



# Article Interactions of piRNAs with the mRNA of Candidate Genes in Esophageal Squamous Cell Carcinoma

Aizhan Rakhmetullina <sup>1,2</sup>, Aigul Akimniyazova <sup>3</sup>, Togzhan Niyazova <sup>4</sup>, Anna Pyrkova <sup>4,5</sup>, Makpal Tauassarova <sup>3</sup>, Anatoliy Ivashchenko <sup>5,\*</sup> and Piotr Zielenkiewicz <sup>1,\*</sup>

- <sup>1</sup> Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02-106 Warsaw, Poland
- <sup>2</sup> Department of Technology of Production of Livestock Products, A. Baitursynov Kostanay Regional University, Kostanay 110000, Kazakhstan
- <sup>3</sup> Higher School of Medicine, Faculty of Medicine and Healthcare, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan
- <sup>4</sup> Faculty of Biology and Biotechnology, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan
- <sup>5</sup> Center for Bioinformatics and Nanomedicine, Almaty 050060, Kazakhstan
- \* Correspondence: a.iavashchenko@gmail.com (A.I.); piotr@ibb.waw.pl (P.Z.)

Abstract: Recently, a database of human piRNAs (piwi-interacting RNAs) was created, which allows the study of the binding of many piRNAs to the mRNAs of genes involved in many diseases, including cancer. In the present work, we identified the piRNAs that can interact with candidate esophageal squamous cell carcinoma (ESCC) genes. The binding of 480 thousand piRNAs with the mRNAs of 66 candidate ESCC genes was studied. Bioinformatic studies found that piRNAs bind only to the mRNAs of nine candidate genes: AURKA, BMP7, GCOM1, ERCC1, MTHFR, SASH1, SIX4, SULT1A1, and TP53. It has been shown that piRNAs can bind to mRNA by overlapping nucleotide sequences in limited 3'UTR and 5'UTR regions called clusters of binding sites (BSs). The existence of clusters of piRNA BSs significantly reduces the proportion of the nucleotide sequences of these sites in the mRNA of target genes. Competition between piRNAs occurs for binding to the mRNA of target genes. Individual piRNAs and groups of piRNAs that have separate BSs and clusters of BSs in the mRNAs of two or more candidate genes have been identified in the mRNAs of these genes. This organization of piRNAs BSs indicates the interdependence of the expression of candidate genes through piRNAs. Significant differences in the ability of genes to interact with piRNAs prevent the side effects of piRNAs on genes with a lack of the ability to bind such piRNAs. Individual piRNAs and sets of piRNAs are proposed and recommended for the diagnosis and therapy of ESCC.

Keywords: esophageal squamous cell carcinoma; gene; mRNA; piRNA; diagnosis

## 1. Introduction

After the sequencing of the human genome, methods of the molecular studies of the regulation of gene expression and the human genome began to develop. piRNAs (piwi-interacting RNAs) were identified more than twenty years ago [1,2]. In addition to these small RNAs, siRNAs (small interfering RNAs) capable of regulating gene expression were created [3,4]. For many years, due to a number of misconceptions, miRNAs and piRNAs were not widely used for practical purposes [5–9]. However, in vivo and in vitro experiments have successfully used siRNAs against undruggable targets for treating cancer and other diseases [10–12]. Therefore, candidate genes responsible for the development of many diseases have been identified [13–17].

This approach allows us to specify disease causes and perform diagnosis and develop therapies more effectively. The identification of candidate genes further requires the identification of factors regulating their expression. It has been shown that piRNAs found in human and animal cells are associated with gene expression, but it has not been established how these molecules can modify this process [6,17,18].



Citation: Rakhmetullina, A.; Akimniyazova, A.; Niyazova, T.; Pyrkova, A.; Tauassarova, M.; Ivashchenko, A.; Zielenkiewicz, P. Interactions of piRNAs with the mRNA of Candidate Genes in Esophageal Squamous Cell Carcinoma. *Curr. Issues Mol. Biol.* 2023, 45, 6140–6153. https://doi.org/ 10.3390/cimb45070387

Academic Editor: Quan Zou

Received: 19 June 2023 Revised: 16 July 2023 Accepted: 21 July 2023 Published: 23 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). The asymptomatic nature of early esophageal cancer and late clinical manifestations result in poor prognosis and limited therapeutic success. Efforts to identify diagnostic/prognostic markers in clinics have not been successful. Thus, there is an urgent need to develop novel non-invasive biomarkers for the early diagnosis of esophageal cancer. Esophageal squamous cell carcinoma (ESCC) involves genes that encode regulators of numerous cellular processes, such as methylenetetrahydrofolate reductase (*MTHFR*), cell cycle control—(tumor protein p53 (*TP53*), Aurora kinase A (*AURKA*), bone morphogenetic protein-7 (*BMP7*), and others. Numerous genetic variations of these genes have been linked to ESCC during the past ten years using candidate gene approaches [19–22].

In previous work, we studied the piRNA interactions with mRNAs for esophageal adenocarcinoma, multiple sclerosis, and SARS-CoV-2 and found that these interactions can be evaluated by quantitative characteristics [23–25]. Many human genes are functionally related, so a single gene can affect several diseases. piRNA is a relatively new field of study, but recent findings have uncovered its crucial role in disease development. Specifically, piRNAs are involved in the regulation of gene expression, which can lead to the abnormal growth and spread of cancer cells [26]. Several publications have established piRNAs' role in carcinogenesis [27–29]. An important fact is the possibility of preserving and transporting small RNA in the human body as part of exosomes and nanosized vesicles through the bloodstream [30–33]. After a database of piRNAs was created [34], it became possible to study the direct interaction of piRNAs with genes. The more than eight million piRNA molecules available in the database are difficult and expensive to study in wet experiments, so it is necessary to use bioinformatic approaches to identify the biological function of piRNAs more efficiently. These studies have shown the promise of this approach in determining the properties of piRNAs and their biological role. This work aimed to identify piRNAs that can interact with candidate esophageal squamous cell carcinoma (ESCC) genes and how this information could be used to diagnose the disease.

#### 2. Materials and Methods

The 66 candidate ESCC genes' nucleotide sequences (Table S1, see Supplementary Materials) were downloaded from the National Center for Biotechnology Information (NCBI) website (accessed on 5 October 2022). The 480 thousand piRNAs' nucleotide sequences were taken from Wang et al. [34]. The MirTarget program was used to predict the piRNA binding sites (BSs) in the mRNA [35]. The following characteristics of piRNA binding to mRNA are predicted by this program: the localization of the piRNA BSs in the 5'-untranslated region (5'UTR), coding domain sequence (CDS), and 3'-untranslated region (3'UTR) of the mRNAs; the schemes of nucleotide interactions between piRNAs and mRNA; and the free energy of the interaction between piRNAs and the mRNA ( $\Delta G$ , kJ/mol); for each BS, the ratio  $\Delta G/\Delta Gm$  (%) was calculated. The  $\Delta Gm$  equals the free energy of piRNA binding with its fully complementary canonical nucleotide sequence. The calculated data were used only to choose piRNAs whose nucleotides interacted with mRNA utilizing canonical (G-C and A-U) and noncanonical (G-U and A-C) nucleotides with a specific  $\Delta G$ value. The MirTarget program studies the hydrogen bonds between piRNAs and mRNA based on the physicochemical properties of nucleotide interactions [35–38]. MirTarget differs from other programs in terms of finding piRNA BSs on mRNA in the following aspects: it takes into account the interaction of piRNA with mRNA over the entire piRNA sequence; it considers noncanonical G-U and A-C pairs; and it calculates the free energy of the interaction of the piRNAs with mRNA. It is notable that under identical circumstances, the G, A, C, and U nucleotides that make up the RNA structure of bacteria, plants, and mammals interact in the same way. Therefore, the physicochemical properties of canonical and noncanonical nucleotide pairs described above do not need additional proof of the previously established physicochemical characteristics of their interaction [36–39].

## 3. Results

In examining the interaction of piRNAs with mRNA genes, criteria were established for the following characteristics. The interaction of piRNAs and mRNA nucleotides must be along the entire length of the piRNAs. The free energy of the interaction must be at least 90% of the maximum  $\Delta G/\Delta Gm$ . Splitting piRNAs into three groups based on the value of  $\Delta G$  and  $\Delta G/\Delta Gm$  of the interaction of piRNAs with mRNA will determine whether the values of  $\Delta G$  and  $\Delta G/\Delta Gm$  within the selected limits can reflect the basic properties of the interaction of piRNAs with mRNA. The property of candidate genes to interact with different numbers of piRNAs is acquired by selecting interrelated genes whose expression should be regulated by piRNAs as universal controllers of genome-wide expression. We determined whether piRNAs interact with specific individual mRNA sites and clarified the role of clusters of piRNAs BSs as a necessity for the coordinated regulation of gene expression with a requirement for compactness of gene regulatory sites. Bioinformatic technologies greatly facilitate the characterization of associations between piRNAs and mRNA interactions.

