IRSTI 34.15.23; 34.15.25; 31.27.31; 76.03.31; 76.29.49

[1*]A.N. Akimniyazova, [2]M. Régnier, [1]Sh.A. Atambayeva,
[1]R.E. Niyazova, [1]A.T. Ivashchenko

[1]al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Ecole Polytechnique, Paris, France
[*]e-mail: akimniyazova@gmail.com

Characteristics of miRNA interaction with mRNA of candidate genes of small intestinal cancer

Abstract: miRNAs is a class of small non-coding RNAs that regulate the expression of genes, and associated with approximately all known physiological and pathological processes, especially cancer. Expression of many genes is regulated by binding of miRNA with mRNA, therefore it is required to identify candidate genes, associated with small intestinal cancer and the extent of their interaction with miRNA. To determine the important miRNAs binding sites in genes, involved in the development of small intestinal cancer, MirTarget program was used. The article presents the results of studying the characteristics of the interaction of miRNAs with mRNAs of 40 candidate genes involved in the development of small intestinal cancer, out of which only 27 genes were targets for miRNAs. 135 miRNAs have binding sites at 5'UTR, CDS, and 3'UTR; the average free energy of binding (ΔG) of miRNAs with mRNAs was - 126 kJ/mole, - 119 kJ/mole and - 109 kJ/mole, respectively. 79 associations of miRNAs and mRNAs of genes with a free energy of interaction more than – 125 kJ/mole are recommended for the diagnosis of small intestinal cancer. *ARID1A, ASXL1, KRAS, NF1, PDXP, PTEN* and *SMAD4* genes are characterized as a candidate target genes for miRNAs having binding sites in 5'UTR of mRNA, while *ARID1A, CDKN2B, EGFR, GNAS, MLL2, MSH6* and *PDXP* are characterized as a candidate genes, having miRNAs binding sites in CDS, and *CDKN2B*, *SMAD4* as a candidate genes, having miRNAs binding sites in 3'UTR. Based on the results obtained, groups of miRNA and mRNA associations of candidate genes are recommended for developing methods for early diagnosis of small intestinal cancer.
Key words: miRNA, mRNA, genes, small intestinal cancer, clusters.

Introduction

Gastrointestinal (GI) tract cancer is one of the three most common oncological diseases in the world with a high mortality rate [1; 2]. Contrary to common belief, an overall and cancer-specific survival of patients with small intestinal tumors are not different from those of patients with stomach cancer, and takes the second place among the leading causes of death from oncology diseases all over the world [3-5]. Understanding of genetic events driving the pathogenesis of small intestinal cancer is of critical importance for devision of new strategies aimed to treat this disease. miRNAs play an important role in carcinogenesis. In recent years, the interaction of miRNA with mRNA of genes, responsible for the development of small intestinal cancer, was actively studied. Recently, a change in the expression of miRNA became an important feature of cancer. Various miRNAs can function as tumor suppressors or oncogenes in cancer cells, while disregulation of some miRNAs can contribute to human cancer [6; 7]. An individual miRNA could potentially alter complex cellular processes, such as cell growth, cell cycle, apoptosis and invasion. Identification of specific miRNAs and their target genes, participating in carcinogenesis allows better understanding the mechanism of regulation of genes expression [8].

Berillo O.A. et al. have previously studied the characteristics of intronic miRNAs and features of their interaction with mRNA [9]. The present study is aimed to identify previously undefined miRNAs binding sites in mRNA of genes involved in the development of small intestinal cancer and the clusters of miRNA binding sites and their properties. Studying miRNA binding sites clusters in mRNA of human genes is valuable for identification of the role of these genes and miRNAs in oncogenesis.

 Printed in Kazakhstan

Materials and methods

The information about the role and function of genes participating in the development of small intestinal cancer was taken from the GenBank databases and publications. mRNA nucleotide sequences of the human genes were derived from GenBank (http://www.ncbi.nlm.nih.gov). 40 mRNAs of genes associated with the development of small intestinal cancer were used in the study. The nucleotide sequences of 3,707 miRNAs were taken from Londin et al. [10].

Search for miRNA's target genes was performed by the MirTarget program, created in our laboratory [11]. This program defines the start of miRNA binding sites in mRNA; localization of binding sites in 5'-untranslated region (5'UTR), protein coding region (CDS), and 3'-untranslated region (3'UTR); free energy of interaction (ΔG, kJ/mole) and scheme of miRNA-mRNA nucleotides (nt) interaction. The $\Delta G/\Delta G_m$ (%) ratio was calculated for each binding site, where ΔG_m is equal to the free energy of miRNA interaction with fully complementary nucleotide sequence. miRNA-mRNA binding sites with $\Delta G/\Delta G_m$ ratio higher than 86% were selected. However, this criterion does not include the length of miRNA, on which ΔG energy also varies, depending on the miRNA lengths. Thus, in miRNAs with the same ΔG/ΔGm value, but varying lengths of 17 nt and 25 nt, correspondingly, the energy of binding of mRNA for miRNA with the length of 25 nt was 1.47 times higher than for miRNA with the length of 17 nt. ΔG/ΔGm value leads to the reduction in the number of false-positive miRNAs with a length of less than 20 nt. The position of binding sites is indicated from the first nucleotide of the 5'UTR in mRNA. The unique property of MirTarget program include consideration of nucleotide interaction in miRNA with mRNA of target genes not only between adenine (A) and uracil (U), guanine (G) and cytosine (C), but also between A and C, G and U via single hydrogen bond [12; 13]. The distance between A-C and G-U is equal to distance value between G-C and A-U.

