International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-3 Number-1 Issue-8 March-2013

A Comprehensive Study on Animal miRNAs: A Computational approach to explore its implications in Biological and Chemical environments

Joyshree Nath
Department of Computer Science, Jogesh Chandra Chaudhuri College, Kolkata, India

Abstract

The major research area is now to estimate microRNA (miRNA) in a quantitative manner. The researchers across the globe are working on quantitative estimation of microRNA (miRNA) from different angles. In the present paper the author has reviewed the recent trends in miRNAs and its implications in Chemical and Biological environment. The miRNAs are non-coding short ribonucleic acid (RNA) molecules, approximately ~25 nucleotides long. MicroRNAs have emerged as powerful regulators of diverse cellular and other biological processes. Many researchers have found different aspects of these miRNAs. But there are quite a number of issues not yet explored. The author will also focus on those genres which have been discovered and well accepted in the animal miRNA domain.

Keywords

Animal miRNAs, Biology, Chemical environments.

1. Introduction

The descriptive, quantitative and qualitative understanding of microRNA (miRNA) is now an emerging area of research. There is much to explore in the field of miRNAs for biological progression. MiRNAs are non-coding short ribonucleic acid molecules help in apoptosis and fat metabolism [5]. Although the most important role of miRNAs are as key post-transcriptional regulators of genetic expressions. For human and other vertebrate cell lines, miRNA genes are involved in tumor suppression, antiviral defense, adipocyte differentiation and susceptibility to cytotoxic T-cells [10]. In C. elegans, lin-4 and let-7 were first discovered as key regulators of developmental timing in early larval developmental transitions [4]. In section 3 the biogenesis and functions of miRNA are discussed. In section 4 the miRNA‟s relation with mRNA target binding and translational repression are discussed. In section 5 the biological implications of miRNAs are discussed and thus the relationship of various miRNAs and various diseases are also discussed. In section 6 the chemical implications of miRNAs are discussed. In section 7 there is a discussion on the algorithms used quantifying and classifying unknown miRNA sequence and nullifying it to be any member of hsa or ptr or mml species straightaway, as proposed by [16] and [17]. Finally in section 8 the conclusion of the present study are discussed.

2.1. Biogenesis miRNA

MicroRNAs (MiRNAs) are a class of highly conserved short non-coding RNAs. They originate in the nucleus and modify mRNA in the cytoplasm. Each miRNA is first transcribed by RNA polymerase II as longer RNA molecule called pri-miRNA. The pri-miRNA is processed in the nucleus itself into hairpin like RNA by a protein complex consisting of the ribonuclease Drosha and an RNA binding protein Pasha. This stem-loop precursor molecule, pre-miRNA, is transported to the cytoplasm by proteins Exportin-5 and Ran-GTP. In the cytoplasm, a complex called TRBP(TAR RNA binding protein) binds to the single-stranded loop structure and a ribonuclease called Dicer cleaves it to form ~22-25 base double-stranded RNA fragment. After this cleavage, Dicer and its TRBP dissociate from the miRNA duplex. The pre-miRNA attaches to the protein complex RISC (RNAi-induced silencing complex). This RISC that performs helps in transcriptional gene silencing.This complex degrades one of the strands, passenger strand(which is degraded), leaving the other RNA strand, the mature strand(which is complementary to the target). These mature miRNAs and the Argonaute protein family form the RISC protein complex [8]. The mature strand bind to its target mRNA(preferably in the

3‟UTR region). mRNA or messenger RNA carries information of the directions for protein synthesis by ribosomes. Complementary base-pairing of the miRNA guides RISC to its corresponding target mRNAs, which are degraded, or translationally inhibited by the Ago proteins [5].

2.2 Function of miRNA

Some of the miRNA functions are as follows:

1) MiRNAs help in apoptosis and fat metabolism [5].
2) MicroRNAs are involved in regulating gene expressions that control different biological processes and, such as development, differentiation, viral infection and growth control, cancer etc. [6].
3) For human and other vertebrate cell lines, miRNA genes are involved in tumor suppression (defect in the biogenesis of miRNA may lead to tumorigenesis, antiviral defense, adipocyte differentiation and susceptibility to cytotoxic T-cells [10].
4) In C. elegans, over expression or under expression of a miRNA either accelerates or delayes differentiation of cells, thus helps in controlling the developmental timing [13].
5) MiRNA miR-27 is important for myogeneis and miRNA miR-9 is important in neurogenesis [13].

