DNA Methylation Based Molecular Subtypes Predict Prognosis in Breast Cancer Patients

Zeng-Hong Wu, MD[1,2] ⓘ, Yun Tang, MD[3], and Yan Zhou, MD[2]

Abstract

Background: Epigenetic changes are tightly linked to tumorigenesis development and malignant transformation' However, DNA methylation occurs earlier and is constant during tumorigenesis. It plays an important role in controlling gene expression in cancer cells.

Methods: In this study, we determining the prognostic value of molecular subtypes based on DNA methylation status in breast cancer samples obtained from The Cancer Genome Atlas database (TCGA).

Results: Seven clusters and 204 corresponding promoter genes were identified based on consensus clustering using 166 CpG sites that significantly influenced survival outcomes. The overall survival (OS) analysis showed a significant prognostic difference among the 7 groups (p<0.05). Finally, a prognostic model was used to estimate the results of patients on the testing set based on the classification findings of a training dataset DNA methylation subgroups.

Conclusions: The model was found to be important in the identification of novel biomarkers and could be of help to patients with different breast cancer subtypes when predicting prognosis, clinical diagnosis and management.

Introduction

Breast cancer (BC) is a tumor that develops from breast tissue and easily metastasizes to bones and lungs. Globally, its incidence rates are very high.[1] It is the second leading cause of death among women,[2] and accounted for 30% of the almost 880,000 new cancer cases in 2018 in the US.[3] However, the asymptomatic early stage leads to late stage diagnosis that exerts serious consequences. Early breast cancer diagnosis and intervention is particularly important.[4] According to BC genome and transcriptome sequences, it is divided into 5 intrinsic molecular subtypes, including Luminal A, Luminal B, HER-2 enriched, Basal-like and Claudin-low.[5] Unfortunately, despite the advances in locoregional, endocrine, chemo and molecular-targeted therapies, BC patients exhibit a low 5-year survival rate.[6] This is attributed to high cytotoxicity and poor

therapeutic effects in the advanced tumor. Epigenetic changes are tightly linked to tumorigenesis development and malignant transformation.[7] However, DNA methylation occurs early and

[1] Department of Infectious Diseases, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[2] Department of Otorhinolaryngology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[3] Department of Critical Care Medicine, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Corresponding Author:
Yan Zhou, Department of Otorhinolaryngology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China.
Email: wzhs134@163.com

is a constant process during tumorigenesis. DNA methylation regulates gene expression in cancer cells. Transcription start site (TSS) promoter of methylation is a promising and sensitive molecular, pan-cancer biomarker that is detectable in tumor tissues and liquid biopsy samples.[8] Therefore, it is important to establish sensitive and specific novel epigenetic biomarkers for the effective prognostic evaluation of survival outcomes as well as inform therapeutic management strategies of BC.

DNA methylation refers to the reaction in which S-adenosine methionine is catalyzed by DNA methyltransferase as a methyl donor to cytosine into 5-methylcytosine (5-mC). DNA methylation mostly occurs at the C-phosphate-G (CpG) site. Approximately 60% to 90% of mammalian CpGs are methylated.[9] Unmethylated CpGs cluster at the core sequence of the structural gene promoter and TSS to form CpG islands.[10] CpG islands are frequent in the 5′ regulatory region of many genes, particularly the promoter area. It has been found that high 5-methylcytosine levels in the gene promoter region of tumor cells can cause gene transcription silencing, which in turn inhibits downstream gene product expression, resulting in abnormal cell function and increasing the risk of cell cancelation. Chromosomal instabilities have been associated with defective mitotic spindle checkpoints in BC.[11] Methylation of several promoter sequences, including *ESR1, APC, HSD17B4, HIC1, RASSF1A, MAD1L1, MAD2L2, MAD2L1, BUB3, BUB1B, BUB1, CDC20*, and *TTK* has been correlated to BC occurrence and development.[11-13] However, the clinical significance of these methylations in BC prognosis, tumor classification and survival outcomes has not been elucidated. Studies have not systematically assessed the impact of DNA methylation on the overall survival outcomes of BC patients. In this study, we constructed a prognostic prediction model that correlates to BC development and progression. The model combines various DNA methylation biomarkers that are based on high-throughput multiomics dates.

Material and Methods

Data Collection

RNA-seq data and comparative clinical data (1053 cases, data format: BCR XML) were identified and downloaded from the level 3 gene expression information (standardized FPKM) of the TCGA-BC cohort. The collected clinico-pathological data included gender, age, BC Stage, TMN classification, survival status and the number of survival days. Methylation information for the Illumina Infinium Human Methylation 27 and 450 BeadChip arrays of samples obtained from 342 and 890 patients, respectively was downloaded from the UCSC Cancer Browser.

Methylation degree at every site is represented by β-values, ranging from 0 (unmethylated) to 1 (fully methylated). If the obtained CpG sites were lacking data for over 70% of the samples, the samples were removed from the analysis. The k-nearest neighbor (KNN) estimation program was used to estimate the reserving sites for which data were not useful. The

ComBat algorithm in the SVA (R package) was used to eliminate batch effects by combining every DNA methylation array data as well as by merging patient data and batch. Unsteady genomic sites, involving CpG and single nucleotide polymorphisms in the sex chromosome were removed. Since DNA methylation in the promoter region affects gene expression, we examined CpGs in the promoter region. The promoter region is determined as 2 kb upstream to 0.5 kb downstream of the transcription start site. In the subsequent analyses, a total of 21,122 methylation sites were involved. The samples were classified 2 groups: a testing set (27 BeadChip) and a training set (450 BeadChip).

