Epigenetic Regulation of Cancer Stem Cell Genes in Triple-Negative Breast Cancer

Naofumi Kagara,* Kelly T. Huynh,* Christine Kuo,*
Hideyuki Okano,[†] Myung Shin Sim,[‡]
David Elashoff,[§] Kelly Chong,*
Armando E. Giuliano,[¶] and Dave S.B. Hoon*

From the Departments of Molecular Oncology* and Biostatistics[‡]
and the Margie Robert Petersen Breast Center,[¶] John Wayne
Cancer Institute, Saint John's Health Center, Santa Monica,
California; the Department of Physiology,[†] Keio University of
Medicine, Tokyo, Japan; and the Department of Medicine,[§] David
Geffen School of Medicine, UCLA, Los Angeles, California

Expression of specific breast cancer stem cells (BCSCs) is seen in aggressive tumors, but their regulation is unclear. Epigenetic changes influence gene expression and are implicated in breast cancer progression. We hypothesized that promoter methylation regulates specific BCSC-related genes [*CD44*, *CD133*, *CD24*, *MSH1* (alias, *Musashi-1*), and *ALDH1*] and that this epigenetic profile can identify aggressive subtypes, such as triple-negative breast cancer (TNBC). Methylation analysis was performed using MassARRAY EpiTYPER sequencing; CpG-rich sites were identified in the promoter regions of BCSC genes, except *ALDH1*. These sites were screened by treatment with 5-aza-2'-deoxycytidine in four TN and five non-TNBC cell lines. The specific regulatory CpG site demonstrating the most significant inverse correlation between CpG site methylation and mRNA expression was identified for *CD44*, *CD133*, and *Musashi-1*, but not for *CD24*. Methylation of *CD44*, *CD133*, and *Musashi-1* was evaluated in 91 American Joint Committee on Cancer stage I to III primary breast cancer tumors, and these sites were significantly hypomethylated in TNBC versus non-TNBC. The IHC staining of primary tumors with the highest and lowest methylation levels revealed the strongest staining in hypomethylated specimens, suggesting that hypomethylation leads to gene activation. We demonstrate that methylation is a significant mechanism regulating *CD44*, *CD133*, and *Musashi-1*, and that gene hypomethylation correlates with TNBC. Assessment of epigenetic changes in BCSC genes may provide a more accurate classification of TNBC and could be developed as potential therapeutic

targets. (*Am J Pathol* 2012, 181:257–267; <http://dx.doi.org/10.1016/j.ajpath.2012.03.019>)

Triple-negative breast cancer (TNBC), defined by lack of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER2) expression, accounts for approximately 12% to 17% of breast cancer (BCa) and is a poor prognostic factor for disease-free and overall survival.¹ Patients with TNBC are at increased risk for early relapse, usually within 5 years of initial diagnosis.² Unlike tumors that overexpress hormone receptors (HRs) or are HER2⁺, TNBCs cannot be treated with hormone therapy or trastuzumab, the anti-HER2 monoclonal antibody (mAb), because they lack these specific targets^{3,4}; as a group, patients with TNBC have worse outcomes after chemotherapy than patients with other BCa subtypes.^{1,3} Recently, advances in the characterization of TNBC demonstrate that it is a more heterogeneous group than previously thought, encompassing several molecular subtypes, including basal-like breast cancer (BLBC), which is enriched in cells with stem cell-like properties.¹

Cancer stem cells (CSCs) are capable of tumorigenesis and differentiation into virulent aggressive progenies. High expression of various CSC genes reportedly correlates with the biological aggressiveness of the tumor, but it is unclear how these CSC genes are regulated. Epigenetic changes can alter gene expression without changes in DNA sequence. One of the most well-characterized epigenetic alterations in cancer is methylation of CpG islands in the promoter region of tumor-related

Supported in part by the Leslie and Susan Gonda (Goldschmied) Foundation (Los Angeles, CA) (D.H.), ABC Foundation (Beverly Hills, CA) (D.H.), Ruth and Martin H. Weil Fund (Los Angeles, CA) (D.H.), and a grant from the Susan G. Komen Breast Cancer Foundation (BCTR0707528 to D.H.).

Accepted for publication March 12, 2012.

N.K. and K.T.H. contributed equally to this work.

Supplemental material for this article can be found at <http://ajp.amjpathol.org> or at <http://dx.doi.org/10.1016/j.ajpath.2012.03.019>.

Address reprint requests to Dave S.B. Hoon, Ph.D. Department of Molecular Oncology, John Wayne Cancer Institute, 2200 Santa Monica Blvd, Santa Monica, CA 90404. E-mail: hoon@jwci.org.

genes, resulting in gene activation or silencing, and has been implicated in tumor progression.^{5,6}

Al-Hajj et al⁷ first demonstrated that the CD44⁺/CD24^{-/low} phenotype was associated with stem cell-like properties. Recent studies showed that CD44⁺/CD24^{-/low} tumor cells in BCa may be related to distant metastases⁸ and are enriched in basal-like/TNBC.^{9,10} CD133, also known as Prominin-1, is considered one of the putative breast CSC (BCSC) biomarkers, and CD133⁺ BCa cells have had high tumor-initiating capacity.¹¹ Aldehyde dehydrogenase-1 (*ALDH1*) has been reported to function as a biomarker of BCSCs, with high expression correlating to poor prognosis.^{12,13} *Musashi-1* (*MSH1*), originally described as a neural tumor cell marker, is an RNA-binding protein that regulates the translation of its target mRNA, mNumb, and other genes involved in cell cycle regulation, proliferation, and apoptosis.^{14–16} Clarke et al¹⁷ demonstrated that human breast epithelial cells with Hoechst dye-effluxing side population properties characteristic of mammary stem cells in mice expressed higher levels of *Musashi-1*, suggesting that *Musashi-1* has a role as a stem cell marker in BCa.

The aim of this study was to determine whether promoter methylation status regulates BCSC-related genes *CD24*, *CD44*, *CD133*, *ALDH1*, and *Musashi-1*. We hypothesized that the methylation status levels of these BCSC genes could serve as surrogate biomarkers for identifying TNBC. Our results demonstrated that epigenetic regulation governs BCSC genes *CD44*, *CD133*, and *Musashi-1* and that hypomethylation of these genes leads to gene activation. In addition, TNBC specimens were hypomethylated for *CD44*, *CD133*, and *Musashi-1* when compared with other receptor subtypes, indicating that TNBC has a particular methylation pattern of these BCSC genes that can be detected and thereby used as biomarker surrogate of their TN status.

Materials and Methods

BCa Cell Lines and Treatment

Four TNBC cell lines (MDA-MB-231, BT-549, BT-20, and HCC1937)^{18–20} and five non-TNBC cell lines (MCF-7, T47D, ZR-75-1, ZR-75-30, and SK-BR-3)^{18–20} were obtained from ATCC (Manassas, VA) and maintained in an appropriate medium (Gemini Bio-Products, West Sacramento, CA) and conditions as recommended. For demethylation studies, cultured cells were treated with 5 μ mol/L 5-aza-2'-deoxycytidine (5-aza; Sigma-Aldrich, St Louis, MO) for 72 hours, with media changed every 24 hours, as previously described.²¹ Control cells were incubated in dimethyl sulfoxide only under the same conditions.

