Research Article

The Properties of Binding Sites of miR-619-5p, miR-5095, miR-5096, and miR-5585-3p in the mRNAs of Human Genes

Anatoly Ivashchenko, Olga Berillo, Anna Pyrkova, Raigul Niyazova, and Shara Atambayeva

National Nanotechnology Laboratory, Al-Farabi Kazakh National University, Almaty 050038, Kazakhstan

Correspondence should be addressed to Anatoly Ivashchenko; a_ivashchenko@mail.ru

Received 14 April 2014; Revised 3 July 2014; Accepted 3 July 2014; Published 4 August 2014

Academic Editor: Vassily Lyubetsky

Copyright © 2014 Anatoly Ivashchenko et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The binding of 2,578 human miRNAs with the mRNAs of 12,175 human genes was studied. It was established that miR-619-5p, miR-5095, miR-5096, and miR-5585-3p bind with high affinity to the mRNAs of the 1215, 832, 725, and 655 genes, respectively. These unique miRNAs have binding sites in the coding sequences and untranslated regions of mRNAs. The mRNAs of many genes have multiple miR-619-5p, miR-5095, miR-5096, and miR-5585-3p binding sites. Groups of mRNAs in which the ordering of the miR-619-5p, miR-5095, miR-5096, and miR-5585-3p binding sites differ were established. The possible functional and evolutional properties of unique miRNAs are discussed.

1. Introduction

MicroRNAs (miRNAs) participate in the regulation of the expression of protein-coding genes at the posttranscriptional stage [1]. miRNAs, as a part of the RNA-induced silencing complex, bind to mRNAs and interfere with translation or promote mRNA destruction [2]. The study of the properties of miRNAs and their influences on the expression of the genes that participate in all key cellular processes of cells was established in the last 20 years. The actions of miRNAs on the cell cycle [3], apoptosis [4], differentiation [5], and growth and development in plants [6] and animals [7] have been shown. Connections between miRNA expression and the development of various diseases have been established. miRNA concentrations change in cancer [8] and cardiovascular diseases [9]. Metabolic disturbances necessarily change miRNA concentrations in cells [10]. It is possible to normalize some processes using miRNAs [11]. The aforementioned roles do not encompass the full list of the biological processes in which miRNAs participate, which proves the importance of their biological functions.

Despite the appreciable successes in the study of miRNA properties, there are obstacles to establishing the target genes of miRNAs. Normally, one miRNA interacts with the mRNA of one gene. However, there are miRNAs that bind to many

mRNAs, and one mRNA can be the target of many miRNAs. These circumstances significantly complicate the study of the properties of miRNAs and their diagnostic and medical applications. There are more than 2,500 miRNAs in the human genome, and they are thought to act on 50% or more of genes. It will be difficult to draw unique conclusions about the participation of miRNAs in specific biological processes, and until those conclusions can be drawn, the connections between the majority of miRNAs and their target genes will remain unknown.

Recently, we found a set of unique miRNAs that have hundreds of target genes and bind to mRNAs with high affinity. The binding sites unique to miRNAs are located in the 5'-untranslated regions (5'UTRs), the coding domain sequences (CDSs) and the 3'-untranslated regions (3'UTRs) of mRNAs. In present work, we studied some unique miRNAs that bind to the mRNAs of several hundred human genes.

2. Materials and Methods

The human gene mRNAs were taken from GenBank (http:// www.ncbi.nlm.nih.gov/) using Lextractor002 script (http:// sites.google.com/site/malaheenee/software). The nucleotide sequences of human miR-619-5p, miR-5095, miR-5096,

5'
GGG.AAGUGCUAAAGCCGCU.A.U.AGG
ACAU.CAAA.UAAAUUG.UG.UGUAC.CA.CA.A.AG. ANAPCI6 281* .CAA.GU.C.AG.A.GAGGGG
.CAA.GU.C.AG.A.GA
ACAAU.C.AAAUUAAUGG.GUGUACCAA.AG. CYB5D2 225*
GA.UCCG.AAUG.GUGGU
GAG. AACUU. AAAA AG U U A
.A.AA.UACAAAA.UAU.UUG.GUGUAC.CCA.AA.A.UG. DNASE1735*
AC. AAG. C. CA. CC. CA A A UG
AAA.AU.CAAAA.UA.UAAUUG.UGG.UAC.CGA.A.A.G PRR5 660*
UCG.AA.CC.CU.G.AAGAUGUGCACA
AGCU.UUUG.AAAAGAACAUUGUGC
UAA.CAU.GGAAAAU.CAAAUUGAAAACUAAG. ZNF7141896**
.AA.AACAUG. ANAPC16 2889***
AGACUUUU.GAAAA.UG.A.UU.A.UU.A.UGUGCAUAA.U. C8otf44 1626*** AC.CAU.UU.A.AGUGUA.A.UG.CYB5RL 3426***
GAA. AAUG. GAUGAUGUGA. AAGCCCG
AGA.AAU.C.A.C.U.UGAAU
UAGUUA.UU.A.AAAAA.U.UGCACCAA.UZNF714 6597***
ACCAU.CAAAUAACAU.UUGUGUAC.CUAC.CU.CAA.AG. ZNF714 6731***