A study of 480 thousand piRNAs binding to the mRNAs of 66 candidate ESCC genes showed that the mRNAs of only nine genes could interact with piRNAs. The number of interacting piRNAs for each gene differed significantly. To elucidate the interaction of piRNAs with mRNAs, the interacting piRNAs were divided into three groups according to the free energy of interaction: piRNAs interacting with a  $\Delta$ G value of -170 kJ/mol or higher (group 1), with a  $\Delta$ G value of -160 kJ/mol to -169 kJ/mol (group 2), and with a  $\Delta$ G value of -140 kJ/mol to -159 kJ/mol (group 3). This approach allowed us to clarify how the interacting groups of piRNAs may vary.

#### 3.1. Characteristics of piRNAs' Interaction with the mRNA of the AURKA Gene

The mRNA of the AURKA gene consists of 566 nt of 5'UTR, 1212 nt of CDS, and 770 nt of 3'UTR regions. The start of the piRNAs' group 1 BSs is localized only in the 5'UTR from 410 nt to 533 nt (Table S2). The BSs of 43 piRNAs were located with overlapping nucleotide sequences, that is, they formed a cluster of BSs and were localized at the end of the 5'UTR. The BSs' cluster was 157 nt long and contained BS piRNAs with lengths of 33 nt and 34 nt. Group 2 consisted of 38 piRNAs whose BSs were localized in the BS cluster starting from 414 nt to 534 nt with a length of 153 nt (Table S2). The piRNAs of group 2 varied in length from 31 nt to 34 nt. In this cluster of BSs also located in the 5'UTR, 38 piRNAs from piR-49581 to piR-363915 were bound. The piRNAs from group 3 were from 27 nt to 33 nt in length. Group 3 consisted of 60 piRNAs whose BSs were localized in a cluster of BSs (410 nt to 536 nt with a length of 156 nt). Consequently, this cluster of BSs also ended at the 5'UTR before the CDS. The piRNAs from this group were 27 nt to 33 nt long. Group 3 included piRNAs in the range from piR-30739 to piR-480279. The BSs of 141 piRNAs of the three groups were in a cluster 156 nt long with 1.1 nt per piRNA, indicating a high density of BSs for piRNAs. Figure 1 shows the nucleotide sequences of four piRNAs that interacted with a cluster of BSs located in the 5'UTR mRNA of the AURKA gene with overlapping nucleotide sequences. From these data, we can see how piRNAs will compete for binding in the BS cluster.

5 ' - GAUUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAUUACA-3 '	5221	nt r	RNA AURKA	
3 ' -CUA <b>G</b> GAGGACGGAGUCGG <mark>G</mark> GGG <mark>UC</mark> CAUCGACC-5 '	522	nt	piR-12852	9
3 ' - GAGGAUAGAGUCGGAGGGUUCAUUGACCCUGAU-5 '	526	nt	piR-12316	2
3 ' -AGGACGGAGUCGGAGGGCUCAUCAACCCUGGUGU-5 '	527	nt	piR-12123	4
3 ' - GUGGAGUCGGAGGACUCAUCGACUCUGAUAU-5 '	530	nt	piR-11965	4

**Figure 1.** Nucleotide sequences of BSs of piRNAs and mRNA of the *AURKA* gene in the region from 522 nt to 560 nt. Note: The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green. The piRNAs' names are followed by the start of their BSs.

## 3.2. Characteristics of piRNAs' Interaction with the mRNA of the BMP7 Gene

The mRNA of the *BMP7* gene includes a 5'UTR of 529 nt, a CDS of 1296 nt, and a 3'UTR of 2207 nt. piRNAs from all three groups bound only in the 3'UTR (Table S3). In group 1, the BSs of 40 piRNAs ranged from 2756 nt to 2889 nt, and the length of the BS cluster was 166 nt. The length of the binding piRNAs was 33 nt and 34 nt. Forty piRNAs from piR-43404 to piR-440219 were bound in the cluster. Within the large cluster was a 56 nt BS cluster for eight piRNAs with a  $\Delta G/\Delta Gm$  value of 98% to 99% and a free energy value of -170 kJ/mol to -178 kJ/mol (Figure 2). This corresponds to the formation of predominantly canonical nucleotide pairs except for two noncanonical pairs in the interaction of piR-121964 and piR-265442 and one noncanonical pair in each of the remaining piRNAs. Group 2 included 67 piRNAs of length 31 nt to 34 nt that bound in a cluster (2729 nt to 2970 nt with length 273 nt). Group 3 included 111 piRNAs 28–34 nt long; the beginning of the BSs was located from 2733 nt to 2981 nt, and the BS cluster length was 279 nt. The BSs of the piRNAs were unevenly distributed in the clusters, with a total of 1.3 nt per piRNA in the cluster from 2729 nt to 3012 nt for 218 piRNAs of the three groups.

#### piRNA; start BS, nt; $\Delta G, kJ/mol; \Delta G/\Delta Gm, \%;$ piRNA length, nt

piR-121964; 2793; -178; 98; 34	piR-265442; 2794; -178; 98; 34
5 ' - UGAGGUCAGGAGUUCGAGACCAGCCUGGCCAACA-3 '	5 ' - GAGGUCAGGAGUUCGAGACCAGCCUGGCCAACAU-3
3 ' - ACUUCAGUCCUUAAGCUCUGGUCGGACCGGUUGU-5 '	3 ' - CUCCAGUCUUCAAGCUCUGGUCAGACCGGUUGUA-5
piR-98746; 2797; -178; 99; 34	piR-102574; 2801; -176; 99; 33
5'-GUCAGGAGUUCGAGACCAGCCUGGCCAACAUGGU-3'	5 ' - GGAGUUCGAGACCAGCCUGGCCAACAUGGUGAA-3 '
3 ' - CAGUCCUCAAGCUCUGGUUGGACCGGUUAUACCA-5 '	3 ' - UCUCAAGCUCUGGUCGGACCGGUUGUACCACUU-5 '
piR-440219: 2801: -176: 99: 34	piR-215165: 2809: -172: 98: 33
pin 110110, 1001, 1,0, 00, 01	F=:: =====; ==; ==; ==; ==;
5'-GGAGUUCGAGACCAGCCUGGCCAACAUGGUGAAA-3'	5'-AGACCAGCCUGGCCAACAUGGUGAAACCCCAUC-3'
5'-GGAGUUCGAGACCAGCCUGGCCAACAUGGUGAAA-3'	5'-AGACCAGCCUGGCCAACAUGGUGAAACCCCAUC-3'
5'-GGAGUUCGAGACCAGCCUGGCCAACAUGGUGAAA-3' 	5'-AGACCAGCCUGGCCAACAUGGUGAAACCCCAUC-3' 
5'-GGAGUUCGAGACCAGCCUGGCCAACAUGGUGAAA-3' 	5'-AGACCAGCCUGGCCAACAUGGUGAAACCCCAUC-3' IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII

**Figure 2.** Nucleotide sequences of piRNAs and the characteristics of their interaction with the mRNA of the *BMP7* gene. The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green.

#### 3.3. Characteristics of piRNAs' Interaction with the mRNA of the ERCC1 Gene

The mRNA of the *ERCC1* gene consists of 146 nt of 5'UTR, 822 nt of CDS, and 2343 nt of 3'UTR in terms of length. The start of the BSs of 47 piRNAs of group 1 with  $\Delta$ G values varying from -170 kJ/mol to -180 kJ/mol was located in the 3'UTR from 2597 nt to 2794 nt (a cluster with a length of 228 nt; Table S4). The length of the piRNAs varied from 32 nt to 34 nt. Forty-nine piRNAs of group 2 bound to the mRNA. The BSs of these piRNAs were also localized in the 3'UTR from 2592 nt to 2796 nt. The BS cluster of 49 piRNAs was 235 nt long, accounting for 10% of the 3'UTR length; in this cluster of BSs, piRNAs with numbers from piR-41630 to piR-365910 bound. The beginning of piRNAs' BSs in group 3 was 2595 nt to 2785 nt, and the length of the cluster of BSs was 221 nt. Consequently, the clusters of BSs of the three groups of piRNAs were close, and each of the 158 piRNAs averaged 1.5 nt. piR-289349, piR-298414, and piR-436360 bound with mRNA of the *ERCC1* gene fully complementary, forming canonical nucleotide pairs only, and the  $\Delta$ G/ $\Delta$ G value was 100% (Table S4). An example of a very compact arrangement of piRNAs' BSs to the mRNA of the ERCC1 gene is shown in Figure 3. In the 35 nt cluster, seven piRNAs were bound predominantly by canonical nucleotide pairs.