Clinical Specimens

Ninety-one patients who underwent surgical resection without neoadjuvant treatment for American Joint Committee on Cancer stage I to III invasive ductal carcinoma between 2003 and 2008 were selected for analysis from our breast cancer database. Western Institutional Review

Board approval was obtained to analyze paraffin-embedded archival tissue (PEAT) specimens of primary tumors. All ER, PR, and HER2 expression was determined by a qualified referral diagnostic laboratory. In this study, HR⁺ specimens were either ER⁺ or PR⁺ and HR⁻ specimens were ER⁻ and PR⁻. According to guidelines by the American Society of Clinical Oncology and College of American Pathologists, HER2 was determined by immunohistochemistry (IHC) and scored from 0 to 3+. A score of 2+ was defined as intermediate and retested by fluorescence *in situ* hybridization (Clariant, Inc., Aliso Viejo, CA); when the fluorescence *in situ* hybridization ratio was ≥ 2.2 , it was considered HER2 positive.²²

RT-qPCR

Total RNA was extracted from cell lines using TRI-Reagent (Molecular Research Center, Inc., Cincinnati, OH), as previously described.²³ RNA was quantified and assessed for purity by UV spectrophotometry and the Quant-iT RiboGreen RNA Assay Kit (Invitrogen, Carlsbad, CA). Total RNA, 1 μ g, was used for reverse transcription, and quantitative RT-PCR (RT-qPCR) was performed as previously described.²³ The PCR reactions were performed on the ABI 7900HT Fast Real-Time PCR System (Applied Biosystems, Carlsbad, CA) at 95°C (15 minutes), followed by 30 cycles of 95°C (30 seconds), 57°C (30 seconds), and 72°C (30 seconds). Primer and probe sequences are listed in Table 1. *GAPDH* was used as a reference gene. Amplification of the target gene was normalized to the reference gene, and each assay was performed in duplicate.

Flow Cytometry Analysis

Cells were trypsinized and washed in PBS with 0.5% fetal bovine serum (Gemini Bio-Products) and stained with phosphatidylethanolamine anti-human CD44 (mouse mAb clone 515; BD Pharmingen, San Diego, CA) or phosphatidylethanolamine anti-human CD133 (human mAb clone 293C3; Miltenyi Biotec, Auburn, CA). Mouse IgG1 κ (BD Pharmingen) and mouse IgG2b (Miltenyi Biotec) were used as isotype controls for anti-CD44 and anti-CD133, respectively, in accordance with the manufacturer's instructions. Ab for *Musashi-1* was not available

Table 1. RT-qPCR Primer Sequences

Gene	Primer	Primer sequence
CD44	Forward	5'-CAGTGAAGGAGCAGCACTT-3'
	Reverse	5'-TGAATGTGCTTGGTCTCTG-3'
	Probe	5'-AGACGAAGACAGTCCCTGGATCACC-3'
CD133	Forward	5'-AGCTACTTGGCTCAGACTGG-3'
	Reverse	5'-TGCACTCTCTTTTCAGGGAGT-3'
	Probe	5'-CCCCGCAGGAGTGAATCTTTTATCA-3'
Musashi-1	Forward	5'-CCAGCCGGAGTTATACAGG-3'
	Reverse	5'-CAGTGAGAGGAATGGCTGTAA-3'
	Probe	5'-TACCAGTTCCCGAATTCGCTGTAG-3'
CD24	Forward	5'-CCTCCCAGAGTACTTCCAAC-3'
	Reverse	5'-AGTGAGACCACGAAGAGACTG-3'
	Probe	5'-AAATCCAACCTAATGCCACCACCAAG-3'

for fluorescence-activated cell sorter (FACS) analysis; therefore, confirmation of protein expression for *Musashi-1* was performed using IHC. Unsorted cells were incubated with Abs for 30 minutes at 4°C in the dark, washed twice, and analyzed by FACS Calibur flow cytometer (BD Biosciences, San Jose, CA) with an acquisition of 10⁴ events for each sample. All data were analyzed by BD CellQuest software (BD Biosciences, Franklin Lakes, NJ).

DNA Extraction and Bisulfite Modification

Genomic DNA from cell lines was isolated using DNAzol (Molecular Research Center, Inc.), according to the manufacturer's recommendation. DNA from the cell lines was then quantified using the PicoGreen kit (Invitrogen) and bisulfite modified with the EpiTect Bisulfite Kit (Qiagen, Valencia, CA). Sections (7 μm thick) of PEAT were stained with H&E and then cancer cell populations were accurately laser capture microdissected (Arcturus Laser Capture Microdissection; Applied Biosystems). On-cap bisulfite modification was then performed as previously described.²⁴

Methylation Analysis

Bisulfite-treated DNA was amplified by PCR using AccuStart TaqDNA Polymerase (Quanta Biosciences, Gaithersburg, MD) at 94°C for 15 minutes, followed by 45 cycles of 94°C (20 seconds), 60°C (30 seconds), and 72°C (60 seconds). Promoter sequences for each gene were obtained using the UCSC Genome Bioinformatics Site (<http://genome.ucsc.edu>, last accessed January 5, 2011). The primer sets were designed using the free online software EpiDesigner (<http://www.epidesigner.com>, last accessed January 5, 2011; Sequenom, San Diego, CA), as previously described.²⁴ The EpiDesigner identifies CpG islands in the promoter region; this information was then used to calculate the percentage of CpG content for each promoter. The primer sequences are

shown in Table 2. A T7 promoter tag for mRNA transcription was added to the reverse primer and a 10-mer tag was added on the forward primer to adjust for melting temperature and mass differences in mass spectrometry. For screening of the promoter region for the regulatory CpG sites in the cell lines, primer sets 1 to 3 for *CD44*; 1 to 4 for *CD24*; 1, 2, and 4 for *CD133*; and 1 and 2 for *Musashi-1* were used (Table 2). The primer sets used for methylation analysis of the PEAT clinical specimens were designed to target specific CpG sites: CD44 CpG 3, CD133 CpG 25, and Musashi-1 CpG 9 (Table 2; primer sets CD44 4, CD133 3, and Musashi-1 3). These sites demonstrated the best inverse correlation between methylation level and mRNA expression by screening analysis of the cell lines. The primer sets for the clinical specimens were designed for shorter PCR products (<150 bp) because older archival tissue DNA can become fragmented.

We used a novel approach for high-throughput quantitative DNA methylation analysis based on matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry, which uses a base-specific cleavage reaction combined with mass spectrometric analysis.^{24,25} The T-cleavage reaction was performed using the Mass Cleave T Cleavage kit following the manufacturer's instructions (Sequenom). The samples were desalted using a clean resin kit and spotted on a 384-pad SpectroCHIP using the nanodispenser, followed by spectral acquisition on MassARRAY. The results were analyzed using EpiTyper software version 1.0 (Sequenom). The analysis quantifies the percentage of methylation corresponding to the CpG sites within the PCR amplicon. All experiments were performed in triplicate.

