DOI 10.18699/SBB-2022-434

piRNAs regulate expression of candidate genes in esophageal adenocarcinoma

Akimniyazova A.N.[2]*, Niyazova T.K.[2], Abdikarim G.M.[2], Sagynbay Sh.[2], Ivashchenko A.T.[1]

[1] *Center for Bioinformatics and Nanomedicine, Almaty, Kazakhstan*

[2] *al-Farabi Kazakh National University, Almaty, Kazakhstan*

** a_ivashchenko@mail.ru*

Key words: piRNA, esophageal adenocarcinoma, candidate genes, diagnostic, treatment

Motivation and Aim: The available information on the biological role of piRNAs (PIWI-interacting RNA) remains at the conjecture stage. The involvement of piRNAs in the regulation of the expression of protein-coding genes is practically not studied. The purpose of this work is to show the possibility of a significant effect of piRNAs on the translation process through the interaction of these molecules with mRNA.

Methods and Algorithms: The nucleotide (nt) sequences of esophageal adenocarcinoma candidate genes were downloaded from NCBI. The nucleotide sequences of 200000 piRNAs were taken from Wang et al [1]. The piRNA binding sites (BSs) in the mRNAs of genes were predicted using the MirTarget program [2]. This program defines the following features of piRNA binding to mRNA: a) the initiation of piRNA binding from the first nucleotide of the mRNAs; c) the schemes of nucleotide interactions between piRNAs and mRNA; d) the free energy (ΔG, kJ/mol) of the interaction between piRNA and the mRNA; e) the ratio ΔG/ΔGm (%) is determined for each site (ΔGm equal to the free energy of piRNA binding with its fully complementary nucleotide sequence).

Results: The protein coding region of the *BTG3* gene includes from 261 to 1151 nucleotides. Binding sites of 33 piRNAs are located from 573 to 665 nt CDS mRNA with overlapping nucleotide sequences. Binding sites piR-12399, piR-37172, piR-44059, piR-126527, piR-129369 start at 626 nt, i.e. approximately in the middle of the complex of binding sites for all piRNA. The length of the CDS is 7.5 times greater than the length of the cluster of binding sites. In the CDS mRNA of the CD55 gene, located from 295 nt to 1617 nt, 18 piRNAs binding sites were identified with the onset from 1420 nt to 1458 nt. That is, the length of the cluster of piRNA binding sites is 63 nt, which is 21 times less than the CDS length. The mRNA of the ERBB2 gene included binding sites for only two piRNAs located in the 5'UTR with overlapping nucleotide sequences. Expression of the ERBB3 gene may depend on 59 piRNAs whose binding sites are located in the 3'UTR from 4934 nt to 5139 nt. These piRNA binding sites formed a cluster from 4934 nt to 5171 nt, 238 nt long. The 3'UTR length was 6.1 times the cluster length. A feature of the cluster is the presence of an identical binding site for seven piRNAs: piR-5300, piR-5303, piR-5358, piR-6236, piR-7637, piR-65119, piR-96686. The mRNA of the *FKBP5* gene contained three clusters of piRNA binding sites located in the 3'UTR. In the first cluster, 27 piRNAs bound from 1366 nt to 1422 nt. The second cluster starts from 4854 nt to 4965 nt and binds 38 piRNAs. This cluster contains five binding sites with the same origin for piRNA at position 4862 nt: piR-56610, piR-62204, piR-75328, piR-92776, piR-102326. These five piRNAs have binding sites in the

third binding sites cluster at 6359 nt which includes 264 piRNAs binding sites from 6356 nt to 7329 nt. Binding sites for seven piRNAs start at position 6397 nt: piR-36976, piR-50827, piR-57845, piR-84920, piR-102738, piR-116971, piR-125671. In the 3'UTR of the mRNA of the *FOXP1* gene, there are only five piR-162173 binding sites, the beginning of which was located two nucleotides apart. This mRNA region contained 20 GU dinucleotide repeats. Nine piRNAs (piR-5300, piR-5303, piR-5358, piR-6236, piR-7637, piR-35905, piR-65119, piR-96686, piR-101606) from the third cluster bound at one position 7115 nt. These include seven piRNAs binding to the mRNA of the *ERBB3* gene. A group of five piRNAs (piR-45567, piR-93145, piR-96134, piR-123755, piR-169382) bound at position 7167 nt. Expression of the *LEP* gene may also depend on piRNAs, because 162 piRNAs can bind to its mRNA, including piR-23387, which is completely complementary due to canonical nucleotides. The binding sites of all piRNAs formed a cluster from 3084 nt to 3385 nt located in the 3'UTR from 562 nt to 3428 nt. The length of the binding sites cluster was 9.8 times less than the length of the 3'UTR. A group of nine piRNAs (piR-5300, piR-5303, piR-5358, piR-6236, piR-7637, piR-35905, piR-65119, piR-96686, piR-101606) with binding sites with 3101 nt coincides with a group of piRNAs that bind to the mRNA of the *FKBP5* gene and includes a group of seven piRNAs that interact with the mRNA of the *ERBB3* gene. The mRNA of the gene contains binding sites for three groups each of five piRs at positions 3087 nt, 3088 nt, and 3246 nt. 46 piRNAs bind to the mRNA of the *SEPP1* gene from 108 nt to 172 nt (65 nt) at a 5'UTR length of 188 nt, i.e., 2.9 times shorter. The length of all 46 binding piRNAs is 1367 nt, and due to the compaction of piRNAs binding sites into clusters, competition between piRNAs for binding in the cluster simultaneously arises. At position 123 nt of the 5'UTR of the mRNA of the *SEPP1* gene, a group of nine piRNAs was identified, identical to the groups of piRNAs that bind to the mRNA of the *FKBP5* and *LEP* genes. The 102 piRNAs bound to the 3'UTR of the *SMAD4* gene mRNA and their binding sites formed two clusters from 4316 nt to 4563 and from 7725 nt to 7778 nt. In the first cluster, three groups of five piRNAs were identified. The first group included piR-36976, piR-50827, piR-84920, piR-116971, piR-125671. The second group consisted of piR-19014, piR-19076, piR-82088, piR-114649, piR-125598, piR-98951. The third group was formed by piR-56309, piR-81517, piR-96803, piR-98951, piR-126140. Both clusters were only 301 nt (54 nt and 247 nt) from the 3'UTR length of 6576 nt, which is 21.8 times the length of the clusters. There were no clusters of binding sites in the mRNA of the *TP53* gene for groups with 5 piRNAs or more. The cluster of 33 piRNAs binding sites with a length of 263 nt was 5.4 times less than the 3'UTR length of 1408 nt. piR-44059 and piR-83728 bound to the mRNA of the *TP53* gene with ΔG kJ/mol equal to –170 and 172 kJ/mol respectively, i.e. they could strongly inhibit the translation process.

Acknowledgements: The study is supported by Center for Bioinformatics and Nanomedicine, Almaty, Kazakhstan.

References

1. Wang J. et al. piRBase: a comprehensive database of piRNA sequences. *Nucleic Acids Res.* 2019;47:175-180. doi: 10.1093/nar/gky1043.
2. Ivashchenko A. et al. MiR-3960 Binding Sites with mRNA of Human Genes. *Bioinformation*. 2014;10:423-427. doi: 10.6026/97320630010423.