### Establishing the Prognostic Related DNA Methylation Signature

CpG sites that meaningfully Affect survival outcomes were used to standardize categorical features. Univariate COX proportional hazard regression model was built based on the methylation levels of single CpG site, T, M, N, age, gender, stage, and survival information. Significant CpGs obtained from univariate COX regression models and clinical factors were used in the establishment of multivariate COX proportional hazard regression models. CpG sites that were significant in both univariate and multivariate Cox regression analyses were chosen as the typical CpG sites. The prognostic signature was demonstrated as risk score = (Coefficient$CpG1$ × expression of CpG1) + (Coefficientm$CpG2$ × expression of CpG2) + ⋯ + (Coefficient $CpGn$ × expression $CpGn$). A time-dependent receiver operating characteristic (ROC) curve was used to assess the prediction accuracy of the prognostic signatures for BC patients.

Recognition of Molecular Subtypes Associated With Prognosis by Consensus Clustering

Consensus clustering was performed by the "ConsensusClusterPlus" (R software package) to determine BC subgroups from the CpG sites with the most changes. The algorithm begins by subsampling a part of the features and items based on a data matrix. Each subsample is allocated into k groups by kmeans and used to define "consensus" clustering by calculating the steadiness of clustering outcomes by using a specific clustering way to randomize data subsets. The classification count was deemed as the area under the Cumulative Distribution Function (CDF) curve without significant alterations. To thoroughly classify BC categories, more categories were used. Color gradients from 0 (white) to 1 (dark blue) were used as consensus values. The matrices were arranged so that for each classified item, identical clusters were next to each other.

OS and Clinical Factors Analyses

K–M plots were used to determine OS outcomes among BC subgroups based on DNA methylation profiles while the log-
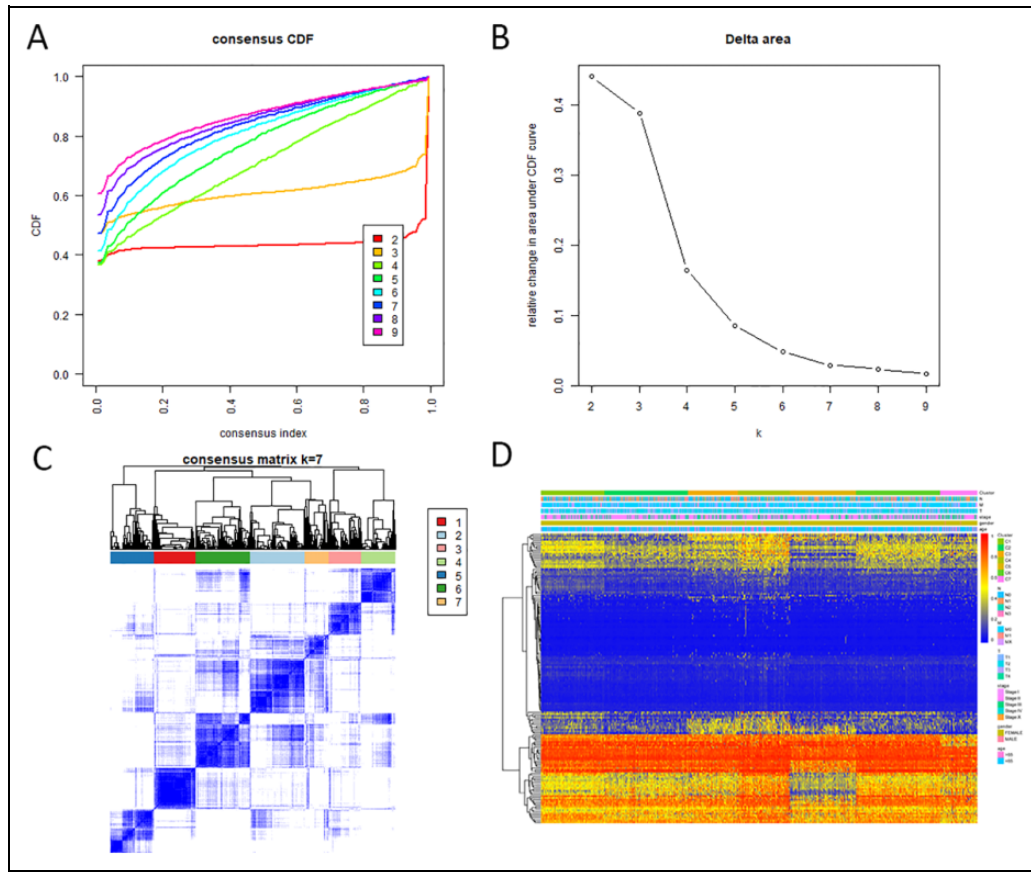
Figure 1. Criteria for selecting number of categories and consensus matrix for DNA methylation classification with the corresponding heat map. A, Consensus among clusters for each category number k. B, Delta area curves for consensus clustering indicating the relative change in area under the CDF curve for each category number k compared to k-1. The horizontal axis represents the category number k and the vertical axis represents the relative change in area under CDF curve. C, Color-coded heatmap corresponding to the consensus matrix for k = 7 obtained by applying consensus clustering. Color gradients represent consensus values from 0-1; white corresponds to 0 and dark blue to 1. D, A heatmap corresponding to the dendrogram in (C) was generated using the heatmap function with DNA methylation classification, TNM stage, clinicopathological stage, and histological type as the annotations.

rank test was used to assess the significance of the differences among the clusters. The Chi-square tests were used to analyze the correlations among clinical characteristics, biological characteristics and DNA methylation clusters. All the tests were double-sided and $p \leq 0.05$ was considered to be statistically significant.

