Research Article

Cancer Epidemiology, Biomarkers & Prevention

miRNA-Processing Gene Methylation and Cancer Risk

Brian T. Joyce[1], Yinan Zheng[1,2], Zhou Zhang[1], Lei Liu[1,3], Masha Kocherginsky[4], Robert Murphy[5], Chad J. Achenbach[6], Jonah Musa[5,7,8], Firas Wehbe[4], Allan Just[9], Jincheng Shen[10], Pantel Vokonas[11], Joel Schwartz[12], Andrea A. Baccarelli[13], and Lifang Hou[1,3]

Abstract

Background: Dysregulation of miRNA and methylation levels are epigenetic hallmarks of cancer, potentially linked via miRNA-processing genes. Studies have found genetic alterations to miRNA-processing genes in cancer cells and human population studies. Our objective was to prospectively examine changes in DNA methylation of miRNA-processing genes and their associations with cancer risk.

Methods: We examined cohort data from the Department of Veterans' Affairs Normative Aging Study. Participants were assessed every 3 to 5 years starting in 1999 through 2013 including questionnaires, medical record review, and blood collection. Blood from 686 consenting participants was analyzed using the Illumina 450K BeadChip array to measure methylation at CpG sites throughout the genome. We selected 19 genes based on a literature review, with 519 corresponding CpG sites. We then used Cox proportional hazards models to examine associations with cancer incidence, and generalized estimating equations to examine associations with cancer prevalence. Associations at false discovery rate < 0.05 were considered statistically significant.

Results: Methylation of three CpGs (*DROSHA*: cg23230564, *TNRC6B*: cg06751583, and *TNRC6B*: cg21034183) was prospectively associated with time to cancer development (positively for cg06751583, inversely for cg23230564 and cg21034183), whereas methylation of one CpG site (*DROSHA*: cg16131300) was positively associated with cancer prevalence.

Conclusions: DNA methylation of *DROSHA*, a key miRNA-processing gene, and *TNRC6B* may play a role in early carcinogenesis.

Impact: Changes in miRNA processing may exert multiple effects on cancer development, including protecting against it via altered global miRNAs, and may be a useful early detection biomarker of cancer. *Cancer Epidemiol Biomarkers Prev; 27(5); 550–7.* ©*2018 AACR.*

Introduction

Over the past 20 years, U.S. cancer incidence and mortality have generally declined, but this trend has been driven to a great extent by changes in lifestyle and cancer screening behaviors that impact the four most common cancers in the United States (lung, breast, prostate, and colorectal; ref. 1). Death rates from other cancers have remained largely stable over the same period of time (1), emphasizing the need for new medical technologies to improve the diagnosis, prognostication, and treatment of these cancers. Studies of epigenetics, and other forms of posttranscriptional gene regulation, are believed to hold great promise for such development. miRNAs are small (~22 nucleotide), noncoding RNAs that bind to messenger RNAs and regulate posttranscriptional gene expression in up to one-third of the human genome by inhibiting gene translation (2, 3). In healthy individuals miRNAs play important roles in regulating cell proliferation, differentiation, and apoptosis while in cancer they exert both oncogenic and tumor-suppressive effects (3). Global downregulation of miRNAs has also been found in most human tumors, which is generally thought to enhance neoplastic transformation and metastasis (4). One mechanism through which this may be accomplished is through altered expression of miRNA-processing genes (5–7). Epigenetic alterations of miRNA-processing genes (8, 9), and in particular changes in the methylation of these genes (10–14), have been implicated in several types of cancer.

[1]Center for Population Epigenetics, Robert H. Lurie Comprehensive Cancer Center and Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois. [2]Institute for Public Health and Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois. [3]Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois. [4]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois. [5]Center for Global Health, Northwestern University Feinberg School of Medicine, Chicago, Illinois. [6]Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois. [7]Health Sciences Integrated Program, Center for Healthcare Studies, Institute of Public Health and Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois. [8]Department of Obstetrics and Gynecology, Faculty of Medical Sciences, University of Jos, Plateau State, Nigeria. [9]Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, New York. [10]Department of Population Health Sciences, University of Utah School of Medicine, Salt Lake City, Utah. [11]VA Normative Aging Study, Veterans Affairs Boston Healthcare System and the Department of Medicine, Boston University School of Medicine, Boston, Massachusetts. [12]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. [13]Departments of Epidemiology and Environmental Health Sciences, Columbia University Mailman School of Public Health, New York, New York.

Corresponding Author: Brian T. Joyce, Center for Population Epigenetics, Robert H. Lurie Comprehensive Cancer Center and Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, 680 N. Lake Shore Drive, Suite 1400, Chicago, IL 60611. Phone: 312-503-5407; Fax: 312-908-9588; E-mail: b-joyce@northwestern.edu

doi: 10.1158/1055-9965.EPI-17-0849

Dysregulation of DNA methylation is likewise ubiquitous across a variety of cancer types and considered a hallmark of cancer (15–17). Disease- and exposure-related methylation changes are detectable in blood, potentially allowing them to serve as biomarkers for cancer and/or the immune response to it (18–20). From a clinical standpoint, biomarkers measurable in blood are attractive for their minimally invasive, cost-effective method of collection and potential utility for screening, risk stratification, and personalized medicine (21). From a biological standpoint, epigenetic alterations of blood leukocyte DNA may be a precursor to cancer, for example, due to their involvement in inflammatory and immune response pathways (19, 22–26). Additional evidence suggests that global miRNA levels in leukocytes may play a role in these pathways as well (27–29). However, few studies have measured epigenetic alterations of miRNA-processing genes in more accessible surrogate tissues such as blood. Exploring epigenetic modifications of these genes over time may thus inform their use as an early detection biomarker of cancer and potentially provide important information on the role global miRNA dysregulation plays in carcinogenesis and/or cancer promotion.

