

Complete Identification of Autophagy-Related lncRNA Prognostic Signature for Breast Cancer

Wenxi Wang

Xiangya Hospital Central South University

Na Li

Xiangya Hospital Central South University Department of Oncology

Lin Shen

Xiangya Hospital Central South University

Qin Zhou

Xiangya Hospital Central South University

Zhanzhan Li (✉ lizhanzhan@csu.edu.cn)

Xiangya Hospital Central South University <https://orcid.org/0000-0002-7965-7608>

Liangfang Shen

Xiangya Hospital Central South University

Research Article

Keywords: breast cancer, autophagy-related lncRNA, risk score, prognosis, immunity

Posted Date: August 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-347298/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: Breast cancer (BC) has a relatively high morbidity and mortality for women. The research about BC prognosis is significant. Autophagy is an essential process for tumor progression, which could play its role with lncRNA, a kind of ncRNA that have regulatory roles in multiple tumors. Therefore, constructing an autophagy-related prognostic model for breast cancer is meaningful.

Methods: We download data from the TCGA and HADb. Pearson correlation analysis was performed to excavate autophagy-related lncRNA. Then with gene expression difference analysis, etc. we explored the relationship between clinical features and the signature. We applied Cytoscape as well as KEGG, etc. to explore expression condition. And the autophagy status of our signature was investigated by GSEA, etc. We also searched the immune distinction by CIBERSORTx to extend our study and preliminarily verified our study in the end.

Results: Firstly, we got an independent autophagy-related lncRNA prognostic model, by which BC patients were divided into high- and low-risk groups. We found that the OS of high-risk group is significantly lower than that of low-risk group, which was identical to those within various clinical subgroups. Then, the KEGG and GO analysis enriched several pathways including autophagy. PCA and GSEA analysis demonstrated the autophagy status. Several distinguishing immune cell types in separated groups were revealed by immunity analysis. Then the verification in the end proved the feasibility of our signature.

Conclusion: In this study, we acquired an independent autophagy-related lncRNA signature involving 12 lncRNAs, which contributes to the prediction of prognosis of BC patients.

Background

According to WHO, cancer is the second leading cause of death globally, and is responsible for an estimated 9.6 million deaths in 2018. Among these, breast cancer (BC) has the second morbidity and the fifth mortality, therefore is detrimental to the public health especially for women. The risk factors of breast cancer include being female, older age, family history, etc.[Sun, et al. 2017] About 5–10% patients are related to inherited or mutated genes, for example, BRCA, p53, PTEN and ESR1[Gage, et al. 2012, Costa, et al. 2020, Duffy, et al. 2018, Dustin, et al. 2019, Xiao, et al. 2020]. And the mutation type would influence the classification, treatment and prognosis of BC[Corso, et al. 2018, Gao, et al. 2019]. Therefore, the research at the gene and molecule level is pivotal and reasonable. Although there have been systematic diagnosis, treatment and prognosis utilized in clinic, reliable risk assessment model is still needed to evaluate the survival prognosis of patients and to conduct further treatment.

Autophagy is known as a kind of mechanism that decompose and utilize the unnecessary or dysfunctional cellular components[Klionsky 2008, Mizushima and Komatsu 2011]. Till now, three pathways of autophagy have been observed, which are macroautophagy, microautophagy and chaperone-mediated autophagy (CMA). All those pathways rely on lysosome but are different in transporting cellular components[Mizushima, et al. 2002, Lee, et al. 2012]. Autophagy is executed by

autophagy-related genes and its process involves abundant signaling pathways. The function of autophagy includes cell metabolism [Reggiori and Klionsky 2002, Poillet-Perez and White 2019], immune response [Cui, et al. 2019, Wu and Adamopoulos 2017], repair of damage [Boya, et al. 2018] and programmed cell death [Ou, et al. 2017]. Thus, autophagy is crucial to the basal homeostasis [He, et al. 2012]. Furthermore, it also plays important roles in cancer, not only as a suppressor, but also a promoter under different circumstances [Mowers, et al. 2018, Levy, et al. 2017]. Hence, the research about autophagy is quite necessary.

