









Article

# Molecular and Functional Characterization of the Somatic PIWIL1/piRNA Pathway in Colorectal Cancer Cells

Assunta Sellitto <sup>1,2,†</sup>, Konstantinos Geles <sup>1,2,†</sup>, Ylenia D'Agostino <sup>1</sup>, Marisa Conte <sup>1</sup>, Elena Alexandrova <sup>1,2</sup>, Domenico Rocco <sup>1</sup>, Giovanni Nassa <sup>1,3</sup>, Giorgio Giurato <sup>1</sup>, Roberta Tarallo <sup>1,3</sup>, Alessandro Weisz <sup>1,3,\*</sup> and Francesca Rizzo <sup>1,3,\*</sup>

- <sup>1</sup> Laboratory of Molecular Medicine and Genomics, Department of Medicine, Surgery and Dentistry “Scuola Medica Salernitana”, University of Salerno, 84081 Baronissi (SA), Italy; assellitto@unisa.it (A.S.); yeles.konstantinos@gmail.com (K.G.); ydagostino@unisa.it (Y.D.); marisaconte30@gmail.com (M.C.); ealexandrova@unisa.it (E.A.); drocco@unisa.it (D.R.); gnassa@unisa.it (G.N.); ggiurato@unisa.it (G.G.); rtarallo@unisa.it (R.T.)
- <sup>2</sup> Genomix4Life srl, Department of Medicine, Surgery and Dentistry “Scuola Medica Salernitana”, University of Salerno, 84081 Baronissi (SA), Italy
- <sup>3</sup> Medical Genomics Program, Department of Onco-Haematology SS. Giovanni di Dio e Ruggi d’Aragona’ University Hospital, University of Salerno, 84131 Salerno, Italy
- \* Correspondence: aweisz@unisa.it (A.W.); frizzo@unisa.it (F.R.); Tel.: +39-089-965043 (A.W.); +39-089-965221 (F.R.)
- † These authors equally contributed to this work.

Received: 18 September 2019; Accepted: 3 November 2019; Published: 5 November 2019



**Abstract:** PIWI-like (PIWIL) proteins and small non-coding piRNAs, involved in genome regulation in germline cells, are found aberrantly expressed in human tumors. Gene expression data from The Cancer Genome Atlas (TCGA), the Genotype-Tissue Expression (GTEx) project, and the European Genome-Phenome Archive (EGA) indicate that the *PIWIL1* gene is ectopically activated in a significant fraction of colorectal cancers (CRCs), where this is accompanied by promoter demethylation, together with germline factors required for piRNA production. Starting from this observation, the PIWIL/piRNA pathway was studied in detail in COLO 205 CRC cells, which express significant levels of this protein, to investigate role and significance of ectopic *PIWIL1* expression in human tumors. RNA sequencing and cell and computational biology led to the demonstration that PIWIL1 localizes in a nuage-like structure located in the perinuclear region of the cell and that a significant fraction of the piRNAs expressed in these cells are methylated, and, therefore, present in an active form. This was further supported by RNA immunoprecipitation, which revealed how several piRNAs can be found loaded into PIWIL1 to form complexes also comprising their target mRNAs. The mature transcripts associated with the PIWIL–piRNA complex encode key regulatory proteins involved in the molecular mechanisms sustaining colorectal carcinogenesis, suggesting that the PIWI/piRNA pathway may actively contribute to the establishment and/or maintenance of clinico-pathological features of CRCs.

**Keywords:** colorectal cancer; small non-coding RNAs; RNA-binding proteins; PIWIL1; piRNAs; nuage compartment; RIP-Seq

## 1. Introduction

Cancer and germ cells share essential similarities, such a high proliferation rate and self-renewal ability, and cancer cells sometimes re-activate cancer testis antigen (CTA) genes, whose expression is usually restricted to the germline and silenced in adult somatic tissues. The expression of germline genes in cancer may reflect the aberrant activation of a silenced developmental program that leads to the escape from cell death, immune evasion, and invasiveness, thus contributing to the molecular mechanisms of carcinogenesis [1]. Among CTAs, *PIWI-like (PIWIL)* genes, belonging to the Argonaute family, have recently attracted attention as they are frequently deregulated in cancer and have been associated with advanced tumor stage and poor prognosis, making them useful biomarkers for cancer diagnosis and possible druggable targets [2,3].

The PIWI-like family includes evolutionarily conserved proteins (PIWIL1-4) with key functions in the regulation of gene expression [4], whose activity is mediated by the association with a specific class of small non-coding RNAs (sncRNAs) of 20–35 nt in length, defined as PIWI-interacting RNAs (piRNAs). In non-pathological conditions, both PIWILs and piRNAs are abundantly expressed in the germline, where they are involved in transposable element silencing [5]. However, several studies revealed that not only transposon transcripts but also mRNAs are targeted by piRNAs, by binding to their 3'UTR and the induction of mRNA degradation [6], mainly during embryogenesis and germ-cell development [7]. Recently, a crucial role for the PIWIL–piRNA complex has also emerged in somatic cells, where growing evidence indicates that both PIWIL and piRNA expression can be reactivated under pathological stimuli [8–11]. In this respect, *in vitro/in vivo* studies demonstrated that PIWIL proteins are not only a marker of malignancy but are also involved in cell cycle regulation, tumorigenesis, drug resistance, and acquisition of self-renewal abilities [12,13].

Colorectal (CRC) is one of the most common human cancers, and still the third leading cause of cancer-related mortality. Identifying new diagnostic and therapeutic targets for CRC therefore remains a priority in cancer research. Interestingly, in the last few years, piRNAs emerged as promising diagnostic and prognostic biomarkers for CRC [14–16], where ectopic presence of PIWIL1 has recently been correlated with tumor differentiation status, infiltration, lymph node invasion and metastasis [17–19]; patients with PIWIL1 overexpression also showed lower overall survival and disease-free survival [20]. Moreover, the expression of PIWIL1 in CRC has been positively correlated with the mRNA level of OCT4, a cancer stem cell marker, indicating that this gene may contribute to the tumor stemness, which in turn is strongly related to its metastatic potential [21]. However, the molecular mechanisms and biological roles of PIWILs and piRNAs have not been explored in detail, so that the cellular pathways and molecular mechanisms in which somatic piRNAs are involved are still poorly known.

We thus investigated here the co-expression and possible functional relationship between PIWILs and piRNAs in CRC. To this end, the PIWIL1/piRNA pathway was characterized in CRC tumor tissues and cell lines, combining RNA sequencing, methylation data, and protein biochemistry. Interestingly, we observed in human CRC cells that PIWIL1 localizes in nuage-like perinuclear granules in the cytoplasmic perinuclear region of the cell, a result that, combined with the identification of mature piRNAs carrying a methyl group at their 3'-end, indicates the presence of an active piRNA pathway in CRC cells. Finally, we demonstrated, in COLO 205 cells, the physical association between PIWIL1 and several piRNAs and their predicted mRNA targets, a result experimentally validated by RNA immunoprecipitation (RIP)-Seq. The mature transcripts associated with the PIWIL–piRNA complex are involved in the control of cell proliferation, differentiation, and survival in cancer cells. These data demonstrate that re-activation of this germline pathway in CRC cells can contribute to the regulation of cellular pathways controlling key functions of the cancerous cell.

## 2. Materials and Methods

### 2.1. Gene Expression Databases and DNA Methylation Data Analysis

The expression level of 4 *PIWIL* genes was evaluated across 31 tumor types listed in the PanCancer Atlas [22] by mining RNA-Seq data from two databases, The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx). Expression bar plots for *PIWIL* genes were obtained using the Gepia2 web server [23]. Transcriptomic data for 53 CRC cell lines were downloaded from the European Genome-Phenome Archive (EGA, <https://ega-archive.org/>) dataset (E-MTAB-2706). To evaluate the relationship between *PIWIL1* expression and DNA methylation of tumor tissues, Infinium HumanMethylation450 data and RNA-Seq data for 275 colon adenocarcinomas and 19 normal tissues (TCGA data) were downloaded from the TCGA database. For the methylation analysis of 8 CRC cell lines, Infinium HumanMethylation450 collected by Barault et al. (GEO accession GSE86078) were used [24].

### 2.2. Cell Culture

Human colon cancer cell lines Caco-2, SW403, SW1417, COLO 205, HT-29, HCT 116, and RKO were supplied by the American Type Culture Collection (ATCC, Rockville, MD, USA); the HT115 cell line was obtained from Sigma-Aldrich, Milan, Italy. All cell lines were cultured following the manufacturer's instructions; culture media were supplemented with fetal bovine serum (HyClone, Cramlington, UK), 100 U/mL penicillin, 100 mg/mL streptomycin, and 250 ng/mL Amphotericin-B. The identity of all cell lines was confirmed by short tandem repeat (STR) profiling; cells were routinely screened for mycoplasma contamination with MycoAlert mycoplasma detection kit (Lonza, Milan, Italy).

### 2.3. Transient Transfections

For the transient overexpression of *PIWIL1*, HCT 116 cells were transfected with the human full-length cDNA clone pCMV6-*PIWIL1* (RC205269) or with control pCMV6-Entry mammalian vector (PS100001) (Origene, Herford, Germany). At 24 h prior to transfection, HCT 116 cells at the exponential growth phase were seeded in 100 mm culture dishes; the next day, plates at 60% confluency were washed and re-fed with culture medium shortly before transfection. A total of 15 µg of DNA was mixed with linear 25 kDa polyethyleneimine (PEI) (Polysciences, Eppenheim, Germany) and incubated for 20 min at room temperature, then the DNA/PEI mixture was added to plates. After 24 h from transfection, HCT 116-pCMV6-*PIWIL1* cells were analyzed by immunofluorescence using rabbit anti-*PIWIL1* (ab12337, Abcam, Cambridge, UK); fluorescence images were collected with a LEICA DM6000 B Confocal Microscope and used to evaluate and quantify transfection efficiency, which was found to be ~20%.

### 2.4. Real-Time qRT-PCR

To generate cDNA with the AffinityScript cDNA Synthesis Kit (Agilent Technologies, Rome, Italy), 1 µg of total RNA was used. cDNAs were diluted to a final concentration of 20 ng per reaction. Real-time qRT-PCR was performed in triplicate using Brilliant II SYBR Master Mixes (Agilent Technologies) on an Mx3005P Instrument (Agilent Technologies); the expression level of *PIWIL* genes was normalized against  $\beta$ -actin mRNA. Specific primer sets are reported in Table 1.

**Table 1.** Real-time qRT-PCR primers.

Gene	Forward (5'-3')	Reverse (5'-3')
<i>PIWIL1</i>	ACGAAGTGCCACAGTTTTTGG	AGTCTTCCTCCAGACTGAGC
<i>PIWIL2</i>	GCCTGGGTGAACTAAAGGA	CCATGATGATGCAAACAACC
<i>PIWIL3</i>	TCAGATGGCAGCAAAATCAC	ACGTTGTGTACCCGTTAGGC
<i>PIWIL4</i>	ATGGCACCGAGATCACCTAT	GCTGAGCCTCACTGTTGTCA

### 2.5. Total Protein Extraction

Cells at exponential growth phase were harvested by scraping in cold phosphate buffered saline (PBS) and collected by centrifugation at  $1000\times g$ . Then, for total protein extraction, pelleted cells were resuspended in 3 volumes with respect to the cell pellet of hypotonic buffer (50 mM Tris-HCl, 0.15% NP40, 10% Glycerol, 180 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 1 mM NaMoO<sub>4</sub>, 0.5 mM NaF) supplemented with 1× Protease Inhibitor Cocktail (Sigma-Aldrich, Milan, Italy), 1 mM DTT and 1 mM PMSF. Upon incubation on ice for 15 min, cell lysate was clarified by centrifugation at  $20,000\times g$  for 15 min at 4 °C, then the salt concentration was adjusted with hypotonic buffer without salts addition (2V with respect to the hypotonic buffer previously added).

