



A Prognostic Model for Breast Cancer Based on Cancer Incidence-Related DNA Methylation Pattern

Zhenchong Xiong^{1†}, Lin Yang^{2†}, Juan Ao^{3†}, Jiarong Yi¹, Xiazi Zouxu¹, Wenjing Zhong¹, Jikun Feng¹, Weiling Huang¹, Xi Wang^{1*} and Zeyu Shuang^{1*}

¹State Key Laboratory of Oncology in South China, Collaborative Innovation Center of Cancer Medicine, Department of Breast Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China, ²Department of Radiation Oncology, Nanfang Hospital, Southern Medical University, Guangzhou, China, ³Department of Neurology, Guangzhou First People's Hospital, Guangzhou, China

OPEN ACCESS

Edited by:

Jitian Li,
Henan Luoyang Orthopedic Hospital
(Henan Provincial Orthopedic
Hospital), China

Reviewed by:

Tianzhu Lu,
Jiangxi Cancer Hospital, China
San-Gang Wu,
First Affiliated Hospital of Xiamen
University, China
Si-Qi Qiu,
Shantou Central Hospital, China

*Correspondence:

Xi Wang
wangxi@systucc.org.cn
Zeyu Shuang
shuangzy@systucc.org.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 13 November 2021

Accepted: 13 December 2021

Published: 03 January 2022

Citation:

Xiong Z, Yang L, Ao J, Yi J, Zouxu X,
Zhong W, Feng J, Huang W, Wang X
and Shuang Z (2022) A Prognostic
Model for Breast Cancer Based on
Cancer Incidence-Related DNA
Methylation Pattern.
Front. Genet. 12:814480.
doi: 10.3389/fgene.2021.814480

Breast cancer (BC) is the most diagnosed cancer and the leading cause of cancer-related deaths in women. The purpose of this study was to develop a prognostic model based on BC-related DNA methylation pattern. A total of 361 BC incidence-related probes (BCIPs) were differentially methylated in blood samples from women at high risk of BC and BC tissues. Twenty-nine of the 361 BCIPs that significantly correlated with BC outcomes were selected to establish the BCIP score. BCIP scores based on BC-related DNA methylation pattern were developed to evaluate the mortality risk of BC. The correlation between overall survival and BCIP scores was assessed using Kaplan–Meier, univariate, and multivariate analyses. In BC, the BCIP score was significantly correlated with malignant BC characteristics and poor outcomes. Furthermore, we assessed the BCIP score-related gene expression profile and observed that genes with expressions associated with the BCIP score were involved in the process of cancer immunity according to GO and KEGG analyses. Using the ESTIMATE and CIBERSORT algorithms, we discovered that BCIP scores were negatively correlated with both T cell infiltration and immune checkpoint inhibitor response markers in BC tissues. Finally, a nomogram comprising the BCIP score and BC prognostic factors was used to establish a prognostic model for patients with BC, while C-index and calibration curves were used to evaluate the effectiveness of the nomogram. A nomogram comprising the BCIP score, tumor size, lymph node status, and molecular subtype was developed to quantify the survival probability of patients with BC. Collectively, our study developed the BCIP score, which correlated with poor outcomes in BC, to portray the variation in DNA methylation pattern related to BC incidence.

Keywords: breast cancer, prognosis, DNA methylation, immune, nomogram

Abbreviations: BC, breast cancer; BCIPs, breast cancer incidence-related probes; c-index, the concordance index; CAFs, cancer-associated fibroblasts; Chr, Chromosome; DMP, differentially methylated probe; DNMTs, DNA methyltransferases; knn, k-nearest neighbors; ER, estrogen receptor; GEO, Gene Expression Omnibus; GO, Gene Ontology; GSEA, Gene Set Enrichment Analysis; HRs, hazard ratios; ICI, immune checkpoint inhibitor; IFNG, IFN- γ signature; IFNG.GS, IFNG hallmark gene set; IFN.ISGs, IFN stimulated genes resistance signature; KEGG, Kyoto Encyclopedia of Genes and Genomes; MDSCs, myeloid-derived suppressor cells; PR, progesterone receptor; TNBC, triple-negative breast cancer; SNPs, single nucleotide polymorphisms; TCGA, The Cancer Genome Atlas.

INTRODUCTION

Breast cancer (BC), the most diagnosed cancer globally, is the leading cause of cancer-related deaths in women (Siegel et al., 2020). As the incidence of BC has continuously increased over recent years, BC has become a major public health problem, and one in eight women would be affected by BC by the age of 85 years in high-income countries (Britt et al., 2020). With a deep understanding of cancer biology, BC is subdivided into four molecular subtypes (luminal A, luminal B, HER2, and TNBC) based on the expression of the estrogen receptor (ER), progesterone receptor (PR), HER2 receptor, and Ki-67 (Abreu et al., 2020). Although BC therapy guided by molecular subtyping greatly reduces BC mortality, cancer recurrence inevitably occurs (Alwan and Tawfeeq, 2019). Investigations of prognostic methods for predicting outcomes in BC may provide clues for improving cancer treatment.

DNA methylation, an epigenetic modification, plays an important role in cancer development (Jones et al., 2016). Although previous studies have shown that BC is a genetic disease, genome-wide variations in DNA methylation have been observed in cancer cells (Flavahan et al., 2017; Widschwendter et al., 2018). DNA methylation induces chromatin structure changes and inhibits gene expression mediated by DNA methyltransferases. Aberrations in DNA methylation, which are affected by environmental, lifestyle-related, and heritable factors, may induce cancer development by silencing tumor suppressors and/or re-activating oncogenes (Baylin and Jones, 2011; Koo et al., 2015; Muvarak et al., 2016). In BC, hypomethylation of stemness- and proliferation-associated genes in circulating tumor cells promotes stemness and metastasis (Gkoutela et al., 2019). Thienpont et al. indicated that tumor hypoxia induces hypermethylation and promotes BC progression by inactivating TET enzymes (Thienpont et al., 2016). Thus, tracking the changes in DNA methylation pattern associated with BC is helpful for developing prognostic methods for BC.

The purpose of our study was to discover DNA methylation pattern and develop a prognostic model based on BC-related DNA methylation pattern. We identified 361 breast cancer incidence-related probes (BCIPs) that were differentially methylated in blood samples from women at a high risk of BC and BC tissues. Twenty-nine of the 361 BCIPs that were significantly correlated with BC outcomes were included to establish the BCIP score. In BC, the BCIP scores were significantly correlated with malignant BC characteristics and poor outcomes. Furthermore, we assessed the BCIP score-related gene expression profile and observed that the BCIP score-related gene profile participated in the process of cancer immunity. BCIP scores were negatively correlated with immune cell infiltration and the immune checkpoint inhibitor (ICI) response in BC tissues. Finally, a nomogram comprising the BCIP score, tumor size, lymph node status, and molecular subtype was developed to quantify the survival probability of patients with BC.

MATERIALS AND METHODS

Data Collection and Processing

For the GSE51057, GSE72308, and TCGA-BRCA DNA methylation datasets, genome-wide methylation data were profiled using Illumina Infinium HumanMethylation450 BeadChips Assay (Illumina 450 K platform). For the GSE57285 DNA methylation dataset, genome-wide methylation data were profiled using Illumina Infinium HumanMethylation27 BeadChips Assay (Illumina 27 K platform). The DNA methylation level of each probe was calculated using β values ranging from 0 (no DNA methylation) to 1 (complete DNA methylation). Probes containing missing values in over half of the samples in each dataset were removed, while missing values of the remaining probes were imputed with the k-nearest neighbors imputation method. Probes located on the sex chromosome and probes containing known single-nucleotide polymorphisms were removed (Price et al., 2013). Finally, 23,614 probes were selected for further investigation. The above process was performed using the R package Champ (Tian et al., 2017).

For gene expression data, mRNA expression data were obtained from the Cancer Genome Atlas (TCGA) database. Background correction and normalization of mRNA expression data were performed using the R package limma (Ritchie et al., 2015). Expression data for protein-encoding genes were included in further analysis.