The primary objective of this study is to prospectively examine the relationship between DNA methylation of miRNA-processing genes and cancer incidence. We will also examine the cross-sectional relationship between miRNA-processing gene methylation and cancer prevalence. Our primary hypothesis is that time to cancer incidence will be positively associated with increased DNA methylation in miRNA-processing genes. We also hypothesize that cancer prevalence will be positively associated with increased DNA methylation of miRNA-processing genes, after adjusting for time between diagnosis and blood draw in addition to other covariates.

Materials and Methods

Study population

In 1963, the Normative Aging Study (NAS) was established by the U.S. Department of Veterans Affairs with an initial cohort of 2,280 healthy male veterans who were living in the Boston area; were of ages 21 to 80; and had no history of hypertension, cardiovascular disease, cancer, diabetes, or other chronic health conditions. From 1963 to 1999, 981 (43%) participants died and 470 (21%) were lost to follow up. Our previous study using NAS data (30) found no significant differences in subject characteristics [e.g., age, body mass index (BMI)] between participants lost to follow up and the 829 still active in the study as of 1999. Participants were recalled for in-person follow up every 3 to 5 years, and beginning in 1999 these visits included 7-mL blood samples for genetic and epigenetic analysis. All participants were asked to fast and abstain from smoking prior to their morning visit. From January 1, 1999, to December 31, 2013, 802 of 829 participants (96.7%) consented to blood donation and 686 were randomly selected for whole-epigenome analysis ($n = 1379$ samples) using the Illumina 450K BeadChip array; our prior analysis also found no differences in characteristics between this randomly selected subset and the full population of 802 participants (31). Median age at first blood draw for these participants was 72 (range: 5–100). Data including blood samples, physical examinations, and questionnaires were collected from these participants at each visit for a total of 1 to 4 visits separated by a median interval of 3.5 years (IQR: 3.1–5.7 years). The NAS was approved by the institutional review boards of all participating institutions, and all participants provided written consent. This analysis excludes 18 participants of nonwhite race (to minimize potential confounding by genetic ancestry) as well as participants with a diagnosis of unspecified malignancy (24 pre-1999 diagnoses and 20 diagnoses occurring between 1999 and 2013), leaving 624 participants for analysis. Of these, 193 (31%) had methylation measured at one time point, 288 (46%) at two, 139 (22%) at three, and four (1%) at four.

Cancer diagnosis

Information on medical history obtained from questionnaires was confirmed via blinded medical record review, and included cancer diagnoses and comorbidities. Of 624 participants included in this analysis 182 (29%) reported prevalent cancers (54 prostate, 82 skin, 16 colorectal, 5 lung, 4 bladder, 21 others) prior to their first blood sample collection and 132 (21%) developed incident cancer (38 prostate, 50 skin, 9 lung, 5 colorectal, 5 bladder, 25 others including leukemia, lymphoma, brain, etc.) during follow up. Median follow-up time from first blood draw to cancer diagnosis/censoring was 10.1 years (IQR: 5.8–12.7 years).

Methylation measurement

We identified 19 candidate miRNA-processing genes through: (i) a PubMed literature search for genes directly involved in miRNA biogenesis and degradation, (ii) a GeneCard search for additional genes modifying the expression of those genes identified in step (i), and (iii) a second PubMed search to identify associations between cancer and either up- or downregulation of any of the genes identified in steps 1 and 2 (Fig. 1). We identified 30 genes as being involved in miRNA processing but without a link to cancer, which were excluded from subsequent analysis. Methylation of each CpG site within and up to 50 kb upstream of each gene of interest was examined (to ensure proper characterization of promoter regions). We ultimately identified 519 CpGs on 19 genes as candidates for analysis; full details on these CpG sites can be found in our Supplementary Materials (Supplementary Table S1).

For methylation analysis, DNA was extracted from the buffy coat using the QIAamp DNA Blood Kit (Qiagen). A total of 500 ng of DNA was used to perform bisulfite conversion using the EZ-96

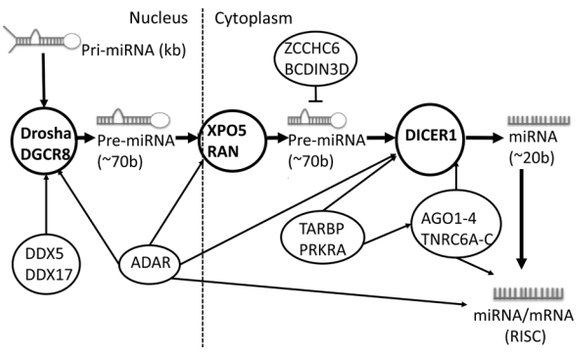

Figure 1.

Roles of candidate genes in miRNA biogenesis. Figure 1 represents the involvement of each candidate gene of interest in distinct step(s) of miRNA processing both within and without of the nucleus. It also labels distinct precursor miRNA products generated up to mRNA silencing in the RNA silencing complex (RISC). Critical miRNA processing genes/steps are bolded.

DNA Methylation Kit (Zymo Research). To limit chip and plate effects a two-stage, age-stratified algorithm was used to randomize samples and ensure similar age distributions across chips and plates. We randomized 12 samples (sampled across all age quartiles) per chip, and then randomized chips to plates. We measured DNA methylation of CpG probes using Illumina Infinium HumanMethylation450 BeadChip (32). Quality control analysis was performed to remove samples with >1% of probes that had a detection $P > 0.05$, or probes with P value > 5% samples with a bead count < 3 or with >1% samples with a detection $P > 0.05$. The remaining samples were preprocessed using Illumina-type background correction (33) and normalized with the dye-bias (34) and BMIQ (35) adjustments to generate methylation β-values, which we then converted to M-values for statistical analysis (36).