Long non-coding RNA (lncRNA) is one kind of non-coding RNAs (ncRNA) that has various functions in almost every activity of cells. More and more researches have shown that lncRNA plays essential roles in cancer progression, such as cell proliferation, apoptosis, metastasis and immune inhibition [Zhao, et al. 2019, Zhao, et al. 2018, Wei and Wang 2017]. With the discovery of autophagy, a large number of lncRNAs were found to be related with the process [Zhou, et al. 2020, Liu, et al. 2018]. Thus, many scientists have shown a deep interest in this field and the signaling pathways about it is being completed gradually. There have been a large number of researches demonstrating the biomarker function of lncRNA. But the majority of these research claims single molecule as potential predicted marker. There are two articles about breast cancer and autophagy-related lncRNA prognostic model. But the research did not involve in-depth and extended exploration. So, in this study, we not only applied data in TCGA database, filtering autophagy-related lncRNAs and construct a predictive signature to be a more systematic biomarker of breast cancer, but also verify its feasibility, universality and correlation with clinical features as well as extent our results to immune status. We believe that with the signature, there could be a new significant way to evaluate the prognosis of BC patients.

Methods And Materials

Acquisition of clinical and transcriptome data

We searched and downloaded the data of transcriptome profiling and clinical information of breast cancer patients in The Cancer Genome Atlas (TCGA) database. Fragments per kilobase of transcript per million (FPKM) from gene expression quantification were screened for our research. Then we separated and rearranged the data of lncRNAs and mRNAs. In addition, clinical information including age, stage, tumor depth (T) phase, lymph node (N) phase, metastasis (M) phase and follow-up outcomes were downloaded and settled for further research. By eliminating some incomplete data, we finally got the chart could be used in the research.

Selection of autophagy-related lncRNA

The autophagy-related gene expression data were downloaded from Human Autophagy Database (HADb). To select autophagy-related lncRNA, we performed Pearson correlation analysis and found autophagy co-expressed lncRNA with the correlation efficiency ($|R|$ greater than 0.3 and P value less than 0.001). The statistical software R (version 3.6.0) were applied. And with the processed transcriptome

data, we have got the information of lncRNA in BC patients. Then we picked the expression data of autophagy-related lncRNA for further research.

Construction of prognostic signature

Univariate Cox regression was used to explore the relationship between single autophagy-related lncRNA expression and prognosis of breast cancer patients with P value less than 0.05. Besides, we could get HR value to differentiated the dangerous or protective character of the lncRNA. Then multivariate Cox regression was performed to exclude interaction of lncRNAs and construct the predicted model for BC. And not only HR and its section, but also regression coefficients (β) and P values were demonstrated. So, we could get the risk scores of every patient with the expression of the lncRNAs and the formula exhibited as follows: risk score = coef(lncRNA1) \times expr (lncRNA1) + coef(lncRNA2) \times expr (lncRNA2) +...+ coef(lncRNAn) \times expr (lncRNAn).

Prediction analysis of the prognosis signature

According to the median score, the patients were separated into high-risk group and low-risk group. Using the Kaplan-Meier analysis, we depicted and compared the survival curve of the two groups. Also, receiver operating characteristic (ROC) curve was depicted and area under curve were calculated to evaluate the efficiency of the signature. In addition, by Cox regression, we explored the influence of several clinical factors and risk score for survival. Age, gender, T, N, M phase were involved. Because the lack of data of male patients, we only got the result of female. And we also detected the independency of the risk score by Cox regression. After that, we divided the samples into different subgroups according to the separation of several clinical factors and depicted the survival curve of those subgroups respectively. Then we searched the differentiated risk scores between the subgroups as well as the relationship among every target lncRNA and clinical factors by heat map and Student's t-test, proving the feasibility and universality of our signature.

Establishment of co-expression network and pathway analysis.

Expression data of autophagy-related lncRNA and mRNA were input into Cytoscape to construct co-expression network of lncRNA-mRNA. Then, a Sankey diagram was depicted to visualize the relationship among expression of autophagy gene, autophagy-related lncRNA and risk factors. Besides, Kyoto Encyclopedia Genes and Genomes (KEGG) and Gene Ontology (GO) analysis were used to excavate signaling pathways and the function of the target lncRNA. Principal component analysis (PCA) was performed to explore the differences of high- and low-risk groups based on autophagy-related lncRNA sets and genome-wide expression profiles. In addition, Gene set enrichment analysis (GSEA) have observed enrichment condition of gene sets in high- and low-risk groups respectively. The version of GSEA is v4.1.0, with $P < 0.05$.