### 2.6. Cytosol/Nucleus Protein Fractionation

For cytosol/nucleus protein fractionation, pelleted cells were resuspended in 3 volumes with respect to the cell pellet of hypotonic buffer (20 mM HEPES pH 7.4, 5 mM NaF, 10 μM NaMoO<sub>4</sub>, 0.1 mM EDTA) supplemented with 1× Protease Inhibitor Cocktail (Sigma-Aldrich, Milan, Italy), 1 mM DTT and 1 mM PMSF; upon incubation on ice for 15 min, 0.5% Triton X-100 was added, and lysate was centrifuged at  $15,000\times g$  for 30 s at 4 °C in order to collect the cytosolic fraction, which was further clarified by centrifugation at  $15,000\times g$  for 15 min. The nuclear pellets were first stratified in sucrose gradient to remove cytosolic contaminants and then resuspended in 1 volume of nuclear lysis buffer (20 mM HEPES, pH 7.4, 25% glycerol, 420 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA) supplemented with 1× Protease Inhibitor Cocktail (Sigma-Aldrich, Milan, Italy), 1 mM DTT, and 1 mM PMSF; nuclei were incubated for 30 min at 4 °C with gentle shaking and then centrifuged at  $15,000\times g$  for 30 min at 4 °C. Salt concentration was adjusted with 1 volume of nuclear lysis buffer without salt addition, supplemented with 1× Protease Inhibitor Cocktail (Sigma-Aldrich, Milan, Italy), 1 mM DTT, and 1 mM PMSF.

### 2.7. Cytosol/Perinucleus/Nucleus Protein Fractionation

Protein fractionation was performed as previously described by Shaiken et al. [25]. Briefly, cells at exponential growth phase were harvested by scraping in cold PBS, collected by centrifugation at  $1000\times g$ , and resuspended in 3 volumes with respect to the cell pellet of isotonic buffer A (40 mM HEPES pH 7.4, 120 mM KCl, 2 mM EGTA, 0.4% Glycerol, 10 mM β-glycerophosphate, and 0.4% NP-40) supplemented with 1× Protease Inhibitor Cocktail (Sigma-Aldrich, Milan, Italy), 1 mM DTT, and 1 mM PMSF; cells were incubated under gentle rotation for 30 min at 4 °C, then nuclei were pelleted by centrifugation at  $1000\times g$  for 5 min; the supernatant containing the cytosolic fraction was clarified at  $10,000\times g$  for 10 min; nuclei were gently washed two times (first with 0.1% NP-40 and then with no-detergent-containing Buffer A), then centrifuged at  $1000\times g$  for 5 min. To isolate the perinuclear fraction, the nuclear pellet was resuspended in 1 volume of Buffer B (10 mM Tris-HCl pH 7.4, 1.5 mM KCl, 0.5% Triton X-100; 0.5% deoxycholate, 2.5 mM MgCl<sub>2</sub>) supplemented with fresh 0.2 M LiCl, 1× Protease Inhibitor Cocktail (Sigma-Aldrich, Milan, Italy), 1 mM DTT, and 1 mM PMSF; after 1 h of incubation at 4 °C under gentle rotation, the extract was centrifuged at  $10,000\times g$  for 10 min to collect the perinuclear fraction; nuclei were resuspended in 0.34 M sucrose and centrifuged at  $2000\times g$  for 10 min in order to remove contaminants; nuclei were finally lysed in 8 M urea and sonicated 3 times (30 sec ON, 30 sec OFF); the nuclear extract was further clarified by centrifugation at  $10,000\times g$  for 10 min.

### 2.8. SDS-Page and Western Blotting (WB)

Protein expression was analyzed by SDS page, in each case loading 30 µg of protein extract, and western blotting (WB) using standard protocols and the following primary antibodies: rabbit anti-PIWIL1 (ab105393, Abcam, Cambridge, UK), rabbit anti-Lamin B2 (ab151735, Abcam, Cambridge, UK), rabbit anti-DDX6 (ab40684, Abcam, Cambridge, UK), mouse anti-β-Actin (ab6276, Abcam, Cambridge, UK), mouse anti-β-Tubulin (T6074, Sigma-Aldrich, Milan, Italy). The specificity of anti-PIWIL1 antibody was tested against the transfected protein on HCT 116 cell lysate.

### 2.9. Immunofluorescence (IF) Assay and Confocal Microscopy

For IF, cells were fixed with 4% paraformaldehyde for 20 min and washed with PBS-Tween; permeabilization was performed in 0.1% Triton-X-100 in PBS for 10 min, then cells were incubated in blocking buffer (0.5% BSA in PBS-Tween) for 1 h. Slides were incubated at 4 °C overnight with rabbit anti-PIWIL1 (ab12337, Abcam, Cambridge, UK). The next day, slides were washed with 0.5% BSA-PBS-Tween and incubated for 45 min at room temperature with Alexa Fluor 488 goat anti-rabbit IgG (Thermo-Fisher, Carlsbad, CA) secondary antibody in PBS. For nuclei staining, 4',6-diamidino-2-phenylindole (DAPI, Sigma Aldrich, St. Louis, MS, USA) was used. Images were collected with a Leica DM6000 B confocal microscope and processed with ImageJ software (<https://imagej.net>).

### 2.10. Sodium-Periodate (NaIO<sub>4</sub>) Treatment/β-Elimination

For the sodium periodate treatment, 20 µg of RNA were mixed with borate buffer pH 8.6 (29.6 mM Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub> 10 H<sub>2</sub>O; 29.6 mM H<sub>3</sub>BO<sub>3</sub>) and 25 mM NaIO<sub>4</sub>, then the reaction was incubated for 40 min at 4 °C in the dark; to quench the unreacted NaIO<sub>4</sub>, glycerol was added (10% final) and samples were incubated for 30 min at 4 °C; then, RNA was ethanol-precipitated and resuspended in RNase-free water. For the β-elimination step, borate buffer pH 9.5 (29.6 mM Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub> 10 H<sub>2</sub>O; 29.6 mM H<sub>3</sub>BO<sub>3</sub>) was added and samples were incubated for 90 min at 45 °C; then, RNA was ethanol-precipitated and resuspended in RNase-free water. The experiment was performed in triplicate. To monitor the efficiency of the reaction, single-stranded (SS) spike-in molecules were chemically synthesized at Exiqon, adapting the sequences described in Locati et al. [26] to our conditions. Two non-methylated (SS-22 and SS-28) and two methylated (mSS-22 and mSS-28) of 22 nt and 28 nt were used:

SS-22: 5'-UAUCAGUGCUACGUGUCUCAGU-3'

SS-28: 5'-CGUAUCGUCUCUGAGUCACUAUCUAC-3'

mSS-22: 5'-AUGCUGAUGAUAGACGCUACUmG-3'

mSS-28: 5'-AGAUAGUACUGAUCUGCUGCGACGAGUmG-3'

Spike-ins were pooled in equimolar concentration, then a final amount of  $0.3 \times 10^{10}$  molecules/µg of RNA was added to each sample before the treatment.

### 2.11. RNA Extraction and Sequencing with Illumina Technologies

Total RNA extraction from cell lines was performed using TRIzol reagent (Invitrogen, Carlsbad, Canada) following the manufacturer's instructions. For small RNA sequencing, indexed libraries were prepared using 1 µg of total RNA with a TruSeq small RNA Sample Prep Kit (Illumina, San Diego, Canada). As a control for germline piRNAs, a human testis RNA pool (Biochain, Newark, Canada) was included in all experiments.

Equimolar concentration of each library was sequenced for 75 cycles on the NextSeq 500 platform (Illumina, San Diego, CA, USA). For RNA sequencing, 1 µg of total RNA was used for library preparation with a TruSeq Stranded Total RNA Sample Prep Kit (Illumina, San Diego, CA, USA) and sequenced at a concentration of 8 pM/lane (paired-end, 2 × 100 cycles) on the HiSeq 2500 platform (Illumina, San Diego, CA, USA). All sequencing experiments were performed at Genomix4life, Italy (<https://www.genomix4life.com/>).



### 2.12. RNA Immunoprecipitation and Sequencing (RIP-Seq)

RNA immunoprecipitation of PIWIL1-associated small RNAs was performed in biological triplicate. Briefly, COLO 205 cells were lysed in polysome lysis buffer as described by Keene et al., 2006 [27], then an aliquot of extract corresponding to  $1 \times 10^6$  cells was taken as input, while 8 mg of protein lysate were mixed with one of the following antibodies: rabbit anti-PIWIL1 (ab105393, Abcam, Cambridge, UK), rabbit anti-PIWIL1 (ab12337, Abcam, Cambridge, UK), rabbit IgG isotype control (02-6102, Thermo Fisher Scientific, Carlsbad, Canada), or without primary antibody addition (No Antibody Control Beads). After overnight incubation, Dynabeads M-280 Sheep Anti-Rabbit IgG (11203-D, Thermo Fisher Scientific, Carlsbad, Canada) were added and the mixture was incubated again for 3 h at 4 °C with gentle rotation. After binding, 1/10 of each sample was taken for WB analysis, while 9/10 were used for the isolation of PIWIL1 co-precipitated RNAs with the mirVana PARIS Isolation Kit (AM1556, Ambion Inc., Austin, TX, USA). For small RNA sequencing analyses, 500 ng of RNA were used for indexed sequencing library preparation with a TruSeq small RNA Sample Prep Kit (Illumina, San Diego, CA, USA).

### 2.13. Bioinformatic Data Analysis

To identify small RNA sequences, preprocessing, alignment, and annotation to small RNAs were performed with the SPORTS1.0 pipeline [28], with the only difference that miRBase was updated to 22.1 and piRBase to the 2.0 release. In brief, SPORTS1.0 performs sequential mapping against the reference genome, miRBase, rRNA, GtRNAdb, Ensembl, Rfam, and finally, piRBase. Unannotated sequences matching to the reference genome were extracted from the final resulting file that includes all reads. The complete list of small RNA databases used, the details of small RNA and methylated piRNA identification, the strategy used to identify PIWIL1-bound piRNAs in the RIP-Seq experiment, and the functional enrichment and pathway analysis are reported in the Supplementary Methods. piRNA target prediction was performed with iSmaRT [29].

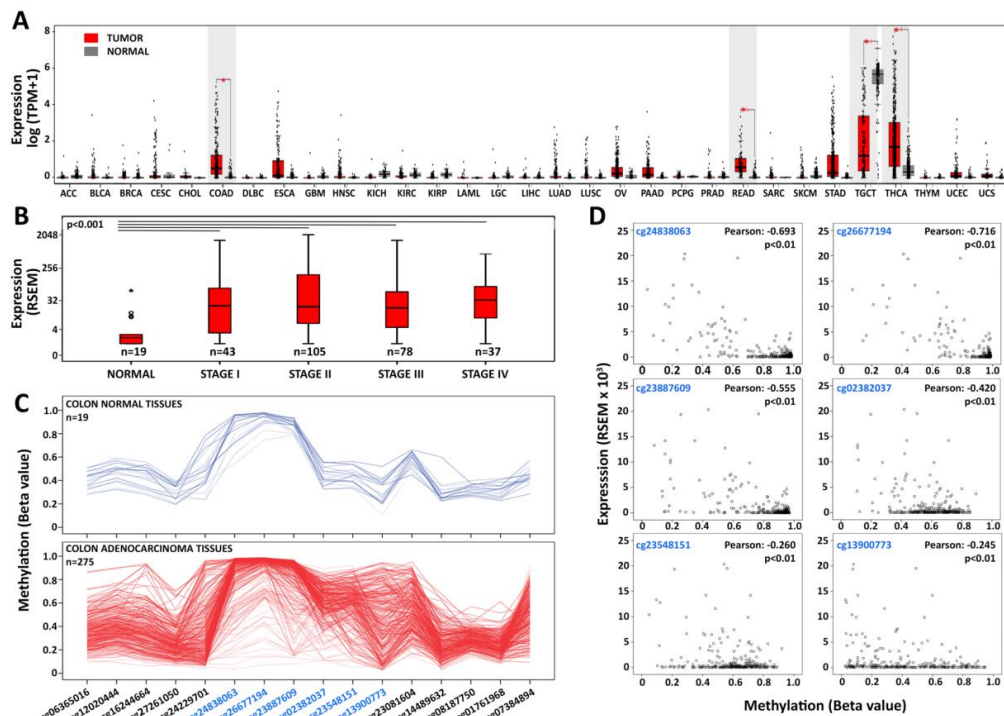
### 2.14. Availability of Supporting Data

Sequencing data were deposited in the EBI ArrayExpress database (<http://www.ebi.ac.uk/arrayexpress>) with the following Accession Numbers: E-MTAB-8115, E-MTAB-8116, E-MTAB-8117, and E-MTAB-8121.

