Cell. Mol. Biol. (ISSN: 1165-158X)

Journal Homepage: www.cellmolbiol.org

Cellular and Molecular Biology

Original Article

Identification of a novel prognostic signature for breast cancer derived from post-translational ubiquitin and ubiquitin-like modification-related genes

Nanyang Zhou[1], Dejia Kong[1], Qiao Lin[1], Xiaojing Yang[2], Dan Zhou[3], Lihua Lou[4], Xiangming Lou[5*]

[1] *Department of Traditional Chinese Medicine, Hangzhou Women's Hospital, Hangzhou 310008, Zhejiang, China*
[2] *Department of Reproductive Center, Hangzhou Women's Hospital, Hangzhou 310008, Zhejiang, China*
[3] *Department of Breast and Thyroid Surgery, Traditional Chinese Medical Hospital of Zhuji 311800 Zhejiang, China*
[4] *Department of Breast Surgery, Zhejiang Provincial Hospital of Chinese Medicine Hangzhou 310006, Zhejiang, China*
[5] *Gynecology Clinic, Hangzhou Women's Hospital, Hangzhou 310008, Zhejiang, China*

Article Info

Article history:

Received: July 31, 2024
Accepted: October 02, 2024
Published: October 31, 2024

Use your device to scan and read the article online

Abstract

Ubiquitin and ubiquitin-like (UUL) modifications play pleiotropic functions and are subject to fine regulatory mechanisms frequently altered in cancer. However, the comprehensive impact of UUL modification on breast cancer remains unclear. Transcriptomic and clinical data of breast cancer were downloaded from TCGA and GEO databases. Molecular subtyping of breast cancer was conducted using the NMF and CIBERSORT algorithms. Prognostic genes were identified via univariate, lasso and multivariate Cox regression analyses. Clinical pathological features, immune cell infiltration, immune therapeutic response and chemotherapy drug sensitivity were compared between groups using the Wilcoxon test. Survival analysis was performed using the Kaplan-Meier method and log-rank test. A total of 63 UUL modification-related genes were differentially expressed, with 29 up-regulated and 34 down-regulated genes. These genes were used to generate two UUL modification patterns that exhibited significant differences in prognostic features and immune cell infiltration. The UUL modification patterns were associated with 2038 differentially expressed genes that were significantly enriched in nuclear division, chromosome segregation, neuroactive ligand-receptor interaction, cell cycle, and other biological processes. Of these genes, 425 were associated with breast cancer prognosis, which enabled the classification of breast cancer into two clusters with significantly distinct prognoses. We developed a prognostic model, UULscore, which comprised nine genes and showed a significant correlation with partial immune cell infiltration. Furthermore, UULscore demonstrated potential predictive value in breast cancer overall survival prediction, immune therapeutic response, and chemotherapy drug sensitivity. UULscore, stage, radiotherapy, and chemotherapy were identified as independent prognostic factors for breast cancer. Based on these factors, a nomogram model was constructed, which demonstrated exceptional prognostic predictive performance. The present study identified two UUL modification-derived molecular subtypes in breast cancer, and have successfully constructed a risk-scoring model that holds potential value in prognosis, immune infiltration, immune therapeutic response, and chemotherapy sensitivity.

Keywords: Ubiquitin, Breast cancer, Tumor microenvironment, Prognosis, Immunotherapy.

1.Introduction

Breast cancer (BC) is the most prevalent malignancy affecting women globally, and has emerged as a leading cause of cancer-related mortality in women [1]. This disease is heterogeneous, with unique molecular, biological, and clinical features. Despite significant strides in prevention, early detection, and select therapies, BC remains challenging to treat due to concerns such as recurrence, distant metastasis, and chemotherapy resistance [2]. Therefore, it is critical to investigate the characteristics of BC patients and explore potential therapeutic targets. Various treatment regimens have been developed, including surgery, chemotherapy, endocrine therapy, and anti-HER2 targeted therapy [3]. However, the extensive tumor heterogeneity limits the broad applicability of standard therapy. Consequently, the generation of reliable prognostic prediction and treatment response evaluation tools is of paramount importance.

Ubiquitin and ubiquitin-like (UUL) modifications are essential cellular mechanisms for protein modification. Ubiquitination involves the covalent attachment of ubiquitin molecules to target proteins, resulting in their tagging for degradation or transport [4]. Conversely, ubiquitin-like modifications entail the attachment of small molecule modifications to target proteins, such as SUMOylation and NEDDylation. These modifications play critical roles in regulating various biological processes, including cell cycle progression, signal transduction, and gene expression [5,6]. Aberrant regulation of UUL modifications is closely linked to cancer initiation, progression,

* Corresponding author.
E-mail address: louxiangming2008@163.com (X. Lou).
Doi: http://dx.doi.org/10.14715/cmb/2024.70.10.25

and treatment resistance [7]. Hence, research into these modifications not only advances our understanding of cellular biology but also provides novel avenues for cancer treatment and targeted therapies. Increasing evidence suggests that transcriptomic features based on UUL modifications can identify cancer heterogeneity and hold potential predictive value in cancer prognosis and immune therapeutic response [8,9]. Recent studies have systematically analyzed mutations in UUL-related genes, identifying valuable gene mutations with potential for BC drug development [10]. However, research on UUL-modified gene-related BC molecular subtyping and prognostic value is still in its infancy, which limits the development of UUL-targeted drugs.

Our study presents a comprehensive investigation into the expression profiles of UUL-modification-related genes in BC. Based on these profiles, we identified distinct molecular subtypes that exhibit significant prognostic differences. We further developed a prognostic risk model and nomogram that utilize the UUL modification pattern to predict clinical outcomes, tumor immune microenvironment, and response to immunotherapy and chemotherapy. This article was previously posted to the medRxiv preprint server on April 27, 2020 (DOI:10.22541/au.169454707.71853863/v1).

2. Material and methods

2.1. Data collection

Transcriptome data and clinical information from the TCGA-BRCA cohort were downloaded from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) database and used as the training set to construct the BC prognostic risk model. Data from the GSE96058 cohort were downloaded from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) database for model validation. A total of 879 BC patients were included in the TCGA-BRCA cohort after excluding samples with follow-up times of less than 30 days and incomplete clinical information, including age, clinical stage, TNM stage, radiotherapy and chemotherapy records. The GSE96058 cohort included data from 3,409 BC patients for model validation. In addition, a total of 560 UUL-modification-related genes were retrieved from the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb) database using key terms such as ubiquitination, SUMOylation, and Neddylation (Supplementary materials: Table S1).

