Research Paper

Subtype recognition and identification of a prognosis model characterized by antibody-dependent cell phagocytosis-related genes in breast cancer

Li Wang[1,*], Menghan Li[2,3,*], Hongyu Yang[4], Fenghuan Dai[2], Ning Xie[2], Linhui Li[3], Meiying Zhu[4], Ran Ding[5,*]

[1]Department of Breast Cancer, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Key Laboratory of Breast Cancer Prevention and Therapy, Tianjin Medical University, Ministry of Education, Tianjin 300060, China
[2]Acupuncture-Moxibustion Clinical Department, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, National Clinical Research Center for Chinese Medicine Acupuncture and Moxibustion, Tianjin 300381, China
[3]Graduate School, Tianjin University of Traditional Chinese Medicine, Tianjin 300381, China
[4]Department of Oncology, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin 300381, China
[5]School of Biomedical Sciences and Engineering, South China University of Technology, Guangzhou International Campus, Guangzhou 511442, China
*Equal contribution

Correspondence to: Li Wang, Ran Ding; **email:** wangli_tjch@tmu.edu.cn; ran_ding_ahtcm@163.com, https://orcid.org/0000-0003-3564-1142
Keywords: antibody-dependent cell phagocytosis, breast cancer, immune checkpoint inhibition, prognostic biomarker, tumor microenvironment
Received: October 10, 2023 **Accepted:** January 23, 2024 **Published:** February 22, 2024

ABSTRACT

Background: Breast cancer (BC) is a heterogeneous tumor with a variety of etiology and clinical features. Antibody-dependent cell phagocytosis (ADCP) is the last step of immune checkpoint inhibition (ICI), and macrophages detect and recognize tumor cells, then destroy and engulf tumor cells. Despite the large number, negative regulators that inhibit phagocytic activity are still a key obstacle to the full efficacy of ICI.
Patients and methods: An ADCP-related risk score prognostic model for risk stratification as well as prognosis prediction was established in the Cancer Genome Atlas (TCGA) cohort. The predictive value of ADCP risk score in prognosis and immunotherapy was also further validated in the TCGA along with International Cancer Genome Consortium cohorts. To promote the clinical application of the risk score, a nomogram was established, with its effectiveness verified by different methods.
Results: In this study, the genes collected from previous studies were defined as ADCP-related genes. In BC patients, two ADCP-related subtypes were identified. The immune characteristics and prognostic stratification were significant different between them.
Conclusions: We identified two subtypes associated with ADCP gene expression in breast cancer. They have significant differences in immune cells, molecular functions, HLA family genes, immune scores, stromal scores, and inflammatory gene expression, which have important guiding significance for the selection of clinical treatment methods. At the same time, we constructed a risk model based on ADCP, and the risk score can be used as a good indicator of prognosis, providing potential therapeutic advantages for chemotherapy and immunotherapy, thus helping the clinical decision-making of BC patients.

INTRODUCTION

The complete antibody of human immunoglobulin (IgG) consists of an antigen-binding fragment (Fab) and a crystal fragment (Fc) that binds to the Fcγ receptor. In the constant region of IgG1, IgG2, IgG3 and IgG4, IgG1 exhibits the highest FcγR binding affinity compared to IgG3, IgG2 and IgG4, particularly in the subclasses characterized by different hinge and CH2 domains [1]. The Fc region binds to FcγR on the surface of immune cells, and antibodies can cause some effector cells to produce strong cytotoxic effects [2]. Therefore, different subtypes possess distinct effector functions, such as ADCP.

Antibody-dependent cell phagocytosis (ADCP) of tumor is a kind of antibody that binds to target cells (such as tumor cells) and effector cells (such as macrophages) at the same time, resulting in the phagocytosis of target cells by effector cells [3]. After phagocytosis, target cells are digested and degraded in effector cells by acidification. ADCP is considered to be the main MOA for a variety of biological drugs such as anti-CD20, CD38, EGFR, HER-2 [4]. At the same time, macrophage-mediated tumor cell phagocytosis by using anti-CD47 antibodies to block the anti-phagocytic CD47-SIRPa interaction has shown promise in preclinical xenotransplantation of various human malignancies [5]. The satisfactory anti-cancer effect of ADCP depends on early antibody administration, and drug resistance develops with cancer progression [6].

BC is a heterogeneous tumor with a variety of etiology and clinical presentations. Studies have revealed the presence of various subtypes involved in the development and progression of BC [7]. Our research introduces the research status of breast cancer subtypes in detail and identifies new subtypes by means of machine learning [8], which opens up new fields for breast cancer research.

A large number of studies have proved the importance of ADCP in tumor treatment. In this study, we determined the ADCP-related subtypes of breast cancer, and comprehensively discussed the significance of ADCP in multi-omics. It provides an important basis for clinical precision treatment.

In this study, genes defined as ADCP-related genes were collected from previous studies. In BC patients, two ADCP-related subtypes were identified. The immune characteristics as well as prognostic stratification were significant different between the two ADCP-related subtypes. Based on these findings, it can be inferred that there may be a significant relationship between ADCP in BC and TME (tumor microenvironment). Risk score could be used as a good indicator for prognosis, potential in offering therapeutic advantages for chemotherapy as well as immunotherapy, thereby aiding in the clinical decision-making process for BC patients.

MATERIALS AND METHODS

Data download and processing

Study cohorts A and B were based on data obtained from two publicly available datasets that derived from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC, https://www.cbioportal.org/study/summary?id=brca_metabric) databases, respectively [9, 10]. Cohort A consisted of 103 paracancerous samples and 903 breast cancer (BC) samples, which were used for gene expression profiling. On the other hand, cohort B included 154 paracancerous samples and 1826 BC samples for gene expression profiling. In order to evaluate the reliability of the prognostic model constructed in this study, cohort A was designated as the training set, while cohort B served as the validation set. This division allowed us to train the model on one dataset and validate its performance on an independent dataset.

Identification of the survival-related ADCP genes

To evaluate the prognostic significance of ADCP genes, Cox regression analysis was further adopted to evaluate the relationship of each gene with survival status in the TCGA cohort. To mitigate the risk of overlooking potentially important genes, a truncated p-value of 0.05 was set as a threshold, and a total of 130 survival-related genes were identified for further analysis.

Consensus clustering analysis of ADCP genes

The unsupervised clustering 'Pam' method was applied to the identification of different molecular subtypes based on survival-related ADCP gene expression [11]. The R package Consensus Cluster Plus performs this process and determines the number of clusters in Queue A, with 1000 times repeated to make sure classification stability [12]. Principal component analysis (PCA) was employed to illustrate the distribution differences. To evaluate the clinical significance of ADCP subtypes, their relationship with prognosis and other clinicopathological features (including age, stage and cancer subtypes) were examined. In different cohorts, the overall survival (OS) of different clusters was compared using the Kaplan-Meier (K-M) survival chart, with $P \leq$

0.05 indicating statistical significance. Besides, the relationship between ADCP subtypes and other clinical variables was visualized using the Sankey plot and drawn using the ggalluvial R package [13]. The heat maps of different ADCP gene expression patterns were plotted by the pheatmap R package [14].

Identification and verification of the key ADCP genes

Based on the TCGA, the 'limma' R package was adopted to identify the differentially expressed genes (DEGs) between group1 and group2 [15]. The fold change cut off was 1, and the adjusted p value was less than 0.05. We used the glmnet R package in the TCGA cohort to use the LASSO Cox regression model to screen out the best prognostic biomarkers in 130 ADCP genes [16]. A total of 39 ADCP genes with non-zero coefficients were screened by 10-fold cross-validation. We overlapped the ADCP genes screened by DEGs and LASSO regression to screen out the key ADCP genes. Risk scoring formula:

$$\text{Risk score} = \sum_{i=1}^{n} \text{coefi} * \text{id} \qquad (1)$$

Where Coefi represents the coefficient and id represents the normalized count for each gene.

After constructing the risk score according to the LASSO regression model, the risk grouping and group were visualized through the Sankey diagram using the ggalluvial R package according to the survival situation.

Univariate Cox regression together with K-M plot were used to analyze the prognostic significance of key ADCP genes. P-values less than 0.05 indicated statistical significance. Subsequently, multivariate Cox regression model was utilized to explore whether ADCP genes could be combined with other clinicopathological features as independent prognostic factors.

Correlation of the ADCP genes with copy-number alterations and immune traits

Gene Set Cancer Analysis (GSCA) is a multi-omics online analysis tool based on TCGA data. It was used to analyze the expression and prognosis of pan-oncogenes in key ADCP genes (http://bioinfo.life.hust.edu.cn/web/GSCALite/) [17]. Results are displayed in the bubble plot by the 'ggplot2' package [18]. To analyze the expression characteristics of key ADCP genes in tumor and ADCP, t-test algorithm was used to compare their expression levels in different risk scores, groups and tumor groups in cohort A, respectively. The CAN (copy-number alterations) ratio as well as methylation level (β value) of key ADCP genes among TCGA-BC samples were verified by GSCA online tool. As a convenient online tool, TISIDB provides convenient access to correlation between genes and immune traits, such as lymphocytes, immunomodulators, and chemokines [19]. The scatter plots of ADCP genes with several tumor-infiltrating lymphocytes (TIL), immunosuppressants and immunostimulants were downloaded from TISIDB.

Construction and verification of risk model

For each patient, the risk score was predicted by the 'predict' function in the survival R package [20]. Patients were then divided into high- and low-risk groups according to the median value. To test the accuracy of the Cox regression model, C-indexes as well as receiver operating characteristic curves (ROC) were adopted. Prognostic factors, including C-indexes representing risk score, were visualized using histogram. ROC analysis was performed by the 'pROC' package [21], with the area under the curve (AUC) values calculated and compared to evaluate the performance of the model. In addition, the K-M survival analyses were conducted for risk scores. Based on training and test datasets, 1-, 3- and 5-year ROC curve were plotted using the R package timeROC [22].