FIGURE 1: miR-619-5p binding sites in the 5'UTRs, CDSs, and 3'UTRs of human genes. Note: for Figures 1 and 2, the symbols *, ** and ** * indicate the position of the origin of the miR-619-5p binding site from the first nucleotide of the 5'UTR in the 5'UTRs, CDSs, and 3'UTRs, respectively.

and miR-5585-3p were taken from the miRBase site (http://mirbase.org/).

The target genes for the tested miRNAs were revealed using the MirTarget program, which was developed in our laboratory. 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 the 5'untranslated regions (5'UTRs), the coding domain sequences (CDSs), and the 3'-untranslated regions (3'UTRs) of the mRNAs; (c) the free energy of hybridization (ΔG , kJ/mole); and (d) the schemes of nucleotide interactions between the miRNAs and the mRNAs. The ratio $\Delta G/\Delta G_m$ (%) was determined for each site (ΔG_m equals the free energy of an miRNA binding with its perfect complementary nucleotide sequence). The miRNA binding sites located on the mRNAs had $\Delta G/\Delta G_m$ ratios of 90% or more. We also noted the positions of the binding sites on the mRNA, beginning from the first nucleotide of the mRNA's 5'UTR. This 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 values (i.e., 3, 2, 1, and 1, resp.) [12, 13].

3. Results and Discussion

3.1. Features of miR-619-5p, miR-5096, miR-5585-3p, and miR-5095. The binding powers between the 2,578 tested hsa-miRNAs and the mRNAs of 12,175 human genes were calculated. Some of these miRNAs have greater numbers of

target genes than others. For example, miR-619-5p, miR-5095, miR-5096, and miR-5585-3p were found to be capable of binding more 600 genes each with value $\Delta G/\Delta G_m$ ratios of 90% or more. These miRNAs were termed unique miRNAs (umiRNAs). Additionally, the binding sites for these unique miRNAs were unusually located in the mRNAs. miR-619-5p, miR-5095, miR-5096, and miR-5585-3p have different miRNA binding site origins, lengths, quantities, and miRNA binding site properties, among other features. Some characteristics of these unique miRNAs are outlined below.

With a length of 22 nucleotides (nt), miR-619-5p is coded in an intron of the slingshot protein phosphatase 1 host gene (*SSH1*), which is located on chromosome 12. We found that miR-619-5p has 1811 binding sites on 1215 target mRNAs. Of those, 1772 miR-619-5p binding sites are located in 3'UTRs, 26 sites are located in 5'UTRs, and 13 sites are located in CDSs. The mRNAs of 197 genes have completely complementary binding sites for miR-619-5p. The mRNAs of 27 genes have four binding sites. Seven genes have five binding sites, and the mRNAs of the *CATAD1, ICA1L, GK5, POLH*, and *PRR11* genes have six miR-619-5p binding sites. The mRNAs of the *OPA3* and *CYP20A1* genes have eight and ten binding sites, respectively. All of these sites are located in 3'UTRs.

With a length of 21 nt, miR-5096 is coded in an intron of the BMP2 inducible kinase host gene (*BMP2K*), which is located on chromosome 4. We found that miR-5096 has 997 binding sites on 832 target mRNAs. Of these, 984 miR-5096 binding sites are located in 3'UTRs, nine sites are located in 5'UTRs, and four sites are located in CDSs. The mRNAs of 42 genes have completely complementary binding sites for miR-5096. The mRNAs of the *IP09* gene have four binding sites. The *PRR11* gene has five binding sites. The mRNAs of the *OPA3* and *CYP20A1* genes have six and eleven miR-5096

3' GCGCCACCAAGUGCGGACAUU 5' mir-5095 5' IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
(a)
3' CGGACUGGUUGUACCACUUUG 5' mir-5096 5' 1
(b)
3' UGGACAUCAGGGUCGAUAAGUC 5' mir-5585-3p 5'

FIGURE 2: miR-5095, miR-5096, and miR-5585 binding sites in the 5'UTRs, CDSs, and 3'UTRs of human genes.

binding sites, respectively. All of these sites are located in 3'UTRs.

With a length of 22 nt, miR-5585-3p is coded in an intron of the transmembrane protein 39b host gene (*TMEM39B*), which is located on chromosome 4. We found that 725 target gene mRNAs have 844 binding sites for miR-5585-3p. Nine of these binding sites are located in 5'UTRs, two sites are located in CDSs, and 833 sites are located in 3'UTRs. The mRNAs of the *CYP20A1* and *GPR155* genes each have four binding sites.

