Intragenic ATM Methylation in Peripheral Blood DNA as a Biomarker of Breast Cancer Risk

Kevin Brennan1, Montserrat Garcia-Closas4,6, Nick Orr6, Olivia Fletcher4, Michael Jones6, Alan Ashworth4, Anthony Swerdlow6, Heather Thorne on behalf of KConFab Investigators7, Elio Riboli3, Paolo Vineis2,8, Dimitrios Trichopoulos13,14, Rudolf Kaaks15, Kay-Tee Khaw16, Robert Brown1,5, and James M. Flanagan1

Abstract

Few studies have evaluated the association between DNA methylation in white blood cells (WBC) and the risk of breast cancer. The evaluation of WBC DNA methylation as a biomarker of cancer risk is of particular importance as peripheral blood is often available in prospective cohorts and easier to obtain than tumor or normal tissues. Here, we used prediagnostic blood samples from three studies to analyze WBC DNA methylation of two ATM intragenic loci (ATMmvp2a and ATMmvp2b) and genome-wide DNA methylation in long interspersed nuclear element-1 (LINE1) repetitive elements. Samples were from a case–control study derived from a cohort of high-risk breast cancer families (KConFab) and nested case–control studies in two prospective cohorts: Breakthrough Generations Study (BGS) and European Prospective Investigation into Cancer and Nutrition (EPIC). Bisulfite pyrosequencing was used to quantify methylation from 640 incident cases of invasive breast cancer and 741 controls. Quintile analyses for ATMmvp2a showed an increased risk of breast cancer limited to women in the highest quintile [OR, 1.89; 95% confidence interval (CI), 1.36–2.64; \( P = 1.64 \times 10^{-5} \)]. We found no significant differences in estimates across studies or in analyses stratified by family history or menopausal status. However, a more consistent association was observed in younger than in older women and individually significant in KConFab and BGS, but not EPIC. We observed no differences in LINE1 or ATMmvp2b methylation between cases and controls. Together, our findings indicate that WBC DNA methylation levels at AT could be a marker of breast cancer risk and further support the pursuit of epigenome-wide association studies of peripheral blood DNA methylation. Cancer Res; 72(9); 2304–13. ©2012 AACR.

Introduction

Dysregulation of epigenetic modification in tumor DNA such as hypermethylation of CpG islands at the promoters of hundreds of genes and global reduction of 5-methylcytosine (5-mC) levels has been observed in almost every cancer type (1). However, the roles of epigenetic modifications as risk factors for cancer and early disease biomarkers are yet to be determined (2).

Epigenetic changes such as DNA methylation could be a biologic indicator of lifetime accumulation of environmental exposures including ageing (3, 4), hormones (5, 6), ionizing radiation (7), alcohol (8), smoking (9, 10), and traffic particles (11). Alternatively, epigenetic changes might modify the effects of genetic susceptibility loci through genotype–epigenotype interactions, via in cis allele–specific methylation (12). Some methylation changes may also reflect parental and early-life exposures that are particularly difficult to measure in epidemiologic studies of adults (2, 13). Although DNA methylation profiles are often tissue- and cell-specific, recent data indicate that epigenetic traits in white blood cells (WBC) are promising candidate risk markers for solid tumors (14–16). Evaluating WBC DNA methylation as a biomarker of risk is of particular interest because peripheral blood DNA is comparatively accessible and is readily available in many large prospective studies.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Corresponding Author: James M. Flanagan, Epigenetics Unit, Division of Cancer, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, 4th Floor IRDB, Hammersmith Campus, Du Cane Road, London W12 0NN, UK. Phone: 020-75941804; Fax: 020-75942129; E-mail: j.flanagan@imperial.ac.uk
doi: 10.1158/0008-5472.CAN-11-3157

©2012 American Association for Cancer Research.

2304 Cancer Res; 72(9) May 1, 2012

American Association for Cancer Research

Downloaded from cancerres.aacrjournals.org on September 8, 2017. © 2012 American Association for Cancer Research.
ATM Intragenic Methylation as a Breast Cancer Risk Marker

Epidemiologic studies. Previous reports of WBC DNA methylation and cancer risk include studies of global DNA methylation levels in repeat regions across the genome (e.g., LINE1, Alu) or 5-mC content in genomic DNA (16–20): studies of genespecific DNA methylation levels in candidate genes (14, 21–29), and genome-wide DNA methylation microarray studies (15, 30–32). Although these studies provide enticing findings, most included relatively small study populations and/or used samples collected after diagnosis, thus raising concerns about reverse causality and the potential confounding influences of active disease or treatment on DNA methylation in blood (25).

