Heliyon 10 (2024) e34524

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Integrative analyses of genes associated with oxidative stress and cellular senescence in triple-negative breast cancer

Lihua Wu [a,1], Hongyan Zheng [a,1], Xiaorong Guo [a], Nan Li [b], Luyao Qin [a], Xiaoqing Li [a], Ge Lou [a,*]

[a] Department of Pathology, The Second Affiliated Hospital of Harbin Medical University, Harbin, 150086, China
[b] Department of Pathology, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, 150001, China

ARTICLE INFO

Keywords:
Triple-negative breast cancer
Bioinformatics
Oxidative stress
Cellular senescence
Tumor immune microenvironment

ABSTRACT

Background: Oxidative stress and cellular senescence (OSCS) have great impacts on the occurrence and progression of triple-negative breast cancer (TNBC). This study was intended to construct a prognostic model based on oxidative stress and cellular senescence related difference expression genes (OSCSRDEGs) for TNBC.

Methods: The Cancer Genome Atlas (TCGA) databases and two Gene Expression Omnibus (GEO) databases were used to identify OSCSRDEGs. The relationship between OSCSRDEGs and immune infiltration was examined using single-sample gene-set enrichment analysis (ssGSEA), ESTIMATE, and the CIBERSORT algorithm. Least absolute shrinkage and selection operator (LASSO) regression analyses, Cox regression and Kaplan-Meier analysis were employed to construct a prognostic model. Receiver operating characteristic (ROC) curves, nomograms, and decision curve analysis (DCA) were used to evaluate the prognostic efficacy. Gene Set Enrichment Analysis (GSEA) Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) were utilized to explore the potential functions and mechanism.

Results: A comprehensive analysis identified a total of 27 OSCSRDEGs, out of which 15 genes selected for development of a prognostic model. A high degree of statistical significance was observed for the riskscores derived from this model to accurately predict TNBC Overall survival. The decision curve analysis (DCA) and ROC curve analysis further confirmed the superior accuracy of the OSCSRDEGs prognostic model in predicting efficacy. Notably, the nomogram analysis highlighted that DMD exhibited the highest utility within the model. In comparison between high and low OSCScore groups, the infiltration abundance of immune cells was statistically different in the TCGA-TNBC dataset.

Conclusion: These studies have effectively identified four essential OSCSRDEGs (CFI, DMD, NDRG2, and NRP1) and meticulously developed an OSCS-associated prognostic model for individuals diagnosed with TNBC. These discoveries have the potential to significantly contribute to the comprehension of the involvement of OSCS in TNBC.

Triple-negative breast cancer (TNBC), a distinct form of breast cancer distinguished by the absence of human epidermal growth

* Corresponding author. Department of Pathology, The Second Affiliated Hospital of Harbin Medical University, No. 246 Xuefu Road, Nangang District, Harbin, China.
E-mail addresses: h5144@hrbmu.edu.cn, louge_009@aliyun.com (G. Lou).
[1] Co-first authors.

https://doi.org/10.1016/j.heliyon.2024.e34524
Received 3 December 2023; Received in revised form 6 July 2024; Accepted 10 July 2024
Available online 14 July 2024

factor receptor 2 (HER2), progesterone receptor (PR) and the estrogen receptor (ER) [1], constitutes approximately 20 % of breast cancer cases [2]. In contrast to other molecular subtypes of breast cancer, TNBC typically manifests notably aggressive biological traits, such as high histological grades, heightened recurrence rates, and early distant metastases [3], all of which collectively contribute to a poorer prognosis. Presently, the primary modalities employed in the management of individuals diagnosed with TNBC encompass surgical intervention, chemotherapy, and radiotherapy [4,5]. However, the efficacy of these approaches is constrained by the absence of well-defined therapeutic targets. Notably, patients with TNBC frequently encounter drug resistance and endure deleterious side effects following chemotherapy, thereby impeding the achievement of successful cancer treatment, and yielding unsatisfactory clinical outcomes [6]. Consequently, it becomes imperative to investigate pivotal genes and pathways that exert a significant influence on the advancement and prognosis of TNBC.

Oxidative stress (OS) refers to the dysregulation of oxidation and antioxidant mechanisms within the human body, resulting in excessive production of reactive nitrogen species (RNS) and reactive oxygen species (ROS). This imbalance subsequently triggers a range of physiological and pathological responses [7]. Excessive levels of ROS have the capacity to oxidize intracellular macromolecules directly or indirectly, including DNA, proteins, and lipids. This oxidative process can lead to gene mutations, protein degradation, and lipid peroxidation. Consequently, it is widely recognized as the primary risk factor for human aging and the development of various aging-related conditions, such as tumors, cardiovascular and cerebrovascular diseases, diabetes, neurodegenerative disorders (such as Alzheimer's disease), and among others. ROS have been found to induce DNA mutations and abnormal cell proliferation, thereby activating proto-oncogenes and facilitating the initiation and progression of tumors [8]. Furthermore, cancer cells typically experience elevated levels of ROS, which further enhance their malignant characteristics such as increased proliferation, evasion of cell death, angiogenesis [9], and invasiveness. Despite the extensive evidence demonstrating the tumor-promoting effects of ROS, it is noteworthy that they may also exert a significant role in the anti-tumor process. Elevated levels of ROS exhibit cytotoxic effects, potentially inducing apoptosis in cancer cells, while oxidative stress typically hampers the survival of cancer cells during metastasis [10,11]. Furthermore, ROS significantly influences the tumor microenvironment (TME) [12], encompassing the development, functionality, and regulation of various immune cells such as T cells, macrophages, dendritic cells (DCs), B cells and natural killer (NK) cells.

Cellular senescence (CS) is commonly characterized as an irreversible cessation of cell proliferation, accompanied by observable alterations in cellular phenotype. This phenomenon is triggered by both intrinsic and extrinsic factors, such as oxidative stress, oncogenic signaling, mitochondrial impairment, critically shortened telomeres, mechanical strain, genotoxic injury, imbalanced nutrient availability, and viral or bacterial infection. The most prominent characteristic of cellular senescence throughout history has been a durable cessation of cell proliferation, which is facilitated by the activation of tumor suppressor genes CDKN2A/p16 and TP53, along with their respective downstream effectors retinoblastoma-1 (RB1) and CDKN1A/p21 family proteins. There is an inherent contradiction between cellular senescence and cancer [13]. On the one hand, CS functions as a robust mechanism for suppressing tumor growth. Increasing evidence suggests that CS is also involved in tissue repair processes, wherein senescent cells are involve in promoting localized fibrosis and the recruiting immune cells to eliminate damaged and senescent cells [14]. On the other hand, senescence is a stress response triggered by cellular damage, which results in tumor cells being long-term arrested and the tumor-immune environment being remodeled. Various proteinases, growth factors, chemokines, and pleiotropic cytokines are activated during this remodeling process, collectively known as the senescence-associated secretory phenotype (SASP) [15–17].

OS is a common driver of CS and could and trigger all the senescence traits. Being double-edged swords, the interplay between OS and CS has been observed to either enhance or impede the development and progression of tumors, by influencing tumor cells or modifying the TME. Consequently, the involvement of OSCS in the onset and advancement of breast cancer, as well as its applicability as a prognostic indicator, remains a subject of contention. A recent study revealed elevated levels of ROS in cell lines of TNBC, highlighting the reliance on ROS for survival as evidenced by the induction of cell death upon antioxidant treatment specifically in TNBC cells, while no such effect was observed in an estrogen receptor positive (ER+) cell line [18]. Nevertheless, the worldwide implications of OSCS on the prognostication of individuals with TNBC, as well as the association between OSCS-related biomarkers and TME in the progression of TNBC are yet to be elucidated. In this investigation, we have effectively identified OSCSRDEGs and conducted an analysis of the somatic mutation of OSCSRDEGs. We explored the expression patterns and prognostic values of OSCS signatures that constructed a prognostic model and evaluated immune characteristic in the context of TNBC. Identification and comprehensive analysis of potential candidate genes will enhance our understanding of gene regulation in OSCS, potentially yielding promising biomarkers and therapeutic targets for TNBC.

1. Materials and methods

1.1. Data download

The TCGA biolinks package [19] was utilized to retrieve the TCGA-BRCA dataset from the Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/). The samples lacking clinical prognosis and survival information were excluded, while breast cancer samples negative for ER, PR, and proto-oncogene Her-2 were filtered based on clinical information. To facilitate subsequent elucidation, we have designated this dataset as the TCGA-TNBC dataset, encompassing 116 TNBC cases and 11 adjacent control samples. The clinical data pertaining to these specimens were procured from the UCSC Xena [20] database, accessible at http://genome.ucsc.edu.

We obtained somatic mutation data (SNP, Single Nucleotide Polymorphism) of TNBC samples from the TCGA website and visualized it using the maftools [21] R package. The R package TCGAbiolinks was used to download and filter the Copy Number Segment data, which were then subjected to GISTIC 2.0 [22] analysis for samples with TNBC, using the default parameters.

Furthermore, we obtained the expression profile datasets GSE57544 [23] and GSE24202 [24] of patients with TNBC from the Gene Expression Omnibus (GEO) database [25] using the R package GEOquery [26], both of which were sourced from Homo sapiens. Additionally, the GSE57544 and GSE24202 data sets were utilized as validation sets for subsequent verification. Further details regarding the dataset information can be found in Supplementary Table S1.

Consequently, we employed "Oxidative stress" and "cellular senescence" as search terms to retrieve 7401 genes associated with OS and 3559 genes associated with CS from the GeneCards [27] database (https://www.genecards.org/). The study only included genes that had expression data available in the three datasets: TCGA-TNBC, GSE57544, and GSE24202. A total of 1877 oxidative stress and cellular senescence related genes (OSCSRGs), as detailed in Supplementary Table S2.

Differential analysis of TNBC/Control groups and further screening of related genes in TCGA-TNBC dataset.

The limma package was utilized to perform an analysis on the differential expression of OSCSRGs between the Cancer group (TNBC) and the Control group (Control) in the TCGA-TNBC dataset. The oxidative stress and cellular senescence related difference expression genes (OSCSRDEGs) exhibiting significant differences between the two groups were identified through a screening process. The screening threshold for OSCSRGs was established at a significance level of p-value.adj < 0.05 and $|\log FC| > 1$. The outcomes of the differential analysis were graphically represented through a volcano plot.

The OSCSRDEGs were subjected to further screening using COX univariate analysis, with a screening threshold of p-value <0.05. The genes that passed this screening were subsequently utilized for subsequent analysis.

1.2. Gene Set Enrichment Analysis (GSEA) between TNBC group samples and control group samples

The clusterProfiler package was utilized to perform GSEA [28] on all OSCSRGs in the TCGA-TNBC dataset. The GSEA was executed with specific parameters, including a seed value of 2022, 1000 computations, a minimum requirement of 10 genes per gene set, a maximum limit of 500 genes per gene set, and the application of the Benjamini-Hochberg (BH) method for p-value correction. The reference gene set used in this study was c2.cp.all.v2022.1.Hs.symbols.gmt [All Canonical Pathways] (3050), obtained from the Molecular Signatures Database 3.0 (MSigDB) [29]. Significant enrichment was determined based on the screening criteria of p-value $<$ 0.05 and q < 0.25.

Effects of oxidative stress and cellular senescence score (OSCSscore) on immune evaluation in TCGA-TNBC dataset.

The utilization of the single-sample gene-set enrichment analysis (ssGSEA) algorithm was implemented to assess the proportional abundance of individual genes within a given dataset sample. The R package GSVA was employed to ascertain potential mechanisms of action, as well as associated biological features and pathways of OSCSRDEGs in TNBC. This allowed for the calculation of the OSCSscore or each sample in the TCGA-TNBC dataset according to the expression matrix of OSCSRDEGs in each specimens using the ssGSEA algorithm.

To enhance comprehension of the impact of immunity-related genes on prognosis, we classified samples from the TCGA-TNBC dataset into Low/High OSCSscore groups, employing the median of OSCSscore as the demarcation threshold. The Immune Score and Stromal Score of the TCGA-TNBC dataset were computed utilizing the estimate package and the ESTIMATE algorithm. Additionally, the ESTIMATE Score was employed to quantify the immune and stromal constituents of the tumor.

1.3. Identification and correlation analysis of immune infiltrating cells in TCGA-TNBC dataset

The ssGSEA algorithm was utilized to assess the proportional representation of immune-cell infiltrates, wherein each specific infiltrated immune cell category was accurately identified, encompassing $CD8^+$ T cells, macrophages, regulatory T cells, and various other subtypes of human immune cells. A relative abundance of immune cells infiltration in each sample was calculated using enrichment scores derived from ssGSEA calculations [30,31]. Our analysis, conducted using the ssGSEA algorithm within the GSVA package of the R software, yielded enrichment scores that represented the levels of immune cell infiltration for various cell types in each sample. Boxplots were utilized to visually represent the discrepancies in the prevalence of 28 immune cells between the two groups. The Spearman algorithm was employed to calculate the correlation among immune cells within specific groups in the TCGA-TNBC dataset, which was then graphically displayed using the R package ggplot2. Additionally, the gene expression matrix of the TCGA-TNBC dataset was integrated to assess the correlation between immune cells and OSCSRDEGs in various groups, and a correlation point plot was generated using the R package ggplot2 for presentation purposes.

The CIBERSORT algorithm [32] was employed to conduct immune infiltration analysis, utilizing linear support vector regression to deconvolute the transcriptome expression matrix. This methodology allows for the estimation of immune cell abundance and composition within heterogeneous cell populations. In the present study, we applied the CIBERSORT algorithm to the gene expression matrix data obtained from the TCGA-TNBC dataset. By integrating this algorithm with the LM22 feature gene matrix, we successfully identified data instances exhibiting immune cell enrichment scores exceeding zero. As a result, we successfully acquired and presented the precise outcomes of the matrix indicating the abundance of immune cell infiltration. The Low/High OSCSscore group was categorized by the median value of OSCSscore, and the proportions of immune cell infiltration in the high and low groups were visually illustrated using stacked bar charts. The discrepancies in the abundance of immune cell infiltration among different groups were evaluated utilizing the Wilcoxon rank sum test and visually depicted by means of boxplots. The Spearman correlation coefficient was employed to ascertain the correlation between immune cells in distinct cohorts of the TCGA-TNBC dataset. This correlation was then visually represented through the utilization of the R package ggplot2. To determine the correlation between the expression of OSCSRDEGs and the infiltration abundance of cells with differing levels of infiltration between the high and low groups in the TCGA-TNBC dataset, the gene expression matrix of the dataset was merged. Following this, a correlation point plot was generated by

the R package ggplot2.

