

Development of prognostic index based on autophagy-related genes analysis in breast cancer

Qing-Guang Lin^{1,*}, Wei Liu^{2,*}, Yu-zhen Mo³, Jing Han¹, Zhi-Xing Guo¹, Wei Zheng¹, Jian-wei Wang¹, Xue-Bin Zou¹, An-Hua Li¹, Feng Han¹

¹Department of Ultrasound, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou 510060, Guangdong, China

²Department of Breast, Guangzhou Red Cross Hospital, Medical College, Jinan University, Guangzhou 510220, Guangdong, China

³Department of Radiotherapy, Guangzhou Red Cross Hospital, Medical College, Jinan University, Guangzhou 510220, Guangdong, China

*Co-first authors

Correspondence to: Feng Han; email: hanfeng@sysucc.org.cn

Keywords: autophagy-related genes, breast cancer, prognosis

Received: October 14, 2019

Accepted: December 25, 2019

Published: January 22, 2020

Copyright: Lin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Background: Autophagy is a self-digesting process that can satisfy the metabolic needs of cells, and is closely related to development of cancer. However, the effect of autophagy-related genes (ARGs) on the prognosis of breast cancer remains unclear.

Results: We first found that 27 ARGs were significantly associated with overall survival in breast cancer. The prognosis-related ARGs signature established using the Cox regression model consists of 12 ARGs that can be divided patients into high-risk and low-risk groups. The overall survival of patients with high-risk scores (HR 3.652, 2.410-5.533; $P < 0.001$) was shorter than patients with low-risk scores. The area under the receiver operating characteristic (ROC) curve for 1-year, 3-year, and 5-year survival rates were 0.739, 0.727, and 0.742, respectively.

Conclusion: The 12-ARGs marker can predict the prognosis of breast cancer and thus help individualized treatment of patients at different risks.

Methods: Based on the TCGA dataset, we integrated the expression profiles of ARGs in 1,039 breast cancer patients. Differentially expressed ARGs and survival-related ARGs were evaluated by computational difference algorithm and COX regression analysis. In addition, we also explored the mutations in these ARGs. A new prognostic indicator based on ARGs was developed using multivariate COX analysis.

INTRODUCTION

Autophagy is a multi-step lysosomal degradation process that promotes nutrient cycling and metabolic adaptation, and has been extensively studied and been proven involved in the development of cancer [1]. However, the function of autophagy in tumors is bilateral, may be cancer-promoting, or may be a tumor

suppressor, depending on the type of tumor and the stage of the tumor [2]. For example, autophagy can remove damaged organelles and/or DNA before canceration to maintain normal cellular structure and metabolic stability, thereby exerting a tumor suppressing effect [3]. To the stage of tumor progression, autophagy is often up-regulated and promotes tumor cell proliferation and invasion by

absorbing nutrients and energy from degrading proteins and organelles [4]. Autophagy is a complex multi-step process that is tightly controlled by a series of autophagy-related genes (ARGs).

The incidence of breast cancer ranks first among women in cancer worldwide, and it is a significant threat to the health of women [5, 6]. It is well known that breast cancer is a group of highly heterogeneous diseases, and the prognosis of individuals varies widely [7]. Clinically, tumor staging, histological grades, and molecular subtypes are used to evaluate the prognostic factors of breast cancer patients. However, these clinicopathological features do not accurately provide information to predict a patient's prognosis. This may lead to inaccurate judgments on the patient's prognosis, and some low-risk patients may receive unnecessary or excessive treatment, while other high-risk patients may face relapse or metastasis due to inadequate treatment. Therefore, there is an urgent need to find new molecular markers to predict the prognosis of breast cancer patients, which is conducive to the precise treatment of patients.

A large number of studies have reported a correlation between autophagy and breast cancer [8–10]. For example, Vera-Ramirez et al. reported that autophagy promotes the therapeutic resistance of breast cancer stem cells and contributes to its survival [11]. Notable, previous studies have focused on the association between single or a few ARGs and breast cancer progression. Currently, studies using large-scale ARGs expression profiles to screen and identify molecular markers for predicting the prognosis of breast cancer are lacking. The purpose of this study was to gain insight into the potential clinical utility of ARGs for prognostic stratification and to facilitate the development of personalized prognostic information for breast cancer patients. We combined ARGs expression profiles with clinical information to systematically analyze the expression status of ARGs and its impact on prognosis.

RESULTS

Identification of differentially expressed ARGs

RNA-seq and clinical data from 1109 breast cancer tissue samples and 113 non-tumor samples were downloaded from TCGA. Of these patients, a total of 1039 patients with primary breast cancer who were followed for more than 1 month were included in the study. The expression values of 232 ARGs were extracted. Considering the criteria for $FDR < 0.05$ and $[\log_2(\text{fold change})] > 1$, we finally obtained 13 up-regulated and 16 down-regulated ARGs (Figure 1A and 1B). A detailed flow chart for the establishment of the prediction model was shown in Supplementary Figure 1.

In addition, a scatter plot was visualized to show the expression pattern of 29 differentially expressed ARGs between breast cancer and non-tumor tissue (Figure 1C). Scatter plot showing expression patterns of 13 down-regulated genes (*CCL2*, *DLC1*, *EGFR1*, *CX3CLI*, *FOS*, *FOXO1*, *GABARAPL1*, *MAP1LC3C*, *NRG1*, *NRG2*, *PPP1R15A*, *PRKN* and *TP63*) and 16 up-regulated genes (*ATG9B*, *BAK1*, *BAX*, *BIRC5*, *CDKN2A*, *CTSD*, *CXCR4*, *EIF4EBP1*, *ERBB2*, *FADD*, *GAPDH*, *IKBKE*, *IL24*, *PARP1*, *PTK6* and *RGS19*).

Functional enrichment of the differentially expressed ARGs

Functional enrichment analysis of 29 differentially expressed ARGs provides a biological understanding of these genes. Top 30 of GO enrichment and top 30 of pathway enrichment are summarized in Figure 2A. GO enrichment shows that the biological process of differential genes is mainly involved in autophagy, apoptosis and endopeptidase regulation. KEGG enrichment shows that pathways of differential genes mainly involve pathways in cancer, protein processing in endoplasmic reticulum, cytokine-cytokine receptor interaction and the like (Figure 2B).

