NEWS
OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
SERIES OF BIOLOGICAL AND MEDICAL
ISSN 2224-5308
Volume 4, Number 328 (2018), 36 – 46

IISTI 34.15.25; 76.29.30; 76.29.49; 76.29.51

A. T. Ivashchenko*[1], R. Ye. Niyazova[1], Sh. A. Atambayeva[1], A. Yu. Pyrkova[1], D. E. Aisina[1], O. Yu. Yurikova[1], A. Kondybayeva[2], A. Akimniyazova[1], D. Bayzhigitova[1], A. A. Bolshoy[3]

[1]Research Institute of Problems of Biology and Biotechnology, al-Farabi Kazakh National University, Almaty, Kazakhstan,
[2]S. D. Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan,
[3]University of Haifa, Haifa, Israel.
E-mail: a_ivashchenko@mail.ru; raygul.nyiyazova@kaznu.kz; shara.atambaeva@kaznu.kz; Anna.Pyrkova@kaznu.kz; dana.aisina03@gmail.com; oksasha1992@gmail.com; dr.kondybayeva@gmail.com; 401052@mail.ru; dianabay@mail.ru; bolshoy@research.haifa.ac.il

miRNA: ACHIEVEMENTS, MISCONCEPTIONS, PERSPECTIVES

Abstract. Among small RNAs miRNAs play an important role, they carry out the regulation of gene expression at the post-transcriptional level. The paper discusses the properties of miRNAs and their interaction with mRNAs. It is shown the role of miRNA in the regulation of the expression of protein-coding genes involved in various metabolic processes and the development of cardiovascular, oncological and neurodegenerative diseases. The features of miRNA binding sites in 5'UTR, CDS and 3'UTR of mRNA of target genes have been established. There were shown the advantages of the MirTarget program prior to known search programs for miRNA binding sites with mRNA. In mRNA of many candidate genes of various diseases, single miRNA binding sites and miRNA binding site clusters are detected in 5'UTR, CDS and 3'UTR of mRNA. There was found miRNA binding sites that encode oligopeptides in proteins in mRNA of transcription factor genes. It was analyzed the interaction of miRNA with mRNA of candidate genes involved in cardiovascular, oncological and neurodegenerative diseases. The properties of unique miRNAs binding sites in mRNAs of several hundred genes were discussed. There were considered the features of the interaction of mRNA with miRNA in the RISC complex. Discussed the role of miRNA in the regulation of gene and genome expression through the interaction of genes involving miRNA host genes. It is proposed the hypothesis of regulation of gene and genome expression involving miRNA. Shown the role of miRNA as an integrating system for the mutual regulation of gene expression in the cell and in the body.

Keywords: miRNA, mRNA, genes, binding sites, bioinformatic programs.

Introduction. During the period of studying miRNA, many original articles and reviews have been published and important properties of the functioning of these molecules have been established and discussed [1-7]. The conducted researches have allowed to identify the features of miRNA properties and their interaction with mRNA. The obtained knowledge was the basis for the use of synthetic molecules of siRNA by which it is possible to completely suppress the translation process or destroy mRNA. For the development of this method of turning off genes, a group of scientists was awarded the Nobel Prize. However, despite significant success in studying the interaction of miRNA with mRNA, with rare exceptions, miRNA cannot be used for practical purposes, in particular, in diagnosis and therapy of diseases. The reasons for the poor performance of miRNA studies are inadequate assumptions about their properties that brake on the identification of the biological role of miRNA and the use of miRNA in biology, biotechnology, medicine, etc.

Small noncoding RNAs (ncRNAs) include transfer RNAs (tRNAs), antisense RNAs (asRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), micro RNAs (miRNAs), Piwi-interacting RNAs (piRNAs), competing endogenous RNAs (ceRNAs), tRNA-derived small RNAs (tsRNA) and small interfering RNAs (siRNAs), which play an important role in the expression of genes and genomes [8, 9]. The miRNAs participate in the regulation of gene expression at the post-transcriptional level, suppressing the translation process. The miRNAs are nanoscale molecules with a length of five to nine nanometers consist of 18 to 27 nucleotides. Consequently, the name of the microRNA contradicts the nanoscale of these molecules. The name of the microRNA has another drawback - it does not reflect the function of these molecules. However, the used "miRNA" abbreviation can be treated as mRNA-inhibitory RNA (miRNA), which corresponds to the function of these molecules. The term "microRNA" should be phased out as an inadequate term. The interaction of mRNA with miRNA is studied using different approaches: how much miRNA binds to one gene; how many genes are targeted by a single miRNA; what are the criteria for predicting sites and the energy of interaction of mRNA and miRNA; whether there are gene any preferences for miRNA binding; what proteins functions encoded by target genes of specific miRNA are; what are functional links of miRNA in the implementation of post-transcriptional regulation of gene expression; how, what and to what extent the synthesis of miRNA is regulated and etc.

The present work is devoted to the consideration of miRNAs properties and their interaction with mRNAs, which in our opinion can substantially clarify the existing problems in the study and application of miRNAs for practical purposes.

Materials and methods. The nucleotide sequences of mRNAs of genes were downloaded from NCBI GenBank (http://www.ncbi.nlm.nih.gov). Nucleotide sequences of miRNAs were downloaded from the miRBase database (http://mirbase.org) and borrowed from the article of Londina E. et al. [10]. miRNA binding sites in 5'-untranslated regions (5'UTRs), coding domain sequences (CDSs) and 3'-untranslated regions (3'UTRs) of several genes were predicted using the MirTarget program [11]. This program defines the following features of binding: a) the origin of the initiation of miRNA binding to mRNAs; b) the localization of miRNA binding sites in 5'UTRs, CDSs and 3'UTRs of the mRNA; c) the free energy of hybridization (ΔG, kJ/mole); and d) the schemes of nucleotide interactions between the miRNA and the mRNA. The ratio ΔG/ΔGm (%) was determined for each site (ΔGm equals the free energy of miRNA binding with its perfect complementary nucleotide sequence). The miRNA binding sites located on mRNAs had ΔG/ΔGm ratios of 90% or more. The program identifies the positions of binding sites on mRNA, beginning from the first nucleotide of the mRNA's 5'UTR. The MirTarget program found hydrogen bonds between adenine (A) and uracil (U), guanine (G) and cytosine (C), G and U, and A and C. The distances between A and C were equal to those between G and C, A and U, and G and U. The numbers of hydrogen bonds in the G-C, A-U, G-U and A-C interactions were found to be 3, 2, 1 and 1, respectively. The free binding energies of these nucleotide pairs were taken as the same ratio, i.e., 3, 2, 1, and 1, respectively.