AUUCUCCUGCCUCAGCCUCUCAAGUAGCUGGGAU-3 ' 26'	4 nt	mRNA ERCC1
JAAGAGGGCGGAGUCGGAGAGCUCAUCGACC-5 ' 26	4 n	piR-257600
JAAGAGGAUGGAGUCGGAGGGUUCAUCGACC-5 ' 26	4 n	piR-272067
JAAGGGGACGGAGUCGGAGGGUUCAUCGACCC-5 ' 26	/5 nt	piR-445186
JAAGAGGACGGAGUCGAAGGGUUCAUCGACCC-5' 26	/5 nt	piR-454666
JAAGAGGACGGAACCGGAGGGUUCAUCGACCCU-5 ' 26	6 n	piR-249982
JAAGAGGACGGAGUCGGAAAAUUCAUCGACCCUA-5' 26	6 nt	piR-282903
-AAGAGGACGGAGUCGGGGGGGUUCGUCGACCCUG-5 ' 26	7 nt	piR-320670

**Figure 3.** Nucleotide sequences of the BSs of piRNAs and the mRNA of the *ERCC1* gene in the region from 2674 nt to 2710 nt. Note: The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green. The piRNAs' names are followed by the start of their BSs.

3.4. Characteristics of piRNAs' Interaction with the mRNA of the GCOM1 Gene

The nucleotide sequence of the *GCOM1* gene consists of 131 nt of 5'UTR, 1653 nt of CDS, and 2879 nt of 3'UTR in terms of length. An analysis of the interaction parameters of the 16 piRNAs of group 1 with the mRNA of the *GCOM1* gene shows that these piRNAs bind from 2458 nt to 2466 nt in the 3'UTR only (Table S5). The beginning of the BS cluster of these piRNAs with a length of 33–34 nt begins at 2458 nt and ends at position 2499 nt, i.e., its length is 41 nt. This cluster of BSs binds piRNAs with numbers from piR-56610 to piR-465348. All of the starts of the BSs of the 20 piRNAs of group 2 are localized by mRNA in the site from 2453 nt to 2486 nt in the 3'UTR with a BS cluster length of 66 nt. In this cluster of BSs, piRNAs with numbers from piR-34813 to piR-458324 bind. The length of the interacting piRNAs varied from 30 to 34 nt. The 14 piRNAs of group 3 bound from 2456 nt to 2488 nt in a cluster with a length of 61 nt. In this cluster of BSs, piRNAs bind with numbers ranging from piR-117518 to piR-468536. On average, 1.3 nt of the 50 piRNAs interact with the mRNA of the *GCOM1* gene in the 66 nt cluster. piRNAs less than 30 nt in length can also interact with the *GCOM1* gene (Figure 4). The  $\Delta G/\Delta Gm$  value was higher than 92%, indicating the preferential formation of canonical nucleotide pairs.

nt

**Figure 4.** Nucleotide sequences of piRNAs and characteristics of their interaction with the mRNA of the *GCOM1* gene. The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green.

#### 3.5. Characteristics of piRNAs' Interaction with the mRNA of the MTHFR Gene

The mRNA of the *MTHFR* gene has a long 3'UTR of 4951 nt, unlike the mRNA of the other candidate ESCC genes studied. The 5'UTR and CDS consist of 229 nt and 1971 nt, respectively. Group 1 piRNAs bound to the 3'UTR in two clusters of BSs (Table S6). The first cluster of BSs contained BSs from 6182 nt to 6385 nt, and the second from 6849 nt to 7051 nt. The distance between the second cluster's beginning and the first cluster's end

was 464 nt. Accordingly, 16 and 29 piRNAs were bound in these clusters of BSs. The length of both clusters was 235 nt. The first cluster bound piRNAs from piR-47672 to piR-456736 and the second cluster from piR-36604 to piR-475473. Group 2 piRNAs also bound in two clusters of BSs: the beginning BSs of the first cluster were from 6190 nt to 6383 nt, and the second cluster was from 6849 nt to 7062 nt. The interval between the end and beginning of the clusters was 466 nt. In the first and second clusters, 12 and 53 piRNAs were bound, respectively. Group 3 piRNAs also bound in two clusters of BSs (Table S6). The first cluster of BSs began at 6177 nt and ended with piRNAs binding at the 6388 nt position. It bound 25 piRNAs, and the second cluster bound 78 piRNAs. The second cluster started at 6848 nt, and the last piRNAs were bound at position 7056 nt. As a result, 53 piRNAs of the three groups were bound in the first cluster and 160 piRNAs were bound in the second cluster. The total length of the first cluster for the three groups of piRNAs was 240 nt and that of the second cluster 246 nt. As a result, the first cluster had 4.5 nt per piRNAs and the second cluster had 1.5 nt. Both clusters of BSs were localized at the end of the 3'UTR mRNA of the MTHFR gene. Figure 5 shows one of several clusters of BSs formed from the BSs of four competing piRNAs located one to three nucleotides apart.

5 ' - GGCUCACACCUGUAAUCCCAGCACUUUGGGAAG-3 '	6861	nt mRNA MTHFR
3 ' - UCGAGUACAGACAUUAGGGUCGUGAAACCCUC-5 '	6861	piR-108567
3 ' - CGAGUGUAGAUAUUAGGGUCGUGAAA - 5 '	6862	piR-101606
3 ' - GGGUGCGGACAUUAGAGUCGUGAAACCUUCC-5 '	6863	piR-100284
3 ' - AGUGCAGACAUUAGGGUUGUGAAACUCU-5 '	6864	piR-114722

**Figure 5.** Nucleotide sequences of the BSs of piRNAs and the mRNA of the *MTHFR* gene in the region from 6861 nt to 6893 nt. Note: The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green. The piRNAs' names are followed by the start of their BSs.

#### 3.6. Characteristics of piRNAs' Interaction with the mRNA of the SASH1 Gene

The mRNA of the *SASH1* gene in the 5'UTR, CDS, and 3'UTR contained 475 nt, 3744 nt, and 3491 nt, respectively. The 63 piRNAs of group 1 had the beginnings of BSs only in the 3'UTR from 5490 nt to 5741 nt (Table S7). The piRNA interval from piR-26543 to piR-459543 indicates the participation of piRNAs from the entire studied range of 480 thousand piRNAs. The 61 piRNAs in group 2 had the beginnings of BSs in the 3'UTR from 5507 nt to 5747 nt. The piRNAs from piR-40007 to piR-355919 were associated with this cluster. The piR-50295 that bound to the mRNA of the *SASH1* gene is fully complementary with a  $\Delta G/\Delta G$  value of 100%. The 73 piRNAs of group 3 had the beginnings of BS in the 3'UTR from 5491 nt to 5741 nt. In this cluster, piRNAs from piR-21000 to piR-466023 were bound. The length of the cluster common to the three groups of piRNAs from 5490 nt to 5777 nt was 287 nt in which 197 piRNAs were bound, hence 1.5 nt per piRNAs. Figure 6 shows diagrams of the interaction of several piRNAs with the mRNA of the *SASH1* gene, indicating their binding using canonical and noncanonical nucleotide pairs.

## 3.7. Characteristics of piRNAs' Interaction with the mRNA of the SIX4 Gene

The mRNAs of the *SIX4* gene in the 5'UTR, CDS, and 3'UTR contained 60 nt, 2346 nt, and 3870 nt, respectively. There were 86 piRNAs in group 1 whose beginning of the BSs was located only in the 3'UTR from 4026 nt to 4266 nt (Table S8). In this interval, piRNAs from piR-38669 to piR-479539 bound to mRNAs. The 83 piRNAs of group 2 bound from 4021 nt to 4259 nt. The corresponding piRNAs were between piR-40921 and piR-414045. There were 105 piRNAs in group 3 with the beginning of their BSs also located in the 3'UTR from 4013 nt to 4269 nt. The length of this BS cluster was 287 nt. The 105 piRNAs included piR-20572 to piR-480249. The number of the three groups of piRNAs was 274, which were bound in a 287 nt cluster with a high density of 1.1 nt per piRNA. Figure 7

