Computational Based Analysis for Internal Ribosome Entry Site (IRES) and Viral Replication in FMDV

Amr Badr1, Ahmed Farouk Al-Sadek2, Alaa Eldin Abdallah Yassin3

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

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

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

Abstract

In the recent years, the big challenge which faces the world is the lack of vaccines and drugs for viruses which appeared widely with different families and serotypes, especially RNA viruses. The high mutation and instability nature of RNA viruses prevent the long life affect of the vaccines, for this reason a lot of researches in RNA viruses structure and function have been conducted to overcome this challenge. The deep understanding of the virus in its life cycle, genome sequencing, second and tertiary structure of the virus affect positively in solving the addressed problem. This paper focuses on foot-and-mouth disease virus (FMDV) which is considered one of the most popular viruses in picornafamily, that is belonging to RNA viruses' class, To be more accurate the paper concerned in IRES region only, which is crucial to exploring and deciphering the initiation mechanism of translation and replication of the virus. It is notable that: the most researches in this area have been conducted in vitro, and few researches in computational. This paper introduce a survey for the computational efforts in IRES region of FMDV and suggests some novel research points which can be investigate in future.

Keywords: *IRES, RNA virus, viral replication, FMDV, IRES database, RNA prediction.*

1. Introduction

Its well known that viruses are contagious agents and cannot replicate without the presence of the host cell, and also have evolved mechanisms to achieve the replication process in an efficient way [1]. The host cell is differ from virus class to another, where viruses can be classified into many types, each of which has its own families of viruses, which in turn have differing replication strategies themselves . Also viruses classified to DNA viruses and RNA viruses, and its notable that , RNA mutation level is higher than DNA mutation level, and DNA viruses are stable while RNA viruses are unstable, so the low rate of mutation in DNA explain why vaccines of DNA viruses work effectively throughout the years and RNA vaccines not like that, and also explain the importance of research in RNA viruses structure and function[2]. RNA viruses can folded in many junctions and protein synthesis of an RNA template started by two different known mechanisms: cap-dependent translation initiation and cap-independent translation initiation, The latter called internal ribosome entry sites (IRESs) that are found in both viral RNAs and cellular mRNAs [3]. This paper produced a review to the research in a specific part of RNA virus, which the replication of the virus start from it, it is the IRES region, focusing this survey on the most popular viruses in picornafamily, which is belonging to RNA viruses' class, is FMDV, where picornaviridae are a large family of human and animal RNA viruses[31], and FMDV is an acute, infection, systemic disease of domestic and wild cloven-hoofed animals including cattle, pigs, sheep, goats, buffalo and various wildlife species and this dangerous dieses considered one of the fastest growing of all animal viruses and one of the most contagious animal diseases, with a large economic impact [5, 6, 7, 8]. We are excited and have a great motivation to accomplish this paper because FMDV utilizes non-canonical translation initiation for viral protein synthesis, by forming a specific RNA structure (IRES)[9, 10, 11] and IRES elements direct cap-independent internal initiation of protein synthesis within mammalian cells[9], so the good understanding of viral RNA sequence conservation and variability and genetic diversity in nature will likely impact our understanding of FMDV infections, host range, and transmission [12].This paper organized the related works of IRES in FMDV in the computational side, explained the different research perspectives and results of each research in this point, showing the consists of IRES region, important domains in this region, IRES structure and how the structure affects the function, popular databases, tools and programs in domain elements that the researcher can use in this point of research.

