Proceedings of the 2nd International Conference on Modern Medicine and Global Health
DOI: 10.54254/2753-8818/33/20240908

Prognostic value of mitophagy gene PINK1 in triple-negative breast cancer

Zengjian Tian[1], Shilong Yu[1] and Yongzhao Zhu[2, 3]

[1]The First School of Clinical Medicine, Ningxia Medical University, Yinchuan, Ningxia, China
[2]Institute of Medical Sciences, General Hospital of Ningxia Medical University, Yinchuan, Ningxia, China

[3]btxnxu05@163.com

Abstract. Mitochondrial autophagy plays a crucial role in the development, progression, and treatment of triple-negative breast cancer (TNBC). A total of 43 genes related to mitochondria were obtained from the scientific literature in Pubmed. To assess the expression of these genes, they were categorized into high and low-scoring groups using the ssGSEA algorithm. Additionally, the infiltration of immune cells in the combined dataset was evaluated using the CIBERSORTx algorithm. To analyze the expression of the selected genes, one-way Cox regression analysis, along with clinical staging variables, was employed. A multifactorial Cox regression model was constructed, and a forest plot was generated to assess the predictive power of the model on the actual outcomes. The accuracy of the model was confirmed by validating the expression of the PINK1 gene and its model through immunohistochemistry. TNBC cancer was categorized into two molecular subtypes (cluster1 and cluster2) based on the MRGs. Finally, a prognostic model was developed using the PINK1 gene, which demonstrated a considerable degree of accuracy. The validity of the PINK1 gene and its model was confirmed through immunohistochemistry, indirectly indicating that PINK1 might serve as a potential therapeutic target and prognostic biomarker for TNBC.

Keywords: mitochondrial autophagy; TNBC; immune cell infiltration; predictive modeling; PINK1.

1. Introduction

Breast cancer, a highly diverse disease characterized by various molecular subtypes, relies on the presence of hormone receptor (HR) and human epidermal growth factor 2 (HER2) to determine its categorization into three distinct subtypes: HR+/HER2-, HER2+, and TNBC cancer (TNBC) [1, 2]. In terms of therapeutic choices for breast carcinogenesis, these subtypes serve as primary guidelines, disregarding the molecular heterogeneity. It is crucial to note that breast cancer stands as the second leading cause of death among women due to its tendency for rapid metastasis and disease recurrence, with TNBC cancer emerging as one of the subtypes linked to an unfavorable prognosis [2, 3]. Given the lack of appropriate targets and the considerable heterogeneity surrounding TNBC cancer, the search for well-suited targets has become a central focus of current research efforts.

DOI: 10.54254/2753-8818/33/20240908

Mitophagy, also known as mitochondrial autophagy, is a form of selective autophagy that helps maintain the functional integrity of mitochondria and cellular homeostasis. Its main function is to remove dysfunctional mitochondria from the cytoplasm in a targeted manner [4]. When external factors like nutritional deficiencies and cellular senescence occur, the mitochondria within the cell undergo depolarization and suffer damage. Damaged mitochondria are then enclosed within autophagosomes and fused with lysosomes. This process helps degrade the damaged mitochondria and ensures the stability of the intracellular environment. The important role of mitochondrial autophagy has been demonstrated in various diseases such as cardiovascular diseases, Alzheimer's disease, and cancer [5-8]. In the case of cancer, there is evidence suggesting that autophagy acts as a mechanism of tumor suppression. Tumor cells themselves experience oxidative and metabolic stresses, leading to mitochondrial abnormalities, gene mutations, and DNA strand breaks. Autophagy acts as a scavenger, removing damaged organelles and reactive oxygen species, thus contributing to tumor suppression [8-10]. However, in the context of cancer pathogenesis, autophagy also provides insights for potential new drug therapies. As cancer progresses to advanced stages, autophagy may serve as a mechanism to meet the metabolic demands of cancer cells and repair intracellular damage caused by the aggressive tumor environment [9, 11, 12]. Recent studies have highlighted the indispensable role of mitochondrial autophagy in progressive tumors, which significantly impacts the survival of tumor patients.

The discovery resulting from this investigation holds significant promise in the identification of novel cancer biomarkers. Presently, there is limited literature on the examination of mitochondrial autophagy in TNBC cancer. Accordingly, this study aims to investigate the biological functions of mitochondrial autophagy in TNBC cancer using a bioinformatics approach. Furthermore, it seeks to develop a prognostic model based on said autophagy. The outcomes of this research possess substantial potential in the detection of fresh cancer biomarkers.

2. Materials and methods

2.1. Data sources

In this study, we utilized the R package TCGAbiolinks to download the expression matrix of the breast cancer (BRCA) dataset (TCGA-BRCA) from the cancer genome atlas (TCGA) specifically screened for the expression matrices of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) to obtain triple negative breast cancer (TNBC) samples. The TNBC dataset (TCGA-TNBC) expression matrix consisted of 116 triple negative breast cancer samples (cancer group, subgroup: TNBC) and 11 paracancer samples (normal group, subgroup: Normal). All samples with matched clinical information were included in this study and were normalized to the Fragments Per Kilobaseper Million (FPKM) format.

The corresponding clinical data were obtained from the UCSC Xena database . The data of the dataset (TCGA-TNBC) were normalized by R package limma. We also selected the "Masked Somatic Mutation" data from the TCGA website as the somatic mutation data of TNBC patients, pre-processed the data using VarScan software, and visualized the somatic mutations using the R package map tools. To analyze the copy number variations (CNV) of TNBC patients, we used the R package TCGAbiolinks to download the "Copy Number Segment" data of patients. The downloaded and processed CNV segments were then analyzed by GISTIC 2.0, which uses the default parameter settings for all GISTIC 2.0 analyses.

We passed the R package GEOquery From the GEO database. And downloaded the TNBC-related dataset GSE38959 from the GEO database, GSE65194 and GSE53752 . Among them, GSE38959 was from Homo Sapiens, and the data platform was GPL4133 Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Feature Number version), which contained microarray gene-expression profiles of 30 TNBC cancer samples and 13 normal Microarray gene-expression profiles of 30 TNBC cancer samples and 13 normal tissue samples; GSE65194 was obtained from Homo Sapiens, and the data platform is GPL570, Affymetrix Human Genome U133 Plus 2.0 Array, containing a total of 41 TNBC cancer samples and 11 normal tissue samples of microarray gene expression profiling data; of which GSE53752 is from Homo Sapiens, the data platform is GPL7264. Agilent-012097 Human 1A

DOI: 10.54254/2753-8818/33/20240908

Microarray (V2) G4110B (Probe Name version), which contains microarray gene expression profile data of 51 TNBC cancer samples and 25 normal tissue samples. All of the above samples were included in this study. The dataset probe names were annotated using the GPL platform files corresponding to the microarrays.

2.2. Analysis of differential expression

To identify potential mechanisms of action and associated biological features and pathways of the differential genes in the cancer and normal groups of TNBC, first,we normalized the datasets TCGA-TNBC, GSE38959, GSE65194, and GSE53752, using the R package limma, and then used the R package sva.The TNBC datasets GSE38959, GSE65194 and GSE53752 were processed to remove the batch effect to obtain the merged GEO dataset, and the datasets before and after the removal of the batch effect were compared by the Distribution Box-Line Plot and PCA (Principal Component Analysis) Plot. We then analyzed the expression profiles of the dataset TCGA-TNBC after treatment by dividing them into the cancer group and the normal group for differential analysis, and obtained the differentially expressed genes (DEGs) among different subgroups of the dataset TCGA-TNBC, in which the genes with logFC > 0.5 and P.adj < 0.05 were the genes with up-regulated expression. The genes with logFC >0.5 and P.adj < 0.05 were differentially expressed genes (up regulated genes), and the genes with logFC < -0.5 and P.adj < 0.05 were differentially expressed genes (down regulated genes), and the results of the differentially expressed genes analysis were demonstrated by volcano plots with R package ggplot2.

We took the intersections of the obtained up-regulated differentially expressed genes and down-regulated differentially expressed genes with MRGs respectively and plotted Venn diagrams to show them. The two parts of the intersections were merged into the differentially expressed genes of TNBC. Subsequently, group comparison plots were drawn with differentially expressed genes in the dataset TCGA-TNBC and the merged dataset respectively, and the genes with the same trend of high and low expression and statistically significant differences were used as the important genes for further analysis in subsequent years, and the expression of the genes in the dataset TCGA-TNBC was demonstrated by heatmap drawn with R package heatmap.

2.3. Mutation analysis

We selected “Masked Somatic Mutation” data from the TCGA website (https://portal.gdc.cancer.gov/) as the somatic mutation data of TNBC cancer (TCGA-TNBC) patients, and used VarScan software to preprocess the data and the maftools R package to analyze the data.The data were pre-processed using VarScan software and visualized using the maftools R package for somatic mutations.

Copy number variation (CNV) is a genomic alteration phenomenon that could lead to copy number abnormalities in one or more genes. To analyze gene copy number variations in TNBC cancer patients, we used the R package TCGAbiolinks to download Masked Copy Number Segment data from patients and used the Hiplot website (https://hiplot-academic.com/advance/gistic2) to analyze the downloaded CNV segments analyzed by GISTIC 2.0. The GISTIC 2.0 analysis results were visualized by the R package Maftools.

