The Binding Sites of miR-619-5p in the mRNAs of Human and Orthologous Genes

Shara Atambayeva 1*, Raigul Niyazova 1, Anatoliy Ivashchenko 1, Anna Pyrkova 1, Ilya Pinsky 1, Aigul Akimniyazova 1 and Siegfried Labeit 2

Abstract

Background: Normally, one miRNA interacts with the mRNA of one gene. However, there are miRNAs that can bind to many mRNAs, and one mRNA can be the target of many miRNAs. This significantly complicates the study of the properties of miRNAs and their diagnostic and medical applications.

Results: The search of 2,750 human microRNAs (miRNAs) binding sites in 12,175 mRNAs of human genes using the MirTarget program has been completed. For the binding sites of the miR-619-5p the hybridization free energy of the bonds was equal to 100% of the maximum potential free energy. The mRNAs of 201 human genes have complete complementary binding sites of miR-619-5p in the 3'UTR (214 sites), CDS (3 sites), and 5'UTR (4 sites). The miRNAs of CATAD1, ICA1L, GKS, POLH, and PRR11 genes have six miR-619-5p binding sites, and the miRNAs of OPA3 and CYP20A1 genes have eight and ten binding sites, respectively. All of these miR-619-5p binding sites are located in the 3'UTRs. The miR-619-5p binding site in the 5'UTR of mRNA of human USP29 gene is found in the mRNAs of orthologous genes of primates. Binding sites of miR-619-5p in the coding regions of mRNAs of C8H8orf44, C8orf44, and ISY1 genes encode the WLMPVIP oligopeptide, which is present in the orthologous proteins. Binding sites of miR-619-5p in the mRNAs of transcription factor genes ZNF429 and ZNF429 encode the AHACNP oligopeptide in another reading frame. Binding sites of miR-619-5p in the 3'UTRs of all human target genes are also present in the 3'UTRs of orthologous genes of mammals. The completely complementary binding sites for miR-619-5p are conservative in the orthologous mammalian genes.

Conclusions: The majority of miR-619-5p binding sites are located in the 3'UTRs but some genes have miRNA binding sites in the 5'UTRs of mRNAs. Several genes have binding sites for miRNAs in the CDSs that are read in different open reading frames. Identical nucleotide sequences of binding sites encode different amino acids in different proteins. The binding sites of miR-619-5p in 3'UTRs, 5'UTRs and CDSs are conservative in the orthologous mammalian genes.

Keywords: miR-619-5p, miRNA, mRNA, Gene, Human, Orthologous genes

Background

miRNAs participate in the regulation of the expression of protein-coding genes at the post-transcriptional stage [1]. miRNAs, as a part of the RNA-induced silencing complex, bind to mRNAs and interfere with translation or promote mRNA destruction [2]. In the last two decades, properties of miRNAs and their influences on the expression of the genes involved in all key cellular processes have been established. The actions of miRNAs on the cell cycle [3], apoptosis [4], differentiation [5], and growth and development of 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 perturbations change miRNA concentrations in cells [10]. The aforementioned roles do not encompass all of the biological processes in which miRNAs participate, which further proves the importance of their biological functions. Despite the significant success in the study of miRNA properties, there are obstacles in identifying the target genes of miRNAs. Normally, one miRNA interacts with the mRNA of one gene. However, there are miRNAs that can bind to many mRNAs, and one miRNA
can be the target of many miRNAs, which significantly complicates 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 believed to act on 60% or more genes. Therefore, it is difficult to draw specific conclusions about the participation of miRNAs in specific biological processes, and until then the connections between the majority of miRNAs and their target genes will remain unknown. Recently, a set of unique miRNAs (umiRNA) were identified that have hundreds of target genes and bind to mRNAs with high affinity [11–14]. The binding sites of these umiRNAs are located in the 3' UTRs, CDSs, and 5' UTRs of mRNAs. Among these umiRNAs, miR-619-5p interacts with the largest number of target genes that have the greatest number of binding sites with complete complementarity of miR-619-5p and mRNAs. It is necessary to identify many miRNA binding sites in the mRNAs of these genes for the control of gene expression. Furthermore, it is important to control the expression of the corresponding gene complexes that are functionally associated with miRNAs. Therefore, we have studied a unique miR-619-5p that binds to the mRNAs of several hundred human and orthologous genes.

Methods
The nucleotide sequences of mRNAs of human genes (Homo sapiens – Hsa) and orthologous genes (Bos mutus - The wild yak (Bmu), Callithrix jacchus – The common marmoset (Cja), Camelus dromedarius – Arabian camel (Cdr), Camelus ferus – The wild Bactrian camel (Cfe), Chlorocebus sabaeus – The green monkey (Csa), Colobus angolensis palliatus – The Angola colobus (Can), Equus caballus – The horse (Eca), Gorilla gorilla – The western gorilla (Ggo), Macaca fascicularis – The crab-eating macaque (Mfa), Macaca mulatta – The rhesus macaque (Mmu), Macaca nemestrina – Pig-tailed macaque (Mne), Mandrillus leucophaeus – The drill (Mle), Nomascus leucogenys - The northern white-cheeked gibbon (Nle), Ovis aries – The sheep (Oar), Pan paniscus – Bonobos (Ppa), Pan troglodytes – The common chimpanzee (Ptr), Papio anubis – The olive baboon (Pan), Pongo abelii - The Sumatran orangutan (Pab), Rhinopithecus roxellana – The golden snub-nosed monkey (Rro) were downloaded from NCBI GenBank (http://www.ncbi.nlm.nih.gov) [15] in FASTA format using Lextractor002 script [11]. Nucleotide sequences of human mature miR-619-5p (GCUGGGA UUAAGGCAUGAGCC) were downloaded from the miRBase database (http://mirbase.org) [16]. The miR-619-5p binding sites in the 5'-untranslated regions (5'UTRs), the coding domain sequences (CDSs) and the 3'-untranslated regions (3'UTRs) of several genes were predicted using the MirTarget program [12]. This program defines the features of binding: a) the localization of miRNA binding sites in the 5'UTRs, the CDSs and the 3'UTRs of the mRNAs; b) the free energy of hybridization (ΔG, kJ/mole). The ratio ΔG/ΔGm (%) was determined for each site (ΔGm equals the free energy of miRNA binding with its perfect complementary nucleotide sequence).

