

RESEARCH ARTICLE

Genome-wide DNA methylation signatures to predict pathologic complete response from combined neoadjuvant chemotherapy with bevacizumab in breast cancer

Ping-Ching Hsu¹, Susan A. Kadlubar¹, Eric R. Siegel², Lora J. Rogers³, Valentina K. Todorova⁴, L. Joseph Su³, Issam Makhoul^{4*}

1 Department of Environmental and Occupational Health, College of Public Health, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States of America, **2** Department of Biostatistics, Colleges of Medicine and of Public Health, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States of America, **3** Department of Epidemiology, College of Public Health, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States of America, **4** Department of Internal Medicine, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States of America

* MakhoulIssam@uams.edu

OPEN ACCESS

Citation: Hsu P-C, Kadlubar SA, Siegel ER, Rogers LJ, Todorova VK, Su LJ, et al. (2020) Genome-wide DNA methylation signatures to predict pathologic complete response from combined neoadjuvant chemotherapy with bevacizumab in breast cancer. *PLoS ONE* 15(4): e0230248. <https://doi.org/10.1371/journal.pone.0230248>

Editor: Qian Tao, Chinese University of Hong Kong, HONG KONG

Received: January 29, 2019

Accepted: February 25, 2020

Published: April 16, 2020

Copyright: © 2020 Hsu et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This study was supported by funds from the Laura F. Hutchins, M.D. Distinguished Chair for Hematology and Oncology (IM).

Competing interests: NO authors have competing interests.

Abstract

Neoadjuvant chemotherapy is given before surgery to patients with primarily unresectable, advanced-stage cancers to render the tumor resectable, and to facilitate breast conservation. Previously, we have reported a prospective phase II trial for women with stage IIA-B/IIIA-B-C breast cancer with improved pathologic complete response (pCR) when using bevacizumab in the neoadjuvant setting. Chemotherapy agents are given intravenously during multiple cycles of systemic treatments. However, the effect of the treatment is only evaluated upon the completion of therapy. Here, data from a clinical trial of 40 breast cancer patients with very aggressive disease and poor prognosis were studied aiming to identify epigenetic signatures in blood-derived DNA at baseline as potential non-invasive markers to predict pCR and to determine if treatment-related changes in epigenetic profiles reflect responsiveness to therapy. We performed genome-wide DNA methylation profiling using blood-derived DNA, and found that pre-treatment methylation status of *BRD9* was predictive of responsiveness to therapy. Post-treatment global methylation differences were also observed between responders and non-responders. Most differentially methylated (DM) CpGs were located in promoter CpG-island regions for responders and in the open-sea region for non-responders. In responders, *DNMT3B* was hypomethylated while most of the other genes were hypermethylated after 4 cycles of treatment. Hypomethylation of *DNMT3B* could potentially lead to the increased methylation of oncogenes and genes responsible for cell growth and proliferation, facilitating responsiveness to the therapy. These results support the possible development of *BRD9* as a biomarker for treatment selection before neoadjuvant therapy with chemotherapy and bevacizumab, and indicate *DNMT3B* as a potential target to improve clinical response. Further prospective validation of these findings is warranted.

Abbreviations: AA, African American; ANOVA, Analysis of variance; BRD9, Bromodomain Containing 9; CpG, Cytosine-guanine dinucleotide; dbSNP, The Single Nucleotide Polymorphism Database; DCB, Docetaxel, cyclophosphamide and bevacizumab; DM, Differentially methylated; DNMT3A, DNA Methyltransferase 3 Alpha; DNMT3B, DNA Methyltransferase 3 Beta; DNMTs, DNA methyltransferases; EA, European American; ER, Estrogen receptor; FDR, False discovery rate; FFPE, Formalin-fixed paraffin-embedded; HER2, Human epidermal growth factor receptor 2; HSD17B4, Hydroxysteroid 17-Beta Dehydrogenase 4; IPA, Ingenuity Pathway Analysis; IRB, Institutional Review Board; LVEF, Left ventricular ejection fraction; NCBI, National Center for Biotechnology Information; NCT, Neoadjuvant chemotherapy; PCA, Principal Component Analysis; pCR, Pathologic complete response; PITX2, Paired Like Homeodomain 2; ROC, Receiver Operating Characteristic; SNPs, Single-nucleotide polymorphisms; SWAN, Sub set-quantile Within Array Normalization.

Trial registration

ClinicalTrials.gov Identifier: [NCT00203502](https://clinicaltrials.gov/ct2/show/study/NCT00203502).

Introduction

In oncology, identification of predictive and prognostic biomarkers of treatment response is an area of intensive research. Genomic profiling has revealed tumor mutations and genetic variants that guide therapeutic decision making. In addition to tumor characteristics, host genetic variability also plays a role in treatment efficacy, and incorporation of genomic information into clinical decision making is a goal of precision medicine.

Neoadjuvant chemotherapy (NCT) is widely used in breast cancer before surgery to decrease tumor volume and facilitate surgical resection. Reduction of tumor volume in many cases allows breast-conserving surgery and the avoidance of mastectomy. In some cases, NCT results in the complete disappearance of the tumor prior to surgery, generating a pathological Complete Response (pCR). Either pCR or marked tumor reduction represent a net benefit to the patient. However, sometimes NCT has little direct effect on the tumor, which means the patient then endures ineffective treatment that can have long-lasting, and potentially irreversible, adverse effects. Clearly, it would be clinically useful in patient management to have a biomarker-based assay that can predict how well or how poorly the breast cancer patient's tumor responds to NC.

Genetic-variation studies have primarily been focused on tumor somatic mutations or on germline single-nucleotide polymorphisms (SNPs). However, it is becoming increasingly appreciated that epigenetic modifications controlling the expression of critical genes also contribute to therapeutic response. The most common epigenetic modification, and the one that has received the most attention to date, is cytosine methylation at cytosine-guanine dinucleotide ("CpG") sites or islands along the DNA sequence. Differences in CpG-island methylation status between different subjects have been shown to be associated with phenotype differences that include both a subject's susceptibility to disease and a disease's susceptibility to treatment. Likewise, changes in CpG methylation over time within the same subject have been associated with normal life-cycle processes ranging from embryogenesis to aging and senescence. Changes in CpG-island methylation have also been related to pathological processes such as carcinogenesis, responsiveness to starvation, gluttony, dietary imbalances, and exposures to pollutants, toxins, phytochemicals, and chemotherapy agents. Therefore, incorporation of differential CpG-island methylation detection into a biomarker-based assay has the potential to improve the prediction of response to NC, and thus refine precision-medicine practices.

Here, we report the results of a correlative study of CpG methylation in prospectively enrolled breast cancer patients that received NCT for their disease. Our purpose was two-fold. First, we sought to assess whether the methylation status of certain CpG-islands at baseline (before NCT started) was associated with achievement of pCR six months after NCT began. Second, we sought to study whether and/or how the methylation status of CpG-islands changes in response to treatment with NCT, and whether the pattern of change in CpG-methylation status was different among subjects who attained pCR compared to subjects who did not.