Calculation of the BCIP Score

GSE51057 (including 177 blood samples from normal women and 146 blood samples from women diagnosed with BC after sample donation) and GSE57285 (including 49 blood samples from normal women and 35 blood samples from women diagnosed with BC after sample donation) were selected to identify differentially methylated probes (DMPs) that correlated with a high risk of BC. In addition, 76 cases with matched tumor and tumor-adjacent breast tissues from TCGA database were enrolled to identify DMPs in BC tissues. CpG probes that were commonly demethylated in blood samples from women with high risk of BC and BC tissues were defined as BCIPs.

DMPs were identified using the R package limma. Differential hyper/hypo-methylation probe was defined according to logF_c value. Hyper-methylation probe are defined as logF_c > 0, p value < 0.05 (blood/cancer sample of BC patients VS blood/cancer sample of non-BC patients). Hypo-methylation probe are defined as logF_c < 0, p value < 0.05 (blood/cancer sample of BC patients VS blood/cancer sample of non-BC patients). The distribution of BCIPs on chromosomes, CpG islands, and TSS regions was investigated using the R package Champ. The hazard ratios (HRs) of BCIPs with respect to OS were evaluated using TCGA-BRCA data.

Univariate Cox regression was used to calculate the HR of each BCIP, and BCIPs significantly correlated cancer survival in BC were included to develop the BCIP score model. The BCIP score model was assessed as follows: BCIP score = [(transformed HR¹ * β value of BCIP¹) + (transformed HR² * β value of BCIP²) + (transformed HRⁿ * β value of BCIPⁿ)]/[abs (transformed HR¹) +

TABLE 1 | Characteristics of patients in the training and validation cohort.

Variate	Training cohort (n = 587) Num (%)	Validation cohort (n = 231) Num (%)
Age		
0–39	37 (6.3)	28 (12.1)
≥40	550 (93.7)	203 (87.8)
Tumor size		
≤2 cm	154 (26.2)	117 (50.6)
>2 cm	433 (73.8)	114 (49.4)
Lymph node status		
Non-metastasis	261 (44.5)	130 (56.3)
Metastasis	326 (55.5)	101 (43.7)
Subtype (IHC)		
Luminal A/B	504 (85.9)	111 (48.1)
HER2	16 (2.7)	55 (23.8)
TNBC	67 (11.4)	65 (28.1)
BCIP score		
Low	425 (72.4)	146 (63.2)
High	162 (27.6)	85 (36.8)
Survival status		
Alive	536 (91.3)	189 (81.8)
Dead	51 (8.7)	42 (18.2)

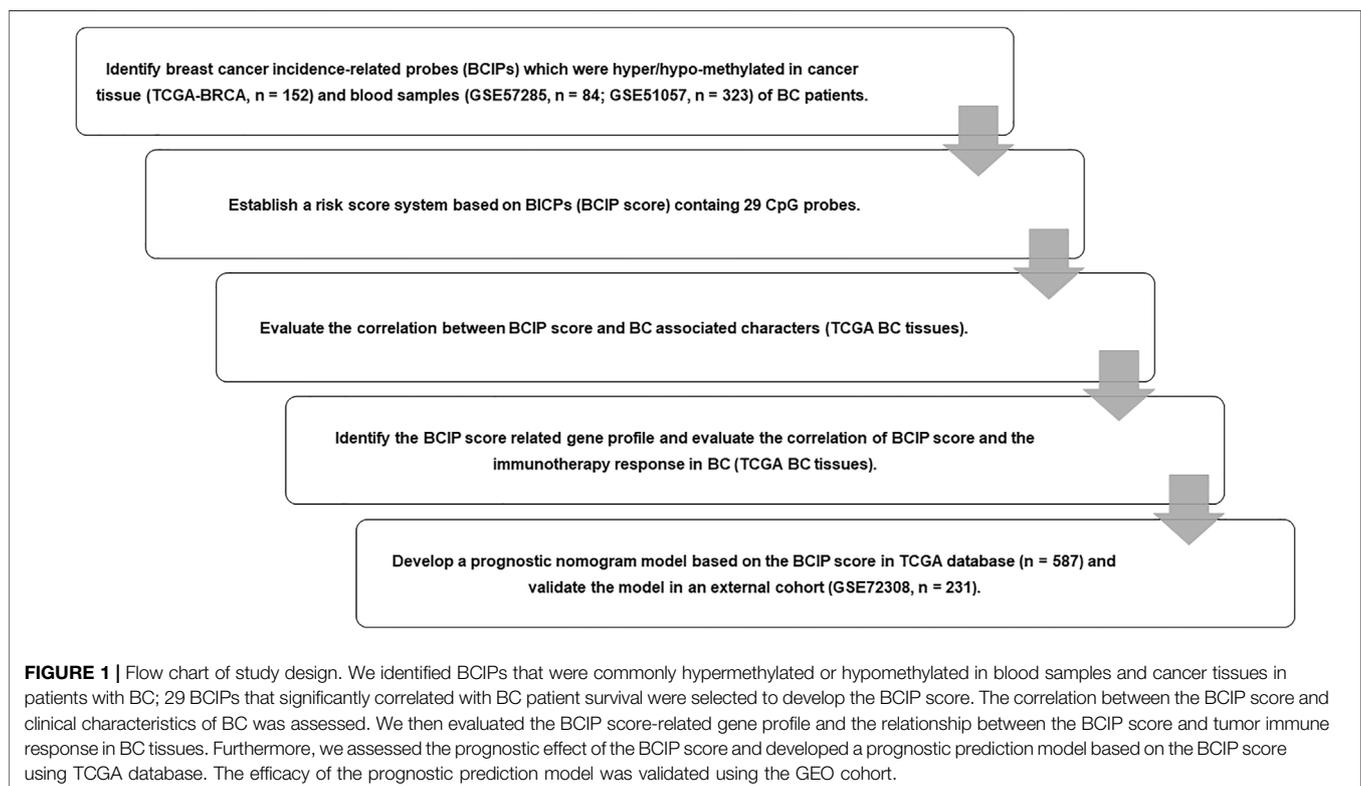
abs (transformed HR^2) + (transformed HR^n)]. The cutoff value of BCIP score was 0.2, identified using x-tile (<https://medicine.yale.edu/lab/rimm/research/software/>). Samples with a BCIP score of <0.2 were assigned to the low BCIP group, while those with a BCIP score of ≥0.2 were assigned to the high BCIP group.

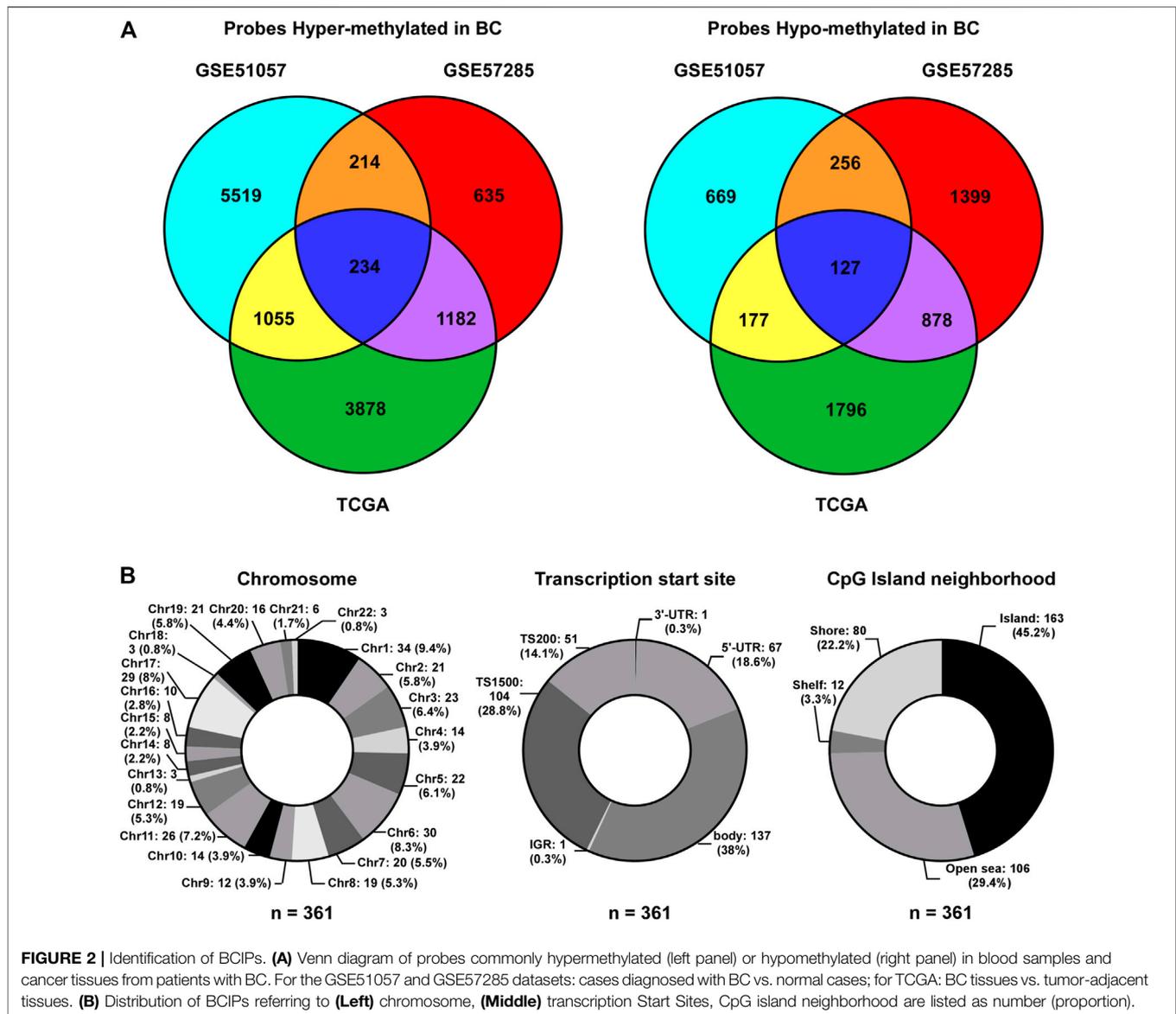
Functional and Clinical Characteristics Analysis of the BCIP Score in BC