Results and discussions. After detecting miRNA and establishing their interaction with mRNA, it became necessary to develop programs for predicting miRNA binding sites in mRNA. This need derives from the fact that more than 6000 miRNA encoded in the human genome can potentially bind to all mRNA of 20000 genes and their isoforms, encoded in the human genome. It has been established that miRNA can bind to mRNA blocking the translation [12]. Therefore, it needs to create programs that establish miRNA binding sites in mRNA and quantitative characteristics of the interaction of these molecules, evaluating the effectiveness of this binding.

The need to take into account the interaction of miRNA with mRNA throughout the entire miRNA nucleotide sequence, rather than the "seed", is due to several factors. For the high specificity of the interaction of miRNA with mRNA, the entire length of miRNA must be taken into account. It is same as applying primers, of at least 20 nucleotides length in polymerase chain reaction, when it is necessary for amplification to specifically choose only one nucleotide sequence among numerous nucleotide sequences. Another reason for using the full miRNA nucleotide sequence is that during the evolution process, a part of miRNA other than "seed" will vary, if it is not a critical site in interactions of miRNA with mRNA. Confirmation of the need to maintain the entire length of miRNA is the high conservatism of miRNA in organisms that have diverged over tens of millions of years of species evolution. According to miRBase,

the nucleotide sequences were identical in miR-200c-5p of human, mouse, rat, miR-216a of human, bull, mouse, miR-574-5p and miR-574-3p of human, mouse, rat, pig, etc.

TargetScan program finds the binding site for miR-3180, miR-3180-3p, miR-3196, miR-6816-5p in mRNA of *TGFB1*, gene coinciding with MirTarget program (table 1). However, with equal probability TargetScan indicates other sites that have an identical "seed" region. Such prediction increases the number of false-positive sites by tens of times, which makes prediction of sites extremely ineffective. Similarly, other programs [13] based on the search for miRNA binding sites in mRNA over the homology of 6-8 nucleotides "seed" at the 5' end of miRNA are also inadequate. The MirTarget program predicts the most likely binding sites for miR-6816-5p in CDS of *TGFB1* mRNA. Other miRNAs have less free binding energy and a smaller value of ΔG/ΔGm, which indicate a weak interaction with mRNA of *TGFB1*. Thus, miRNA-mRNA interaction schemes presented in table 1 clearly show the inadequacy of the TargetScan program and other programs based on the use of "seed" [13].

Table 1 – Schemes of miRNA interaction in mRNA of *TGFB1* gene

Program MirTarget	Program TargetScan

TGFB1, miR-6816-5p, 2051, CDS,-113, 90, 21
5'-GCUGAGGUCCCGCCCCGCCCCG-3'
  || | || |||||||||||||
3'-CGUC-CCUGGACGGGGCGGGGU-5'
	TGFB1, miR-6816-5p, 3'UTR, 21
	5'-...NNGGUCCCGCC**CCGCCCC**G-3'
	  |||||||
	3'- CGUCCCUGGACGG**GGCGGGG**U-5'

TGFB1, miR-3196, 2054, CDS, -98, 88, 18
5'-GAGGUCCCGCCCCGCCCCG-3'
  |||||||| | ||||||||
3'-CUCCGGGG-ACGGCGGGGC-5'
	TGFB1, miR-3196, 3'UTR, 18
	5'-..NNGGUCCCGCCCCGCC**CCGCCCC**G-3'
	  ||||| |||||||
	3'- CUCCGGGGAC-**GGCGGGG**C-5'

TGFB1, miR-3180-3p, 471, 5'UTR, -104, 80, 22
5'-AGCCCUCGGGAGUCGCCGACCCG-3'
  |||||||||||| | ||| ||||
3'-CCGGAGGCCUUC-GAGGCGGGGU-5'
	TGFB1, miR-3180-3p, 3'UTR, 22
	5'-..NNGGUCCCGCCCCGCC**CCGCCCC**G-3'
	  |||||||
	3'- CCGGAGGCCUUCGA**GGCGGGG**U-5'

TGFB1, miR-3180-5p, 227, 5'UTR, -106, 74, 25
5'-CCACUGCGGGGAGGAGGGGGAGGAGG-3'
  | || |||||| |||| | |||||
3'-GCUG-CACCCCGCCUCGCAGACCUUC-5'
	TGFB1, miR-3180 , 3'UTR, 19
	5'-..NNGGUCCCGCCCCGCC**CCGCCCC**G-3'
	  |||||||
	3'- GAGGCCUUCGA**GGCGGGG**U-5'

Note: Gene; miRNA; the beginning of binding site; the miRNA region; the free energy change (ΔG, kJ/mole); the ΔG/ΔGm (%); length of miRNA (nt). In bold type highlighted the "seed" nucleotides.

One of the first misconceptions in the study of miRNA was the assumption that miRNA binds only (or predominantly) in the 3'UTR mRNA of human genes [14]. However, miRNAs do not have the property of distinguishing binding sites in 5'UTR, CDS and 3'UTR. miRNA interact with mRNA on the basis of physicochemical properties of these molecules. Therefore, the interaction site can be located in any mRNA region and as yet unknown prohibitions on the location of such sites throughout the mRNA nucleotide sequence. The conditions for the successful interaction of miRNA with mRNA are energy characteristics and conformational properties of this interaction. The proposed assumption of preferential miRNA binding in the 3'UTR contradicts several established properties of miRNA and mRNA. The proposed assumption of preferential miRNA binding in the 3'UTR contradicts several established properties of miRNA and mRNA. Considering that known miRNAs (miRBase) have differences in GC-content in the same range as in human genes, it was logical to assume that the probability of binding miRNA in 5'UTR, CDS and 3'UTR of mRNA will correlate with the GC-content of genes and corresponding miRNAs.