IHC Data

Sections (4 μm thick) of the PEAT were deparaffinized in xylene and dehydrated in ethanol. Antigen retrieval was achieved by placing slides in citrate buffer (pH 6.0) or in 1 mmol/L EDTA solution (pH 9.0) at 100°C in a water bath

Table 2. Primers for MassARRAY Methylation Analysis

Gene	n	Forward primer sequence*	Reverse primer sequence*	Length (bp)	No. of CpGs covered
<i>CD44</i>	1	5'-AGTATGTGTGTGGAGAGAGGTGTTT-3'	5'-AATTCACCTTTAACCTCTCCTTTC-3'	378	4
	2	5'-AAAGGAGAGGTTAAAGGTTGAATTT-3'	5'-AAACACACCCCAACAAAAAACTA-3'	307	12
	3	5'-ATAGTTTTTTTTTGTGTTGGGTGTGTT-3'	5'-CAAACAACCTCACTTAACCTCCAATCC-3'	419	13
	4	5'-TTTGGGTTTTATAGGATGTTGGGATA-3'	5'-CCCTCACTCCCCACTATAAACAC-3'	106	3
<i>CD133</i>	1	5'-GGTGAGTATGTTTAAGGAATTTTTTTT-3'	5'-TCACTATACACCCCAATACAATAA-3'	412	19
	2	5'-TATTTGTGAGGGGTTAGGGAGG-3'	5'-CACTCCTTCCACTATACTAAAAATATACAA-3'	106	3
	3	5'-TATTTAGATTAAAAAGTTTGGGTTGGA-3'	5'-TCCCTAACCCCTCAACAAATAATAC-3'	115	4
	4	5'-AAATTTTGTAGTTTGTAGTGGTGGTT-3'	5'-AAACCATCCCTAAAAATTTCTTTTAC-3'	247	11
<i>Musashi-1</i>	1	5'-GGAGAGAATTAGGGGAGATTTTTTA-3'	5'-TCTTCCCTCTCAATCCCTACTATC-3'	490	10
	2	5'-GAGATTGGGGTTTTTTTAAATTTTG-3'	5'-AACAAACCATACTACCCCTCC-3'	416	11
	3	5'-AGTATGTGGGATTGGGGAGTAT-3'	5'-TCTCAAATCCCTACTATCCAAAATTA-3'	92	2
<i>CD24</i>	1	5'-TTAATTTTGAGGGGATTTTTT-3'	5'-TCCACCAAAATCTAAAAATAATAA-3'	295	9
	2	5'-GAGGGGTTTTTTTAGGGTTG-3'	5'-TCCTAAACAAATACATTACCACTCAA-3'	523	20
	3	5'-TTTGTGTTGGAGTAAGTGATTTGTT-3'	5'-AAAAAATCCCTCAAAATTAACCC-3'	527	21
	4	5'-GTTTAGTAGGATGTTGGGTGTTTG-3'	5'-TCTAAATAACAATACACTTACTCCAAA-3'	539	13

*The tags AGGAAGAGAG and CAGTAATACGACTCACTATAGGGAGAAGGCT were added to the 5' ends of forward and reverse primers, respectively, to adjust for mass differences in mass spectrometry.

for 10 to 20 minutes, based on the manufacturer's recommendations. The sections were incubated with anti-human CD44 Ab [clone 156-3C11, mouse mAb (1:100); Thermo Scientific, Waltham, MA], anti-human CD133 Ab [clone AC133, mouse mAb (1:10); Miltenyi Biotec], anti-human Musashi-1 Ab [clone EP1302, rabbit mAb (1:300); Abcam], anti-human cytokeratin (CK) 5/6 Ab [clone D5/16B4, mouse mAb (1:100); Dako, Carpinteria, CA], and anti-human CK14 Ab [clone VP-C410, mouse mAb (1:100); Vector Laboratories, Burlingame, CA] at 4°C overnight. The Tyramide Signal Amplification Kit (Invitrogen) was used for staining of Musashi-1, the CSA II Biotin-Free Catalyzed Amplification System (Dako) for staining CD133, CK5/6, and CK14, and the LSAB+System-HRP (Dako) and diaminobenzidine peroxidase substrate kit (Vector Laboratories) for staining CD44 on PEAT sections. All slides were counterstained with hematoxylin.

Biostatistical Analysis

The Student's *t*-test was used to compare methylation rates with respect to various clinicopathological parameters for each BCSC gene. The Pearson's correlation coefficient for each CpG site in the promoter region of the BCSC genes was calculated to determine which CpG site demonstrated the strongest correlation between methylation status and mRNA expression. Tukey's test was used with a 5% composite type I error to determine whether the difference in the methylation status of the BCSC genes was significantly different between TNBC and other BCa subtypes. The receiver operating characteristic (ROC) curve and area under the curve were calculated for predicting TNBC, as previously described.²⁶

Logistic regression models based on the methylation status of *CD44*, *CD133*, and/or *Musashi-1* were constructed to predict the probability of a BCa being TN versus non-TN. The Akaike information criterion was used in comparing nonnested models, and the likelihood ratio test was used for the comparison of nested models. An Akaike information criterion is a measure of the goodness of fit of an estimated statistical model.²⁷ Models were compared to determine whether prediction of the TNBC phenotype should be based on the methylation status of one, two, or three BCSC genes. $P < 0.05$ was considered statistically significant. All analyses used SAS 9.1.3 (SAS Inc., Cary, NC), and ROC curves were constructed using Splus (TIBCO Software, Palo Alto, CA).

Results

Methylation Status of BCSC Genes

Sequence analysis of *CD44*, *CD24*, *CD133*, *ALDH1*, and *Musashi-1* promoter regions using Sequenom's Epidesigner software revealed CG-rich regions in *CD44*, *CD24*, *CD133*, and *Musashi-1* (CG% = 66%, 66%, 68%, and 72%, respectively), but not in *ALDH1* (CG% = 37%) (see Supplemental Figure S1 at <http://ajp.amjpathol.org>). Therefore, *ALDH1* was not further assessed in the methylation analysis because it does not appear to be directly regulated by promoter region CpG islands.

To compare methylation versus expression of individual BCSC genes, we analyzed the methylation status of each CpG site in the CSC gene promoter regions of BCa cell lines by MassARRAY EpiTYPER analysis (Sequenom). mRNA expression of *CD44*, *CD24*, *CD133*, and

Figure 1. Pearson's correlation coefficient: mRNA expression versus individual CpG island methylation. Pearson's correlation analysis demonstrating the most significant inverse correlation between methylation status and mRNA expression at CpG site 3 of *CD44* (-0.925), CpG site 25 of *CD133* (-0.792), and CpG site 9 of *Musashi-1* (-0.875). The promoter region of *CD24* was uniformly unmethylated in all cell lines, regardless of mRNA expression levels.