Most research on DNA methylation in cancer has focused on gene promoter CpG islands (CGI). Reasons for this include the historical identification of methylation at the promoters of tumor suppressor genes in a wide variety of cancers and the mechanistic association of methylation with transcriptional repression at these loci (33). However, recent data suggest that methylation in regions around CpG islands or "shores" and intragenic sequences also appears to be important in tissue-specific expression and may be an important contributor to interindividual variation in gene expression (34). For example, using a differential methylation hybridization microarray analysis of candidate genes in WBC DNA, we showed that the majority of methylation variability in 55 genes was associated with intragenic repetitive elements (14). In addition, an intragenic differentially methylated region (DMR) within the ATM gene correlated with gene expression and was significantly more methylated in postdiagnostic blood samples from 190 bilateral breast cancer cases compared with 190 controls (14). However, global DNA methylation as measured by long interspersed nuclear element-1 (LINE1) repetitive element methylation was not significantly associated with risk. In the present study, we followed up these findings to test the hypothesis that prediagnostic intragenic and global LINE1 repetitive element methylation is prospectively associated with the risk of breast cancer using 3 studies with prediagnostic blood samples: the Kathleen Cuningham Foundation Consortium for Research into Familial Breast cancer (KConFab) study, a prospective cohort of families at high risk of breast cancer; and 2 general population prospective cohorts: the Breakthrough Generations Study (BGS) and the European Prospective Investigation into Cancer and Nutrition (EPIC).

Materials and Methods

Study populations

Study participants were drawn from 3 large studies with blood samples collected before breast cancer diagnosis (Table 1). All contributing studies have appropriate ethical approval for sample and data collection. The first study was performed by KConFab (35). From 1997 to May 2011, KConFab collected peripheral blood samples from 12,747 members of 1,395 families (~8.8 samples per family) in Australia and New Zealand. Families had an average of 3 verified (5.4 unverified) breast cancers per family. At the time of sampling, there were 12,747 blood samples collected from breast cancer family members of whom 4,305 had a verified prior cancer diagnosis. All remaining healthy subjects were screened for subsequent incident cancer cases which were identified and confirmed in 1 of 4 ways, primarily by clinical pathology report, then doctor’s notes, cancer council registry verification, or death certificate. Incident cases of invasive breast cancer were selected for this study from all individuals with a breast cancer diagnosis more than 1 month after blood sample collection (n = 171). Five cases of non-white ethnicity were excluded from the analyses resulting in 166 invasive cases that were compared with 225 healthy unrelated controls without a family history of breast cancer drawn from “best friends” of subjects enrolled in KConFab. Incident breast cancer cases had blood samples taken, on average, 45 months before diagnosis (range, 1–140 months).

Information on breast cancer risk factors including hormonal and reproductive factors, cigarette smoking, and alcohol drinking was collected from questionnaires at enrollment. In addition, pathology data, including grade (12% grade I, 33% grade II, 40% grade III), nodal status (37% node positive), estrogen receptor (ER; 63% ER positive, 24% ER negative), progesterone receptor (PR; 55% PR positive), HER2 status (27% HER2 positive), and BRCA1/2 mutation status (23% BRCA1 mutant, 14% BRCA2 mutant, 63% non-BRCA1/non-BRCA2), were available for breast cancer cases.

The second study was based on the BGS, a large general population cohort consisting of approximately 110,000 women enrolled in the United Kingdom from 2003 to 2011 (36). Study participants were sampled from a nested case-control study of all incident cases of breast cancer diagnosed in the BGS up to June 2010 and controls individually matched on recruitment source, year of completion of the baseline questionnaire at enrollment, ethnicity (white only), date of birth within 12 months, availability of blood sample, and duration that the blood sample was in the mail. Breast cancer cases were self-reported in a follow-up questionnaire about 2.5 years after enrollment or notified by study participants by phone or letter. Self-reported diagnoses were confirmed through an electronic linkage with England/Wales/Scotland/Northern Ireland cancer registrations (or by the general practitioner for a small number of cases who could not be successfully linked). Checks against U.K. cancer registrations were also made for those BGS participants known to have died by the time of the 2.5-year follow-up or who otherwise failed to respond to the follow-up but had given permission for such follow-up. A random sample of 257 case-control pairs of 534 pairs identified as of December 2010 was selected for the methylation study. Four controls who were subsequently found to have had prevalent breast cancer at study entry, 3 cases whose blood was collected after diagnosis, and one case with ductal carcinoma in situ (DCIS) were excluded from the analyses, resulting in a total of 253 cases and 253 controls available for analyses. Blood samples, from incident cases, were taken on average 18 months before diagnosis (range, 0.03–59; Table 1). Extensive information on breast cancer risk factors was collected from a baseline questionnaire at enrollment. Pathology information from all cases was not available at the time of these analyses, available data represent only 25% of cases and included morphology (78% ductal, 12% lobular, 10% other), grade (14% grade I, 50% grade II, 36% grade III), nodal status (33% node positive), ER status (64% ER positive, 16% ER negative), and HER2 status (19% HER2 positive, 40% HER2 negative, 41% HER2 unknown).

www.aacjournals.org Cancer Res; 72(9) May 1, 2012 2305
positive). An additional random sample of 92 participants in the BGS was selected from women enrolled in 2004 with 2 blood samples/questinnaires collected at baseline and first follow-up (~6 years after baseline). Additional inclusion criteria included the following: 35–84 years at enrollment, free of breast cancer before second blood collection, not know to have a relative in the study, blood samples received at processing laboratory less than 1 day after collection, expected amount of blood receipt at the laboratory, no reported problems at laboratory less than 1 day after collection, expected amount of blood collection or processing (e.g., lipemic, hemolyzed, clotted samples), and time between the receipt of each sample between 5.5 and 6.5 years. Paired samples from these women were analyzed to evaluate the stability of the ATMmvp2a and LINE1 methylation markers over time.

