Carcinogenesis, 2015, Vol. 36, No. 10, 1094–1102

doi:10.1093/carcin/bgv105
Advance Access publication July 25, 2015
Original Manuscript

ORIGINAL MANUSCRIPT

PIWI-interacting RNA 021285 is involved in breast tumorigenesis possibly by remodeling the cancer epigenome

Alan Fu[1,3], Daniel I.Jacobs[1], Aaron E.Hoffman[2], Tongzhang Zheng[1] and Yong Zhu[1,*]

[1]Department of Environmental Health Sciences, Yale University School of Public Health, New Haven, CT 06520, USA and [2]Department of Epidemiology, Tulane School of Public Health and Tropical Medicine and Tulane Cancer Center, New Orleans, LA 70112, USA

[3]Present Address: Department of Epidemiology, UCLA Fielding School of Public Health, Los Angeles, CA 90095, USA

*To whom correspondence should be addressed. Tel: +1 203 785 4844; Fax: +1 203 737 6023; Email: yong.zhu@yale.edu

Abstract

Although PIWI-interacting RNAs (piRNAs) account for the largest class of the small non-coding RNA superfamily, virtually nothing is known of their function in human carcinogenesis. Once thought to be expressed solely in the germ line where they safeguard the genome against transposon-induced insertional mutations, piRNAs are now believed to play an active role in somatic gene regulation through sequence-specific histone modification and DNA methylation. In the current study, we investigate the role of piRNA-021285 (piR-021285) in the regulation of the breast cancer methylome. Genotypic screening of a panel of single-nucleotide polymorphism (SNP)-containing piRNAs revealed a significant association between SNP rs1326306 G>T in piR-021285 and increased likelihood for breast cancer in a Connecticut-based population (441 cases and 479 controls). Given nascent but compelling evidence of piRNA-mediated gene-specific methylation in the soma, a genome-wide methylation screen was then carried out using wild type (WT) and variant piR-021285 mimic-transfected MCF7 cells to explore whether the observed association could be attributed in part to piR-021285-induced methylation at cancer-relevant genes. We found significant methylation differences at a number of experimentally implicated breast cancer-related genes, including attenuated 5′ untranslated region (UTR)/first exon methylation at the proinvasive *ARHGAP11A* gene in variant mimic-transfected cells. Follow-up functional analyses revealed both concurrent increased *ARHGAP11A* mRNA expression and enhanced invasiveness in variant versus WT piR-021285 mimic-transfected breast cancer cell lines. Taken together, our findings demonstrate the first evidence supporting a role of piRNAs, a novel group of non-coding RNA, in human tumorigenesis via a piRNA-mediated epigenetic mechanism, which warrants further confirmation and investigation.

Introduction

The 2006 discovery of PIWI-interacting RNAs (piRNAs) unveiled the newest member of the small non-coding RNA superfamily (1–4). Roughly 26–31 nucleotides in length and purportedly existing in over 20 000 unique species in the human genome, piRNAs are functionally distinct from their microRNA (miRNA) cousins and play a critical role in safeguarding the germ line against transposon-induced insertional mutations (1,2,5–9). Compelling new evidence suggests that piRNAs may also have essential functions in the soma, where they have been shown to induce histone modification and DNA methylation in a sequence-specific manner (10–12). This observation highlights the intriguing possibility of piRNA-mediated epigenetic control of cancer-related processes in human somatic cancers. While miRNAs have been implicated in almost every cancer type, virtually nothing is known about the far more abundant piRNAs in the carcinogenic process.

Received: January 27, 2015; **Revised:** July 1, 2015; **Accepted:** July 17, 2015

Abbreviations

cDNA	complementary DNA
DMEM	Dulbecco's Modified Eagle's Medium
FDR	false discovery rate
FBS	fetal bovine serum
miRNA	microRNA
piRNA	PIWI-interacting RNA
SNP	single-nucleotide polymorphism
UTR	untranslated region
WT	wild type
YNHH	Yale-New Haven Hospital

Although the advent of piRNA-dependent epigenetic regulation holds broad implications for the study of cancer, current research of the PIWI/piRNA axis in human cancers is largely restricted to expression profiling of PIWI proteins and a narrow selection of individual piRNAs. Aberrant expression of PIWI family proteins, which are required for piRNA biogenesis and function, has been observed in a number of malignancies, including cancers of the brain (13), esophagus (14), pancreas (15), liver (16), gastrointestinal tract (17–19) and breast (20). More recent studies have also found differential expression of individual piRNAs, including piR-823 in gastric cancer (21), piR-Hep1 in liver cancer (22), piR-651 in six different cancer types (23) and seven distinct piRNAs in breast cancer (23,24). Beyond expression profiling, only two studies have ventured to investigate piRNA function in a cancer model, finding a potential growth modulating effect for piR-651 (23) and piR-823 (21) in gastric cancer cells. However, no study, to our knowledge, has attempted to establish an empirical connection between a human cancer and piRNA-associated epigenetic phenomena.

Given the emergent, yet compelling evidence of piRNA-mediated gene-specific methylation in the soma, it is feasible that certain piRNAs may be responsible for the epigenetic regulation of cancer-relevant pathways. Polymorphisms found within the functional mature piRNA may thus be able to alter its capacity to methylate target genes, leading, in turn, to the dysregulation of cancer-related processes. The current study aimed at testing the possible role of piRNAs in human breast cancer in a molecular epidemiologic study with functional analyses.

Materials and methods

Breast cancer study population

Study subjects consisted of participants previously enrolled in a Connecticut breast cancer case–control study, which was approved by the Connecticut Department of Public Health, the National Cancer Institute, and the Institutional Review Boards at Yale University. Details of the study population, including participant characteristics and recruitment details, have been described previously (25). Briefly, breast cancer patients were identified from hospital records in Tolland County and Yale-New Haven Hospital (YNHH) in New Haven County, Connecticut, by the Rapid Case Ascertainment Shared Resource at the Yale Cancer Center. Cases were incident and histologically confirmed (International Classification of Diseases for Oncology, 174.0–174.9) and carcinomas were staged according to the tumor node metastasis (TNM) tumor staging system. Patients and were alive at the time of the interview, were between the ages of 30 and 80, and had no previous history of cancer apart from non-melanoma skin cancer. Controls at Tolland County younger than 65 were identified through random digit dialing while those over 65 were identified through Health Care Finance Administration files. YNHH controls were identified through computerized files as patients who underwent breast-related surgery at YNHH and had histologically confirmed benign breast disease. Permission to contact subjects was obtained from the hospital, as well as the personal physician, for all cases. Potential participants were then contacted first by letter, and then by telephone, if necessary. Following an initial interview, subjects were asked to complete a standardized questionnaire and blood samples were collected into sodium-heparinized tubes for immediate DNA isolation and subsequent analyses. Informed consent was collected from all participants. Participation rates were 77% for YNHH cases, 71% for YNHH controls, 74% for Tolland County cases and 61% for Tolland County controls. A total of 441 cases and 479 controls had DNA samples available for the current study. Supplementary Table 1, available at *Carcinogenesis* Online presents the distribution of selected baseline characteristics for all participants.

piRNA genotyping

A panel of 30 single-nucleotide polymorphism (SNP)-containing human piRNA sequences was identified by aligning piRNABank (http://pirnabank.ibab.ac.in/) piRNA genomic coordinate data with SNP coordinate data compiled in the HapMap database (HapMap genome browser release 22). All SNPs possessed a minor allele frequency ≥ 0.10. Genotyping was performed at Yale University's W.M. Keck Foundation Biotechnology Research Laboratory using the Sequenom MassARRAY multiplex genotyping platform (Sequenom, San Diego, CA) according to the manufacturer's protocol. Duplicate samples from 100 study subjects and 40 replicate samples from each of two blood donors were interspersed throughout each batch, and the concordance rates for QC samples was >99%. All genotyping scores, including quality control data, were rechecked by different laboratory personnel and the assay accuracy was confirmed. Of the genotyped SNPs, rs1326306 in hsa_piR_021285 (piR-021285) yielded the greatest statistically significant association with breast cancer and was selected for follow-up functional analyses.

