

Research Article

Chromatin Regulator-Related Gene Signature for Predicting Prognosis and Immunotherapy Efficacy in Breast Cancer

Dongxu Feng,¹ Wenbing Li,¹ Wei Wu,¹ Ulf Dietrich Kahlert,² Pingfa Gao,¹ Gangfeng Hu,¹ Xia Huang,¹ Wenjie Shi ^{2,3} and Huichao Li ⁴

¹Department of General Surgery, Chongming Hospital Affiliated to Shanghai University of Medicine and Health Sciences, Chongming District, Shanghai 202150, China

²University Clinic for General, Visceral, Vascular-and Transplantation Surgery, Faculty of Medicine, Otto-von-Guericke-University, Magdeburg 39120, Germany

³University Hospital for Gynaecology, Pius-Hospital, University Medicine Oldenburg, Oldenburg 26121, Germany

⁴Department of Thyroid Surgery, The Affiliated Hospital of Qingdao University, Qingdao 266555, Shandong, China

Correspondence should be addressed to Wenjie Shi; wenjie.shi@uni-oldenburg.de and Huichao Li; lihuichao@qdu.edu.cn

Received 3 September 2022; Revised 22 October 2022; Accepted 24 November 2022; Published 30 January 2023

Academic Editor: Feng Jiang

Copyright © 2023 Dongxu Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Many studies have found that chromatin regulators (CRs) are correlated with tumorigenesis and disease prognosis. Here, we attempted to build a new CR-related gene model to predict breast cancer (BC) survival status. **Methods.** First, the CR-related differentially expressed genes (DEGs) were screened in normal and tumor breast tissues, and the potential mechanism of CR-related DEGs was determined by function analysis. Based on the prognostic DEGs, the Cox regression model was applied to build a signature for BC. Then, survival and receiver operating characteristic (ROC) curves were performed to validate the signature's efficacy and identify its independent prognostic value. The CIBERSORT and tumor immune dysfunction and exclusion (TIDE) algorithms were used to assess the immune cells infiltration and immunotherapy efficacy for this signature, respectively. Additionally, a novel nomogram was also built for clinical decisions. **Results.** We identified 98 CR-related DEGs in breast tissues and constructed a novel 6 CR-related gene signature (ARID5A, ASCL1, IKZF3, KDM4B, PRDM11, and TFF1) to predict the outcome of BC patients. The prognostic value of this CR-related gene signature was validated with outstanding predictive performance. The TIDE analysis revealed that the high-risk group patients had a better response to immune checkpoint blockade (ICB) therapy. **Conclusion.** A new CR-related gene signature was built, and this signature could provide the independent predictive capability of prognosis and immunotherapy efficacy for BC patients.

1. Introduction

Breast cancer (BC) is a common cancer in the world [1]. Although the widespread use of adjuvant chemotherapy and hormonal drugs has reduced mortality in BC patients, there are still individual differences in treatment outcomes and different clinical benefits for BC patients [2]. Fortunately, with the continuous updating of new therapies, the use of cancer biomarkers has become an aid in BC diagnosis, prognosis, treatment response prediction, and disease monitoring during and after treatment [3]. Nowadays, many researchers have tried to use various bioinformatics techniques to identify the

biomarkers or build the risk model in BC and achieved good research results, such as the machine learning model in predicting the immune subtype [4] and the eight-lncRNA prognostic model [5]. Therefore, in order to provide a more effective prediction of survival in tumor patients, continuous efforts are needed to identify new prognostic key molecules and explore their prognostic values.

Chromatin regulators (CRs) are essential upstream regulatory factors in epigenetics that can cooperate to connect the organizational scales of chromatin from nucleosome assembly to the establishment of functional chromatin domains [6]. CRs can highly regulate chromatin structure by

four broad classes of nuclear factors, including histone variants, histone chaperones, chromatin remodelers, and histone modifiers [7]. In human cancer pathogenesis, it has been found that the function of CRs is frequently disrupted by genetic mutations and epigenetic alterations, resulting in perturbed gene expression profiles. The role of CRs in cancer is complex and highly specific [8]. To begin with, the carcinogenic effects of chromatin regulators are well established. Meanwhile, recent new studies have demonstrated their tumor suppressive properties in the regulation of multiple cellular processes [9]. For example, MLL3/4, one of the chromatin regulators, may exert broad tumor suppressor effects in various cancers [10]. In addition, growing evidence has shown that CRs and tumor prognosis are closely related. In cervical cancer, 57 overexpressed chromatin regulators were identified to have prognostic significance [11]. Similarly, CR-related genes were found to be significantly associated with postoperative outcomes in astrocytomas [12]. However, CRs affecting prognosis of breast cancer is still lacking more understanding and research.

The tumor microenvironment (TME), consisting of extracellular matrix, stromal cells, and immune cells [13], is important in cancer initiation, progression, and therapy [14]. With the emergence of new technologies, immunotherapy has become a widely concerned research direction for cancer treatment, which aims to use the immune system as a tool for treating oncological diseases [15]. The effect of immunotherapy is associated with the TME, especially the tumor-infiltrating immune cells (TICs) [16]. Currently, there are 3 types of immunotherapies that target tumor-specific T cells, including immune checkpoint blockade (ICB), adoptive cellular therapies, and cancer vaccines [17]. In immune checkpoint blockade therapy, multiple cancers have been treated with two different antiprogrammed cell death protein 1 (PD1) drugs (nivolumab and pembrolizumab) [18] and anticytotoxic T lymphocyte-associated antigen 4 (CTLA4) (ipilimumab) [19]. The use of therapeutic antibodies does provide clinical benefit to a small patient population, but adverse effects associated with immune checkpoint blockade complicate immunotherapy and limits its use in cancer patients [20]. Therefore, in order to promote the efficacy of ICB therapy, it is necessary to identify and explore the predictive biomarkers for immune checkpoint-blocking therapies [21, 22].

In this work, we investigated the potentiality of CR-related DEGs as prognostic markers in BC through several bioinformatic analyses. Furthermore, CR-related DEGs were successfully employed to construct a six-gene prognostic model and a new CR-related gene signature for predicting patients' outcome and immunotherapy efficacy.

2. Materials and Methods

2.1. Data Collection. The training set was downloaded from the The Cancer Genome Atlas (TCGA) database ($n = 1167$; 113 normal samples vs. 1054 tumor samples). The validate set was obtained from GSE20685 ($n = 327$) [23]. Considering that a patient's follow-up for less than 30 days is too short to evaluate the prognosis, we excluded those patients for

a more accurate evaluation. The complete clinical information of patients is shown in Supplementary Table 1. Lastly, there were 870 chromatin regulator (CR)-related genes obtained from previous research [6].

2.2. Identification of CR-Related DEGs. First, the data of the training set were normalized with the log₂ transformation. Next, the expression profiles of 853 CRs were extracted from the normalized TCGA-BRCA matrix based on the obtained CR names. In the training set, DEGs related to CRs were performed by the "limma" package. Differential expression was set as the adjusted p value less than 0.05, and the absolute value of log₂FC (fold change) was greater than 1. The result of the analysis was visualized in a volcano map utilizing the "ggplot2" R package.

2.3. Biology Function Enrichment. ClusterProfiler was conducted to analyze gene oncology and the KEGG pathway, and this procedure aims at discovering the potential mechanism of the CR-related DEGs. In addition, the enrichment analysis results were visualized via circos diagrams.

2.4. Signature Model Construction and Validation. Batch Cox regression was applied to screen the prognosis-related DEGs of CRs in the training set. And the important candidate genes were analyzed by the multivariate Cox analysis for identifying independent prognosis risk factors. Based on the multivariate regression results, we then developed an optimal signature to predict prognosis based on the coefficients (Coef). The model was constructed as the risk formula: risk score = A gene \times Coef + B gene \times Coef + . . . + X gene \times Coef. The forest map was plotted by the "forestplot" R package. The area under curve of ROC will be used to evaluate the model in the training and validation set, respectively.