The third study was provided by the EPIC cohort, a large general population cohort consisting of about 520,000 individuals with standardized lifestyle and personal history questionnaires, anthropometric data and blood samples collected for DNA extraction (37). Study participants were sampled in 2 groups including a group of premenopausal women (145 cases and 145 controls) and postmenopausal women (139 cases and 146 controls), with menopausal status defined at the time of blood collection. Controls were individually matched on age at baseline, recruitment centre, and date and time of blood collection.

Pathology information on cases included morphology (73% ductal, 14% lobular, 13% other), grade (15% grade I, 65% grade II, 20% grade III), stage, ER (79% ER positive, 21% ER negative), PR (62% PR positive, 38% PR negative), and HER2 status (20% HER2 positive).

Laboratory methods
DNA samples were extracted from whole blood using Qiagen DNA Blood Mini Kits in KConFab. DNA samples from BGS and EPIC were extracted from buffy coats using DNA Blood Mini Kits (Qiagen), except for 29 cases and 15 controls in BGS extracted using Nucleon Genomic DNA Extraction Kit (Tepnel, Life Sciences). Five hundred nanograms of DNA (KConFab) or 250 ng (BGS and EPIC studies) from each subject was bisulfite-converted using EZ–96 DNA Methylation-Gold kit according to the manufacturer’s protocol (Zymo Research). Methylation analysis of LINE1 was analyzed using commercially available LINE1 primers (Qiagen). Primers and PCR conditions for ATMMvp2a and ATMMvp2b regions were as described previously (14). Methylation values were calculated as an average of all high-quality CpG sites (determined as “passed” by the quality control thresholds within the Pyro Q-CpG Software; Qiagen). The Pyro Q-CpG Software has inbuilt overall quality assessment for each sample which flags any sequence that deviates from the expected pattern. Any sample failing quality control was removed from the analysis. The number of samples failing in each assay were ATMMvp2a (55 of 1,436 subjects), ATMMvp2b (56 of 1,436 subjects), and LINE1 (87 of 1,436 subjects). In addition, a commercially available fully methylated genomic DNA sample was used as a positive control (Zymo Research) and in-house whole genome amplified genomic DNA (Genomiphi, GE Healthcare) used as an unmethylated negative control. The percentage of cells with methylated DNA at each of the loci was calculated as the average of 3 (ATMMvp2a) or 4 (LINE1 and ATMMvp2b) CpG sites and was

Table 1. Characteristics of the study populations

<table>
<thead>
<tr>
<th></th>
<th>KConFab Controls</th>
<th>KConFab Cases</th>
<th>BGS Controls</th>
<th>BGS Cases</th>
<th>EPIC Controls</th>
<th>EPIC Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>225</td>
<td>166</td>
<td>253</td>
<td>253</td>
<td>291</td>
<td>248</td>
</tr>
<tr>
<td>Age at blood draw, mean (range), y</td>
<td>60 (33–83)</td>
<td>50 (21–91)</td>
<td>54 (23–82)</td>
<td>54 (23–82)</td>
<td>52 (33–76)</td>
<td>52 (33–76)</td>
</tr>
<tr>
<td>Family history, n (%)</td>
<td>0 (0)</td>
<td>166 (100)</td>
<td>51 (20)</td>
<td>69 (27)</td>
<td>24 (14)</td>
<td>23 (18)</td>
</tr>
<tr>
<td>Yes</td>
<td>225 (100)</td>
<td>0 (0)</td>
<td>201 (80)</td>
<td>183 (73)</td>
<td>144 (86)</td>
<td>108 (82)</td>
</tr>
<tr>
<td>No</td>
<td>37 (17)</td>
<td>101 (65)</td>
<td>135 (61)</td>
<td>134 (62)</td>
<td>145 (50)</td>
<td>127 (50)</td>
</tr>
<tr>
<td>Prenopausal</td>
<td>183 (83)</td>
<td>54 (35)</td>
<td>87 (39)</td>
<td>83 (38)</td>
<td>146 (50)</td>
<td>121 (50)</td>
</tr>
<tr>
<td>Postmenopausal</td>
<td>45 (1–140)</td>
<td>18 (1–59)</td>
<td>55 (24–108)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time from blood collection to diagnosis, mean (range), mo</td>
<td>52 (29–88)</td>
<td>55 (23–84)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Cases were individually matched to controls for recruitment source, year of completion of the baseline questionnaire at enrolment, ethnicity, availability of blood sample, date of birth within 12 months, and duration that the blood sample was in the mail.**

**Cases and controls were selected within strata of menopausal status (pre and post) and individually matched on age, recruitment centre, and date and time of blood collection.**

---

Published OnlineFirst February 28, 2012; DOI: 10.1158/0008-5472.CAN-11-3157

Downloaded from cancerres.aacrjournals.org on September 8, 2017. © 2012 American Association for Cancer Research.
used as the measure of methylation for each subject. On the basis of previous experimental results, the range for a typical assay is 90% to 98% for the positive control and 1% to 6% for the negative control. Further quality assurance was conducted with blinded duplicate samples (12 pairs, 2 pairs within each plate) with median differences of 3.5%, 4.1%, and 1.8% in the BGS duplicates and 2.1%, 1.7%, and 1.3% in the positive and negative controls on each plate for ATMmvp2a, ATMmvp2b, and LINE1, respectively. The intraclass correlation (ICC) for the 12 blinded duplicates was 0.56 [95% confidence interval (CI), 0.26–0.96], 0.37 [95% CI, 0–0.76], and 0 [95% CI, 0–0.61] for ATMmvp2a, ATMmvp2b, and LINE1, respectively. It should be noted that the low ICCs for ATMmvp2b and LINE1 are due to a relatively low between subject variation, rather than large assay variation.