With a length of 21 nt, miR-5095 is coded in an intron of the sterol carrier protein 2 host gene (*SCP2*), which is located on chromosome 1. We found that 655 target gene mRNAs have 734 binding sites. Fourteen of these binding sites are located in 5'UTRs, eight sites are located in CDSs, and 712 sites are located in 3'UTRs. The mRNAs of two genes have completely complementary binding sites for miR-5095. The mRNAs of the *OPA3* and *SPN* genes each have four binding sites.

3.2. miRNA Binding Sites in 5' UTRs, CDSs, and 3' UTRs. The miR-619-5p, miR-5095, miR-5096, and miR-5585-3p binding sites in the 5' UTRs, CDSs, and 3' UTRs of several genes were predicted using the MirTarget program. Multiple miRNA binding sites were revealed to be in the 5' UTRs of several genes. For example, miR-619-5p has two binding sites in each of the 5' UTRs of the ANAPC16, CYB5D2, and PRR5 mRNAs and three binding sites in the DNASE1 mRNA (Figure 1).

The mRNAs of some genes have binding sites for miR-619-5p, miR-5095, miR-5096, and miR-5585-3p within their 5'UTRs and 3'UTRs or CDSs and 3'UTRs. For example, the 5'UTRs and 3'UTRs of the *ATAD3C* and *CYB5RL* genes have miR-619-5p binding sites. The CDSs and 3'UTRs of the *C8orf44, ISY1*, and *ZNF714* genes have miR-619-5p binding sites.

The 5'UTR and 3'UTR of the ANAPCI6 gene have miR-5095, miR-5096, and miR-5585-3p binding sites (Figure 2). The 5'UTR and 3'UTR of the ATAD3C gene have miR-5095 and miR-619-5p binding sites. The 5'UTRs and 3'UTRs of the *Cl4orf182* and *CYB5RL* genes have miR-5096 and miR-619-5p binding sites, respectively.

miR-5095 and miR-619-5p binding sites were found in the CDS and 3'UTR of the *ISY1* gene. The CDS and 3'UTR of the *ZNF714* gene have binding sites for miR-5096 and miR-619-5p, and the *C8orf44* mRNA has only an miR-619-5p binding site.

The nucleotide sequences of the miR-619-5p binding sites mRNAs of the OPA3 and SPN genes each have four binding locate in the CDSs of the *C8orf44*, *ISY1*, and *ZNF714* genes'. The sites code for the following oligopeptides

ENHWKGRARWLMPVIPALWEAKAGRS C8orf44, LFEKERQVRWLMPVIPALWEAEAGGS ISY1, KHRKIQQGMVAHACNPNTLRGLGEQI ZNF714.

The first two oligopeptides are coded in one open reading frame (ORF), and the amino acids the miR-619-5p binding site codes for are highly conserved (highlighted in bold). The homologous oligonucleotide of the miR-619-5p binding site in *ZNF714* mRNA codes for an oligopeptide in other ORF. The presence of miR-619-5p binding sites in the CDSs

	3' CCGAGUACGGACAUUAGGGUCG	5′ mir-619-5p	
5'	111111111111111111111111111111111111111	3'	
UGGCCGGGAGUGGUGGCU	CAUGCCUGUAAUCCCAGCACUUUGGG	AUGCUGAGGCGGGUGGAUCACCUGAGGUCAGGAG	CYP20A1 2539
C	CA	.G.UCC.A	CYP20A1 2676
.AUCAA	UAC.CA	.GAA.AUUACCUG	CYP20A1 2954
CA		.GCACAC.	CYP20A1 4709
CACA	CA	.GCAAAUG	CYP20A1 5724
CCAGGC.A	CA	.GA	CYP20A1 6031
.AA	UUAC.G	.GAA.AA.AG.UACCU	CYP20A1 6165
. AA. UGGCCU. U. G A	UUC	.GAU.A.A.AA.AAAU.AAC.UGA	CYP20A1 8845
		.GU.CA.A.UC.UAU	CYP20A1 9054
CAUU	G.A.	.GCAU.ACUUUGGGUCU.	CYP20A1 9770
		.GCCU	OPA3 1192
		.G.UCUCGUUAG	OPA3 1644
		.GA.CAGA.A.G.UCU.UU.AG.CCAG.	OPA3 2902
GA.UGACCCAG.CCA	CCA.	.GCAUACUAUU	OPA3 4388
		.GCAGU	OPA3 5063
AC	CAUU	.GCCCACG	OPA3 6171
		.GUAA.AG.UACC.G	OPA3 6305
		.GA.CAUAUA	OPA3 7199
GU.U.U.CA	C	.GA.UCUU	GK5 3808
CAUCACA	CAUU	.GCAACU.CC	GK5 5278
.AU.U.CUA	UAG.CA	.GAA.A.C.G.UACC.CU	GK5 6112
		.GCA.AUA.C	GK5 6355
AAUAA.CU.	CAU.	.GCAAG.UC	GK5 7088
AAUA	C	.GCA	GK5 7230