KEGG and GO Enrichment Analyses

The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a resource for exploring high-level gene functions and associating genomic data from large-scale molecular datasets. Gene ontology (GO) function analysis (biological processes (BP), cellular components (CC) and molecular functions (MF)) is a powerful bioinformatics tool for analyzing biological processes and annotating genes. To explore the functions of the identified motifs in DNA and Protein gene groups, biological analyses were performed using GO and KEGG pathways of the R language ggplot2 package for visualization of the figures.

Results

Prognostic Methylation Sites Signatures

Table S1 shows the potential prognostic methylation sites from 21,122 sites linked to OS outcomes among patients in the training dataset. The univariate Cox regression analysis revealed that 2167 CpG sites were potential DNA methylation markers for OS in BC patients (Tables S2). The multivariate Cox regression analysis of the methylation sites using stage, gender, T, M, N, and age as covariates revealed 166 independent prognosis-related CpG sites (Table S3).

Consensus Clustering to Recognize Obvious DNA Methylation Prognosis Subgroups

A consensus cluster of 166 possible prognostic methylation sites were used to determine the different DNA methylation molecular subgroups associated with BC prognosis. The area under the CDF curve and the consensus matrix were used to establish cluster numbers. The CDF curve started to smoothen

after category 6 (Figure 1A and B). To enhance the prognostic value of the BC classifications, we selected more cluster counts. The consensus matrix (Figure 1C) for k = 7 illustrates a 7-block pattern. A heatmap relevant to the dendrogram in Figure 1C with T, M, N, gender, stage, DNA methylation subgroup and age as the annotations is shown in Figure 1D.

The OS analysis showed a significant prognostic difference among the 7 groups ($p \leq 0.05$). Cluster 7 exhibited the worst prognosis while Clusters 5/6 exhibited the best prognosis (Figure 2A). The analyzed intra-cluster proportions of 7 clusters based on T, N, M, stage, gender, and age were as shown in Figure 2B-G, respectively. The association trend between features and specific clusters were as follows; Clusters 5/7 with lower T grade, Clusters 3/6/7 with advanced stage, Clusters 2/6/7 with lower N grade, Clusters 3/7 with higher M grade, Cluster 7 with older age, Cluster 4/7 with more males. These findings suggest that each clinical factor was correlated with different intra-cluster rates.

Features of DNA Methylation Clustering

The genomic annotation of the 166 CpG sites was used to identify 204 corresponding promoter genes (Tables S4). Functional enrichment analysis of the 204 promoter genes was performed using the "clusterProfiler" R software package. GO analysis results showed that: changes in BP were significantly enriched in the regulation of T cells functions, immune response-regulating cells and cell-cell adhesion; changes in MF were enriched in the symporter activity, protein binding, kinase activity while changes in CC were enriched in the membrane region, immunological synapse, membrane raft, membrane microdomain and receptor complex. KEGG pathways were enriched in the signaling pathways regulating stem cell pluripotency, Hematopoietic cell lineage, T cell receptor signaling pathway, Renal cell carcinoma, PPAR signaling pathway, and EGFR tyrosine kinase inhibitor resistance (Figure 3 and Tables S5). The heatmap of expression levels of the methylated genes identified in the subgroup is shown in Figure 4A. Gene expression levels varied among the subgroups, indicating that DNA methylation levels can be used to reflect the expression of these genes. The constructed protein-protein interaction (PPI) network identified 4 hub genes (*LTB4 R, TSFM, SLC24A1* and *FYN*) using the plug-in Molecular Complex Detection (MCODE) of Cytoscape (Figure 4B).

The cluster-specific methylation sites were screened using the methylation sites as cluster features. When 1 cluster was selected as a single group, and the other 6 clusters as a different group, then, $|\log 2FC| > 1$ combined $p \leq 0.05$ were used as a screening criterion to analyze the differences among the 7 clusters (Table S6). Cluster 4 exhibited the highest number of specific sites, with all of them being hypermethylated. The methylation level was highest in all clusters (Figure 2H).

## The BC Prognosis Prediction Model

Cluster 5 was selected as the hub cluster because it contained the highest count of specific methylation sites (11), of which cg24194539 were hypermethylated while the others were hypomethylated. We used Akaike information criterion (AIC) gain the goodness of data fitting but prevents overfitting as much as possible. Thus, the formula for our model was: Risk Score = 1.54*cg01986577+1.01*cg22920417+1.45*cg24194539+2.35*cg27076139. The median risk score was then used as the threshold for dividing the cohort into high-risk and low-risk groups. The prognostic analysis showed significant differences between the 2 sets ($p \leq 0.001$) (Figure 5A). The prognosis of the hypermethylated group was poor, suggesting that these specific methylation sites could be used as prognostic markers. The ROC outcomes from the risk scores of each subject are shown in Figure 5C. The area under the curve (AUC) was 0.757, suggesting that our model operated normally. The samples were then sorted by risk scores to determine if methylation levels systematically changed with the risk score (Figure 5B). As the risk score increased, methylation levels at specific sites significantly increased.

Finally, a prognostic model was used to estimate patient outcomes in the testing set. The testing dataset samples were also divided into low-risk and high-risk groups by the cut-off score. The prognosis was found to be significantly different between the 2 groups (Figure 6A, $p = 0.03$). The AUC was 0.601 in the testing samples, indicating that the model functioned well (Figure 6B). These results were consistent with the outcomes from the training dataset, and prove that our prediction model is stable and accurate.

Discussion

Breast cancer is a global public-health concern issue that is associated with high cancer-related mortality rates. To enhance the clinical management of BC, it is important to establish new clinical biomarkers that can promote the prognostic assessment, molecular subtyping, recurrence prediction, and early intervention or chemotherapy. In addition, molecular based procedures do not require large tissue samples, which increases patient allowance and lowers the need for surgical procedures. Among all the molecular features, DNA methylation at the CpG site plays a vital role in epigenetic functions by inhibiting the activity of DNA fragments as well as gene transcription.

