Breast Cancer: Targets and Therapy

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

A Prognostic Risk Signature of Two Autophagy-Related Genes for Predicting Triple-Negative Breast Cancer Outcomes

Bing Yu[1,*], Zhimei Xing[2,*], Xiaoxuan Tian[2], Rui Feng[1]

[1]Department of Breast Surgery, Tianjin Central Hospital of Obstetrics and Gynecology, Tianjin, 300100, People's Republic of China; [2]State Key Laboratory of Component-Based Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, 300193, People's Republic of China

*These authors contributed equally to this work

Correspondence: Rui Feng, Department of Breast Surgery, Tianjin Central Hospital of Obstetrics and Gynecology, No. 156 of Nankai SAN Road, Nankai District, Tianjin, 300100, People's Republic of China, Tel +86 022-58287167, Email dr.fr@126.com

Background: Triple-negative breast cancer (TNBC) is recognized as the most aggressive molecular subtype of breast cancer. Recent studies have highlighted the complex role of autophagy in the pathogenesis of TNBC.

Methods: In this study, we evaluated 18,330 genes, including 1111 autophagy-related genes, (ARGs), across 579 TNBC samples from online databases. Differentially expressed ARGs in TNBC were identified using high-throughput RNA-seq data from the Cancer Genome Atlas (TCGA). Prognostic factors were examined through Cox regression and multivariate Cox analyses, with predictive efficacy assessed using receiver operating characteristic (ROC) curves. A nomogram integrating the risk signature with clinicopathological factors, such as TNM stage, was developed. Immunohistochemical analysis of clinical samples was also conducted.

Results: EIF4EBP1 and NPAS3 were significantly correlated with prognostic outcomes in patients with TNBC. Multivariate Cox regression analysis demonstrated that the expression levels of these two genes were accurate predictors of disease progression in TNBC samples from TCGA and the GSE31519 dataset. The efficacy of this predictive model was validated using ROC curve analysis and calibration plots, confirming its ability to accurately estimate the 1-, 2-, and 3-year survival rates for individuals with TNBC. Additionally, EIF4EBP1 and NPAS3 expression influenced drug sensitivity in TNBC cell lines, with notably lower NPAS3 expression in TNBC tissues, particularly in Stage III cases. This study is the first to report NPAS3 expression in patients with TNBC.

Conclusion: The autophagy-related genes EIF4EBP1 and NPAS3 may serve as independent prognostic factors for individuals with TNBC.

Keywords: autophagy, nomogram, prognosis, risk signature, triple-negative breast cancer

Introduction

Breast cancer significantly contributes to cancer-related mortality globally, particularly among women.[1] Among the diverse subtypes of breast cancer, triple negative breast cancer (TNBC) is characterized by the lack of estrogen, progesterone, and HER2 receptors. TNBC is associated with early onset, high invasiveness, poor prognosis and elevated rates of metastasis, recurrence, and mortality.[2,3] Autophagy, as a process of self-digestion, is closely related to multiple biological processes in cancer. As summarized by Ashrafizadeh et al,[4] autophagy exhibits a dualistic role in TNBC, potentially either reducing tumor cell viability or serving as a cytoprotective mechanism. Depending on the environment, autophagy induction may stimulate cancer cell growth ablation and induce autophagic cell death. Conversely, autophagy may attenuate drug sensitivity and facilitate apoptosis.

Although extensive literature highlights the clinical relevance of autophagy-related genes (ARGs) in breast cancer, identifying crucial ARGs pertinent to TNBC prognosis remains elusive. Furthermore, prognostic tools integrating these genes with clinicopathological risk factors are lacking.[5–10] Due to the complex role of autophagy in tumor biology and

Received: 5 June 2024
Accepted: 17 August 2024
Published: 2 September 2024

the resistance of TNBC to conventional treatments, ARGs have garnered significant attention in TNBC research. Autophagy may both promote and inhibit tumor progression in TNBC, offering new avenues for treatment and prognostic assessment for this challenging cancer. The absence of common hormone and HER2 receptors in TNBC renders traditional targeted therapy strategies ineffective, positioning ARGs as a promising research field. Moreover, the specific expression and regulatory mechanisms of ARGs provide opportunities for developing personalized healthcare strategies targeting TNBC. Therefore, in-depth research on ARGs can help reveal the unique molecular characteristics of TNBC, promoting the development of new treatment strategies and prognostic tools.

In this study, gene expression datasets of TNBC breast cancer sourced from the Cancer Genome Atlas (TCGA) were analyzed to identify ARGs with prognostic significance. Two signature genes were ultimately identified and verified within the GEO GSE31519 dataset. The risk score, derived from their expression levels, was used to categorize patients with TNBC into high-risk and low-risk groups. Furthermore, utilizing the prognostic model, the area under the receiver operating characteristic (ROC) curve (AUC) values for predicting 1-year overall survival, improved from 0.936 (based solely on TNM stage data) to 0.975 (incorporating both TNM stage data plus ARGs-related risk scores).

Materials and Methods

Data Collection

The catalog of human ARGs was compiled from various sources, including the Human Autophagy Database (http://autophagy.lu/clustering/index.html), the AUTOPHAGY DATABASE (http://autophagy.info/), and the Human Autophagy Modulators Database (http://hamdb.scbdd.com/). After removing redundant information, a total of 1180 autophagy-related genes were retained (see Supplementary Table 1). Transcriptome data and clinicopathological information for patients were sourced from UCSC Xena (https://xenabrowser.net/). An initial cohort of 1247 patients diagnosed with breast cancer, for whom expression data for 20,530 genes were available, was compiled. (see Supplementary Table 2). Subsequently, 112 TNBC samples and 10 paired normal samples were selected according to the clinical matrix. Genes with zero-expression in at least 80% of the samples were excluded to eliminate the potential interference from low-expression genes. Consequently, a total of 18,330 genes, including 1111 ARGs were retained for further analysis. To validate the findings, the gene expression dataset of 579 TNBC samples was collected from the Gene Expression Omnibus (GEO) database (GSE31519) in the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/geo/).

Identification and Enrichment Analysis of Differentially Expressed Autophagy-Related Genes (DE-ATGs)

RNA-Seq data sourced from TCGA were normalized, and differential expression analysis was conducted between TNBC tumor samples and their corresponding normal counterparts using the LIMMA package in R software.[11] The Benjamini and Hochberg method was subsequently employed to adjust *P*-values for multiple testing. Differentially expressed genes were screened based on the following criteria: |log2 (Fold change) |>1 and an adjusted *P* value < 0.05 (tumor tissues vs normal tissues). Differentially expressed ARGs were highlighted among these genes. Subsequent enrichment analyses, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses, were conducted using the ClusterProfiler package.[12] Terms with an adjusted *P* value < 0.2 and a nominal *P* value < 0.05 were considered statistically significant.

Construction of an ARG-Related Prognostic Model

To assess which differentially expressed ARGs could function as prognostic indicators for patients with TNBC, a univariate Cox regression model was used to identify significant predictors, with a significance threshold of *P*-values < 0.05. Subsequently, the multivariate Cox analysis was conducted to further screen these predictors and generate the risk score (R score). The R score was calculated as a linear combination of the regression model coefficients (beta) multiplied by the mRNA expression levels of the corresponding genes. The correlation among prognostic factors within the TCGA breast cancer samples was calculated through TIMER2.[13] To evaluate this prognostic model, patients with TNBC were categorized into low-risk and high-risk groups based on the mean R score. Survival differences between

the two groups were assessed using Kaplan–Meier analysis and compared using log-rank statistical methods. ROC curves were generated to assess the predictive accuracy of the model for 1-, 2-, 3-, and 5-year outcomes. The BRCA mutant TNBC samples were screened out by SNP genotyping. Wilcoxon rank sum test was used to examine significant differences between various age groups (< 60 and ≥ 60 years old), different lymph node metastatic statuses, pathologic sizes, pathologic stages, histological types, and BRCA genotypes. Spearman correlation analysis was conducted to examine the correlation between the R score and tumor node metastasis (TNM) stages. Differential expression analysis of ARG prognostic predictors between low-risk and high-risk groups was performed using the students' *t*-test function "*t*-test" in R. The Human Protein Atlas (HPA) was used for validating the different expressions of selected ARGs at the protein level.[14]

Validation of the Autophagy-Associated Score in GEO Cohorts

Among the 579 TNBC samples from the GSE31519 dataset, a subset of 383 samples with available survival time data was selected for further analysis. The procedures for calculating the R score, categorizing patients into low-risk and high-risk groups, assessing survival differences, and evaluating model prediction accuracy were consistent with the methodology described in Construction of an ARG-Related Prognostic Model.

Construction of the Nomogram

Multivariate Cox regression analysis was utilized to investigate the effects of age, TNM stage, BRCA1/2 mutant, and the R score on the survival of patients with TNBC. Based on this analysis, the TNM stage and R score were selected for inclusion in the nomogram, which was constructed using the "rms" and "survival" packages within the R software environment. The performance of the nomogram was evaluated using ROC curve analysis to determine the AUC values. Calibration curves were generated to evaluate the consistency between actual and predicted survival outcomes. The parameters were set as m = 30 and B = 100 for 1-year survival rate prediction calibration, while m = 40 was the parameter for 2-year, 3-year, and 5-year calibration. In cases where TNM data were unavailable in the GEO dataset, histological grade was utilized (grades 1 and 2 were set as level 1, and grade 3 as level 2) in conjunction with the R score, to construct nomograms for predicting 1-year, 2-year, 3-year, and 5-year survival rates.