(a)

3' CGGACUGGUUGUACCACUUUG 5' mir-5096	
5'	
UGAGGCGGGUGGAGCACCUGAGGUCAGGAAUUCGAGAUCAGCCUGGCCAACAUGGUGAAACCCCCAUCUCUACUGAAAA	CYP20A1 2886
AUCAGUCUAAAA.	CYP20A1 2477
	CYP20A1 2613
CUUC.A.GGCG.AUGUGU.A	CYP20A1 2750
CAUAUUGCAUGGA	CYP20A1 3596
CCGACAUGCAGAUCA.GAG.UCA.AGUC.U	CYP20A1 3942
CACAUCGACAA	CYP20A1 4783
CAAA.UUGG.UCAA	CYP20A1 5798
CAUAGACAA	CYP20A1 6105
CAA.UUCG.UGGG	CYP20A1 8604
CA.A.UC.U.UAUGAAGG	CYP20A1 9128
CCUUUGGCC	OPA3 1266
.UU.AA.UCGGU.GGAGGCCAUAGGA	OPA3 1734
CAUACUUAUUGACCACAG	OPA3 4462
CAGUUU.GGACAAUU.GA	OPA3 5137
CC.A.U.CGGCAGAA.	OPA3 6245
CAUAUUAGCCGGA.A	OPA3 7273
AUUC.CAA	PRR11 1540
CACUUUCAC	PRR11 3428
CAA.AUACGUC	PRR11 4271
CACAUUGAGCUUGUGA	PRR11 5271
CUU.UUGUCCGG	PRR11 5615

(b)

FIGURE 3: The nucleotide parts of 3'UTRs having multiple miR-5095 (a) and miR-5585-3p (b) binding sites. Note: *—one adenine was deleted before cytosine.

of three genes with different functions and the evolutionary conservation of these sites testify to the importance of the role of miRNA in the regulation of the expression of these genes. The nucleotide sequences of the parts of the *C8orf44*, *ISY1*, and *ZNF714* mRNAs that contain the miR-619-5p binding sites in the CDSs are homologous among themselves and homologous with the binding sites located in the 5'UTRs and 3'UTRs.

In mRNA oligonucleotides, the miR-619-5p, miR-5095, miR-5096, and miR-5585-3p binding sites and the nucleotide sequences immediately downstream of the binding sites are highly homologous. It is possible that these downstream sequences are binding sites for other miRNAs. The sequences of the 5'UTRs, CDSs, and 3'UTRs upstream of the miR-5096 and miR-5585-3p binding sites are also homologous, likely for the same reason. The sequences of the mRNAs upstream of

,

3'	GCGCCACCAAG	UGCGGACAUU 5' n	nir-5095
5'	111111111111		3'
AGGAAAGAGUGGCCAGGCGUGGUG	GCUCACACCUG	UAAUCUCAGCACUU	UGGGAGGCCGAGGCGGGCCAAUCCCCUGAGGUCGG OPA3 6165
.AA.CG.U.CUGA.C	G	CU	A. OPA3 1186
.A.UCCUU.G.A.UGCA	A	CU	
CAGUUUUA	G	U	AAUAGGA.UA. OPA3 7193
GAGGC.CAGU.C	G		AUAGAUGA.GUCAG.A SPN 3911
.C.UGGAGCA.AU.CACA	UG	C	ACAAGGUG.UUCUA. SPN 5281
AU.GAGCAA	*GU	C	A. UGGAA. SPN 6012
GAA.G.ACA		C	A.U.AGAUGA.GCCAGUU SPN 6627
.CGUUCAGUUGC	G	C	
.CUGU.A.CGAC	G	C	UAUGGAA. PDDC13442
C.UUCUUGGCA		C.C	GGA.GA.GUCAG.A PDDC11677
GUACCGGC	UG	CCG	UUAGUA.UCC.A. CACNG8 3212
UUA.GCUGAAGCA	UG	C	UUGGA.UA. CACNG8 5000
U.AAU.A.CUGU	G	GCUAUAC.	CUAGGUG.UUC.A. CACNG8 7663

,

(a)

3' UGGACAUCAGGGUCGAUAAGUC 5' mir-5585-3p

5'	3'
UGGGCAUGGUGGCACAUGCCUGUAAUCCCAGCUACUCAGGAGGCU	JGAGGCAGGAGAAUCAUUUGAACCUGGGAGGCGGAGCUUG CYP20A1 2961
GUU.A	UGGCGA.CAAG.UCAAGGC CYP20A1 3364
CCAC.U.GU.GCACCAC.GAU	GAGGGCG.UCAA.UCAGC.A CYP20A17510
CAUGCGCAGAA	UGGUGCGCAGUAG CYP20A19912
UGGGG	
CAUGA.UGGUUA.	UGU.C.CCGCG <i>GPR155</i> 4375
CAUGUGUGCACC	AUGCA.CAAG GPR155 5020
CA.A.GUGGGCAAG	
CGAGGGCAGG.	ACCAAG ORAI2 2703
CU	A
UGGCUCGUG	CCCA.AAG.A. ORAI2 3946
CG	
CAUGG	
CG	UGCCAA <i>CCDC142</i> 3706

(b)

FIGURE 4: The nucleotide parts of 3'UTRs having multiple miR-619-5p (a) and miR-5096 (b) binding sites.