2.4. GSEA

GSEA (Gene Set Enrichment Analysis) is used to assess the distribution trend of genes of a predefined gene set in a gene list sorted by phenotypic relevance to determine their contribution to the phenotype. In this way,We obtained the data from the Molecular Signatures Database (MSigDB) database. We obtained the gene set “c2.cp.v7.2.symbols” from the Molecular Signatures Database (MSigDB), and analyzed the gene set using the R package clusterProfiler . Enrichment analysis was performed on all genes in the cancer and normal groups of the TCGA-TNBC dataset using the following parameters for this GSEA: seed of 2022, number of calculations of 100,000, minimum number of genes in each gene set of 5, maximum number of genes in each gene set of 500, p-value correction by Benjamini-Hochberg (BH), significant enrichment filtering criteria by Benjamini-Hochberg (BH), and significant enrichment

DOI: 10.54254/2753-8818/33/20240908

filtering criteria by Benjamini-Hochberg (BH). and the screening criteria for significant enrichment were P.adj < 0.05 and FDR value (q.value) < 0.25.

2.5. GO and KEGG

Gene Ontology (GO) and The Kyoto Encyclopedia of Genes and Genomes (KEGG) are commonly used for conducting large-scale functional enrichment studies, which include biological process (BP), molecular function (MF), and cellular component (CC).Analysis is an essential method in these studies. The analysis involved the use of a widely used database for storing information about genomes, biological pathways, diseases, and drugs. To perform the analysis, we utilized the R package cluster Profiler and conducted GO and KEGG annotation analysis of MRGs. The entry screening criteria for statistical significance were set as P < 0.05 and FDR value (q.value) < 0.2. We applied P-value correction using the Benjamini-Hochberg (BH) method.

2.6. Consistency clustering

Consistency clustering (Consensus Clustering) is a resampling-based algorithm for identifying each sample and its subgroup number and verifying the reasonableness of the clusters. Consensus clustering is multiple iterations over subsamples of the data set and provides an indication of cluster stability and parameter decisions by utilizing subsampling to induce sampling variability. We used the R package Consensus Cluster Plus.

2.7. GSVA

Gene Set Variation Analysis (GSVA) is a non-parametric, unsupervised analysis method that evaluates the gene set enrichment results of microarrayed nuclear transcriptomes by converting gene expression matrices across samples into gene set expression matrices across samples. From there, it is used to assess whether different pathways are enriched among different samples. We obtained the "h.all.v7.4.symbols.gmt" gene set from the MSigDB database and performed GSVA analysis on the dataset TCGA-TNBC at the gene expression level to calculate the difference in functional enrichment between the two sets of samples, and the screening criterion for significant enrichment was P < 0.05.

2.8. MRGs scores for ssGSEA algorithm

The single-sample gene-set enrichment analysis (ssGSEA) algorithm quantifies the relative abundance of each gene in the dataset samples. We used the R package GSVA via the ssGSEA algorithm based on the MRGs scores of each sample in the dataset TCGA-TNBC, and the scores could be used to group the cancer group samples into high and low scores and to perform correlation analyses, among other things.

2.9. CIBERSORTx immunoinfiltration analysis

CIBERSORTx (https://cibersortx.stanford.edu/) is a tool used for deconvolution of expression matrices of human immune cell subtypes.In this study, the CIBERSORTx algorithm was used to evaluate the immune cell infiltration status of the merged dataset from the TNBC group of the TCGA-TNBC dataset. The results were presented using stacked bar graphs, which demonstrate the proportion of different immune cells in the samples. Additionally, boxplot plots were used to illustrate the difference in immune cell infiltration abundance between samples from the high and low MRGs scores groups of the TCGA-TNBC dataset.

2.10. Cox model.

To investigate the clinical prognostic value of these screened MRGs for TNBC, we performed a one-way Cox regression analysis of the expression of screened MRGs in the dataset TCGA-TNBC in combination with clinical staging variables. The variables with P < 0.1 were selected to be included in the multifactorial Cox regression analysis, and the multifactorial Cox regression model was constructed and demonstrated by drawing a forest plot. Based on the results of the multifactorial Cox regression analysis, we constructed a line graph (nomogram) to predict the survival probability of TNBC patients

DOI: 10.54254/2753-8818/33/20240908

at 1, 3, and 5 years. The scales are set to characterize the individual variables within the multifactor regression model, and the total score is calculated to predict the probability of an event occurring. Finally, we use calibration curves to evaluate the accuracy and discriminatory power of the plots. We used the R package "rms" to construct column plots and calibration curves. Decision curve analysis (DCA) is a simple method for evaluating clinical prediction models, diagnostic tests and molecular markers. We used the R package ggDCA. DCA plots were drawn to assess the effect of nomogram modeling for predicting 1, 3 and 5-year survival outcomes in ESCA patients.

2.11. PINK1 expression in TNBC

2.11.1. Patient sample information. Specimens were collected from the General Hospital of Ningxia Medical University during the period of January 1, 2016, to December 31, 2016. The present study confirmed the pathological diagnosis as triple-negative breast cancer. The pathological information was comprehensive, and the patients did not have any significant underlying diseases or other malignant tumors affecting the systemic condition. The study included a total of 50 female participants, ranging in age from 29 to 83 years, all of whom had invasive ductal breast cancer confirmed by histological examination. The average age of the subjects was 59 years, with a median age of 51 years. For each of the 50 paraffin specimens, a specific tumor area along with neighboring tissues was chosen. A cylindrical core tissue sample measuring 0.6mm in diameter was extracted from the visible section of every paraffin block. These core specimens were then meticulously placed into new recipient paraffin blocks (20x35mm) to form a tissue microarray (TMA). All patients who possessed paraffin specimens underwent surgical procedures, which included both axillary lymph node dissection and modified radical mastectomy. The follow-up duration ranged from 1 to 60 months. Individuals lacking a definitive histopathological diagnosis or whose samples had insufficient cancer cells were excluded from the analysis. Comprehensive data regarding clinicopathological parameters and patients with long-term follow-up were meticulously collected. Informed consent was acquired and signed by all patients, and the Ethics Committee of the General Hospital of Ningxia Medical University approved the study.

2.11.2. Immunohistochemistry (IHC). To assess the expression of PINK1, we conducted immunohistochemical staining (IHC). Initially, we deparaffinized the samples twice in xylene for 15 minutes each and then rehydrated them using an alcoholic solution with a progressively reduced concentration gradient. To prevent any non-specific staining, a blocking agent of 3% H2O2 was applied for 1 minute at room temperature. Subsequently, the samples were incubated with a rabbit monoclonal antibody specific to human PINK1, and the specimens were then washed with PBS for 5 minutes. Further, we utilized secondary antibodies (biotin-conjugated goat anti-rabbit, ab207995, 1:4000) and incubated them for 1 hour at 37°C. Finally, we employed 3,3'-diaminobenzidine for chromogen detection. The IHC staining related to Pink1 was interpreted based on staining intensity (3 for strong, 2.1 for mild, and 0 for negative) as well as the percentage of positive cells (4 for ≥66%, 3 for ≥33% and <66%, 2 for ≥10% and <33%, 1 for <10%, and 0 for negative).

To calculate the PINK1 IHC scores, the product of intensity and proportion was obtained. The examined cases were divided into two categories based on their IHC scores: a high expression group (8-12) and a low expression group (0-6). After tabbing and staining, the final tissue microarray consisted of 55 spots that could be evaluated (55 TNBC tissue and 5 paracancerous tissue). Two pathologists, blind to any clinicopathological information about the samples, evaluated the stained tissue microarrays. Differences in interpretation were resolved by using a confocal microscope with 8x magnification (Leica TCS SP40 CARS), which allowed for simultaneous re-examination of the specimens by both pathologists. Eventually, the expression level of PINK1 in the specimens was determined, and survival analysis was conducted.

DOI: 10.54254/2753-8818/33/20240908

2.12. Statistical analysis

All data processing and analysis in this study were done through R software (Version 4.1.2), and continuous variables were presented as mean ± standard deviation. Comparisons between the two groups of continuous variables were made using the Wilcoxon rank sum test (i.e., Wilcoxon rank sum test, Wilcoxon rank sum test), and the statistical significance of normally distributed variables was estimated by independent STUDENT t-test. The chi-square test or Fisher exact test was used to compare and analyze the statistical significance between two groups of categorical variables. subject operating characteristic curve (ROC). Based on R package pROC. If not specified, results were calculated by Spearman correlation analysis of correlation coefficients between different molecules and all P-statistics were two-sided, with P-values less than 0.05 as the criterion for statistical significance.