Analysis of Immune Cell Composition from Gene Expression Data

CIBERSORT was used to estimate the relative proportions of 9 aggregated immune cell populations and 22 subpopulations within TNBCs stratified by high-or low-risk scores.[15] The relative fractions of each immune cell population were compared. The LM22 signature immune panel from CIBERSORT was used to analyze the expression matrix. Comparisons of relative fractions were conducted using the Wilcoxon rank-sum test. Furthermore, for TCGA breast cancer samples, the relationship between selected autophagy-related genes (ARGs) and various cell types was examined using TIMER2 and GEPIA2021.[16]

Correlation Between Selected ARGs and Drug Sensitivity

Drug sensitivity data for four TNBC cell lines (BT-483, MDA-MB-231, BT-20, and MDA-MB-468) were obtained from the Genomics of Drug Sensitivity in Cancer (version GDSC1). The RMA-normalized basal expression profiles for all the cell lines were downloaded from Sanger & CCLE.

Immunohistochemistry Staining of Human TNBC Tissue Samples

To determine the expression profiles of key ARGs, tissue microarrays designated for immunohistochemical (IHC) staining were analyzed. These arrays included 54 samples, including 49 from TNBC tissues and 5 from adjacent normal tissues. The samples were collected from 37 patients with stages I–II (A/B) and 12 patients with stages III (A/B/C). The tissue microarrays were obtained from Bioaitech (cat.no. F551101, Xi'an, China) (see Supplementary Table 3).

The tissue specimens were initially fixed in 10% formalin, then embedded in paraffin and sectioned. Optimal tissue sections were selected for degreasing and immunohistochemistry staining. The antibody used in this study was NPAS3 (Bioss, bs-11905R). Staining results were evaluated by randomly observing 5 to 10 fields of view and averaging the

findings. The expression grade was calculated as the product of the intensity score (ranging from 0 to 4) and the frequency of positive cells (also graded from 0 to 4).

Results

Identification of Differentially Expressed ARGs

Upon analyzing 112 TNBC samples and 10 paired normal controls, 1411 up-regulated and 2733 down-regulated genes were identified. The list of differentially expressed genes is detailed in Supplementary Table 4. The volcano plot and heatmap depicting the expression patterns of the 4144 differentially expressed genes are presented in Figure 1A and B, respectively. As demonstrated in Figure 2A and detailed in Supplementary Table 5, a total of 202 ARGs exhibited significant differential expression between breast cancer tissues and adjacent normal tissues. Among these, 72 ARGs were up-regulated, while 130 were down-regulated (see Figure 2B). The boxplot illustrating the expression levels of the 202 ARGs is presented in Supplementary Figure 1.

Functional Enrichment Analysis of Differentially Expressed ARGs

Enrichment analysis revealed that the 202 differentially expressed ARGs were associated with 1304 Gene Ontology (GO) terms (see Supplementary Table 6), and clustered into three ontologies: biological process (1241 GO terms), cellular component (21 GO terms), and molecular function (42 GO terms). The top 20 GO terms within each ontology, along with their associated genes, are illustrated in Figures 3A–C. Figure 3A demonstrates that the biological processes involving the differentially expressed ARGs are primarily related to autophagy and autophagic mechanisms. Within the cellular components, the ARGs were mainly enriched in the outer membrane, organelle outer membrane, and mitochondrial outer membrane (Figure 3B). Furthermore, the molecular function analysis revealed a significant enrichment of ARGs related to protein serine/threonine kinase activity, as shown in Figure 3C. Figure 3D shows the result of the KEGG enrichment analysis. The 202 ARGs were enriched into 98 pathways, mainly the PI3K-Akt signaling pathway (see Supplementary Table 7).

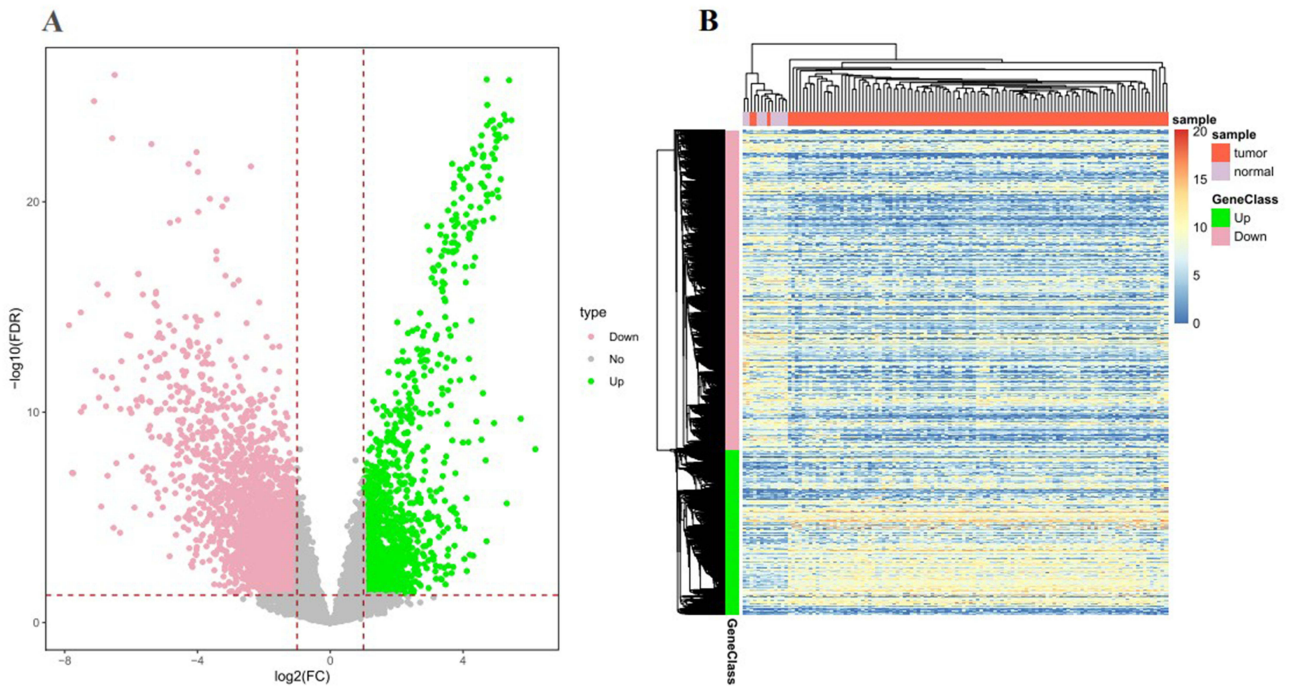

Figure 1 Identification of 4144 differentially expressed genes, featuring (**A**) a volcano map and (**B**) a heatmap. These genes were found to exhibit differential expression in breast cancer tissues in comparison to adjacent normal tissues. In the volcano map, green dots represent genes significantly up-regulated in breast cancer tissues, while pink dots denote those significantly down-regulated. genes, and gray dots represent genes with no difference. Gray dots indicate genes showing no discernible difference in expression between the two tissue types.

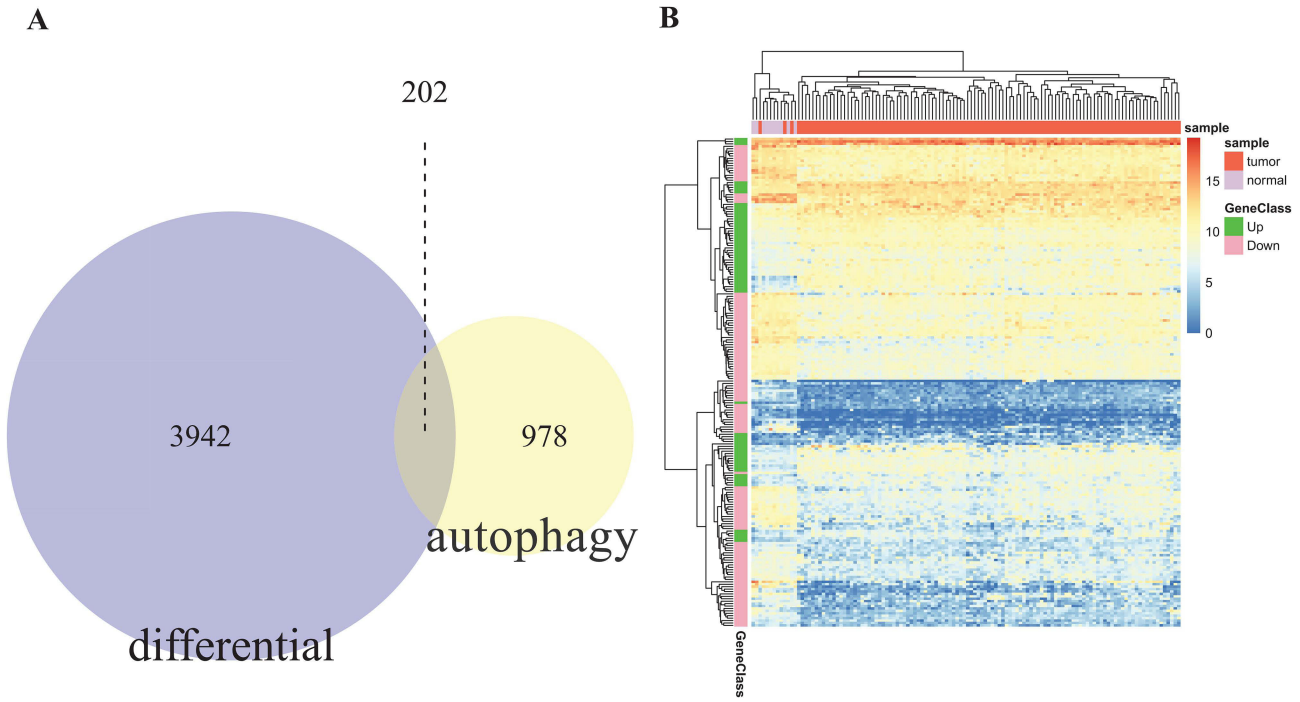

Figure 2 Identification of 202 differentially expressed ARGs, including (**A**) a Venn diagram depicting their distribution. Additionally, (**B**) a heatmap displays the expression patterns of these 202 ARGs in breast cancer tissues compared to adjacent normal tissues. Within the heatmap, green dots represent genes significantly up-regulated in breast cancer tissues, while pink dots signify those significantly down-regulated.