Results
The search of 2,750 human microRNAs (miRNAs) binding sites in 12,175 mRNAs of human genes using the MirTarget program has been completed. The miRNAs have different miRNA binding site origins, lengths, quantities, and properties. The list of miR-619-5p target genes and the positions of binding sites are outlined in Table 1. miR-619-5p is 22 nucleotides in length and is coded by an intron of the slingshot protein phosphatase 1 (SSH1) gene, which is located on chromosome 12 [17, 18]. miRNAs of 201 genes have complete complementary binding sites for miR-619-5p (ΔGm = 100%). Therefore, the energy of interaction of miR-619-5p with mRNA of all the genes listed in the table is the same and equal to ΔG = −121 kJ/mole.

The miRNAs of 201 human genes have complete complementary binding sites of miR-619-5p in the 3'UTR (214 sites), CDS (3 sites), and 5'UTR (4 sites). The miRNAs of 27 genes have four binding sites, seven genes have five binding sites, and CATADI, ICAI1, GKS, POLH, and PRR11 genes have six miR-619-5p binding sites. The miRNAs of OPA3 and CYP20A1 genes have eight and ten binding sites, respectively. All of these sites are located in the 3'UTRs of mRNAs. The target genes of the miR-619-5p carry out one or more different functions and are involved in the development of various diseases (Table 1).

The miRNAs of the C17orf75, C8orf44, CIAO1, CPM, CYP20A1, DCAF10, FBKPI4, RAB3BP, SYNJ2BP, VHL genes have two complete complementary binding sites for miR-619-5p, and the mRNA of the CACNG8 gene has three such binding sites. This indicates a stronger dependence of the expression of these genes on miR-619-5p.

One of the methods to establish the credibility of the presence of miRNA binding site in the mRNA is to verify this site in the mRNAs of orthologous genes. In finding the miRNA binding sites raises the question of the level of reliability of the found sites. One effective way to establish the credibility of the binding sites is to establish binding sites in the orthologous genes and the identification of orthologous miRNA. Location of binding site in the protein coding region facilitates its conservation in evolution, especially if the corresponding oligopeptide plays an important role in the function of the protein. miR-619-5p binding sites with complete complementarity (ΔGm = 100%) to the mRNAs of the four genes are located in the 5'UTRs (Table 2).
| Gene     | Site, nt | Disease or function | PMID   | Gene     | Site, nt | Disease or function | PMID   |
|----------|----------|---------------------|--------|----------|----------|---------------------|--------|
| ACSL6    | 4639     | prostate cancer     | 19064571 | MRPS25   | 1609     | uncharacterized     | 26302410 |
| ADAL     | 2041     | proliferation       | 23645737 | MSH3     | 4139     | carcinoma           | 24934723 |
| ADAM17   | 3466     | breast cancer       | 22967992 | NANO1    | 3219     | retinoblastoma      | 25100735 |
| AGMAT    | 2207     | renal carcinoma     | 14648699 | NCMAP    | 2259     | uncharacterized     |        |
| AK1      | 1449     | hypertension        | 23863634 | NDUF4F7  | 1697     | leukemia            | 24292274 |
| AKT2     | 4571     | neuroblastoma       | 23468863 | NDUF4C2  | 1646     | colon cancer        | 25804238 |
| ALDH3A2  | 2617     | detoxification      | 9829906  | NLI      | 4215     | Parkinson's D.      | 25378390 |
| ANKRD16  | 2075     | atopic asthma       | 17075290 | NRP2     | 2075     | atopic asthma       |        |
| AP5B1    | 4316     | differentiation     | 15146197 | NLS1     | 3063     | kinetochore-protein | 16585270 |
| ARFX     | 2642     | development         | 20565723 | NYK3     | 7447     | hepatocarcinoma     | 26883180 |
| ARHGEF39 | 1307     | tumorogenesis       | 22327280 | OPTN     | 2352     | glaucoma            | 26302410 |
| ARL11    | 1033     | tumorogenesis       | 18337727 | PAG1     | 8156     | prostatic cancer    | 21092590 |
| ATK1     | 2991     | schizophrenia       | 19165527 | PAQR5    | 4435     | ovarian cancer       | 21761364 |
| ATP1A2   | 4410     | tumorogenesis       | 23497007 | PARK2    | 3729     | Parkinson's D.      | 26860075 |
| BCL2L15  | 2650     | apoptosis           | 16690252 | PBD2     | 2077     | hepatocarcinoma     | 26594708 |
| BPNT1    | 1128     | ovarian cancer      | 20628624 | PCGF5    | 5089     | Alzheimer's D.      | 16385451 |
| C10orf40 | 523      | uncharacterized     | 18613    | PCSK5    | 1095     | tumorogenesis        | 21094132 |
| C17orf75 | 2895     | uncharacterized     | 19355679 | PDAP1    | 1926     | proliferation       | 23555679 |
| C17orf75 | 3672     | uncharacterized     | 3221     | PCDC4    | 3221     | tumorogenesis        | 26871813 |
| C21orf58 | 2668     | uncharacterized     | 11707072 | PEX2     | 3056     | cerebellar ataxia   | 21392394 |
| C4orf91  | 2068     | uncharacterized     | 1476     | PGP2     | 1476     | liver cirrhosis     | 25687677 |
| C6orf170 | 4113     | uncharacterized     | 20159594 | PINK2    | 3345     | tumorogenesis        | 26677064 |
| C8orf44  | 336**    | uncharacterized     | 1991     | PMLA1    | 1991     | childhood obesity   | 19390624 |
| C8orf44  | 1626     | uncharacterized     | 1876     | POONL1   | 1876     | uncharacterized     | 12479732 |
| C9orf85  | 871      | uncharacterized     | 4679     | POUDT1   | 4679     | hepatocarcinoma     | 27030260 |
| CACNB2   | 4301     | hypertension        | 25966706 | POLH     | 5550     | ovarian cancer       | 25831546 |
| CACNB2   | 338      | uncharacterized     | 252858   | PPM1K    | 2192     | diabetes mellitus   | 2344828 |
| CACNBG8  | 5006     | uncharacterized     | 5156     | PPP1R12B | 5156     | childhood asthma    | 23640410 |
| CACNBG8  | 7535     | uncharacterized     | 998      | PRRG4    | 998      | Parkinson's D.      | 19772629 |
| CALHM1   | 2896     | Alzheimer's D.      | 26944452 | PSMB2    | 2925     | proteolysis         | 21660142 |
| CCBE1    | 3321     | ovarian cancer      | 19935792 | PTCDC3   | 4116     | osteosarcoma        | 19427859 |
| CCDC114  | 261*     | dyskinesia          | 23506398 | PTK6     | 2233     | tumorogenesis       | 27311570 |
| CD109    | 6841     | bladder cancer      | 20946523 | QKFIIR1  | 1949     | metabolic S.        | 16648250 |
| CD36     | 4042     | atherosclerosis     | 16515687 | RAB11FIP1| 4928     | cell transport      | 26790954 |
| CD68     | 1398     | carcinomas          | 21113139 | RAB13P   | 3975     | tumorogenesis        | 12007189 |
| CDAN1    | 4296     | erythropoiesis      | 19336738 | 7022     | Parkinson's D.      |        |
| CCHR3    | 4878     | asthma              | 25848009 | RAB7L1   | 1693     | Parkinson's D.      | 26914237 |
| CEP68    | 4934     | cervical cancer     | 17507516 | RBBP9    | 1818     | tumorogenesis        | 21933118 |
| CHST5    | 2946     | colon carcinoma     | 12107080 | RGS3     | 205**    | cardiovascular D.   | 24375609 |
| CHST6    | 2979     | dystrophy           | 20592220 | RPS6KA6  | 7136     | tumorogenesis        | 26732474 |
| CHST6    | 3876     |                     | 5871     | SCN11A   | 5871     | neurophathy         | 25791876 |
| CIAO1    | 2416     | tumorogenesis       | 21955663 | SEPT1    | 4033     | hepatocarcinoma     | 20419844 |
| CIAO1    | 3814     |                     | 1575     | SEPT14   | 1575     | Parkinson's D.      | 27115672 |
| CLECT19A | 1747     | lectin              | 12075309 | STG2     | 3142     | lymphopoiesis       | 2158125 |