Identification of prognostic ARGs

To analyze ARGs involved in breast cancer progression, we screened for ARGs that were significantly associated with prognosis. The forest map of the hazard ratio indicates that most of these genes are protective factors (Figure 3A). Both GO and KEGG analysis showed that these genes are closely related to autophagy-related biological processes and signaling pathways (Figure 3B and 3C). Given the important clinical implications of these ARGs, we examined the genetic alterations of these genes and found that mRNA up-regulation and fusion are the two most common types of mutations (Figure 4). A total of 22 genes have a mutation rate $\geq 5\%$, of which *EIF4EBP1* is the most frequently mutated gene (20%).

Prognostic ARGs with significant significance after the above univariate analysis were further included in the subsequent multivariate analysis. A total of 12 genes were significantly associated with prognosis after multivariate analysis. The expression patterns of these 12 genes were shown in supplementary Figure 2. Protein-protein interaction (PPI) network analysis indicated that *EIF4G1*, *CASP8* and *MAP1C3CA* are the three core genes of these 12 genes (Supplementary Figure 3). Based on the results of multivariate Cox regression analysis, we constructed autophagy prognostic index (API) to divide breast cancer patients into two groups with discrete clinical outcomes for overall survival (OS).

[Expression level of *CASP8** (-0.5681)] + [Expression level of *EIF4G1* * (0.3535)] + [Expression level of *MAP1LC3A* * (-0.2183)] + [Expression level of *HSPB8* * 0.1302] + [Expression level of *NRG1* * (-0.6381)] + [Expression level of *SERPINA1** (-0.1772)] + [Expression level of *EIF4EBP1* * (0.1618)] + [Expression level of *BAG1** (-0.2884)] + [Expression level of *CASP4* * (0.3839)] + [Expression level of *AMBRA1* * (-0.4073)] + [Expression level of *ATG4A* * 0.6891] + [Expression level of *IFNG* * (-0.9043)].

Figure 5 showed distribution of prognostic index in TCGA dataset (Figure 5A), survival status of patients in different groups (Figure 5B) and heatmap of the

expression profile of the included ARGs (Figure 5C). To determine the performance of the API in predicting clinical outcomes in breast cancer patients, K-M survival curves were plotted to analyze different survival times between high-risk and low-risk groups. K-M analysis showed that the survival rate of patients in the high-risk group was significantly lower than that in the low-risk group (Figure 5D). Univariate analysis showed that ARI was significantly associated with patient prognosis (Figure 6A). In addition, after adjusting for clinicopathological features such as age, tumor subtype, tumor stage, tumor size, and lymph node metastasis, API remained an independent prognostic indicator for breast cancer patients in multivariate

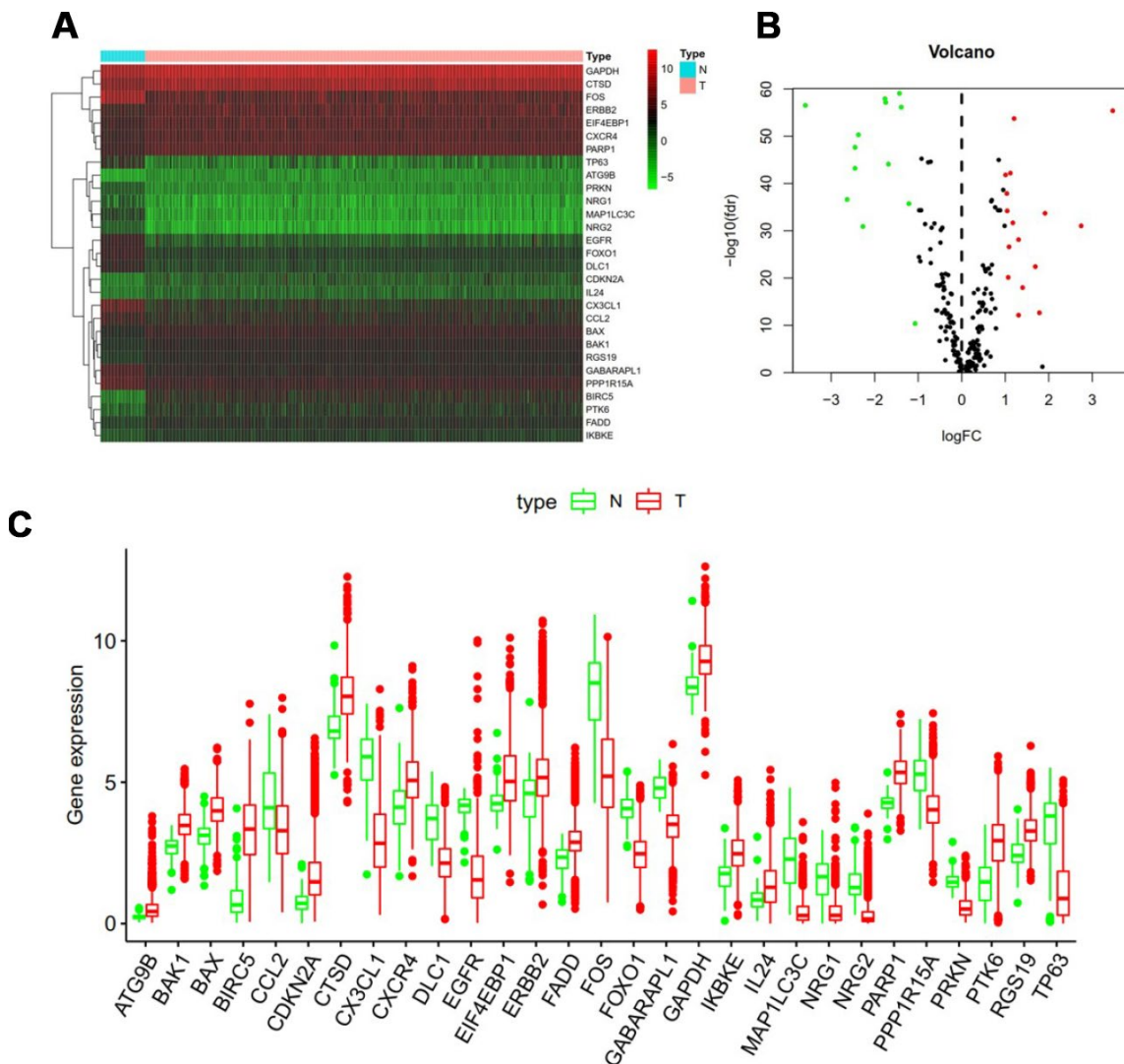


Figure 1. Differentially expressed autophagy-related genes. Heat map (A) and volcano map (B) show differentially expressed genes between breast cancer and normal tissues, with red dots representing significantly up-regulated genes, green dots representing significantly down-regulated genes, and black dots representing no differences gene. (C) Expression patterns of 29 autophagy-related genes (ARGs) in breast cancer types and paired non-tumor samples. Each red box plot represents a different tumor sample and blue represents a non-tumor sample.