Statistical analysis
The Wilcoxon test for matched pairs was used for BGS and EPIC and the Mann–Whitney U test was used for KConFab, to test for differences in median methylation levels between cases and controls. Levels of methylation across studies were standardized using Z-scores that were categorized in quintiles on the basis of their distribution in the combined control population. Logistic regression was used to estimate ORs and 95% CIs for individuals in the second, third, fourth, and fifth methylation quintiles, compared with individuals in the first (lowest) quintile. Analyses of combined data from all studies were adjusted by age in 5-year categories and study. Age at blood draw, age at menarche, parity, age at menopause, alcohol consumption, body mass index, oral contraceptive and hormone replacement use, and family history of breast cancer were considered as potential confounders. Analyses were stratified by age at blood drawn using tertiles of the combined control population (21–49, >49–59, >59–91), family history, and time from blood collection to diagnosis to evaluate effect modification by these variables. Estimates from conditional logistic regression models for individually matched pairs in BGS (n = 241 pairs after quality control) and EPIC (n = 221 pairs after quality control) were similar to estimates from unconditional logistic models adjusted or unadjusted by matching factors. Only findings from the unconditional logistic analyses are presented to avoid loss of data from exclusion of pairs with one member excluded because of missing methylation data or other reasons (see Study populations). Heterogeneity of estimates by study was tested by including an interaction term for the biomarker and an indicator variable for study in the logistic model. Random-effect meta-analyses of estimated ORs from all studies in this report and a previously published study were conducted in R using the "metafor" package (38). Polytomous logistic regression models with categories of methylation levels as the outcome variable were used to test for associations between methylation levels and the breast cancer risk factors specified above, adjusted for age. B-spline quadratic logistic regression models fitted in the "bs" R package were used to explore the relationship between continuous measures of methylation levels and breast cancer risk. All statistical tests were conducted using R (v 2.12.0).

Results
For the ATMmvp2a locus, we observed significantly higher median methylation in the familial samples from KConFab (81.8% vs. 76.9%, \( P = 4.87 \times 10^{-10} \); Table 2) and marginally higher median methylation in the population-based cases from BGS than in controls (76.8% vs. 76.4%, \( P = 0.02 \)). In the EPIC cohort, we observed no significant differences in median methylation levels in cases compared with controls (75.7% vs. 76.1%, \( P = 0.40 \)). In BGS and KConFab, we observed an upward shift in the distribution of methylation in cases compared with controls, which was not observed in EPIC (Supplementary Fig. S1). No significant

<table>
<thead>
<tr>
<th>Table 2. Distribution of methylation levels in ATMmvp2a, ATMmvp2b, and LINE1 in breast cancer cases and controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assay</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>ATMmvp2a</td>
</tr>
<tr>
<td>EPIC</td>
</tr>
<tr>
<td>KConFab</td>
</tr>
<tr>
<td>ATMmvp2b</td>
</tr>
<tr>
<td>EPIC</td>
</tr>
<tr>
<td>KConFab</td>
</tr>
<tr>
<td>LINE1</td>
</tr>
<tr>
<td>EPIC</td>
</tr>
<tr>
<td>KConFab</td>
</tr>
</tbody>
</table>

**NOTE:** Bold signifies \( P < 0.05 \).
Abbreviation: IQR, interquartile range.

\(^{a}\)Differences in numbers of cases and controls within each study with total numbers are due to missing data (failed quality control) on methylation markers.

\(^{b}\)Wilcoxon matched pairs test for BGS and EPIC and Mann–Whitney U test for KConFab.
differences were found for methylation levels at the ATMmvp2b locus (131 bp downstream from ATMmvp2a) or LINE1 in any of the studies (Table 2). Quartzine analyses for the ATMmvp2a locus, adjusted by age at blood collection in 5-year categories, showed a significantly increased risk of breast cancer for women in the highest quintile compared with the lowest quintile in the BGS and KConFab studies, but not in EPIC (Table 3). Further adjustment by age as continuous variable and conditional logistic analyses for paired samples individually matched in BGS and EPIC showed similar results (Supplementary Table S1). Analyses of combined data from all studies adjusting by study and age at blood collection indicated that women in the highest quintile (>6.3% methylation above study mean methylation) were at 1.9-fold increased risk of breast cancer compared with women in the lowest quintile (OR, 1.89; 95% CI, 1.36–2.64; P = 1.64 × 10^-3; Table 3). While the overall difference in median levels between cases and controls was small (1.1%), the difference in median methylation between the highest quintile (86%) and lowest (65%), where the association with cancer status is observed, was large (21%). A quadratic B-spline regression model of continuous levels of methylation at ATMmvp2a and breast cancer risk confirmed a threshold association, rather than a linear association, between methylation levels and breast cancer risk (Supplementary Fig. S2).