1.4. Construction of OSCSRGs prognostic model

To develop the prognostic model of OSCSRDEGs in the TCGA-TNBC dataset, we utilized the glmnet package [33] to performe Least Absolute Shrinkage and Selection operator (LASSO) regression [34], with the parameter set.seed (2022) and 10-fold cross-validation. To prevent overfitting, 1000 cycles were executed. A LASSO regression algorithm is widely used in diagnostic models, especially in prognostic diagnostics. The overfitting issue during curve fitting can be addressed by regularizing linear regression. This is achieved by introducing a penalty term (lambda × absolute value of slope), thereby enhancing the model's capacity for generalization.

$$\text{risk}Score = \sum_{i} Coefficient\ (hub\ gene_i) * mRNA\ Expression\ (hub\ gene_i)$$

The penalty coefficient (lambda) of each prognostic OSCSRDEG was extracted in LASSO regression diagnostic model. The median RiskScore with prognostic OS clinical information was employed as the critical value to divide the TNBC patient samples into Low RiskScore and High RiskScore groups, denoted as the Low/High RiskScore group. The prognostic model was evaluated by integrating the clinical prognostic data from the TCGA-TNBC dataset. The Kaplan-Meier (KM curve) analysis was conducted using survival package specifically to compare the overall survival (OS) of high and low risk groups.

1.5. Gene Set Variation Analysis (GSVA) analysis between high and low risk groups of prognostic model

GSVA [35] is a non-parametric unsupervised analytical method employed to evaluate the enrichment of gene sets in the nuclear transcriptome of microarray data. This approach entails the transformation of the gene expression matrix from multiple samples into a gene expression matrix among samples, facilitating the assessment of potential enrichment of distinct pathways across diverse samples. The "c2.cp.all.v2022.1.Hs.symbols.gmt" reference gene set was obtained from the Molecular Signatures Database (MSigDB) database (https://www.gsea-msigdb.org/). GSVA analysis was conducted on the gene expression matrix of the high and low risk groups of the OSCSRDEGs prognostic model. The subsequent step involved the calculation of functional differences among the enriched pathways in the GSVA enrichment analysis results, specifically focusing on the samples categorized into high and low expression groups within the OSCSRDEGs prognostic model. This study has selected the top 10 pathways exhibiting the highest log2FC values and the bottom 10 pathways displaying the lowest log2FC values from the pool of pathways that satisfy the criterion of p-value.adj <0.05.

1.6. Functional enrichment analysis (GO), pathway enrichment (KEGG) analysis

Gene Ontology (GO) [36] analysis is a common approach in large-scale functional enrichment studies, encompassing biological process (BP), cellular component (CC), and molecular function (MF). The Kyoto Encyclopedia of Genes and Genomes (KEGG) [37], serves as a utilized repository for housing comprehensive information pertaining to drugs, diseases, biological pathway, and genomes. In this study, the R package cluster Profiler [38] was employed to perform GO and KEGG enrichment analysis on prognostic OSCSRDEGs within the prognostic model of the TCGA-TNBC dataset. A screening criterion of p-value <0.05 and $q < 0.25$ was utilized. The items that satisfied the screening criterion were deemed to possess statistical significance, with the p-value correction method employed being Benjamini-Hochberg (BH).

1.7. Differential expression analysis of OSCSRGs

The OSCSRDEGs Lasso prognostic model's RiskScore was employed to classify samples from the TCGA-TNBC and GSE57544 datasets into low and high groups, with the partitioning determined by the median RiskScore. Subsequently, the Mann-Whitney U test (Wilcoxon rank sum test) was employed to examine and verify the expression disparities of OSCSRDEGs, between the low and high groups within the TCGA-TNBC and GSE57544 datasets, respectively. The R package ggplot2 was employed to present the outcomes of differential analysis, enabling the creation of group comparison maps.

1.8. Construction of multivariate cox prognostic model

All the identified genes were included and utilized to construct a multivariate Cox prognostic model. Subsequently, the R package rms [39] was employed for developing a nomogram. By assigning a specific scale to score and represent each variable in the multivariate regression model, the nomogram ultimately calculates the total score for predicting the probability of event occurrence.

The decision curve analysis (DCA) [40] is a straightforward method of assessing clinical prediction models, diagnostic tests, and molecular markers. In this study, the DCA chart was utilized to appraise the precision and discriminatory ability of the model. DCA plots were generated using the R package ggDCA for evaluation of the prognostic efficacy of a multivariate Cox model.

1.9. Statistical analysis

The data processing and analysis in this article were conducted by R software (Version 4.1.2). The independent Student t-test was used to compare continuous variables between two groups. For variables that were not normally distributed, the differences were

analyzed using the Mann-Whitney *U* test (Wilcoxon rank sum test). The survival package in the R programming language was utilized for conducting survival analysis. Kaplan-Meier survival curves were employed to visually represent the disparities in survival rates, while the Log-rank test was employed to evaluate the statistical significance of the differences in survival time between the two groups. The outcomes were computed using Spearman's correlation analysis when not explicitly stated. All statistical p-values were considered two-sided, and p-values below 0.05 were deemed indicative of statistical significance.

2. Results

2.1. Data set processing

A flowchart was developed to provide a comprehensive depiction of our study (Fig. 1). The triple-negative breast cancer related datasets (TCGA-TNBC, GSE57544, GSE24202) were initially normalized using the "normalizeBetweenArrays" function from the R package limma package. The figure demonstrates that the expression levels of these three datasets (TCGA-TNBC (Fig. 2A–B), GSE57544 (Fig. 2C–D), and GSE24202 (Fig. 2E–F)) exhibit a tendency towards consistency following standardization.

2.2. GSEA analysis between cancer group and control group (TNBC/control)

We took the intersection of oxidative stress and cell senescence related genes, and only retained the genes with expression data in three datasets (TCGA-TNBC, GSE57544 and GSE24202). A total of 1877 OSCSRGs were obtained using the limma package (Fig. 3A). Among these genes, 469 met the criteria of $|\log FC| > 1$ and p-value.adj < 0.05, indicating significant differences in expression.

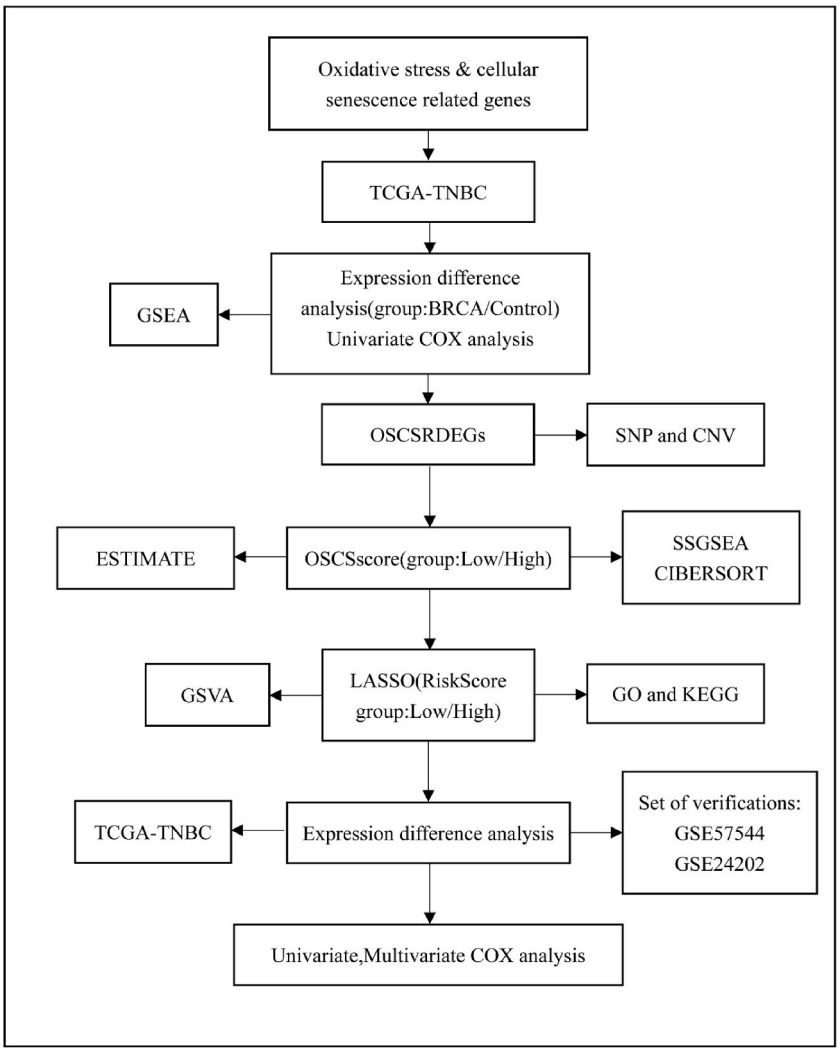

Fig. 1. The flowchart of this study.

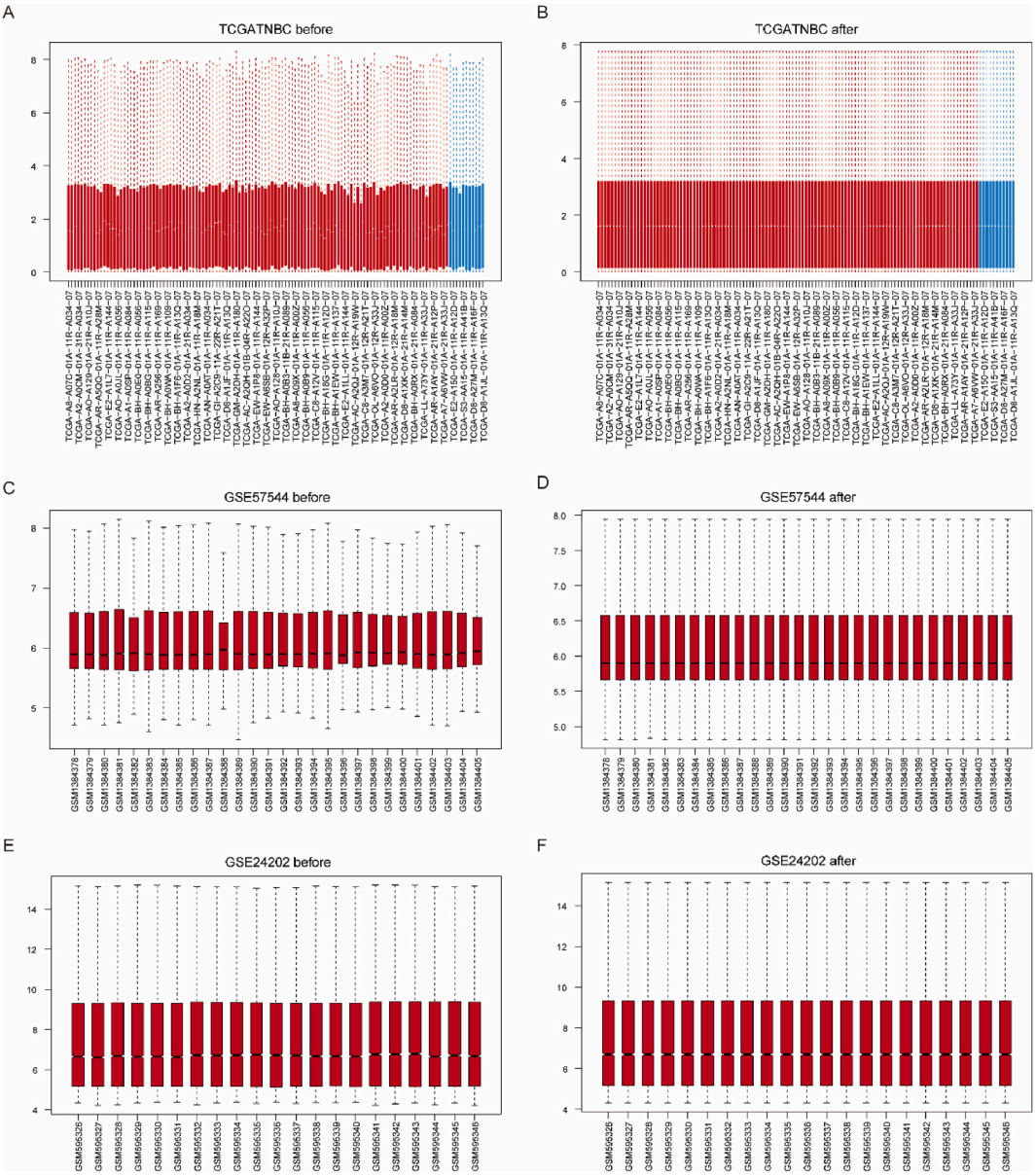

Fig. 2. Boxplots of the TCGA-TNBC, GSE57544, and GSE24202 datasets before and after standardized processing.

Specifically, in the high RiskScore group, which is characterized by high expression levels (as opposed to the low RiskScore group with low expression levels), there were 239 upregulated genes with a positive logFC. Conversely, the high RiskScore (high expression in the low RiskScore group, logFC negative) group exhibited low expression levels in 230 genes. The set of 469 genes was denoted as OSCSRDEGs. The outcomes of the differential analysis were visually displayed in a volcano plot (Fig. 3B).

To evaluate the influence of gene expression levels on TNBC within the TCGA-TNBC dataset, we utilized the GSEA technique. This approach facilitated the investigation of the correlation between gene expression and various biological processes, cellular components, and molecular functions associated with 1879 genes. A mountain plot (Fig. 3C) was used to visualize the results (p-value > 0.05 and $q > 0.25$), which confirmed significant enrichment. The results showed that these genes were found to be enriched in pathways such as Oncogene Induced Senescence, Biological Oxidations, Cellular Senescence, Regulation of TP53 Expression and Degradation, DNA Damage Telomere Stress Induced Senescence, Oxidative Stress Response, and other pathways as depicted in Fig. 3D–I and detailed in Supplementary Table S3.