analysis (HR = 3.105, 95% CI = 1.988-4.848; $P < 0.001$; Figure 6B). The area under the curve of the corresponding receiver operating characteristic (ROC) curve for 1 year, 3 years, and 5 years of survival is 0.739, 0.727, and 0.742, respectively. This indicated that the prognostic index based on ARGs has a certain potential in survival prediction (Figure 6C).

Clinical utility of prognostic signature

Relationship between ARGs prognostic index and clinical features were subsequently analyzed. Significant increases in risk score were in larger tumor size (Figure 7A), lymph node metastasis (Figure 7B), late clinical stage (Figure 7C), HER2 subtype and luminal B subtype (Figure 7D).

DISCUSSION

Although numerous studies have demonstrated that autophagy is involved in the malignant progression of breast cancer, a comprehensive analysis of ARGs have not been conducted to explore its clinical significance. To analyze breast cancer prognosis-related genes from the perspective of autophagy, we screened and identified 12 prognostic ARGs. Our results suggested that a prognostic model based on 12 ARGs can be used for prognostic stratification in breast cancer patients,

thereby helping to develop individualized treatment options based on patient risk.

We identified a group of ARGs that predict the prognosis of breast cancer patients. Most of these genes have been reported in previous studies to be closely related to the prognosis of breast cancer or other malignancies [12]. Loss of CASP8 protein expression is associated with a poor prognosis in children with medulloblastoma [13]. Muhammad JS et al. [11] reported that *Helicobacter pylori*-induced *MAP1LC3* methylation silencing may impair autophagy and promote gastric cancer. Overexpression of EIF4G1 is associated with tumor progression and poor prognosis in nasopharyngeal carcinoma [14]. HSPB8 promotes cancer cell growth and is associated with poor prognosis in patients with gastric cancer [15]. Overexpression of phosphorylated EIF4EBP1 is closely associated with tumor recurrence and worse survival outcomes of cervical cancer [16]. Afentakis M et al. reported that the incidence of distant recurrence in women with higher BAG1 expression was reduced by 30% compared with women with low expression of breast cancer [17]. The circulating NRG1 reported by De Iuliis F et al. may be a biomarker for the prognosis of breast cancer patients [18]. Boccellino M et al found that SERPINA1 may be a useful biomarker for early detection of lung cancer and monitoring its evolution [19]. ATG4A has been reported to promote tumor metastasis by

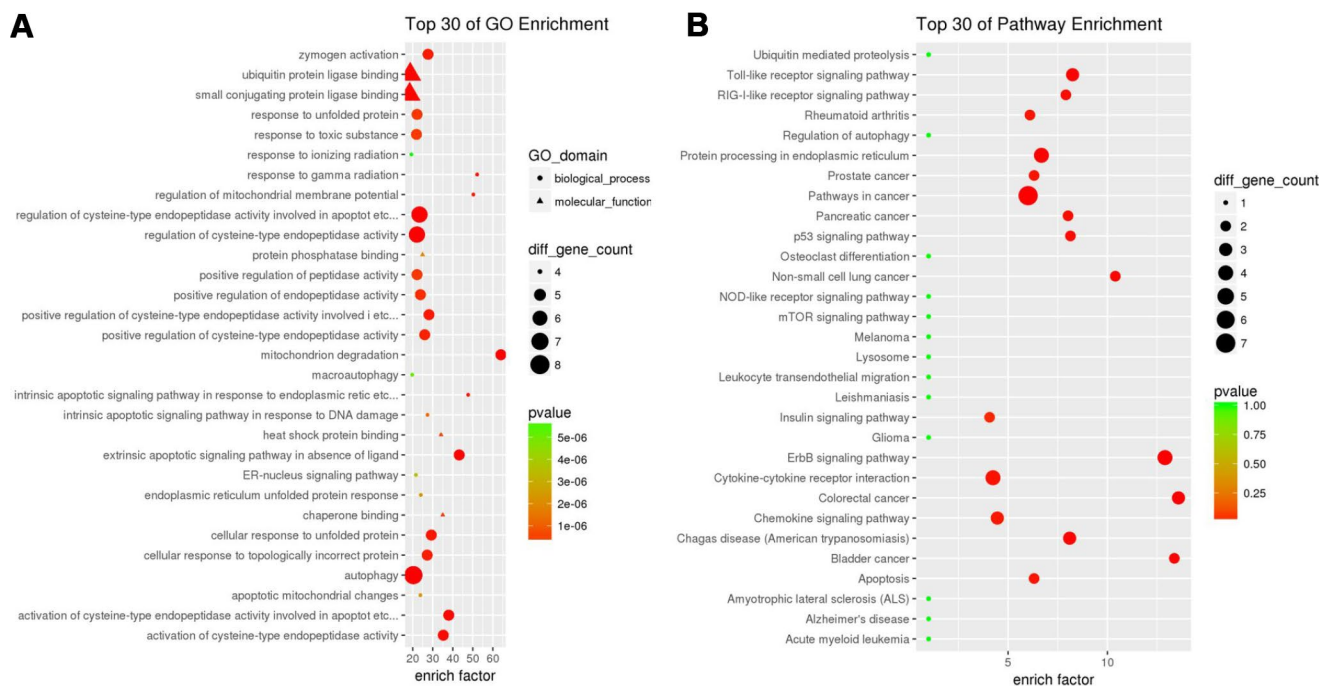


Figure 2. Gene functional enrichment of differentially expressed ARGs. (A) GO analysis shows the biological processes and molecular functions involved in differential genes. **(B)** KEGG shows the signaling pathway involved in differential ARGs.

inducing epithelial-mesenchymal transition and stem cell-like properties of gastric cells [20]. In triple-negative breast cancer, high expression of IFNG was found to be associated with better disease-free survival. The focus of this study was on the relationship between the mRNA expression of ARGs and the prognosis of breast cancer

patients. It is known that the genetic alternation of genes is likely to affect the expression levels of their mRNAs. Gene amplification is often positively correlated with up-regulation of mRNA expression. For example, we noted that gene amplification and mRNA upregulation are the most common genetic variants for the *EIF4EBP1* gene.

