

## Genome-Wide DNA Methylation Profiling of CpG Islands in Breast Cancer Identifies Novel Genes Associated with Tumorigenicity

Victoria K. Hill<sup>1</sup>, Christopher Ricketts<sup>1</sup>, Ivan Bieche<sup>2</sup>, Sophie Vacher<sup>2</sup>, Dean Gentle<sup>1</sup>, Cheryl Lewis<sup>3</sup>, Eamonn R. Maher<sup>1</sup>, and Farida Latif<sup>1</sup>

### Abstract

Epigenetic profiling of tumor DNAs may reveal important new theranostic targets to improve prognosis and treatment of advanced cancer patients. In this study, we performed a genome-wide profile of DNA methylation patterns in sporadic breast tumors by using the HumanMethylation27 BeadChips to assess relationships between DNA methylation changes and patient tumor characteristics. The arrays identified 264 hypermethylated loci/genes present in genomic CpG islands. Hierarchical clustering based on methylation levels divided the specimens into three distinct groups, within which certain clinical features also clustered. Statistically significant differences were determined between overall methylation levels of these clusters and estrogen receptor and progesterone receptor (ER/PR) status ( $P = 0.001$ ), tumor relapse ( $P = 0.035$ ), and lymph node metastasis ( $P = 0.042$ ). We identified several individual methylated genes associated with clinical features, including six genes (*RECK*, *SFRP2*, *UAP1L1*, *ACADL*, *ITR*, and *UGT3A1*) that showed statistical significance between methylation and relapse-free survival. Notably, the *RECK* gene in this group has been associated in other cancers with poorest prognosis. Among the leading relapse-associated genes and the genes associated with ER/PR status, we sequenced an independent set of paired normal/tumor breast DNA samples to confirm tumor specificity of methylation. Further, we carried out quantitative real-time reverse transcriptase PCR to confirm reduced expression in methylated tumors. Our findings suggest the utility for the DNA methylation patterns in these genes as clinically useful surrogate markers in breast cancer, as well as new molecular pathways for further investigation as therapeutic targets. *Cancer Res*; 71(8); 2988–99. ©2011 AACR.

### Introduction

Epigenetic silencing of tumor-related genes due to hypermethylation has recently emerged as one of the pivotal alterations in cancer development (1, 2). Hence, cancer is not only a polygenetic disease, but also a polyepigenetic disease. There are 3 main clinical uses for the DNA methylation patterns associated with cancer development and progression. First, DNA methylation patterns that are cancer cell specific and occur at an early stage of cancer develop-

ment makes them ideal biomarkers for early cancer detection. Second, DNA methylation patterns can also be used for prognostic purposes such as survival and/or relapse. Third, unlike gene mutations, hypermethylation of gene promoters can be reversed by treatment with demethylating agents and are regarded as optimal targets for what is now known as epigenetic therapy. The clinical use of demethylating agents such as 5-aza-2'-deoxycytidine (decitabine) and 5-azacytidine (Vidaza) have had the greatest success in the field of hematologic malignancies. In 2004, FDA approved the use of 5-azacytidine for the treatment of myelodysplastic syndromes (1).

It is estimated that more than 1 million women are diagnosed with breast cancer globally every year, and a substantial number of these women will die from the disease (3). The economic cost of treating breast cancer patients is substantial. Because of the heterogeneous nature of the disease, clinical progression and response to therapy is difficult to predict with the current prognostic factors and hence treatment regimens are not as effective as they should be. In this study we used the HumanMethylation27 BeadChips to explore associations between breast tumor methylation profiles and important tumor characteristics such as tumor stage, metastasis, relapse, survival, and hormone receptor status.

**Authors' Affiliations:** <sup>1</sup>Medical and Molecular Genetics, School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, Birmingham, United Kingdom; <sup>2</sup>Laboratory of Oncogenetics, Institut Curie—Hôpital René Huguenin, FNCLCC, St-Cloud, France; and <sup>3</sup>Department of Surgery, University of Texas Southwestern Medical Center, Dallas, Texas

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

**Corresponding Author:** Farida Latif, Medical and Molecular Genetics, School of Clinical and Experimental Medicine, College of Medical and Dental Sciences, University of Birmingham, Birmingham B15 2TT, United Kingdom. Phone: 44-0-121-414-4452; Fax: 44-0-121-414-2538; E-mail: f.latif@bham.ac.uk

doi: 10.1158/0008-5472.CAN-10-4026

©2011 American Association for Cancer Research.

## Materials and Methods

### Cell lines and patient DNA samples

We used a total of 9 breast cancer cell lines (HCC1806, HCC1419, HCC1937, HCC1143, HCC1395, MCF-7, T47-D, MDA-MB-231, and HTB19) during this study. All sporadic breast tumor samples were ductal carcinomas, we used one cohort of 39 tumors (tumor characteristics described in ref. 4 and in Fig. 3) and a further independent cohort of 20 tumors with matched normal breast tissue DNA. None of the primary tumor samples were microdissected. All samples were acquired with ethical consent.

### Infinium array

We used the Illumina Infinium HumanMethylation27 BeadChips (Illumina). The BeadChip contains 27,578 highly informative CpG sites covering more than 14,000 human RefSeq genes. This allows researches to interrogate all these sites per sample at a single nucleotide resolution. Bisulfite modification of DNA and chip processing and data analysis were carried out by following the manufacturer's manual and was done at the Wellcome Trust Centre for Human Genetics Genomics Lab, Oxford, UK. Array results were run through BeadStudio software (Illumina).

### List analysis

Background normalized average  $\beta$  values were considered hypermethylated if no normal had a  $\beta$  value greater than 0.25 and if either (a) had a  $\beta$  value greater than 0.5 in more than 20% of samples, or (b) had a difference of at least 0.4 between sample  $\beta$  value and mean normal  $\beta$  value in more than 20% of samples. For subsequent list analysis by Fisher's exact tests, samples were given a value of 0 or 1 representing unmethylated or methylated, respectively, as determined by the above criteria.

### Clustering

Hierarchical clustering was carried out in Cluster3. When clustering all samples for all probes, Euclidean distance and single linkage were used, when clustering the 291 hypermethylated loci for the 39 tumors, Euclidean distance and average linkage were used. Clustering results were viewed by using the JavaTree program.

### Statistics

Assessment of overall methylation levels with clinical features were analyzed by ANOVA in MedCalc, significance was taken to be  $P < 0.05$ . Association of single loci with clinical features were assessed by Fisher's exact test, all  $P$  values were corrected by false discovery rate (FDR) and taking significance to be  $P < 0.05$ . Kaplan-Meier analysis was carried out in MedCalc taking into account Bonferroni correction when assessing significance, samples were considered significant if  $P < 0.05$ . Student's  $t$  test was used to determine whether there was significance between quantitative expression of normals and tumors, again significance was taken as  $P < 0.05$ .

### DAVID analysis

The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to identify groups of genes within the selected list of 263 that were within groups based on related functions and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways relevant to carcinogenesis by using the functional annotation chart and pathways analysis functions within the functional annotation tool.

### Methylation analysis of individual genes

Combined bisulfite restriction analysis (CoBRA) was used to assess the methylation status of genes in cell lines and a separate tumor/normal cohort. All DNA were modified by using an EpiTect kit (Qiagen) according to the manufacturers' instructions. All CoBRA PCR were semi-nested, except *ACADL*, which was fully nested because of unsatisfactory amplification by using seminested primers. CoBRA primer sequences are shown in Supplementary Table S1. Primary PCRs were touchdown PCRs with a gene-specific final annealing temperature. Secondary PCRs used 5  $\mu$ L primary product as starting material and were also touchdown PCRs with a longer cycle number (45 instead of 40). PCR products were digested with BstUI (Fermentas UK) overnight at 37°C prior to visualization on a 2% agarose gel.