DNA methylation markers can be used for prognosis and predicting therapeutic outcomes. DNA methylation changes occur early during tumorigenesis and are important in regulating gene expression levels in cancer cells. Therefore, epigenetic changes alone, or in combination with other standard biomarkers can be detected during the early diagnosis of BC. Methylated *PTCH* promoter in MCF-7 cells and BC samples was correlated with low *PTCH* expression levels.[14] Furthermore, DNA hypermethylation of *MINT17, MINT31, RARβ2* and *RASSF1A* occurs in the early stages of BC development, and may be predictors of malignancy. Abnormally methylated
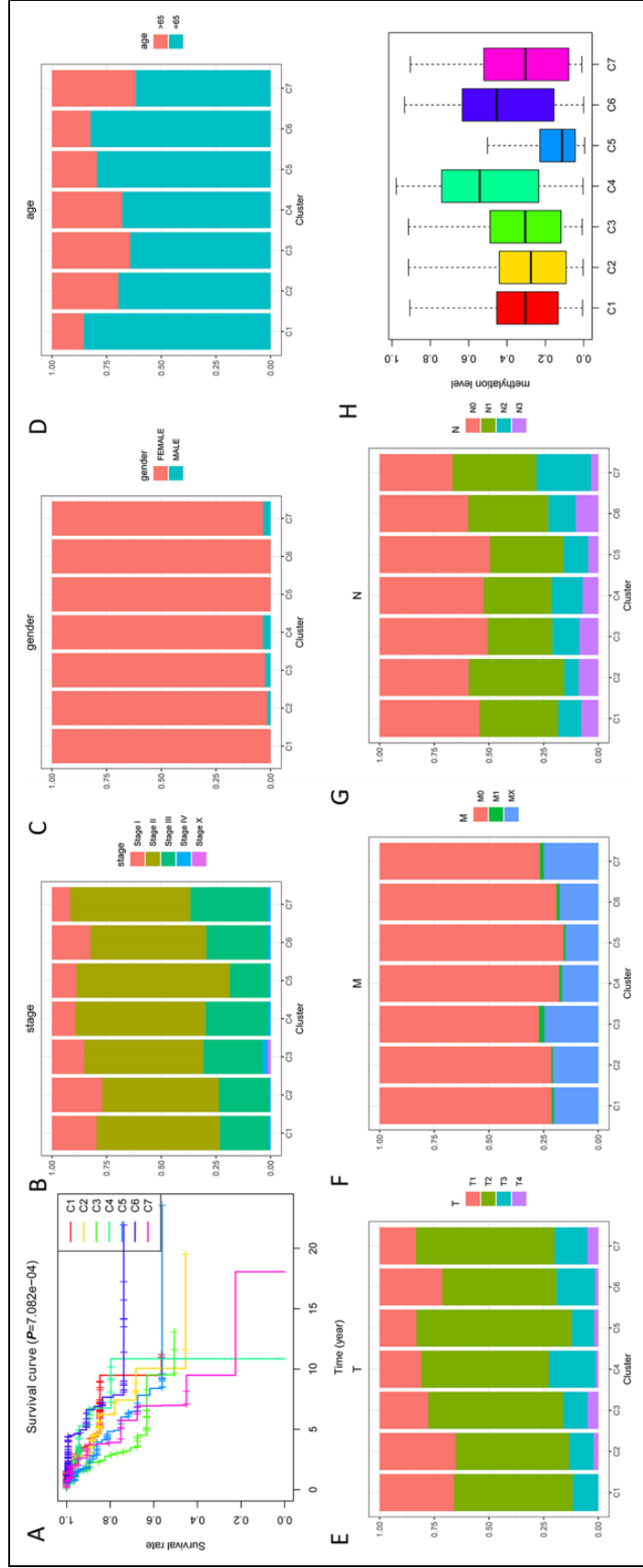
Figure 2. Comparison of prognosis and clinical factors between the DNA methylation clusters. (A) Survival curves for each DNA methylation subtype in the training set. Stage (B), gender (C), age (D), topography score (E), metastasis (F), and lymphocyte infiltration (G) distributions for each DNA methylation subtype in the training set. (H) Box plot of CpG methylation levels of the 7 Clusters.
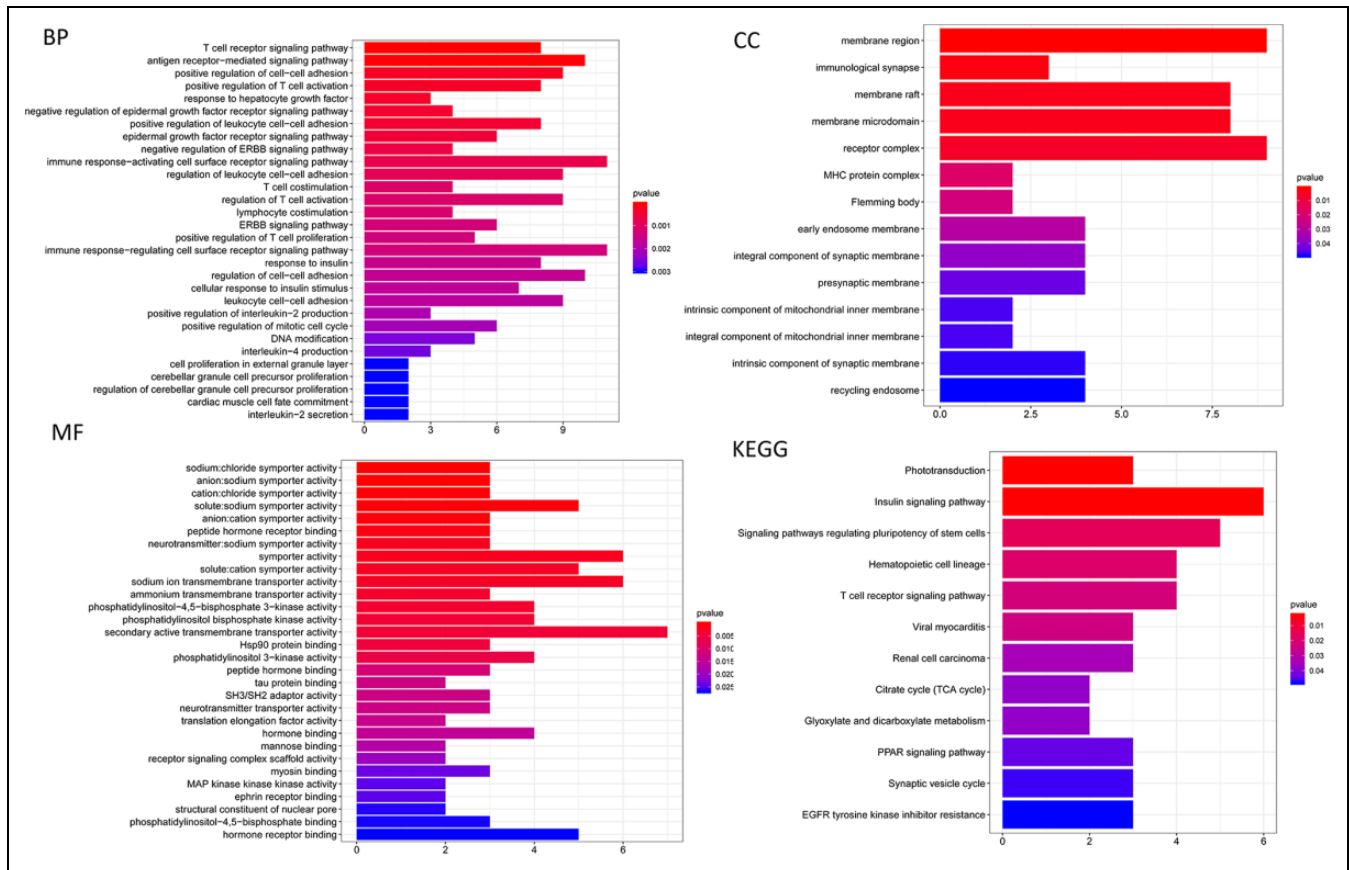
Figure 3. The function of the identified CpG sites corresponding promotor genes using gene ontology (GO) enrichment and KEGG pathway analysis.
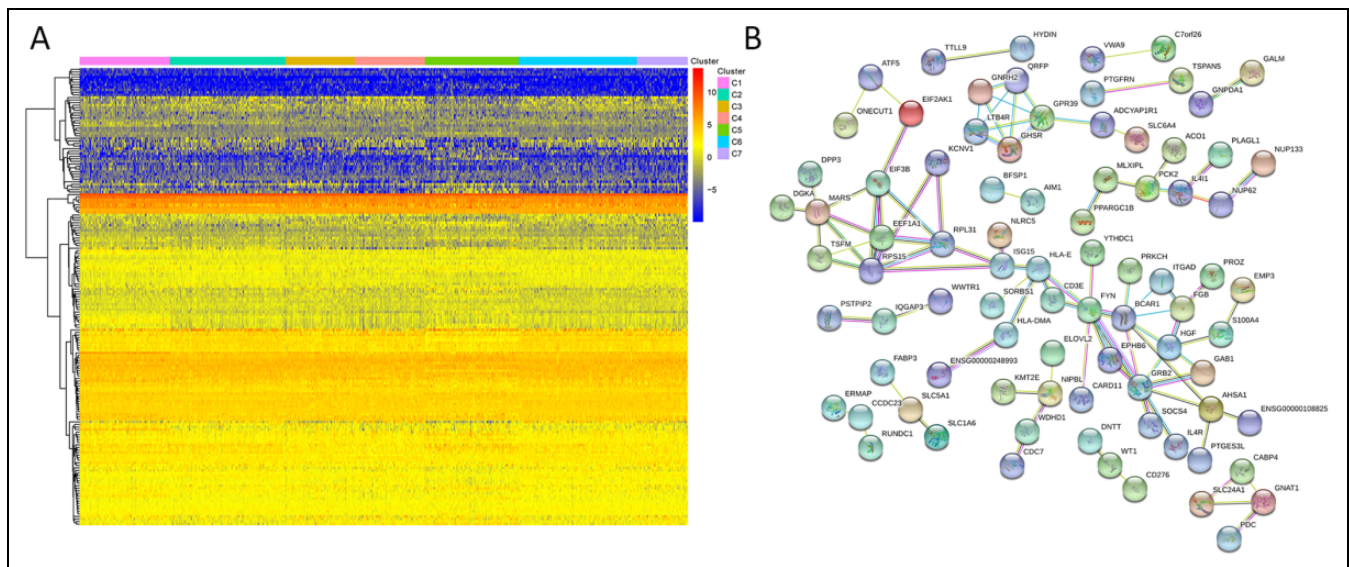