Missing data

Of the 481,462 CpG sites from 1379 samples (total 663,936,098) processed, 10,406,437 (1.6%) CpG measurements were missing after quality control and data preprocessing. Of the 519 CpG sites of interest from 1,379 samples (725,354 values total), 11,094 values (1.5%) were missing. All 519 CpG sites of interest had less than 10% missing data, and 94.0% of them had less than 5% missing data across the 1,379 samples. Using the full dataset, we used the K Nearest Neighbors method to impute missing values (37). Assuming missingness completely at random, we took advantage of the correlated structure of methylation data to estimate the missing methylation values and impute them in the final dataset. The univariate distributions of all significant cancer-associated CpG sites were examined for outliers via visual inspection and cross-checked against the methylation values imputed in the primary dataset. For one CpG site (cg15827285) significantly associated with cancer incidence, we identified six outliers that were the results of imputation (i.e., missing in the original dataset). With these observations excluded this model was rerun, and the resulting nonsignificant results are not reported (data available upon request).

Statistical methods

To evaluate the stability of our methylation markers of interest across individuals and across time we calculated the median and interquartile range (IQR) of all methylation values at the first blood draw as well as the three-year intraclass coefficient (ICC; Supplementary Table S1). For the ICC calculation, we restricted our sample to the first two visits per each participant (excluding participants with only a single visit) and only to visits between two and four years apart (339/466 or 73% of all participants with at least two observations). ICCs were then calculated separately for each CpG site using a mixed linear regression model only adjusting for technical variations and cell-type abundancies. A descriptive analysis using data at the first blood draw evaluated differences across subject characteristics by cancer status (cancer-free vs. prevalent cancer vs. incident cancer vs. unspecified) using χ^2 and Kruskal–Wallis tests for categorical and continuous variables, respectively. For the cancer incidence analysis, we examined the relationship between time-dependent miRNA processing gene methylation and time to cancer diagnosis via Cox proportional hazards models. Methylation measures obtained after a diagnosis of cancer were excluded from this analysis. As with our prior analyses, all multivariate models included three principal components that we previously showed could adjust for DNA CpG methylation processing batch, white blood cell counts (WBC), and proportion neutrophils to account for technical variation and potential disease-related alterations in white blood cell types (31) based on the method by Houseman and colleagues (38). On the basis of our previous work with methylation and cancer in this cohort (19), we additionally included age, BMI, education, smoking status, smoking pack-years, and alcohol consumption (dichotomized as reporting 0–1 vs. 2+ drinks per day on average) as covariates.

For the cancer prevalence analysis, we used generalized estimating equations (GEE) to examine associations between cancer prevalence and DNA CpG methylation according to the following model equation:

$$Y_{ijk} = \beta_{0k} + \beta_{1k} t_{ij} + \beta_{2k} I(t_{ij} > D_i) + \ldots + e_{ijk}$$

where Y_{ijk} is the the kth methylation marker of subject i at time t_{ij} (i.e., each methylation marker was examined individually, in its own model), β_{0k} is the intercept, β_{1k} is the regression coefficient for temporal trend, β_{2k} is the regression coefficient for $I(t_{ij} > D_i)$ to denote the effect of cancer diagnosis (D_i) on methylation marker (i.e., an indicator variable for whether the methylation marker was measured after cancer diagnosis), and e_{ijk} the residual error term. Various linear combinations of coefficients in this model represent the average increases in methylation (as measured in standardized units) over time and for participants who are cancer-free versus those who have already developed cancer. All cancer prevalence models adjusted for the same covariates as above. All incidence and prevalence analyses adjusted for multiple testing using the Benjamini–Hochberg false discovery rate (FDR; ref. 39), based on the number of tests performed per each candidate gene.

Finally, we conducted several sensitivity analyses. First, for all analyses of *DICER1* and *DGCR8* methylation, we examined the effects of including SNPs on those genes that are available in the NAS dataset (rs1640299, rs3742330, and rs13078). Next, we performed two analyses using data from The Cancer Genome Atlas (TCGA). We first examined Pearson correlations between methylation at each of our significant CpG sites and RNA levels in both skin and prostate cancer tissue (the most common cancer types in the NAS). We then conducted a case–control analysis of each of our significant CpG sites using TCGA prostate tissue (due to lack of healthy skin tissue, skin cancer could not be analyzed) via unpaired t test (so as to maximize sample size for comparison). We conducted all NAS analyses in SAS v.9.4 and all TCGA analyses in R v.3.4.2, and changes were considered statistically significant if the FDR-corrected P value was less than 0.05 (40). Both FDR and Bonferroni-corrected P values are reported in the complete results, available in our Supplementary Materials.

Results

Select participant characteristics are presented in Table 1. Briefly, age varied by cancer status ($P < 0.0001$), with incident cancer cases younger and prevalent cases older than those remaining cancer-free. There were no other significant variations in participant characteristics. For the Cox proportional hazards models, results significant at $P < 0.05$ are presented in Table 2; complete results are available in Supplementary Table S2. Prior to adjusting for multiple tests a total of 44 (19 hypermethylated, 25 hypomethylated) CpG sites on 16 genes were significantly associated with time to cancer diagnosis. After FDR adjustment a total of three CpG sites (1 hypermethylated, 2 hypomethylated) remained significantly associated with time to cancer diagnosis at FDR < 0.05: cg23230564 on *Drosha* (HR = 0.66, FDR = 0.05),