BCIP scores of TCGA-BRCA tissues (76 cases with matched tumor and tumor-adjacent breast tissues and 699 cases with unmatched tumor tissues) were calculated using DNA methylation data. The correlation between the BCIP score and BC-related characteristics (tumor size, oncogene copy number variation, and oncogene expression) was assessed using linear regression and Spearman's correlation coefficient. The correlation between gene mRNA expression and the BCIP score was analyzed using Spearman's correlation. The gene expression profile associated with BCIP score was identified based on Spearman's coefficient (cutoff value: 0.1), and functional study of the related gene expression was performed using the GO and KEGG databases. The above procedure was performed using R software.

Correlation Analysis Between the BCIP Score and Immune Microenvironment in BC

The degree of infiltration of immune cells and stromal cells in each sample was assessed using the ESTIMATE algorithm (Yoshihara et al., 2013). The proportion of 22 immune cells in each tissue was evaluated using the CIBERSORT algorithm (<http://cibersort.stanford.edu/>) (Gentles et al., 2015). Correlations between the BCIP score and ESTIMATE and CIBERSORT scores were assessed using linear regression and Spearman's correlation coefficient.





Correlation Analysis Between the BCIP Score and Cancer Immunotherapy Response

Four biomarkers were used to assess the response to immunotherapy—CD274, CD8, IFN- γ signature (IFNG) (Ayers et al., 2017), and IFNG hallmark gene set (IFNG.GS) (Benci et al., 2019). Three biomarkers were used to assess resistance to immunotherapy—IFN-stimulated gene resistance signature (ISG.RS) (Benci et al., 2019), myeloid-derived suppressor cells (MDSCs), and cancer-associated fibroblasts (CAFs) (Joyce and Fearon, 2015). IFNG was calculated by averaging six genes (IFNG, STAT1, IDO1, CXCL9, CXCL10, and HLA-DRA). IFNG.GS was calculated as previously reported (Benci et al., 2019). CD274, CD8, ISG.RS, MDSCs, and CAFs were assessed using a web application (<http://tide.dfci.harvard.edu>).

Establishment and Validation of the Nomogram

A total of 587 BC cases with DNA methylation data, clinical characteristics, and complete follow-up in TCGA database were enrolled as the training cohort, while 231 BC cases with data of DNA methylation, clinical characteristics, and complete follow-up in the GSE72308 dataset were selected as the external validation cohort. The clinical and pathological characteristics of patients with BC are listed in **Table 1**. Tumor size was defined as ≤ 2 cm or >2 cm. Lymph node status was defined as non-metastasis or lymph node metastasis. The molecular subtypes were defined based on the IHC assessment of ER, PR, and HER2 as follows: luminal A/B (ER+/PR+/HER2-, ER-/PR+/HER2-, ER+/PR-/HER2-, ER+/PR+/HER2+, ER-/PR+/HER2+, ER+/PR-/HER2+),