shows the nucleotide sequences of the piRNAs and the mRNA of the *SIX4* gene in one of the clusters, clearly indicating a fully complementary interaction between their nucleotides. These results clearly indicate high competition among 15 piRNAs for binding in this cluster. If any piRNA is present in concentrations an order of magnitude higher than the others, it will dominate in the suppression of mRNA translation.

```
piRNA; start BS, nt; region mRNA; \Delta G, kJ/mol; \Delta G/\Delta Gm, ; piRNA length, nt
piR-111484; 5477; 3'UTR; -149; 97; 31
                                  piR-101898; 5490; 3'UTR; -170; 93; 34
5'-UUUUUUUUUUUGAAAUGGAGUUUCGCUCUUG-3'
                                  5 ' - AAUGGAGUUUCGCUCUUGUCACCCAGGCUGGAGU-3 '
  3'-AAAAAAAAAAAACUUUGCCUCAAAGCGAGAAC-5'
                                  3'-CUACCUCAAAGCAAAAACAGCGGGUCCGACCUCA-5'
piR-108578; 5516; 3'UTR; -163; 92; 32
                                  piR-112090; 5561; 3'UTR; -149; 89; 31
5'-GCUGGAGUGCAAUGGCACGAUCUUGGCUCACU-3'
                                  5'-CCCAGGUUCAAGUGAUUCUCCUGCCUUAGCC-3'
  3 '-CGACCUCACGUUACCGCGCUAGAGUCGAGUGG-5 '
                                  3'-GGACCCAAGUUUGCUAAGAAGACGGAGUCGG-5'
piR-110734; 5566; 3'UTR; -163; 90; 34
                                  piR-112665; 5579; 3'UTR; -161; 95; 30
5'-GUUCAAGUGAUUCUCCUGCCUUAGCCUCCCGAGU-3'
                                  5'-UCCUGCCUUAGCCUCCCGAGUAGCUGGGAU-3'
  3 ' - CAAGUUCGUCAAGAAGACGGAGUCGGAGGGUUCA-5 '
                                  3 '-GGGACGGAGUCGGAGGGCUCAUCGACCCUA-5 '
                                  piR-143064; 5738; 3'UTR; -140; 88; 30
piR-148974; 5663; 3'UTR; -159; 89; 32
5'-GUUUCACCAUGUUGGCCAGGCUGGUCUCAAAC-3'
                                  5'-AGUGCUGGGAUUACAGGUGUGAGCCACCGC-3'
  3'-CAAAGUGGUGCGACCGGUCUGACCAGGGCUUG-5'
                                  3'-UCACGACUCUGAUGUCCGUACUUGGUAACG-5'
```

**Figure 6.** Nucleotide sequences of piRNAs and the characteristics of their interaction with the mRNA of the *SASH1* gene. The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green.

AUUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGACUACA-3 '	4114	nt	mRNA SIX4
UAAGAGGACAGAGUCGGAGGGCUCAUUGACCC-5 '	4114	nt	piR-432591
UAAGAGGGCGGAGUCGGAGGACUCAUCGACUCUA-5 '	4115	nt	piR-219325
UAAGAGGACGGAACCGGAGGGUUCAUCGACCCU-5 '	4115	nt	piR-249982
-AGAAGGACGGAGUUGGAGGGUUCAUCGACCCUG-5 '	4115	nt	piR-341211
-AAGAGGACAGAGUCGGAGGACUUAUCGACCCUGA-5 '	4116	nt	piR-69165
-AAGAGGACGGAGUCGGGGGGUUCGUCGACCCUG-5 '	4116	nt	piR-197774
-AAGAGGACGGAGUCGGGGGGCUCAUCGACCCUA-5 '	4116	nt	piR-320670
-AAGAGGACGGAGUCGGGGGGCUCAUCGACCCUA-5 '	4116	nt	piR-424861
-AAGAGGACGAGGUCGGAGGACUCAUCGACCCUG-5 '	4116	nt	piR-479539
- AGAGGGCGGAGUCGGAGGGCUCGUCGGCCCUAA-5 '	4117	nt	piR-347859
3 ' -GAGGGCGGAGUCGGAGGGCUCAUUGACCCUGG-5 '	4118	nt	piR-321692
3 ' - AGGACAGAGUCGGAGGGUUCAUCGACCCUAAUGU-5 '	4119	nt	piR-325698
3 ' - GGACGGAGUCGGAGGGCUCAUCGACUCUAGUGU-5 '	4120	nt	piR-67883
3 ' - GGACGGAGUCGGGGGGUUCAUCGACCCUGAUGU-5 '	4120	nt	piR-200086
3 ' - GGGCGGAGUCGGAGGGUUCAUCGACUCUGGUGU-5 '	4120	nt	piR-302301

**Figure 7.** Nucleotide sequences of the BS cluster in the 3'UTR of the mRNA of the *SIX4* gene from 4114 nt and the nucleotide sequences of 15 piRNAs. Note: The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green. The piRNAs' names are followed by the start of their BSs.

### 3.8. Characteristics of piRNAs' Interaction with the mRNA of the SULT1A1 Gene

The nucleotide sequence of the *SULT1A1* gene consisted of 473 nt 5'UTR, 888 nt CDS, and 231 nt 3'UTR in terms of length. An analysis of the interaction of the eight piRNAs from group 1 with the mRNA of the *SULT1A1* gene shows that these piRNAs started binding from 1483 nt to 1510 nt only in the 3'UTR (Table S9). In this cluster of BSs, piRNAs with numbers ranging from piR-10749 to piR-475473 bound. All of the beginnings of ten piRNAs BSs from group 2 were localized in the mRNA at the 1482 nt to 1521 nt site in the 3'UTR with a BS cluster length of 72 nt. In this cluster of BSs, piRNAs with numbers from piR-200742 to piR-472871 bound. The 11 piRNAs of group 3 were bound from 1489 nt to 1517 nt in a cluster with a length of 69 nt. In this cluster of BSs, piRNAs with numbers ranging from piR-104710 to piR-447258 bound. On average, one of the 29 piRNAs interacting with the mRNA of the *SULT1A1* gene in the 71 nt cluster was 2.4 nt. The BS cluster of all piRNAs was located at the end of the 3'UTR. The small number of piRNAs that bound to the mRNA of the *SULT1A1* gene is probably due to the short length of the 3'UTR. The low density of piRNAs' BSs in the mRNA of the gene revealed only one cluster for the binding of three piRNAs, with a significant shift in the beginning of the BSs (Figure 8).

5'-GAGCCCAGGAGUUCGAGACCAACCUGGGCAACAUAGU-3'	1507	nt	mRNA
3 ' - <b>U</b> UCG <b>A</b> GUCCUCAAGUUCUGGUCGGACCCGUUGUA-5 '	1507	nt	piR-107049
3 ' - UCCUUAAGUUCUGGUUGGACCCGUUGUGUCA-5 '	1513	nt	piR-104710
3 ' - UCAAACUCUGGUCAAACCCGUUGUAUCA-5 '	1516	nt	piR-113554

**Figure 8.** Nucleotide sequences of the BSs of piRNAs and the mRNA of the *SULT1A1* gene in the region from 1507 nt to 1543 nt. Note: The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green. The piRNAs' names are followed by the start of their BSs.

## 3.9. Characteristics of piRNAs' Interaction with the mRNA of the TP53 Gene

The mRNAs of the TP53 gene mRNAs in the 5'UTR, CDS, and 3'UTR contained 197 nt, 1041 nt, and 1409 nt, respectively. In group 1, there were 19 piRNAs whose beginning BSs were located only in the 3'UTR from 2293 nt to 2512 nt (Table S10). In this interval, mRNAs bound piRNAs from piR-44059 to piR-455123. The 30 piRNAs of group 2 bound from 2292 nt to 2516 nt. The corresponding piRNAs were between piR-40367 and piR-291404. Group 3 had 42 piRNAs beginning with their BSs in the 3'UTR from 2290 nt to 2533 nt. The 42 piRNAs included piR-32165 through piR-365438. The number of piRNAs of the three groups was 91, bound in a cluster 274 nt long with a density of 3.0 nt per piRNA. It should be noted that there were portions of the cluster that did not bind piRNAs between 2301 nt and 2382 nt (length 81 nt) and between 2390 nt and 2456 nt (length 66 nt); the cluster was reduced by 147 nt, and the working part of the cluster was 127 nt with 1.4 nt per piRNA, which is comparable to the results obtained for other candidate genes. Two BSs at 163 nt intervals of ten piRNAs binding at positions 2293 nt and 2456 nt were identified in the mRNA of the TP53 gene. These are piR-101480 and piR-69443 from group 1 piRNAs; piR-106322, piR-53597, and piR-62868 from group 2 piRNAs; and piR-32165, piR-37715, piR-52128, piR-57508, and piR-32194 from group 3. This is explained by the presence of two sites in the 3'UTR mRNA of the TP53 gene with identical nucleotide sequences binding these piRNAs. The relatively small number of piRNAs interacting with the mRNA of the gene (Table S10) is reflected in the absence of large clusters of BSs in the mRNA of the *TP53* gene. However, the binding of piRNAs to the mRNA was efficient (Figure 9).

piRNA; start BS,nt; region mRNA; $\Delta$	AG,kJ/mol; ΔG/ΔGm,%; piRNA length,nt
piR-107204; 2341; 3'UTR; -140; 90; 28	piR-111518; 2342; 3'UTR; -159; 89; 31
5 ' -AUCUUGGCUUACUGCAGCCUUUGCCUCC-3 '	5'-UCUUGGCUUACUGCAGCCUUUGCCUCCCCGG-3'
3 ' - CAGAACCGAGUGACGUUGGAGACGGAGG-5 '	3 ' -AGAGCCGAGUGACGUUGGAGGCGGAGGGGGCC-5 '
piR-134281; 2456; 3'UTR; -157; 92; 31	piR-112987; 2463; 3'UTR; -153; 94; 31
5 ' - UAGAGAUGGGGUCUCACAGUGUUGCCCAGGC-3 '	5'-GGGGUCUCACAGUGUUGCCCAGGCUGGUCUC-3'
3 ' -AUCUCUGCCCCAGAGCGUUACAACGAGUCCG-5 '	3'-CCUCAAAGUGUUACAACAGGUCCGACCAGAA-5'

**Figure 9.** Nucleotide sequences of piRNAs and the characteristics of their interaction with the mRNA of the *TP53* gene. Note: The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green.

