Global DNA methylation and one-carbon metabolism gene polymorphisms and the risk of breast cancer in the Sister Study

Lisa A.DeRoo1,2,*, Sophia C.E.Bolick3,*, Zongli Xu1, David M.Umbach4, David Shore5, Clarice R.Weinberg4, Dale P.Sandler1 and Jack A.Taylor1,3,6

1Epidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA, 2Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway NO-5020, 3Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA, 4Biostatistics Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC 27709, USA and 5Westat, Research Triangle Park, NC 27703, USA

*To whom correspondence should be addressed. Tel: +1 919 541 4631; Fax: +1 919 541 2511; Email: taylor@niehs.nih.gov

Global decrease in DNA methylation is a common feature of cancer and is associated with genomic and chromosomal instability. Retrospective case–control studies have reported that cancer patients have lower global methylation levels in blood DNA than do controls. We used prospectively collected samples and a case–cohort study design to examine global DNA methylation and incident breast cancer in 294 cases and a sample of 646 non-cases in the Sister Study, a study of 50 884 women aged 35–74 years who had not been diagnosed with breast cancer at the time of blood draw. Global methylation in DNA from peripheral blood was assessed by pyrosequencing of the LINE-1 repetitive element. Quartiles of LINE-1 methylation levels were associated with the risk of breast cancer in a dose-dependent fashion (P, trend = 0.002), with an increased risk observed among women in the lowest quartile compared with those in the highest quartile (hazard ratio = 1.75; 95% confidence interval 1.19, 2.59).

We also examined 22 polymorphisms in 10 one-carbon metabolism genes in relation to both LINE-1 methylation levels and breast cancer. We found three single-nucleotide polymorphisms in those genes associated with LINE-1 methylation: SLCT19A1 (rs1051266); MTRR (rs10380) and MTHFR (rs1537514), one of which was also associated with breast cancer risk: MTHFR (rs1537514). PON1 (rs757158) was associated with breast cancer but not methylation.

Introduction

Breast cancer is one of the most common cancers in women, with an estimated 226 870 new cases and 39 510 deaths in the USA in 2012 (www.seer.cancer.gov). Both decreased and increased methylation is common in cancer tissue depending on genomic location. Although some gene promoter regions become overmethylated and transcriptionally silenced, the level of global methylation in the cancer genome decreases by 20–60%. This loss of methylation, particularly at long interspersed repeat sequences (LINE-1) and other repetitive elements, is associated with increased genomic instability (1–4). The ~500 000 copies of LINE-1 throughout the human genome (5,6) are normally highly methylated (7,8), and these sequences have been used as a surrogate for estimating global DNA methylation levels (6). In breast tumor tissue, LINE-1 undermethylation is correlated with tumor size, histologic grade and stage of disease (9) and is associated with decreased survival and increased risk of recurrence in breast cancer patients (10). Global undermethylation has been observed even in low-grade tumors, suggesting that demethylation is an early event in breast carcinogenesis (11).

Global methylation differences between cancer cases and controls have been demonstrated using DNA extracted from blood or serum (12). Studies of bladder (13,14) and head and neck cancers (15) have shown that blood DNA from cancer cases has reduced LINE-1 DNA methylation. Although breast cancer case–control studies have not shown differences in LINE-1 methylation of blood DNA (16–19), undermethylation has been shown using other assays of global methylation (16,19,20).

DNA methylation is dependent on DNA methyltransferases that make use of methyl groups donated by S-adenosyl-l-methionine (21). Methyl group availability is dependent on one-carbon metabolism from available folate stores, and polymorphisms in one-carbon metabolism genes affect levels of both folate and DNA methylation (22). Common polymorphisms in genes related to one-carbon metabolism are associated with both breast cancer risk (23–29) and DNA methylation levels in breast tissue of healthy women (24,26).

Retrospective case–control designs are potentially problematic because methylation is measured in samples collected after case diagnosis, so that any observed differences in methylation levels may be due to the disease itself or to treatment for it. Use of a prospective design in which samples are collected before diagnosis is advantageous because it avoids possible postdiagnosis effects. We examined global methylation and incident breast cancer using DNA from prospectively collected blood samples of women enrolled in a large, national cohort study. Global methylation was assessed by pyrosequencing of LINE-1. We also assessed associations between polymorphisms in one-carbon metabolism genes and both LINE-1 methylation levels and breast cancer risk.