In numerous publications, sites of miRNA interaction with mRNA were studied only in the 3'UTR. This is because practically all programs for detecting miRNA binding sites are predicted only in this mRNA region. Several publications show the interaction of miRNA with mRNA in 5'UTR and CDS [15-19].

Using the MirTarget program, we have found binding sites in 5'UTR, CDS and 3'UTR of many animal and plant genes. It is shown, that miRNAs on average have a large free energy of interaction in the 5'UTR, because GC-content of miRNA and mRNA binding sites have greater importance, than in the

CDS. In turn, on the average free energy of miRNA interaction in the CDS is larger than miRNA interaction in the 3'UTR mRNA. The share of binding sites in 5'UTR, CDS and 3'UTR mRNA sites accounts for 20%, 54% and 26%, respectively. Based on 1000 nucleotides of 5'UTR, CDS, and 3'UTR length, the binding site density in these sites is 24.2, 4.3 and 5.1 sites on 1000 nucleotides, respectively. This distribution of binding sites in 5'UTR, CDS and 3'UTR has biological significance. Since binding of miRNA to mRNA can lead to stop translation and to separation of abortive polypeptides, then for energy saving is beneficial to stop translation at beginning of this process by binding of miRNA in the 5'UTR. For the same reason, miRNA binding sites in the CDS are generally located at beginning of CDS.

Another property of location of miRNA binding sites in 5'UTR, CDS, and 3'UTR is the optimization of their localization in sites with overlapping of binding sites of different miRNAs. This allows several and even dozens of miRNAs to interact with mRNA in a short area. This feature is particularly important when localizing of miRNA binding sites in the CDS, as these sites may encode oligopeptides is not involved in the functioning of protein.

Polysites increase the probability of binding miRNA with mRNA. Typically, among many sites there is a site that interacts with miRNA more effectively than other sites. Polysites of miRNA binding sites encode polyaminoacids that have functional significance, for example, interact directly with DNA, or with a protein bound to DNA [20].

miRNAs are inhibitors of translation reaction, and a biochemical approach to evaluating the action of inhibitors is applicable to them. Below it is shown a diagram of interaction of mRNA with miRNA included in RISC:

$$\text{miRNA} + \text{RISC} \leftrightarrow \text{miRISC} + \text{mRNA} \leftrightarrow \text{miRISC} \equiv \text{mRNA} \rightarrow \text{RISC} + \text{res-mRNA},$$

where miRISC is the association of all proteins of the RISC complex with miRNA; miRISC ≡ mRNA is miRISC complex with mRNA due to hydrogen bonds; res-mRNA is restricted mRNA. The diagram shows the following processes. The miRNA binds to a group of RISC proteins, forming miRISC. Then, miRISC binds to mRNA via hydrogen bonds and blocks protein synthesis, or miRISC cuts mRNA, which is further destroyed by cytoplasmic restriction enzymes. Binding stage of miRISC with mRNA is reversible and in the absence of their interaction, mRNA can again serve as a template for protein synthesis. It follows from this scheme that different effects can be observed on the ratio of concentrations of miRNA and mRNA. Assume that miRNA is complementary to the binding site in mRNA, that is, it has a high affinity for mRNA. Despite this, at low concentrations of miRNA compared to mRNA, the complex will have little effect on protein synthesis, since it will block the small quantity of mRNA. If concentration of miRNA is comparable or greater than the concentration of mRNA, protein synthesis will be slowed down or completely inhibited. With an average affinity of interaction of miRNA with mRNA, effect of complete inhibition of protein synthesis can be achieved with miRNA concentrations much greater than mRNA. Therefore, when calculating the probability of the degree of inhibition of gene expression by miRNA, it is not sufficient to know the affinity of miRNA for mRNA and the ratio of their concentrations.

In addition, it is necessary to take into account degree of intramolecular interaction of miRNA binding sites with other sites of mRNA. As a rule, intramolecular interactions are weaker than those of miRNA with mRNA, but cases of almost or completely complementary intramolecular interaction of these sites are known. In this case, energy is needed to break bindings of miRNA with mRNA comparable to binding energy of miRNA to mRNA. Therefore, calculation of probability of binding miRNA to mRNA only on basis of known programs for prediction of binding sites is not adequate, since it does not take into account intramolecular interactions in mRNA.

Considered variants of conditions for interaction of miRNA with mRNA are realized in cells. It is known that the concentration of miRNA can vary hundreds of times in cells [21]. Synthesis of mRNA, depending on the functional state of cell, can also vary hundreds of times [22]. In addition, gene expression and miRNA synthesis are tissue-specific [23]. Even in experiments of study of miRNA effect on protein synthesis concentrations of miRNA and mRNA are often not indicated. An important factor in studying of miRNA interaction with mRNA under *in vivo* is difficult-to-take effect of intronic miRNA (inmiRNA), which, as a rule, is synthesized coherently with the expression of host gene. Of all human

miRNA intergenic miRNA (ig-miRNA) constitute 40%, intronic - 52%, exonic - 5%, and the rest of 3'UTR and 5'UTR.

Study of ig-miRNA and in-miRNA in a comparative sense is necessary because ig-miRNA precursors are transcribed from DNA with the help of their promoters, and in-miRNAs mainly ripen from pre-mRNA and only a part of their precursors are transcribed directly from DNA. In addition, expression of in-miRNA depends on expression of host genes in introns in which they are located, whereas ig-miRNA is expressed independently.

Prospects of miRNA application in directed regulation of metabolism and diagnosis of diseases and therapy. Possibilities of using of miRNAs for these purposes of medicine are enormous, because miRNAs are endogenous, physiological regulators of biological processes, they escape from the action of the protective mechanisms of body (immune system, proteases, nucleases, etc.). Diagnostic method using miRNA can be based on miRNA associations and their target genes that interact almost complementarily. Confirmation of stability of such association is the preservation of orthologous associations in animals of related species. We have shown that such associations are conservative for tens of millions of years after the divergence of species.