Study-adjusted analyses stratified by age at blood collection suggested a weaker ATMmvp2a risk association when methylation was measured in samples collected from women older than 59 years (Table 4). Similar results were obtained when stratified by age at diagnosis (data not shown). However, age-specific estimates within study revealed that the weaker association was driven by the EPIC cohort that showed no increased risk in this age subgroup (Supplementary Fig. S3). Overall analyses showed some evidence for heterogeneity of estimates between studies (Table 3; P value for test for heterogeneity = 0.07). This evidence was limited to women in the older age group and there was no evidence for study

Table 3. Association between methylation levels in ATMmvp2a and breast cancer risk in BGS, EPIC, KConFab, and combined analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Quintile</th>
<th>Methylation range</th>
<th>Controls</th>
<th></th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Freq.</td>
<td>N</td>
<td>Freq.</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>BGS</td>
<td>Q1</td>
<td>3.4%–68.0%</td>
<td>50</td>
<td>0.20</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>68.0%–74.1%</td>
<td>40</td>
<td>0.16</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>74.1%–77.6%</td>
<td>54</td>
<td>0.22</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>77.6%–81.0%</td>
<td>56</td>
<td>0.23</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>81.0%–91.7%</td>
<td>48</td>
<td>0.19</td>
<td>79</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>248</td>
<td></td>
<td></td>
<td>249</td>
</tr>
<tr>
<td>EPIC</td>
<td>Q1</td>
<td>53.6%–69.7%</td>
<td>60</td>
<td>0.21</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>69.7%–74.8%</td>
<td>61</td>
<td>0.22</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>74.8%–78.6%</td>
<td>58</td>
<td>0.19</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>78.6%–82.4%</td>
<td>49</td>
<td>0.17</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>82.4%–97.5%</td>
<td>55</td>
<td>0.19</td>
<td>51</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>283</td>
<td></td>
<td></td>
<td>235</td>
</tr>
<tr>
<td>KConFab</td>
<td>Q1</td>
<td>19.0%–70.2%</td>
<td>38</td>
<td>0.18</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>70.2%–75.4%</td>
<td>47</td>
<td>0.22</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>75.4%–79.1%</td>
<td>36</td>
<td>0.17</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>79.1%–83.0%</td>
<td>43</td>
<td>0.20</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>83.0%–100%</td>
<td>46</td>
<td>0.22</td>
<td>67</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>210</td>
<td></td>
<td></td>
<td>156</td>
</tr>
<tr>
<td>Combined</td>
<td>Q1</td>
<td>71.3%–6.5%</td>
<td>148</td>
<td>0.20</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>Q2</td>
<td>6.5%–12.2%</td>
<td>148</td>
<td>0.20</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>12.3%–2.5%</td>
<td>148</td>
<td>0.20</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>Q4</td>
<td>2.5%–6.3%</td>
<td>148</td>
<td>0.20</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>Q5</td>
<td>6.3%–23.3%</td>
<td>149</td>
<td>0.20</td>
<td>197</td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>741</td>
<td></td>
<td></td>
<td>640</td>
</tr>
</tbody>
</table>

NOTE: Bold, statistically significant results, P < 0.05.

aCutoff values determined by quintiles of methylation range.

bORs within each study are adjusted by 5-year age categories, with further adjustment by menopausal status in EPIC to account for stratified sampling. ORs in combined analyses are adjusted by 5-year age categories and study (with EPIC defined by 2 categories of menopausal status to account for stratified sampling).

Published OnlineFirst February 28, 2012; DOI: 10.1158/0008-5472.CAN-11-3157
Table 4. Association between methylation levels in ATMmvp2a and breast cancer risk in combined analysis stratified by age at blood draw