Material and methods

Study design & study participants

Detailed study methods for subject accrual have been described previously [1, 2]. Briefly, the study included 40 women with very aggressive disease and poor prognosis enrolled after providing written informed consent in NCT00203502 (ClinicalTrials.gov), a prospective single-

arm, single-institution phase II study approved by the Institutional Review Board (IRB) of the University of Arkansas for Medical Sciences between September 2005 and January 2008. Enrolled participants consisted of women older than 18 years with histologically confirmed stage II or III breast cancer, who did not have inflammatory BC (T4d) or any recent arterial thromboembolic events, uncontrolled hypertension, recent major surgery, or baseline left ventricular ejection fraction (LVEF) <50%. The participants were offered chemotherapy combination with docetaxel (75 mg/m²), cyclophosphamide (500 mg/m²), and bevacizumab (15 mg/kg) every 3 weeks×4 cycles followed by doxorubicin (60 mg/m²) every 3 weeks×4 cycles before undergoing surgery for their cancer. After surgery, patients were evaluated for pCR as the primary clinical endpoint, defined as the absence of invasive cancer in the breast specimen. Of the 40 patients enrolled, one withdrew consent and one did not undergo surgery after NCT but she was included in the intent to treat analysis. Sixteen (41%) of the 39 evaluable subjects attained pCR, significantly higher than the null-hypothesis rate of 25% ($P = 0.0204$). We have previously reported that the planned sample size of 40 subjects gave NCT00203502 sufficient power to detect a hypothesized increase of 20 percentage points in pCR [2]. For this report, we undertook retrospective power analyses of post-treatment changes in methylation status detectable at FDR-adjusted $p < 0.05$ in the responder and non-responder subgroups. The results of that retrospective power analysis appear below, and show that the subgroups had adequate sample sizes for the planned methylation analysis.

Sample collection, storage, and DNA isolation

Blood samples were collected from patients at baseline, after 4 cycles of docetaxel, cyclophosphamide and bevacizumab (DCB), and just before infusion at each of the 8 preoperative chemotherapy cycles. Blood components were separated, aliquotted into cryovials, and stored at -80°C for later analyses. Genomic DNA was extracted from buffy coats of the blood samples collected using Maxwell (Promega, Fitchburg, WI, USA). Concentration and quality of the extracted genomic DNA (gDNA) was measured using the Nanodrop 8000 instrument (Thermo Scientific, Waltham, MA, USA).

Infinium methylation EPIC beadChip analysis

Following bisulfite treatment of 1 µg genomic DNA using the EZ DNA Methylation kit (Zymo Research, Irvine, CA), the bisulfite-converted DNA was hybridized onto the Infinium Methylation EPIC BeadChip (Illumina, San Diego, CA), following the Illumina Infinium HD Methylation protocol in the Genomics Core Facility at UAMS. The Methylation EPIC BeadChip covers over 850,000 CpG sites, and has increased genome coverage of regulatory regions and higher reproducibility and reliability compared to previous versions [3]. Whole genome amplification, hybridization, staining and scanning steps for all samples were performed, the Illumina iScan SQ scanner was used to create images of the single arrays, and the intensities of the images were extracted using the Methylation module (v.1.9.0) of the GenomeStudio (v.2011.1) software (Illumina). Raw intensity data as IDAT files were imported into GenomeStudio for the computation of detection p -value of the probes, and all further steps including data import, normalization, filtering and analyses were performed using the methylation pipeline in Partek Genomics Suite™ 6.6 (Partek Inc., St. Louis, MO).

Methylation data analysis

The data was normalized using the Sub set-quantile Within Array Normalization (SWAN)[4]. For analysis, probes were excluded if detection p values were > 0.01 ($n = 130$), or if SNPs were present in the target CpG based on NCBI dbSNP Build 138 ($n = 59$). Probes on the X and Y

chromosomes ($n = 19,572$), and probes with polymorphic targets & cross-hybridization potential [5] ($n = 318,809$) were also excluded. The final dataset contained 527,348 probes across all samples. Percent methylation values for each CpG site (β -values) and logit-transformed ratios of methylated to unmethylated probe intensities (M -values) were extracted for further analysis [6]. Since race plays an important role in genome-wide DNA methylation [5], data from one Asian and one Hispanic participant was excluded from the analysis.

For pattern identification in DNA methylation, unsupervised analysis including unsupervised hierarchical clustering [7] and Principal Component Analysis (PCA) were used. T -tests and chi-square (X^2) tests were performed to evaluate differences between non-responders and responders. Analysis of variance (ANOVA) adjusting for race, tumor type, and estrogen receptor (ER) status with Fisher's Least Significant Difference contrast method were used to assess the differentially methylated (DM) CpG sites univariately for their potential to predict the subsequent response to therapy. The resulting P -values were adjusted for multiple testing with the false discovery-rate (FDR) procedure of Benjamini and Hochberg [8], and significance was granted with FDR P -value < 0.05 to identify significantly differential DNA methylation. In subgroup analysis, we examined post-treatment changes from baseline within each responder group, again using ANOVA adjusting for race, tumor type, and estrogen receptor (ER) status with Fisher's Least Significant Difference contrast method. Classical Receiver Operating Characteristic (ROC) curve analysis was used to evaluate the performance of a single CpG site as a biomarker.

To determine whether the responders and non-responders had adequate sample sizes for the subgroup analysis of post-treatment change in methylation, a retrospective power analysis was conducted using the FDR-control method of Jung *et al.* [9]. We define a truly positive CpG site to be one that shows a standardized change of ≥ 1.5 SD in post-treatment methylation level compared to baseline. To calculate power retrospectively for multiple 2-sample t -tests with FDR set equal to 5%, we hypothesize that 1.00% (5,274) of the 527,348 CpG sites in the final data set satisfy the ≥ 1.5 -SD criterion of being truly positive. Under this hypothesis, a sample size of $N = 16$ responders provides the multiple 2-sample t -tests at FDR = 5% with 89.77% power per test to detect a truly positive CpG site, and yields a binomial expected value ± 1 SD of $4,734 \pm 22.0$ positive test results (true discoveries) from the 5,274 CpG sites hypothesized to be truly positive. Under the same hypothesis, a sample size of $N = 21$ non-responders provides the multiple 2-sample t -tests at FDR = 5% with 99.01% power per test to detect a truly positive CpG site, and yields a binomial expected value ± 1 SD of $5,222 \pm 7.2$ positive test results (true discoveries) from the 5,274 CpG sites hypothesized to be truly positive. These retrospective power calculations demonstrate that both the responders and the non-responders had sample sizes large enough to power adequately the subgroup analysis of methylation response to treatment.

Pathway analysis

The genes corresponding to the significant loci were projected onto knowledge-based networks for their potential biological implications using Ingenuity Pathway Analysis (IPA[®], www.qiagen.com/ingenuity). IPA employs information obtained from the literature to assemble and extrapolate known interactions, signaling, as well as the relationships between the molecules. The analysis here was restricted to genes or molecules that have been observed directly in human breast or immune cell lines in the literature. Network analysis was generated *de novo* based on the mapped CpG sites to explore potential molecular events and mechanisms impacted by the DM CpGs. The P -score [computed as $-\log_{10}(P\text{-value})$] reflected the probability of the identified genes in a list of biological functions stored in the IPA knowledge base.

Results

Demographics of the study participants

Samples for CpG methylation analysis were available for 37 patients; their demographics and disease characteristics are shown in Table 1. The two groups differed significantly at $p < 0.05$ in race (African American, AA; and European American, EA) and estrogen receptor (ER) status. Although differences in the type of tumor did not reach statistical significance ($p = 0.07$), a tendency was observed for most of the lobular and poorly differentiated breast cancer patients to be non-responders. Thus, race, ER status, and tumor type were included as covariates to adjust for in the subsequent analysis.

Table 1. Demographics and disease characteristics of study participants.

Total 37	21 Non-responders (0)	16 Responders (1)	<i>p</i>
Race			0.02
Black	3	9	
White	18	7	
Age	46.4 ± 12.9	44.6 ± 12.9	0.68
Tumor type			0.07
Ductal	13	15	
Lobular	5	1	
Poorly diff	3	0	
ER			0.03
Positive	14	4	
Negative	7	12	
PR			0.63
Positive	8	4	
Negative	13	12	
HER2			1.00
Positive	5	3	
Negative	16	13	
Triple negative			0.32
Yes	6	8	
No	15	8	
Tumor grade			0.70
I	1	1	
II	8	4	
III	12	11	
Both cycles complete			0.66
Yes	16	14	
No	5	2	
FU in years (yrs)	2.8 ± 1.4	3.5 ± 0.9	0.13
Tumor 1st diameter (cm)	6.2 ± 2.8	4.9 ± 1.9	0.13
Tumor 2nd diameter (cm)	6 ± 2.8	4.9 ± 2.3	0.29
Tumor max diameter (cm)	6.5 ± 2.8	5.1 ± 2.0	0.11

^a*p*-values represent differences in non-responders vs. responders for each characteristic. Continuous variables were evaluated by two-sample t-tests, and chi square (X^2) tests were used to investigate the differences in distributions of categorical variables from non-responders to responders.