Samples were cloned and sequenced to validate the array and confirm CoBRA results. PCR products were cloned into the pGEM-T easy vector (Promega) according to the manufacturers' instructions. Up to 12 colonies were picked per sample and sequenced following single-colony PCR by using the following primers: forward, 5'-TAATACGACTCACTA-TAGGG-3'; reverse, 5'-ACACTATAGAATACTCAAGC-3'. Methylation indexes were calculated as a percentage of the number of methylated CpGs out of the total CpGs sequenced.

### Cell line expression analysis

Breast cell line genomic DNA was demethylated by treatment with 5  $\mu$ mol/L 5'-aza-2'-deoxycytidine (5-azadc) over 5 days. Daily media changes were made with the addition of fresh 5-azadc each time. All cell lines were maintained in Dulbecco's modified Eagle's medium (Sigma-Aldrich) supplemented with 2 mmol/L glutamine and 10% fetal calf serum at 37°C, 5% CO<sub>2</sub>. RNA bee (AMS Biotechnology) was used to extract total RNA from cell lines according to the manufacturers' instructions. One microgram total RNA was converted to cDNA by using Superscript III (Invitrogen) and random hexamer primers (Fermentas UK). Reverse transcriptase (RT)-PCR was then carried out by a touchdown PCR program with a final annealing temperature of 56°C. All expression primer sets are shown in Supplementary Table S2.

### Real-time RT-PCR

Quantitative real-time RT-PCR was carried out as described previously (4). Briefly, cDNA was made from total RNA extracted from frozen normal and tumor breast tissues. We quantified transcripts of the *TBP* (TATA box-binding protein) gene as the endogenous RNA control. Results, expressed as  $N$ -fold differences in target gene expression relative to the *TBP* gene (termed *Ntarget*), were determined by the following formula:  $Ntarget = 2^{\Delta Ct_{sample}}$ , where the  $\Delta Ct$  value of the

sample was determined by subtracting the average Ct value of the target gene from the average Ct value of the *TBP* gene. The *Ntarget* values of the samples were subsequently normalized such that the mean ratio of the normal breast samples would equal a value of 1. Primer sequences provided on request. PCR was carried out by using the SYBR Green PCR Core Reagents kit (Perkin-Elmer Applied Biosystems). The thermal cycling conditions comprised an initial denaturation step at 95°C for 10 minutes and 50 cycles at 95°C for 15 seconds and 65°C for 1 minute. Experiments were conducted with duplicates for each data point.

## Results

HumanMethylation27 BeadChip data were obtained from 39 sporadic breast tumors and 4 matched breast tumor/normal pairs. Comparison of overall DNA methylation between normal breast DNA samples showed little variance, with high correlation coefficient values ( $R^2$ ) ranging from 0.974 to 0.9302 (Supplementary Fig. S1). Additionally, hierarchical clustering (Euclidean, single linkage) of all the probes grouped all 4 matched normals together. Of 4 of the matched tumor/normal pairs, 3 showed low correlations ( $R^2 = 0.724, 0.834,$  and  $0.909$ ), indicative of varying levels of aberrant methylation, whereas the fourth sample showed a correlation similar to normals ( $R^2 = 0.9381$ ). This suggested that altered methylation levels could be important in a large number of the sporadic breast tumor cohorts analyzed whereas for some tumors other genetic factors may be involved.

### Identification of hypermethylated CpG loci

To identify the CpG loci showing the most significant tumor-specific changes in methylation all probes with one or more normal sample showing a  $\beta$  value greater than 0.25 were excluded ( $n = 11,211$ ). Remaining probes were only considered hypermethylated if more than 20% of the tumor samples had either (a)  $\beta$  values of 0.5 or more ( $n = 261$  probes, 242 loci) or (b) a difference between the tumor sample  $\beta$  value and the mean normal  $\beta$  value of 0.4 or more ( $n = 202$  probes, 182 loci). All imprinted and X chromosome genes were removed from the analysis and a final selection of 291 probes representing 264 genes was chosen (Supplementary Table S3; Supplementary Fig. S2). The selected genes were analyzed by PubMeth, identifying 34 genes out of the 264 (12.9%) that have been previously shown to be methylated in cancer, including 8 genes in breast cancer representing 10.3% of the genes in PubMeth methylated at a frequency of more than 20% in breast cancer (Supplementary Table S4).

Functional analysis of this selection by the DAVID bioinformatics resource (<http://david.niaid.nih.gov>) showed the presence and enrichment of genes involved in functions related to tumorigenesis: cell adhesion (24 genes), regulation of cell proliferation (21 genes), negative regulation of cell death (15 genes), cell migration (12 genes), cell-cell adhesion (11 genes), regulation of cell cycle (10 genes), and tumor suppressor (6 genes; Supplementary Fig. S3). Similarly, DAVID KEGG pathway analysis showed genes present from relevant pathways: pathways in cancer (12 genes), cell adhesion molecules

(6 genes), and the Wnt signaling pathway (7 genes; Supplementary Fig. S3).

Further confirmation and analysis of the 291 selected probes was performed by bisulfite sequencing of (a) selected genes previously known to be associated with breast cancer, (b) analysis of most highly methylated genes, and (c) analysis to identify methylation specific to clinical features.

### Bisulfite sequencing confirmation of probe methylation

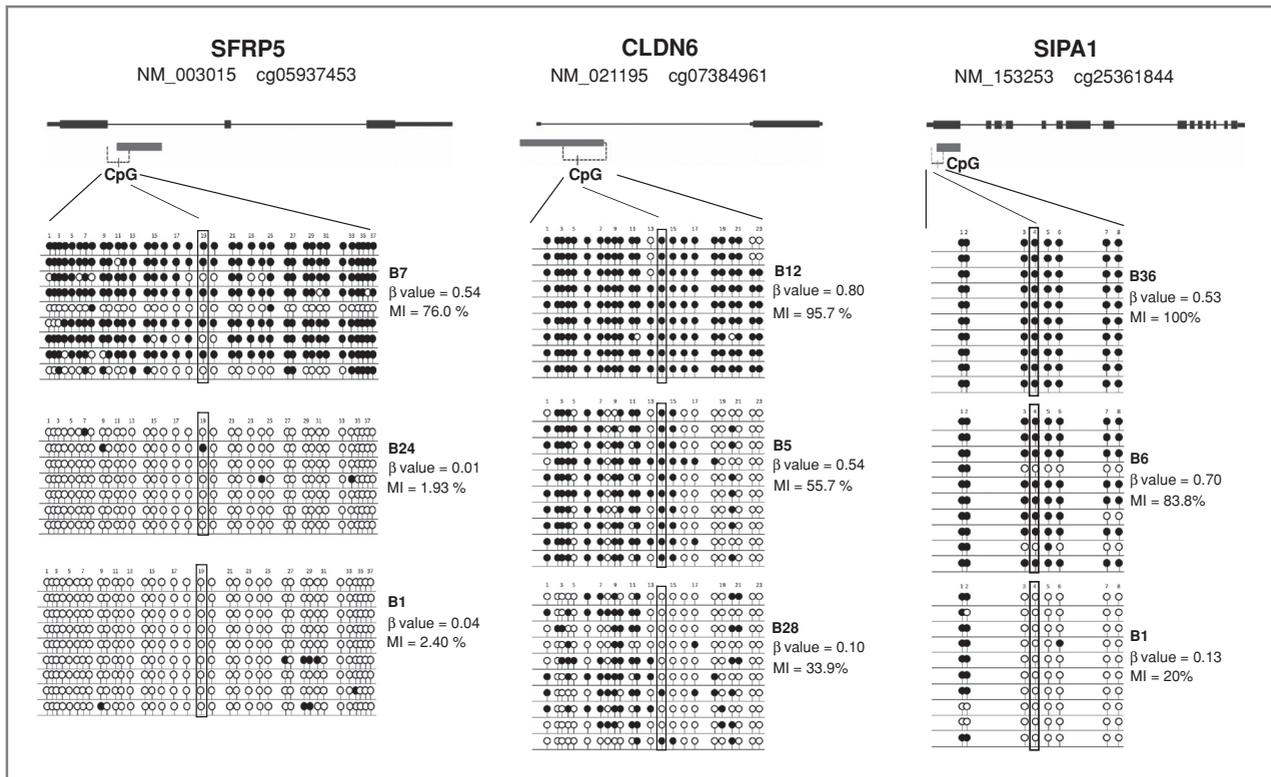
For 3 selected hypermethylated genes previously known to be associated with breast cancer (*SFRP5*, *CLDN6*, and *SIPAI*; refs. 5–8), bisulfite sequencing was done on a region surrounding the CpG island probe for examples of high and low  $\beta$  values. Sequencing showed that  $\beta$  values greater than 0.5 tended to be both highly methylated at the probe position and be indicative of wider CpG island methylation, whereas low  $\beta$  values showed little methylation at the probe position, although in some cases, a greater level of methylation was observed in the surrounding region compared with the probe (Fig. 1). These results confirm the correlation of the  $\beta$  value with methylation levels at the specific CpG being measured by the array.

