Research article

Identification of exosome-related gene signature as a promising diagnostic and therapeutic tool for breast cancer

Qitong Chen [a,b], Qin Zhou [a,b,*]

[a] *Department of General Surgery, The Second Xiangya Hospital of Central South University, Changsha, Hunan, China*
[b] *Clinical Research Center for Breast Disease in Hunan Province, Changsha, Hunan, China*

A R T I C L E   I N F O

A B S T R A C T

Background: Exosomes are promising tools for the development of new diagnostic and therapeutic approaches. Exosomes possess the ability to activate signaling pathways that contribute to the remodeling of the tumor microenvironment, angiogenesis, and the regulation of immune responses. We aimed to develop a prognostic score based on exosomes derived from breast cancer.
Materials and methods: Training was conducted on the TCGA-BRCA dataset, while validation was conducted on GSE20685, GSE5764, GSE7904, and GSE29431. A total of 121 genes related to exosomes were retrieved from the ExoBCD database. The Cox proportional hazards model is used to develop risk score model. The GSVA package was utilized to analyze single-sample gene sets and identify exosome signatures, while the WGCNA package was utilized to identify gene modules associated with clinical outcomes. The clusterProfiler and GSVA R packages facilitated gene set enrichment and variation analyses. Furthermore, CIBERSORT quantified immune infiltration, and a correlation between gene expression and drug sensitivity was assessed using the TIDE algorithm.
Results: An exosome-related prognostic score was established using the following selected genes: *ABCC9*, *PIGR*, *CXCL13*, *DOK7*, *CD24*, and *IVL*. Various immune cells that promote cancer immune evasion were associated with a high-risk prognostic score, which was an independent predictor of outcome. High-risk and low-risk groups exhibited significantly different infiltration abundances (p < 0.05). By conducting a sensitivity comparison, we found that patients with high-risk scores exhibited more favorable responses to immunotherapy than those with low-risk scores.
Conclusion: The exosome-related gene signature exhibits outstanding performance in predicting the prognosis and cancer status of patients with breast cancer and guiding immunotherapy.

1. Introduction

Recent GLOBOCAN data [1] reveals that breast cancer is increasingly prevalent among women worldwide, emphasizing its status as a major public health concern and the urgent need for enhanced research and treatment. Accurate diagnosis, outcome prediction, and treatment selection for breast cancer rely on histological and phenotypic distinctions among tumors [2]. The rapid advancement of high-throughput technology in multi-omics has resulted in the identification of biomarkers that hold tremendous clinical significance

for early screening, differential diagnosis, precise treatment, and prognosis of breast cancer [3]. However, despite these advances, numerous patients with breast cancer still experience disease progression after their initial diagnosis [3]. Hence, an urgent need exists to acquire a profound understanding of the fundamental mechanisms implicated in breast cancer progression and the establishment of efficacious therapeutic approaches.

Exosomes, with a size range of 30–150 nm, are extracellular vesicles enclosed by a membrane released by different cells, such as immune cells, stem cells, and cancer cells [4,5]. Initially discovered in the 1980s as potential cellular debris [6,7], exosomes have since been recognized as pivotal mediators of intercellular communication [8,9]. Exosomes play a significant role in transferring information and signaling molecules between cells, thereby activating pathways that promote cell proliferation, angiogenesis, and immune evasion, while also suppressing antitumor immune responses [10–12]. Specifically, exosomes from stromal or cancer cells within the tumor microenvironment (TME) are crucial in drug resistance development [13,14]. By transferring signaling molecules like miRNAs, proteins, and transcription factors [15], these vesicles can alter gene expression in recipient cells and induce drug resistance. Exosomes exert a significant influence on the behavior of infiltrating immune cells and are crucial in enhancing the effectiveness of immunotherapies as well as facilitating tumor progression [16,17]. The phenomenon of exosome-mediated drug resistance, significant across various cancer types, is notably observed in breast, lung, and colorectal cancers, yet it is not confined to these [18]. This represents a substantial challenge, particularly in the context of chemotherapy and immunotherapy [19]. As such, exosomes have become promising targets for developing novel diagnostic and therapeutic strategies to modify the tumor-associated immune environment and enhance responsiveness to immunotherapy.

Infiltrating immune cells within the TME modulate immune responses and promote tumor cell migration and invasion [14]. This process contributes to the formation of premetastatic niches in distant organs for disseminated tumor cells [14]. These cells exert antitumor effects by releasing cytotoxic substances, recognizing and eliminating tumor cells, and producing cytokines [20,21]. However, tumor cells have developed mechanisms to evade immune detection, such as releasing immunosuppressive factors, altering immune cell activity and function, and inducing immune cell exhaustion [20,22]. The dynamic interplay between immune cells and tumor cells is central to understanding tumor progression and developing effective immunotherapeutic strategies. By dissecting these interactions, researchers can gain essential insights into creating new therapeutic approaches against cancer.

Recent advancements in cancer prognostics underscore the pivotal role of exosomes in tumor microenvironment modulation across various cancers [23–29]. Studies, including groundbreaking work on head-neck squamous cell carcinoma [30], have elucidated the prognostic significance of exosome-related genes (ERGs) in forecasting treatment responses and patient prognoses. These models leverage bioinformatics and statistical analyses to establish prognostic signatures correlating with survival rates, immune infiltration, and therapeutic responses. Despite significant advancements, the exploration of exosome-related biomarkers in breast cancer prognostics remains underexplored. Our study sought to assess the association between ERGs and breast cancer prognosis, diagnostic accuracy, and patient responsiveness to immunotherapy. To construct an exosome risk model, we initially collected high-throughput data of patients with breast cancer from TCGA and GEO databases. Subsequently, we focused on exosome-associated genes using the ExoBCD database, which identifies these genes through detailed analyses and literature reviews, underlining their significance in breast cancer. The investigation concluded with the identification of prognostic ERGs, after employing a series of analytical techniques. Hub genes were utilized in a preliminary step to evaluate the differences in immune cell infiltration rates and immunotherapy responsiveness. Receiver operating characteristic (ROC) curves and survival analyses were used to determine and validate the diagnostic, therapeutic, and prognostic values of these new biomarkers. Distinct from existing models, our research highlights the unique molecular landscape of breast cancer, providing a bespoke prognostic tool that enhances precision medicine for this prevalent disease.

Table 1
List of dataset Information.

	TCGA-BRCA	GSE20685	GSE5764	GSE7904	GSE29431
Platform	TCGA	GPL570	GPL570	GPL570	GPL570
Species	*Homo sapiens*	*Homo sapiens*	*Homo sapiens*	*Homo sapiens*	*Homo sapiens*
Tissue	Breast	primary breast cancer	mastectomy specimen	human breast tissue	Breast tissue
Samples in normal group	113	/	20	7	12
Samples in BRCA group	1109	327	10	43	54
Reference	/	Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization.	Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis.	X chromosomal abnormalities in basal-like human breast cancer.	/

2. Methods

2.1. Data acquisition

The TCGAbiolinks [31] was utilized for to acquire TCGA-BRCA dataset. Samples with incomplete clinical information were excluded from the analysis. The study analyzed data from 1,109 patients within the TCGA-BRCA breast cancer dataset, which consisted of 1,109 tumor tissue samples in the BRCA group and 113 samples of adjacent normal tissue in the normal group. Log2 transformed RNA-Seq read counts were used to calculate fragments per kilobase per million (FPKM). The clinical data of 1,109 breast cancer were obtained from the UCSC Xena database [32] (https://xena.ucsc.edu/). cBioPortal for Cancer Genomics (https://www.cbioportal.org/) provided tumor mutation burden (TMB) and microsatellite instability (MSI) metrics [33].

The gene expression datasets (Table 1), GSE20685 [34], GSE5764 [35], GSE7904 [36], and GSE29431, were downloaded from the Gene Expression Omnibus (GEO) database using the GEOquery R package [37]. The GSE20685 (BRCA tissue, n = 327), GSE5764 (normal tissue, n = 20; invasive ductal carcinomas, n = 5; invasive papillary carcinoma, n = 5), GSE7904 (BRCA tissue, n = 43; normal tissue, n = 7), and GSE29431 (BRCA tissue, n = 54; normal tissue, n = 12) datasets served as the validation sets. The methodology for this study is depicted in a flowchart in Supplementary Fig. 1.

2.2. Exosome-related genes and differentially expressed genes

121 ERGs were identified from the ExoBCD database [38] (https://exobcd.liumwei.org/; Table S1). ExoBCD database was instrumental in our gene selection process, employing comprehensive analyses, TCGA case validations, and literature reviews to pinpoint genes critical for breast cancer's exosomal processes. Through ExoBCD, these genes were chosen for their strong linkage to diagnostic, prognostic, and therapeutic relevance in the disease context. Thereafter, the differentially expressed genes (DEGs) between tumors and adjacent tissues in TCGA-BRCA dataset were determined using the limma [39] R package. Count data were normalized and analyzed using the limma package. DEGs were selected based on |logFC| > 1 and p.adjust <0.05. The study divided DEGs into two categories: upregulated genes, which had a logFC >1 and an adjusted p-value <0.05, and downregulated genes, characterized by a logFC < −1 and an adjusted p-value <0.05.

2.3. Phenotype score

An analysis of single-sample gene set enrichment (ssGSEA) was performed using the GSVA [40] R package. Each sample was analyzed using the ssGSEA to determine its exosome-related enrichment scores. Tumor samples in TCGA-BRCA dataset were divided into two groups, high and low, based on the median phenotype scores.

2.4. Weighted gene co-expression network analysis (WGCNA)

Weighted gene co-expression network analysis (WGCNA) [41] was performed to assess the inter-gene co-expression relationships and subsequently categorizing co-expressed genes into modules. Within each module, gene expression levels were consistent, but showed substantial differences when compared across various modules. Of note, this concept allows the transformation of complicated high-throughput data into multiple basic modules, and the dimensionality reduction process is partially accomplished. The association between these gene co-expression modules and clinical phenotypes should then be determined, followed by the biological importance of this module.

To obtain co-expression modules from TCGA-BRCA tumor samples, DEGs and exosome-related scores were obtained. The WGCNA [42] R package was used to identify the co-expressed modules, with parameters "Soft-power = 5, RsquaredCut = 0.9, minModuleSize = 50, mergeCutHeight = 0.2, mincutHeight = 0.2." Key genes were extracted from the overlap between module genes and DEGs, prioritizing those with the greatest correlation to exosome scores.

2.5. Mutation analysis and Cox regression analysis

With a group comparison graph and the Pearson correlation coefficient, we identified and compared the differentially expressed key genes between breast cancer and adjacent normal tissues. Somatic mutation data from the TCGA-BRCA dataset were preprocessed using VarScan software (https://varscan.sourceforge.net/), and the maftools package [43] enabled visualization of mutations. TCGA-BRCA dataset was analyzed to determine the expression of essential genes using univariate Cox regression analysis [44]. Genes with a significance level of $p < 0.05$ were selected for further investigation.

2.6. Risk model construction

We conducted regularization and dimensionality reduction analyses using the glmnet package [45]. LASSO regression was used in conjunction with ten-fold crossvalidation [46], with "2021″ as the chosen random seed number. We selected key genes to build a risk model, which was later used to construct a prognostic score for OS [44,47]. Using the 'rms' R package in conjunction with the multivariate Cox regression analysis findings, we developed a nomogram [48], illustrating the functional interactions between

independent variables. This allowed for event probability prediction by applying a unique scoring scale to each model variable. Based on the median prognostic score of the multivariate Cox regression model, patients were categorized into high- and low-risk groups. Using the same coefficients, a risk model was established for the GSE20685 dataset. Multivariate Cox models for high- and low-risk groups were assessed using time-dependent ROC curves for the TCGA-BRCA dataset. The prognostic score was calculated using the following equation:

$$prognostic\ score = \sum_i beta\ (hub\ gene_i) \times x\ (hub\ gene_i)$$

2.7. Functional enrichment analysis

Analysis of Gene Ontologies (GO) [49] is a common technique for large-scale functional enrichment analyses. Kyoto Encyclopedia of Genes and Genomes (KEGG [50]) database provides information on biological systems and pathways. Using the ClusterProfiler R package [51], GO and KEGG enrichment was executed. Our study adopted *P*-values or FDR below 0.05 as thresholds for statistical significance. Corrections to *P*-values were carried out using Benjamini-Hochberg (BH) methods. We proceeded to visualize the enriched pathways via the Pathview [52] R package.