| Age range, y | Quintiles | Controls | | | Cases | | | Case proportions by study (K, B, E) | OR* (95% CI) | P |
|-------------|-----------|---------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 21–49       | Q1        | 45      | 0.21            | 44               | 0.17            | 0.20, 0.27, 0.52 | 1.00            | 0.007           |                   |                  |
|             | Q2        | 51      | 0.23            | 39               | 0.15            | 0.28, 0.26, 0.46 | 0.68 (0.37–1.25) | 0.21            |
|             | Q3        | 44      | 0.20            | 41               | 0.16            | 0.27, 0.32, 0.41 | 0.94 (0.51–1.72) | 0.83            |
|             | Q4        | 38      | 0.18            | 44               | 0.17            | 0.43, 0.18, 0.39 | 1.00 (0.54–2.17) | 0.99            |
|             | Q5        | 39      | 0.18            | 90               | 0.35            | 0.40, 0.33, 0.27 | 2.07* (1.16–3.68) | 0.01            |
| Total       |          | 217     | 0.23, 0.28, 0.38 | 258              | 0.23, 0.28, 0.38 |                   |                  |
| >49–59      | Q1        | 54      | 0.20            | 27               | 0.13            | 0.19, 0.41, 0.41 | 1               |
|             | Q2        | 55      | 0.20            | 43               | 0.21            | 0.05, 0.49, 0.47 | 1.57 (0.84–2.91) | 0.15            |
|             | Q3        | 62      | 0.23            | 48               | 0.24            | 0.19, 0.38, 0.44 | 1.56 (0.85–2.86) | 0.15            |
|             | Q4        | 50      | 0.19            | 33               | 0.16            | 0.09, 0.48, 0.42 | 1.26 (0.66–2.41) | 0.48            |
|             | Q5        | 49      | 0.18            | 53               | 0.26            | 0.19, 0.49, 0.32 | 2.25* (1.21–4.17) | 0.01            |
| Total       |          | 270     | 0.14, 0.45, 0.41 | 204              | 0.14, 0.45, 0.41 |                   |                  |
| >59–91      | Q1        | 49      | 0.19            | 33               | 0.19            | 0.18, 0.36, 0.45 | 1.00            |
|             | Q2        | 42      | 0.17            | 31               | 0.17            | 0.10, 0.48, 0.42 | 1.00 (0.51–1.97) | 0.99            |
|             | Q3        | 42      | 0.17            | 26               | 0.15            | 0.08, 0.62, 0.31 | 0.84 (0.45–1.58) | 0.61            |
|             | Q4        | 60      | 0.24            | 54               | 0.31            | 0.26, 0.53, 0.21 | 0.80 (0.45–1.45) | 0.50            |
|             | Q5        | 61      | 0.24            | 54               | 0.31            | 0.39, 0.43, 0.19 | 1.39* (0.87–2.26) | 0.27            |
| Total       |          | 254     | 0.23, 0.47, 0.50 | 178              | 0.23, 0.47, 0.50 |                   |                  |

*NOTE: Bold, statistically significant results, P < 0.05.

*Study proportions for KConFab (K), BGS (B), and EPIC (E) contributing to each quintile in each age group.

*ORs within each age group are adjusted by study. Test for heterogeneity of effects by age group in combined analysis, P = 0.3109 for Q5 versus Q1.

*P values for study heterogeneity for ORs comparing Q5 versus Q1 within age subgroups: 21–49 years (P = 0.51), 50–59 years (P = 0.73), and 60–91 years (P = 0.09).

heterogeneity within the younger age subgroups [P value for study heterogeneity by age subgroups 21–49 (P = 0.51), 50–59 (P = 0.72), 60–91 years (P = 0.09)]. We observed a significant association between ATMmvp2a methylation levels and increasing age at blood collection in controls (Spearman’s r = 0.15, P = 0.0015), but not in cases (Spearman’s r = –0.02, P = 0.43), that was most significant in the EPIC cohort (r = 0.11, P = 0.007) compared with KConFab (r = 0.06, P = 0.26) and BGS (P = 0.02, P = 0.40). This underlying age association may account for the apparent cross-over risk association with ATM methylation by age at blood collection seen in EPIC (Supplementary Fig. S3). Analyses by menopausal status at blood collection, adjusted by study and age, showed similar risk estimates for pre- and postmenopausal women (Supplementary Table S2).

Adjustment by breast cancer risk factors (age at menarche, menopausal status at blood drawn, parity, age at menopause, alcohol consumption, body mass index, oral contraceptive and hormone replacement use, and family history of breast cancer) did not result in appreciable changes in relative risk estimates for any of the markers across the 3 studies in this report. Consistently, these risk factors were not significantly associated with ATMmvp2a methylation levels in any of the 3 control populations (data not shown). In the familial cases from KConFab, we found no significant associations between ATMmvp2a methylation levels and BRCA1/2 mutation status, tumor pathology (morphology, grade, nodal, ER, PR, and HER2 status; data not shown). Similarly, in EPIC and BGS samples, we found no significant association with available data on tumor pathology (data not shown).

Analyses stratified by time from blood collection to diagnosis ([< 1 vs > 1 year) showed no significant differences in effect estimates using the combined data (Supplementary Table S3). Consistently, linear regression analyses of methylation levels and time from blood collection to diagnosis, adjusted by age, showed no significant associations (KConFab, P = 0.97; BGS, P = 0.10; EPIC, P = 0.28), over the range of 0 to 11 years studied here.

A biomarker for risk that is observed only at one time point (as is the case with most case–control studies) would ideally be stable over time. We measured the stability of the ATMmvp2a marker and LINE1 by assessing methylation in a control population where blood samples were taken 6 years apart from the same individuals in the BGS cohort (n = 92 pairs). Using conditional logistic regression, we observed no significant change in either ATMmvp2a (median change = 0.19%, P = 0.51) or LINE1 (median change = 0.27%, P = 0.69) over 6 years. The ATM variation between individuals is much larger than within individuals at 2 time points (ICC = 0.57; between time points correlation, R² = 0.79) and there are no significant
differences between measurements in the 2 time points (paired \( t \) test, \( P = 0.24 \); signed test for H0: median difference = 0, \( P = 0.74 \)). These data show that the ATM methylation is stable for at least 6 years.