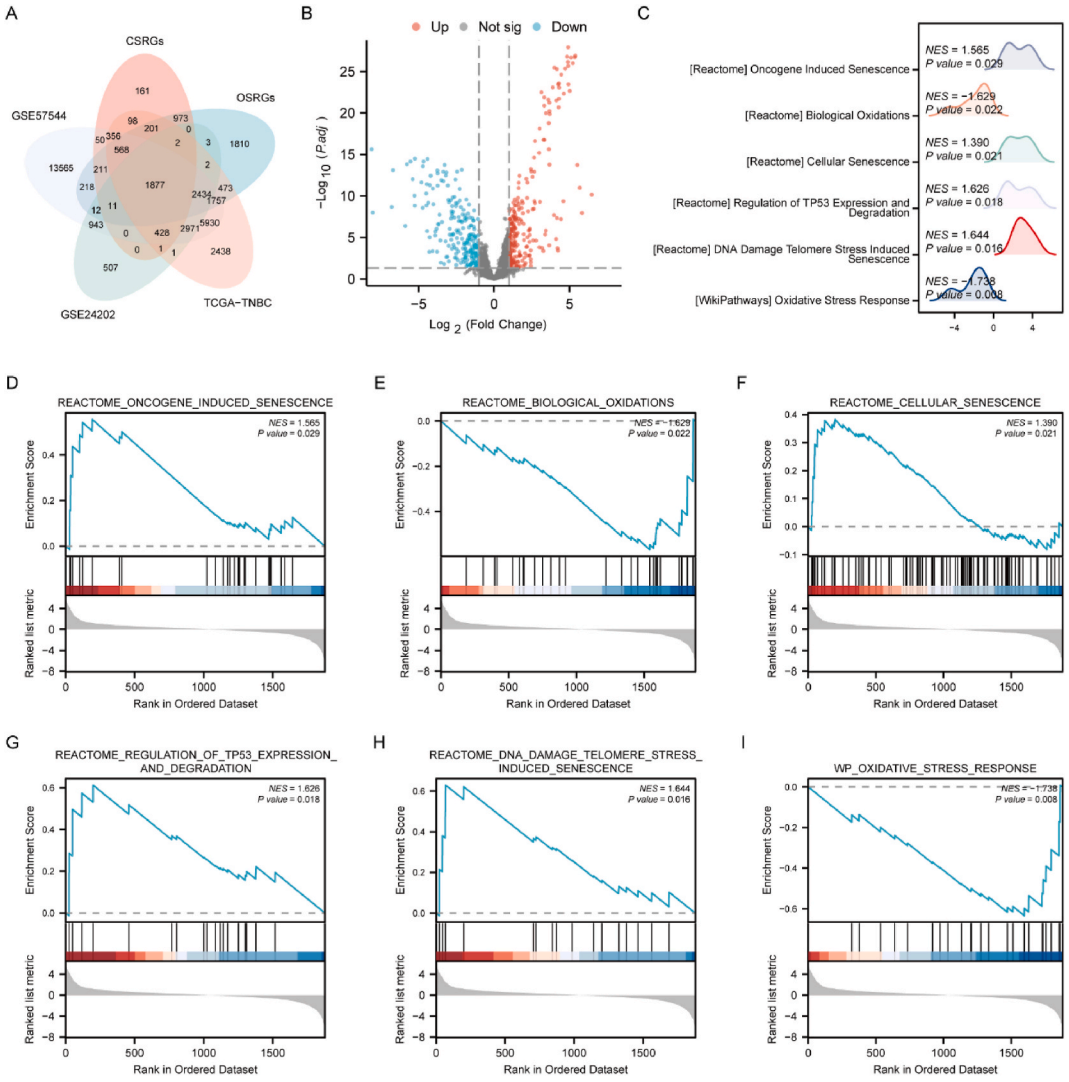

Fig. 3. (A) Venn diagram of the OSCSRGs. (B) Volcano plot of the OSCSRGs. (C–I) The six main biological characteristics for GSEA enrichment.

2.3. Mutation analysis of OSCSRDEGs in TNBC tissues

A total of 469 OSCSRDEGs were subjected to further screening using Univariate COX analysis. Only genes with a p-value < 0.05 were retained, resulting in the identification of 27 OSCSRDEGs. These genes include IVL, ALDOA, ITPKA, CRAT, HGF, ADCY6, KLF10, CSF3, NRP1, CFI, MAP1A, LEPR, ENDOG, DMD, FADS2, NDRG2, AKAP12, SLC12A5, NRTN, PHGDH, CFH, GLDC, ERG, VWF, SRC, WDR77, and SMARCA4.

The statistical analysis and visualization of the mutation analysis results of 27 OSCSRDEGs in the TCGA-TNBC data set, which consists of TNBC individuals with somatic mutations (Single Nucleotide Polymorphisms, SNPs), were utilized by the R package maftools. The study's findings unveiled that the TCGA-TNBC data set exhibited four primary categories of somatic mutations, namely Missense Mutation, Nonsense Mutation, Frameshift Deletion Mutation (Frame_shift_del), and Frameshift Insertion Mutation (Frame_Shift_Ins). Among these, missense mutations were found to be the most prevalent (Fig. 4A). Furthermore, the mutation types observed in 27 OSCSRDEGs among TNBC patients primarily consisted of SNPs, with a minor occurrence of deletions (DEL) and insertions (INS). The prevalence of the single nucleotide variant (SNV) C > A was observed to be the highest among TNBC patients, with subsequent occurrences of C > T, C > G, T > C, T > A, T > G, and so forth (Fig. 4A). Fig. 4A illustrates that within the TCGA-TNBC dataset, DMD exhibited the highest number of somatic mutations among the 27 OSCSRDEGs, encompassing three distinct mutation types. Furthermore, the three OSCSRDEGs with the greatest number of somatic mutations in the TCGA-TNBC dataset were MAP1A, SLC12A5, and VWF, respectively. The prevailing mutation type observed was missense mutation. Additionally, Fig. 4B presents a comprehensive display of all somatic mutations (SNPs) of the 27 OSCSRDEGs in TNBC patients.

Subsequently, we conducted an analysis of 27 OSCSRDEGs in the TCGA-TNBC dataset, which consisted of TNBC patients with

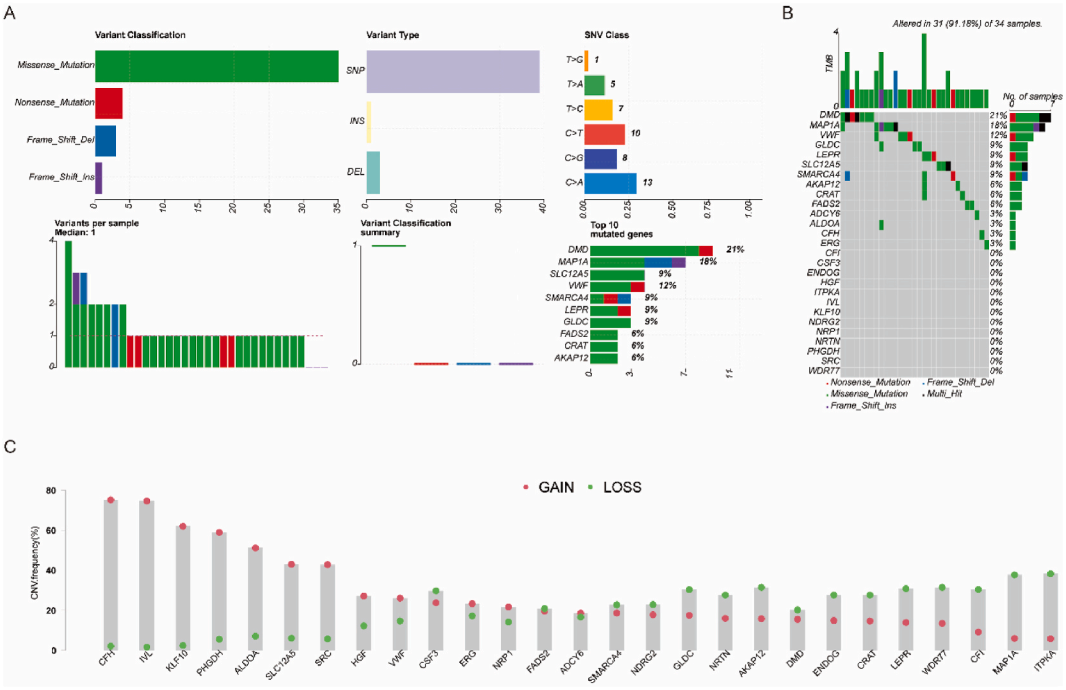

Fig. 4. (A) Somatic mutations (SNPS) of OSCSRDEGs in TNBC. (B) The proportion of somatic mutations of OSCSRDEGs in TNBC. (C) CNVs of OSCSRDEGs in TNBC.

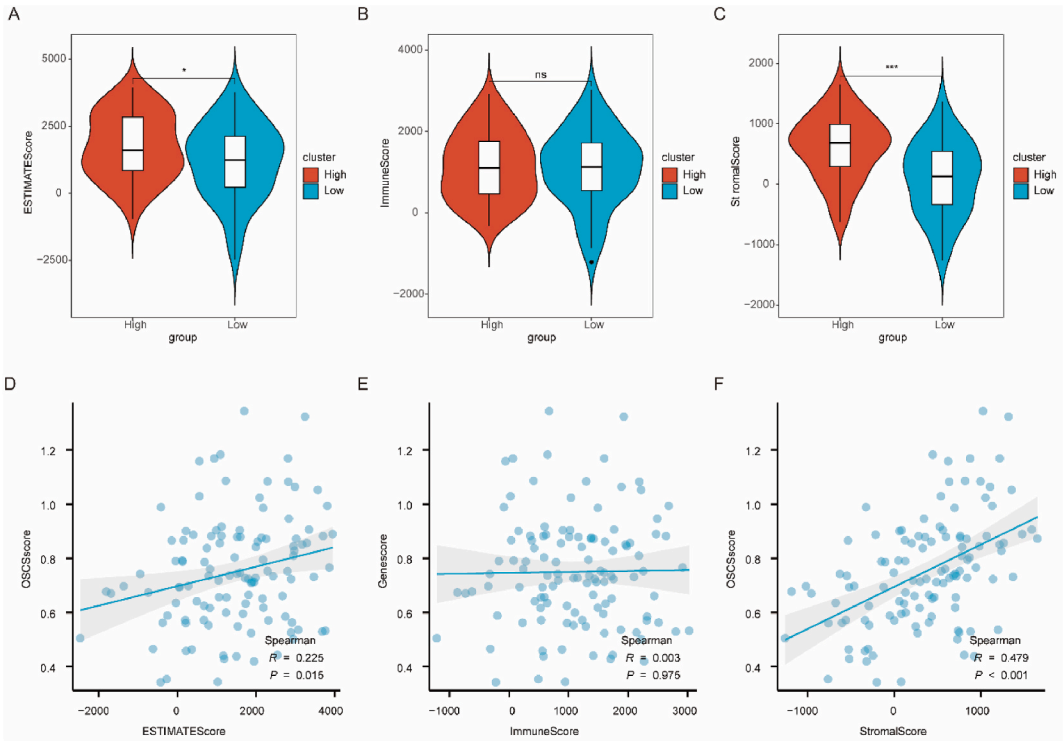

Fig. 5. (A–C) Violin plot of ESTIMATE Score, Immune Score, and Stromal Score between high and low score groups. (D–E) Scatter plot of correlation between ESTIMATE Score, Immune Score, Stromal Score, and OSCSscore.

CNVs. The CNV data of TNBC patients was downloaded and merged, and GISTIC2.0 was employed for analysis. The results were then visualized (Fig. 4C). The findings revealed a significant number of amplifications and deletions in the samples of TNBC patients for the 27 OSCSRDEGs. Notably, CFH, IVL, and KLF10 were identified as the most amplified OSCSRDEGs. Conversely, ITPKA, MAP1A, and WDR77 were the top three OSCSRDEGs with the highest number of deletions (Fig. 4C).

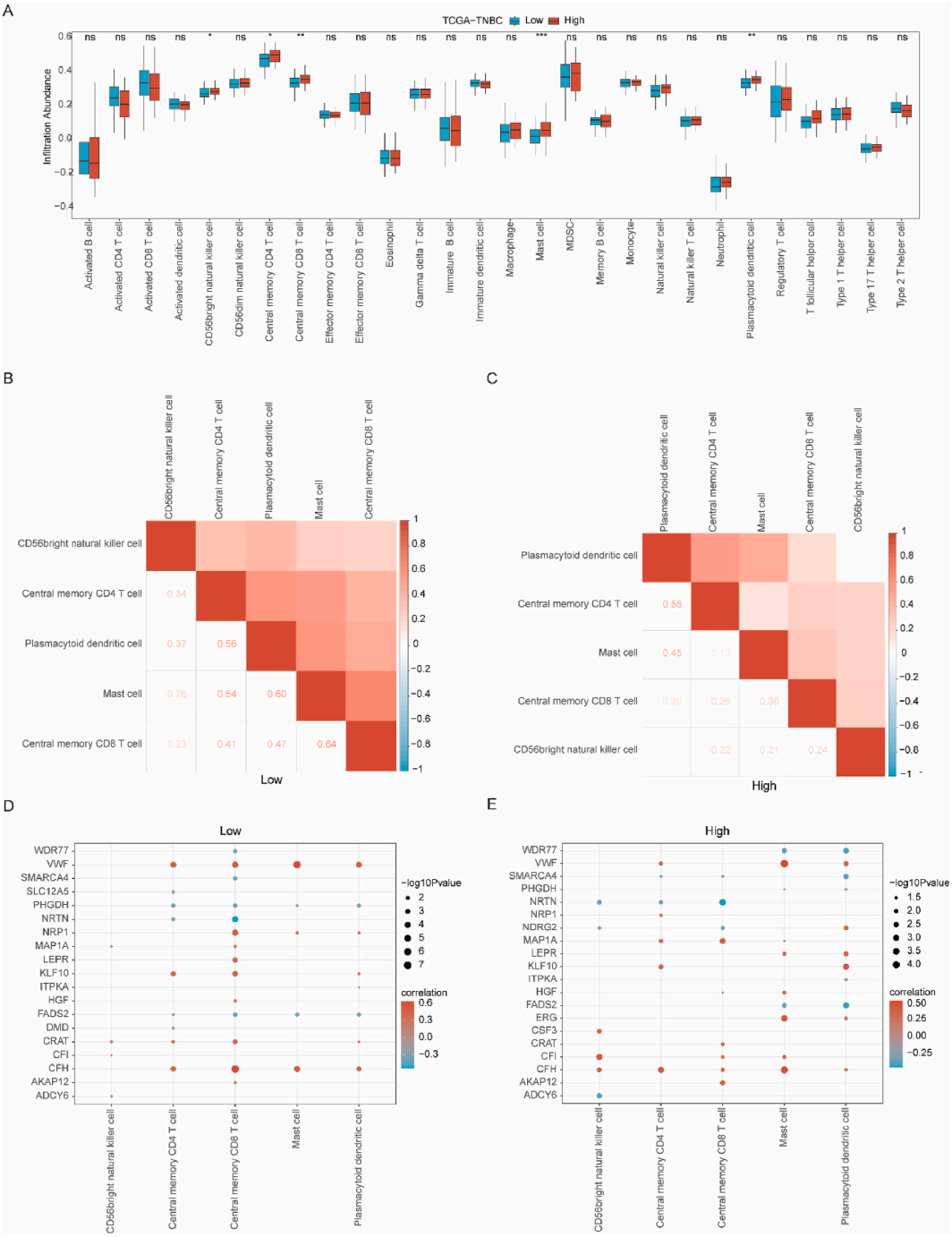

Fig. 6. (A) The comparison figure of ssGSEA immune infiltration analysis results in low/high OSCSscore group. (B–C) Correlation analysis results of immune cell infiltration abundance. (D–E) Dot plot of correlation between immune cells and 27 OSCSRDEGs.