2.2. UUL molecular subtyping

Differential expression analysis was performed using the limma package to identify UUL modification-related genes that show significant differential expression in BC, with a false discovery rate (FDR) < 0.05 and absolute log fold change (|logFC|) > 1 as cutoffs. To identify the characteristic expression patterns of differentially expressed UUL modification-related genes in BC, the non-negative matrix factorization (NMF) algorithm was utilized with the NMF package, setting the method to 'brunet' and maxIter to 500. Kaplan-Meier (K-M) survival analysis and log-rank test were used to evaluate the overall survival differences among different UUL patterns. In addition, principal component analysis (PCA) was used to test the discriminative ability of UUL patterns and to evaluate the immune infiltrate between different UUL modification patterns.

2.3. Consensus clustering

Differential expression analysis was performed using the limma package to identify genes that show significant differential expression between different UUL-modified patterns in BC, with an FDR < 0.05 and |logFC| > 1 as cutoffs. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of differentially expressed genes (DEGs) was performed using the clusterprofiler package, and the results were visualized using a bubble plot. To identify DEGs associated with BC prognosis, univariate Cox regression analysis was conducted. Prognostic DEGs with $p < 0.05$ were selected, and consensus clustering analysis was subsequently performed using the ConsensusClusterPlus package. To identify the optimal clustering threshold, we set the maximum number of clusters to six, employed the partitioning around medoids (PAM) clustering algorithm, and utilized the proportion of ambiguous clustering (PAC) method. K-M survival analysis and log-rank test were used to evaluate the overall survival differences among different clusters derived from different UUL modification patterns.

2.4. Construction and evaluation of the UUL risk scoring model

To refine the prognostic DEGs and prevent overfitting, Lasso Cox regression analysis was performed using the glmnet package and genes with non-zero coefficients were selected. To determine a robust set of independent prognostic genes, a backward stepwise elimination process was implemented using the coxph function from the survival package. Once the iterative process stabilized, a risk score was calculated for each patient by applying a weighted sum of the gene expression values using the coefficients from the final multivariate Cox regression model. The UUL risk score was calculated using the following formula: UULscore = $\sum_{i=1}^{n} coef_i * exp(gene)_i$. The UULscore was calculated for each patient in the TCGA-BRCA and GSE96058 cohorts, and the cohorts were divided into high_UULscore and low_UULscore groups using the median value. Kaplan-Meier survival analysis and log-rank test were used to evaluate the overall survival differences among different UUL patterns. Receiver operating characteristic (ROC) curve analysis was used to validate the effectiveness of the risk model.

2.5. Immune infiltration, immunotherapy response, and chemotherapy sensitivity analysis

The CIBERSORT algorithm was used to evaluate the immune cell infiltration status of each patient and to compare it between different groups. CIBERSORT package is commonly used to calculate the infiltration of 22 immune cell types based on transcriptome data [11]. Two methods were used to evaluate the response to immunotherapy: tumor immune dysfunction and exclusion (TIDE) score and immunophenoscore (IPS). The TIDE score was obtained by normalizing the transcriptome data and inputting it into the TIDE website (http://tide.dfci.harvard.edu/) to compare the scores between different groups. The IPS score was calculated using The Cancer Immunome Atlas (TCIA, https://tcia.at/home) database. Chemotherapy drug sensitivity was evaluated using the pRRophetic package, which analyzed the sensitivity of commonly used chemotherapy drugs in TCGA-BRCA and compared it between different groups.

2.6. Nomogram construction and evaluation

Univariate and multivariate Cox regression analyses were performed to identify independent prognostic factors for BC patients by analyzing the relationship between the UULscore, other clinical and pathological features, and BC patient prognosis. The rms package was used to construct a nomogram to predict 1-, 3-, and 5-year overall survival in BC patients, and the consistency index (c-index) was used to evaluate the accuracy of the nomogram. The calibration curve and receiver operating characteristic (ROC) curves of the nomogram was analyzed for performance correction, and the rmda package was used to analyze the decision curve of different prognostic strategies for predicting overall survival in BC patients.

2.7. Statistical analysis

All statistical analyses and result visualizations were performed using R software version 4.2.0. Wilcoxon test was used to compare the differences between groups, with $p < 0.05$ indicating statistical significance.

3.Results

3.1. UUL modification-related genes derived two distinct molecular subtypes of BC

The workflow diagram is illustrated in Figure 1. Utilizing differential expression analysis of BC and its adjacent tissues, aberrant expression of 63 UUL-related genes was observed in BC, with 29 genes exhibiting abnormally high expression and 34 genes displaying abnormally low expression (Figure 2A, Supplementary materials: Table S2). To investigate the interaction between these differentially expressed UUL-related genes and tumor characteristics, we employed NMF clustering with k = 2 as the optimal value of k, determined via comprehensive correlation coefficient calculation, to classify patients in the element queue into two UUL modification patterns, termed UULpattern A and UULpattern B (Figure 2B, C). Principal component analysis enabled an easy differentiation of the two UUL modification patterns, where transcriptional profiles of UUL-related genes showed significant differences between the two different patterns (Figure 2D). In Pattern A, expression levels of UUL modification-related genes were higher than those in UULpattern B, indicating an enhancement of UUL modification in UULpattern A (Figure 2E). Clinical characteristics of the two UUL modification patterns exhibited no significant differences in gender, age, TNM category, chemotherapy, and clinical stage, as shown in Figure 2E. Survival analysis showed that prognosis of UULpattern A was better than that of UULpattern B (Figure 2F). Additionally, immune infiltration was analyzed between different patterns, where results demonstrated differential infiltration of 10 immune cells, including naive B cells, resting CD4+ memory T cells, monocytes, resting mast cells, and resting dendritic cells, significantly enriched in UULpattern A, while activated memory CD4+ T cells, follicular helper T cells, regulatory T cells (Treg), resting NK cells, M0 and M1 macrophages were significantly enriched in UULpattern B (Figure 2G). These findings imply an association between UUL modification and immune cell infiltration.

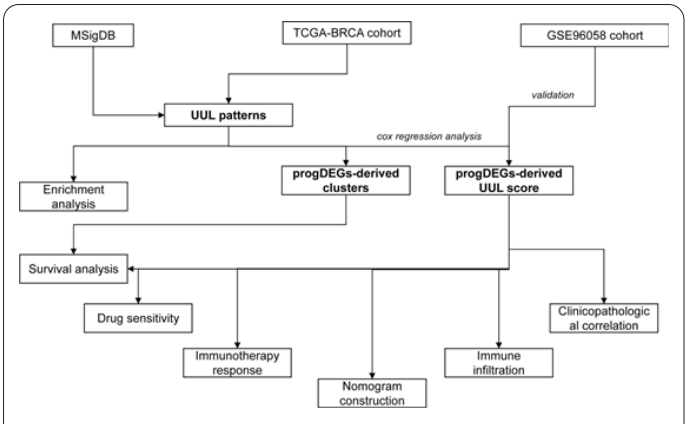

Fig. 1. Workflow for constructing UUL modification-derived molecular subtypes of BC and the prognosis risk model.