5'UTR 3' CCGAGUACGGACAUUAGGGUCG 3' GCGCCACCAAGUGCGGACAUU 5'	miR-5095
1 ANAPC16 147 GAGGUGCAGUGGCUCACGCCUGUAAUCCCAGCACU	JUGGGAGGCCAAGGCAGGUGGAUCACCUGAGGUCGGGAGUUCAAGACCA
2 CAMKK2159 CGC.GGAAA.	CAAUU.
3 DNASE1 602 UGGA	GCACUG
3' CGGACUGGUUGUACCACUUUG 5' miR-5095	miR-5585-3p 3' UGGACAUCAGGGUCGAUAAGUC 5'
1 GCCUGACCAACAUGGAGAAACCCCCAUCUCUACUAGAAAUACAAAA-U	
2GAU	A.UCC.GCA
3 UG	GUG

FIGURE 5: miR-619-5p, miR-5095, miR-5096, and miR-5585-3p binding sites located in 5'UTRs.

the miR-5095 binding sites are not homologous, and no other miRNA binding sites are identified in these sequences. The 10-nucleotide sequences upstream of the miR-619-5p binding sites contain homologous nucleotide sequences because they are the miR-5095 binding sites.

3.3. Multiple miRNA Binding Sites in the mRNAs of Target Genes. The mRNAs of some genes have multiple umiRNA

binding sites. The nucleotide sequences with lengths of 95 nt that contain multiple miR-619-5p, miR-5096, miR-5095, and miR-5585-3p binding sites are given in Figures 3 and 4. These results testify to the high degree of homology between the umiRNA binding sites in the mRNAs of different genes. In addition to these binding sites, many other nucleotide sequences of the mRNAs are also homologous. It is possible that the nucleotide sequences adjacent to the binding sites are binding sites for other miRNAs.

		3'CCGAGUACGGACAUUAGGGUCG miR-619-5p
		3' GCGCCACCAAGUGCGGACAUU miR-5095
		5′
1	GPR155 3365	CGGGCAUGGUGGCUCACGCCUGUAUCCCAGUACUUUGGGAGGCCGAGUCAGGUGGAUCACCUGAGGUCAGGAGUUCAAGACCA
2	ABHD11 1038	GUUCUAG.G.AG.UC
3	AGMAT 2207	$U\ldots A\ldots C\ldots \ldots . U\ldots \ldots U\ldots \ldots C\ldots \ldots G\ldots \ldots G U G \ldots U \ldots U \ldots \ldots U \ldots \ldots G \ldots \ldots G \ldots \ldots G U G U$
4	BRCA16412	UAGUGCAUG
5	C5ofr28 2452	A.GUU
6	CHST6 2979	.AGUUUUUUU
7	DCAF104559	.AGCAUC
8	DES11 1977	.UGCU
9	ENAH 8575	GUGCAUG
10	FAM126B 4290	C
11	GNE 2797	UGCA
12	IL17RD 8011	GC
	KCTD20 4742	.AG
	<i>KIAA1191</i> 1403	UGAGUAC.U
	METTL6 1188	U UGC
	NME6 910	C
	NUMBL 3040	UAGUUGUGUUU
	OTUD6A 1266	GC
	PLEKHA2 3417	
		.AGCA
	SLC25A15 2783 TMEM120B 3626	C
21		
21		miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5'
	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5' 3'
	CGGACUGGUUGUAC	miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5'
	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5' 3'
1	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5' GCGAAACCCCGUCUCUACUAUAAAAUACAAAAAUUAGUUGGGUGUGGGUGG
1 2	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5'
1 2 3 4	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5'
1 2 3	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5'
1 2 3 4 5 6	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5' IIIIIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
1 2 3 4 5 6 7	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5' IIIIIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
1 2 3 4 5 6	CGGACUGGUUGUAC	CACUUUG 5' miR-5095 miR-5885-3p UGGACAUCAGGGUCGAUAAGUC 5' IIIIIII IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
1 2 3 4 5 6 7 8 9	CGGACUGGUUGUAC	$\begin{array}{cccccccccccccccccccccaccaccaccaccaccacc$
1 2 3 4 5 6 7 8 9 10	CGGACUGGUUGUAC	$\begin{array}{cccccccccccccccccccaccaccaccaccaccaccac$
1 2 3 4 5 6 7 8 9 10 11	CGGACUGGUUGUAC	$ \begin{array}{c} \text{miR-5885-3p} \ \text{UGGACAUCAGGGUCGAUAAGUC} 5' \\ \\ \\ GCGAAACCCCGUCUCUACUAUAAAAUACAAAAAUUAGUUGGUGUGGUGGUGGCGGUGCCUGUAAUCCAGGAGG \\UAUACACCU.CU.CACAU. \\AUACAG.GCC.UCAU.ACAU.ACAU. \\AUACAG.GCC.UCAU.ACAG.UUU.G. \\U.A.A.A.A.A.AAAAUACAGA.GCC.C.U.U.CA.A.A.G.UU.A.A.A.AAAAUACAGAGG \\U.A.A.A.A.AAAAAAAAAAAAAAAAAAAAAAA$
1 2 3 4 5 6 7 8 9 10 11 12	CGGACUGGUUGUAC	$\begin{array}{cccccccccccccccccccccaccaccaccacccccccc$
1 2 3 4 5 6 7 8 9 10 11 12 13	CGGACUGGUUGUAC	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
1 2 3 4 5 6 7 8 9 10 11 12 13 14	CGGACUGGUUGUAC	$\begin{array}{cccccccccccccccccccccaccaccaccacccccccc$
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	CGGACUGGUUGUAC	$ \begin{array}{c} \text{miR-5885-3p} \ \text{UGGACAUCAGGGUCGAUAAGUC} 5' \\ \\ \\ GCGAAACCCCGUCUCUACUAUAAAAUACAAAAAUUAGUUGGUGUGGUGGUGGCGGUGCCUGUAAUCCCAGGAGG \\$
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16	CGGACUGGUUGUAC	$ \begin{array}{c} \text{miR-5095} \\ \text{miR-5095} \\ \text{miR-5885-3p} \ \text{UGGACAUCAGGGUCGAUAAGUC 5'} \\ HIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII$
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17	CGGACUGGUUGUAC	$ \begin{array}{c} miR-5885-3p \ UGGACAUCAGGGUCGAUAAGUC 5' \\ \\ GCGAAACCCCGUCUCUACUAUAAAAUACAAAAAUUAGUUGGUGUGGUGGUGGGUG$
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18	CGGACUGGUUGUAC	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19	CGGACUGGUUGUAC	$\begin{array}{ccccccccccccccccccccccaccaccccccccccc$
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20	CGGACUGGUUGUAC	$\begin{array}{cccccccccccccccccccccccccccccccccccc$