2.8. GSEA

This study used clusterProfiler to perform Gene Set Enrichment Analysis (GSEA) [53]. In TCGA-BRCA dataset, gene enrichment analyses were performed on genes in the high- and low-risk groups of the LRRGs risk model. The following parameters were used for GSEA: the random seed used in this study was set to 2022. To ensure statistical robustness, gene set permutations were performed 1, 000 times for each analysis, and each gene set contained 10–500 genes. The reference gene collection was c2.cp.v7.2.symbols.gmt obtained from the Molecular signature database 3.0 (MsigDB, https://www.gsea-msigdb.org/gsea/msigdb) [54]. Genes with values less than 0.05 were considered significantly enriched.

2.9. Protein-protein and mRNA-miRNA interaction network

The Protein-Protein Interaction (PPI) network comprises individual proteins that interact with one another to participate in various biological processes, including signaling, gene expression regulation, material metabolism, and energy transduction. LRRGs with low confidence were selected from the STRING database (https://string-db.org/) [55], using a minimum interaction score of 0.150. For visualization purposes, the network was rendered in Cytoscape (https://cytoscape.org/). Importantly, closely interconnected regions in the PPI network are indicative of molecular complexes involved in specialized biological processes.

The Encyclopedia of RNA Interactomes (ENCORI) [56] database (http://starbase.sysu.edu.cn/) version 3.0 of the starBase database has identified millions of interactions among miRNAs, ncRNAs, mRNA, and RBP by analyzing thousands of CLIP-seq and other high-throughput sequencing datasets. The database also offers several interfaces for visualizing miRNA target information. The ENCORI database was utilized to predict the interactions between miRNAs and hub genes. Subsequently, the mRNA-miRNA interaction network was illustrated using Cytoscape software.

2.10. Immune infiltration analysis of TCGA-BRCA dataset

We integrated TCGA-BRCA RNA matrix with the LM22 characteristic gene matrix using R code from the CIBERSORT (https://cibersortx.stanford.edu/) [57] website. After eliminating the data with null immune cell enrichment scores, an immune cell infiltration abundance matrix was obtained. The stacked histogram displays the fraction of immune cell infiltration in TCGA-BRCA dataset samples, whereas the boxplot illustrates the difference. Immune cells were compared across various groups using the Pearson correlation coefficient. By applying the ssGSEA algorithm, we assessed the abundance of infiltrating immune cell types and classified them into distinct subtypes. ssGSEA enrichment scores provided insight into the relative presence of these immune cells [58,59].

Consensus Clustering [60] can categorize samples into subtypes based on many omics datasets to identify new disease subtypes or perform comparisons of different subtypes. Utilizing ConsensusClusterPlus [61], we conducted cluster analysis to delineate molecular subtypes from the expression profiles of phenotype-related genes in the TCGA-BRCA dataset. The following parameters were used for the clustering algorithm: reps = 1000; pItem = 0.8; pFeature = 1; and method, 'euclidean'.

2.11. Statistical analysis

Statistical analyses involved the Student's t-test and Mann-Whitney *U* test for normal and non-normal variables, respectively, with the Kruskal-Wallis test for multi-group comparisons. Categorical differences were examined using the chi-square or Fisher's exact test. A survival analysis was conducted using Kaplan-Meier curves and the log-rank test in the R 'survival' package, which considered two-sided p-values below 0.05 as statistically significant discrepancies. We performed all data processing and analyses with the open-source R software, version 4.1.3.

3. Results

3.1. Screening of key genes via differential expression analysis and WGCNA

Differentially expressed genes between tumor and adjacent normal tissues in the TCGA-BRCA dataset were identified using the limma R package. 6,786 DEGs were identified, each exhibiting an absolute log fold change ($|logFC|$) greater than 1 and an adjusted p-value below 0.05. Under this threshold, 2,899 genes were upregulated ($logFC > 1$ and p.adjust < 0.05) and 3,887 genes were
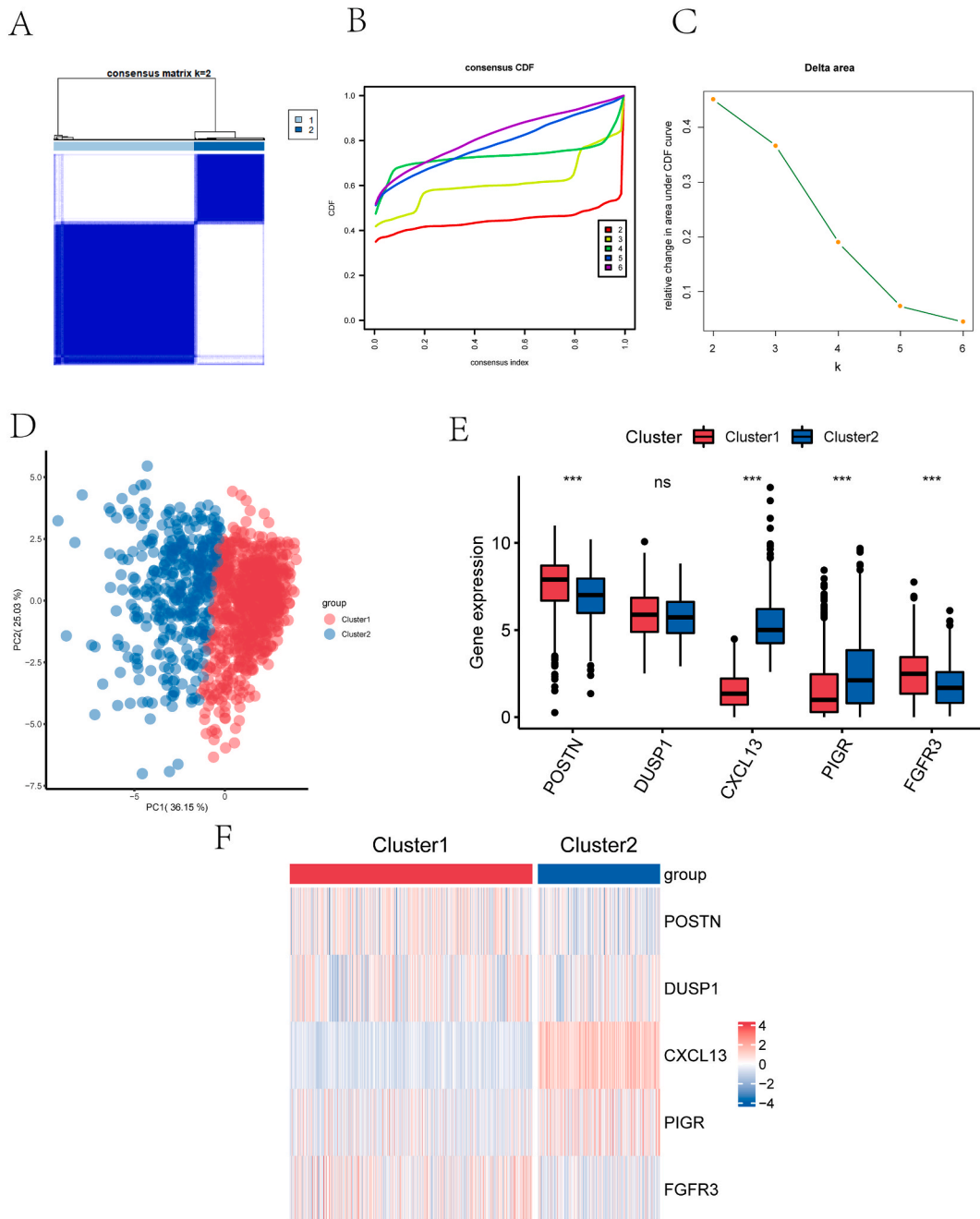
Fig. 1. Clustering molecular subtype. **(A)** Consensus cluster heatmaps (k = 2) in TCGA-BRCA dataset; **(B)** Consensus cluster cumulative distribution function (CDF) plot; **(C)** Delta area under CDF of the consensus cluster; **(D)** Principal component analysis (PCA) plot of cluster 1 and cluster 2; **(E)** Boxplot of the comparison of key genes in cluster 1 and cluster 2; **(F)** Heatmap illustrating the expression of key genes in different subtypes. ns, non significant, p ≥ 0.05; *, p < 0.05; **, p < 0.01; ***, p < 0.001.

downregulated (logFC < -1 and p.adjust <0.05). A volcano map was generated based on the results of differential analysis of this dataset (Supplementary Fig. 2A). We obtained 30 differentially expressed exosome-related genes (ERDEGs) by intersecting DEGs and ERGs. The expression of ERDEGs is shown as a heatmap (Supplementary Fig. 2B).

Using the expression of 30 phenotype-related genes, we calculated the exosome phenotype score for each individual in TCGA-BRCA cohort using ssGSEA. The TCGA-BRCA dataset underwent WGCNA to identify co-expression modules. The optimal soft threshold was 6 (Supplementary Fig. 2C), and the genes in TCGA-BRCA dataset were clustered into 11 modules (Supplementary Fig. 2D), among which the pink module had the highest correlation with the exosome score (Supplementary Fig. 2E). The differentially expressed ERGs in TCGA-BRCA dataset intersected with the green, magenta, brown, and pink module genes to obtain five key genes (*POSTN*, *DUSP1*, *CXCL13*, *PIGR*, and *FGFR3*) (Supplementary Fig. 2F).

3.2. Mutation analysis of key genes

The boxplot displays the expression differences in five key genes (*POSTN, DUSP1, CXCL13, PIGR*, and *FGFR3*) between the tumor and adjacent normal groups (Supplementary Fig. 3A). All key genes were identified to exhibit statistically significant group differences ($p < 0.05$). An expression correlation heatmap was generated for the five key genes (Supplementary Fig. 3B). A substantial (R = 0.31, p < 0.05) positive association was observed between *DUSP1* and *PIGR*.