We previously reported a quartile analysis of the ATMmp2b region in bilateral patients with breast cancer who showed a 3-fold higher chance of being a breast cancer case compared with healthy controls in the highest quartile (14). We have now reanalyzed these data for both loci (ATMmp2a and ATMmp2b) using a quintile analysis for direct comparison with the present study. This analysis shows that women in the highest quintile of methylation had increased bilateral breast cancer risk for both ATMmp2a (age-adjusted OR, 1.90; 95% CI, 1.00–3.62; \( P = 0.05 \)) and ATMmp2b (age-adjusted OR, 3.07; 95% CI, 1.58–5.93; \( P = 8.8 \times 10^{-5} \); Supplementary Table S4). A meta-analysis of OR estimates including the 3 additional studies and the study of bilateral breast cancer cases, suggested that the methylation levels in the ATMmp2a marker are associated with a 1.89-fold OR (95% CI, 1.27–2.82; \( P = 1.7 \times 10^{-5} \), random-effects model) for women in the fifth quintile compared with the first quintiles (Fig. 1), with no significant heterogeneity between studies (\( P = 0.15 \)).

**Discussion**

Our findings indicate that high levels of methylation in the ATM DMR might be a biomarker of breast cancer risk. To our knowledge, this is the first report to identify a significant association between breast cancer risk and gene-specific methylation in WBC DNA measured in prediagnostic blood samples from cases in prospective cohorts and using pyrosequencing which is a highly quantitative method.

Previous studies investigating peripheral blood methylation variability in relation to cancer risk have focused mainly on rare epimutations or genome-wide methylation levels (25, 39). Only a few studies have investigated gene-specific DMRs and all of these have been carried out in small retrospective studies or used nonquantitative methods (26–29). This includes our initial report of the association between ATM hypermethylation in WBC DNA and increased breast cancer risk using retrospectively collected blood samples from bilateral patients with breast cancer (14). While findings for ATMmp2a were consistent with those in this report, the strong association found for ATMmp2b in the bilateral study was not replicated. Differences in findings could be due to the different designs and study populations, for example, use of pre- versus post-diagnostic blood samples, source of control populations, and inclusion of incident versus prevalent cases, or bilateral versus mostly unilateral breast cancers.

The main strength of the current report is the inclusion of 3 independent study populations. Although we found no significant heterogeneity in ATMmp2a risk associations across studies, the evidence was strongest for KConFab and weaker for the BGS and, particularly, for the EPIC cohort. The stronger association in KConFab could be explained by the inclusion of cases with very strong family histories, and/or to the choice of controls that were selected from best friends of KConFab participants, and had no family history of breast cancer. In contrast, BGS and EPIC are nested case–control studies within 2 general population cohorts, thus ensuring that cases and controls come from the same source population. We observed a weak correlation between increasing methylation and increasing age at blood draw in the control populations, as has been shown for numerous methylation markers in blood (3, 4). The age correlation was particularly strong in the EPIC cohort, which might explain that the risk association in this cohort was not seen in women with blood collected at older ages. It is important to note that breast cancer cases in KConFab were younger than controls; therefore, if there was no real association with risk, an association between increasing age at blood collection and increasing methylation would result in an observed inverse association, rather than a positive association that we found. The fact that the ATM risk association was most robust and consistent across studies for women with bloods collected in younger ages suggests that this finding may be more applicable to women in this group or to breast cancer cases with earlier onset. However, additional data are needed to confirm this potential effect modification by age as differences were not statistically significant in our analyses. We can speculate that if ATM hypermethylation is associated with an endogenous increased risk of breast cancer, then it is plausible that it would be more frequent in earlier onset breast cancer cases, similar to genetic mutations, and we would thus have an ascertainment bias toward finding this association in younger than in older women. Larger sample sizes from studies with long follow-up will be required to separate out the associations among women with blood drawn at a young age from those associated with earlier onset breast cancer.

The use of prediagnostic samples from incident breast cancer cases in this report excludes the possibility that methylation variability in WBC DNA was influenced by the presence of clinical cancer or treatment in these patients. Furthermore, we found no association between time from blood collection to diagnosis and level of ATM methylation, indicating that

![Figure 1. Meta-analysis of ATM methylation in 4 case–control studies.](image-url)
findings are unlikely to be explained by preclinical disease effects (range of time from blood collection to diagnosis in our studies was <1 month to 11 years). Furthermore, using serial blood samples taken 6 years apart, we have shown that this marker appears to be stable over time. Thus, data suggest that ATM hypermethylation in WBC DNA represents a stable marker of predisposition rather than an early tumorigenic event that increases with increasing tumor burden.

The strongest association between ATM methylation and breast cancer risk was observed in the KConFab study which is a familial breast cancer cohort that includes an overrepresentation of BRCA1 and BRCA2 mutation carriers. We have previously shown that BRCA1 mutant tumors have much lower levels of gene-specific hypermethylation than non-BRCA1/2 or BRCA2 mutant tumors (40). This suggests that BRCA1 mutations may lead to an overall disturbance in methylation patterns. To support this hypothesis, a recent study has shown that young girls with a strong family history of breast cancer have significantly lower levels of WBC DNA methylation of ALU and LINE1 repetitive elements than young girls without a strong family history (41). However, in contrast to this finding, we found no differences between methylation levels of LINE1, or ATM in carriers of BRCA1 (n = 39), BRCA2 (n = 23) mutations compared with noncarriers (n = 104) in the KConFab adult cases (data not shown). In addition, we found no association between methylation in LINE1 or ATM and family history of breast cancer in BGS and EPIC controls (data not shown), thus not supporting a strong association between methylation and family history. Further work on larger numbers of mutation carriers will be required to determine whether these do indeed have an overall aberrant methylation defect.