Table 1. Participant characteristics at first blood draw by cancer status

	Total	Cancer-free	Prevalent cancer	Incident cancer	Unknown (excluded)	*P*
N		310	182	132	44	
Age (years)	72.5 ± 6.9	72.1 ± 6.8	74.3 ± 6.7	70.7 ± 6.6	72.4 ± 7.1	<0.0001
BMI	28.1 ± 4.1	28.4 ± 4.3	27.7 ± 3.6	28.2 ± 4.2	27.3 ± 3.4	0.25
Education						
≤High school	174 (26%)	87 (28%)	44 (24%)	27 (21%)	16 (36%)	0.43
Some college	328 (49%)	149 (48%)	89 (49%)	71 (54%)	19 (43%)	
College grad	166 (25%)	74 (24%)	49 (27%)	34 (26%)	9 (21%)	
Smoking						
Never	193 (29%)	78 (25%)	58 (32%)	41 (31%)	16 (36%)	0.10
Current	27 (4%)	16 (5%)	2 (1%)	8 (6%)	1 (2%)	
Former	448 (67%)	216 (70%)	122 (67%)	83 (63%)	27 (61%)	
Pack-years	21.1 ± 25.2	21.3 ± 24.5	21.6 ± 26.8	21.8 ± 25.6	15.0 ± 21.7	0.30
Alcoholic drinks						
0–1/day	537 (80%)	257 (83%)	139 (76%)	110 (83%)	31 (71%)	0.09
2+/day	131 (20%)	53 (17%)	43 (24%)	22 (17%)	13 (30%)	

Table 2. Cox model results where $P < 0.05$ ($n = 442$)

CpG	Gene	HR	*P*	FDR
cg23230564	*DROSHA*	0.66	<0.01	0.05[a]
cg06751583	*TNRC6B*	1.62	<0.01	0.04[a]
cg21034183	*TNRC6B*	0.56	<0.01	0.04[a]
cg01780585	*AGO1*	0.46	0.02	0.41
cg24492446	*AGO2*	2.16	0.01	0.48
cg11598062	*AGO2*	0.67	0.01	0.48
cg09815962	*AGO2*	1.50	0.02	0.48
cg11130692	*AGO2*	0.45	0.02	0.48
cg14650175	*AGO2*	1.74	0.02	0.48
cg15936375	*AGO2*	1.56	0.04	0.56
cg00405484	*AGO2*	0.62	0.04	0.56
cg10010182	*AGO2*	0.76	0.04	0.56
cg08444833	*AGO2*	0.72	0.04	0.56
cg15717330	*AGO2*	1.32	0.05	0.56
cg06007675	*AGO2*	1.26	0.05	0.56
cg03988279	*AGO3*	0.74	0.04	0.43
cg05930400	*AGO4*	0.71	0.04	0.44
cg15107202	*BCDIN3D*	0.78	0.03	0.42
cg03373091	*DDX17*	2.04	<0.01	0.06
cg02917540	*DDX17*	1.90	0.01	0.06
cg13604020	*DDX17*	0.67	0.03	0.18
cg17829984	*DDX5*	0.64	0.02	0.34
cg01901579	*DICER1*	1.35	0.03	0.33
cg23488578	*DICER1*	0.73	0.03	0.33
cg13997647	*DICER1*	0.55	0.04	0.33
cg00580354	*PRKRA*	0.70	<0.01	0.07
cg27480241	*PRKRA*	0.62	0.01	0.07
cg13151361	*PRKRA*	0.75	0.05	0.38
cg17629322	*RAN*	1.24	0.02	0.39
cg22265988	*TARBP2*	0.78	0.03	0.45
cg06420129	*TNRC6A*	0.77	0.03	0.36
cg08811227	*TNRC6A*	1.73	0.04	0.36
cg27407707	*TNRC6A*	0.58	0.04	0.36
cg01949902	*TNRC6B*	0.62	0.02	0.12
cg15462501	*TNRC6B*	0.68	0.03	0.19
cg14523804	*TNRC6C*	1.53	0.01	0.21
cg00142402	*TNRC6C*	1.64	0.01	0.21
cg12072376	*TNRC6C*	0.74	0.01	0.21
cg07787614	*TNRC6C*	1.29	0.02	0.21
cg26740249	*TNRC6C*	1.22	0.02	0.21
cg00423030	*TNRC6C*	1.28	0.03	0.22
cg18748888	*TNRC6C*	2.12	0.05	0.29
cg15646497	*XPO5*	0.51	0.01	0.11
cg01849007	*XPO5*	1.48	0.01	0.11

[a]Statistically significant at FDR < 0.05.

and cg06751583 and cg21034183 on *TNRC6B* (HR = 1.62 and 0.56, respectively, FDR = 0.04 for each). When using the more stringent Bonferroni correction for multiple testing, two of these three CpG sites remained significantly associated with time to cancer diagnosis while cg06751583 became nonsignificant ($P = 0.08$).

For the GEE models of prevalent cancers, results significant at $P < 0.05$ are presented in Table 3; complete results can be found in Supplementary Table S3. Briefly, prior to FDR adjustment 30 (14 hypermethylated, 16 hypomethylated) CpG sites on 16 genes were associated with cancer prevalence. After adjusting for multiple tests, one CpG site remained significantly associated with cancer prevalence at FDR < 0.05: cg16131300 on *Drosha* ($\beta = 0.055$, FDR = 0.04). This CpG remained significantly associated with prevalent cancers after Bonferroni correction. For both the cancer incidence and the cancer prevalence analyses of *DGCR8* and *DICER1*, we found no CpG sites associated with cancer regardless of whether data on SNPs were included in our models (results available upon request). Supplementary Table S4 shows the correlations between methylation of our significant CpG sites and RNA expression in TCGA cancer tissue. In skin cancer, methylation at cg16131300 in *DROSHA* ($r = 0.31$, $P < 0.01$) as well as cg06751583 ($r = 0.25$, $P = 0.01$) and cg21034183 ($r = -0.21$, $P = 0.03$) in *TNRC6B* was significantly and positively correlated with RNA expression. In prostate cancer methylation of cg23230564 ($r = 0.27$, $P < 0.01$) and cg16131300 ($r = 0.12$, $P = 0.01$) in *DROSHA* was significantly and positively correlated with RNA expression, as was methylation of cg06751583 in *TNRC6B* ($r = 0.30$, $P < 0.01$). We also observed a marginally significant positive correlation between methylation at cg21034183 in *TNRC6B* and RNA expression ($r = -0.08$, $P = 0.08$). Finally, our case–control analysis examining TCGA prostate cancer tissue was only able to validate one of our four significant findings. We found that methylation of cg23230564 in *DROSHA* was significantly lower in cancer tissue compared with normal (normal mean methylation M = 2.19 in tumor and 2.51 in normal tissue; $P < 0.01$).