Selection of ARG-Related Prognostic Markers in the TCGA TNBC Dataset

Using the described methodology, univariate Cox regression identified 13 genes with significant associations with the prognosis of patients with TNBC. Multivariate Cox analysis subsequently highlighted two ARGs as prognostic markers: EIF4EBP1 (P = 0.028, HR =1.623) and NPAS3 (P = 0.024, HR = 0.530). Expression data for these genes enabled the stratification of samples into high and low expression groups, with thresholds set at the lower quartile values. Survival differences between these groups were assessed using Kaplan–Meier analysis and compared through log-rank statistical methods. Figure 4 illustrates distinct differences in survival probabilities between the high and low expression groups for both EIF4EBP1 and NPAS3. Specifically, EIF4EBP1 emerged as a risk factor for TNBC prognosis, while NPAS3 exhibited a protective effect. The hazard ratios and coefficients of the genes selected by univariate and multivariate Cox regression are shown in Supplementary Tables 8 and 9, respectively. Additionally, a negative correlation was observed between the expression levels of EIF4EBP1 and NPAS3 in TCGA breast cancer basal-like samples (Spearman correlation coefficient = −0.277, P = 1.04e-04), as depicted in Supplementary Figure 2.

Development of the Prognostic R Score Model Based on EIF4EBP1 and NPAS3 Genes

Utilizing coefficients from multivariate Cox regression analysis for EIF4EBP1 and NPAS3 genes, a prognostic R score model was developed to classify patients with TNBC into distinct risk categories based on their clinical outcomes. The R score was calculated using the formula:

R score = [expression level of EIF4EBP1* (0.485)] + [expression level of NPAS3* (−0.634)]

The R scores for all TCGA samples included in the study were computed and are detailed in Supplementary Table 10. The mean value of the R score of patients with TNBC was taken as the threshold to divide samples into high-or low-risk groups. As illustrated in Figure 5A, the survival rate of patients in the high-risk group was significantly lower compared to those in the low-risk group (P = 0.02). The AUC for 1-, 2-, 3-, and 5-years of survival were calculated as 0.75, 0.789, 0.76, and 0.572, respectively. These results indicate the potential of the prognostic indicators derived from these genes in predicting survival, particularly within the initial three years (see Figure 5B). Furthermore, survival status (see

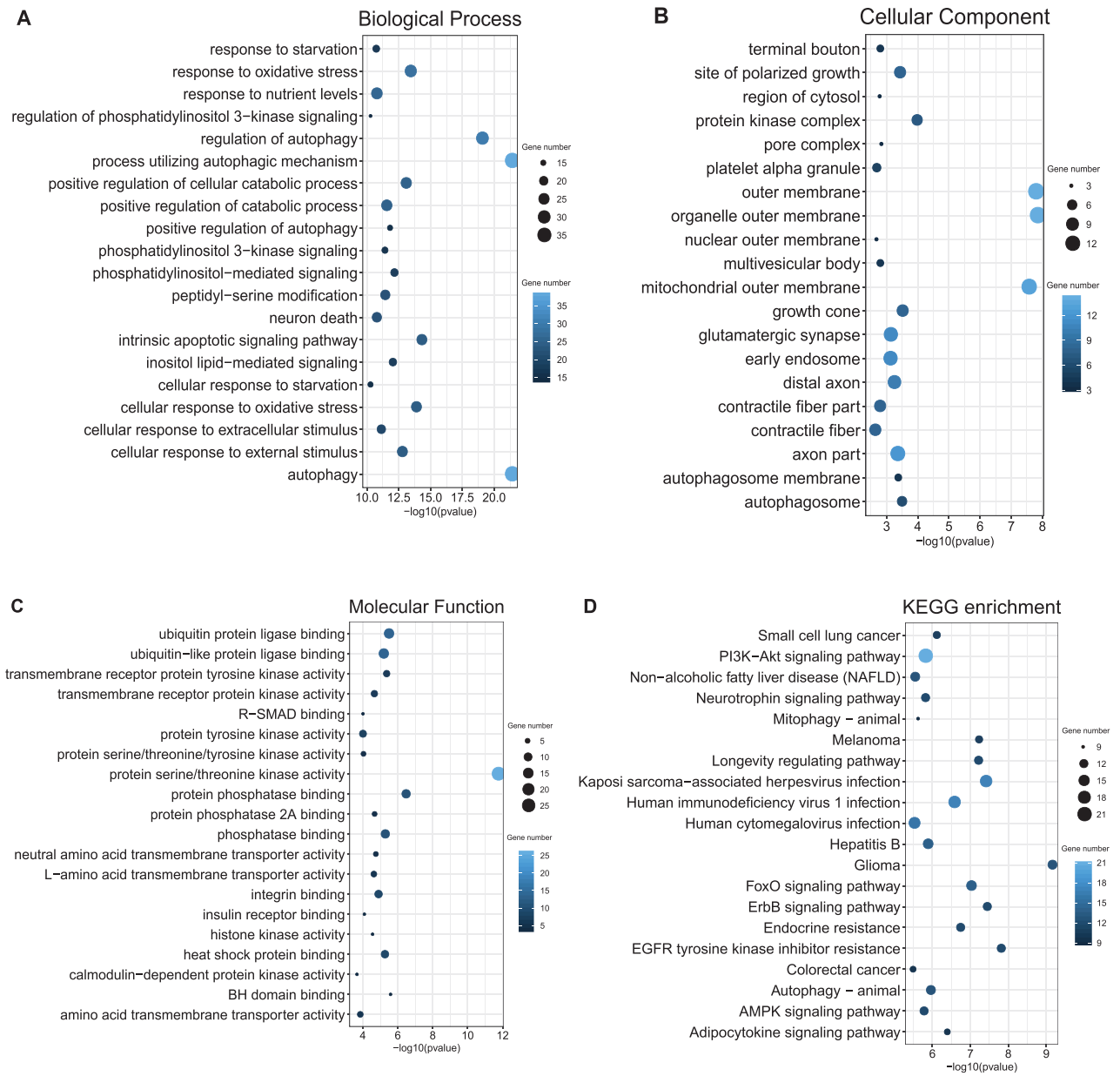

Figure 3 Gene functional enrichment analysis of differentially expressed ARGs, encompassing ontologies annotation across (**A**) biological processes, (**B**) cellular components, (**C**) molecular function; (**D**) KEGG pathway analysis, which elucidates the signaling pathways involved in differential ARGs.

Figure 5C) and the distribution of R scores among different patient groups (refer to Figure 5D) were analyzed. The expression profiles of EIF4EBP1 and NPAS3 (see Figures 5E–F) corroborated the findings depicted in Figure 4. Specifically, patients in the high-risk group showed increased expression of EIF4EBP1, whereas those in the low-risk group exhibited higher expression of NPAS3.

Validation of the Prognostic Risk Model in the GEO GSE31519 Dataset

A total of 383 samples with documented survival times and clinical outcomes were obtained from the GSE31519 dataset. Figure 6A illustrates a significant correlation between lower R scores and prolonged survival time as well as reduced mortality rates ($P = 0.016$). Furthermore, as depicted in Figure 6B, the ROC curve indicates that the prognostic indicators based on these genes have a certain potential for survival prediction, especially for the initial three years. The R scores for all samples in the GSE31519 dataset were computed and are detailed in Supplementary Table 11.