2.5. Signature Prognostic Value Evaluation. Other clinical variables also affect patient clinical prognosis, and when compared with risk score, if risk score is also associated with patients outcomes, here, we use Cox regression model to demonstrate it, and we also conduct subgroup analyses of risk score and survival status.

2.6. Immune Cell Infiltration and Immunotherapy Efficacy Estimation. To evaluate the association between this signature and tumor-infiltrating immune cells, CIBERSORT was used to analyze the proportion of various TICs between high-score and low-score patients in the TCGA training set. We also use the tumor immune dysfunction and exclusion (TIDE) indicator to estimate ICB therapy response in BC patients. Among them, lower TIDE scores mean that tumor cells have less chance with immune escape, indicating a higher response rate to ICBs therapy. Finally, we assessed the correlation between this signature and immunotherapeutic markers and ICB-related genes via Wilcoxon test.

2.7. Predictive Nomogram Establishment and Evaluation. A nomogram, a visual scoring system, was established in the training set by integrating the risk score of CR-related gene signature and clinicopathological variables to be used to evaluate the OS of the BC patients. At the same time, the bootstrap method was used to calculate the concordance index (C-index), ROC, and calibration curves. Besides, we further utilized the Kaplan–Meier analysis for OS, DSS, and PFS to prove the clinical prognostic value of the nomogram. All the analyses and plots were done by the R software and suitable R packages.

3. Results

3.1. CR-Related DEGs and Function Enrichment. In the training dataset, a total of 98 CR-related DEGs (Figure 1(a)) (Supplementary Table 2) were screened from 1154 BC tissue samples and 113 normal tissue samples. We predicted the biological mechanisms of 98 CR-related DEGs. GO enrichment analysis showed that the most highly enriched molecular functions (MF) of these DEGs were histone binding, transcription corepressor activity, and histone kinase activity. Cellular components (CC) of these DEGs were covalent chromatin modification, histone modification, and DNA replication. The biological process (BP) of these DEGs was remarkably involved in covalent chromatin modification, histone modification, and DNA replication. Moreover, they are also enriched in the cell cycle, homologous recombination, and lysine degradation by KEGG enrichment analysis (Figure 1(b)).

3.2. Prognostic Signature Construction Based on the CR-Related Gene. We used 98 CR-related DEGs to build the prognostic signature in the TCGA training set. First, ten important CR-related genes associated with OS in BC patients were selected (Figure 1(c)). Then, six hub CR-related genes (ARID5A, ASCL1, IKZF3, KDM4B, PRDM11, and TFF1) were identified as associated with the OS and contributed to the risk signature (Figure 1(d)) (Table 1). The risk score = $(-0.201073277) * ARID5A + 0.1231206 * ASCL1 + (-0.268713644) * IKZF3 + (-0.197882736) * KDM4B + (-0.538035757) * PRDM11 + (-0.050431061) * TFF1$.

3.3. Evaluation and Validation of Signature Efficacy. The high-risk group was defined as risk score more than the median value, others were defined as low-risk groups, respectively, in the training dataset. Then, the risk score distribution between the two groups was compared (Figure 2(a)). Patients with high risk have a poor outcome while the low-risk patients always suggest a better prognosis ($p = 1.179e - 04$) (Figure 2(b)). Moreover, the area under the ROC curve (AUC) was 0.713, which proved the predictive efficacy of our signature for predicting the OS in the training set (Figure 2(c)), which was also validated in the GEO. The risk score distribution in the two groups is shown in Figure 2(d). In addition, patients with low risk also have a better prognosis ($p = 7.955e - 04$) (Figure 2(e)), and the

results further supported the above conclusion: AUC is equal to 0.821 (Figure 2(f)).

3.4. Applicability of the CR-Related Gene Signature. We integrated the signature with other clinical risk factors in the total data set to further assess the independent prognostic value of the risk model for BC. As shown in the TCGA training dataset, age, TNM stage, ER status, and risk score were associated with the OS ($p < 0.05$), and risk score was an independent prognostic factor ($p < 0.05$) (Figures 3(a) and 3(b)). Additionally, in the GEO validation dataset, analysis results also support the abovementioned conclusions (Figures 3(c) and 3(d)).

3.5. Subgroup Analysis to Evaluate Gene Signature. We performed survival subgroup analysis in the TCGA training set to demonstrate that the signature is associated with clinical features. First of all, the clinical patients were classified into two groups, such as age ≤ 60 vs. >60 groups, the T1/2 stage vs. T3/4 stage, the N (-) stage vs. N (+) stage (N0 and N1–N3, respectively), and the M0 stage vs. M1 stage. Subgroup analysis suggests that subgroup patients also have survival rate differences in age, T stage, N1–N3 stage, and M0 stage between high- and low-risk groups. However, patients with the N0 or M1 stage has no significance ($p = 0.080$ and $p = 0.486$, respectively) (Figures 4(a)–4(h)).

3.6. Immune Cell Infiltration and Immunotherapy Efficacy Estimation for the Signature. As for the relationship between the signature and TME, 19 of the 22 TICs were significantly related to the risk score (Figure 5(a)). Here, the CD8⁺ T cells, resting memory CD4⁺ T cells, regulatory T cells (Tregs), gamma delta T cells, follicular helper T cells, M1 macrophages, memory B cells, naive B cells, activated NK cells, monocytes, resting myeloid dendritic cells, and activated mast cells were negatively correlated with the risk score, while the M0 macrophages, M2 macrophages, resting NK cells, activated myeloid dendritic cells, resting mast cells, eosinophils, and neutrophils were positively correlated with the risk score. Meanwhile, the proportion of 12 TICs did differ significantly between the two groups ($p < 0.05$, Figure 5(b)). The proportion of M0 and M2 macrophages was significantly higher in the high-risk group, while the proportion of plasma cells, naive B cells, CD8⁺ T cells, gamma delta T cells, resting memory CD4⁺ T cells, resting NK cells, and resting mast cells was significantly higher in the low-risk group.

Regarding the CR-related signature's potential for predicting response to immunotherapy in BC patients, the result showed that the low-risk group had a higher TIDE score ($p < 0.001$, Figure 5(c)), meaning a poor response to immunotherapy. Likewise, 13 immune checkpoint molecules (CTLA4, PDL1, PD1, PDL2, LAG3, TIM3, CD86, BTLA, ICOS, CD96, CD160, TIGIT, and IDO1) are also positive with low-risk group gene expression ($p < 0.05$, Figure 5(d)). Taken together, these results illustrated the importance of the risk score in breast cancer immunotherapy prediction.