3.1 MiRNA and its effects on translational repression and gene silencing

It is seen that many prior works in bio informatics suggest that miRNAs lead to gene silencing and target mRNA degradation, and the latter being caused by the former reason. Different mRNA are translated at different paces into protein sequences and miRNAs are negative regulators of mRNA translation From stable and enduring mRNA molecule comes more and more synthesized proteins. Now here mRNA degradation means the method by which it is determined which of the genes are to be expressed in a cell among the thousands of them available and the amount of proteins to be translated from each mRNA. miRNAs inhibit the mRNA translation by deadenylation of mRNAs, thus removing rhe the cap and poly(A) tails of the target mRNAs. MiRNA helps RISC with transcript specificity to facilitate the transcript silencing or in other words transcript destruction. Thus it is experimented and seen that miRNAs inhibit the different levels of peptide or protein synthesis by reducing the corresponding mRNAs. Experiments have shown that the ultimate gene silencing or regulation of gene expression or mRNA degradation is a successive effect after the translational repression of miRNA [2], [3]. During the biogenesis of miRNAs mature ones associate with Argonaute (Ago) proteins to form the core of the RISC. In mammals, the miRNA mediated gene inhibition occurs by site-specific cleavage with the help of Ago2. Also this phenomenon occurs with the help of rigorous mRNA degradation and translational inhibition (a process where protein products are decreased more than that of the mRNA molecules) [11], as also mentioned earlier. Argonaute proteins have many functions in the miRNA biogenesis. They help in the much talked about mRNA degradation. The Ago proteins regulate miRNA abundance post-transcriptionally, and also regulate the expressions and functions of mature miRNAs [5].

Also the protein group GW182 help in this gene silencing along with the Argonaute protein family, leading to an ultimate deadenylation of the mRNAs. It is seen mRNAs without poly-A tails are subject to translational repression and that is brought about in some cases by the protein Ago1 along with a target mRNA degradation being brought about by GW182 [11]. The miRNAs helps in translational repression of some target mRNAs by translational inhibition and thus reducing the target mRNA levels [2], [11]. Having said all this, the exact procedure of how miRNAs inhibit the mRNA molecules is still to be known.

3.2. MiRNAs and their mRNA Targets

Identifying the correct mRNA targets is necessary to get the correct implication of miRNA sequences over disease. It also may happen that a single miRNA target more than one mRNAs or a single target mRNA may be bonded by many miRNAs. Within the past few years lots of animal miRNA targets are identified and validated by various target prediction tools. Few of the target prediction rules that are followed while making a target prediction algorithm are checking base pairing pattern and alignment of the miRNA and target mRNA(e.g-miRanda,TargetScan and TargetScanS, PicTar), Thermodynamic stability of miRNA -mRNA duplex(e.g.-DIANA-microT, RNAHybrid), Comparative sequence analysis to check conservation(e.g.-Watanabe et al., 2007), Checking for the presence of multiple target sites(e.g.-Rajewsky, 2006) [14]. The most functional mRNA: miRNA pairing resides in the 3‟UTR region. Genes that are considered to have evolved through selected

elimination of the targeting sites are called anti-targets [13].

4.1. MiRNAs involved in different cardiovascular activities

MiRNAs miR-1-1 and miR-1-2, are specifically expressed in cardiac and skeletal muscle cells. The regulated gene expression of miR-1 helps in balancing between the differentiation and proliferation of cardiac progenitor cells during cardiogenesis. These progenitor cells help in regenerating the cardiovascular system during disease. MiRNAs miR-133a1 and miR-133a2 help in cardiac development. MiRNA miR-21(including cardiac fibroblasts) increases during cardiac hypertrophy. MiR-21 suppreses cardiac fibrosis and improves cardiac function. For smooth muscle cells A regulation of miRNAs miR-145 and miR-143 are very much important [12].

4.2. MiRNAs involved in different carcinogenic activities

MicroRNA-21 was one of the first microRNAs to be found in relation with carcinogenic activities. More than 50% of miRNAs are located at different genomic areas that are disrupted or amplified in various cancers [13]. Clear cell carcinoma makes up about more than 50% of Renal cell and is associated with high mortality rate. MiR-200c and MiR-141 are down-regulated in this clear cell carcinoma. miR-21 as for cardiovascular activities also, is an important factor here and it is very much upregulated in cancers of different kinds like pancreatic, breast, throat, lung and many more [13].