We evaluated somatic mutations in five key genes in TCGA-BRCA dataset. Supplementary Figs. 2C–D shows the mutation analysis results of the five key genes in the breast cancer patient samples. Missense and nonsense mutations were the most prevalent forms of
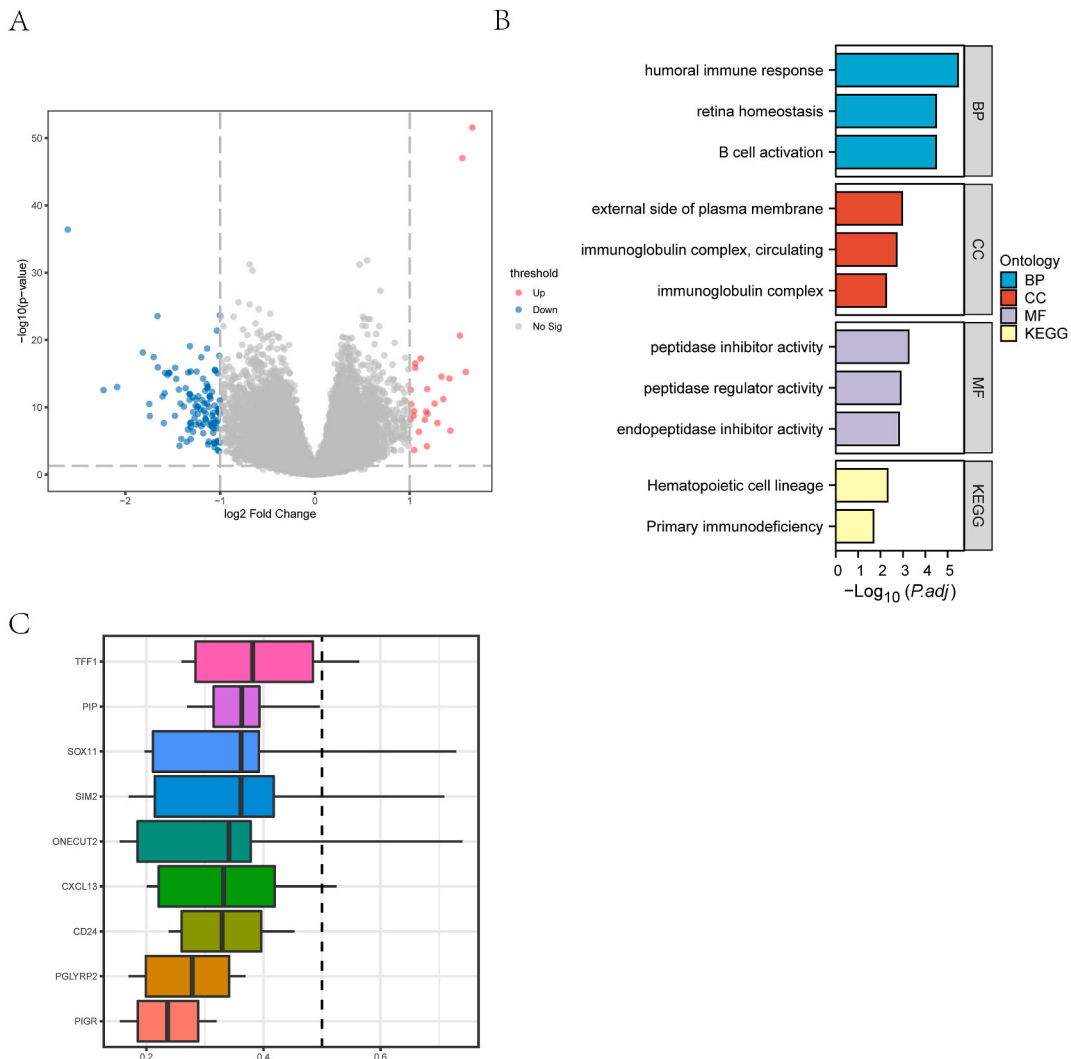
Fig. 2. Differential expression analysis and GO and KEGG analyses based on the high- and low-risk groups. **(A)** Volcano plot of differential expression analysis; **(B)** Bar plot of GO functional annotation and KEGG pathway analysis in TCGA-BRCA dataset; **(C)** Boxplot of the semantic similarity of hub genes.

somatic mutations in TCGA-BRCA dataset. Notably, missense mutations are the most common in patients with breast cancer. Furthermore, the most prevalent SNV mutation in patients with cancer was from C to T, followed by C to G. (Supplementary Fig. 3C). Supplementary Fig. 3D shows that three of the five key genes in TCGA-BRCA dataset had somatic mutations, with *POSTN* and *PIGR* having the highest mutation rates and accounting for 1 % of breast cancer patient mutation samples.

3.3. Clustering and risk prediction model construction

We used the R package, ConsensusClusterPlus, to cluster molecular subtypes based on five key genes (*POSTN*, *DUSP1*, *CXCL13*, *PIGR,* and *FGFR3*) and obtained two subtypes: Cluster 1 and Cluster 2 (Fig. 1A–C). Cluster 1 comprised 738 samples whereas cluster 2 comprised 371 samples. The PCA plot indicated that the distributions of the principal components of the two subtypes were distinct (Fig. 1D). The group comparison boxplot revealed that among the five key genes, *DUSP1* did not significantly differ between the two subtypes, whereas *POSTN*, *CXCL13*, *PIGR*, and *FGFR3* exhibited substantial differences (Fig. 1E). A heatmap, created using the pheatmap R package, depicted the differential expression of five key genes between two subtype groups (Fig. 1F), *POSTN* and *FGFR3* were found to be abundantly expressed in Cluster 1 whereas *CXCL13* and *PIGR* were highly expressed in Cluster 2.

Based on the DEGs of tumors and adjacent normal tissues, a risk scoring system was established to quantify the prognostic risk of each breast cancer patient using prognostic score. Univariate Cox analysis of the DEGs was conducted and six genes (*ABCC9*, *PIGR*, *CXCL13*, *DOK7*, *CD24*, and *IVL*) were identified (Supplementary Fig. 4A). LASSO analysis facilitated the selection of pivotal candidate genes for the prognostic risk model, as detailed in Supplementary Figs. 4B–C, yielding an optimal model incorporating six essential variables: *ABCC9*, *PIGR*, *CXCL13*, *DOK7*, *CD24*, and *IVL*. These six genes were integrated into multivariate Cox regression analysis, and the prognostic score was computed using multivariate regression coefficients and a gene expression matrix. The division of TCGA-BRCA tumor samples into high- and low-risk groups was determined by the median prognostic score.

Kaplan-Meier curves were generated for the six exosome-related differentially expressed genes (*ABCC9*, *PIGR*, *CXCL13*, *DOK7*, *CD24*, and *IVL*, Supplementary Figs. 4D–4I). *ABCC9* ($p = 0.015$), *PIGR* ($p = 0.008$), *DOK7* ($p = 0.0028$), and *CD24* were identified as ERGs associated with prognostic significance ($p < 0.0001$).

3.4. DEGs of the high- and low-risk groups and functional enrichment analysis

Based on the high- and low-risk groups, differential analysis was performed on the RNA-seq data from TCGA-BRCA dataset. Seventeen DEGs exhibited upregulation (logFC >1 and p.adjust <0.05), while ninety-four DEGs were downregulated (logFC < −1 and p.adjust <0.05). Differential expression analysis results are represented in a volcano plot (Fig. 2A).

We evaluated the functional enrichment of the 111 DEGs using GO and KEGG annotations (Fig. 2B, Table 2). The following terms were enriched according to GO enrichment analyses: B cell activation, retinal homeostasis, and humoral immune response for BP; immunoglobulin complex, circulating, and external side of the plasma membrane for CC; and endopeptidase inhibitor activity, peptidase regulator activity, and peptidase inhibitor activity for MF (Fig. 2B). KEGG analysis revealed that the hematopoietic cell lineage and primary immunodeficiency pathway were enriched (Fig. 2B). The R package, GOSemSim, was used to estimate the semantic similarity of the GO terms and gene clusters. The TOP5 positive and negative DEGs were considered as hub genes (*PIGR*, *PIP*, *TFF1*, *CXCL13*, *PGLYRP2*, *ONECUT2*, *SOX11*, *CD24P4*, *SIM2*, and *CD24*). Boxplots of the hub genes are presented in Fig. 2C. Based on the results, *TFF1* had the highest functional similarity among the 10 hub genes.

GSEA was performed to demonstrate biological processes (summarized in Table 3), cellular components, and molecular functions of the DEGs (Supplementary Fig. 5A). DEGs were significantly enriched in oxidative phosphorylation (Supplementary Fig. 5B), the

Table 2
GO and KEGG enrichment analysis results of differentially expressed genes in high - and low-risk groups.

ONTOLOGY	Description	pvalue	p.adjust	qvalue
BP	humoral immune response	2.3294E-09	3.4638E-06	2.962E-06
BP	B cell activation	4.6626E-08	3.3129E-05	2.833E-05
BP	retina homeostasis	8.5941E-08	3.3129E-05	2.833E-05
BP	humoral immune response mediated by circulating immunoglobulin	8.9118E-08	3.3129E-05	2.833E-05
BP	B cell receptor signaling pathway	1.6504E-07	4.9082E-05	4.1972E-05
CC	external side of plasma membrane	8.3521E-06	0.00108577	0.00099346
CC	immunoglobulin complex, circulating	2.9354E-05	0.00190803	0.00174581
CC	immunoglobulin complex	0.00012838	0.00556323	0.00509024
CC	blood microparticle	0.00061497	0.01998658	0.01828732
CC	tertiary granule lumen	0.00213506	0.05551143	0.05079184
MF	peptidase inhibitor activity	2.6326E-06	0.00055285	0.0004711
MF	peptidase regulator activity	1.2063E-05	0.00126667	0.00107937
MF	endopeptidase inhibitor activity	2.131E-05	0.00149169	0.00127111
MF	endopeptidase regulator activity	3.4452E-05	0.00180873	0.00154127
MF	glycosaminoglycan binding	0.00011225	0.00471436	0.00401725
KEGG	Hematopoietic cell lineage	7.5916E-05	0.00485859	0.0043152
KEGG	Primary immunodeficiency	0.00064816	0.02074111	0.01842138

GO，Gene Ontology; BP，biological process; CC，cellular component; MF，molecular function; KEGG，Kyoto Encyclopedia of Genes and Genomes.

Table 3
GSEA analysis of TGCA-BRCA.

ID	NES	p.adjust	qvalue
WP_OXIDATIVE_PHOSPHORYLATION	−1.85535702	0.02932436	0.02282698
WP_IL18_SIGNALING_PATHWAY	−1.58655817	0.02932436	0.02282698
WP_APOPTOSIS	−1.78367559	0.0370099	0.02880964
REACTOME_REGULATION_OF_TP53_ACTIVITY_THROUGH_PHOSPHORYLATION	1.71621257	0.0370099	0.02880964
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY	−1.59908052	0.04035079	0.03141029
WP_REGULATORY_CIRCUITS_OF_THE_STAT3_SIGNALING_PATHWAY	−1.61447182	0.04037724	0.03143088
KEGG_MAPK_SIGNALING_PATHWAY	−1.34749707	0.04668053	0.03633755
REACTOME_INITIAL_TRIGGERING_OF_COMPLEMENT	−2.71823284	0.02932436	0.02282698
REACTOME_CD22_MEDIATED_BCR_REGULATION	−2.71766712	0.02932436	0.02282698
REACTOME_COMPLEMENT_CASCADE	−2.66616791	0.02932436	0.02282698
REACTOME_SCAVENGING_OF_HEME_FROM_PLASMA	−2.66316589	0.02932436	0.02282698
REACTOME_FCGR_ACTIVATION	−2.66299217	0.02932436	0.02282698

TCGA，The cancer genome atlas; BRCA，Breast invasive carcinoma; GSEA：Gene Set Enrichment Analysis.

IL18 signaling pathway (Supplementary Fig. 5C), apoptosis (Supplementary Fig. 5D), regulation of TP53 activity through phosphorylation (Supplementary Fig. 5E), the toll-like receptor signaling pathway (Supplementary Fig. 5F), regulatory circuits of the STAT3 signaling pathway (Supplementary Fig. 5G), and MAPK signaling pathway (Supplementary Fig. 5H).

3.5. Protein-protein interaction analysis and mRNA-miRNA network construction

Using the STRING database, we predicted and constructed 111 DEGs interaction networks (Supplementary Fig. 6A) to explore the physical interaction relationships, common protein domains, and gene interactions between them. Setting thresholds of K-core = 2, Node Score Cutoff = 2, and maximum depth = 100, we utilized the MCODE plug-in to identify core gene clusters within the protein interaction network (Supplementary Fig. 6A). The following three main gene clusters were identified via the analysis: (*TCL1A, CXCL13, BLK, PAX5, CR2, TNFRSF13B, CCL21, FCER2, MS4A1, CCL19, CD19, CD79A, CD24*, score = 7.167), (*TFF1, TFF3, AGR2*, AGR3, score = 4), and (*PIP, SCGB2A2, SCGB1D2*, score = 3). The top10 genes were obtained consecutively using the MCC, DMNC, and EPC algorithms of the CytoHubba plug-in. Seven core genes (*BLK, PAX5, CR2, TNFRSF13B, FCER2, MS4A1*, and *CD79A*, Supplementary Fig. 6B) were selected as hub genes for subsequent studies by intersecting the core gene clusters acquired by MCODE with the genes measured using different methods of the cytoHubba plug-in.

Using the ENCORI database's mRNA-miRNA prediction tool, miRNAs associated with seven hub genes were pinpointed (Supplementary Fig. 6C). The resulting mRNA-miRNA interaction network, featuring five hub genes (*PAX5, CR2, TNFRSF13B, MS4A1*, and *CD79A*), 82 miRNA molecules, and 96 pairings, was visualized via Cytoscape. The specific mRNA-miRNA interactions are shown in Table S2.