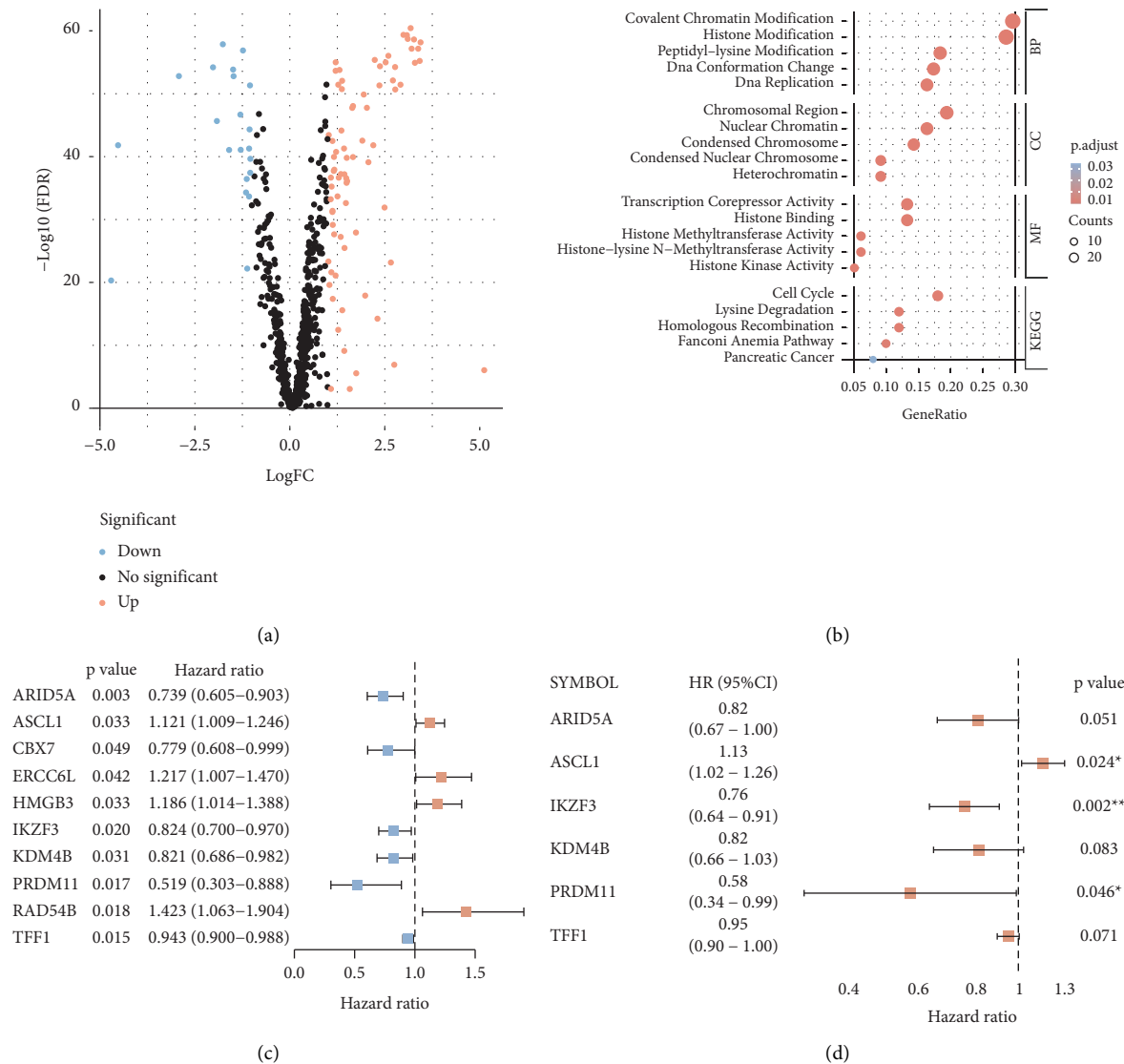


FIGURE 1: Differential expression analysis and the construction of CR-related prognosis signature in the TCGA training set. Volcano map of the CR-related DEGs. Red represents the up-regulated genes; blue represents the down-regulated genes ($|\log_{2}FC| > 1$, FDR q value < 0.05) (a). GO enrichment analysis (consisting of BP, CC and MF) and KEGG pathway enrichment analysis for the CR-related DEGs ($p < 0.05$, FDR q value < 0.05) (b). Univariate Cox regression analysis selected 10 CR-related genes correlated with OS (c). Multivariate Cox regression analysis identified a 6-gene (ARID5A, ASCL1, IKZF3, KDM4B, PRDM11, and TFF1) prognostic signature (d). * $p < 0.05$ and ** $p < 0.01$.

TABLE 1: Six CR-related prognostic genes significantly associated with OS in breast cancer patients.

Gene	Multivariate cox regression analysis				
	Coef	HR	HR 0.95L	HR 0.95 H	p value
ARID5A	-0.201073277	0.8178525	0.668022015	1.001288425	0.051475659
ASCL1	0.1231206	1.131020814	1.016273871	1.258723773	0.02408809
IKZF3	-0.268713644	0.764362104	0.642458524	0.909396333	0.002434433
KDM4B	-0.197882736	0.820466059	0.655829072	1.026432927	0.083334851
PRDM11	-0.538035757	0.583894036	0.344413023	0.989893597	0.045751465
TFF1	-0.050431061	0.950819475	0.900181977	1.004305459	0.070902358

CRs, chromatin regulators; OS, overall survival; Coef, β coefficient; HR, hazard ratio.

3.7. Nomogram for Clinical Decision. A visualized nomogram was built and was used to predict OS probability at three periods (1, 3, and 5 years) (Figure 6(a)). The

concordance C-index was 0.734, which illustrated a good ability in predicting OS for BC patients. Furthermore, we found that the calibration curves indicated that the predicted