Results and discussion

The search of genes responsible for the development of small intestinal cancer was performed by the fragmented data because there is no available unified database of genes. To create the database of genes, we took as a basis the information available in the NCBI (National Center for Biotechnology Information) and through a search of PubMed. Table 1 presents the information about the candidate genes involved in the development of small intestinal cancer. The list of candidate genes was formed from publications based on laboratory research.

Table 1 – Candidate genes of small intestinal cancer

Gene	PMID	Gene	PMID	Gene	PMID	Gene	PMID
*APC**	29575536	*EGFR*	26892442	*MME**	25759539	*PIK3CA*	28617917
ARID1A	28617917	*ERBB2*	24797764	*MRC1*	26530135	*PMS2**	25029614
*ARID2**	28617917	*FBXW7*	28617917	*MSH2**	25759539	*PROM1*	21064103
ASXL1	28617917	*GNAS*	28617917	*MSH6*	25029614	*PTEN*	28617917
ATM	28617917	*KIT**	18246046	*MUC2**	25759539	*SMAD4*	28617917
*BRAF**	26892442	*KRAS*	26892442	*MUC5AC*	25759539	*SOAT1*	25987131
*CD34**	25264210	*LRP1B*	28617917	*MUC6*	25759539	*SOX9*	28617917
*CDKN2A**	28617917	*MDM2*	28617917	*NF1*	28617917	*TP53*	27546842
CDKN2B	28617917	*MLH1**	25029614	*PDGFRA**	25264210	*UGT1A1**	24114122
CTNNB1	27546842	*MLL2*	28617917	*PDXP*	21297586	*VEGFA*	27546842
Note: * – indicates mRNAs, which are not targets for miRNA with chosen criteria							

Genes that are not targets of miRNAs with $\Delta G/\Delta G_m$ value higher than 86%, show that their expression level is independent of miRNAs. It was found that 33% of 40 candidate genes are not regulated by miRNA, and therefore their expression could not be suppressed.

As a result of the analysis of the binding schemes of miRNA with mRNA of the genes *ASXL1, GNAS* and *LRP1B* complete complementarity of binding sites was revealed (Table 2).

The binding sites TJU_CMC_MD2.ID00061.3p-miR and TJU_CMC_MD2.ID03064.3p-miR in

mRNA of the *ASXL1* gene are located at 5'UTR at a value of $\Delta G/\Delta G_m$ of 100%, the value of ΔG was -138 kJ/mole and -115 kJ/mole, respectively.

The binding site of TJU_CMC_MD2.ID02093.5p-miR, consisting of 22 nt, in mRNA of the *GNAS* gene is located in the CDS with free binding energy of -127 kJ/mole.

Table 3 shows the schemes of miRNAs interaction with mRNAs of genes in 5'UTR, CDS, and 3'UTR. Many publications describe other widely used programs for identification of miRNA binding sites in mRNA of target genes, having binding sites of 6-8 nt long (seed) regions. This opinion has been published in many articles and contains inaccurate conclusions, leading towards the big number of false-positive binding sites. Table 2 and 3 show the advantage of MirTarget program [14].

Table 2 – The scheme and characteristics of the interaction of TJU_CMC_MD2.ID00061.3p-miR, TJU_CMC_MD2.ID03064.3p-miR, TJU_CMC_MD2.ID02093.5p-miR and TJU_CMC_MD2.ID03151.3p-miR with mRNA of *ASXL1, GNAS* and *LRP1B* genes

Gene, miRNA, site, region of mRNA, characteristics of binding	Gene, miRNA, site, region of mRNA, characteristics of binding

ASXL1, TJU_CMC_MD2.ID00061.3p-miR,
340 nt,5'UTR, -138kJ/mole, 100%,22 nt
5' - CCACCGCCGCCGCCGCCCCAGCC - 3'
  |||||||||||||||||| |||
3' - GGUGGCGGCGGCGGCGGGG-CGG - 5'
	GNAS, TJU_CMC_MD2.ID02093.5p-miR,
	1626 nt,CDS, -127 kJ/mole, 100%, 22 nt
	5' - UGACGCCCCAGCCGAUCCCGAC - 3'
	  ||||||||||||||||||||||
	3' - ACUGCGGGGUCGGCUAGGGCUG - 5'

ASXL1, TJU_CMC_MD2.ID03064.3p-miR,
348 nt,5'UTR,-115kJ/mole,100%, 18nt
5' - GCCGCCGCCCCAGCCCCGC - 3'
  ||||||||||| |||||||
3' - CGGCGGCGGGG-CGGGGCG - 5'
	LRP1B, TJU_CMC_MD2.ID03151.3p-miR,
	893 nt,5'UTR,-108kJ/mole,100%, 17 nt
	5' - CCCCUGCCGCCGCCGCCG - 3'
	  |||| ||||||||||||
	3' - GGGG-CGGCGGCGGCGGC - 5'

Table 3 – The scheme and characteristics of the interaction of miRNAs with mRNAs of genes participating in the development of small intestinal cancer

Gene, miRNA, site, region of mRNA, characteristics of binding	Gene, miRNA, site, region of mRNA, characteristics of binding