3.6. Analysis of immune infiltration

To examine immune infiltration, we calculated the infiltration abundance of 22 immune cells using the CIBERSORT algorithm. The bar plot depicts the proportion of abundant immune cell infiltration in TCGA-BRCA dataset (Fig. 3A). Using the Mann-Whitney *U* test, the differences in immune cell infiltration between the groups were statistically examined, and the results are displayed in a grouped boxplot (Fig. 3B). There were statistically significant differences (p < 0.05) in the infiltration abundances of the 10 types of immune cells in TCGA-BRCA dataset between the high- and low-risk groups. The 10 types of immune cells included naïve B cells, CD8 T cells, follicular helper T cells, regulatory (Tregs) T cells, gamma delta T cells, activated NK cells, monocytes, macrophages M0, macrophages M2, and dendritic cells. We computed the correlation between the abundance of these 10 types of immune cells (Fig. 3C and D).

We also calculated the correlation between hub genes (*BLK, PAX5, CR2, TNFRSF13B, FCER2, MS4A1*, and *CD79A*) and immune cells using ssGSEA and the Pearson correlation coefficient (Supplementary Fig. 7A). Several immune cell types were correlated with the hub genes (Supplementary Figs. 7B–I). *CD79A* expression was significantly and positive correlated with activated B cells (R = 0.89, p < 0.01) and immature B cells (R = 0.84, p < 0.01). Similarly, *MS4A1* demonstrated significant positive correlations with activated B cells (R = 0.81, p < 0.01) and immature B cells (R = 0.80, p < 0.01), as did *TNFRSF13B* with activated B cells (R = 0.7, p < 0.01) and immature B cells (R = 0.67, p < 0.01), and *BLK* with both activated B cells and immature B cells (R = 0.69, p < 0.01 for each).
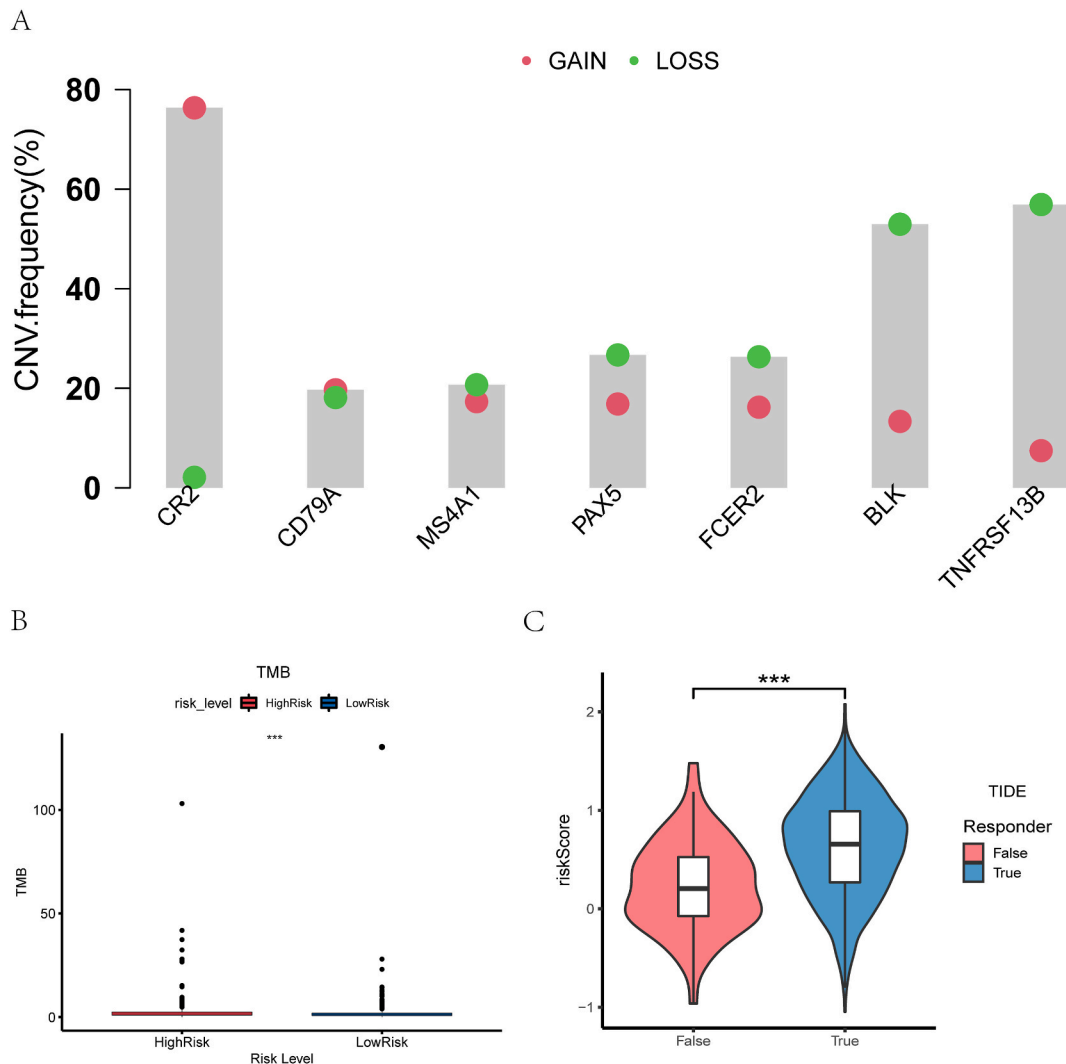
3.7. Mutational analysis of hub genes, and TMB and TIDE analyses

We investigated the copy number variations (CNVs) of seven hub genes (*BLK, PAX5, CR2, TNFRSF13B, FCER2, MS4A1, and CD79A*) in TCGA-BRCA dataset. We downloaded and combined the CNV data of patients with breast cancer and analyzed and visualized the results using GISTIC 2.0. (Fig. 4A). Seven hub genes demonstrated a high frequency of amplifications and deletions in breast cancer patient samples, with *CR2* being the most amplified and *TNFRSF13B* and *BLK* having the highest frequency of deletions.

An analysis of TCGA-BRCA dataset samples revealed significant differences in Tumor Mutation Burden (TMB) between different risk groups, illustrated in Fig. 4B. The TMB of breast cancer patient samples in TCGA-BRCA dataset exhibited statistically significant changes across the groups (p < 0.05). Fig. 4C illustrates how the TIDE algorithm was applied to the assessment of immunotherapy

Fig. 3. Analysis of immune infiltration. **(A–B)** Stacked histogram of CIBERSORT immune infiltration analysis in TCGA-BRCA dataset (A), Boxplots of comparison (B); **(C–D)** Correlation analysis of the infiltration abundance of 10 types of immune cells in the high-risk group (C) and low-risk group (D) in TCGA-BRCA dataset.

efficacy. The prognostic score of the responder group was significantly higher than that of the non-responder group (p < 0.01).

3.8. Risk model construction

Univariate analysis revealed a significant association between overall survival (OS) and prognostic score, age, and stage (p < 0.001 for each), as illustrated in Fig. 5A. Multivariate Cox regression analysis of significant factors from univariate Cox regression revealed that prognostic score (p < 0.001), Stage IV (p < 0.001), and age (p < 0.001) were attributes of OS (Fig. 5B). The time-dependent ROC curve of TCGA-BRCA data risk model was constructed using the R package, timeROC. The analysis revealed that the area under the

A

B

C

Fig. 4. Gene mutation analysis, and TMB and TIDE analyses. **(A)** CNV; **(B)** Boxplot of TMB between the high- and low-risk groups. **(C)** Boxplot of prognostic score of TIDE analysis between the responder and non-responder groups. (ordinate) prognostic score; (abscissa) grouping, (red) non-responder group, (blue) responder group. Asterisks indicates significance. TMB, tumor mutation burden; TIDE, tumor immune dysfunction and exclusion; CNV, copy number variation; TCGA, The cancer genome atlas; BRCA, Breast invasive carcinoma. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

curve (AUC) values for 1-, 3-, and 5-year OS were 0.604, 0.709, and 0.673, respectively (Fig. 5C). The results demonstrate the model's strong predictive performance. The Kaplan-Meier curve showed significantly lower OS for breast cancer patients in the high-risk group (p < 0.0001, Fig. 5D). We developed a nomogram (Fig. 5E) that integrated the prognostic score, stage, and age to comprehensively predict the OS status of patients with breast cancer.

The risk factor distribution for TCGA-BRCA risk model is displayed in Supplementary Fig. 8A. In the high-risk group, there was a higher incidence of deaths and increased expression of *IVL* and *CD24*; while the low-risk group demonstrated higher expression levels of *DOK7*, *CXCL13*, and *PIGR*. Incorporating prognostic score, age, and stage into the Decision Curve Analysis (DCA) resulted in 3-year and 5-year DCA curves that differed notably from the reference lines, underscoring the model's high level of precision and reliability (Supplementary Figs. 8B–D).

3.9. Efficacy assessment and external validation of the risk model

The risk score prediction model for the GSE20685 dataset was computed and created according to the multivariate Cox regression coefficient of TCGA-BRCA dataset. To establish whether the prognostic score and various clinicopathological variables were independent prognostic factors, we examined the GSE20685 dataset using Cox regression analysis. Based on univariate analysis, prognostic score (p < 0.001), age (p < 0.5), stage_T (p < 0.001), stage_N (p < 0.001) and stage_M (p < 0.001) were correlated with OS
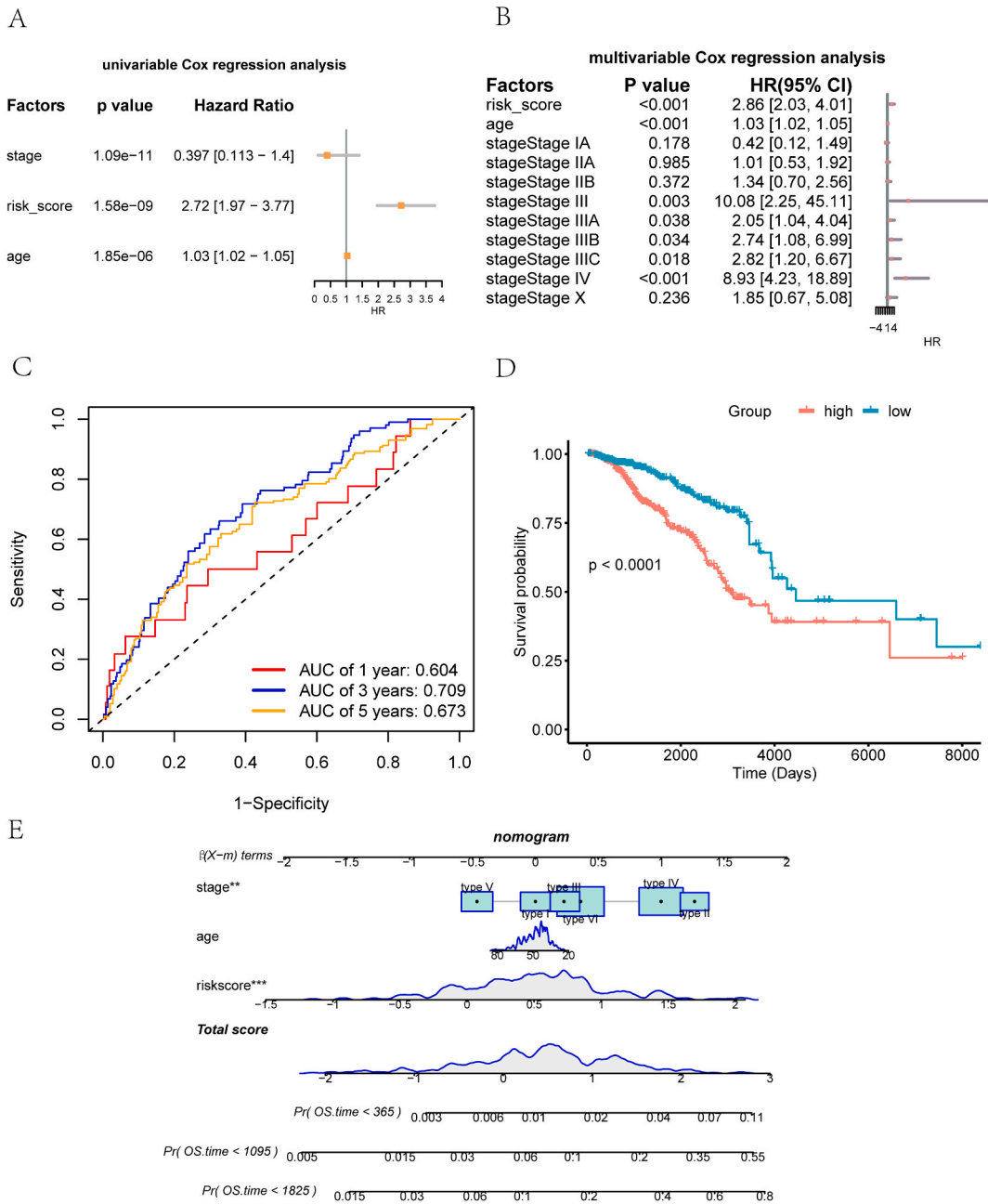
Fig. 5. Risk model for prediction of prognosis. **(A–B)** Forest plot; **(C)** Time-dependent ROC curves for breast cancer risk prediction model, the colors represent the discrimination accuracy (AUC value), with red indicating 1-year survival, blue indicating 3-year survival, and yellow indicating 5-year survival; **(D)** Kaplan-Meier curves; **(E)** Nomogram constructed using the prognostic score combined with age, stage, and clinicopathological features predicts the 1-, 2-, and 3-year OS rates of patients with breast cancer. ROC, Receiver operating characteristic; AUC, area under curve; OS, overall survival. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(Supplementary Fig. 9A). Multivariate Cox regression analysis of the significant factors revealed that the prognostic score (p < 0.001) and stage_N was significantly associated with OS (Supplementary Fig. 9B). The time-dependent ROC curve of the GSE20685 dataset risk model was constructed using the R package, timeROC. The analysis revealed that the AUC values for 1-, 3-, and 5-year OS were 0.755, 0.78, and 0.747, respectively (Supplementary Fig. 9C), indicating that the model has a good predictive ability. Compared with patients in the low-risk group, patients in the high-risk group had significantly shorter OS (p = 0.00038; Supplementary Fig. 9D). The 1-year, 3-year and 5-year DCA curves differed between the two reference lines, indicating that the model was precise and reliable (Supplementary Figs. 9E–G).