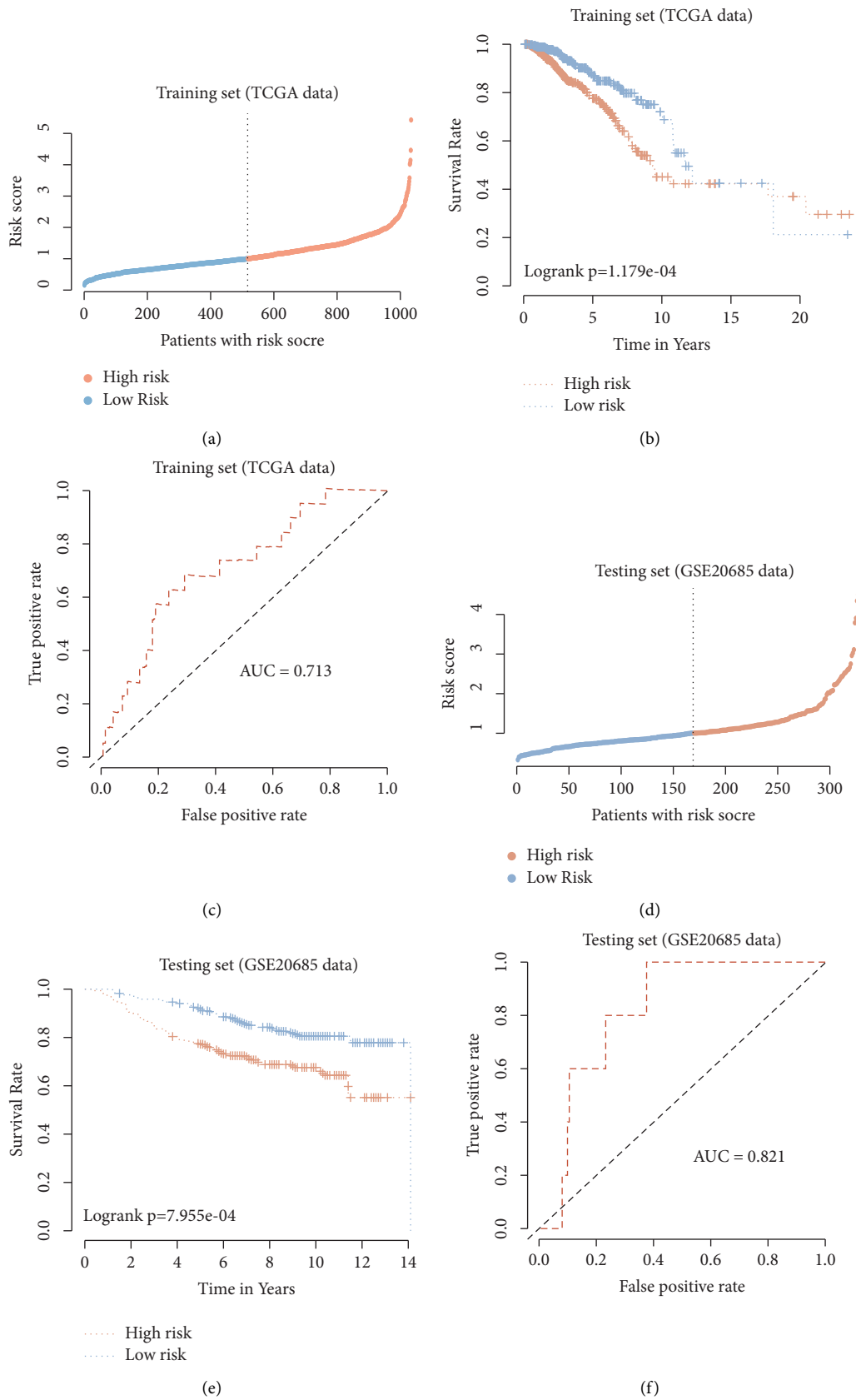


FIGURE 2: Validation of the CR-related prognostic signature in BC patients. Risk score distribution for patients in low- and high-risk groups from the TCGA training set (a). Kaplan–Meier survival analysis of OS between the low- and high-risk groups from the TCGA training set (b). AUC in ROC analysis for risk scores predicting the OS from the TCGA training set (c). Risk score distribution for BC patients in low- and high-risk groups from the GEO validation set (d). Kaplan–Meier survival analysis of OS between the low- and high-risk groups from the GEO validation set (e). AUC in ROC analysis for risk scores predicting the OS from the GEO validation set (f).

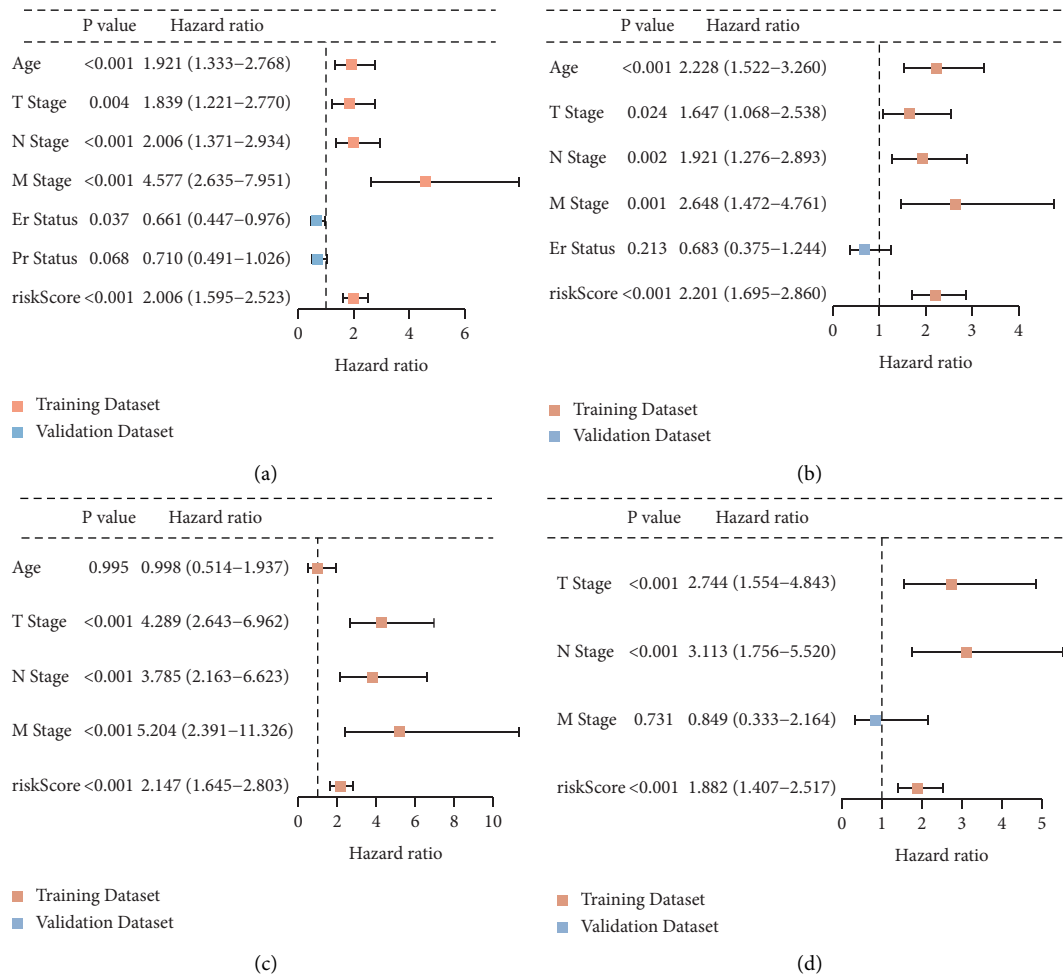


FIGURE 3: Independent prognostic analysis of the clinical traits and risk score in BC patients. Univariate and multivariate Cox regression analyses of the OS in the TCGA training set (a) and (b). Univariate and multivariate cox regression analyses of the OS in the GEO validation set (c) and (d).

curve is close to the ideal curve (Figure 6(b)). At the same time, the ROC curves of this nomogram also showed a good accuracy to individually predict OS for BC patients (AUC = 0.788, 0.731, and 0.713, respectively) (Figures 6(c)–6(e)). Lastly, we also assessed the prognostic value of the nomogram, finding that it was remarkably associated with OS, DSS, and PFS ($p < 0.05$) (Figure 6(f)).

4. Discussion

A significantly poor prognosis for the stage IV female BC patients who were diagnosed between 2007 and 2013 can be observed [24]. To improve the survival rate of BC patients, an increasing number of genetic signatures have been established. For example, Shi et al. [25] developed a five-mRNA signature model based on the ceRNA network for predicting the survival of BC. However, few studies have focused on prognostic signatures based on key genes in CRs that have been shown to have potential as prognostic markers. What's more, studies have shown that epigenetic factors and mechanisms can be involved in regulating the TME and the ICBs response. In lung adenocarcinomas,

ASF1A deficiency could sensitize lung adenocarcinomas to anti-PD-1 therapy by inducing immunogenic M1-like macrophage differentiation and enhancing T cell activation of the TME [26]. Then, in checkpoint-blocked refractory mouse melanoma, histone demethylase LSD1 depletion enhanced tumor immunogenicity and T-cell infiltration in poorly immunogenic tumors and elicited a significant response to anti-PD-1 treatment [27]. Thus, exploring and evaluating the CR-related gene expression in BC patients is important. In our study, we constructed a CR-related gene signature, a useful tool to predict patients' outcomes and immunotherapy sensitivity.

The CR-related gene signature has 6 hub genes; ASCL1, as one of the 6 genes, was positively associated with outcome, while the levels of ARID5A, IKZF3, KDM4B, PRDM11, and TFF1 were negatively associated with survival. As the results demonstrated, the model could predict the patient's outcome.

The ROC curves also confirmed the favorable predictive performance of this signature. Besides, the independent prognostic analysis determined that this signature, age, and TNM stage were independent predictors for BC prognosis.