Discussion

To our knowledge, this is the first study examining prospective relationships between alterations of DNA methylation of miRNA-processing genes in blood leukocyte DNA and cancer

Table 3. GEE results where $P < 0.05$ ($n = 624$)

CpG	Gene	β	*P*	FDR
cg16131300	*DROSHA*	0.05	<0.01	0.04[a]
cg09786420	*ADAR*	0.08	0.01	0.45
cg27026509	*AGO1*	0.04	0.04	0.57
cg00175844	*AGO2*	−0.05	0.01	0.25
cg00405484	*AGO2*	−0.08	<0.01	0.25
cg01980793	*AGO2*	−0.08	0.02	0.38
cg04476876	*AGO2*	0.04	0.03	0.38
cg07281647	*AGO2*	−0.04	0.02	0.38
cg11130692	*AGO2*	−0.04	0.03	0.38
cg14301764	*AGO2*	−0.07	<0.01	0.09
cg14484434	*AGO2*	−0.05	0.03	0.38
cg14835938	*AGO2*	−0.08	0.05	0.47
cg17050632	*AGO2*	−0.08	0.04	0.42
cg21208682	*AGO2*	−0.09	0.02	0.38
cg25071674	*AGO2*	0.06	0.03	0.38
cg02927292	*AGO3*	−0.05	0.04	0.67
cg07875848	*BCDIN3D*	−0.04	0.01	0.19
cg27416957	*DDX17*	0.04	0.04	0.68
cg01795697	*DDX5*	−0.03	0.05	0.81
cg26372423	*DGCR8*	−0.02	0.05	0.71
cg23126376	*DICER1*	0.08	0.03	0.67
cg00460704	*PRKRA*	−0.07	0.04	0.78
cg01060409	*RAN*	0.08	0.01	0.17
cg02785870	*TNRC6A*	0.05	0.05	0.62
cg03156546	*TNRC6A*	0.1	0.01	0.25
cg14311088	*TNRC6B*	0.04	0.02	0.38
cg00175441	*TNRC6C*	0.08	0.01	0.46
cg05736847	*TNRC6C*	0.06	0.03	0.46
cg18211066	*TNRC6C*	−0.11	0.03	0.46
cg14454015	*ZCCHC6*	0.04	0.04	0.31

[a]Statistically significant at FDR < 0.05.

risk. These results suggest relationships between DNA methylation of two miRNA-processing genes (*DROSHA* and *TNRC6B*) and cancer. DNA methylation of three CpGs, one on *Drosha* and two on *TNRC6B*, was prospectively associated with time from sample collection to cancer diagnosis. Methylation of another CpG (also on *Drosha*) was associated with cancer prevalence in this cohort, suggesting that these genes may play a role in cancer and that these CpG sites may be useful for cancer detection.

cg23230564 is located in an mRNA-encoding area of the *Drosha* gene body. Both up- (10) and downregulation (41, 42) of *Drosha* have been found in various cancer cells. However, in humans SNPs in *Drosha* have been associated with risk of breast (43, 44) and bladder (45) cancer as well as lung cancer recurrence (46). Emerging evidence suggests that methylation of CpG sites in the gene body region can stimulate transcriptional elongation (47). Our analysis of TCGA data found that increased methylation at this CpG site was associated with increased *DROSHA* RNA expression, and found lower methylation at this CpG site in prostate cancer tissue compared with normal. If confirmed in other cohorts, this protective finding would be consistent with our hypothesized protective effect of greater miRNA processing gene expression on cancer. Future research should explore prospective cross-talk between cg23230564 methylation and miRNA expression levels in blood from cancer patients to explore the potential mechanistic explanations for this association outlined above. Paired blood–tissue studies could also examine the extent to which DNA methylation of cg23230564 correlates between blood and tumor tissue. In addition, this CpG site demonstrated only modest variation in methylation at the first blood draw (IQR = 0.13) and strong ICC (0.68), suggesting that it is not highly variable in the population and is relatively stable over time, meaning it may be useful as a potential biomarker.

Another protective association was identified at cg21034183 on *TNRC6B*, located in an mRNA-encoding area of the 5′UTR of *TNRC6B*. Conversely DNA methylation of the other CpG site identified on *TNRC6B*, cg06751583, was positively associated with time to cancer incidence. This CpG site is within the TSS200 region, suggesting that it affects gene expression. The differential effects between these two CpG sites may be reflective of hydroxymethylation at one CpG site and methylcytosine methylation at the other, resulting in opposite effects on gene expression and therefore (potentially) disease risk. However this explanation is not supported by our TCGA analysis, which found DNA methylation at both sites to be positively associated with gene expression in cancer tissue. Our case–control analysis also could not validate our findings. One potential resolution to these contradictions comes from an *in vivo* study (48) that identified downregulated H3K36me3 around cg21034183 in cancers. H3K36 has been shown *in vitro* to inhibit gene expression (49–52). As no significant histone modifications have been identified in the region of cg06751583, a mechanism through H3K36me3 may explain these contradictory results. With a low IQR (0.08) at the first blood draw and a high ICC (0.70), cg21034183 may have sufficient between- and intra-individual stability to be a useful biomarker, compared with cg06751583, which was stable between individuals at the first blood draw (IQR 0.04), but not over time (ICC = 0.08). However the lack of validation in our TCGA case–control analysis suggests that our results should be interpreted with caution and that further studies of methylation at these loci in cancer are necessary.