3.2. BC clusters derived from UUL modification patterns

To investigate the underlying biological distinctions among distinct UUL modification patterns, we identified 2038 genes with differential expression and subsequently performed GO and KEGG enrichment analyses. Our results, as depicted in Figure 3A and B, indicated significant enrichment of these genes in biological processes such as nuclear division and chromosome segregation, as well as in cell pathways such as neuroactive ligand-receptor interaction, cell cycle, and cytokine-cytokine receptor. Univariate Cox regression analysis revealed that 425 of these genes were closely associated with the prognosis

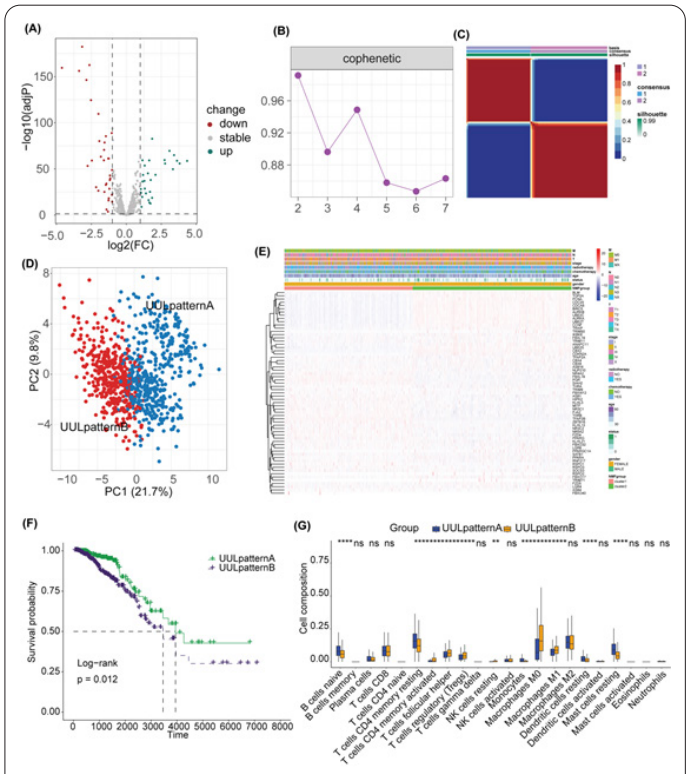

Fig. 2. Molecular subtyping of BC based on UUL modification-related genes. A: Volcano plot of differentially expressed UUL modification-related genes in BC. B: Composite correlation coefficient obtained from NMF subtyping based on 63 differentially expressed UUL modification-related genes. C: Connectivity matrix of BC patients in the meta-cohort analyzed by NMF when k=2. D: Principal component analysis of transcriptome profiles of two UUL modification patterns, showing significant differences in transcriptome between the two patterns. E: Heatmap showing the relationship between clinical features and UUL modification patterns. F: Kaplan-Meier survival analysis revealed significant differences in prognosis between the two UUL modification patterns. G: Analysis and comparison of 22 immune cell infiltrations between the two UUL modification patterns.

of BC (Supplementary materials: Table S3). The consensus clustering of these DEGs related to prognosis yielded two highly consistent subtypes of BC, UULclusterA and UULclusterB (Figure 3C). Survival analysis demonstrated a significant difference in prognosis between the two subtypes, with UULclusterA displaying significantly higher overall survival than UULclusterB (Figure 3D).

3.3. Construction and evaluation of the UUL scoring model

To develop a risk-scoring model based on UUL modification, we performed lasso and multivariate Cox regression analysis on the aforementioned prognostic-related DEGs. As depicted in Figure 4A-B, lasso reduced the number of prognostic-related genes from 425 to 27. Subsequent multivariate Cox regression analysis revealed that 9 of these genes were independent prognostic factors for BC (Figure 4C). As a result, we constructed a UUL modification-related scoring model: UULscore = -0.1704988 * apolipoprotein D (*APOD*) + 1.3180226 * calcitonin related polypeptide beta (*CALCB*) - 1.3242989 * cadherin related family member 4 (*CDHR4*) + 1.4503083 * MAF bZIP transcription factor A (*MAFA*) + 0.9647131 * neuroligin 1 (*NLGN1*) - 0.3708044 * Purkinje cell protein 2 (*PCP2*) + 0.9165657 * SLIT and NTRK like family member 3 (*SLITRK3*) + 0.2776859S * speedy/RINGO cell cycle regulator family member C (*SPDYC*) + 0.2978558 * serine peptidase inhibitor Kazal type 8 (*SPINK8*). Using the median, the TCGA-BRCA cohort was divided into high_UULscore and low_UULscore groups (Figure 4D). Survival analysis demonstrated that the low_UULscore group had a better prognosis than the high_UULscore group (Figure 4E, p<0.0001). ROC analysis revealed that the AUC values of UULscore in predicting 1-, 3-, and 5-year overall survival of BC were 0.814, 0.741, and 0.766, respectively (Figure 4F). The GSE96058 cohort was also divided into high_UULscore and low_UULscore groups, exhibiting

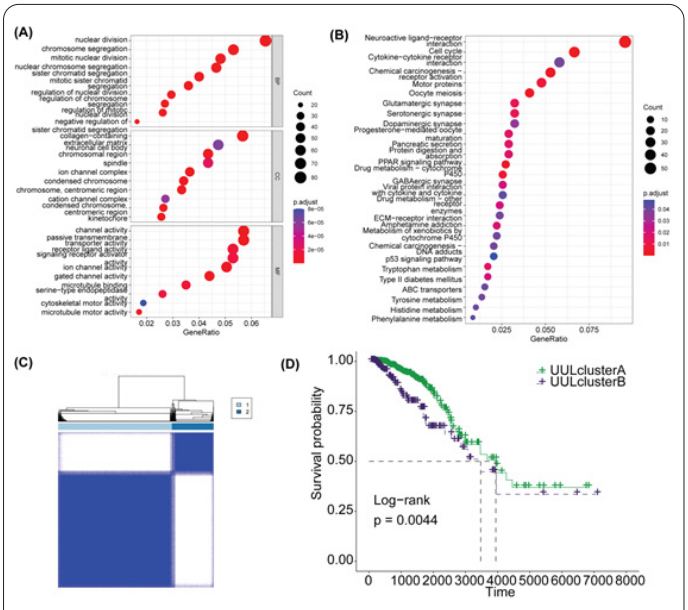

Fig. 3. Enrichment analysis and clustering analysis of DEGs between UUL modification patterns. A: GO enrichment bubble plot for DEGs between UUL modification patterns. B: KEGG pathway enrichment bubble plot for DEGs between UUL modification patterns. C: Consensus clustering analysis based on prognostic DEGs identified two optimal and consistent groups. D: Kaplan-Meier survival analysis revealed significant differences in prognosis between the two groups.