Figure 4. Gene annotations of 204 methylated sites. A, Cluster analysis heat map for annotated genes associated with the 166 CpG site. B, The PPI network for annotated genes associated with the 166 CpG site.

genes can be used as non-invasive biomarkers for the early diagnosis, treatment choice, therapeutic outcomes, and potential applications of new therapies. In this study, we identified

166 independent prognosis-associated CpG sites and a total of 204 corresponding promotor genes. We then constructed the PPI network and identified 4 hub genes (*LTB4 R, TSFM,*
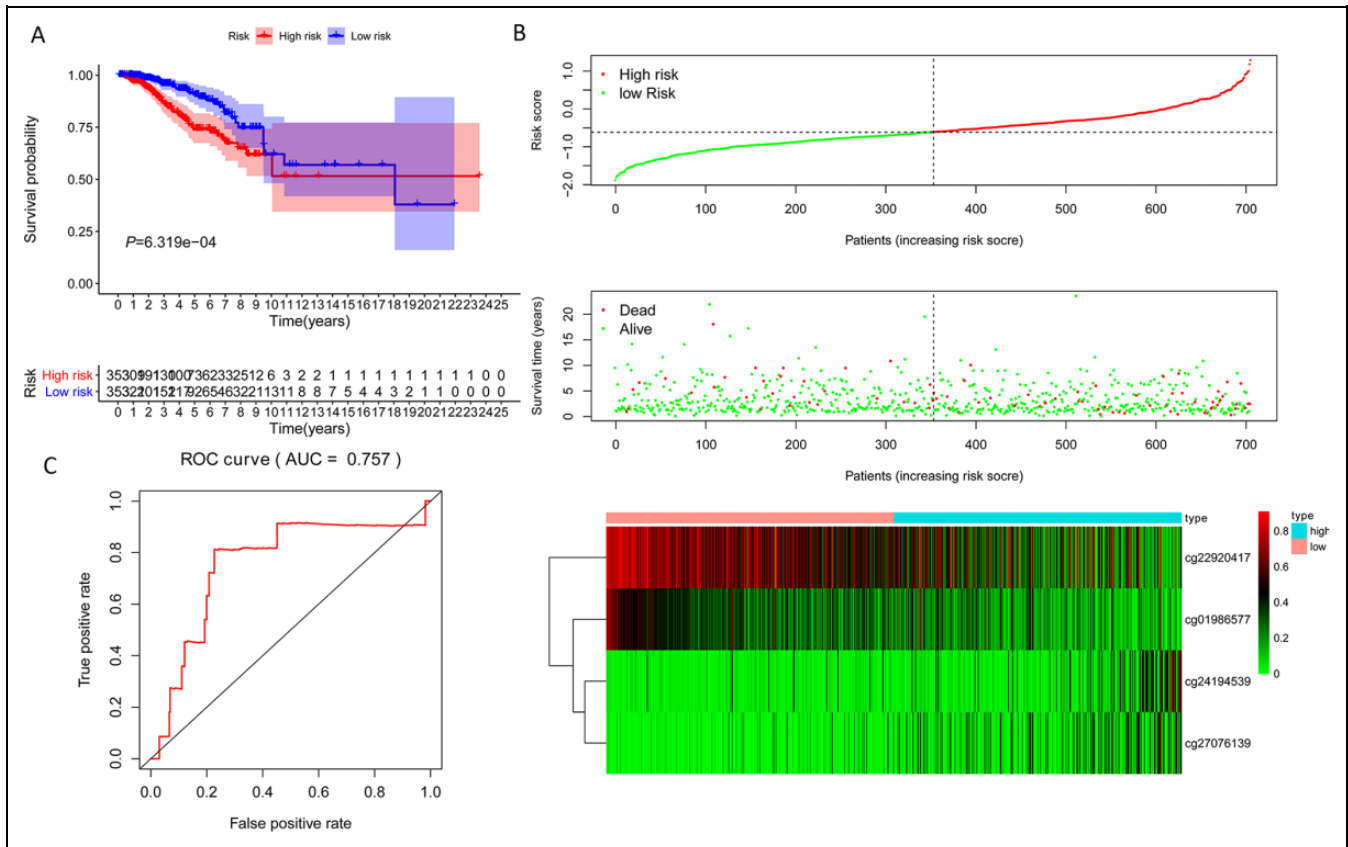
Figure 5. Construction of the prognosis prediction model for training set BC patients. A, Analysis of prognostic differences after classification in the training set. B, The horizontal axis represents the samples, and the vertical axis represents risk scores (top), overall survival (middle), and methylation site (bottom). C, ROC curves of prognostic predictors in BC patients.
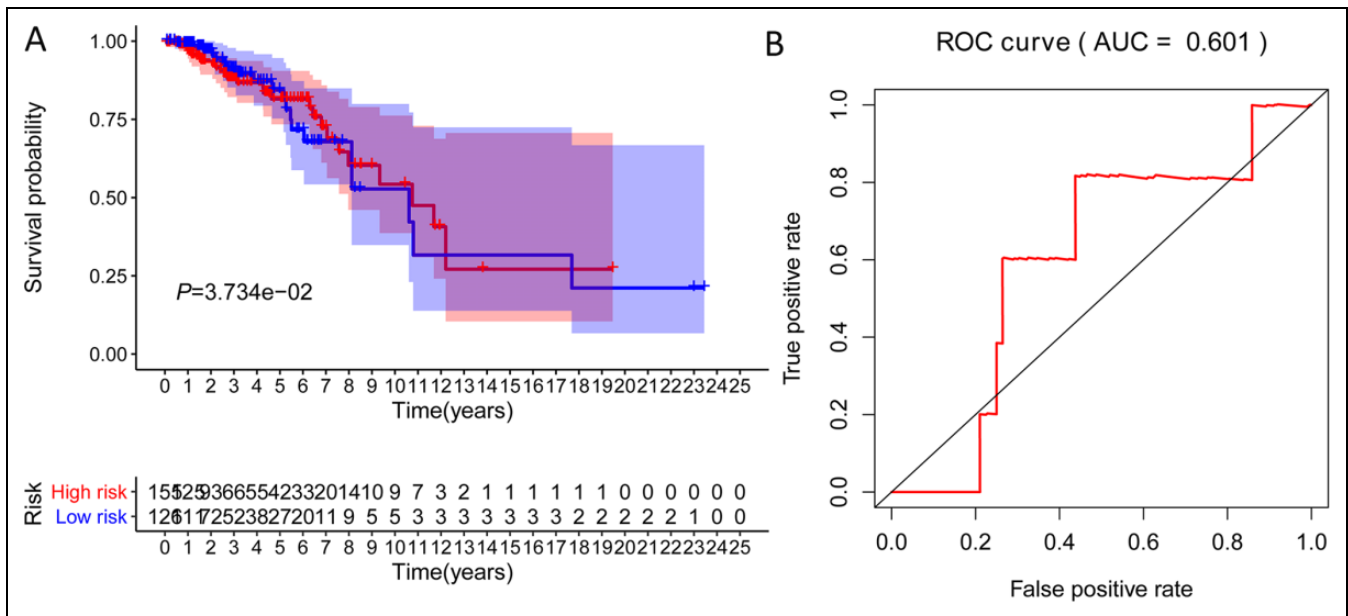
Figure 6. Stability of the prognosis prediction model in testing set BC patients. A, Analysis of prognostic differences after classification in the testing set. B, ROC curves of prognostic predictors in testing set BC patients.