2. IRES importance and crucial function

May this paper need to answer an important dual question before the reader decide to go ahead of complete reading all this survey paper or not, what is IRES? , and why this region especially in RNA the researchers gave it extra attention? Yes it's good and logic question, and the answer will orient the importance of this review paper. Before the answer we want to memorized that the initiation of translation is the key step of protein synthesis, and this operation is performed through two basic mechanisms govern translation initiation in eukaryotic mRNAs, the cap-dependent initiation mechanism that operates in the great majority of mRNAs and the internal ribosome entry site (IRES)-dependent mechanism[4]. From previous fact, we can shortly answer the question above as , because IRES region is crucial to exploring and deciphering the initiation mechanism of translation and the replication of the virus, where IRES contains some crucial domains which necessary to complete replication process [9,13], and also contains some conserved motifs which are essential for IRES activity to start its functions in initiation and replication [14,15]. Added to the previous answer , the different mechanisms used by IRESs are reflected in their structural diversity, which motivate some researchers to develop a model linking IRES function to structure and the understanding of the 3D structural aspects of RNA junctions in IRES's domains are essential to decode the mechanism of IRES-driven translation, where the global conformation of the IRES second structure is ineligible to accomplish this task [9]. Biological researches concluded that the viral IRES is known and characterized by its ability to drive cap-independent translation, but the origin of the IRES and the context in which it functions can vary dramatically due to the relation between function and structure[3,16], and thus although the secondary structure of IRES has been predicted, little information on the tertiary structure are available[17] , and this lead us to say that : the knowing of picornavirus IRES-interacting proteins is essential to drive internal initiation of translation, and consequently make a great progress [1]. Also the creation of specialist database for IRES to store the novel IRES serotypes, where the list of unknown IRESes is certainly still very large, can be considered a good indicator to the importance of IRES and explain the reason of researcher's interest [18] and because IRES found in genetically distant mRNAs seem to be organized in different RNA structures, the definition of the structural requirements for IRES activity is challenging [19], so IRES progress and other RNA research like RNA tertiary motifs discovery and applications of graph theory approaches to RNA structure and function, can be combined to impact many problems in RNA structure and function and achieve a progress in our life at all [20] . RNA is the bimolecular cousin to DNA and protein[21], and the suggestion of "Anything that DAN can do, RNA can do better ", and the behaviour of RNA that mimic DANA[21] can explain the importance of RNA in biology dogma, however though recently there are many 3D prediction programs but computational programs cannot predict multiple RNA junction structures well[22].

3. IRES composition in FMDV

Internal ribosome entry site (IRES) elements were first identified about 25 years ago within the 5′ untranslated region of picornavirus RNAs [23]. Viral IRESs are classified into four major structural groups, which vary in size (from 460 or 450 – to 270 nt). IRES region have different and complex secondary structures and distinct requirements for cellular proteins to allow them to function, each group from those four uses a different mechanism for initiation. Although sequence variation between those IRESs can reach 50%, a similar overall structure is maintained by compensatory base changes in helical elements [24, 25].

The laboratory tests discovered that, the composition of IRES region differs from virus family to another, for an example the hepatitis C virus which is belonging to Flavivirus family, internal ribosome entry site (IRES) element contains four domains, but in our reviewed family, picornavirus as FMDV, IRES consists of ~ 450 nt and can fold in multiple stem-loops organized in five domains as presented in figure 1. IRES domains differ in size and functions, where some domains are bigger than others; also recent biochemical data have suggested that some domains produce some activities more important than others, because IRES domains contain loops, junctions, hosted proteins to directed translation as (eIFs) and (ITAFs) factors, and also contain motifs like GNRA and RAAA motif [9], where these motifs are critical for IRES function[10, 26] , and some of those motifs are responsible for factors interaction, to enable and help the initiation factors to do their activities with IRES[13], so the difference of size and contents of each domain in IRES reflect the difference in their functions.

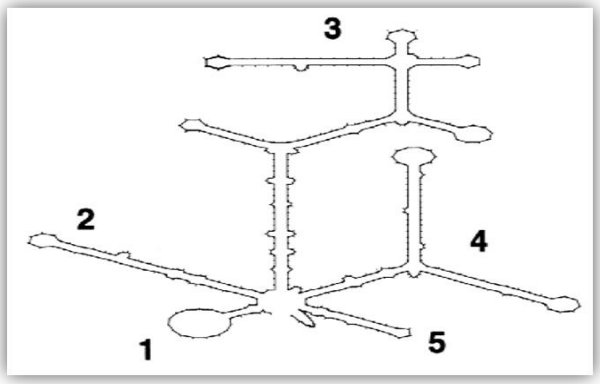

Fig. 1 Simple figure to the five domains of IRES region in FMDV, explain the differ in size and shape between domains [13,25].

4. IRES importance domains

IRES in FMDV is a relatively long RNA region, exceed ~450 nucleotides that are predicted to fold in five domains (named 1 to 5). The central domain in IRES is domain 3 in picornavirus [14], and domain 2 in Flavivirus. Domain 3 in IRES consists of 2 regions, apical region and basal region, but the two regions not in the same level of importance. Apical region is the functional region which includes the conserved motifs of IRES region. As shown previously, that domain 3 of the FMDV IRES is essential to establish functional interactions with eIF4GI [13, 27], also domain 5 is important because it holds the preferential binding site for eIF3, although this complex initiation factor can establish multiple contacts with the IRES structure [13].

Steady RNA–RNA interactions between separated domains (1–2, 3, 4–5, or HH) of the IRES structure in FMDV are performed and all the domains were able to interact with the full-length IRES as well as with domain 3 , also formation of domain 3 homodimer competed with formation of hetero complexes with other domains and under a specific degree of temperature complexes between domain 3 and 4-5 reaches to 50% [25]. Figure 2 presents domain 3 in IRES region.