3. Results

3.1. Aa Analysis of differential expression

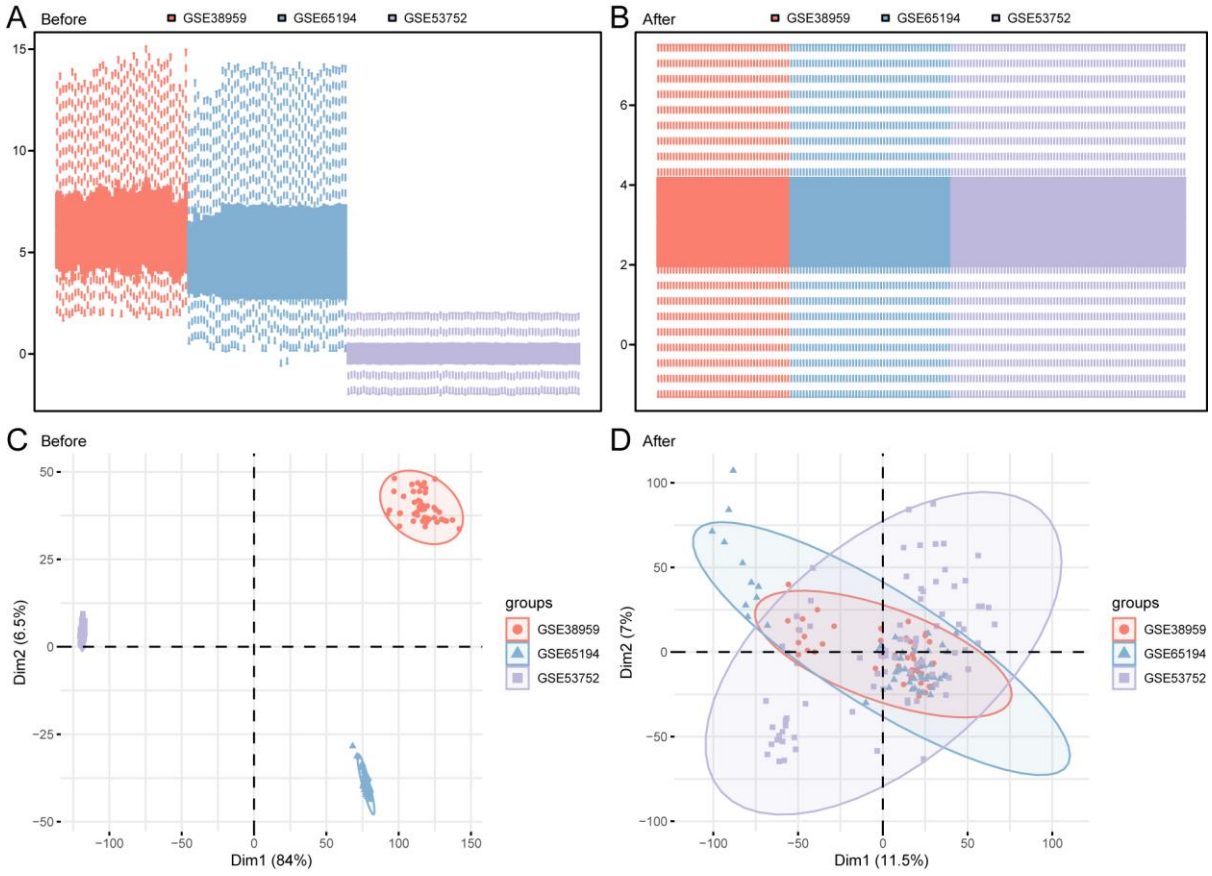

Figure 1. Boxplots and PCA plots of sample distribution before and after data pooling

A line plot of the sample distribution for the data set is shown in Box (A) and after (B). C-D presents the PCA plot before (C) and after (D) for the data set. PCA stands for principal component analysis.

DOI: 10.54254/2753-8818/33/20240908

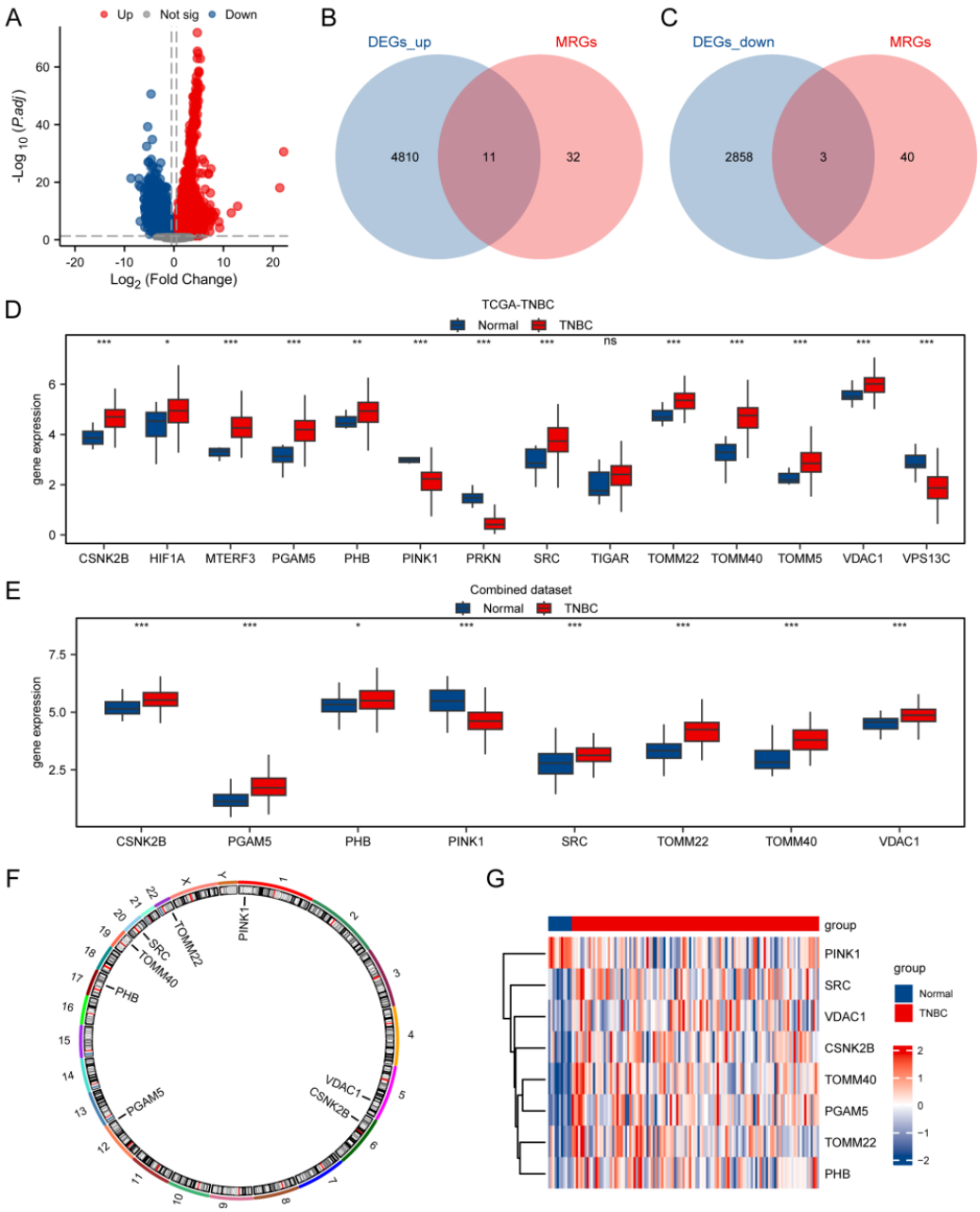

Figure 2. Differential expression of MRGs in the dataset TCGA-TNBC

(A). Volcano plots were used to analyze the differential expression between the cancer group (subgroup: TNBC) and the normal group (subgroup: Normal) in the TCGA-TNBC dataset. (B). Venn plots were generated to show the intersection of up-regulated differentially expressed genes identified from the TCGA-TNBC dataset with MRGs (mitophagy related genes). (C). Displayed in Venn plots is the overlap of down-regulated differentially expressed genes obtained from the TCGA-TNBC dataset with MRGs. (D). The comparison of gene grouping within the TCGA-TNBC dataset is presented. (E). The comparison of gene grouping within the merged dataset is presented. (F). Chromosomal localization information for MRGs is provided. (G). A heatmap is used to demonstrate the expression pattern of MRGs in the cancer group (subgroup: TNBC) and the normal group (subgroup: Normal) within the TCGA-TNBC dataset from tcga (The Cancer Genome Atlas), focused on tNBC (triple-negative breast cancer).

DOI: 10.54254/2753-8818/33/20240908

3.2. *Mutation analysis of MRGs in TNBC patients*

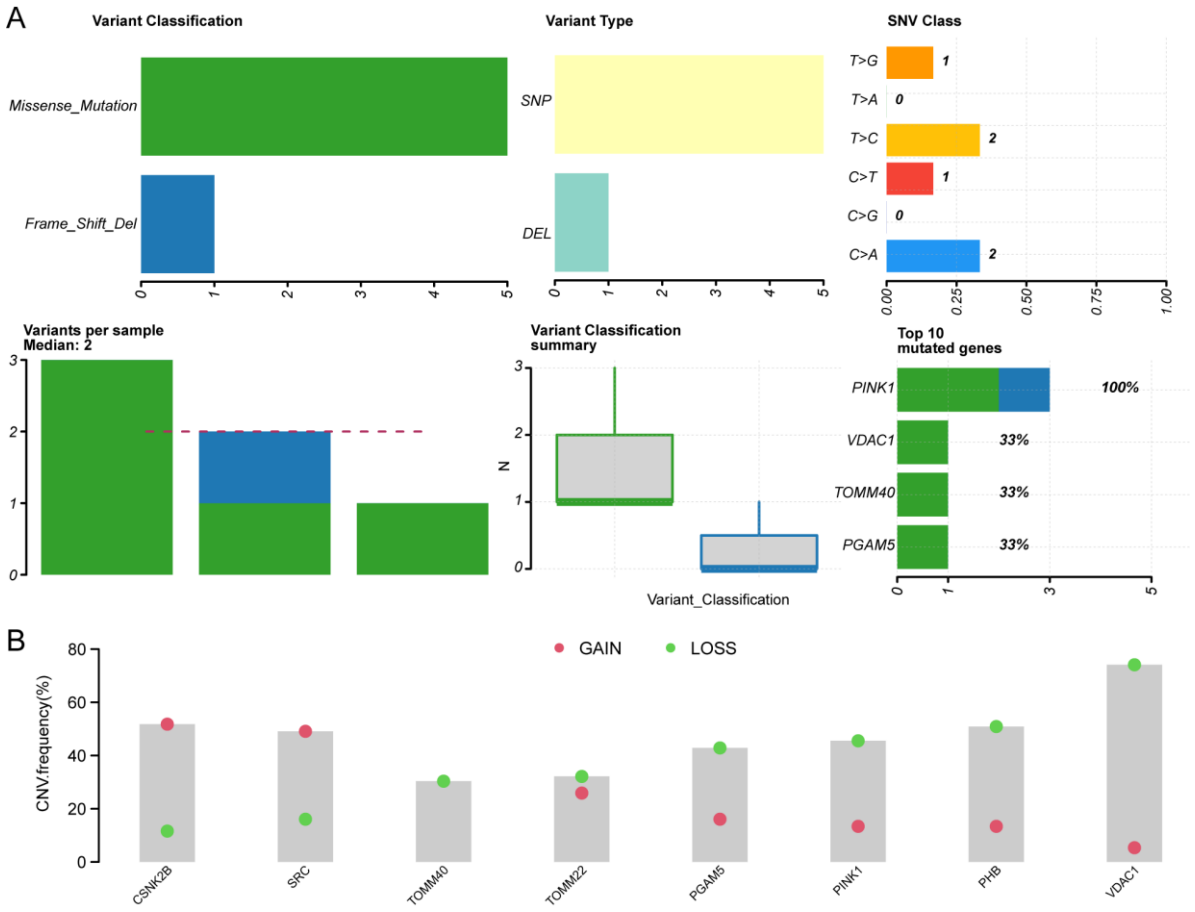

Figure 3. Mutation analysis of MRGs in TNBC patients with

(A).Demonstration of somatic mutations in MRGs in patients with TNBC. (B). Demonstration of copy number variations in MRGs in patients with TNBC.TNBC, triple negative breast cancer. SNP, single nucleotide polymorphism; CNV, copy number variations.