Table 2 – Schemes of interaction of miRNA with mRNA of candidate genes subtypes of breast cancer

Triple-negative (Basal) subtype	
ATM, miR-619-5p, 9793, 3'UTR, -119, 98 5' – GGCUCACGCCUGUAAUCCCAGC – 3' 3' – CCGAGUACGGACAUUAGGGUCG – 5'	*AXL*, miR-1273g-3p, 3'UTR, -115, 98 5' – CCCAGGCUGGAGUGCAGUGGU – 3' 3' – GAGUCCGACCUCACGUCACCA – 5'
IAPP, miR-5096, 876, 3'UTR, -113, 100 5' – GCCUGACCAACAUGGUGAAAC – 3' 3' – CGGACUGGUUGUACCACUUUG – 5'	*CEACAM5*, miR-5095, 3229, 3'UTR, -115, 98 5' – CGCGGUGGCUCACGCCUGUAA – 3' 3' – GCGCCACCAAGUGCGGACAUU – 5'
ERBB3, miR-619-5p, 5104, 3'UTR, -121, 100 5' – GGCUCAUGCCUGUAAUCCCAGC – 3' 3' – CCGAGUACGGACAUUAGGGUCG – 5'	*IL11*, miR-1273f, 1466, 3'UTR, -102, 98 5' – CACUGCAACCUCCACCUCC – 3' 3' – GUGACGUUGGAGGUAGAGG – 5'
MAGEA10, miR-1273e, 2188, 3'UTR, -110,95 5' – UCCGCCUCCUGGGUUCAAGCGA – 3' 3' – AGGUGAAGGACCCAAGUUCGUU – 5'	*MAGEA10*, miR-1273e, 2188, 3'UTR, -110, 95 5' – UCCGCCUCCUGGGUUCAAGCGA – 3' 3' – AGGUGAAGGACCCAAGUUCGUU – 5'
Her2 subtype	
GTF2E1,miR-1273g-3p, 1720, 3'UTR,-108,93 5' – CCCAGGCUGGAGUGCAAUGGC – 3' 3' – GAGUCCGACCUCACGUCACCA – 5'	*MAZ*, miR-3960, CDS, -118, 93 5' – CCCCCGCCUCCGCCGCCACU – 3' 3' – GGGGGCGGAGGCGGCGGCGG – 5'
ADAM17, miR-619-5p, 3466, 3'UTR,-121,100 5' – CCCAGGCUGGAGUGCAGUGGU – 3' 3' – GAGUCCGACCUCACGUCACCA – 5'	*ERBB3*, miR-619-5p, 5104, 3'UTR, -121, 100 5' – GGCUCAUGCCUGUAAUCCCAGC – 3' 3' – CCGAGUACGGACAUUAGGGUCG – 5'
Luminal A,B subtype	
HMGA2, miR-3960, 512, 5'UTR, -108, 86 5' – CCUCCACCUCCACCGCCACC – 3' 3' – GGGGGCGGAGGCGGCGGCGG – 5'	*MAPT*, miR-6756-3p, 3'UTR, -98, 85 5' – CUGGGCAGAGGGGAGAGGAA – 3' 3' – GACCCGUCCCUCCUUCCCCU – 5'
MCM7, miR-4433b-5p, 248, 5'UTR, -100, 85 5' – GCGGGAGCGGGGGUGGGGUGC – 3' 3' – UGUCCUCACCCCCACCCUGUA – 5'	*MCM7*, miR-670-3p, 2769, CDS, -89, 86 5' – CUCUGGAUGAAUAUGAGGAGC – 3' 3' – AGGACUUACUUAUACUCCUUU – 5'
Notes. Gene; miRNA; the beginning of binding site; the miRNA region; the free energy change (ΔG, kJ/mole); the ΔG/ΔGm (%); length of miRNA (nt).	

Table 2 shows the interaction characteristics of some miRNAs with mRNAs of genes. A high degree of homology of nucleotide sequences of miRNAs in binding sites of mRNAs of different candidate breast cancer genes is seen from the schemes. This can serve as an example of associations for use in the diagnosis of breast cancer.

We discovered a new property of miRNA-mRNA interaction: several miRNA binds in a short region of mRNA sequence that contains binding sites for these miRNAs.

Using MirTarget program, it was found that ten miRNAs can bind in the 5'UTR of mRNA in a region from 110 nucleotides (nt) to 148 nt (table 3). The beginning of miR-9-25082-3p and miR-1-1819-3p binding sites coincides and locates from 110 nt. The free energy of interaction of these miRNAs is equal to -121 kJ/mole and -123 kJ/mole, with ΔG/ΔGm equal to 85% and 86%, respectively. From 112 nt, miR-9-20317-3p and miR-X-48174-3p binding sites started, which interacted with mRNA with a value of free energy of interaction -129 kJ/mole and -121 kJ/mole, with ΔG/ΔGm equal to 87% and 85%, respectively. miR-17-39416-3p binding site starts from 113 nt and in this miRNA 92% of nucleotides are complementary to mRNA, with a free interaction energy of -121 kJ/mole. Binding sites for next pair of miR-5-15733-3p and miR-7-20203-3p are located started from 115 nt. The free energy of interaction of these miRNAs is -127 kJ/mole and -121 kJ/mole, with the value of ΔG/ΔGm equal to 86% and 90%, respectively. The miR-9-27797-5p has two binding sites in the 5'UTR of mRNA at 118 nt and 124 nt positions. The free energy of interaction of this miRNA is -121 kJ/mole and -127 kJ/mole, with ΔG/ΔGm equal to 85% and 90%. The presence of two miR-9-27797-5p binding sites provides for it an increased probability of interaction with mRNA of *MMP2* gene. miR-12-17092-3p and miR-9-24743-3p binding sites are located from 124 nt and from 125 nt. ΔG/ΔGm value is 89%, and a free energy of miR-12-17092-3p and miR-9-24743-3p interaction with mRNA is -123 kJ/mole and -127 kJ/mole.

