International Journal of Biological, Life and Agricultural Sciences
ISSN: 2415-6612
Vol:5, No:11, 2011

Characteristics of Intronic and Intergenic Human miRNAs and Features of their Interaction with mRNA

Assel S. Issabekova, Olga A. Berillo, Vladimir A. Khailenko, Shara A. Atambayeva, Mireille Regnier, Anatoly T. Ivachshenko

***Abstract*—**Regulatory relationships of 686 intronic miRNA and 784 intergenic miRNAs with mRNAs of 51 intronic miRNA coding genes were established. Interaction features of studied miRNAs with 5'UTR, CDS and 3'UTR of mRNA of each gene were revealed. Functional regions of mRNA were shown to be significantly heterogenous according to the number of binding sites of miRNA and to the location density of these sites.

***Keywords*—**5'UTR, 3'UTR, CDS, miRNA, target mRNA

I. Introduction

AFTER discovery of miRNAs in the genome of nematode these regulators of protein-coding genes expression have been studied intensively and effectively [1]. miRNAs were found in almost all living systems: viruses, bacteria, unicellular eukaryotes, higher plants and animals [2]. The list of biological phenomena where miRNAs are involved is constantly widening: it was shown to participate in the development of organisms, differentiation, metabolism, biological rythms, proliferation, apoptosis and stress [3-6]. Association of miRNA expression with such diseases as diabetes, hepatitis, cardio-vascular disorders and etc. were revealed in studies. Significant role of miRNA in the protection against infection and pathogenicity of the infectious agents has been established [7-8]. Moreover, miRNA was shown to be involved in the development of cancer of esophagus, stomach, colon, breast, lungs, prostate, pancreatic gland, ovary and etc. [9, 10]. Historically, interest in studying miRNA was due to its involvement in expression of genes on post-transcriptional stage by affecting 3'-untranslated region (3'UTR) of mRNA.

A. S. Issabekova is with the Biology and Biotechnology Department, al-Farabi Kazakh National University, 050040, Almaty, Kazakhstan (e-mail: venerah07@mail.ru).
O. A. Berillo is with the Biology and Biotechnology Department, al-Farabi Kazakh National University, 050040, Almaty, Kazakhstan (e-mail: devolia18@mail.ru).
V. A. Khailenko is with SRI Problems of Biology and Biotechnology, al-Farabi Kazakh National University, 050040, Almaty, Kazakhstan (e-mail: malaheenee@gmail.com).
S. A. Atambayeba is with the Biology and Biotechnology Department, al-Farabi Kazakh National University, 050040, Almaty, Kazakhstan (e-mail: shara.atambaeva@kaznu.kz).
M. Ragnier is with INRIA, Paris, France (e-mail: mireille.ragnier@inria.fr).
A. T. Ivachshenko is with the Biology and Biotechnology Department, al-Farabi Kazakh National University, 050040, Almaty, Kazakhstan (e-mail: a.ivashchenko@mail.ru).

The majority of binding-sites searching programs has been directed to predicting sites in this part of mRNA. However, binding sites of miRNA with mRNA were also found in 5'-untranslated regions (5'UTR) [12-13] and in protein-coding sequences (CDS) of mRNA [14]. Despite being revealed, studies on effect of miRNAs on these sites are scarce but the results of these studies explicitly demonstrate effective binding of miRNA with mRNA on 5'UTR and CDS. Another feature of miRNA is that it is encoded in different regions of DNA. According to their origin, miRNA can be classified to several types. There are intergenic miRNAs (ig-miRNAs) that are encoded in intergenic regions of DNA. Primary transcript of miRNA (pri-miRNA) has a sequence that includes one or several pre-miRNAs that are processed to form separate molecules. Further, double stranded sequence with 2 nucleotides flanking on its 3' and 5' ends is formed in cytoplasm by the cleavage of pre-miRNA. One of these strands becomes active after forming RISC (RNA-induced silencing complex) and binds to target mRNA. Depending on the degree of miRNA-mRNA interaction during binding of RISC to mRNA and some other characteristics, this formed associate leads either to blocking of protein synthesis or to the cleavage of mRNA.

Intronic miRNA (in-miRNA) are encoded in introns of protein coding genes and are either transcribed from DNA autonomously from transcription of pre-mRNA in the presence of its own promoter, or are cleaved from mRNA as pre-miRNA. Further processing of pre-miRNA is the same as for ig-miRNA.