## 3.10. piRNAs That Bind to the mRNA of Two or More Candidate Genes

It is essential to know which piRNAs bind to which genes and how this knowledge can be adequately used to regulate gene expression with piRNAs. We analyzed the interaction of piRNAs in nine candidate ESCC genes. Finding BSs of piRNAs in the mRNAs of alternative candidate genes was considered positive when the  $\Delta G$  value was -160 kJ/mol or higher and the  $\Delta G/\Delta Gm$  value was 90% or higher.

Table S11 shows the search results for piRNAs that bind to the mRNAs of the *AURKA*, *ERCC1*, *SASH1*, and *SIX4* genes. Only piR-55670 and piR-93385 bound to the mRNA of the *AURKA* and *BMP7* genes. Only piR-55670, piR-93385, piR-89432, piR-218175, piR-55670, and piR-93385 bound to the mRNA of the *AURKA* and *MTHFR* genes. The piR-67883, piR-200086, and piR-306330 bound to the mRNAs of the *AURKA* and *TP53* genes. An example of the interaction between the piRNAs in the *AURKA* and *SIX4* genes is shown in Figure 10.

The piRNAs interacting with the mRNAs of the *AURKA* and *SIX4* genes bound in the clusters of BSs with similar nucleotide sequences in the mRNA of both genes. In the mRNA of the *AURKA* gene, piRNAs bound in the 5'UTR from 523 nt to 534 nt, and in the mRNA of the *SIX4* gene, piRNAs bound in the 3'UTR from 4115 nt to 4126 nt with the same difference in BS beginning at about 11 nt. The value of the free binding energy of each piRNA in both mRNAs was close and high. The results demonstrate that one of the functions of the 5'UTR and 3'UTR is the binding of piRNAs regulating mRNA translation.

As a result, the mRNAs of all candidate genes binding piRNAs had BSs in two or more mRNAs. These data indicate the interconnection of candidate genes through the interaction of piRNAs, which should be considered both in developing diagnostic methods for ESCC and in creating therapeutic drugs using piRNAs. For using piRNAs as markers of ESCC, it is logical to use the genes that are targeted with the highest free energy of interaction between their mRNA and piRNAs, as well as the genes whose mRNAs bind mainly due to canonical nucleotide pairs. Figure 11 shows the interaction schemes of such associations of piRNAs and mRNA candidate genes. piR-289349, piR-298414, and piR-436360 can completely block protein synthesis at a concentration equal to or greater than that of the mRNA of *ERCC1*. Similarly, piR-50295 can completely inhibit the translation of the mRNA of the *SASH1* gene.

AURKA; piR-67883; 528; 5'UTR; -176; 97; 33

AURKA; piR-69165; 524; 5'UTR; -166; 92; 33

-CCUGCCUCAGCCUCCCGAGUAGCUGCGAUUACA-3

GGGCUCAUCGACUCU

-5

AU-3

5 1

5'-UUCUCCUGCCUCACCCUCCCGAGUAGCUGGGAU-3'

3'-AGAAGGACGGAGUUGGAGGGUUCAUCGACCCUG-AURKA;piR-197774;524;5'UTR;-172;94;34

5'-UUCUCCUGCCUCAGCCUCCCGAGUAGCUGCGAUU-3

-AAGAGGACAGAGUCGGAGGACUUAUCGACCCUGA-

CAGCCUCCCGAGUAGCUGGGAUUACAG

-AUUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAU-3

3'-CGGAGUCGGAGGGCUCAUCGACCCUGAUGUCCAU-5

3'-UAAGAGGGCGGAGUCGGAGGACUCAUCGACUCUA-5 AURKA;piR-249982;523;5'UTR;-172;95;33

AGCCUCCCGAGUAGCUGGGAUUACAGGC

-UGCCUCAGCCUCCCGAGUAGCUGGGAUUACAGGC-3

UCAGCCUCCCGAGUAGCUGGGAUUACAGGCAUC-3

3'-GUGGAGUCGGAAGACUCAUCAACCCUGAUGUCCG-5

GGAGGGUUCAUCGACCCUAAUGUCCGC

-CCUGCCUCAGCCUCCCGAGUAGCUGGGAUUACA-3

AURKA; piR-288286; 534; 5'UTR; -168; 93; 33

AURKA; piR-302301; 528; 5'UTR; -168; 91; 33

GGGCGGAGUCGGAGGGUUCAUCGACUCUG

AURKA; piR-306330; 530; 5'UTR; -178; 95; 34

AURKA; piR-320670; 524; 5'UTR; -172; 92; 33

5'-UUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAU-3'

3'-AAGAGGACGGAGUCGGGGGGUUCGUCGACCCUG-5

UGCCUCAGCCUCCCGAGUAGCUGGGAUU-3

-UCCUGCCUCAGCCUCCCGAGUAGCUGGGAUUACA-3

UCCUGCCUCAGCCUCCCGAGUAGCUGGG

3'-AGGACAGAGUCGGAGGGUUCAUCGACCCUAAUGU-5

3'-UGGGUCGGAGGGCUCGUCGACCCUGAUGUCCG-5'

3'-AGAGGGCGGAGUCGGAGGGCUCGUCGGCCCUAA-5'

5'-UUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAU-3'

......

3'-AAGAGGACGGAGUCGGGGGGCUCAUCGACCCUA-5'

..........

3'-AAGAGGACGAGGUCGGAGGACUCAUCGACCCUG-5'

-UUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAU-3

AURKA; piR-479539; 524; 5'UTR; -170; 93; 33

AURKA; piR-424861; 524; 5'UTR; -180; 98; 33

-UCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAUU-3

AURKA: piR-347859:525:5'UTR:-176:93:33

-CAAGAGGACGGAGUCGGAGGGCUCGUCGGCCCU-5

AURKA; piR-341211; 523; 5'UTR; -176; 93; 33

AURKA; piR-344301; 531; 5'UTR; -174; 92; 33 5'-GCCUCAGCCUCCCGAGUAGCUGGGAUUACAGGC-

AURKA; piR-321692; 526; 5'UTR; -168; 92; 32

3'-GAGGGCCGGAGUCGGAGGGCUCAUUGACCCUGG-5 AURKA;piR-325698;527;5'UTR;-178;98;34

5 ' -UGCCUCAGCCUCCCGAGUAGCUGGGAUUACAGGC-

ACAGAGUCGGAGGGUUCAUCGACCCUGAUGUCCG

5'-UCA

3'-GAGUCGGAGGGCUCGUCGACCCUGAUGUCCGUG-5'

AURKA; piR-274756; 530; 5'UTR; -168; 91; 34

3'-UAAGAGGACGGAACCGGAGGGUUCAUCGACCCU-5

AURKA; piR-267379; 533; 5'UTR; -174; 93; 33

3'-GGACGGAGUCGGGGGGUUCAUCGACCCUGAUGU-5'