3.10. Hub genes and prediction of cancer status

The six exosome-related prognostic DEGs in TCGA-BRCA dataset exhibited statistically significant differences in expression between tumor and normal tissues (*ABCC9* (p < 0.001), *PIGR* (p < 0.001), *CXCL13* (p < 0.001), *DOK7* (p < 0.001), *CD24* (p < 0.001), and *IVL* (p < 0.001), Supplementary Fig. 10A). The ROC curves of the six exosome-related DEGs in TCGA-BRCA dataset are shown in Supplementary Figs. 10B–G. The expression of *ABCC9* (AUC = 0.832, Supplementary Fig. 10B), *CXCL13* (AUC = 0.726, Supplementary Fig. 10D), *DOK7* (AUC = 0.738, Supplementary Fig. 10E), *PIGR* (AUC = 0.839, Supplementary Fig. 10G) in TCGA-BRCA dataset demonstrated an association with the occurrence of breast cancer, despite a weaker correlation between the expression of *CD24* (AUC = 0.699, Supplementary Fig. 10C) and *IVL* (AUC = 0.635, Supplementary Fig. 10F) and the incidence of breast cancer.

For the GSE5764, GSE7904, and GSE29431 datasets devoid of prognostic information, we validated the ROC for cancer status prediction. In the GSE5764 dataset, *IVL* (AUC = 0.860, Supplementary Fig. 11E), *PIGR* (AUC = 0.860, Supplementary Fig. 11F), and *CXCL13* (AUC = 0.7, Supplementary Fig. 11C) could effectively distinguish invasive ductal carcinoma from normal tissues, whereas *ABCC9* (AUC = 0.6, Supplementary Fig. 11A), *CD24* (AUC = 0.61, Supplementary Fig. 11B), and *DOK7* (AUC = 0.56, Supplementary Fig. 11D) could not perform this distinction. In the GSE7904 dataset, *IVL* (AUC = 0.761, Supplementary Fig. 11K), *PIGR* (AUC = 0.987, Supplementary Fig. 11L), and *CD24* (AUC = 0.709, Supplementary Fig. 11H) distinguished breast cancer from normal tissue, whereas *ABCC9* (AUC = 0.605, Supplementary Fig. 11G), *DOK7* (AUC = 0.618, Supplementary Fig. 11J), and *CXCL13* (AUC = 0.568, Supplementary Fig. 11I) could not distinguish breast cancer from normal tissue. In the GSE29431 data set, *ABCC9* (AUC = 0.824, Supplementary Fig. 11M), *CD24* (AUC = 0.877, Supplementary Fig. 11N), and *CXCL13* (AUC = 0.781, Supplementary Fig. 11O) could well distinguish breast cancer from normal tissue, while *IVL* (AUC = 0.669), *PIGR* (AUC = 0.574), and *DOK7* (AUC = 0.596) could not adequately distinguish breast cancer from normal tissue (Supplementary Fig. 11P).

4. Discussion

Breast cancer is a prevalent and heterogeneous disease, primarily diagnosed in females. With approximately 2.3 million new cases reported each year, breast cancer is the leading cause of cancer-related mortality in the world [62]. Although comprehensive interventions can control the growth of breast cancer cells, recurrence and metastasis are common and result in treatment failure. Failure could be attributed to a small fraction of cells breaking away from the original tumor, migrating to nearby or distant organs, and forming secondary tumors. Understanding the mechanisms underlying metastasis is crucial for developing new and effective treatments for breast cancer.

Exosomes are critical regulators of intercellular communication [5] and regulators of oncogenaration, cancer advancement and immune system by altering the TME [4]. Exosomes present in the peripheral blood are highly durable and thus hold great promise as a source of tumor-derived materials for investigating tumor behavior. Monitoring exosomes and analyzing their protein and RNA content could enable their use in diagnostic applications [63]. In view of this, the molecular composition of breast cancer cells' exosomes holds great promise for non-invasive diagnosis and treatment of breast cancer. The ability to track and analyze exosomes from breast cancer cells can provide a wealth of information regarding tumor behavior and aid in the development of personalized treatment strategies.

Our goal was to develop a prognostic score for tumor-derived exosomes. The prognostic score offers insights into immunological states and predicts survival outcomes in patients with breast cancer. Our predictive model, validated on the GSE20685 dataset, demonstrated strong long-term prognostic power with AUCs of 0.755, 0.78, and 0.747 for 1-, 3-, and 5-year OS (Supplementary Fig. 9C). In comparison, a distinct study [64] focused on short-term outcomes reported AUCs of 0.68, 0.745, and 0.714 for 100, 200, and 300 days, emphasizing the variability in predictive timelines. Another research [65] showed promising long-term predictions for 2, 3, and 5 years with AUCs of 0.727, 0.691, and 0.695, while a further study [66] validated its risk score for breast cancer survival over similar periods, achieving AUCs of 0.714, 0.676, and 0.729. Our model's focus on longer-term predictions aligns better with clinical needs in breast cancer, emphasizing its utility in providing meaningful prognostic insights. Furthermore, According to Cox regression analysis, prognostic scores were independent prognostic factors (P < 0.001). Six novel tumor-derived exosome genes were identified in this study, including ATP Binding Cassette subfamily C member 9 (*ABCC9*), Polymeric Immunoglobulin Receptor (*PIGR*), C-X-C Motif Chemokine Ligand 13 (*CXCL13*), Docking Protein 7 (*DOK7*), CD24 Molecule (*CD24*), and involucrin (*IVL*), which significantly affected the prognostic score. These results offer valuable insights into the diagnostic and prognostic potential of exosomes in breast cancer.

As a member of the superfamily of ATP-binding cassette (ABC) transporters, *ABCC9* serves as a crucial component in the transportation of a diverse range of molecules across both the extracellular and intracellular membranes, including sugars, proteins, and lipids [67]. Specifically, the *ABCC9* gene encodes the muscle plasma protein activator K_{ATP} channel subunit SUR2, which is expressed in the heart, blood vessels, and pancreatic islet cells. It plays a significant role in regulating potassium ion channels on the cell membrane, associated with heart diseases and insulin secretion disorders. As part of its role in regulating K_{ATP} channel activity, *ABCC9* collaborates with potassium channel proteins [67,68]. ABC transporters can cause drug resistance by effluxing the drug out of cells, thereby reducing the bioavailability and effectiveness of cancer cells [69]. *PIGR* or the polymeric immunoglobulin receptor is a glycoprotein associated with exosomes and is important for the development and progression of tumors [70,71]. The *PIGR* gene encodes the polymeric Ig receptor, part of the immunoglobulin A receptor, and is highly expressed in mucosal epithelial cells. It is involved in mucosal immune defense functions, including the adhesion and transport of anti-pathogenic microbes [72]. The function and regulation of *PIGR* may be closely related to the immune defense system of the body [72]. Recent studies have shown that *PIGR* can serve as a biomarker for good outcomes in luminal breast cancer [73], and the current study suggests that high *PIGR* expression is associated with better survival. *CXCL13* has been identified as a key mediator of cellular interactions and a hallmark of allergic

inflammation, as reported in previous studies [74]. The *CXCL13* gene encodes chemokine 13 (C-X-C motif chemokine ligand 13) and is expressed in lymphoid tissue, playing a significant role in immune regulation, especially in lymph node formation and B-cell attraction [75]. Research suggests *CXCL13* plays an important role in shaping the immunoactive TME and enhancing PD-1 checkpoint blockade efficacy in ovarian cancer [76]. Although some researchers have suggested that the co-expression of *CXCL13* and *CXCR5* could promote disease progression and metastasis in breast tumors through chemotaxis of these cells [77], other studies have shown that CXCL13+ cells can respond effectively to anti-PD-L1 therapies [78]. Our findings highlight a correlation between *CXCL13* and better outcomes. These results emphasize the complex and multifaceted role of *CXCL13* in cancer development and treatment. Several studies have explored the correlation between *DOK7* and cancer. High *DOK7* expression has been associated with favorable prognosis in acute myeloid leukemia [79]. Conversely, reduced *DOK7* expression has been linked to poor survival in lung cancer [80] and the acquisition of tamoxifen resistance in breast cancer [81]. The *DOK7* gene encodes the kinase binding protein Dok-7, which is critical for the development and function maintenance of the neuromuscular junction [82]. This gene is expressed in nerve and muscle tissues and is involved in the aggregation of acetylcholine receptors and neuromuscular transduction signal transmission [82]. Based on recent studies, *DOK7* inhibits the proliferation, migration, and invasion of breast cancer cells through the PI3K/PTEN/AKT pathway [83]. Thus, *DOK7* may be an advantageous marker for breast cancer. The protein cluster of differentiation (CD) is known for its multifaceted functions and anti-inflammatory properties. The *CD24* gene encodes the cell surface antigen CD24, which is widely distributed in various tissues and cells [84]. According to recent research, *CD24* is overexpressed in various cancer cell lineages [85] and is an oncogenic marker [86]. The accumulation of *CD24* in the cytoplasm associated with tumor progression, lymph node positivity, and reduced patient survival [84]. Therapeutic blockade of *CD24* using monoclonal antibodies has been investigated in in vitro and in vivo animal models with promising results, including increased phagocytosis of cancer cells, tumor reduction, and improved surviva [87]. Oncogenic effectiveness was confirmed in this study. The Involucrin (*IVL*) gene encodes the intermediate filament protein 1 in keratinocyte-forming cells, which is highly expressed in the skin and other epithelial tissues [88]. It is involved in the formation and maintenance of the cytoskeleton and the structural stability of the stratum corneum [88]. *IVL* is localized in the cytoplasm and is crosslinked to membrane proteins by transglutaminase [89]. The expression of *IVL* in squamous cell carcinoma was found to be associated with poor prognosis [90]. *IVL* may play an important role in the prognosis and treatment of melanoma [91] and head and neck cancer [92].