Nomogram construction

To further investigate the prognostic significance of ADCP genes, the TCGA datasets were utilized to generate risk plots between high- and low-risk groups. To improve the predictive accuracy, a nomogram incorporating TNM stage, Cancer type and group was developed using survival and rms package [23]. To demonstrate the consistency between the predicted 1-, 3- and 5-year endpoint events and the corresponding actual outcomes, calibration plots were generated using rms package.

Evaluation of drug sensitivity

To investigate the molecular characteristics associated with drug sensitivity/resistance, the Genomics of Drug Sensitivity in Cancer (GSDC) database (https://www.cancerrxgene.org/) [24], a public resource for identifying cancer cell biomarkers, was utilized. The pRRophetic package was adopted to determine the sensitivity data of two BC groups towards various drugs [25]. Furthermore, the drug sensitivity of BC patients with different phenotypes was predicted from gene expression data.

Quantitative reverse transcriptase PCR (qRT-PCR)

The MCF-10A (human breast cancer paracancerous) and MDA-MB-453 (human breast cancer) cell lines were obtained (Shangcheng North Na Chuanglian Biotechnology Co., Ltd.). Control groups were created by combining MCF-10A with MDA-MB-453 cells to compare the expression of DEFB1, SIAH2, and SYT1 between normal and breast cancer cell lines. Total RNA was extracted using the Redzol kit (Beijing SBS Gene Technology Co., Ltd.), and qRT-PCR was performed using the SYBR® Premix Ex Taq™ II Kit. The relative mRNA expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method, with β-actin as the internal reference gene. The forward primer sequences were as follows: (1) DEFB1: F-5′-CTCAGGTGGTAACTTTCTCA-3′; (2) SIAH2: F-5′-CACTTGACAGGCTGTTGCAC-3′; (3) SYT1: F-5′-CGCTTCGGCAGCACATATACTAA AATTGGAAC-3′. (4) GAPDH: F-5′-GAGTCAACGG ATTTGGTCGT-3′. The reverse primer sequences were as follows: (1) DEFB1: R-5′-AAGCACTCCGGGTG ATTCAG-3′; (2) SIAH2: R-5′-AAGCACTCCGGGT GATTCAG-3′; (3) SYT1: R-5′-TTGGTCAGCACAG ATCATCG-3′. (4) GAPDH: R-5′-GATCTCGCTCCTG GAAGATG-3′.

Western blot (WB)

The WB analysis was performed following previously established protocols and the manufacturer's recommendations. The concentrations of extracted proteins were determined using the Bradford method. Proteins were separated on a 10% SDS-polyacrylamide gel, transferred to PVDF membranes, and blocked with 5% skimmed milk. Primary antibodies against DEFB1 (1:500; catalog number ab115813; Abcam), SIAH2 (1:500; catalog number ab31234; Abcam), SYT1 (1:500; catalog number ab302627; Abcam), and GAPDH (1:500; catalog number ab181602; Abcam) were incubated with the membranes overnight at 4° C. Afterward, HRP-labeled secondary antibody was added and the membranes were washed with TBST. WB analysis was conducted using chemiluminescence, and the protein bands were visualized using a film.

Statistical analysis

Statistical analyses were conducted using R software (version 4.2.1) [26]. A two-sided P-value less than 0.05 was considered statistically significant. The K-M survival curve was constructed using the survival curve function 'ggsurvplot' in the R package survminer (0.4.2 version) [27], and the bilateral log-rank test was adopted to estimate the difference in OS between different groups. The cut-off value of ADCP genes, differentiating the high-risk group from the low-risk group, was determined as the median or calculated using the surv cutpoint function implemented in the R package survminer (0.4.2 version) [28]. The LASSO Cox regression model was used to identify the ADCP genes that exhibited a significant association with OS, and multivariate Cox regression analysis was performed using the identified ADCP genes to calculate hazard ratios (HR), 95 % confidence intervals (CI), and corresponding p-values to construct ADCP genes. Additionally, multivariate Cox regression analysis was used to evaluate the independent prognostic value of clinical indicators, including the ADCP genes, and a nomogram visual risk prediction map was formed after scoring each factor using the rms package of R software. To evaluate the consistency of nomogram, the calibration plot was generated using the rms package of R software [29].

Availability of data and material

Publicly available datasets were analyzed in this study. These data can be found here: https://portal.gdc.cancer.gov/, https://www.cbioportal.org/study/summary?id=brca_metabric. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Consent for publication

All authors approved the publication of the article.

RESULTS

Survival-related ADCP gene

To investigate the prognostic significance of ADCP genes, a gene signature was established using data from cohort A. The workflow for the construction of gene signature is depicted in Supplementary Table 1. Based on univariate Cox regression, 130 prognostic genes were screened out (p-value < 0.01, Supplementary Table 1).

Consensus clustering analysis of ADCP genes

Based on the expression levels of survival-related ADCP genes in the TCGA database, two distinct regulatory patterns were identified using an unsupervised clustering method. A total of 515 cases were classified into ADCP-related cluster 1 and 491 cases were classified into ADCP-related cluster 2 (Figure 1A). PCA analysis demonstrated that the patients can be classified into two distinct parts, providing further evidence of the existence of two significantly different subtypes (Figure 1B). To

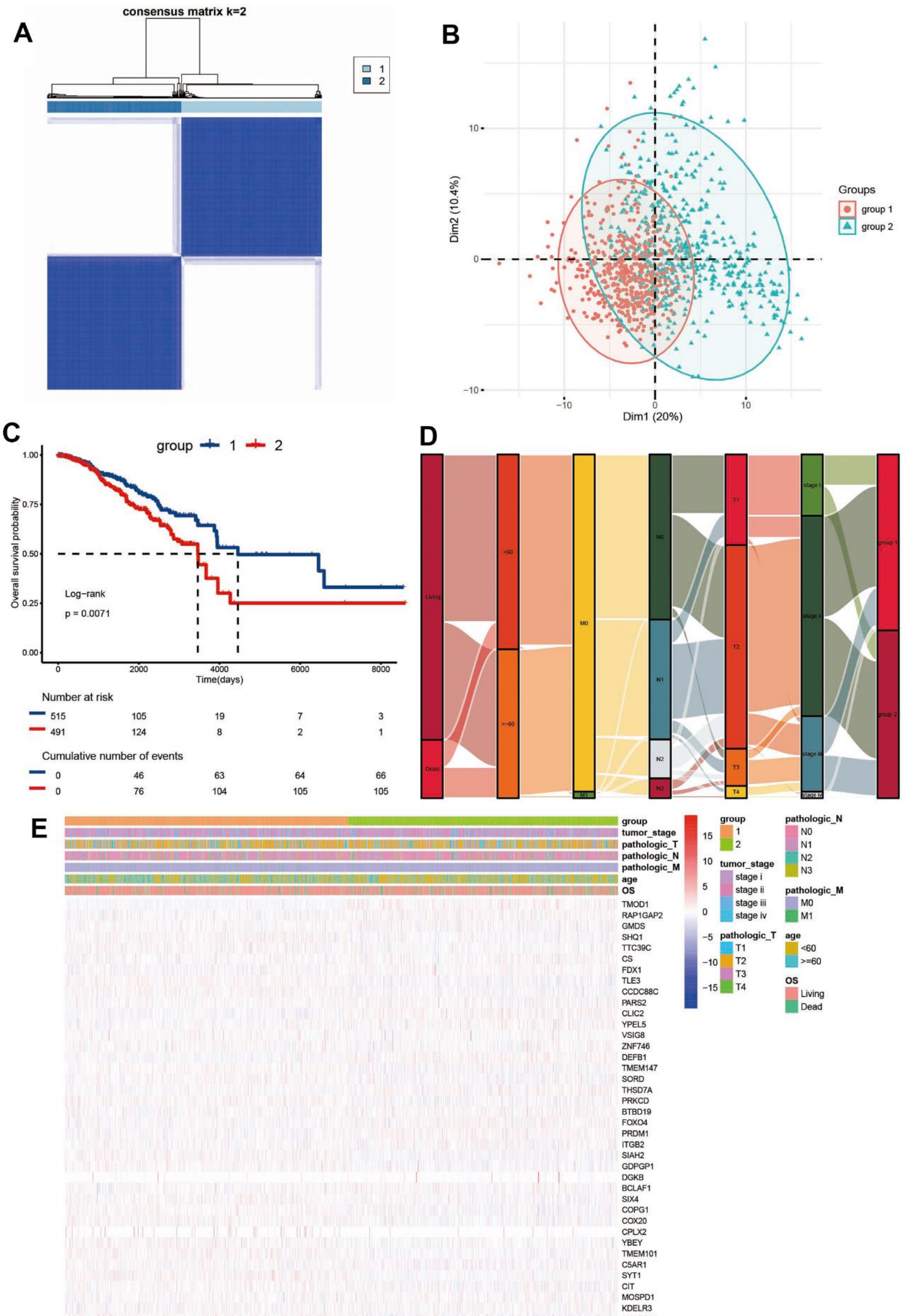

Figure 1. Consensus clustering of ADCP-related genes in BC. (**A**) The consensus matrices and (**B**) PCA analysis were performed to assess the stability of clustering and explore coagulation subtypes. (**C**) Kaplan-Meier curves and (**D**) an alluvial diagram demonstrated the association between ADCP groups, tumor stage, age, grade, and survival state. (**E**) Unsupervised clustering of all coagulation-related genes in TCGA cohorts.

evaluate the survival difference between the two clusters, an analysis was conducted on a diverse range of patients from a variety of cohort A (TCGA) with complete survival information were analyzed. Subsequently, the relationship of these two clusters with various clinical features were analyzed (Figure 1D). The survival advantage of cluster 1 was higher compared to that of cluster 2. In cluster 1, the proportion of age less than 60, M0, N0-1, T1-2, Stagei-ii was higher. These suggest that ADCP genes may affect tumor development through some potential mechanisms. The heat map depicts the transcriptomic characteristics of different expressed ADCP genes in the two ADCP subtypes (Figure 1E).