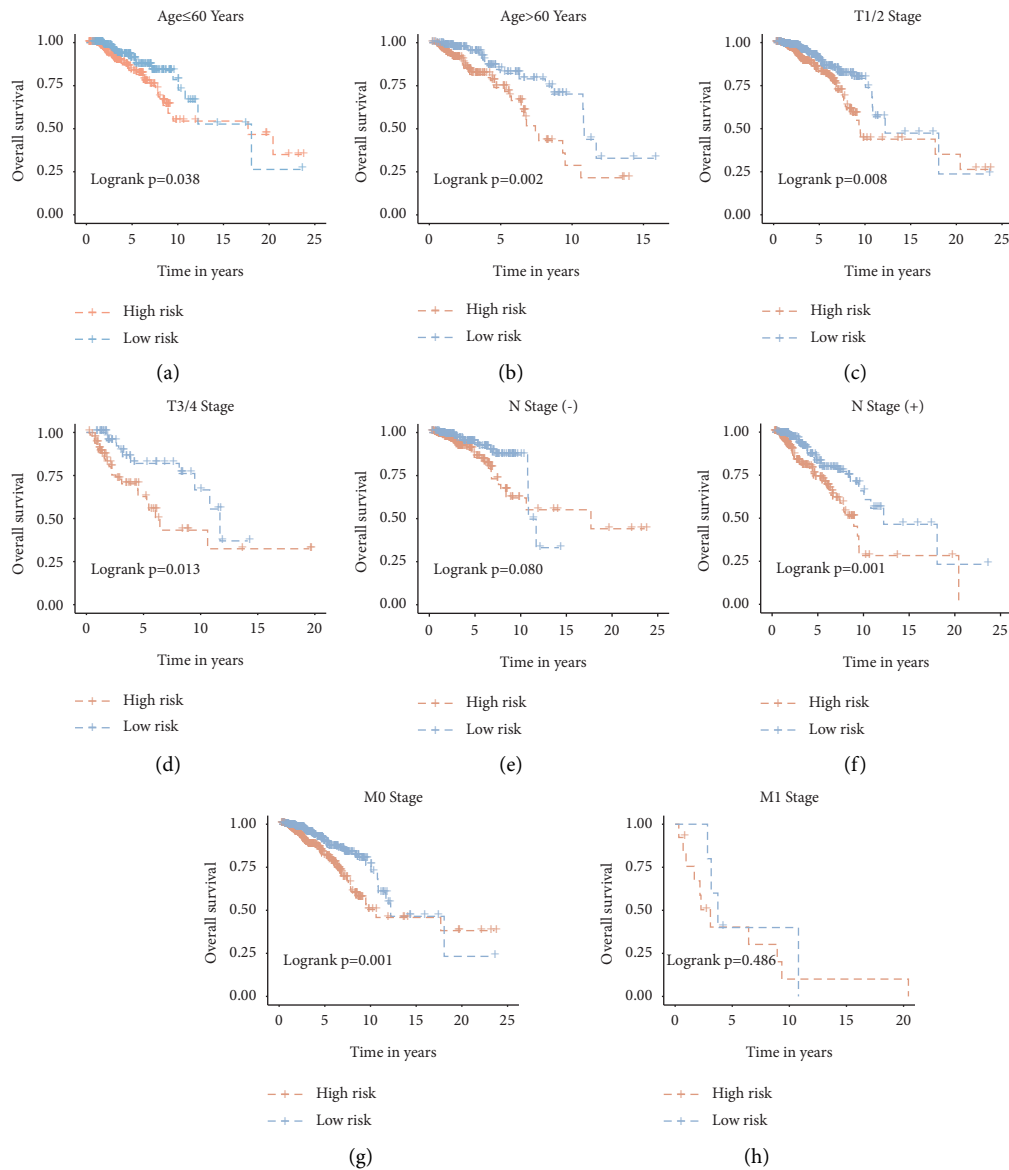


FIGURE 4: Kaplan–Meier survival subgroup analysis in BC patients from the TCGA training set based on the 6-gene signature stratified by clinical characteristics. The OS differences between low- and high-risk group stratified by age (≤ 60 years and > 60 years, respectively) (a) and (b). The OS differences between low- and high-risk group stratified by T stage (T1-2 and T3-4, respectively) (c) and (d). The OS differences between low- and high-risk group stratified by N ($-$) and N ($+$) stage. (N0 and N1–3, respectively) (e) and (f). The OS differences between low- and high-risk group stratified by M stage (M0 and M1, respectively) (g) and (h).

As expected, survival subgroup analysis also suggests the effective prediction ability of the signature in subgroups. Next, we found that patients with high-risk conditions will obtain a better response to immunotherapy and were more suitable for ICB therapy. Moreover, a nomogram model consisting of clinical factors and signature was established, which showed good power and accuracy with a high AUC value and C-index in BC patients. Therefore, the CR-related gene signature was a reliable model to predict prognosis and immunotherapy efficacy, which might have potential implications in clinical practice for BC.

All six genes involved in the signature model are associated with chromatin regulation or tumorigenesis.

Achaete-scute complex homolog 1 (ASCL1) is a key regulator of neuroendocrine differentiation [28]. ASCL1 was highly expressed in classic small cell lung cancer (SCLC); additionally, it was a key driver of tumorigenesis in classic SCLC and correlated with the survival and development of lung cancers with neuroendocrine (NE) features [29]. Jiang et al. [30] confirmed that when ASCL1 was successfully overexpressed with SV40 large T-antigen, it could synergistically inhibit retinoblastoma protein and p53 to promote the development of aggressive adenocarcinoma with NE characteristics; however, in the developing mouse lung, knockout of ASCL1 resulted in specific ablation of lung NE cells. Moreover, in recent studies, ASCL1 demonstrated that