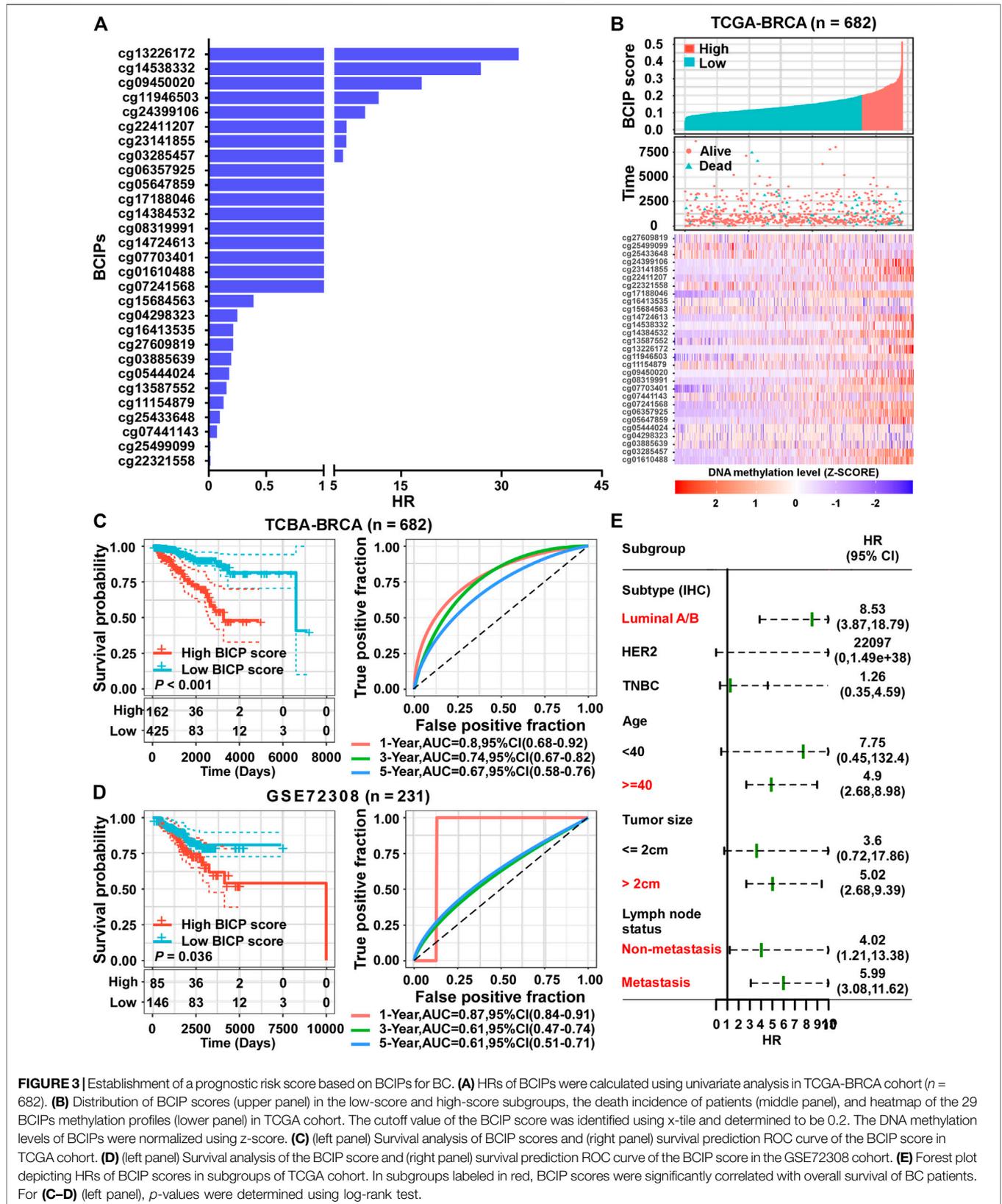
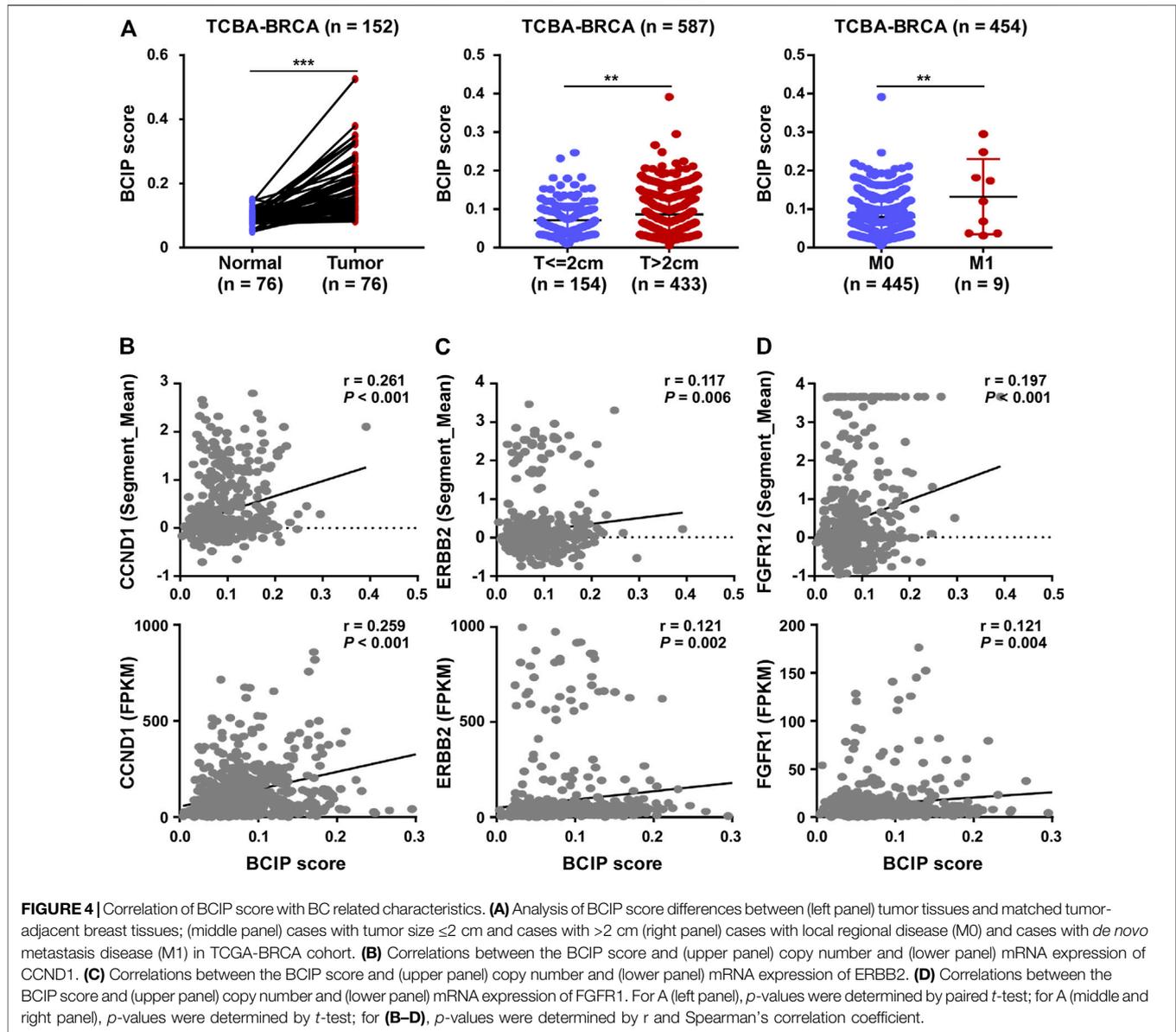


FIGURE 3 | Establishment of a prognostic risk score based on BCIPs for BC. **(A)** HRs of BCIPs were calculated using univariate analysis in TCGA-BRCA cohort (n = 682). **(B)** Distribution of BCIP scores (upper panel) in the low-score and high-score subgroups, the death incidence of patients (middle panel), and heatmap of the 29 BCIPs methylation profiles (lower panel) in TCGA cohort. The cutoff value of the BCIP score was identified using x-tile and determined to be 0.2. The DNA methylation levels of BCIPs were normalized using z-score. **(C)** (left panel) Survival analysis of BCIP scores and (right panel) survival prediction ROC curve of the BCIP score in TCGA cohort. **(D)** (left panel) Survival analysis of the BCIP score and (right panel) survival prediction ROC curve of the BCIP score in the GSE72308 cohort. **(E)** Forest plot depicting HRs of BCIP scores in subgroups of TCGA cohort. In subgroups labeled in red, BCIP scores were significantly correlated with overall survival of BC patients. For **(C–D)** (left panel), p-values were determined using log-rank test.



HER2 (ER-/PR-/HER2+), and TNBC (ER-/PR-/HER2-). Age was defined as 0–39 and ≥ 40 years.

BC features, which were significantly correlated with BC survival in the multivariate Cox regression, were selected to establish the nomogram model. The patients' survival probability was assumed by summing the scores of the variates, with a higher score corresponding to a higher mortality risk. The efficiency of the model was evaluated with regard to discrimination and calibration. The concordance index (c-index) was used to quantify discrimination ranging from 0 to 1 (<0.5, absolute discordance; 0.5, equal concordance to chance; and 1, best concordance). The calibration curve was used to compare the predicted survival probability with the observed survival probability at 3, 5, and 10 years in the training and validation cohorts.

RESULTS

Identification of BCIPs

To identify CpG probes associated with BC incidence, two GEO datasets (GSE51057 and GSE57285), including 226 blood samples from healthy women (177 cases in GSE51057 and 49 cases in GSE57285) and 181 blood samples from women diagnosed with BC after sample donation (146 cases in GSE51057 and 35 cases in GSE57285; **Figure 1**), were used. The 448 and 383 CpG probes were commonly hypermethylated or hypomethylated in samples from women with BC compared to those in samples from healthy women (**Figure 2A**). Further, we included 76 paired tumor-adjacent breast tissues and tumor tissues from TCGA-BRCA database and identified 9,327 differentially methylated probes (tumor vs. tumor-adjacent;