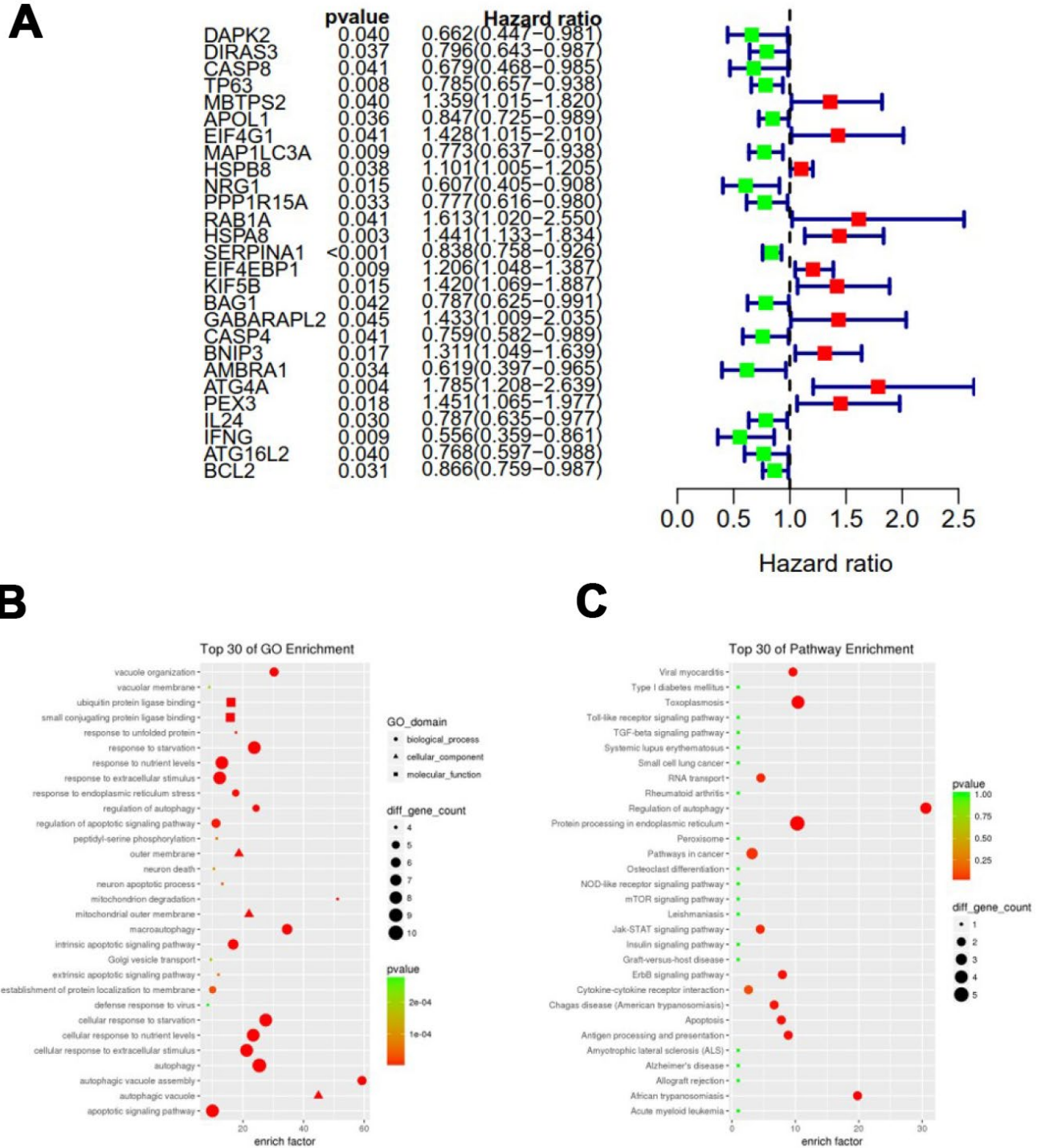


Figure 3. Expression profile and prognostic value of ARGs. (A) Risk ratio forest plot showed the prognostic value of the gene; (B) GO analysis revealed the biological processes and molecular functions involved in 27 prognostic-related ARGs; (C) KEGG shows the signaling pathways involved in 27 prognostic-related ARGs.

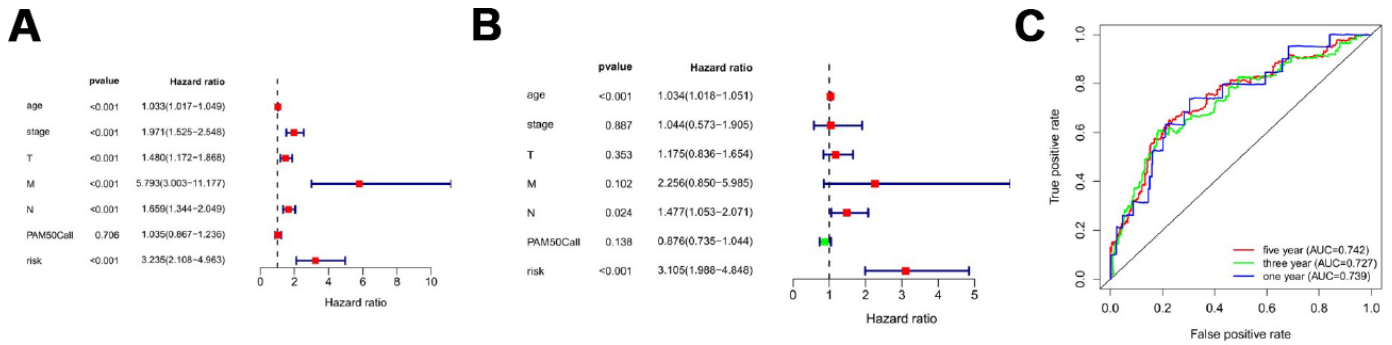


Figure 6. Prognostic indicators based on ARGs show good predictive performance. A forest plot of univariate (A) and multivariate (B) Cox regression analysis in breast cancer. (C) Survival-dependent receiver operating characteristic (ROC) curves validate the prognostic significance of ARGs-based prognostic indicators.