Current work studied characteristics of human miRNA and their features of interaction with 5'UTR, CDS, 3'UTR. mRNAs of genes encoding in-miRNA are selected as targets for in-miRNA and ig-miRNA. These genes are interesting for two reasons; firstly, as a target for all types of miRNA, secondly, as targets for in-miRNA encoded by them. For example, the key enzyme of miRNA biogenesis Diser has three binding sites for let-7 (miRNA) and because it is involved in the last stage of miRNA processing, there is a feedback mechanism of their production [13]. Prevalent number of studies on miRNA was accomplished to reveal the interaction of separate miRNAs with a specific mRNA and also on their related expression. We tried to reveal consistent pattern and features of interaction of 1450 miRNAs (686 in-miRNA and 784 ig-miRNAs) and 51 mRNA genes encoding miRNAs.

II. Materials and Methods

mRNA nucleotide sequences datasets of the genes containing miRNA in introns are accessed at GenBank (http://www.ncbi.nlm.nih.gov) build 37.2. miRNA and their pre-miRNA nucleotide sequences are accessed at miRBase database (http://www.mirbase.org). The program Finder 2.2 (http://sites.google.com/site/malaheenee/software/mirnafinder) has been developed to find in-miRNAs. RNAHybrid 2.1 was used for computing the value of hybridization free energy (ΔG) (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/). Sites with the greatest ΔG value were selected relatively to ΔG value of binding of miRNA with completely complementary site of mRNA. Binding sites of miRNA and target mRNA were determined based on ΔG value and its standard deviation from each site with significance of $p < 0.002$. Density of sites in mRNA, 5'UTR, CDS and 3'UTR were calculated as a number of sites divided on length of these regions and multiplied to 10^3 (s/l).

III. Results And Discussion

Intronic and intergenic miRNAs that interact with corresponding genes were revealed for each gene using the developed program. All genes were grouped according to the features of interaction of 5'UTR of mRNAs with in-miRNAs and ig-miRNAs. All genes having density of binding sites in 5'UTR higher than density of binding sites in gene are referred to genes with high affinity to 5'UTR. The average density of binding sites of in-miRNA and ig-miRNA were 18.3, 41.5, 18.1 and 15.2 s/l for mRNA, 5'UTR, CDS and 3'UTR respectively. The highest number of sites in 5'UTR is 162.5 s/l for AATK gene. The average densities of binding sites of in-miRNAs with mRNAs and their 5'UTR, CDS and 3'UTR were 7.4, 14.9, 6.8 and 7.4 s/l respectively. The average densities of binding sites of ig-miRNA with whole of mRNA and their 5'UTR, CDS and 3'UTR were 10.6, 26.6, 11.4 and 7.8 s/l respectively.

mRNAs of different genes are significantly different according to the number of binding sites with miRNA. For in-miRNA the average numbers of sites on mRNA of gene ranged from 1.4 s/l *(TNFAIP6)* to 21.9 s/l *(BBC3)* sites with compared length of 1423 and 1827 nucleotides correspondingly. For ig-miRNA the average number of sites of mRNA of the same genes was 2.1 and 44.3 s/l respectively. Therefore, expression of *BBC3* gene is under strong control of miRNA, while miRNA is less involved in regulation of *TNFAIP6* gene. Average data of mRNAs of genes with increased density of binding sites in their 5' UTR with in-miRNAs and ig-miRNAs were shown in Table 1. mRNAs of *BRE, EPCAM, EPHB2, HDAC4, MAP7D2, PTPRJ, SLIT2, SLIT3* genes have high conjunction between number of sites for in-miRNAs and ig-miRNAs in 5'UTR, CDS and 3'UTR of mRNA. Correlation coefficient between number of binding equals $r=0.9$ and $p<1.9e-10$. Therefore, similarity of mRNA of these genes according to their affinity to in-miRNAs and ig-miRNAs is not random. Some of genes with high affinity to 5'UTR have increased density of binding sites in CDS and 3'UTR (*AATK, DNMT3A, EGFL7, GIPR, HNF4A, IGF2, LFN6, LRP1, MCM7*). *TNFAIP6* gene has a low binding sites number but the most of them localized in 5'UTR. mRNAs of some studied genes (*CDH13*, *EIF4H, EVL, IGF1R, MRE11A, NOTCH1, TNKS*) do not have any sites for in-miRNAs and ig-miRNAs in their 5'UTR (Table 2). Probably, this significant difference of affinity of in-miRNA to 5'UTR as compared to CDS and 3'UTR is because of biological function of these genes.