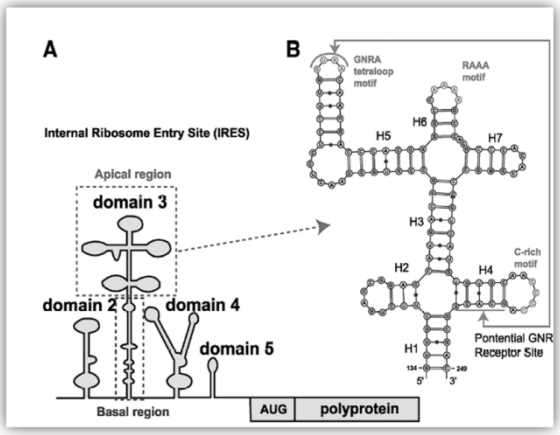

Fig. 2 Domain 3 of IRES region in FMDV [9].

5. IRES conserved motifs and junctions

previous laboratory researches are proved that, eukaryotic translation initiation factors (eIF) play an essential role in picornavirus IRES-dependent initiation, and the different factors which involved in IRES operation are related in function with some motifs, which necessary in RNA folding operation[13]. Specifically, the apical region of domain 3 includes two conserved motifs, GNRA and RAAA motifs, where GNRA motif is common in folded RNA. Conserved motifs may facilitate RNA–RNA or RNA–protein interactions required to maintain the tertiary structure needed for proper recognition of the IRES element by the translational machinery [25]. RNA junctions are "secondary-structure elements formed when three or more helices come together" [28], and RNA junctions provide a hub for different double-stranded helical arms to come together [29]. RNA junctions are placed in different family types, because nucleotide length of RNA in a single strand and junctions have identified structural patterns in coaxial stacking and that define the different family types [9].

6. Different studies in domain elements

we can categorize the different researches in IRES region into three categories based on the structure level of IRES region, where some researches done on the first structure (1D) of the IRES region but not for IRES only but to RNA at all, and the analysis of the results introduced IRES region as a very important region in RNA, the second category done on the second structure (2D) of IRES, and the last category deal with the third structure (3D) of IRES region, but the research of this category are very rare with respect to the second one, we summarize the previous work of IRES in FMDV in figure 3. and in the following sub-section we will present the different researches done in the three categories.

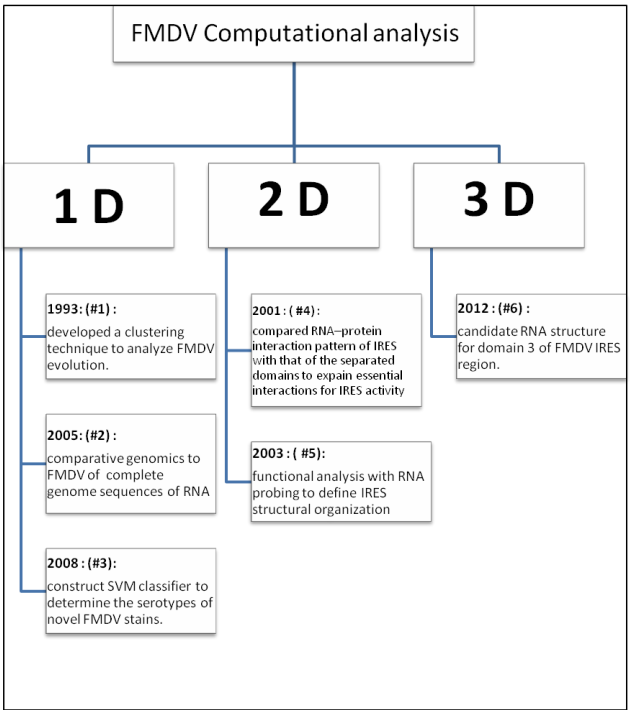

Fig. 3 Tree for the computational research in FMDV.

i. First structure (1D) studies in FMDV

• At 2005 Carrillo and et.al produced a comparative genomics to FMDV of complete genome sequences of RNA followed by cDNAs of 133 isolates of the seven serotypes of FMDV(A, O, C, Asia1, SAT1, SAT2, SAT3), 30 from 133 are not previously fully sequenced [30], the samples are obtained from several databases represents more than 40 countries and the isolated samples from year 1947 to 2001, they aim from this analysis to produce a deep analysis to RNA of FMDV, because the good understand of viral RNA sequence conservation, variability and genetic diversity in nature will likely impact our understanding of FMDV infections, host range, and transmission. This research includes a lot of results of their analysis as 58 % of the amino acids encoded by FMDV isolates are invariant , so the 42 % residues are critical for virus biology , but here we focused on IRES region results only, where they concluded that: FMDV IRES sequences (positions 640 to 1151) showed 70 to 100% nt identity in pair-wise comparisons, 47% invariant nt, and numerous invariant motifs in the predicted secondary structure domains, thus the approximately 500-nt FMDV IRES, which is responsible for cap-independent polyprotein translation is predicted to contain four structural domains (domain 2 to 5) .