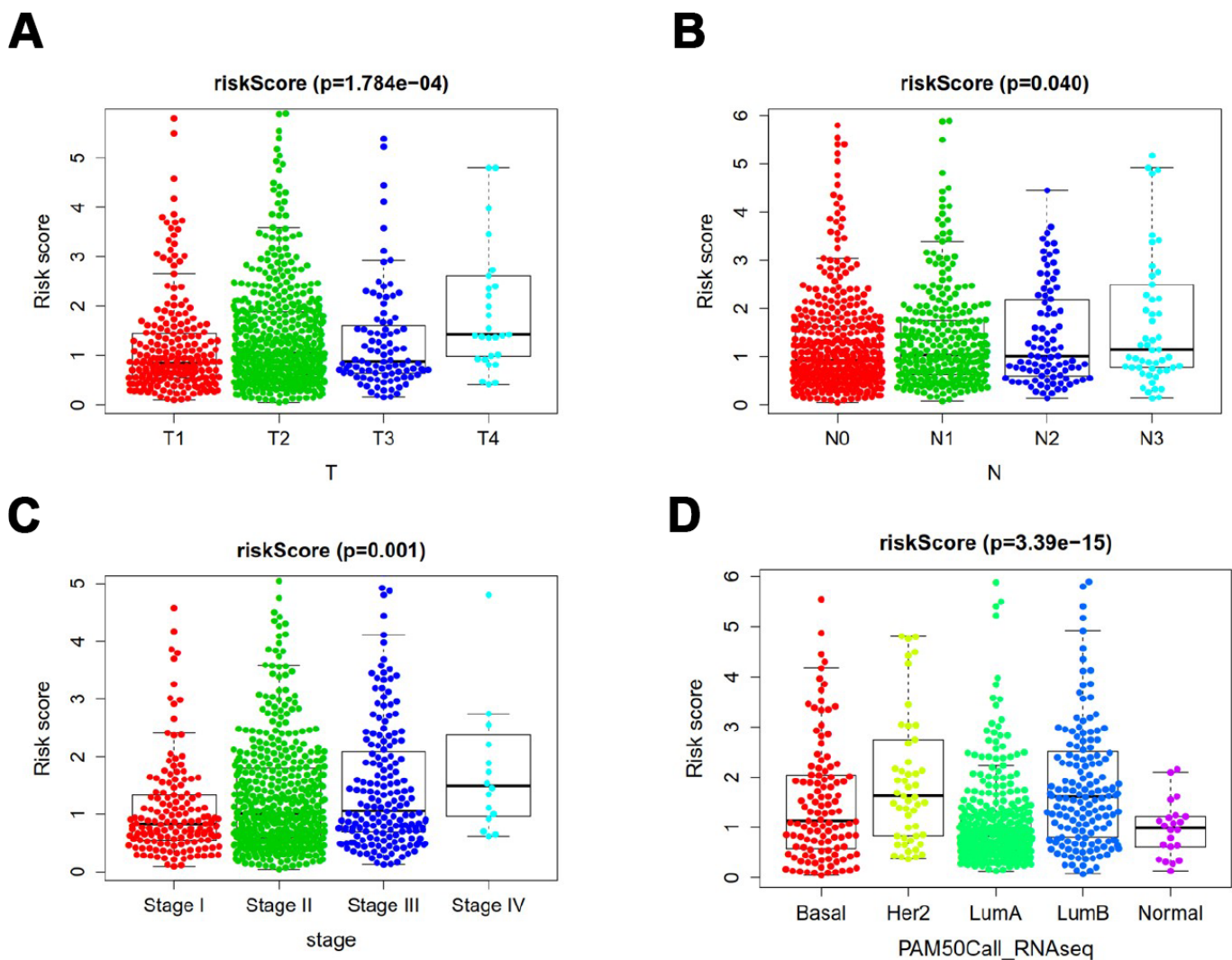


Figure 7. Clinicopathological significance of the prognostic index of breast cancer. P values were at different (A) tumor size, (B) lymph node metastasis (C) tumor stage, and (D) tumor subtypes.

clinical features. The genes identified in the autophagy pathway also offer new possibilities for breast cancer therapeutic intervention. Based on the molecular features of autophagy, we constructed a new risk scoring model that can effectively assess the prognosis of breast cancer patients. However, prospective studies are needed to further validate the findings of this study to aid clinical personalized treatment.

MATERIALS AND METHODS

Autophagy related genes (ARGs)

A total of 232 ARGs were extracted from the HADB database (Human Autophagy Database, <http://autophagy.lu/clustering/index.html>), which provides a complete, up-to-date list of human genes involved in autophagy.

TCGA data acquisition

Our study included only 1039 breast cancer patients who were followed up for at least one month from the TCGA database, with a follow-up time ranging 1 to 283 months. ARGs associated with patient survival were identified using univariate Cox regression for subsequent model construction.

Functional analysis

The Bohao Online Enrichment Tool (<http://enrich.shbio.com/>) was used to perform functional enrichment of differentially expressed ARGs. Gene Ontology (GO) and the Kyoto Gene and Genomic Encyclopedia (KEGG) were used to assess relevant functional categories. GO and KEGG enrichment pathways with *p* and *q* values less than 0.05 are considered to be significant categories.

Construction of ARGs related prognostic model

Prognosis-related genes were constructed using multivariate cox regression. After incorporating the expression values for each particular gene, a risk score formula for each patient was constructed and weighted by its estimated regression coefficients in a multivariate cox regression analysis. According to the risk scoring formula, the median risk score was used as the cut-off point, and the patients were divided into low-risk group and high-risk group. Survival differences between the two groups were assessed by Kaplan-Meier and compared using log-rank statistical methods. Multivariate cox regression analysis and stratified analysis were used to examine the role of risk scores in predicting patient outcomes. ROC curves were used to study the accuracy of model predictions.

