

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

Kevin Brennan¹, Montserrat Garcia-Closas^{4,6}, Nick Orr⁴, Olivia Fletcher⁴, Michael Jones⁶, Alan Ashworth⁴, Anthony Swerdlow⁶, Heather Thorne on behalf of KConFab Investigators⁷, Elio Riboli³, Paolo Vineis^{2,8}, Miren Dorronsoro⁹, Francoise Clavel-Chapelon¹⁰, Salvatore Panico¹¹, N. Charlotte Onland-Moret¹², Dimitrios Trichopoulos^{13,14}, Rudolf Kaaks¹⁵, Kay-Tee Khaw¹⁶, Robert Brown^{1,5}, 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; $P = 1.64 \times 10^{-4}$]. 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 ATM 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

Authors' Affiliations: ¹Epigenetics Unit, Department of Surgery and Cancer, ²MRC-HPA Centre, ³School of Public Health, Imperial College; ⁴Breakthrough Breast Cancer Research Centre, ⁵Section of Medicine, Institute for Cancer Research, London; ⁶Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, United Kingdom; ⁷Peter MacCallum Cancer Center, Melbourne, Victoria, Australia; ⁸HuGeF Foundation, Torino, Italy; ⁹Public Health Division of Gipuzkoa, IIS Institute BioDonostia, Basque Health Department; CIBER Epidemiología y Salud Pública, CIBERESP, Gipuzkoa, Spain; ¹⁰INSERM, Centre for Research in Epidemiology and Population Health, Institut Gustave Roussy, University Paris-Sud, Paris, France; ¹¹Department of Clinical and Experimental Medicine, Federico II University, Naples, Italy; ¹²Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands; ¹³Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts; ¹⁴Hellenic Health Foundation, Athens, Greece; ¹⁵Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany; and ¹⁶Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge, United Kingdom

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.

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 gene-specific 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 Cunningham 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 provided 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 enrolment, 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 (84% ER positive, 16% ER negative), and HER2 status (19% HER2

Table 1. Characteristics of the study populations

	KConFab		BGS ^a		EPIC ^b	
	Controls	Cases	Controls	Cases	Controls	Cases
N	225	166	253	253	291	248
Age at blood draw, mean (range), y	60 (33–83)	50 (21–91)	54 (23–82)	54 (23–82)	52 (33–76)	52 (33–76)
Family history, <i>n</i> (%)						
Yes	0 (0)	166 (100)	51 (20)	69 (27)	24 (14)	23 (18)
No	225 (100)	0 (0)	201 (80)	183 (73)	144 (86)	108 (82)
Menopausal status at blood draw, <i>n</i> (%)						
Premenopausal	37 (17)	101 (65)	135 (61)	134 (62)	145 (50)	127 (50)
Postmenopausal	183 (83)	54 (35)	87 (39)	83 (38)	146 (50)	121 (50)
Time from blood collection to diagnosis, mean (range), mo	—	45 (1–140)	—	18 (<1–59)	—	55 (24–108)
Age at diagnosis, mean (range)	—	52 (29–88)	—	55 (23–84)	—	57 (37–80)

^aCases 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.

^bCases 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.

positive). An additional random sample of 92 participants in the BGS was selected from women enrolled in 2004 with 2 blood samples/questionnaires 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 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. DCIS cases were excluded from analyses (*n* = 36), resulting in a total of 248 cases and 291 controls available for analyses (Table 1). We did not have precise ethnicity data on these individuals; however, the majority (80%) of individuals were provided from Italy, Spain, and the Netherlands and the remainder (20%) from France, Germany, United Kingdom, and Greece. Blood samples from cases were taken on average 55 months before diagnosis (range, 24–108); Table 1). Extensive information on cancer risk factors, including extensive alcohol, smoking and dietary data, family history, and hormonal factors, was collected from a baseline questionnaire at enroll-

ment. 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

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 cases than controls in the familial samples from KConFab (81.8% vs. 76.9%, $P = 4.87 \times 10^{-6}$; 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 2. Distribution of methylation levels in ATMmvp2a, ATMmvp2b, and LINE1 in breast cancer cases and controls