SLC24A1 and *FYN*). Leukotriene B4 receptor (*LTB4 R*), is a lipid mediator that acts as a potent chemoattractant for inflammatory leukocytes and can enhance macrophage and fibroblast accumulation.[15] The *LTB4/BLT1* signaling pathway associated with pSmad3 L has been found to be constitutively activated in BC cells and is correlated with TGF-β1-resistant cell growth.[16] Ts translation elongation factor, mitochondrial (*TSFM*), encodes a mitochondrial translation elongation factor. *TSFM* mutations lead to an infantile fatal multisystem disorder with cardiomyopathy.[17] Mutations in this gene are linked to oxidative phosphorylation deficiency-3 syndrome that suppresses the expression levels of oxidative phosphorylation complexes in BC.[18] Solute carrier family 24 member 1 (*SLC24A1*), encodes a member of the potassium-dependent sodium/calcium exchanger protein family. Defects in *SLC24A1* lead to night blindness.[19] In our study, functional enrichment analysis of 204 genes revealed that changes in MF were mainly enriched in sodium/anion/cation: chloride/ symporter activity. Therefore, we hypothesize that *SLC24A1* regulates DNA methylation during BC development through the sodium/calcium exchanger protein. FYN proto-oncogene, Src family tyrosine kinase (*FYN*), belongs to the protein-tyrosine kinase oncogene family. It encodes a membrane-associated tyrosine kinase that has been implicated in cell growth regulation. The FYN gene is involved in the regulation of epigenetic modifications in mammalian gametes.[20] FYN mutations play an important role in T cell receptor-NF-κB signaling.[21] This study is the first to comprehensively analyze the DNA methylation roles of *LTB4 R, TSFM, SLC24A1* and *FYN* in BC and provides important information to future studies.

We performed functional enrichment analysis of the 204 genes. KEGG pathways were found to be mainly enriched in signaling pathways regulating stem cell pluripotency, Hematopoietic cell lineage, T cell receptor signaling pathway, renal cell carcinoma, PPAR signaling pathway, EGFR tyrosine kinase inhibitor resistance. The majority of breast cancer patients are resistant to EGFR tyrosine-kinase inhibitors such as gefitinib. The AKT and MAPK signaling pathways have been implicated in this resistance.[22] Breast cancer cells can establish a pluripotent program with enhanced stemness activity that is associated with resistance to sex hormone stimulation or deprivation.[23] Hematopoiesis presents a novel opportunity for limiting bone metastasis in breast cancer patients. This can be achieved through the hematopoietic myeloid/osteoclast progenitor cell lineage potential, and biomarkers that stratify responses among patients with high recurrence risks.[24] T cell receptor signaling pathway responsiveness is a key determinant of intratumoral immunosuppressive potential and clinical outcomes.[25] Therefore, we hypothesized that the 204 promotor genes play a critical role in these signaling pathways and should be further studied. Even though methylation may play a critical role in BC, specific methylation sequences in the promoter regions that affect genes have not been established. The clinical significance of these gene methylations that are linked to tumor classification, survival, and prognosis should be studied in a large sample of BC patients. In this study, we developed a classification system that integrates some DNA methylation biomarkers to assess the prognosis of treatment effects and to help in informing treatment choices. This model can be used to detect novel biomarkers, classify molecular subtypes and to precisely establish medical targets for BC. Moreover, the model is also important in the prognostic prediction, diagnosis and development of treatment strategies for patients with distinct BC epigenetic subtypes.

The molecular signatures in our study may show DNA methylation changes and provide potential biomarkers for cancer treatment and prediction of therapeutic outcomes. However, these signatures should be proven to be viable in studies using independent cohorts and predictive DNA methylation functional experiments. The limitations of our study include; first, our outcomes have not been clinically validated, and second, our results do not provide an accurate clinical data due to the relatively small sample size of patients. Therefore, our prognostic prediction model needs improvement.

Conclusion

In conclusion, from the TCGA database and other bioinformatics methods, we established prognostic specific methylation sites and build a prognostic prediction model for BC patients. The model is important in the identification of novel biomarkers and can help patients with different BC subtypes to predict prognosis, clinical diagnosis and management.

Abbreviations

BC   Breast cancer
TCGA   The Cancer Genome Atlas database
OS   overall survival
CpG   C-phosphate-G.