FIGURE 6: miR-619-5p, miR-5095, miR-5096, and miR-5585-3p binding sites located in the 3'UTRs.

3.4. The Arrangements of the Locations of umi-RNA Binding Sites. The mRNAs that were targeted by miR-619-5p, miR-5096, miR-5095, and miR-5585-3p were established. The 5'UTRs of three target genes contain these miRNA-binding sites (Figure 5). The degree of homology of the nucleotide sequences in these genes is high not only in the binding sites of the studied miRNAs but also across all mRNA 150 nt sequences. The distance between the miR-5095 and miR-5096 binding sites is 57–59 nt and that between the miR-5096 and miR-5585-3p binding sites is 46-47 nt. The miR-5095 and miR-619-5p binding sites partially overlap.

The greatest numbers of miR-619-5p, miR-5096, miR-5095, and miR-5585-3p binding sites are located in the 3'UTRs, and it is, therefore, possible that many target genes have umiRNAs binding sites. The data about the locations of the miR-619-5p, miR-5096, miR-5095, and miR-5585-3p binding sites and the degrees of homology of the corresponding nucleotide sequences in the mRNAs of 21 genes are presented in Figure 6. The distances between the miR-5095 and miR-5096 binding sites are all 57–60 nt. The distances between the miR-5095 and miR-5096 binding sites in the mRNAs of 78 genes averaged 58.6 \pm 0.9 nt. Thus, the distances between miR-5095 and miR-5096 binding sites are highly conserve. The distances between the miR-5096 and miR-5585-3p binding sites are all 46–49 nt. The distances between the miR-5096 and miR-5585-3p binding sites in the mRNAs of 325 genes average 47.3 \pm 1.1.

3'UTR

3'CCGAGUACGGACAUUAGGGUCG miR-619-5p
3'UGGACAUCAGGGUCGAUAAGUC miR-5885-3p
5'
UUAGCCGGGUGUGGUGGCACACGCCUGUAGUCCCAGCUACUCGGGAGGC AARS2 3377
UAAAUUGUUA AMZ1 2088
AA.ACA.UUGUAAAA
AAAAA
A.UAUUA
.GCAUU.U.UAAAA
GUU.AG.CCUU.CUU.CUUAUA
CAAU.U.U CLN8 1823
UGAACUAADHFR 3098
AGAC.CAU.GUAAA DPH3 2387
ACAUUAA
UUCAUUCAEIF2AK4 5249
U
ACA.AUGUAGCA <i>F11R</i> 2001

FIGURE 7: 3'UTRs with miR-619-5p binding sites upstream of the miR-5585-3p binding sites. Note: Symbol " \approx " indicates nucleotide sequence equaled to 84 nt, which is not shown here.