Statistical analysis of genetic association

Case–control analyses were performed using the SAS statistical software, version 9.2 (SAS Institute, Cary, NC). The allelic distribution of the SNP was tested in control subjects by goodness-of-fit Chi-square for compliance with Hardy–Weinberg equilibrium, and no departure from equilibrium was detected. Odds ratios and 95% confidence intervals were determined by unconditional multivariate logistic regression, adjusted for the following covariates: age (continuous), race, family history of breast cancer in a first-degree relative, menopausal status and study site. Associations were interpreted with respect to a Bonferroni-corrected statistical significance threshold in order to account for the multiple comparisons inherent in the genetic association analysis.

Cell culture

MCF-7 and MDA-MB-231 breast cancer cell lines were purchased directly from the American Type Culture Collection (ATCC, Manassas, VA) in November 2012 and December 2013, respectively. All ATCC cell lines are tested for contaminants and authenticated prior to shipment using several characterization approaches including morphology assessment, karyotyping, cytochrome C oxidase variant analysis and short tandem repeat DNA fingerprinting. Cells were not reauthenticated as they were passaged in our laboratory for fewer than 6 months after resuscitation. Cells were maintained as monolayer cultures in 25 cm^2 polystyrene flasks in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) (Life Technologies, Carlsbad, CA). Cells were incubated at 37°C in a humid atmosphere containing 5% CO_2 and subcultured every 2–3 days.

Confirmation of endogenous piR-021285 expression in the nucleus

Nuclear RNA was extracted from MCF-7 cells using the SurePrep Nuclear or Cytoplasmic RNA Purification Kit (Fisher Scientific, Pittsburgh, PA). Nuclear RNA was then poly-A tagged and converted to complementary DNA (cDNA) using the NCode miRNA first-strand cDNA Synthesis Kit (Life Technologies). Relative expression of piR-021285 was quantitated using the $2^{-\Delta\Delta Ct}$ SYBR Green qPCR method with normalization to miR-16. piR-021285- and miR-16-specific forward primers (piR-021285: TAAGAATAAAAGCTGTTGAATGTTGCCTGT, miR-16: TAGCAGCACGTAA ATATTGGCG) were purchased from Integrated DNA Technologies (San

Jose, CA) and paired with a universal reverse primer (Life Technologies) in each reaction. All qPCR reactions were performed in technical triplicate using the SYBR Fast Universal qPCR Kit (Kapa Biosystems, Woburn, MA) on a 7500 Fast Real-Time PCR System (Life Technologies, Carlsbad, CA) with the following conditions: Denaturation at 95°C (3 min) followed by 40 cycles of denaturation at 95°C (3 s) and annealing/elongation at 60°C (30 s). Results of the confirmation can be found in Supplementary Figure 1, available at *Carcinogenesis* Online.

Genome-wide methylation array

MCF-7 cells were transfected in biological duplicate with either a single-stranded piR-021285 mimic [wild type (WT): UAAGAAUAAAAGCUGUU GAAUGUUGCCUGU, variant: UAAGAAUAAAAUCUGUUGAAUGUUGCCUGU] (Integrated DNA Technologies) or a single-stranded scrambled RNA control (QIAGEN, Valencia, CA), followed by genomic DNA isolation and genome-wide methylation measurements. Briefly, approximately 25 000 cells suspended in 10% FBS-supplemented DMEM were added to a pre-incubated mixture of LipofectAMINE RNAiMAX (1 µl) (Life Technologies) and RNA oligo in each well of a 12-well plate to a final oligo concentration of 25 nM. Cells were incubated at 37°C and 5% CO_2 for 72 h and then subjected to a second, forward, transfection at a final oligo concentration of 25 nM following aspiration of old media. Cells were then incubated for an additional 72 h prior to DNA isolation using the DNeasy Blood & Tissue Kit (QIAGEN). About 750 µg of DNA from each biological replicate of WT and variant piR-021285- and scrambled RNA control-transfected cells were submitted to Yale University's W.M. Keck Foundation Biotechnology Research Laboratory for genome-wide DNA methylation measurements using Illumina's (San Diego, CA) Infinium HumanMethylation450 array platform. The array is designed to accommodate methylation measurements at 485 477 CpG sites distributed throughout the genome (26). 31% (150 254 sites) of the array's CpG content are in CpG islands, 23% (112 072 sites) are in CpG shores, 10% (47 161) are in CpG shelves and 36% (176 277 sites) are defined as 'Open Sea'. 29% (140 004 sites) of CpG content are located in proximal gene promoters, defined as CpG sites within 200 bp (13% or 62 625 sites) or between 201 and 1500 bp (16% or 77 379 sites) upstream of the transcriptional start site, 10% (49 525 sites) are in the 5′ UTR, 2% (10 810 sites) are in the first exon, 31% (150 212 sites) are in the gene body, 3% (15 383 sites) are in the 3′ UTR and 25% (119 830 sites) correspond to intergenic regions. The Illumina Custom Model, as implemented in the Illumina GenomeStudio software, was used to assess the statistical significance of methylation differences between piR-021285-transfected and control DNA samples at each CpG site. This model generates *P* values for differential methylation of each CpG site based on the average methylation indices β (a value ranging from 0 to 1, where 0 represents a completely unmethylated site and 1 a completely methylated site) and their variance across replicate probes. To correct for the increased likelihood of false positive differences resulting from multiple testing, the Benjamini–Hochberg method was used to control for the false discovery rate (FDR) (27), defined as the expected ratio of erroneous rejections of the null hypothesis to the total number of rejected hypotheses. Methylation differences at CpG sites were considered statistically significant at a FDR-corrected *P* value (*Q* value) <0.05. As an additional quality control measure, CpG sites with intensity values of less than 500 for both piR-021285-transfected and control samples were filtered out to minimize the contribution from background signals.

Confirmation of piR-021285 delivery into nucleus

Nuclear RNA was extracted from piR-021285- and scrambled RNA control-transfected MCF-7 cells at 48, 72 and 96 h post-transfection and subsequently converted to cDNA, as described above. Relative expression of piR-021285 was quantitated using the $2^{-\Delta\Delta Ct}$ SYBR Green qPCR method with normalization to miR-16, as described previously, at each time point post-transfection. Results of the confirmation can be found in Supplementary Figure 2, available at *Carcinogenesis* Online.

Gene ontology

Differentially methylated CpG sites were paired to known genes according to their location within the following genomic regions: proximal promoter (up to 1500 bp upstream of transcriptional start site), 5′ UTR, first exon, gene body (exon or intron, excluding the first exon) and 3′ UTR. Breast cancer-related genes harboring differentially methylated CpG sites were assigned to functional domains (apoptosis, cellular motility, cell proliferation and DNA damage/repair) according to experimentally verified findings in peer-reviewed published works. A search of the Google Scholar (Google, Mountain View, CA) online publication database was conducted using the gene symbol and 'breast cancer' as keywords.