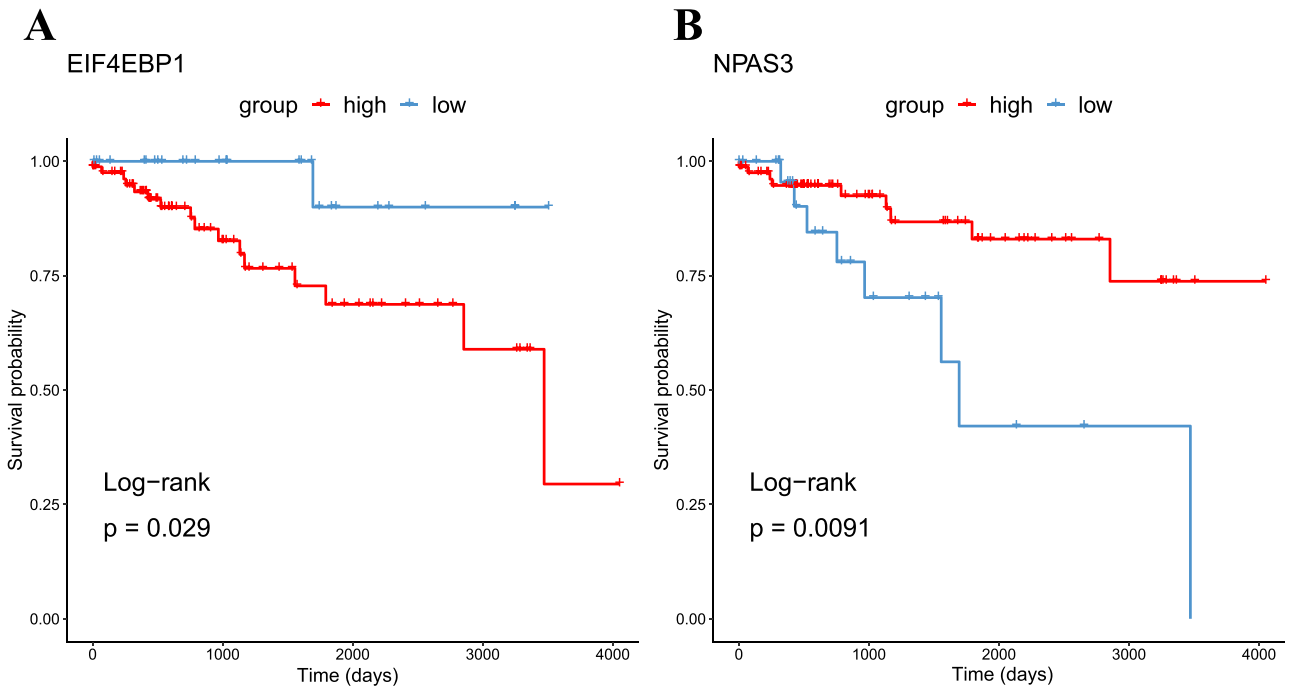

Figure 4 The overall survival probability in TNBC patients with different EIF4EBP1 (**A**) or NPAS3 (**B**) expression level.

ARG Prognostic Signature Independent of Clinical Features for Patients with TNBC

The relationship between the ARG prognostic index and various clinical characteristics, including age (Figure 7A), tumor size (Figure 7B), lymph node metastasis (Figure 7C), pathologic tumor stage (Figure 7D), and histological subtype of TNBC (Figure 7E), was analyzed. The analysis did not reveal significant associations between autophagy-related gene (ARG)-based risk scores and the clinical characteristics of TNBC. Although the deficient BRCA1/2 function may provide a survival advantage to carriers, statistical analysis indicated that BRCA1/2 mutations were associated with higher R scores ($P = 0.007$) as shown in Figure 7F.[17] Additionally, Spearman correlation analysis revealed no significant relationship between TNM stages and the risk score ($P = 0.215$, correlation coefficient = 0.118). These results suggest that the prognostic index, based on genetic factors, may be independent of these clinical features, potentially offering benefits for patient prognosis through a model that integrates genetic and clinical data.

Construction and Verification of the Nomogram

Regression analysis, which included variables such as age, TNM stage, BRCA1/2 mutation status, and R score, indicated significant correlations between TNM stage and R score, and patient survival (Supplementary Table 12). A nomogram was subsequently developed to predict 1-, 2-, 3-, and 5-year overall survival based on TNM stage and R score (c-index = 0.85) (see Figure 8A). The accuracy of the nomogram was assessed using ROC curve analysis. The AUC values for 1-year overall survival of TNM, R score, and the nomogram were 0.936, 0.75, and 0.975, respectively (Figure 8B). These data suggest that the developed nomogram provides accurate predictions for 1-year overall survival of the patients with TNBC, outperforming predictions based on either the TNM stage or R score individually. As shown in Supplementary Figure 3A–C, the AUC values for the 2-, 3-, and 5-year overall survival of the nomogram were all > 0.6, respectively. The 1-, 2-, and 3-year calibration curves indicated good agreements between the prediction and observation of overall survival outcomes (refer to Figure 8C–E). However, both the comprehensive nomogram and TNM stages demonstrated insufficient accuracy in predicting the 5-year survival rate (Supplementary Figure 3D and E). In general, the comprehensive nomogram constructed using TNM stage and R score effectively predicted 1, 2, and 3-year overall survival in TCGA patients with TNBC. However, validation efforts using the GEO GSE31519 dataset, which included histological grades due to the lack of TNM data, yielded a c-index of 0.59, indicating limited practical utility of the nomogram in this dataset (Supplementary Figure 4).

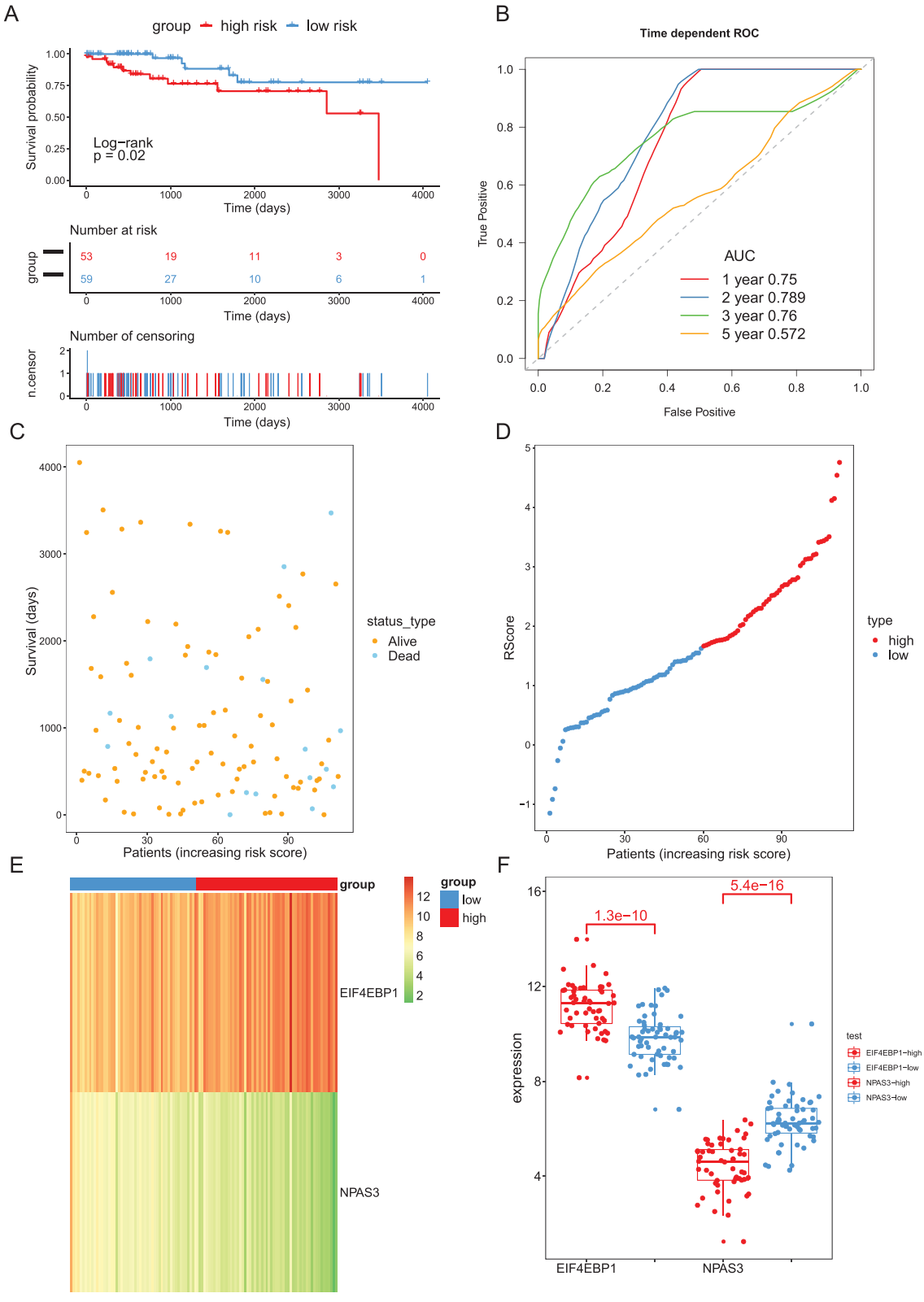

Figure 5 Establishment and evaluation of the prognostic R score model, which comprises EIF4EBP1 and NPAS3. (**A**) Survival rate comparison between high-and low-R score patients; (**B**) survival-dependent receiver operating characteristic (ROC) curves validating the prognostic significance of EIF4EBP1 and NPAS3; (**C**) survival status of patients in different risk groups; (**D**) distribution of the R score among patients in different risk groups; (**E**) heatmap illustrating the expression profiles of EIF4EBP1 and NPAS3 in different risk groups; and (**F**) significant difference analysis of EIF4EBP1 and NPAS3 between different risk groups.