Assay	Study	Control	Case	Control	Case	<i>P</i> ^b
		<i>n</i> ^a	<i>n</i> ^a	Median (IQR)	Median (IQR)	
ATMmvp2a	BGS	248	249	76.4 (70.2–80.2)	76.8 (70.9–82.7)	0.02
	EPIC	283	235	76.1 (70.5–80.6)	75.7 (70.0–80.8)	0.40
	KConFab	210	156	76.9 (71.6–81.5)	81.8 (75.8–86.5)	4.87×10^{-6}
ATMmvp2b	BGS	234	248	91.0 (87.0–94.8)	91.4 (85.6–95.0)	0.61
	EPIC	287	240	92.2 (87.3–95.2)	92.3 (88.3–95.7)	0.36
	KConFab	208	162	92.6 (87.2–96.3)	92.3 (82.4–96.5)	0.24
LINE1	BGS	242	241	79.0 (77.9–80.1)	79.0 (78.1–79.9)	0.96
	EPIC	263	232	75.1 (73.9–76.3)	75.2 (73.9–76.3)	0.89
	KConFab	218	153	76.0 (74.3–78.0)	76.6 (75.2–77.6)	0.20

NOTE: Bold signifies $P < 0.05$.

Abbreviation: IQR, interquartile range.

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

^bWilcoxon matched pairs test for BGS and EPIC and Mann–Whitney *U* test for KConFab.

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

Study	Quintile ^a	Methylation range	Controls		Cases		OR ^b (95% CI)	P
			N	Freq.	N	Freq.		
BGS	Qi1	3.4%–68.0%	50	0.20	35	0.14	1.00	
	Qi2	68.0%–74.1%	40	0.16	46	0.18	1.56 (0.85–2.88)	0.15
	Qi3	74.1%–77.6%	54	0.22	47	0.19	1.15 (0.66–2.16)	0.54
	Qi4	77.6%–81.0%	56	0.23	42	0.17	1.00 (0.55–1.81)	0.99
	Qi5	81.0%–91.7%	48	0.19	79	0.32	2.31 (1.31–4.06)	3.7 × 10⁻³
	Totals		248		249			
EPIC	Qi1	53.6%–69.7%	60	0.21	49	0.21	1.00	
	Qi2	69.7%–74.8%	61	0.22	51	0.22	1.02 (0.60–1.74)	0.96
	Qi3	74.8%–78.6%	58	0.19	46	0.20	0.97 (0.56–1.67)	0.97
	Qi4	78.6%–82.4%	49	0.17	38	0.16	0.95 (0.54–1.68)	0.95
	Qi5	82.4%–97.5%	55	0.19	51	0.22	1.13 (0.66–1.94)	0.76
	Totals		283		235			
KConFab	Qi1	19.0%–70.2%	38	0.18	20	0.13	1.00	
	Qi2	70.2%–75.4%	47	0.22	16	0.10	0.55 (0.25–1.25)	0.15
	Qi3	75.4%–79.1%	36	0.17	22	0.14	1.11 (0.51–2.44)	0.80
	Qi4	79.1%–83.0%	43	0.20	31	0.20	1.40 (0.67–2.95)	0.37
	Qi5	83.0%–100%	46	0.22	67	0.43	3.06 (1.53–6.10)	1.5 × 10⁻³
	Totals		210		156			
Combined	Qi1	–71.3% to –6.5%	148	0.20	104	0.16	1.00	
	Qi2	–6.5 to –1.2%	148	0.20	113	0.18	1.08 (0.76–1.54)	0.64
	Qi3	–1.2 to 2.5%	148	0.20	115	0.18	1.09 (0.77–1.54)	0.64
	Qi4	2.5%–6.3%	148	0.20	111	0.17	1.06 (0.74–1.51)	0.75
	Qi5	6.3%–23.3%	149	0.20	197	0.31	1.89 (1.36–2.64)	1.6 × 10⁻⁴
	Totals		741		640			

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

^aCutoff values determined by quintiles of Z-scores based on the distribution in the combined control population. Z-score cutoff values are –0.83, –0.24, 0.19, and 0.64.

^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).

differences were found for methylation levels at the ATMmvp2b locus (131 bp downstream from ATMmvp2a) or LINE1 in any of the studies (Table 2).