Analysis of mutant profiling

We used water plot to investigate mutant information of BC patients and compared the results between high- and low-risk groups. The variant classification, type, SNV class were calculated to find the rule of autophagy-related gene mutant, and gene cloud was depicted to visualize the mutant frequencies.

Differences of immune status in Groups divided by risk score

By CIBERSORTx (<https://cibersortx.stanford.edu/>), the data from TCGA were input. Then the predictive results of various infiltrated immune cells were used to compare the immune difference between the high- and low-risk groups. The analysis could investigate the association of the autophagy-related lncRNA signature and immunity, providing a new way of research.

Verification of target lncRNA in tumor and normal tissues

Gene Expression Profiling Interactive Analysis compared the lncRNA expression data of tumor tissues and normal tissues. Then with heat map and box plot, we could clearly get the significantly distinct lncRNA and primarily verify the feasibility of our signature.

Statistical Analysis

R software was used for our statistical analysis. Pearson correlation analysis contributed to the relationship analysis. Kaplan-Meier analysis and log-rank test were performed for survival curve while Student's t-test for correlation among lncRNA and several clinical factors. Univariate and multivariate Cox Regression were implemented to construct the prognostic model and evaluate the independency of clinical factors and risk score. The significance was confirmed as $P < 0.05$.

Results

Acquisition of TCGA data and construction of a prognostic signature

1222 samples were downloaded from TCGA database, in which 1109 belonged to BC patients and 113 were from normal tissues. By excluding incomplete samples, we kept 1089 patients. A total of 14142 lncRNA data were collected from tumor tissues. According to the Human Autophagy Database, we downloaded autophagy regulatory genes and screened 1270 autophagy related lncRNAs by pearson correlation analysis. Then, by comparing overall survival (OS), the univariate Cox regression models identified 51 autophagy-related lncRNAs (Table S1, Fig. 1a), 12 of which were finally screened to construct prognostic signature by a multivariate Cox regression analyses (Table 1, Fig. S1). Among these, lncRNA AC008105.3, Z68871.1 and AL109811.3 are risk factors for the prognosis. While AL731571.1, AC061992.1, AL136368.1, MAPT-AS1, SEMA3B-AS1, LINC01871, ST7-AS1, AL122010.1 and AC121761.2 are protective factors. With the accuired regression coefficients, the confirmed signature formula is: risk score = $(0.286 * \text{ExpAC008105.3}) + (0.461 * \text{ExpZ68871.1}) + (-0.452 * \text{ExpAL731571.1}) + (-0.270 * \text{ExpAC061992.1}) + (-0.515 * \text{ExpAL136368.1}) + (-0.285 * \text{ExpMAPT-AS1}) + (-0.078 * \text{ExpSEMA3B-}$

AS1) + (-0.333*ExpLINC01871) + (-0.322*ExpST7-AS1) + (-0.292*ExpAL122010.1) + (0.197*ExpAL109811.3) + (-0.377*ExpAC121761.2). Then we calculated and recorded the risk score of every sample. And according to the median risk score, the patients were divided into high- (n = 544) and low- risk groups (n = 545) (Fig. 1c). Then by assembling the patients' survive time and drawing survival curve, we found the OS of the high risk group is significantly worse than the low risk group (Fig. 1b). Meanwhile, as the risk score increases, the mortality of the patient increases to some extent (Fig. 1d). To verify the significance of the signature, we compared some clinical factors including age, gender, stage, T, N and M phase, and risk scores, respectively. By univariate Cox regression (Fig. e), we found that risk score is significant prognostic factors related to worse OS, while multivariate Cox regression (Fig. f) showed only age and risk score were independent prognostic factors. Besides, the area under the ROC curve (AUC) equals to 0.882, which indicates its great predictive effect. All those results revealed that the signature is meaningful and is able to predict the prognosis of BC.