RNA extraction and qPCR expression change confirmation

Total RNA from WT and variant piR-021285 mimic- and scrambled RNA negative control-transfected MCF7 cells was isolated concurrently with genomic DNA using the miRNeasy Mini Kit (QIAGEN) per the instructions of the manufacturer. Isolated mRNA was subsequently converted to cDNA using the AffinityScript Multi-Temp cDNA Synthesis Kit (Agilent, Santa Clara, CA) per the instructions of the manufacturer. cDNA was then amplified using custom-designed gene-specific primers for *AATF*, *ARHGAP11A*, *CDC7*, *DCL1*, *PIP4K2B* and *THAP10* and the housekeeping gene *GAPDH*. All qPCR reactions were performed in technical triplicate using the SYBR Fast Universal qPCR Kit (Kapa Biosystems, Woburn, MA) on a 7500 Fast Real-Time PCR System (Life Technologies, Carlsbad, CA) with the following conditions: Denaturation at 95°C (3 min) followed by 45 cycles of denaturation at 95°C (3 s), annealing at 60°C (30 s) and elongation at 72°C (30 s). Relative gene-specific cDNA abundance between transfections was assessed using the $2^{-\Delta\Delta Ct}$ method with normalization to *GAPDH*. Statistical differences in expression were assessed using the unpaired Student's *t*-test on ΔCt values. Gene-specific primer sequences can be found in Supplementary Table 2, available at *Carcinogenesis* Online.

Cell proliferation

Following transfection with either the WT or variant single-stranded piR-021285 mimic or a single-stranded scrambled RNA, MCF-7 and MDA-MB-231 cells were measured at fixed 24-h intervals to gauge treatment-specific differences in proliferation. Approximately 1500 MCF-7 and 2500 MDA-MB-231 cells suspended in 10% FBS-supplemented phenol red-free DMEM were reverse transfected using 0.1 µl of RNAiMAX at a final oligo concentration of 25 nM in each well of an optically clear 96-well plate. Daily cellular proliferation measurements were carried out using the CellTiter 96 AQueous One Solution Cell Proliferation Assay (MTS) kit (Promega, Madison, WI) starting on the day of the transfection and at 24, 48, 72, 96, 120 and 144-h post-transfection. About 20 µl of MTS dye were added to each well and allowed to incubate with transfected cells in a 37°C CO_2 incubator for 2 h. Following incubation, MTS-stained cells were quantified using an Epoch Microplate Spectrophotometer (Biotek, Winooski, VT). A repeated-measures, mixed-design analysis of variance model was used to assess whole curve differences in proliferation rate between transfection groups by measuring the interaction between transfection type and time. All statistical analyses were performed using the SAS statistical software, version 9.2.

Apoptosis

MCF-7 and MDA-MB-231 cells were reverse transfected in duplicate with either a piR-021285 mimic or a scrambled RNA control in T25 flasks in 5 ml of 10% FBS-supplemented DMEM media. RNAiMAX and RNA oligo concentrations used are identical to those described previously. About 72 h post-transfection, methyl methanesulfonate (Sigma–Aldrich, St Louis, MO) was added to one replicate of each transfection duplicate (0.015% by volume in media) for 1-h treatment period followed by 16 h of recovery in fresh media. Cells were then harvested and lysed by freeze-thaw in cell lysis buffer (Promega, Madison, WI). Soluble protein was isolated and incubated in Ac-DEVD-pNA colorimetric substrate and reaction buffers supplied by the CasPACE Assay System (Promega), as per the instructions of the manufacturer. Caspase-3 activity was quantified in technical triplicate using an Epoch Microplate Spectrophotometer. The unpaired Student's *t*-test was used to assess differences in 405 absorbance values between transfection groups.

Cell invasion

About 50 000 MCF-7 and MDA-MB-231 cells were reverse transfected with either a piR-021285 mimic or a scrambled RNA control in each well of a

24-well plate in 500 µl of 10% FBS-supplemented DMEM media. RNAiMAX and RNA oligo concentrations used are identical to those used to measure cell proliferation (see above). About 50 000 cells were replaced in matrigel-coated 8 micron porous inserts supplied by the BD BioCoat Matrigel Invasion Chamber kit (BD Biosciences, San Jose, CA) set atop a 24-well plate 48 h post-transfection. Cells within inserts were incubated in 500 µl of serum-free DMEM atop wells containing 750 µl of 20% Gibco® non-heat inactivated certified FBS-supplemented DMEM for 48 h (MDA-MB-231) or 72 h (MCF-7), at which point media and non-invading cells within inserts were removed. Invading cells were then fixed by incubating inserts in 100% methanol and stained using 1% toluidine blue stain in 1% sodium borate solution. The relative invasiveness of transfected cells was then quantified by counting the number of invading cells within four equally spaced visual fields under 40× magnification. The unpaired Student's *t*-test was used to assess statistical differences in cell count between transfection groups.

Cell migration

MCF-7 and MDA-MB-231 cells were reverse transfected with either a piR-021285 mimic or a scrambled RNA control in each well of a collagen-coated six-well plate (Life Technologies) in 2 ml of 10% FBS-supplemented DMEM media. RNAiMAX and RNA oligo concentrations used are identical to those described previously. Transfected cells were allowed to grow to confluence at 48 h post-transfection, at which point parallel scratches were introduced along the monolayer surface using a 200 µl pipette tip. 10% FBS-containing DMEM was replaced with fresh serum-free DMEM media after washing transfected cells with 1× phosphate-buffered saline. Cells were allowed to incubate in serum-free DMEM for an additional 24 h, at which point the relative motility of transfected cells was measured by quantifying the ratios of the pre- to post-incubation gap widths. Gap measurements were performed at three different fields per transfection. The unpaired Student's *t*-test was used to assess statistical differences in pre- to post-incubation gap width ratios between different transfection groups.

Results

piR-021285 genetic variant is associated with increased breast cancer risk

A previous *in silico* survey revealed a G>T germline SNP (rs1326306) that accounts for a G>U substitution at base 12 of piR-021285. Results of genotyping uncovered a significant association between the variant T allele (minor allele frequency = 0.22 in CEU Caucasian population) and increased risk for breast cancer in our predominantly Caucasian study population after adjustment for multiple comparisons (Bonferroni-adjusted P_{trend} = 0.024). Individuals harboring the G/T or T/T genotype were 1.57 times were more likely (95% CI: 1.20–2.06) to have breast cancer compared to those homozygous for the WT G allele after adjustment for age, race, family history of breast cancer in a first-degree relative, menopausal status and study site (Table 1). In addition, likelihood of disease was significantly correlated with allele zygosity (P_{trend} < 0.001); relative to homozygous WT individuals, heterozygous individuals were 1.51 times more likely to have breast cancer (95% CI: 1.14–2.00) while homozygous variant individuals were 2.01 times more likely (95% CI: 1.15–3.52).