Table 1: Positions of miR-619-5p binding sites and disease or function of target genes.
Table 1  Positions of miR-619-5p binding sites and disease or function of target genes (Continued)

| Gene   | Position | Disease/Function                  | Accession Number | Disease/Function                  | Accession Number |
|--------|----------|-----------------------------------|------------------|-----------------------------------|------------------|
| CLTC   | 7006     | pancreatic cancer                 | SH3GLB1          | prostate cancer                   | 4856             |
| CORO2A | 2227     | colon cancer                      | SLC15A2          | hepatocarcinoma                   | 4333             |
| COX18  | 1264     | tumorogenesis                     | SLC17A5          | cardiovascular D                  | 2389             |
| CPM    | 2698     | renal carcinoma                   | SLC26A2          | colorectal cancer                 | 5066             |
| CPM    | 4996     |                                    | SLC26A4          | hearing loss                      | 4210             |
| CPT2   | 2557     | sudden death                      | SLC28A2          | chronic hepatitis C               | 2196             |
| CYB5RL | 3426     | transcription                      | SLC7A11          | tumorogenesis                     | 6304             |
| CYP20A1| 2539     | tumorogenesis                     | SLC7A14          | breast cancer                     | 8487             |
| CYP20A1| 4709     |                                    | SNX22            | liver-disease                     | 902              |
| CYP27C1| 3823     | self-rated health                 | SOWAHC           | retrotransposon                   | 3417             |
| CYP2W1 | 2176     | colorectal cancer                 | SPATA13          | colorectal cancer                 | 5020             |
| DAP3   | 1842     | breast cancer                     | SPATA5           | microcephaly                      | 5648             |
| DCAF10 | 3305     | lung cancer                       | SPATS2           | breast cancer                     | 3332             |
| DCAF10 | 4559     |                                    | SNAI2BP          | tumorogenesis                     | 5287             |
| DCLRE1C| 2966     | Omenn syndrome                    | STAC2            | inherited ataxias                 | 2241             |
| DDOST  | 1782     | hyperglycermia                    | SYNS2BP          | breast cancer                     | 1298             |
| DHODH  | 1709     | melanoma                          | SYNS2BP          | tumorogenesis                     | 4175             |
| DHR59  | 1281*    | tumorogenesis                     | TCEB1            | uncharacterized                   | 1964             |
| DNAL1  | 4925     | dyskinesia                        | TGD6             | uncharacterized                   | 3439             |
| DSCR6  | 1706     | Down syndrome                     | TMEM156          | uncharacterized                   | 1593             |
| ERBB3  | 5104     | tumorogenesis                     | TMEM19           | uncharacterized                   | 3510             |
| FADS6  | 1777     | liver disease                     | TMEM213          | uncharacterized                   | 875              |
| FAM161A| 2785     | retinal disease                   | TMEM214          | uncharacterized                   | 1190             |
| FAM227A| 4981     | cancer                            | TMEM50B          | uncharacterized                   | 1026             |
| FAM84B | 3626     | tumorogenesis                     | TMEM56           | nicotin dependence               | 1243             |
| FBUM1  | 2126     | breast cancer,                    | TMF1             | prostate cancer                   | 4736             |
| FBXL22 | 1411     | cardiomyopathy                    | TMCOD2           | bladder cancer                    | 7816             |
| FBXO27 | 1535     | leukemia                          | TNRFS10A         | cancer                           | 1621             |
| FGOD4  | 7619     | cancer                            | TNRFS10D         | cancer                           | 1532             |
| FKBP14 | 1515     | ovarian cancer                    | TOP3A            | leukaemia                        | 3814             |
| FKBP14 | 2129     |                                    | TRPGL1           | uncharacterized                   | 1754             |
| FKBP5  | 7114     | schizophrenia                     | TRPM2            | ischemia                         | 1885             |
| FXN    | 3288     | metabolic disease                 | TRPM7            | neuroblastoma                     | 8079             |
| GDPD1  | 1559     | phosphodiesterase                 | TRPM7            | carcinoma                         | 8221             |
| GEMIN8 | 2172     | neuropathy                        | TXNDC15          | thrombosis                        | 2460             |
| GGT6   | 1956     | ovarian cancer                    | TWYS             | schizophrenia                     | 3692             |
| GKS    | 3808     | glioblastoma                      | UACA             | lung cancer                       | 6120             |
| GKS    | 6355     | glioblastoma                      | UACA             | thyroid diseases                  | 6120             |
| GLB1L  | 2224     | phosphatase                       | UBIAD1           | cancer                           | 2881             |
| GOLGA3 | 7240     | immune disease                    | UBXN2A           | colon cancer                      | 1665             |
| GP2    | 1877     | crohn disease                     | UBXN2A           | colon cancer                      | 1665             |
| GPR65  | 3309     | tumorogenesis                     | UQCRB            | colorectal cancer                 | 1269             |
| GPR65  | 3309     | immune diseases                   | USP29            | protease                          | 2*               |
| GPR82 | 2664    | uncharacterized                   | VHL              | tumorogenesis                     | 3764             |
| GPRIN2 | 6676     | schizophrenia                     | VHL              | uncharacterized                   | 3898             |
Before the 5’ end and after the 3’ end of miR-619-5p binding site, nucleotides are not homologous. The mRNAs of RGS3 and USP29 orthologous genes have binding sites in H. sapiens, N. leucogenys, P. abelii, M. leucophaeus, C. angolensis palliatus, G. gorilla, and R. roxellana.