Table 1
Multivariate Cox regression of autophagy-related lncRNAs

lncRNAs	Beta	SE	HR	HR.95L	HR.95H	P value
AC008105.3	0.286	0.181	1.33	0.93	1.90	0.114
Z68871.1	0.461	0.167	1.59	1.14	2.20	0.006
AL731571.1	-0.452	0.201	0.64	0.43	0.94	0.024
AC061992.1	-0.270	0.143	0.76	0.58	1.01	0.059
AL136368.1	-0.515	0.252	0.60	0.37	0.98	0.041
MAPT-AS1	-0.285	0.109	0.75	0.61	0.93	0.009
SEMA3B-AS1	-0.078	0.036	0.92	0.86	0.99	0.029
LINC01871	-0.333	0.092	0.72	0.60	0.86	0.000
ST7-AS1	-0.322	0.194	0.72	0.50	1.06	0.097
AL122010.1	-0.292	0.106	0.75	0.61	0.92	0.006
AL109811.3	0.197	0.088	1.22	1.02	1.45	0.026
AC121761.2	-0.377	0.205	0.69	0.46	1.02	0.066

Stratified Analysis Of Prognosis Signature Based On Clinical Features

Clinical characteristics are important factors to evaluate the patients' condition and predict the OS. So, we next investigated the relationship between risk scores and survival time under the same clinical condition. The clinical information acquired from TCGA includes age, gender, stage, T, N and M phase. Then we

divided the clinical features into several subgroups, which are age (≤ 65 and > 65), stage (I&II and III&IV), gender (only female was investigated because of the lack of data for male), T (T1-2 and T3-4), N (N0 and N1-3), and M (M0 and M1) phase, respectively. The results revealed that all the subgroups showed a significant superior survival time for low-risk groups compared to high-risk groups with P values < 0.05 (Fig. 2). So, our findings indicated that the signature is suitable for multiple clinical factors, thus is universal under different circumstances and important in prognosis of BC.

Associated Relation Between The Clinical Factors And Autophagy-related Lncrnas

According to the assigned subgroups mentioned above, we further explored the expression of the 12 autophagy-related lncRNAs and the distribution of the clinical factors in high- and low-risk groups, which were represented by heat map (Fig. 3a). We found that between the two groups, stage ($p < 0.001$), N ($p < 0.05$) phase, M ($p < 0.05$) phase and survival condition ($p < 0.001$) were distinguished. Then by comparing the risk scores within the corresponding subgroups, we discovered that the patients with lower age (Fig. 3b), earlier T phase (Fig. 3c) or earlier N phase (Fig. 3d) generally had lower risk scores with P values equals to 0.0048, 0.026 and 0.0041, respectively, in accordance with the fact and our prediction. However, there was no significant difference between the subgroups of M phase. Hence, we could draw a brief conclusion that in BC patients, the autophagy-related lncRNA signature may be associated with some clinical characteristics.

To further explore the relationship between the expression of the respective 12 autophagy-related lncRNAs and clinical factors, we ulteriorly divided some clinical characteristics including stage (I, II, III and IV), T (1,2,3 and 4) and N (0,1,2 and 3) phase. Then we recorded the expression level of the lncRNAs in different subgroups and found that lncRNA AL731571.1, LINC01871, ST7-AS1 and SEMA3B-AS1 expressed differently in the subgroups about age (Fig. 4a). In regard to the stage, lncRNA AL136368.1, MAPT-AS1 and AL122010.1 were significantly distinguishing (Fig. 4b). And for T phase, lncRNA AL136368.1, MAPT-AS1, SEMA3B-AS1, LINC01871, ST7-AS1, AL122010.1 and AL109811.3 were associated with the tumor depth of BC (Fig. 4c). In addition, lncRNA SEMA3B-AS1 was correlated with lymph node metastasis (Fig. 4d). While no lncRNA was found to express discrepantly in the groups divided by M phase (Fig. 4e). In sum, lncRNA AL731571.1, AL136368.1, MAPT-AS1, SEMA3B-AS1, LINC01871, ST7-AS1, AL122010.1 and AL109811.3 are approximately suppressors for BC, which are consistent with the results before. Therefore, these lncRNAs were proved to be correlated with some clinical factors and thus might be available biomarkers of BC.