Table 3 – Characteristics of miRNAs interaction in the 5'UTR mRNA of *MMP2* gene

miRNA	Position, nt	ΔG, kJ/mole	ΔG/ΔGm, %	Length, nt
miR-9-25082-3p	110	-121	85	24
miR-1-1819-3p	110	-123	89	23
miR-9-20317-3p	112	-129	87	24
miR-X-48174-3p	112	-121	85	24
miR-17-39416-3p	113	-121	92	22
miR-5-15733-3p	115	-127	86	24
miR-7-20203-3p	115	-121	90	22
miR-9-27797-5p	118	-121	85	24
miR-9-27797-5p	124	-127	90	24
miR-12-17092-3p	124	-123	89	22
miR-9-24743-3p	125	-127	89	23

At the 5'UTR length of 312 nt, binding sites of ten miRNAs are located compactly in a region of 38 nt. Such a compact arrangement of binding sites of several miRNAs facilitates their preservation in the process of evolution. Overlap of miRNA binding sites nucleotide sequences suggests their competition at inhibition of mRNA translation, since one miRNA in the RISC complex interferes with interaction of remaining miRNAs with this site. As a result, the control of mRNA translation is reliably ensured by several miRNAs, which seems to be necessary to suppress the increased synthesis of MM2 proteinase. It should be noted that the location of translation inhibitory miRNA binding sites in the 5'UTR allows cell to save energy on abortive proteins synthesis comparing with miRNA binding occurring in protein coding region or in the 3'UTR with protein synthesis interrupting in these regions.

Binding sites for many miRNAs have been identified in mRNA of *ZFHX3* gene (table 4). Binding sites of miR-15-36707-5p and miR-5-15548-3p are located in the 5'UTR with arranged location of nucleotide sequences and if these miRNAs are present in the cell simultaneously, they will compete for the binding site. The effect of each of the miRNAs will depend on the ratio of their concentrations, and

Table 4 – Characteristics of miRNAs binding sites in the 5'UTR mRNA of *ZFHX3* gene

miRNA	Position, nt	ΔG, kJ/mole	ΔG/ΔGm, %	Length, nt
miR-15-36707-5p	26	-125	88	23
miR-5-15548-3p	31	-123	88	23
miR-19-30988-5p	187	-125	87	23
miR-16-36024-3p	189	-123	87	23
miR-17-39126-5p	195	-123	97	21
miR-4-11437-3p	267	-123	88	23
miR-9-20317-3p	278	-132	89	24
miR-1-1819-3p	285	-125	91	23
miR-2-5973-3p	297	-123	89	24

overall the expression of the *ZFHX3* gene will be determined by the total concentration of miR-15-36707-5p and miR-5-15548-3p, since they have close free interaction energies (ΔG are equal to -125 kJ/mole and -123 kJ/mole, respectively) with mRNA of *ZFHX3* gene.

Binding sites of miR-19-30988-5p, miR-16-36024-3p and miR-17-39126-5p also form a clusters with arranged location of nucleotide sequences (table 4). Binding sites of miR-4-11437-3p, miR-9-20317-3p, miR-1-1819-3p and miR-2-5973-3p form another cluster of multiple binding sites. For these two multi-sites, the same reasoning applies as for miR-15-36707-5p and miR-5-15548-3p which binding sites located in front of them. In general, the expression of *ZFHX3* gene will depend on nine miRNAs that bind in the 5'UTR.

A cluster of miRNA binding sites was revealed in the CDS of mRNA of *ALK* gene (table 5). The miRNA binding sites from 3,387 nt to 3,424 nt are formed a cluster. In the site with length of 37 nt there are binding sites for nine miRNAs: miR-1281, miR-11-29785, miR-13-35476-3p, miR-17-39011-3p, miR-7-20459-3p, miR-9-25099-3p, miR-6792-3p, miR-1-2802-3p, miR-22-40302-3p, miR-X-48174-3p. There are 4-5 binding sites for miR-1281 from 3389 nt to 3421 nt, five binding sites for miR-9-25099-3p from 3387 nt to 3421 nt, three binding sites for miR-11-29785 from 3391 nt to 3425 nt, two binding sites for miR-13-35476-3p from 3394 nt to 3420 nt, three binding sites for miR-7-20459-3p from 3395 nt to 3424 nt. The oligopeptide EWAGGGGGGGGA is conserved in the human ALK protein and 54 animal species, including rat, mouse and rabbit. The mRNA nucleotide sequences adjacent to the binding sites of nine studied miRNAs are variable, which are reflected in the variability of amino acids of flanking oligopeptide EWAGGGGGGGGA.

Table 5 – Characteristics of miRNA interaction with CDS mRNA of *ALK* gene associated with the development of NSCLC

miRNA	Position, nt	ΔG, kJ/mole	ΔG/ΔGm, %	Length, nt
miR-9-25099-3p	3387 ÷ 3399 (5)	-104 ÷ -108	82 ÷ 85	22
miR-17-39011-3p	3388 ÷ 3394 (2)	-110 ÷ -113	84 ÷ 85	23
miR-11-29785	3391 ÷ 3404 (3)	-102 ÷ -106	86 ÷ 89	21
miR-6792-3p	3391	-110	90	22
miR-1-2802-3p	3395	-113	90	22
miR-13-35476-3p	3394 ÷ 3398 (2)	-110	85	22
miR-22-40302-3p	3395	-117	89	22
miR-7-20459-3p	3395 ÷ 3404 (3)	-98 ÷ -102	82 ÷ 86	20
miR-X-48174-3p	3394	-125	88	24

Cluster of miRNA binding sites were identified in the 3'UTR mRNA of *FOXP1* gene (table 6). Binding sites of five miRNAs were located at a region 30 nt in length. The total length of five binding sites is 113 nt, which is almost four times the length of the cluster.

The organization of binding sites in clusters allows gene to significantly reduce its length and preserve dependence on the influence of many miRNAs (tables 3-6).