2.4. Construction of OSCSscore and effects of OSCSscore on immune evaluation in TCGA-TNBC

The Low/High OSCSscore group was divided according to the median of OSCSscore, and the scores of ESTIMATE Score, Immune Score, and Stromal Score were presented through violin plots (Fig. 5A–C). The results suggest that there were statistically significant variations in the ESTIMATE Score (Fig. 5A) and Stromal Score (Fig. 5C) observed in the TCGA-TNBC dataset among different sample groups (Low/High OSCSscore group) ($P < 0.05$). Specifically, the high OSCSscore group exhibited significantly higher scores in both the Stromal Score and ESTIMATE Score compared to the low OSCSscore group. Furthermore, an examination was conducted to assess the relationship between ESTIMATE Score (Fig. 5D), Immune Score (Fig. 5E), Stromal Score (Fig. 5F), and OSCSscore. The findings revealed a noteworthy positive correlation between Stromal Score, ESTIMATE Score, and OSCSscore. However, no significant correlation was observed between Immune Score and OSCSscore. In all the figures in this study, the symbol ns was equivalent to p-value $\geq$ 0.05, which had no statistical significance. The asterisk (*) denotes a statistically significant p-value < 0.05, the double asterisk (**) signifies a highly statistically significant p-value < 0.01, and the triple asterisk (***) indicates a highly statistically significant p-value < 0.001.

Difference analysis of ssGSEA immune characteristics between high-score and low-score groups in TCGA-TNBC dataset.

In the TCGA-TNBC dataset, TNBC samples were categorized into Low/High OSCSscore groups, based on their median OSCSscores. To examine the discrepancy in immune infiltration among the Low/High OSCSscore groups of TNBC disease samples in the TCGA-

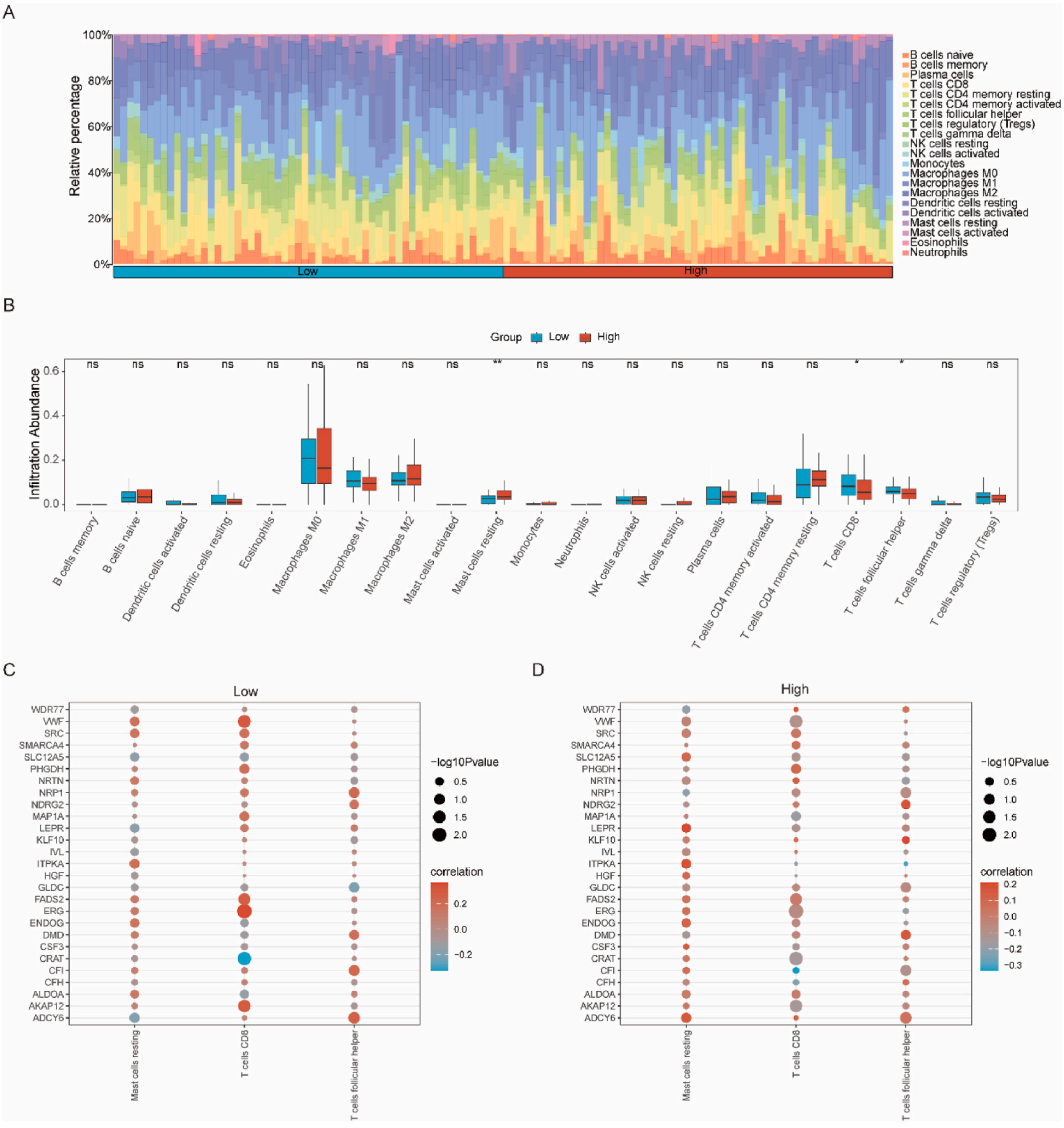

Fig. 7. (A–B) CIBERSORT immune infiltration analysis results are shown in stacked bar chart and group comparison chart. (C–D) Dot plot of correlation between immune cells and OSCSRDEGs.

TNBC dataset, the ssGSEA algorithm was utilized to calculate the abundance of 28 immune cells in both high-score and low-score groups of samples. Subsequently, the Mann-Whitney *U* test was employed to evaluate the degree of variation in infiltration of the 28 immune cells between the Low/High OSCSscore group. Fig. 6A illustrates the results using a group comparison plot. The results of this study demonstrate statistically significant differences in the levels of infiltration of five immune cell types between the Low/High OSCSscore group (p-value <0.05) in the TCGA-TNBC dataset. These immune cell types include Plasmacytoid dendritic cell, Central memory CD4 T cell, CD56bright natural killer cell, Central memory CD8 T cell, and Mast cell.

We proceeded to perform additional calculations and presented the correlation between the abundance of five immune cell infiltrations, along with the statistical disparities, in the samples belonging to the Low/High OSCSscore group (Fig. 6B–C). The findings indicated a positive correlation between the majority of the five immune cell infiltration abundances in both the low OSCSscore group (Fig. 6B) and high OSCSscore group (Fig. 6C) samples of the TCGA-TNBC dataset. Notably, the most significant correlation was observed between Mast cells and Central memory CD8 T cells in the low OSCSscore group (Fig. 6B). The group with high OSCS scores exhibited a significant correlation between Plasmacytoid dendritic cells and Central memory CD4 T cells, as depicted in Fig. 6C.

Simultaneously, we conducted a correlation analysis between the abundance of 5 immune cell infiltration and the expression levels of 27 OSCSRDEGs in the patient samples from the TCGA-TNBC dataset. The patients were divided into low OSCSscore group (Fig. 6D) and high OSCSscore group (Fig. 6E). A significance level of p-value < 0.05 was employed for screening, and the results were presented using a correlation plot (Fig. 6D–E). The findings indicated that within the low-score group of the TCGA-TNBC data set (Fig. 6D), notable associations were observed between 5 distinct cell types and 19 specific genes, with a predominance of positive correlations. Similarly, within the high-score group of the TCGA-TNBC data set (Fig. 6E), significant relationships were identified between 5 types of cells and 20 types of cells, with a majority of these associations being positive in nature.

Difference analysis of CIBERSORT immune characteristics between high-score and low-score groups in TCGA-TNBC dataset.

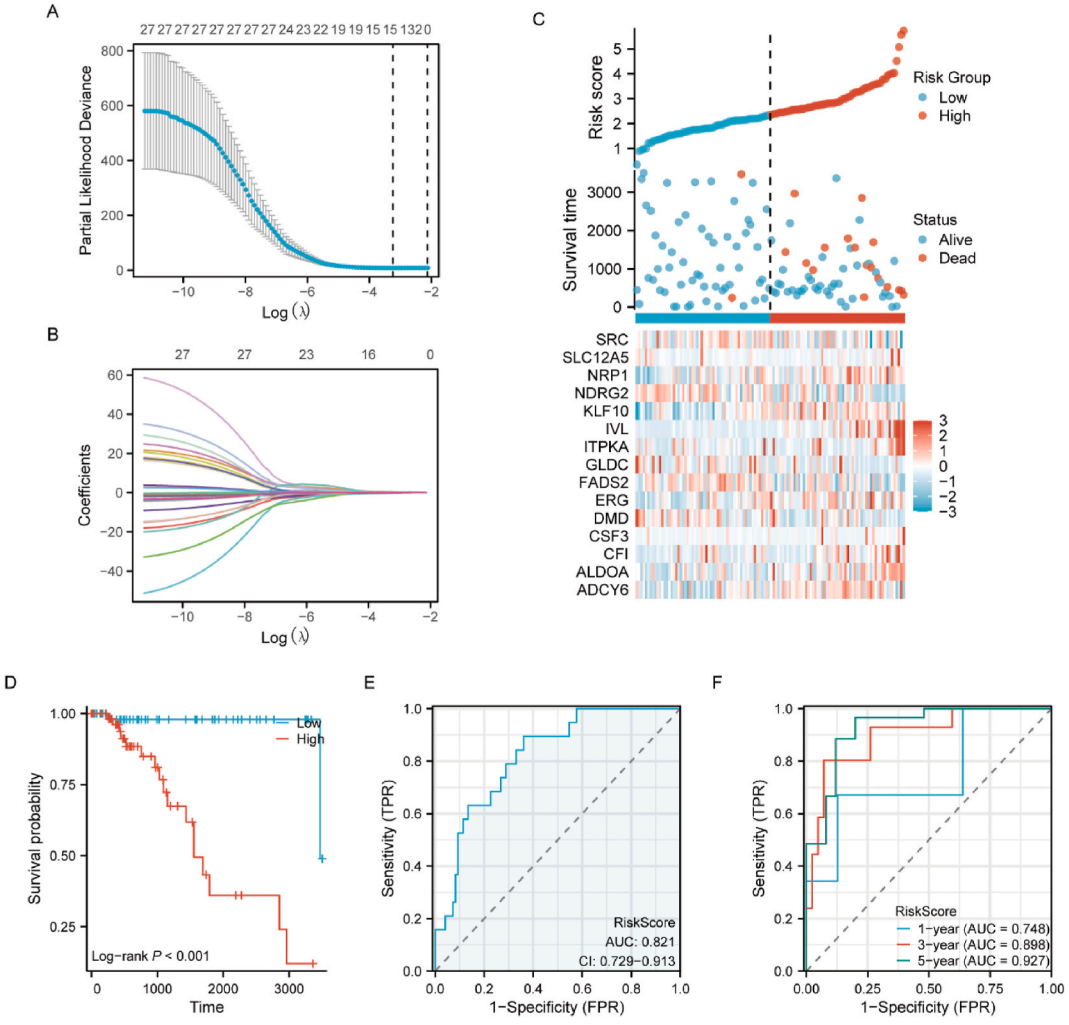

Fig. 8. (A) Prognostic model diagram of OSCSRDEGs. (B) Variable trajectory diagram. (C) Risk factors diagram. (D–F) KM curve, ROC curve, and time-dependent ROC between RiskScore of OSCSRDEGs prognostic model and survival outcome in TCGA-TNBC dataset.

In the TCGA-TNBC dataset, the TNBC samples were categorized into Low/High OSCSscore groups based on the median OSCSscore. The CIBERSORT algorithm was employed to compute the infiltration abundance of 22 immune cells in the Low/High OSCSscore groups. A stacked bar chart was utilized to visually represent the proportion of immune cells in the infiltration degree of the dataset samples (Fig. 7A). The findings revealed that out of the 21 immune cells present in the TCGA-TNBC dataset, their infiltration abundance was not uniformly zero in the TNBC samples. The following immune cell types were identified: B cells memory, B cells naive, Dendritic cells activated, Dendritic cells resting, Eosinophils, Macrophages M0, Macrophages M1, Macrophages M2, Mast cells activated, Mast cells resting, Monocytes, Neutrophils, NK cells activated, NK cells resting, Plasma cells, T cells CD4 memory activated, T cells CD4 memory resting, T cells CD8, T cells follicular helper, T cells gamma delta, and T cells regulatory (Tregs).