ARID1A, TJU_CMC_MD2.ID00457.3p-miR,
416 nt,CDS, -125 kJ/mole, 92%, 22 nt
5' - **A**ACCCGCCGCCGCCGCCGCCCU - 3'
  || ||||||||||||||||||
*3' – **C**UCGGCGGCGGCGGCGGCGGGA – 5'*
	ASXL1, TJU_CMC_MD2.ID01895.5p-miR,
	400 nt, 5'UTR, -132 kJ/mole, 89%, 24 nt
	5' - CGCG**U**GGAG**CU**GCCGCCGCCGCCG - 3'
	  ||| ||||||||||||||||||
	3' - GCGG**G**CCUC**AG**CGGCGGCGGCGGC – 5'

ATM, TJU_CMC_MD2.ID03006.5p-miR,
9778 nt,3'UTR,-121 kJ/mole, 89%, 24 nt
5' - **C**GGGCUGGGC**G**CAGCGGCUCACGC - 3'
  ||||||||||||| |||||||||
3' - **A**CCCGACCCG**U**GUCCCCGAGUGUA – 5'
	CDKN2B, TJU_CMC_MD2.ID00470.5p-miR,
	1750 nt, 3'UTR,-108 kJ/mole, 89%, 23 nt
	5' - ACACACAC**A**CA**C**ACACAC**AC**ACA - 3'
	  ||||||||||||||||||||||
	3' - UGUGUGUG**C**GU**A**UGUGUG**CA**UGU – 5'

LRP1B, TJU_CMC_MD2.ID02429.3p-miR,
896 nt,5'UTR, -127 kJ/mole, 94%, 23 nt
5' - CU**G**CCGC**C**GCC**G**CCGC**C**GCCGGG - 3'
  ||||||||||||||||||||||
3' - GA**U**GGCG**A**CGG**U**GGCG**A**CGGCCC - 5'
	MUC6, TJU_CMC_MD2.ID01401.3p-miR,
	7883 nt,3'UTR, -108 kJ/mole, 91%, 21 nt
	5' - C**G**CCCC**C**GCUCCCA**G**CCAGGC - 3'
	  ||||||| ||||||||||||
	3' - G**U**GGGG**A**AGAGGGU**U**GGUCCG – 5'

TP53, TJU_CMC_MD2.ID01838.5p-miR,
2460 nt, 3'UTR,-110 kJ/mole,88%, 24 nt
5' - **G**AUGG**G**GUCUCACAGUGUUGC**C**CA - 3'
  || |||||||||| ||||||||||
3' - **U**UUCC**U**CAGAGUGACACAACG**A**GU – 5'
	VEGFA, TJU_CMC_MD2.ID03097.3p-miR,
	887 nt, CDS, -123 kJ/mole, 88%, 24 nt
	5' - **C**UCC**A**GCC**G**CGCGCGCUCC**C**CAGG - 3'
	  |||||| ||||||||||||||
	3' - **A**AGG**C**CCG**U**GCGCGCGAGG**A**GUCC – 5'

In bold type highlighted the non-canonical nucleotide (A-C, G-U) pairs	

The free energy (ΔG) of miRNAs binding sites in different mRNAs varied from -108 to -132 kJ/mole. These changes occurred in connection with several nucleotide substitutions of G-C to G-U pairs.

The results show the study of miRNAs interaction with mRNA of 11 genes in 5'UTR, each of which binds from one to several miRNAs (Table 4). Some of these mRNAs bind six or more miRNAs. mRNA of *ASXL1* gene in a region from 172 nt to 197 nt and form the cluster of this binding site with whole length of 26 nt and average ΔG = -129 kJ/mole. TJU_CMC_MD2.ID00021.5p-miR and TJU_CMC_MD2.ID01895.5p-miR have binding sites from 400 nt, but most likely TJU_CMC_MD2.ID01895.5p-miR will occupy this site cause of the energy of interaction of this miRNA due to the higher concentration of GC-pairs.

mRNAs of most genes containing two or more miRNA binding sites with overlapping of their nucleotide sequences form clusters. mRNA of *ARID1A* gene contain 2 clusters of binding sites located in 5'UTR. TJU_CMC_MD2.ID01257.3p-miR, TJU_CMC_MD2.ID00414.3p-miR and TJU_CMC_MD2.ID02428.3p-miR form a cluster from 295 to 332 nt with a length of 38 nt. The free energy of interaction of these miRNAs average ΔG value is equal to -111 kJ/mole. TJU_CMC_MD2.ID02106.3p-miR, TJU_CMC_MD2.ID01778.3p-miR, TJU_CMC_MD2.ID00296.3p-miR, TJU_CMC_MD2.ID01804.3p-miR, TJU_CMC_MD2.ID01702.3p-miR, TJU_CMC_MD2.ID02592.5p-miR and TJU_CMC_MD2.ID03065.3p-miR form a cluster with a length equal to 58 nt and average ΔG = -142 kJ/mole. TJU_CMC_MD2.ID02751.3p-miR has a binding site from 206 nt and ΔG/ΔGm equal to 92%.

19 miRNAs form a cluster of binding sites in 5'UTR mRNA of *LRP1B* gene with a whole length equal to 36 nt and average free energy of interaction equal to -131 kJ/mole. The full length of this 19 miRNAs 553 nt. The formation of such cluser of binding sites indicates the great ability of this gene to compaction, which serves the formation of competition for this binding site.

mRNA of *GNAS* gene contain the cluster of binding sites from 31 nt to 70 nt with an average free energy of interaction equal to -104 kJ/mole. Three miRNAs formed a cluster of binding sites in mRNA of *PDXP* gene in position from 17 nt to 53 nt with an average free energy of interaction equal to -127 kJ/mole. *KRAS* gene had four binding sites, of which three sites formed a cluster with overlapped nucleotide sequences. mRNA of *CTNNB1, SOAT1* and *SOX9* genes have binding sites for single miRNAs.