Table 1 Positions of miR-619-5p binding sites and disease or function of target genes (Continued)

| Species | Gene   | Position of site, nt | Nucleotide sequence |
|---------|--------|----------------------|---------------------|
| Hsa     | GTPBP10| 1873                 | prostate cancer     | 27409348 GWA2 3366 colon cancer 15580307 |
| Hsa     | H6PD   | 5754                 | tumorogenesis       | 15221007 WDR73 1736 microcephaly 25466283 |
| Hsa     | HM13   | 1745                 | glioblastoma        | 28198167 XIAP 5681 ovarian cancer 26779627 |
| Hsa     | IFIT3  | 1864                 | pancreatic cancer   | 25650658 YAE1D1 1548 oral cancer 23318452 |
| Hsa     | ISY1   | 686**                | uncharacterized     | 25667785 ZBTB24 4842 hepatocarcinoma 27730394 |
| Hsa     | IYD    | 1658                 | hypothyroidism.     | 18765512 ZC3H12D 4821 Acute lung injury 26059755 |
| Nle     | KIAA1456| 2536               | colorectal cancer   | 24743840 ZDHHC20 3390 tumorogenesis 20334580 |
| Pab     | KIF11  | 3598                 | tumorogenesis       | 28011472 ZFP30 3463 hypertension 19851296 |
| Mle     | KLHL23 | 2570                 | tumorogenesis       | 23676014 ZNF114 1827 transcription factor 8467795 |
| Nle     | KPN1a  | 5711                 | breast cancer       | 26052702 ZNF197 3446 thyroid cancer 12682018 |
| Nle     | KREME1 | 2199                 | schizophrenia       | 20153141 ZNF320 5534 glioblastoma 11536051 |
| Nle     | KREME1 | 2792                 | schizophrenia       | 20153141 ZNF429 2081** transcription factor 7865130 |
| Nle     | LAX1   | 2057                 | uncharacterized     | 20153141 ZNF445 8820 transcription factor 16368201 |
| Nle     | LILRA6 | 3224                 | myopathy            | 25250574 ZNF626 4620 liver diseases 18255255 |
| Nle     | LIMD1  | 3931                 | breast cancer       | 27656835 ZNF549 3736 transcription factor 16344560 |
| Nle     | LMS5   | 5711                 | cancer              | 27500440 ZNF557 4791 transcription factor 15851553 |
| Nle     | LMOD3 | 3224                 | myopathy            | 25250347 ZNF626 4620 liver diseases 18255255 |
| Nle     | METTL6 | 3933                 | Alzheimer's D       | 2281374 ZNF667 3240 transcription factor 17397802 |
| Nle     | MR1    | 3664                 | hepatocarcinoma     | 26823810 ZNF708 5415 transcription factor 15057824 |
| Nle     | MREG   | 1540                 | pulmonary D         | 20463177 ZNF84 4920 transcription factor 11856868 |