The CpG site significantly associated with cancer prevalence, cg16131300, is located within an mRNA-encoding area of the 5′UTR region of *Drosha* 62 kb away from cg23230564. Given that methylation at cg16131300 was associated with cancer only postdiagnostically, this may suggest a possible relationship to cancer prognosis or progression rather than development. However our TCGA validation analysis failed to validate this finding; this may reflect tissue-specific differences in methylation of cg16131300 or a relation to cancer progression. In addition the low interindividual variation (IQR = 0.01) coupled with the low ICC (0.04) suggests that this CpG site may be dynamic over time. This may include a sensitivity to cancer development, reflected in its low consistency over time and the association with postdiagnosis cancer only. However, the lack of validation in our TCGA case–control analysis suggests that this is not the case, thus our results should be interpreted with caution until they can be validated. Some studies have associated *Drosha* with cancer-related death in humans (53–55). Another possibility is that methylation of *Drosha* is a mechanism by which tumors evade the immune response, as demonstrated *in vitro* by miRNA-based reprogramming of dendritic cells by lung tumors (56). Other *in vitro* studies have also proposed a role for *Drosha* expression (29) and/or numerous miRNAs in immune system functioning (57–59). *Drosha* deletion in T cells has also been found to produce severe inflammation in mice (60), raising the possibility that epigenetic silencing of *Drosha* in blood leukocytes may be a method by which cancer promotes inflammation. Future studies should also explore *Drosha* expression in relation to inflammation, the immune response, and cancer progression.

This analysis has several limitations of note. First, the NAS population is not representative. These subjects are older, white, and male; live in Boston; and have a history of military service, all of which may affect the results of the above analyses. Thus, validation in other, more diverse populations is necessary before these findings can be applicable to the general population. While we attempted a simple validation using prostate cancer tissue data from TCGA, the lack of available prospectively collected DNA methylation data from blood collected prior to cancer diagnosis is a serious limitation to our findings. As the two datasets are very heterogeneous in terms of data collection methods, population characteristics, etc. future research to replicate our findings is warranted. Furthermore, DNA methylation and gene expression changes in cancer tissue may not reflect the biological dysregulation measured by disease-related changes in blood DNA methylation (e.g., inflammation and immune response). Thus validation of our findings may be more appropriate in a cohort with available blood DNA data. Second, cancer data from the NAS are limited. Lack of information on mode of detection for prostate cancer, and lack of sufficient sample size for most other cancer types present, prevented a more detailed exploration of specific cancers. Because of the biological heterogeneity of different cancer types, these findings may not be generalizable to all cancers and should be verified in specific cancer types. However, this limitation was somewhat obviated by the use of blood leukocyte DNA for epigenetic analysis, as it reflects immune and inflammatory pathways and systemic exposures common to many different cancer types. Thus, all-cancer incidence may be an appropriate outcome for this analysis.

In conclusion, these findings suggest that methylation changes to two miRNA processing genes, *Drosha* and *TNRC6B*, may be important early events in cancer. Methylation changes at specific CpG sites on these genes may also reflect the alteration of miRNA biogenesis and/or inflammatory pathways that in turn affect cancer development. Alternatively, the proper functioning of these pathways in participants who did not develop cancer despite their advanced age and the lengthy follow up of this cohort may point to the important role of miRNAs in tumor suppression elucidated in other, prior research. Extensive additional research is necessary to validate these findings in other, more diverse cohorts (e.g., younger, female, and racial/ethnic minority populations, and with greater numbers of site-specific cancers). With sufficient validation and further study, these and other epigenetic events occurring early in cancer development could potentially be used in the future to develop an early detection biomarker for cancer. If confirmed these relationships could also reveal important information about a potential host of cancer-promoting (or suppressing) effects at the cellular level. Taken together, these findings indicate that epigenetic alterations of miRNA processing genes are a rich potential field of study, and future research should attempt to elucidate their additional potential clinical applications in cancer research and clinical care.

Disclosure of Potential Conflicts of Interest

M. Kocherginsky has reported receiving royalties for U.S. Patent US8710035 B2. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: B.T. Joyce, L. Liu, J. Schwartz, A. Baccarelli, L. Hou
Development of methodology: B.T. Joyce, L. Liu, A. Baccarelli
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A. Just, P. Vokonas, J. Schwartz, A. Baccarelli
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): B.T. Joyce, Y. Zheng, Z. Zhang, L. Liu, M. Kocherginsky, R. Murphy, A. Just, P. Vokonas, J. Schwartz, A. Baccarelli
Writing, review, and/or revision of the manuscript: B.T. Joyce, Z. Zhang, L. Liu, M. Kocherginsky, R. Murphy, C. Achenbach, J. Musa, F. Wehbe, A. Just, A. Baccarelli, L. Hou
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): B.T. Joyce, R. Murphy, F. Wehbe, J. Shen, P. Vokonas, L. Hou
Study supervision: B.T. Joyce, F. Wehbe, P. Vokonas, A. Baccarelli, L. Hou

Acknowledgments

The Normative Aging Study is supported by the Epidemiology Research and Information Center of the U.S. Department of Veterans Affairs (NIEHS R01-ES015172) and is a research component of the Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC). L. Hou received additional support from the Northwestern University Robert H. Lurie Comprehensive Cancer Center Rosenberg Research Fund. L. Hou, R. Murphy, and C. Achenbach received additional support from the National Cancer Institute (1U54CA221205-01). A. Baccarelli and J. Schwartz received additional support from the National Institute of Environmental Health Sciences (NIEHS) R01-ES021733, R01-ES015172, R01ES025225, P30-ES009089, P30-ES00002. L. Liu received support from the Agency for Health Research and Quality (R01 HS020263). A. Just received support from the NIEHS (R0-ES023450). F. Wehbe received support from the Robert H. Lurie Comprehensive Cancer Center (3 P30 CA060553).