PTEN is onco-suppressor gene. Byun et al. [15] observed that PTEN inactivation by deletion alone is sufficient to initiate developing intestinal tumors, including adenocarcinomas. mRNA of *PTEN* gene contain the cluster of binding sites from 531 to 558 nt with a total length of 28 nt and average ΔG = -120 kJ/mole. The binding sites for TJU_CMC_MD2.ID01315.3p-miR and TJU_CMC_MD2.ID01377.3p-miR in position from 705 to 726 nt form the cluster of binding sites with a whole length of 22 nt. That is why suppression of translation of this mRNA provides the development of this disease.

The average free energy of binding of all miRNAs with mRNAs in the 5'UTR region equals -126 kJ/mole. 27 miRNAs bound with mRNAs of corresponding target genes, and the number of miRNA associations with mRNA having free energy of interaction greater than -125 kJ/mole is 57. These associations are recommended as markers for early diagnosis of small intestinal cancer.

Some binding sites are located in overlapping mRNA nucleotide sequences in protein-coding regions (Table 5). Presence of multiple binding sites for one and/or several miRNAs in one mRNA increases the probability of their interaction, and as a consequence, translation of such mRNAs is reduced [16].

When the binding energy of one miRNA site decreases, this can be compensated by other sites. Several binding sites with mRNA can enhance the inhibitory effect.

mRNA of *CDKN2B* (ΔG = -132 kJ/mole), *EGFR* (ΔG = -127 kJ/mole), *ERBB2 (*ΔG = -113 kJ/mole), *FBXW7* (ΔG = -108 kJ/mole), *MSH6* (ΔG = -134 kJ/mole), *PIK3CA* (ΔG = -100 kJ/mole) and *SOX9* (ΔG = -119 kJ/mole) genes are bound by one miRNAs.

Nine miRNAs have binding sites in mRNA of *MLL2* gene. Kantidakis et al. [17] identified that *MLL2* knockdown affects adhesion-related processes and suppresses cell growth. These binding sites are located without overlapping on protein-coding region of mRNA of *MLL2* gene, and we could say that the expression of this gene will be suppressed. The average free energy of interaction of these miRNAs is equal to -114 kJ/mole.

Table 4 – Characteristics of the binding sites of miRNA and mRNA in 5'UTR of genes involved in the development of SIC

Gene	miRNA	Start of binding site, nt	ΔG, kJ/mole	ΔG/ΔGm, %	Length, nt
ARID1A	TJU_CMC_MD2.ID02751.3p-miR	206	-125	92	23
	TJU_CMC_MD2.ID01257.3p-miR	295	-113	93	20
	TJU_CMC_MD2.ID00414.3p-miR	303	-108	93	20
	TJU_CMC_MD2.ID02428.3p-miR	310	-113	91	22
ASXL1	TJU_CMC_MD2.ID00522.5p-miR	172	-125	89	23
	TJU_CMC_MD2.ID01804.3p-miR	173	-138	87	25
	TJU_CMC_MD2.ID02187.5p-miR	174	-123	89	23
	TJU_CMC_MD2.ID00021.5p-miR	400	-117	92	20
	TJU_CMC_MD2.ID01895.5p-miR	400	-132	89	24
CTNNB1	TJU_CMC_MD2.ID00477.5p-miR	77	-113	95	20
GNAS	TJU_CMC_MD2.ID02111.3p-miR	31	-108	91	21
	TJU_CMC_MD2.ID00724.5p-miR	49	-100	92	21
KRAS	TJU_CMC_MD2.ID01310.3p-miR	17÷29 (2)	-121÷ -123	92÷94	22
	TJU_CMC_MD2.ID03332.3p-miR	37	-132	89	24
	TJU_CMC_MD2.ID00805.3p-miR	96	-113	91	21
NF1	TJU_CMC_MD2.ID02352.5p-miR	157	-123	91	24
	TJU_CMC_MD2.ID01787.3p-miR	297	-115	89	23
	TJU_CMC_MD2.ID01491.3p-miR	349	-125	91	23
PDXP	TJU_CMC_MD2.ID01018.3p-miR	17	-125	89	24
	TJU_CMC_MD2.ID01622.3p-miR	23	-132	91	23
	TJU_CMC_MD2.ID00714.3p-miR	29	-123	88	24
PTEN	TJU_CMC_MD2.ID02079.5p-miR	75	-115	92	20
	TJU_CMC_MD2.ID02611.3p-miR	486	-125	91	22
	TJU_CMC_MD2.ID01310.3p-miR	531	-119	90	22
	TJU_CMC_MD2.ID02430.3p-miR	533	-108	96	18
	TJU_CMC_MD2.ID02761.3p-miR	533	-132	89	24
	TJU_CMC_MD2.ID03037.3p-miR	536	-121	90	22
	TJU_CMC_MD2.ID01315.3p-miR	705	-115	92	20
	TJU_CMC_MD2.ID01377.3p-miR	708	-110	96	18
	TJU_CMC_MD2.ID02429.3p-miR	790	-121	89	23
SMAD4	TJU_CMC_MD2.ID00577.3p-miR	160÷161 (2)	-104÷-106	92÷94	20
	TJU_CMC_MD2.ID00961.3p-miR	248	-127	90	23
SOAT1	TJU_CMC_MD2.ID03036.3p-miR	46	-115	89	23
SOX9	TJU_CMC_MD2.ID01969.3p-miR	294	-113	96	20
Note: Here and in the tables below the number of miRNAs binding sites is indicated in parentheses					