^bOne Hispanic and one Asian (both non-responders) were excluded from the Race comparison.

<https://doi.org/10.1371/journal.pone.0230248.t001>

DM CpG sites at baseline predictive of pathologic response

First, we aimed to identify germline DM CpGs at baseline that can serve as non-invasive markers to predict their pCR status following NCT. From our 4-way analysis of variance (ANOVA) model controlling for race, ER status, tumor type, and pCR, 25,321 CpG sites were significantly different between responders and non-responders at baseline (unadjusted $p < 0.05$), but only one CpG passed multiple-testing adjustments (FDR $p < 0.05$, Fig 1a). This CpG site (cg06362249)

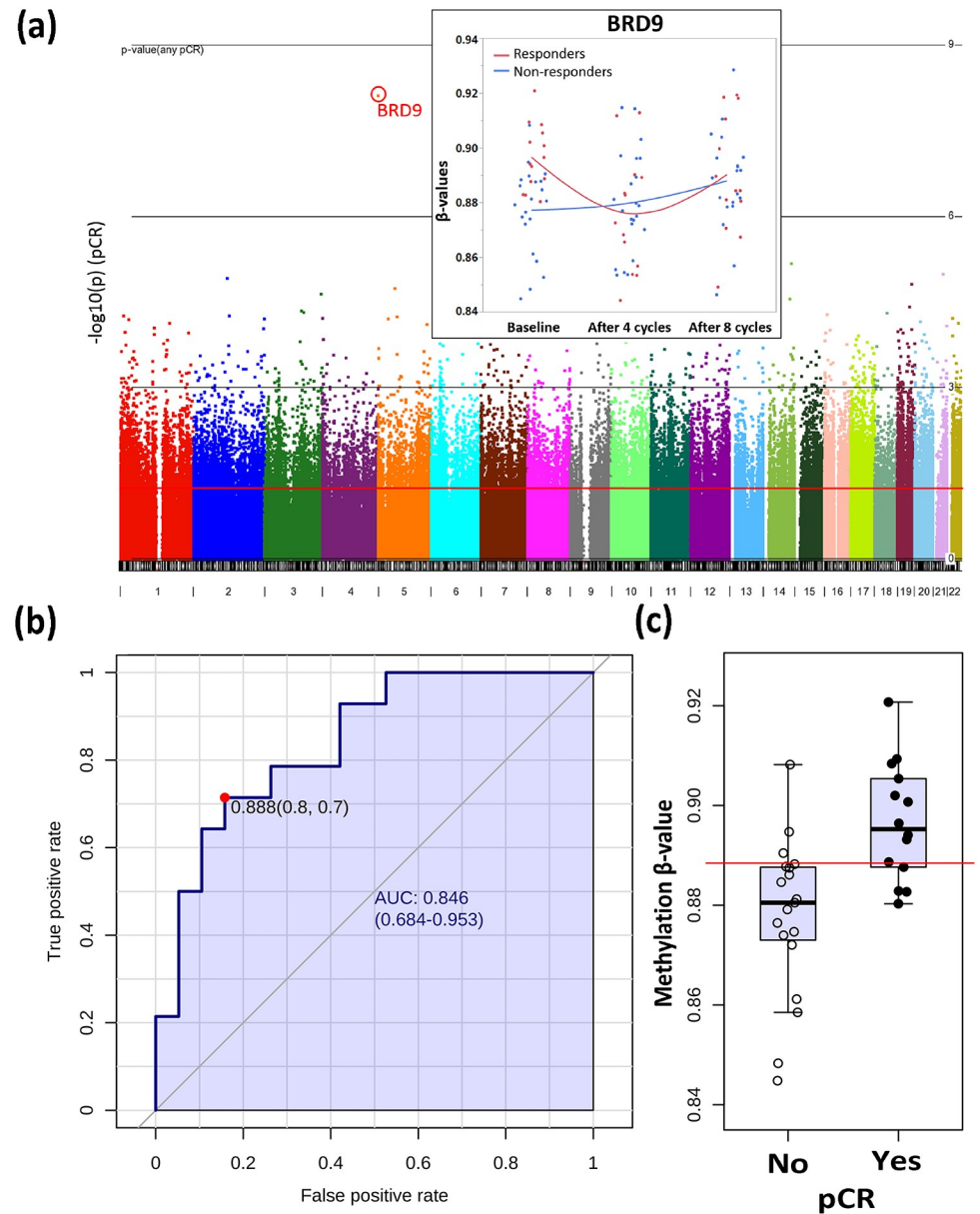


Fig 1. BRD9 (cg06362249) predicts cPR at baseline. (a) Manhattan plot on the chromosomal distribution of differentially methylated CpGs at baseline based on responsiveness to therapy. (b) ROC curve analysis of BRD9. The sensitivity (true positive rate) is on the y-axis, and the specificity (one minus the false positive rate) is on the x-axis with the area under the curve (AUC) of 0.842 (95% CI: 0.708–0.976). (c) Box-plot on the methylation levels of BRD9 between responders (Yes) and non-responders (No). X axis represents β values. A horizontal line in red indicates the optimal cutoff of 0.889 between two groups.

<https://doi.org/10.1371/journal.pone.0230248.g001>

binds to three *BRD9* transcripts (Chromosome 5, NR_02763, NM_001009877, NM_023924) in the gene body, was hypermethylated among responders at baseline compared to non-responders (Fold change = 1.4; unadjusted $p = 7.78e-009$, FDR $p = 0.004$), decreased its methylation level by 1.24 fold after 4 cycles of treatment among responders (unadjusted $p = 0.02$), but did not change appreciably after treatment among non-responders (Fold change = 1.0; unadjusted $p = 0.69$). The area under the curve (AUC, Fig 1b) for cg06362249 is 0.842 (95%CI: 0.708–0.976) in the Receiver Operating Characteristic (ROC) analysis with the optimal cutoff of 0.889 in β value (% methylation, see Fig 1b and 1c) between responders and non-responders. The methylation levels of cg06362249 were not distinguishable between responders and non-responders after 4 cycles and 8 cycles of treatment (Fig 1a).

Responder- & non-responder-specific CpG sites

To determine the effect of treatment on global methylation profiles specific to the pCR group, separate analyses were performed using the 4-way ANOVA model controlling for race, ER status, treatment cycle, and tumor type for responders and non-responders. There were 5,246 CpGs (FDR $p < 0.05$, Fig 2a) among responders that were DM between baseline and the 4th treatment cycle, 4,984 (95%) of which are projected to be true discoveries under a 5% FDR. A projection of 4,984 true discoveries among responders is slightly higher than this group's estimate of $4,734 \pm 22.0$ true discoveries from the retrospective power analysis, suggesting that the proportion of CpG sites defined to be "truly positive" (see above) is consistent among responders with the hypothesized rate of 1%. In contrast, there were only 1,201 CpGs (Fig 2b) among non-responders that were DM between baseline and the 4th treatment cycle, 1,141 (95%) of which are projected to be true discoveries under a 5% FDR. A projection of 1,141 true