The immune landscape of ADCP subtypes

To explore the disparity in pathway enrichment analysis between the two clusters of ADCP in cohort A, GSEA was performed, with different immune infiltration patterns identified within the two subtypes. The enrichment histogram showed that cluster 1 significantly enriched hormone metabolism (including early estrogen response, late estrogen response) and inflammatory signal regulation pathways (allograft rejection, inflammatory response, kras signal and il6 jak stat3 signal, tnfa signal conduction through NFKB) (Figure 2A). At the same time, GSEA confirmed that there were differences in the immune pathways of ADCP clusters. The results showed that DEGs with higher expression levels in cluster 1 were significantly enriched in early estrogen response, allograft rejection pathway, inflammatory response, interferon response, kras signaling pathway and tnfa signal transduction through NFKB (Figure 2B). Given the strong correlation between ADCP subtypes and immune activity, the TME of the two clusters in cohort A was investigated (Figure 2C). Cluster 1 subtype is featured with high infiltration of Natural Kill cells (resting) and Macrophages, while cluster 2 subtype exhibits elevated infiltration levels of B cells, Plasma cells, T cells, Natural Kill cells, Dendritic cells, Mast cells and Neutrophils. Based on TCGA expression profile, the stromal score, immune score as well as ESTIMATE score of malignant tumor tissues were calculated through ESTIMATE algorithm. ESTIMATE produces a matrix score that measures the presence of tumor-associated matrices, as well as an immune score that reflects the level of immune cell infiltration. These scores are combined to produce an index called an 'estimated score' which provides a comprehensively estimation tumor purity. Compared with cluster 1, samples in cluster 2 also showed significantly higher estimated scores (Wilcoxon test, $P < 0.05$, Figure 2D). This trend was also observed for matrix scores as well as immune scores (Wilcoxon test, $P < 0.05$). Additionally, we studied the association between the two subtypes and major histocompatibility complex (MHC) and T cell stimulators. In addition to HLA-C, the expression levels of MHC as well as T cell stimulators exhibited a tendency to be higher in cluster 2 (Figure 3E, 3F).

Recognition of key ADCP genes

Using the limma algorithm, a total of 432 degs (Differentially expressed gene) between group1 and group2 are identified under the filtering threshold of FDR q-value < 0.01 and the absolute value of logFC > 1 (Figure 3A and Supplementary Table 2). Prognostic models were constructed for 1006 BC patients with OS information in cohort A. LASSO Cox regression analysis was utilized to determine the best prognostic features based on 130 survival-related ADCP genes. After the variables were included in the LASSO Cox regression model with the smallest lambda, the genes of 39 ADCP-related features were chosen to construct the ADCP-related risk scoring model. Three key ADCP genes (DEFB1, SIAH2 and SYT1) by overlapped DEGs from cohort A and the model ADCP genes were identified (Figure 3B). The relationship of the expression levels with 3 ADCP-related signatures together with OS are also presented in the forest plot (Figure 3D). The expression of three key ADCP genes is of significance for survival (Figure 3E).

Multidimensional analysis of key ADCP genes

We studied the relationship of CNV with immune infiltration in BRCA. Additionally, we explored the association of gene methylation with immune infiltration. CNV and methylation of key ADCP genes were closely associated with the infiltration of key immune cells including T and B cells (Figure 4A, 4B). At the same time, based on cohort A, the gene expression levels of 3 key ADCP genes were analyzed. In the data set, only DEFB1 was down-regulated in group1, while SIAH2 and SYT1 were up-regulated in group2 (Wilcoxon test, $P > 0.05$). Only DEFB1 was up-regulated in the high-risk group, while in the low-risk group, SIAH2 together with SYT1 were down-regulated (Wilcoxon test, $P > 0.05$). Only DEFB1 was up-regulated in the tumor group; in contrast, SIAH2 and SYT1 were down-regulated in the low-risk group (Wilcoxon test, $P > 0.05$, Figure 4C). Among the three key ADCP genes, the frequency of copy number alterations (CNA) of SIAH2 was higher compared to that of the others. Specifically, CNA deletions were predominant among all types (Figure 4D). Methylation analysis showed that the beta value of SIAH2 was higher in the tumor group than in the normal group, and DEFB1 was the opposite (Figure 4E). It is well known that gene expression is negatively correlated with the

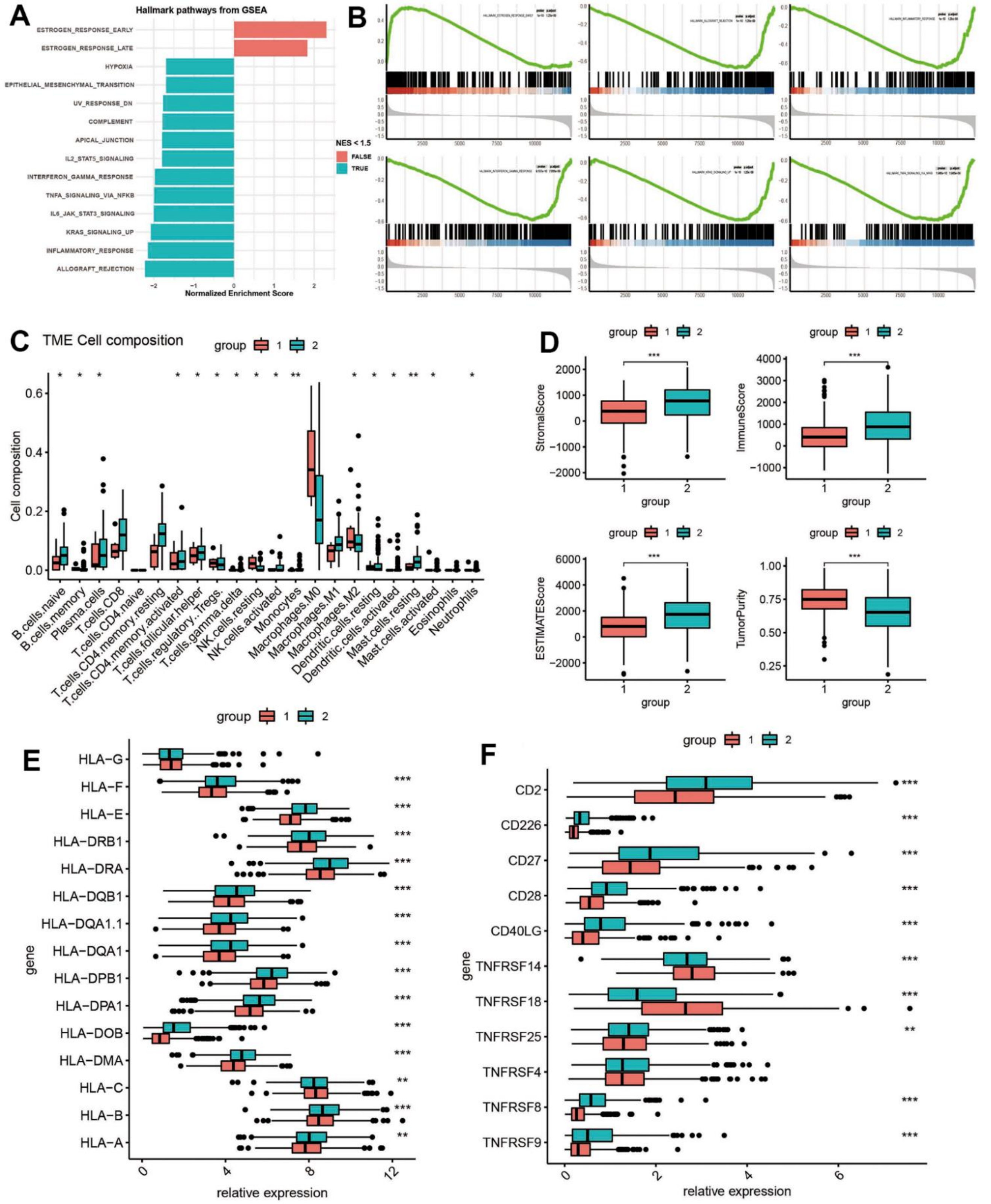

Figure 2. Clinical significance and immune landscape of ADCP groups in the TCGA cohort. (**A**) GSVA and Gene Set Enrichment Analysis (GSEA) revealed the activation or inhibition of biological pathways and (**B**) significant enrichment in immune-associated processes. (**C**) Immune cell infiltration, (**D**) stromal and immune scores, and gene expression of HLA and MHC gene sets were analyzed between ADCP groups. (**E**) The expression box plot of HLA family genes between the two subtypes. (**F**) The expression box plot of inflammatory genes between the two subtypes. Statistical significance at the level of * <0.05, ** <0.01, and *** <0.001.

level of methylation, while CNA has a positive effect on gene expression. In summary, CNA and methylation may be may be important in the up-regulation of SYT1 expression in BC.

ADCP-related signatures for the prognostic prediction of BC

LASSO algorithm was used to establish a risk model. Finally, 39 genes related to prognosis were identified, and used to construct models based on risk scores using the training (n = 1006) and test (n = 1980) data sets of cohorts A and B patients, respectively. Survival analysis showed that a higher risk score of the training and the test sets was related with a lower survival rate ($p < 0.0001$) (Figure 5A, 5B). A time-varying ROC curve was generated for the assessment of the sensitivity of the model. The 3-year, 5-year, and 10-year AUCs were found to be 0.743,0.754, and 0.79 in the training set, respectively (Figure 5C); in contrast, they were 0.546,0.678, and 0.716 in the test set, respectively (Figure 5C).