Table 5 – Characteristics of the binding sites of miRNA and mRNA in CDS of genes involved in the development of SIC

Gene	miRNA	Start of binding site, nt	ΔG, kJ/mole	ΔG/ΔGm, %	Length, nt
ARID1A	TJU_CMC_MD2.ID02057.3P-MIR	385	-110	95	19
	TJU_CMC_MD2.ID01753.3p-miR	749	-104	91	21
	TJU_CMC_MD2.ID02294.5p-miR	851	-129	88	24
	TJU_CMC_MD2.ID00061.3p-miR	852	-125	91	22
	TJU_CMC_MD2.ID01473.3p-miR	1093	-125	89	23
	TJU_CMC_MD2.ID02986.5p-miR	1387	-117	92	21
	TJU_CMC_MD2.ID03167.3p-miR	1399	-123	91	23
	TJU_CMC_MD2.ID01610.5p-miR	1404	-110	91	21
	TJU_CMC_MD2.ID01508.5p-miR	1459	-129	90	23
	TJU_CMC_MD2.ID01388.5p-miR	4587	-119	89	23
	TJU_CMC_MD2.ID01565.5p-miR	4916	-115	93	21
CDKN2B	TJU_CMC_MD2.ID02899.3p-miR	412	-132	89	24
EGFR	TJU_CMC_MD2.ID02344.3p-miR	1779	-127	88	24
ERBB2	TJU_CMC_MD2.ID00692.3p-miR	3815	-113	90	22
FBXW7	TJU_CMC_MD2.ID02514.3p-miR	1243	-108	93	22
GNAS	TJU_CMC_MD2.ID02093.5p-miR	1589÷1590 (2)	-117÷ -123	92÷97	22
	TJU_CMC_MD2.ID00377.3p-miR	1617	-119	90	22
	TJU_CMC_MD2.ID02093.5p-miR	1625÷1626 (2)	-117÷ -127	92÷100	22
	TJU_CMC_MD2.ID00377.3p-miR	1653	-119	90	22
	TJU_CMC_MD2.ID02093.5p-miR	1662	-117	92	22
	TJU_CMC_MD2.ID03100.3p-miR	1830	-110	91	22
LRP1B	TJU_CMC_MD2.ID02669.5p-miR	14051	-96	92	20
MLL2	TJU_CMC_MD2.ID01491.3p-miR	3300	-125	91	23
	TJU_CMC_MD2.ID01753.3p-miR	5258	-104	91	21
	TJU_CMC_MD2.ID02409.3p-miR	5878	-102	91	21
	TJU_CMC_MD2.ID02562.3p-miR	5967	-115	90	22
	TJU_CMC_MD2.ID02664.5p-miR	7226	-104	92	20
	TJU_CMC_MD2.ID01013.3p-miR	7483	-123	88	24
	TJU_CMC_MD2.ID03483.3p-miR	11514	-115	90	22
	TJU_CMC_MD2.ID01562.3p-miR	12237	-117	90	22
	TJU_CMC_MD2.ID02045.3p-miR	13623	-123	89	23
MSH6	TJU_CMC_MD2.ID00156.5p-miR	246	-134	90	24
PDXP	TJU_CMC_MD2.ID01242.3p-miR	172÷173 (2)	-121÷ -125	88÷91	24
	TJU_CMC_MD2.ID01877.3p-miR	386	-123	88	24
	TJU_CMC_MD2.ID01610.5p-miR	646	-110	91	21
PIK3CA	TJU_CMC_MD2.ID00724.5p-miR	3043	-100	92	21

Continuation of Table 5

Gene	miRNA	Start of binding site, nt	ΔG, kJ/mole	ΔG/ΔGm, %	Length, nt
SOX9	TJU_CMC_MD2.ID00978.5p-miR	1065	-119	90	22
VEGFA	TJU_CMC_MD2.ID03238.3p-miR	775	-113	91	22
	TJU_CMC_MD2.ID03097.3p-miR	887	-123	88	24
MUC5AC	TJU_CMC_MD2.ID00481.5p-miR	668	-113	90	22
	TJU_CMC_MD2.ID02341.3p-miR	2037	-113	90	23
	TJU_CMC_MD2.ID02620.5p-miR	3277	-110	90	22
	TJU_CMC_MD2.ID00967.5p-miR	4672	-115	90	22
	TJU_CMC_MD2.ID00967.5p-miR	5791	-115	90	22
MUC6	TJU_CMC_MD2.ID02841.5p-miR	1226	-108	93	20
	TJU_CMC_MD2.ID00611.5p-miR	1574	-115	89	23
	TJU_CMC_MD2.ID02888.5p-miR	2918	-121	89	23
	TJU_CMC_MD2.ID00564.5p-miR	6488	-110	90	22
	TJU_CMC_MD2.ID02507.3p-miR	7353	-115	89	24