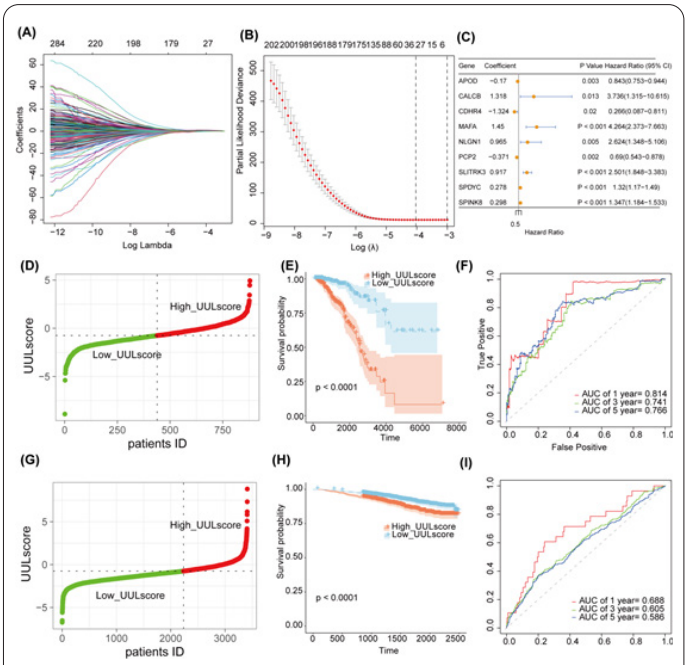

Fig. 4. Construction and evaluation of UUL modification-derived risk score model. A: LASSO regression analysis of 425 prognostic DEGs. B: Cross-validation method was used to select optimal genes. C: Forest plot of multivariate Cox regression analysis. D: Distribution of UULscore values in TCGA-BRCA cohort, divided into high_UULscore and low_UULscore groups based on the median value. E: Kaplan-Meier survival analysis showed significant differences in prognosis between the high_UULscore and low_UULscore groups in the TCGA-BRCA cohort. F: The 3-, 5- and 10-year ROC curves of the UULscore in the TCGA-BRCA cohort. G: Distribution of UULscore values in the GSE96058 cohort, divided into high_UULscore and low_UULscore groups. H: Kaplan-Meier survival analysis showed significant differences in prognosis between the high_UULscore and low_UULscore groups in the GSE96058 cohort. I: The 3-, 5- and 10-year ROC curves of the UULscore in the GSE96058 cohort.

the same prognostic features (Figure 4G,H). ROC analysis showed that the AUC values of UULscore in predicting 1-, 3-, and 5-year overall survival of BC were 0.688, 0.605, and 0.586, respectively (Figure 4I).

3.4. Expression and K-M survival curves of model-related genes

Figure 5A presents a heatmap of the expression levels of the model-related genes, showing that *CALCB*, *MAFA*, *SPDYC*, *NLGN1*, *SLITRK3*, and *SPINK8* are highly expressed in the high-risk group, whereas *APOD*, *CDHR4*, and *PCP2* are lowly expressed in the same group. Further, we divided the samples into high-expression and low-expression groups based on the median values of these genes' expressions and conducted K-M survival curve analyses. The results indicated that patients with high expression of *CDHR4* and *PCP2* have significantly better overall survival compared to those with low expression, whereas high expression of SPDYC is associated with significantly poorer overall survival (Figure 5B).

3.5. UULscore predicts immunotherapy response and chemotherapy sensitivity

Figure 6A and 6B depict the relationship between UUL modification-based molecular subtypes. UULclusterA was observed to almost entirely comprise UULpatternA, whereas the majority of patients in the Low_UULscore

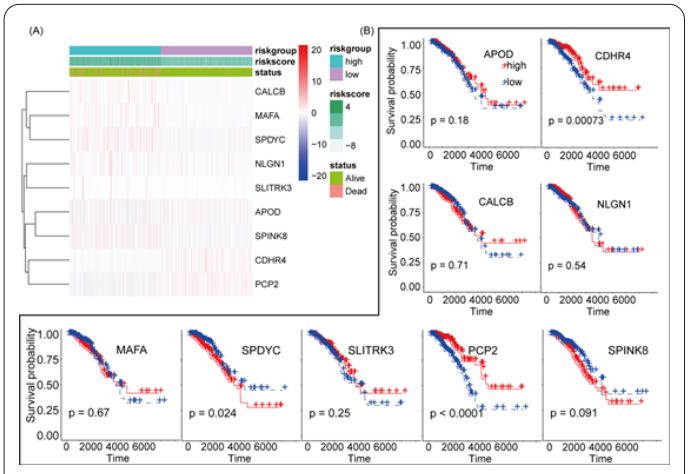

Fig. 5. Expression Heatmap and K-M Survival Analysis of Model-Related Geness. A: Heatmap of the expression levels of the model-related genes. B: Kaplan-Meier survival curve analysis of the model-related genes.

group were categorized under the UULclusterB subtype. Additionally, UULpatternB and UULclusterB exhibited significantly higher UULscores relative to their respective counterparts, UULpatternA and UULclusterA (Figure 6C and 6D). Subsequently, we conducted an analysis of immune infiltration status across different UULscore groups, which revealed that the high_UULscore group exhibited higher levels of infiltration by activated memory CD4+ T cells, follicular helper T cells, resting NK cells, M0 and M1 macrophages, and activated dendritic cells. Conversely, the low_UULscore group exhibited higher levels of infiltration by naive B cells, monocytes, resting dendritic cells, and resting mast cells (Figure 6E). Correlation analysis demonstrated a positive correlation between UULscore and resting NK cells, M0 macrophages, and activated CD4+ memory T cells, and a negative correlation with naive B cells and resting mast cells (Figure 6F). Drug sensitivity analysis indicated that patients in the low_UULscore group exhibited lower sensitivity to metformin, while those in the high_UULscore group exhibited lower sensitivity to cisplatin and paclitaxel (Figure 6G-I). Analysis of immune response revealed that TIDE scores were higher in the low_UULscore group relative to the high_UULscore group (Figure 6J). Furthermore, in cases where CTLA4 and PD1 were both negative, IPS scores were higher in the low_UULscore group than in the high_UULscore group (Figure 6K). These findings underscore the critical role of UULscore in assessing the tumor immune microenvironment, immunotherapy response, and drug sensitivity.