### Analysis of breast tumor hypermethylated genes

The selected genes were further filtered into 2 groups; the top 10 most highly methylated genes with true CpG islands (amenable to CoBRA; Supplementary Fig. S4a) and genes with more than a single selected probe (Supplementary Fig. S5). Selected genes were then analyzed for methylation in breast cancer cell lines, expression in breast cancer cell lines after treatment with 5-azadC and demonstration of tumor-specific methylation in an additional independent set of 20 tumor/normal matched pairs, where appropriate.

Of the 3 most highly methylated genes with true CpG islands (*ABCA3*, *COX7A1*, and *SST*), *COX7A1* and *SST* showed increased expression after 5-azadC treatment in methylated breast tumor cell lines (Supplementary Fig. S4b). In an independent cohort of breast tumor/normal pairs, *COX7A1* showed increased methylation in 70.6% (12 of 17) of tumors compared with corresponding normal tissue and a further 5 showed equal amounts of low-level methylation in tumor and corresponding normal tissue. *SST* showed tumor-specific methylation in 31.6% (6 of 19) of pairs and increased methylation in a further 31.6% (6 of 19) of tumors compared with corresponding normal tissue DNA. One tumor showed equal amounts of methylation in the tumor and normal and no methylation in the remaining tumors (Supplementary Fig. S4b).

A total of 24 genes had more than a single selected probe, 4 of which (17%) had been previously shown to be methylated in breast cancer (*APC*, *CCND2*, *GSTP1*, and *PTPRO*; Supplementary Fig. S5; refs. 9–12). Further 2 genes (*CDKL2* and *ZNF154*) from the above 24 genes were analyzed in this study. Both *CDKL2* and *ZNF154* showed increased expression after 5-azadC treatment in methylated breast tumor cell lines and also showed altered methylation patterns in matched tumor/normal-paired DNA (Fig. 2). *CDKL2* and *ZNF154* showed tumor-specific methylation in 25% (4 of 16) and 56.3% (9 of



**Figure 1.** Validation of the HumanMethylation27 BeadChip with clone sequencing. Clone sequencing results are shown for 3 samples each, for each of the 3 genes, *SFRP5*, *CLDN6*, and *SIPA1*. In each case, the Illumina ID (ILMNID) is given for the probe being validated and the location of the CpG being assessed in the array shown on a gene schematic and highlighted on the clone sequencing results. For each sample sequenced, the  $\beta$  value is given and the methylation index (MI) for the overall region as a percentage of methylated CpGs. For each sample, each line represents a single allele, black circles represent a methylated CpG, and each white circle represents an unmethylated CpG.

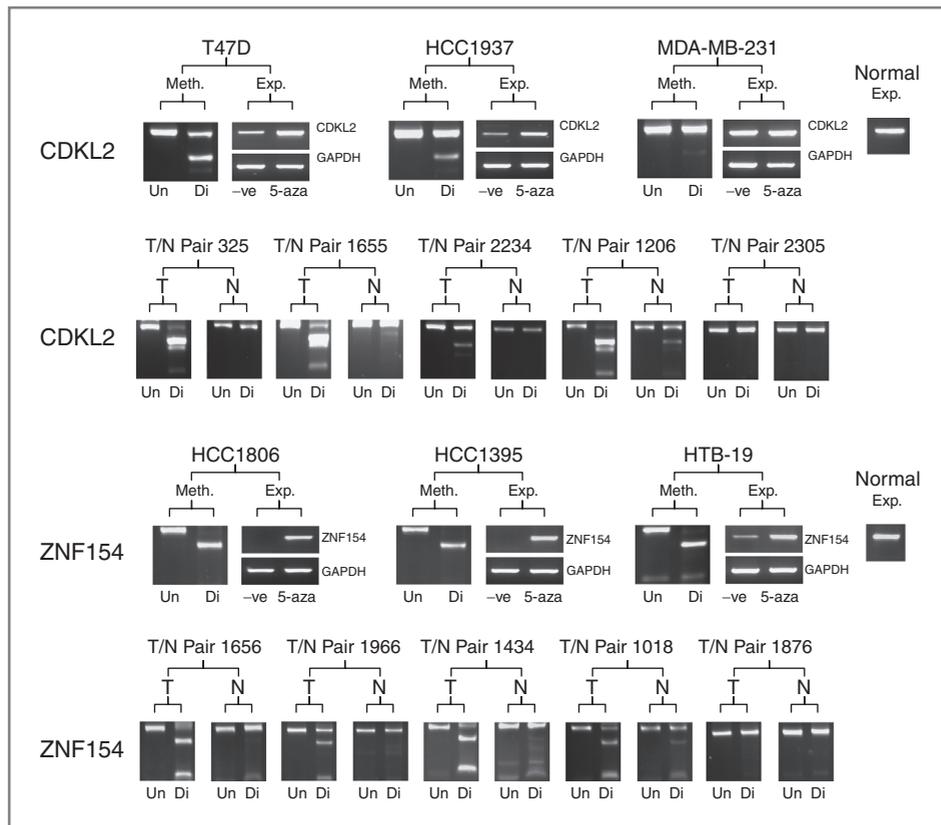
16), respectively, and increased methylation in tumors compared with low-level methylation of matched normal samples in 12.5% (2 of 16) and 25% (4 of 16), respectively. *CDKL2* and *ZNF154* also showed equal methylation in tumor and normal DNA in 18.8% (3 of 16) and 6.3% (1 of 16) of samples. No methylation was detected in the remainder of the tumors.

#### Identification of gene hypermethylation specific to clinical features

Initially, to investigate whether hypermethylation seemed to be associated with any clinical feature, hierarchical clustering was conducted on the 291 selected probes (Fig. 3). This split the sporadic tumor samples into 3 distinct methylation-based clusters: high methylation (18 tumors; mean  $\beta$  value of 0.433), intermediate methylation (7 tumors; mean  $\beta$  value of 0.317), and low methylation (12 tumors; mean  $\beta$  value of 0.176), with 2 outliers (Fig. 3). The high methylation cluster associated with samples that relapsed (high, 12 of 18; intermediate, 1 of 7; and low, 3 of 12) and retained both the estrogen and progesterone receptor (ER/PR; high, 14 of 18; intermediate, 1 of 7; and low, 2 of 12), whereas the low methylation cluster associated with the triple-negative tumors (low, 7 of 12; intermediate, 2 of 7; and high, 0 of 18; Fig. 3).