AURKA; piR-200135; 531; 5'UTR; -183; 97; 34

AURKA; piR-219325; 523; 5'UTR; -176; 95; 34

5 ' -AUUCUCCUGCCUCAGCCUCCCGAGUAGCUGG

CCUCAGCCUCCCGAGUAGCUGGGAUUACA-3

AURKA; piR-200086; 528; 5'UTR; -174; 94; 33

SIX4; piR-67883; 4120; 3'UTR; -174; 95; 33 5 '-CCUGCCUCAGCCUCCCGAGUAGCUGGGACUACA-3 ' 3'-G0 SIX4;piR-69165;4116;3'UTR;-170;94;33 AGAAGGACGGAGUUGGAGGGUUCAUCGACCCUG-5 SIX4; piR-197774; 4116; 3'UTR; -176; 97; 34 TICCUCCCUCACCUCCCCACUACCUCCCACU-3 -AAGAGGACAGAGUCGGAGGACUUAUCGACCCUGA-5 SIX4; piR-200086; 4120; 3'UTR; -178; 97; 33 5'-CCUGCCUCAGCCUCCCGAGUAGCUGGGACUACA-3' 3'-GGACGGAGUCGGGGGGUUCAUCGACCCUGAUGU-5' piR-200135;4123;3'UTR;-187;99;34 SIX4: 5'-GCCUC GCCUCCCGAGUAGCUGGGACUACAG ....... 3 '-CGGAGUCGGAGGGCUCAUCGACCCUGAUGUCCAU-5 ' SIX4;piR-219325;4115;3'UTR;-174;94;34 5 '-AUUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAC-3 3 '-UAAGAGGCCGGAGUCGGAGGACUCAUCGACUCUA-5 SIX4; piR-249982; 4115; 3'UTR; -172; 95; 33 5 ' - AUUCUCCUGCCUCAGCCUCCCGAGUAGCUGG 3'-UAAGAGGACGGAACCGGAGGGUUCAUCGACCCU-5' SIX4; piR-267379; 4125; 3'UTR; -183; 98; 33 AGUAGCUGGG UACA 3 '-GAGUCGGAGGGCUCGUCGACCCUGAUGUCCGUG-5 ' SIX4: piR-274756;4122;3'UTR;-172;93;34 5 '-UGCCUCAGCCUCCCGAGUAGCUGGGACUACAGGC-3 ' 3 '-GUGGAGUCGGAAGACUCAUCAACCCUGAUGUCCG-5 ' SIX4;piR-288286;4126;3'UTR;-170;94;33 5 '-UCAGCCUCCCCAGUAGCUGGGACUACAGGCACC-3 ' 3'-AGUCGG AGGGUUCAUCGACCCUAAUGUCC -51 SIX4;piR-302301;4120;3'UTR;-172;93;33 -CCUGCCUCAGCCUCCCGAGUAGCUGGGACUACA-3 GUCGGAGGGUUCAUCGACUCUGGUG SIX4;piR-306330;4122;3'UTR;-183;98;34 5 '-UGCCUCAGCCUCCCGAGUAGCUGGGACUACAGGC-3 ACAGAGUCGGAGGGUUCAUCGACCCUGAUGUCCG SIX4; piR-320670; 4116; 3'UTR; -176; 94; 33 5'-UUCUCCUCCUCAGCCUCCCCAGUAGCUGGGAC-3' 3 ' -AAGAGGACGGAGUCGGGGGGUUCGUCGACCCUG-5 ' SIX4; piR-321692; 4118; 3'UTR; -172; 94; 32 CUCCCUCACCCUCCCCACUACCUCCCACU-3 -GAGGGCGGAGUCGGAGGGCUCAUUGACCCUGG SIX4; piR-325698; 4119; 3'UTR; -176; 97; 34 -UCCUGCCUCAGCCUCCCGAGUAGCUGGGACUACA-3 3'-AGGACAGAGUCGGAGGGUUCAUCGACCCUAAUGU-5' piR-341211;4115;3'UTR;-176;93;33 SIX4; 5 ' - AUUCUCCUGCCUCAGCCUCCCGAGUAGCUGGG -CAAGAGGACGGAGUCGGAGGGCUCGUCGGCCCU-5 SIX4; piR-344301; 4123; 3'UTR; -178; 94; 33 GCCUCAGCCUCCCGAGUAGCUGGGACUA **\_\_\_\_** 3'-UGGGUCGGAGGGCUCGUCGACCCUGAUGUCCG-5' piR-347859;4117;3'UTR;-174;92;33 STX4 --UCUCCUGCCUCAGCCUCCCGAGUAGCUGGG 3 '-AGAGGCCGGAGUCGGAGGGCUCGUCGCCCUAA-5 ' SIX4; piR-424861; 4116; 3'UTR; -178; 97; 33 5'-UUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAC-3 3'-AAGAGGACGGAGICGGGGGGCUCAUCGACCCUA-5' SIX4; piR-479539; 4116; 3'UTR; -174; 95; 33 -UUCUCCUGCCUCAGCCUCCCGAGUAGCUGGGAC-3 3'-AAGAGGACGAGGUCGGAGGACUCAUCGACCCUG-5'

Gene; piRNA; start BS,nt; region mRNA;  $\Delta G, kJ/mol; \Delta G/\Delta Gm, \%;$  piRNA length,nt

**Figure 10.** Schemes of the interaction of identical piRNAs with the mRNAs of the *AURKA* and *SIX4* genes. Note: The mRNA nucleotides are highlighted in red. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green.

Gene; piRNA; start BS, nt; region mR	NA; $\Delta G, kJ/mol; \Delta G/\Delta Gm, \%; piRNA length, nt$
ERCC1;piR-289349;2726;3'UTR;-161;100;32	ERCC1; piR-298414; 2727; 3'UTR; -155; 100; 31
5'-CACGCCCAGCCAAUUUUUGUAUUUUAGUAGA-3'	5'-ACGCCCAGCCAAUUUUUGUAUUUUAGUAGA-3'
111111111111111111111111111111111111111	111111111111111111111111111111111111111
3'-GUGCGGGUCGGUUAAAAACAUAAAAAUCAUCU-5'	3'-UGCGGGUCGGUUAAAAAACAUAAAAAUCAUCU-5'
ERCC1;piR-436360;2611;3'UTR;-144;100;26	SASH1;piR-50295;5746;3'UTR;-168;100;30
5 ' - GCCUAGGCUGGAGUGCAGUGGCAUGA-3 '	5 '-GAUUACAGGUGUGAGCCACCGCACCUGGCC-3 '
	111111111111111111111111111111111111111
3 '-CGGAUCCGACCUCACGUCACCGUACU-5 '	3 '-CUAAUGUCCACACUCGGUGGCGUGGACCGG-5 '
AURKA; piR-126340; 410; 5'UTR; -180; 94; 34	ERCC1;piR-88365;2659;3'UTR;-180;98;33
5 ' -GAAGGCCGGGCGCGGUGGCUCAUGCCCGUAAUCC-3 '	5 ' - CCUCCUGGGUUCAAGCGAUUCUCCUGCCUCAGC-3 '
111111111111111111111111111111111111111	111111111111111111111111111111111111111
3 ' - CUUCCGACCCGCACCACCGAGUGCGGACAUUAGG-5 '	3 '-GGAGGACCCAAGCUCGCUAAGAGGACGGAGUCG-5 '
GCOM1;piR-87408;2460;3'UTR;-180;96;34	SASH1;piR-94289;5698;3'UTR;-180;93;34
5 ' -UGGGGUCUUGCUCUGUCGCCCAGGCUGGAGUGCA-3 '	5 ' -ACUCCUGACCUCAGGUGGUCCACCCGCCUCAGCC-3 '
3 ' - GCCUCAGAACGAGACAGUGGGUCCGACCUCACGU-5 '	3 ' -UGAGGGCUGGAGUCCACUAGGUGGGUGGAGCCGG-5 '
SASH1;piR-220013;5497;3'UTR;-180;99;34	SASH1;piR-245383;5717;3'UTR;-180;94;34
5 ' -UUUCGCUCUUGUCACCCAGGCUGGAGUGCAAUGG-3 '	5 ' - CCACCCGCCUCAGCCUCCCAAAGUGCUGGGAUUA-3 '
3 ' -AAAGUGAGAACAGUGGGUCCGACCUCACGUUACC-5 '	3 '-GGCGGGUGGAGUCGGAGGGUUUCACGGCCCUAAU-5 '
SASH1;piR-427055;5654;3'UTR;-180;99;34	SIX4;piR-252003;4032;3'UTR;-180;94;34
5 ' -AGAGACAGGGUUUCACCAUGUUGGCCAGGCUGGU-3 '	5 ' -AUGGAGUCUCGCUCUGUUGCCCAGGCUGGAGUGC-3 '
	1
3 ' -UUUCUGUCCCAAAGUGGUACAACCGGUCCGACCA-5 '	3 ' -UGCCUCAGAGCGAGACAGCGAGUCCGACCUCACG-5 '
SIX4;piR-252240;4035;3'UTR;-180;94;34	SIX4; piR-121234; 4119; 3'UTR; -183; 97; 34
5 ' -GAGUCUCGCUCUGUUGCCCAGGCUGGAGUGCAGU-3 '	5 ' - UCCUGCCUCAGCCUCCCGAGUAGCUGGGACUACA-3 '
3 ' -CUCAGAGUGAGACAGCGGGUCCGACCUCACGCCA-5 '	3 ' -AGGACGGAGUCGGAGGGCUCAUCAACCCUGGUGU-5 '
SIX4;piR-267379;4125;3'UTR;-183;98;33	SIX4;piR-248674;4130;3'UTR;-180;97;33
5 ' -CUCAGCCUCCCGAGUAGCUGGGACUACAGGCAC-3 '	5 ' - CCUCCCGAGUAGCUGGGACUACAGGCACCUGCC-3 '
3 ' -GAGUCGGAGGGCUCGUCGACCCUGAUGUCCGUG-5 '	3 '-GGAGGGCUCAUCAACCCUGAUGUCCGCGGACGG-5 '
SIX4;piR-204309;4133;3'UTR;-180;92;34	SIX4;piR-89068;4243;3'UTR;-180;93;34
5 ' - CCCGAGUAGCUGGGACUACAGGCACCUGCCACCA-3 '	5 ' - CUCGUGAUCUGCCCGCCUCGGCCUCCCAAAGUGC-3 '
111111111111111111111111111111111111111	
3 '-GGGCUCAUCGACCCUGAUACCCGCGGGCGGUGGU-5 '	3 '-GAGUACUGGGUGGGGGGGGGGGGGGGGGGUUUCACG-5 '

**Figure 11.** Schemes of the interaction between piRNAs and the mRNAs of candidate genes with  $\Delta G/\Delta Gm$  equal to 92% to 100% and a  $\Delta G$  value equal to -180 kJ/mol and higher. Canonical pairs of piRNA and mRNA nucleotides are shown in violet, whereas noncanonical pairs are marked in green.