DOI: 10.54254/2753-8818/33/20240908

3.3. *Correlation analysis and prognostic analysis of MRGs*

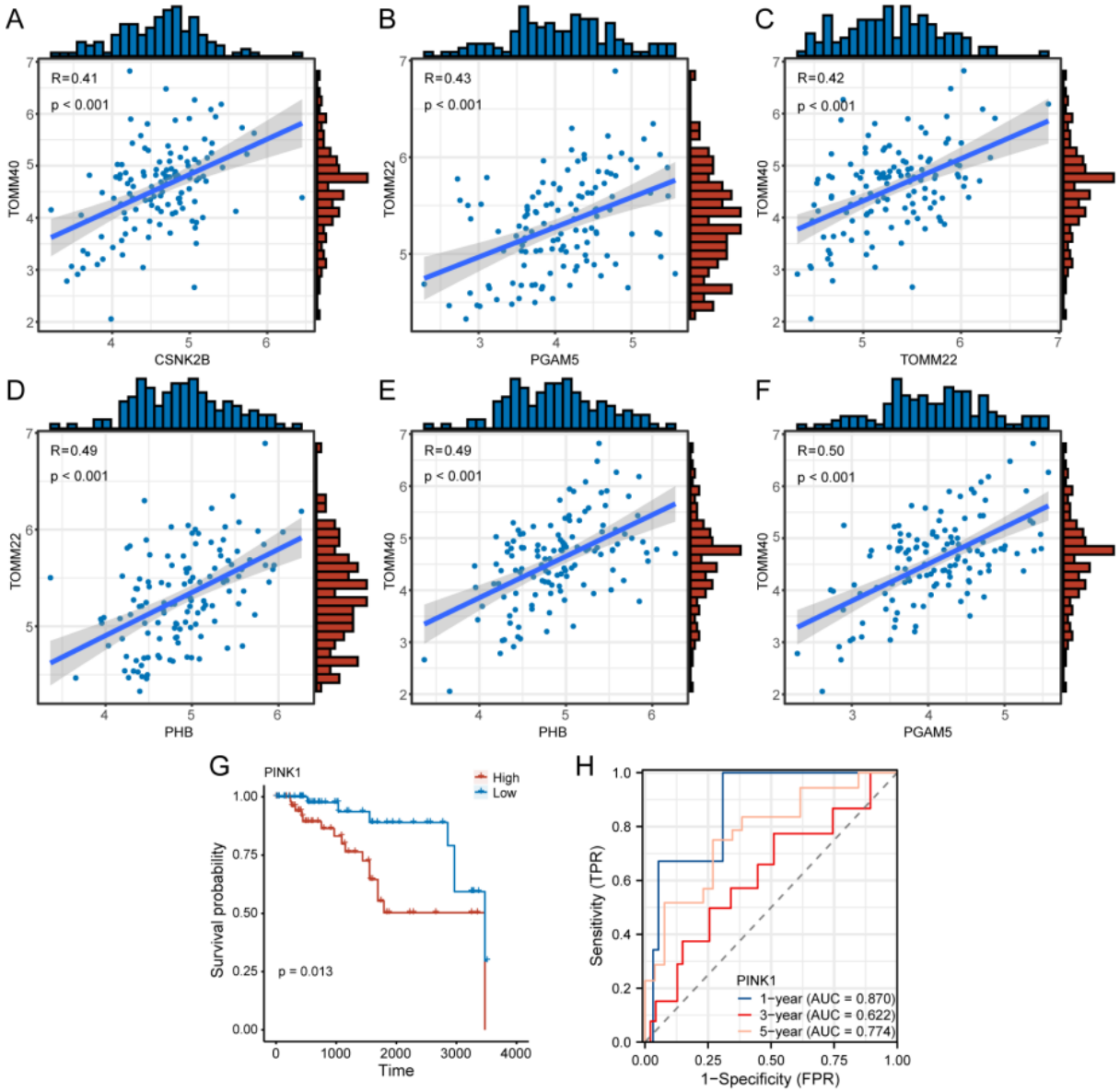

Figure 4. Analysis of the correlation and prognosis of MRGs in the TCGA-TNBC dataset.

(A). Correlation scatter plot of gene CSNK2B with gene TOMM40 is shown. (B). Correlation scatter plot of gene PGAM5 with gene TOMM22 is shown.(C). Correlation scatter plot of gene TOMM22 with gene TOMM40 is shown. (D). Correlation scatter plot of gene PHB with gene TOMM22 is shown. (E). Correlation scatter plot of gene PHB with gene TOMM40 is shown. (F). Correlation scatter plot of gene PGAM5 with gene TOMM40 is shown. (G). KM survival curves illustrated for the PINK1 gene in subgroups with low and low expression levels. Displayed correlation scatterplot. (H). Time-dependent ROC curves showcasing the performance of PINK1 in predicting outcomes in the TCGA-TNBC dataset. (I). Time-dependent ROC curves evaluating the predictive ability of PINK1 in the TCGA-TNBC dataset.

DOI: 10.54254/2753-8818/33/20240908

3.4. GSEA

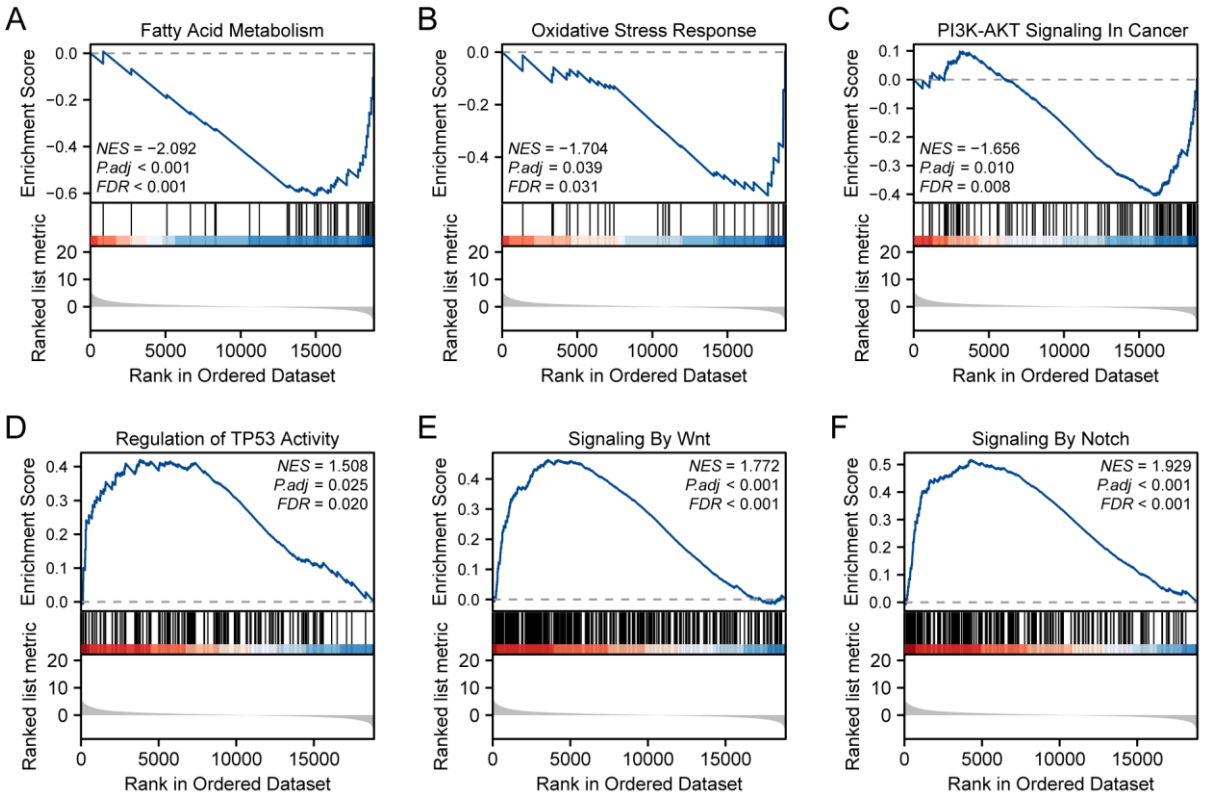

Figure 5. GSEA

In the TCGA-TNBC dataset, there was a notable enrichment of genes in various pathways, including Fatty Acid Metabolism (A), Oxidative Stress Response (B), PI3K-AKT Signaling in Cancer (C), Regulation of TP53 Activity (D), Signaling By Wnt (E), and Signaling By Notch (F). To identify these enrichments, the screening criteria for GSEA enrichment analysis used were a P.adj < 0.05 and FDR value (q.value) < 0.25. TCGA refers to the cancer genome atlas, while TNBC stands for triple negative breast cancer. GSEA represents Gene Set Enrichment Analysis, and MRGs corresponds to mitophagy related genes.