3.6. Clinical-pathological correlation and nomogram model construction of UULscore

In order to enhance the clinical applicability of UULscore, we first analyzed its relationship with clinical-pathological features (Figure 7A-D). The results demonstrated that UULscore was associated with clinical stage and prognosis, with higher UULscore scores observed in dead patients and later clinical stages. Furthermore, UULscore scores were found to be independent of radiotherapy or chemotherapy. Multivariate Cox regression analysis revealed that UULscore, stage, radiotherapy, and chemotherapy were independent prognostic factors for BC (Figure 7E). Consequently, we developed a nomogram based on these independent prognostic factors for clinical BC prognosis assessment (Figure 7F). Calibration curve analysis demonstrated the predictive performance of the nomogram for 1-, 3-, and 5-year overall survival prediction (Figure 7G), while decision curve analysis demonstrated that the performance of the nomogram in BC prognosis assessment was superior to that of other individual prognostic factors (Figure 7H). The ROC curves showed that the nomogram predicted 1-, 3-, and 5-year overall survival probabilities for BC patients with area under the curve (AUC) values of 0.849, 0.82, and 0.798, respectively (Figure 7I). These results indicated that the nomogram exhibits excellent performance in predicting the prognosis of BC patients.

4. Discussion

Ubiquitination and sumoylation play important roles in cancer pathogenesis. These modifications regulate diverse biological processes, such as protein stability, interaction, transport, and signal transduction, thereby influencing cellular growth, differentiation, and apoptosis [4]. Dysregulation of ubiquitination and sumoylation modifications in cancer leads to abnormal biological features, including aberrant cell proliferation, apoptosis, and metastasis, which

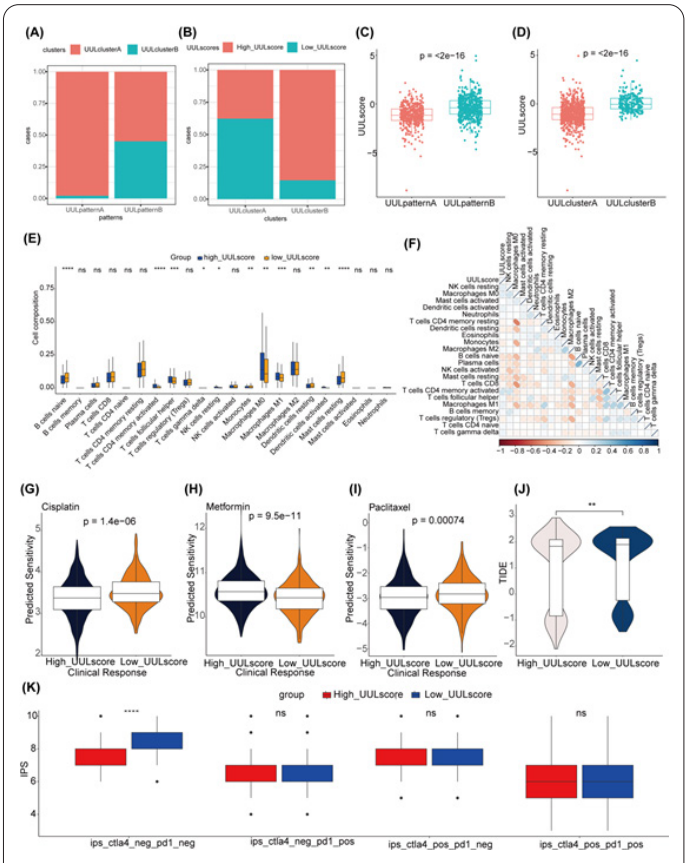

Fig. 6. Correlation analysis between UULscore and immune cell infiltration, immune therapy response, and chemotherapy sensitivity based on the TCGA-BRCA cohort. A: Distribution of UULclusters patients in UULpatterns. B: Distribution of High_UULscore and Low_UULscore group patients in UULclusters. C: Comparison of UULscores values in different UULpatterns. D: Comparison of UULscore values in different UULclusters. E: Evaluation and comparison of 22 immune cell infiltrations in High_UULscore and Low_UULscore groups. F: Correlation analysis between UULscore and immune cell infiltration. G: Drug sensitivity analysis of cisplatin, H: metformin, and I: paclitaxel in High_UULscore and Low_UULscore groups. J: Comparison of TIDE scores between High_UULscore and Low_UULscore groups. K: Comparison of IPS scores between High_UULscore and Low_UULscore groups.

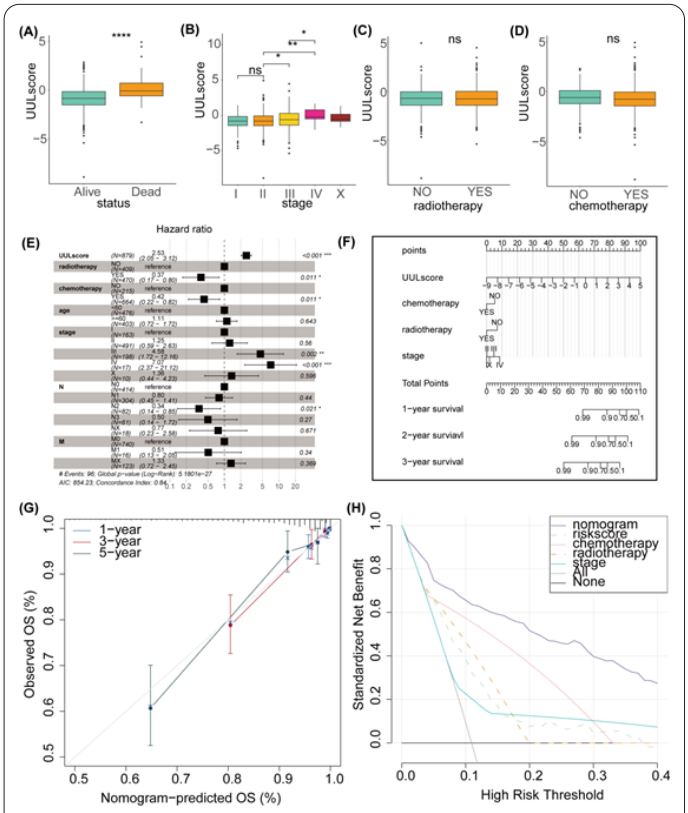

Fig. 7. Clinical-pathological correlation of UULscore and construction and evaluation of the nomogram. A: Comparison of UULscore among BC patients with different survival outcomes. B: Comparison of UULscore among BC patients with different clinical stages. C: Comparison of UULscore between BC patients who received radiotherapy and those who did not. D: Comparison of UULscore between BC patients who received chemotherapy and those who did not. E: Forest plot of multivariate Cox regression analysis between UULscore and clinical-pathological features. F: Construction of a nomogram based on UULscore, stage, radiotherapy, and chemotherapy for predicting 1-, 3-, and 5-year overall survival of breast cancer patients. G: Calibration curves of the nomogram for 1-, 3-, and 5-year overall survival prediction in breast cancer patients. H: Performance comparison of the nomogram and other strategies for predicting the prognosis of breast cancer patients. I: ROC curve analysis evaluated the performance of the nomogram in predicting 1-, 3-, and 5-year overall survival for breast cancer patients.