ANOVAs were carried out by using either the mean  $\beta$  value (methylation) or the disease-free survival for each sample to compare between clinical features. The clusters showed no statistically significant variation between disease-free survival ( $P = 0.162$ ) and no statistically significant variation was observed between methylation and tumor stage ( $P = 0.214$ ) or menopause ( $P = 0.414$ ; Supplementary Fig. S6). Statistically significant variation was observed between methylation and tumor relapse ( $P = 0.035$ ), ER/PR status ( $P = 0.001$ ), and lymph node metastasis ( $P = 0.042$ ; Fig. 4).

To identify individual gene loci whose methylation status correlated with clinical features, Fisher's exact tests with FDR correction were conducted for all 291 selected probes between groups defined by a single clinical feature. This identified 9 methylated probes (9 genes) significantly associated with tumor relapse, 19 methylated probes (18 genes) significantly associated with presence of both estrogen and progesterone receptors, and 2 probes specifically methylated in tumors with lymph node metastasis (Supplementary Table S5). Probes specifically methylated in samples that relapsed should show strong correlations with disease-free survival and thus Kaplan-Meier analysis was done. The 9 identified probes/genes (*RECK*, *ACADL*, *SFRP2*, *ITR*, *UGT3A1*, *UAP1L1*, *SYDE1*, *UGT3A2*, and *PRKCB1*) all showed significant association with relapse-free



**Figure 2.** CoBRA and expression results for CDKL2 and ZNF154. For CDKL2 and ZNF154, CoBRA results are shown for selected cell lines (top) and tumor/normal pairs (bottom). In each case, undigested product (Un) is shown next to digested product (Di). T, tumor; N, normal; Meth., CoBRA result for cell lines; Exp., expression result for cell lines. Expression results pre- (-ve) and posttreatment (5-aza) with 5-azadc are also shown next to the corresponding cell line methylation results. In each case, results are shown for the target gene and glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Expression of the target gene in normal breast tissue is also shown.

survival and 6 remained significant after Bonferroni correction (underlined genes; Fig. 5; Supplementary Table S6).

Selected genes associated with tumor relapse and the presence of both hormone receptors were further analyzed for breast cancer cell line methylation, expression in breast tumor cell lines after treatment with 5-azadc, demonstration of tumor-specific methylation in an additional 20 tumor/matched normal pairs and decreased expression in the original sporadic primary tumors (shown by HumanMethylation27 array to be methylated) by quantitative real-time RT-PCR, where appropriate.

For tumor relapse association, the 3 genes with the highest association with disease-free survival were selected; *ACADL* ( $P < 0.0009$ ), *RECK* ( $P = 0.0009$ ), and *SFRP2* ( $P = 0.0009$ ; Fig. 6). All the 3 genes first showed increased expression or were reexpressed (after 5-azadc treatment) in methylated breast tumor cell lines and then showed methylation in tumor/normal-paired samples at a frequency of 30% (6 of 20), 22.2% (4 of 18), and 57.1% (8 of 14) of samples for *ACADL*, *RECK*, and *SFRP2*, respectively [Fig. 6A and B (ii and iii)], via CoBRA. CoBRA showed no methylation in any of the corresponding normal breast tissues for *RECK* and some methylation in adjacent normal tissues of methylated tumors in *ACADL* (2 of 6) and *SFRP2* (4 of 8), in all cases methylation observed in the associated normal tissue was less than that observed in the tumor tissue. Bisulfite sequencing of *ACADL* and *RECK* showed tumor-specific methylation throughout the selected area of the CpG island [Fig. 6A and B (iv); Supple-

mentary Fig S7]. Quantitative real-time RT-PCR showed significant downregulation of *ACADL* ( $P = 0.043$ ) and *RECK* ( $P > 0.0001$ ) in methylated primary sporadic tumors compared with a panel of normal breast tissues [Fig. 6A and B (v)]. In general, expression in unmethylated tumors was found to be higher than that in methylated tumors for both *ACADL* and *RECK* (Supplementary Fig. S8).

For association between methylation and the presence of both hormone receptors ( $ER^+/PR^+$ ), 3 genes were selected (*TNFRSF10D*, *C1orf114*, and *COL1A2*) for analysis that were also present within the list of genes with multiple hypermethylated probes (Supplementary Fig. S5). Again, all 3 genes first showed increased expression in methylated breast tumor cell lines after 5-azadc treatment and then showed specific methylation in tumors of tumor/normal-paired samples. *COL1A2* showed tumor-specific methylation in 76.5% (13 of 17) of the paired samples (Fig. 7A). *TNFRSF10D* showed tumor-specific methylation in 50% (8 of 16), increased methylation in the tumor compared with normal tissue in 18.8% (3 of 16) and 12.5% (2 of 16) with similar levels in the tumor and normal tissue. *C1orf114* showed tumor-specific methylation in 35.3% (6 of 17), and increased methylation in the tumor compared with normal tissue in 42.2% (7 of 17) of tumors. To confirm the above CoBRA methylation results we selected *C1orf114* for bisulfite sequencing. Cloning and bisulfite sequencing of *C1orf114* showed tumor-specific methylation throughout the selected area of the CpG island, hence confirming the CoBRA results (Fig. 7B). To discern further biological