DOI: 10.54254/2753-8818/33/20240908

3.5. GO and KEGG

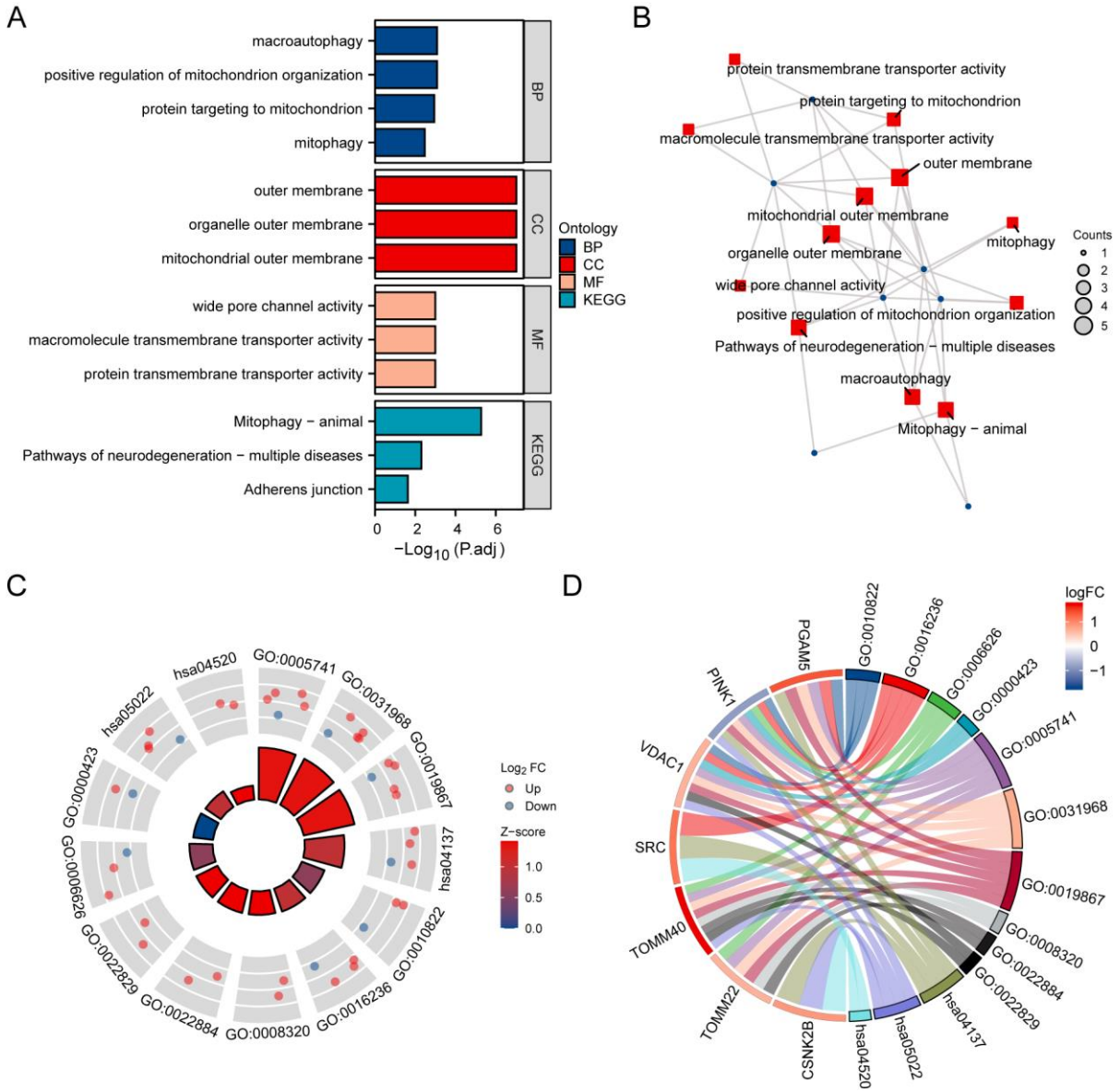

Figure 6. Functional enrichment analysis (GO) and pathway enrichment analysis (KEGG)

(A). The results of MRGs GO and KEGG enrichment analysis are presented in a bar graph. (B). A divergence network diagram demonstrates the outcomes of the GO and KEGG enrichment analysis for MRGs. (C). A circle diagram illustrates the results of GO enrichment analysis of MRGs based on co-logFC. (D). A chord diagram visually represents the findings of GO and KEGG enrichment analysis for MRGs using co-logFC. In the network diagram (B), specific genes are denoted by blue dots, while specific pathways are represented by red squares. The circle diagram (C) showcases up-regulated genes (logFC > 0) as red dots and down-regulated genes (logFC < 0) as blue dots. The screening criteria for GO and KEGG enrichment entries were set as P.adj < 0.05 and FDR value (q.value) < 0.2. GO stands for Gene Ontology, and KEGG refers to Kyoto Encyclopedia of Genes and Genomes. MRGs denotes mitophagy related genes.

DOI: 10.54254/2753-8818/33/20240908

3.6. *Consistent clustering to construct disease subtypes of TNBC*

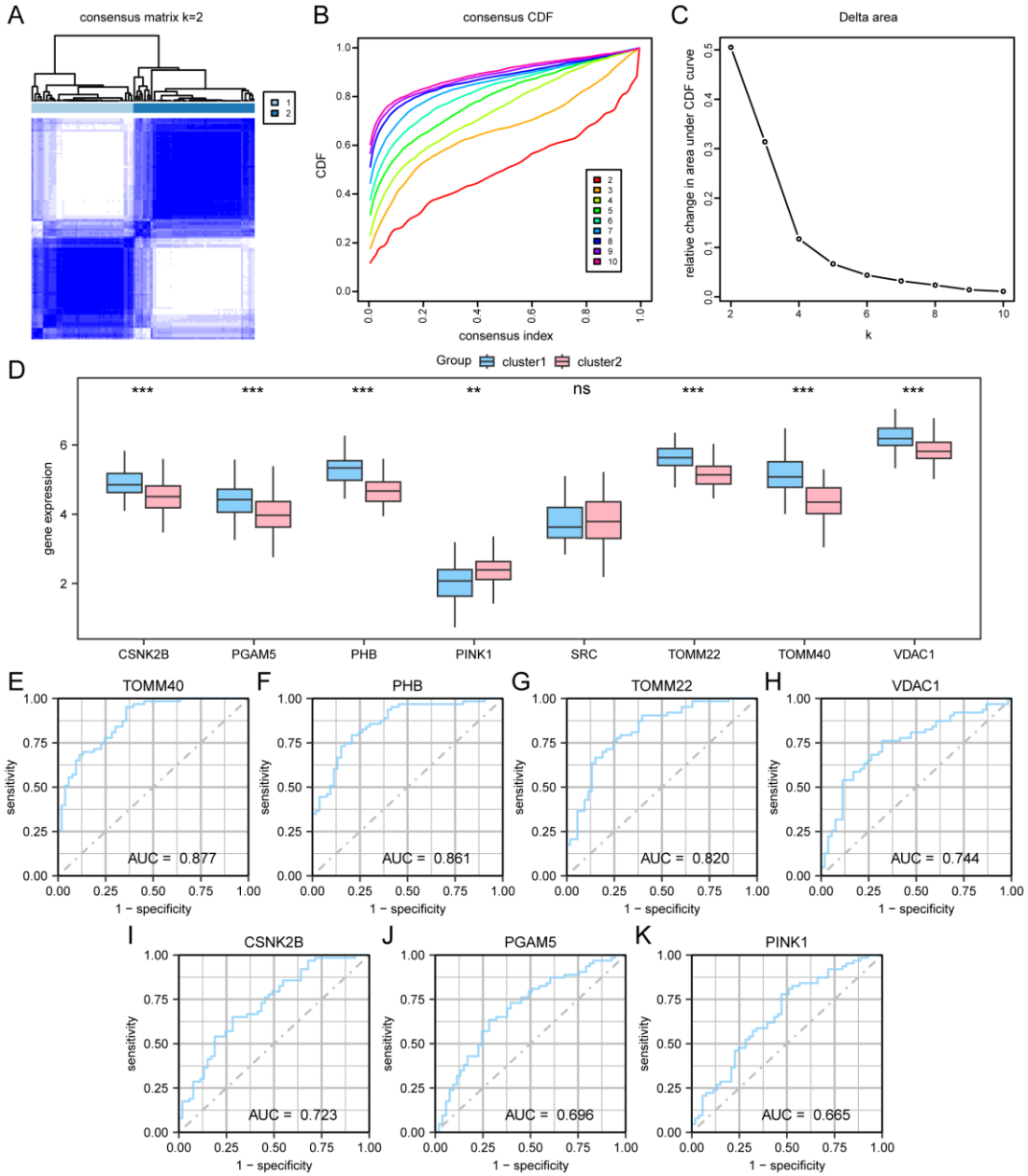

Figure 7. Utilizing Consistency Clustering to Establish Disease Subtypes of Triple-Negative Breast Cancer (TNBC)

(A). Representation of consistent clustering (K=2) outcomes for TNBC diseases. (B-C). The cumulative distribution function (CDF) plot illustrating the consistency clustering results for varying TNBC disease subtypes (B), along with the Delta plot showcasing the area under the CDF curve (C). D. Group comparison plot displaying the expression of mitophagy related genes (MRGs) in the TCGA-TNBC dataset, categorized based on two TNBC disease subtypes (cluster1 and cluster2).(E-J). The ROC curve plot demonstrating the results in the TCGA-TNBC dataset when organized according to disease subtype groups (cluster1 and cluster2). Various group comparison plots are included. Symbols: “ns” represents $P \geq 0.05$, indicating no statistically significant difference; “**” represents $P < 0.01$, denoting high statistical significance; “***” represents $P < 0.001$, signifying very high statistical significance.