Table 6 – Characteristics of miRNAs interaction in the 3'UTR mRNA of *FOXP1* gene

miRNA	Position, nt	ΔG, kJ/mole	ΔG/ΔGm, %	Length, nt
miR-10-29282-3p	5952	-104	89	23
miR-10-29282-3p	5970	-104	89	23
miR-19-42814-5p	5953	-106	91	23
miR-19-42814-5p	5955	-104	89	23
miR-6-17605-3p	5960	-110	93	21

in-miRNAs, which co-expressed with the host gene, can be considered as agents realizing the interaction between genes.

In our opinion, the main mechanism of signal transfer from regulator gene to target gene is through miRNA. These molecules are co-expressed together with the host gene (regulator gene) and directly effect on target gene expression at the translation level. Thus, rapid signal transmission within the cell, between cells and tissues is achieved, since miRNAs are much faster than proteins leaving the cell and circulating in the body interacting with virtually all tissues. Thus, miRNAs can serve as integral regulators of the expression of genes and genomes, depending on their physicochemical properties and host genes. There are some miRNAs, which can effectively regulate the expression of hundreds of genes (unique miRNAs). Probably, therefore, the targets of such miRNAs are predominantly transcription factor genes and genes of signaling systems proteins. Transmission of the signal from regulatory gene to target gene by miRNAs is not limited, since one miRNA can interact with any number of target genes having a binding sites in their mRNAs. It was found that some genes have from one to several dozen binding sites in mRNA. In order to reduce the proportion of these sites in mRNA, the binding sites are clustered. Thus in mRNA region of about 100 nucleotides in length, two to several dozens of binding sites can be located with nucleotide sequence overlap. A great variety of the effectiveness of miRNA-mRNA interaction is achieved because of specificity and selectivity of their interaction, ratio of miRNAs concentrations and concentrations of mRNA relative to miRNA. Competitiveness of miRNA binding to mRNA results in the fact that a more strongly binding miRNA disables the influence of other miRNAs, that is, the effect of regulator genes. Either miRNA presented at a higher concentration competitively eliminates the effect of miRNAs having similar binding characteristics to mRNA of target gene. This effect may change by increased expression of other miRNAs. There are cases of an increase (decrease) in the expression of miRNAs in tens and hundreds of times.

Existing systems of regulation of gene expression suppose their regulation in the cell. Generally, such regulation is represented in form of schemes in which regulator gene (or its product) affects target gene (or its product). This relationship of genes in the regulatory system of genome (gene) expression is not biologically appropriate for a number of reasons. The proteins synthesized in cell, with a few exceptions, do not leave the cell and signaling between genes remains intracellular. Regulatory proteins that left the cell to interact with target genes or target proteins have difficulty penetrating cells containing targets. Therefore, the time of signal transmission of regulator to target is long enough. The transmission of signal from regulatory gene to target gene is limited by ability of proteins to interact with several proteins.

Conclusion. It is shown for the first time that miRNA binding sites can be located in the form of clusters. That is, nucleotide sequences of binding sites of several miRNAs are localized in the mRNA region, which is many times shorter than the sum of nucleotide sequences of all miRNAs. This is achieved by arranged location of miRNAs binding sites with maintaining of high specificity. Such compact localization of miRNA binding sites allows economical using of the nucleotide sequence of mRNA. Since the cluster organization of binding sites is observed in 5'UTR, CDS and 3'UTR, it allows particularly to have such regions that encode not necessarily functionally important oligopeptides of protein in the protein coding region. Because in some cases, binding sites in the CDS encode oligopeptides with significantly different lengths in different species while maintaining a functionally complete protein.

The paper shows that widely used programs for predicting miRNA binding sites in mRNA based on 6-8 nucleotide sites (seed) in miRNA are inadequate, since many false positive sites are predicted. In a comparative aspect, these programs highlight the advantages of the MirTarget program used by us for the prediction of binding sites with quantitative characteristics of the interaction of miRNA with mRNA. A hypothesis of the regulation of expression of genes and genomes involving miRNA is proposed. The role of miRNA as an integrating system for the mutual regulation of gene expression in the cell and in the body is shown.

Authors' contributions. Data for tables 2 were provided by Aisina D.E., for tables 3, 4 and 6 by Kondybayeva A., for table 5 by Yurikova O.Yu. All authors involved in drafting the manuscript, read and approved the final version of the manuscript.

The study was carried out with the financial support of the Ministry of Education and Science of the Republic of Kazakhstan within the framework of grant №0118PK00034.

REFERENCES

[1] Rink C., Khanna S. MicroRNA in ischemic stroke etiology and pathology // Physiol Genomics. 2011. 43(10). P.521-8. doi: 10.1152/physiolgenomics.00158.2010.

[2] Uddin A., Chakraborty S. Role of miRNAs in lung cancer // J Cell Physiol. 2018. doi: 10.1002/jcp.26607.

[3] Chan J.J., Tay Y. Noncoding RNA:RNA Regulatory Networks in Cancer // Int J Mol Sci. 2018. 19(5). P. E1310. doi: 10.3390/ijms19051310.

[4] Long H., Wang X., Chen Y., Wang L., Zhao M., Lu Q. Dysregulation of microRNAs in autoimmune diseases: pathogenesis, biomarkers and potential therapeutic targets // Cancer Lett. 2018. P. S0304-3835(18)30277-5. doi: 10.1016/j.canlet.2018.04.016.

[5] Zhang C., Ji Q., Yang Y., Li Q., Wang Z. Exosome: Function and Role in Cancer Metastasis and Drug Resistance // Technol Cancer Res Treat. 2018. P. 1533033818763450. doi: 10.1177/1533033818763450.

[6] .Zhang X., Hamblin M.H., Yin K.J. Noncoding RNAs and Stroke // Neuroscientist. 2018. P. 1073858418769556. doi: 10.1177/1073858418769556.

[7] Kaur H., Sarmah D., Saraf J., Vats K., Kalia K., Borah A., Yavagal D.R., Dave K.R., Ghosh Z., Bhattacharya P. Noncoding RNAs in ischemic stroke: time to translate // Ann N Y Acad Sci. 2018. doi: 10.1111/nyas.13612.