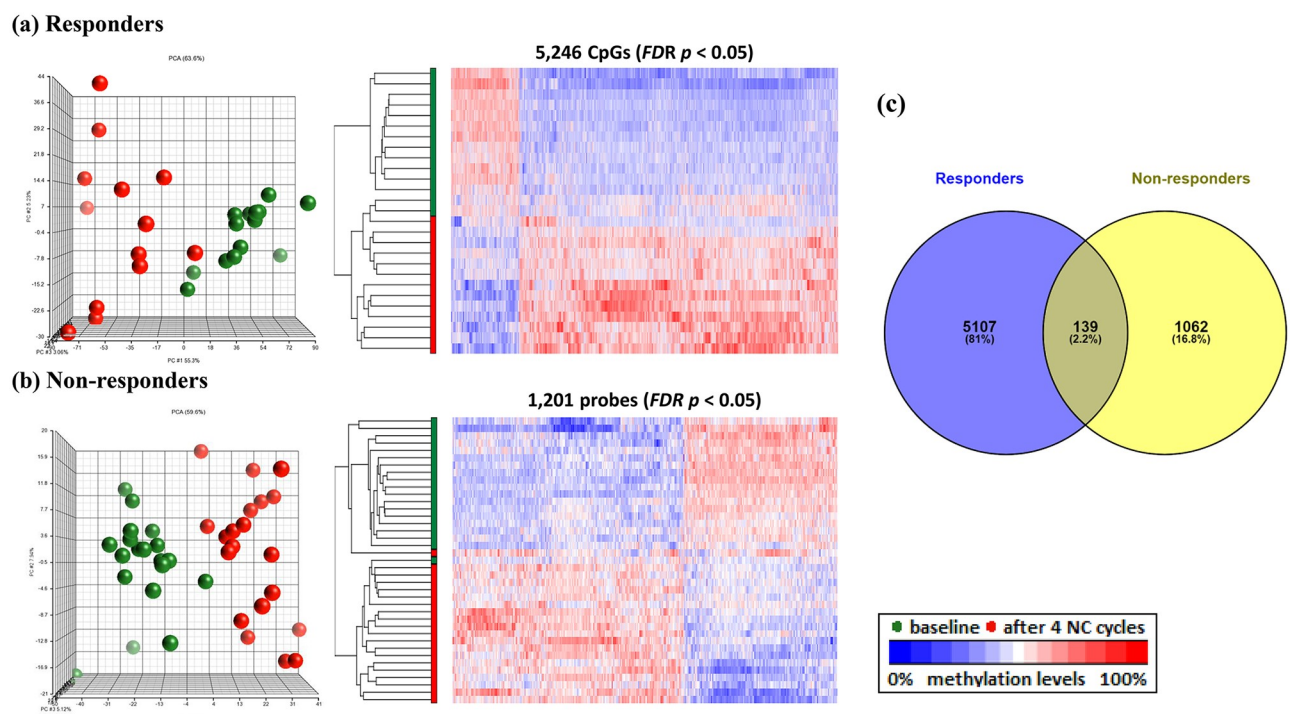
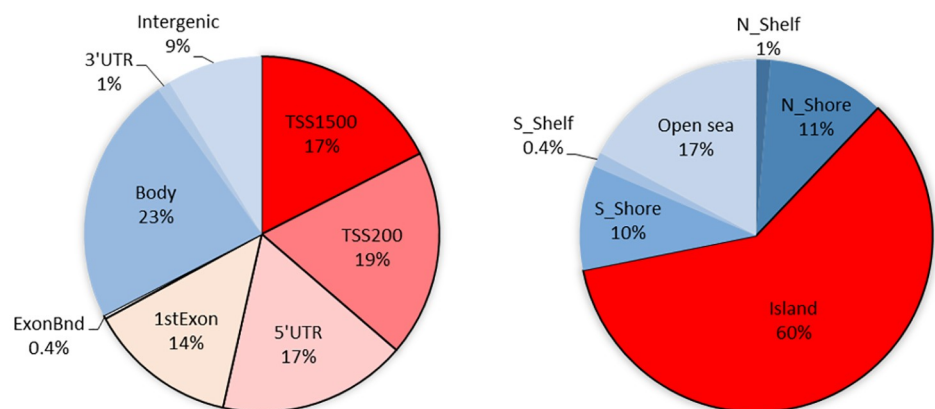


Fig 2. Visualization of the top CpG sites (FDR $p < 0.05$) in principle component analysis and hierarchical clustering distinguishing methylation levels between baseline and after 4 cycles of NCT treatments among responders (a) and non-responders (b). (c) Venn diagram of the DM CpGs showed that 97.8% of DM CpGs were specific to responders or non-responders.

<https://doi.org/10.1371/journal.pone.0230248.g002>

discoveries among non-responders is more than 78% lower than this group's estimate of $5,222 \pm 7.2$ true discoveries from the retrospective power analysis, indicating that the proportion of "truly positive" CpG sites among non-responders is much lower than the hypothesized rate of 1%. When we inspected the DM CpGs among responders, 67% were localized in the promoter regions that correspond to regions TSS1500 (17%), TSS200 (19%), 5'UTR (17%), 1stExon (14%), and ExonBnd (0.4%) (Fig 3a left). In contrast, 46% of the DM CpGs among non-responders were localized in the promoter regions, including TSS1500 (12%), TSS200 (11%), 5'UTR (15%), 1stExon (7%), and ExonBnd (0.4%) (Fig 3b, left). Further analysis on the distribution of DM CpGs by CpG-island regions revealed more changes in the CpG islands (60%, Fig 3a, right) among responders than the surrounding regions including N Shelf (1%, 4 kb upstream of CpG island), N_Shore (11%, 0–2 kb upstream of the CpG islands), S_Shore (10%, 0–2 kb downstream of the CpG islands), S_Shelf (0.4%, 2–4 kb downstream of the CpG islands), and open sea (17%), while for non-responders, more changes were observed in the open sea (47%, Fig 3b, right) compared to islands (29%). Subsequent analysis using ANOVAs with Bonferroni adjustment selected 19 top CpG sites with the most significant change between

(a) Responders



(b) Non-responders

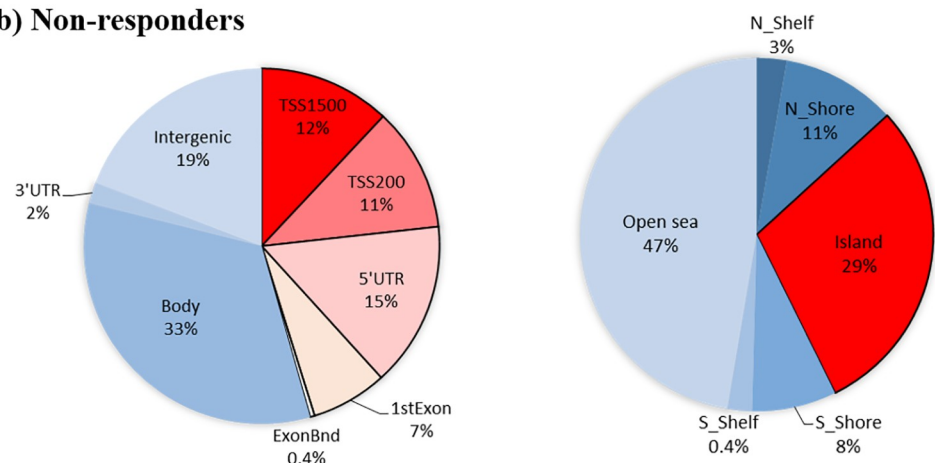


Fig 3. Distribution of differentially methylated CpGs by functional locations (left) and by CpG island regions (right). (a) 67% of DM CpGs among responders were located in the promoter regions that correspond to TSS1500, TSS200, 5'UTR, 1st Exon, and ExonBnd, with 60% of CpGs linked to CpG islands, whereas (b) 46% of DM CpGs among non-responders were located in the promoter regions and 29% linked to CpG islands.

<https://doi.org/10.1371/journal.pone.0230248.g003>

baseline and after 4 cycles of NCT treatment, as presented in Table 2. All 10 of the top CpG sites from responders belonged to CpG islands within named genes. In contrast, 8 of the 9 top CpG sites from non-responders belonged instead to “open sea” regions, 3 of which were located in intergenic regions.