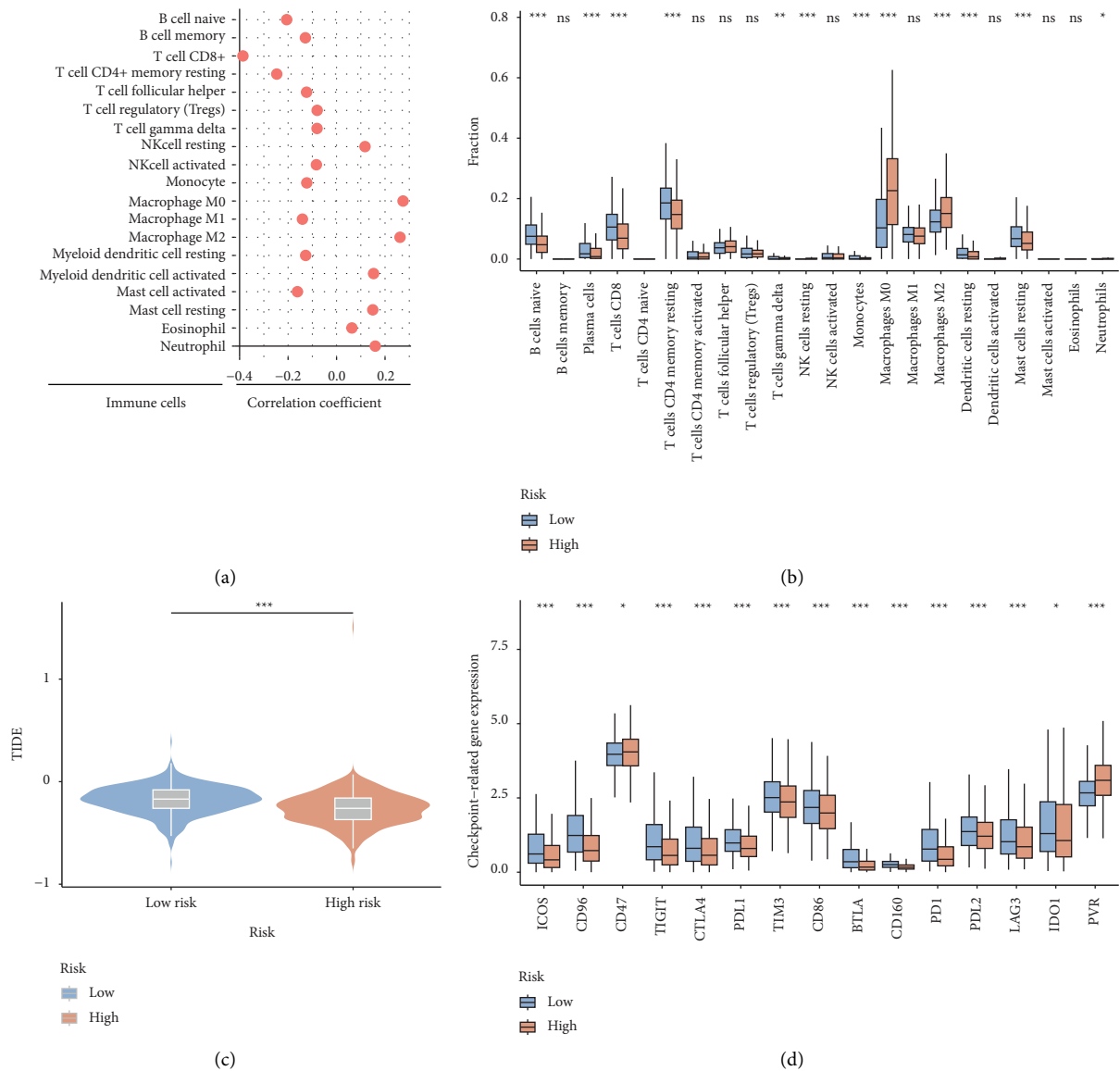


FIGURE 5: The estimation of immune cells infiltration and immunotherapy prediction for the prognostic model in the TCGA training set. Ns, not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. The correlation analysis between the risk score and 19 tumor-infiltrating immune cells evaluated by CIBERSORT algorithm ($p < 0.05$) (a). The differences of 22 tumor-infiltrating immune cells between low- and high-risk groups evaluated by CIBERSORT algorithm (b). Violin plot showing the differential TIDE scores between the high-risk and low-risk groups (c). Box plot showing the expression of immune checkpoint-related markers in low- and high-risk groups (d).

it is involved in the differentiation, cell proliferation, and E-cadherin expression of NE cells in SCLC cell lines by regulating the Wnt signaling pathway [31]. Our work also found that ASCL1 was a risk factor for BC prognosis patients, implying that it might promote breast cancer tumorigenesis. ARID5A (AT-rich interactive domain-containing protein 5a) is one of the Arid family of proteins and possesses the ability to bind nucleic acids, which exist in the nucleus under normal conditions [32]. Meanwhile, ARID5A has been shown to mainly regulate inflammatory and autoimmune disease development by regulating the expression of Interleukin-6 (IL-6) mRNA [33]. Subsequent research revealed that ARID5A could regulate IL-6 mRNA stability through NF- κ B and MAPK signaling pathways [34]. As for

IKZF3 (Aiolos), it belongs to the family of Cys2-His2 zinc finger proteins, which is a lymphocyte lineage transcription factor necessary for the survival of the malignant [35]. For multiple myeloma, immunomodulatory drugs, such as thalidomide could activate E3-ubiquitin ligases and induce degradation of key transcription factors to exert direct antimyeloma effects and promote the survival of myeloma [36]. Then, IKZF3 was a frequently mutated tumor suppressor gene in acute lymphoblastic leukemia (ALL), and its deletion could block lymphocytic lineage differentiation and increase the susceptibility to developing ALL [37]. Consistently, we also found that IKZF3 was a protective factor for BC and facilitated the prognosis of BC patients. Lysine-specific histone demethylase 4B (KDM4B) is a histone

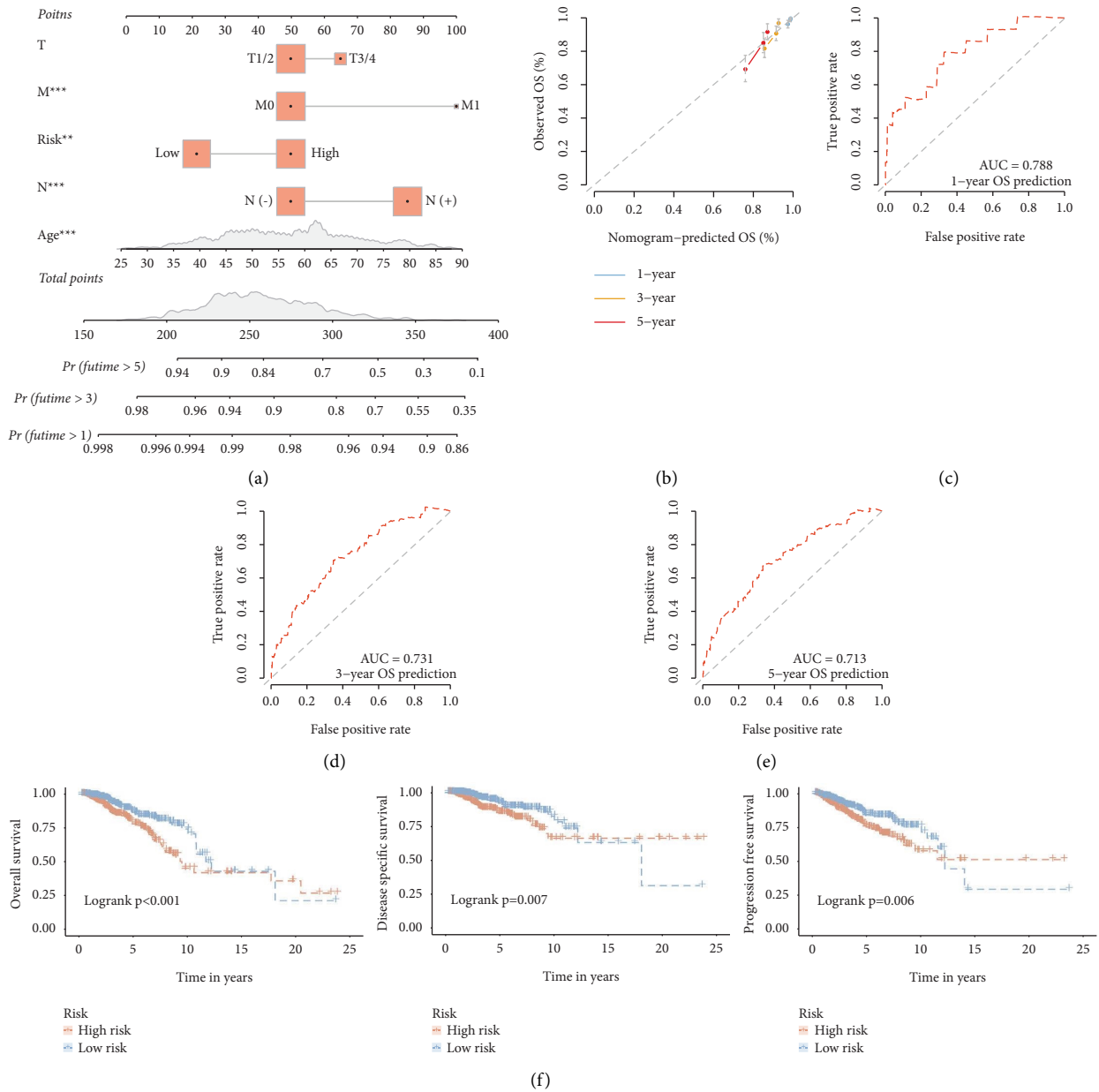


FIGURE 6: Construction and evaluation of a nomogram in the TCGA training set. The nomogram consists of age, T stage, N stage, M stage and risk score to predict the probability of 1-, 3-, and 5-year OS in BC patients (a). Calibration curves of 1-, 3- and 5- year OS in BC patients predicted by the Nomogram (b). ROC curves of 1-, 3-, and 5-year OS predicted by the nomogram (c)–(e). Kaplan–Meier survival curves for OS, DSS and PFS in BC patients based on the nomogram (f).