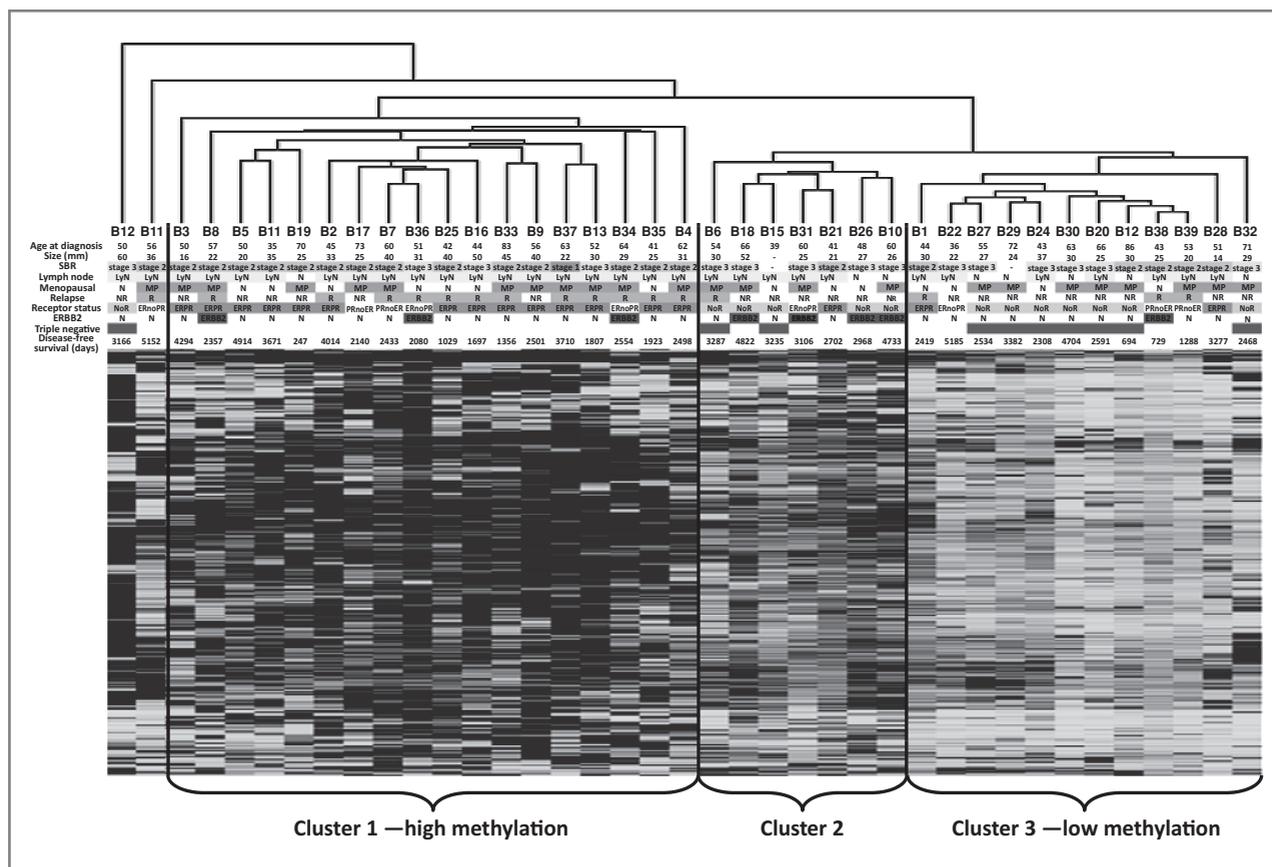


Figure 3. Hierarchical clustering splits tumor samples into 3 main groups. The dendrogram can be seen aligned with sample information showing clinical status of age at diagnosis (years), tumor size (mm), SBR (Scarff-Bloom-Richardson) stage (1, 2, or 3), lymph node status (LyN, lymph node positive; N, lymph node negative), menopause status (MP, menopausal; N, not menopausal), relapse status (R, relapse; NR, no relapse), ER/PR status (ERPR, estrogen and progesterone-positive; ERnoPR, estrogen-positive only; PRnoER, progesterone-positive only; noR, estrogen and progesterone-negative), ERBB2 status (ERBB2, ERBB2 overexpression; N, normal expression), triple negative, and survival (disease-free survival in days). The 3 cluster groups, showing high, medium, and low methylation are separated with thick black lines and labeled underneath. Two outliers can be seen on the extreme left of the clustering image. Black and white shading refers to  $\beta$  values greater than 0.5 and less than 0.5, respectively.

relevance of the methylation found in the above 3 genes, we selected *TNFRSF10D* gene for expression analysis in breast tumor samples. Quantitative real-time RT-PCR analysis

showed significant downregulation of *TNFRSF10D* ( $P = 0.016$ ) in methylated primary sporadic tumors compared with a panel of normal breast controls (Fig. 7C). *TNFRSF10D*

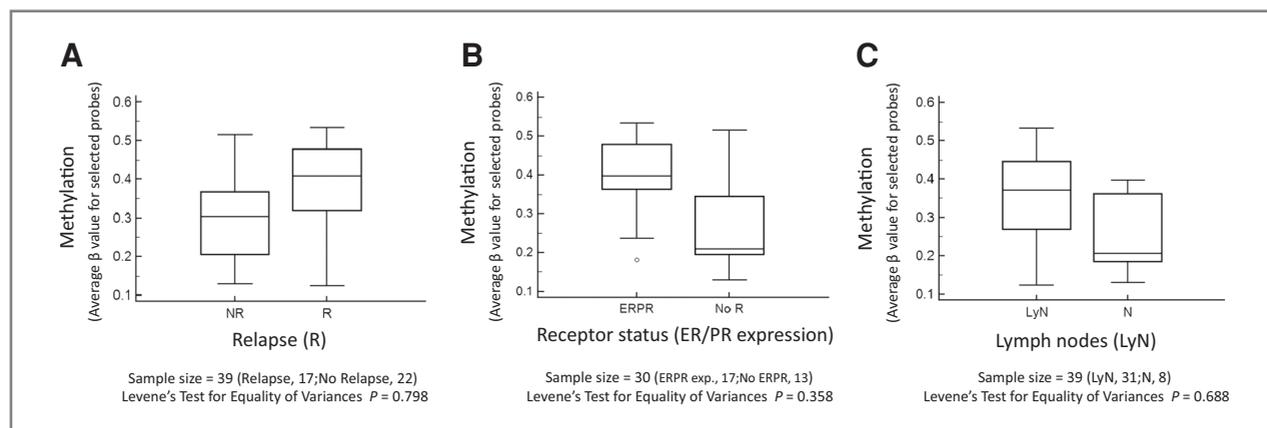
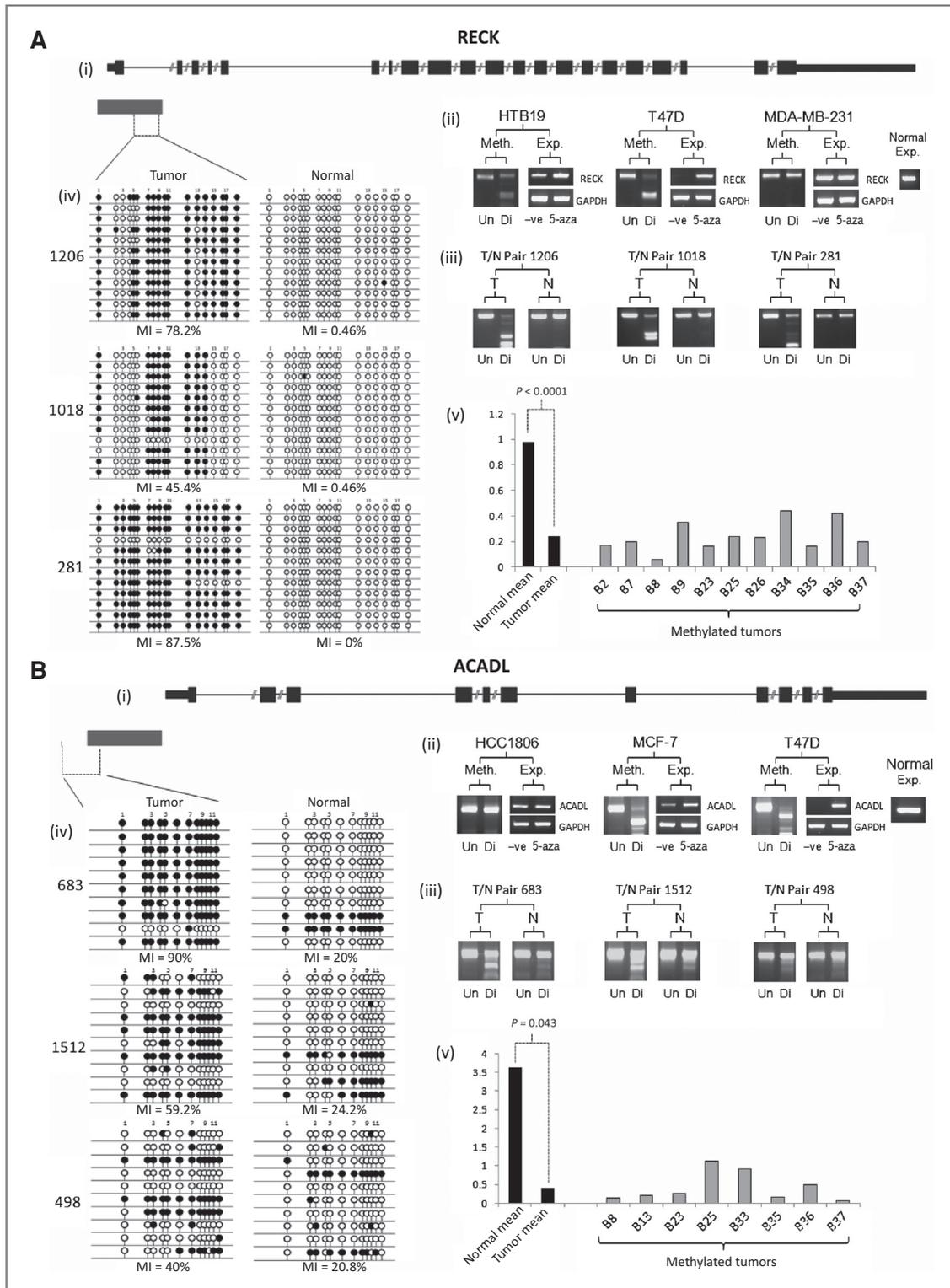
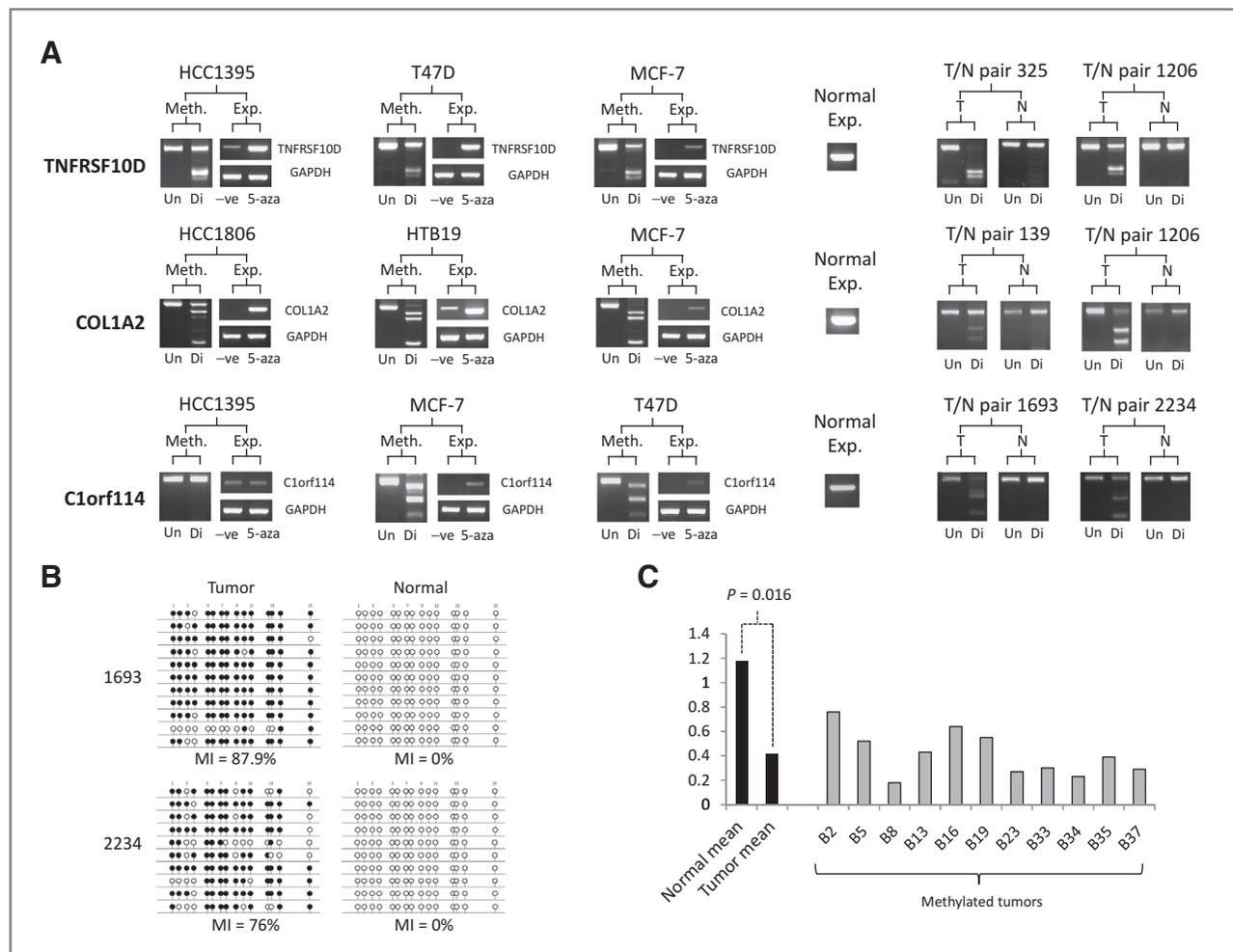