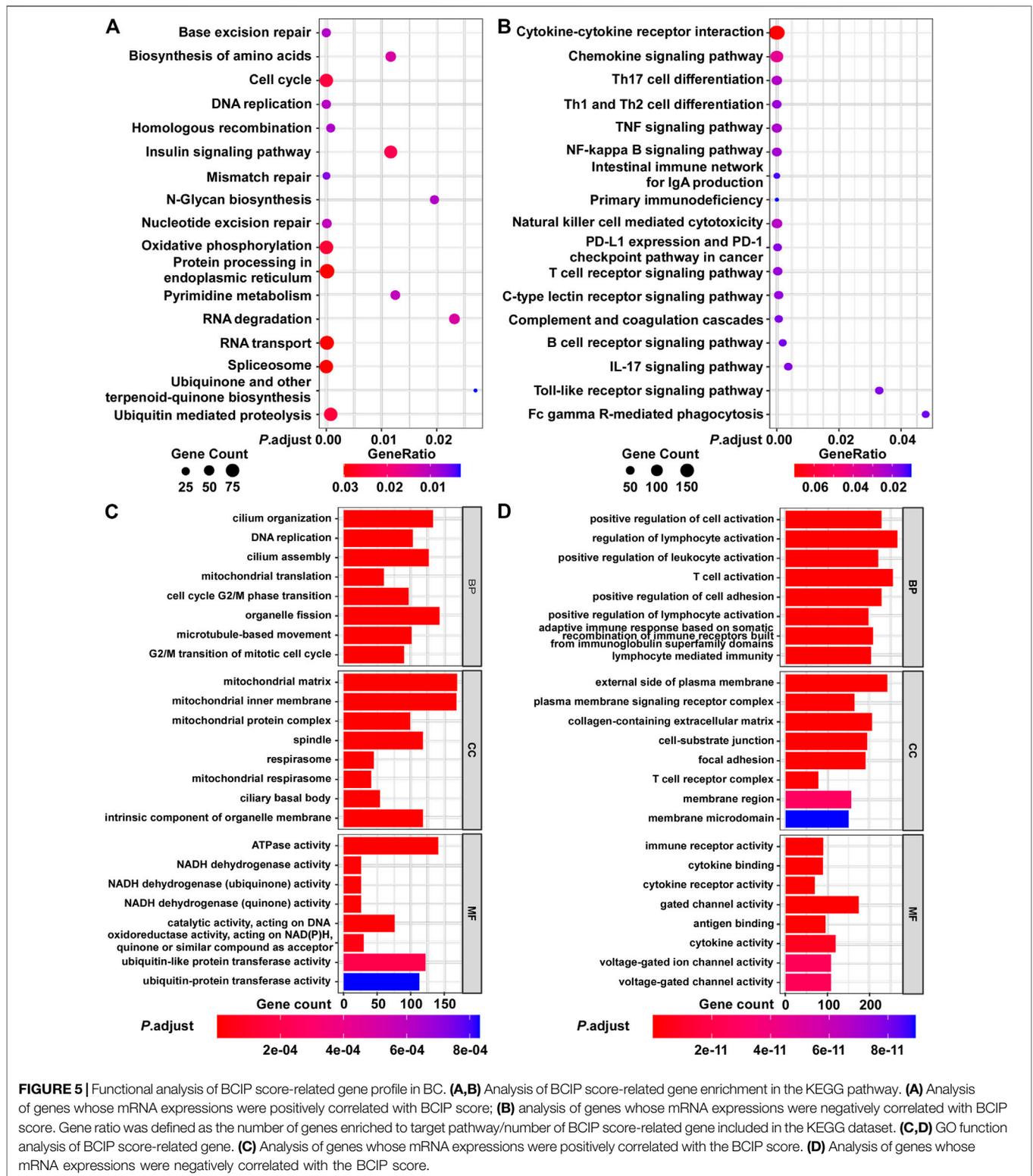
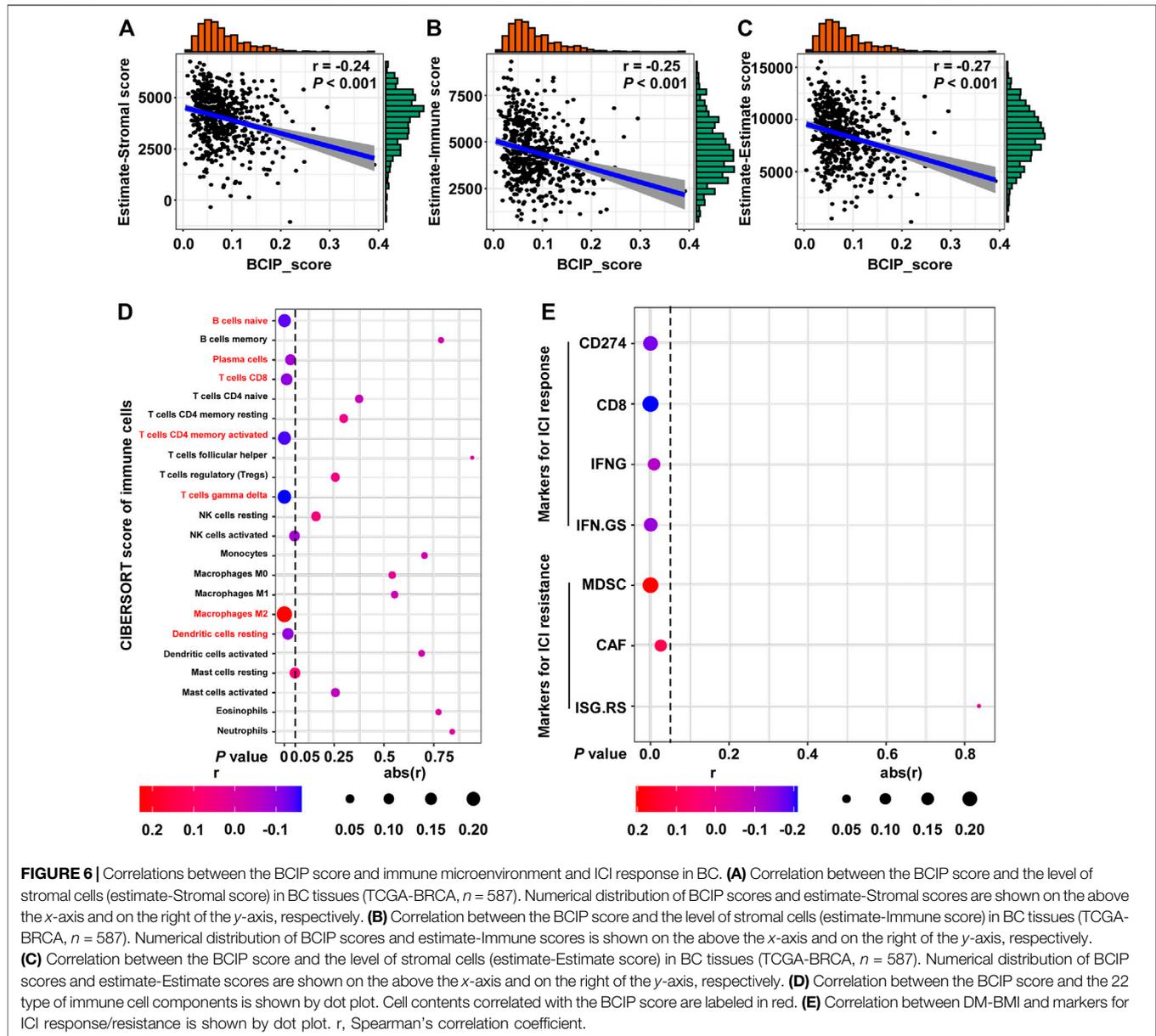


FIGURE 5 | Functional analysis of BCIP score-related gene profile in BC. **(A,B)** Analysis of BCIP score-related gene enrichment in the KEGG pathway. **(A)** Analysis of genes whose mRNA expressions were positively correlated with BCIP score; **(B)** analysis of genes whose mRNA expressions were negatively correlated with BCIP score. Gene ratio was defined as the number of genes enriched to target pathway/number of BCIP score-related gene included in the KEGG dataset. **(C,D)** GO function analysis of BCIP score-related gene. **(C)** Analysis of genes whose mRNA expressions were positively correlated with the BCIP score. **(D)** Analysis of genes whose mRNA expressions were negatively correlated with the BCIP score.

hypermethylated: 6,349; hypomethylated: 2,978). By integrating the results from the GEO and TCGA cohorts, 234 hypermethylated and 127 hypomethylated CpG probes were identified as BCIPs in blood samples from women with

high risk of BC and BC tissues (**Figure 2A**). Next, we assessed the distribution of BCIPs based on chromosomes, transcription start sites, and CpG islands. Among the 22 pairs, chromosomes (Chr) 1, 6, 11, and 17 were the most



common regions for BCIP distribution; 45.2% were distributed in CpG islands and 42.9% were distributed in the promoter regions (TS200 and TS1500; **Figure 2B**).

Establishment of a Prognostic Risk Score Based on BCIPs for BC

Using a univariate Cox regression model, we evaluated the association between methylation levels of BCIPs and overall survival in TCGA-BRCA cohort. DNA methylation levels of 29/361 BCIPs significantly correlated with OS in BC, and these probes were selected to establish the BCIP score model (**Figure 3A** and **Supplemental Table S1**). Using X-Tile analysis, the cutoff value for the BCIP score was set at 0.2; patients with a

BCIP score ≤ 0.2 were assigned to the low score group, while patients with a BCIP score of >0.2 were assigned to the high score group (**Figure 3B**). In TCGA cohort, patients with low BCIP scores had better survival rates than those in the high score group (**Figure 3C**). The ROC curve showed that the BCIP score exhibited a high predictive efficacy for OS in BC (**Figure 3C**). Likewise, a high BCIP score predicted a high mortality risk for BC in the GEO cohort (**Figure 3D**). Further, subgroup analysis indicated that the BCIP score was a negative prognostic factor in subgroups of BC with luminal A/B subtype, old age (≥ 40 years), larger tumor size, and lymph node metastasis status (**Figure 3E** and **Supplemental Table S2**). These data show that the BCIP score is an efficient prognostic model for BC.