DOI: 10.54254/2753-8818/33/20240908

3.7. GSVA

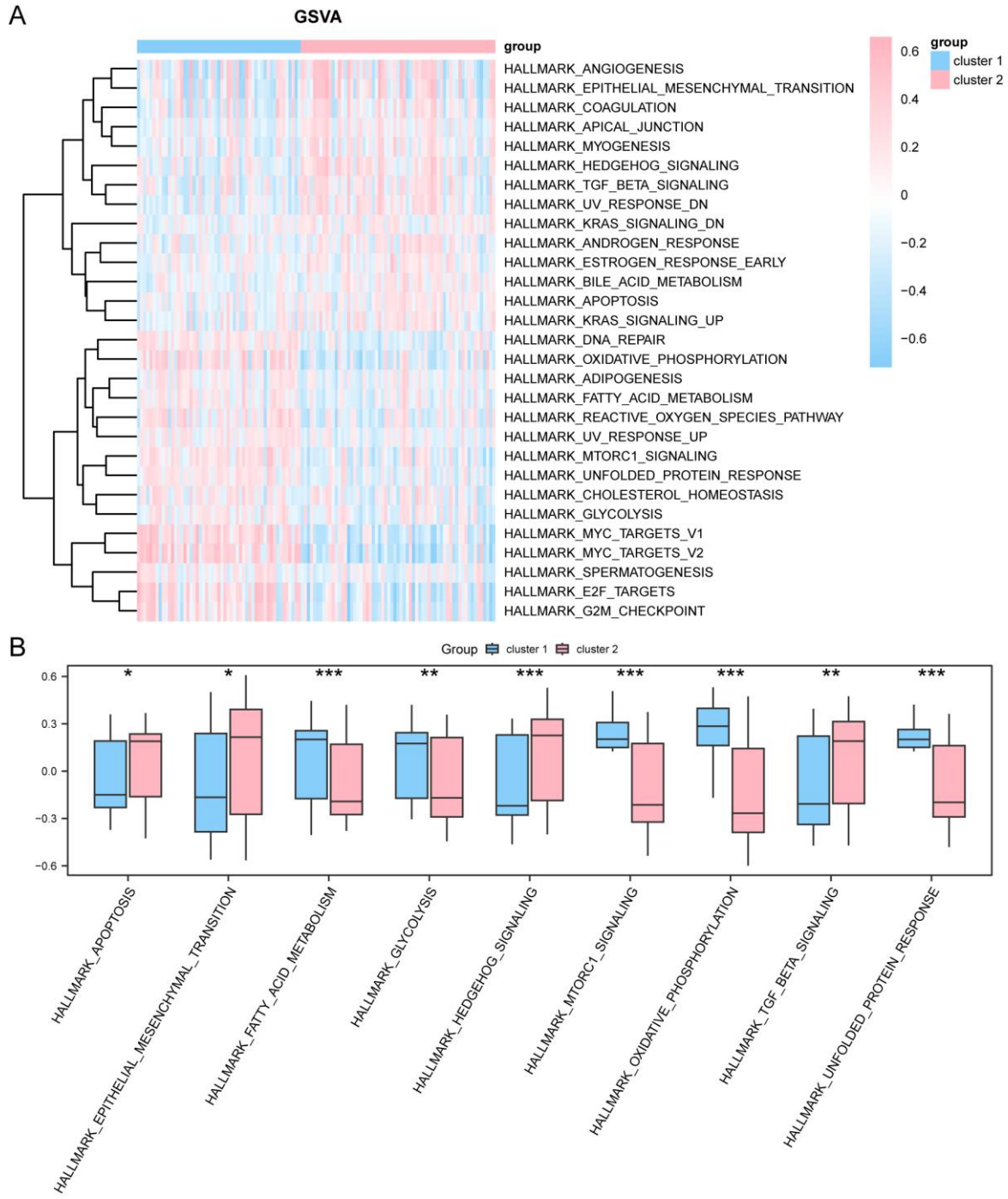

Figure 8. The GSVA of the MRGs

(A). Presentation of gene set functional scores in GSVA of the TCGA-TNBC dataset using a heatmap. (B). Comparative grouping plot illustrating enriched pathways and expression differences between cluster1 and cluster2 in GSVA of the TCGA-TNBC dataset. * denotes statistical significance at $P < 0.05$, ** denotes high statistical significance at $P < 0.01$, and *** denotes highly statistical significance at $P < 0.001$. MRGs refer to mitophagy related genes. TCGA represents the cancer genome and TNBC stands for triple negative breast cancer. GSVA stands for Gene Set Variation Analysis.

DOI: 10.54254/2753-8818/33/20240908

3.8. Constructing MRGs scores

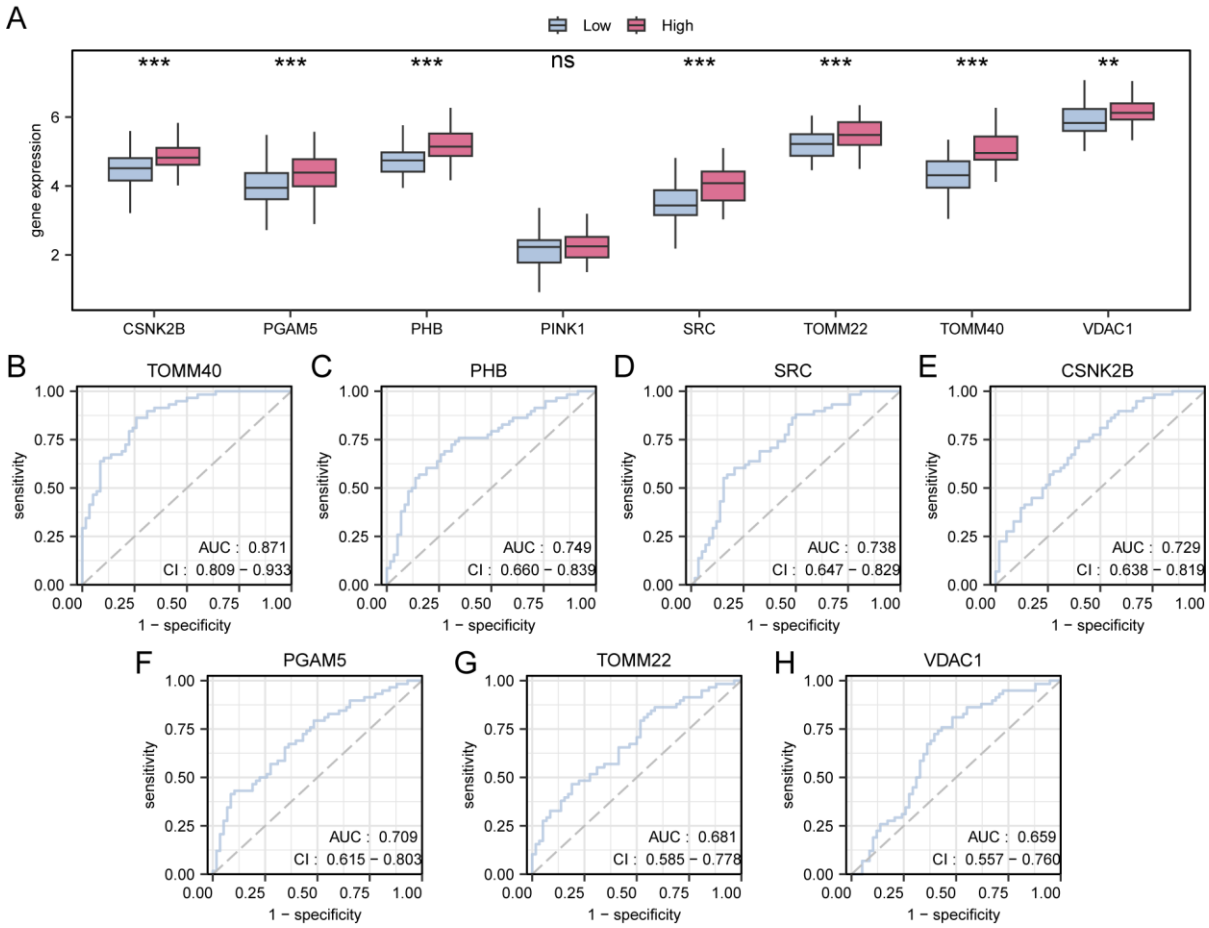

Figure 9. Constructing MRGs scores and Immunotherapy analysis.

A comparison of subgroup MRGs presentations using graphical methods between groups with high and low scores of MRGs in the TCGA-TNBC dataset, denoted as Graph A to Graph H. The ROC curve results for MRGs presentation based on groups with high and low scores of MRGs in the TCGA-TNBC dataset are summarized as B to H. The symbols "ns" represent $P \geq 0.05$ indicating no statistical significance, while symbols "**" and "***" represent $P < 0.01$ and $P < 0.001$, respectively, signifying high statistical significance. The prediction quality improves as the AUC value approaches 1. An AUC range of 0.5 to 0.7 suggests low accuracy, whereas an AUC range of 0.7 to 0.9 indicates TCGA-TNBC (the cancer genome atlas) with MRGs (mitophagy related genes) related. ROC (receiver operating characteristic) curves are utilized, and the AUC (area under the curve) is calculated along with the CI (confidence interval).