Discussion and suggestion of this research: make the comparison not only on all genome of the FMDV, but also focus the research in biology and computational field on the IRES region only, to analyze and note the difference, similarity between the different IRESs of the different FMDV serotypes. And also from this survey we noted that the research in the 1D of IRES in computational side are rare and the efforts concentrated in biology side only.

• Identify the presence of new serotype or recombinants of multiple serotype viruses, which are very common case for FMDV can help in vaccines and understanding the evolution of the virus , where evolution of the virus are important issue from a long time ago. At 1993 JOAQUIN DOPAZO and et.al developed a clustering technique to analyze viral evolution named "Split decomposition" to analyze of FMDV evolution[31]. At 2008 Guohui and et.al [32] produced a computational genotyping method to determine the serotypes of novel FMDV stains, through constructing SVM classifier and applying this classifier with 98.45% accuracy, on 129 FMDV strains with known serotypes as training strains to select the most serotype-specific nucleotide strings using advanced feature selection method to extract this serotype-specific nucleotide strings instated of using whole genomes of the strains. The big advantages of the proposed method that its doesn't involve the high complexity stage of multiple sequence alignments, and thus supports high throughput genotyping and facilitate the disease control process once the genotype is defined, also computational genotyping is much faster and much cheaper than the wet-lab based biological experiments, specially by knowing that the dissimilarity of nucleic acid in the protein coding region among different serotypes of FMDV can be up to 54%[30].

Discussion and suggestion of this research: using one of the optimal algorithms as PSO to extract features from RNA of the FMDV may enhance the prediction of the new serotypes of the virus, especially if we add some additional factors out of the sequence like the temperature and humidity of the country witch the tested serotypes isolated from it. also we suggested that to apply the proposed classifier on different data to test the over-vetting of this classifier.

ii. Second structure (2D) studies in FMDV

Researchers give the 2D structures a big importance and this type of research is the biggest number of research between the other types mentioned in this paper. some studies concern on a special twist of secondary structural elements named pseudoknot, others concern on different loops existed in IRES regions and some studies focused on the important motifs involved in IRES region to demonstrate the critical role of this motifs in IRES activity and function

A. Pseudoknot is RNA structure that is found in almost all classes of RNA and several viral [33,34,35], and pseudoknot can be considered super-secondary structural elements [21, 36], because pseudoknot motifs are produced from a special twist of secondary structural elements. The interplay between the loops and the helical stems determine the pseudoknot stability and folding, and because RNA pseudoknots play an pivotal and indispensable role in the structures and functions of many RNAs as translation process of many viruses like FMDV, pseudoknots are now prevalent motif with diverse functions in various biological processes and directly related to the folding stability and the conformational changes[37] which related to RNA region specially the viral pseudoknots and their role in virus gene expression and genome replication[34] and also the affects of pseudoknot on IRES activity [35].

• At 2006 Song Cao and et al. produced a model to predict pseudoknot folding, and they validated their model through large experimental tests both for the original structures and for the folding thermodynamics. They discuss the big challenge which face the folding prediction for the pseudoknot, which are the energy parameters and entropy parameters, and produced a promising solution for this challenge though computing the free energy landscapes and the base pairing probability at different temperatures. They concluded that the pseudoknot free energy can be computed as the sum of the free energies of the stems, loops and possibly the coaxial-stacking[37]. Some other research provides a computational algorithm using dynamic programming to predict the structure of pseudoknotted RNAs [38], and others developed a model to estimate the loop entropies, then they performed an

analysis for the asymmetry in the pseudoknot structure. [39].

Discussion and suggestion of this research: IRES includes a predicted pseudoknot interaction near the AUG start codon, but the results of previous studies in structure scope have been conflicting [40], so we have a challenge to produce a good prediction to this critical part in IRES region, and the suggestion is to mimic the model developed to predict IRES and produced a good results to predict pseudoknots, where the two regions are special region in RNA.