Quintile 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; $P = 1.64 \times 10^{-4}$; 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 4. Association between methylation levels in ATMmpv2a 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) ^a	OR ^b (95% CI)	P
		N	Freq.	N	Freq.			
21–49	Qi1	45	0.21	44	0.17	0.20, 0.27, 0.52	1.00	
	Qi2	51	0.23	39	0.15	0.28, 0.26, 0.46	0.68 (0.37–1.25)	0.21
	Qi3	44	0.20	41	0.16	0.27, 0.32, 0.41	0.94 (0.51–1.72)	0.83
	Qi4	38	0.18	44	0.17	0.43, 0.18, 0.39	1.00 (0.54–2.17)	0.99
	Qi5	39	0.18	90	0.35	0.40, 0.33, 0.27	2.07^c (1.16–3.68)	0.01
	Total	217		258		0.33, 0.28, 0.38		
>49–59	Qi1	54	0.20	27	0.13	0.19, 0.41, 0.41	1	
	Qi2	55	0.20	43	0.21	0.05, 0.49, 0.47	1.57 (0.84–2.91)	0.15
	Qi3	62	0.23	48	0.24	0.19, 0.38, 0.44	1.56 (0.85–2.86)	0.15
	Qi4	50	0.19	33	0.16	0.09, 0.48, 0.42	1.26 (0.66–2.41)	0.48
	Qi5	49	0.18	53	0.26	0.19, 0.49, 0.32	2.25^c (1.21–4.17)	0.01
	Total	270		204		0.14, 0.45, 0.41		
>59–91	Qi1	49	0.19	33	0.19	0.18, 0.36, 0.45	1.00	
	Qi2	42	0.17	31	0.17	0.10, 0.48, 0.42	1.00 (0.51–1.97)	0.99
	Qi3	42	0.17	26	0.15	0.08, 0.62, 0.31	0.84 (0.45–1.58)	0.61
	Qi4	60	0.24	34	0.19	0.26, 0.53, 0.21	0.80 (0.45–1.45)	0.50
	Qi5	61	0.24	54	0.31	0.39, 0.43, 0.19	1.39 ^c (0.87–2.26)	0.27
	Total	254		178		0.23, 0.47, 0.30		

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

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

^bORs within each age group are adjusted by study. Test for heterogeneity of effects by age group in combined analysis, $P = 0.3109$ for Qi5 versus Qi1.

^c P values for study heterogeneity for ORs comparing Qi5 versus Qi1 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 ATMmpv2a methylation levels and increasing age at blood collection in controls (Spearman's $\rho = 0.15$, $P = 0.0015$), but not in cases (Spearman's $\rho = -0.02$, $P = 0.43$), that was most significant in the EPIC cohort ($\rho = 0.11$, $P = 0.007$) compared with KConFab ($\rho = 0.06$, $P = 0.26$) and BGS ($\rho = 0.02$, $P = 0.40$). 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 ATMmpv2a methylation levels in any of the 3 control populations (data not shown). In the familial cases from KConFab, we found no significant associations between

ATMmpv2a methylation levels and *BRCAl/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 ATMmpv2a 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 ATMmpv2a (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, $R^2 = 0.79$) and there are no significant

differences between measurements in the 2 time points (paired *t* test, $P = 0.24$; signed test for H_0 : 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 *ATM*mvp2b 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 (*ATM*mvp2a and *ATM*mvp2b) 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 *ATM*mvp2a (age-adjusted OR, 1.90; 95% CI, 1.00–3.62; $P = 0.05$) and *ATM*mvp2b (age-adjusted OR, 3.07; 95% CI, 1.58–5.93; $P = 8.8 \times 10^{-4}$; 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 *ATM*mvp2a marker are associated with 1.89-fold OR (95% CI, 1.27–2.82; $P = 1.7 \times 10^{-3}$, 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

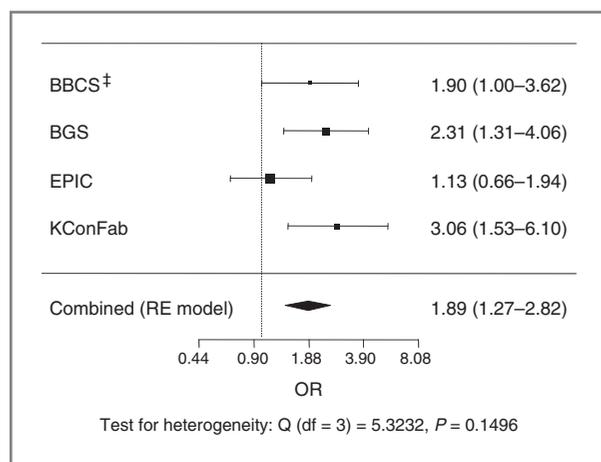


Figure 1. Meta-analysis of *ATM* methylation in 4 case-control studies. A, methylation levels at *ATM*mvp2a were separated into control quintile ranges (using identical ranges from combined analysis), for the previously published report on the British Breast Cancer Study (BBCS[‡]; ref. 14) and the 3 new case-control studies. ORs of Q₅ versus Q₁ are shown. Combined overall analysis was determined using a random-effects (RE) model.