DOI: 10.54254/2753-8818/33/20240908

3.9. *CIBERSORTx immunoinfiltration analysis*

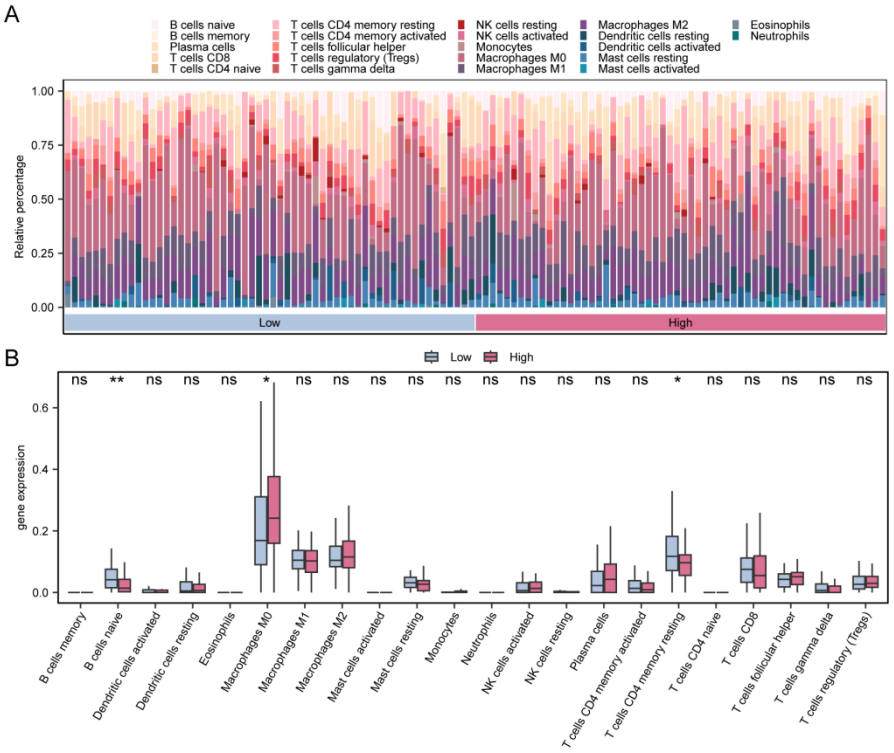

Figure 10. Illustrates an analysis of immune infiltration in the TCGA-TNBC dataset, comparing groups with high and low scores of MRGs (mitophagy related genes).

A stacked bar graph is used to present the results of CIBERSORTx immunoinfiltration analysis, showcasing the differences between the two groups. Moreover, a subgroup comparison graph further demonstrates the variations in immune infiltration. The comparison of CIBERSORTx immunoinfiltration analysis results leaves us with two crucial observations - first, there is no statistically significant difference in groups with high and low MRGs scores, denoted by the symbol "ns" ($P \geq 0.05$). However, when P values are below 0.05, indicated by the symbol "*", the difference becomes statistically significant. Furthermore, if P values drop below 0.001, depicted as "**", the difference is considered highly statistically significant. It is worth mentioning that TCGA refers to the Cancer Genome Atlas, while TNBC stands for triple-negative breast cancer.

DOI: 10.54254/2753-8818/33/20240908

3.10. Cox model and PINK1 expression in TNBC

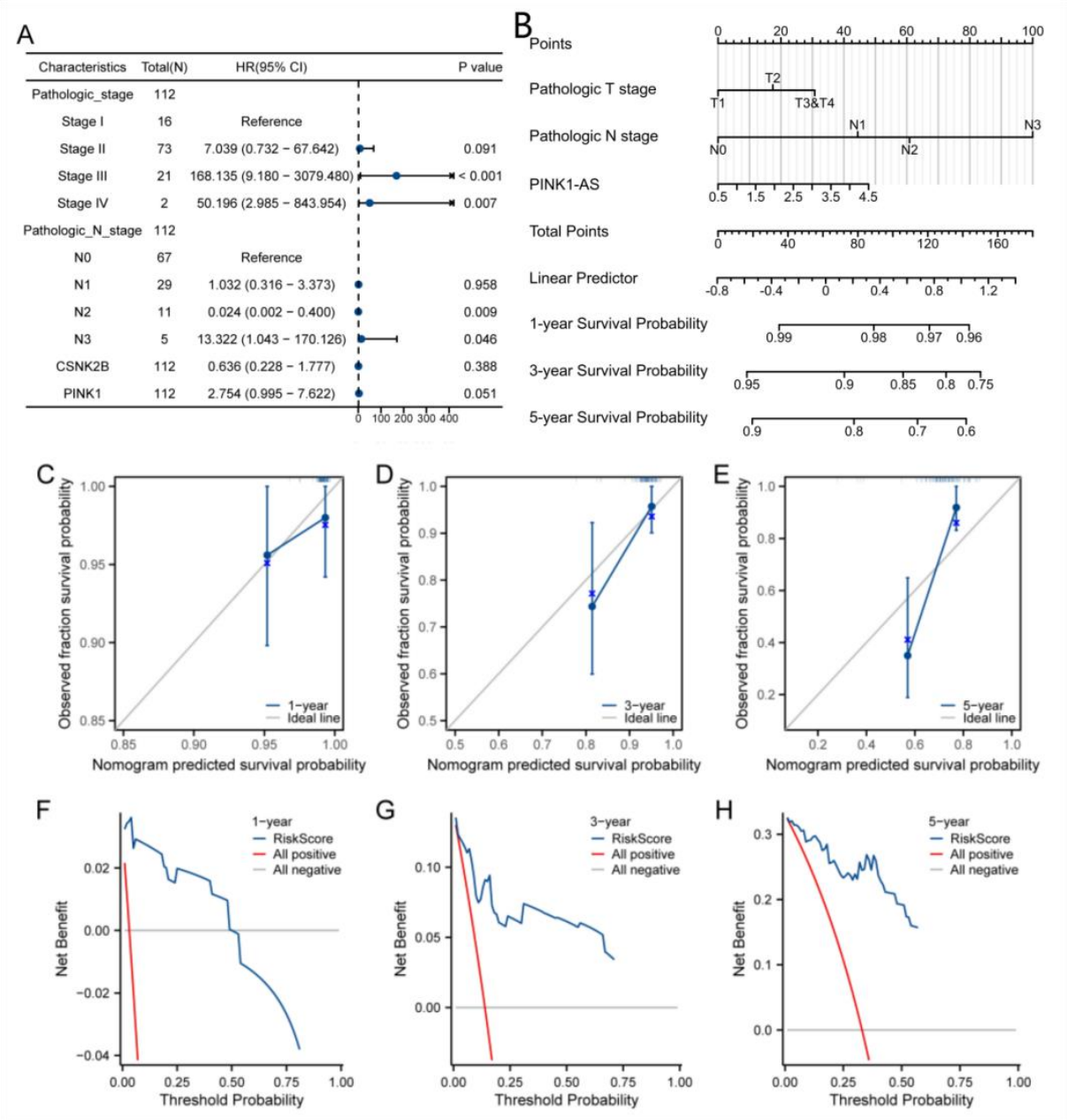

Figure 11. Prognostic clinical correlation analysis

The plot of the forest (A) and the line plot for the column (B) was conducted to perform multifactor Cox regression analysis.1-year (C), 3-year (D), and 5-year (E) calibration curve plots were generated to analyze the nomogram for multifactor Cox regression model.(F), (G), and (H) represent the 1-year, 3-year, and 5-year DCA plots, respectively, for the multifactor Cox regression models.

DOI: 10.54254/2753-8818/33/20240908

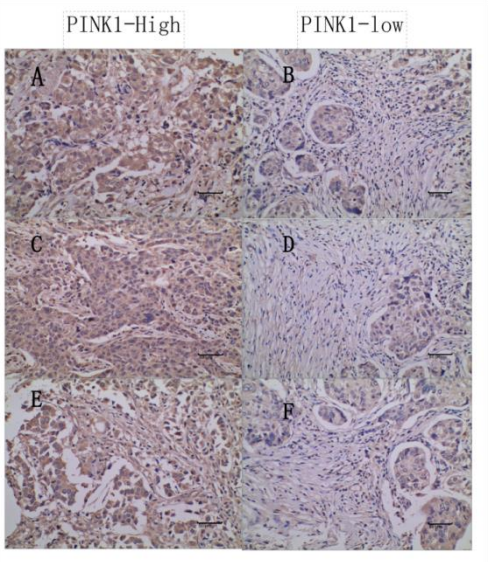

Figure 12. PINK1 expression in TNBC

The immunohistochemical results ACE is PINK1 low expression group, BDF is PINK1 high expression group.

4. Discussion

In this research, we examined the expression of 11 genes that were significantly up-regulated in TNBC cancer (CSNK2B, HIF1A, MTERF3, PGAM5, PHB, SRC, TIGAR, TOMM5, TOMM22, TOMM40, and VDAC1), and also analyzed 3 genes that were significantly down-regulated (PINK1, PRKN, and VPS13C). After conducting thorough validation of the dataset, we identified 8 central genes (CSNK2B, PGAM5, PHB, PINK1, TOMM22, TOMM40, SRC, VDAC1) that may potentially contribute to the progression of TNBC cancer. The correlation analysis demonstrated a positive relationship among these eight genes, indirectly suggesting that they may have similar biological functions or mechanisms of action. Importantly, we confirmed these findings through KEGG and GO correlation enrichment analysis, which revealed that the biological functions of these genes primarily involve the regulation of mitochondria and mitochondrial autophagy.