Pathway analysis

In order to explore potential molecular events and mechanisms impacted by the neoadjuvant chemotherapy, we applied a statistical-significance threshold of FDR $p < 0.05$ coupled with a scientific-significance threshold of absolute fold change > 1.5 signifying methylation changes of at least 50% from the therapy. This yielded 651 responder-specific DM CpGs and 355 non-responder-specific CpGs (S1 Fig) that were included in separate IPA analyses queried with Ingenuity’s knowledge-based networks. Direct relationships focusing on interaction networks observed from all data sources were considered in the analysis, and demonstrated 10 network affected including the top networks affected among responders that were associated with Cellular Development, Cellular Growth and Proliferation, Organismal Development (Score = 49), with 14 molecules involved in breast cancer (*ABCBI*, *ALK*, *BCL2*, *CCND1*, *CDK1*, *DNMT3B*, *FOXA1*, *HIST2H2BE*, *ITPR1*, *JAK1*, *MDM2*, *MFN2*, *RUNX3*, *USP6NL*, Fig 4a). Top canonical pathways involved in the network were PI3K/Akt signaling and glucocorticoid receptor signaling, with *DNMT3B* decreased in methylation levels after 4 cycles of NCT while almost all the other genes in the network were increased in the methylation levels after the therapy (Fig 4a). Among non-responders, the top network

Table 2. Responder-specific & non-responder specific CpG sites after 4 cycles of treatment.

Probeset ID	UCSC RefGene Name	CHR	Relation to UCSC CpG Island	Unadjusted <i>p</i>	Bonferroni <i>p</i>	Fold Change*	Methylation level
Responders-specific CpG sites							
cg27424117	CDC25B	20	Island	4.2E-09	0.002	-1.4	↑ after treatment
cg25542733	WBSR27	7	Island	7.3E-09	0.004	-1.7	↑ after treatment
cg13940160	ADNP	20	Island	2.1E-08	0.011	-1.7	↑ after treatment
cg13494334	NEK9	14	Island	2.8E-08	0.015	-1.8	↑ after treatment
cg20161984	COQ2	4	Island	3.0E-08	0.016	-1.3	↑ after treatment
cg04925841	SOX5	12	Island	3.7E-08	0.019	-1.6	↑ after treatment
cg19377384	ACPL2	3	Island	4.9E-08	0.026	-1.5	↑ after treatment
cg05368942	RBM15	1	Island	5.4E-08	0.028	-1.4	↑ after treatment
cg12743270	TLE2	19	Island	5.5E-08	0.029	-1.7	↑ after treatment
cg02709211	RB1CC1	8	Island	6.2E-08	0.033	-1.4	↑ after treatment
Non-responders-specific CpG sites							
cg09805347	(Intergenic region)	2	Open sea	1.2E-08	0.006	3.3	↓ after treatment
cg17420619	ARHGAP1	11	Open sea	1.6E-08	0.008	-1.4	↑ after treatment
cg26692003	IQSEC1	3	Open sea	1.6E-08	0.009	-3.3	↑ after treatment
cg02402539	(Intergenic region)	19	Open sea	4.2E-08	0.022	-1.4	↑ after treatment
cg26709300	YPEL3	16	N_Shore	5.5E-08	0.029	-1.5	↑ after treatment
cg01680062	RUNX1	21	Open sea	7.1E-08	0.038	2.1	↓ after treatment
cg24481263	LINC01258	4	Open sea	8.1E-08	0.043	2.1	↓ after treatment
cg00889210	MKL2	16	Open sea	8.8E-08	0.046	-2.3	↑ after treatment
cg17227564	(Intergenic region)	19	Open sea	8.8E-08	0.047	1.5	↓ after treatment

*Fold changes represent differences from baseline to the 4th cycles of treatment for each characteristic. Unadjusted *P*-values characterize the statistical significance of the fold changes, while Bonferroni *P*-values are used to select the CpG sites shown in the table. All selected CpG sites have Bonferroni $P < 0.05$.

<https://doi.org/10.1371/journal.pone.0230248.t002>

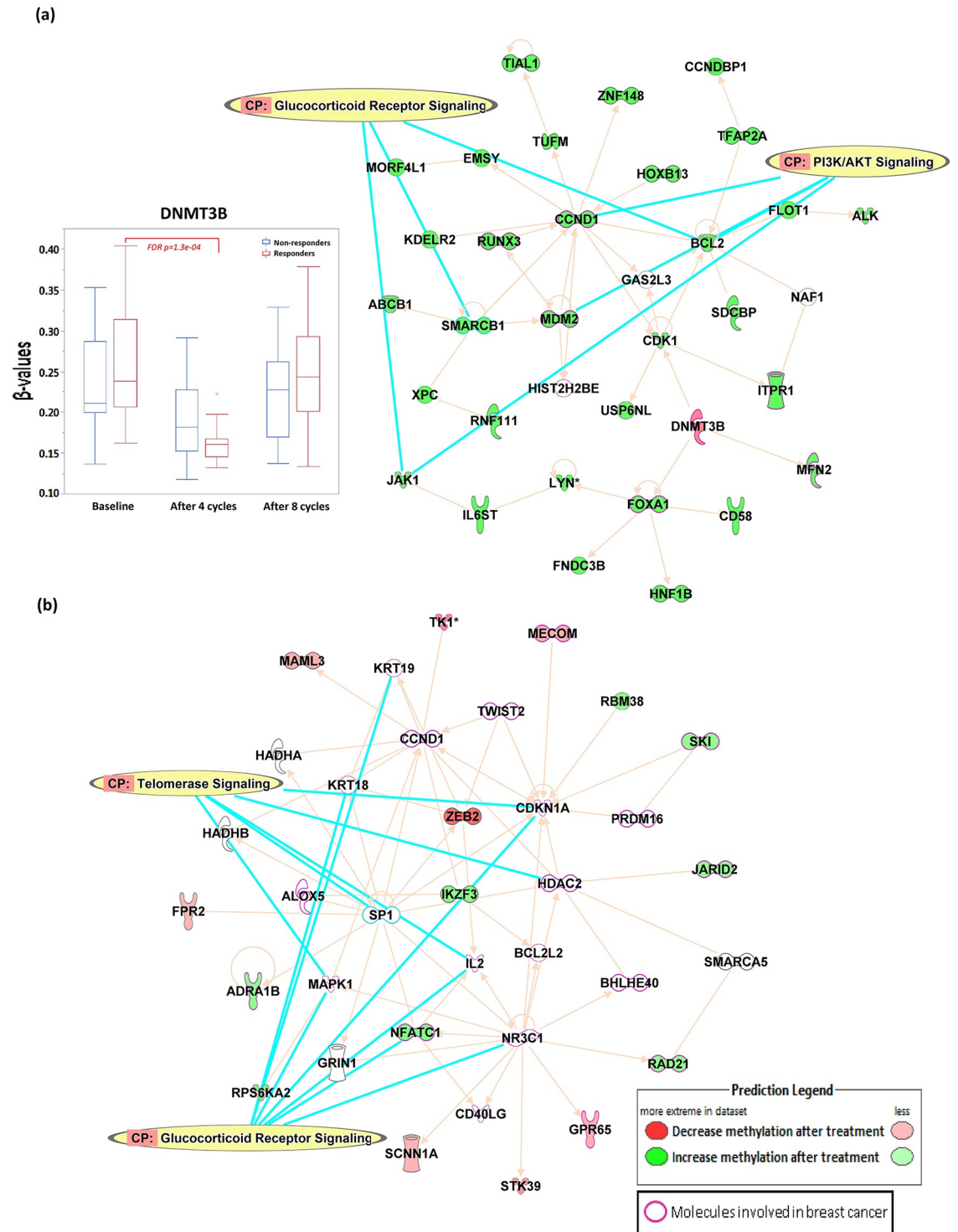


Fig 4. Ingenuity pathway analysis of DM CpGs after 4 cycles of NCT treatment among (a) responders and (b) non-responders.