Materials and methods

Study population

The Sister Study is a prospective cohort study of environmental and genetic risk factors for breast cancer among 50 884 US and Puerto Rican women (www.SisterStudy.niehs.nih.gov). Eligible participants were between the ages of 35–75 years and had a full or half-sister with breast cancer but had not been diagnosed with the disease themselves. Women completed extensive baseline interviews including family history of cancer, reproductive and medical history and lifestyle and environmental factors. Trained examiners obtained blood samples during a home visit. Enrollment took place in 2003–09 and active follow-up will continue for at least 10 years to ascertain breast cancer and other diagnoses. The study was approved by the Institutional Review Boards of the National Institute of Environmental Health Sciences, National Institutes of Health and the Copernicus Group. All women provided written informed consent.

Study sample

Women in this analysis were selected from the 29 026 participants fully enrolled in the Sister Study by June 2007. Our case–cohort design initially included 328 women who developed incident breast cancer between the time of their blood draw and May 2008 and a random sample of 717 women. Of the random sample, 12 women were among those who developed breast cancer during the study period, resulting in 705 non-cases. LINE-1 methylation assays were conducted on DNA extracted from 1015 whole blood samples and 18 blood clots. Six low-quality samples were excluded, resulting in 1027 samples with high-quality measurements. Methylation levels differ by race and ethnicity (30), so we restricted analyses to non-Hispanic white women, resulting in the exclusion of 87 women and yielding an analysis sample of 294 cases and 646 non-cases. Medical records and pathology reports were used to verify breast cancer diagnosis. When medical records were unavailable (n = 39), self-reported data were used; the accuracy of self-report in our data was high for cancer verification (98%).
DNA extraction and bisulfite conversion

Genomic DNA was extracted using an automated process (Autopure LS; Gentra Systems) and DNA concentrations were determined using a Nanodrop ND-1000 spectrophotometer. Three DNA aliquots from each participant were bisulfite converted in independent “batches” of bisulfite conversion reactions for a total of 49 independent bisulfite conversion batches. Samples of DNA (500ng) were denatured by adding 5.5 μl 2M NaOH and incubating at 37°C for 10 minutes. The cytosine deamination reaction occurred by adding 30 μl freshly prepared 10mM hydroxionone, followed by 520 μl 3M sodium bisulfite, pH 5.0, and incubating the sample at 50°C overnight. The samples were desalted using the Wizard DNA Clean-Up System (Promega) and the resin was washed twice with 80% isopropanol. Column-purified DNA was desulfonated by adding 5.5 μl 3M NaOH for 5 minutes. Samples were neutralized by the addition of 33 μl 10M NH4Cl. The DNA precipitated in ethanol overnight at −20°C. DNA was pelleted by centrifuging for 30 minutes at 4°C, followed by a wash with 70% ethanol and centrifugation at 4°C for 10 minutes. The bisulfite-converted DNA was air-dried and resuspended in 10 μl diethylpyrocarbonate H2O. At least three different members of a set of 10 control samples were prepared in each batch to facilitate adjustment for batch-to-batch variability (batch effects). These control samples included peripheral blood lymphocyte DNA (Promega), the MCF-7 cell line (ATCC) and blood from eight healthy female donors. The Promega peripheral blood lymphocyte DNA was included as a control sample in all batches, whereas other control samples were included in overlapping subsets. Both case and non-case samples were present in each batch, with a relative proportion of roughly 2:1 non-cases to cases.

PCR amplification and pyrosequencing

Bisulfite-converted DNA was amplified with primers designed to recognize a consensus LINE-1 sequence (modified from ref. 6): L-1-F, 5′-TTTTTTGAGTTAGGTGGG-3′; L-1-R, 5′-Biotin-TCTCTAAATAAATACCCAAAA-3′ and L-1-Seq, 5′-AGTATGTTGGGGATAAGTGT-3′ (IDT). The 25 μl PCR reaction mix contained 1x reaction buffer (Qagen), 1.5 mM MgCl2, 800μM dNTPs, 5 pmol of forward and reverse primers, 0.8 μl HotStarTaq polymerase (Qagen) and 1 μl bisulfite-converted DNA. PCR cycling parameters were as follows: 95°C hot start for 15 minutes, 45 cycles of amplification with 95°C for 20 s, 50°C for 20 s, 72°C for 20 s and a final extension at 72°C for 5 minutes. Following PCR, the biotin-labeled DNA product was bound to streptavidin-coated Sepharose beads (GE Healthcare), purified and denatured using 0.2 M NaOH to a single-stranded template. Pyrosequencing primers at 0.3 μmol/l were annealed to the single-stranded template and pyrosequencing was carried out using PSQ HS 96 System (Biotage). Percent methylation at four CpG sites within the amplicon was quantified using the PSQ Software (Biotage).