Table 3. Assessment of BCSC Gene mRNA Expression after 5-aza Treatment

Cell line	<i>CD44</i>	<i>CD133</i>	<i>Musashi-1</i>
MDA-MB-231	↑	↑	↑
BT-549	NC	↑	↑
BT-20	↑	↑	↑
HCC1937	↑	↑	↑
MCF7	↑	↑	↑
T47D	↑	↑	↑
ZR75-1	↑	↑	NC
SKBR3	↑	NC	↑
ZR75-30	NC	NC	NC

Average values from triplicate experiments were normalized to the reference gene, *GAPDH*.

NC, no change; ↑, any increase greater than a onefold response.

Musashi-1 was assessed by RT-qPCR in the BCa cell lines. For each individual BCSC gene, mRNA expression was analyzed in conjunction with the methylation status of specific CpG islands to identify which site(s) regulated gene expression. We identified and prioritized the most active methylated CpG site(s) for individual BCSC genes that best correlated with expression. We found that CpG site 3 of *CD44* (*CD44* CpG 3), CpG site 25 of *CD133* (*CD133* CpG 25), and CpG site 9 of *Musashi-1* (*Musashi-1* CpG 9) showed the most significant inverse correlation between methylation status and mRNA expression (Pearson's correlation coefficient = -0.925 , -0.792 , and -0.875 , respectively) (Figure 1). In contrast, the promoter region of *CD24* was uniformly unmethylated in all cell lines regardless of mRNA expression levels (Figure 1).

Up-Regulation of BCSC Genes by Demethylation

When the BCa cell lines ($n = 9$) were demethylated with 5-aza, the change in mRNA expression was variable among cell lines. Up-regulation of mRNA expression of the BCSC genes after 5-aza treatment occurred in all of the cell lines, except for BT-549 and ZR75-30 for *CD44*, SKBR3 and ZR75-30 for *CD133*, and ZR75-1 and ZR75-30 for *Musashi-1* (Table 3). For *CD44*, the average fold increase ranged from 1.2-fold with MDA-MB-231 to 6.4-fold with T47D. Up-regulation of mRNA expression for *CD133* ranged from an average of 1.3-fold with HCC1937 to 19.4-fold with MDA-MB-231 and from 1.2-fold with MCF7 to 12.1-fold with MDA-MB-231 for *Musashi-1*. Overall, the increase in mRNA expression after 5-aza treatment was most robust in cell lines with the lowest baseline mRNA expression. The results suggested that the BCSC genes responded variably to 5-aza treatment for individual cell lines. However, overall, there was a consistent up-regulation of gene expression. Demethylation of T47D and MDA-MB-231 BCa cell lines with 5-aza was associated with increased hypomethylation of *CD44* CpG 3, *CD133* CpG 25, and *Musashi-1* CpG 9 (Figure 2; see also Supplemental Figure S2 at <http://ajp.amjpathol.org>).

The effect of 5-aza treatment on BCSC protein re-expression was also assessed. In concordance with the RT-qPCR results, FACS analyses showed a marked increase of *CD44* and *CD133* expression in 5-aza-treated T47D cells compared with the control cells (Figure 3A). To confirm protein expression, IHC analysis revealed in-

Figure 2. Epigrams demonstrating methylation of individual CpG sites in breast cancer cell lines before and after 5-aza treatment. Demethylation of T47D and MDA-MB-231 BCa cell lines with 5-aza is associated with increased hypomethylation of *CD44* CpG 3 (A), *CD133* CpG 25 (B), and *Musashi-1* CpG 9 (C), as demonstrated by MassARRAY. Open circles, sites without CpG islands.

Figure 3. Assessment of BCSC gene activation after 5-aza treatment. **A:** The protein expression of CD44 and CD133 on T47D cells untreated or treated with 5-aza is evaluated by FACS analysis. Dotted, untreated cells; bold line, 5-aza-treated cells; and gray area, isotype controls. **B:** Musashi-1 expression is evaluated by IHC of MDA-MB-231 cells before and after 5-aza treatment.

creased expression of Musashi-1 in 5-aza-treated MDA-MB-231 cells (Figure 3B).

DNA Methylation of BCSC Genes in BCa Tissues

Of the 91 PEAT primary tumor specimens, 22 were HR⁺/HER2⁻, 23 were HR⁺/HER2⁺, 14 were HR⁻/HER2⁺, and 32 were TN. MassARRAY primer sets covering the regulatory CpG site that demonstrated the most significant inverse correlation between mRNA expression and methylation level for *CD44* (CpG 3), *CD133* (CpG 25), and *Musashi-1* (CpG 9) were used in analysis of the clinical PEAT specimens. The methylation status of these BCSC genes was correlated to clinicopathological parameters (Table 4). By univariate analysis, methylation status was significantly correlated to HR and HER2 status for *CD133* ($P = 0.003$ and $P \leq 0.0001$, respectively) and *Musashi-1* ($P = 0.01$ and $P = 0.04$, respectively) and only to HER2 status for *CD44* ($P = 0.0002$), but all three were significantly correlated to TN status (*CD44*, $P = 0.0008$; *CD133*, $P = 0.0006$; *Musashi-1*, $P = 0.0009$). By multivariate analysis, all three BCSC genes significantly correlated with TNBC (Table 5).

Table 4. Comparison of Patient Clinicopathological Parameters with BCSC Gene Methylation Status

Parameters	N	Methylation level					
		<i>CD44</i>		<i>CD133</i>		<i>Musashi-1</i>	
		Mean $\pm$ SD (%)	P value	Mean $\pm$ SD (%)	P value	Mean $\pm$ SD (%)	P value
Age (years)							
<50	30	40.5 $\pm$ 20.5		24.2 $\pm$ 23.9		24.5 $\pm$ 17.1	
≥ 50	61	38.3 $\pm$ 22.2		26.2 $\pm$ 24.8		21.0 $\pm$ 13.7	
Tumor size (cm)							
≤ 2	52	38.2 $\pm$ 21.1		21.7 $\pm$ 22.4		21.9 $\pm$ 12.6	
>2	39	39.7 $\pm$ 22.2		30.7 $\pm$ 26.2		23.1 $\pm$ 17.5	
Lymph node metastasis							
Negative	58	36.9 $\pm$ 20.3		24.3 $\pm$ 23.9		21.4 $\pm$ 13.4	
Positive	33	42.3 $\pm$ 23.2		27.8 $\pm$ 25.5		24.1 $\pm$ 17.2	
AJCC stage							
I	41	37.7 $\pm$ 20.1		23.2 $\pm$ 23.1		21.2 $\pm$ 12.6	
II	34	36.7 $\pm$ 22.7		28.1 $\pm$ 25.3		23.7 $\pm$ 17.1	
III	16	46.3 $\pm$ 22.0		26.1 $\pm$ 26.7		22.7 $\pm$ 15.6	
HR							
Negative	46	35.5 $\pm$ 20.7		18.2 $\pm$ 19.9	0.003	18.6 $\pm$ 11.6	0.01
Positive	45	42.3 $\pm$ 21.9		33.1 $\pm$ 26.4		26.3 $\pm$ 16.8	
HER2							
Negative	54	32.2 $\pm$ 17.6	0.0002	17.5 $\pm$ 18.9	≤ 0.0001	18.8 $\pm$ 9.1	0.04
Positive	37	48.5 $\pm$ 23.1		37.3 $\pm$ 26.9		27.6 $\pm$ 19.5	
TN							
No	59	44.3 $\pm$ 22.3	0.0008	32.6 $\pm$ 26.6	0.0006	26.1 $\pm$ 16.9	0.0009
Yes	32	28.8 $\pm$ 15.6		12.5 $\pm$ 11.4		15.5 $\pm$ 5.7	
Ki-67							
<20	19	46.4 $\pm$ 25.8		23.5 $\pm$ 24.6		24.1 $\pm$ 16.5	
≥ 20	72	36.9 $\pm$ 19.9		26.1 $\pm$ 24.5		22.0 $\pm$ 14.4	
P53							
Negative	40	38.4 $\pm$ 20.1		25.0 $\pm$ 24.1		23.9 $\pm$ 15.4	
Positive	51	39.2 $\pm$ 22.6		26.0 $\pm$ 24.9		21.3 $\pm$ 14.4	