Construction Of Lncrna-mrna Co-expression Network And Function Analysis

Among all the functions of lncRNA, regulatory role in gene expression is one of the most important, also known as lncRNA-mRNA co-expression. As is shown in Fig. 5a, 27 mRNAs were found to be correlated with the autophagy-related lncRNAs, 19 of which were regulated by more than one autophagy-related lncRNAs. In addition, a Sankey diagram demonstrated the corresponding relationship among lncRNA, mRNA and risk factors (Fig. 5b). The co-expression network and Sankey diagram both showed that the center and major of the signature could be Z68871.1, AL109811.3, AC008105.3 and LINC01871. Except for LINC01871, the others are risk factors, thus we could roughly recognize that the signature is in direct proportion to risk, which is consistent with the nature of our signature. Then, to further search the correlated pathways, we applied and visualized KEGG enrichment analysis (Fig. 5c), which revealed that autophagy, neurodegeneration and shigellosis were the top 3 enriched pathways. Almost 40% of the co-expressed mRNAs located in autophagy-related signaling pathway. There were 88 pathways enriched in autophagy signaling pathway, including 65 pathways that exist in animal (Table S3). Besides, by GO analysis, we got 58 genes. And the main genes included ubiquitin-like protein ligase binding, ubiquitin protein ligase binding and protein serine/threonine kinase activity, etc. (Fig. S2, Table S2), which all act on cell autophagy process.

The essence of the samples compared in high- and low-risk group.

Since KEGG and GO analysis had found the signaling pathways regulated by autophagy-related lncRNAs, further research should be done to confirm the essence of the correlated genes. So, with the autophagy-related lncRNA set and whole gene expression profiles, we have applied principal component analysis. According to the distribution of the gene in high- and low-risk groups presented in Fig. 6a and 6b respectively, significant difference was observed between the two groups, showing a distinguishing autophagy-related gene profiles between the two groups. However, with genome-wide expression profiles, PCA demonstrated no distinct separation of the groups (Fig. 6c). Then to search for the pathways of the differentially expressed gene sets, we used GSEA analysis and found 43 enriched gene sets ((cutoff NOM P value < 0.05), within which 12 were enriched in high-risk group and 31 in low-risk group. Then we enumerated the top 10 gene sets of the two groups respectively (Fig. 6d, 6e) and observed extracellular matrix (ECM) receptor interaction, TGF- β signaling pathway, adhesion junction, focal adhesion and o-glycan biosynthesis were enriched signaling pathways in high-risk group and autoimmune thyroid disease, arachidonic acid metabolism, intestinal immune network for IgA production, alpha-Linolenic acid metabolism and ether lipid metabolism in high-risk group.

Mutation profiles difference in high- and low-risk groups

Since the research about expression of the autophagy-related lncRNA and pathway enrichment analysis have been finished, we wanted to deeply explore the diversity in mutation profile. So, 498 patients have been selected to high-risk group while 475 patients to low-risk group. With waterfall plots, we could get the information of mutated genes in each sample, as well as the proportion and type of mutation in total. (Fig. 7a, b) With more statistics of the data, we found that in both groups, missense mutation accounted for the largest fraction of mutation. Meanwhile, single nucleotide polymorphism (SNP) was most

common type of mutation and the transversion C>T, C>G, C>A were the top 3 SNV classes. More details of variants per sample and variant classification summaries in both groups were shown simultaneously. We could find that the median variants per sample in high-risk group is 32, comparing 26 in low-risk group (Fig. 7c, d), indicating that more mutants occur in high-risk group. According to the mutation rate, the top 10 mutated genes in high-risk group are TP53, PIK3CA, TTN, MUC16, GATA3, MAP3K1, CDH1, KMT2C, MUC4 and USH2A, and in low-risk group are PIK3CA, TP53, TTN, CDH1, MAP3K1, MUC16, GATA3, MUC4, KMT2C and PTEN. Most top 10 mutated genes in both groups are the same except for USH2A in high-risk group and PTEN in low-risk group. Finally, we used gene cloud demonstrating the frequency of the mutation, making it more visualized according to the size of the gene (Fig. 7e, f).

Immunology status based on risk score.

Immune has become the hot spot ever since it was discovered. It is known to have close relationship with multiple tumor progression and to play crucial roles in tumor cell death escape. The GSEA analysis has enriched several immune-related pathways, so we compared immune status between separated risk groups by CIBERSORTx, trying to figure out the impact of the autophagy-related signature on immunity. Heat map (Fig. 8a) and box plot (Fig. 8b) revealed that 9 of the infiltrated immune cell types had significant difference between high- and low-risk groups. Among them, we found that in high-risk group, the majority of immune cells increased, including CD8⁺ T cell (P <0.001), activated CD4⁺ B memory cell (P <0.001), macrophages M1 (P <0.001), Treg cell (p <0.01), T follicular helper cell (P <0.05), activated NK cell (P <0.05) and resting dendritic cell (P <0.05), accompanied with higher immune scores. Those cells most function as suppressors for tumor. Thus, it is revealed that in high-risk group, immune system may be more vibrant. Interestingly, macrophages M2 had higher amount in low-risk group, while macrophages M1 showed more numbers in high-risk group. Macrophage M2 is recognized as tumor associated macrophages that promotes tumor growth while M1 could prevent autophagy of cells. The results may be in accordance with the possible activated immunity of the patients with higher risk, fighting for cancer.