4.3. MiRNAs involved in different carcinogenic activities

MiRNAs somewhat control the nervous system by over expression or under expression and thus either help in brain development or formation of various neural diseases [13]. MiRNAs have neurodegenerating implications, especially the miR-1 and miR-133 miRNAs are linked in this matter. Other miRNAs like miR-132, miR-134, miR-138, miR-124 are involved in different types and levels of neural development. Over expression of miR-134 may lead to neural defect and that of miR-124 may lead to neurogenesis. Defective expression of miRNAs lead to Alzheimer's disease, Parkinson's disease etc. [13].

4.4. MiRNAs involved in different carcinogenic activities

MiRNAs are very crucial in the genre of differentiating adipocytes. MiRNAs miR-155,miR-221,and miR-222, are found to be negative initiators of differentiation. The let-7 class of miRNAs that regulate diabetes, adipocyte differentiation,etc. Psychiatric diseases like Schizophrenia are genetic disorder that happens due to differential expression of miRNAs [13].

5. MiRNAs involved in chemical activities

It is seen through experiments that miRNAs can regulate gene expression along with some other factors, without involving DNA sequence changes. Here also as earlier lies the same thing that the correct and appropriate mechanism of the alteration of DNA sequences by the chemical factors is still not known. The effect of the different environmental chemicals can be so much so hostile as to change or modify the DNA sequence itself. Thus miRNAs regulates final gene expression after combining with its target mRNAs and along with the effect of the environmental chemicals. Thus these chemicals can very well regulate miRNA expressions. Some metals like arsenic, cadmium, mercury etc. (that are found naturally and excessively in the environment) are causes to different sorts of diseases and also increasing and modifying the miRNA expression It has been accepted that the etiology roles of heavy metals are dependent on both the genetic background and amount and duration of exposures. Experimental and observational data have also linked altered miRNA expression with exposure to arsenic, cadmium, and aluminum, as detailed below. The miRNAs could even be down regulated due to such metal effect leading to apoptosis or very less cell development. Also metals like cadmium increases the chances of heart diseases. One such miRNA is miR-146a, that gets increased in its expression level due to the effect of such heavy metals like aluminum or cadmium and has a carcinogenic and adverse effect on the blood vessels or in the nervous system. Also there are few other miRNAs whose ultimate expression gets changed on getting attacked by any such harmful chemicals. For example there are miR-21, miR-222, miR-10, miR-26, miR-124, miR-191and many more whose gene expression gets affected due to environmental pollution. Thus due to

presence of harmful chemicals in the surrounding environment, the miRNAs get either highly over.
expressed or highly under expressed and thus possibly give rise to different genres of diseases [15].

6. Computational Algorithms proposed by Nath et al.[16, 17]

To classify and to quantify pre_mature miRNAs Nath et al[16-17] introduced the statistical and fractal features. The results are shown in section-8.

(i) Extracting the dataset and Generating Indicator Matrices

The pre-mature miRNA sequences of the three organisms Homo sapiens (hsa), Macaca mulatta (mml) and Pan troglodytes (ptr) were extracted from the miRNA database, miRBase, version 17.

The miRNA consists of 4 basic nucleotides, i.e, A=Adenine, C=Cytosine, U=Uracil, G=Guanine. So a set is taken to define the four, V and xЄV. The indicator matrix of an N-length miRNA string is taken as N×N sparse symmetric, binary matrix and denoted as,
Mhk= fxh(xk) xh,xk
ЄS,h,k=1,2,3…N

Mhk for a respective miRNA sequence is a four threshold matrix, namely 0, 1, 2 and 3. The matrix Mhk, was decomposed into four binary matrices A1, A2, A3 and A4 as follows [18]:

$$A1_{hk}=\begin{cases}1, & \text{where } xh = xk: xh, xk \in S\\ 0, & \text{otherwise}\end{cases}$$

$$A2_{hk}=\begin{cases}1, & \text{where } xh \neq xk; xh, xk \in \{G,U\} or \{A,C\}\\ 0, & \text{otherwise}\end{cases}$$

$$A3_{hk}=\begin{cases}1, & \text{where } xh \neq xk; xh, xk \in \{U,C\} or \{A,G\}\\ 0, & \text{otherwise}\end{cases}$$

And

$$A4_{hk}=\begin{cases}1, & \text{where } xh \neq xk; xh, xk \in \{C,G\} or \{A,U\}\\ 0, & \text{otherwise}\end{cases}$$

(ii) DNA Walk of miRNAs

DNA walk is defined on a 1-D real signal {Yn}, as a series ΣYn, Y{0,1,2,3} which is the cumulative sum on the miRNA sequence.