Only significant P values are shown.
AJCC, American Joint Committee on Cancer.

Table 5. BCSC Gene Methylation Status as a Predictor of HR, HER2, and TN Status

Methylated BCSC	Dependent variable					
	HR positive		HER2 positive		TN	
	Univariate OR	Multivariate OR	Univariate OR	Multivariate OR	Univariate OR	Multivariate OR
<i>CD44</i>	4.67 (0.63–34.36)*		50.43 (5.01–507.4) [†]	44.60 (3.62–549.43) [‡]	0.01 (<0.001–0.21) [‡]	0.009 (<0.001–0.22) [‡]
<i>CD133</i>	15.82 (2.29–109.39) [‡]	15.82 (2.29–109.39) [‡]	37.25 (5.05–274.59) [†]	33.1 (3.93–278.74) [‡]	0.006 (<0.001–0.13) [‡]	0.01 (<0.001–0.39) [‡]
<i>Musashi-1</i>	61.1 (1.97–>999.99) [‡]		81.81 (2.87–>999.99) [‡]		<0.001 (<0.001–0.006) [‡]	<0.001 (<0.001–0.28) [‡]

Data in parentheses are 95% CIs.

*Not statistically significant.

[†] $P < 0.001$.

[‡] $P < 0.05$.

OR, odds ratio.

Expression of BCSC Genes

We performed IHC analysis of CD44, CD133, and Musashi-1 protein expression on PEAT BCa specimens with the highest and lowest methylation level for each CSC gene ($n = 15$ in each category) to determine whether methylation status is related to gene expression. Hypomethylated tumor specimens demonstrated strongly positive IHC staining of CD44, CD133, and Musashi-1 (Figure 4A), whereas hypermethylated specimens showed negative or weak staining (Figure 4B) (Spearman rank correlation analysis: CD44, $P < 0.0001$; CD133, $P = 0.02$; Musashi-1, $P = 0.03$).

Basal Biomarker Expression in BCa Tissues

Because TNBC and BLBC share many similar features, we investigated the expression of basal biomarkers in our TNBC PEAT specimens by IHC staining with anti-human CK5/6 and anti-human CK14 Abs. Of 32 TNBC specimens, 28 (88%) were positive for either CK5/6 or CK14.

Methylation of BCSC Genes in TNBC

Methylation levels of *CD44*, *CD133*, and *Musashi-1* were significantly lower in the 32 TNBC specimens than in the

59 non-TNBC specimens (Figure 5A; $P = 0.004$, $P = 0.012$, and $P = 0.019$, respectively; see also Supplemental Figure S3 at <http://ajp.amjpathol.org>). When the specimens were further divided into subtypes based on HR and HER2 receptor status, the TNBC subtype had the lowest methylation rate for all three BCSC genes (Figure 5B). Tukey's honestly significant difference test showed a statistically significant difference for *CD44* and *CD133* methylation of TN versus HR⁺/HER2⁺ and HR⁺/HER2⁺ specimens. For *Musashi-1*, this significance was only seen between TNBC and HR⁺/HER2⁺ specimens.

Next, we developed univariate and multivariate logistic regression models that included BCSC genes *CD44*, *CD133*, and *Musashi-1* singly and collectively. Although hypomethylation of each gene was predictive of the TN phenotype, *CD133* was the best individual predictive biomarker by the Akaike information criterion. The likelihood ratio test confirmed that two-factor models (ie, *CD44/CD133*, *CD44/Musashi-1*, and *CD133/Musashi-1*) were better predictors of TNBC than any individual BCSC gene, but overall, a model that included all three BCSC genes was the best at predicting the TN subtype, with an area under the curve of 0.84 (Figure 6, A–D).

With a median follow-up of 37.5 months (range, 5 to 88 months), of the 91 patients, there were a total of 8 (8.8%) recurrences and 5 (5.5%) deaths. There was no statisti-

Figure 4. IHC analysis of BCSC genes in clinical breast cancer specimens. The clinical samples with the lowest and highest methylation rate of *CD44*, *CD133*, and *Musashi-1* were evaluated by IHC with each Ab. The numerical value of the methylation status of each gene promoter is represented in each photograph, with the MassARRAY methylation results of the corresponding CpG site represented. Representative tumors are shown that were hypomethylated (A) and demonstrated strongly positive staining of CD44, CD133, and Musashi-1 (Spearman rank correlation analysis: $P < 0.0001$, $P = 0.02$, and $P = 0.03$, respectively) and hypermethylated (B) and showed negative or weak staining. Scale bar = 100 μ m.

Figure 5. MassARRAY analyses of BCSC genes in breast cancer tissues. The methylation status of *CD44*, *CD133*, *Musashi-1*, and combined was evaluated by MassARRAY and is represented in the TN and non-TN groups (**A**) and in four subtypes (**B**) based on HR and HER2 status (HR⁺/HER2⁻, HR⁺/HER2⁺, HR⁻/HER2⁺, and HR⁻/HER2⁻). Vertical lines under each subtype represent those subtypes with statistically significant differences in the methylation status for each individual CSC gene and then as a combination of all three genes by Tukey's test. Bars and dotted bars indicate the mean value and the SD, respectively. * $P < 0.05$.

cally significant correlation between increased risk of recurrence or death in the TN group compared with non-TN patients. The methylation status of the BCSC genes in the primary tumors did not correlate with incidence of sentinel lymph node metastasis (H&E positive).