• One of the important points of research in understanding the structure of IRES region and discovering its function through knowing the critical motifs placed in the central domain of IRES - domain 3 - which play an important role in the stability of IRES in second structure level and affect on the main function of IRES which is initiation translation. At 2003, OLGA and ENCARNACIO´ N, showed the importance of motifs in IRES activity according to organizing the structural of a viral IRES depending on the GNRA motif. They proved their theory through laboratory experiments by using the methodology of substitution the nucleotides in different regions in IRES domain 3, and they remodel IRES 2D structure of domain 3 after the substitution process, and then test the predicted structure in vitro to observe the affects in activity of IRES domain 3 according to the change in its 2D structure, which give us a good indicator to the importance of motifs part in domain 3 IRES region. This research concluded that a UNCG motif does not functionally substitute the GNRA motif and the GNRA motif dictates the organization and stability of domain 3. It is noticeable through this methodology of how difficult the selection of motifs which want to remodel without help from the computational analysis to reduced the infinite structure can be modeled for this region, and thus this research team done through using phylogenetic analysis of the different FMDV serotypes to remodel RNA structure, where RNA structure were aligned and then plot in the RNA secondary structure using the most co variation among the 26 FMDV IRES sequences to fit the results of biochemical probing obtained from the vitro phase[14].

Discussion and suggestion of this research: if the computational analysis done on big data than used in this research may the result be more efficient and more accurate, also we can suggest to apply this methodology on anther motifs of IRES, like pseudoknots and loops.

• Some other researchers have focused on the interaction between IRES region and the different related initiation factors and motifs, and the affect which happened if any mutation occurred on one of those factors which affected significantly in the IRES activity and function. This type of researches are important because during the last years, it has become increasingly clear that different IRESs use distinct strategies to interact with the translational machinery. At 2001 SONIA LÓPEZ and et.al identified novel RNA–protein interactions between the foot-and-mouth disease virus (FMDV) IRES and three translation initiation factors. The research team to study the RNA determinants responsible for the interaction between the FMDV IRES and the components of the translation machinery, they compared the RNA–protein interaction pattern of the IRES with that of the separated domains and they presented the binding relation between different IRES domains and other components and explain those interactions are essential or not essential for IRES activity[13].

Discussion and suggestion of this research: the previous studies in this point of research mostly done on the wet-lab without the help of computer science approaches, so the suggest to use the prediction program to predict the interaction relation between the IRES domains and other components.

iii. Tertiary structure (3D) studies in FMDV, IRES

The only and most specific research in IRES of FMDV at third structural level, was done by Segun Jung and Tamer Schlick at 2012, where the research focus on the most important domain in IRES, Domain 3 . The aim of this great effort was to candidate RNA structure for domain 3 of FMDV IRES region, they were motivated because the computational programs till now can't predict multiple RNA junction structures well [9]. at 2011 Tamer Schlick and et.al produced RNA prediction program named, Junction-Explorer [41],based on data mining to predict and enumerate all possible combination of junction topologies with more than 70% prediction accuracy, when we input the 2D structure of RNA junction to the program. The prediction process for the apical region of domain 3 in IRES passed two main phases using the divide-and-conquer approach:

phase 1 : they begin with IRES 2D which pre-probed before by RNA probing [14], they partitioned RNA into subsystems and then modeled each one with Junction-Explorer program to four-way junctions classification using knowledge-based topology, and at the end of this phase they use MC-Sym program[36] to model the 3D candidates after applying experimental date as a constraint to refine topology candidates.

phase 2 : the objective of this phase to perform molecular dynamics (MD) simulations to investigate structural properties of all 3D candidates of the previous phase, using the most energetically favorable and stable conformational states to propose this structure as an appropriate tertiary structure for the apical region in FMDV IRES domain 3.

7. IRES databases

The continuously growing of list of discovering novel IRES , and the huge amount of still unknown, adding to that the importance of IRES region as paper showed,

explain the need to establishing a special databases for this region [18]. One of the most specialized database to IRES region is "IRESite database" , where IRESite db presents up to 92 biologically relevant aspects of every experiment, as the nature of an IRES element, its functionality/defectivity, origin, size, sequence, structure and other elements(http://www.iresite.org/)[42], IRESite stored >600 records, and the published team developed a tool for the examination of viral and cellular internal ribosome entry sites at 2010 [42], and within this database, the IRESes are classified in several categories based on the function of the gene in which they are found, how the genes are regulated, and with which factors the IRES elements interact [35], a portion of the IRESite statistics available at the web site under the 'record counts'. Also the researchers can collect the needed data of IRES from general databases like GenBanks[43], and RAG[44].