Figure 4. ANOVA analysis shows overall methylation levels are significantly different in relapsing, ER/PR-expressing, and lymph node-positive patients. ANOVA analysis shows significant differences in overall methylation levels (as determined by the mean  $\beta$  value) for (A) relapse ( $P = 0.035$ ), (B) ER/PR status ( $P = 0.001$ ), and (C) lymph node metastasis ( $P = 0.042$ ).





**Figure 6.** RECK and ACADL tumor specificity and primary tumor expression. For both *RECK* (A) and *ACADL* (B), a gene schematic showing the location of the CoBRA primers is shown (i). CoBRA results are shown for selected cell lines (Meth.) along with expression (Exp.) with (5-aza) and without (-ve) treatment with 5-azaDC for both the target gene and GAPDH (ii). CoBRA results are shown for selected tumor/normal pairs (iii) along with clone sequencing (iv). Quantitative real-time expression results are shown for methylated primary tumors (v). All CoBRA results are shown with undigested product (Un) run next to digested product (Di). Each line in the clone sequencing results represents a single allele, each black circle a methylated CpG, and each white circle an unmethylated CpG. Methylation indexes (MI) were calculated as a percentage of methylated CpGs out of the total number of CpGs analyzed. T, tumor; N, normal.



**Figure 7.** ER/PR receptor status and individual loci methylation. **A**, CoBRA results are shown for cell lines and tumor/normal pairs along with expression analysis  $\pm$  5-azadC treatment for *TNFRSF10D*, *COL1A2*, and *C1orf114*. Un, undigested; Di, digested; 5-aza, treatment with 5-azadC; -ve, without treatment with 5-azadC. **B**, *C1orf114* clone sequencing results are shown for tumor normal pairs, black circles representing methylated CpGs and white circles representing unmethylated CpGs; methylation indexes (MI) were calculated as a percentage of methylated CpGs out of the total number of CpGs analyzed. **C**, quantitative real-time PCR results are shown for *TNFRSF10D*. Significance between the mean normal expression and mean methylated tumor expression is given (Student's *t* test).

showed higher expression levels in unmethylated tumors compared with methylated tumors in general, although not in every case, suggesting alternative mechanisms of down-regulation (Supplementary Fig. S8).

## Discussion

The use of gene expression data from patient tumor samples to determine better treatment options is becoming increasingly common in clinical practice. The instability of mRNA makes expression profiling a challenging endeavor for utilization in routine clinical practice outside of major hospitals or commercial laboratories. Although DNA methylation patterns are more stable and can be detected by using a number of high-throughput and sensitive techniques which require little patient material. Using a high-throughput quantitative methylation approach, we established aberrant methy-

lation patterns for a clinically well-characterized cohort of breast tumors. The generated methylation patterns showed strong differentiation between tumor and normal DNA, with many tumor DNA samples showing a far greater number of hypermethylated CpG loci compared with normal samples, as well as differing patterns in relation to specific clinical features. The hypermethylated CpG loci identified in our study include both previously reported and novel methylated genes that fall into a wide range of functional categories. Tumors with lymph node metastasis, ER<sup>+</sup>/PR<sup>+</sup> tumors and patients that had relapsed showed overall higher methylation versus patients with none of the above. There were 3 distinct methylation profiles generated within the tumor samples, ranging from the high methylation to intermediate to low methylation clusters. Although patient survival was not significantly different among the clusters ( $P > 0.05$ ), statistically significant ( $P = < 0.05$ ) variation among the clusters was found

for tumor relapse, ER/PR status and lymph node metastasis. Examination of individual loci/genes revealed that (i) methylation of 6 genes (*RECK*, *SFRP2*, *ITR*, *UGT3A1*, *ACADL*, *UAP1L1*) significantly associated with worst survival; (ii) methylation of 18 genes/loci (*B3GAT1*, *TNFRSF10D*, *GRIA4*, *C10RF114*, *DNAJC6*, *C6ORF145*, *CORIN*, *LAMA2*, *ITR*, *MCAM*, *C12ORF34*, *NPY*, *COLIA2*, *RSNL2*, *ZNF660*, *ADAMTSL1*, *DPP4*, *MAT1A*) associated with ER<sup>+</sup>/PR<sup>+</sup> receptor status; and (iii) methylation of 2 genes (*LDHD* and *ZNF660*) associated with lymph node-positive tumor status.