ABCA6, ATF2, *CCAR1, PRKG1, TNFAIP6* genes do not have binding sites with in-miRNAs and ig-miRNAs in their 3'UTR can also be attributed to non-random phenomenon. Results of studies suggested that widely recognized notion that miRNA affects mainly 3'UTR and that all regulatory effects of these miRNAs are explained from this point of view are not correct. mRNAs that have a low number of binding sites with in-miRNAs and ig-miRNAs in 5'UTR were revealed and average data for all miRNA were presented in Table 3.

The highest density of binding sites of miRNAs with mRNAs is typical for 5'UTR. Given biological role of miRNA as a regulator of translation, then it is preferential to regulate translation in 5'UTR. Because it is faster and more energy efficient to block translation in the beginning of the process without binding to ribosome. Interaction of miRNA with CDS suggests cessation of elongation of protein when reaching their point of interaction. When miRNA binds to 3'UTR, protein can be synthesized fully or can be blocked on the last stages of elongation in case when RISC complex will prevent this.

Difference in binding features of miRNA with mRNA of different genes discussed above suggests of non-random selection of determined regions of mRNA for interaction with miRNA. It can be suggested that in the evolution process, genes acquired such a 3D configuration that allows their 5'UTR, CDS and 3'UTR to interact with miRNAs.

It should be noted that some of the in-miRNAs have preference to interact with 5'UTR. For example, miR-1268b interacts with 5'UTR of mRNAs of 11 genes, miR-4486 – with 7; miR-1228* - with 6; miR-1908, miR-3173, miR-3178 and miR-4258 – with 5 genes.

Some mRNAs have regions with high density of miRNA binding. For example, 3'UTR of mRNA of *AAKT* gene in the region that begins with 3089-3095 nt has 8 sites, 5'UTR of mRNA of *MCM2* gene with the region that begins 245-262 nt - 7 sites, while 5'UTR of mRNA of *HDAC4* gene has 4 regions for binding of 4-7 miRNAs.

miRNAs vary by the quantity of binding sites with mRNAs of different genes. Among in-miRNAs, miR-1268b has the highest content of binding sites with mRNAs of 51 genes – 47 sites which localized in mRNA of 19 genes and from that *AAKT, HDACY* and *NOTCH1* genes have 5 sites. miR-4308 and miR-4486 have 26 sites respectively in mRNA of 21 and 18 genes. The majority of miRNAs have one binding site with mRNA of one gene.

The most of the ig-miRNAs as well as in-miRNAs, interact with some of mRNAs in several sites. For example, CDS of

TABLE I
BINDING CHARACTERISTICS OF MIRNAS WITH MRNAS OF GENES WITH HIGH AFFINITY OF THEIR 5'UTR TO MIRNAS

Gene	Gene length, nt	5'UTR length, nt	CDS length, nt	3'UTR length, nt	Sites/ gene	Sites/ 5'UTR	Sites/ CDS	Sites/ 3'UTR
AATK	5312	80	4125	1107	49.3	163	50.9	35.2
ABCF1	3464	95	2538	831	10.7	21.3	12.2	4.8
ANTXR1	5893	356	1695	3842	6.8	14.1	12.4	3.6
ATF2	2111	262	1518	331	9.0	22.9	8.6	0.0
BCAS1	3475	338	1755	1382	12.4	20.8	13.1	9.4
BID	2490	324	726	1440	14.5	24.7	22.0	8.3
BIRC6	15703	134	14574	995	6.8	22.5	7.6	1.0
BRE	1891	177	1248	466	19.0	62.3	15.2	12.9
CCAR1	3858	119	3453	286	8.3	8.4	9.0	0.0
DMD	13993	244	11058	2691	6.3	12.3	7.1	2.2
DNMT3A	4380	338	2739	1303	14.6	71.0	29.2	8.4
DTL	4412	314	2193	1905	6.6	19.1	5.5	5.8
EBF3	4361	59	1656	2646	10.1	68.1	12.7	7.2
EGFL7	1529	315	822	392	45.8	66.8	35.3	51.0
EPCAM	1718	358	945	415	19.8	78.3	5.3	2.4
EPHB2	4866	145	2961	1760	22.0	103.8	16.9	23.9
ERBB4	11923	98	3927	7898	6.0	71.5	8.2	4.1
FBXW7	3896	149	2124	1623	5.9	6.7	7.5	3.7
FOXP1	6222	526	2034	3662	13.2	49.5	15.7	6.6
GIPR	2024	99	1401	524	29.6	40.8	35.0	13.4
HDAC4	8976	792	3255	4929	31.6	132.6	31.0	15.8
HNF4A	1600	89	1254	257	29.4	123.9	18.4	50.6
HUWE1	14735	402	13125	1207	14.2	32.3	13.9	11.6
IGF2	5165	752	543	3870	30.0	37.2	14.7	30.8
LFNG	2374	17	1140	1217	33.7	58.8	38.6	28.8
LRP1	14905	466	13635	804	23.0	57.9	20.8	38.5
NR2F2	5110	1224	1244	2642	15.3	35.1	20.1	3.8
MCM7	2900	1117	1632	151	21.7	28.7	19.0	32.9
MAP2K4	3743	69	1200	2475	9.4	58.0	14.2	5.7
MAP7D2	4151	117	2321	1713	14.9	68.4	21.5	2.3
MTUS1	6419	474	3813	2132	8.0	25.4	6.6	6.6
PRKG1	3824	1614	2016	194	6.3	8.7	4.5	0.0
PTK2	4453	230	3159	1064	6.7	34.8	6.7	0.9
PTPRJ	7849	355	4013	3481	13.8	109.9	10.0	8.0
SDCCAG8	2604	156	2141	307	12.7	51.3	10.3	9.8
SLIT2	4950	204	4589	157	10.5	58.8	8.7	0.0
SLIT3	5380	420	4571	389	29.0	116.7	22.5	10.3
SPATA13	8457	322	3833	4302	16.3	46.6	21.7	9.3
TNFAIP6	1423	76	834	513	2.1	13.2	4.8	0.0
Average	5414	354	3398	1161	16.5	52.5	16.3	12.0