Genotyping

We selected 22 single-nucleotide polymorphisms (SNPs) from 10 genes in the one-carbon metabolism pathway using the SNPinfo GenePipe tool, which prioritizes SNPs based on previous genome-wide association study P values, linkage disequilibrium and functional prediction scores (Xu et al. 31). We used P values from the Cancer Genetic Markers of Susceptibility (http://cancer.gov/) breast cancer genome-wide association study (32) as input for SNPinfo. Genotyping was conducted by the National Institute of Environmental Health Sciences Molecular Genetics Core Facility using a custom-designed Illumina GoldenGate Genotyping Assay. A total of 20 HapMap trios were genotyped to evaluate parent-parent-child error. Illumina BeadStudio genotyping software (version 1.6.3) was used to call genotypes. Individual genotypes with more than 99.5% call rate were included in the models. The following variables were assessed as possible confounders: number of one-carbon metabolism genotypes and LINE-1 methylation levels were assessed using a linear mixed effect model adjusting for batch and subject as random effects in non-cases. All statistical analyses were performed using SAS 9.13 (Cary, NC).

Results

Cases were more likely than non-cases to be aged 60 years or older and postmenopausal at enrollment, but the two groups were similar with regard to education level, body mass index, ever use of oral contraceptives and number of pregnancies (Table 1). By design, all participants had at least one sister diagnosed with breast cancer, but participants who developed breast cancer were more likely to have had their mother or an additional sister diagnosed as well. For cases, the age at diagnosis was 57.9 years (standard deviation = 9.3 years) and the average time between blood draw and breast cancer diagnosis was 487 days (range: 5–1389). For non-cases, the average time between blood draw and the end of the follow-up period (May 2008 for this analysis) was 902 days (range 351–1675).

The modified proportional hazard analysis revealed an inverse association between quartile of LINE-1 methylation and the risk of breast cancer in a dose-dependent fashion (P trend = 0.002), with an increased risk observed among women in the lowest quartile of methylation compared with those in the highest quartile (HR = 1.75; 95% CI: 1.19, 2.59) (Figure 1). LINE-1 methylation levels were associated with the timing of blood collection relative to diagnosis; for each year closer to diagnosis, methylation levels decreased by an average of 0.20% (P = 0.03).

Of 22 SNPs in 10 genes involved in one-carbon metabolism, 3 were significantly associated with LINE-1 methylation: a non-synonymous polymorphism in exon 2 SLC19A1 (rs1051266, P = 0.013), a gene involved in folate transport; a non-synonymous polymorphism in MTHFR (rs10380, P = 0.045) (Table II) and a 3′-untranslated region polymorphism in MTHFR (rs1537514, P = 0.029). This latter SNP was also one of two SNPs found to be associated with breast cancer risk: MTHFR (rs1537514) was associated with a reduced risk of breast cancer (HR = 0.69; 95% CI: 0.48, 0.99), whereas a synonymous SNP in PON1 (rs757158) was associated with an increased risk of breast cancer (HR = 1.25; 95% CI: 1.01, 1.54).

Discussion

Using a prospective case–cohort study of 940 women, we found statistically significant LINE-1 undermethylation in the peripheral blood DNA of women who developed breast cancer compared with those who remained breast cancer-free throughout our study period. Undermethylation of LINE-1 was present in blood samples collected before clinical detection, consistent with the idea that global demethylation is an early event in breast carcinogenesis. Among cases,
Global DNA methylation and breast cancer risk

undermethylation was greater in women who provided blood closer to the time of diagnosis, indicating that undermethylation of LINE-1 may be a marker of active breast cancer. Our use of samples collected and stored before cancer diagnosis is a particular strength of this study because it avoids a number of potential biases inherent in case–control designs including case–control differences in sample collection and processing and the possibility that the changes observed in cases were introduced by responses to the disease itself or by biopsy, surgery, treatment or other postdiagnostic life events. Pyrosequencing has been established as a precise assay for assessing LINE-1 methylation levels in peripheral blood (35) and a particular strength of our study is that three independent replicate measures were made on each sample, with three independent bisulfite conversions, PCR and pyrosequencing reactions.