TABLE 2 | Univariate and multivariate analysis for patients with BC in TCGA cohort.

Variate	Univariate			Multivariate		
	HR	95% CI	p Value	HR	95% CI	p Value
Age			0.339			
Age	1					
0–39	1.799	0.54–5.989				
≥40			0.04			0.045
Tumor size	1			1		
≤2 cm	2.467	1.044–5.83		2.41	1.021–5.688	
>2 cm			<0.001			<0.001
Lymph node status	1			1		
Non-metastasis	4.18	2.087–8.371		3.959	1.997–7.848	
Subtype (IHC)						<0.001
Luminal A/B	1		<0.001	1		
HER2	3.39	0.789–14.561		3.445	0.801–14.816	
TNBC	6.034	3.016–12.073		5.561	2.833–10.918	
BCIP score			<0.001			<0.001
Low	1			1		
High	6.618	3.628–12.073		6.748	3.701–12.304	

Functional and Clinical Characteristic Analysis of the BCIP Score in BC

Compared to tumor-adjacent breast tissues, BC tissues exhibited a significantly higher BCIP score (Figure 4A). In BC, patients with a larger tumor size and *de novo* metastatic disease had a higher BCIP score (Figure 4A). By assessing the gene copy number in BC tissues, the BCIP score was associated with an increased copy number of several oncogenes (including *CCND1*, *ERBB2*, and *FGFR1*; Figures 4B–D). Consistently, the BCIP score was positively correlated with the mRNA levels of *CCND1*, *ERBB2*, and *FGFR1* (Figures 4B–D).

Next, we evaluated the biological significance of the BCIP score in BC. By assessing the correlation between the BCIP score and gene expression, we identified the BCIP score-related gene expression profile. The results of correlation analysis are shown in Supplementary Table S3. Genes with expression correlated with the BCIP score were included in GSEA analysis using the KEGG and GO databases. KEGG and GO analyses showed that genes with expression that positively correlated with the BCIP score were significantly enriched in pathways involving cell cycle regulation, DNA replication, DNA repair (base excision repair, mismatch repair, nucleotide excision repair, and homologous recombination), and energy metabolism (Figures 5A,C). Interestingly, both KEGG and GO analyses showed that genes with expressions that negatively correlated with the BCIP score were significantly involved in cancer-immunity-related pathways (including antigen binding, T cell differentiation and activation, the PD-1 checkpoint pathway, NK cell-mediated cytotoxicity, and immune receptor activity) (Figures 5B,D).

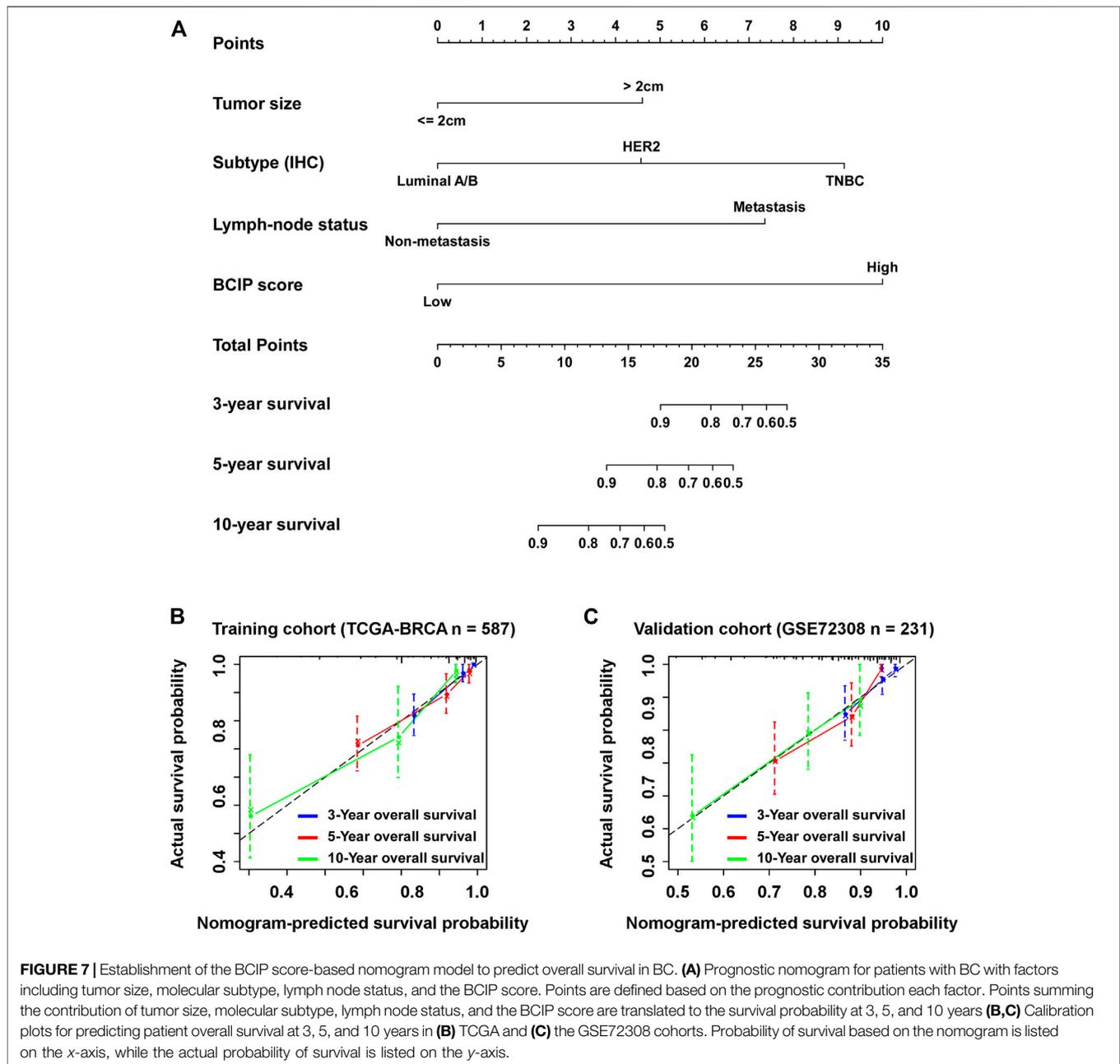
Correlations Between the BCIP Score and Immune Microenvironment and ICI Response in BC

Further, we used the ESTIMATE algorithm to evaluate the correlation between the BCIP score and degree of immune cell infiltration in BC tissues. The BCIP score was significantly

correlated with decreased levels of immune and stromal cell content, indicating that the BCIP score was negatively correlated with immune cell infiltration in BC tissues (Figures 6A–C). The CIBERSORT algorithm was used to evaluate the association between the BCIP score and 22 immune cell contents in BC tissues (TCGA-BRCA cohort). The BCIP score was negatively correlated with the level of several immune cells with antitumor activity (including plasma, CD8⁺ T, CD4⁺ T, and gamma delta T cells; Figure 6D). In addition, the level of resting dendritic cells, which play a critical role in antigen phagocytosis and processing, decreased in BC tissues with increased BCIP scores (Figure 6D). In contrast, the level of M2 macrophages, which promote tumor progression, was positively correlated with the BCIP score (Figure 6D). These results indicate that the BCIP score was correlated with poor antitumor immunity. As a previous study showed that ICIs (Shah et al., 2012) significantly improved cancer survival through T cell immunity in BC, we further investigated whether the BCIP score correlated with the ICI response in BC. Four markers for ICI sensitivity and three markers of ICI resistance were selected to evaluate the ICI response in BC tissues. The BCIP score was negatively correlated with ICI-sensitive markers (CD274, CD8, IFNG, and IFN. GS; Figure 6E). In contrast, two of the three ICI resistance markers (MDSC and CAF) were positively correlated with the BCIP score in BC (Figure 6E). Collectively, the BCIP score was a negative marker of immune cell infiltration and ICI response in BC.