Statistical analysis

Survival curves were generated by the Kaplan-Meier method and compared by log-rank test. Multivariate analysis was performed using the cox proportional hazard model. All statistical analyses were performed using the R language (version 3.6). All statistical tests were bilateral, with *p* < 0.05 being statistically significant.

AUTHOR CONTRIBUTIONS

FG, QGL and WL designed the study. Data collection were conducted with QGL, WL and YZM. QGL, WL, YZM, JH, ZXN, WZ, JWW, XBJ, AHL and FH analyzed and interpreted these data. QGL and WL wrote the manuscripts. The final draft read and approved by all authors.

CONFLICTS OF INTEREST

The authors declare that there is no conflicts of interest.

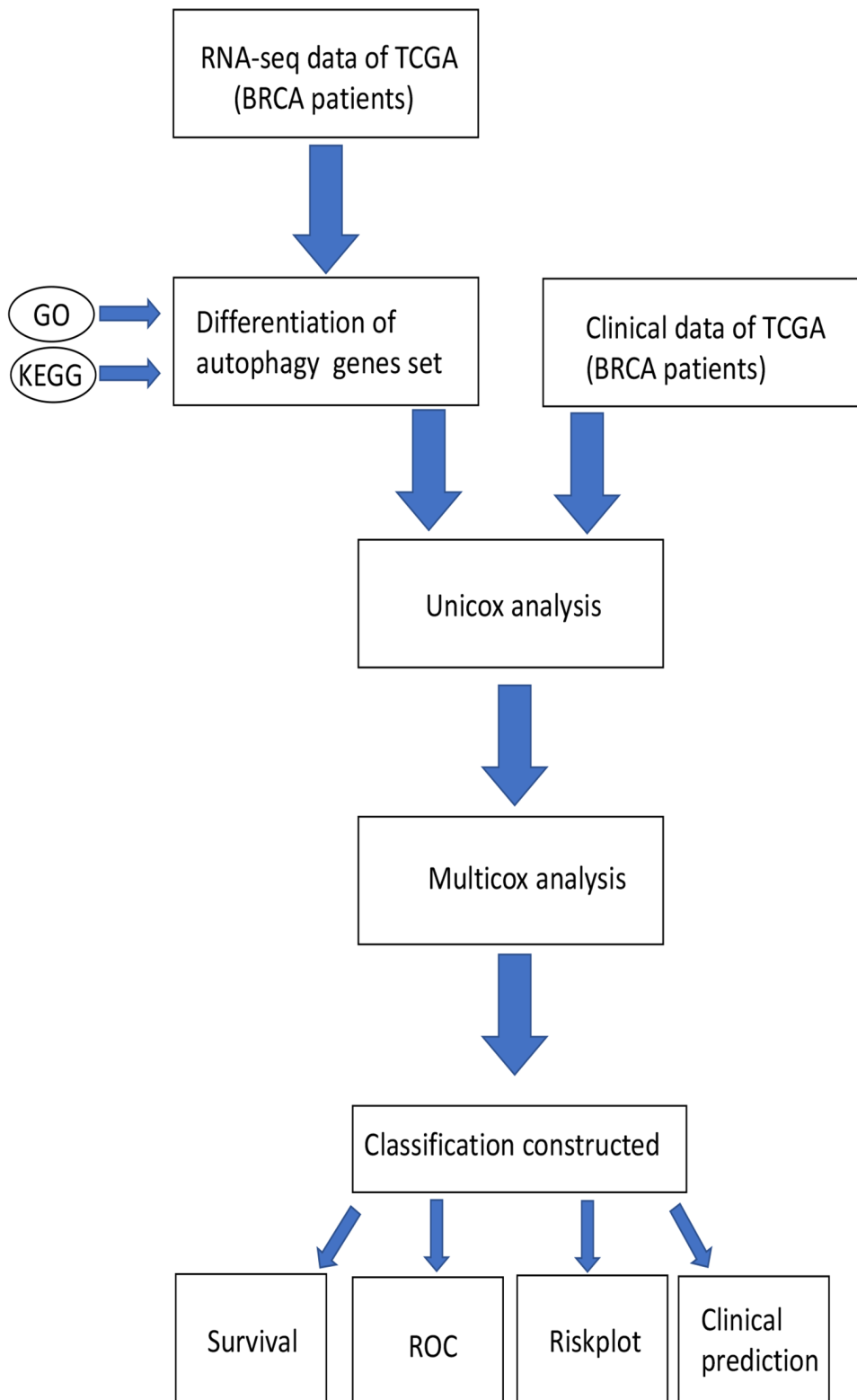
REFERENCES

1. Amaravadi RK, Kimmelman AC, Debnath J. Targeting Autophagy in Cancer: Recent Advances and Future Directions. *Cancer Discov.* 2019; 9:1167–81. <https://doi.org/10.1158/2159-8290.CD-19-0292> PMID:[31434711](https://pubmed.ncbi.nlm.nih.gov/31434711/)
2. Jiang GM, Tan Y, Wang H, Peng L, Chen HT, Meng XJ, Li LL, Liu Y, Li WF, Shan H. The relationship between autophagy and the immune system and its applications for tumor immunotherapy. *Mol Cancer.* 2019; 18:17. <https://doi.org/10.1186/s12943-019-0944-z> PMID:[30678689](https://pubmed.ncbi.nlm.nih.gov/30678689/)
3. Hönscheid P, Datta K, Muders MH. Autophagy: detection, regulation and its role in cancer and therapy response. *Int J Radiat Biol.* 2014; 90:628–35. <https://doi.org/10.3109/09553002.2014.907932> PMID:[24678799](https://pubmed.ncbi.nlm.nih.gov/24678799/)
4. Cheong H. Integrating autophagy and metabolism in cancer. *Arch Pharm Res.* 2015; 38:358–71. <https://doi.org/10.1007/s12272-015-0562-z> PMID:[25614051](https://pubmed.ncbi.nlm.nih.gov/25614051/)
5. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019; 69:7–34. <https://doi.org/10.3322/caac.21551> PMID:[30620402](https://pubmed.ncbi.nlm.nih.gov/30620402/)
6. DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J Clin.* 2017; 67:439–48. <https://doi.org/10.3322/caac.21412> PMID:[28972651](https://pubmed.ncbi.nlm.nih.gov/28972651/)