demethylase for H3K9me3 [38]. According to the genome-wide analysis, KDM4B might be a cancer-specific regulator of alternative splicing by regulating additional alternative splicing-related genes involved in tumorigenesis [39]. In breast cancer, studies have revealed that KDM4B not only antagonizes H3K9 tri-methylation in peripheral heterochromatin and affects H3K4/H3K9 methylation but also plays a role in estrogen receptor α -regulated breast cancer development and mammary epithelial cells proliferation [40, 41]. At the same time, KDM4B was the first identified androgen receptor (AR)-regulated demethylase with effects on AR signaling and turnover and might be a therapeutic

target for prostate cancer [42]. The PR-domain (PRDM) family of genes and the putative transcriptional regulators belong to the SET domain family of histone methyltransferases, which can directly catalyze histone lysine methylation or work by recruiting transcriptional cofactors [43]. Some of the PRDMs are deregulated in cancer and act as tumor suppressors or oncogenes, especially in hematologic malignancies and solid cancers [44]. In diffuse large B-cell lymphomas (DLBCLs), the study showed that the overexpression of PRDM11 (PR-domain family member Prdm11) could induce apoptosis in the E μ -Myc mouse model, then, the DLBCLs patients with low levels of

PRDM11 correlate with shorter overall survival [45]. Additionally, PRDM11 was also identified as a novel locus associated with forced vital capacity, which could be a new target for lung diseases in a genome-wide association analysis [46]. TFF1 (trefoil factor 1), one of the trefoil factor family (TFF), is a small molecule peptide and prevalent in the mucosal environment [47]. In human gastric cancer, it is widely accepted that TFF1 is markedly low-expressed and functions as a gastric tumor suppressor [48]. In BC, although the serum and tissue levels of TFF1 are typically overexpressed [49], many clinical studies have also reported that TFF1 deficiency increases tumorigenicity of human breast cancer cells, and TFF1 expression in BC has an effect on good clinical outcomes for patients [50, 51]. Taken together, all 6 CR-related genes have been reported as taking part in the development of tumors, playing a role in cancer suppression or carcinogenesis. However, the prognostic role in tumors is still less elucidated, and more studies and investigations are needed to understand their prognostic value and mechanisms.

To our knowledge, this is the first time to establish and validate the chromatin regulator-related gene prognostic signature using a large sample size and a high AUC value for breast cancer. Regardless, several limitations can be further improved in the study. In the first place, although our results showed the predictive potentiality and clinical value of our signature, the potential mechanisms of these 6 CR-related genes in BC still require more in-depth experimental investigation. Secondly, the data and information from a total of 1,362 BC patients in the public database used to build the prognostic signature and validate the predictive efficiency of this model are inadequate; therefore, prospective clinical studies are supposed to further confirm our findings.

In conclusion, we identified CR-related DEGs and their predictive ability of prognosis in breast cancer. After that, a novel 6-gene signature model using CR-related DEGs was developed and validated to predict the OS and immunotherapeutic sensitivity for BC patients. Furthermore, a nomogram integrating this novel gene signature and clinical factors was constructed to accurately predict the prognosis for breast cancer, which might provide individualized treatment and aid clinical decision-making for BC patients through prospective validation experiments in the future.

Data Availability

The data used in this study can be obtained from the public databases, namely, The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO).

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

DF, WL, WW, and UDK analyzed the data and wrote the original draft. PG, GH, and XH reviewed and edited the article. WS and HL acquired fund, developed methodology,

administered the project, collected resources, and supervised the study. DF, WL, WS, and HL contributed equally to this study.

Supplementary Materials

Supplementary Table 1: clinical information of BC patients in this study. Supplementary Table 2: chromatin regulators-related differentially expressed genes in breast cancer. (*Supplementary Materials*)

References

- [1] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2022," *CA: A Cancer Journal for Clinicians*, vol. 72, no. 1, pp. 7–33, 2022.
- [2] E. Schwich and V. Rebmann, "The inner and outer qualities of extracellular vesicles for translational purposes in breast cancer," *Frontiers in Immunology*, vol. 9, p. 584, 2018.
- [3] A. Braden, R. Stankowski, J. Engel, and A. Onitilo, "Breast cancer biomarkers: risk assessment, diagnosis, prognosis, prediction of treatment efficacy and toxicity, and recurrence," *Current Pharmaceutical Design*, vol. 20, no. 30, pp. 4879–4898, 2014.
- [4] Z. Chen, M. Wang, R. L. De Wilde et al., "A machine learning model to predict the triple negative breast cancer immune subtype," *Frontiers in Immunology*, vol. 12, Article ID 749459, 2021.
- [5] Z. Liu, M. Li, Q. Hua, Y. Li, and G. Wang, "Identification of an eight-lncRNA prognostic model for breast cancer using WGCNA network analysis and a Cox-proportional hazards model based on L1-penalized estimation," *International Journal of Molecular Medicine*, vol. 44, no. 4, pp. 1333–1343, 2019.
- [6] J. Lu, J. Xu, J. Li et al., "FACER: comprehensive molecular and functional characterization of epigenetic chromatin regulators," *Nucleic Acids Research*, vol. 46, no. 19, pp. 10019–10033, 2018.
- [7] A. L. Steunou, M. Cramet, D. Rossetto et al., "Combined action of histone reader modules regulates NuA4 local acetyltransferase function but not its recruitment on the genome," *Molecular and Cellular Biology*, vol. 36, no. 22, pp. 2768–2781, 2016.
- [8] M. A. Morgan and A. Shilatifard, "Chromatin signatures of cancer," *Genes & Development*, vol. 29, no. 3, pp. 238–249, 2015.
- [9] X. S. Shu, L. Li, and Q. Tao, "Chromatin regulators with tumor suppressor properties and their alterations in human cancers," *Epigenomics*, vol. 4, no. 5, pp. 537–549, 2012.
- [10] C. H. Yeh, X. T. Bai, R. Moles et al., "Mutation of epigenetic regulators TET2 and MLL3 in patients with HTLV-I-induced acute adult T-cell leukemia," *Molecular Cancer*, vol. 15, no. 1, p. 15, 2016.
- [11] A. M. Paul, M. R. Pillai, and R. Kumar, "Prognostic significance of dysregulated epigenomic and chromatin modifiers in cervical cancer," *Cells*, vol. 10, p. 2665, 2021.
- [12] K. Zhang, H. Zhao, K. Zhang, C. Hua, X. Qin, and S. Xu, "Chromatin-regulating genes are associated with post-operative prognosis and isocitrate dehydrogenase mutation in astrocytoma," *Annals of Translational Medicine*, vol. 8, no. 23, p. 1594, 2020.