To calculate the DNA walk of a miRNA string the summation of the matrix cell value for each nucleotide with the other nucleotides is calculated [17].
The definition follows as,

$$a_n \stackrel{def}{=} \sum_{i=1}^{n} f(A, x_i),$$
$$g_n \stackrel{def}{=} \sum_{i=1}^{n} f(G, x_i),$$
$$c_n \stackrel{def}{=} \sum_{i=1}^{n} f(C, x_i),$$
$$u_n \stackrel{def}{=} \sum_{i=1}^{n} f(U, x_i),$$

[where,
i=1,2,3…N; x Є V the value of x at position i] [17,18]
The values of the nucleotides generated for the above formulae come from the following table [17]:

Table-1

f	A	U	C	G
A	0	3	1	2
U	3	0	2	1
C	1	2	0	3
G	2	1	3	0

$W_n \stackrel{def}{=} \sin a_n^2 - \sin g_n^2$ and $V_n \stackrel{def}{=} \sin u_n^2 - \sin c_n^2$

(iii) Complexity of miRNA strings

Non-repetitiveness of a string refers to its complexity unlike its periodicity and patchiness. The complexity of a string of every miRNA string for every species was calculated [17,18].

(iv) Hurst Exponent of miRNA strings

The Hurst exponent for each miRNA string was calculated, taking A=0, C=1, G=2, and U=3:
The mean and Y(i,x)(Y itself being the miRNA string) was calculated as,

$m_{x,n} = \frac{1}{n}\sum_{i=1}^{n} x_i$, where x is the miRNA string and

$Y(i,x) = \sum_{j=1}^{i}\{x_j - m_{x,n}\}$

Therefore the Hurst exponent :$(\frac{n}{2})H = \frac{R(n)}{S(n)}$, for a miRNA string comes from:
R(n)=maxY(i,n) minY(i,n) 1

$$S(n) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - n)^2}^{n}$$

The formulae are taken from [17] and [18].

(v) Variance of miRNA Strings

Another mathematical parameter is the value of variance for each string. For a given miRNA string sequence of N length, the variance at distance N-k(where k is 0,1,2…N-1) was calculated [17, 18].

(vi) Poly-String Mean and Standard Deviation of miRNA Strings

For each of the nucleotide of the miRNA strings of each of those three species there were 4 poly -string mean and 4 poly-string standard deviation values

were obtained. Let total number of poly -strings of different size k_1, k_2, k_3,…k_n of nucleotide sequence length N are m_1, m_2, m_3,… m_n. Then poly-string mean (P_m^N) and poly-string standard deviation (P_{SD}^N) of N are defined as

$$P_m{}^N = \frac{\sum_{i=1:n} m_i k_i}{\sum_{1:n} m_i} \text{ and } P_{SD}{}^N = \sqrt{\frac{1}{n}\sum_{i=1}^{n} m_i (k_i - P_m{}^N)^2}$$

[16, 19].

7. MiRNAs involved in different cardiovascular activities

A. Generating Indicator Matrices

For each of the species a set of four binary matrices A1, A2, A3 and A4 for each of the miRNA sequences were extracted and the corresponding fractal dimension values were also calculated [17]. The values are:

For hsa the values of A1,A2,A3,A4 varied in between an average value of (1.44065-1.68023),
For mml the values of A1,A2,A3,A4 varied in between an average value of (1.41198-1.67725),
For ptr the values of A1,A2,A3,A4 varied in between an average value of (1.41653-1.67951).

B. DNA Walk of miRNAs

The range of the Fractal dimensions of the miRNA sequences of hsa, mml, ptr are (1.89608-1.94513), (1.92149-1.94491), (1.94083-1.94513) respectively [17].

C. Complexity and Variance of miRNA strings

Since the values in the denominator for the factorials of all nucleotides multiplied into huge numbers the total complexity tended to converge towards infinity. Also in the case of variance of each miRNA string, the measures for every nucleotide string converged to an approximated zero value [17].

D. Hurst Exponent of miRNA strings

The range of the Hurst exponents values for the miRNA sequences of these three species are(as taken from [17]):

Specie s	Mi RNA sequence of has	MiRNA sequence of mml	MiRNA sequence of ptr
Hurst Expon ent	(0.0309,0.1 756)	(0.0425,0.1 113)	(0.0295,0.1 238)

E. Poly-String Mean & Standard Deviation of miRNA Strings

The resulting poly-string mean and poly-string standard deviation of each nucleotide are given as [17]:

mml-ACUG/CAUG, hsa-UGAC/AUGC, ptr-AGUC/ACUG. The set before „/" depicts poly-string mean and the one after depicts poly -string standard deviation. This is used to classify the unknown miRNA string and predict the near most genre of the species of that miRNA string.