The association between breast cancer risk and methylation of an intragenic repetitive element in ATM is consistent with a previous study in which the majority of cancer-associated CpG sites were found in intragenic sequences (42). Furthermore, we have previously shown that greater methylation variability occurs within gene bodies than within promoters (14). Several hypotheses have been proposed for how intragenic DNA methylation regulates transcription via mechanisms including transcription rate, nucleosome positioning, alternate start sites, replication timing, or chromatin marks (42). Therefore, these data raise the possibility that variation in gene expression and indeed cancer susceptibility may be more likely to be influenced by the more variable DNA methylation in intragenic repetitive elements than less variable methylation at CpG islands.

The mechanism by which methylation at ATMvmp2a could increase risk is not known. We have previously shown in cancer cell lines that hypermethylation of the nearby locus ATMvmp2b is associated with reduced ATM expression (14), but expression analysis in cell lines and blood samples is needed to examine the effect of ATMvmp2a hypermethylation on ATM expression. Interestingly, expression of ATM in WBCs of patients with breast cancer and controls, at the time of mammographic screening, showed significantly reduced expression in patients with breast cancer (n = 51) compared with controls (n = 31) in a recent study (ref. 43; Supplementary Fig. S4). These data support the hypothesis that reduced expression of ATM, detectable in WBCs, may be linked to breast cancer susceptibility.

Genome-wide DNA hypomethylation has been linked to bladder cancer risk in 3 relatively large case–control studies (285–775 cases; refs. 16–18). In addition, genome-wide hypomethylation has been linked to the risk of head and neck squamous cell carcinoma (19), whereas 2 relatively small studies (40 cases vs. 40 controls and 179 cases vs. 180 controls) suggested an association with breast cancer (20, 39). In our previous study, we showed no significant difference between 190 breast cancer cases and 190 controls using the LINE1 repetitive element in WBC DNA methylation (14). Consistent with these findings, we found no evidence of association between LINE1 methylation and breast cancer risk in any of the 3 studies in this report. The lack of association between methylation levels in LINE1, and also ATMvmp2b, might be explained by the low between-subject variation observed in the populations under study relative to the total variation (thus resulting in low ICCs for these 2 assays). Inconsistencies between our findings and those reporting associations between blood genome-wide hypomethylation and cancer risk may be due to smaller sample sizes (20, 39), the use of postdiagnostic DNA samples in previous studies (20, 39), different cancer types (16–18), and different methods using other repetitive elements (20, 39) or total 5-mC (16–18).

In conclusion, our findings on the association between ATM hypermethylation in WBC DNA before diagnosis and the risk of developing breast cancer provide further support for the investigation of common epigenetic variability as risk markers for breast and other cancers. In addition to candidate-gene studies, epigenome-wide association studies (EWAS) are now possible with the development of high-throughput technology that allows high-resolution analysis on a genomic level (15, 32, 44). However, adequately powered studies with blood samples collected before diagnosis will be critical for the success of both candidate- and epigenome-wide approaches to discover epigenetic biomarkers of cancer risk.

Disclosure of Potential Conflicts of Interest
R. Brown is a member of Generations Study Oversight committee and has employment (other than primary affiliation: e.g., consulting) in Institute Cancer Research as a professor. No potential conflicts of interests were disclosed by the other authors.

Acknowledgments
The authors thank Wei Dai and Charlotte Wilhelm-Benartzi for helpful discussions; Heather Thorne, Eveline Niedermayr, all the KConFab research nurses, and staff, the heads, and staff of the Family Cancer Clinic; the Clinical Follow Up Study for their contributions to this resource: the many families who contribute to KConFab; and the study participants, study staff, and the doctors, nurses and other health care staff and data providers who have contributed to the study.

Grant Support
This work was funded by Breast Cancer Campaign fellowship to J.M. Flanagan and Cancer Research UK program CS16/A6699 to R. Brown, J.M. Flanagan and K. Brennan are funded by Breast Cancer Campaign. The Clinical Follow Up Study was funded by NHMRC grants 145684, 288704, and 454008. KConFab is supported by grants from the National Breast Cancer Foundation, the National Health and Medical Research Council (NHMRC) and by the Queensland Cancer Fund, the
Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-
8.
7.
11.
Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T,
15.
Meaburn EL, Schalkwyk LC, Mill J. Allele-speci
c methylation in the
4.
Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, et al.
4.
5.
1.
2.
9.
10.
11.
Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH, et al. Rapid DNA methylation changes after exposure to traffic parti-
12.
13.
14.
15.
16.
17.
18.
Intragenic ATM Methylation in Peripheral Blood DNA as a Biomarker of Breast Cancer Risk

Kevin Brennan, Montserrat Garcia-Closas, Nick Orr, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-11-3157

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2012/02/28/0008-5472.CAN-11-3157.DC1

Cited articles
This article cites 44 articles, 9 of which you can access for free at:
http://cancerres.aacrjournals.org/content/72/9/2304.full#ref-list-1

Citing articles
This article has been cited by 10 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/72/9/2304.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.