Notes: * - 5’UTR, **- CDS; others – 3’UTR, D - disease

Table 2 Variation of positions and nucleotide sequences of miR-619-5p binding sites in the 5’UTRs of mRNAs of mammal genes

| Species | Gene | Position of site, nt | Nucleotide sequence |
|---------|------|---------------------|---------------------|
| Hsa     | CCDC114 | 261                | GCCUGCCUCCUCGCUUGUAUCUCACGACGUUGG |
| Hsa     | DHR59  | 1281                | GCCUGCCGCLGCCUGCAUGGCUUGUAUCCACGACGUUGG |
| Hsa     | RGS3   | 205                 | GCCGCUGCCGCUGCAUGGCUUGUAUCCACGACGUUGG |
| Pab     | RGS3   | 1                   | GCCGCUGCCGCUUGCUUGUAUCCACGACGUUGG |
| Nle     | RGS3   | 205                 | GCCGCUGGAUCCGCUGCUUGUAUCCACGACGUUGG |
| Hsa     | USP29  | 2                   | CGUGACCAAGGCGCUAGCUGCUUGUAUCCACGACGUUGG |
| Pab     | USP29  | 52                  | CGUGACCAAGGCGCUAGCUGCUUGUAUCCACGACGUUGG |
| Nle     | USP29  | 52                  | CGUGACCAAGGCGCUAGCUGCUUGUAUCCACGACGUUGG |
| Mle     | USP29  | 47                  | CGUGACCAAGGCGCUAGCUGCUUGUAUCCACGACGUUGG |
| Can     | USP29  | 98                  | CGUGACCAAGGCGCUAGCUGCUUGUAUCCACGACGUUGG |
| Ggo     | USP29  | 100                 | CGUGACCAAGGCGCUAGCUGCUUGUAUCCACGACGUUGG |
| Rro     | USP29  | 52                  | CGUGACCAAGGCGCUAGCUGCUUGUAUCCACGACGUUGG |

Notes: In the table 2-5 the bold type indicates the binding site of miR-619-5p
have miR-619-5p binding sites in the 5'UTRs and 3'UTRs, and C8orf44, ISY1, and ZNF714 have miR-619-5p binding sites in the CDSs and 3'UTRs.

The nucleotide sequences of miR-619-5p binding sites are located in the CDSs of the C8orf44, C8H8orf44, ISY1, ZNF429, and ZNF714 genes and encode the following oligopeptides (Table 3). C8H8orf44, C8orf44, and ISY1 genes encode the WLMPVIP oligopeptide, which is also present in the orthologous proteins of P. anubis, P. anubis, P. paniscus, and P. troglodytes. The mRNA of transcription factor ZNF429 and ZNF429 genes binding sites are encoded the AHACNP oligopeptide in the another reading frame. The first two oligopeptides are encoded in one open reading frame (ORF) and the amino acid sequences are highly conserved. The homologous oligonucleotide of the miR-619-5p binding site in the mRNA of ZNF714 gene codes for an oligopeptide in a different ORF.

The presence of miR-619-5p binding sites in the CDSs of five genes with different functions and the evolutionary conservation of these sites signify the role of miRNA in the regulation of the expression of these genes. The nucleotide sequences of specific regions of mRNAs of C8H8orf44, C8orf44, ISY1, ZNF429, and ZNF714 genes that contain miR-619-5p binding sites in the CDSs are homologous among themselves and to the binding sites located in the 5'UTRs and 3'UTRs.

The miRNA binding sites in the coding region, as opposed to the 3'UTR and 5'UTR, clearly demonstrate the relationship between miRNA and mRNA by their conserved amino acid sequences in orthologous proteins. miRNA binding site can be translated by two open reading frames that encode WLTPVIPA and AHACNPS oligopeptides. In the third reading frame, the miR-619-5p binding site has a stop codon. However, in the genes studied, no such sequence was found. In the absence of complete complementarity between miR-619-5p and its binding site, miR-619-5p uses a site containing the corresponding mutation in the CDS for the regulation of gene expression. Thus, a single miRNA binding site in the mRNA of various genes may correspond to three different oligopeptides. Generally, one out of these three oligopeptides is present in the proteins encoded by the orthologous genes.

ISY1 orthologous genes in H. sapiens, P. troglodytes, and N. leucogenys encode a protein containing QVRWLMPVIPALWEAEAGGSQA oligopeptide sequence (Table 4).

However, the RAB43 gene, which is paralogous to human ISY1, lacks the nucleotide sequence encoding the QVRWLMPVIPALWEAEAGGSQA oligopeptide. Additionally, ISY1 gene in the genomes of other animals also lacks the nucleotide sequence encoding this oligopeptide (Table 4).

**Table 4** Amino acid sequences coding in miR-619-5p binding sites in the mRNA of ISY1 gene of orthologous genes

| Species | Gene       | Amino acid sequence                      |
|---------|------------|-----------------------------------------|
| Hsa     | PWRELFEKQVRWLMPVIPALWEAEAGGSQA LPPPRKTRAELMKA |
| Prt     | PWRELFEKQVRWLMPVIPALWEAEAGGSQA LPPPRKTRAELMKA |
| Nle     | PWRELFEKQARWLTPVIPALWEAEAGGSQA LPPPRKTRAELMKA |
| Hsa*    | PGVRELFEKE     |
| Bmu     | PGVRELFEKE     |
| Cja     | PGVRELFEKE     |
| Cfa     | PGVRELFEKE     |
| Cdr     | PGVRELFEKE     |
| Eca     | PGVRELFEKE     |
| Ggg     | PGVRELFEKE     |
| Mmu     | PGVRELFEKE     |
| Nle     | PGVRELFEKE     |
| Cja     | PGVRELFEKE     |
| Cfa     | PGVRELFEKE     |
| Cdr     | PGVRELFEKE     |
| Eca     | PGVRELFEKE     |
| Ggg     | PGVRELFEKE     |
| Mmu     | PGVRELFEKE     |
| Nle     | PGVRELFEKE     |
| Oar     | PGVRELFEKE     |
| Pab     | PGVRELFEKE     |
| Ppa     | PGVRELFEKE     |
| Rho     | PGVRELFEKE     |

* RAB43 - human ISY1 paralog gene
Nucleotide sequences of miR-619-5p binding sites in the mRNAs of ADAM17, ALDH3A2, and ARL11 orthologous genes are shown in Table 5.