TJU_CMC_MD2.ID00522.5p-miR, TJU_CMC_MD2.ID01804.3p-miR, TJU_CMC_MD2.ID02052.5p-miR, TJU_CMC_MD2.ID02187.5p-miR, TJU_CMC_MD2.ID02692.5p-miR, TJU_CMC_MD2.ID01323.3p-miR, TJU_CMC_MD2.ID00457.3p-miR, TJU_CMC_MD2.ID02084.3p-miR, TJU_CMC_MD2.ID02064.5p-miR and TJU_CMC_MD2.ID02538.3p-miR formed a cluster of binding sites with a whole length equals 33 nt and average ΔG = -128 kJ/mole. The full length of this 10 miRNAs equals 249 nt; TJU_CMC_MD2.ID01704.5p-miR, TJU_CMC_MD2.ID01810.3p-miR, TJU_CMC_MD2.ID02430.3p-miR and TJU_CMC_MD2.ID02761.3p-miR formed cluster in segment with a whole length of 47 nt and average ΔG = -124 kJ/mole; TJU_CMC_MD2.ID02986.5p-miR, TJU_CMC_MD2.ID03167.3p-miR and TJU_CMC_MD2.ID01610.5p-miR form a cluster of binding sites with a whole length of 39 nt and average free energy of interaction equal to -117 kJ/mole. The formation of cluster of binding sites in this mRNA notes the ability of this gene to compaction, and thereby causes the competition of these miRNAs for binding site. Also, mRNA of *ARID1A* gene has the binding sites for single miRNAs. TJU_CMC_MD2.ID02057.3p-miR have a binding site in position from 385 nt. Kim et al. [18] established that the expression level of *ARID1A* could be used as a prognostic marker in small intestinal cancer, because the low or loss of expression of this gene is correlated considerably with a high-grade tumors. Suppression of expression of this gene by miRNA promotes the development of this disease.

mRNA of *GNAS* gene has multiple binding sites of TJU_CMC_MD2.ID02093.5p-miR in position from 1589 nt to 1612 nt with an average ΔG = -120 kJ/mole, and from 1625 nt to 1648 nt with an average ΔG = 122 kJ/mole. Also this mRNA have a cluster of binding sites from 1653 to 1684 nt with a whole length of 32 nt and average ΔG = -118 kJ/mole.

Kumagai et al. [19] indicate that mucin core proteins (*MUC5AC*, *MUC6* and *MUC2*) expressed in the cytoplasm of the tumor cells. 4 miRNAs have binding sites in mRNA of *MUC5AC* gene. All binding sites have the same ΔG/ΔGm value, equal to 90% and relatively equal ΔG value. It is important to mention the significance of free energy of interaction. TJU_CMC_MD2.ID00967.5p-miR with a length of 22 nt has multiple binding sites in *MUC5AC* gene with ΔG = -115 kJ/mole. mRNA of *MUC6* gene have 5 binding sites in CDS. The higher ΔG is observed in association with TJU_CMC_MD2.ID02888.5p-miR (ΔG = -121 kJ/mole).

The average free energy of binding of all miRNAs with mRNAs in the CDS equals -119 kJ/mole. The number of miRNA associations with mRNA having free energy of interaction greater than -125 kJ/mole is 22. All of them can serve as markers in the development of methods for early diagnosis of small intestinal cancer.

Table 6 shows the binding sites of some miRNAs with mRNA genes involved in the development of small intestinal cancer in 3'UTR region.

From 12 target genes in 3'UTR, binding sites for 26 miRNAs were identified (Table 6). mRNA of *SMAD4* gene contained multiple binding sites for TJU_CMC_MD2.ID00470.5p-miR and TJU_CMC_MD2.ID02299.5p-miR in the same region. Moreover, TJU_CMC_MD2.ID00470.5p-miR, which binds to mRNA with free energy of -108 kJ/mole, can have the highest efficiency of regulation of *SMAD4* gene expression. *SMAD4* plays the role of tumor growth suppressor.

Table 6 – Characteristics of the binding sites of miRNA and mRNA in 3'UTR of genes involved in the development of SIC