Table 1. Genetic association of piR-021285 SNP rs1326306 with breast cancer[a]

Genotype	Cases *N* (%)	Controls *N* (%)	Crude OR (95% CI)	Adjusted OR (95% CI)[b]
G/G	205 (47.8)	274 (59.7)	Reference	Reference
G/T	187 (43.6)	161 (35.1)	1.55 (1.18, 2.05)	1.51 (1.14, 2.00)
T/T	37 (8.6)	24 (5.2)	2.06 (1.20, 3.55)	2.01 (1.15, 3.52)
G/T or TT	224 (52.2)	185 (40.3)	1.62 (1.24, 2.11)	1.57 (1.20, 2.06)
P_{trend}			<0.001	<0.001

[a]Subjects with incomplete genotype information were not included in the analysis.
[b]Adjusted for age (continuous), race, family history of breast cancer in a first-degree relative, menopause status and study site.

DNA methylation differences found at breast cancer-related genes between WT and variant piR-021285 mimic-transfected MCF7 cells

To explore whether piR-021285 can induce DNA methylation changes at breast cancer-related genes and whether these changes can be modulated by the presence of the variant allele, genome-wide methylation signatures were measured following transfection of MCF7 cells with either WT or variant piR-021285 mimics or a scrambled RNA negative control. In total, 485 477 CpG sites were analyzed across the genome, revealing FDR-qualified (Q < 0.05) methylation differences at 45 genes between WT and variant piR-021285 mimic-transfected cells (Figure 1A). Interestingly, no gene was associated with more than one differentially methylated CpG site. A brief description of each gene, including known breast cancer-relevant functions, if any, is given in Supplementary Table 3, available at *Carcinogenesis* Online.

About 19 of the 45 differentially methylated genes are linked to breast cancer via experimental evidence or genetic association. Six of these breast cancer-related genes have been implicated by direct experimental observation in one or more of the following 'classic' cancer-related processes: apoptosis, cell motility, cell proliferation and DNA damage and repair (Figure 1B). These genes include the putative breast cancer oncogene *AATF*, the expression of which has been demonstrated to enhance MCF7 breast cancer cell viability (28); the proapoptotic and antiproliferative *ARHGAP11A* gene, which has been found to be upregulated in migratory breast tumor cells (29,30); *CDC7*, the inhibition of which results in S phase retardation, accumulation of DNA damage and induction of apoptosis in certain cancer cell lines (31) and the overexpression of which has been linked with increased gene amplification in breast carcinomas (32); *DLC1*, which has been found to inhibit breast cancer cell growth and *in vivo* tumorigenicity in mice (33,34); *PIP4K2B*, the overexpression of which has been associated with proliferative advantage in breast cancer cells (35); and the DNA damage-responsive *THAP10* gene, which has been found to be frequently silenced by promoter hypermethylation in breast tumors (36).

An examination of the methylation profiles of these 'functionally relevant' genes under disparate transfection conditions reveals the influence of a possible variant-associated piR-021285 methylome shift (Figure 1C). While WT piR-021285 mimic-induced methylation is evident at four of six genes (*AATF*, *ARHGAP11A*, *PIP4K2B* and *THAP10*), induction is wholly absent at these genes in variant piR-021285 mimic-transfected cells. Conversely, there is also apparent evidence of variant-associated gain-of-function, as manifested in induced methylation at *CDC7* in variant piR-021285 mimic-transfected cells, which is absent in the WT transfection.

To check what effect these methylation differences might have on gene expression, relative mRNA levels of transfected MCF7 cells were measured by qPCR. Consistent with expectation, lower *ARHGAP11A* 5′ UTR/first exon methylation, higher *CDC7* proximal promoter methylation and lower *DLC1* gene body methylation were concurrent with higher *ARHGAP11A* (2.60-fold; *P* = 0.023) and lower *CDC7* (0.37-fold; *P* = 0.050 (borderline significant)) and *DLC1* (0.41-fold; *P* < 0.001) mRNA expression in variant