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 *ATM*mvp2a were consistent with those in this report, the strong association found for *ATM*mvp2b 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 *ATM*mvp2a 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

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 *ATM*mvp2a could increase risk is not known. We have previously shown in cancer cell lines that hypermethylation of the nearby locus *ATM*mvp2b is associated with reduced *ATM* expression (14), but expression analysis in cell lines and blood samples is needed to examine the effect of *ATM*mvp2a 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 *ATM*mvp2b, 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 Clinics; 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 C536/A6689 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 454508. 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

Cancer Councils of New South Wales, Victoria, Tasmania, and South Australia, and the Cancer Foundation of Western Australia. The authors thank Breakthrough Breast Cancer and the Institute of Cancer Research for support and funding of the Breakthrough Generations Study. The ICR also acknowledges NHS funding to the NIHR Biomedical Research Centre. The EPIC cohort is supported by the Europe Against Cancer Program of the European Commission (SANCO).

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 22, 2011; revised February 17, 2012; accepted February 19, 2012; published OnlineFirst February 28, 2012.

References

- Jones PA, Baylín SB. The epigenomics of cancer. *Cell* 2007;128:683–92.
- Bjornsson HT, Fallin MD, Feinberg AP. An integrated epigenetic and genetic approach to common human disease. *Trends Genet* 2004;20:350–8.
- Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res* 2010;20:440–6.
- Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res* 2010;20:434–9.
- Li S, Hursting SD, Davis BJ, McLachlan JA, Barrett JC. Environmental exposure, DNA methylation, and gene regulation: lessons from diethylstilbestrol-induced cancers. *Ann N Y Acad Sci* 2003;983:161–9.
- Kovalchuk O, Burke P, Besplug J, Slovack M, Filkowski J, Pogribny I. Methylation changes in muscle and liver tissues of male and female mice exposed to acute and chronic low-dose X-ray-irradiation. *Mutat Res* 2004;548:75–84.
- Ma S, Liu X, Jiao B, Yang Y, Liu X. Low-dose radiation-induced responses: focusing on epigenetic regulation. *Int J Radiat Biol* 2010;86:517–28.
- Christensen BC, Kelsey KT, Zheng S, Houseman EA, Marsit CJ, Wrensch MR, et al. Breast cancer DNA methylation profiles are associated with tumor size and alcohol and folate intake. *PLoS Genet* 2010;6:e1001043.
- Vineis P, Chuang SC, Vaissiere T, Cuenin C, Ricceri F, Johansson M, et al. DNA methylation changes associated with cancer risk factors and blood levels of vitamin metabolites in a prospective study. *Epigenetics* 2011;6:195–201.
- Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet* 2011;88:450–7.
- Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH, et al. Rapid DNA methylation changes after exposure to traffic particles. *Am J Respir Crit Care Med* 2009;179:572–8.
- Meaburn EL, Schalkwyk LC, Mill J. Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics* 2010;5:578–82.
- Feinberg AP, Irizarry RA, Fradin D, Aryee MJ, Murakami P, Aspelund T, et al. Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med* 2010;2:49ra67.
- Flanagan JM, Munoz-Alegre M, Henderson S, Tang T, Sun P, Johnson N, et al. Gene-body hypermethylation of ATM in peripheral blood DNA of bilateral breast cancer patients. *Hum Mol Genet* 2009;18:1332–42.
- Marsit CJ, Koestler DC, Christensen BC, Karagas MR, Houseman EA, Kelsey KT. DNA methylation array analysis identifies profiles of blood-derived DNA methylation associated with bladder cancer. *J Clin Oncol* 2011;29:1133–9.
- Moore LE, Pfeiffer RM, Poscablo C, Real FX, Kogevinas M, Silverman D, et al. Genomic DNA hypomethylation as a biomarker for bladder cancer susceptibility in the Spanish Bladder Cancer Study: a case-control study. *Lancet Oncol* 2008;9:359–66.
- Cash HL, Tao L, Yuan JM, Marsit CJ, Houseman EA, Xiang YB, et al. LINE-1 hypomethylation is associated with bladder cancer risk among nonsmoking Chinese. *Int J Cancer* 2012;130:1151–9.
- Wilhelm CS, Kelsey KT, Butler R, Plaza S, Gagne L, Zens MS, et al. Implications of LINE1 methylation for bladder cancer risk in women. *Clin Cancer Res* 2010;16:1682–9.
- Hsiung DT, Marsit CJ, Houseman EA, Eddy K, Furniss CS, McClean MD, et al. Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* 2007;16:108–14.
- Cho YH, Yazici H, Wu HC, Terry MB, Gonzalez K, Qu M, et al. Aberrant promoter hypermethylation and genomic hypomethylation in tumor, adjacent normal tissues and blood from breast cancer patients. *Anticancer Res* 2010;30:2489–96.
- Cui H, Onyango P, Brandenburg S, Wu Y, Hsieh CL, Feinberg AP. Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. *Cancer Res* 2002;62:6442–6.
- Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science* 2003;299:1753–5.
- Cruz-Correa M, Cui H, Giardiello FM, Powe NR, Hylind L, Robinson A, et al. Loss of imprinting of insulin growth factor II gene: a potential heritable biomarker for colon neoplasia predisposition. *Gastroenterology* 2004;126:964–70.
- Jirtle RL. IGF2 loss of imprinting: a potential heritable risk factor for colorectal cancer. *Gastroenterology* 2004;126:1190–3.
- Ito Y, Koessler T, Ibrahim AE, Rai S, Vowler SL, Abu-Amero S, et al. Somatically acquired hypomethylation of IGF2 in breast and colorectal cancer. *Hum Mol Genet* 2008;17:2633–43.
- Widschwendter M, Apostolidou S, Raum E, Rothenbacher D, Fiegl H, Menon U, et al. Epigenotyping in peripheral blood cell DNA and breast cancer risk: a proof of principle study. *PLoS One* 2008;3:e2656.
- Iwamoto T, Yamamoto N, Taguchi T, Tamaki Y, Noguchi S. BRCA1 promoter methylation in peripheral blood cells is associated with increased risk of breast cancer with BRCA1 promoter methylation. *Breast Cancer Res Treat* 2011;129:69–77.
- Snell C, Krypyuk M, Wong EM, Loughrey MB, Dobrovic A. BRCA1 promoter methylation in peripheral blood DNA of mutation negative familial breast cancer patients with a BRCA1 tumour phenotype. *Breast Cancer Res* 2008;10:R12.
- Wong EM, Southey MC, Fox SB, Brown MA, Dowty JG, Jenkins MA, et al. Constitutional methylation of the BRCA1 promoter is specifically associated with BRCA1 mutation-associated pathology in early-onset breast cancer. *Cancer Prev Res (Phila)* 2010;4:23–33.
- Wang X, Zhu H, Snieder H, Su S, Munn D, Harshfield G, et al. Obesity related methylation changes in DNA of peripheral blood leukocytes. *BMC Med* 2010;8:87.
- Pedersen KS, Bamlet WR, Ober AL, de Andrade M, Matsumoto ME, Tang H, et al. Leukocyte DNA methylation signature differentiates pancreatic cancer patients from healthy controls. *PLoS One* 2011;6:e18223.
- Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Gayther SA, Apostolidou S, et al. An epigenetic signature in peripheral blood predicts active ovarian cancer. *PLoS One* 2009;4:e8274.
- Feinberg AP. The epigenetics of cancer etiology. *Semin Cancer Biol* 2004;14:427–32.
- Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009;41:178–86.
- Mann GJ, Thorne H, Balleine RL, Butow PN, Clarke CL, Edkins E, et al. Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. *Breast Cancer Res* 2006;8:R12.