To show biological relevance for these methylated genes, we next examined the relationship between gene expression and gene methylation for 2 genes most associated with poor survival (*RECK* and *ACADL*). In both cases, expression analysis showed that methylated breast tumors had lower gene expression versus normal breast tissue samples. We confirmed tumor-specific and/or increased methylation in tumors compared with corresponding normal breast tissue for *RECK* and *ACADL* in an independent set of 20 breast normal/tumor-paired DNA. *RECK* is a membrane-anchored metalloprotease regulator, expression of *RECK* is frequently reduced in various tumor types and reduced expression correlates with worst prognosis. Forced expression of *RECK* in malignant cells and in mouse xenograft models leads to reduced angiogenesis, invasion and reduced metastasis (13, 14). *RECK*<sup>-/-</sup> mice die around embryonic day 10.5 with reduced tissue integrity and defects in various organs including blood vessels and central nervous system. *RECK* has been shown to be methylated in NSCLC, gastric cancer, oral squamous cell carcinoma, colon cancer, and melanoma. *RECK* methylation is associated with poor prognosis and with lymph node metastasis in oral squamous cell carcinoma and NSCLC cancer, respectively (15, 16). But to our knowledge this is first report of *RECK* methylation in breast cancer and that methylated *RECK* is significantly associated with relapse and worst survival.

Effect of nutrition and lifestyle choices have been recognized to influence breast cancer risk and survival but there is a lack of studies investigating these breast cancer risk factors and tumor DNA methylation. A very recent study showed that not only breast tumor size but also alcohol and folate intake were associated with DNA methylation profiles (17).

Hence, it is interesting that we found *ACADL* (acyl-CoA dehydrogenase, long chain) methylation is associated with poor survival in our cohort of breast cancer patients. The protein encoded by the *ACADL* gene belongs to a family of mitochondrial flavoenzymes that are involved in fatty acid and branched-chain amino acid metabolism. This protein is 1 of the 4 enzymes that catalyze the initial step of mitochondrial  $\beta$ -oxidation of straight-chain fatty acid. Yanagitani and colleagues, showed that transgenic mice expressing retinoic acid receptor- $\alpha$ -dominant-negative form in hepatocytes exhibited severe fatty acid deposition (steatohepatitis) and liver tumors (18). The expression of mitochondrial enzymes involved in fatty acid  $\beta$ -oxidation, including CAPT II, VLCAD, *ACADL*, and HCD was greatly reduced in the transgenic mice compared with the wild-type mice, thus hepatic steatosis was induced in transgenic mice due to impaired  $\beta$ -oxidation of fatty acids in mitochondria. In addition, Wnt signaling was activated in

transgenic mice. Histologic and biochemical abnormalities were reversed and occurrence of liver tumors was inhibited by feeding the transgenic mice with a high retinoic acid diet. The authors suggest that retinoic acid plays an important role in preventing hepatocellular carcinoma in association with fatty acid metabolism and Wnt signaling (18).

We also showed that *SFRP2* methylation is significantly associated with poor survival. Secreted frizzled proteins (SFRP) act as soluble modulators of Wnt signaling. There are 5 SFRP genes, and the gene family has been shown to be methylated in various cancers including breast cancer. Methylation of *SFRP1* and *SFRP5* in breast cancer is associated with unfavorable prognosis (5, 19). *SFRP2* methylation has previously been reported in breast, bladder, colon, liver, lung, esophagus, stomach cancer and leukemia. Although detection of *SFRP2* methylation in fecal DNA has been shown to be a marker for colorectal cancer and for detection of colorectal polyps (20–22), in breast cancer *SFRP2* methylation was not associated with clinical patient outcome or other clinicopathologic factors, but loss of SFRP2 protein expression was weakly linked to poor patient overall survival (23). Overexpression of SFRP2 in mammary cells leads to inhibition in cell proliferation. A recent article linked methylation of *sFRP2* and *sFRP5* with adverse clinical outcome in patients with normal karyotype acute myeloid leukemia (24). In this report we have shown that methylation of *SFRP2* in breast cancer is associated with poor patient survival.

The remaining 3 genes [*UGT3A1* (UDP glycosyltransferase-3 family, polypeptide A1); *UAP1L1* (UDP-N-acetylglucosamine pyrophosphorylase 1-like 1); *ITR/GPRI80* (G protein-coupled receptor 180)] associated with survival in our cohort of breast cancer patients are novel genes and to our knowledge epigenetic inactivation of these genes have not been described previously in cancer; further work will be required to understand their role in breast cancer development.

In a recent genome-wide DNA methylation study in Korean breast cancer patients, hyper/hypomethylation of 5 genes was linked to ER<sup>+</sup>/PR<sup>+</sup> status (25). We have identified 18 hypermethylated genes associated with ER<sup>+</sup>/PR<sup>+</sup> status in Caucasian breast cancer patients. Li and colleagues identified *FAM124B* and *ST6GALNAC1* hypermethylation associated with ER<sup>+</sup>/PR<sup>+</sup> status in Korean breast cancer patients (25). These 2 genes were not linked to ER/PR status in our cohort of breast tumor cases from Caucasian individuals. This could be because of the ethnical differences between the 2 sets of breast cancer patients; in addition, we used a high-stringency criterion to identify differentially methylated genes in tumor versus normal DNA. Hence, according to our selection criteria the above 2 genes were not differentially methylated in tumor versus normal DNA.

We identified *TNFRSF10D* (*DCR2*) gene among the 18 genes/loci significantly hypermethylated in ER<sup>+</sup>/PR<sup>+</sup> breast tumors compared with receptor-negative breast tumors. *TNFRSF10D* (*DCR2*), the protein encoded by this gene is a member of the TNF-receptor superfamily. TNF-related apoptosis-inducing ligand TRAIL can induce apoptosis in cancer cells but has little effect on normal cells, hence it is a promising anticancer agent. However, cancer cells are often resistant to TRAIL-induced

cell death. It was recently shown that depletion of adenine nucleotide translocase-2 in breast cancer cells sensitized these cells to TRAIL-induced apoptosis by activating *c-jun* NH kinase and modulating the expression of TRAIL receptors, including downregulation of DCR2 expression both in *in vitro* and *in vivo* studies (26). *DCR2* methylation has been previously reported in several cancers including NSCLC, breast cancer, and prostate cancer (27). *COLIA2* methylation was also associated with ER/PR receptor status in our breast cancer cohort, was cancer specific, and could be reversed in methylated breast tumor cell lines following treatment with 5-azadC. Using computational protocols, Loss and colleagues recently showed *COLIA2* methylation in breast tumor cell lines (28). In medulloblastoma *COLIA2* methylation was found to be subgroup specific, and in bladder cancer *COLIA2* gene inactivation due to methylation was shown to contribute to cell proliferation and migration of bladder cancer cells (29, 30).

In summary, we have identified a set of genes hypermethylated in breast cancer that are linked to patient clinical

outcome, confirmation of our results in a larger breast cancer cohort and subsequent biological analysis of the proteins encoded by these genes may lead to a better understanding of breast cancer pathogenesis and development of clinical applications.

### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

### Grant Support

F. Latif's laboratory was funded by Breast Cancer Campaign and Cancer Research UK.

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 November 5, 2010; revised January 28, 2011; accepted February 18, 2011; published OnlineFirst March 1, 2011.