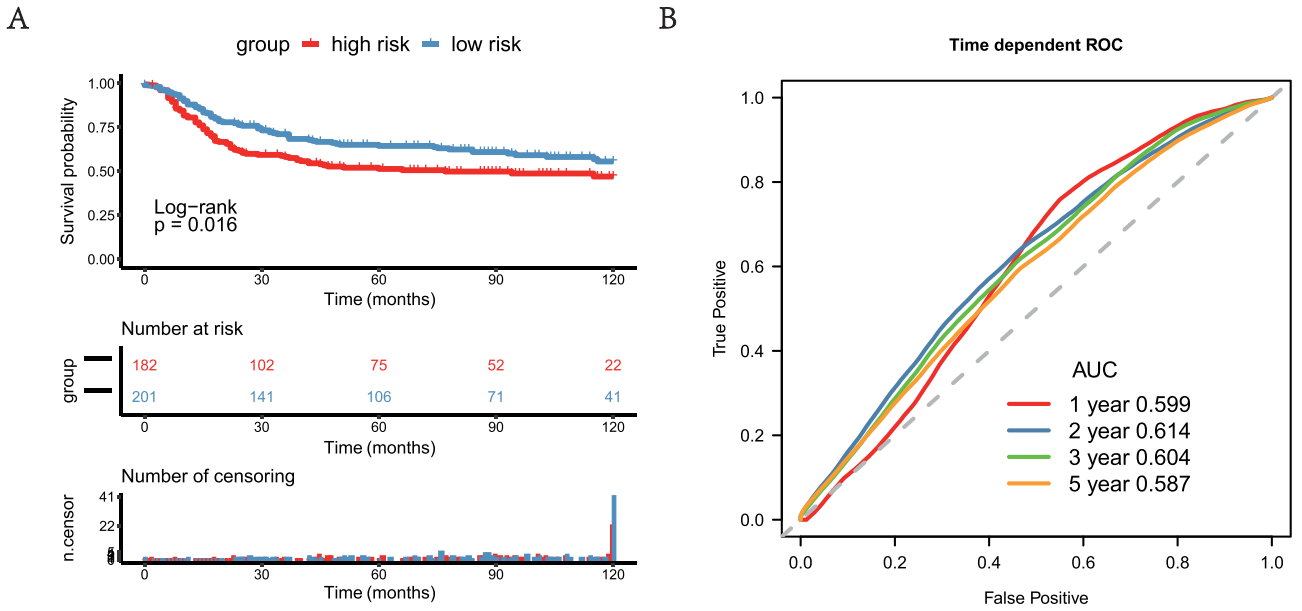

Figure 6 Validation of the prognostic R score model in the GSE31519 dataset. (**A**) Survival rate comparisons between patients with high and low R scores; (**B**) survival-dependent receiver operating characteristic (ROC) curves validating the prognostic significance of EIF4EBP1 and NPAS3.

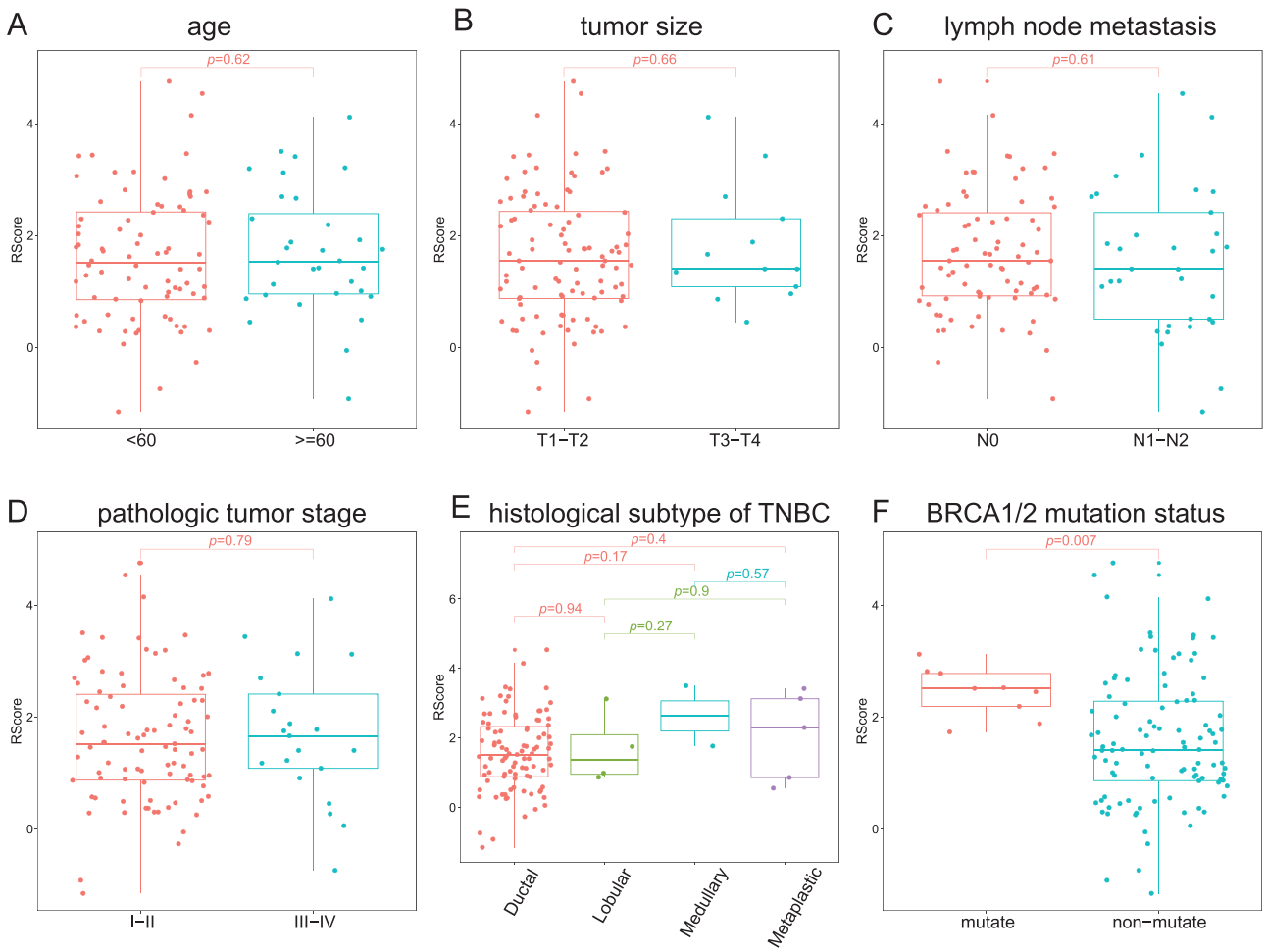

Figure 7 Clinicopathological significance of the prognostic index of TNBC. *P* values at different (**A**) age, (**B**) tumor size, (**C**) lymph node metastasis, (**D**) pathologic tumor stage, (**E**) histological subtype of TNBC, and (**F**) BRCA1/2 mutation status.

https://doi.org/10.2147/BCTT.S475007

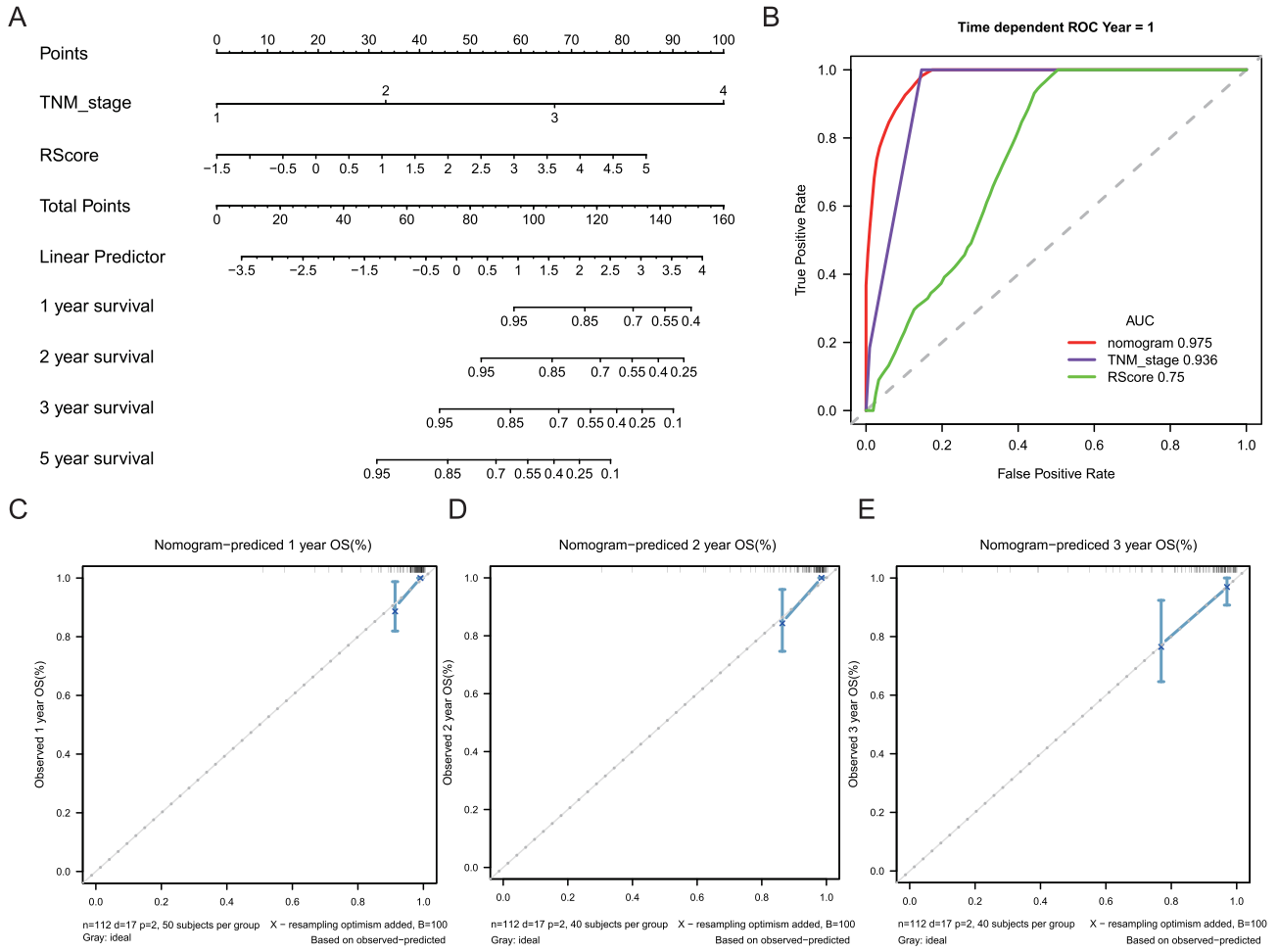

Figure 8 Construction of a nomogram and ROC curve analysis of prognosis for patients with TNBC from the TCGA database. (**A**) The nomogram is based on TNM stages and the R score; and (**B**) ROC curve analyses based on the nomogram (red), TNM stages (purple), and R score (green), showing 1-year overall survival and the corresponding AUC values; and (**C-E**) Calibration curve for 1-year, 2-year, and 3-year prediction.