Gene	miRNA	The beginning of binding site, nt	ΔG, kJ/mole	ΔG/ΔGm, %	Length, nt
ATM	TJU_CMC_MD2.ID03006.5p-miR	9778	-121	89	24
	TJU_CMC_MD2.ID00367.5p-miR	11069	-110	90	22
CDKN2B	TJU_CMC_MD2.ID00470.5p-miR	1746÷1752 (4)	-108	89	23
	TJU_CMC_MD2.ID02299.5p-miR	1746÷1752 (4)	-100	90	22
KRAS	TJU_CMC_MD2.ID03224.5p-miR	3163	-121	92	23
MDM2	TJU_CMC_MD2.ID01404.5p-miR	2154	-110	90	23
	TJU_CMC_MD2.ID01815.5p-miR	2257	-106	89	23
	TJU_CMC_MD2.ID01360.3p-miR	2476	-104	91	21
	TJU_CMC_MD2.ID01838.5p-miR	3214	-110	88	24
MLL2	TJU_CMC_MD2.ID02459.5p-miR	16748	-110	91	21
	TJU_CMC_MD2.ID03342.3p-miR	16931	-119	92	23
	TJU_CMC_MD2.ID00670.3p-miR	16971	-102	91	21
	TJU_CMC_MD2.ID03270.3p-miR	17062	-123	88	25
	TJU_CMC_MD2.ID02547.5p-miR	17065	-113	91	21
	TJU_CMC_MD2.ID01254.5p-miR	17780	-110	91	21
	TJU_CMC_MD2.ID00564.5p-miR	17894	-113	91	22
NF1	TJU_CMC_MD2.ID02733.5p-miR	10571	-89	91	21
PROM1	TJU_CMC_MD2.ID02076.5p-miR	3221	-104	94	19
SMAD4	TJU_CMC_MD2.ID01838.5p-miR	4291	-113	90	24
	TJU_CMC_MD2.ID01404.5p-miR	4349	-115	93	23
	TJU_CMC_MD2.ID02732.3p-miR	7721	-123	91	23
	TJU_CMC_MD2.ID00470.5p-miR	7744÷7752 (5)	-108	89	23
	TJU_CMC_MD2.ID02299.5p-miR	7744÷7752 (5)	-100	90	22
	TJU_CMC_MD2.ID00106.5p-miR	7825	-106	91	22
	TJU_CMC_MD2.ID01592.3p-miR	8227	-117	89	23
SOAT1	TJU_CMC_MD2.ID00868.5p-miR	4885	-123	88	25
	TJU_CMC_MD2.ID01815.5p-miR	4953	-106	89	23
	TJU_CMC_MD2.ID01404.5p-miR	5523	-110	90	23
SOX9	TJU_CMC_MD2.ID00509.3p-miR	2153	-98	92	21
TP53	TJU_CMC_MD2.ID02379.3p-miR	1397	-119	89	24
	TJU_CMC_MD2.ID01838.5p-miR	2459÷2460 (2)	-110÷ -115	88÷92	24
	TJU_CMC_MD2.ID00785.5p-miR	2520	-113	90	23
MUC6	TJU_CMC_MD2.ID01401.3p-miR	7883	-108	91	21

Interestingly, TJU_CMC_MD2.ID00470.5p-miR and TJU_CMC_MD2.ID02299.5p-miR have multiple binding sites in mRNA of *CDKN2B* gene in the same region with ΔG value equal to -108 kJ/mole and -100 kJ/mole, respectively. TJU_CMC_MD2.ID00470.5p-miR will have the advantage of binding for this site. miRNA with higher ΔG value is more prevalent takes this binding site and therefore inhibit the expression of this gene.

TP53 is a transcription factor that regulates the cell cycle and performs the function of onco-suppressor. Binding sites in mRNA of this gene were identified. TJU_CMC_MD2.ID01838.5p-miR has multiple consistently located binding sites with an average ΔG value equals -113 kJ/mole.

The results obtained show that interactions of the examined miRNA and mRNA can serve as the basis for the choice of miRNA and mRNA associations for the diagnosis of small intestinal cancer.

Seven miRNAs have binding sites in mRNA of *MLL2* gene. These binding sites are scattered throughout all 3'UTR of mRNA, and we could say that the expression of this gene will be suppressed. Free energy of interaction of these miRNAs equals -113 kJ/mole.

mRNAs of *KRAS, NF1, PROM1, SOX9, MUC6* genes have binding sites for single miRNAs. This evidence have the positive phenomena, since typically each miRNA has one or several target genes, and conversely, one gene can be a target for one or more miRNA. That is why, these associations can possibly serve as a marker for early diagnosis of this disease.

The average free energy of binding of miRNAs with mRNAs in 3'UTR equals -109 kJ/mole. Two miRNAs (TJU_CMC_MD2.ID00470.5p-miR and TJU_CMC_MD2.ID02299.5p-miR) have multiple binding sites in mRNA of *CDKN2B* and *SMAD4* genes and therefore are also suggested as associations for the diagnosis of small intestinal cancer.

Thus, miRNAs that participate in the formation of fully complementary binding sites have been identified. With this binding, mRNA degradation may occur to a greater extent than its inhibition. There were revealed slightly different binding sites with the same miRNA located in different mRNA target genes. This shows the effect of resistance to point mutations (single nucleotide polymorphism) in miRNA binding site. Revealed closely related miRNAs having similar sequences and capable of binding to the same site with different hybridization energy.

Conclusion

In this paper, we studied the characteristics of the interaction of miRNA with mRNA of genes involved in the development of small intestinal cancer. 79 key associations of miRNAs and mRNAs were identified that have had a free energy of interaction of -125 kJ/mole or more that can be recommended both for the diagnostic markers of violations of the expression of these genes, and for the development of methods for the treatment of these disease at the molecular level.

The average free energy of miRNA binding in mRNA of genes involved in small intestinal cancer development is greater in 5'UTR and CDS compared to 3'UTR, which suggests preferential binding of miRNA to 5'UTR and CDS of the studied genes. *ARID1A*, *ASXL1*, *KRAS*, *NF1*, *PDXP*, *PTEN* and *SMAD4* genes was selected as candidate target genes for miRNAs having binding sites in 5'UTR of mRNA. *ARID1A*, *CDKN2B, EGFR, GNAS, MLL2, MSH6* and *PDXP* are candidate genes, having miRNAs binding sites in CDS, *CDKN2B* and *SMAD4* are candidate genes, having miRNAs binding sites in 3'UTR.

In summary, mRNA of these genes binds with miRNA with high affinity and, therefore, if the concentrations of the corresponding miRNAs are the same or exceed the concentrations of mRNAs, the effect of mRNA translation inhibition will be significant. Given that miRNAs are endogenous regulators of gene expression, these methods are quite feasible. The results obtained may be able to provide insights into the pathogenesis mechanism and pave the way for the development of new diagnostic markers and therapeutic targets for patients with small intestinal cancer.