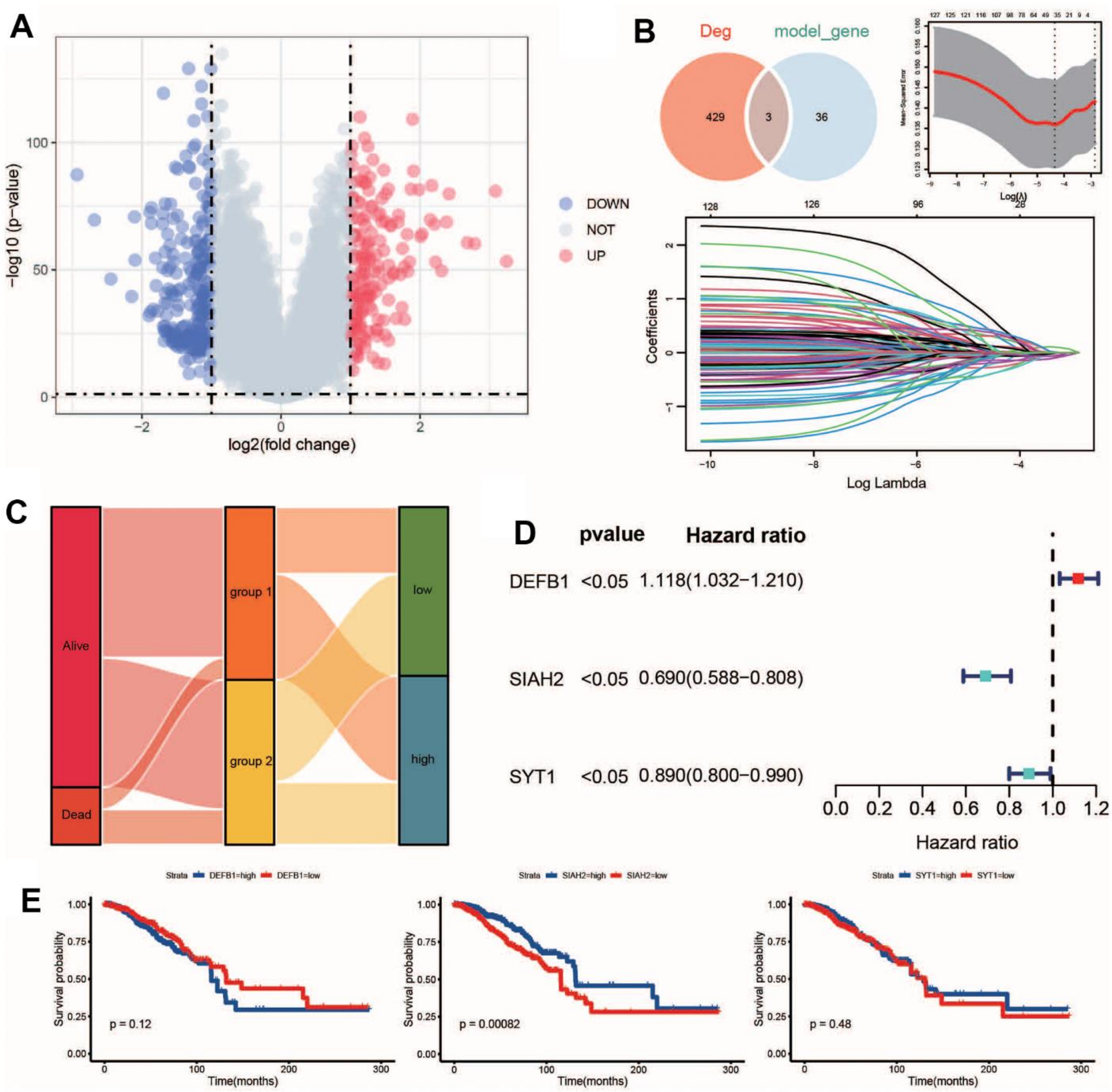

Figure 3. Identification of key ADCP targets. (**A**) Volcano plots and (**B**) LASSO regression model were used to identify differentially expressed genes (DEGs) and select key CRGs. (**C**) An alluvial diagram illustrated the changes in risk groups, ADCP groups, and survival state. (**D**) Univariate Cox regression analysis and (**E**) Kaplan-Meier curves evaluated the associations between infiltrate levels of ADCP targets and overall survival.

The ADCP group served as an independent prognostic factor in BC

Since ADCP genes are significantly associated with high malignancy and advanced tumors of BC, univariate and multivariate Cox regression analysis were carried out to determine the prognostic significance of ADCP genes for BC patients. ADCP group, age, TNM stage, stage together with risk score were included as covariates. Results showed that ADCP group, age, TNM stage, stage as well as risk score were independent prognostic factors for BC patients (Figure 6A, 6B). we constructed a nomogram by combining independent prognostic factors, serving as a clinically relevant quantitative method tool for predicting the mortality of individual BC patient (Figure 6C). Based on the c-exponential curve of different variables over time in the TCGA cohort, nomogram performed best compared to other single factors (Figure 6D). Add up the scores of each prognostic parameter and assign a total score to each patient. The higher the total score, the worse the prognosis of patients. The modal diagram has similar performance to the ideal model (Figure 6E). In addition, the DCA curve also indicated that the nomogram had good stability and reliability net benefit curve in age compared with other clinical factors (Figure 6F).

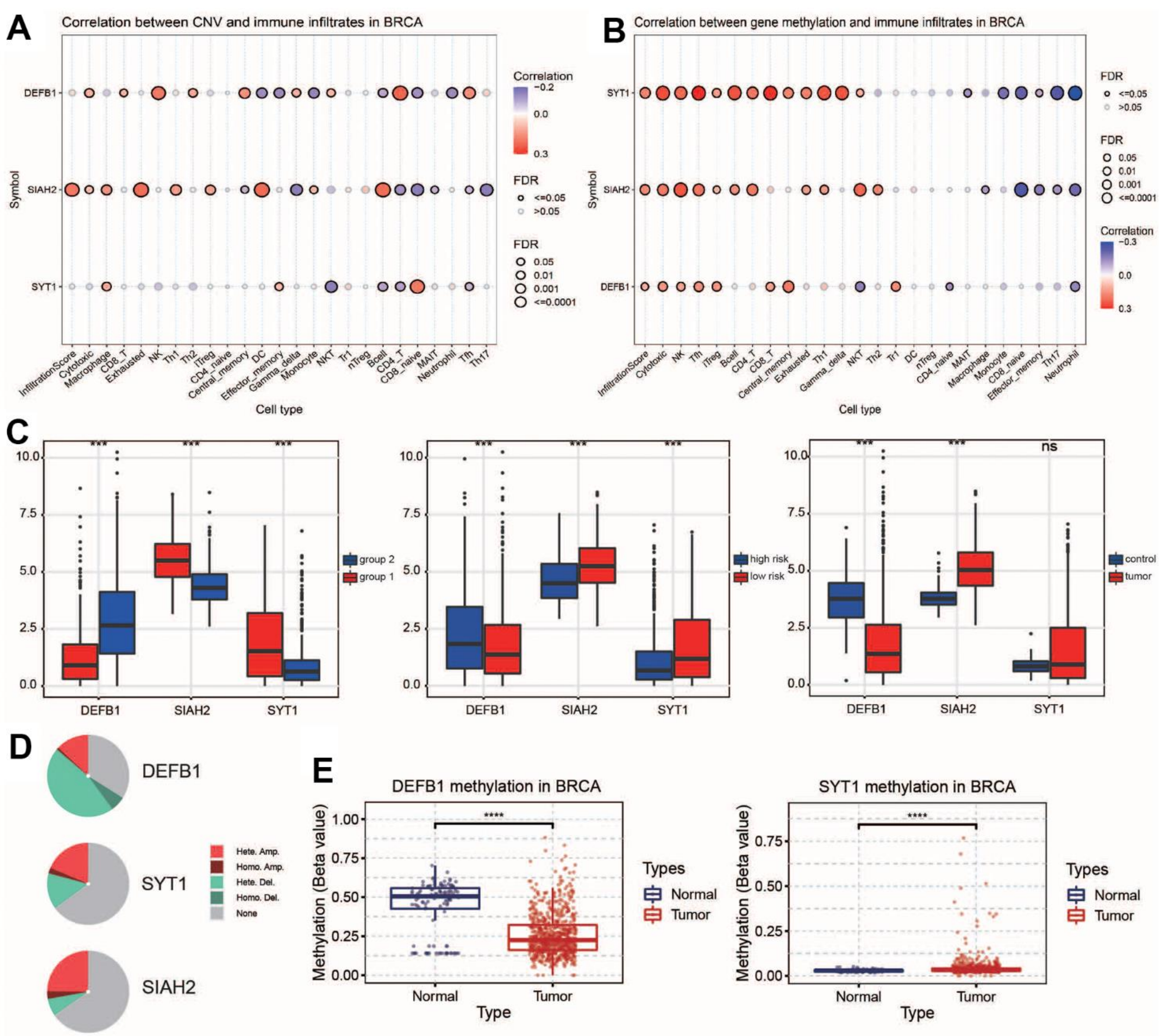

Figure 4. Validation of key ADCP genes in expression level. (**A**, **B**) Correlation analyses were conducted between CNV/methylation and immune infiltrates in BC. (**C**) Expression levels and (**D**, **E**) CNA percentages of key ADCP genes were compared between ADCP groups, risk groups, and cancer types. Methylation levels of key ADCP genes in BC and normal samples were also examined. Statistical significance at the level of ns ≥ 0.05, *** <0.001 and **** <0.0001.

ADCP groups for the prediction of the chemotherapeutic response

We evaluated the chemotherapy response and drug resistance of patients in the ADCP group. Figure 7 shows the sensitivity of two ADCP subtypes to six anticancer drugs (AZD8055, A.443654, AMG.706, AKT.inhibitor.VIII, ABT.888, ATRA). Results showed that the IC50 level of group2 was higher compared to group1 (Figure 7A–7F), and small molecule drugs with therapeutic effects on BC could be found according to the results of drug sensitivity. Three-dimensional structural tomography of AZD8055, A.443654, AMG.706, AKT.inhibitor.VIII, ABT.888 and ATRA was found in PubChem (Figure 7G).