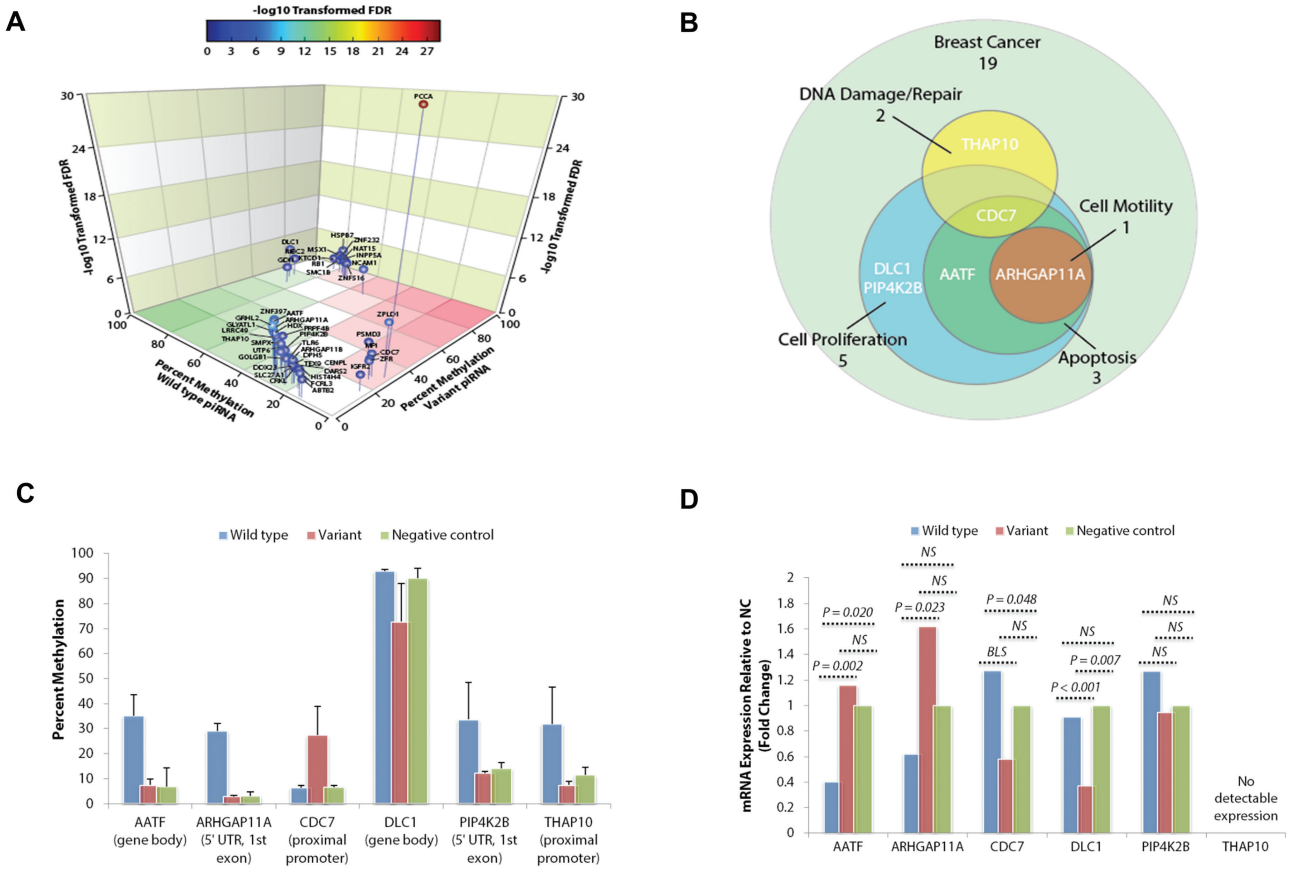

Figure 1. DNA methylation differences at breast cancer-related genes found between WT and variant piR-021285 mimic-transfected MCF7 cells. (**A**) A total of 45 genes exhibited FDR-qualified statistically significant methylation changes ($Q < 0.05$) at one or more associated CpG sites between WT and variant piR-021285 mimic-transfected MCF7 cells. (**B**) Previous literature implicate 19 of the 45 differentially methylated genes in breast cancer or breast cancer-related processes. Six of these genes (*AATF*, *ARHGAP11A*, *CDC7*, *DLC1*, *PIP4K2B* and *THAP10*) have been experimentally implicated in several 'classic' cancer-related cellular and molecular processes, namely, apoptosis, cell motility, cell proliferation and DNA damage and repair. (**C**) Methylation profiles of differentially methylated genes with functional relevance to apoptosis, cell motility, cell proliferation and/or DNA damage and repair in WT and variant piR-021285 mimic- and scrambled RNA negative control-transfected MCF7 cells. Methylation profiles under disparate transfection conditions reveal the influence of a possible variant-associated piR-021285 methylome shift. While WT piR-021285 mimic-induced methylation is evident at *AATF*, *ARHGAP11A*, *PIP4K2B* and *THAP10*, induction is absent at these genes in variant piR-021285 mimic-transfected cells. Conversely, variant-associated gain-of-function is evident in the form of induced methylation at *CDC7* in variant piR-021285 mimic-transfected cells, which is absent in the WT transfection. (**D**) Relative mRNA expression levels of differentially methylated genes with functional relevance in 'classic' cancer-related processes in piR-021285 mimic- and scrambled RNA-transfected MCF7 cells. Consistent with the regions and direction of methylation differences, we observed higher *ARHGAP11A* (2.60-fold; $P = 0.023$) and lower *CDC7* (0.37-fold; $P = 0.050$ (borderline significant)) and *DLC1* (0.41-fold; $P < 0.001$) expression in variant versus WT mimic-transfected cells. Higher *AATF* expression (2.85-fold; $P = 0.020$) was observed in the variant versus WT transfection despite lower methylation in the gene body. No expression differences were found for *PIP4K2B* and expression of *THAP10* could not be detected. Expression fold change values were normalized based on expression in the negative control.

versus WT mimic-transfected cells (Figure 1D). Interestingly, higher *AATF* expression (2.85-fold; $P = 0.020$) was observed in variant versus WT mimic-transfected cells despite lower gene body methylation. In addition, a significant expression difference in *ARHGAP11A* was not observed between WT mimic- and scrambled RNA-transfected cells despite increased 5′ UTR/first exon methylation in the WT transfection. Expression differences for *PIP4K2B* were not found between the variant and WT transfections and no *THAP10* expression was detected.

No detectable differences in proliferation and apoptotic incidence between WT and variant piR-021285 mimic-transfected MCF7 and MDA-MB-231 cells

Five genes with experimentally established roles in cell proliferation were found to be differentially methylated between WT and variant piR-021285 mimic-transfected MCF7 cells, including pro-proliferative *AATF* (decreased methylation at CpG site in gene body in variant transfection), *CDC7* (increased methylation in proximal promoter) and *PIP4K2B* (decreased methylation in 5′ UTR/first exon) and antiproliferative *ARHGAP11A* (decreased methylation in 5′ UTR/first exon) and *DLC1* (decreased methylation in gene body). These methylation differences were concurrent with increased mRNA expression of *AATF*, *ARHGAP11A* and *DLC1* in variant piR-021285-transfected cells. These observations suggest a capacity for piR-021285 to affect the proliferative rate of breast cancer cells via variant-associated methylation changes at these genes. To test this hypothesis, the proliferative rate of MCF7 and MDA-MB-231 cells were measured over a 144-h period following individual transfections with the WT or variant piR-021285 mimic or scrambled RNA control. Contrary to expectation, no differences in proliferation were detectable between any of the transfections in either cell line (Figure 2A).

As *AATF* (antiapoptotic), *ARHGAP11A* (proapoptotic) and *CDC7* (antiapoptotic) are also experimentally linked to apoptosis, we decided to probe for evidence of variant piR-021285-inducible changes in apoptotic potential in MCF7 and MDA-MB-231 cells.