### References

1. Esteller M. Epigenetics in cancer. *N Engl J Med* 2008;358:1148–59.
2. Jones PA, Baylin SB. The epigenomics of cancer. *Cell* 2007;128:683–92.
3. Coughlin SS, Ekwueme DU. Breast cancer as a global health concern. *Cancer Epidemiol* 2009;33:315–8.
4. Hill VK, Hesson LB, Dansranjav T, Dallol A, Bieche I, Vacher S, et al. Identification of 5 novel genes methylated in breast and other epithelial cancers. *Mol Cancer* 2010;9:51.
5. Veeck J, Geisler C, Noetzel E, Alkaya S, Hartmann A, Knüchel R, et al. Epigenetic inactivation of the secreted frizzled-related protein-5 (SFRP5) gene in human breast cancer is associated with unfavorable prognosis. *Carcinogenesis* 2008;29:991–8.
6. Osanai M, Murata M, Chiba H, Kojima T, Sawada N. Epigenetic silencing of claudin-6 promotes anchorage-independent growth of breast carcinoma cells. *Cancer Sci* 2007;98:1557–62.
7. Wu Q, Liu Y, Ren Y, Xu X, Yu L, Li Y, et al. Tight junction protein, claudin-6, downregulates the malignant phenotype of breast carcinoma. *Eur J Cancer Prev* 2010;19:186–94.
8. Gaudet MM, Hunter K, Pharoah P, Dunning AM, Driver K, Lissowska J, et al. Genetic variation in SIPA1 in relation to breast cancer risk and survival after breast cancer diagnosis. *Int J Cancer* 2009;124:1716–20.
9. Virmani AK, Rathi A, Sathyanarayana UG, Padar A, Huang CX, Cunnigham HT, et al. Aberrant methylation of the adenomatous polyposis coli (APC) gene promoter 1A in breast and lung carcinomas. *Clin Cancer Res* 2001;7:1998–2004.
10. Evron E, Umbricht CB, Korz D, Raman V, Loeb DM, Niranjani B, et al. Loss of cyclin D2 expression in the majority of breast cancers is associated with promoter hypermethylation. *Cancer Res* 2001;61:2782–7.
11. Esteller M, Corn PG, Urena JM, Gabrielson E, Baylin SB, Herman JG. Inactivation of glutathione S-transferase P1 gene by promoter hypermethylation in human neoplasia. *Cancer Res* 1998;58:4514–8.
12. Ramaswamy B, Majumder S, Roy S, Ghoshal K, Kutay H, Datta J, et al. Estrogen-mediated suppression of the gene encoding protein tyrosine phosphatase PTPRO in human breast cancer: mechanism and role in tamoxifen sensitivity. *Mol Endocrinol* 2009;23:176–87.
13. Takahashi C, Sheng Z, Horan TP, Kitayama H, Maki M, Hitomi K, et al. Regulation of matrix metalloproteinase-9 and inhibition of tumor invasion by the membrane-anchored glycoprotein RECK. *Proc Natl Acad Sci U S A* 1998;95:13221–6.
14. Oh J, Takahashi R, Kondo S, Mizoguchi A, Adachi E, Sasahara RM, et al. The membrane-anchored MMP inhibitor RECK is a key regulator of extracellular matrix integrity and angiogenesis. *Cell* 2001;107:789–800.
15. Long NK, Kato K, Yamashita T, Makita H, Toida M, Hatakeyama D, et al. Hypermethylation of the RECK gene predicts poor prognosis in oral squamous cell carcinomas. *Oral Oncol* 2008;44:1052–8.
16. Chang HC, Cho CY, Hung WC. Downregulation of RECK by promoter methylation correlates with lymph node metastasis in non-small cell lung cancer. *Cancer Sci* 2007;98:169–73.
17. 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.
18. Yanagitani A, Yamada S, Yasui S, Shimomura T, Murai R, Murawaki Y, et al. Retinoic acid receptor alpha dominant negative form causes steatohepatitis and liver tumors in transgenic mice. *Hepatology* 2004;40:366–75.
19. Veeck J, Niederacher D, An H, Klopocki E, Wiesmann F, Betz B, et al. Aberrant methylation of the Wnt antagonist SFRP1 in breast cancer is associated with unfavourable prognosis. *Oncogene* 2006;25:3479–88.
20. Müller HM, Oberwalder M, Fiegl H, Morandell M, Goebel G, Zitt M, et al. Methylation changes in faecal DNA: a marker for colorectal cancer screening? *Lancet* 2004;363:1283–5.
21. Oberwalder M, Zitt M, Wöntner C, Fiegl H, Goebel G, Zitt M, et al. SFRP2 methylation in fecal DNA—a marker for colorectal polyps. *Int J Colorectal Dis* 2008;23:15–9.
22. Nagasaka T, Tanaka N, Cullings HM, Sun DS, Sasamoto H, Uchida T, et al. Analysis of fecal DNA methylation to detect gastrointestinal neoplasia. *J Natl Cancer Inst* 2009;101:1244–58.
23. Veeck J, Noetzel E, Bektas N, Jost E, Hartmann A, Knüchel R, et al. Promoter hypermethylation of the SFRP2 gene is a high-frequency alteration and tumor-specific epigenetic marker in human breast cancer. *Mol Cancer* 2008;7:83.
24. Griffiths EA, Gore SD, Hooker C, McDevitt MA, Karp JE, Smith BD, et al. Acute myeloid leukemia is characterized by Wnt pathway inhibitor promoter hypermethylation. *Leuk Lymphoma* 2010;51:1711–9.
25. Li L, Lee KM, Han W, Choi JY, Lee JY, Kang GH, et al. Estrogen and progesterone receptor status affect genome-wide DNA methylation profile in breast cancer. *Hum Mol Genet* 2010;19:4273–7.
26. Jang JY, Jeon YK, Choi Y, Kim CW. Short-hairpin RNA-induced suppression of adenine nucleotide translocase-2 in breast cancer cells restores their susceptibility to TRAIL-induced apoptosis by

- activating JNK and modulating TRAIL receptor expression. *Mol Cancer* 2010;9:262.
27. Shivapurkar N, Toyooka S, Toyooka KO, Reddy J, Miyajima K, Suzuki M, et al. Aberrant methylation of trail decoy receptor genes is frequent in multiple tumor types. *Int J Cancer* 2004;109:786–92.
28. Loss LA, Sadanandam A, Durinck S, Nautiyal S, Flaucher D, Carlton VE, et al. Prediction of epigenetically regulated genes in breast cancer cell lines. *BMC Bioinformatics* 2010;11:305.
29. Anderton JA, Lindsey JC, Lusher ME, Gilbertson RJ, Bailey S, Ellison DW, et al. Global analysis of the medulloblastoma epigenome identifies disease-subgroup-specific inactivation of COL1A2. *Neuro Oncol* 2008;10:981–94.
30. Mori K, Enokida H, Kagara I, Kawakami K, Chiyomaru T, Tatarano S, et al. CpG hypermethylation of collagen type I alpha 2 contributes to proliferation and migration activity of human bladder cancer. *Int J Oncol* 2009;34:1593–602.

# Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

## Genome-Wide DNA Methylation Profiling of CpG Islands in Breast Cancer Identifies Novel Genes Associated with Tumorigenicity

Victoria K. Hill, Christopher Ricketts, Ivan Bieche, et al.

*Cancer Res* Published OnlineFirst March 1, 2011.

<b>Updated version</b>	Access the most recent version of this article at: doi: <a href="https://doi.org/10.1158/0008-5472.CAN-10-4026">10.1158/0008-5472.CAN-10-4026</a>
<b>Supplementary Material</b>	Access the most recent supplemental material at: <a href="http://cancerres.aacrjournals.org/content/suppl/2011/03/01/0008-5472.CAN-10-4026.DC1">http://cancerres.aacrjournals.org/content/suppl/2011/03/01/0008-5472.CAN-10-4026.DC1</a>

**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](mailto:pubs@aacr.org).

**Permissions** To request permission to re-use all or part of this article, use this link <http://cancerres.aacrjournals.org/content/early/2011/04/08/0008-5472.CAN-10-4026>. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.