Acknowledgments

The work was performed within the framework of the grant project AP05132460 "Development of test systems for early diagnosis of cardiovascular, oncological and neurodegenerative diseases based on miRNA associations and their target genes", financed by the Ministry of Education and Science of the Republic of Kazakhstan.

References

1. Syngal S., Brand R.E., Church J.M., Giardiello F. M., Hampel H. L., Burt R.W. (2015) ACG Clinical guideline, genetic testing and management of hereditary gastrointestinal cancer syndromes. *Am J Gastroenterol*., vol. 110, no. 2, pp. 223-262.

2. Torre L.A., Bray F., Siegel R.L., Ferlay J., Lortet-Tieulent J., Jemal A. (2015) Global cancer statistics, 2012. *CA Cancer J Clin*., vol. 65, pp. 87-108.

3. Guller U., Tarantino I., Cerny T., Ulrich A., Schmied B.M., Warschkow R. (2017) Revisiting a dogma: similar survival of patients with small bowel and gastric GIST. A population-based propensity score SEER analysis. *Gastric Cancer*. vol. 20, no. 1, pp. 49-60.

4. Giuliano K., Nagarajan N., Canner J., Najafian A., Wolfgang C., Schneider E., Meyer C., Lennon AM., Johnston F.M., Ahuja N. (2017) Gastric and small intestine gastrointestinal stromal tumors: Do outcomes differ? *J Surg Oncol*., vol. 115, no. 3, pp. 351-357.

5. Ye H., Xin H., Zheng Q., Shen Q., Dai W., Wu F., Zheng C., Chen P. (2018) Prognostic role of the primary tumour site in patients with operable small intestine and gastrointestinal stromal tumours: a large population-based analysis. *Oncotarget*, vol. 9, no. 8, pp. 8147-8154.

6. Lee Y. S., Dutta A. (2009) MicroRNAs in cancer. *Annu Rev Pathol.,* vol. 4, pp. 199-227.

7. Grammatikakis I., Gorospe M., Abdelmohsen K. (2013) Modulation of cancer traits by tumor suppressor microRNAs. *Int J Mol Sci.,* vol. 14, no. 1, pp. 1822-1842.

8. Rath S.N., Das D., Konkimalla V.B., Pradhan S.K. (2016) In silico study of miRNA based gene regulation, involved in solid cancer, by the assistance of Argonaute protein. *Genomics Inform.,* vol. 14, no. 3, pp. 112-124.

9. Berillo O., Régnier M., Ivashchenko A. (2013) Binding of intronic miRNAs with mRNAs of genes coding intronic microRNAs and proteins participating in tumourigenesis. *Comput Biol Med.,* vol. 43, no 10, pp. 1374-1381.

10. Londin E., Lohera P., Telonisa A.G., Quanna K., et al. (2015) Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs. *PNAS USA*, vol. 112, no. 10., pp. 1106-1115.

11. Ivashchenko A., Berillo O., Pyrkova A., Niyazova R., Atambayeva S. (2014) 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*., vol. 2014, pp. 1-8.

12. Kool E.T. (2001) Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annu Rev Biophys Biomol Struct.,* vol. 30, pp. 1-22.

13. Leontis N.B., Stombaugh J., Westhof E. (2002) The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res*., vol. 30, no. 16, pp. 3497-3531.

14. Ivashchenko A. T., Niyazova R. Ye., Atambayeva Sh. A., Pyrkova A. Yu., Aisina D. E., Yurikova O. Yu., Kondybayeva A., Akimniyazova A., Bayzhigitova D., Bolshoy A.A. (2018) miRNA: achievements, misconceptions, perspectives. *News of NAS RK,* vol 4, no. 328, pp. 36-46.

15. Byun D., Ahmed N., Nasser S., Shin J., Al-Obaidi S., Goel S., Corner G., Wilson A., Flanagan D., Williams D., Augenlicht L., Vincan E, Mariadason J. (2011) Intestinal epithelial-specific PTEN inactivation results in tumor formation. *Am J Physiol Gastrointest Liver Physiol*., vol. 301, no.5, pp. 856-864.

16. Niyazova R., Atambayeva Sh., Akimniyazova A., Pinsky I., Alybaeva A., Faye B., Ivashchenko A. (2015) Features of mir-466-3p binding sites in mRNA genes with different functions. *Int J Biol Chem.*, vol. 8, no. 2, pp. 44-51.

17. Kantidakis T., Saponaro M., Mitter R., Horswell S., Kranz A., Boeing S., Aygün O., Kelly G.P., Matthews N., Stewart A., Stewart A.F., Svejstrup J.Q. (2016) Mutation of cancer driver MLL2 results in transcription stress and genome instability. *Genes Dev*., vol.30, no. 4, pp. 408-420.

18. Kim M.J., Gu M.J., Chang H.K., Yu E. (2015) Loss of ARID1A expression is associated with poor prognosis in small intestinal carcinoma. *Histopathology*, vol. 66, no. 4, pp. 508-516.

19. Kumagai R., Kohashi K., Takahashi S., Yamamoto H., Hirahashi M., Taguchi K., Nishiyama K., Oda Y. (2015) Mucinous phenotype and CD10 expression of primary adenocarcinoma of the small intestine. *World J Gastroenterol.,* vol. 21, no. 9, pp. 2700-2710.