<https://doi.org/10.1371/journal.pone.0230248.g004>

out of 20 networks affected was associated with Tissue Morphology, Organismal Development, Embryonic Development (Score = 25) with 23 molecules involved in breast cancer (*ADRA1B*, *ALOX5*, *BCL2L2*, *BHLHE40*, *CCND1*, *CD40LG*, *CDKN1A*, *GPR65*, *HDAC2*, *IL2*, *JARID2*, *KRT18*, *KRT19*, *MAPK1*, *MECOM*, *NFATC1*, *NR3C1*, *PRDM16*, *RAD21*, *RPS6KA2*,

SKI, *TK1*, *TWIST2*, Fig 4b). Top canonical pathways involved in the network were glucocorticoid receptor signaling and telomerase signaling. Full list of networks affected from both responders and non-responders are available in the S1 and S2 Tables.

Discussion

Neoadjuvant chemotherapy is administered to patients with operable and locally advanced breast cancers to facilitate breast conservation, or to render inoperable breast tumors viable candidates for surgery. However, the evaluation of clinical response to NCT is delayed until the completion of the therapy. Therefore, biomarkers to predict pCR are urgently needed to avoid unnecessary toxicity to the patient from NCT. Here, we aimed to identify genomic signatures of clinical response among breast cancer patients with very aggressive disease and poor prognosis who underwent NCT using a genome-wide methylation approach. Methylation levels of bromodomain-containing protein 9 (*BRD9*) in the gene body were significantly different at baseline between responders and non-responders, and were the strongest predictors of clinical response in this study. We also identified 5,246 DM CpG sites among responders, with the majority localized in the CpG islands of promoter regions, while 1,201 DM CpG sites were associated with non-responders. The majority of these sites were localized in the intergenic and open-sea regions. Finally, pathway analysis revealed that the DM CpGs among responders were mapped to canonical pathways including PI3K/Akt signaling and glucocorticoid receptor signaling, with decreasing methylation of *DNMT3B*, a major gene responsible for methylation, in responders after 4 cycles of NCT.

BRD9 protein is an epigenetic "reader" of acetylated lysines on post-translationally modified histone proteins that facilitates gene expression. Although little is known about the function of BRD9, copy number gains of the *BRD9* gene in lung cancer patients has been reported and was associated with lung cancer susceptibility [10, 11]. It is over-expressed in cervical cancer [12], promotes *MYC* expression and cell proliferation in acute myeloid leukemia [13]. Protein-engineering and inhibitor studies targeting BRD9 have shown promise in overcoming epigenetically-defined drug resistance [14], and single or combinatorial effects of BRD9 inhibitors with doxorubicin or carboplatin also resulted in anti-proliferative effects in five malignant rhabdoid tumor cell lines [15]. Risdom et al. revealed that epigenetic plasticity mediated by the chromatin modifier BRD4 drives survival of triple-negative breast cancer cells after targeted therapy treatment [16]. In our study, baseline methylation levels of *BRD9* in the gene body were significantly higher among responders compared to non-responders. Gene body methylation has been shown to increase the potential for both somatic and germline mutations [17], and might involve in the recruitment of DNMTs [18, 19]. Functionally, hypermethylation of *BRD9* reduces the amount of available *BRD9* readers, thus preventing recognition of transcription regulators to the acetylated histones, further resulting in the reduced level of transcription. Conversely, lower levels of methylation among non-responders increases the transcription of *BRD9*, thus facilitating gene transcription. Whether the differences affects structural change of the BRD9 gene, binding affinity of the gene, or expression of the gene will need to be further studied.

While chemotherapy profoundly impacts DNA methylation landscape in cancer patients [20, 21], resistance to treatment continues to be a major obstacle for the success of cancer therapy. In our study, increased in the CpG site methylation levels among responders were all mapped to CpG islands including *CDC25B*, *WBSR27*, *ADNP*, *NEK9*, *COQ2*, *SOX5*, *ACPL2*, *RBM15*, *TLE2*, *RB1CC1* (Table 2). Those genes have been shown to have key roles in G2/M cell-cycle progression (*CDC25B*), methyltransferase (*WBSR27*), determination of the cell fate (*SOX5*), and cell growth (*RB1CC1*), which the expression levels would expect to be inhibited after treatment. On the other hand, CpG changes among non-responders were mainly

mapped to open sea or intergenic regions. Lund et al. reported that the resistance of cisplatin on ovarian cancer cells is associated with loss of hypermethylation in a high number of CpG sites, and are primarily localized in the intergenic regions of the genome [22]. Gene expression can be regulated by the methylation of intergenic regions [23–25], which contain key regulatory elements including enhancers, silencers, and noncoding RNAs. Our results support the role of DNA methylation-associated cancer chemotherapy resistance, and identified putative genes that might be associated with drug resistance. Further investigation in larger studies is needed in order to elucidate the complex molecular mechanisms involved in drug resistance.

Patterns of DNA methylation are established and maintained by DNA methyltransferases (DNMTs) [26]. It is well-known that global DNA hypomethylation in cancer is an early epigenetic abnormality in cancer compared to their normal-tissue counterparts [27, 28], and may result in the accumulation of epigenetic abnormalities leading to chromosomal instability, increased transcription from transposable elements, and an elevated mutation rate due to mitotic recombination [29]. Thus, DNA methylation has been a target of anticancer therapies in the past few decades [30], with much effort being focused on the development of 5-azacytosides or non-nucleoside DNMT inhibitors to prevent the translation of oncogenic proteins [31, 32]. However, not all breast tumors overexpress DNMTs [33] and little is known about the specificity of DNMTs. In fact, many of the DNMTs are involved in specific and non-specific DNA-methylation machineries [34], suggesting a tight regulation over the DNA methylome to maintain genome and epigenome integrities. Furthermore, DNMT inhibitors have been reported to increase the proliferation of pluripotent stem cell [35] or reprogramming through epigenetic modifications [36], as well as to have immune-suppressive effects [37]. In our study, NCT increased the methylation levels for 4,323 CpG sites and decreased methylation levels of 923 CpGs among responders, potentially through the decreased methylation (increased gene expression) of *DNMT3B* as indicated in Fig 4a, suppressing the expression of certain resistance genes thus leading to pathological complete response. On the other hand, methylation levels of 721 CpGs were increased among non-responders and 480 CpGs were decreased in methylation levels (Fig 2), mostly among the open sea regions. No significant change in methylation was observed for *DNMT3B* among non-responders. Further studies are needed in order to fully understand the function and mechanism underlying the observed differences.

Other attempts at identifying epigenetic markers including targeted approach by examining gene-specific methylation markers including the use of *PITX2* methylation to predict anthracycline-based chemotherapy for triple-negative breast cancer and high-risk breast cancer [38, 39], to examine changes in the methylation levels in estrogen receptor- α promoter regions among triple-negative breast cancer patients [40], as well as global approach to identify HSD17B4 methylation as predictor of HER2-positive breast cancer to trastuzumab and chemotherapy [41]. Another global approach by Gampenrieder et al. reported the use of DNA-methylation signatures in archival formalin-fixed paraffin-embedded (FFPE) specimens to predict Bevacizumab efficacy in HER2-negative metastatic breast cancer by the Illumina Infinium HumanMethylation450 BeadChip [42]. Albeit using the older version of the beadchip, the study identified 9-gene and 3-gene DNA-methylation signatures that were predictive for responders with longer progression-free survival [42]. The study focused on patients with metastatic breast cancers treated with different chemotherapy backbones (taxane or capecitabine) alone or in combination with bevacizumab, which may have different DNA methylation profiles than patients with locally advanced breast cancer treated with neoadjuvant therapy containing bevacizumab in our study. However, their study has a significant age difference (median age 61) compared to ours (median age 43), and did not account for racial differences in the analysis. Since both age [43] and race [44, 45] play important roles in predicting cancer risk, and are related to the DNA-methylation changes in breast cancer tissues [46–48], our

study do provide more insights by addressing variability in race, ER status, and tumor type with higher significance cutoff (FDR $p < 0.05$, $p < 1.1 \times 10^{-9}$) and more coverage of CpGs in the new EPIC Beadchips.