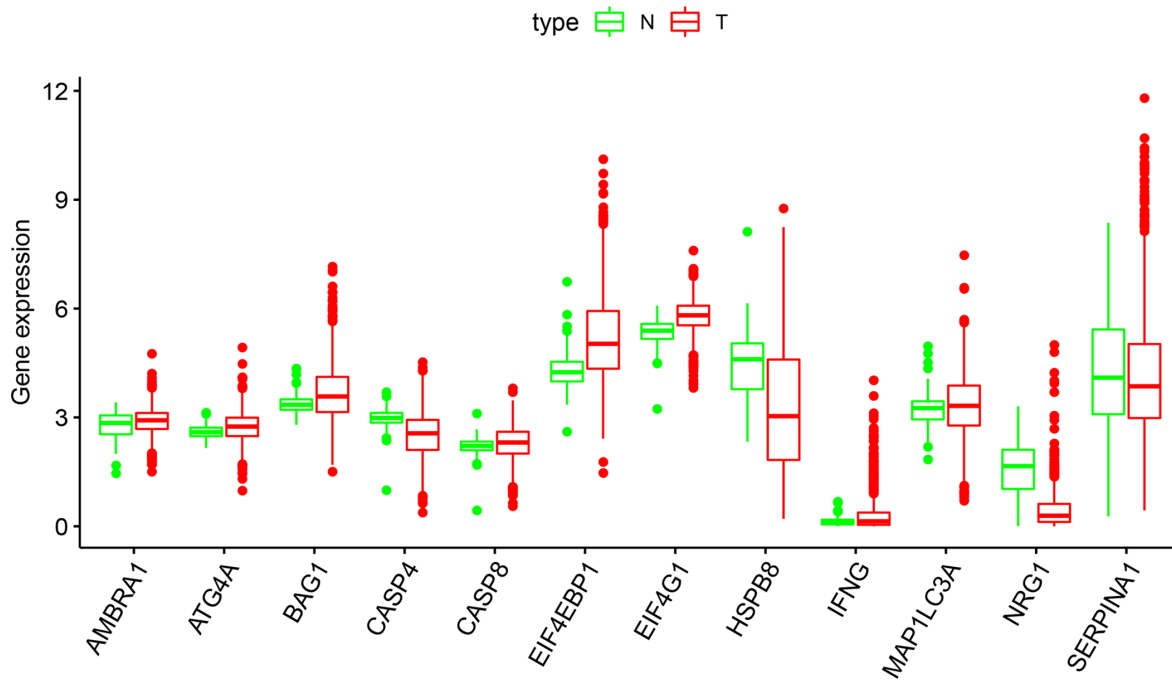
7. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *Lancet*. 2011; 378:1812–23.
[https://doi.org/10.1016/S0140-6736\(11\)61539-0](https://doi.org/10.1016/S0140-6736(11)61539-0)
PMID:[22098854](https://pubmed.ncbi.nlm.nih.gov/22098854/)
8. Huo Y, Cai H, Teplova I, Bowman-Colin C, Chen G, Price S, Barnard N, Ganesan S, Karantza V, White E, Xia B. Autophagy opposes p53-mediated tumor barrier to facilitate tumorigenesis in a model of PALB2-associated hereditary breast cancer. *Cancer Discov*. 2013; 3:894–907.
<https://doi.org/10.1158/2159-8290.CD-13-0011>
PMID:[23650262](https://pubmed.ncbi.nlm.nih.gov/23650262/)
9. Tang Z, Lin MG, Stowe TR, Chen S, Zhu M, Stearns T, Franco B, Zhong Q. Autophagy promotes primary ciliogenesis by removing OFD1 from centriolar satellites. *Nature*. 2013; 502:254–57.
<https://doi.org/10.1038/nature12606>
PMID:[24089205](https://pubmed.ncbi.nlm.nih.gov/24089205/)
10. Chang CH, Bijian K, Wernic D, Su J, da Silva SD, Yu H, Qiu D, Asslan M, Alaoui-Jamali MA. A novel orally available seleno-purine molecule suppresses triple-negative breast cancer cell proliferation and progression to metastasis by inducing cytostatic autophagy. *Autophagy*. 2019; 15:1376–90.
<https://doi.org/10.1080/15548627.2019.1582951>
PMID:[30773992](https://pubmed.ncbi.nlm.nih.gov/30773992/)
11. Vera-Ramirez L, Vodnala SK, Nini R, Hunter KW, Green JE. Autophagy promotes the survival of dormant breast cancer cells and metastatic tumour recurrence. *Nat Commun*. 2018; 9:1944.
<https://doi.org/10.1038/s41467-018-04070-6>
PMID:[29789598](https://pubmed.ncbi.nlm.nih.gov/29789598/)
12. Yeong J, Lim JC, Lee B, Li H, Ong CC, Thike AA, Yeap WH, Yang Y, Lim AY, Tay TK, Liu J, Wong SC, Chen J, et al. Prognostic value of CD8 + PD-1+ immune infiltrates and PDCD1 gene expression in triple negative breast cancer. *J Immunother Cancer*. 2019; 7:34.
<https://doi.org/10.1186/s40425-019-0499-y>
PMID:[30728081](https://pubmed.ncbi.nlm.nih.gov/30728081/)
13. Pingoud-Meier C, Lang D, Janss AJ, Rorke LB, Phillips PC, Shalaby T, Grotzer MA. Loss of caspase-8 protein expression correlates with unfavorable survival outcome in childhood medulloblastoma. *Clin Cancer Res*. 2003; 9:6401–9.
PMID:[14695141](https://pubmed.ncbi.nlm.nih.gov/14695141/)
14. Tu L, Liu Z, He X, He Y, Yang H, Jiang Q, Xie S, Xiao G, Li X, Yao K, Fang W. Over-expression of eukaryotic translation initiation factor 4 gamma 1 correlates with tumor progression and poor prognosis in nasopharyngeal carcinoma. *Mol Cancer*. 2010; 9:78.
<https://doi.org/10.1186/1476-4598-9-78>
PMID:[20398343](https://pubmed.ncbi.nlm.nih.gov/20398343/)
15. Shen J, Li M, Min L. HSPB8 promotes cancer cell growth by activating the ERK-CREB pathway and is indicative of a poor prognosis in gastric cancer patients. *Oncol Rep*. 2018; 39:2978–86.
<https://doi.org/10.3892/or.2018.6376>
PMID:[29693129](https://pubmed.ncbi.nlm.nih.gov/29693129/)
16. Benavente S, Vergés R, Hermosilla E, Fumanal V, Casanova N, García A, Ramón Y Cajal S, Giralt J. Overexpression of phosphorylated 4E-BP1 predicts for tumor recurrence and reduced survival in cervical carcinoma treated with postoperative radiotherapy. *Int J Radiat Oncol Biol Phys*. 2009; 75:1316–22.
<https://doi.org/10.1016/j.ijrobp.2009.01.004>
PMID:[19362787](https://pubmed.ncbi.nlm.nih.gov/19362787/)
17. Afentakis M, Dowsett M, Sestak I, Salter J, Howell T, Buzdar A, Forbes J, Cuzick J. Immunohistochemical BAG1 expression improves the estimation of residual risk by IHC4 in postmenopausal patients treated with anastrozole or tamoxifen: a TransATAC study. *Breast Cancer Res Treat*. 2013; 140:253–62.
<https://doi.org/10.1007/s10549-013-2628-5>
PMID:[23852223](https://pubmed.ncbi.nlm.nih.gov/23852223/)
18. De Iulius F, Salerno G, Taglieri L, Lanza R, Cardelli P, Scarpa S. Circulating neuregulin-1 and galectin-3 can be prognostic markers in breast cancer. *Int J Biol Markers*. 2017; 32:e333–36.
<https://doi.org/10.5301/ijbm.5000262>
PMID:[28430337](https://pubmed.ncbi.nlm.nih.gov/28430337/)
19. Boccellino M, Pinto F, Ieluzzi V, Giovane A, Quagliuolo L, Fariello C, Coppola M, Carlucci A, Santini M, Ferati K, Bexheti-Ferati A, Giordano A, Di Domenico M. Proteomics analysis of human serum of patients with non-small-cell lung cancer reveals proteins as diagnostic biomarker candidates. *J Cell Physiol*. 2019; 234:23798–806.
<https://doi.org/10.1002/jcp.28948> PMID:[31180588](https://pubmed.ncbi.nlm.nih.gov/31180588/)
20. Yang SW, Ping YF, Jiang YX, Luo X, Zhang X, Bian XW, Yu PW. ATG4A promotes tumor metastasis by inducing the epithelial-mesenchymal transition and stem-like properties in gastric cells. *Oncotarget*. 2016; 7:39279–92.
<https://doi.org/10.18632/oncotarget.9827>
PMID:[27276686](https://pubmed.ncbi.nlm.nih.gov/27276686/)