3' GCGCCACCAAGUGCGGACAUU miR-5095	3' CCGAGUACGGACAUUAGGGUCG miR-619-5p
3'CCGAGUACGGACAUUAGGGUCG miR-619-5p	3' UGGACAUCAGGGUCGAUAAGUC miR-5885-3p
5′	$ $
GAGGUGCAGUGGCUCACGCCUGUAAUCCCAGCACUUUGGGAGGCCAA	\approx UUAGCCAGGCAUGGUGGUGCAUGCCUGUAAUCCCAGCUACUCGGGAG ANAPC16 2889
CA.G	≈GCAA C7orf55 860
C.UG	≈UGGA C11orf58 2266
CUCG	≈GCGDESII 1977
UGC.UGA	\approx GGGCDNM1L 3407
UGC.UGG.	\approx GUGA. FAM217B 3710
UGGUUUG.	\approx GCA.GGCCCA1L 2980
UGC.UGUUG.	\approx CG. UA. NDUFAF7 1697
CGCA.GUG.	\approx CUUUAG
CC.UG	≈G
CGCU.G	\approx GUGCAPGM2L1 7269
UGCAUGUG.	≈GG. PTCD1 3990
CCAUGGUG.	\approx AUGCCAAA. RHOBTB3 4033
CCA.G	\approx GG. SLC35E3 1262
UGGAUG.	\approx UG.AZNF430 2049

FIGURE 8: 3'UTRs in which the miR-619-5p binding site was downstream of the miR-5095 binding site and upstream of the miR-5585-3p binding site.

The degree of homology of the nucleotide sequences containing the miR-619-5p, miR-5096, miR-5095, and miR-5585-3p binding sites is high. These areas contain binding sites for miRNAs other than the studied umiRNAs. Other miRNA binding sites are not present in all genes, and these binding sites have lower affinities (data not shown). It is possible that there are conserved domains in the nucleotide sequences of mRNAs.

3.5. Variability in the Arrangement of umiRNA Binding Site Locations. The miR-619-5p binding site is located at a distance of six nucleotides downstream of the beginning of the miR-5095 site in the majority of genes containing arranged umiRNA sites. However, in another group of mRNAs, the beginnings of the miR-619-5p binding sites are located at distances of seven nucleotides upstream of the miR-5585-3p binding sites (Figure 7).

There is another group of genes in which the miR-619-5p binding sites are downstream of the miR-5095 sites and upstream of the miR-5585-3p sites (Figure 8). The distances between the positions of the two miR-619-5p binding sites in the mRNAs of these genes are constant at 112 nt. The nucleotide sequences of the mRNAs with miR-619-5p, miR-5095, and miR-5585-3p binding sites are highly homologous, which testifies to the strength of the selection pressure on these nucleotide sequences.

4. Conclusion

The detection of a large number of binding sites of miR-619-5p, miR-5095, miR-5096, and miR-5585 in the mRNAs of the genes studied here presumably indicates new functional opportunities. It is possible that these umiRNAs are coordinators of gene expression that participate in many major biological processes. The influences of miRNAs on the expression of genes that code for transcription factors [14, 15] and proteins that participate in the cellular cycle [3], apoptosis [4], stress responses [16], and so forth, have previously been shown.

This study established that the binding sites of umiRNAs and miRNAs are in 3'UTRs, CDSs, and 5'UTRs. Highly conserved miR-619-5p, miR-5095, miR-5096, and miR-5585 binding sites in a large number of genes indicate the emergence of these sites in the early stages of human evolution. We have shown previously that miRNA binding sites located in CDS are conserved in target orthologous genes of organisms that diverged some hundred million years ago [17, 18]. SSH1, BMP2K, TMEM39B, and SCP2 host genes and target genes of miR-619-5p, miR-5095, miR-5096, and miR-5585 have independent evolution origin. Revealed miRNA binding sites have arranged localization in mRNA of multiple target genes participating in different metabolic processes. Host genes often coexpress with their intragenic miRNAs; therefore, an expression of SSH1, BMP2K, TMEM39B, and SCP2 genes can be connected with target genes via their miRNAs. Thus, it predicted that host genes and target genes are interconnected between themselves. Arranged localization of binding sites suggests an interconnected evolution of miRNAs and their target genes. Conservatism of the arranged localization of miRNA binding sites in 3'UTRs, 5'UTRs, and CDSs also demonstrates the common origin of the binding sites and the evolution of relationship of miRNAs with their target genes.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank A. Moldagaliyeva and S. Sagaydak for their help in preparing the materials for analysis. The authors would also like to thank Dr. V. Khaylenko for writing the Lextractor002 script. This study was supported by a grant from the Ministry of Education and Science, Kazakhstan Republic.