## 3. Results

### 3.1. The Germline-Specific PIWIL1 Gene is Aberrantly Expressed in Colorectal Adenocarcinomas

To investigate the expression of *PIWIL1* in human normal and cancerous tissues, we mined RNA-Seq data in The Cancer Genome Atlas (TCGA) and/or the Genotype-Tissue Expression (GTEx) databases, including 9736 tumoral (T) and 8587 non-tumoral (N) samples organized in 31 cohorts. Results showed that in non-pathological tissues, *PIWIL1* is almost exclusively detectable in the germline (Figure 1A). In tumors, instead, this gene is significantly up-regulated ( $FC > 1.5$ ;  $p$  value  $< 0.05$ ), compared to matched normal tissues, in colon adenocarcinoma (COAD), rectum adenocarcinoma (READ), and thyroid carcinomas (THCA) when considering both dataset (Figure 1A), or in COAD and THCA in the TCGA data (Figure S1A).



**Figure 1.** The germline-specific *PIWIL1* gene is ectopically activated in colorectal cancer. **(A)** Integrative analysis of The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression GTEx data showed that *PIWIL1*, abundantly expressed in the germline (TCGC, normal), is aberrantly activated in cancer and significantly upregulated (highlighted in grey) in colon (COAD), rectum (READ), and thyroid carcinomas (THCA) with respect to the normal tissue ( $\log_2$  (median (Tumor)–median (Normal))  $> 0.5$ ;  $p$  value  $< 0.01$ ). Expression levels are displayed in  $\log_2$ -transformed (TPM (transcripts per million) + 1). ACC: adrenocortical carcinoma; BLCA: bladder urothelial carcinoma; BRCA: breast invasive carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL: cholangio carcinoma; COAD: colon adenocarcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; ESCA: esophageal carcinoma; GBM: glioblastoma multiforme; HNSC: head and neck squamous cell carcinoma; KICH: kidney chromophobe; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LAML: acute myeloid leukemia; LGG: brain lower grade glioma; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; OV: ovarian serous cystadenocarcinoma; PAAD: pancreatic adenocarcinoma; PCPG: pheochromocytoma and paraganglioma; PRAD: prostate adenocarcinoma; READ: rectum adenocarcinoma; SARC: sarcoma; SKCM: skin cutaneous melanoma; STAD: stomach adenocarcinoma; TGCT: testicular germ cell tumor; THCA: thyroid carcinoma; THYM: thymoma; UCEC: uterine corpus endometrial carcinoma; UCS: uterine carcinosarcoma. **(B)** Boxplots showing the expression of *PIWIL1* across different AJCC (American Joint Committee on Cancer) stages of TCGA colon adenocarcinomas. **(C)** DNA methylation profile of CpGs in the promoter region of *PIWIL1* in colon normal tissues (in blue) and adenocarcinomas (in red); CpG sites highlighted in blue constitute a CpG island and are strongly methylated in colon normal tissues, while lower methylation levels are observed in several tumors. **(D)** Statistically significant negative correlation between *PIWIL1* expression and DNA methylation at CpG sites hypomethylated in tumors; beta values are plotted on the  $x$ -axis and RSEM (RNA-Seq by expectation-maximization) gene expression values on the  $y$ -axis.

On the other hand, *PIWIL2* and *PIWIL4* show low expression in many normal tissues and are upregulated in several cancer types, while *PIWIL3* is almost undetectable in all tissues considered, except germline (Figure S1). Moreover, the expression of *PIWIL1* seems to represent an early event in CRC carcinogenesis since the gene is significantly upregulated here ( $p$  value  $< 0.001$ ), compared to normal tissues, starting from the earliest stages (I–II) of CRC progression (Figure 1B). DNA

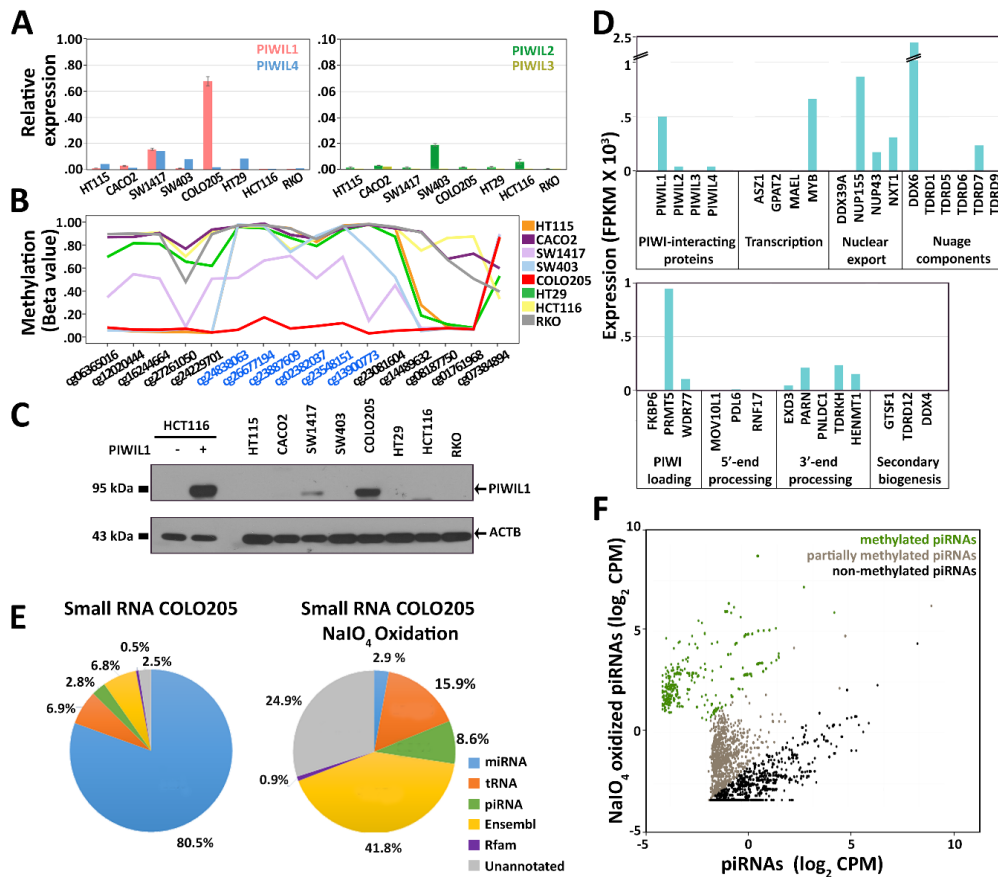
methylation, histone modifications, and chromatin remodeling are important epigenetic regulators of gene expression. Among them, however, an aberrant DNA methylation pattern is a hallmark event in colorectal carcinogenesis, characterized by either global hypomethylation or paradoxical gene-specific hypermethylation [30], which in turn drive neoplastic transformation and cancer progression via oncogene activation and tumor suppressor gene silencing. It has been shown that also *PIWIL1* expression may be linked to its epigenetic asset since, DNA hypermethylation at CpG islands located in the promoter region and first exon of the gene has been verified to occur during germ cell development and to be maintained in somatic cells, all conditions where its expression is not anymore required [31]. CpG island methylation is also the mechanism for *PIWIL1* down-regulation in testicular tumors [32]. Interestingly, we noticed a relationship between *PIWIL1* expression and DNA methylation. Analyzing DNA methylation data from 275 colon adenocarcinoma and 19 normal cases from TCGA, a group of CRC tissues (~14%) shows loss of DNA methylation in a CpG island in the promoter region of *PIWIL1* (Figure 1C, Table S1A). Specifically, we identified 6 CpG sites (cg24838063, cg26677194, cg23887609, cg02382037, cg23548151, cg13900773) in the CpG island located in the promoter and first exon of *PIWIL1*, strongly demethylated in cancers. Moreover, we observed a significant negative correlation ( $p$  value < 0.01 and Pearson correlation < -0.2) between *PIWIL1* promoter methylation involving all six CpGs and *PIWIL1* mRNA levels (Figure 1D). Loss of DNA methylation and re-expression is a peculiar feature of cancer testis antigens (CTAs), germline-specific genes that are normally downregulated in somatic differentiated tissues. Interestingly, *PIWIL1* has been listed among very highly expressed cancer testis genes (EECTGs), considered epigenetic drivers of cancer as when expressed, they confer a selective growth advantage to the transformed cell; they therefore may represent potential targets for new cancer therapies [33]. In line with these findings, Xie et al. [34] recently reported a similar mechanism for *PIWIL1* re-expression in lung cancer, where gene re-activation is attributed to DNA hypomethylation. Several studies explored the effects of DNA methylation inhibitors, such as 5-aza-2-deoxycytidine (5-AZA-CdR), and on gene reactivation in cancer, and two studies have investigated, in particular, their effects in several CRC cell lines [30,35]. We analyzed the results of these studies focusing on *PIWIL1* and found that in Caco2, HCT 116, HT29, and RKO cells, the promoter of this gene is not demethylated by treatment with 5-AZA-CdR, and, consequently, expression of the corresponding mRNA is not enhanced ([30,35] and data not shown). On the other hand, Yagi et al. showed a slightly positive effect on *PIWIL1* expression of a combined treatment with 5-AZA-CdR and Trichostatin A (TSA, a histone deacetylase inhibitor) in HCT 116 cells, while upregulation of the gene was observed using only 5-AZA-CdR in the SW480 cell line, which already expresses this gene at low levels ([27] and data not shown). Based on these results, it is not possible to exclude that *PIWIL1* silencing might involve a combination of DNA methylation and repressive histone modifications.

### 3.2. *PIWIL* and piRNA Expression in CRC Cell Lines

Interestingly, a recent genome-wide CRISPR–Cas9 loss-of-function genetic screening, prompted by the Achilles project of the Broad Institute to create a catalog of genes essential for cell survival [36], showed that *PIWIL1* knockout has an effect on survival in 55 cell lines from different tumor types, including CRC (Supplementary Table S2).

To investigate the relationship between PIWILs and piRNAs in cancer, we focused on CRC cell lines as model systems. To this end, we evaluated expression of *PIWIL* genes in the European Genome-Phenome Archive (EGA) [37], a database containing RNA-Seq data from a collection of 53 cell lines isolated from CRCs. Interestingly, the molecular heterogeneity of *PIWIL* expression observed in CRC biopsies was well recapitulated in CRC cell lines (Figure S2). According to gene expression data, we selected 8 CRC cell lines for further study. Real-time qRT-PCR was used to evaluate the expression level of all *PIWIL* genes in the 8 cell lines (Figure 2A). Results show that *PIWIL1* transcripts were highly abundant in COLO 205 and medium-low expressed in SW-1417 cells, while *PIWIL2* and *PIWIL4* were expressed at low levels in 6 out of 8 cell lines, and *PIWIL3* was undetectable in all cases, with HCT 116 and RKO cells negative for all *PIWIL* transcripts.