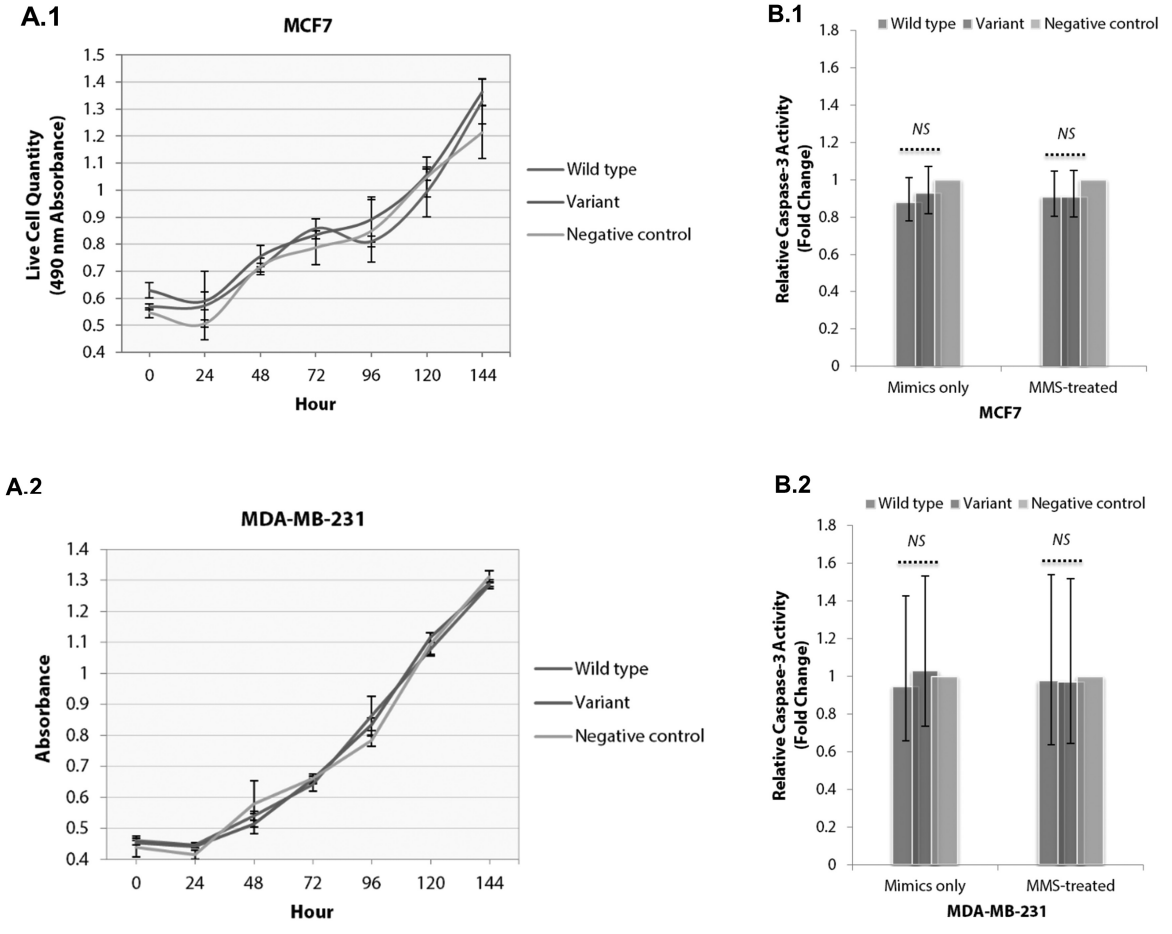

Figure 2. No evidence of differences in proliferation and apoptotic incidence between WT and variant piR-021285 mimic-transfected MCF7 and MDA-MB-231 cells. (**A**) Despite methylation and expression differences at proliferation-relevant genes between the WT, variant and scrambled RNA negative control transfections, no significant proliferation differences were detectable in MCF7 (A.1) and MDA-MB-231 cells (A.2) over a 144-h period. (**B**) As with proliferation, no differences in apoptotic incidence (as approximated by caspase-3 activity) were detected between the WT, variant or scrambled RNA transfections in either MCF7 (B.1) or MDA-MB-231 cells (B.2), both when piR-021285 mimics were administered as standalone treatments and concurrently with a mutagenic agent (methyl methanesulfonate). Fold change values were normalized based on caspase-3 activity in the negative control. Error bars represent two-sided 95% confidence intervals.

piR-021285-associated differences in apoptotic incidence were measured with and without methyl methanesulfonate treatment in order to gauge both the intrinsic effect of piR-021285 challenge and piR-021285 responsiveness to mutagen-induced DNA damage. As with our proliferation measurements, no significant differences were detected between the WT, variant, or scrambled RNA transfections in either cell line (Figure 2B), yielding no evidence that piR-021285 is able to independently induce apoptosis or to protect against or amplify apoptotic signals resulting from mutagenic DNA damage in breast cancer cells.

Increased invasiveness in variant piR-021285 mimic-transfected MCF7 and MDA-MB-231 cells

ARHGAP11A encodes a poorly characterized Rho GTPase-activating protein that enhances the invasiveness of colon cancer cells *in vivo* (37) and which is upregulated in xenograft-derived migratory MDA-MB-231 breast tumor cells (29). In variant piR-021285 mimic-transfected MCF7 cells, *ARHGAP11A* demonstrated both attenuated methylation at a CpG site within the 5′ UTR/first exon and 2.6-fold higher mRNA expression relative to WT mimic-transfected cells. The direction of these differences suggest the possibility that the variant piR-021285 may be able to confer an invasive edge to breast cancer cells where it exists as the predominant piR-021285 species. We tested this hypothesis by measuring the relative motility of WT and variant piR-021285 mimic-transfected MCF7 and MDA-MB-231 cells through a FBS gradient using a conventional matrigel-coated 8 μm pore invasion chamber. Consistent with the idea of a variant-associated enhancement in breast cancer invasiveness, variant piR-021285 mimic-transfected cells exhibited increased tendency to migrate from serum-free to serum-containing media compared to WT mimic-transfected cells (MCF7: $P = 0.004$, MDA-MB-231: $P = 0.022$; Figure 3A). To test whether this increased migratory capacity persisted in the absence of a serum-derived chemotactic gradient, wound closure rates were compared between variant and WT mimic-transfected MCF7 and MDA-MB-231 cells using a simple scratch assay. Despite the contrast in invasion rates, no differences in closure rates were detected (Figure 3B), suggesting that piR-021285-associated changes in motility likely arose from extracellular interactions with chemoattractants found in serum.

Discussion

In contrast to the voluminous body of experimental and observational evidence that underpin the miRNA-cancer connection, knowledge of piRNAs in the carcinogenic process remains in

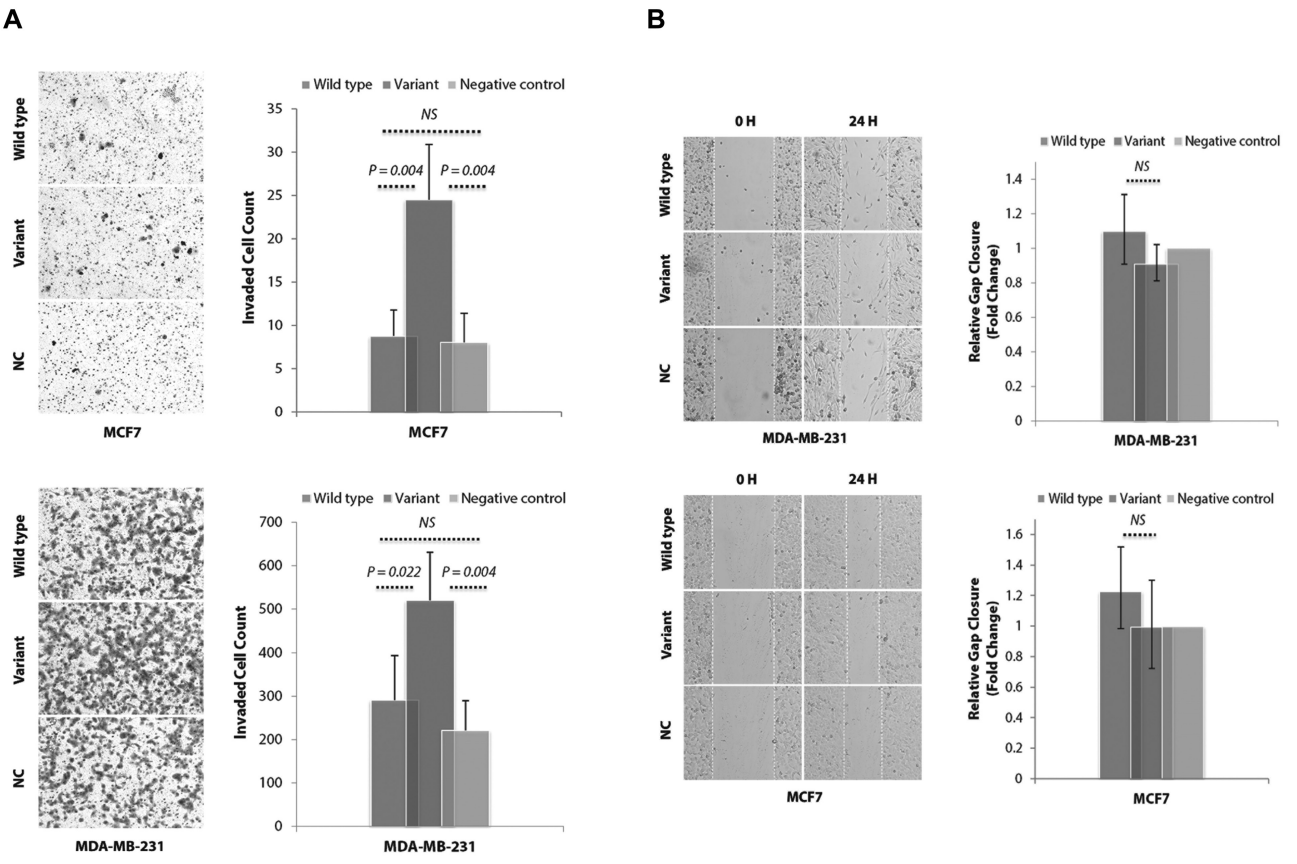

Figure 3. Increased invasiveness observed in variant piR-021285 mimic-transfected MCF7 and MDA-MB-231 cells in the presence of serum gradient. (**A**) piR-021285 mimic-transfected MCF7 and MDA-MB-231 cells were allowed to from serum-free to FBS-containing media over a course of 48 h (MDA-MB-231) or 72 h (MCF7), at which point the number of invaded cells were quantified. Consistent with variant piR-021285-associated attenuated *ARHGAP11A* 5′ UTR/first exon methylation and heightened mRNA expression, variant mimic-transfected MCF7 and MDA-MB-231 cells exhibited greater tendency to invade compared to WT mimic-transfected cells (MCF7: $P = 0.004$, MDA-MB-231: $P = 0.022$). (**B**) Wound closure rates were compared between variant and WT mimic-transfected MCF7 and MDA-MB-231 cells using a simple scratch assay to determine whether the variant-associated increased migratory capacity persisted in the absence of a FBS-derived chemotactic gradient. No differences in closure rates were found in either cell line, which suggests that observed changes in motility arose from extracellular interactions with chemoattractants present in serum. Fold change values were normalized based on the negative control. Error bars represent two-sided 95% confidence intervals.