References

- E. Doxakis, "Principles of miRNA—target regulation in metazoan models," *International Journal of Molecular Sciences*, vol. 14, no. 8, pp. 16280–16302, 2013.
- [2] G. Tang, "siRNA and miRNA: an insight into RISCs," *Trends in Biochemical Sciences*, vol. 30, no. 2, pp. 106–114, 2005.
- [3] Q. Luo, X. Li, J. Li et al., "MiR-15a is underexpressed and inhibits the cell cycle by targeting *CCNE*1 in breast cancer," *International Journal of Oncology*, vol. 43, no. 4, pp. 1212–1218, 2013.
- [4] X. Li, Y.-T. Chen, S. Josson et al., "MicroRNA-185 and 342 inhibit tumorigenicity and induce apoptosis through blockade of the SREBP metabolic pathway in prostate cancer cells," *PLoS ONE*, vol. 8, no. 8, Article ID e70987, 2013.
- [5] N. Qian, W. Liu, W. Lv et al., "Upregulated microRNA-92b regulates the differentiation and proliferation of EpCAM-positive

fetal liver cells by targeting C/EBP β ," *PLoS ONE*, vol. 8, no. 8, Article ID e68004, 2013.

- [6] K. Rogers and X. Chen, "Biogenesis, turnover, and mode of action of plant microRNAs," *Plant Cell*, vol. 25, no. 7, pp. 2383– 2399, 2013.
- [7] Y. H. Ling, J. P. Ding, X. D. Zhang et al., "Characterization of microRNAs from goat (*Capra hircus*) by Solexa deepsequencing technology," *Genetics and Molecular Research*, vol. 12, no. 2, pp. 1951–1961, 2013.
- [8] Q. J. Guo, J. N. Mills, S. G. Bandurraga et al., "MicroRNA-510 promotes cell and tumor growth by targeting peroxiredoxin1 in breast cancer," *Breast Cancer Research*, vol. 15, no. 4, article R70, 2013.
- [9] A. Magenta, S. Greco, C. Gaetano, and F. Martelli, "Oxidative stress and microRNAs in vascular diseases," *International Journal of Molecular Sciences*, vol. 14, no. 9, pp. 17319–17346, 2013.
- [10] S. Swaminathan, K. Suzuki, N. Seddiki et al., "Differential regulation of the Let-7 family of microRNAs in CD4⁺ T cells alters IL-10 expression," *Journal of Immunology*, vol. 188, no. 12, pp. 6238–6246, 2012.
- [11] D. V. Glazkova, A. S. Vetchinova, E. V. Bogoslovskaia, I. A. Zhogina, M. L. Markelov, and G. A. Shipulin, "Downregulation of human CCR5 receptor gene expression using artificial microRNAs," *Molecular Biology*, vol. 47, no. 3, pp. 475–485, 2013.
- [12] E. T. Kool, "Hydrogen bonding, base stacking, and steric effects in DNA replication," *Annual Review of Biophysics and Biomolecular Structure*, vol. 30, pp. 1–22, 2001.
- [13] N. B. Leontis, J. Stombaugh, and E. Westhof, "The non-Watson-Crick base pairs and their associated isostericity matrices," *Nucleic Acids Research*, vol. 30, no. 16, pp. 3497–3531, 2002.
- [14] Q. Cui, Z. Yu, Y. Pan, E. O. Purisima, and E. Wang, "MicroR-NAs preferentially target the genes with high transcriptional regulation complexity," *Biochemical and Biophysical Research Communications*, vol. 352, no. 3, pp. 733–738, 2007.
- [15] L. Yan, M. Kang, Z. Qin, W. Zhang, Y. Li, and H. Ou, "An intronic miRNA regulates expression of the human endothelial nitric oxide synthase gene and proliferation of endothelial cells by a mechanism related to the transcription factor SP-1," *PLoS ONE*, vol. 8, no. 8, Article ID e70658, 2013.
- [16] K. Cawley, S. E. Logue, A. M. Gorman et al., "Disruption of microRNA biogenesis confers resistance to ER stress-induced cell death upstream of the mitochondrion," *PLoS ONE*, vol. 8, no. 8, Article ID e73870, 2013.
- [17] A. Bari, S. Orazova, and A. Ivashchenko, "Mir156- and mir171binding sites in the protein-coding sequences of several plant genes," *BioMed Research International*, vol. 2013, Article ID 307145, 7 pages, 2013.
- [18] A. T. Ivashchenko, A. S. Issabekova, and O. A. Berillo, "MiR-1279, miR-548j, miR-548m, and miR-548d-5p binding sites in CDSs of paralogous and orthologous PTPN12, MSH6, and ZEB1 genes," *BioMed Research International*, vol. 2013, Article ID 902467, 10 pages, 2013.