**Figure 2.** Characterization of main molecular features of the piRNA pathway in colorectal cancer (CRC) cell lines. (A) The expression level of *PIWIL* mRNAs in 8 CRC cell lines was determined by real-time qRT-PCR. The highest expression for *PIWIL1* was observed in COLO 205 and SW-1417. *PIWIL2* and *PIWIL4* were identified in some cell lines but at lower levels, while *PIWIL3* was undetectable in all cases. (B) DNA methylation profile of *PIWIL1* promoter in 8 CRC cell lines; a marked hypomethylation, correlated to *PIWIL1* expression, was observed for two cell lines (COLO 205 and SW-1417), while a strong methylation was present in the absence of *PIWIL1*. CpG sites highlighted in blue constitute a CpG island. (C) Western blotting analysis performed onto total protein lysate from 8 CRC cell lines confirmed the presence of abundant *PIWIL1* levels in COLO 205 and medium-abundant levels in SW-1417; as a control to test the specificity of the antibody, total protein lysate from HCT 116 cells overexpressing *PIWIL1* or transfected with empty vector were loaded in the same gel;  $\beta$ -actin was used as a loading control. (D) Gene expression (indicated as fragment per kilobase million, FPKM) was measured for known germline components of the piRNA pathway genes by RNA-Seq; genes were grouped according to their role in the piRNA pathway. (E) Pie-chart representing the relative abundance of small RNA categories identified in non-treated (left) and sodium-periodate ( $\text{NaIO}_4$ )/ $\beta$ -elimination oxidized RNA (right) from COLO 205; small RNA-Seq analysis confirmed the presence of piRNAs in the COLO 205 cell line and their resistance to the sodium periodate/ $\beta$ -elimination treatment. (F) Scatterplot representing the effect of the sodium periodate ( $\text{NaIO}_4$ )/ $\beta$ -elimination treatment on the piRNAs population;  $\log_2$  CPM (count per million) values are displayed per each COLO 205 piRNA in treated ( $y$ -axis) and non-treated ( $x$ -axis) RNA; the treated (T) vs. non-treated (NT)  $\log_2$  ratio of expression values ( $p$  value < 0.01) was used to identify 767 methylated piRNAs ( $\log_2(T/NT) > 2$ ), 1514 partially methylated piRNAs ( $-2 < \log_2(T/NT) < 2$ ), and 2078 non-methylated piRNAs ( $\log_2(T/NT) < -2$ ) show that *PIWIL1* transcripts were highly abundant in COLO 205 and medium-low expressed in SW-1417 cell lines, while *PIWIL2* and *PIWIL4* genes were expressed at low levels in 6 out of 8 cell lines and *PIWIL3* was undetectable in all cases with HCT 116 and RKO cells negative for all *PIWIL* transcripts.

In line with CRC data from TCGA, we observed a demethylation of the CpG island in the *PIWIL1* promoter region in the two cell lines expressing the gene, SW-1417 and COLO 205 (Figure 2B, Table S1B). Loss of DNA methylation was also proportional to the level of expression of the gene; a direct relationship between the two data suggesting the existence of a common epigenetic mechanism responsible for *PIWIL1* re-expression in both CRC tissues and cell lines. WB analysis on total protein lysates from the selected cell lines confirmed the presence of PIWIL1 in SW-1417 and COLO 205 (Figure 2C).

Our attention was then focused on piRNAs, and small RNA-Seq was applied to evaluate the presence and abundance of these small non-coding RNAs in CRC cell lines. As a control for piRNA identification, human testis RNA was sequenced in parallel. Since piRNAs map to thousands of genomic regions, reads corresponding to miRNAs, rRNA-derived small RNAs, tRNA-derived small RNAs, and small non-coding RNAs from the Ensembl and Rfam databases were filtered out before piRNA annotation to avoid misidentifications (Table S3).

Nevertheless, we identified a high number of piRNAs (~6000, with expression value > 0.1 count per million, CPM) in all colon cell lines and ~15,000 in the germline (Table S4). Surprisingly, we found that piRNAs are transcribed in colon cancer cells independently from *PIWIL* expression, as a comparable number of piRNAs were identified in all cell lines analyzed. Previous studies reported the presence of low levels of piRNAs transcribed outside the germline, in somatic tissues [9]. We found that the piRNA population identified is expressed at relatively low levels also in CRC cells, where it represents a small percentage of all reads assigned to small RNAs, <5% compared to ~17% in human testis (Figure S3, Figure S4, Table S3). The lack of correlation between *PIWIL* expression and piRNA levels suggests that basal piRNA transcription may occur in CRC cells—and possibly in all tumors—independently of the presence and abundance of PIWIL proteins. As a consequence, small amounts of piRNAs are always detectable, with their relative number and abundance being linked to the cell type, rather than to the presence/absence of PIWIL proteins. In fact, specific piRNA expression profiles can be observed in different cell lines, as shown by a hierarchical clustering analysis (single-linkage, Pearson distance) for 2500 piRNAs (>1.55 CPM) sorted by their abundance in the COLO 205 cell line (Figure S5, [9]). Based on the assumption that piRNAs are engaged in the PIWI-induced silencing complex (piRISC) to regulate gene expression, we decided to focus on the PIWIL1–piRNA complex and analyzed it in detail in COLO 205 cells, where the highest amount of PIWIL1 was detected in the absence of other PIWIL proteins (see below).

### 3.3. Characterization of PIWI–piRNA Pathway in COLO 205

Since piRNA maturation from precursors, their interaction with PIWIL proteins, and consequent downstream activities are mediated by a number of cofactors, identified and thoroughly studied in the germline and conserved across a wide range of animals [5], the expression of key components of the PIWI–piRNA pathway was evaluated in COLO 205 cells by RNA-Seq.

Results summarized in Figure 2D and Table S5 show the expression of several human homologs of piRNA-related genes previously described in the animal germline cells, including several factors involved in primary piRNA biogenesis and nuclear exporting. These include, together with factors involved in piRNA loading in PIWIL proteins and intermediate processing, DDX6, an RNA helicase, and the Tudor domain-containing protein TDRD7, belonging to the piRNA processing machinery in the germline, where it has been shown in multiprotein structures of the perinucleus called nuage [38]. On the other hand, it was not possible to detect expression of any factor involved in the secondary piRNA biogenesis pathway.

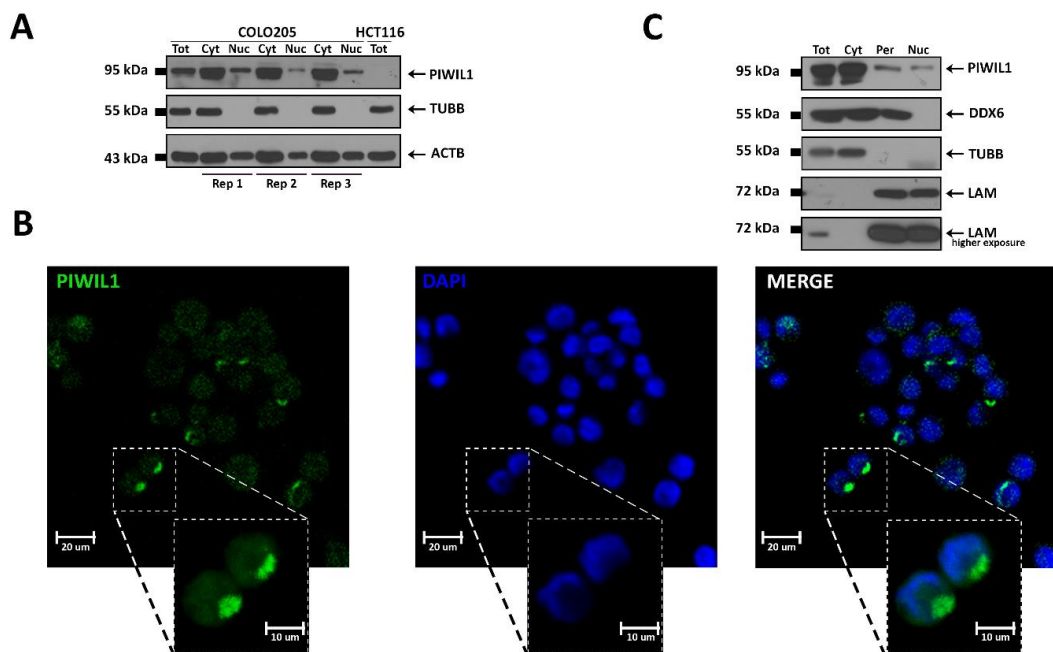
These results support a growing body of evidence indicating that genes of this pathway are expressed in different types of cancer, so this pathway is active and could play a key role in cancer cells. Recently, a *PIWIL1* knock-out experiment in COLO 205 cells showed a mild global effect on gene expression [39]. Differential expression analysis of these data accompanied by pathway analysis reveals, however, that several cellular processes and signaling pathways are significantly affected by *PIWIL1* deficiency in these cells (Table S6).

It is noteworthy that we also noticed, among others, the presence in COLO 205 of HENMT1, the methyltransferase responsible for 3'-end 2'-*O*-methylation of piRNAs, a typical feature of these small RNAs shared among all organisms [40]. Based on this result, the 3'-methylation status of piRNAs was investigated in triplicate by sodium periodate/ $\beta$ -elimination treatment of total RNA extracted from COLO 205, followed by RNA sequencing (Table S3). As an internal control, to monitor the efficiency of the treatment, a pool of 4 synthetic molecules, two carrying the 3'-end methyl group characteristic of piRNAs and two with a free 3'-OH, as miRNAs and other small RNAs were spiked-in in each sample before processing. RNAs lacking the 3'-end methyl group were cleaved at the terminal ribose, preventing their cloning during library preparation and, therefore, detection by subsequent sequencing, while methylated molecules were incorporated and sequenced. Results showed that depletion efficiency for non-methylated RNA molecules, calculated from the spike-in abundance, was close to 98% (Figure S6). Sodium periodate oxidation had a strong impact on library composition, dramatically depleting non-methylated molecules such as miRNAs (from 80% to 3%), while a 3-fold increase was observed in piRNA reads (from 3% to 9%). As a control, the experiment was performed in parallel on a human testis RNA sample, where piRNA reads underwent to a 2.15-fold increase after the treatment (from 17.10% to 36.82%, Figure S4). Specifically, most of COLO 205 piRNAs (54.6%) were still detectable after the treatment, but the relative abundance and composition of the piRNA population in the library was strongly impacted. In fact, several abundant piRNAs were depleted by the treatment, while many less abundant molecules were efficiently sequenced (Figure S7). This last result can be explained considering both variable methylation levels of individual piRNAs and different library incorporation efficiency for 3'-end methylated and non-methylated molecules, since non-methylated RNAs are cloned with high efficiency, contrary to methylated ones that are, instead, incorporated with much lower efficiency [41]. On the other hand, methylated piRNAs are better detected after the treatment, due to the massive removal of non-methylated molecules and consequent reduction of competition during the enzymatic steps (Figure S7). Sequencing data showed that many piRNAs present in this cell line are methylated, including both completely and partly methylated ones. According to the ratio between piRNA detection levels before (NT) and after (T) treatment, we grouped those expressed in COLO 205 cells in three categories: 767 highly methylated ( $\log_2(T/NT) > 2$ ), 1514 partially methylated ( $-2 < \log_2(T/NT) < 2$ ), and 2078 non-methylated ( $\log_2(T/NT) < -2$ ) (Figure 2F and Table S7).

Other categories were enriched by sodium periodate/ $\beta$ -elimination treatment, including tRNAs (from 7% to 16%) and those listed as Ensembl RNAs (snRNA, snoRNA, lincRNA, IncRNA, etc.), suggesting the presence of methylated species also in one or more of these classes of sncRNAs (Figure 2E), as reviewed by Dimitrova et al. [42]. It is also interesting to note that a strong enrichment upon treatment (from 2% to 30%, 468 methylated ( $\log_2(T/NT) > 2$ ) and 225 partially methylated ( $-2 < \log_2(T/NT) < 2$ ) molecules, Figure 2E and Table S7) was found also for unannotated reads, a category that comprises RNAs detectable by sequencing but not yet listed in current databases, some of which could represent still-uncharacterized piRNAs.