In the present study, the enrichment of immune-related pathways was observed. B cell activation, humoral immune response, immunoglobulin complex, and endopeptidase inhibitor activity were enriched according to the GO enrichment analyses (Fig. 2B). Kyoto Encyclopedia of Genes and Genomes analysis revealed that the primary immunodeficiency pathway was enriched (Fig. 2B). GSEA indicated that the DEGs were implicated in immune-related pathways, including the IL18 signaling pathway (Supplementary Fig. 5C), apoptosis (Supplementary Fig. 5D), Toll-like receptor signaling pathway (Supplementary Fig. 5F), regulatory circuits of the STAT3 signaling pathway (Supplementary Fig. 5G), and the Mapk signaling pathway (Supplementary Fig. 5H).

The immune system is responsible for defending the body against tumors. Several studies have emphasized the importance of the tumor immune microenvironment in cancer development [63,93]. The identification of a new approach for categorizing patients suitable for immunotherapy is challenging. Therefore, we examined the relationship between prognostic score and tumor immune microenvironment. When the potential efficacy of immunotherapy was evaluated, the prognostic score of the responder group was found to be substantially higher ($p < 0.01$). This therapeutic strategy based on biomarkers can be utilized in the future to generate a variety of intriguing combination therapy strategies.

This study presented certain limitations. The research utilized data samples sourced from various origins, potentially introducing elements of selection and sample bias. Moreover, validating our findings through clinical specimens would have strengthened the study. The retrospective design and in silico analysis constrained our capacity to perform detailed analyses of senescence and exhaustion phenotypes at the single-cell level. Not conducting crosstalk analysis limited our insight into gene-pathway interactions and potential therapeutic implications. Future research should concentrate on this aspect to enhance comprehension of exosomal biomarkers and their utility in diagnosis and prognosis.

5. Conclusion

A model that predicts breast cancer prognosis based on exosomes associated with immune infiltration has been established and validated. Overall, the ERG signature shows outstanding performance in predicting the prognosis and cancer status of patients with breast cancer. These findings provide novel insights into the treatment of breast cancer.

Data availability statement

The Cancer Genome Atlas Breast Cancer (TCGA-BRCA) dataset, Gene Expression Omnibus (GEO) datasets GSE20685, GSE5764, GSE7904, and GSE29431. Access to the TCGA-BRCA dataset can be obtained through the National Cancer Institute's Genomic Data Commons. GEO datasets are publicly available and can be accessed through the GEO repository by searching the respective accession numbers. These datasets were used under the terms of access provided by their respective repositories.

CRediT authorship contribution statement

Qitong Chen: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Data curation. **Qin Zhou:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Formal analysis, Data

curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Appendix ASupplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e29551.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, Epub 2021/02/05, CA A Cancer J. Clin. 71 (3) (2021 May) 209–249, https://doi.org/10.3322/caac.21660. Cited in: Pubmed; PMID 33538338.

[2] N. Harbeck, F. Penault-Llorca, J. Cortes, M. Gnant, N. Houssami, P. Poortmans, et al., Breast cancer, Epub 20190923, Nat. Rev. Dis. Prim. 5 (1) (2019 Sep 23) 66, https://doi.org/10.1038/s41572-019-0111-2. Cited in: Pubmed; PMID 31548545.

[3] B. Kinnel, S.K. Singh, G. Oprea-Ilies, R. Singh, Targeted therapy and mechanisms of drug resistance in breast cancer, The authors declare that they have no competing interests. Epub 20230219, Cancers 15 (4) (2023 Feb 19), https://doi.org/10.3390/cancers15041320. Cited in: Pubmed; PMID 36831661.

[4] C. Piombino, I. Mastrolia, C. Omarini, O. Candini, M. Dominici, F. Piacentini, et al., The role of exosomes in breast cancer diagnosis, The authors declare no conflict of interest associated with this manuscript. Epub 20210318, Biomedicines 9 (3) (2021 Mar 18), https://doi.org/10.3390/biomedicines9030312. Cited in: Pubmed; PMID 33803776.

[5] J.C. Akers, D. Gonda, R. Kim, B.S. Carter, C.C. Chen, Biogenesis of extracellular vesicles (EV): exosomes, microvesicles, retrovirus-like vesicles, and apoptotic bodies, Epub 20130302, J. Neuro Oncol. 113 (1) (2013 May) 1–11, https://doi.org/10.1007/s11060-013-1084-8. Cited in: Pubmed; PMID 23456661.

[6] C. Harding, P. Stahl, Transferrin recycling in reticulocytes: pH and iron are important determinants of ligand binding and processing, Biochem. Biophys. Res. Commun. 113 (2) (1983 Jun 15) 650–658, https://doi.org/10.1016/0006-291x(83)91776-x. Cited in: Pubmed; PMID 6870878.

[7] B.T. Pan, R.M. Johnstone, Fate of the transferrin receptor during maturation of sheep reticulocytes in vitro: selective externalization of the receptor, Cell 33 (3) (1983 Jul) 967–978, https://doi.org/10.1016/0092-8674(83)90040-5. Cited in: Pubmed; PMID 6307529.

[8] T. Groot Kormelink, S. Mol, E.C. de Jong, M.H.M. Wauben, The role of extracellular vesicles when innate meets adaptive, Epub 20180403, Semin. Immunopathol. 40 (5) (2018 Sep) 439–452, https://doi.org/10.1007/s00281-018-0681-1. Cited in: Pubmed; PMID 29616308.

[9] M.P. Bebelman, M.J. Smit, D.M. Pegtel, S.R. Baglio, Biogenesis and function of extracellular vesicles in cancer, Epub 20180221, Pharmacol. Ther. 188 (2018 Aug) 1–11, https://doi.org/10.1016/j.pharmthera.2018.02.013. Cited in: Pubmed; PMID 29476772.

[10] H. Peinado, M. Aleckovic, S. Lavotshkin, I. Matei, B. Costa-Silva, G. Moreno-Bueno, et al., Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET, Nat. Med. 18 (6) (2012 Jun) 883–891, https://doi.org/10.1038/nm.2753. Cited in: Pubmed; PMID 22635005.

[11] M. Yanez-Mo, P.R. Siljander, Z. Andreu, A.B. Zavec, F.E. Borras, E.I. Buzas, et al., Biological properties of extracellular vesicles and their physiological functions, Epub 20150514, J. Extracell. Vesicles 4 (2015) 27066, https://doi.org/10.3402/jev.v4.27066. Cited in: Pubmed; PMID 25979354.

[12] A. Bobrie, M. Colombo, G. Raposo, C. Thery, Exosome secretion: molecular mechanisms and roles in immune responses, Epub 20110630, Traffic 12 (12) (2011 Dec) 1659–1668, https://doi.org/10.1111/j.1600-0854.2011.01225.x. Cited in: Pubmed; PMID 21645191.

[13] L.B. Kennedy, A.K.S. Salama, A review of cancer immunotherapy toxicity, Epub 20200116, CA A Cancer J. Clin. 70 (2) (2020 Mar) 86–104, https://doi.org/10.3322/caac.21596. Cited in: Pubmed; PMID 31944278.

[14] Y.W. Yi, Therapeutic implications of the drug resistance conferred by extracellular vesicles derived from triple-negative breast cancer cells, The author declares no conflict of interest. Epub 20230212, Int. J. Mol. Sci. 24 (4) (2023 Feb 12), https://doi.org/10.3390/ijms24043704. Cited in: Pubmed; PMID 36835116.

[15] H.D. Zhang, L.H. Jiang, J.C. Hou, S.L. Zhong, L.P. Zhu, D.D. Wang, et al., Exosome: a novel mediator in drug resistance of cancer cells, Epub 2018/10/13, Epigenomics 10 (11) (2018 Nov) 1499–1509, https://doi.org/10.2217/epi-2017-0151. Cited in: Pubmed; PMID 30309258.

[16] D. Hanahan, L.M. Coussens, Accessories to the crime: functions of cells recruited to the tumor microenvironment, Cancer Cell 21 (3) (2012 Mar 20) 309–322, https://doi.org/10.1016/j.ccr.2012.02.022. Cited in: Pubmed; PMID 22439926.

[17] A. Byrne, P. Savas, S. Sant, R. Li, B. Virassamy, S.J. Luen, et al., Tissue-resident memory T cells in breast cancer control and immunotherapy responses, Epub 20200228, Nat. Rev. Clin. Oncol. 17 (6) (2020 Jun) 341–348, https://doi.org/10.1038/s41571-020-0333-y. Cited in: Pubmed; PMID 32112054.

[18] C. Zhang, Q. Ji, Y. Yang, Q. Li, Z. Wang, Exosome: function and role in cancer metastasis and drug resistance, Epub 2018/04/24, Technol. Cancer Res. Treat. 17 (2018 Jan 1) 1533033818763450, https://doi.org/10.1177/1533033818763450. Cited in: Pubmed; PMID 29681222.

[19] C. Lyu, H. Sun, Z. Sun, Y. Liu, Q. Wang, Roles of exosomes in immunotherapy for solid cancers, Epub 2024/02/02, Cell Death Dis. 15 (2) (2024 Feb 1) 106, https://doi.org/10.1038/s41419-024-06494-z. Cited in: Pubmed; PMID 38302430.

[20] H.X. Li, S.Q. Wang, Z.X. Lian, S.L. Deng, K. Yu, Relationship between tumor infiltrating immune cells and tumor metastasis and its prognostic value in cancer, Epub 2023/01/09, Cells 12 (1) (2022 Dec 23), https://doi.org/10.3390/cells12010064. Cited in: Pubmed; PMID 36611857.

[21] M.V. Dieci, F. Miglietta, V. Guarneri, Immune infiltrates in breast cancer: recent updates and clinical implications, Epub 2021/01/28, Cells 10 (2) (2021 Jan 23), https://doi.org/10.3390/cells10020223. Cited in: Pubmed; PMID 33498711.

[22] Chen H., Sun Y., Yang Z., Yin S., Li Y., Tang M., et al., Metabolic heterogeneity and immunocompetence of infiltrating immune cells in the breast cancer microenvironment (Review) Oncol. Rep. 45 (3) (2021 Mar) 846–856 Cited in: Pubmed; PMID 33650671, doi:10.3892/or.2021.7946.

[23] G. Cui, C. Wang, J. Liu, K. Shon, R. Gu, C. Chang, et al., Development of an exosome-related and immune microenvironment prognostic signature in colon adenocarcinoma, eng. Epub 2022/10/01, Front. Genet. 13 (2022) 995644, https://doi.org/10.3389/fgene.2022.995644. Cited in: Pubmed; PMID 36176299.

[24] M. Ding, Q. Xu, X. Jin, Z. Han, H. Jiang, H. Sun, et al., Novel exosome-related risk signature as prognostic biomarkers in glioblastoma, eng. Epub 2023/03/04, Front. Immunol. 14 (2023) 1071023, https://doi.org/10.3389/fimmu.2023.1071023. Cited in: Pubmed; PMID 36865549.

[25] C. Li, Z. Zhang, E. Peng, J. Peng, Role of an exosomes-related lncRNAs signature in tumor immune microenvironment of gastric cancer, eng. Epub 2022/04/26, Front. Cell Dev. Biol. 10 (2022) 873319, https://doi.org/10.3389/fcell.2022.873319. Cited in: Pubmed; PMID 35465325.