We further conducted an analysis on the disparity in the proportion of infiltration abundance of 21 immune cells between the Low/High OSCSscore group within the TCGA-TNBC dataset. The outcomes were visually presented through a group comparison plot (Fig. 7B). The findings indicated statistically significant variations in the infiltration abundance of three immune cells between the

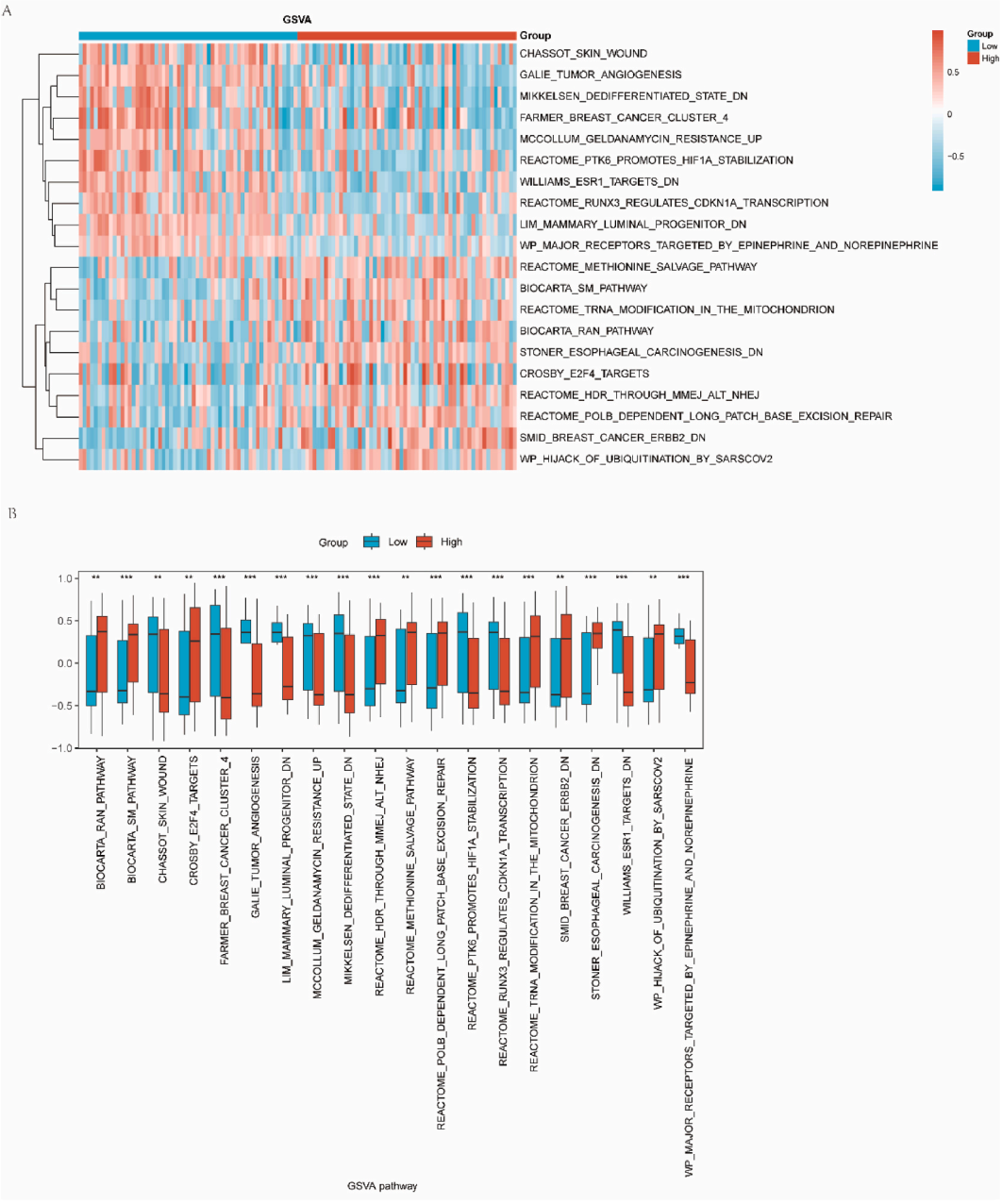

Fig. 9. Complex numerical heat map and group comparison map of GSVA enrichment analysis results between high and low risk groups in the TCGA-TNBC dataset.

Low/High OSCSscore group in the TCGA-TNBC dataset (Pp-value <0.05). These three immune cell types include T cells follicular helper, T cells CD8, and Mast cells resting (Fig. 7B).

Next, the "Spearman" algorithm was employed to assess and present the correlation between the abundance of three distinct types of immune cell infiltration and the statistical variances observed in the Low/High OSCSscore group samples within the TCGA-TNBC dataset. Additionally, the expression levels of 27 OSCSRDEGs were considered in this analysis (Fig. 7C–D). The results suggest that in the TCGA-TNBC dataset, there is a significant positive correlation observed between the infiltration abundance of three immune cells and the expression levels of 27 key genes for both the low OSCSscore group (Fig. 7C) and the high OSCSscore group (Fig. 7D).

2.5. Construction of OSCSRDEGs prognostic model

To assess the prognostic value of 27 OSCSRDEG,the expression levels of these 27 OSCSRDEGs in the TCGA-TNBC dataset were then combined with prognostic information (OS, OS.time) for conducting LASSO regression analysis to establish a prognostic model (Fig. 8A). During the entire process, we identified and retained 15 out of 27 OSCSRDEGs (including ADCY6, ALDOA, CFI, CSF3, DMD, ERG, FADS2, GLDC, ITPKA, IVL, KLF10, NDRG2, NRP1, and SRC) as potential markers for TNBC. based on the minimum partial likelihood deviation. Furthermore, the LASSO regression results were visually represented through the LASSO variable trajectory map (Fig. 8B). In order to further validate the prognostic significance of 15 OSCSRDEGs in the TCGA-TNBC dataset, the sample grouping in the developed OSCSRDEGs prognostic model was visually represented through a risk factor map (Fig. 8C). The risk factor map is composed of three components. Firstly, the data set samples were categorized based on the median of RiskScore predicted by the OSCSRDEGs prognostic model, resulting in risk grouping. In addition, a dot plot was employed to visually depict the survival time and survival outcome of clinical samples in the TCGA-TNBC dataset. Finally, the expression of OSCSRDEGs in the prognostic model, which was developed by the researchers, was visually presented in the high and low risk groups of samples from the TCGA-TNBC dataset using a heat map.

$$\begin{aligned}\text{RiskScore} = {} & \text{ERG} * 0.48100989 + \text{IVL} * 0.388510938 + \text{ALDOA} * 0.380802265 + \text{ADCY6} * 0.27004406 + \text{CFI} * 0.239930279 \\ & + \text{KLF10} * 0.203280145 + \text{ITPKA} * 0.192615021 + \text{NRP1} * 0.149238819 + \text{CSF3} * 0.137898021 + \text{SLC12A5} \\ & * 0.096493715 + \text{FADS2} * -0.018734706 + \text{SRC} * -0.070486597 + \text{GLDC} * -0.072744399 + \text{NDRG2} * -0.134742786 \\ & + \text{DMD} * -0.235234897\end{aligned}$$

The high and low risk groups were determined based on the median RiskScore of the OSCSRDEGs prognostic model. We then combined this information with survival information (OS.event) to conduct a prognostic KM curve analysis. The finding demonstrated that the risk score derived from the OSCSRDEGs prognostic model exhibited a statistically difference in predicting OS of TNBC patients in the TCGA-TNBC dataset (p-value < 0.001, Fig. 8D). The results shown in the ROC curve (Fig. 8E) demonstrated that the OSCSRDEGs prognostic model exhibited a certain level of accuracy in predicting patient OS, as evidenced by an AUC value of 0.821. The time-dependent ROC analysis (Fig. 8F) revealed that the prognostic model exhibited a suboptimal accuracy in predicting the 1-year survival outcome (AUC = 0.748), a moderate accuracy in predicting the 3-year survival outcome (AUC = 0.898), and a high accuracy in predicting the 5-year survival outcome (AUC = 0.927).

2.6. GSVA analysis between high and low risk groups of prognostic models

Gene expression was compared between high and low risk groups (divided by median) in the TCGA-TNBC dataset using GSVA. Specifically, we selected the top 10 pathways with the highest logFC and the bottom 10 pathways with the lowest logFC, all of which met the criterion of p-value. Adj <0.05, from the GSVA enrichment Analysis results. Specifically, the Biocarta Sm Pathway, Trna Modification in the Mitochondrion, Stoner Esophageal Carcinogenesis Dn, Smid Breast Cancer Erbb2 Dn, Polb Dependent Long Patch Base Excision Repair, Biocarta Ran Pathway, Hdr Through Mmej Alt Nhej, Crosby E2f4 Targets, Methionine Salvage Pathway, Hijack of Ubiquitination by Sarscov2, Farmer Breast Cancer Cluster 4, Mccollum Geldanamycin Resistance Up, Chassot Skin Wound , Major Receptors Targeted by Epinephrine And Norepinephrine, Runx3 Regulates Cdkn1a Transcription, Ptk6 Promotes Hif1a Stabilization, Lim Mammary Luminal Progenor Dn, Mikkelsen Dedifferentiated State Dn, Williams Esr1 Targets Dn, and Galie Tumor Angiogenesis. A heatmap was generated using the R package pheatmap to visualize the specific results of the differential analysis (Fig. 9A, refer to Supplementary Table S4 for detailed information). Additionally, the Mann-Whitney *U* test was utilized to evaluate the extent of disparity in these 20 pathways among distinct groups within the TCGA-TNBC dataset, and the findings were depicted in a comparative diagram (Fig. 9B). The findings of this study indicate that there was a significant statistical difference (Pp-value <0.01) in the expression of these 20 pathways among various groups within the TCGA-TNBC dataset.

Functional enrichment (GO) and pathway enrichment (KEGG) analysis of the prognostic model between high and low risk groups.

To investigate the biological processes and molecular functions of 15 OSCSRDEGs in the TCGA-TNBC dataset, an initial step involved conducting GO and KEGG enrichment analysis. Enrichment items were screened based on the criteria of p-value < 0.05 and q < 0.25, and pathways meeting these criteria were deemed statistically significant.

The results of the study revealed that a collective count of 15 OSCSRDEGs exhibited significant enrichment in diverse BP, such as the inhibition of extrinsic apoptotic signaling pathway, suppression of inflammatory pathway response to antigenic stimulus, downregulation of peptidyl-serine phosphorylation, regulation of inflammatory response to antigenic stimulus, and control of mitochondrial depolarization. Additionally, these OSCSRDEGs were also found to be enriched in several CC, such as actin cytoskeleton, cell

projection membrane, cell-substrate junction, endocytic vesicle lumen, and semaphorin receptor complex. Additionally, the identified genes were found to exhibit enrichment in various MF, including growth factor receptor binding, lyase activity, phospholipase activator activity, phosphorus-oxygen lyase activity, and protein kinase C binding, among others (as detailed in Supplementary Table S5). Furthermore, a subset of 15 differentially expressed genes were predominantly enriched in the KEGG pathway associated with GABAergic synapse. The outcomes of the functional enrichment analyses, both in terms of GO and KEGG pathways, were visually represented using bar graphs (refer to Fig. 10A). The GO gene functional enrichment analysis results included a ring network diagram (Fig. 10B–E) that visually depicted the BP pathway, CC pathway, MF pathway, and KEGG pathway.

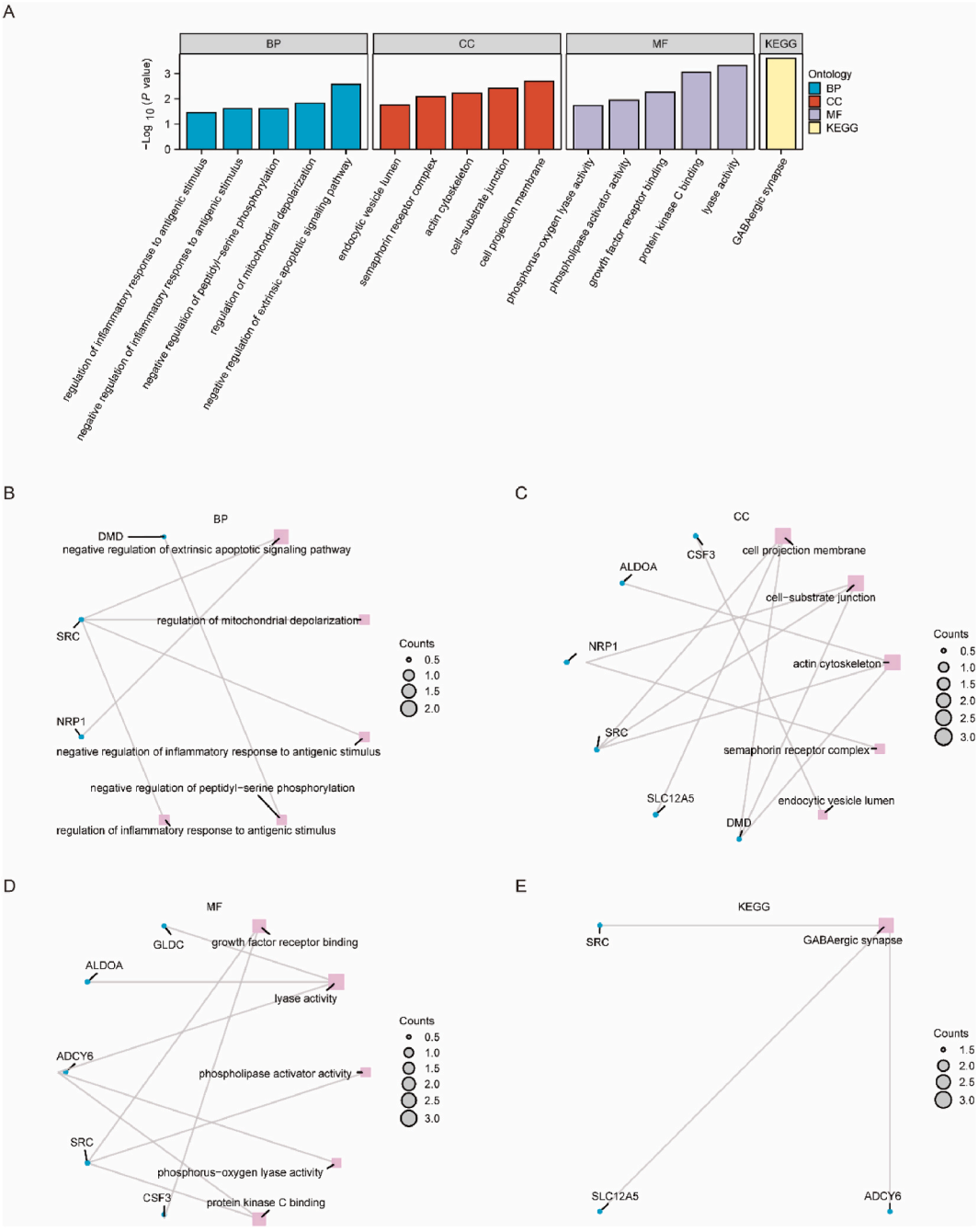

Fig. 10. (A) The bar chart of GO and KEGG enrichment analysis results. (B–E) Loop network diagram of BP pathway, CC pathway, and MF pathway in GO functional enrichment analysis results and loop network diagram of KEGG enrichment results.