- [13] F. Chen, X. Zhuang, L. Lin et al., "New horizons in tumor microenvironment biology: challenges and opportunities," *BMC Medicine*, vol. 13, no. 1, p. 45, 2015.
- [14] D. F. Quail and J. A. Joyce, "Microenvironmental regulation of tumor progression and metastasis," *Nature Medicine*, vol. 19, no. 11, pp. 1423–1437, 2013.
- [15] S. A. Rosenberg, "Cancer immunotherapy comes of age," *Nature Clinical Practice Oncology*, vol. 2, no. 3, p. 115, 2005.
- [16] Y. Zhang and Z. Zhang, "The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications," *Cellular and Molecular Immunology*, vol. 17, no. 8, pp. 807–821, 2020.
- [17] A. D. Waldman, J. M. Fritz, and M. J. Lenardo, "A guide to cancer immunotherapy: from T cell basic science to clinical practice," *Nature Reviews Immunology*, vol. 20, no. 11, pp. 651–668, 2020.
- [18] E. J. Lipson, P. M. Forde, H. J. Hammers, L. A. Emens, J. M. Taube, and S. L. Topalian, "Antagonists of PD-1 and PD-L1 in cancer treatment," *Seminars in Oncology*, vol. 42, no. 4, pp. 587–600, 2015.
- [19] J. Larkin, V. Chiarion-Sileni, R. Gonzalez et al., "Combined nivolumab and ipilimumab or monotherapy in untreated melanoma," *New England Journal of Medicine*, vol. 373, pp. 23–34, 2015.
- [20] M. A. Postow, R. Sidlow, and M. D. Hellmann, "Immune-related adverse events associated with immune checkpoint blockade," *New England Journal of Medicine*, vol. 378, no. 2, pp. 158–168, 2018.
- [21] S. L. Topalian, J. M. Taube, R. A. Anders, and D. M. Pardoll, "Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy," *Nature Reviews Cancer*, vol. 16, no. 5, pp. 275–287, 2016.
- [22] G. T. Gibney, L. M. Weiner, and M. B. Atkins, "Predictive biomarkers for checkpoint inhibitor-based immunotherapy," *The Lancet Oncology*, vol. 17, no. 12, pp. e542–e551, 2016.
- [23] K. J. Kao, K. M. Chang, H. C. Hsu, and A. T. Huang, "Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization," *BMC Cancer*, vol. 11, no. 1, p. 143, 2011.
- [24] J. Ferlay, M. Colombet, I. Soerjomataram et al., "Cancer incidence and mortality patterns in Europe: estimates for 40 countries and 25 major cancers in 2018," *European journal of cancer*, vol. 103, pp. 356–387, 2018.
- [25] W. Shi, D. Hu, S. Lin, and R. Zhuo, "Five-mRNA signature for the prognosis of breast cancer based on the ceRNA network," *BioMed Research International*, vol. 2020, Article ID 9081852, 17 pages, 2020.
- [26] F. Li, Q. Huang, T. A. Luster et al., "In vivo epigenetic CRISPR screen identifies *Asf1a* as an immunotherapeutic target in *k*-mutant lung adenocarcinoma," *Cancer Discovery*, vol. 10, no. 2, pp. 270–287, 2020.
- [27] W. Sheng, M. W. LaFleur, T. H. Nguyen et al., "LSD1 ablation stimulates anti-tumor immunity and enables checkpoint blockade," *Cell*, vol. 174, no. 3, pp. 549–563.e19, 2018.
- [28] F. F. Vasconcelos and D. S. Castro, "Transcriptional control of vertebrate neurogenesis by the proneural factor *Ascl1*," *Frontiers in Cellular Neuroscience*, vol. 8, p. 412, 2014.
- [29] B. A. Westerman, S. Neijenhuis, A. Poutsma et al., "Quantitative reverse transcription-polymerase chain reaction measurement of HASH1 (ASCL1), a marker for small cell lung carcinomas with neuroendocrine features," *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, vol. 8, no. 4, pp. 1082–1086, 2002.
- [30] T. Jiang, B. J. Collins, N. Jin et al., "Achaete-scute complex homologue 1 regulates tumor-initiating capacity in human small cell lung cancer," *Cancer Research*, vol. 69, no. 3, pp. 845–854, 2009.
- [31] Y. Tenjin, S. Kudoh, S. Kubota et al., "Ascl1-induced Wnt11 regulates neuroendocrine differentiation, cell proliferation, and E-cadherin expression in small-cell lung cancer and Wnt11 regulates small-cell lung cancer biology," *Laboratory Investigation*, vol. 99, no. 11, pp. 1622–1635, 2019.
- [32] J. Iwahara and R. T. Clubb, "Solution structure of the DNA binding domain from Dead ringer, a sequence-specific AT-rich interaction domain (ARID)," *The EMBO Journal*, vol. 18, no. 21, pp. 6084–6094, 1999.
- [33] K. Masuda, B. Ripley, R. Nishimura et al., "Arid5a controls IL-6 mRNA stability, which contributes to elevation of IL-6 level in vivo," *Proceedings of the National Academy of Sciences*, vol. 110, no. 23, pp. 9409–9414, 2013.
- [34] K. K. Nyati, K. Masuda, M. M. U. Zaman et al., "TLR4-induced NF- κ B and MAPK signaling regulate the IL-6 mRNA stabilizing protein Arid5a," *Nucleic Acids Research*, vol. 45, no. 5, pp. 2687–2703, 2017.
- [35] Q. L. Sievers, G. Petzold, R. D. Bunker et al., "Defining the human c2h2 zinc finger degrome targeted by thalidomide analogs through crbn," *Science*, vol. 362, 2018.
- [36] G. Lu, R. E. Middleton, H. Sun et al., "The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins," *Science*, vol. 343, pp. 305–309, 2014.
- [37] S. Winandy, P. Wu, and K. Georgopoulos, "A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma," *Cell*, vol. 83, no. 2, pp. 289–299, 1995.
- [38] B. D. Fodor, S. Kubicek, M. Yonezawa et al., "Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells," *Genes & Development*, vol. 20, no. 12, pp. 1557–1562, 2006.
- [39] L. Duan, Z. Chen, J. Lu et al., "Histone lysine demethylase KDM4B regulates the alternative splicing of the androgen receptor in response to androgen deprivation," *Nucleic Acids Research*, vol. 47, no. 22, pp. 11623–11636, 2019.
- [40] L. Shi, L. Sun, Q. Li et al., "Histone demethylase JMJD2B coordinates H3K4/H3K9 methylation and promotes hormonally responsive breast carcinogenesis," *Proceedings of the National Academy of Sciences*, vol. 108, no. 18, pp. 7541–7546, 2011.
- [41] L. Gaughan, J. Stockley, K. Coffey et al., "KDM4B is a master regulator of the estrogen receptor signalling cascade," *Nucleic Acids Research*, vol. 41, no. 14, pp. 6892–6904, 2013.
- [42] Y. W. Chin and S. Y. Han, "KDM4 histone demethylase inhibitors for anti-cancer agents: a patent review," *Expert Opinion on Therapeutic Patents*, vol. 25, no. 2, pp. 135–144, 2015.
- [43] F. Di Tullio, M. Schwarz, H. Zorgati, S. Mzoughi, and E. Guccione, "The duality of PRDM proteins: epigenetic and structural perspectives," *FEBS Journal*, vol. 289, no. 5, pp. 1256–1275, 2022.
- [44] C. K. Fog, G. G. Galli, and A. H. Lund, "PRDM proteins: important players in differentiation and disease," *BioEssays*, vol. 34, no. 1, pp. 50–60, 2012.
- [45] C. K. Fog, F. Asmar, C. Côme et al., "Loss of PRDM11 promotes MYC-driven lymphomagenesis," *Blood*, vol. 125, no. 8, pp. 1272–1281, 2015.
- [46] D. W. Loth, M. S. Artigas, S. A. Gharib et al., "Genome-wide association analysis identifies six new loci associated with

- forced vital capacity,” *Nature Genetics*, vol. 46, no. 7, pp. 669–677, 2014.
- [47] L. Thim, “A new family of growth factor-like peptides. ”Trefoil’ disulphide loop structures as a common feature in breast cancer associated peptide (pS2), pancreatic spasmolytic polypeptide (PSP), and frog skin peptides (spasmolysins),” *FEBS Letters*, vol. 250, no. 1, pp. 85–90, 1989.
- [48] M. Soutto, Z. Chen, A. M. Katsha et al., “Trefoil factor 1 expression suppresses *Helicobacter pylori*-induced inflammation in gastric carcinogenesis,” *Cancer*, vol. 121, no. 24, pp. 4348–4358, 2015.
- [49] M. H. Elnagdy, O. Farouk, A. K. Seleem, and H. A. Nada, “TFF1 and TFF3 mRNAs are higher in blood from breast cancer patients with metastatic disease than those without,” *Journal of oncology*, vol. 2018, Article ID 4793498, pp. 1–8, 2018.
- [50] E. Buache, N. Etique, F. Alpy et al., “Deficiency in trefoil factor 1 (TFF1) increases tumorigenicity of human breast cancer cells and mammary tumor development in TFF1-knockout mice,” *Oncogene*, vol. 30, no. 29, pp. 3261–3273, 2011.
- [51] L. K. Dunnwald, M. A. Rossing, and C. I. Li, “Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients,” *Breast Cancer Research*, vol. 9, no. 1, p. R6, 2007.