Discussion

To our knowledge, our study is the first to report that in BCa, CSC genes are regulated by the methylation status of promoter CpG regions and that hypomethylation activates BCSC genes *CD44*, *CD133*, and *Musashi-1* (*MSI1*), leading to a clinically aggressive phenotype of BCa. By using the MassARRAY system, we performed rapid, semiquantitative sequencing analysis of the entire promoter region and high-throughput analysis of multiple methylation sites.^{28–32} Primary breast tumors with the lowest methylation level demonstrated the strongest IHC staining, thereby corroborating that hypomethylation of the gene promoter region activates these BCSC genes. We found that, after demethylation with 5-aza, the BCa cell lines demonstrated a variable response in mRNA expression of the BCSC genes, likely reflecting the heterogeneous nature of BCa and responsiveness to 5-aza. Up-regulation of mRNA expression of the BCSC genes after 5-aza treatment was not always the most robust in the TNBC cell lines, possibly because of several factors, including variability in drug concentration, uptake into the cell, and drug saturation.²¹ In addition, drug toxicity or another mechanism regulating mRNA expression, such as histone modification, may be a plausible explanation why some cell lines did not demonstrate increased mRNA expression after 5-aza treatment. However, most cell lines responded to 5-aza treatment, resulting in hypomethylation of the BCSC genes, leading to increased

Figure 6. ROC analysis of the BCSC genes for patients with TNBC. ROC curves toward TNBC showing a combination of all three BCSC genes (*CD44*, *CD133*, and *Musashi-1*) (**A**) and in individual genes *CD44* (**B**), *CD133* (**C**), and *Musashi-1* (**D**). The area under the curve (AUC) and P values for each BCSC individually and in combination are shown.

gene expression, as supported by the IHC (tissue) and FACS (cell) results.

Our results also strongly support the cellular plasticity and adaptability of BCSC genes. Others have reported that partially or terminally differentiated cells dedifferentiate to generate multipotent progenitor cells with concurrent reversal of epigenetic alterations.³³ This suggests that BCSC genes may turn on and off in response to *de novo* events and signals, and implies that a differentiated cancer cell can become self-renewing when these genes are expressed. Thus, the regulation of BCSC genes appears to be a dynamic process in which passive or active (enzyme-driven) demethylation leads to tumor progression and more aggressive subtypes, such as TNBC.³⁴

Musashi-1, originally described as a neural stem cell marker,^{14,16} has recently emerged as a key regulator of BCSC functions.³⁵ However, it is not known how *Musashi-1* is regulated, and, to our knowledge, our study is the first to report an innovative mechanism of hypomethylation-activating *Musashi-1* expression and function as a BCSC gene. *Musashi-1* expression in breast epithelial cells reportedly activates Notch and Wnt signaling and mammary stem cell proliferation,³⁵ and Notch expression is up-regulated in CD44⁺ breast cancer cells.³⁶ Hao et al³⁷ showed that Notch-induced transcriptional activity is highest in ER α ⁺ and HER2⁺ cells.

Interestingly, we found that the promoter methylation of *CD44*, *CD133*, and *Musashi-1* was lower in TNBC than in other subtypes (Figure 5A; see also Supplemental Figure S3 at <http://ajp.amjpathol.org>). This is concordant with the biologically aggressive nature of TNBCs, poor overall prognosis, and limited therapeutic targets.³⁸ Often used interchangeably, TNBCs and BLBCs are not equivalent.³⁹ TNBC is determined by IHC, whereas BLBC is a molecular phenotype defined by cDNA microarray analysis.⁴⁰ The literature varies, with studies reporting approximately 50% to 90% of TNBCs as being basal-like, and approximately 80% of BLBCs as triple negative.^{41,42} BLBCs express high levels of CK5/6, 14, and 17 and/or epidermal growth factor receptor.^{40,43} In our study, 88% of TNBCs stained positive for CK5/6 or CK14. The lack of 100% concordance between TNBC and BLBC is explained by the fact that TNBC comprises other molecular subtypes of BCa, including claudin-low and interferon-rich groups.¹ Prat et al⁴⁴ reported that claudin-low tumors were enriched in stem cell-like features, including higher expression of *CD44*. BLBC also shares a similar protein expression pattern with stem/progenitor cells in the breast,⁴⁵ frequently expressing the CD44⁺/CD24[−] phenotype, indicating that BLBC and, therefore, TNBC contain more CSCs than other subtypes. A recent study by D'Anello et al⁴⁶ reported that *p53* inactivation in BLBC tumors leads to loss of methylation at *IL6*, which, in turn, triggers an autocrine loop that causes loss of methylation at *IL6*, *CD44*, and *CD133*, thereby enhancing their expression.⁴⁶ The analysis of Bloushtain-Qimron et al⁴⁷ of epigenetic programming in embryonic and mammary epithelial cells demonstrated that *CD44* was the most hypomethylated and highly expressed transcription factor with stem cell-like function, including *FOXC1*. Hypomethylation of *FOXC1* caused differentiated mammary

epithelial cells to assume a progenitor-like phenotype,⁴⁷ and our group recently reported that *FOXC1* was consistently overexpressed in BLBC and correlated with poor survival.⁴⁸

In our study, prediction of TNBC was more accurate with three BCSC genes than with one or two genes. A combination of genes accommodates a tumor's intrinsic heterogeneity, as demonstrated by de Maat et al,⁵ who reported that methylation analysis of multiple genomic loci predicted distant and local recurrence of early-stage rectal cancers. We initially analyzed five BCSC genes but subsequently excluded *ALDH1* because its promoter region did not have defined CpG islands. We also excluded *CD24* because the promoter methylation status did not correlate with mRNA expression, possibly indicating that another mechanism, such as histone modification, is regulating the increase in mRNA expression or that the regulatory CpG site is far upstream from the promoter. Given that *CD24* and *ALDH1* have been identified as BCSC genes with possible clinical significance, it is likely that these two genes are regulated by other mechanisms not identified.

Although TNBCs initially respond to standard neoadjuvant chemotherapy, particularly taxane and anthracycline-based treatment, patients with residual disease still exhibit decreased disease-free and overall survival.⁴⁹ In this study, although patients with TNBC did not have worse outcomes, most study patients were early stage (Table 4), the follow-up was limited to a median of approximately 3 years, and we did not account for differences in adjuvant treatment that the patients may have received. Therefore, our findings suggest that the methylation status of these BCSC genes might still be useful for identifying aggressive tumors and help tailor treatment accordingly. Prior studies have suggested CSC genes as possible therapeutic targets in several cancers.^{50–52} CSCs are resistant to both chemotherapy and radiation,³⁶ but targeted therapies that disrupt the tumor microenvironment, induce apoptosis, or alter the tumorigenicity of the CSCs could be effective for TNBC and other poor prognosis subgroups.