2.7. Differential expression analysis of OSCSRDEGs

The Wilcoxon rank sum test was utilized to examine the expression differences of 15 OSCSRDEGs within the TCGA-TNBC dataset and the GSE57544 dataset between the low and high Lasso risk score groups. The findings indicated that 11 OSCSRDEGs (ADCY6, ALDOA, CFI, DMD, ERG, FADS2, GLDC, IVL, KLF10, NDRG2, NRP1) differ significantly between high and low risk groups in the TCGA-TNBC dataset (Fig. 11A). Similarly, in the GSE57544 dataset, 5 OSCSRDEGs (CFI, CSF3, DMD, NDRG2, NRP1) demonstrated significant difference (Fig. 11B). Four OSCSRDEGs (CFI, DMD, NDRG2, and NRP1) exhibited statistically significant differences in both the TCGA-TNBC and GSE57544 datasets. Specifically, CFI and NRP1 were found to be up-regulated in the high-risk group samples of both datasets, while the expression of DMD and NDRG2 was down-regulated in the high-risk group samples compared to the low-risk group in both datasets.

Furthermore, a correlation heat map (Fig. 11C–E) was generated to analyze the expression patterns of 15 OSCSRDEGs across three distinct datasets, namely TCGA-TNBC, GSE57544, and GSE24202. In the TCGA-TNBC dataset, a predominance of negative correlations was observed among the expression levels of 15 OSCSRDEGs. Significantly, the expression of KLF10 displayed the most robust positive correlation with the expression of NRP1, whereas the expression of CFI exhibited the most pronounced negative correlation with the expression of FADS2. Additionally, a strong negative correlation was also observed between ADCY6 expression and SRC expression. In the GSE57544 dataset, a majority of the correlations observed between the expression levels of 15 OSCSRDEGs exhibited negativity. Notably, the strongest positive correlation was found between the expression of CFI and NRP1, while the strongest negative correlations were observed between the expression of NDRG2 and CFI, as well as between the expression of NDRG2 and NRP1. Similarly, in the GSE24202 dataset, a majority of the correlations between the expression levels of the 15 OSCSRDEGs were negative. The expression of IVL and CSF3 exhibited the strongest positive correlation, whereas the expression of SRC and CFI displayed the most prominent negative correlation.

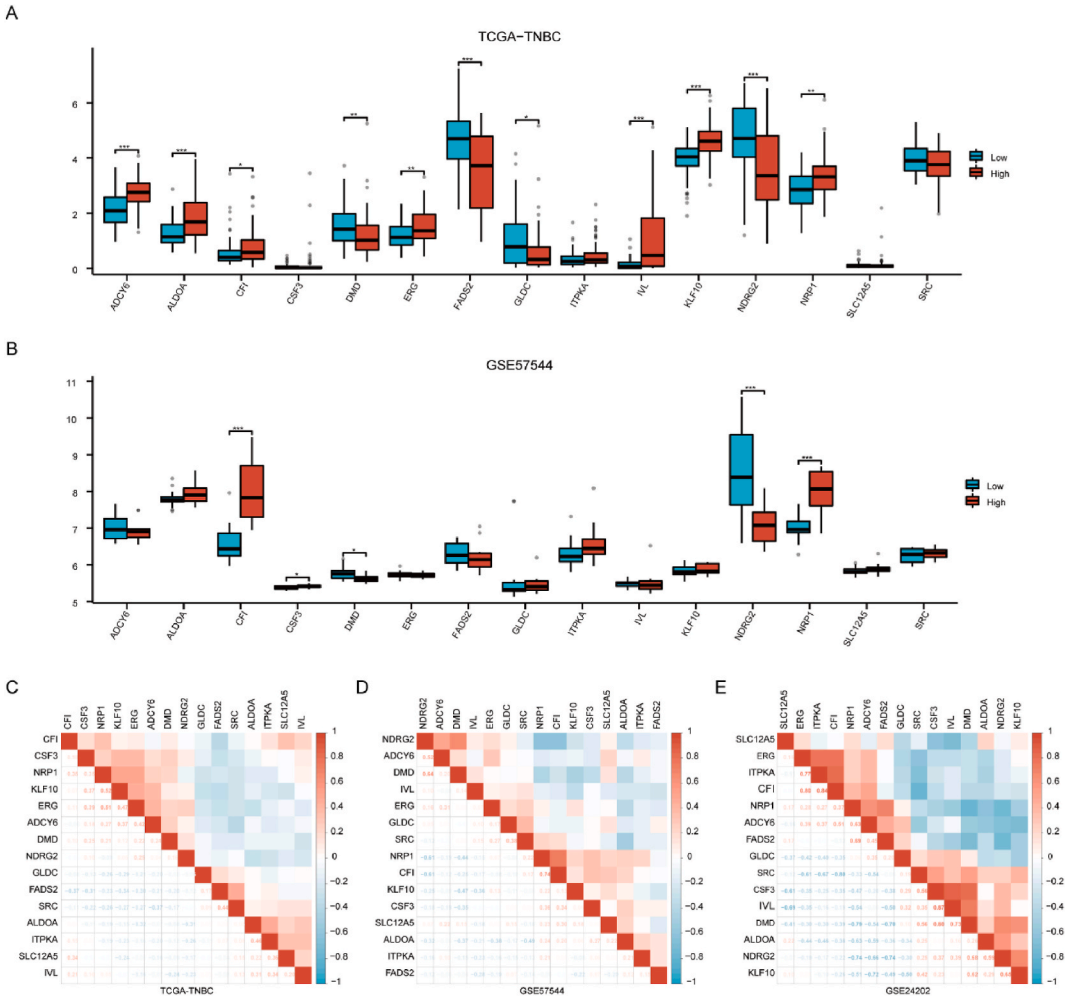

Fig. 11. (A–B) Grouping comparison of 15 OSCSRDEGs in TCGA-TNBC dataset and GSE57544 dataset. (C–E) Heat map of correlation between expression levels of 15 OSCSRDEGs in TCGA-TNBC dataset, GSE57544 dataset, and GSE24202 dataset.

2.8. Construction of multivariate cox prognostic model for TCGA-TNBC dataset

The four OSCSRDEGs (CFI, DMD, NDRG2, NRP1) exhibiting significant differences in risk scores between groups in both TCGA-TNBC and GSE57544 datasets were incorporated into a multivariate Cox regression analysis to construct a prognostic model. (The results of Cox regression can be found in Supplementary Table S6, while a baseline data table is provided in Supplementary Table S7). The outcomes of the multivariate Cox regression were compiled and visually displayed as a forest plot (Fig. 12A).

$$\text{RiskScore} = \text{CFIHigh} * 0.537418656 + \text{DMDHigh} * \text{-}1.368539695 + \text{NDRG2High} * \text{-}0.629547269 + \text{NRP1High} * 1.090492236$$

Subsequently, a nomogram analysis was conducted to assess the prognostic efficacy of the model, and the resulting nomogram was depicted in Fig. 12B. The nomogram was constructed based on a multivariate regression analysis, wherein each variable in the model was assigned a specific score to represent its respective influence. The cumulative score was then utilized to predict the likelihood of event occurrence. Fig. 12B illustrates that the DMD gene exhibited the highest utility within the multivariate Cox prognostic mode, followed by NRP1, NDRG2, and CFI.

Finally, DCA was employed to assess and demonstrate the clinical efficacy of the developed multivariate Cox regression model over a span of 1 year (Fig. 12C), 3 years (Fig. 12D), and 5 years (Fig. 12E). The X-axis of the DCA plot corresponds to the Probability Threshold or Threshold Probability, while the Y-axis represents the net benefit. Evaluation of the results can be conducted by examining the extent of stability above the All-positive line and All-negative line, as well as the portion of the All-positive line and All-negative line. A larger area indicates a more favorable model performance. The results of the study indicated that the clinical predictive effectiveness of our developed multivariate Cox regression model was ranked in the following order: 3 years > 5 years > 1 year.

3. Discussion

TNBCs are associated with unfavorable patient survival due to their highly aggressive biological characteristics and the absence of efficient targeted therapeutic strategies. There is accumulating evidence indicating that OSCS are involved in the occurrence, progression, and metastasis of breast cancer [41,42]. Lei et al. found that the overexpression of C/EBPβ in TNBC cell lines leads to a significant reduction in ROS levels, accompanied by enhanced migration and invasion capabilities, thereby facilitating metastasis [43]. Furthermore, the potential efficacy of redox-based anticancer therapies in TNBC has exhibited promising outcomes [44]. Regarding the impact of aging on TNBC, a primary focus lies in recognizing senescence as a pivotal determinant of treatment resistance in TNBC [45]. However, the specific mechanism through which OSCS contribute to the pathogenesis and progression of TNBC remains insufficiently understood.

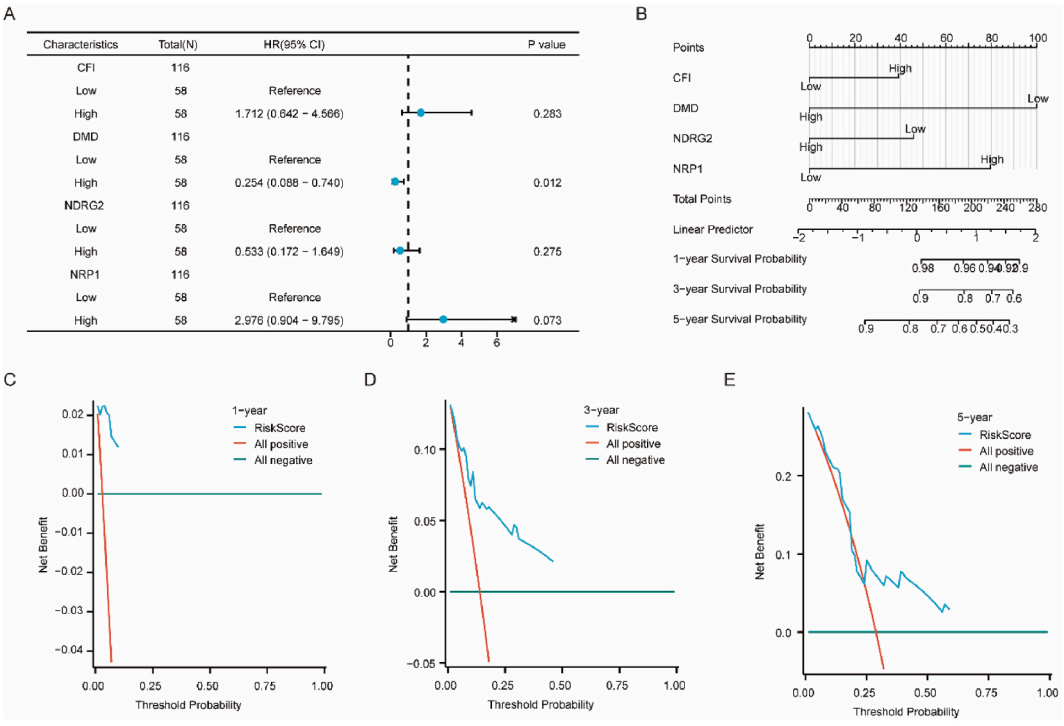

Fig. 12. (A) Forest plot of the multivariate Cox regression mode. (B) Nomogram of multivariate Cox regression model. (C–E) DCA plots at 1 year, 3 years, and 5 years from the multivariate Cox regression model.

In the present study, RNA-sequencing data were acquired from the TCGA and GEO databases, in conjunction with the corresponding clinical information. We conducted a comprehensive screening of OSCSRDEGs and subsequently validated their expression in TNBC using two independent external datasets. Additionally, an assessment of the somatic mutation landscape of OSCSRDEGs was performed, and the relationship between immune infiltration levels and the expression of 27 OSCSRDEGs, as determined by OSCSscore in TNBC, was explored. The results of our study indicate significant disparities between the Low/High OSCSscore groups, specifically highlighting notable variations in five distinct types of immune cells: Central memory CD8 T cell, Central memory CD4 T cell, Mast cell, CD56 bright natural killer cell, and Plasmacytoid dendritic cell. Our objective in analyzing these mutations was to identify potential driver genes and pathways that contribute to the pathogenesis of TNBC. The underlying mechanisms of OSCSRDEGs in TNBC were also investigated by GSEA, GSVA, GO, and KEGG pathway enrichment analyses. According to our prognostic model, the OSCSRDEGs (15 genes) exhibited significant enrichment primarily in the negative regulation of the extrinsic apoptotic signaling pathway, regulation of mitochondrial depolarization, endocytic vesicle lumen, regulation of inflammatory response to antigenic stimulus, and phosphorus-oxygen lyase activity. Ultimately, a prognostic model was developed using OSCSRDEGs, showcasing its efficacy in accurately evaluating the prognosis of patients diagnosed with TNBC. Significantly, our study successfully identified four OSCSRDEGs (CFI, DMD, NDRG2, and NRP1), which offer promising prospects as biomarkers and therapeutic targets for TNBC.

The prognostic power of the model was assessed through nomogram analysis, which indicated that DMD exhibited the highest utility, followed by NRP1, NDRG2, and CFI. Additionally, our analysis of these mutations aimed to uncover potential driver genes and pathways that contribute to the pathogenesis of TNBC. DMD was found to possess the highest occurrence of somatic mutations, encompassing three distinct mutation types, as evidenced by our study. DMD, the gene responsible for Duchenne muscular dystrophy [46], is a considerably large gene consisting of 79 exons that encode the vital dystrophin protein (Dp). The primary protein product of DMD is the complete dystrophin protein (Dp427), which plays a critical role in preserving skeletal muscle integrity through its connection between the actin cytoskeleton and the extracellular matrix [47]. In addition to skeletal muscle, the DMD gene product Dp71 is highly prevalent and widespread, with at least 14 splice isoforms identified [48]. Increasing evidence suggests that DMD contributes to the development of various types of cancer, including sarcomas, carcinomas, melanomas, lymphomas, and leukemia [49]. The imbalance between short and long gene products may have a significant impact on tumorigenesis. Notably, in sporadic breast cancer, the frequency of DMD alterations was found to be higher than that of BRCA1 and BRCA2. Moreover, the present study's survival analyses have demonstrated that patients with breast cancer who possess DMD alterations exhibit considerably inferior survival rates in comparison to their counterparts [50]. Additionally, a separate investigation has identified DMD as one of the seven genes displaying enriched mutations in metastatic TNBCs, while its presence is not observed in primary tumors [51]. These findings align with our own research, which has revealed a heightened frequency of DMD genetic alterations and a significantly diminished OS for breast cancer patients with DMD mutations. DQ. Cai et al. [52] conducted a comprehensive investigation into the existence of a chemoresistance-associated tumor-infiltrating exhausted-like $CD8^+$ T cell signature (TILCD8TSig) in breast cancer. Their findings indicated a higher frequency of mutations in PIK3CA, DMD, and APOB within the cr-TILCD8Sig-high subgroup compared to the cr-TILCD8Sig-low subgroup for genomic alterations. This study provided evidence suggesting an association between the DMD mutation and tumor-infiltrating lymphocytes (TILs), which aligns with the results of our own research.