These orthologous genes are characterized by highly conserved nucleotide sequence GGCTCATGCCTGTAATCCCAGC of miR-619-5p binding sites. This shows that the interaction of miR-619-5p with mRNAs of these genes is conserved during evolution. Some of the human miR-619-5p target genes and their corresponding orthologous genes have two miR-619-5p binding sites in their mRNAs.

Table 6 shows the nucleotide sequences of two miR-619-5p binding sites in the 3' UTR of mRNAs of ERBB3, FBLIM1, and FKBP14 orthologous genes.

Table 7 shows the degree of conservation of miR-619-5p binding sites in the 201 mRNAs of target genes. All mRNAs with complete complementarity to miR-619-5p binding sites (ΔG/ΔGm is 100%) were divided into four groups, and the frequency of occurrence of nucleotides was determined in each group. The results suggest that miR-619-5p binding sites are highly conserved. The binding site GGCTCATGCCTGTAATCCCAGC does not change and in each of the four gene groups the observed variability of nucleotides on the right and left is high.

**Discussion**

Here we have identified many miRNAs binding sites in the mRNAs of 201 human genes which indicates that miRNAs act as coordinators of gene expression by participating in many biological processes. Previous studies have shown the influences of miRNAs on the expression of genes that encode the transcription factors [19, 20] and on the expression of proteins that participate in the cellular cycle [3, 21–23], apoptosis [4, 24–26], and stress responses [27]. It was shown the role of the mir-619-5p in the regulation of different pathological processes [28]. It was investigated the correlations between the expression of MALAT1 and miR-619-5p, in addition to the association between the clinicopathological features and survival outcomes of patients with stage II and III colorectal cancer tumors [28]. It was observed, that hsa-miR-619-5p and hsa-miR-1184 microRNA expression significantly increased in prostatic cancer. MicroRNA-gene-net analysis indicated that miR-619-5p and other some

| Species | Gene | Position, nt | Nucleotide sequence |
|---------|------|--------------|---------------------|
| Hsa     | ADAM17 | 3466 | TGGGAGTGGTGGCTCATGCCTGTAATCCCAGCCTTGAGAGG |
| Cat     | ADAM17 | 3485 | GGGGCAGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Mmul    | ADAM17 | 3491 | GGGGCAGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Mne     | ADAM17 | 3438 | GGGGCAGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Ptr     | ADAM17 | 3449 | TGGGAGTGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Rro     | ADAM17 | 3425 | GGGGCAGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Hsa     | ALDH3A2 | 2617 | GGGGCAGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Cja     | ALDH3A2 | 3444 | CCGGCGTGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Gga     | ALDH3A2 | 2712 | CCGGCGTGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Mmul    | ALDH3A2 | 2509 | CCGGACATGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Mne     | ALDH3A2 | 2504 | CCGGACATGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Nle     | ALDH3A2 | 2714 | TGGGAGTGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Pab     | ALDH3A2 | 2297 | TGGGAGTGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Ppa     | ALDH3A2 | 2715 | CCGGACATGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Ptr     | ALDH3A2 | 2711 | TGGGAGTGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Rro     | ALDH3A2 | 2727 | CCGGACATGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Hsa     | ARL11  | 1033 | TGGGAGTGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Cat     | ARL11  | 1642 | CAGATGCACTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Mfa     | ARL11  | 1698 | CAGATGCACTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Mmul    | ARL11  | 1747 | CAGATGCACTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Mne     | ARL11  | 1024 | TGGGAGTGGTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Mne     | ARL11  | 1471 | CAGATGCACTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Ptr     | ARL11  | 1353 | CAGATGCACTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
| Rro     | ARL11  | 1254 | CAGATGCACTGGCTCATGCCTGTAATCCCAGCCTTTGGAGG |
Table 6. Variation of nucleotide sequences of two miR-619-5p binding sites in the 3’UTR of mRNAs of ERBB3, FBLIM1, and FKBP14 of orthologs

| Species | Gene | Position, nt | Nucleotide sequence |
|---------|------|--------------|---------------------|
| Hsa     | ERBB3 | 4950         | CGGGCATGGTGCTCATGCTGTGTAATTCAGCAGTTTGGAG |
| Hsa     | ERBB3 | 5104         | TGCCGATGCGTGCATGCTGTGTAATTCAGCAGCTTTGGAG |
| Csa     | ERBB3 | 4989         | CGGGCATGGTGCTCATGCTGTGTAATTCAGCAGTTTGGAG |
| Csa     | ERBB3 | 5149         | TGGGCATGGTGCTCATGCTGTGTAATTCAGCAGTTTGGAG |
| Mfa     | ERBB3 | 5114         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Mfa     | ERBB3 | 5269         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Mmu     | ERBB3 | 5114         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Mmu     | ERBB3 | 5269         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Mne     | ERBB3 | 5112         | CGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Mne     | ERBB3 | 5267         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Pan     | ERBB3 | 5106         | CGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Pan     | ERBB3 | 5274         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Ppr     | ERBB3 | 5105         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Ppr     | ERBB3 | 5243         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Mne     | FBUM1 | 1938         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Mne     | FBUM1 | 5267         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |
| Pab     | FKBP14| 1514         | CAGGCACCGTGCTCACGGCTGTGTAATTCAGCAGTTTGGAG |
| Pab     | FKBP14| 2128         | TGGGCATGCGTGCATGCTGTGTAATTCAGCAGTTTGGAG |

Notes: The black type indicates the binding site of miR-619-5p

Table 7. Variation of nucleotide sequences of mRNA region with miR-619-5p binding sites (See Additional file 1, 2, 3 and 4)

| From | To     | Nucleotide sequence |
|------|--------|---------------------|
| CSL6 | COX18  | GGGTCATGCTGGCTGTGTAATTCAGCAGTTTGGAG |
| GK5  | HM13   | GGGTCATGCTGGCTGTGTAATTCAGCAGTTTGGAG |
| IFIT3| SLC26A4| GGGTCATGCTGGCTGTGTAATTCAGCAGTTTGGAG |
| LC28A2 | ZNF841 | GGGTCATGCTGGCTGTGTAATTCAGCAGTTTGGAG |
miRNAs had the most important and extensive regulatory function for Qi-stagnation syndromes and Qi-deficiency syndromes in coronary heart disease [29].