Gene expression level verification via qRT-PCR and WB

We verified the mRNA and protein levels of DEFB1, SIAH2 and SYT1 in BC cell lines and adjacent cell lines by qRT-PCR and WB. Results of qRT-PCR is

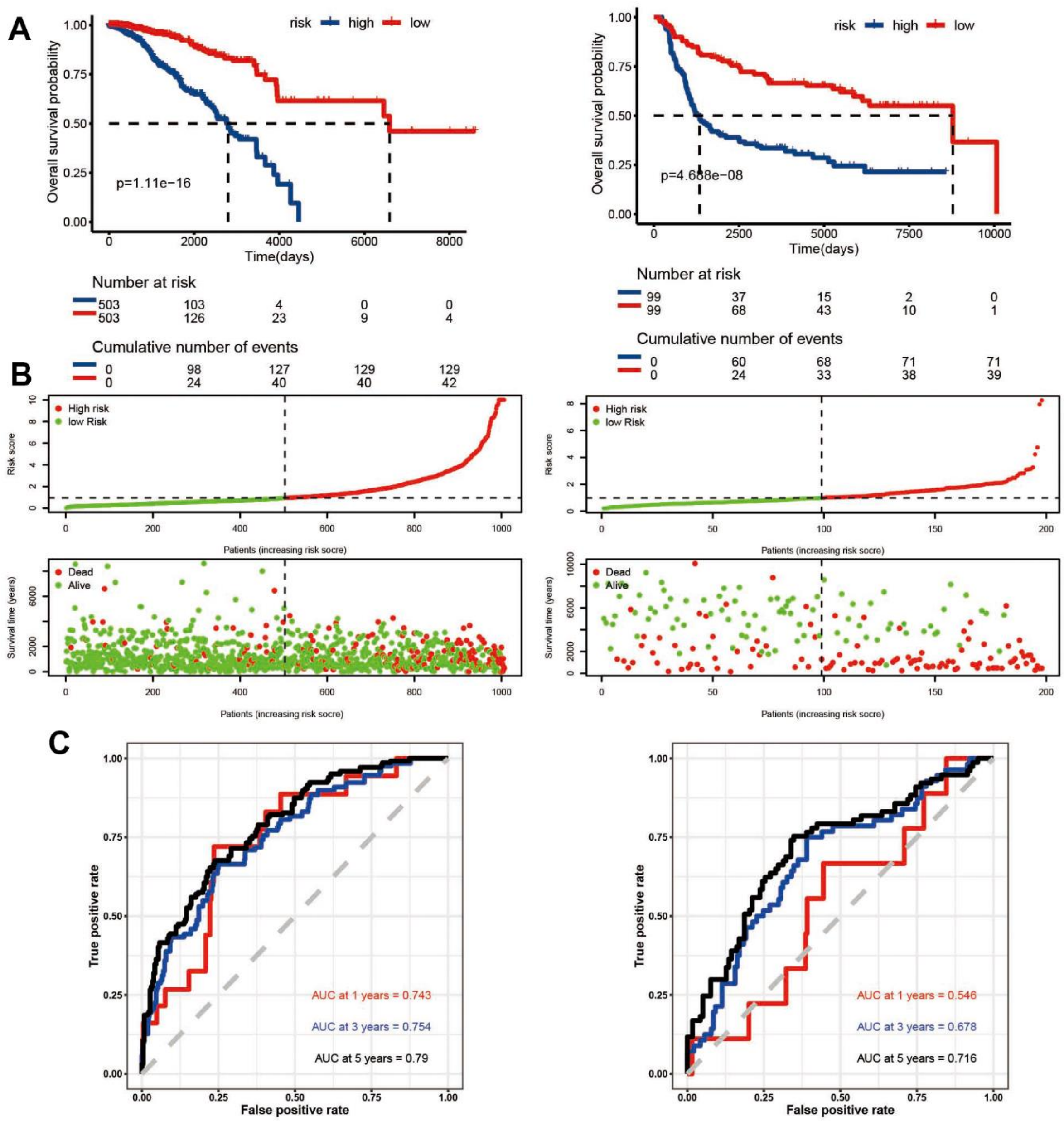

Figure 5. Survival evaluation of ADCP-related risk scoring model in TCGA and METABRIC datasets. (**A**) Kaplan-Meier survival curves, (**B**) patient subgrouping based on risk scores, and (**C**) ROC curves assessed the performance of the ADCP-related risk scoring model in predicting overall survival at 1, 3, and 5 years.

shown in Figure 8A. In contrast with the normal control, DEFB1 level in BC cell line MDA-MB-453 was lower, but the SIAH2 and SYT1 levels were higher. The protein expression levels of the three key ADCP-genes based on WB analysis were consistent with the results of qRT-PCR experiments (Figure 8B, 8C).

DISCUSSION

Traditional cancer treatment methods such as chemotherapy are non-selective and increase the energy of the immune system, leading to serious side effects and tumor recurrence [30]. Since most tumor-associated

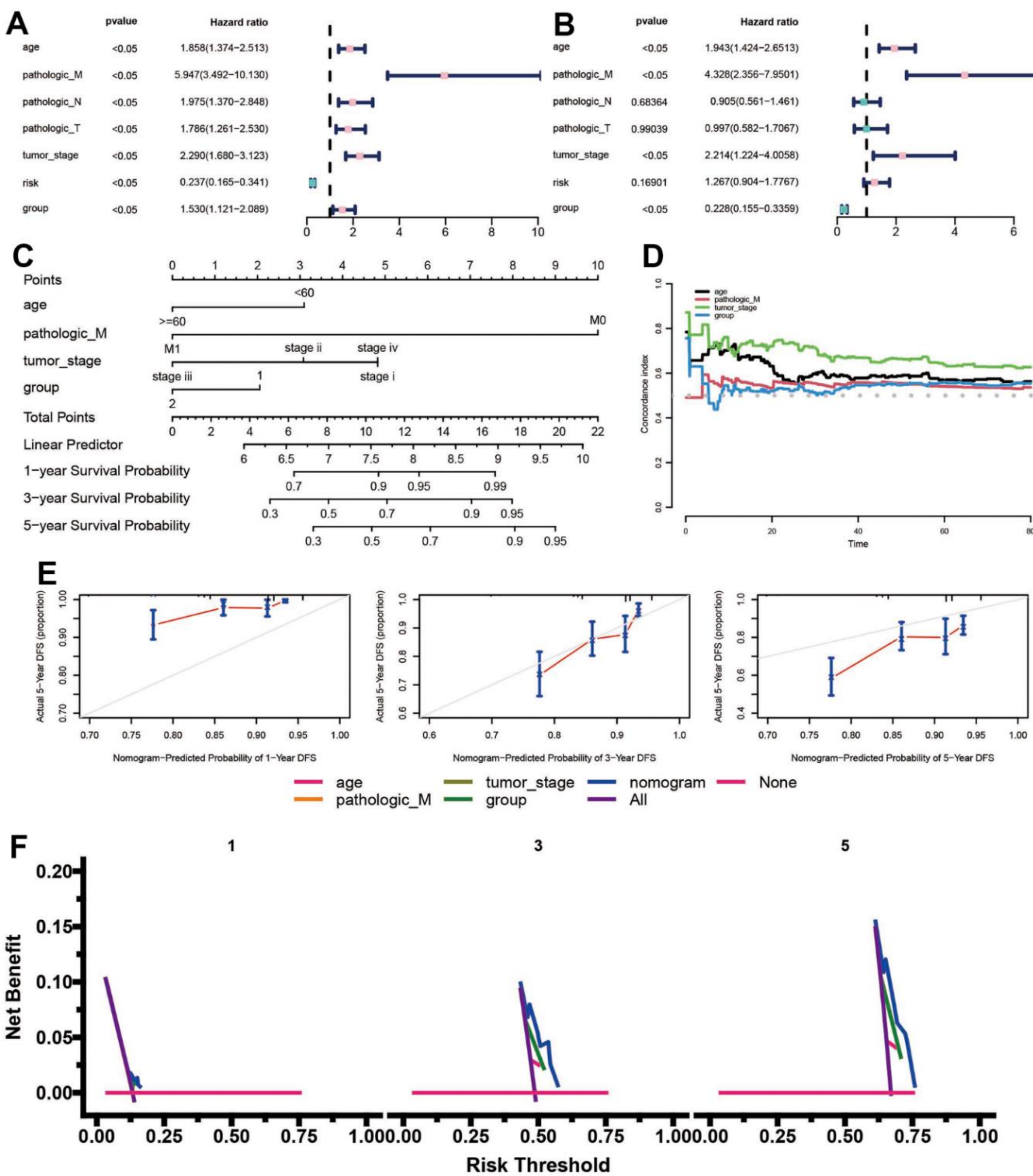

Figure 6. A nomogram was developed to predict the probability of 1-, 3-, and 5-year overall survival (OS) in the training cohort. (**A**, **B**) Univariate and multivariate analyses were conducted, including the ADCP-related risk scoring model and clinical factors. (**C**) The comprehensive nomogram provided predictions for BC patients' OS probabilities in the TCGA dataset. (**D**) The time-dependent c-index plot compared the performance of the nomogram with other clinical factors. (**E**) Calibration plots assessed the accuracy of the nomogram's predictions for 1-, 3-, and 5-year OS in the TCGA cohort. (**F**) Decision curve analysis evaluated the clinical utility of the nomogram and other factors for 1-, 3-, and 5-year risk assessment.