promote cancer initiation and progression [4]. Substantial research has been conducted to investigate the prognostic and therapeutic response prediction of ubiquitination and sumoylation-related genes in cancer. For example, neural precursor cell expressed, developmentally down-regulated 9 (*NEDD9*), an important ubiquitination-related gene, is associated with poor prognosis and therapy resistance in breast cancer [12]. Furthermore, studies indicate that the expression levels of sumoylation-related genes, such as small ubiquitin like modifier 1 (*SUMO1*), are related to prognosis and treatment response in several cancer types, including lung [13], and colorectal cancer [14]. Additionally, ubiquitination or sumoylation genes may have potential value in predicting prognosis and treatment response in cancer molecular subtyping [8,9]. Our study focused on a systematic bioinformatics analysis of ubiquitination and sumoylation-related genes in breast cancer to elucidate their transcriptomic features and potential value in predicting prognosis, immune infiltration, immunotherapy response, and chemotherapy drug sensitivity.

Initially, 63 UUL modification-associated genes exhibiting differential expression in breast cancer were identified, suggesting a dysregulated state of UUL modification in this malignancy. Several of these genes have been shown to participate in the initiation and progression of breast cancer. For example, high expression of ubiquitin conjugating enzyme E2 C (*UBE2C*) was associated with poor prognosis in breast cancer [15], influencing tumor proliferation through the AKT/mTOR signaling pathway [16]. F-box protein 32 (*FBXO32*) exhibited tumor-suppressive activity in breast cancer by directly targeting KLF transcription factor 4 (*KLF4*), leading to its ubiquitination and degradation, thereby inhibiting colony formation in vitro and primary tumor development and growth in vivo [17]. Tripartite motif containing 9 (*TRIM9*)-mediated ubiquitination of pyruvate kinase M1/2 (*PKM2*) driven aerobic glycolysis, promoting the progression of triple-negative breast cancer [18]. R-spondin 2 (*RSPO2*) mediated Wnt signaling, playing a critical role in breast tumor progression [19]. Tripartite motif containing 11 (*TRIM11*) acted as a regulator of ERα, enhancing its stability through monoubiquitination, which promotes breast cancer cell proliferation [20]; it also facilitated breast cancer cell proliferation by modulating glycolytic metabolism and the AKT/GLUT1 signaling pathway [21]. The chromobox 4 (*CBX4*) regulated *hTERT*-mediated transcription of cadherin 1 (*CDH1*), promoting migration and invasion of breast cancer cells [22]. Overexpression of ubiquitin conjugating enzyme E2 S (*UBE2S*) in breast cancer reduced the levels of Numb and enhances cellular malignancy [23]. In breast cancer, cell division cycle 20 (*CDC20*) bound to and promoted the proteasomal degradation of *BTG3*-associated nuclear protein in a D-box motif-dependent manner, facilitating cell migration and invasion [24].

Utilizing these differentially expressed UUL modification-associated genes, molecular subtyping of breast cancer was achieved, thereby elucidating the background molecular characteristics and mechanisms of the pronounced heterogeneity of this disease. Survival analysis demonstrated significant differences in prognostic features between molecular subtypes based on UUL modification, which may be attributed to alterations in immune cell infiltration. Subtype analysis revealed significant enrichment of naive B cells, resting CD4+ memory T cells, monocytes, resting mast cells, and resting dendritic cells in molecular subtypes with favorable prognoses, while activated memory CD4+ T cells, Tfh cells, Treg, resting NK cells, M0 and M1 macrophages were significantly enriched in subtypes with unfavorable prognoses. Tumor-associated macrophages (TAMs) have been shown to be associated with adverse prognoses in breast cancer and resistance [25]. Follicular helper T cells have been reported to have either positive or negative prognostic effects in various human cancers. Recent research has established the critical role of CD8+/Tfh crosstalk in the anti-tumor immune response generated by immunotherapy [26], which could account for the relationship between UUL modification-derived UULscore and immune response. Treg cells are implicated in poor prognoses in breast cancer due to their ability to regulate the immune response of CD8+ T cells and NK cells by secreting immunosuppressive cytokines and inhibiting APC maturation [27]. Dendritic cells have been demonstrated to be associated with breast cancer prognosis and survival, with patients exhibiting high dendritic

cell counts having longer progression-free survival than those with low infiltrative lesions [28]. To further explore the genes and mechanisms underlying the significant prognostic differences between molecular subtypes, differential gene identification and enrichment analysis were conducted. The analysis identified 2,038 genes with differential expression, which were significantly associated with nuclear division, chromosome segregation, and cell cycle regulation.

In addition, a prognostic correlation analysis was conducted on DEGs between molecular subtypes derived from UUL modification, and a breast cancer risk score model, UULscore, was constructed. This model comprises nine genes, including three protective genes (*APOD*, *CDHR4*, and *PCP2*) and six risk genes (*CALCB*, *MAFA*, *NLGN1*, *SLITRK3*, *SPDYC*, and *SPINK8*). The prognostic value of *APOD* has been demonstrated in various cancers, including cervical cancer [29] and prostate cancer [30]. Recent studies revealed that ApoD-positive patients have better overall survival, and its expression can independently predict breast cancer prognosis irrespective of ERα and AR expression [31]. *NLGN1* is a major component of excitatory glutamatergic synapse complexes, and recent research has identified it as a novel adverse prognostic marker for colon and rectal cancer [32]. It can promote colorectal cancer progression through the regulation of the APC/β-catenin pathway [33]. *SLITRK3* expression is an important predictor of recurrence and metastasis in gastrointestinal stromal tumors. *SLITRK3* combined with NIH staging has strong predictive and prognostic value and is a feasible biomarker for clinical use in GIST [34]. Meanwhile, high *SLITRK3* expression is associated with a poor outcome in LUSC, and it can activate *NTRK3* to promote suppressed cancer stem cell phenotypes [35]. However, despite the basic and clinical research conducted on some genes, the role of other genes and their value in future targeted therapies have not yet been fully elucidated. Therefore, these results suggest that further research can be conducted based on these prognostic genes to deepen our understanding of the mechanisms of breast cancer tumorigenesis and targeted drug development.