## 4. Discussion

The free energy value of the interaction of piRNAs with mRNA determines the probability of their interaction with mRNA. Depending on the free energy of interaction of piRNAs with the mRNA of a gene, groups of piRNAs differ little in the BS interval and their position in the mRNA region, which is due to the selection during the evolution of the site for binding piRNAs with different lengths and nucleotide compositions. Therefore, changing piRNAs' length and nucleotide composition leads to the selection of optimal piR-NAs for the desired effect in regulating the expression of one or more genes. The piRNAs that we identified regulating two or more candidate genes formed both single bonds in regulating the expression of these genes and associations of several genes regulated by a group of piRNAs. When studying the interaction of piRNAs with mRNA genes, criteria were established for the following characteristics. 1. The piRNAs' interaction with the mRNAs of candidate ESCC genes occurs along the entire length of the piRNAs. 2. The free interaction energy ( $\Delta G$ ) must be at least 90% of the maximum  $\Delta Gm$  value. 3. The division of piRNAs into three groups according to the value of  $\Delta G$  and  $\Delta G/\Delta G$ m of the interaction of piRNAs with mRNA allowed us to determine that these characteristics do not fundamentally change the basic properties of the interaction of piRNAs with mRNA within the chosen limits.

The above results indicate non-random binding of piRNAs in a narrow interval of mRNA length only in the 3'UTR. This allows the human genome to control gene expression by many piRNAs, keeping only a small portion of the gene conserved. Another property of piRNAs is the presence of BSs of some piRNAs in the mRNA of more than one gene.

This fact indicates the general dependence of a group of candidate genes on the regulatory influence of the same piRNAs. This important property must be known for each piRNA when using piRNAs as disease markers and therapeutic agents. Ignorance of this property can lead to side effects when using piRNAs.

The organization of BSs into clusters reduces the mRNA region responsible for regulating expression by piRNA molecules, which must be kept conservative to preserve the regulation of genome expression balanced during evolution. The binding of piRNAs in clusters leads to competition between them for the ability to exert a regulatory effect on the expression of candidate genes. At the same time, the competition between piRNAs ensures the stability of their regulation of the target gene expression because an increase or decrease in the concentration of any piRNA will have a compensatory effect on other piRNAs. That is, a significant effect of one piRNA requires considerable increases in its concentration to exceed the total impact of other piRNAs.

Using piRNAs for disease diagnosis requires compliance with several criteria that will ensure the validity of the proposed piRNAs as markers. These are as follows: 1. piRNAs must bind to mRNAs with sufficiently high free energy to affect the expression of one or more candidate genes; 2. piRNAs must be expressed in a manner that is ordinarily comparable to the expression of candidate target genes; 3. Measurements of piRNAs and candidate gene concentrations should be made simultaneously, directly indicating which piRNAs correlate with the target gene's expression level. The studies conducted show specific associations of piRNAs and target genes that need to be studied. The findings contribute to determining the expression level of piRNAs and their target genes under normal and disease conditions for the diagnosis and therapy of diseases.

**Supplementary Materials:** The following supporting information can be downloaded at https: //www.mdpi.com/article/10.3390/cimb45070387/s1, Table S1: List of candidate genes for ESCC; Table S2: Characteristics of the piRNA interaction with the mRNA of the *AURKA* gene; Table S3: Characteristics of the piRNA interaction with the mRNA of the *BMP7* gene; Table S4: Characteristics of the piRNA interaction with the mRNA of the *ERCC1* gene; Table S5: Characteristics of the piRNA interaction with the mRNA of the *ERCC1* gene; Table S5: Characteristics of the piRNA interaction with the mRNA of the *GCOM1* gene; Table S6: Characteristics of the piRNA interaction with the mRNA of the *MTHFR* gene; Table S7: Characteristics of the piRNA interaction with the mRNA of the *SASH1* gene; Table S8: Characteristics of the piRNA interaction with the mRNA of the *SIX4* gene; Table S9: Characteristics of the piRNA interaction with the mRNA of the *SIX4* gene; Table S9: Characteristics of the piRNA interaction with the mRNA of the *SIX4* gene; Table S10: Characteristics of the piRNA interaction with the mRNA of the *SULT1A1* gene; Table S10: Characteristics of the piRNA interaction with the mRNA of the *SULT1A1* gene; Table S10: Characteristics of the piRNA interaction with the mRNA of the *SULT1A1* gene; Table S10: Characteristics of the piRNA interaction with the mRNA of the *SULT1A1* gene; Table S10: Characteristics of the piRNA interaction with the mRNA of the *SULT1A1* piRNAs that bind to the mRNA of the *AURKA*, *ERCC1*, *SASH1* and *SIX4* genes.

**Author Contributions:** Conceptualization, A.I., A.A. and T.N.; methodology, A.I., A.P. and P.Z.; software, A.P. and A.I.; validation, A.A., A.R., T.N. and M.T.; investigation, T.N., A.A., A.R. and M.T.; resources, A.A., A.R. and T.N.; data curation, A.P. and M.T.; writing—original draft preparation, A.I. and P.Z.; writing—review and editing, A.A., A.R. and T.N; visualization, A.A. and A.R.; supervision, A.I. and P.Z.; project administration, P.Z. and A.I.; funding acquisition, P.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Polish Ministry of Science and Higher Education, under the project: DIR/WK/2018/06.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

## References

- Kim, V.N. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. *Genes Dev.* 2006, 20, 1993–1997.
   [CrossRef]
- Aravin, A.A.; Sachidanandam, R.; Girard, A.; Fejes-Toth, K.; Hannon, G.J. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* 2007, *316*, 744–747. [CrossRef]