Dr. Joyce would like to thank the members of his dissertation committee from the University of Illinois at Chicago for their feedback on this work, including Dr. Garth Rauscher, Dr. Maria Argos, Dr. Alan Diamond, and Dr. Larisa Nonn.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 27, 2017; revised December 6, 2017; accepted February 2, 2018; published first February 23, 2018.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin 2015;65:5–29.
2. Davalos V, Esteller M. MicroRNAs and cancer epigenetics: a macrorevolution. Curr Opin Oncol 2010;22:35–45.
3. Babashah S, Soleimani M. The oncogenic and tumour suppressive roles of microRNAs in cancer and apoptosis. Eur J Cancer 2011;47:1127–37.
4. Melo SA, Esteller M. A precursor microRNA in a cancer cell nucleus: get me out of here! Cell Cycle 2011;10:922–5.
5. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, et al. The Microprocessor complex mediates the genesis of microRNAs. Nature 2004;432:235–40.
6. Chiang HR, Schoenfeld LW, Ruby JG, Auyeung VC, Spies N, Baek D, et al. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev 2010;24:992–1009.
7. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. Nature 2004;432:231–5.
8. Leibowitz-Amit R, Sidi Y, Avni D. Aberrations in the micro-RNA biogenesis machinery and the emerging roles of micro-RNAs in the pathogenesis of cutaneous malignant melanoma. Pigment Cell Melanoma Res 2012;25:740–57.
9. Zheng ZH, Sun XJ, Fu WN, Guan Y, Gao F, Wang Y, et al. Decreased expression of DICER1 in gastric cancer. Chin Med J 2007;120:2099–104.