Immunotherapy is an emerging cancer treatment modality that involves activating the patient's intrinsic immune system to attack malignant cells [36]. It encompasses a range of techniques, including immune checkpoint inhibitors, T cell stimulators, and CAR-T cell therapy, that function through distinct mechanisms to activate or augment the patient's immune response, ultimately enhancing their ability to combat cancer. Nevertheless, not all patients respond favorably to immunotherapy, and the heterogeneity of response may be attributed to various factors, such as the patient's immunological status, the tumor's immune microenvironment, as well as the type and dose of immunotherapy [37]. Therefore, the investigation of immunotherapy efficacy and the immune microenvironment's characteristics is of paramount importance. The status of immune cell infiltration is crucial in understanding the tumor's immune microenvironment and devising effective treatment strategies [38]. Currently, transcriptome-based molecular subtyping of tumors has been widely acknowledged for its potential predictive value regarding immune cell infiltration and immunotherapeutic response. Analysis has demonstrated that both molecular subtypes based on 63 UUL-modified genes and risk-scoring models derived from UUL modification patterns are associated with immune cell infiltration, with UULscore being positively correlated with activated CD4+ T cells and negatively correlated with B cells and mast cells. Thus, these findings may pave the way for a deeper exploration of the regulatory mechanisms that govern tumor immune cell infiltration. Differences in immunotherapeutic response across distinct UULscore subgroups are clinically significant in guiding immunotherapy. Furthermore, analyses have revealed that different UULscores exhibit varying degrees of sensitivity to certain chemotherapeutic agents, indicating the guiding significance of the risk model in selecting chemotherapy drugs and suggesting the potential role of relevant genes in drug responses, thereby providing novel directions for future drug mechanism research and development.

Given the good predictive performance of the UULscore in breast cancer prognosis, the development of a commercial multigene detection panel would facilitate its clinical implementation. Large-scale prospective studies are warranted to further evaluate the potential clinical utility of the UULscore in conjunction with clinical-pathological characteristics for prognostic assessment in BC. However, there remain some limitations that must be addressed. Firstly, the retrospective cohorts sourced from public databases used to develop our model may introduce sample selection bias, which could limit its generalizability. Additional validation through prospective cohort studies is required to confirm its performance. Secondly, although we have identified several promising genes, their expression profiles and underlying mechanisms in breast cancer warrant further validation and investigation. Furthermore, the prognostic performance of the UULscore is derived from the UUL modification pattern; whether it operates independently of other modification patterns or synergizes with them to enhance predictive performance for breast cancer prognosis requires further exploration.

5. Conclusion

To summarize, our study initially identified two molecular subtypes of breast cancer derived from UUL modification-related genes. Based on DEGs between these subtypes, we constructed the UULscore model for breast cancer prognosis risk assessment, which has potential value in predicting breast cancer prognosis, immune therapy response, and chemotherapy drug sensitivity.

Conflict of interests

The author has no conflicts with any step of the article preparation.

Consent for publications

The author read and approved the final manuscript for publication.

Ethics approval and consent to participate

No human or animals were used in the present research.

Informed consent

The authors declare that no patients were used in this study.

Availability of data and material

The data that support the findings of this study are available from the corresponding author upon reasonable request

Authors' contributions

Nanyang Zhou, Dejia Kong, Qiao Lin, Xiaojing Yang, Dan Zhou and Lihua Lou designed, extracted, analyzed, and interpreted the data from databases. Nanyang Zhou and Xiangming Lou wrote the manuscript and made substantial contributions to the conception of the work and substantively revised it.

Funding

None.