Authors' Note

W.Z.H. and T.Y. designed and analyzed the research study; W.Z.H. and Z.Y. wrote and revised the manuscript, W.Z.H. collected the data and all authors contributed to and approved the final version of manuscript. RNA-seq data and corresponding clinical data were acquired from the data portal for TCGA (https://portal.gdc.cancer.gov/). As the work is a bioinformatics analysis article, ethical approval was not necessary and all the data were retrieved from the free online databases. This article does not contain studies with human or animal subjects. There are no human subjects in this article and informed consent is not applicable.

Declaration of Conflicting Interests

Funding

ORCID iD

Zeng-Hong Wu, MD ⓘ https://orcid.org/0000-0003-4977-6478

Supplemental Material

Supplemental material for this article is available online.

References

1. Urooj T, Wasim B, Mushtaq S, et al. Cancer cell-derived secretory factors in breast cancer-associated lung metastasis: their mechanism and future prospects. *Curr Cancer Drug Targets*. 2020;20(3):168-186.

2. DeSantis CE, Ma J, Goding Sauer A, et al. Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J Clin*. 2017;67(6):439-448.

3. Miller KD, Fidler-Benaoudia M, Keegan TH, et al. Cancer statistics for adolescents and young adults, 2020. *CA Cancer J Clin*. 2020 Sep 17.

4. Yedjou CG, Sims JN, Miele L, et al. Health and racial disparity in breast cancer. *Adv Exp Med Biol*. 2019;1152:31-49.

5. Rainey L, Eriksson M, Trinh T, et al. The impact of alcohol consumption and physical activity on breast cancer: the role of breast cancer risk. *Int J Cancer*. 2020;147(4):931-939.

6. Prat A, Pineda E, Adamo B, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast*. 2015; 24(Suppl 2):S26-35.

7. Zamora A, Alves M, Chollet C, et al. Paclitaxel induces lymphatic endothelial cells autophagy to promote metastasis. *Cell Death Dis*. 2019;10(12):956.

8. Manoochehri M, Wu Y, Giese NA, et al. SST gene hypermethylation acts as a pan-cancer marker for pancreatic ductal adenocarcinoma and multiple other tumors: toward its use for blood-based diagnosis. *Mol Oncol*. 2020;14(6):1252-1267.

9. Ehrlich M, Gama-Sosa MA, Hang LH, et al. Amour and distribution of 5-metylcytoinine in human DNA form different types of tissues of cells. *Nucleic Acids Res*. 1982;10(8):2709-27221.

10. Gu Y, Zhang CW, Wang L, et al. Association analysis between body mass index and genomic DNA methylation across 15 major cancer types. *J Cancer*. 2018;9(14):2532-2542.

11. Yuan B, Xu Y, Woo J H, et al. Increased expression of mitotic checkpoint genes in breast cancer cells with chromosomal instability. *Clinical Cancer Res*. 2006, 12(2):405-410.

12. Widschwendter M, Siegmund KD, Müller HM, et al. Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. *Cancer Res*. 2004, 64(11):3807.

13. Fiegl H, Mueller HM, Widschwendter A, et al. P13 DNA methylation in serum of breast cancer patients: an independent prognostic marker. *EJC Supplements*. 2004;2(1):44-44.

14. Wolf I, Bose S, Desmond J C, et al. Unmasking of epigenetically silenced genes reveals DNA promoter methylation and reduced expression of PTCH in breast cancer. *Breast Cancer Res Tr*. 2007; 105(2):139-155.

15. Kamata M, Amano H, Ito Y, et al. Role of the high-affinity leukotriene B4 receptor signaling in fibrosis after unilateral ureteral obstruction in mice. *PLoS One*. 2019;14(2): e0202842.

16. Jeon W K, Choi J, Park S J, et al. The proinflammatory LTB4/ BLT1 signal axis confers resistance to TGF-β1-induced growth inhibition by targeting Smad3 linker. *Oncotarget*. 2015;6(39): 41650-41666.

17. Ahola S, Isohanni P, Euro L, et al. Mitochondrial EFTs defects in juvenile-onset leigh disease, ataxia, neuropathy, and optic atrophy. *Neurology*. 2014;83(8):743-751.

18. Putignani L, Raffa S, Pescosolido R, et al. Alteration of expression levels of the oxidative phosphorylation system (OXPHOS) in breast cancer cell mitochondria. *Breast Cancer Res Tr*. 2008; 110(3):439-452.

19. Neuillé M, Malaichamy S, Vadalà M, et al. Next-generation sequencing confirms the implication of SLC24A1 in autosomal-recessive congenital stationary night blindness. *Clin Genet Int J Genet Med*. 2016;89(6):690-699.

20. Ou XH, Zhu CC, Sun SC. Effects of obesity and diabetes on the epigenetic modification of mammalian gametes. *J Cell Physiol*. 2019;234(6):7847-7855.

21. Kataoka K, Nagata Y, Kitanaka A, et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. *Nat Genet*. 2015; 47(11):1304-1315.

22. Normanno N, Campiglio M, Maiello MR, et al. Breast cancer cells with acquired resistance to the EGFR tyrosine kinase inhibitor gefitinib show persistent activation of MAPK signaling. *Breast Cancer Res Tr*. 2008, 112(1):25-33.

23. Nasr M, Farghaly M, Elsaba T, et al. Resistance of primary breast cancer cells with enhanced pluripotency and stem cell activity to sex hormonal stimulation and suppression. *Int J Biochem Cell Biol*. 2018;105:84-93.

24. Ubellacker JM, Baryawno N, Severe N, et al. Modulating bone marrow hematopoietic lineage potential to prevent bone metastasis in breast cancer. *Cancer Res*. 2018;78(18): 5300-5314.

25. Wang L, Simons DL, Lu X, et al. Connecting blood and intratumoral Treg cell activity in predicting future relapse in breast cancer. *Nat Immunol*. 2019;20(9):1220-1230.