N-myc downstream-regulated gene 2 (NDRG2), a constituent of the NDRG family, has been recognized as a stress-responsive gene that experiences downregulation in reaction to N-myc gene expression and encounters transcriptional activation when exposed to diverse cellular stress stimuli, including hypoxia, DNA damage, or endoplasmic reticulum stress (ERS). Moreover, the roles of NDRG2 as a tumor suppressor further contribute to the inhibition of tumor growth, augmentation of sensitivity to anticancer drugs, and prevention of metastasis in different tumor types [53]. Moreover, our research findings are in alignment with previous studies indicating that the upregulation of NDRG2 enhances apoptosis, thereby inhibiting the sensitivity of tumor cells to anticancer therapy [54]. Lee A et al. have demonstrated the significant role of NDRG2 expression in suppressing PD-L1 expression and restoring T cell proliferation activity that is inhibited by PD-L1 in TNBC cells [55]. In order to elucidate the molecular mechanism underlying the anti-tumor phenotype associated with NDRG2, further investigation into the interactions between NDRG2 and upstream genes located in the OSCS pathway is necessary. The present study has uncovered a robust inverse correlation between the expression of NDRG2 and the expression of CFI and NRP1 in TNBC. Consequently, further extensive experimental investigations are necessary to elucidate the potential regulatory impacts of NDRG2.

In terms of enrichment analysis, an initial screening was conducted on 469 OSCSRDEGs, followed by an examination of the association between their expression levels and molecular functions, cellular components, and various biological processes using the TCGA-TNBC dataset. The findings revealed a significant enrichment of genes related to Oncogene Induced Senescence, Biological Oxidations, Cellular Senescence, Regulation of TP53 Expression and Degradation, DNA Damage Telomere Stress Induced Senescence, and Oxidative Stress Response within the low and high RiskScore groups of the OSCSRDEGs prognostic model in the TCGA-TNBC dataset. Furthermore, to explore the underlying functions and mechanisms of the prognostic model between the high and low Lasso risk score groups, we conducted GSVA, GO, and KEGG analyses. The findings indicated that the 15 OSCSRDEGs exhibited significant enrichment in various biological processes, including negative regulation of extrinsic apoptotic signaling pathway, regulation of mitochondrial depolarization, endocytic vesicle lumen, regulation of inflammatory response to antigenic stimulus, and phosphorus-oxygen lyase activity, among others. ROS have been identified as regulators of the expression and function of various crucial signaling factors implicated in cellular proliferation, apoptosis, and differentiation. In the presence of an excessive amount of ROS, the apoptotic signaling pathway is triggered to facilitate regular cell death on a frequent basis [56]. Recent research has extensively demonstrated that oxidative stress can induce cellular apoptosis through both mitochondria-dependent and mitochondria-independent pathways [57]. Mitochondrial dysfunction is posited as a contributing factor to the aging process, with a strong association to OSCS. An increasing body of evidence indicates that ROS triggers alterations in mitochondrial dynamics and

hastens the buildup of oxidized by-products. The excessive presence of ROS leads to oxidative stress, thereby activating the PINK1/Parkin-dependent pathway, which in turn induces the formation of Mitochondrial-derived vesicles (MDV), mitochondrial depolarization, and mitophagy [58]. Nevertheless, excessive activation or suppression of mitochondrial quality control (MQC) can lead to the accumulation of oxidized proteins, hasten abnormal energy metabolism, and trigger senescence caused by mitochondrial dysfunction [59].

The clinical characteristics of the patients were assessed through the application of diverse regression techniques to develop robust models and mitigate the potential bias arising from reliance on a single screening method. Specifically, LASSO regression analysis was employed to establish the prognostic model for OSCSRDEGs, which demonstrated a highly significant statistical association between the riskscores derived from this model and the prediction of OSl for patients with TNBC in the TCGA-TNBC dataset. The results of the ROC curve analysis revealed that the prognostic model of OSCSRDEGs exhibited a higher level of accuracy in predicting prognosis, with its efficacy showing gradual improvement over time. To evaluate the prognostic relevance of four OSCSRDEGs (CFI, DMD, NDRG2, NRP1) exhibiting notable distinctions between high and low risk score cohorts in both TCGA-TNBC and GSE57544 dataset, we developed a multivariate Cox prognostic model. The analysis of the nomogram indicated that DMD displayed the highest utility within the model, followed by NRP1, NDRG2, and CFI. Furthermore, the application of DCA demonstrated that the multivariate Cox regression model that we developed demonstrated superior clinical predictive performance across three-year, five-year, and one-year time spans. In summary, we conducted an analysis of clinical data and gene expression to develop a prognostic model that incorporates a range of OSCSRDEGs. The expression levels of these genes are closely linked to the biological behavior of TNBC and overall patient survival. Our prognostic model demonstrates excellent predictive capability and has significant implications for guiding the prognosis of TNBC patients.

There is a mounting body of evidence regarding the impact of OSCS on TME. Immune cells and stromal cells represent two prominent non-cancerous constituents within the tumor microenvironment (TME) and have been proposed as valuable entities for tumor diagnosis and prognostic assessment. To further investigate the impact and underlying mechanisms of OSCS on the immune microenvironment of TNBC, we developed an OSCSscore based on 27 OSCSRDEGs. Subsequently, we performed an analysis to compare the inherent differences in tumor infiltrating lymphocytes between the high-score group and low-score group in the TCGA-TNBC dataset. Initially, it was observed that the Stromal Score and ESTIMATE Score exhibited a significant increase in the high OSCSscore group compared to the low OSCSscore group. These was a noteworthy positive correlation between Stromal Score, ESTIMATE Score, and OSCSscore. These findings imply that OSCSRDEGs possess the ability to regulate and modify the tumor microenvironment, particularly the stromal components. Furthermore, the TCGA-TNBC dataset analyzed using the ssGSEA algorithm revealed statistically significant variations in the abundance of five immune cell types, namely CD56bright natural killer cell, Mast cell, Central memory CD4 T cell, Central memory CD8 T cell, and Plasmacytoid dendritic cell, between the Low/High OSCSscore group. Additionally, significant correlations were observed between these five cell types and OSCSRDEGs. Notably, dendritic cells (DCs), responsible for antigen processing and presentation to T cells in response to pathogens and cancer cells, can be further classified into conventional DCs (cDCs) and plasmacytoid DCs. Based on prior research [60], it has been observed that the levels of cDC and pDC infiltration, as determined by transcriptional profiling, are higher in TNBC compared to other breast cancer subtypes. Furthermore, a significant positive association has been found between pDC infiltration in TNBC and the proportion of $CD4^+$ memory T cells and $CD8^+$ T cells, as well as the IFN-γ score and cytolytic activity score. These findings align with our own study results. The analysis of CIBERSORT immune characteristics ultimately demonstrated statistically significant differences in the infiltration abundance of three immune cells, specifically Mast cells resting, T cells CD8, and T cells follicular helper, between the Low/High OSCSscore group. Among these immune cells, the mast cell warrants particular attention, as previous research [61] has suggested its association with heightened oxidative stress. Furthermore, the presence of mast cells in tumor sites contributes to the formation of the tumor immune microenvironment (TIME). Mast cells (MCs) have been a subject of debate within the breast cancer stromal compartment. Certain scholars posit that mast cell infiltration indicates a favorable prognosis for breast cancer [62], whereas other research groups have found that a higher density of mast cells in breast cancer tissue is associated with a worsened prognosis [63]. A comprehensive investigation reveals that MCs are significantly influenced by hormone receptors and HER2 status in breast cancer, with their levels notably elevated in aggressive subtypes of the disease, thereby serving as an unfavorable prognostic factor [64]. The findings of our study indicate that OSCS may have an impact on the prognosis of TNBC through its influence on mast cell function and status.

A valuable insight into the association between OSCS and TNBC has been provided by this study. However, it is important to acknowledge certain inherent limitations. TNBC encompasses a diverse range of malignancies with distinct histologic, genomic, and immunologic features. Although the primary focus of this study was on the significant role of OSCS in TNBC, it did not explore the correlation between the expression of OSCSRDEGs and prognosis in patients with different histological subtypes of TNBC. Additional research is necessary to explore the prognostic importance of OSCSRDEGs in the clinical management of TNBC, emphasizing the need for comprehensive investigations encompassing larger cohorts and diverse histopathological subtypes. Additionally, our analysis has shed light on potential mechanisms underlying distinct OSCSRDEGs in TNBC; however, the exact mechanisms remain elusive and necessitate further in vitro and in vivo experimental verification, as well as clinical research, to substantiate our findings.

In conclusion, our study conducted a comprehensive integrated analysis of OSCSRDEGs in TNBC, resulting in the successful identification of four crucial OSCSRDEGs (CFI, DMD, NDRG2, and NRP1). Additionally, we developed a precise OSCS-related prognostic model for TNBC patients. Moreover, we have elucidated the underlying mechanisms of OSCSRDEGs in TNBC and evaluated the association between their expression and the characterization of tumor microenvironment infiltration. These findings hold promise in providing therapeutic targets and valuable guidance for monitoring the effectiveness of personalized clinical therapy. However, further validation through future studies is necessary before implementing this prognostic model into medical practice.

Data availability statement

The datasets analyzed in this study are accessible from public repositories, as detailed in the "Materials and Methods" section. All data relevant to the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Fund

This study was supported by the Fundamental Research Funds for the Provincial Universities (Grant No. 2019-KYYWF-0349) to Ge Lou, the Research Project of Health Commission of Heilongjiang Province (Grant No. 2018–273) to Ge Lou.

Ethics approval

Not applicable to this study. Since all the data were retrieved from the online databases and published literature, it was confirmed that all written informed consent had already been obtained.

CRediT authorship contribution statement

Lihua Wu: Writing – original draft, Formal analysis. **Hongyan Zheng:** Validation, Data curation. **Xiaorong Guo:** Visualization, Software. **Nan Li:** Methodology, Investigation. **Luyao Qin:** Resources, Investigation. **Xiaoqing Li:** Visualization, Validation. **Ge Lou:** Writing – review & editing, Writing – original draft, Investigation.

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.

Acknowledgements

We thank all our authors for their efforts.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e34524.

References

[1] A.C. Garrido-Castro, N.U. Lin, K. Polyak, Insights into molecular classifications of triple-negative breast cancer: improving patient selection for treatment, Cancer Discov. 9 (2019) 176–198, https://doi.org/10.1158/2159-8290.CD-18-1177.
[2] L. Yin, J.J. Duan, X.W. Bian, S.C. Yu, Triple-negative breast cancer molecular subtyping and treatment progress, Breast Cancer Res. 22 (2020) 61, https://doi.org/10.1186/s13058-020-01296-5.
[3] S. Nofech-Mozes, M. Trudeau, H.K. Kahn, R. Dent, E. Rawlinson, P. Sun, S.A. Narod, W.M. Hanna, Patterns of recurrence in the basal and non-basal subtypes of triple-negative breast cancers, Breast Cancer Res. Treat. 118 (2009) 131–137, https://doi.org/10.1007/s10549-008-0295-8.
[4] G. Bianchini, J.M. Balko, I.A. Mayer, M.E. Sanders, L. Gianni, Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease, Nat. Rev. Clin. Oncol. 13 (2016) 674–690, https://doi.org/10.1038/nrclinonc.2016.66.
[5] F. Derakhshan, J.S. Reis-Filho, Pathogenesis of triple-negative breast cancer, Annu. Rev. Pathol. 17 (2022) 181–204, https://doi.org/10.1146/annurev-pathol-042420-093238.
[6] M. Nedeljkovic, A. Damjanovic, Mechanisms of chemotherapy resistance in triple-negative breast cancer-how we can rise to the challenge, Cells 8 (2019), https://doi.org/10.3390/cells8090957.
[7] A.M. Pisoschi, A. Pop, F. Iordache, L. Stanca, G. Predoi, A.I. Serban, Oxidative stress mitigation by antioxidants - an overview on their chemistry and influences on health status, Eur. J. Med. Chem. 209 (2021) 112891, https://doi.org/10.1016/j.ejmech.2020.112891.
[8] P.L. de Sa Junior, D.A.D. Camara, A.S. Porcacchia, P.M.M. Fonseca, S.D. Jorge, R.P. Araldi, A.K. Ferreira, The roles of ROS in cancer heterogeneity and therapy, Oxid. Med. Cell. Longev. (2017) 2467940, https://doi.org/10.1155/2017/2467940.
[9] Y.J. Huang, G.X. Nan, Oxidative stress-induced angiogenesis, J. Clin. Neurosci. 63 (2019) 13–16, https://doi.org/10.1016/j.jocn.2019.02.019.
[10] J.G. Gill, E. Piskounova, S.J. Morrison, Cancer, oxidative stress, and metastasis, Cold Spring Harbor Symp. Quant. Biol. 81 (2016) 163–175, https://doi.org/10.1101/sqb.2016.81.030791.
[11] J.D. Hayes, A.T. Dinkova-Kostova, K.D. Tew, Oxidative stress in cancer, Cancer Cell 38 (2020) 167–197, https://doi.org/10.1016/j.ccell.2020.06.001.
[12] C.L. Kuo, A. Ponneri Babuharisankar, Y.C. Lin, H.W. Lien, Y.K. Lo, H.Y. Chou, V. Tangeda, L.C. Cheng, A.N. Cheng, A.Y. Lee, Mitochondrial oxidative stress in the tumor microenvironment and cancer immunoescape: foe or friend? J. Biomed. Sci. 29 (2022) 74, https://doi.org/10.1186/s12929-022-00859-2.
[13] D. Hanahan, Hallmarks of cancer: new dimensions, Cancer Discov. 12 (2022) 31–46, https://doi.org/10.1158/2159-8290.CD-21-1059.
[14] W. Xue, L. Zender, C. Miething, R.A. Dickins, E. Hernando, V. Krizhanovsky, C. Cordon-Cardo, S.W. Lowe, Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas, Nature 445 (2007) 656–660, https://doi.org/10.1038/nature05529.
[15] B. Wang, J. Kohli, M. Demaria, Senescent cells in cancer therapy: friends or foes? Trends Cancer 6 (2020) 838–857, https://doi.org/10.1016/j.trecan.2020.05.004.