8. Tools and programs in domain elements

Different programs are used in IRES alignments, prediction and simulation, those programs valid to IRES and any RNA sequences, and can be categorize to three categories, sequence alignments programs, second structure programs and third structure programs.

a) Sequence alignments programs and Tools
In the past, biologists are looking at databases manually, by printing its contents on a paper, taping the printout to the office wall, writing down their own query sequence on a piece of paper, and spending more hours manually scanning the wall. With the huge amount of sequences now available, this process can't act as before, it's differ, now we can use computers to do this difficult task by some bioinformatics tool and programs. BLAST (Basic Local Alignment Search Tool) the most popular sequence alignment tool [45] which used in a lot of bioinformatics researches in general not only in IRES sequence alignment studies .Clustal program is one of the famous programs used in domain elements which overcome other programs as T-coffee, where T-coffee give accurate results but only with small sets of sequences which leads to high computational costs. Clustal is a general purpose multiple sequence alignment program [46] include graphical alignment facility. EBI Clustal site gets millions of multiple alignments job per year and the big features included in Clustal latest version (ClustalW , ClustalX 2.0) to faster alignment of the big sized data sets.

b) Predicting second structure of RNA program
RNA junctions are important structural elements of three or more helices in the organization of the global structure of RNA molecules, the coaxial stacking of helices is a common motif among junctions. Junction-Explorer can be considered one of the commonly used program in junction prediction for IRES region and for RNA at all. Given an RNA secondary structure, the Junction-Explorer web server can identify and locate the junctions on the RNA secondary structure. Junction-Explorer employs a machine learning algorithm called random forests (RFs) for prediction [47] where in training phase : use various geometric and energetic parameters as 'feature vectors' ,which contain information on :free energies, loop sizes between junctions and adenine content and the prediction phase: search for all possible combinations of the multiple four-way junctions to produce combined structures(2D), add to the ability to remove pseudoknots from the structure before prediction.

c) Building tertiary structure of RNA program
There are many useful 3D prediction programs [22], MC-Fold | MC-Sym pipeline is a sophisticated web-hosted service for RNA secondary and tertiary structure prediction [48, 49]. The first phase of this pipeline MC-Fold responsible for predicts secondary structures, and the second phase of the pipeline MC-Sym builds tertiary structures from MC-Fold's secondary structures, so the pipeline here means that the input sequence to MC-Fold outputs secondary structures that are direct input to MC-Sym, which outputs tertiary structures. the paper surveyed this pipeline because the research effort concluded by Segun Jung and Tamar Schlick [9] use this pipeline, With the fact that this program is a great benefit, but it has some limitations as operating system, where MC-Fold | MC-Sym run with Unix but with Windows no.
Also some researches collect their needed data form big general database like GenBanks, which include data for DNA, RNA, Protein [50], where the researchers establish their own data sets, as Guohui Lin team create FMDV serotypes and strains data set from NCBI GenBank [32].

Acknowledgements
we want to acknowledge Central Lab for Agricultural Experts Systems and the Faculty of Computers and Information, Cairo University for the great support in this research.

References

[1]. Marcelo Lo´pez-Lastra, Pablo Ramdohr, Alejandro Letelier, Maricarmen Vallejos, Jorge Vera-Otarola and Fernando Valiente-Echeverrı´a, "Translation initiation of viral mRNAs" , Rev. Med. Virol. 2010; 20: 177–195, 2010. DOI: 10.1002/rmv.649.

[2]. E. Domingo and J. J. Holland , "RNA VIRUS MUTATIONS AND FITNESS FOR SURVIVAL" , Annual Review of Microbiology, Vol. 51: 151-178, 1997,DOI: 10.1146/annurev.micro.51.1.151.

[3]. Filbin ME, Kieft JS,"Toward a structural understanding of IRES RNA function" , Curr Opin Struct Biol 2009, DOI 10.1016/j.sbi.2009.03.005.

[4]. Kathryn Sandberg, Susan E. Mulroney "RNA Binding Proteins: New Concepts in Gene Regulation", chapter: Sung Key Jang, Eckard Wimmer, "The Role of RNA-Binding Proteins in IRES-Dependent Translation", pp:1-33,2002, ISBN: 978-1-4419-4935-6.

[5]. Institute for Animal Health, Pirbright Laboratory, "Foot-and-mouth disease: scientific problems and recent progress", 1st annual report (2003) prepared for DEFRA, Science Directorate, 5th June.

[6]. G. J. Belsham, "Translation and Replication of FMDV RNA", CTMI (2005) 288:43--70, Springer-Verlag.

[7]. Sonia LoÂpez de Quinto, Margarita SaÂ iz, Diana de la Morena, Francisco Sobrino and EncarnacioÂn MartõÂnez-Salas, "IRES-driven translation is stimulated separately by the FMDV 3¢-NCR and poly(A) sequences",2002.