Comprehensive Analysis of Genes in ARG-Related Risk Scores for Patients with TNBC

The protein expression patterns of EIF4EBP1 and NPAS3 were examined using the Human Protein Atlas (HPA) database (see Figure 9). The analysis revealed that the EIF4EBP1 protein was moderately expressed in normal breast tissues but exhibited high expression levels in cancerous tissues (see Figure 9A). Conversely, NPAS3 exhibited moderate expression in normal tissues but generally lower or undetectable expression in tumor tissues (see Figure 9B). To understand the contribution of different immune cell subpopulations to our ARGs-related prognostic model for TNBC, immune cell fractions were assessed using CIBERSORT (see Figure 10). A significant increase in neutrophil fractions ($P = 0.033$) was observed in the tumor microenvironment of patients with TNBC with a higher R scores. Higher overall expression of native B cells ($P = 0.007$) was also observed in patients with a lower risk score. All relative fractions and P-values obtained are documented in Supplementary Table 13. Furthermore, the expression of EIF4EBP1 and NPAS3 within neutrophils or native B cells was compared between breast tumor and normal samples (Supplementary Figure 5). EIF4EBP1 and NPAS3 were found to be down regulated in native B cells (P =5.19e-5 and 4.12e-4, respectively), with no significant effects observed in neutrophils. Additionally, the correlation between ARGs expression levels and immune cell infiltration levels in breast cancer basal-like samples was explored (see Supplementary Figure 6). A negative correlation was identified between EIF4EBP1 expression and the infiltration level of native B cells (Rho = −0.0232, P = 2,04e-03), while a positive correlation was observed between NPAS3 and the proportion of native B cells (Rho = 0.156,

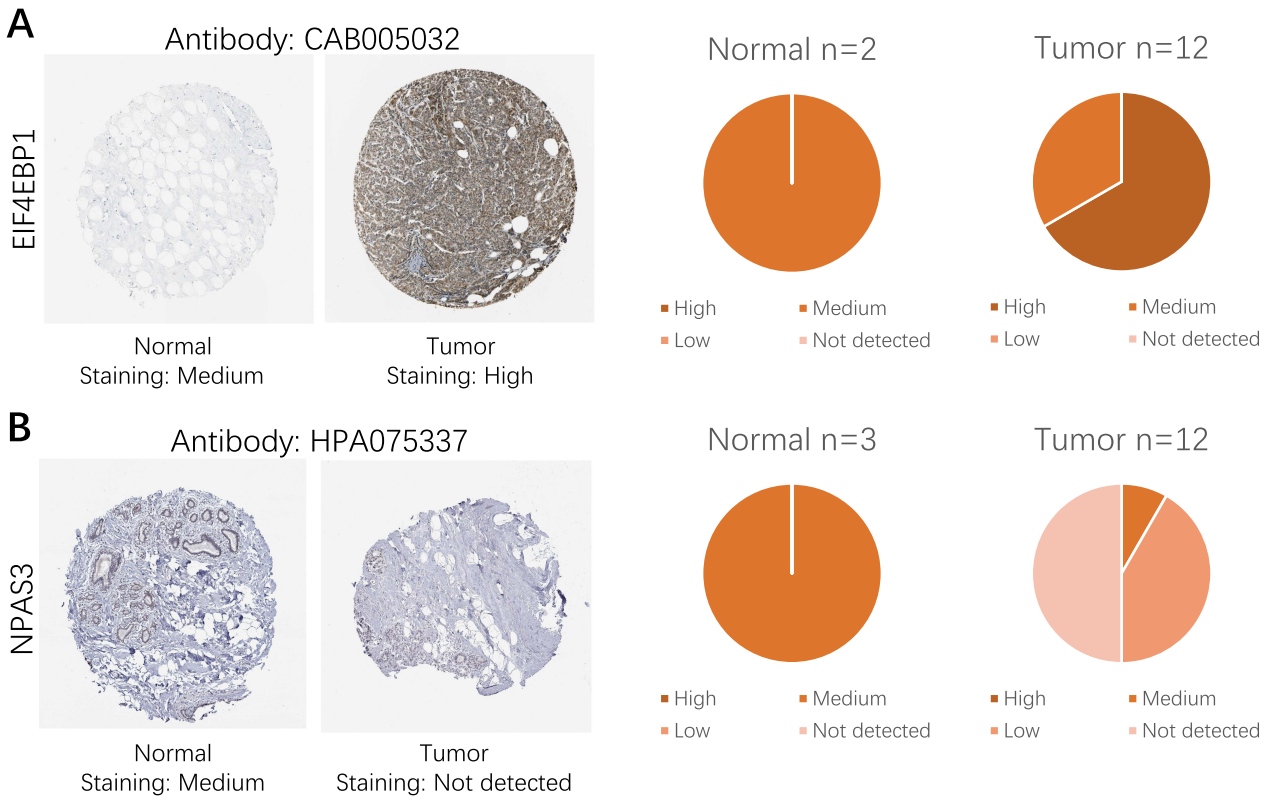

Figure 9 Analysis of the protein expression of EIF4EBP1 (**A**) and NPAS3 (**B**) in the HPA database, determined by immunohistochemistry using antibodies indicated in the HPA database.

P = 3.96e-02). Conversely, no significant correlation was observed in neutrophils. Further analysis investigated the prognostic potential of autophagy-related genes by exploring the correlation between EIF4EBP1 and NPAS3 expression levels and drug tolerance using the GDSC1 dataset. The relationship between the expression of these two genes and the IC_{50} values of several targeted drugs in the TNBC cell lines (BT-483, MDA-MB-231, BT-20, and MDA-MB-468) was analyzed. As a negative correlation between gene expression and IC_{50} value suggests increased drug sensitivity in the cell lines, our results indicated that higher expression of EIF4EBP1 increased the sensitivity of TNBC cell lines to drugs such as AZD6482, Enzastaurin, CAP-232, ZSTK474, TL-2-105, PI-103, SB505124, NSC319726, Fulvestrant, Apitolisib, Trichostatin A, and 5-Fluorouracil ($P < 0.01$; Figure 11A). No significant correlation was observed between NPAS3 expression and drug sensitivity (see Figure 11B).

NPAS3 Gene Acts as a Protective Factor to Validate the Model

The NPAS3 gene was used as a protective factor to validate the prognostic model. Immunohistochemical analysis of 54 clinical samples indicated significantly lower expression of NPAS3 in TNBC tissues compared to normal tissues, with the most pronounced difference observed in Stage III cases (see Figure 12). This study is the first to report on NPAS3 expression levels in patients with TNBC.

Discussion

TNBC, which constitutes 15% to 25% of breast cancer cases, presents significant challenges due to its lack of response to hormonal therapy and HER2-targeted treatments. There is substantial interest in understanding the mechanisms underlying TNBC progression. Autophagy mechanisms, which offer new perspectives for TNBC, were explored, resulting in the identification of two key prognostic autophagy-related genes (ARGs) that may serve as potential therapeutic targets for TNBC treatment. Initially, 202 differentially expressed ARGs between TNBC and non-tumor tissues were screened. Given the involvement of these genes in TNBC initiation, KEGG analysis revealed that, in comparison with general

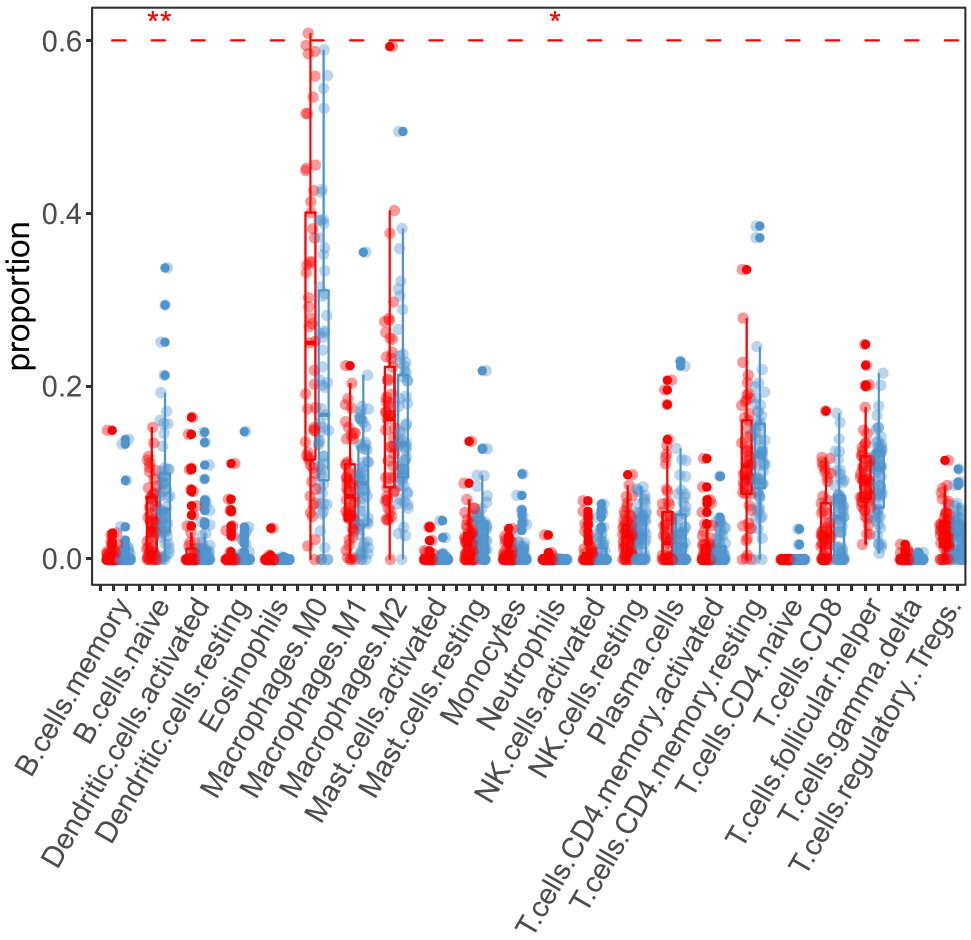

Figure 10 Relative fractions of 22 leukocyte subtypes (LM22 signature) evaluated by CIBERSORT in patients with TNBC according to high or low risk scores. Blue dots represent cases with lower risk scores, while the red ones indicate the samples with higher risk scores. * $P < 0.05$; ** $P < 0.01$.