[16] L. Cuollo, F. Antonangeli, A. Santoni, A. Soriani, The senescence-associated secretory phenotype (SASP) in the challenging future of cancer therapy and age-related diseases, Biology 9 (2020), https://doi.org/10.3390/biology9120485.
[17] L. Chibaya, J. Snyder, M. Ruscetti, Senescence and the tumor-immune landscape: implications for cancer immunotherapy, Semin. Cancer Biol. 86 (2022) 827–845, https://doi.org/10.1016/j.semcancer.2022.02.005.
[18] F.L. Sarmiento-Salinas, A. Delgado-Magallon, J.B. Montes-Alvarado, D. Ramirez-Ramirez, J.C. Flores-Alonso, P. Cortes-Hernandez, J. Reyes-Leyva, I. Herrera-Camacho, M. Anaya-Ruiz, R. Pelayo, L. Millan-Perez-Pena, P. Maycotte, Breast cancer subtypes present a differential production of reactive oxygen species (ROS) and susceptibility to antioxidant treatment, Front. Oncol. 9 (2019) 480, https://doi.org/10.3389/fonc.2019.00480.
[19] A. Colaprico, T.C. Silva, C. Olsen, L. Garofano, C. Cava, D. Garolini, T.S. Sabedot, T.M. Malta, S.M. Pagnotta, I. Castiglioni, M. Ceccarelli, G. Bontempi, H. Noushmehr, TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data, Nucleic Acids Res. 44 (2016) e71, https://doi.org/10.1093/nar/gkv1507.
[20] M.J. Goldman, B. Craft, M. Hastie, K. Repecka, F. McDade, A. Kamath, A. Banerjee, Y. Luo, D. Rogers, A.N. Brooks, J. Zhu, D. Haussler, Visualizing and interpreting cancer genomics data via the Xena platform, Nat. Biotechnol. 38 (2020) 675–678, https://doi.org/10.1038/s41587-020-0546-8.
[21] A. Mayakonda, D.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, Genome Res. 28 (2018) 1747–1756, https://doi.org/10.1101/gr.239244.118.
[22] C.H. Mermel, S.E. Schumacher, B. Hill, M.L. Meyerson, R. Beroukhim, G. Getz, GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers, Genome Biol. 12 (2011) R41, https://doi.org/10.1186/gb-2011-12-4-r41.
[23] B. Weigelt, C.K. Ng, R. Shen, T. Popova, M. Schizas, R. Natrajan, O. Mariani, M.H. Stern, L. Norton, A. Vincent-Salomon, J.S. Reis-Filho, Metaplastic breast carcinomas display genomic and transcriptomic heterogeneity [corrected], Mod. Pathol. 28 (2015) 340–351, https://doi.org/10.1038/modpathol.2014.142.
[24] J.H. Taube, J.I. Herschkowitz, K. Komurov, A.Y. Zhou, S. Gupta, J. Yang, K. Hartwell, T.T. Onder, P.B. Gupta, K.W. Evans, B.G. Hollier, P.T. Ram, E.S. Lander, J. M. Rosen, R.A. Weinberg, S.A. Mani, Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 15449–15454, https://doi.org/10.1073/pnas.1004900107.
[25] T. Barrett, D.B. Troup, S.E. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, I.F. Kim, A. Soboleva, M. Tomashevsky, R. Edgar, NCBI GEO: mining tens of millions of expression profiles–database and tools update, Nucleic Acids Res. 35 (2007) D760–D765, https://doi.org/10.1093/nar/gkl887.
[26] S. Davis, P.S. Meltzer, GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor, Bioinformatics 23 (2007) 1846–1847, https://doi.org/10.1093/bioinformatics/btm254.
[27] G. Stelzer, N. Rosen, I. Plaschkes, S. Zimmerman, M. Twik, S. Fishilevich, T.I. Stein, R. Nudel, I. Lieder, Y. Mazor, S. Kaplan, D. Dahary, D. Warshawsky, Y. Guan-Golan, A. Kohn, N. Rappaport, M. Safran, D. Lancet, The GeneCards suite: from gene data mining to disease genome sequence analyses, Curr Protoc Bioinformatics 54 (1) (2016) 30 31–31 30 33, https://doi.org/10.1002/cpbi.5.
[28] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, J.P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 15545–15550, https://doi.org/10.1073/pnas.0506580102.
[29] A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdottir, P. Tamayo, J.P. Mesirov, Molecular signatures database (MSigDB) 3.0, Bioinformatics 27 (2011) 1739–1740, https://doi.org/10.1093/bioinformatics/btr260.
[30] P. Charoentong, F. Finotello, M. Angelova, C. Mayer, M. Efremova, D. Rieder, H. Hackl, Z. Trajanoski, Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade, Cell Rep. 18 (2017) 248–262, https://doi.org/10.1016/j.celrep.2016.12.019.
[31] D.A. Barbie, P. Tamayo, J.S. Boehm, S.Y. Kim, S.E. Moody, I.F. Dunn, A.C. Schinzel, P. Sandy, E. Meylan, C. Scholl, S. Frohling, E.M. Chan, M.L. Sos, K. Michel, C. Mermel, S.J. Silver, B.A. Weir, J.H. Reiling, Q. Sheng, P.B. Gupta, R.C. Wadlow, H. Le, S. Hoersch, B.S. Wittner, S. Ramaswamy, D.M. Livingston, D. M. Sabatini, M. Meyerson, R.K. Thomas, E.S. Lander, J.P. Mesirov, D.E. Root, D.G. Gilliland, T. Jacks, W.C. Hahn, Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1, Nature 462 (2009) 108–112, https://doi.org/10.1038/nature08460.
[32] A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, C.D. Hoang, M. Diehn, A.A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles, Nat. Methods 12 (2015) 453–457, https://doi.org/10.1038/nmeth.3337.
[33] S. Engebretsen, J. Bohlin, Statistical predictions with glmnet, Clin. Epigenet. 11 (2019) 123, https://doi.org/10.1186/s13148-019-0730-1.
[34] W. Cai, M. van der Laan, Nonparametric bootstrap inference for the targeted highly adaptive least absolute shrinkage and selection operator (LASSO) estimator, Int. J. Biostat. (2020), https://doi.org/10.1515/ijb-2017-0070.
[35] S. Hanzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7, https://doi.org/10.1186/1471-2105-14-7.
[36] G. Yu, Gene Ontology semantic similarity analysis using GOSemSim, Methods Mol. Biol. 2117 (2020) 207–215, https://doi.org/10.1007/978-1-0716-0301-7_11.
[37] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (2000) 27–30, https://doi.org/10.1093/nar/28.1.27.
[38] G. Yu, L.G. Wang, Y. Han, Q.Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (2012) 284–287, https://doi.org/10.1089/omi.2011.0118.
[39] S.Y. Park, Nomogram: an analogue tool to deliver digital knowledge, J. Thorac. Cardiovasc. Surg. 155 (2018) 1793, https://doi.org/10.1016/j.jtcvs.2017.12.107.
[40] T. Tataranni, C. Piccoli, Dichloroacetate (DCA) and cancer: an overview towards clinical applications, Oxid. Med. Cell. Longev. (2019) 8201079, https://doi.org/10.1155/2019/8201079.
[41] R. Farahzadi, B. Valipour, E. Fathi, S. Pirmoradi, O. Molavi, S. Montazersaheb, Z. Sanaat, Oxidative stress regulation and related metabolic pathways in epithelial-mesenchymal transition of breast cancer stem cells, Stem Cell Res. Ther. 14 (2023) 342, https://doi.org/10.1186/s13287-023-03571-6.
[42] M. Fane, A.T. Weeraratna, How the ageing microenvironment influences tumour progression, Nat. Rev. Cancer 20 (2020) 89–106, https://doi.org/10.1038/s41568-019-0222-9.
[43] K. Lei, S.S. Kang, E.H. Ahn, C. Chen, J. Liao, X. Liu, H. Li, L.E. Edgington-Mitchell, L. Jin, K. Ye, C/EBPbeta/AEP signaling regulates the oxidative stress in malignant cancers, stimulating the metastasis, Mol. Cancer Therapeut. 20 (2021) 1640–1652, https://doi.org/10.1158/1535-7163.MCT-21-0019.
[44] E. Hatem, S. Azzi, N. El Banna, T. He, A. Heneman-Masurel, L. Vernis, D. Baille, V. Masson, F. Dingli, D. Loew, B. Azzarone, P. Eid, G. Baldacci, M.E. Huang, Auranofin/vitamin C: a novel drug combination targeting triple-negative breast cancer, J. Natl. Cancer Inst. 111 (2019) 597–608, https://doi.org/10.1093/jnci/djy149.
[45] B. de Paula, R. Kieran, S.S.Y. Koh, S. Crocamo, E. Abdelhay, D. Munoz-Espin, Targeting senescence as a therapeutic opportunity for triple-negative breast cancer, Mol. Cancer Therapeut. 22 (2023) 583–598, https://doi.org/10.1158/1535-7163.MCT-22-0643.
[46] S. McAvoy, S. Ganapathiraju, D.S. Perez, C.D. James, D.I. Smith, DMD and IL1RAPL1: two large adjacent genes localized within a common fragile site (FRAXC) have reduced expression in cultured brain tumors, Cytogenet. Genome Res. 119 (2007) 196–203, https://doi.org/10.1159/000112061.
[47] F. Muntoni, S. Torelli, A. Ferlini, Dystrophin and mutations: one gene, several proteins, multiple phenotypes, Lancet Neurol. 2 (2003) 731–740, https://doi.org/10.1016/s1474-4422(03)00585-4.
[48] M. Naidoo, K. Anthony, Dystrophin Dp71 and the neuropathophysiology of Duchenne muscular dystrophy, Mol. Neurobiol. 57 (2020) 1748–1767, https://doi.org/10.1007/s12035-019-01845-w.
[49] L. Jones, M. Naidoo, L.R. Machado, K. Anthony, The Duchenne muscular dystrophy gene and cancer, Cell. Oncol. 44 (2021) 19–32, https://doi.org/10.1007/s13402-020-00572-y.
[50] L.N. Luce, M. Abbate, J. Cotignola, F. Giliberto, Non-myogenic tumors display altered expression of dystrophin (DMD) and a high frequency of genetic alterations, Oncotarget 8 (2017) 145–155, https://doi.org/10.18632/oncotarget.10426.
[51] C.H. Saravia, C. Flores, L.J. Schwarz, L. Bravo, J. Zavaleta, J. Araujo, S. Neciosup, J.A. Pinto, Patterns of mutation enrichment in metastatic triple-negative breast cancer, Clin. Med. Insights Oncol. 13 (2019) 1179554919868482, https://doi.org/10.1177/1179554919868482.

[52] D.Q. Cai, D. Cai, Y. Zou, X. Chen, Z. Jian, M. Shi, Y. Lin, J. Chen, Construction and validation of chemoresistance-associated tumor- infiltrating exhausted-like CD8+ T cell signature in breast cancer: cr-TILCD8TSig, Front. Immunol. 14 (2023) 1120886, https://doi.org/10.3389/fimmu.2023.1120886.
[53] K.W. Lee, S. Lim, K.D. Kim, The function of N-myc downstream-regulated gene 2 (NDRG2) as a negative regulator in tumor cell metastasis, Int. J. Mol. Sci. 23 (2022), https://doi.org/10.3390/ijms23169365.
[54] G. Kim, S. Lim, K.D. Kim, N-Myc downstream-regulated gene 2 (NDRG2) function as a positive regulator of apoptosis: a new insight into NDRG2 as a tumor suppressor, Cells 10 (2021), https://doi.org/10.3390/cells10102649.
[55] A. Lee, S. Lim, J. Oh, J. Lim, Y. Yang, M.S. Lee, J.S. Lim, NDRG2 expression in breast cancer cells downregulates PD-L1 expression and restores T cell proliferation in tumor-coculture, Cancers 13 (2021), https://doi.org/10.3390/cancers13236112.
[56] S. Sajadimajd, M. Khazaei, Oxidative stress and cancer: the role of Nrf2, Curr. Cancer Drug Targets 18 (2018) 538–557, https://doi.org/10.2174/1568009617666171002144228.
[57] K. Sinha, J. Das, P.B. Pal, P.C. Sil, Oxidative stress: the mitochondria-dependent and mitochondria-independent pathways of apoptosis, Arch. Toxicol. 87 (2013) 1157–1180, https://doi.org/10.1007/s00204-013-1034-4.
[58] K. Todkar, L. Chikhi, V. Desjardins, F. El-Mortada, G. Pepin, M. Germain, Selective packaging of mitochondrial proteins into extracellular vesicles prevents the release of mitochondrial DAMPs, Nat. Commun. 12 (2021) 1971, https://doi.org/10.1038/s41467-021-21984-w.
[59] Y. Guo, T. Guan, K. Shafiq, Q. Yu, X. Jiao, D. Na, M. Li, G. Zhang, J. Kong, Mitochondrial dysfunction in aging, Ageing Res. Rev. 88 (2023) 101955, https://doi.org/10.1016/j.arr.2023.101955.
[60] M. Oshi, S. Newman, Y. Tokumaru, L. Yan, R. Matsuyama, P. Kalinski, I. Endo, K. Takabe, Plasmacytoid dendritic cell (pDC) infiltration correlate with tumor infiltrating lymphocytes, cancer immunity, and better survival in triple negative breast cancer (TNBC) more strongly than conventional dendritic cell (cDC), Cancers 12 (2020), https://doi.org/10.3390/cancers12113342.
[61] S. Gangemi, P.L. Minciullo, D. Magliacane, S. Saitta, S. Loffredo, A. Saija, M. Cristani, G. Marone, M. Triggiani, Oxidative stress markers are increased in patients with mastocytosis, Allergy 70 (2015) 436–442, https://doi.org/10.1111/all.12571.
[62] A. Glajcar, J. Szpor, A. Pacek, K.E. Tyrak, F. Chan, J. Streb, D. Hodorowicz-Zaniewska, K. Okon, The relationship between breast cancer molecular subtypes and mast cell populations in tumor microenvironment, Virchows Arch. 470 (2017) 505–515, https://doi.org/10.1007/s00428-017-2103-5.
[63] J. Sang, D. Yi, X. Tang, Y. Zhang, T. Huang, The associations between mast cell infiltration, clinical features and molecular types of invasive breast cancer, Oncotarget 7 (2016) 81661–81669, https://doi.org/10.18632/oncotarget.13163.
[64] E. Carpenco, R.A. Ceausu, A.M. Cimpean, P.N. Gaje, L. Saptefrati, V. Fulga, V. David, M. Raica, Mast cells as an indicator and prognostic marker in molecular subtypes of breast cancer, In Vivo 33 (2019) 743–748, https://doi.org/10.21873/invivo.11534.