TABLE II
BINDING CHARACTERISTICS OF MIRNAS WITH MRNAS OF GENES WITHOUT AFFINITY OF THEIR 5'UTR TO MIRNAS

Gene	Gene length, nt	5'UTR length, nt	CDS length, nt	3'UTR length, nt	Sites/ gene	Sites/ 5'UTR	Sites/ CDS	Sites/ 3'UTR
ABCA6	5296	175	4854	267	4.2	0.0	4.5	0.0
CDH13	3828	119	2141	1566	9.4	0.0	11.2	7.7
EIF4H	2546	8	746	1791	17.3	0.0	8.0	21.2
EVL	1842	87	1257	498	25.0	0.0	27.8	22.1
IGF1R	11242	50	4104	7088	15.0	0.0	13.4	16.1
MRE11A	5141	189	2127	2825	5.3	0.0	8.5	3.2
NOTCH1	9294	0	7669	1625	25.2	0.0	26.5	18.5
TNKS	9599	5	3984	5610	8.9	0.0	17.1	3.0
Average	6111	78	3337	2697	13.8	0.0	14.6	11.5

TABLE III
BINDING CHARACTERISTICS OF MIRNAS WITH MRNAS OF GENES WITH LOW AFFINITY OF THEIR 5'UTR TO MIRNAS

Gene	Gene length, nt	5'UTR length, nt	CDS length, nt	3'UTR length, nt	Sites/ gene	Sites/ 5'UTR	Sites/ CDS	Sites/ 3'UTR
AKT2	5263	262	1446	3555	33.6	19.1	22.1	39.4
BBC3	1827	164	786	877	66.2	6.1	72.5	71.8
BIRC7	1322	173	897	252	40.9	28.9	46.8	27.8
LRRC4	3707	1608	1962	138	16.5	11.8	16.8	65.2
Average	3030	552	1273	1206	39.3	16.5	39.6	51.1

mRNA of *NOTCH1* gene has five sites of interaction with miR-1268, miR-3665, miR-4266, miR-4455 and miR-4472. 5'UTR of mRNA of *HDAC4* gene has five binding sites for miR-1268, miR-4443, miR-4466, miR-4483 and six sites with miR-4674. miR-4508 has 32 binding sites in 18 genes and from these 15 genes are located in 5'UTR, which is non-random. mRNA of *AAKT* gene has 5 binding sites for ig-miRNAs: miR-1268, miR-4266, miR-4455, miR-4466, miR-4472 and miR-4492.The number of studied ig-miRNAs (784) is 1.14 times more than in-miRNAs (686). However, the average number of sites of ig-miRNAs on mRNA of one gene is 1.61±0.09 more ($p<0.001$) than the average number of sites of in-miRNAs. The number of binding sites of miRNAs with mRNAs suggests of the strength of control of gene expression from the corresponding miRNAs. While the number of genes that are controlled by particular miRNAs suggest functional importance of these miRNAs. Analysis of the effect of in-miRNAs on 51 mRNAs of their coding genes showed that only eight of them can effectively regulate translation of own mRNA. These are miR-558 (*BIRC6*), miR-1228* (*LRP1*), miR-1250 (*AATK)*, miR-1469 (*NR2F2*), miR-2467-3p (*HDAC4*), miR-3182 (*CDH13*), miR-3196 (*BIRC7*), miR-4297 (*EBF3*), genes that they encoding are shown in the brackets. It has to be mentioned that miR-1469 acts on 2 sites, while miR-1228* acts on 4 sites of mRNA encoding genes. From the above mentioned eight miRNAs, miR-558, miR-1228*, miR-1469 and miR-3182 together with miR-1250 act on mRNA of *AATK* gene. This example shows how gene expression can occur with the help of in-miRNA.