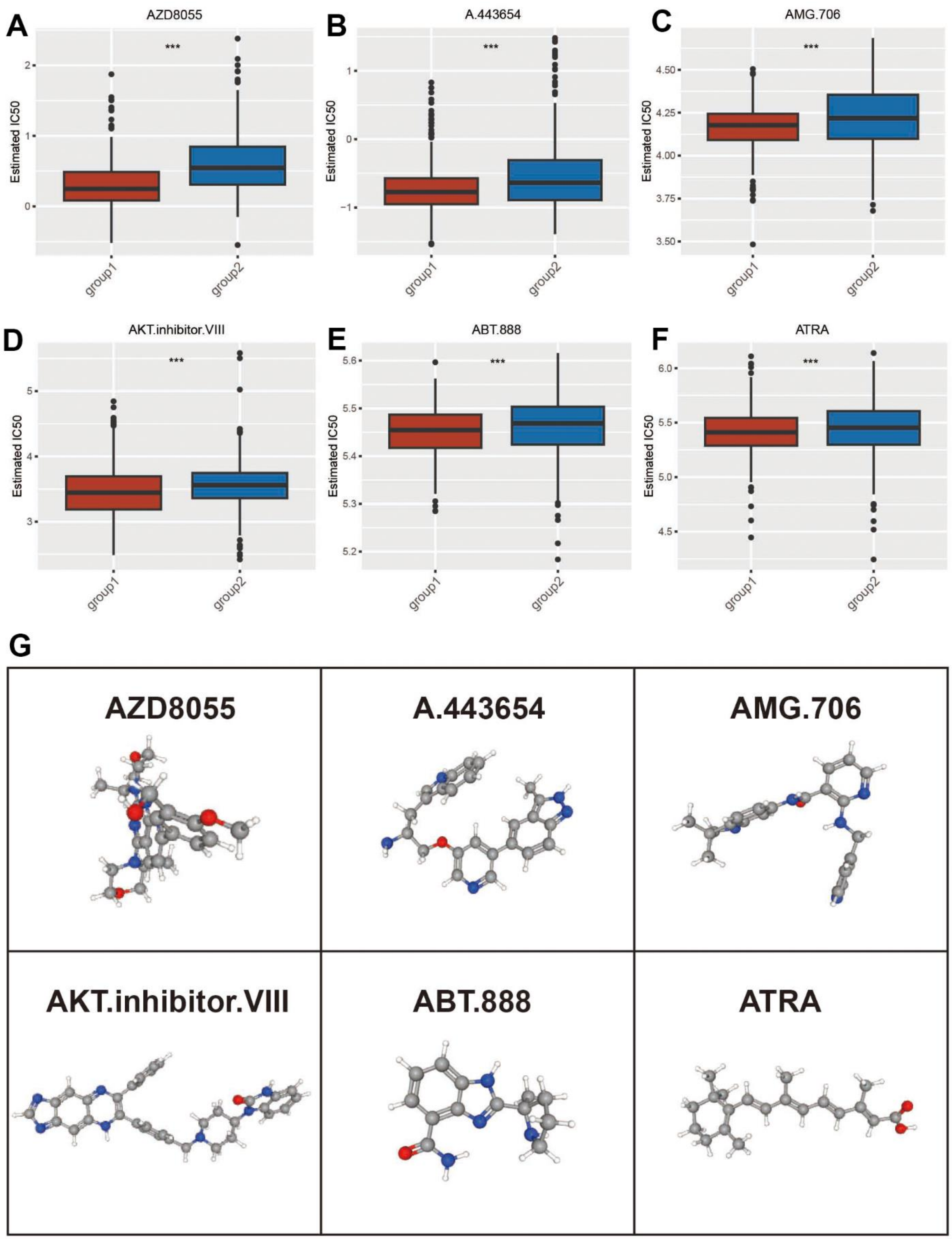

Figure 7. The ADCP-related groups were evaluated for their ability to predict therapeutic benefit. (**A**–**F**) The chemotherapy response of two ADCP-related groups to six common chemotherapy drugs was analyzed. (**G**) The 3D structure tomographs of six candidate small-molecule drugs for BC were examined. Statistical significance at the level of *** <0.001.

antigens are abundantly expressed in cancer cells and some normal cells, it is a challenge for the human immune system to produce cancer antigen-targeted antibodies in cancer patients [31]. The rise of ADCP-related research has revolutionized cancer treatment. Changes, through ADCP-related targets can obviously improve the OS rate of cancer patients [32–36].

Despite the extensive understanding of the intercellular process of phagocytosis, specific genetic changes within cancer cells that drive / expel ADCP remains have not been fully determined. Recently, Kamber et al. [37] identified a collection of previously unknown cancer-intrinsic genes that play a role in ADCP based on an unbiased CRISPR/Cas9 screen. In this investigation, Unsupervised consistent clustering was adopted to cluster gene expression patterns of Immune check-point, immune microenvironment, inflammation related molecules, immune cells as well as stromal elements derived from BC datasets and ADCP cancer-intrinsic

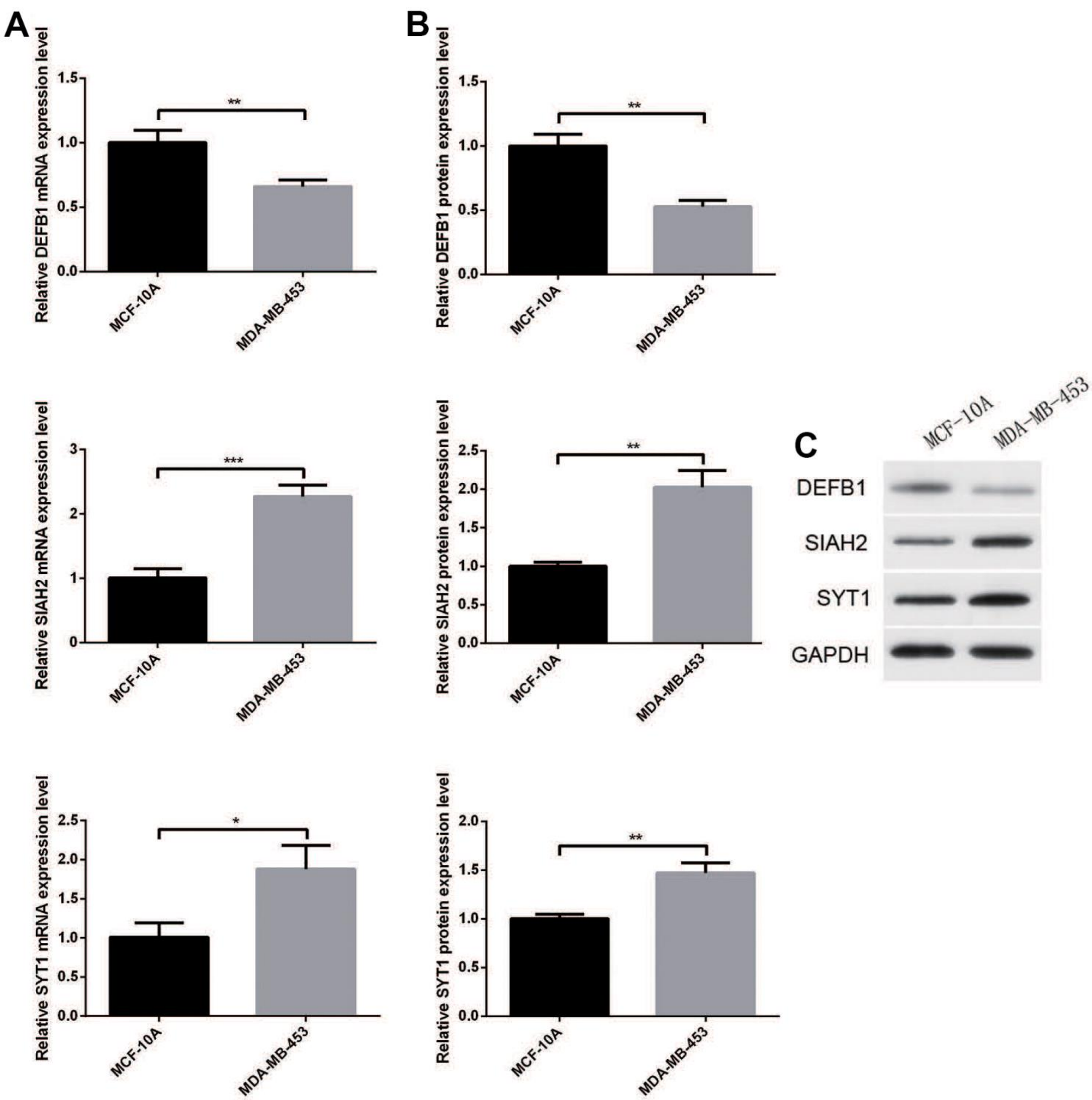

Figure 8. (**A**) RT-qPCR and (**B**, **C**) Western blotting analysis of DEFB1, SIAH2, and SYT1 expression in BC cell lines and adjacent cell lines. Statistical significance at different levels was reported for the results. Statistical significance at the level of * <0.05, ** <0.01, and *** <0.001.

genes. In this study, we successfully discovered a new ADCP-related BC category comprising 51.2% of BC patients. The ADCP-BC group, in contrast to the other group, showed lower immunological and stromal enrichment scores, suggesting a reduced presence of immune cells and stromal components. The ADCP-BC group exhibited distinct characteristics, including obvious immune cell infiltration, increased expression of multiple inflammatory genes and HLA family gene. At the same time, the three ADCP-related targets identified also showed significant differences in different groups and cancer-adjacent samples. This shows a potential advantage for them as targets for clinical treatment. Our findings have significant translational implications in the field. In addition to its prognostic value, our identified signature demonstrates a notable association with drug sensitivity, emphasizing its potential clinical relevance and applicability in therapeutic interventions. Based on these findings, it could be inferred that the ADCP group may benefit more from AZD8055, A.443654, AMG.706, AKT .inhibitor. VIII, ABT.888 and ATRA than from ICI, suggesting the combination of ZD8055, A.443654, AMG.706, AKT. inhibitor. VIII, ABT.888 and ATRA with ICI may be as potential therapeutic drug.