its infancy. Although it was believed that the primary role of piRNAs is to silence active transposable elements in the germ line and thereby guard the genome against transposon-induced insertional mutations, mounting evidence now point towards an active role for piRNAs in somatic gene regulation through, among other processes, sequence-specific DNA methylation. This has been perhaps most convincingly demonstrated in *Aplysia* neurons, where an endogenously expressed piRNA was found capable of inducing methylation at the *CREB2* gene promoter (11). In humans, the evidence is more circumstantial, with the finding of a strong correlation between expression of a 'piRNA-like' 28-nt transcript antisense to the *KIR3DL1* promoter and *KIR3DL1* promoter methylation in CD56+ NK cells (10). Our recent epigenome-wide analysis further demonstrated piRNA mediated gene specific DNA methylation in human cells (38). Despite the emergent nature of the evidence, the prospect that piRNA-mediated DNA methylation may represent a general somatic phenomenon holds vast implications for our understanding of epigenetic control of the oncogenic process.

Our finding of piR-021285-responsive gene methylation represents the first evidence of a piRNA-mediated epigenetic mechanism capable of affecting cancer-related processes in the human breast. Consistent with the observed association between piR-021285 SNP rs1326306 and breast cancer risk, methylation differences were found at multiple breast cancer-related genes between WT and variant piR-021285 mimic-transfected MCF7 cells. While we note that our findings are limited to one cell line, piR-021285's evident capacity to influence the methylation of multiple genes in MCF7 implies that, like oncomiRs, certain piRNAs may possess multiple gene targets and their disruption can have reverberating consequences on multiple downstream cancer-relevant pathways.

Closer examination of the nature of the methylation differences between WT and variant piR-021285 mimic-transfected cells yields additional insights of interest. Despite concurrent differences in mRNA expression at a number of breast cancer-relevant genes, no gene exhibited statistically significant methylation differences at more than one CpG site. This finding implies that piR-021285 may be capable of discriminating between and targeting individual CpG sites for methylation, which nonetheless may be sufficient to provoke downstream phenoptypic changes in cancer-relevant pathways. Furthermore, while previous evidence of piRNA-associated somatic methylation is constrained to the promoter (10,11), we observed methylation differences at a number of gene body-associated CpG sites, which, in the case of the antiproliferative *DLC1* gene, was directly correlated with mRNA expression. This suggests that piRNA-associated epigenetic activation, in addition to silencing, may be actively present in human breast carcinogenesis. Of additional interest is the finding of both attenuation and

enhancement of methylation in variant piR-021285 mimic-transfected cells, suggesting that rs1326306, located 12 bases from the 5′ end, may be situated in a region responsible for target recognition rather than biogenic processing or PIWI affinity, in which case more unidirectional changes would be expected. Moreover, the degree of these differences in a subset of breast cancer-related genes suggests that this region is highly sensitive to polymorphic disruption, such that a single base substitution can result in the abolishment of methylation at existing targets and the simultaneous establishment of novel targets.

Of the six 'functionally relevant' cancer-related genes exhibiting variant-responsive methylation, four displayed concurrent differences in expression at the mRNA level. Although exhibiting attenuated methylation in the variant transfection, *AATF* displayed increased mRNA expression, which is at odds with the conventional understanding of gene body methylation and transcriptional activity being directly correlative (39). However, this relationship may be confounded by the putative role gene body methylation plays in the regulation of splicing (39–41) or crosstalk with histone or post-transcriptional modifiers. In spite of expression differences in apoptosis and proliferation-affiliated genes, no detectable differences in either proliferative rate or apoptotic incidence were observed between WT and variant mimic-transfected cell lines. A number of possible scenarios could account for the lack of phenotypic response, including undetected phenotypic changes at other genes with counteracting effects, expression differences being below the biological threshold required to elicit downstream effects and interference from post-transcriptional crosstalk. Of significance is the finding of functional consistency between attenuated 5′ UTR/first exon methylation at the pro-invasive *ARHGAP11A* and increased *ARHGAP11A* expression and cell invasiveness in variant mimic-transfected breast cancer cells. To our surprise, this enhancement in migratory capacity dissipated in the absence of a FBS-derived chemotactic gradient, implying that piR-021285-associated invasiveness may be dependent upon and directed by extracellular interactions with chemoattractants found in serum. piR-021285's evident inability to elicit a measureable impact on apoptosis or proliferation, classic "initiator" processes in cancer, suggests that piR-021285 may be a late actor in the carcinogenic process. Although this may seem difficult to reconcile given piR-021285's association with breast cancer risk, we must bear in mind the empirical nature of the diagnostic process, during which cases were assigned based on what was observable rather than biological. It is feasible, then, that variant piR-021285-associated breast cancers may be easier to detect due to their more aggressive nature.

It is important to note that a number of caveats limit the scope of our interpretations. With regard to piR-021285 expression in breast tissue, whereas we observed endogenous expression of this piRNA in MCF7 cells, future work will be required to elucidate the levels of piR-021285 expression in primary tumor and normal breast specimens and whether the variant allele may modify expression of the piRNA. Additionally, although methylation differences were observed at a number of genes between WT and variant piR-021285-transfected cells, it is uncertain what proportion of these differences was the result of direct gene–piRNA interactions. Future investigations may help yield more insight into the mechanistic nature of piR-021285 as an effector of gene-specific methylation in the context of breast cancer. Furthermore, although it is generally understood that gene body methylation is not associated with transcriptional repression, the exact downstream effects and the biological context in which they occur are more ambiguous. Affiliated developments include transcriptional activation (39,42) and altered splicing (39–41), as well as increased susceptibility to C>T transition mutations at methylated cytosines (43). Additional functional analyses may be able to offer more insights into the phenotypic nature of piR-021285-associated gene body methylation. Lastly, although the dependence of piR-021285-associated invasiveness on FBS-derived chemoattractants is in itself an interesting and potentially therapeutically valuable observation, future investigations should examine whether this phenomenon persists *in vivo* or with human sera.

Taken together, our findings implicate piR-021285 as a potential modulator of human breast cancer invasiveness, a function possibly linked to piR-021285-dependent 5′ UTR/first exon methylation of the proinvasive *ARHGAP11A* gene. Apparent associations with breast cancer susceptibility may be an artifact of increased piR-021285-associated tumor aggressiveness, a characteristic that can translate to higher detection rates. To our knowledge, our finding of piR-021285-associated methylation of breast cancer-related genes represents the first evidence of piRNA-mediated epigenetic control of cancer-relevant processes in human breast cancer. The precise implication this holds for the design of novel epigenetic anticancer therapeutics merits further investigation.