Almost 50% of patients with advanced ER α ⁺ BCa do not respond to first-line treatment with tamoxifen, and those who initially respond may acquire resistance to endocrine therapy during subsequent treatment.^{50–52} Acquired endocrine resistance might reflect the presence of an ER α [−] BCSC population that can be found within ER α ⁺ cancers that are selected out after endocrine therapy. We speculate that hypomethylation of the BCSC genes can activate this ER α [−] subpopulation of cells; if so, these BCSC genes might serve as targets to reverse drug resistance. Similarly, women with HER2⁺ metastatic BCa will initially respond when treated with trastuzumab (Herceptin); however, most will develop acquired resistance within months or years of treatment, and 15% of patients with early BCa treated with trastuzumab exhibit *de novo* resistance.^{53,54} Oliveras-Ferraro et al⁵³ recently found that in the trastuzumab-resistant cell line JIMT-1, multiple passages selected out for a population containing 80% CD44^{pos}/CD24^{neg/low} cells, which resembled the same cells found in the TN cell line MDA-MB-

231. In HER2⁺ patients, resistance may emerge to trastuzumab treatment because BCSC gene activation by hypomethylation may be involved in changes that lead to the down-regulation of the HER2 receptor, although this is yet to be elucidated.

Our study showed that gene promoter methylation is one of the mechanisms regulating BCSC genes and that certain BCSC genes are significantly hypomethylated in TNBC. These pilot studies may allow us to develop a new molecular classification system based on clinically significant epigenetic changes of BCSC genes in primary tumors. It may be the existence of this BCSC subpopulation, even within HR⁺ or HER2⁺ tumors, that, when activated by gene promoter hypomethylation, contributes to therapy resistance and, therefore, may serve as targets for new treatments.

Acknowledgments

We thank Tung Nguyen for his technical expertise on the methylation studies and Kana Sato for editorial support.

References

- Foulkes WD, Smith IE, Reis-Filho JS: Triple-negative breast cancer. *N Engl J Med* 2010, 363:1938–1948
- Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA: Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res* 2007, 13:4429–4434
- Anders CK, Carey LA: Biology, metastatic patterns, and treatment of patients with triple-negative breast cancer. *Clin Breast Cancer* 2009, 9(Suppl 2):S73–S81
- Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, Collichio F, Ollila DW, Sartor CI, Graham ML, Perou CM: The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clin Cancer Res* 2007, 13:2329–2334
- de Maat MF, van de Velde CJ, van der Werff MP, Putter H, Umetani N, Klein-Kranenburg EM, Turner RR, van Krieken JH, Bilchik A, Tollenaar RA, Hoon DS: Quantitative analysis of methylation of genomic loci in early-stage rectal cancer predicts distant recurrence. *J Clin Oncol* 2008, 26:2327–2335
- Herman JG, Baylin SB: Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med* 2003, 349:2042–2054
- Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF: Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 2003, 100:3983–3988
- Abraham BK, Fritz P, McClellan M, Hauptvogel P, Athelogou M, Brauch H: Prevalence of CD44+/CD24–/low cells in breast cancer may not be associated with clinical outcome but may favor distant metastasis. *Clin Cancer Res* 2005, 11:1154–1159
- Honeth G, Bendahl PO, Ringner M, Saal LH, Gruvberger-Saal SK, Lovgren K, Grabau D, Ferno M, Borg A, Hegardt C: The CD44+/CD24– phenotype is enriched in basal-like breast tumors. *Breast Cancer Res* 2008, 10:R53
- Klingbeil P, Natrajan R, Everitt G, Vatcheva R, Marchio C, Palacios J, Buerger H, Reis-Filho JS, Isacke CM: CD44 is overexpressed in basal-like breast cancers but is not a driver of 11p13 amplification. *Breast Cancer Res Treat* 2010, 120:95–109
- Wright MH, Calcagno AM, Salcido CD, Carlson MD, Ambudkar SV, Varticovski L: Brca1 breast tumors contain distinct CD44+/CD24– and CD133+ cells with cancer stem cell characteristics. *Breast Cancer Res* 2008, 10:R10
- Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S, Schott A, Hayes D, Birnbaum D, Wicha MS, Dontu G: ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* 2007, 1:555–567
- Morimoto K, Kim SJ, Tanei T, Shimazu K, Tanji Y, Taguchi T, Tamaki Y, Terada N, Noguchi S: Stem cell marker aldehyde dehydrogenase 1-positive breast cancers are characterized by negative estrogen receptor, positive human epidermal growth factor receptor type 2, and high Ki67 expression. *Cancer Sci* 2009, 100:1062–1068
- Kaneko Y, Sakakibara S, Imai T, Suzuki A, Nakamura Y, Sawamoto K, Ogawa Y, Toyama Y, Miyata T, Okano H: Musashi1: an evolutionally conserved marker for C progenitor cells including neural stem cells. *Dev Neurosci* 2000, 22:139–153
- Nakamura M, Okano H, Blendy JA, Montell C: Musashi, a neural RNA-binding protein required for Drosophila adult external sensory organ development. *Neuron* 1994, 13:67–81
- Okano H, Imai T, Okabe M: Musashi: a translational regulator of cell fate. *J Cell Sci* 2002, 115:1355–1359
- Clarke RB, Spence K, Anderson E, Howell A, Okano H, Potten CS: A putative human breast stem cell population is enriched for steroid receptor-positive cells. *Dev Biol* 2005, 277:443–456
- Subik K, Lee J-F, Baxter L, Strzepek T, Costello D, Crowley P, Xing L, Hung M-C, Bonfiglio T, Hicks DG, Tang P: The expression patterns of ER, PR, HER2, CK 5/6, EGFR, Ki-67 and AR by immunohistochemical analysis in breast cancer cell lines. *Breast Cancer (Auckl)* 2010, 4:35–41
- Holliday D, Speirs V: Choosing the right cell line for breast cancer research. *Breast Cancer Res* 2011, 13:215
- Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe J-P, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo W-L, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Graw JW: A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 2006, 12:515–527
- Mori T, Kim J, Yamano T, Takeuchi H, Huang S, Umetani N, Koyanagi K, Hoon DS: Epigenetic up-regulation of C-C chemokine receptor 7 and C-X-C chemokine receptor 4 expression in melanoma cells. *Cancer Res* 2005, 65:1800–1807
- Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF: American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 2007, 25:118–145
- Nicholl MB, Elashoff D, Takeuchi H, Morton DL, Hoon DS: Molecular upstaging based on paraffin-embedded sentinel lymph nodes: ten-year follow-up confirms prognostic utility in melanoma patients. *Ann Surg* 2011, 253:116–122
- Yoshimura T, Nagahara M, Kuo C, Turner RR, Soon-Shiong P, Hoon DS: Lymphovascular invasion of colorectal cancer is correlated to SPARC expression in the tumor stromal microenvironment. *Epigenetics* 2011, 6:1001–1011
- Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ: Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. *Nucleic Acids Res* 2007, 35:e119
- Asaga S, Kuo C, Nguyen T, Terpenning M, Giuliano AE, Hoon DS: Direct serum assay for microRNA-21 concentrations in early and advanced breast cancer. *Clin Chem* 2011, 57:84–91
- Akaike H: A new look at the statistical model identification. *IEEE Trans Automat Contr* 1974, 19:716–723
- Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, Cantor CR, Field JK, van den Boom D: Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A* 2005, 102:15785–15790
- Nair SS, Coolen MW, Stirzaker C, Song JZ, Statham AL, Strbenac D, Robinson MW, Clark SJ: Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveal CpG sequence coverage bias. *Epigenetics* 2011, 6:34–44
- Smith JF, Mahmood S, Song F, Morrow A, Smiraglia D, Zhang X, Rajput A, Higgins MJ, Krumm A, Petrelli NJ, Costello JF, Nagase H, Plass C, Held WA: Identification of DNA methylation in 3' genomic