DEFB1 has not been specifically reported in the context of phagocytosis. However, it has received significant attention within the field of cancer research. Defensins form a family of microbicidal and cytotoxic peptides made by neutrophils. Members of the defensin family are highly similar in protein sequence. This gene encodes defensin, beta 1, an antimicrobial peptide implicated in the resistance of epithelial surfaces to microbial colonization. This gene maps in close proximity to defensin family member, defensin, alpha 1 and has been implicated in the pathogenesis of cystic fibrosis. Results of a previous study suggested that the DEFB1 gene, which encodes human ß-defensin-1 (HBD-1), plays a role in innate immune responses and may act as a potential tumor suppressor in urological cancers [38]. Furthermore, it has been observed that DNA methylation pattern within non-CpG island promoter region of DEFB1 can influence epigenetic silencing of DEFB1 in tumor cells [39]. In colorectal cancer, SIAH2 has been identified as an oncogene [40]. SIAH2 (Siah E3 Ubiquitin Protein Ligase 2) is a Protein Coding gene. Diseases associated with SIAH2 include Dyskeratosis Congenita, Autosomal Dominant 3. Among its related pathways are Class I MHC mediated antigen processing and presentation and Nervous system development. Gene Ontology (GO) annotations related to this gene include ligase activity and transcription corepressor activity. It promoted various aggressive behaviors of colorectal cancer cells, such as proliferation, migration, invasion, as well as colony formation [41]. Interestingly, high mRNA levels of SIAH2 may be correlated with elevated Estrogen Receptor (ER) mRNA levels and improved progression-free survival (PFS) after initial tamoxifen [42, 43]. Studies have shown that cytoplasmic proteins interact with SYT1 on the endoplasmic reticulum and then are spatially localized in the SEC22B + vesicles of liver cancer cells. This gene encodes a member of the synaptotagmin protein family. The synaptotagmins are integral membrane proteins of synaptic vesicles that serve as calciumsensors in the process of vesicular trafficking and exocytosis. The encoded protein participates in triggering neurotransmitter release at the synapse in response to calcium binding. Therefore, SEC22B on the vesicles is secreted on the PM by Q-SNAREs (SNAP23, SNX3 and SNX4) [44]. In addition, inhibition of the interaction between protein kinase Cδ (PKCδ) and SYT1 by PKCδ antibody can reduce the secretion and tumorigenicity of PKCδ. PKCδ is a cytoplasmic protein specifically secreted by hepatocellular carcinoma [44]. The results revealed the significance of ER-PM contact sites in facilitating the secretion of cytoplasmic protein, establishing a foundation for targeted therapies of liver cancer [45]. As mentioned above, chemotherapy drugs as well as targeted drugs are the primary therapeutic approaches for treating BC [46]. In our study, a significant increase of ADCP targets was observed in BC patients, suggesting that ADCP targets can serve as an indicator for predicting the efficacy of BC patients. Furthermore, the IC50 values of certain chemotherapeutic drugs was compared with those of targeted drugs between ADCP groups, and results showed that BC patients in the low ADCP group may present with a more favorable response to the drugs.

This study had certain limitations. The cohorts were obtained from high-throughput sequencing platforms in distinct public datasets, making the presence of intratumor/intrapatient tumor heterogeneity inevitable. Previous investigations have found that tumor heterogeneity might exert an influence on the efficacy of immunotherapy or chemotherapy. Unfortunately, owing to limitations in the available data, we were obliged to overlook the remarkable heterogeneity observed in cases of BC. Secondly, although immune interaction and survival effects associated with the inflammation pathway and ADCP-targets were observed among BC patients, their underlying biological/medical mechanisms remain obscure. Thus, it is imperative to conduct extensive large-scale prospective studies along with functional/mechanistic experiments to validate and elucidate the effect of the inflammation pathway on BC. Thirdly, the median cutoff of survival-related ADCP genes was adopted to stratify BC samples into the high and low survival-related groups. Despite that, the optimal cutoff for the survival-related ADCP genes may be a superior stratification strategy for BC patients. Lastly, due to incomplete clinicopathological information, we organized and adjusted certain clinical data for survival and Cox regression analyses, which may introduce the possibility of potential biases and uncertainties in determining the prognostic significance of the ADCP-related group.

The demand for precision medicine in cancer treatment is urgent. Our article has hints for the classification and targets of breast cancer. These targets can help the development of chemotherapy drugs in computer-aided drug design. Our research content expands the research of ADCP in the field of breast cancer, and also provides reference for other types of cancer. The research on ADCP for other cancers can learn from this research method.

CONCLUSIONS

In summary, we identified a subtype and key pathogenic factor associated with antibody-dependent cellular phagocytosis in the pathogenesis of breast cancer, accounting for about 51.2 % of BC patients, showing better therapeutic effect on tumor drug treatment and unique immune molecular characteristics of tumor immune microenvironment. Our results offered novel insights into the molecular mechanisms underlying ADCP therapy, contributing to the development of personalized immunotherapy strategies.

AUTHOR CONTRIBUTIONS

LW, RD designed the study, performed the major data analysis, and drafted the manuscript. LW, ML provided funding source, designed, oversaw, and supervised the project and edited, reviewed, and finalized the paper. All the authors were involved in experimental studies. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Ran Ding for his help with the preparation of data analysis in this paper.

CONFLICTS 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.

FUNDING

This work was supported by Tianjin Municipal Education Commission Scientific Research Project (Natural Science) (Grant Nos. 2021KJ150).

REFERENCES

1. Yu J, Song Y, Tian W. How to select IgG subclasses in developing anti-tumor therapeutic antibodies. J Hematol Oncol. 2020; 13:45. https://doi.org/10.1186/s13045-020-00876-4 PMID:32370812

2. Vankemmelbeke M, McIntosh RS, Chua JX, Kirk T, Daniels I, Patsalidou M, Moss R, Parsons T, Scott D, Harris G, Ramage JM, Spendlove I, Durrant LG. Engineering the Human Fc Region Enables Direct Cell Killing by Cancer Glycan-Targeting Antibodies without the Need for Immune Effector Cells or Complement. Cancer Res. 2020; 80:3399–412. https://doi.org/10.1158/0008-5472.CAN-19-3599 PMID:32532823

3. Cao X, Chen J, Li B, Dang J, Zhang W, Zhong X, Wang C, Raoof M, Sun Z, Yu J, Fakih MG, Feng M. Promoting antibody-dependent cellular phagocytosis for effective macrophage-based cancer immunotherapy. Sci Adv. 2022; 8:eabl9171. https://doi.org/10.1126/sciadv.abl9171 PMID:35302839

4. Musolino A, Gradishar WJ, Rugo HS, Nordstrom JL, Rock EP, Arnaldez F, Pegram MD. Role of Fcγ receptors in HER2-targeted breast cancer therapy. J Immunother Cancer. 2022; 10:e003171. https://doi.org/10.1136/jitc-2021-003171 PMID:34992090

5. Bahri M, Kailayangiri S, Vermeulen S, Galopin N, Rossig C, Paris F, Fougeray S, Birklé S. SIRPα-specific monoclonal antibody enables antibody-dependent phagocytosis of neuroblastoma cells. Cancer Immunol Immunother. 2022; 71:71–83. https://doi.org/10.1007/s00262-021-02968-7 PMID:34023958

6. Saltarella I, Desantis V, Melaccio A, Solimando AG, Lamanuzzi A, Ria R, Storlazzi CT, Mariggiò MA, Vacca A, Frassanito MA. Mechanisms of Resistance to Anti-CD38 Daratumumab in Multiple Myeloma. Cells. 2020; 9:167. https://doi.org/10.3390/cells9010167 PMID:31936617

7. Ding R, Liu Q, Yu J, Wang Y, Gao H, Kan H, Yang Y. Identification of Breast Cancer Subtypes by Integrating Genomic Analysis with the Immune Microenvironment. ACS Omega. 2023; 8:12217–31. https://doi.org/10.1021/acsomega.2c08227

PMID:37033796

8. Ding R, Wang Y, Fan J, Tian Z, Wang S, Qin X, Su W, Wang Y. Identification of immunosuppressive signature subtypes and prognostic risk signatures in triple-negative breast cancer. Front Oncol. 2023; 13:1108472. https://doi.org/10.3389/fonc.2023.1108472 PMID:37377907

9. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn). 2015; 19:A68–77. https://doi.org/10.5114/wo.2014.47136 PMID:25691825

10. Liu Y, Zhu XZ, Xiao Y, Wu SY, Zuo WJ, Yu Q, Cao AY, Li JJ, Yu KD, Liu GY, Wu J, Sun T, Cui JW, et al. Subtyping-based platform guides precision medicine for heavily pretreated metastatic triple-negative breast cancer: The FUTURE phase II umbrella clinical trial. Cell Res. 2023; 33:389–402. https://doi.org/10.1038/s41422-023-00795-2 PMID:36973538

11. Xue Y, Guo Y, Luo S, Zhou W, Xiang J, Zhu Y, Xiang Z, Shen J. Aberrantly Methylated-Differentially Expressed Genes Identify Novel Atherosclerosis Risk Subtypes. Front Genet. 2020; 11:569572. https://doi.org/10.3389/fgene.2020.569572 PMID:33381146

12. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010; 26:1572–3. https://doi.org/10.1093/bioinformatics/btq170 PMID:20427518

13. Brunson JC. ggalluvial: Layered Grammar for Alluvial Plots. J Open Source Softw. 2020; 5:2017. https://doi.org/10.21105/joss.02017 PMID:36919162

14. Ren Q, Zhang P, Zhang X, Feng Y, Li L, Lin H, Yu Y. A fibroblast-associated signature predicts prognosis and immunotherapy in esophageal squamous cell cancer. Front Immunol. 2023; 14:1199040. https://doi.org/10.3389/fimmu.2023.1199040 PMID:37313409