To our knowledge, the NIH GDC Data Portal that holds all TCGA database has only data from older chip versions (Illumina Human Methylation 450 and Illumina Human Methylation 27), which does not contain the BRD9 probe we identified from the study. Our samples were collected in 2005–2008 without RNA stabilization reagent so we were unable to validate the results using gene expression assays. However, we do want to acknowledge that the data from Illumina Infinium HumanMethylation BeadChips, especially the Illumina EPIC BeadChips is highly robust and reproducible and is a significant improvement over the HM450 array for human studies [3, 49]. Despite limitations in validation and sample size, this study highlights the use of methylation signatures with the use of stringent statistical methods to correct for multiple comparisons to predict clinical response of breast cancer patients to NCT. In this regard, a larger-scale effort to define the usefulness of the markers may have the potential to greatly expand the clinical utility and to be evaluated in future clinically relevant studies.

Supporting information

S1 Fig. Venn diagram representing DM CpGs from responders and non-responders.
(DOCX)

S1 Table. Top networks affected from 651 responder-specific CpGs.
(XLSX)

S2 Table.
(XLSX)

Acknowledgments

The authors thank all the patients, the coinvestigators, and the personnel of the Cancer Clinical Trials Office at the WPRCI.

Author Contributions

Conceptualization: Ping-Ching Hsu, Susan A. Kadlubar, Issam Makhoul.

Data curation: Ping-Ching Hsu, Eric R. Siegel, Issam Makhoul.

Formal analysis: Ping-Ching Hsu, Issam Makhoul.

Funding acquisition: Issam Makhoul.

Investigation: Issam Makhoul.

Methodology: Susan A. Kadlubar, Lora J. Rogers, Valentina K. Todorova, Issam Makhoul.

Resources: Issam Makhoul.

Software: Ping-Ching Hsu.

Supervision: Issam Makhoul.

Visualization: Ping-Ching Hsu, Eric R. Siegel.

Writing – original draft: Ping-Ching Hsu, Susan A. Kadlubar.

Writing – review & editing: Ping-Ching Hsu, L. Joseph Su, Issam Makhoul.

References

1. Makhoul I, Griffin RJ, Siegel E, Lee J, Dhakal I, Raj V, et al. High-circulating Tie2 Is Associated With Pathologic Complete Response to Chemotherapy and Antiangiogenic Therapy in Breast Cancer. *Am J Clin Oncol*. 2016; 39(3):248–54. <https://doi.org/10.1097/COC.000000000000046> PMID: 24577164
2. Makhoul I, Klimberg VS, Korourian S, Henry-Tillman RS, Siegel ER, Westbrook KC, et al. Combined neoadjuvant chemotherapy with bevacizumab improves pathologic complete response in patients with hormone receptor negative operable or locally advanced breast cancer. *Am J Clin Oncol*. 2015; 38(1):74–9. <https://doi.org/10.1097/COC.0b013e31828940c3> PMID: 23563210
3. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. 2016; 17(1):208. <https://doi.org/10.1186/s13059-016-1066-1> PMID: 27717381
4. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol*. 2012; 13(6):R44. <https://doi.org/10.1186/gb-2012-13-6-r44> PMID: 22703947
5. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genom Data*. 2016; 9:22–4. <https://doi.org/10.1016/j.gdata.2016.05.012> PMID: 27330998
6. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 2010; 11:587. <https://doi.org/10.1186/1471-2105-11-587> PMID: 21118553
7. Quackenbush J. Computational analysis of microarray data. *Nat Rev Genet*. 2001; 2(6):418–27. <https://doi.org/10.1038/35076576> PMID: 11389458
8. Benjamini Y H Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*. 1995; 57(1):289–300.
9. Jung SH, Bang H, Young S. Sample size calculation for multiple testing in microarray data analysis. *Biostatistics*. 2005; 6(1):157–69. <https://doi.org/10.1093/biostatistics/kxh026> PMID: 15618534
10. Liu Y, Lusk CM, Cho MH, Silverman EK, Qiao D, Zhang R, et al. Rare Variants in Known Susceptibility Loci and Their Contribution to Risk of Lung Cancer. *J Thorac Oncol*. 2018.
11. Kang JU, Koo SH, Kwon KC, Park JW, Kim JM. Gain at chromosomal region 5p15.33, containing TERT, is the most frequent genetic event in early stages of non-small cell lung cancer. *Cancer Genet Cytogenet*. 2008; 182(1):1–11. <https://doi.org/10.1016/j.cancergencyto.2007.12.004> PMID: 18328944
12. Scotto L, Narayan G, Nandula SV, Subramaniam S, Kaufmann AM, Wright JD, et al. Integrative genomics analysis of chromosome 5p gain in cervical cancer reveals target over-expressed genes, including Drosha. *Mol Cancer*. 2008; 7:58. <https://doi.org/10.1186/1476-4598-7-58> PMID: 18559093
13. Hohmann AF, Martin LJ, Minder JL, Roe JS, Shi J, Steurer S, et al. Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition. *Nat Chem Biol*. 2016; 12(9):672–9. <https://doi.org/10.1038/nchembio.2115> PMID: 27376689
14. Crawford TD, Vartanian S, Cote A, Bellon S, Duplessis M, Flynn EM, et al. Inhibition of bromodomain-containing protein 9 for the prevention of epigenetically-defined drug resistance. *Bioorg Med Chem Lett*. 2017; 27(15):3534–41. <https://doi.org/10.1016/j.bmcl.2017.05.063> PMID: 28606761
15. Kramer KF, Moreno N, Fruhwald MC, Kerl K. BRD9 Inhibition, Alone or in Combination with Cytostatic Compounds as a Therapeutic Approach in Rhabdoid Tumors. *Int J Mol Sci*. 2017; 18(7).
16. Risom T, Langer EM, Chapman MP, Rantala J, Fields AJ, Boniface C, et al. Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer. *Nat Commun*. 2018; 9(1):3815. <https://doi.org/10.1038/s41467-018-05729-w> PMID: 30232459
17. Rideout WM 3rd, Coetzee GA, Olumi AF, Jones PA. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science*. 1990; 249(4974):1288–90. <https://doi.org/10.1126/science.1697983> PMID: 1697983
18. Hahn MA, Wu X, Li AX, Hahn T, Pfeifer GP. Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. *PLoS One*. 2011; 6(4):e18844. <https://doi.org/10.1371/journal.pone.0018844> PMID: 21526191
19. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012; 13(7):484–92. <https://doi.org/10.1038/nrg3230> PMID: 22641018
20. Yao S, Hu Q, Kerns S, Yan L, Onitilo AA, Mislleh J, et al. Impact of chemotherapy for breast cancer on leukocyte DNA methylation landscape and cognitive function: a prospective study. *Clin Epigenetics*. 2019; 11(1):45. <https://doi.org/10.1186/s13148-019-0641-1> PMID: 30867049
21. Flanagan JM, Wilson A, Koo C, Masrour N, Gallon J, Loomis E, et al. Platinum-Based Chemotherapy Induces Methylation Changes in Blood DNA Associated with Overall Survival in Patients with Ovarian