References

1. Afifi AM, Saad AM, Al-Husseini MJ et al (2020) Causes of death after breast cancer diagnosis: a us population-based analysis. Cancer 126:1559-1567. https://doi.org/10.1002/cncr.32648
2. DeSantis CE, Ma J, Gaudet MM et al (2019) Breast cancer statistics, 2019. Ca Cancer J Clin 69:438-451. https://doi.org/10.3322/caac.21583
3. Trapani D, Ginsburg O, Fadelu T et al (2022) Global challenges and policy solutions in breast cancer control. Cancer Treat Rev 104:102339. https://doi.org/10.1016/j.ctrv.2022.102339
4. Cockram PE, Kist M, Prakash S et al (2021) Ubiquitination in the regulation of inflammatory cell death and cancer. Cell Death Differ 28:591-605. https://doi.org/10.1038/s41418-020-00708-5
5. Vaughan RM, Kupai A, Rothbart SB (2021) Chromatin regulation through ubiquitin and ubiquitin-like histone modifications. Trends Biochem Sci 46:258-269. https://doi.org/10.1016/j.tibs.2020.11.005
6. Wirth M, Schick M, Keller U et al (2020) Ubiquitination and ubiquitin-like modifications in multiple myeloma: biology and therapy. Cancers (Basel) 12. https://doi.org/10.3390/cancers12123764
7. Hwang JT, Lee A, Kho C (2022) Ubiquitin and ubiquitin-like proteins in cancer, neurodegenerative disorders, and heart diseases. Int J Mol Sci 23. https://doi.org/10.3390/ijms23095053
8. Song G, Zhang Y, Li H et al (2021) Identification of a ubiquitin related genes signature for predicting prognosis of prostate cancer. Front Genet 12:778503. https://doi.org/10.3389/fgene.2021.778503
9. Lin Y, Wang Q, Lin Y et al (2021) An immunohistochemical panel of three small ubiquitin-like modifier genes predicts outcomes of patients with triple-negative breast cancer. Gland Surg 10:1067-1084. https://doi.org/10.21037/gs-21-37
10. Fuentes-Antrás J, Alcaraz-Sanabria AL, Morafraile EC et al (2021) Mapping of genomic vulnerabilities in the post-translational ubiquitination, sumoylation and neddylation machinery in breast cancer. Cancers (Basel) 13. https://doi.org/10.3390/cancers13040833
11. Chen B, Khodadoust MS, Liu CL et al (2018) Profiling tumor infiltrating immune cells with cibersort. Methods Mol Biol 1711:243-259. https://doi.org/10.1007/978-1-4939-7493-1_12
12. Hu Z, Wei F, Su Y et al (2023) Histone deacetylase inhibitors promote breast cancer metastasis by elevating nedd9 expression. Signal Transduct Target Ther 8:11. https://doi.org/10.1038/s41392-022-01221-6
13. Ke C, Zhu K, Sun Y et al (2019) Sumo1 promotes the proliferation and invasion of non-small cell lung cancer cells by regulating nf-κb. Thorac Cancer 10:33-40. https://doi.org/10.1111/1759-7714.12895
14. Cheng XH, Xu TT, Zhou LB et al (2022) Sumo1-modified dna methyltransferase 1 induces dna hypermethylation of vwc2 in the development of colorectal cancer. Neoplasma 69:1373-1385. https://doi.org/10.4149/neo_2022_220817N841
15. Kariri Y, Toss MS, Alsaleem M et al (2022) Ubiquitin-conjugating enzyme 2c (ube2c) is a poor prognostic biomarker in invasive breast cancer. Breast Cancer Res Treat 192:529-539. https://doi.org/10.1007/s10549-022-06531-5
16. Lu ZN, Song J, Sun TH et al (2021) Ube2c affects breast cancer proliferation through the akt/mtor signaling pathway. Chin Med J (Engl) 134:2465-2474. https://doi.org/10.1097/CM9.0000000000001708
17. Zhou H, Liu Y, Zhu R et al (2017) Fbxo32 suppresses breast cancer tumorigenesis through targeting klf4 to proteasomal degradation. Oncogene 36:3312-3321. https://doi.org/10.1038/onc.2016.479
18. Wu H, Jiao Y, Guo X et al (2024) Mettl14/mir-29c-3p axis drives aerobic glycolysis to promote triple-negative breast cancer progression though trim9-mediated pkm2 ubiquitination. J Cell Mol Med 28:e18112. https://doi.org/10.1111/jcmm.18112
19. Tocci JM, Felcher CM, García SM et al (2020) R-spondin-mediated wnt signaling potentiation in mammary and breast cancer development. Iubmb Life 72:1546-1559. https://doi.org/10.1002/iub.2278
20. Tang J, Luo Y, Tian Z et al (2020) Trim11 promotes breast cancer cell proliferation by stabilizing estrogen receptor α. Neoplasia 22:343-351. https://doi.org/10.1016/j.neo.2020.06.003
21. Song W, Wang Z, Gu X et al (2019) Trim11 promotes proliferation and glycolysis of breast cancer cells via targeting akt/glut1 pathway. Onco Targets Ther 12:4975-4984. https://doi.org/10.2147/OTT.S207723
22. Sanyal S, Mondal P, Sen S et al (2020) Sumo e3 ligase cbx4 regulates htert-mediated transcription of cdh1 and promotes breast cancer cell migration and invasion. Biochem J 477:3803-3818. https://doi.org/10.1042/BCJ20200359
23. Guo Y, Chen X, Zhang X et al (2023) Ube2s and ube2c confer a poor prognosis to breast cancer via downregulation of numb. Front Oncol 13:992233. https://doi.org/10.3389/fonc.2023.992233
24. Paul D, Ghorai S, Dinesh US et al (2017) Cdc20 directs proteasome-mediated degradation of the tumor suppressor smar1 in higher grades of cancer through the anaphase promoting complex. Cell Death Dis 8:e2882. https://doi.org/10.1038/cddis.2017.270
25. Qiu SQ, Waaijer S, Zwager MC et al (2018) Tumor-associated macrophages in breast cancer: innocent bystander or important player? Cancer Treat Rev 70:178-189. https://doi.org/10.1016/j.ctrv.2018.08.010
26. Niogret J, Berger H, Rebe C et al (2021) Follicular helper-t cells restore cd8(+)-dependent antitumor immunity and anti-pd-l1/pd-1 efficacy. J Immunother Cancer 9. https://doi.org/10.1136/jitc-2020-002157
27. Burugu S, Asleh-Aburaya K, Nielsen TO (2017) Immune infiltrates in the breast cancer microenvironment: detection, characterization and clinical implication. Breast Cancer 24:3-15. https://doi.org/10.1007/s12282-016-0698-z
28. Szpor J, Streb J, Glajcar A et al (2021) Dendritic cells are associated with prognosis and survival in breast cancer. Diagnostics (Basel) 11. https://doi.org/10.3390/diagnostics11040702
29. Zhang Y, Qin Y, Li D et al (2022) A risk prediction model mediated by genes of apod/apoc1/sqle associates with prognosis in cervical cancer. Bmc Womens Health 22:534. https://doi.org/10.1186/s12905-022-02083-4
30. Allina DO, Andreeva YY, Zavalishina LE et al (2016) [Estimation of the diagnostic potential of apod, ptov1, and epha4 for prostatic neoplasms]. Arkh Patol 78:9-14. https://doi.org/10.17116/patol20167859-14
31. Jankovic-Karasoulos T, Bianco-Miotto T, Butler MS et al (2020) Elevated levels of tumour apolipoprotein d independently predict poor outcome in breast cancer patients. Histopathology 76:976-987. https://doi.org/10.1111/his.14081

32. Yu Q, Wang X, Yang Y et al (2021) Upregulated nlgn1 predicts poor survival in colorectal cancer. Bmc Cancer 21:884. https://doi.org/10.1186/s12885-021-08621-x
33. Pergolizzi M, Bizzozero L, Maione F et al (2022) The neuronal protein neuroligin 1 promotes colorectal cancer progression by modulating the apc/β-catenin pathway. J Exp Clin Cancer Res 41:266. https://doi.org/10.1186/s13046-022-02465-4
34. Wang CJ, Zhang ZZ, Xu J et al (2015) Slitrk3 expression correlation to gastrointestinal stromal tumor risk rating and prognosis. World J Gastroenterol 21:8398-8407. https://doi.org/10.3748/wjg.v21.i27.8398
35. Bollig-Fischer A, Bao B, Manning M et al (2021) Role of novel cancer gene slitrk3 to activate ntrk3 in squamous cell lung cancer. Mol Biomed 2:26. https://doi.org/10.1186/s43556-021-00051-2
36. Zhang Y, Zhang Z (2020) The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. Cell Mol Immunol 17:807-821. https://doi.org/10.1038/s41423-020-0488-6
37. Vitale I, Shema E, Loi S et al (2021) Intratumoral heterogeneity in cancer progression and response to immunotherapy. Nat Med 27:212-224. https://doi.org/10.1038/s41591-021-01233-9
38. Mao X, Xu J, Wang W et al (2021) Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives. Mol Cancer 20:131. https://doi.org/10.1186/s12943-021-01428-1