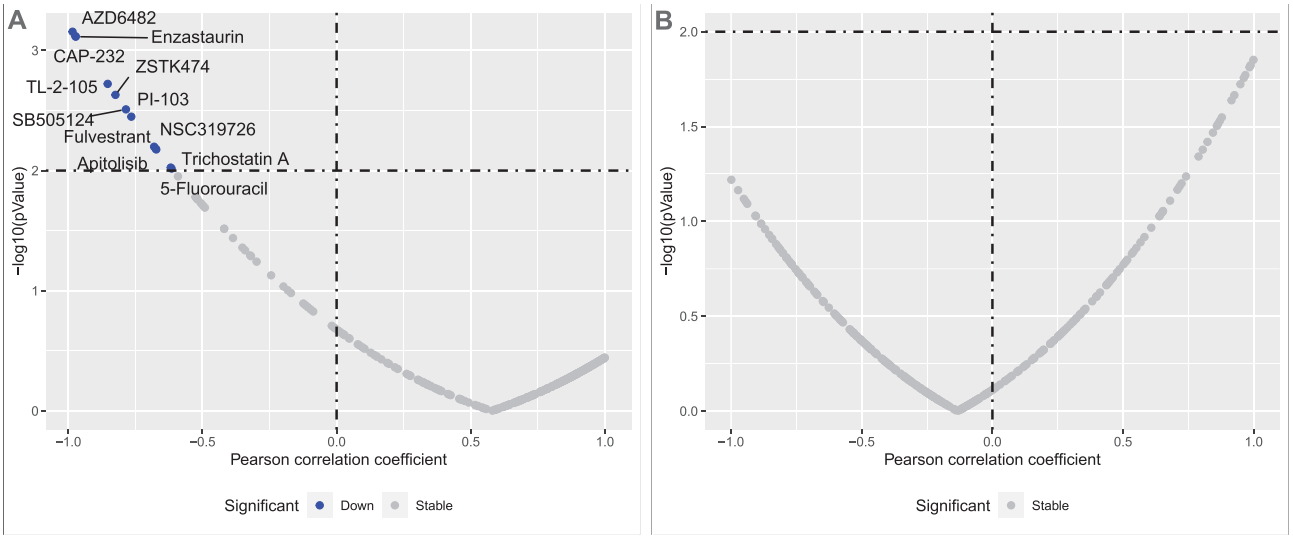

Figure 11 Correlation between the expression status of selected ARGs and drug sensitivity of TNBC cell lines, showing the correlation between the expression status of (**A**) EIF4EBP1 and (**B**) NPAS3 genes relative to the sensitivity of four TNBC cell lines to various drugs. The blue dots represent drugs that negatively correlated with the expression of genes ($P < 0.01$) based on their IC_{50} values, and the gray dots represent drugs that did not show any significant correlation with the expression of genes ($P > 0.01$).

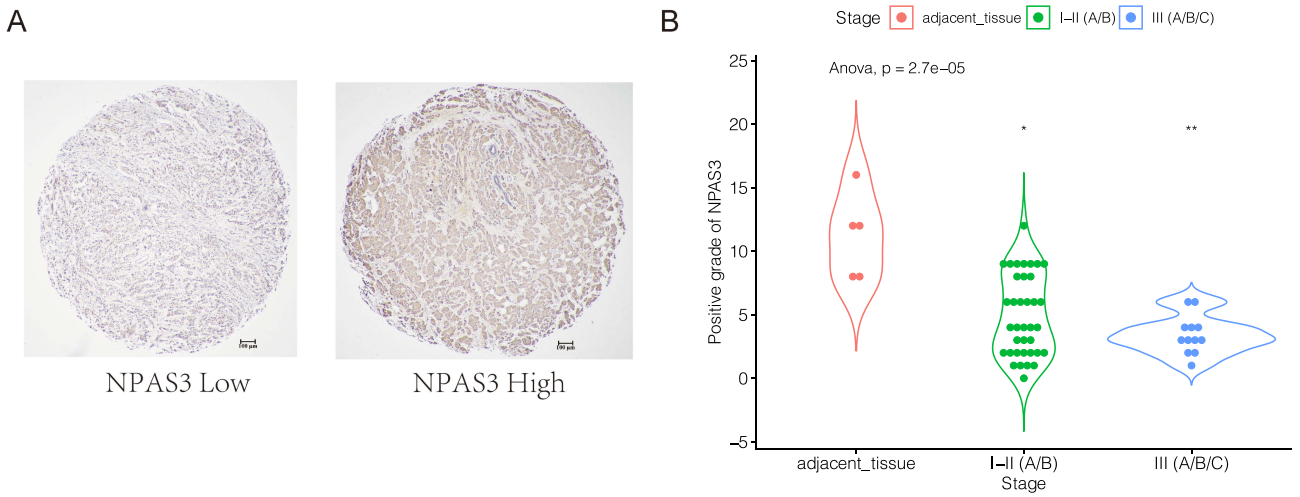

Figure 12 Immunohistochemical analysis of 54 clinical samples. (**A**) Typical Immunohistochemical results of clinical samples. (**B**) The significantly reduced NPAS3 expression in TNBC tissues compared to normal tissues, with a pronounced decrease observed in Stage III cases.

breast cancer, ARGs differentially expressed in TNBC were predominantly enriched in the PI3K-Akt signaling pathway.[8] Considering the emergence of PI3K-Akt pathway inhibitors as promising therapeutic drugs for TNBC, further exploration of the role of ARGs enriched in this pathway is warranted for a deeper understanding of TNBC pathogenesis.[18]

TNBC and basal-like subtypes represent different molecular classes of breast cancer with significant overlap. In basal-like breast cancer, terms such as EGFR tyrosine kinase inhibitor resistance and apoptosis were significantly enriched.[10] The current findings, which highlight the enrichment of ARGs in the PI3K-Akt pathway and their association with TNBC, extend previous research by offering a detailed examination of the role of ARGs in TNBC pathogenesis. While previous studies have underscored the importance of ARGs, this research specifically focuses on the PI3K-Akt pathway as a potential therapeutic target and contrasts the molecular landscape of TNBC with basal-like subtypes. This comparative analysis offers new insights, suggesting that ARGs may have distinct functions and therapeutic implications in different TNBC molecular classes.

Additionally, GO analysis revealed distinct molecular function enrichments in TNBC compared to basal-like breast cancer, notably highlighting the enrichment of ARGs in protein serine/threonine kinase activity in patients with TNBC. Subsequently, using univariate and multivariate Cox regression, ARGs signatures, including the EIF4EBP1 and NPAS3 genes, were constructed. Noteworthy is the absence of NPAS3 from previous reports, despite utilizing the same TCGA-BRCA database and similar approaches, due to the insufficient autophagy related gene databases.[5–10] Moreover, key prognostic ARGs for basal-like breast cancer, such as PARK2, FOS, BAX, and IFNG, were not significantly associated with prognosis following univariate Cox regression, highlighting a disparity between subtypes. This incongruent result between the two subtypes underscores the significance of our autophagy-related risk model for patients with TNBC.

Consistent with earlier findings, EIF4EBP1 emerged as an unfavorable survival predictor for patients with TNBC.[8] Unexpectedly, NPAS3, traditionally linked to neurogenesis, also emerged as a prognostic signature for TNBC.[19,20] Notably, this marks the first instance where NPAS3, a member of the bHLH-PAS gene family, has been highlighted as a tumor suppressive factor in driving the progression of breast cancer. While other members of the bHLH-PAS family, such as ARNT and SIM1, have been extensively investigated in cancer biology,[21,22] this study is pioneering in its identification of NPAS3's role. The different expressions of these two genes in breast cancer and normal tissue were also validated at the protein level by immunohistochemical analysis.

Our risk score model, based on the expression levels of EIF4EBP1 and NPAS3, exhibited reliable prediction capacity for the prognosis of TNBC samples in the TCGA and GSE31519 datasets. Furthermore, our prognostic index appeared to be independent of various clinical features such as age, tumor size, lymph node metastasis, pathologic tumor stage, histological subtype, and TNM stages, thereby offering a new perspective for cancer treatment. Individuals with BRCA gene mutations may exhibit increased sensitivity to chemotherapy and potentially possess a survival advantage due to

defects in homologous recombination repair (HRR) or increased immune activation.[23–25] However, our analysis revealed that individuals with BRCA1/2 mutations tend to have higher risk scores.