[8]. Jeffery S. Kieft, Kaihong Zhou, Angie Grech, Ronald Jubin and Jennifer A.Doudna, "Crystal structure of an RNA ternary domain essential to HVC IRES-modified translation initiation", npg: Natural Publishing Group, April 2002, doi: 10.1038/nsb 781.

[9]. 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.

[10]. MORWENNA E.M. ROBERTSON, RACHAEL A. SEAMONS, and GRAHM J.BELSHAM, "A selection system for functional internal ribosome entry site (IRES) elements: Analysis of the requirement for a conserved GNRA tetraloop in the encephalomyocarditis virus IRES", RNA (1999), 5:1167–1179+ Cambridge University Press.

[11]. jeffrey S.Kieft, Kaihong Zhou, Angie Grech, Ronald Jubin and Jennifier A.Doudna, "Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation" , 2002.

[12]. Paul Lawrence, Michael LaRocco, Barry Baxt and Elizabeth Rieder, "Examination of soluble integrin resistant mutants of foot-and-mouth disease virus", V. Journal 2013, 10:2 doi:10.1186/1743-422X-10-2,2013.

[13]. 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.

[14]. 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.

[15]. CARL C.CORRELL and KERREN SWINGER, "Common and distinctive features of GNRA tetraloops based on a GUAA tetraloop structure at 1.4 Å resolution", rna journal, doi:10.1261/rna.2147803, RNA (2003) 9:2003, pp: 355–363.

[16]. P. E. URWIN, E. I. ZUBKO and ATKINSON, "The biotechnological application and limitation of IRES to deliver multiple defence genes to plant pathogens", Physiological and Molecular Plant Pathology (2002) 61, 103-108, doi: 10.1006/pmpp.2002.0421.

[17]. LUCY P. BEALES, DAVID J. ROWLANDS, and ANDREAS HOLZENBURG, "The internal ribosome entry site (IRES)of hepatitis C virus visualized by electron microscopy", RNA 2001 7: 661-670.

[18]. Sophie Bonnal, Christel Boutonnet, Leonel Prado-Lourenc¸o and Ste´phan Vagner, "IRESdb: the Internal Ribosome Entry Site database", Nucleic Acids Research, 2003, Vol. 31, No. 1 427–428 DOI: 10.1093/nar/gkg003.

[19]. Encarnacio´ n Martı´nez-Salas," The impact of RNA structure on picornavirus IRES activity", Trends in Microbiology Vol.16 No.5, 2008.

[20]. Laing,C. and Schlick,T, "Computational approaches to RNA structure prediction, analysis, and design", DOI 10.1016/ j.sbi.2011.03.015, Elsevier Struct. Biol: 21, 306–318, 2011.

[21]. Tamar Schlick, "Molecular Modeling and Simulation An Interdisciplinary Guide" ,2nd edition, 216-217, 2010.

[22]. Christian Laing and Tamar Schlick, "Computational approaches to RNA structure prediction, analysis, and design", Curr Opin Struct Biol. 2011 June ; 21(3): 306–318. doi:10.1016/j.sbi.2011.03.015.

[23]. Graham J. Belsham, "Divergent picornavirus IRES elements" , Elsevier: Virus Research 139 (2009) 183–192.

[24]. Yingpu Yu, Irina S. Abaeva, Assen Marintchev, Tatyana V. Pestova and Christopher U. T. Hellen, "Common conformational changes induced in type 2 picornavirus IRESs by cognate trans-acting factors", Nucleic Acids Research, 2011, 1–15, doi:10.1093/nar/gkr045.

[25]. R Ramos and E Martínez-Salas, "Long-range interaction between structural domains of the aphthovirus internal ribosome entry site (IRES)", RNA 1999 5: 1374-1383.

[26]. Andrew M. Borman, Philippe Le Mercier, Marc Girard and Katherine M. Kean, "Comparison of picornaviral IRES-driven internal initiation of translation in cultured cells of different origins", Nucleic Acids Research, 1997, Vol. 25, No. 5 925–932.

[27]. S López de Quinto and E Martínez-Salas," Interaction of the eIF4G initiation factor with the aphthovirus IRES is essential for internal translation initiation in vivo" , RNA 2000 6: 1380-1392, 2000.

[28]. Laing,C., Jung,S., Iqbal,A. and Schlick,T., "Tertiary motifs revealed in analyses of higher-order RNA junctions", J Mol Biol. 2009 October 16; 393(1): 67–82. doi:10.1016/j.jmb.2009.07.089.

[29]. Lilley,D.M., Clegg,R.M., Diekmann,S., Seeman,N.C., VonKitzing,E. and Hagerman,P.J., "A nomenclature of junctions and branchpoints in nucleic acids", Nucleic Acids Research, 1995, Vol. 23, No. 17 3363-3364.