15. Liu S, Wang Z, Zhu R, Wang F, Cheng Y, Liu Y. Three Differential Expression Analysis Methods for RNA Sequencing: limma, EdgeR, DESeq2. J Vis Exp. 2021. https://doi.org/10.3791/62528 PMID:34605806

16. Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin Epigenetics. 2019; 11:123. https://doi.org/10.1186/s13148-019-0730-1 PMID:31443682

17. Liu CJ, Hu FF, Xia MX, Han L, Zhang Q, Guo AY. GSCALite: a web server for gene set cancer analysis. Bioinformatics. 2018; 34:3771–2. https://doi.org/10.1093/bioinformatics/bty411 PMID:29790900

18. Ito K, Murphy D. Application of ggplot2 to Pharmacometric Graphics. CPT Pharmacometrics Syst Pharmacol. 2013; 2:e79. https://doi.org/10.1038/psp.2013.56 PMID:24132163

19. Ru B, Wong CN, Tong Y, Zhong JY, Zhong SSW, Wu WC, Chu KC, Wong CY, Lau CY, Chen I, Chan NW, Zhang J. TISIDB: an integrated repository portal for tumor-immune system interactions. Bioinformatics. 2019; 35:4200–2. https://doi.org/10.1093/bioinformatics/btz210 PMID:30903160

20. Wang L, Wang D, Yang L, Zeng X, Zhang Q, Liu G, Pan Y. Cuproptosis related genes associated with Jab1 shapes tumor microenvironment and pharmacological profile in nasopharyngeal carcinoma. Front Immunol. 2022; 13:989286. https://doi.org/10.3389/fimmu.2022.989286 PMID:36618352

21. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011; 12:77. https://doi.org/10.1186/1471-2105-12-77 PMID:21414208

22. Lu L, Wang H, Fang J, Zheng J, Liu B, Xia L, Li D. Overexpression of OAS1 Is Correlated With Poor Prognosis in Pancreatic Cancer. Front Oncol. 2022; 12:944194. https://doi.org/10.3389/fonc.2022.944194 PMID:35898870

23. Liu TT, Li R, Huo C, Li JP, Yao J, Ji XL, Qu YQ. Identification of CDK2-Related Immune Forecast Model and ceRNA in Lung Adenocarcinoma, a Pan-Cancer Analysis. Front Cell Dev Biol. 2021; 9:682002. https://doi.org/10.3389/fcell.2021.682002 PMID:34409029

24. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, Ramaswamy S, Futreal PA, Haber DA, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013; 41:D955–61. https://doi.org/10.1093/nar/gks1111 PMID:23180760

25. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic

response from tumor gene expression levels. PLoS One. 2014; 9:e107468. https://doi.org/10.1371/journal.pone.0107468 PMID:25229481

26. Giorgi FM, Ceraolo C, Mercatelli D. The R Language: An Engine for Bioinformatics and Data Science. Life (Basel). 2022; 12:648. https://doi.org/10.3390/life12050648 PMID:35629316

27. Messori A. Synthetizing Published Evidence on Survival by Reconstruction of Patient-Level Data and Generation of a Multi-Trial Kaplan-Meier Curve. Cureus. 2021; 13:e19422. https://doi.org/10.7759/cureus.19422 PMID:34786276

28. Li R, Yin YH, Ji XL, Liu X, Li JP, Qu YQ. Pan-Cancer Prognostic, Immunity, Stemness, and Anticancer Drug Sensitivity Characterization of N6-Methyladenosine RNA Modification Regulators in Human Cancers. Front Mol Biosci. 2021; 8:644620. https://doi.org/10.3389/fmolb.2021.644620 PMID:34150845

29. Wu J, Zhang H, Li L, Hu M, Chen L, Xu B, Song Q. A nomogram for predicting overall survival in patients with low-grade endometrial stromal sarcoma: A population-based analysis. Cancer Commun (Lond). 2020; 40:301–12. https://doi.org/10.1002/cac2.12067 PMID:32558385

30. Kumar AR, Devan AR, Nair B, Vinod BS, Nath LR. Harnessing the immune system against cancer: current immunotherapy approaches and therapeutic targets. Mol Biol Rep. 2021; 48:8075–95. https://doi.org/10.1007/s11033-021-06752-9 PMID:34671902

31. Swain SM, Shastry M, Hamilton E. Targeting HER2-positive breast cancer: advances and future directions. Nat Rev Drug Discov. 2023; 22:101–26. https://doi.org/10.1038/s41573-022-00579-0 PMID:36344672

32. Rashid MH. Full-length recombinant antibodies from *Escherichia coli*: production, characterization, effector function (Fc) engineering, and clinical evaluation. MAbs. 2022; 14:2111748. https://doi.org/10.1080/19420862.2022.2111748 PMID:36018829

33. van de Donk NWCJ, Usmani SZ. CD38 Antibodies in Multiple Myeloma: Mechanisms of Action and Modes of Resistance. Front Immunol. 2018; 9:2134. https://doi.org/10.3389/fimmu.2018.02134 PMID:30294326

34. Vozella F, Fazio F, Lapietra G, Petrucci MT, Martinelli G, Cerchione C. Monoclonal antibodies in multiple myeloma. Panminerva Med. 2021; 63:21–7. https://doi.org/10.23736/S0031-0808.20.04149-X PMID:32955188

35. Ye D, Liu J, Zhou A, Zou Q, Li H, Fu C, Hu H, Huang J, Zhu S, Jin J, Ma L, Guo J, Xiao J, et al. Tislelizumab in Asian patients with previously treated locally advanced or metastatic urothelial carcinoma. Cancer Sci. 2021; 112:305–13. https://doi.org/10.1111/cas.14681 PMID:33047430

36. Gogesch P, Dudek S, van Zandbergen G, Waibler Z, Anzaghe M. The Role of Fc Receptors on the Effectiveness of Therapeutic Monoclonal Antibodies. Int J Mol Sci. 2021; 22:8947. https://doi.org/10.3390/ijms22168947 PMID:34445651

37. Kamber RA, Nishiga Y, Morton B, Banuelos AM, Barkal AA, Vences-Catalán F, Gu M, Fernandez D, Seoane JA, Yao D, Liu K, Lin S, Spees K, et al. Inter-cellular CRISPR screens reveal regulators of cancer cell phagocytosis. Nature. 2021; 597:549–54. https://doi.org/10.1038/s41586-021-03879-4 PMID:34497417

38. Li S, Li H, Xu Y, Ning W, Hu S, Wei S, Song H, Sun J, Ziebolz D, Schmalz G, Hu X, Liu M. Implications of Human Antimicrobial Peptide Defensin Beta-1 in Clinical Oral Squamous Cell Carcinoma Patients via an Integrated Bioinformatics Approach. Comput Math Methods Med. 2022; 2022:2203615. https://doi.org/10.1155/2022/2203615 PMID:35222682

39. Lee J, Han JH, Jang A, Kim JW, Hong SA, Myung SC. DNA Methylation-Mediated Downregulation of DEFB1 in Prostate Cancer Cells. PLoS One. 2016; 11:e0166664. https://doi.org/10.1371/journal.pone.0166664 PMID:27835705

40. Hu Y, He Y, Liu W, Yu S, Wei Y, Bai S, Su Y, Xiao B. SIAH2 regulates colorectal cancer tumorigenesis via PI3K/ATK signaling pathway. Tissue Cell. 2022; 78:101878. https://doi.org/10.1016/j.tice.2022.101878 PMID:35926257

41. Liu Q, Luo Q, Feng J, Zhao Y, Ma B, Cheng H, Zhao T, Lei H, Mu C, Chen L, Meng Y, Zhang J, Long Y, et al. Hypoxia-induced proteasomal degradation of DBC1 by SIAH2 in breast cancer progression. Elife. 2022; 11:e81247. https://doi.org/10.7554/eLife.81247 PMID:35913115

42. Fan L, Peng G, Hussain A, Fazli L, Guns E, Gleave M, Qi J. The Steroidogenic Enzyme AKR1C3 Regulates Stability of the Ubiquitin Ligase Siah2 in Prostate Cancer Cells. J Biol Chem. 2015; 290:20865–79. https://doi.org/10.1074/jbc.M115.662155 PMID:26160177

43. Interiano RB, Yang J, Harris AL, Davidoff AM. Seven In

Absentia Homolog 2 (SIAH2) downregulation is associated with tamoxifen resistance in MCF-7 breast cancer cells. J Surg Res. 2014; 190:203–9. https://doi.org/10.1016/j.jss.2014.02.018 PMID:24656476

44. Yamada K, Motohashi S, Oikawa T, Tago N, Koizumi R, Ono M, Tachibana T, Yoshida A, Yoshida S, Shimoda M, Oka M, Yoneda Y, Yoshida K. Extended-synaptotagmin 1 engages in unconventional protein secretion mediated via $SEC22B^{+}$ vesicle pathway in liver cancer. Proc Natl Acad Sci U S A. 2022; 119:e2202730119. https://doi.org/10.1073/pnas.2202730119 PMID:36044553

45. Pan T, Liu J, Xu S, Yu Q, Wang H, Sun H, Wu J, Zhu Y, Zhou J, Zhu Y. ANKRD22, a novel tumor microenvironment-induced mitochondrial protein promotes metabolic reprogramming of colorectal cancer cells. Theranostics. 2020; 10:516–36. https://doi.org/10.7150/thno.37472 PMID:31903135

46. Dandawate PR, Subramaniam D, Jensen RA, Anant S. Targeting cancer stem cells and signaling pathways by phytochemicals: Novel approach for breast cancer therapy. Semin Cancer Biol. 2016; 40–41:192–208. https://doi.org/10.1016/j.semcancer.2016.09.001 PMID:27609747

SUPPLEMENTARY MATERIALS

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1, 2.

Supplementary Table 1. Results of single factor regression analysis.

Supplementary Table 2. Results of gene differential expression analysis.