[26] H. Shao, L. Yao, Y. Tao, X. Huang, Identification and verification of an exosome-related gene risk model to predict prognosis and evaluate immune infiltration for colorectal cancer, eng. Epub 2023/10/06, Medicine (Baltim.) 102 (40) (2023 Oct 6) e35365, https://doi.org/10.1097/md.0000000000035365. Cited in: Pubmed; PMID 37800824.

[27] Z. Yang, X. Li, C. Pan, Y. Li, L. Lin, Y. Jin, et al., A comprehensive study based on exosome-related immunosuppression genes and tumor microenvironment in hepatocellular carcinoma, eng. Epub 2022/12/23, BMC Cancer 22 (1) (2022 Dec 22) 1344, https://doi.org/10.1186/s12885-022-10463-0. Cited in: Pubmed; PMID 36550445.

[28] Y. Yue, J. Tao, D. An, L. Shi, A prognostic exosome-related long non-coding RNAs risk model related to the immune microenvironment and therapeutic responses for patients with liver hepatocellular carcinoma, eng. Epub 2024/01/31, Heliyon 10 (2) (2024 Jan 30) e24462, https://doi.org/10.1016/j.heliyon.2024.e24462. Cited in: Pubmed; PMID 38293480.

[29] F. Zhao, Z. Li, Z. Dong, Z. Wang, P. Guo, D. Zhang, et al., Exploring the potential of exosome-related LncRNA pairs as predictors for immune microenvironment, survival outcome, and microbiotain landscape in esophageal squamous cell carcinoma, eng. Epub 2022/07/27, Front. Immunol. 13 (2022) 918154, https://doi.org/10.3389/fimmu.2022.918154. Cited in: Pubmed; PMID 35880180.

[30] Y. You, Z. Du, G. Xu, Z. Tian, M. Xiao, Y. Wang, Identification of exosome-related genes associated with prognosis and immune infiltration features in head-neck squamous cell carcinoma, Epub 2023/06/28, Biomolecules 13 (6) (2023 Jun 7), https://doi.org/10.3390/biom13060958. Cited in: Pubmed; PMID 37371537.

[31] L. Laugier, A.F. Frade, F.M. Ferreira, M.A. Baron, P.C. Teixeira, S. Cabantous, et al., Whole-genome cardiac DNA methylation fingerprint and gene expression analysis provide new insights in the pathogenesis of chronic Chagas disease cardiomyopathy, eng. Epub 2017/06/03, Clin. Infect. Dis. 65 (7) (2017 Oct 1) 1103–1111, https://doi.org/10.1093/cid/cix506. Cited in: Pubmed; PMID 28575239.

[32] M.J. Goldman, B. Craft, M. Hastie, K. Repečka, F. McDade, A. Kamath, et al., Visualizing and interpreting cancer genomics data via the Xena platform, eng. Epub 2020/05/24, Nat. Biotechnol. 38 (6) (2020 Jun) 675–678, https://doi.org/10.1038/s41587-020-0546-8. Cited in: Pubmed; PMID 32444850.

[33] J. Gao, B.A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S.O. Sumer, et al., Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal, eng. Epub 2013/04/04, Sci. Signal. 6 (269) (2013 Apr 2) pl1, https://doi.org/10.1126/scisignal.2004088. Cited in: Pubmed; PMID 23550210.

[34] K.J. Kao, K.M. Chang, H.C. Hsu, A.T. Huang, Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization, eng. Epub 2011/04/20, BMC Cancer 11 (2011 Apr 18) 143, https://doi.org/10.1186/1471-2407-11-143. Cited in: Pubmed; PMID 21501481.

[35] G. Turashvili, J. Bouchal, K. Baumforth, W. Wei, M. Dziechciarkova, J. Ehrmann, et al., Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis, eng. Epub 2007/03/29, BMC Cancer 7 (2007 Mar 27) 55, https://doi.org/10.1186/1471-2407-7-55. Cited in: Pubmed; PMID 17389037.

[36] A.L. Richardson, Z.C. Wang, A. De Nicolo, X. Lu, M. Brown, A. Miron, et al., X chromosomal abnormalities in basal-like human breast cancer, eng. Epub 2006/02/14, Cancer Cell 9 (2) (2006 Feb) 121–132, https://doi.org/10.1016/j.ccr.2006.01.013. Cited in: Pubmed; PMID 16473279.

[37] S. Davis, P.S. Meltzer, GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor, 7. eng. Epub 2007/05/15, Bioinformatics 23 (14) (2007 Jul 15) 1846, https://doi.org/10.1093/bioinformatics/btm254. Cited in: Pubmed; PMID 17496320.

[38] X. Wang, Z. Chai, G. Pan, Y. Hao, B. Li, T. Ye, et al., ExoBCD: a comprehensive database for exosomal biomarker discovery in breast cancer, Briefings Bioinf. 22 (3) (2021 May 20), https://doi.org/10.1093/bib/bbaa088. Cited in: Pubmed; PMID 32591816.

[39] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, eng. Epub 2015/01/22, Nucleic Acids Res. 43 (7) (2015 Apr 20) e47, https://doi.org/10.1093/nar/gkv007. Cited in: Pubmed; PMID 25605792.

[40] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, eng. Epub 2013/01/18, BMC Bioinf. 14 (7) (2013 Jan 16), https://doi.org/10.1186/1471-2105-14-7. Cited in: Pubmed; PMID 23323831.

[41] B. Zhang, S. Horvath, A general framework for weighted gene co-expression network analysis, 4:Article17. eng. Epub 2006/05/02, Stat. Appl. Genet. Mol. Biol. (2005), https://doi.org/10.2202/1544-6115.1128. Cited in: Pubmed; PMID 16646834.

[42] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, eng. Epub 2008/12/31, BMC Bioinf. 9 (2008 Dec 29) 559, https://doi.org/10.1186/1471-2105-9-559. Cited in: Pubmed; PMID 19114008.

[43] A. Mayakonda, D.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, eng. Epub 2018/10/21, Genome Res. 28 (11) (2018 Nov) 1747–1756, https://doi.org/10.1101/gr.239244.118. Cited in: Pubmed; PMID 30341162.

[44] T. Emura, S. Matsui, H.Y. Chen, compound.Cox: univariate feature selection and compound covariate for predicting survival, Epub 2018/12/12, Comput. Methods Progr. Biomed. 168 (2019 Jan) 21–37, https://doi.org/10.1016/j.cmpb.2018.10.020. Cited in: Pubmed; PMID 30527130.

[45] S. Engebretsen, J. Bohlin, Statistical predictions with glmnet, eng. Epub 2019/08/25, Clin. Epigenet. 11 (1) (2019 Aug 23) 123, https://doi.org/10.1186/s13148-019-0730-1. Cited in: Pubmed; PMID 31443682.

[46] W. Cai, M. van der Laan, Nonparametric bootstrap inference for the targeted highly adaptive least absolute shrinkage and selection operator (LASSO) estimator, eng. Epub 2020/08/11, Int. J. Biostat. (2020 Aug 10), https://doi.org/10.1515/ijb-2017-0070. Cited in: Pubmed; PMID 32772002.

[47] C.T. Yeh, G.Y. Liao, T. Emura, Sensitivity analysis for survival prognostic prediction with gene selection: a copula method for dependent censoring, Epub 2023/03/30, Biomedicines 11 (3) (2023 Mar 6), https://doi.org/10.3390/biomedicines11030797. Cited in: Pubmed; PMID 36979776.

[48] S.Y. Park, Nomogram: an analogue tool to deliver digital knowledge, eng. Epub 2018/01/27, J. Thorac. Cardiovasc. Surg. 155 (4) (2018 Apr) 1793, https://doi.org/10.1016/j.jtcvs.2017.12.107. Cited in: Pubmed; PMID 29370910.

[49] G. Yu, Gene Ontology semantic similarity analysis using GOSemSim, eng. Epub 2020/01/22, Methods Mol. Biol. 2117 (2020) 207–215, https://doi.org/10.1007/978-1-0716-0301-7_11. Cited in: Pubmed; PMID 31960380.

[50] H. Ogata, S. Goto, K. Sato, W. Fujibuchi, H. Bono, M. Kanehisa, KEGG: Kyoto Encyclopedia of genes and genomes, eng. Epub 1998/12/10, Nucleic Acids Res. 27 (1) (1999 Jan 1) 29–34, https://doi.org/10.1093/nar/27.1.29. Cited in: Pubmed; PMID 9847135.

[51] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, eng. Epub 2012/03/30, OMICS A J. Integr. Biol. 16 (5) (2012 May) 284–287, https://doi.org/10.1089/omi.2011.0118. Cited in: Pubmed; PMID 22455463.

[52] W. Luo, C. Brouwer, Pathview: an R/Bioconductor package for pathway-based data integration and visualization, 1. eng. Epub 2013/06/07, Bioinformatics 29 (14) (2013 Jul 15) 1830, https://doi.org/10.1093/bioinformatics/btt285. Cited in: Pubmed; PMID 23740750.

[53] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, eng. Epub 2005/10/04, Proc. Natl. Acad. Sci. U.S.A. 102 (43) (2005 Oct 25) 15545–15550, https://doi.org/10.1073/pnas.0506580102. Cited in: Pubmed; PMID 16199517.

[54] A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdottir, P. Tamayo, J.P. Mesirov, Molecular signatures database (MSigDB) 3.0, Epub 20110505, Bioinformatics 27 (12) (2011 Jun 15) 1739–1740, https://doi.org/10.1093/bioinformatics/btr260. Cited in: Pubmed; PMID 21546393.

[55] C. von Mering, M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, B. Snel, STRING: a database of predicted functional associations between proteins, eng. Epub 2003/01/10, Nucleic Acids Res. 31 (1) (2003 Jan 1) 258–261, https://doi.org/10.1093/nar/gkg034. Cited in: Pubmed; PMID 12519996.

[56] J.H. Li, S. Liu, H. Zhou, L.H. Qu, J.H. Yang, starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data, eng. Epub 2013/12/04, Nucleic Acids Res. 42 (Database issue) (2014 Jan) D92–D97, https://doi.org/10.1093/nar/gkt1248. Cited in: Pubmed; PMID 24297251.

[57] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, et al., Robust enumeration of cell subsets from tissue expression profiles, eng. Epub 2015/03/31, Nat. Methods 12 (5) (2015 May) 453–457, https://doi.org/10.1038/nmeth.3337. Cited in: Pubmed; PMID 25822800.

[58] P. Charoentong, F. Finotello, M. Angelova, C. Mayer, M. Efremova, D. Rieder, et al., Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, eng. Epub 2017/01/05, Cell Rep. 18 (1) (2017 Jan 3) 248–262, https://doi.org/10.1016/j.celrep.2016.12.019. Cited in: Pubmed; PMID 28052254.

[59] D.A. Barbie, P. Tamayo, J.S. Boehm, S.Y. Kim, S.E. Moody, I.F. Dunn, et al., Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1, 12. eng. Epub 2009/10/23, Nature 462 (7269) (2009 Nov 5) 108, https://doi.org/10.1038/nature08460. Cited in: Pubmed; PMID 19847166.

[60] G. Brière, É. Darbo, P. Thébault, R. Uricaru, Consensus clustering applied to multi-omics disease subtyping, eng. Epub 2021/07/08, BMC Bioinf. 22 (1) (2021 Jul 6) 361, https://doi.org/10.1186/s12859-021-04279-1. Cited in: Pubmed; PMID 34229612.

[61] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, eng. Epub 2010/04/30, Bioinformatics 26 (12) (2010 Jun 15) 1572–1573, https://doi.org/10.1093/bioinformatics/btq170. Cited in: Pubmed; PMID 20427518.

[62] R.L. Siegel, K.D. Miller, H.E. Fuchs, A. Jemal, Cancer statistics, 2022, Epub 2022/01/13, CA A Cancer J. Clin. 72 (1) (2022 Jan) 7–33, https://doi.org/10.3322/caac.21708. Cited in: Pubmed; PMID 35020204.