Establishment of a BCIP Score-Based Nomogram Model to Predict Overall Survival in BC

Univariate and multivariate Cox regression analyses showed that tumor size, molecular subtype, lymph node status, and the BCIP score were significant prognostic factors for BC (Table 2). After including the above variables, we developed a comprehensive prognostic nomogram based on TCGA-BRCA cohort



(Figure 7A). Factors correlated with high mortality risk (larger tumor size, TNBC subtype, metastatic lymph node status, and high BCIP score) were scored higher than those correlated with low mortality risk (smaller tumor size, luminal A/B subtype, non-lymph node metastasis, and low BCIP score). The c-index of this model was 0.831 (95% CI: 0.774–0.888) in TCGA-BRCA cohort. Furthermore, data from the GSE72308 dataset were selected for external validation of the nomogram. The c-index of the model was 0.734 (95% CI: 0.665–0.803) in the external validation cohort. The calibration curve was constructed to evaluate the accuracy of model prediction and indicated that the BCIP score-based nomogram exhibited good consistency in the prediction of 3-, 5-, and 10-year survival

probabilities in both TCGA-BRCA and the GSE72308 cohorts (Figures 7B,C).

DISCUSSION

For years, prognostic prediction in patients with BC has been mostly based on the pathological features of the tumor (including tumor size, lymph node status, distant metastatic status, and molecular subtype) (Harbeck and Gnant, 2017). BC therapy guided by these prognostic indicators has significantly improved cancer survival and avoids the therapeutic side effects caused by overtreatment in BC (DeSantis et al.,

2019). With advances in BC treatment, more precise prognostic methods are required (Denkert et al., 2017). Several prognostic biomarkers, including gene sequencing, gene copy number, and circulating tumor cells, have been used in clinical practice for BC (Lord and Ashworth, 2016; Tsai et al., 2018; Radovich et al., 2020). For instance, the 21-gene assay (Oncotype Dx) has been used to identify patients with low recurrence risk who could be exempted from chemotherapy in T1b/c and T2, HR+, HER2-, and lymph node-negative BC (Giuliano et al., 2017). Detection of *BRCA1/2* mutation status is helpful in identifying the potential benefits of PARP inhibitors (Tarsounas and Sung, 2020). Recently, Fackler et al. identified hypermethylation signatures that were correlated with cancer recurrence in TNBC (Fackler et al., 2020). In this study, we identified DNA methylation signatures related to BC incidence and developed a nomogram based on DNA methylation to calculate the 3-, 5-, and 10-year survival probabilities for BC.

An increasing amount of data has shown that aberrations in DNA methylation correlated with breast malignancy development is a prerequisite for the transformation of normal cells into BC cells, and these changes in DNA methylation accumulate in malignant cells, inducing an enhanced ability for proliferation and self-renewal (Zhuang et al., 2012; Wienken et al., 2016). As DNA methylation is regulated by environmental factors, genetic predisposition, and individual lifestyle, variation in DNA methylation pattern can be a reflection of the individual response to exposure to BC risk (Widschwendter et al., 2018). Identification of cancer incidence-related DNA methylation pattern is of great value for prognostic prediction in BC. Hence, we developed the BCIP score based on BC-related DNA methylation pattern and observed that the BCIP score was significantly correlated with poor outcomes in patients with BC.

In BC tissues, the BCIP score not only acted as a prognostic biomarker but also significantly correlated with aggressive BC features (such as larger tumor size, distant metastatic disease, and oncogene amplification). In addition to being a reflection of biological changes, changes in DNA methylation play a critical role in tumor progression through the regulation of gene expression. By assessing the BCIP score-related gene expression profile, we discovered that genes with expressions that were negatively correlated with the BCIP score were also significantly involved in the cancer immunity-related pathway. As a high BCIP score correlated with increased mortality, an aberration of cancer immunity might account for the poor outcome of patients with a high BCIP score.

In our study, we observed that BCIP scores were negatively correlated with the degree of T cell infiltration in BC tissues, indicating that the T-cell-mediated immune response (cellular immunity) was aberrantly inactivated. Cellular immunity is the main type of tumor immune response, which is mediated by the direct killing effect of T cells and release of cytokines by T cells (Joyce and Fearon, 2015). When activated by tumor antigens, T cells become immunoreactive and acquire the ability to recognize and kill tumor cells. However, T cell activation is often blocked by immune checkpoint molecules (including PD-1/PDL1, CTLA-4, and TIM-3) in individuals with cancer (Joyce and Fearon, 2015; Voorwerk et al., 2019). In BC, ICIs targeting PD1/PDL1 have been proven to improve survival in some patients, while most patients exhibit a poor response to ICIs

(Voorwerk et al., 2019). To improve the efficacy of ICIs, effective biomarkers are required to identify potential beneficiaries. As our data showed that the BCIP score was negatively correlated with ICI response markers, the BCIP score is a potential biomarker to predict the sensitivity of BC to ICIs.

It is well known that DNA methylation correlates with suppressed gene expression, indicating that DNA methylation alterations can be exploited in cancer diagnosis. Compared with other genetic approaches (such as mutational analysis), DNA methylation-based approaches present advantages with regard to their clinical application (Heyn and Esteller, 2012). For instance, DNA methylation detection mostly focuses on a specific promoter region containing CpG islands, while mutational studies can cope with large regions since point mutations are located throughout the length of the gene (Heyn and Esteller, 2012). Moreover, alterations in DNA methylation were detected in a higher proportion of tumor tissue than genetic alterations, leading to higher sensitivity in prognostic analysis (Esteller et al., 2001). In our study, we developed a BCIP scoring system based on BC-related DNA methylation variation and observed that the BCIP score was positively correlated with mortality in patients with BC. Further, we adjusted for the confounding effects of tumor size, lymph node status, and molecular subtypes using Cox regression and found that BCIP was an independent prognostic factor for BC. By integrating the BCIP score, tumor size, lymph node status, and molecular subtypes, we established a nomogram model that could accurately quantify the survival probability of patients with BC. Based on our nomogram, patients with BC and a low mortality risk could be identified and might be exempted from aggressive and excessive medical treatment. However, patients with a high mortality risk should undergo more intensive surveillance for cancer recurrence.

Collectively, our study identified the variation in DNA methylation pattern related to BC incidence and developed a BCIP score model depicting BC-related DNA methylation variation. In BC, the BCIP score was significantly correlated with malignant BC characteristics and poor outcomes. Furthermore, we observed that the BCIP score was negatively correlated with immune cell infiltration and ICI response markers in BC tissues. Finally, a nomogram comprising the BCIP score, tumor size, lymph node status, and molecular subtype was developed to quantify the survival probability of patients with BC.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

ZS and XW designed the overall project; ZX, LY, and JA wrote the manuscript; JF, WZ, and WH revised and polished the manuscript. ZX, ZS, JY, and XZ performed the statistical analysis of the data; All the authors have read and approved the final manuscript.

FUNDING

This work was supported by project grants from the National Natural Science Foundation of China (81802663 and 82002557).