[8] Wang C., Jing Q. Non-coding RNAs as biomarkers for acute myocardial infarction // Acta Pharmacol Sin. 2018. doi: 10.1038/aps.2017.205.

[9] Balatti V., Pekarsky Y., Croce C.M. Role of the tRNA-Derived Small RNAs in Cancer: New Potential Biomarkers and Target for Therapy // Adv Cancer Res. 2017. N 135. P. 173-187. doi: 10.1016/bs.acr.2017.06.007.

[10] Londina E., Lohera P., Telonisa A.G., Quanna K., et al. Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs // PNAS USA. 2015. 112(10). P. 1106-1115.

[11] Ivashchenko A., Berillo O., Pyrkova A., Niyazova R., Atambayeva S. MiR-3960 binding sites with mRNA of human genes // Bioinformation. 2014. 10(7). P. 423-427.

[12] Bartel D.P. MicroRNAs: genomics, biogenesis, mechanism, and function // Cell. 2004. 116(2). P. 281-297.

[13] Peterson.S.M., Thompson J.A., Ufkin M.L., Sathyanarayana P., Liaw L., B.Congdon. Common features of microRNA target prediction tools // Frontiers in Genetics. 2014. doi: 10.3389/fgene.2014.00023

[14] Atambayeva Sh., Niyazova R., Ivashchenko A., Pyrkova A., Pinsky I., Akimniyazova A., Labeit S. The binding sites of miR-619-5p in the mRNAs of human and orthologous genes // BMC Genomics. 2017. N 18. P. 428. Doi: 10.1186/s12864-017-3811-6.

[15] Niyazova R., Berillo O., Atambayeva Sh., Pyrkova A., Alybayeva A., Ivashchenko A. miR-1322 binding sites in paralogous and orthologous genes // BioMed Research Int. 2015.P.1-7. Doi: dx.doi.org/10.1155/2015/962637.

[16] Ivashchenko A., Berillo O., Pyrkova A., Niyazova R. Binding Sites of miR-1273 Family on the mRNA of Target Genes // BioMed Res Int. 2014. Doi: dx.doi.org/10.1155/2014/620530.

[17] Ivashchenko A., Berillo O., Pyrkova A., Niyazova R., Atambayeva Sh. The Properties of Binding Sites of miR-619-5p, miR-5095, miR-5096, and miR-5585-3p in the mRNAs of Human Genes // BioMed Res Int. 2014. Doi: dx.doi.org/10.1155/2014/720715.

[18] BariA., Orazova S., Ivashchenko A. Mir156- and mir171-binding sites in the protein-coding sequences of several plant genes // BioMed Res Int. 2013. Doi: dx.doi.org/10.1155/2013/307145.

[19] Ivashchenko A.T, Issabekova A.S., Berillo O.A. MiR-1279, miR-548j, miR-548m, and miR-548d-5p binding sites in CDSs of paralogous and orthologous PTPN12, MSH6, and ZEB1 genes // BioMed Res Int. 2013. Doi: dx.doi.org/10.1155/2013/902467.

[20] Stavast C.J., Leenen P.J.M., Erkeland S.J. The Interplay between Critical Transcription Factors and MicroRNAs in the Control of Normal and Malignant Myelopoiesis // Cancer Lett. 2018. P. S0304-3835(18)30271-4. doi: 10.1016/j.canlet.2018.04.010.

[21] de Ronde M.W.J., Ruijter J.M., Lanfear D., Bayes-Genis A., Kok M.G.M., Creemers E.E., Pinto Y.M., Pinto-Sietsma S.J. Practical data handling pipeline improves performance of qPCR-based circulating miRNA measurements // RNA. 2017. 23(5). P. 811-821. doi: 10.1261/rna.059063.116.

[22] Krishna M.S., Aneesh Kumar A., Abdul Jaleel K.A. Time-dependent alterations in mRNA, protein and microRNA during in vitro adipogenesis // Mol Cell Biochem. 2018. doi: 10.1007/s11010-018-3307-y.

[23] Minatel B.C., Martinez V.D., Ng K.W., Sage A.P., Tokar T., Marshall E.A., Anderson C., Enfield K.S.S., Stewart G.L., Reis P.P., Jurisica I., Lam W.L. Large-scale discovery of previously undetected microRNAs specific to human liver // Hum Genomics. 2018. 12(1). P. 16. doi: 10.1186/s40246-018-0148-4.

А. Т. Иващенко*[1], Р. Е. Ниязова[1], Ш. А. Атамбаева[1], А. Ю. Пыркова[1], Д. Е. Айсина[1], О. Ю. Юрикова[1], А. Кондыбаева[2], А. Акимниязова[1], Д. Байжигитова[1], А. А. Большой[3]

[1]Әл-Фараби атындағы ҚазҰУ Биология және биотехнология проблемалары ғылыми-зерттеу институты, Алматы, Қазақстан,
[2]С. Д. Асфендияров атындағы Қазақ ұлттық медицина университеті, Алматы, Қазақстан,
[3]Хайфа университеті, Хайфа, Израиль