SUPPLEMENTARY MATERIALS

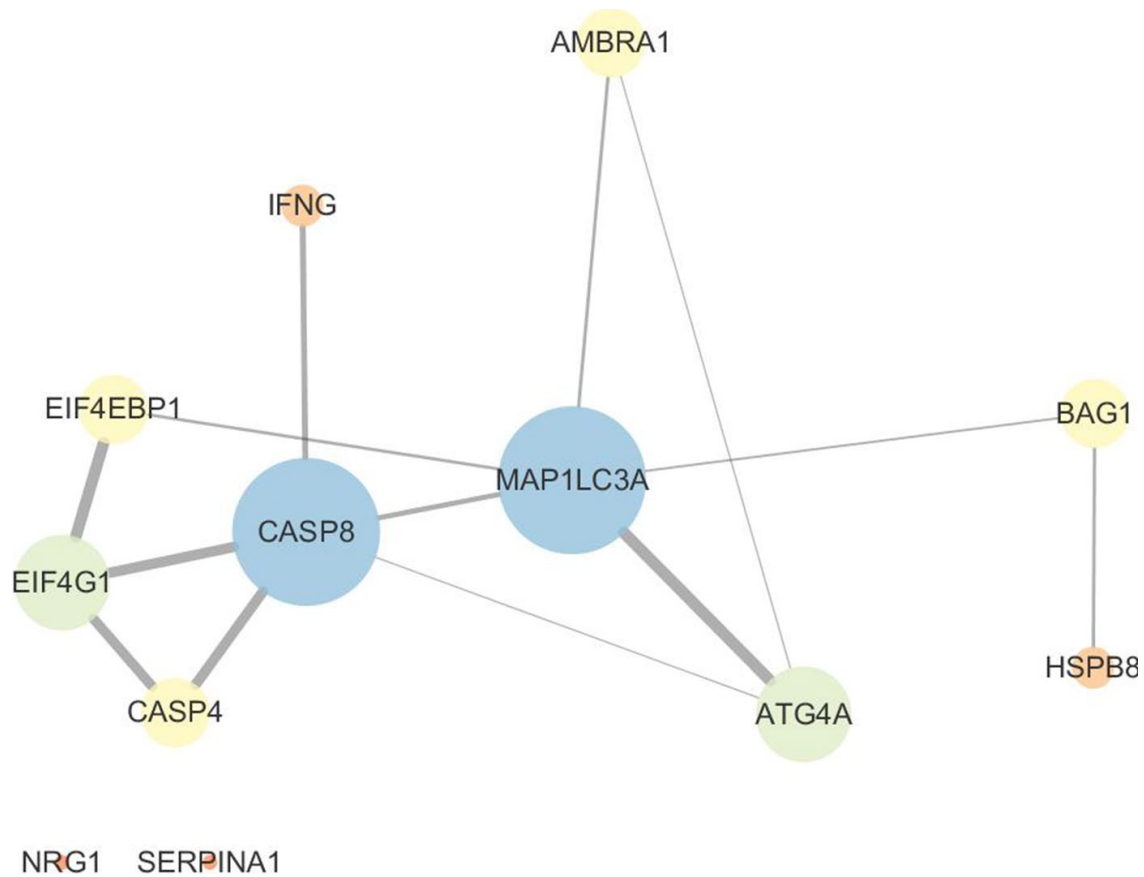
Supplementary Figures



Supplementary Figure 1. Flowchart for identifying 12 ARGs signature associated with breast cancer survival.



Supplementary Figure 2. The expression patterns of these 12-ARGs.



Supplementary Figure 3. Protein-protein interaction (PPI) network of 12-ARGs.