Supplementary material

Supplementary Tables 1–3 and Figures 1 and 2 can be found at http://carcin.oxfordjournals.org/

Funding

D.I.J. was supported by the Yale-National Cancer Institute (NCI) predoctoral training grant T32 CA105666. Yale University.

Acknowledgements

We would like to thank Christopher Castaldi and Irina Tikhonova at Yale University's W.M. Keck Biotechnology Resource Laboratory for performing the Illumina HumanMethylation450 array. We also thank Donghai Liang and Alison Fritz at Yale University for their laboratory and research assistance.
Conflict of Interest Statement: None declared.

References

1. Aravin, A. et al. (2006) A novel class of small RNAs bind to MILI protein in mouse testes. Nature, 442, 203–207.
2. Girard, A. et al. (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature, 442, 199–202.
3. Grivna, S.T. et al. (2006) A novel class of small RNAs in mouse spermatogenic cells. Genes Dev., 20, 1709–1714.
4. Watanabe, T. et al. (2006) Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev., 20, 1732–1743.
5. Bamezai, S. et al. (2012) Concise review: The Piwi-piRNA axis: pivotal beyond transposon silencing. Stem Cells, 30, 2603–2611.
6. Kim, V.N. (2006) Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. Genes Dev., 20, 1993–1997.
7. Siomi, M.C. et al. (2011) PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol., 12, 246–258.
8. Brennecke, J. et al. (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. Science, 322, 1387–1392.
9. Aravin, A.A. et al. (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science, 318, 761–764.
10. Cichocki, F. et al. (2010) Cutting edge: KIR antisense transcripts are processed into a 28-base PIWI-like RNA in human NK cells. J. Immunol., 185, 2009–2012.

11. Rajasethupathy, P. et al. (2012) A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. Cell, 149, 693–707.
12. Ross, R.J. et al. (2014) PIWI proteins and PIWI-interacting RNAs in the soma. Nature, 505, 353–359.
13. Sun, G. et al. (2011) Clinical significance of Hiwi gene expression in gliomas. Brain Res., 1373, 183–188.
14. He, W. et al. (2009) Expression of HIWI in human esophageal squamous cell carcinoma is significantly associated with poorer prognosis. BMC Cancer, 9, 426.
15. Grochola, L.F. et al. (2008) The stem cell-associated Hiwi gene in human adenocarcinoma of the pancreas: expression and risk of tumour-related death. Br. J. Cancer, 99, 1083–1088.
16. Zhao, Y.M. et al. (2012) HIWI is associated with prognosis in patients with hepatocellular carcinoma after curative resection. Cancer, 118, 2708–2717.
17. Wang, Y. et al. (2013) Honokiol protects rat hearts against myocardial ischemia reperfusion injury by reducing oxidative stress and inflammation. Exp. Ther. Med., 5, 315–319.
18. Zeng, Y. et al. (2011) HIWI expression profile in cancer cells and its prognostic value for patients with colorectal cancer. Chin. Med. J. (Engl)., 124, 2144–2149.
19. Li, D. et al. (2012) Piwil2 modulates the proliferation and metastasis of colon cancer via regulation of matrix metallopeptidase 9 transcriptional activity. Exp. Biol. Med. (Maywood), 237, 1231–1240.
20. Liu, J.J. et al. (2010) Piwil2 is expressed in various stages of breast cancers and has the potential to be used as a novel biomarker. Int. J. Clin. Exp. Pathol., 3, 328–337.
21. Cheng, J. et al. (2012) piR-823, a novel non-coding small RNA, demonstrates *in vitro* and *in vivo* tumor suppressive activity in human gastric cancer cells. Cancer Lett., 315, 12–17.
22. Law, P.T. et al. (2013) Deep sequencing of small RNA transcriptome reveals novel non-coding RNAs in hepatocellular carcinoma. J. Hepatol., 58, 1165–1173.
23. Cheng, J. et al. (2011) piRNA, the new non-coding RNA, is aberrantly expressed in human cancer cells. Clin. Chim. Acta., 412, 1621–1625.
24. Huang, G. et al. (2013) Altered expression of piRNAs and their relation with clinicopathologic features of breast cancer. Clin. Transl. Oncol., 15, 563–568.
25. Zheng, T. et al. (2000) Risk of female breast cancer associated with serum polychlorinated biphenyls and 1,1-dichloro-2,2'-bis(p-chlorophenyl)ethylene. Cancer Epidemiol. Biomarkers Prev., 9, 167–174.
26. Bibikova, M. et al. (2011) High density DNA methylation array with single CpG site resolution. Genomics, 98, 288–295.
27. Benjamini, Y. et al. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. Royal Stat. Soc. Series B (Methodological), 57, 289–300.
28. Sharma, M. (2013) Apoptosis-antagonizing transcription factor (AATF) gene silencing: role in induction of apoptosis and down-regulation of estrogen receptor in breast cancer cells. Biotechnol. Lett., 35, 1561–1570.
29. Patsialou, A. et al. (2012) Selective gene-expression profiling of migratory tumor cells *in vivo* predicts clinical outcome in breast cancer patients. Breast Cancer Res., 14, R139.
30. Xu, J. et al. (2013) RhoGAPs attenuate cell proliferation by direct interaction with p53 tetramerization domain. Cell Rep., 3, 1526–1538.
31. Sawa, M. et al. (2009) Drug design with Cdc7 kinase: a potential novel cancer therapy target. Drug Des. Dev. Ther., 2, 255–264.
32. Choschzick, M. et al. (2010) Overexpression of cell division cycle 7 homolog is associated with gene amplification frequency in breast cancer. Hum. Pathol., 41, 358–365.
33. Yuan, B.Z. et al. (2003) DLC-1 gene inhibits human breast cancer cell growth and *in vivo* tumorigenicity. Oncogene, 22, 445–450.
34. Plaumann, M. et al. (2003) Analysis of DLC-1 expression in human breast cancer. J. Cancer Res. Clin. Oncol., 129, 349–354.
35. Luoh, S.W. et al. (2004) Overexpression of the amplified Pip4k2beta gene from 17q11-12 in breast cancer cells confers proliferation advantage. Oncogene, 23, 1354–1363.
36. De Souza Santos, E. et al. (2008) Silencing of LRRC49 and THAP10 genes by bidirectional promoter hypermethylation is a frequent event in breast cancer. Int. J. Oncol., 33, 25–31.
37. Kagawa, Y. et al. (2013) Cell cycle-dependent Rho GTPase activity dynamically regulates cancer cell motility and invasion *in vivo*. PLoS One, 8, e83629.
38. Fu, A. et al. (2015) Epigenome-wide analysis of piRNAs in gene-specific DNA methylation. RNA Biol., 11, 1301–1312.
39. Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet., 13, 484–492.
40. Laurent, L. et al. (2010) Dynamic changes in the human methylome during differentiation. Genome Res., 20, 320–331.
41. Shukla, S. et al. (2011) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature, 479, 74–79.
42. Hellman, A. et al. (2007) Gene body-specific methylation on the active X chromosome. Science, 315, 1141–1143.
43. Rideout, W.M. 3rd et al. (1990) 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. Science, 249, 1288–1290.