In our investigation, it was observed that a significant connection existed between the overexpression of PINK1 and an unfavorable prognosis in TNBC cancer. PINK1 exhibited a high incidence of mutation in tumor development, predominantly in the form of missense mutation. Interestingly, these somatic missense mutations actually facilitated the creation of novel antigens, thereby offering potential targets for immunotherapy. Several studies, including the Human Protein Atlas, have noted varying levels of PINK1 expression in different tumor types, with low expression often associated with an unfavorable prognosis. Interestingly, the interaction between Parkin and HIF-1α was found to promote the latter's ubiquitination at position K477, subsequently leading to its degradation. Consequently, this process effectively hindered breast cancer metastasis [13]. However, in our study, we identified an association between PINK1 and adverse prognosis in TNBC cancer. This finding may be attributed to the crucial role played by PINK1 in mediating chemoresistance. Consequently, the elevated expression of PINK1 in TNBC was found to be linked to poor prognosis and adversely impacted the overall survival of TNBC patients. Furthermore, we validated the accuracy of PINK1 and its model using immunohistochemistry and PCR techniques. The findings from these experiments indirectly supported the notion that PINK1 holds potential as a drug target and prognostic biomarker for TNBC.

Recent research has revealed that TP53 is responsible for encoding the essential transcription factor p53. This transcription factor plays a crucial role in responding to various stressors like radiation and chemotherapy. In response to these stressors, it triggers the transcription of genes involved in critical

DOI: 10.54254/2753-8818/33/20240908

processes such as cell cycle arrest, apoptosis, metabolism, DNA repair, and cellular senescence [14]. It is worth noting that TP53 mutations are more prevalent in TNBC cancer with a staggering rate of 80% compared to non-TNBC cancer [15]. This finding aligns perfectly with our analysis using Gene Set Enrichment Analysis (GSEA), which showcases the highly diverse nature of TNBC cancers. In TNBC, TP53 mutations predominantly occur within the DNA-binding domain, rendering mutant p53 proteins transcriptionally inactive upon stress stimuli [16]. Additionally, studies have highlighted the impact of dysregulated Wnt/β-cyclin signaling on TNBC patients. Those with disturbed Wnt/β-cyclin signaling are at a higher risk of developing secondary metastases to the lungs and brain. Moreover, it has been demonstrated that nuclear accumulation of β-cyclins significantly promotes various detrimental effects such as cell migration, colony formation, stem-like features, and chemoresistance. Notably, in vitro experiments using TNBC cells and a mouse cancer model have provided evidence suggesting the significant role of Wnt signaling in driving TNBC tumorigenesis [17].

By detecting different molecular subtypes of TNBC, it is possible to personalize treatment strategies. We classified TNBC cancer into two distinct molecular subtypes using MRGs, and accurately categorized the seven different genes, excluding SRC, into these two subtypes. To evaluate the variation among the genes of these subtypes, we performed a GSVA analysis on the entire dataset of TNBC genes. Our results demonstrated that cluster2 exhibited significant enrichment in apoptosis, inter-epithelial transformation, and the hedgehog pathway, while cluster1 showed notable enrichment in the TGF pathway. Moreover, cluster1 displayed significant enrichment in fatty acid metabolism, glycolysis, and the mTORC1 signaling pathway. The hedgehog signaling pathway plays a critical role in the development of tumors, progression, drug resistance, and metastasis [18]. In relation to the TGF-β-mediated inter-epithelial transformation, which is recognized to promote invasion and metastasis in pancreatic cancer, our suggestion is that cluster2 may exhibit a higher degree of metastasis and invasiveness. Additionally, the mTORC1 signaling pathway is essential for cellular growth and value-added processes. The outcomes of this investigation indicate that abnormal activation of the mTORC1 signaling pathway is associated with lymphatic metastasis and a poorer prognosis, which also suggests that cluster1 is linked to an unfavorable prognosis. The molecular subtyping based on MRGs aims to further investigate its significance in the context of TNBC [19].

In our research, we utilized the ssGSEA algorithm to acquire the MRGs score for each sample within the cancer groups. This score effectively represents the expression level of MRGs in the respective samples. Subsequently, we partitioned these samples into two categories based on their scores, namely high and low scoring groups. The primary purpose of this categorization was to assess any disparities in immune infiltration between these two groups. Notably, the analysis revealed notable distinctions in B cell naive, Macrophages M0, and T cells CD4 memory resting between the high and low scoring groups. The outcomes of our investigation demonstrated that the high scoring group exhibited a greater degree of Macrophages M0 infiltration in comparison to the low scoring group. This observation provides evidence supporting the involvement of mitochondrial autophagy in the promotion of tumor growth through macrophage infiltration. Previously, Laurence Buisseret and colleagues have conducted research establishing that tumor-infiltrating lymphocytes-B hold the ability to present antigens to tumor-infiltrating lymphocytes-CD4+, thereby leading to enhanced anti-tumor immune responses. Furthermore, tumor-infiltrating CD4+ and tumor-infiltrating lymphocyte-B have demonstrated a positive correlation with a favorable prognosis in TNBC cancer [20]. Based on the findings of our study, it can be inferred that the mechanism involving mitochondrial autophagy potentially facilitates tumor progression by reducing the infiltration of tumor-infiltrating lymphocytes-CD4+ and tumor-infiltrating lymphocytes-B.

5. Conclusion

To summarize, this investigation is the foremost effort to establish and validate the model of mitochondrial autophagy, using the mitochondrial autophagy-associated gene PINK1 as an independent prognostic determinant for TNBC patients. Our findings indicate that the model of mitochondrial autophagy plays a vital role in the advancement of tumors and exhibits a significant association with the

DOI: 10.54254/2753-8818/33/20240908

prognosis of TNBC. These outcomes may empower clinicians to determine tailored treatments for TNBC patients based on the varying expressions of genes associated with mitochondrial autophagy. Additionally, these findings provide essential evidence for further exploring the significance of mitochondrial autophagy in TNBC and furnish a theoretical foundation for precise, targeted therapies in the future.

References

[1] Dixon-Douglas J, Loi S. 2023. Immunotherapy in Early-Stage Triple-Negative Breast Cancer: Where Are We Now and Where Are We Headed? Curr Treat Options Oncol 24:1004-20
[2] Guven DC, Yildirim HC, Kus F, Erul E, Kertmen N, et al. 2023. Optimal adjuvant treatment strategies for TNBC patients with residual disease after neoadjuvant treatment. Expert Rev Anticancer Ther 23:1049-59
[3] Waks AG, Winer EP. 2019. Breast Cancer Treatment: A Review. JAMA 321:288-300
[4] Ferro F, Servais S, Besson P, Roger S, Dumas JF, Brisson L. 2020. Autophagy and mitophagy in cancer metabolic remodelling. Semin Cell Dev Biol 98:129-38
[5] Bravo-San Pedro JM, Kroemer G, Galluzzi L. 2017. Autophagy and Mitophagy in Cardiovascular Disease. Circ Res 120:1812-24
[6] Lizama BN, Chu CT. 2021. Neuronal autophagy and mitophagy in Parkinson's disease. Mol Aspects Med 82:100972
[7] White E, Mehnert JM, Chan CS. 2015. Autophagy, Metabolism, and Cancer. Clin Cancer Res 21:5037-46
[8] Karantza-Wadsworth V, Patel S, Kravchuk O, Chen G, Mathew R, et al. 2007. Autophagy mitigates metabolic stress and genome damage in mammary tumorigenesis. Genes Dev 21:1621-35
[9] Eskelinen EL. 2011. The dual role of autophagy in cancer. Curr Opin Pharmacol 11:294-300
[10] Mah LY, Ryan KM. 2012. Autophagy and cancer. Cold Spring Harb Perspect Biol 4:a008821
[11] Choi AM, Ryter SW, Levine B. 2013. Autophagy in human health and disease. N Engl J Med 368:651-62
[12] Sridhar S, Botbol Y, Macian F, Cuervo AM. 2012. Autophagy and disease: always two sides to a problem. J Pathol 226:255-73
[13] Liu Y, Jiang Y, Wang N, Jin Q, Ji F, et al. 2017. Invalidation of mitophagy by FBP1-mediated repression promotes apoptosis in breast cancer. Tumour Biol 39:1010428317708779
[14] Kastenhuber ER, Lowe SW. 2017. Putting p53 in Context. Cell 170:1062-78
[15] Cancer Genome Atlas N. 2012. Comprehensive molecular portraits of human breast tumours. Nature 490:61-70
[16] Soussi T. 2011. TP53 mutations in human cancer: database reassessment and prospects for the next decade. Adv Cancer Res 110:107-39
[17] Dey N, Barwick BG, Moreno CS, Ordanic-Kodani M, Chen Z, et al. 2013. Wnt signaling in triple negative breast cancer is associated with metastasis. BMC Cancer 13:537
[18] Wang J, Cui B, Li X, Zhao X, Huang T, Ding X. 2023. The emerging roles of Hedgehog signaling in tumor immune microenvironment. Front Oncol 13:1171418
[19] Pan Z, Zhang H, Dokudovskaya S. 2023. The Role of mTORC1 Pathway and Autophagy in Resistance to Platinum-Based Chemotherapeutics. Int J Mol Sci 24
[20] Garaud S, Buisseret L, Solinas C, Gu-Trantien C, de Wind A, et al. 2019. Tumor infiltrating B-cells signal functional humoral immune responses in breast cancer. JCI Insight 5