Disruption of the usually heavily methylated LINE-1 sites increases the chance of undesired mitotic recombination. Such recombination can lead to the reactivation of transposable elements and their integration at random sites in the genome, resulting in mutagenesis and genomic instability (36). Several studies of global methylation in blood DNA for other cancer types including colon, head and neck, bladder and stomach have found an increased risk of cancer for those in the lowest quantile of global methylation compared with those in the highest (12). Studies of breast cancer; however, have yielded mixed results. Choi et al. (16) found decreased methylation of 5-methyldeoxycytosine in the blood DNA of women with breast cancer compared with controls; likewise both Wu et al. (20) and Cho et al. (19) found decreased methylation of Satellite2. However, Xu et al. (17) found increased global methylation among cases using the luminometric methylation assay. Previous studies of LINE-1 (16–19) did not find an association with breast cancer, but some studies measured LINE-1 in only a small number of cases (<50) (16, 19). In a separate epigenome-wide study of Sister Study blood DNA samples, we have also reported that cases were relatively undermethylated at CpG islands near gene transcription start sites (37).

This study contributes to a growing literature examining methylation in blood DNA showing global undermethylation in the blood DNA of those who develop cancer compared with controls (12). However, little is known about the correlation between global methylation in blood and that in target tissues (17). Evidence based on limited studies supports small tissue-specific differences in global methylation levels (38). In a study comparing DNA global methylation in 40 breast cancer patients, tumor tissue and white blood cells were significantly correlated for LINE-1 methylation and both were relatively undermethylated relative to control blood DNA, although

### Table I. Baseline characteristics of participants

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Non-cases (n = 646), n (%)</th>
<th>Cases (n = 294), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35–39</td>
<td>26 (4.0)</td>
<td>8 (2.7)</td>
</tr>
<tr>
<td>40–49</td>
<td>176 (27.2)</td>
<td>73 (25.5)</td>
</tr>
<tr>
<td>50–59</td>
<td>246 (38.1)</td>
<td>101 (34.3)</td>
</tr>
<tr>
<td>60–69</td>
<td>162 (25.1)</td>
<td>86 (29.3)</td>
</tr>
<tr>
<td>70–74</td>
<td>36 (5.6)</td>
<td>24 (8.2)</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or less</td>
<td>80 (12.4)</td>
<td>44 (15.0)</td>
</tr>
<tr>
<td>Some college</td>
<td>210 (32.6)</td>
<td>76 (25.9)</td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>355 (55.0)</td>
<td>174 (59.1)</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Body mass index</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal or underweight (&lt;24.9)</td>
<td>275 (42.6)</td>
<td>120 (41.0)</td>
</tr>
<tr>
<td>Overweight (25.0–29.9)</td>
<td>212 (32.9)</td>
<td>84 (28.7)</td>
</tr>
<tr>
<td>Obese (≥30.0)</td>
<td>158 (24.5)</td>
<td>89 (30.4)</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Number of births (including stillbirths)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>114 (17.7)</td>
<td>62 (21.1)</td>
</tr>
<tr>
<td>1</td>
<td>96 (14.9)</td>
<td>39 (13.3)</td>
</tr>
<tr>
<td>2</td>
<td>251 (38.9)</td>
<td>102 (34.7)</td>
</tr>
<tr>
<td>3 or more</td>
<td>184 (28.5)</td>
<td>91 (31.0)</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Number of sisters diagnosed with breast cancer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>586 (90.7)</td>
<td>255 (86.7)</td>
</tr>
<tr>
<td>2 or more</td>
<td>60 (9.3)</td>
<td>39 (13.3)</td>
</tr>
<tr>
<td><strong>Mother diagnosed with breast cancer</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>126 (20.0)</td>
<td>79 (27.6)</td>
</tr>
<tr>
<td>Missing</td>
<td>16</td>
<td>8</td>
</tr>
</tbody>
</table>

Fig. 1. Associations between quartile of LINE-1 methylation and the risk of breast cancer. Methylation levels were categorized into quartiles based on the distribution among controls. HRs are adjusted by attained age.

Fig. 1. Associations between quartile of LINE-1 methylation and the risk of breast cancer. Methylation levels were categorized into quartiles based on the distribution among controls. HRs are adjusted by attained age.
the methylation difference between case and control blood cells was not statistically significant (19).

Blood may be more susceptible than other tissues to methylation changes in response to environmental exposures such as chemical pollutants, diet and other lifestyle factors, which may in turn be associated with an increased risk of cancer. These epigenetic alterations may be acquired gradually, resulting in observable differences only after long-term exposure (39). In cross-sectional studies, LINE-1 methylation in blood DNA was lower among smokers (40), inversely associated with arsenic levels in toenails (40) and higher among women with greater physical activity over the life course (41). DNA methylation is a plausible mechanism through which the cumulative effects of environmental exposures may influence cancer risk or it may be a marker of these effects.