Finally, an effective nomogram for patients with TNBC, based on TNM stages and ARGs risk score, was developed. The finding that individuals with BRCA1/2 mutations tend to have higher risk scores introduces a novel perspective, contrasting with some prior studies that suggest a survival advantage for BRCA-related TNBC due to defects in homologous recombination repair. Additionally, the development of a predictive nomogram incorporating both TNM stages and ARGs risk scores represents an advancement in personalized medicine for TNBC. This integrated approach differs from previous ARGs-centric models, offering a more comprehensive strategy for risk assessment and potentially enhancing clinical decision-making.

Interestingly, elevated expression levels of EIF4EBP1 implied a higher risk in the prognostic model. The sensitivity of TNBC cell lines to various drug candidates such as AZD6482, Enzastaurin, CAP-232, ZSTK474, TL-2-105, PI-103, SB505124, NSC319726, Fulvestrant, Apitolisib, Trichostatin A, and 5-Fluorouracil was also found to increase. Due to the negative expression of ER, PR, and HER2, patients with TNBC exhibit insensitivity to endocrine therapy and targeted therapy.[26] Therefore, this research holds promise for guiding clinical medication strategies. However, several limitations are evident in this study. Firstly, the data used were obtained from several public databases. Although the ARG-related prognostic risk model was constructed and validated separately using the ATGC and GSE31519 datasets, the nomogram proved impractical for GSE31519 due to the absence of TNM data. In summary, while the identification of EIF4EBP1 and NPAS3 as significant autophagy-related genes for prognostic prediction in patients with TNBC offers valuable insights, their clinical implications remain uncertain and warrant validation through future clinical trials. Furthermore, although our nomogram demonstrated efficacy in predicting 1-, 2-, and 3-year overall survival, its accuracy in predicting 5-year survival was comparatively lower, highlighting the need for further research to improve its predictive capabilities by incorporating additional data sources.

This study, despite its contributions, has limitations that must be acknowledged. The reliance on retrospective data and the potential for selection bias within the datasets used may impact the generalizability of the results. Additionally, the cross-sectional nature of the data limits the ability to establish causality and to determine the applicability of these findings to broader patient populations.

Future research should focus on prospective validation of these findings in larger, diverse cohorts to ensure the clinical applicability of the ARG-based risk model. Furthermore, integrating these biomarkers into routine diagnostic panels could facilitate the development of personalized treatment strategies, ultimately improving patient outcomes in TNBC.

Conclusion

This study identified EIF4EBP1 and NPAS3 as crucial autophagy-related genes with potential for independent prognostic prediction in patients with TNBC. The expression levels of these genes were significantly associated with disease progression and survival outcomes. Based on these findings, a nomogram was developed that combines these genes with TNM staging to enhance prognostic accuracy. These findings provide a deeper comprehension of the molecular complexities underlying TNBC and hold potential implications for guiding future therapeutic approaches.

Abbreviations

TNBC, Triple negative breast cancer; TCGA, The Cancer Genome Atlas; ARGs, autophagy-related genes; IHC, immunohistochemical; TNM, Tumor node metastasis; AUC, Area under ROC curve; ROC, Receiver Operating Characteristic; GEO, Gene Expression Omnibus; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Data Sharing Statement

The original data used to support the findings of this study was obtained from public databases. Further inquiries can be directed to the corresponding author.

Ethics Approval and Consent to Participate

This study was conducted with approval from the Ethics Committee of Tianjin Central Hospital of Obstetrics and Gynecology on February 21, 2021 (2021KY024). This study was conducted in accordance with the declaration of Helsinki.

Acknowledgments

This work is supported by the grants from Tianjin Central Hospital of Gynecology Obstetrics, China. This paper has been uploaded to ResearchSquare as a preprint: https://www.researchsquare.com/article/rs-1493379/v1.

Disclosure

The authors declare that they have no competing interests.

References

1. Xu S, Liu Y, Zhang T, et al. The global, regional, and national burden and trends of breast cancer from 1990 to 2019: results from the global burden of disease study 2019. *Front Oncol*. 2021;2021:11.
2. Lee KL, Kuo YC, Ho YS, Huang YH. Triple-negative breast cancer: current understanding and future therapeutic breakthrough targeting cancer stemness. *Cancers*. 2019;11(9):1334. doi:10.3390/cancers11091334
3. Won KA, Spruck C. Triple-negative breast cancer therapy: current and future perspectives (Review). *Int j Oncol*. 2020;57(6):1245–1261. doi:10.3892/ijo.2020.5135
4. Ashrafizadeh M, Mohammadinejad R, Tavakol S, Ahmadi Z, Sahebkar A. New insight into triple-negative breast cancer therapy: the potential roles of endoplasmic reticulum stress and autophagy mechanisms. *Anti Can Ag Med Chem*. 2021;21(6):679–691. doi:10.2174/1871520620666200619180716
5. Lai J, Chen B, Mok H, Zhang G, Ren C, Liao N. Comprehensive analysis of autophagy-related prognostic genes in breast cancer. *J Cell & Mol Med*. 2020;24(16):9145–9153. doi:10.1111/jcmm.15551
6. Li X, Jin F, Li Y. A novel autophagy-related lncRNA prognostic risk model for breast cancer. *J Cell & Mol Med*. 2021;25(1):4–14. doi:10.1111/jcmm.15980
7. Lin QG, Liu W, Mo YZ, et al. Development of prognostic index based on autophagy-related genes analysis in breast cancer. *Aging-Us*. 2020;12(2):1366–1376. doi:10.18632/aging.102687
8. J-X D, Chen C, Luo Y-H, et al. Establishment and validation of a novel autophagy-related gene signature for patients with breast cancer. *Gene*. 2020;2020:762.
9. Li XL, Zhang HC, Liu JJ, Li P, Sun Y. Five crucial prognostic-related autophagy genes stratified female breast cancer patients aged 40-60 years. *BMC Bioinf*. 2021;22(1). doi:10.1186/s12859-021-04503-y
10. Han B, Zhang H, Zhu Y, et al. Subtype-specific risk models for accurately predicting the prognosis of breast cancer using differentially expressed autophagy-related genes. *Aging-Us*. 2020;12(13):13318–13337. doi:10.18632/aging.103437
11. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47–e47. doi:10.1093/nar/gkv007
12. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J Integ Biol*. 2012;16(5):284–287. doi:10.1089/omi.2011.0118
13. Li T, Fu J, Zeng Z, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res*. 2020;48(W1):W509–W14. doi:10.1093/nar/gkaa407
14. Uhlen M, Fagerberg L, Hallstroem BM, et al. Tissue-based map of the human proteome. *Science*. 2015;347(6220). doi:10.1126/science.1260419
15. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*. 2015;12(5):453–457. doi:10.1038/nmeth.3337
16. Li C, Tang Z, Zhang W, Ye Z, Liu F. GEPIA2021: integrating multiple deconvolution-based analysis into GEPIA. *Nucleic Acids Res*. 2021;49(W1):W242–W6. doi:10.1093/nar/gkab418
17. Sporikova Z, Koudelakova V, Trojanec R, Hajduch M. Genetic markers in triple-negative breast cancer. *Clin Breast Can*. 2018;18(5):E841–E50. doi:10.1016/j.clbc.2018.07.023
18. Cipriano E, Mesquita A. Emerging therapeutic drugs in metastatic triple-negative breast cancer. *Breast Can*. 2021;15:11782234211002491.
19. Moreira F, Kiehl T-R, So K, et al. NPAS3 demonstrates features of a tumor suppressive role in driving the progression of astrocytomas. *Am J Pathol*. 2011;179(1):462–476. doi:10.1016/j.ajpath.2011.03.044
20. Di CF, Dal BM, Baboci L, De ME, Polano M, Toffoli G. The molecular and microenvironmental landscape of glioblastomas: implications for the novel treatment choices. *Front Neurosci*. 2020;14:603647. doi:10.3389/fnins.2020.603647
21. Opitz CA, Holfelder P, Prentzell MT, Trump S. The complex biology of aryl hydrocarbon receptor activation in cancer and beyond. *Biochem Pharmacol*. 2023;216:115798. doi:10.1016/j.bcp.2023.115798
22. Ebata T, Yamashita S, Takeshima H, et al. DNA methylation marker to estimate ovarian cancer cell fraction. *Med Oncol*. 2022;39(5):78. doi:10.1007/s12032-022-01679-y
23. Copson ER, Maishman TC, Tapper WJ, et al. Germline BRCA mutation and outcome in young-onset breast cancer (POSH): a prospective cohort study. *Lancet Oncol*. 2018;19(2):169–180. doi:10.1016/S1470-2045(17)30891-4
24. Jiang T, Shi W, Wali VB, et al. Predictors of chemosensitivity in triple negative breast cancer: an integrated genomic analysis. *PLoS Med*. 2016;13(12):e1002193. doi:10.1371/journal.pmed.1002193

25. Qureshi Z, Jamil A, Altaf F, Siddique R, Safi A. Efficacy and Safety of BRCA-targeted Therapy (Polyadenosine Diphosphate-ribose Polymerase Inhibitors) in Treatment of BRCA-mutated breast cancer: a systematic review and meta-analysis. *Am J Clin Oncol*. 2024. doi:10.1097/COC.0000000000001120
26. Vagia E, Mahalingam D, Cristofanilli M. The landscape of targeted therapies in TNBC. *Cancers*. 2020;12(4):916. doi:10.3390/cancers12040916