REFERENCES

- Abreu, M., Cabezas-Sainz, P., Pereira-Veiga, T., Falo, C., Abalo, A., Morilla, I., et al. (2020). Looking for a Better Characterization of Triple-Negative Breast Cancer by Means of Circulating Tumor Cells. *Jcm* 9 (2), 353. doi:10.3390/jcm9020353
- Alwan, N. A. S., and Tawfeeq, F. N. (2019). Comparison of Clinico-Pathological Presentations of Triple-Negative versus Triple-Positive and HER2 Iraqi Breast Cancer Patients. *Open Access Maced J. Med. Sci.* 7 (21), 3534–3539. doi:10.3889/oamjms.2019.808
- Ayers, M., Lunceford, J., Nebozhyn, M., Murphy, E., Loboda, A., Kaufman, D. R., et al. (2017). IFN- γ -related mRNA Profile Predicts Clinical Response to PD-1 Blockade. *J. Clin. Invest.* 127 (8), 2930–2940. doi:10.1172/JCI91190
- Baylin, S. B., and Jones, P. A. (2011). A Decade of Exploring the Cancer Epigenome - Biological and Translational Implications. *Nat. Rev. Cancer* 11 (10), 726–734. doi:10.1038/nrc3130
- Benci, J. L., Johnson, L. R., Choa, R., Xu, Y., Qiu, J., Zhou, Z., et al. (2019). Opposing Functions of Interferon Coordinate Adaptive and Innate Immune Responses to Cancer Immune Checkpoint Blockade. *Cell* 178 (4), 933–948. doi:10.1016/j.cell.2019.07.019
- Britt, K. L., Cuzick, J., and Phillips, K.-A. (2020). Key Steps for Effective Breast Cancer Prevention. *Nat. Rev. Cancer* 20 (8), 417–436. doi:10.1038/s41568-020-0266-x
- Denkert, C., Liedtke, C., Tutt, A., and von Minckwitz, G. (2017). Molecular Alterations in Triple-Negative Breast Cancer-The Road to New Treatment Strategies. *The Lancet* 389 (10087), 2430–2442. doi:10.1016/s0140-6736(16)32454-0
- DeSantis, C. E., Ma, J., Gaudet, M. M., Newman, L. A., Miller, K. D., Goding Sauer, A., et al. (2019/2019). Breast Cancer Statistics, 2019. *CA A. Cancer J. Clin.* 69 (6), 438–451. doi:10.3322/caac.21583
- Esteller, M., Corn, P. G., Baylin, S. B., and Herman, J. G. (2001). A Gene Hypermethylation Profile of Human Cancer. *Cancer Res.* 61 (8), 3225–3229. doi:10.1016/S0165-4608(00)00403-9
- Fackler, M. J., Cho, S., Cope, L., Gabrielson, E., Visvanathan, K., Wilsbach, K., et al. (2020). DNA Methylation Markers Predict Recurrence-free Interval in Triple-Negative Breast Cancer. *NPJ Breast Cancer* 6, 3. doi:10.1038/s41523-020-0145-3
- Flavahan, W. A., Gaskell, E., and Bernstein, B. E. (2017). Epigenetic Plasticity and the Hallmarks of Cancer. *Science* 357 (6348). doi:10.1126/science.aal2380
- Gentles, A. J., Newman, A. M., Liu, C. L., Bratman, S. V., Feng, W., Kim, D., et al. (2015). The Prognostic Landscape of Genes and Infiltrating Immune Cells across Human Cancers. *Nat. Med.* 21 (8), 938–945. doi:10.1038/nm.3909
- Giuliano, A. E., Connolly, J. L., Edge, S. B., Mittendorf, E. A., Rugo, H. S., Solin, L. J., et al. (2017). Breast Cancer-Major Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual. *CA: A Cancer J. Clinicians* 67 (4), 290–303. doi:10.3322/caac.21393
- Gkoutela, S., Castro-Giner, F., Szczerba, B. M., Vetter, M., Landin, J., Scherrer, R., et al. (2019). Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding. *Cell* 176 (1-2), 98–112. doi:10.1016/j.cell.2018.11.046
- Harbeck, N., and Gnant, M. (2017). Breast Cancer. *The Lancet* 389 (10074), 1134–1150. doi:10.1016/S0140-6736(16)31891-8
- Heyn, H., and Esteller, M. (2012). DNA Methylation Profiling in the Clinic: Applications and Challenges. *Nat. Rev. Genet.* 13 (10), 679–692. doi:10.1038/nrg3270
- Jones, P. A., Issa, J.-P. J., and Baylin, S. (2016). Targeting the Cancer Epigenome for Therapy. *Nat. Rev. Genet.* 17 (10), 630–641. doi:10.1038/nrg.2016.93

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.814480/full#supplementary-material>

- Joyce, J. A., and Fearon, D. T. (2015). T Cell Exclusion, Immune Privilege, and the Tumor Microenvironment. *Science* 348 (6230), 74–80. doi:10.1126/science.aaa6204
- Koo, G.-B., Morgan, M. J., Lee, D.-G., Kim, W.-J., Yoon, J.-H., Koo, J. S., et al. (2015). Methylation-dependent Loss of RIP3 Expression in Cancer Represses Programmed Necrosis in Response to Chemotherapeutics. *Cell Res* 25 (6), 707–725. doi:10.1038/cr.2015.56
- Lord, C. J., and Ashworth, A. (2016). BRCAness Revisited. *Nat. Rev. Cancer* 16 (2), 110–120. doi:10.1038/nrc.2015.21
- Muvarak, N. E., Chowdhury, K., Xia, L., Robert, C., Choi, E. Y., Cai, Y., et al. (2016). Enhancing the Cytotoxic Effects of PARP Inhibitors with DNA Demethylating Agents - A Potential Therapy for Cancer. *Cancer Cell* 30 (4), 637–650. doi:10.1016/j.ccell.2016.09.002
- Price, E. M., Cotton, A. M., Lam, L. L., Farré, P., Emberly, E., Brown, C. J., et al. (2013). Additional Annotation Enhances Potential for Biologically-Relevant Analysis of the Illumina Infinium HumanMethylation450 BeadChip Array. *Epigenetics & Chromatin* 6 (1), 4. doi:10.1186/1756-8935-6-4
- Radovich, M., Jiang, G., Hancock, B. A., Chitambar, C., Nanda, R., Falkson, C., et al. (2020). Association of Circulating Tumor DNA and Circulating Tumor Cells after Neoadjuvant Chemotherapy with Disease Recurrence in Patients with Triple-Negative Breast Cancer. *JAMA Oncol.* 6, 1410. doi:10.1001/jamaoncol.2020.2295
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Shah, S. P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., et al. (2012). The Clonal and Mutational Evolution Spectrum of Primary Triple-Negative Breast Cancers. *Nature* 486 (7403), 395–399. doi:10.1038/nature10933
- Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer Statistics, 2020. *CA A. Cancer J. Clin.* 70 (1), 7–30. doi:10.3322/caac.21590
- Tarsounas, M., and Sung, P. (2020). The Antitumorigenic Roles of BRCA1-BARD1 in DNA Repair and Replication. *Nat. Rev. Mol. Cel Biol* 21 (5), 284–299. doi:10.1038/s41580-020-0218-z
- Thienpont, B., Steinbacher, J., Zhao, H., D'Anna, F., Kuchnio, A., Ploumaki, A., et al. (2016). Tumour Hypoxia Causes DNA Hypermethylation by Reducing TET Activity. *Nature* 537 (7618), 63–68. doi:10.1038/nature19081
- Tian, Y., Morris, T. J., Webster, A. P., Yang, Z., Beck, S., Feber, A., et al. (2017). ChAMP: Updated Methylation Analysis Pipeline for Illumina BeadChips. *Bioinformatics* 33 (24), 3982–3984. doi:10.1093/bioinformatics/btx513
- Tsai, M., Lo, S., Audeh, W., Qamar, R., Budway, R., Levine, E., et al. (2018). Association of 70-Gene Signature Assay Findings with Physicians' Treatment Guidance for Patients with Early Breast Cancer Classified as Intermediate Risk by the 21-Gene Assay. *JAMA Oncol.* 4 (1), e173470. doi:10.1001/jamaoncol.2017.3470
- Voorwerk, L., Slagter, M., Horlings, H. M., Sikorska, K., van de Vijver, K. K., de Maaker, M., et al. (2019). Immune Induction Strategies in Metastatic Triple-Negative Breast Cancer to Enhance the Sensitivity to PD-1 Blockade: the TONIC Trial. *Nat. Med.* 25 (6), 920–928. doi:10.1038/s41591-019-0432-4
- Widschwendter, M., Jones, A., Jones, A., Evans, I., Reisel, D., Dillner, J., et al. (2018). Epigenome-based Cancer Risk Prediction: Rationale, Opportunities and Challenges. *Nat. Rev. Clin. Oncol.* 15 (5), 292–309. doi:10.1038/nrclinonc.2018.30
- Wienken, M., Dickmanns, A., Nemajerova, A., Kramer, D., Najafova, Z., Weiss, M., et al. (2016). MDM2 Associates with Polycomb Repressor Complex 2 and

Enhances Stemness-Promoting Chromatin Modifications Independent of P53. *Mol. Cell* 61 (1), 68–83. doi:10.1016/j.molcel.2015.12.008

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Zhuang, J., Jones, A., Lee, S.-H., Ng, E., Fiegl, H., Zikan, M., et al. (2012). The Dynamics and Prognostic Potential of DNA Methylation Changes at Stem Cell Gene Loci in Women's Cancer. *Plos Genet.* 8 (2), e1002517. doi:10.1371/journal.pgen.1002517

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Xiong, Yang, Ao, Yi, Zouxu, Zhong, Feng, Huang, Wang and Shuang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.