### 3.4. PIWIL1 Localizes in a Nuage-Like Structure of COLO 205 Cells

Since PIWIL proteins can perform both cytoplasmic and nuclear functions [43], we investigated the subcellular localization of PIWIL1 in COLO 205 (Figure 3), SW1417, and *PIWIL1*-transfected HCT 116 cells (Figure S8). Fractionated protein extraction revealed that most of the protein concentrates in the cytoplasm (Figure 3A and Figure S8A–C), a result confirmed by immunocytochemistry, which revealed how a sizeable fraction of the protein is concentrated in a discrete cytoplasmic region leaning against the nucleus (Figure 3B and Figure S8D). This discrete region is more evident in COLO 205 cells, which express a high level of PIWIL1, compared to SW1417, where the protein is much less abundant (compare Figure 3B and Figure S8D). These results were also confirmed in *PIWIL1*-transfected HCT 116 cells where the IF signal, diffused throughout the cytoplasm, is clearly evident around the nucleus (Figure S8D).



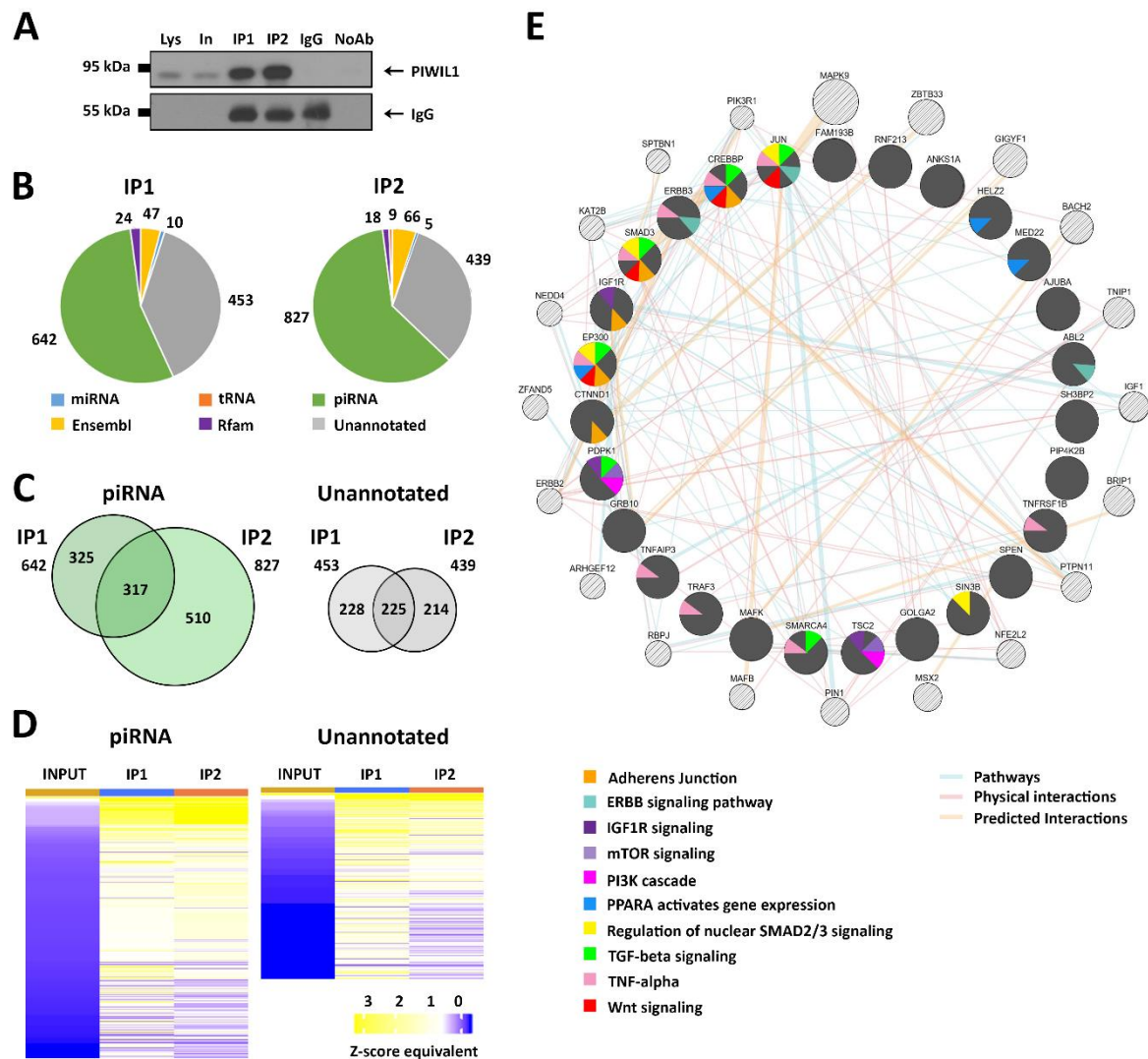
**Figure 3.** Characterization of PIWIL1 subcellular distribution. (A) Cell fractionation performed in triplicate in COLO 205 cells revealed the presence of PIWIL1 in both cytoplasmic and nuclear compartments; as a comparison, total protein lysate from COLO 205 and from HCT 116 (negative control) were run on the same gel;  $\beta$ -actin (ACTB) was used as a loading control, while  $\beta$ -tubulin (TUBB) was used to verify the absence of cross-contamination between cytosolic and nuclear fractions. (B) Localization of PIWIL1 protein (green) in COLO 205 cells by immunocytochemical analysis revealed its presence in both compartments, nucleus and cytosol, with especially bright staining in a discrete region located in proximity of the nucleus (blue, DAPI staining). (C) Cytosol/perinucleus/nucleus fractionation showed the presence of PIWIL1 in the perinuclear region of COLO 205 cells; in the same fraction, DDX6 protein, a marker for the human germline nuage compartment, was detected; TUBB was used to verify the absence of cytoplasmic contamination of perinuclear and nuclear fractions and  $\beta$ -lamin (LAM) as a control for perinuclear and nuclear extraction. As a comparison, total protein lysate from COLO 205 was loaded on the same gel.

In germline, PIWIL proteins and their interactors localize in proximity of the nucleus in a structure called nuage, where piRNA precursor transcripts are exported from the nucleus and biogenesis is thought to occur. To test the possibility that these somatic cells might contain a similar structure, we used a more stringent cell fractionation protocol [25] to isolate the perinuclear region of COLO 205, which was found to contain approximately 20% of the total protein content (Figure S9), and confirmed that ~10% cellular PIWIL1 is localized in the perinuclear region, ~4% in the nucleus, and the remainder in the cytosolic fraction (Figure 3C and Figure S9). Importantly, in the perinuclear fraction of COLO 205, we detected by WB also the presence of the nuage component DDX6 [44] (Figure 3C), which was also present in the cytoplasmic but not in the nuclear fraction. This result is of particular interest since it provides, for the first time, evidence that the perinuclear region of colon cancer cell lines contains a structure with similar localization and protein markers of the germline nuage compartment, the site of piRNA maturation and main actions in germline cells.

### 3.5. Identification of RNAs Associated with PIWIL1 in COLO 205 Cells

To identify and characterize the PIWIL1–piRNA complexes present in the COLO 205 cell line, RNA immunoprecipitation (RIP) was performed with PIWIL1-specific antibodies (Abs) on total protein extracts from exponentially growing cultures, followed by library preparation and sequencing. PIWIL1-bound RNAs were identified by performing independent reactions with two Abs (IP1 and IP2), each raised against a different PIWIL1 epitope; as a negative control, the protocol was carried out in parallel with non-specific IgGs as well as without antibody addition, and the whole experiment was performed on biological triplicates. As shown in Figure 4A, PIWIL1 was efficiently immunopurified by both Abs (Figure 4A). After data filtering and normalization, RNA molecules enriched in either IP1 or IP2 immunoprecipitated samples with respect to the input lysates ( $\log_2$  FC > 2, FDR < 0.01) were identified. To exclude non-specific interactions, RNAs that resulted in being enriched also in the “No Antibody” control samples with respect to the input lysates were excluded from the analysis (Table S8). The two antibodies provided very similar results, showing that more than 50% of the PIWIL1-interacting RNAs identified correspond to known piRNAs, while most of the remaining RNAs are unannotated, suggesting that they could also include unclassified somatic piRNAs from COLO 205 cells (Figure 4B,C). The enrichment factor, calculated for each piRNA in immunoprecipitated samples with respect to its concentration in the input lysates, ranged between 4 and 64, and a similar behavior was observed for unannotated RNAs (Figure 4D and Table S8). The other small RNAs found associated with PIWIL1 by RIP include some miRNAs and sncRNAs included in the Ensembl and Rfam databases, comprising some (in particular, snoRNAs) that are considered piRNA precursor transcripts [45,46]. In order to focus on the most reliable data, the results considered consisted only of molecules immunopurified with both antibodies, represented by 583 small RNA species associated with PIWIL1, including 317 piRNAs and 225 unannotated small RNAs, some of which could represent new somatic piRNAs (Figure 4D).





**Figure 4.** Identification of PIWIL1-bound RNAs. (A) RNA immunoprecipitation (RIP) was performed in the cytosolic fraction of COLO 205 cells with two different antibodies (IP1 and IP2); the presence of PIWIL1 protein was assessed by western blotting in the cytosolic lysate (Lys), Input (In), IP1, and IP2; PIWIL1 was absent in negative controls (IgG and NoAb); IgG were used as a control for IP loading. (B) Pie chart representing the number of enriched molecules per category in both IP1 and IP2 samples; (C) Venn diagram summarizing the number of piRNAs (left) and unannotated molecules (right) identified with both antibodies; (D) heatmaps representing Z-score expression values for 317 piRNAs (on the left) and 225 unannotated RNAs (on the right) identified in PIWIL1 RIP-Seq with two different antibodies (IP1 and IP2), compared to the input RNA. (E) Network analysis and functional enrichment for 106 validated interactions revealed a group of 27 master-regulator genes targeted by the PIWIL1–piRNA complex and 20 correlated genes strongly related with molecular pathways crucial for CRC pathogenesis. Relations between genes (173 links, Table S9E) and association to pathways are represented with a color code. Abbreviations: ERBB, avian erythroblastosis oncogene B; IGF1R, insulin like growth factor 1 receptor; mTOR, mechanistic target of rapamycin kinase; PI3K, phosphoinositide 3-kinase; PPARA, peroxisome proliferator activated receptor alpha; SMAD2/3, mothers against decapentaplegic homolog 2/3; TGF, transforming growth factor; TNF, tumor necrosis factor; Wnt, wingless/int-1.

### 3.6. Identification of mRNAs Targeted by the PIWIL1–piRNA Complex

To investigate the biological function of the PIWIL1–piRNA complex in colon cancer, we performed a target prediction analysis for 317 PIWIL1-bound piRNAs. As a result, we identified 3288 putative target transcripts expressed in the COLO 205 cell line, as determined by RNA-Seq. Interestingly, the number of piRNA target sequence was greater in the 3'UTR of the RNA (2673) compared to the 5'UTR (499) or the Coding Sequences (CDS) region (524) (data not shown). To search for biological processes involving the PIWIL1–piRNA complex, we performed a pathway analysis based on the known functions encoded by the 3288 mRNAs. Results showed a significant enrichment in functional categories involved in the cell cycle, including gene expression, circadian clock, G2/M transition, and metabolism, some of which become dysregulated in the same cell line upon PIWIL1 deprivation (highlighted in bold in Table S9A). In addition, this functional analysis revealed several transcripts encoding key components of signal transduction cascades frequently found deregulated in CRC, including the TGF- $\beta$ , Wnt, mTOR, and PI3K pathways. Interestingly, several other processes potentially regulated by the piRNA pathway emerged, the roles of which are still poorly understood in CRC; among these, it is worth mentioning the signaling pathway comprising of histone deacetylase (HDAC) class II molecules, TNF-alpha, and PPAR-alpha (Figure S10, Table S9A).