One concern is the possibility that differences in the cell populations making up peripheral blood may account for methylation differences between cases and controls. However, LINE-1 measurements in purified granulocyte and lymphocyte samples were not significantly different in one study (40). Fifty CpGs shown previously to be differentially methylated according to leucocyte cell type (42) were not found to be significant predictors of breast cancer in a case–cohort analysis of our samples (37), providing some reassurance that our findings were not due to shifts in leucocyte subpopulations.

We examined four polymorphisms of methyletetrahydrofolate reductase (MTHFR) in relation to LINE-1 methylation levels and the risk of breast cancer. MTHFR plays a key role in one-carbon metabolism by converting 5,10-methylenetetrahydrofolate to 5-methyl tetrahydrofolate, the primary circulating form of folate. MTHFR is a rate-limiting enzyme in the pathway and thus regulates the levels of methyl groups available for methylation reactions. We found that the SNP in the 3′-untranslated region of MTHFR (rs1537514) was associated with increased LINE-1 methylation and reduced risk of breast cancer. Based on predictions in SNPinfo (31), this SNP is also a putative microRNA-binding site. We did not find significant associations for the C/T nucleotide polymorphism at position 677 that was previously associated with lower genome-wide methylation levels (22) and altered MTHFR enzyme activity (43,44) nor the A/C polymorphism at nucleotide 1298 that was previously associated with MTHFR enzyme activity (45).

We saw a significant association in non-cases between LINE-1 methylation and rs10308, a missense polymorphism (His595Tyr) in methionine synthase reductase (MTRR), a gene involved in regeneration of methionine synthase required for one-carbon metabolism; this SNP has been associated with risk of pancreatic cancer (46). We also found a significant association between LINE-1 methylation and rs1051266, a missense polymorphism (His27Arg) in the solute carrier family 19 (SLC19A1), a gene involved in folate uptake and transport; this SNP has been associated with risk of colorectal adenomas (47). Finally, we found a significant association between breast cancer risk and a SNP (rs757158) in a putative transcription factor-binding site upstream of the paraoxonase 1 (PON1) gene although previous studies found no association between this SNP and breast cancer risk (48–50).

Our study had some potential limitations. Although LINE-1 undermethylation was observed in blood collected before breast cancer diagnosis, our relatively short follow-up time (average of 1.3 years before diagnosis for cases) limited our ability to determine how far in advance of clinical detection these differences are observable. Our sample size for the analysis of polymorphisms was small. Because of the number of tests involved, our findings on one-carbon metabolism genotypes are particularly subject to type 1 (false positive) error. We chose to report P values without correction for multiple comparisons because the SNPs in this study were selected based on previous genome-wide association study P values, linkage disequilibrium and functional prediction scores. These SNP findings should be interpreted with caution because none of the associations would have reached the same level of significance after adjustment for multiple comparisons. Our findings were based on women with a family history of cancer and may not generalize to other women; however, such a history would seem unlikely to influence the relationships between SNPs and methylation or methylation and breast cancer.

In conclusion, using prospectively collected blood samples, we found a statistically significant association between LINE-1 undermethylation and an increased risk of breast cancer in non-Hispanic white women in the Sister Study. We also found some evidence that particular SNPs in one-carbon metabolism genes influence both LINE-1 methylation levels and breast cancer status. Undermethylation of LINE-1 was present in blood samples collected before clinical detection, consistent with the idea that global undermethylation is associated with increased cancer risk.
Global DNA methylation and breast cancer risk

Funding
Intramural Program of the National Institutes of Health, National Institute of Environmental Health Sciences (Z01 ES044005, Z01 ES049033).

Acknowledgements
We thank Dr Laurannell Burch, Re Bai and Raymond Thomas of the National Institute of Environmental Health Sciences Molecular Genetics Core for technical assistance and Karen Baldwin, Gleta Carwell and Alexandra J. White of the National Institute of Environmental Health Sciences Laboratory of Molecular Genetics for assistance with the laboratory assays. We also thank Drs Bonnie Joubert and Donna D. Baird for helpful comments on an earlier draft of this paper.

Conflict of Interest Statement: None declared.

References


Received July 26, 2013; revised September 11, 2013; accepted October 4, 2013