One or several umiRNAs regulating the expression of hundreds of genes can create a system of interconnected processes in cells and organisms. Such role of these umiRNAs is possible because they circulate in the blood and have access to nearly all cells of an organism [30–32]. Our results provide the basis for studying the systemic roles of unique and normal miRNAs in the regulation of gene expression in human cells. The expression of many target genes is regulated by umiRNAs does not allow individual miRNAs of target genes to be expressed in more degree than others. The greater expression of one mRNA, the larger number of umiRNAs bind to this mRNA. This allows one umiRNA to maintain a certain balance of expression of the corresponding target genes. If umiRNA expression changes, such system is vulnerable. This will cause the development of pathology in the cell, tissue or body.

Conclusions
The majority of miR-619-5p binding sites are located in the 3’UTRs of mRNAs of target genes. Some genes have miRNA binding sites in the 5’UTRs of mRNAs. It is necessary to maintain nucleotide sequences of the binding site of umiRNA in the CDSs of several genes. Different genes have binding sites for miRNAs that are read in different open reading frames. Therefore, identical nucleotide sequences encode different amino acids in different proteins. In encoded proteins, these sites encode conservative oligopeptides. The binding sites of miR-619-5p in 3’UTRs, 5’UTRs and CDSs are conservative in the orthologous mammalian genes.

Additional files

Additional file 1: Figure S1. Variation of nucleotide sequences of mRNA region with miR-619-5p binding sites of genes from CSL6 to COX18 (Conservative binding sites are in bold) (PDF 218 kb)

Additional file 2: Figure S2. Variation of nucleotide sequences of mRNA region with miR-619-5p binding sites of genes from GKS to HM13 (Conservative binding sites are in bold) (PDF 106 kb)

Additional file 3: Figure S3. Variation of nucleotide sequences of mRNA region with miR-619-5p binding sites of genes from IFIT3 to SLC26A4 (Conservative binding sites are in bold) (PDF 139 kb)

Additional file 4: Figure S4. Variation of nucleotide sequences of mRNA region with miR-619-5p binding sites of genes from LC28A2 to ZNF841 (Conservative binding sites are in bold). The data given in the Additional files 1, 2, 3 and 4 demonstrate the variability of the nucleotides before and after the binding sites of miR-619-5p, which is shown in the Weblogo schemes in the table 8. (PDF 151 kb)

Abbreviations
CDSs: Coding domain sequences; miRNAs: Micrornas; ORF: Open reading frame; Ummrna: Unique miRNA

Acknowledgements
We thank LEDUCQ Foundation and the European Union project «Muscle Stress Reliefs». Also we thank PhD Berillo O. for her help in the collection of data.

Funding
This study was supported by a grant (N0491/F04) from the Ministry of Education and Science, Kazakhstan Republic, SRI of Biology and Biotechnology Problems, Al-Farabi Kazakh National University, and Institute for Anaesthesiology and Intensive Operative Care Medical Faculty Mannheim, Mannheim, Germany.

Availability of data and materials
The data sets supporting the results of this article are included within the article and its additional files and publicly available.

Authors’ contributions
SA, RN and AI conceived of the study and drafted the manuscript. SA, RN, AI, SL, AP, IP and AA made substantial contributions to acquisition of data, to interpretation and modification of the data. All authors involved in drafting the manuscript, read and approved the final version of the manuscript.

Competing interests
The authors declares that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1SRI Of Biology and Biotechnology Problems, Al-Farabi Kazakh National University, Almaty, Kazakhstan. 2Institute for Anaesthesiology and Intensive Operative Care Medical Faculty Mannheim, Mannheim, Germany.

Received: 19 March 2017 Accepted: 22 May 2017
Published online: 01 June 2017

References
1. Doxakis E. Principles of miRNA-Target Regulation in Metazoan Models. Int J Mol Sci. 2013;14:16280–902.
2. Tang G. siRNA and miRNA: an insight into RISCs. Trends Biochem Sci. 2005;30:106–14.
3. Luo Q, Li X, Li J, Kong X, Zhang J, Chen L, Huang Y, Fang L. MiR-15a is underepressed and inhibits the cell cycle by targeting CCNE1 in breast cancer. Int J Oncol. 2013;43:1212–8.
4. Li X, Chen YT, Josson S, Mukhopadhyay NK, Kim J, Freeman MR, Huang WC. MicroRNA-185 and 342 inhibit tumorigenicity and induce apoptosis through blockade of the SREBP metabolic pathway in prostate cancer cells. PLoS One. 2013;8:e70987.
5. Qian NS, Liu WH, Lv WP, Xiang X, Su M, Raut V, Chen YL, Dong JH. Upregulated MicroRNA-92b regulates the differentiation and proliferation of EpCAM-positive fetal liver cells by targeting C/EBPβ. PLoS One. 2013;8:e68004.
6. Poethig RS. Vegetative phase change and shoot maturation in plants. Curr Top Dev Biol. 2013;105:125–52.
7. Ling YH, Deng JP, Zhang XD, Wang LJ, Zhang YH, Li YS, Zhang ZJ, Zhang XR. Characterization of microRNAs from goat (Capra hircus) by Solexa deep-sequencing technology. Genet Mol Res. 2013;12:1951–61.
8. Knyazev EN, Fomicheva KA, Mikhalenko DS, et al. Plasma Levels of hsa-miR-619-5p and hsa-miR-1184 Differ in Prostatic Benign Hyperplasia and Cancer. Bull Exp Biol Med. 2016;161:108–11.
9. Hou J, Wang J, Lin C, Fu J, Ren J, Li L, Guo H, Han X, Liu J. Circulating MicroRNA Profiles Differ between Qi-Stagnation and Qi-Deficiency in Coronary Heart Disease Patients with Blood Stasis Syndrome. Evid Based Complement Alternat Med. 2014;2014:926962.
10. Swaminathan S, Suzuki K, Seddiki N, Kaplan W, Cowley MJ, Hood CL, et al. Differential regulation of the Let-7 family of microRNAs in CD4+ T cells alters IL-10 expression. J Immunol. 2012;188:238–46.

