017+S1355838201010433.

[3] Amr Badr, Ahmed ElSadek and Alaa.Yassin: "Computational Based Analysis for Internal Ribosome Entry Site (IRES) and Viral Replication in FMDV", International Journal of Computer Science Issues (IJCSI), Volume 12, Issue 4, July 2015.

[4] OLGA FERNANDEZ-MIRAGALL and ENCARNACION MARTINEZ-SALAS, "Structural organization of viral IRES depends on the integrity of the GNRA motif", rna journal, doi:10.1261/rna.5950603, RNA (2003) 9:1333–1344.

[5] Du,Z., Ulyanov,N.B., Yu,J., Andino,R. and James,T.L. (2004) " NMR structures of loop B RNAs from the stem-loop IV domain of the enterovirus internal ribosome entry site: a single C to U substitution drastically changes the shape and flexibility of RNA" . Biochemistry, 43, 5757–5771.

[6] https://www.ncbi.nlm.nih.gov/

[7] http://www.genome.jp/tools/clustalw/

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

[8] http://weblogo.berkeley.edu/logo.cgi
[9] http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form
[10] Watson JD, Crick FH ,*"Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid*" ,April 1953,Nature 171 (4356): 737–738.
[11] Mondal M, Mukherjee S, Halder S, Bhattacharyya D, "*Stacking geometry for non-canonical G:U wobble base pair containing dinucleotide sequences in RNA: dispersion-corrected DFT-D study*", Biopolymers. 2015 Jun;103(6):328-38. doi: 10.1002/bip.22616.
[12] Martínez-Salas E1, Sáiz JC, Dávila M, Belsham GJ, Domingo E," A single nucleotide substitution in the internal ribosome entry site of foot-and-mouth disease virus leads to enhanced cap-independent translation in vivo", J Virol. 1993 Jul;67(7):3748-55.