[30]. C. Carrillo, E. R. Tulman, G. Delhon, Z. Lu, A. Carreno,Vagnozzi, G. F. Kutish and D. L. Rock, "Comparative genomics of foot and mouth diseases viruses", J. Virol. 2005, 79(10):6487. DOI:10.1128/JVI.79.10.6487-6504.

[31]. Dopazo J, Dress A, Haeseler AV: Split decomposition: a technique to analyze viral evolution. Proc Natl Acad Sci U S A 1993,90(21):10320-4.

[32]. Guohin Lin, Zhipeng Cai, Junfeng Wu, Xin-Feng Wan, Lizhe Xu and Randy Goebel, "Identifying a few foot-and-mouth disease virus signature nucleotide strings for computational genotyping", BMC Bioinformatics 2008, 9:279 doi:10.1186/1471-2105-9-279.

[33]. David W.Staple and Samuel E.Butcher, "Pseudoknots: RNA Structures with Diverse Functions", PLOS Biology, June 2005, volume 3, isuue 6, e213, DOI: 10.1371/journal.0030213.

[34]. Ian Brierley , Simon Pennell & Robert J. C. Gilbert, " Viral RNA pseudoknots: versatile motifs in gene expression and replication", Nature Reviews Microbiology 5, 598–610 (1 August 2007) | doi:10.1038/nrmicro1704.

[35]. Stephen D. Baird, Marcel Turcotte, Robert G. Korneluk and MARTIN HOLCIK," Searching for IRES ", RNA 2006 12: 1755-1785.

[36]. "http://major.iric.ca/MajorLabEn/MC-Tools.html", accessed at 5/20/2014.

[37]. Song Cao and Shi-Jie Chen, "Predicting RNA pseudoknot folding thermodynamics", Nucleic Acids Research, 2006, Vol. 34, No. 9 (2634–2652), doi:10.1093/nar/gk1346.

[38]. Rivas,E. and Eddy,S.R. "A dynamic programming algorithm for RNA structure prediction including pseudoknots" Article No. jmbi.1998.2436, J. Mol. Biol. (1999) 285, 2053- 2068.

[39]. Aalberts,D.P. and Hodas,N.O. (2005)" Asymmetry in RNA pseudoknots: observation and theory", 2210–2214 Nucleic Acids Research, 2005, Vol. 33, No. 7
doi:10.1093/nar/gki508.

[40]. Katherine E. Berry, Shruti Waghray and Jennifer A. Doudna," The HCV IRES pseudoknot positions the initiation codon on the 40S ribosomal subunit ", RNA 2010 16: 1559-1569, doi:10.1261/rna.2197210, 2010.

[41]. Christian Laing, Dongrong Wen, Jason T. L. Wang and Tamar Schlick "Predicting coaxial helical stacking in RNA junctions", Nucleic Acids Research, 2011, 1–12,doi:10.1093/nar/gkr629.

[42]. IRESite: the database of experimentally verified IRES structures (www.iresite.org), 2006.

[43]. "http://www.ncbi.nlm.nih.gov/genbank/", accessed at 7/20/2014

[44]. "http://www.biomath.nyu.edu/rag/home", accessed at 7/20/2014

[45]. Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang
Zheng Zhang, Webb Miller and David J. Lipman,"Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Research, 1997, Vol. 25, No. 17 3389–3402.

[46]. Larkin MA1, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H,Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG," ClustalW and ClustalX version 2.0", Oxford University Press, 2007.

[47]. Dongrong Wen, Christian Laing, Jason T. L. Wang and Tamar Schlick," Junction-Explorer Help File".

[48]. Marc Parisien and Franc¸ois Major, "The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data", Vol 452|6 March 2008| doi:10.1038/nature06684.

[49]. "http://www.major.iric.ca/MC-Pipeline/", accessed at 5/20/2014.

[50]. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW.,"GenBank.", Nucleic Acids Res. 2011 Jan;39(Database issue):D32-7. doi: 10.1093/nar/gkq1079.

Amr Badr is a professor in department of Computer Science, Faculty of Computers and Information, Cairo University.

Ahmed Farouk Al-Sadek is the head of Bioinformatics unit and Agricultural Expert System Development in Central Laboratory for Agricultural expert System, CLAES and also works as program leader of computer science department in Faculty of computer science, October University for modern science.

Alaa Eldin Abdallah Yassin is a member in research staff of Bioinformatics unit, in Central Laboratory for Agricultural expert System, CLAES, and also PHD student in department of Computer Science, Faculty of Computers and Information, Cairo University.