- Cancer. Clin Cancer Res. 2017; 23(9):2213–22. <https://doi.org/10.1158/1078-0432.CCR-16-1754> PMID: 27663594
22. Lund RJ, Huhtinen K, Salmi J, Rantala J, Nguyen EV, Moulder R, et al. DNA methylation and Transcriptome Changes Associated with Cisplatin Resistance in Ovarian Cancer. *Sci Rep*. 2017; 7(1):1469. <https://doi.org/10.1038/s41598-017-01624-4> PMID: 28473707
 23. Thomas RM, Sai H, Wells AD. Conserved intergenic elements and DNA methylation cooperate to regulate transcription at the il17 locus. *J Biol Chem*. 2012; 287(30):25049–59. <https://doi.org/10.1074/jbc.M112.351916> PMID: 22665476
 24. Lu L, Zhu G, Zhang C, Deng Q, Katsaros D, Mayne ST, et al. Association of large noncoding RNA HOTAIR expression and its downstream intergenic CpG island methylation with survival in breast cancer. *Breast Cancer Res Treat*. 2012; 136(3):875–83. <https://doi.org/10.1007/s10549-012-2314-z> PMID: 23124417
 25. Schlesinger F, Smith AD, Gingeras TR, Hannon GJ, Hodges E. De novo DNA demethylation and non-coding transcription define active intergenic regulatory elements. *Genome Res*. 2013; 23(10):1601–14. <https://doi.org/10.1101/gr.157271.113> PMID: 23811145
 26. Robertson KD. DNA methylation and human disease. *Nat Rev Genet*. 2005; 6(8):597–610. <https://doi.org/10.1038/nrg1655> PMID: 16136652
 27. Ehrlich M. DNA hypomethylation in cancer cells. *Epigenomics*. 2009; 1(2):239–59. <https://doi.org/10.2217/epi.09.33> PMID: 20495664
 28. De Smet C, Loriot A. DNA hypomethylation in cancer: epigenetic scars of a neoplastic journey. *Epigenetics*. 2010; 5(3):206–13. <https://doi.org/10.4161/epi.5.3.11447> PMID: 20305381
 29. Feinberg AP, Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*. 1983; 301(5895):89–92. <https://doi.org/10.1038/301089a0> PMID: 6185846
 30. Yang X, Lay F, Han H, Jones PA. Targeting DNA methylation for epigenetic therapy. *Trends Pharmacol Sci*. 2010; 31(11):536–46. <https://doi.org/10.1016/j.tips.2010.08.001> PMID: 20846732
 31. Stresemann C, Lyko F. Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. *Int J Cancer*. 2008; 123(1):8–13. <https://doi.org/10.1002/ijc.23607> PMID: 18425818
 32. Li LH, Olin EJ, Buskirk HH, Reineke LM. Cytotoxicity and mode of action of 5-azacytidine on L1210 leukemia. *Cancer Res*. 1970; 30(11):2760–9. PMID: 5487063
 33. Girault I, Tozlu S, Lidereau R, Bieche I. Expression analysis of DNA methyltransferases 1, 3A, and 3B in sporadic breast carcinomas. *Clin Cancer Res*. 2003; 9(12):4415–22. PMID: 14555514
 34. Hervouet E, Peixoto P, Delage-Mourroux R, Boyer-Guittaut M, Cartron PF. Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. *Clin Epigenetics*. 2018; 10:17. <https://doi.org/10.1186/s13148-018-0450-y> PMID: 29449903
 35. Zhang Z, Wu WS. Sodium butyrate promotes generation of human induced pluripotent stem cells through induction of the miR302/367 cluster. *Stem Cells Dev*. 2013; 22(16):2268–77. <https://doi.org/10.1089/scd.2012.0650> PMID: 23534850
 36. Katz LS, Geras-Raaka E, Gershengorn MC. Reprogramming adult human dermal fibroblasts to islet-like cells by epigenetic modification coupled to transcription factor modulation. *Stem Cells Dev*. 2013; 22(18):2551–60. <https://doi.org/10.1089/scd.2013.0134> PMID: 23627894
 37. Hu Y, Cui Q, Gu Y, Sheng L, Wu K, Shi J, et al. Decitabine facilitates the generation and immunosuppressive function of regulatory gamma delta T cells derived from human peripheral blood mononuclear cells. *Leukemia*. 2013; 27(7):1580–5. <https://doi.org/10.1038/leu.2012.345> PMID: 23192015
 38. Aubele M, Schmitt M, Napieralski R, Paepke S, Ettl J, Absmaier M, et al. The Predictive Value of PITX2 DNA Methylation for High-Risk Breast Cancer Therapy: Current Guidelines, Medical Needs, and Challenges. *Dis Markers*. 2017; 2017:4934608.
 39. Absmaier M, Napieralski R, Schuster T, Aubele M, Walch A, Magdolen V, et al. PITX2 DNA-methylation predicts response to anthracycline-based adjuvant chemotherapy in triple-negative breast cancer patients. *Int J Oncol*. 2018; 52(3):755–67. <https://doi.org/10.3892/ijo.2018.4241> PMID: 29328369
 40. Xu J, Sun T, Guo X, Wang Y, Jing M. Estrogen receptor-alpha promoter methylation is a biomarker for outcome prediction of cisplatin resistance in triple-negative breast cancer. *Oncol Lett*. 2018; 15(3):2855–62. <https://doi.org/10.3892/ol.2017.7637> PMID: 29456719
 41. Fujii S, Yamashita S, Yamaguchi T, Takahashi M, Hozumi Y, Ushijima T, et al. Pathological complete response of HER2-positive breast cancer to trastuzumab and chemotherapy can be predicted by HSD17B4 methylation. *Oncotarget*. 2017; 8(12):19039–48. <https://doi.org/10.18632/oncotarget.15118> PMID: 28186977
 42. Gampenrieder SP, Rinnerthaler G, Hackl H, Pulverer W, Weinhaeusel A, Illic S, et al. DNA Methylation Signatures Predicting Bevacizumab Efficacy in Metastatic Breast Cancer. *Theranostics*. 2018; 8(8):2278–88. <https://doi.org/10.7150/thno.23544> PMID: 29721079

43. White MC, Holman DM, Boehm JE, Peipins LA, Grossman M, Henley SJ. Age and cancer risk: a potentially modifiable relationship. *Am J Prev Med.* 2014; 46(3 Suppl 1):S7–15. <https://doi.org/10.1016/j.amepre.2013.10.029> PMID: 24512933
44. Akinyemiju T, Wiener H, Pisu M. Cancer-related risk factors and incidence of major cancers by race, gender and region; analysis of the NIH-AARP diet and health study. *BMC Cancer.* 2017; 17(1):597. <https://doi.org/10.1186/s12885-017-3557-1> PMID: 28854891
45. Ozdemir BC, Dotto GP. Racial Differences in Cancer Susceptibility and Survival: More Than the Color of the Skin? *Trends Cancer.* 2017; 3(3):181–97. <https://doi.org/10.1016/j.trecan.2017.02.002> PMID: 28718431
46. Euhus DM, Bu D, Milchgrub S, Xie XJ, Bian A, Leitch AM, et al. DNA methylation in benign breast epithelium in relation to age and breast cancer risk. *Cancer Epidemiol Biomarkers Prev.* 2008; 17(5):1051–9. <https://doi.org/10.1158/1055-9965.EPI-07-2582> PMID: 18483325
47. Jung M, Pfeifer GP. Aging and DNA methylation. *BMC Biol.* 2015; 13:7. <https://doi.org/10.1186/s12915-015-0118-4> PMID: 25637097
48. Song MA, Brasky TM, Marian C, Weng DY, Taslim C, Dumitrescu RG, et al. Racial differences in genome-wide methylation profiling and gene expression in breast tissues from healthy women. *Epigenetics.* 2015; 10(12):1177–87. <https://doi.org/10.1080/15592294.2015.1121362> PMID: 26680018
49. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics.* 2016; 8(3):389–99 <https://doi.org/10.2217/epi.15.114> PMID: 26673039