Based on the consideration that PIWIL1 is in direct contact with piRNAs and their target in the piRISC complex, we reasoned that the RIP-Seq protocol also allows the simultaneous isolation of piRNAs and their target RNAs. To verify the presence of genes enriched in the piRISC, we mapped the longest reads (46 to 75 nucleotides long) generated from all small RNA-Seq datasets to the Ensemble transcripts database (<http://www.ensembl.org>). Then, following the same approach used for PIWIL1-bound piRNA identification, we identified a significant physical interaction between the PIWIL1–piRNA complex and 106 protein-coding transcripts (Table S9B) through sequence complementarity with 90 different piRNAs (Table S9C). The specific fold-enrichment calculated between the input and PIWIL1 immunoprecipitated samples varied between 1.5 and 8.3 for IP1 and 4.8 for IP2 Abs (Table S9B). These transcripts are most likely piRNA targets, and, importantly, they have been identified in the context of the piRISC in the absence of any crosslinking step, thereby avoiding possible artifacts produced by this treatment. To further analyze the gene network directly engaged by PIWIL1–piRNA complexes in COLO 205 cells, starting from the 106 experimentally validated target transcripts, we built a functional interaction network based on a combination of various functional descriptors shared between the partners, such as physical and predicted interactions and functional pathways, using the GeneMANIA platform (<http://www.genemania.org>) [47]. Results show 27/106 genes (dark gray in Figure 4F), as well as 20 additional related genes (light gray) predicted by GeneMANIA, mapped to an interconnected network showing 173 interactions (Table S9D and S9E). In this context, the most significantly affected functional pathways identified include, among others, TNF-alpha, SMAD2/3, TGF-beta, and Wnt signaling (Figure 4F), suggesting that specific PIWIL1–piRNA–mRNA interactions may be controlling essential cellular functions in CRC cells by post-transcriptional gene regulation mechanisms similar to those identified for PIWIL and piRNAs in germline cells.

## 4. Discussion

An increasing number of studies have demonstrated that germline PIWIL proteins and associated piRNAs are dysregulated in cancers and can be used as diagnostic and prognostic biomarkers [15,48,49]. However, the biological role of the piRNA pathway in pathological conditions is still undefined, principally due to the lack of studies focused on the identification of the molecular targets of the PIWIL–piRNA complex in cancer cells. Interestingly, *PIWIL* genes have been included in the list of cancer testis antigens (CTAs), germline genes whose expression is re-activated in cancer, which are generating great interest in the scientific community as possible druggable targets. In particular, *PIWIL1* has been recently listed among the extremely high-expressed cancer testis genes (EECTGs) and are considered as a potential epi-drivers in tumors [33].

Based on a meta-analysis of TCGA and GTEx data, we found that *PIWIL1* presents unique features among *PIWILs*, as *PIWIL2* and *PIWIL4* are frequently deregulated in cancers but are also expressed in several normal tissues, while *PIWIL1* is almost absent in non-transformed tissues. Furthermore, ectopic expression of this germline-specific gene was almost exclusively detected in CRC, where it is significantly upregulated in 11% of tumors, with respect to normal samples. Based on DNA methylation data, the expression of *PIWIL1* in human colorectal cancers and cell lines is directly associated with the hypomethylation of 6 CpG sites in its promoter region, an event frequently linked to CTA re-expression in pathological conditions. These results allow us to hypothesize that in normal human colon cells, the majority of CpGs in the *PIWIL1* promoter are methylated, leading to gene silencing, whereas loss of DNA methylation, which frequently occurs during malignant transformation for delays in maintenance of methylation during accelerated cell division, could in this case drive abnormal *PIWIL1* expression in a subgroup of CRCs, as also observed in lung adenocarcinomas [34].

During the last few years, piRNA expression profiles have been investigated in several human tumors [10,48], but the association of these small RNAs with *PIWILs* and their involvement in CRC are still poorly investigated. Interestingly, by directly comparing *PIWIL* and piRNA expression levels, we noticed that the presence/absence of *PIWIL* gene expression is not directly related to a specific piRNA expression pattern. Indeed, small RNA-Seq analyses in 8 CRC cell lines revealed that a low percentage of piRNAs (<7%) is detectable in all cases, independently from the expression patterns of *PIWIL* genes, a result supported by gene knock-down experiments in somatic tissues that demonstrated how *PIWIL1* depletion does not significantly affect piRNA levels [50,51]. It is thus possible to assume that *PIWIL1* is not required for piRNA biosynthesis and/or maturation in CRC and, more generally, in somatic tissues. On the other hand, this Argonaute protein is associated with a subset of piRNAs (317 molecules) and other unannotated small RNAs, as demonstrated by *PIWIL1* RIP-Seq in CRC cells (Figure 4C). This evidence, combined with the detection of several members of the piRNA biosynthetic and effector pathway in these same cells (Figure 2D) and the finding that a large fraction of the expressed piRNAs (>50%, Figure 2F) are represented by biologically active methylated molecules, indicates a role for these regulatory RNAs in CRC. Furthermore, the fact that piRNAs are produced in CRC by primary processing and not amplified through the secondary amplification pathway explains their relatively low abundance in these cells, compared to germline ones, in agreement with what was previously demonstrated in other somatic cells [52–54].

Across the animal kingdom, *PIWIL* proteins and associated piRNAs have been found in both nuclei and cytoplasm, suggesting that the piRNA pathway may function in both these cellular compartments [55] and, therefore, may exert multiple roles. Using cell fractionation and immunocytochemistry, we found that *PIWIL1* is primarily localized in the cytoplasm, but, surprisingly, we also observed the existence of a nuage-like structure in the perinuclear space, where the protein is much more concentrated than in the rest of the cell, in analogy with what occurs in the germline nuage compartment, where piRNAs are exported and loaded onto *PIWIL* proteins. Importantly, in the perinuclear fraction of COLO 205 cells, we also detected DDX6, an RNA helicase responsible for piRNAs biogenesis and marker of the germline nuage [44]. Since impairment of the piRNA pathway in animals led to disruption of the nuage compartment [56], together with the evidence that these subcellular structures may serve to increase the local concentration of specific proteins of the pathway and of piRNAs, this compartmentalization observed for *PIWIL1* represents a strong indication that the piRNA pathway is functional in these cells. This is further supported by experimental identification of a physical interaction between *PIWIL1* and 317 piRNAs, as well as 225 unannotated small RNAs that might represent new somatic piRNAs in CRC cells. It is noteworthy that Yin et al. [57] have recently identified 35 piRNAs deregulated in CRC biopsies, with respect to paired normal mucosa, and found a correlation between the expression of some of these RNAs and tumor differentiation and lymph node metastasis. Interestingly, 4 of these piRNAs were also found in our dataset: piR-hsa-28319 (hsa\_piR\_020450 in piRNABank), identified among *PIWIL1*/bound molecules in the RIP-Seq experiment, piR-26039 (hsa\_piR\_018780) and piR-17560 (hsa\_piR\_012753), enriched in *PIWIL1* IP but classified as

snoRNA-derived sncRNA (Ensembl molecules), and piR-28317 (has-piR-23317), identified by small RNA-Seq of COLO 205 but not found associated with PIWIL1 by RIP. This evidence supports the results reported here and further suggests that this pathway may indeed exert important functional roles in CRC.

In recent years, an increasing number of post-transcriptional regulation mechanisms for primary piRNAs, beyond transposon silencing, have been described as piRNAs that can guide PIWI proteins to destabilize specific mRNA targets via RNA–RNA interactions by a miRNA-like mechanism [58], but they can also stabilize mRNAs [59]. Interestingly, piRNA target prediction and RNA immunoprecipitation confirmed the interaction between the PIWIL1–piRNA complex and at least 106 COLO 205 transcripts. In this context, interaction network analysis identified that 27 such transcripts are encoded by genes involved in key pathways that are frequently found deregulated during CRC carcinogenesis and cancer progression. Among these, 8 tumor-suppressor genes are worth mentioning (EP300, CREBBP, CTNND1, SMAD3, SMARCA4, SPEN, TNFAIP3, and TSC2), as well as genes that control normal cell proliferation and differentiation, such as IGF1R, JUN, and ERBB3, to suggest the interesting possibility that the PIWIL1–piRNA machinery may participate in tumorigenesis by directly regulating the activity of these genes. Indeed, PIWIL1 expression in CRC has recently been correlated with tumor differentiation, infiltration, lymph node invasion, and metastasis [17] and has been tightly associated with *BRAF* gene mutation (V600E) and poor prognosis [60]. It is worth noting that Herr et al. have observed that the treatment of COLO 205 cells with *BRAF* inhibitors induces a strong downregulation of PIWIL1, suggesting that the activation of PIWIL1 in this cell line may be responsible for their less differentiated and a stem-like signature and, therefore, for their metastatic behavior, and may contribute to the aggressive phenotype of *BRAF*-mutated tumors [61]. The PIWIL/piRNA-induced regulation of gene expression is a complex mechanism, finely regulated in multiple points, with functions performed in both nuclear and cytoplasmic compartments [49]. Here we focused our attention on the molecular interaction between PIWIL1/piRNA complexes and target mRNAs in the cytosolic fraction of COLO 205; however, it is interesting to note that there is also evidence for a PIWIL1-mediated epigenetic regulation of gene expression in CRC [62]. Moreover, Siddiqi et al. [63] demonstrated, in human sarcoma, that HIWI (PIWIL1) mediated tumorigenesis is associated with global DNA-hypermethylation and silencing of tumor suppressor genes; it is interesting to note, among them, is present SMAD3, also identified here among the molecular target of PIWIL1/piRNA complexes. Furthermore CREBBP, EP300, ERBB3, IGF1R, JUN, SMAD3, SMARCA4 and TSC2 genes, identified here as targets of PIWIL1/piRNA complexes, were found differentially expressed between consensus molecular subtype (CMS) groups in primary CRCs and metastases [64], thus suggesting the involvement of the PIWIL1–piRNA system as an regulator of the complex genetic network controlling early as well as advanced stages of CRC progression.

## 5. Conclusions

The present study confirmed recent findings of aberrant PIWIL proteins and piRNA expression in cancer but provides, for the first time, a molecular view of the involvement of these small regulatory RNAs and their pathway in CRC. We demonstrated that among *PIWIL* genes, *PIWIL1* presents some unique features, such as germline specificity and ectopic activation in CRC due to a loss of DNA methylation in the promoter region of the gene, and all peculiar features of CTAs, which are a promising target for the development of cancer-specific immunotherapies. CRC piRNAs present the molecular features required by these molecules to be functionally active, such as the methyl group at their 3'-end. Importantly, this study shows, for the first time in cancer cells, the presence of a nuage-like structure, proposed to be central for recruitment of functional components of the PIWI–piRNA pathway and selection and processing of target mRNAs. Finally, we identified a PIWIL1-associated piRNA population involved in gene regulation events that may control basic features of colon cancer, providing in this way, a useful resource for further functional studies to decipher the role and importance



of piRNA-mediated gene regulatory events in view of identifying new biomarkers and potential therapeutic targets.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4409/8/11/1390/s1>. Figure S1: Expression of *PIWIL2*, *PIWIL3*, and *PIWIL4* genes in cancer tissues; Table S1: Methylation and gene expression data of TCGA colon cancer tissues (S1A) and CRC cell lines (S1B); Figure S2: Molecular heterogeneity of *PIWIL* genes in tumors is maintained in CRC cell lines; Table S2: *PIWIL1* CRISPR–Cas9 loss-of-function screening; Table S3: Mapping statistics of samples analyzed by small RNA sequencing; Figure S3: Small RNA profiling of CRC cell lines; Figure S4: Effect of the sodium periodate ( $\text{NaIO}_4$ )/ $\beta$ -elimination treatment in human testis RNA; Table S4: piRNA expression in human colon cancer cell lines and testis; Figure S5: Relative abundance of CRC piRNAs compared to the germline; Table S5: Expression of piRNA pathway genes in COLO 205; Figure S6: Efficiency of the sodium periodate ( $\text{NaIO}_4$ )/ $\beta$ -elimination treatment; Figure S7: Heatmaps showing the effect of the  $\text{NaIO}_4$  treatment on the piRNA population in COLO 205; Table S6: Pathway analysis of *PIWIL1* loss of function in COLO 205 cells; Table S7: Relative mean ratios and count per million (CPM) values of samples treated with sodium periodate/ $\beta$ -elimination in the COLO 205 cell line; Table S8: Enrichment values for RIP-Seq experiment; Figure S8: Characterization of *PIWIL1* subcellular distribution; Figure S9: Evaluation of subcellular distribution of *PIWIL1* in the COLO 205 cell line; Figure S10: Pathway analysis of putative piRNAs mRNA targets; Table S9: Analysis of *PIWIL1*–piRNAs target genes in COLO 205; Supplementary Methods.