- 3. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. Potent and specific genetic interference by doublestranded RNA in Caenorhabditis elegans. *Nature* **1998**, *391*, 806–811. [CrossRef]
- Timmons, L.; Tabara, H.; Mello, C.C.; Fire, A.Z. Inducible Systemic RNA Silencing in *Caenorhabditis elegans*. Mol. Biol. Cell 2003, 14, 2972–2983. [CrossRef]
- 5. Singh, G.; Mallick, B. Predicting sequence and structural features of effective piRNA target binding sites. *J. Mol. Recognit.* 2022, 35, e2949. [CrossRef]
- 6. Xiong, Q.; Zhang, Y.; Li, J.; Zhu, Q. Small Non-Coding RNAs in Human Cancer. Genes 2022, 13, 2072. [CrossRef]
- Wang, X.; Gou, L.T.; Liu, M.F. Noncanonical functions of PIWIL1/piRNAs in animal male germ cells and human diseases. *Biol. Reprod.* 2022, 107, 101–108. [CrossRef]
- Zhang, J.; Chen, S.; Liu, K. Structural insights into piRNA biogenesis. *Biochim. Biophys. Acta Gene Regul. Mech.* 2022, 1865, 194799. [CrossRef]
- Priyadarshini, M.; AlHarbi, S.; Frøkjær-Jensen, C. Acute and inherited piRNA-mediated silencing in a rde-3 ribonucleotidyltransferase mutant. *microPubl. Biol.* 2022, 2022, 1–6. [CrossRef]
- Feng, Y.; Lin, Y.; Jiang, Z.; Wu, L.; Zhang, Y.; Wu, H.; Yuan, X. Insulin-like growth factor-2 mRNA-binding protein 3 promotes cell migration, invasion, and epithelial-mesenchymal transition of esophageal squamous cell carcinoma cells by targeting zinc finger E-box-binding homeobox 1 mRNA. *Mol. Carcinog.* 2023, 62, 503–516. [CrossRef]
- 11. Sinha, A.; Bhattacharjee, R.; Bhattacharya, B.; Nandi, A.; Shekhar, R.; Jana, A.; Saha, K.; Kumar, L.; Patro, S.; Panda, P.K.; et al. The paradigm of miRNA and siRNA influence in Oral-biome. *Biomed. Pharmacother.* **2023**, *159*, 114269. [CrossRef]
- 12. Lee, Y.R.; Tsai, H.P.; Yeh, C.S.; Fang, C.Y.; Chan, M.W.Y.; Wu, T.Y.; Shen, C.H. RNA Interference Approach Is a Good Strategy against SARS-CoV-2. *Viruses* 2022, *15*, 100. [CrossRef]
- 13. Chattopadhyay, T.; Gupta, P.; Nayak, R.; Mallick, B. Genome-wide profiling of dysregulated piRNAs and their target genes implicated in oncogenicity of tongue squamous cell carcinoma. *Gene* **2023**, *849*, 146919. [CrossRef]
- 14. Belkozhayev, A.; Niyazova, R.; Wilson, C.; Jainakbayev, N.; Pyrkova, A.; Ashirbekov, Y.; Akimniyazova, A.; Sharipov, K.; Ivashchenko, A. Bioinformatics Analysis of the Interaction of miRNAs and piRNAs with Human mRNA Genes Having di- and Trinucleotide Repeats. *Genes* **2022**, *13*, 800. [CrossRef]
- 15. AmeliMojarad, M.; Amelimojarad, M. piRNAs and PIWI proteins as potential biomarkers in Breast cancer. *Mol. Biol. Rep.* **2022**, 49, 9855–9862. [CrossRef]
- 16. Ray, S.K.; Mukherjee, S. Piwi-interacting RNAs (piRNAs) and colorectal carcinoma: Emerging non-invasive diagnostic biomarkers with potential therapeutic target based clinical implications. *Curr. Mol. Med.* **2022**, *23*, 300–311. [CrossRef]
- 17. Riquelme, I.; Pérez-Moreno, P.; Letelier, P.; Brebi, P.; Roa, J.C. The Emerging Role of PIWI-Interacting RNAs (piRNAs) in Gastrointestinal Cancers: An Updated Perspective. *Cancers* **2021**, *14*, 202. [CrossRef]
- Heng, B.; Xie, X.; Zeng, W.; Li, H.; Shi, L.; Ye, W.; Wu, F. PIWI-Interacting RNA Pathway Genes: Potential Biomarkers for Clear Cell Renal Cell Carcinoma. *Dis. Markers* 2022, 2022, 3480377. [CrossRef]
- Urabe, Y.; Kagemoto, K.; Hayes, C.N.; Nakamura, K.; Masuda, K.; Ono, A.; Tanaka, S.; Arihiro, K.; Chayama, K. Genomic characterization of early-stage esophageal squamous cell carcinoma in a Japanese population. *Oncotarget* 2019, 10, 4139–4148. [CrossRef]
- Chen, X.X.; Zhong, Q.; Liu, Y.; Yan, S.M.; Chen, Z.H.; Jin, S.Z.; Xia, T.L.; Li, R.Y.; Zhou, A.J.; Su, Z.; et al. Genomic comparison of esophageal squamous cell carcinoma and its precursor lesions by multi-region whole-exome sequencing. *Nat. Commun.* 2017, *8*, 524. [CrossRef]
- 21. Xu, G.; Tang, S.; Yang, J.; Chen, K.; Kang, J.; Zhao, G.; Feng, F.; Yang, X.; Zhao, L.; Lu, Q.; et al. BMP7 expression in esophageal squamous cell carcinoma and its potential role in modulating metastasis. *Dig. Dis. Sci.* **2013**, *58*, 1871–1879. [CrossRef]
- Du, R.; Huang, C.; Liu, K.; Li, X.; Dong, Z. Targeting AURKA in Cancer: Molecular mechanisms and opportunities for Cancer therapy. *Mol. Cancer* 2021, 20, 15. [CrossRef]
- Kamenova, S.; Sharapkhanova, A.; Akimniyazova, A.; Kuzhybayeva, K.; Kondybayeva, A.; Rakhmetullina, A.; Pyrkova, A.; Ivashchenko, A. piRNA and miRNA Can Suppress the Expression of Multiple Sclerosis Candidate Genes. *Nanomaterials* 2022, 13, 22. [CrossRef]
- 24. Akimniyazova, A.N.; Niyazova, T.K.; Yurikova, O.Y.; Pyrkova, A.Y.; Zhanuzakov, M.A.; Ivashchenko, A.T. piRNAs may regulate expression of candidate genes of esophageal adenocarcinoma. *Front. Genet.* **2022**, *13*, 1069637. [CrossRef]
- 25. Akimniyazova, A.; Yurikova, O.; Pyrkova, A.; Rakhmetullina, A.; Niyazova, T.; Ryskulova, A.G.; Ivashchenko, A. In Silico Study of piRNA Interactions with the SARS-CoV-2 Genome. *Int. J. Mol. Sci.* 2022, 23, 9919. [CrossRef]
- 26. Ameli Mojarad, M.; Ameli Mojarad, M.; Shojaee, B.; Nazemalhosseini-Mojarad, E. piRNA: A promising biomarker in early detection of gastrointestinal cancer. *Pathol. Res. Pract.* 2022, 230, 153757. [CrossRef]
- 27. Jian, Z.; Han, Y.; Li, H. Potential roles of PIWI-interacting RNAs in lung cancer. Front. Oncol. 2022, 12, 944403. [CrossRef]
- 28. Dabi, Y.; Bendifallah, S.; Suisse, S.; Haury, J.; Touboul, C.; Puchar, A.; Favier, A.; Daraï, E. Overview of non-coding RNAs in breast cancers. *Transl. Oncol.* 2022, 25, 101512. [CrossRef]
- 29. AmeliMojarad, M.; AmeliMojarad, M.; Wang, J. The function of novel small non-coding RNAs (piRNAs, tRFs) and PIWI protein in colorectal cancer. *Cancer Treat. Res. Commun.* 2022, *31*, 100542. [CrossRef]
- Goh, T.X.; Tan, S.L.; Roebuck, M.M.; Teo, S.H.; Kamarul, T. A Systematic Review of Extracellular Vesicle-Derived Piwi-Interacting RNA in Human Body Fluid and Its Role in Disease Progression. *Tissue Eng. Part C Methods* 2022, 28, 511–528. [CrossRef]

- 31. Guan, S.; Zhang, Z.; Wu, J. Non-coding RNA delivery for bone tissue engineering: Progress, challenges, and potential solutions. *iScience* **2022**, *25*, 104807. [CrossRef]
- Sun, J.; Yang, X.; Wang, T.; Xing, Y.; Chen, H.; Zhu, S.; Zeng, J.; Zhou, Q.; Chen, F.; Zhang, X.; et al. Evaluating the Effects of Storage Conditions on Multiple Cell-Free RNAs in Plasma by High-Throughput Sequencing. *Biopreserv. Biobank.* 2022, 21, 242–254. [CrossRef]
- Cho, O.; Kim, D.W.; Cheong, J.Y. Screening Plasma Exosomal RNAs as Diagnostic Markers for Cervical Cancer: An Analysis of Patients Who Underwent Primary Chemoradiotherapy. *Biomolecules* 2021, 11, 1691. [CrossRef]
- Wang, J.; Shi, Y.; Zhou, H.; Zhang, P.; Song, T.; Ying, Z.; Yu, H.; Li, Y.; Zhao, Y.; Zeng, X.; et al. piRBase: Integrating piRNA annotation in all aspects. *Nucleic Acids Res.* 2021, 50, 265–272. [CrossRef]
- 35. Ivashchenko, A.; Berillo, O.; Pyrkova, A.; Niyazova, R.; Atambayeva, S. MiR-3960 binding sites with mRNA of human genes. *Bioinformation* **2014**, *10*, 423–427. [CrossRef]
- 36. Friedman, R.A.; Honig, B.A. Free Energy Analysis of Nucleic Acid Base Stacking in Aqueous Solution. *Biophys. J.* **1995**, *69*, 1528–1535. [CrossRef]
- 37. Garg, A.; Heinemann, U.A. Novel Form of RNA Double Helix Based on G·U and C·A+ Wobble Base Pairing. *RNA* **2018**, *24*, 209–218. [CrossRef]
- Leontis, N.B.; Stombaugh, J.; Westhof, E. The Non-watson-crick Base Pairs and Their Associated Isostericity Matrices. Nucleic Acids Res. 2002, 30, 3497–3531. [CrossRef]
- Davis, E.; Caiment, F.; Tordoir, X.; Cavaillé, J.; Ferguson-Smith, A.; Cockett, N.; Georges, M.; Charlier, C. RNAi-Mediated Allelic Trans-interaction at the Imprinted Rtl1/Peg11 Locus. *Curr. Biol.* 2005, 15, 743–749. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.