36. Swerdlow AJ, Jones ME, Schoemaker MJ, Hemming J, Thomas D, Williamson J, et al. The Breakthrough Generations Study: design of a long-term UK cohort study to investigate breast cancer aetiology. *Br J Cancer* 2011;105:911–7.
37. Riboli E, Kaaks R. The EPIC Project: rationale and study design. *European Prospective Investigation into Cancer and Nutrition. Int J Epidemiol* 1997;26 Suppl 1:S6–14.
38. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:1–48.
39. Choi JY, James SR, Link PA, McCann SE, Hong CC, Davis W, et al. Association between global DNA hypomethylation in leukocytes and risk of breast cancer. *Carcinogenesis* 2009;30:1889–97.
40. Flanagan JM, Cocciardi S, Waddell N, Johnstone CN, Marsh A, Henderson S, et al. DNA methylome of familial breast cancer identifies distinct profiles defined by mutation status. *Am J Hum Genet* 2010;86:420–33.
41. Wu HC, John EM, Ferris JS, Keegan TH, Chung WK, Andrusis I, et al. Global DNA methylation levels in girls with and without a family history of breast cancer. *Epigenetics* 2011;6:29–33.
42. Shenker N, Flanagan JM. Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research. *Br J Cancer* 2012;106:248–53.
43. Labreche HG, Nevins JR, Huang ES. Integrating factor analysis and a transgenic mouse model to reveal a peripheral blood predictor of breast tumors. *BMC Med Genomics* 2011;4:61.
44. Rakyen VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 2011;12:529–41.