IV. Conclusion

Results of current studies suggest of an important role of in-miRNA and ig-miRNA in the post-transcriptional regulation of genes, although this study was performed only with 51 genes. An increased number of miRNA target genes can help to obtain new qualities and laws of gene expression regulation by miRNA in humans.

References

[1] M. Isik, H. C. Korswagen, E. Berezikov, "Expression pattern of intronic microRNA in Caenorhabditis elegans," Silence, vol. 1, pp. 5-14, 2010.
[2] B. Zhang, X. Pan, G. P. Cobb, T.A. Anderson, "Plant microRNA: A small regulatory molecule with big impact," Developmental Biology, vol. 289, pp. 3-16, 2006.
[3] E. Erson, E.M. Petty, "MicroRNAs in development and disease," Clin. Genet., vol. 70, pp. 296-306, 2008.
[4] T. Eskildsen, H. Taipaleenmäki, J. Stenvang, B. M. Abdallah, N. Ditzel, A. Y. Nossent et al., "MicroRNA-138 regulates osteogenic differentiation of human stromal (mesenchymal) stem cells in vivo," PNAS, vol. 108, pp. 6139-6144, 2011.
[5] A. V. Sirotkin, M. Laukova, D. Ovcharenko, P. Brenaut, M. Mlyncek, "Identification of microRNAs controlling human ovarian cell proliferation and apoptosis," J. Cell Physiol, vol. 223, pp. 49-56, 2010.
[6] A. Detzer, C. Engel, W. Wunsche, G. Sczakiel, "Cell stress is related to re-localization of Argonaute 2 and to decreased RNA interference in human cells," Nucleic Acids Research, vol. 39, pp. 2727-2741, 2011.
[7] Y. Zhao, H. Xu, Y. Yao, L. P. Smith, L. Kgosana, J. Green et al., "Critical role of the virus-encoded microRNA-155 ortholog in the induction of Marek's disease lymphomas," PLoS Pathog., vol. 7 (2), e1001305, 2011.
[8] R. K. Jangra, M. Yi, S.M. Lemon, "Regulation of hepatitis C virus translation and infectious virus production by the microRNA miR-122," J. Virol., vol. 84, pp. 6615-6625, 2010.
[9] B. Zhang, X. Pan, T. A. Anderson, "MicroRNA: A new player in stem cells," J. Cell. Physiol., vol. 209, pp. 266-269, 2006.
[10] E. Moussay, K. Wang, J.-H. Cho, K van Moer, S. Pierson, J. Paggetti et al., "MicroRNA as biomarkers and regulators in B-cell chronic lymphocytic leukemia," PNAS, vol. 108, pp. 6573-6578, 2011.
[11] A. Grimson, K. K. Fahr, W. K. Johnston, P. Garrett-Engele, L. P. Lim, D. P. Bartel, "MicroRNA targeting specificity in mammals: determinants beyond seed pairing," Mol. Cell., vol. 27, pp. 91-105, 2007.
[12] F. Grey, R. Tirabassi, H. Meyers, G. Wu, S. McWeeney, L. Hook et al., "A viral microRNA down-regulates multiple cell cycle genes through mRNA 5'UTRs," PloS Pathog., vol. 6, e1000967, 2010.
[13] I. Lee, S. S. Ajay, J. I. Yook, H. S. Kim, S. H. Hong, N. H. Kim et al., "New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites," Genome Research, vol. 19, pp. 1175-1183, 2009.
[14] A. E. Erson, E. M. Petty, "MicroRNAs in development and disease," Clin. Genet, vol. 74, pp. 296-306, 2008.