8. Conclusion and Future Scope

In this paper the author has discussed the different facts on animal miRNAs that is gathered over the past few years and experimented and seen to be correct by various bio informatics scientists. This paper will help in understanding collectively about miRNA in different arenas, such as, biogenesis, functions, biological implications, chemical implications etc. The researchers can get to know from this paper about miRNAs and what tremendous effect it has along with its target messenger RNAs in the environment and in people"s life. More and more scientists can be motivated on doing more work in relation to the miRNA field. Also there is a final discussion on some previous work done on quantifying pre-mature miRNAs using fractal and statistical parameters. With these parameters set biologists can rule out any unknown miRNA sequence that is not falling under the values set by [16] or [17] as any candidate for the miRNA of the species of hsa, mml or ptr.

References

[1] Quantification of M i-RNAs and Their Networks in the light of Integral Value Transformations. , Sk. S. Hassan, P. Pal Choudhury, A. Goswami, N. De Sarkar, V. Fangal (2011) (Under review-PloS-ONE).
[2] http://www.nature.com/cr/journal/v22/n9/ full/ cr201280a.html?WT.ec_id=CR-201209.
[3] http://www.ehow.com/ facts_6196816_degradation-mrna_.html
[4] Human M icroRNA Targets, Bino John, Anton J.
[5] Winter J, Jung S, Keller S, et al. M any roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol 2009;11: 228–234.
[6] Prediction of both conserved and non-conserved

microRNA targets in animals, Xiaowei Wang and Issam M . El Naqa, Oxford Journals Life Sciences & M athematics & Physical Sciences, Bioinformatics, Volume 24, Issue3,Page. 325-332(2008).

[7] Got target? Computational methods for microRNA target prediction, Hyeyoung M in and SungrohYoon, ExpM ol M ed. April 30; 42(4): 233-244.

[8] A Comprehensive Study of Target Prediction Algorithms for Animal M icroRNAs(miRNAs), International Journal of Computer plications(IJCA)(0975-8887,USA), Joyshree Nath, Asoke Nath, Vol-40, No 15(Feb),(2012).

[9] M icroRNA target prediction by expression analysis of host genes, Vincenzo AlessandroGennarino, M arco Sardiello, Raffaella Avellino, Nicola M eola, VincenzaM aselli, SantoshAnand, Luisa Cutillo, Andrea Ballabio, and SandroBanfi,Genome Res, v.19(3); M ar 2009.

[10] Advances in microRNAs: implications for therapists., M arquez RT, M cCaffrey Ther. 2008;19:27–38.

[11] Gu S, Kay M A. How do miRNAs mediate translational repression? Silence. 2010;1:11.

[12] Bauersachs regulation of Res 109: 334- 347.

[13] D. Sayed and M . Abdellatif, "M icroRNAs in development and disease," Physiological Reviews, vol. 91, no. 3, pp . 827–887, 2011.

[14] A Comprehensive Study of Target Prediction Algorithms for Animal M icroRNAs(miRNAs), International Journal of Computer Applications(0975-8887, USA), Vol-40, No15(Feb),2012.

[15] http://www.sciencedirect.com/science/ article/ pii/S0027510711001138.

[16] A new algorithm for Quantitative deciphering of pre-mature M iRNAs using some Statistical Parameters, Joyshree Nath, Asoke Nath, Proceedings of IEEE International Conference WICT-2012 held at IIITM -K, Trivandrum Oct 30 to Nov 1, 2012, Page No. 595-601(2012).

[17] A M odified Algorithm for Quantifying of Pre-mature M iRNAs Some Fractal Parameters, Joyshree Nath, Asoke Nath , International Journal of Advanced Computer Research(ISSN(print):2249- 7277 ISSN(online): 2277-7970), Volume-2, Number-4 Issue-6, Page-202-207, Dec(2012).

[18] Quantification of M i-RNAs and Their Networks in the light of Integral Value Transformations. , Sk. S. Hassan, P. Pal Choudhury, A. Goswami, N. De Sarkar, V. Fangal (2011) (Under review-PloS-ONE).

[19] ww3.appliedbiosystems.com /cms/groups/ mcb_marketing/documents/generaldocuments/.

Joyshree Nath was born on 17th August 1986. After finishing M .Tech(IT) from C.U. now she is taking classes of Computer Science in BSc. In a part-time basis. She is also approved academic counselor of IGNOU M CA programme. She is involved in research work in Quantum Computing, Bioinformatics and symmetric key cryptography.