miRNA: ЖЕТІСТІКТЕР, ПРОБЛЕМАЛАР, ПЕРСПЕКТИВАЛАР

Аннотация. Кіші молекулалы РНК-ның арасында miRNA-дары маңызды рөл атқарады, олар посттранскрипциялық деңгейде гендер экспрессиясын реттейді. Мақалада miRNA-дың қасиеттері мен олардың mRNA-мен өзара әрекеттестігі талқыланады. miRNA-дың түрлі метаболизм процестеріне қатысатын белок-кодтайтын гендердің экспрессиясын реттеуде және жүрек-қан тамырлар, онкологиялық және нейродегенеративті аурулардың дамуындағы рөлі көрсетілген. miRNA-дың нысана гендердің mRNA 5'UTR, CDS және 3'UTR-да байланыстыру сайттарының ерекшеліктері анықталды. mRNA-мен miRNA байланыстыратын сайттарды іздеу белгілі бағдарламаларымен салыстырғанда MirTarget бағдарламасының артықшылықтары көрсетілген. Түрлі аурулардың көптеген кандидатты гендерінің mRNA-ның 5'UTR, CDS және 3'UTR-да бірыңғай miRNA байланысу сайттары және miRNA байланыстыру сайттар кластерлері анықталады. Транскрипция факторларының гендерінің mRNA-ның белоктарындағы олигопептидтерді кодтайтын miRNA байланыстыру сайттар анықталды. miRNA-ның жүрек-қан тамырлар, онкологиялық және нейродегенеративті ауруларға қатысатын кандидатты гендердің mRNA-мен өзара әрекеттесуі талданды. Бірнеше жүздеген гендердің mRNA-да уникалды miRNA-ның байланыстыру сайттар қасиеттері талқыланды. RISC кешенінде mRNA-мен miRNA-ның өзара әрекеттесудің ерекшеліктері қарастырылды. miRNA хост гендерді қамтитын гендердің өзара әрекеттесуі арқылы ген мен геномның экспрессиясын реттеудегі miRNA рөлі талқыланады. miRNA-ды қамтитын ген мен геномның экспрессиясын реттелуінің гипотезасы ұсынылады. miRNA-ның клетка және ағзадағы гендік экспрессиясын өзара реттеу үшін интеграторлық жүйе ретінде рөлі көрсетілген.

Түйін сөздер: miRNA, mRNA, гендер, байланыстыру сайттар, биоинформатикалық бағдарламалар.

А. Т. Иващенко*[1], Р. Е. Ниязова[1], Ш. А. Атамбаева[1], А. Ю. Пыркова[1], Д. Е. Айсина[1], О. Ю. Юрикова[1], А. Кондыбаева[2], А. Акимниязова[1], Д. Байжигитова[1], А. А. Большой[3]

[1]Научно-исследовательский институт проблем биологии и биотехнологии, Казахский национальный университет им. аль-Фараби, Алматы, Казахстан
[2]Казахский национальный медицинский университет им. С. Д. Асфендиярова, Алматы, Казахстан
[3]Хайфский университет, Хайфа, Израиль

miRNA: ДОСТИЖЕНИЯ, ЗАБЛУЖДЕНИЯ, ПЕРСПЕКТИВЫ

Аннотация. Среди малых RNA важную роль играют miRNA, которые осуществляют регуляцию экспрессии генов на посттранскрипционном уровне. В работе рассмотрены свойства miRNA и их взаимодействие с mRNA. Показана роль miRNA в регуляции экспрессии белок-кодирующих генов участвующих в различных процессах метаболизма и развитии сердечно-сосудистых, онкологических и нейродегенеративных заболеваний. Установлены особенности сайтов связывания miRNA в 5'UTR, CDS и 3'UTR mRNA генов-мишеней. Показано преимущество программы MirTarget перед известными программами поиска сайтов связывания miRNA с mRNA. В mRNA многих кандидатных генов различных заболеваний выявлены одиночные

сайты связывания miRNA и кластеры сайтов связывания miRNA в 5'UTR, CDS и 3'UTR mRNA. В mRNA генов транскрипционных факторов обнаружены полисайты связывания miRNA которые кодируют олигопептиды в составе белков. Анализируются взаимодействие miRNA с mRNA кандидатных генов, участвующих в сердечно-сосудистых, онкологических и нейродегенеративных заболеваниях. Обсуждаются свойства уникальных miRNA имеющих сайты связывания в mRNA нескольких сот генов. Рассмотрены особенности взаимодействия mRNA с miRNA в составе комплекса RISC. Обсуждается роль miRNA в регуляции экспрессии генов и генома посредством взаимодействия генов с участием miRNA хозяйских генов. Предложена гипотеза регуляции экспрессии генов и геномов с участием miRNA. Показана роль miRNA как интегрирующей системы взаиморегуляции экспрессии генов в клетке и организме.

Ключевые слова: miRNA, mRNA, гены, сайты связывания, биоинформатика.

Information about authors:

Ivashchenko Anatoliy Timofeevich – d.b.s., professor, al-Farabi Kazakh National University (Faculty of Biology and Biotechnology), Research Institute of Problems of Biology and Biotechnology, E-mail: a_ivashchenko@mail.ru.

Niyazova Raigul Yesengeldievna – c.b.s., professor, al-Farabi Kazakh National University (Faculty of Biology and Biotechnology), Research Institute of Problems of Biology and Biotechnology, E-mail: raygul.nyiyazova@kaznu.kz.

Atambayeva Shara Alpysbayevna – c.b.s., docent, al-Farabi Kazakh National University (Faculty of Biology and Biotechnology), Research Institute of Problems of Biology and Biotechnology, E-mail: shara.atambaeva@kaznu.kz.

Pyrkova Anna Yurievna – c.ph.-m.s., docent, al-Farabi Kazakh National University (Phizico-Technical Faculty), Research Institute of Problems of Biology and Biotechnology, E-mail: Anna.Pyrkova@kaznu.kz.

Aisina Dana Evgenyevna – PhD student, al-Farabi Kazakh National University (Faculty of Biology and Biotechnology), Research Institute of Problems of Biology and Biotechnology, E-mail: dana.aisina03@gmail.com.

Yurikova Oksana Yuryevna – PhD student, al-Farabi Kazakh National University (Faculty of Biology and Biotechnology), Research Institute of Problems of Biology and Biotechnology, E-mail: oksasha1992@gmail.com.

Aida Kondybayeva - PhD student, S.D. Asfendiyarov Kazakh National Medical University, E-mail: dr.kondybayeva@gmail.com

Akimniyazova Aigul – Master degree student, al-Farabi Kazakh National University (Faculty of Biology and Biotechnology), Research Institute of Problems of Biology and Biotechnology, E-mail: 401052@mail.ru.

Bayzhigitova Diana – Master degree student, al-Farabi Kazakh National University (Faculty of Biology and Biotechnology), Research Institute of Problems of Biology and Biotechnology, E-mail: dianabay@mail.ru.

Bolshoy A.A. – PhD, associate professor, Department of Evolutionary and Environmental Biology, University of Haifa, Israel, Haifa, E-mail: bolshoy@research.haifa.ac.il