- regions that are associated with upregulation of gene expression in colorectal cancer. *Epigenetics* 2007, 2:161–172
31. Gloss BS, Patterson KI, Barton CA, Gonzalez M, Scurry JP, Hacker NF, Sutherland RL, O'Brien PM, Clark SJ: Integrative genome-wide expression and promoter DNA methylation profiling identifies a potential novel panel of ovarian cancer epigenetic biomarkers. *Cancer Lett* 2012, 318:76–85
 32. Bonazzi VF, Nancarrow DJ, Stark MS, Moser RJ, Boyle GM, Aoude LG, Schmidt C, Hayward NK: Cross-platform array screening identifies COL1A2, THBS1, TNFRSF10D, and UCHL1 as genes frequently silenced by methylation in melanoma. *PLoS One* 2011, 6:e26121
 33. Hayashi K, Surani MA: Resetting the epigenome beyond pluripotency in the germline. *Cell Stem Cell* 2009, 4:493–498
 34. Nguyen T, Kuo C, Nicholl MB, Sim MS, Turner RR, Morton DL, Hoon DS: Down-regulation of microRNA-29c is associated with hypermethylation of tumor-related genes and disease outcome in cutaneous melanoma. *Epigenetics* 2011, 6:388–394
 35. Wang XY, Yin Y, Yuan H, Sakamaki T, Okano H, Glazer RI: Musashi1 modulates mammary progenitor cell expansion through proliferin-mediated activation of the Wnt and Notch pathways. *Mol Cell Biol* 2008, 28:3589–3599
 36. Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, Nikolskaya T, Serebryiskaya T, Beroukhim R, Hu M, Halushka MK, Sukumar S, Parker LM, Anderson KS, Harris LN, Garber JE, Richardson AL, Schnitt SJ, Nikolsky Y, Gelman RS, Polyak K: Molecular definition of breast tumor heterogeneity. *Cancer Cell* 2007, 11:259–273
 37. Hao L, Rizzo P, Osipo C, Pannuti A, Wyatt D, Cheung LW, Sonenshein G, Osborne BA, Miele L: Notch-1 activates estrogen receptor-alpha-dependent transcription via IKKalpha in breast cancer cells. *Oncogene* 2010, 29:201–213
 38. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, Symmans WF, Gonzalez-Angulo AM, Hennessy B, Green M, Cristofanilli M, Hortobagyi GN, Pusztai L: Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol* 2008, 26:1275–1281
 39. Bidard FC, Conforti R, Boulet T, Michiels S, Delaloge S, Andre F: Does triple-negative phenotype accurately identify basal-like tumour? an immunohistochemical analysis based on 143 "triple-negative" breast cancers. *Ann Oncol* 2007, 18:1285–1286
 40. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Borresen-Dale AL, Brown PO, Botstein D: Molecular portraits of human breast tumours. *Nature* 2000, 406:747–752
 41. Bertucci F, Finetti P, Cervera N, Esterni B, Hermitte F, Viens P, Birnbaum D: How basal are triple-negative breast cancers? *Int J Cancer* 2008, 123:236–240
 42. Kim MJ, Ro JY, Ahn SH, Kim HH, Kim SB, Gong G: Clinicopathologic significance of the basal-like subtype of breast cancer: a comparison with hormone receptor and Her2/neu-overexpressing phenotypes. *Hum Pathol* 2006, 37:1217–1226
 43. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lønning P, Borresen-Dale AL: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001, 98:10869–10874
 44. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM: Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* 2010, 12:R68
 45. Korsching E, Jeffrey SS, Meinerz W, Decker T, Boecker W, Buerger H: Basal carcinoma of the breast revisited: an old entity with new interpretations. *J Clin Pathol* 2008, 61:553–560
 46. D'Anello L, Sansone P, Storci G, Mitrugno V, D'Uva G, Chieco P, Bonafe M: Epigenetic control of the basal-like gene expression profile via interleukin-6 in breast cancer cells. *Mol Cancer* 2010, 9:300
 47. Bloushtain-Qimron N, Yao J, Snyder EL, Shipitsin M, Campbell LL, Mani SA, Hu M, Chen H, Ustyansky V, Antosiewicz JE, Argani P, Halushka MK, Thomson JA, Pharoah P, Porgador A, Sukumar S, Parsons R, Richardson AL, Stampfer MR, Gelman RS, Nikolskaya T, Nikolsky Y, Polyak K: Cell type-specific DNA methylation patterns in the human breast. *Proc Natl Acad Sci U S A* 2008, 105:14076–14081
 48. Ray PS, Wang J, Qu Y, Sim MS, Shamoni J, Bagaria SP, Ye X, Liu B, Elashoff D, Hoon DS, Walter MA, Martens JW, Richardson AL, Giuliano AE, Cui X: FOXC1 is a potential prognostic biomarker with functional significance in basal-like breast cancer. *Cancer Res* 2010, 70:3870–3876
 49. Harris LN, Broadwater G, Lin NU, Miron A, Schnitt SJ, Cowan D, Lara J, Bleiweiss I, Berry D, Ellis M, Hayes DF, Winer EP, Dressler L: Molecular subtypes of breast cancer in relation to paclitaxel response and outcomes in women with metastatic disease: results of CALGB 9342. *Breast Cancer Res* 2006, 8:R66
 50. Du L, Wang H, He L, Zhang J, Ni B, Wang X, Jin H, Cahuzac N, Mehrpour M, Lu Y, Chen Q: CD44 is of functional importance for colorectal cancer stem cells. *Clin Cancer Res* 2008, 14:6751–6760
 51. Rappa G, Fodstad O, Loric A: The stem cell-associated antigen CD133 (Prominin-1) is a molecular therapeutic target for metastatic melanoma. *Stem Cells* 2008, 26:3008–3017
 52. Zhou BB, Zhang H, Damelin M, Geles KG, Grindley JC, Dirks PB: Tumour-initiating cells: challenges and opportunities for anticancer drug discovery. *Nat Rev Drug Discov* 2009, 8:806–823
 53. Oliveras-Ferreras C, Vazquez-Martin A, Martin-Castillo B, Cufi S, Del Barco S, Lopez-Bonet E, Brunet J, Menendez JA: Dynamic emergence of the mesenchymal CD44(pos)CD24(neg/low) phenotype in HER2-gene amplified breast cancer cells with de novo resistance to trastuzumab (Herceptin). *Biochem Biophys Res Commun* 2010, 397:27–33
 54. Spector NL, Blackwell KL: Understanding the mechanisms behind trastuzumab therapy for human epidermal growth factor receptor 2-positive breast cancer. *J Clin Oncol* 2009, 27:5838–5847