10. Catto JW, Miah S, Owen HC, Bryant H, Myers K, Dudziec E, et al. Distinct microRNA alterations characterize high- and low-grade bladder cancer. Cancer Res 2009;69:8472–81.
11. Kitagawa N, Ojima H, Shirakihara T, Shimizu H, Kokubu A, Urushidate T, et al. Downregulation of the microRNA biogenesis components and its association with poor prognosis in hepatocellular carcinoma. Cancer Sci 2013;104:543–51.
12. Leaderer D, Hoffman AE, Zheng T, Fu A, Weidhaas J, Paranjape T, et al. Genetic and epigenetic association studies suggest a role of microRNA biogenesis gene exportin-5 (XPO5) in breast tumorigenesis. Int J Mol Epidemiol Genet 2011;2:9–18.
13. Melo SA, Ropero S, Moutinho C, Aaltonen LA, Yamamoto H, Calin GA, et al. A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. Nat Genet 2009;41:365–70.
14. Nemlich Y, Greenberg E, Ortenberg R, Besser MJ, Barshack I, Jacob-Hirsch J, et al. MicroRNA-mediated loss of ADAR1 in metastatic melanoma promotes tumor growth. J Clin Invest 2013;123:2703–18.
15. Laird PW. Cancer epigenetics. Hum Mol Genet 2005;14Spec No 1:R65–76.
16. Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. Biochim Biophys Acta 2007;1775:138–62.
17. Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, et al. Increased methylation variation in epigenetic domains across cancer types. Nat Genet 2011;43:768–75.
18. Brennan K, Flanagan JM. Is there a link between genome-wide hypomethylation in blood and cancer risk? Cancer Prev Res 2012;5:1345–57.
19. Joyce BT, Gao T, Liu L, Zheng Y, Liu S, Zhang W, et al. Longitudinal study of DNA methylation of inflammatory genes and cancer risk. Cancer Epidemiol Biomarkers Prev 2015;24:1531–8.
20. Koestler DC, Marsit CJ, Christensen BC, Accomando W, Langevin SM, Houseman EA, et al. Peripheral blood immune cell methylation profiles are associated with nonhematopoietic cancers. Cancer Epidemiol Biomarkers Prev 2012;21:1293–302.
21. Dietrich D, Meller S, Uhl B, Ralla B, Stephan C, Jung K, et al. Nucleic acid-based tissue biomarkers of urologic malignancies. Crit Rev Clin Lab Sci 2014;51:173–99.
22. Backdahl L, Bushell A, Beck S. Inflammatory signalling as mediator of epigenetic modulation in tissue-specific chronic inflammation. Int J Biochem Cell Biol 2009;41:176–84.
23. Campion J, Milagro FI, Goyenechea E, Martinez JA. TNF-alpha promoter methylation as a predictive biomarker for weight-loss response. Obesity (Silver Spring) 2009;17:1293–7.
24. Nile CJ, Read RC, Akil M, Duff GW, Wilson AG. Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. Arthritis Rheum 2008;58:2686–93.
25. Wilson AG. Epigenetic regulation of gene expression in the inflammatory response and relevance to common diseases. J Periodontol 2008;79:1514–9.
26. Griewank KG, Ugurel S, Schadendorf D, Paschen A. New developments in biomarkers for melanoma. Curr Opin Oncol 2013;25:145–51.
27. Asirvatham AJ, Gregorie CJ, Hu Z, Magner WJ, Tomasi TB. MicroRNA targets in immune genes and the Dicer/Argonaute and ARE machinery components. Mol Immunol 2008;45:1995–2006.
28. Tang X, Tang R, Xu Y, Wang Q, Hou Y, Shen S, et al. MicroRNA networks in regulatory T cells. J Physiol Biochem 2014;70:869–75.
29. Rupp LJ, Brady BL, Carpenter AC, De Obaldia ME, Bhandoola A, Bosselut R, et al. The microRNA biogenesis machinery modulates lineage commitment during alphabeta T cell development. J Immunol 2014;193:4032–42.
30. Hou L, Joyce BT, Gao T, Liu L, Zheng Y, Penedo FJ, et al. Blood telomere length attrition and cancer development in the normative aging study cohort. EBioMedicine 2015;2:591–6.
31. Zheng Y, Joyce BT, Colicino E, Liu L, Zhang W, Dai Q, et al. Blood epigenetic age may predict cancer incidence and mortality. EBioMedicine 2016;5:68–73.
32. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics 2011;98:288–95.
33. Triche TJ, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA methylation beadarrays. Nucleic Acids Res 2013;41:e90–e.
34. Davis S, Du P, Bilke S, Triche T, Bootwalla M. methylumi: handle Illumina methylation data. R package version 2012;2. https://www.bioconductor.org/packages/methylumi
35. Teschendorff AE, Jones A, Fiegl H, Sargent A, Zhuang JJ, Kitchener HC, et al. Epigenetic variability in cells of normal cytology is associated with the risk of future morphological transformation. Genome Med 2012;4:24.
36. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics 2010;11:587.
37. Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. Missing value estimation methods for DNA microarrays. Bioinformatics 2001;17:520–5.
38. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 2012;13:86.
39. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analysing methylation data from Illumina's HumanMethylation450 platform. Bioinformatics 2015;32:286–8.
40. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B (Methodological) 1995;57:289–300.
41. Caramuta S, Lee L, Ozata DM, Akcakaya P, Georgii-Hemming P, Xie H, et al. Role of microRNAs and microRNA machinery in the pathogenesis of diffuse large B-cell lymphoma. Blood Cancer J 2013;3:e152.
42. Dedes KJ, Natrajan R, Lambros MB, Geyer FC, Lopez-Garcia MA, Savage K, et al. Down-regulation of the miRNA master regulators Drosha and Dicer is associated with specific subgroups of breast cancer. Eur J Cancer 2011;47:138–50.
43. Jiang Y, Chen J, Wu J, Hu Z, Qin Z, Liu Xa, et al. Evaluation of genetic variants in microRNA biosynthesis genes and risk of breast cancer in Chinese women. Int J Cancer 2013;133:2216–24.
44. Khan S, Greco D, Michailidou K, Milne RL, Muranen TA, Heikkinen T, et al. MicroRNA related polymorphisms and breast cancer risk. PLoS One 2014;9:e109973.
45. Yuan L, Chu H, Wang M, Gu X, Shi D, Ma L, et al. Genetic variation in DROSHA 3′UTR regulated by hsa-miR-27b is associated with bladder cancer risk. PLoS One 2013;8:e81524.
46. Pu X, Roth JA, Hildebrandt MA, Ye Y, Wei H, Minna JD, et al. MicroRNA-related genetic variants associated with clinical outcomes in early-stage non-small cell lung cancer patients. Cancer Res 2013;73:1867–75.
47. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 2012;13:484–92.
48. Zhu X, He F, Zeng H, Ling S, Chen A, Wang Y, et al. Identification of functional cooperative mutations of SETD2 in human acute leukemia. Nat Genet 2014;46:287–93.
49. Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, et al. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. Cell 2005;123:593–605.
50. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 2005;123:581–92.
51. Joshi AA, Struhl K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. Mol Cell 2005;20:971–8.
52. Bartke T, Vermeulen M, Xhemalce B, Robson SC, Mann M, Kouzarides T. Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell 2010;143:470–84.
53. Diaz-Garcia CV, Agudo-Lopez A, Perez C, Lopez-Martin JA, Rodriguez-Peralto JL, de Castro J, et al. DICER1, DROSHA and miRNAs in patients with non-small cell lung cancer: implications for outcomes and histologic classification. Carcinogenesis 2013;34:1031–8.
54. Li X, Tian X, Zhang B, Chen J. Polymorphisms in microRNA-related genes are associated with survival of patients with T-cell lymphoma. Oncologist 2014;19:243–9.
55. Lin J, Horikawa Y, Tamboli P, Clague J, Wood CG, Wu X. Genetic variations in microRNA-related genes are associated with survival and recurrence in patients with renal cell carcinoma. Carcinogenesis 2010;31:1805–12.

56. Pyfferoen L, Mestdagh P, Vergote K, De Cabooter N, Vandesompele J, Lambrecht BN, et al. Lung tumours reprogram pulmonary dendritic cell immunogenicity at the microRNA level. Int J Cancer 2014;135: 2868–77.
57. Okada H, Kohanbash G, Lotze MT. MicroRNAs in immune regulation–opportunities for cancer immunotherapy. Int J Biochem Cell Biol 2010; 42:1256–61.
58. Belver L, de Yebenes VG, Ramiro AR. MicroRNAs prevent the generation of autoreactive antibodies. Immunity 2010;33:713–22.
59. Liang TJ, Qin CY. The emerging role of microRNAs in immune cell development and differentiation. APMIS 2009;117:635–43.
60. Chong MM, Rasmussen JP, Rudensky AY, Littman DR. The RNAseIII enzyme Drosha is critical in T cells for preventing lethal inflammatory disease. J Exp Med 2008;205:2005–17.