11. Ivashchenko A, Berillo O, Pyrkova A, Niyazova R, Atambayeva S. MiR-3960 binding sites with mRNA of human genes. Bioinformation. 2014;10(7):423–7.

12. Ivashchenko A, Berillo O, Pyrkova A, Niyazova R, Atambayeva S. 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;2014:720715.

13. Ivashchenko A, Berillo O, Pyrkova A, Niyazova R. Binding Sites of miR-1273 Family on the mRNA of Target Genes. Biomed Res Int. 2014;2014:620530.

14. Ivashchenko A, Berillo O, Pyrkova A, Niyazova R, Atambayeva S. The binding sites of unique microRNAs in the human mRNAs. J Biotechnol. 2014;185:125–37.

15. National Center for Biotechnology Information. http://www.ncbi.nlm.nih.gov

16. Griffiths-Jones S, Grocock R, von Dongen S, Bateman A, Enright A. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. 2006;34:140–4.

17. Cummins JM, He Y, Leary RJ, Pagliarini D, Diaz Jr LA, Sjoblom T, Barad O, Bentwich Z, Szafranska AE, Labouvier E, Raymond CK, Roberts BS, Juhi H, Kiroter KW, Vogelstein B, Velculescu VE. The colorectal microRNAome. Proc Natl Acad Sci U S A. 2006;103:3687–92.

18. Ple H, Landry P, Benham A, Coarfa C, Gunaratne PH, Provost P. The repertoire and features of human platelet microRNAs. PLoS One. 2012;7:e50746.

19. Cui Q, Yu Z, Pan Y, Punisima EQ, Wang E. MicroRNAs preferentially target the genes with high transcriptional regulation complexity. Biochem Biophys Res Commun. 2007;352(9):733–8.

20. Yan L, Kang M, Qin Z, Zhang W, Li Y, Ou H. An intrinsic 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. 2013;8:e70658.

21. Wang P, Zou F, Zhang X, Li H, Dulak A, Tomko Jr RJ, Lazo JS, Wang Z, Zhang L, Yu J. microRNA-21 negatively regulates Cdc25A and cell cycle progression in colon cancer cells. Cancer Res. 2013;69:8157–65.

22. Cui X, Witalison EE, Chumanevich AP, Chumanevich AA, Poudyal D, et al. The induction of microRNA-16 in colon cancer cells by protein arginine deiminase inhibition causes a p53-dependent cell cycle arrest. PLoS One. 2013;8:e33791.

23. Wang Y, Zheng X, Zhang Z, Zhou J, Zhao G, et al. MicroRNA-149 inhibits proliferation and cell cycle progression through the targeting of ZBTB2 in human gastric cancer. PLoS One. 2012;7:e41693.

24. Wang Y, Lee CG. MicroRNA and cancer focus on apoptosis. J Cell Mol Med. 2009;13:12–23.

25. Li C, Hashimi SM, Good DA, Cao S, Duan W, Plummer PN, Mellick AS, Wei MQ. Apoptosis and microRNA aberrations in cancer. Clin Exp Pharmacol Physiol. 2012;39:739–46.

26. Lima RT, Busaca S, Almeida GM, Gaudino G, Fennell DA, Vasconcelos MH. MicroRNA regulation of core apoptosis pathways in cancer. Eur J Cancer. 2011;47:163–74.

27. Cawley K, Logue SE, Gorman AM, Zeng Q, Patterson J, Gupta S, Samali A. Disruption of microRNA biogenesis confers resistance to ER stress-induced cell death upstream of the mitochondrion. PLoS One. 2013;8:e73870.

28. Qiu G, Zhang X, Liu P, Wu W, Zhang J, Bai S, et al. Dysregulation of MALAT1 and miR-619-5p as a prognostic indicator in advanced colorectal carcinoma. Oncol Lett. 2016;12(6):5036–42.

29. Hou J, Wang J, Lin C, Fu J, Ren J, Li L, Guo H, Han X, Liu J. Circulating MicroRNA Profiles Differ between QI-Stagnation and Q1-Deficiency in Coronary Heart Disease Patients with Blood Stasis Syndrome. Hindawi Publishing Corporation Evidence-Based Complementary and Alternative Medicine. 2014;2014:1–9.

30. Kumar S, Keerthana R, Pazhanimuthu A, Perumal P. Overexpression of circulating miRNA-21 and miRNA-146a in plasma samples of breast cancer patients. Indian J Biochem Biophys. 2013;50:210–4.

31. Smith-Nikos T, Slack F. MicroRNAs circulate around Alzheimer’s disease. Genome Biol. 2013;14(7):125.

32. Reshmi G, Chanda SS, Babu VI, Babu PS, Santhi WS, Ramachandran S, Lakshmhi S, Nair AS, Pillai MR. Identification and analysis of novel microRNAs from fragile sites of human cervical cancer: computational and experimental approach. Genomics. 2011;97:333–40.