**Author Contributions:** Conceptualization, A.W. and F.R.; investigation, A.S., Y.D., M.C., E.A., and D.R.; software and formal analysis, K.G. and G.G.; writing—original draft preparation, A.S. and F.R.; writing, review and editing, A.W, G.N., G.G., R.T., and F.R.

**Funding:** This research was funded by the Italian Association for Cancer Research, grant number IG-17426, Italian Ministries of Education, University and Research (Flagship Project InterOmics), University of Salerno (Fondi FARB 2017), Regione Campania, Progetto GENOMAeSALUTE (POR Campania FESR 2014/2020, azione 1.5; CUP:B41C17000080007), Regione Campania (“La Campania lotta contro il cancro” project Rare-Plat-Net, CUP: B63D18000380007), Genomix4Life and the European Union’s Horizon 2020 research and innovation programme Evocell ITN project, Marie Skłodowska-Curie grant number 766053. AS and KG are PhD Students of the Research Doctorates in Biomedical Sciences and Technologies of the University Roma Tre and in Molecular Oncology, Immunology and Experimental Development of Innovative Therapies of the University of Catanzaro “Magna Graecia”, respectively.

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

## References

1. Simpson, A.J.G.; Caballero, O.L.; Jungbluth, A.; Chen, Y.-T.; Old, L.J. Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer* **2005**, *5*, 615–625. [[CrossRef](#)] [[PubMed](#)]
2. Gordeeva, O. Cancer-testis antigens: Unique cancer stem cell biomarkers and targets for cancer therapy. *Semin. Cancer Biol.* **2018**, *53*, 75–89. [[CrossRef](#)]
3. Han, Y.-N.; Li, Y.; Xia, S.-Q.; Zhang, Y.-Y.; Zheng, J.-H.; Li, W. PIWI Proteins and PIWI-Interacting RNA: Emerging Roles in Cancer. *Cell. Physiol. Biochem.* **2017**, *44*, 1–20. [[CrossRef](#)] [[PubMed](#)]
4. Rojas-Ríos, P.; Simonelig, M. piRNAs and PIWI proteins: Regulators of gene expression in development and stem cells. *Development* **2018**, *145*, dev161786.
5. Ozata, D.M.; Gainetdinov, I.; Zoch, A.; O’Carroll, D.; Zamore, P.D. PIWI-interacting RNAs: Small RNAs with big functions. *Nat. Rev. Genet.* **2019**, *20*, 89–108. [[CrossRef](#)] [[PubMed](#)]
6. Peng, L.; Song, L.; Liu, C.; Lv, X.; Li, X.; Jie, J.; Zhao, D.; Li, D. piR-55490 inhibits the growth of lung carcinoma by suppressing mTOR signaling. *Tumour Biol.* **2016**, *37*, 2749–2756. [[CrossRef](#)] [[PubMed](#)]
7. Rouget, C.; Papin, C.; Boureux, A.; Meunier, A.-C.; Franco, B.; Robine, N.; Lai, E.C.; Péliisson, A.; Simonelig, M. Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature* **2010**, *467*, 1128–1132. [[CrossRef](#)] [[PubMed](#)]
8. Hashim, A.; Rizzo, F.; Marchese, G.; Ravo, M.; Tarallo, R.; Nassa, G.; Giurato, G.; Santamaria, G.; Cordella, A.; Cantarella, C.; et al. RNA sequencing identifies specific PIWI-interacting small non-coding RNA expression patterns in breast cancer. *Oncotarget* **2014**, *5*, 9901–9910. [[CrossRef](#)]
9. Martinez, V.D.; Vucic, E.A.; Thu, K.L.; Hubaux, R.; Enfield, K.S.; Pikor, L.A.; Becker-Santos, D.D.; Brown, C.J.; Lam, S.; Lam, W.L. Unique somatic and malignant expression patterns implicate PIWI-interacting RNAs in cancer-type specific biology. *Sci. Rep.* **2015**, *5*, 10423. [[CrossRef](#)]
10. Ng, K.W.; Anderson, C.; Marshall, E.A.; Minatel, B.C.; Enfield, K.S.S.; Saprunoff, H.L.; Lam, W.L.; Martinez, V.D. Piwi-interacting RNAs in cancer: Emerging functions and clinical utility. *Mol. Cancer* **2016**, *15*, 5. [[CrossRef](#)] [[PubMed](#)]



11. Rizzo, F.; Rinaldi, A.; Marchese, G.; Coviello, E.; Sellitto, A.; Cordella, A.; Giurato, G.; Nassa, G.; Ravo, M.; Tarallo, R.; et al. Specific patterns of PIWI-interacting small noncoding RNA expression in dysplastic liver nodules and hepatocellular carcinoma. *Oncotarget* **2016**, *7*, 54650–54661. [[CrossRef](#)] [[PubMed](#)]
12. Oh, S.-J.; Kim, S.-M.; Kim, Y.-O.; Chang, H.-K. Clinicopathologic Implications of PIWIL2 Expression in Colorectal Cancer. *Korean J. Pathol.* **2012**, *46*, 318–323. [[CrossRef](#)] [[PubMed](#)]
13. Liu, W.; Gao, Q.; Chen, K.; Xue, X.; Li, M.; Chen, Q.; Zhu, G.; Gao, Y. Hiwi facilitates chemoresistance as a cancer stem cell marker in cervical cancer. *Oncol. Rep.* **2014**, *32*, 1853–1860. [[CrossRef](#)] [[PubMed](#)]
14. Vyčtytilova-Faltejskova, P.; Stitkovcova, K.; Radova, L.; Sachlova, M.; Kosarova, Z.; Slaba, K.; Kala, Z.; Svoboda, M.; Kiss, I.; Vyzula, R.; et al. Circulating PIWI-Interacting RNAs piR-5937 and piR-28876 Are Promising Diagnostic Biomarkers of Colon Cancer. *Cancer Epidemiol. Biomark. Prev.* **2018**, *27*, 1019–1028. [[CrossRef](#)]
15. Weng, M.; Wu, D.; Yang, C.; Peng, H.; Wang, G.; Wang, T.; Li, X. Noncoding RNAs in the development, diagnosis, and prognosis of colorectal cancer. *Transl. Res.* **2017**, *181*, 108–120. [[CrossRef](#)]
16. Koduru, S.V.; Tiwari, A.K.; Hazard, S.W.; Mahajan, M.; Ravnic, D.J. Exploration of small RNA-seq data for small non-coding RNAs in Human Colorectal Cancer. *Int. J. Genom.* **2017**, *5*, 16–31. [[CrossRef](#)]
17. Sun, R.; Gao, C.-L.; Li, D.-H.; Li, B.-J.; Ding, Y.-H. Expression Status of PIWIL1 as a Prognostic Marker of Colorectal Cancer. *Dis. Markers* **2017**, *2017*, 1–7. [[CrossRef](#)]
18. Raeisossadati, R.; Abbaszadegan, M.R.; Moghbeli, M.; Tavassoli, A.; Kihara, A.H.; Forghanifard, M.M. Aberrant expression of DPPA2 and HIWI genes in colorectal cancer and their impacts on poor prognosis. *Tumour Biol.* **2014**, *35*, 5299–5305. [[CrossRef](#)]
19. Wang, H.L.; Chen, B.B.; Cao, X.G.; Wang, J.; Hu, X.F.; Mu, X.Q.; Chen, X.B. The clinical significances of the abnormal expressions of Piwil1 and Piwil2 in colonic adenoma and adenocarcinoma. *Onco. Targets Ther.* **2015**, *8*, 1259–1264. [[CrossRef](#)]
20. Zeng, Y.; Qu, L.-K.; Meng, L.; Liu, C.-Y.; Dong, B.; Xing, X.-F.; Wu, J.; Shou, C.-C. HIWI expression profile in cancer cells and its prognostic value for patients with colorectal cancer. *Chin. Med. J.* **2011**, *124*, 2144–2149.
21. Litwin, M.; Dubis, J.; Arczynska, K.; Piotrowska, A.; Frydlewicz, A.; Karczewski, M.; Dziegiel, P.; Witkiewicz, W. Correlation of HIWI and HILI Expression with Cancer Stem Cell Markers in Colorectal Cancer. *Anticancer Res.* **2015**, *35*, 3317–3324.
22. Hutter, C.; Zenklusen, J.C. The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. *Cell* **2018**, *173*, 283–285. [[CrossRef](#)]
23. Tang, Z.; Li, C.; Kang, B.; Gao, G.; Li, C.; Zhang, Z. GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **2017**, *45*, W98–W102. [[CrossRef](#)] [[PubMed](#)]
24. Barault, L.; Amatu, A.; Siravegna, G.; Ponzetti, A.; Moran, S.; Cassingena, A.; Mussolin, B.; Falcomata, C.; Binder, A.M.; Cristiano, C.; et al. Discovery of methylated circulating DNA biomarkers for comprehensive non-invasive monitoring of treatment response in metastatic colorectal cancer. *Gut* **2018**, *67*, 1995–2005. [[CrossRef](#)] [[PubMed](#)]
25. Shaiken, T.E.; Opekun, A.R. Dissecting the cell to nucleus, perinucleus and cytosol. *Sci. Rep.* **2014**, *4*, 4923. [[CrossRef](#)]
26. Locati, M.D.; Terpstra, I.; De Leeuw, W.C.; Kuzak, M.; Rauwerda, H.; Ensink, W.A.; Van Leeuwen, S.; Nehrlich, U.; Spaink, H.P.; Jonker, M.J.; et al. Improving small RNA-seq by using a synthetic spike-in set for size-range quality control together with a set for data normalization. *Nucleic Acids Res.* **2015**, *43*, e89. [[CrossRef](#)]
27. Keene, J.D.; Komisarow, J.M.; Friedersdorf, M.B. RIP-Chip: The isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat. Protoc.* **2006**, *1*, 302–307. [[CrossRef](#)] [[PubMed](#)]
28. Shi, J.; Ko, E.-A.; Sanders, K.M.; Chen, Q.; Zhou, T. SPORTS1.0: A Tool for Annotating and Profiling Non-coding RNAs Optimized for rRNA- and tRNA-derived Small RNAs. *Genom. Proteom. Bioinform.* **2018**, *16*, 144–151. [[CrossRef](#)] [[PubMed](#)]
29. Panero, R.; Rinaldi, A.; Memoli, D.; Nassa, G.; Ravo, M.; Rizzo, F.; Tarallo, R.; Milanese, L.; Weisz, A.; Giurato, G. iSmaRT: A toolkit for a comprehensive analysis of small RNA-Seq data. *Bioinformatics* **2016**, *33*, 938–940.