[63] X.X. Li, L.X. Yang, C. Wang, H. Li, D.S. Shi, J. Wang, The roles of exosomal proteins: classification, function, and applications, All authors declare that there is no conflict of interest. Epub 20230204, Int. J. Mol. Sci. 24 (4) (2023 Feb 4), https://doi.org/10.3390/ijms24043061. Cited in: Pubmed; PMID 36834471.

[64] H. Wang, R. Wang, L. Luo, J. Hong, X. Chen, K. Shen, et al., An exosome-based specific transcriptomic signature for profiling regulation patterns and modifying tumor immune microenvironment infiltration in triple-negative breast cancer, Epub 2023/12/21, Front. Immunol. 14 (2023) 1295558, https://doi.org/10.3389/fimmu.2023.1295558. Cited in: Pubmed; PMID 38124743.

[65] Q. Guo, K. Pan, P. Qiu, Z. Liu, J. Chen, J. Lin, Identification of an exosome-related signature associated with prognosis and immune infiltration in breast cancer, Epub 2023/10/25, Sci. Rep. 13 (1) (2023 Oct 24) 18198, https://doi.org/10.1038/s41598-023-45325-7. Cited in: Pubmed; PMID 37875600.

[66] P. Qiu, Q. Guo, J. Lin, K. Pan, J. Chen, M. Ding, An exosome-related long non-coding RNAs risk model could predict survival outcomes in patients with breast cancer, Epub 2022/12/25, Sci. Rep. 12 (1) (2022 Dec 24) 22322, https://doi.org/10.1038/s41598-022-26894-5. Cited in: Pubmed; PMID 36566321.

[67] K. Ando, L. Tong, D. Peng, E. Vazquez-Liebanas, H. Chiyoda, L. He, et al., KCNJ8/ABCC9-containing K-ATP channel modulates brain vascular smooth muscle development and neurovascular coupling, Epub 2022/05/20, Dev. Cell 57 (11) (2022 Jun 6) 1383–1399 e7, https://doi.org/10.1016/j.devcel.2022.04.019. Cited in: Pubmed; PMID 35588738.

[68] A. Akrouh, S.E. Halcomb, C.G. Nichols, M. Sala-Rabanal, Molecular biology of K(ATP) channels and implications for health and disease, Epub 2009/09/30, IUBMB Life 61 (10) (2009 Oct) 971–978, https://doi.org/10.1002/iub.246. Cited in: Pubmed; PMID 19787700.

[69] E.L. Giddings, D.P. Champagne, M.H. Wu, J.M. Laffin, T.M. Thornton, F. Valenca-Pereira, et al., Mitochondrial ATP fuels ABC transporter-mediated drug efflux in cancer chemoresistance, Epub 2021/05/16, Nat. Commun. 12 (1) (2021 May 14) 2804, https://doi.org/10.1038/s41467-021-23071-6. Cited in: Pubmed; PMID 33990571.

[70] Y. Liu, Y. Hu, L. Deng, The underlying roles of exosome-associated PIGR in fatty acid metabolism and immune signaling in colorectal cancer, Epub 2022/09/27, JAMA Oncol. 2022 (2022) 4675683, https://doi.org/10.1155/2022/4675683. Cited in: Pubmed; PMID 36157233.

[71] S.K. Tey, S.W.K. Wong, J.Y.T. Chan, X. Mao, T.H. Ng, C.L.S. Yeung, et al., Patient pIgR-enriched extracellular vesicles drive cancer stemness, tumorigenesis and metastasis in hepatocellular carcinoma, Epub 2021/12/20, J. Hepatol. 76 (4) (2022 Apr) 883–895, https://doi.org/10.1016/j.jhep.2021.12.005. Cited in: Pubmed; PMID 34922977.

[72] F.E. Johansen, C.S. Kaetzel, Regulation of the polymeric immunoglobulin receptor and IgA transport: new advances in environmental factors that stimulate pIgR expression and its role in mucosal immunity, Epub 2011/10/01, Mucosal Immunol. 4 (6) (2011 Nov) 598–602, https://doi.org/10.1038/mi.2011.37. Cited in: Pubmed; PMID 21956244.

[73] B. Xiao, L. Chen, Y. Ke, J. Hang, L. Cao, R. Zhang, et al., Identification of methylation sites and signature genes with prognostic value for luminal breast cancer, Epub 2018/04/13, BMC Cancer 18 (1) (2018 Apr 11) 405, https://doi.org/10.1186/s12885-018-4314-9. Cited in: Pubmed; PMID 29642861.

[74] Y. Kwon, Y. Choi, M. Kim, M.S. Jeong, H.S. Jung, D. Jeoung, HDAC6 and CXCL13 mediate atopic dermatitis by regulating cellular interactions and expression levels of miR-9 and SIRT1, Epub 2021/10/01, Front. Pharmacol. 12 (2021) 691279, https://doi.org/10.3389/fphar.2021.691279. Cited in: Pubmed; PMID 34588978.

[75] D.F. Legler, M. Loetscher, R.S. Roos, I. Clark-Lewis, M. Baggiolini, B. Moser, B cell-attracting chemokine 1, a human CXC chemokine expressed in lymphoid tissues, selectively attracts B lymphocytes via BLR1/CXCR5, Epub 1998/03/28, J. Exp. Med. 187 (4) (1998 Feb 16) 655–660, https://doi.org/10.1084/jem.187.4.655. Cited in: Pubmed; PMID 9463416.

[76] M. Yang, J. Lu, G. Zhang, Y. Wang, M. He, Q. Xu, et al., CXCL13 shapes immunoactive tumor microenvironment and enhances the efficacy of PD-1 checkpoint blockade in high-grade serous ovarian cancer, Epub 2021/01/17, J Immunother Cancer 9 (1) (2021 Jan), https://doi.org/10.1136/jitc-2020-001136. Cited in: Pubmed; PMID 33452206.

[77] S. Biswas, S. Roy Chowdhury, G. Mandal, S. Purohit, A. Gupta, A. Bhattacharyya, RelA driven co-expression of CXCL13 and CXCR5 is governed by a multifaceted transcriptional program regulating breast cancer progression, Epub 2018/12/16, Biochim. Biophys. Acta, Mol. Basis Dis. 1865 (2) (2019 Feb 1) 502–511, https://doi.org/10.1016/j.bbadis.2018.12.002. Cited in: Pubmed; PMID 30553016.

[78] Y. Zhang, H. Chen, H. Mo, X. Hu, R. Gao, Y. Zhao, et al., Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer, Epub 2021/10/16, Cancer Cell 39 (12) (2021 Dec 13) 1578–1593 e8, https://doi.org/10.1016/j.ccell.2021.09.010. Cited in: Pubmed; PMID 34653365.

[79] L. Zhang, R. Li, K. Hu, Y. Dai, Y. Pang, Y. Jiao, et al., Prognostic role of DOK family adapters in acute myeloid leukemia, Epub 2018/10/24, Cancer Gene Ther. 26 (9–10) (2019 Sep) 305–312, https://doi.org/10.1038/s41417-018-0052-z. Cited in: Pubmed; PMID 30348947.

[80] G. Chen, H. Yu, L. Satherley, C. Zabkiewicz, J. Resaul, H. Zhao, et al., The downstream of tyrosine kinase 7 is reduced in lung cancer and is associated with poor survival of patients with lung cancer, Epub 2017/04/11, Oncol. Rep. 37 (5) (2017 May) 2695–2701, https://doi.org/10.3892/or.2017.5538. Cited in: Pubmed; PMID 28393246.

[81] E. Gowdini, S.A. Aleyasin, N. Ramezani, N. Nafisi, M. Tutuni, DOK7 CpG hypermethylation in blood leukocytes as an epigenetic biomarker for acquired tamoxifen resistant in breast cancer, Epub 2022/11/14, J. Hum. Genet. 68 (1) (2023 Jan) 33–38, https://doi.org/10.1038/s10038-022-01092-3. Cited in: Pubmed; PMID 36372800.

[82] K. Okada, A. Inoue, M. Okada, Y. Murata, S. Kakuta, T. Jigami, et al., The muscle protein Dok-7 is essential for neuromuscular synaptogenesis, Epub 2006/06/24, Science 312 (5781) (2006 Jun 23) 1802–1805, https://doi.org/10.1126/science.1127142. Cited in: Pubmed; PMID 16794080.

[83] C. Yue, Y. Bai, Y. Piao, H. Liu, DOK7 inhibits cell proliferation, migration, and invasion of breast cancer via the PI3K/PTEN/AKT pathway, Epub 2021/02/09, JAMA Oncol. 2021 (2021) 4035257, https://doi.org/10.1155/2021/4035257. Cited in: Pubmed; PMID 33552156.

[84] P. Altevogt, M. Sammar, L. Huser, G. Kristiansen, Novel insights into the function of CD24: a driving force in cancer, Epub 2020/08/14, Int. J. Cancer 148 (3) (2021 Feb 1) 546–559, https://doi.org/10.1002/ijc.33249. Cited in: Pubmed; PMID 32790899.

[85] G. Kristiansen, K.J. Winzer, E. Mayordomo, J. Bellach, K. Schluns, C. Denkert, et al., CD24 expression is a new prognostic marker in breast cancer, Clin. Cancer Res. 9 (13) (2003 Oct 15) 4906–4913. Epub 2003/10/29. Cited in: Pubmed; PMID 14581365.

[86] X. Fang, P. Zheng, J. Tang, Y. Liu, CD24: from A to Z, Epub 2010/02/16, Cell. Mol. Immunol. 7 (2) (2010 Mar) 100–103, https://doi.org/10.1038/cmi.2009.119. Cited in: Pubmed; PMID 20154703.

[87] A.A. Barkal, R.E. Brewer, M. Markovic, M. Kowarsky, S.A. Barkal, B.W. Zaro, et al., CD24 signalling through macrophage Siglec-10 is a target for cancer immunotherapy, Epub 2019/08/02, Nature 572 (7769) (2019 Aug) 392–396, https://doi.org/10.1038/s41586-019-1456-0. Cited in: Pubmed; PMID 31367043.

[88] R.L. Eckert, H. Green, Structure and evolution of the human involucrin gene, Epub 1986/08/15, Cell 46 (4) (1986 Aug 15) 583–589, https://doi.org/10.1016/0092-8674(86)90884-6. Cited in: Pubmed; PMID 2873896.

[89] E.A. Brettmann, C. de Guzman Strong, Recent evolution of the human skin barrier, Epub 2018/05/23, Exp. Dermatol. 27 (8) (2018 Aug) 859–866, https://doi.org/10.1111/exd.13689. Cited in: Pubmed; PMID 29787621.

[90] S. Pandey, T.M. Soland, I.H. Bjerkli, L.P. Sand, F.C. Petersen, D.E. Costea, et al., Combined loss of expression of involucrin and cytokeratin 13 is associated with poor prognosis in squamous cell carcinoma of mobile tongue, Epub 2021/08/03, Head Neck 43 (11) (2021 Nov) 3374–3385, https://doi.org/10.1002/hed.26826. Cited in: Pubmed; PMID 34338386.

[91] Y. Han, X. Li, J. Yan, C. Ma, X. Wang, H. Pan, et al., Bioinformatic analysis identifies potential key genes in the pathogenesis of melanoma, Epub 2020/11/13, Front. Oncol. 10 (2020) 581985, https://doi.org/10.3389/fonc.2020.581985. Cited in: Pubmed; PMID 33178610.

[92] B. Liu, G. Huang, H. Zhu, Z. Ma, X. Tian, L. Yin, et al., Analysis of gene co-expression network reveals prognostic significance of CNFN in patients with head and neck cancer, Epub 2019/03/01, Oncol. Rep. 41 (4) (2019 Apr) 2168–2180, https://doi.org/10.3892/or.2019.7019. Cited in: Pubmed; PMID 30816522.

[93] S. Yin, Z. Chen, D. Chen, D. Yan, Strategies targeting PD-L1 expression and associated opportunities for cancer combination therapy, Epub 2023/04/15, Theranostics 13 (5) (2023) 1520–1544, https://doi.org/10.7150/thno.80091. Cited in: Pubmed; PMID 37056572.