30. Yagi, K.; Akagi, K.; Hayashi, H.; Nagae, G.; Tsuji, S.; Isagawa, T.; Midorikawa, Y.; Nishimura, Y.; Sakamoto, H.; Seto, Y.; et al. Three DNA methylation epigenotypes in human colorectal cancer. *Clin. Cancer Res.* **2010**, *16*, 21–33. [[CrossRef](#)]
31. Navarro, A.; Tejero, R.; Viñolas, N.; Cordeiro, A.; Marrades, R.M.; Fuster, D.; Caritg, O.; Moises, J.; Muñoz, C.; Molins, L.; et al. The significance of PIWI family expression in human lung embryogenesis and non-small cell lung cancer. *Oncotarget* **2015**, *6*, 31544–31556. [[CrossRef](#)] [[PubMed](#)]
32. Ferreira, H.J.; Heyn, H.; Garcia del Muro, X.; Vidal, A.; Larriba, S.; Munoz, C.; Villanueva, A.; Esteller, M. Epigenetic loss of the PIWI/piRNA machinery in human testicular tumorigenesis. *Epigenetics* **2014**, *9*, 113–118. [[CrossRef](#)] [[PubMed](#)]
33. Wang, C.; Gu, Y.; Zhang, K.; Xie, K.; Zhu, M.; Dai, N.; Jiang, Y.; Guo, X.; Liu, M.; Dai, J.; et al. Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. *Nat. Commun.* **2016**, *7*, 10499. [[CrossRef](#)] [[PubMed](#)]
34. Xie, K.; Zhang, K.; Kong, J.; Wang, C.; Gu, Y.; Liang, C.; Jiang, T.; Qin, N.; Liu, J.; Guo, X.; et al. Cancer-testis gene PIWIL1 promotes cell proliferation, migration, and invasion in lung adenocarcinoma. *Cancer Med.* **2018**, *7*, 157–166. [[CrossRef](#)] [[PubMed](#)]
35. Li, H.; Chiappinelli, K.B.; Guzzetta, A.A.; Easwaran, H.; Yen, R.-W.C.; Vatapalli, R.; Topper, M.J.; Luo, J.; Connolly, R.M.; Azad, N.S.; et al. Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacitidine in common human epithelial cancers. *Oncotarget* **2014**, *5*, 587–598. [[CrossRef](#)] [[PubMed](#)]
36. Meyers, R.M.; Bryan, J.G.; McFarland, J.M.; Weir, B.A.; Sizemore, A.E.; Xu, H.; Dharia, N.V.; Montgomery, P.G.; Cowley, G.S.; Pantel, S.; et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **2017**, *49*, 1779–1784. [[CrossRef](#)]
37. Lappalainen, I.; Almeida-King, J.; Kumanduri, V.; Senf, A.; Spalding, J.D.; Ur-Rehman, S.; Saunders, G.; Kandasamy, J.; Caccamo, M.; Leinonen, R.; et al. The European Genome-phenome Archive of human data consented for biomedical research. *Nat. Genet.* **2015**, *47*, 692–695. [[CrossRef](#)]
38. Zhang, F.; Wang, J.; Xu, J.; Zhang, Z.; Koppetsch, B.S.; Schultz, N.; Vreven, T.; Meignin, C.; Davis, I.; Zamore, P.D.; et al. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. *Cell* **2012**, *151*, 871–884. [[CrossRef](#)]
39. Genzor, P.; Cordts, S.C.; Bokil, N.V.; Haase, A.D. Aberrant expression of select piRNA-pathway genes does not reactivate piRNA silencing in cancer cells. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 11111–11112. [[CrossRef](#)]
40. Vagin, V.V.; Wohlschlegel, J.; Qu, J.; Jónsson, Z.; Huang, X.; Chuma, S.; Girard, A.; Sachidanandam, R.; Hannon, G.J.; Aravin, A.A. Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genome Res.* **2009**, *23*, 1749–1762. [[CrossRef](#)]
41. Ebhardt, H.A.; Thi, E.P.; Wang, M.-B.; Unrau, P.J. Extensive 3' modification of plant small RNAs is modulated by helper component-proteinase expression. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13398–13403. [[CrossRef](#)] [[PubMed](#)]
42. Dimitrova, D.G.; Teyssset, L.; Carré, C. RNA 2'-O-Methylation (Nm) Modification in Human Diseases. *Genes* **2019**, *10*, 117. [[CrossRef](#)] [[PubMed](#)]
43. De Fazio, S.; Bartonicek, N.; Di Giacomo, M.; Abreu-Goodger, C.; Sankar, A.; Funaya, C.; Antony, C.; Moreira, P.N.; Enright, A.J.; O'Carroll, D. The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature* **2011**, *480*, 259–263. [[CrossRef](#)] [[PubMed](#)]
44. Gomes Fernandes, M.; He, N.; Wang, F.; Van Iperen, L.; Eguizabal, C.; Matorras, R.; Roelen, B.A.J.; Chuva De Sousa Lopes, S.M. Human-specific subcellular compartmentalization of P-element induced wimpy testis-like (PIWIL) granules during germ cell development and spermatogenesis. *Hum. Reprod.* **2018**, *33*, 258–269. [[CrossRef](#)] [[PubMed](#)]
45. Zhong, F.; Zhou, N.; Wu, K.; Guo, Y.; Tan, W.; Zhang, H.; Zhang, X.; Geng, G.; Pan, T.; Luo, H.; et al. A SnoRNA-derived piRNA interacts with human interleukin-4 pre-mRNA and induces its decay in nuclear exosomes. *Nucleic Acids Res.* **2015**, *43*, 10474–10491. [[CrossRef](#)] [[PubMed](#)]
46. He, X.; Chen, X.; Zhang, X.; Duan, X.; Pan, T.; Hu, Q.; Zhang, Y.; Zhong, F.; Liu, J.; Zhang, H.; et al. An Lnc RNA (GAS5)/SnoRNA-derived piRNA induces activation of TRAIL gene by site-specifically recruiting MLL/COMPASS-like complexes. *Nucleic Acids Res.* **2015**, *43*, 3712–3725. [[CrossRef](#)] [[PubMed](#)]
47. Franz, M.; Rodríguez, H.; Lopes, C.; Zuberi, K.; Montojo, J.; Bader, G.D.; Morris, Q. GeneMANIA update 2018. *Nucleic Acids Res.* **2018**, *46*, W60–W64. [[CrossRef](#)]

48. Weng, W.; Li, H.; Goel, A. Piwi-interacting RNAs (piRNAs) and cancer: Emerging biological concepts and potential clinical implications. *Biochim. et Biophys. Acta (BBA) Bioenerg.* **2019**, *1871*, 160–169. [[CrossRef](#)]
49. Liu, Y.; Dou, M.; Song, X.; Dong, Y.; Liu, S.; Liu, H.; Tao, J.; Li, W.; Yin, X.; Xu, W. The emerging role of the piRNA/piwi complex in cancer. *Mol. Cancer* **2019**, *18*, 123. [[CrossRef](#)]
50. Kabayama, Y.; Toh, H.; Katanaya, A.; Sakurai, T.; Chuma, S.; Kuramochi-Miyagawa, S.; Saga, Y.; Nakano, T.; Sasaki, H. Roles of MIWI, MILI and PLD6 in small RNA regulation in mouse growing oocytes. *Nucleic Acids Res.* **2017**, *45*, 5387–5398. [[CrossRef](#)]
51. Viljetic, B.; Diao, L.; Liu, J.; Krsnik, Z.; Wijeratne, S.H.R.; Kristopovich, R.; Dutre-Clarke, M.; Kraushar, M.L.; Song, J.; Xing, J.; et al. Multiple roles of PIWIL1 in mouse neocortico-genesis. *bioRxiv* **2017**, 106070. [[CrossRef](#)]
52. Rizzo, F.; Hashim, A.; Marchese, G.; Ravo, M.; Tarallo, R.; Nassa, G.; Giurato, G.; Rinaldi, A.; Cordella, A.; Persico, M.; et al. Timed regulation of P-element-induced wimpy testis-interacting RNA expression during rat liver regeneration. *Hepatology* **2014**, *60*, 798–806. [[CrossRef](#)] [[PubMed](#)]
53. Zamore, P.D. Somatic piRNA biogenesis. *EMBO J.* **2010**, *29*, 3219–3221. [[CrossRef](#)]
54. Théron, E.; Dennis, C.; Brassat, E.; Vaury, C. Distinct features of the piRNA pathway in somatic and germ cells: From piRNA cluster transcription to piRNA processing and amplification. *Mob. DNA* **2014**, *5*, 28. [[CrossRef](#)] [[PubMed](#)]
55. Klattenhoff, C.; Theurkauf, W. Biogenesis and germline functions of piRNAs. *Development* **2008**, *135*, 3–9. [[CrossRef](#)]
56. Klattenhoff, C.; Xi, H.; Li, C.; Lee, S.; Xu, J.; Khurana, J.S.; Zhang, F.; Schultz, N.; Koppetsch, B.S.; Nowosielska, A.; et al. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. *Cell* **2009**, *138*, 1137–1149. [[CrossRef](#)]
57. Yin, J.; Qi, W.; Ji, C.; Zhang, D.; Xie, X.; Ding, Q.; Jiang, X.; Han, J.; Jiang, H. Small RNA sequencing revealed aberrant piRNA expression profiles in colorectal cancer. *Oncol. Rep.* **2019**, *42*, 263–272. [[CrossRef](#)]
58. Gou, L.-T.; Dai, P.; Yang, J.-H.; Xue, Y.; Hu, Y.-P.; Zhou, Y.; Kang, J.-Y.; Wang, X.; Li, H.; Hua, M.-M.; et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res.* **2014**, *24*, 680–700. [[CrossRef](#)]
59. Vourekas, A.; Zheng, Q.; Alexiou, P.; Maragkakis, M.; Kirino, Y.; Gregory, B.D.; Mourelatos, Z. Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nat. Struct. Mol. Boil.* **2012**, *19*, 773–781. [[CrossRef](#)]
60. Popovici, V.; Budinska, E.; Tejpar, S.; Weinrich, S.; Estrella, H.; Hodgson, G.; Van Cutsem, E.; Xie, T.; Bosman, F.T.; Roth, A.D.; et al. Identification of a poor-prognosis BRAF-mutant-like population of patients with colon cancer. *J. Clin. Oncol.* **2012**, *30*, 1288–1295. [[CrossRef](#)]
61. Herr, R.; Kohler, M.; Andrlova, H.; Weinberg, F.; Moller, Y.; Halbach, S.; Lutz, L.; Mastroianni, J.; Klose, M.; Bittermann, N.; et al. B-Raf inhibitors induce epithelial differentiation in BRAF-mutant colorectal cancer cells. *Cancer Res.* **2015**, *75*, 216–229. [[CrossRef](#)] [[PubMed](#)]
62. Yang, L.; Bi, L.; Liu, Q.; Zhao, M.; Cao, B.; Li, D.; Xiu, J. Hiwi Promotes the Proliferation of Colorectal Cancer Cells via Upregulating Global DNA Methylation. *Dis. Markers* **2015**, *2015*, 383056. [[CrossRef](#)] [[PubMed](#)]
63. Siddiqi, S.; Terry, M.; Matushansky, I. Hiwi mediated tumorigenesis is associated with DNA hypermethylation. *PLoS ONE* **2012**, *7*, e33711. [[CrossRef](#)] [[PubMed](#)]
64. Piskol, R.; Huw, L.Y.; Sergin, I.; Kljin, C.; Modrusan, Z.; Kim, D.; Kljavin, N.M.; Tam, R.; Patel, R.; Burton, J.; et al. A Clinically Applicable Gene-Expression Classifier Reveals Intrinsic and Extrinsic Contributions to Consensus Molecular Subtypes in Primary and Metastatic Colon Cancer. *Clin. Cancer Res.* **2019**, *25*, 4431–4442. [[CrossRef](#)] [[PubMed](#)]