Feasibility verification of the 12 autophagy-related lncRNAs in tumor and normal tissues.

Until now, we have almost finished the construction and exploration of the prognostic model. To primarily verify the function of the 12 target lncRNAs in tumor, we compared the expression of those lncRNAs between 1109 tumor tissues and 113 normal tissues. By heat map (Fig. 9a) and box plot (Fig. 9b), we found the differentially expressed lncRNA Z68871.1, AL731571.1, AC061992.1, AL136368.1, LINC01871, ST7-AS1, AL122010.1, AL109811.3 and AC121761.2. Among them, lncAC061992.1 was significantly highly expressed in tumor tissues, and the rest in normal tissues by mean expression levels. However, our signature recognized lncAL109811.3 as a risk factor while AC061992.1 as protective factor in prognosis. Hence, we could find that lncAC061992.1 is associated with the occurrence of breast cancer but is a protector for the prognosis while lncAL109811.3 acts an opposite way.

Discussion

Currently, there have been comprehensive and standardized procedures for the diagnosis and treatment of breast cancer. Although the clinical classification of TNM stage and molecular subtype is relatively clear and widely used, further research about prediction of survival for BC patients is still needed. Therefore, the establishment of prognosis model has positive significance for the prediction of clinical survival. Many scientists have reported the function of lncRNA to the oncogenesis as well as development of breast cancer. For example, Sun, et al. [Sun, et al. 2020] reported that low expression of lncRNA-MALAT1 contributed to high 5-year overall survival of BC patients and is proved to be an independent prognosis factor. There is another research about lncRNA-miRNA-mRNA axis, claiming the ceRNA function of lncPCNAP1 in BC [Yu, et al. 2020]. And autophagy plays a crucial role in removing useless or even harmful wastes from cells and recycling energy. Previous studies revealed that autophagy may interact with lncRNA [Gu, et al. 2018], circRNA [Liang, et al. 2020] and cancer antigen [Liu, et al. 2018] and promotes the progression of BC. In this study, bioinformatics method was used to conduct a series of information statistics and analysis, thus an autophagy-related prognostic model of breast cancer was established, which has certain significance.

In this study, we firstly collected patients' information through TCGA database and acquired all the lncRNA recorded. Then, autophagy-related genes data were obtained from HADb. By Pearson correlation analysis and Cox regression, we filtered 12 autophagy-related lncRNAs which were associated with OS of BC. So prognostic formula was acquired with the regression coefficients. Then with the univariate and multivariate Cox regression analysis, we finally constructed an independent prognostic model.

Since the signature was established, we further detected its relation with some clinical characteristics. Based on the grouping of patients' risk scores and the stratification of clinical features, we can not only get the OS difference of the BC patients in high- and low-risk groups, but also find the survival probability of the high- and low-risk groups and distinctions of risk score under different clinical subgroups. Moreover, we also observed that there were different expressions of several autophagy-related lncRNAs among clinical subgroups, in which T phase subgroups have the most associated lncRNAs while M phase subgroups contained no significant lncRNA. In addition, lncRNA AL136368.1, MAPT-AS1, SEMA3B-AS1, LINC01871, ST7-AS1 and AL122010.1 have been reported to be related with more than one clinical characteristic. In all, the results indicated that our signature is associated with certain clinical features and thus might be used to predict the patients' clinical condition.

Co-expression of lncRNA-mRNA is another key promoting the research process. We not only got the network of co-expressed lncRNA-mRNA and risk type, but also acquired signaling pathway enriched by KEGG and GO analysis. We found 4 main lncRNAs, Z68871.1, AL109811.3, AC008105.3 and LINC01871, 3 of which were risk factors for BC thereby were the basis of the signature. In addition, several pathways were enriched, and autophagy is the top one, indicating the close relation between our signature and the autophagy process. Moreover, the GSEA analysis screened 12 and 31 signaling pathway in high- and low-risk groups respectively. The pathways were correlated with metabolism, cell adhesion, autophagy and immune, etc. Then, PCA analysis demonstrated that it might be the discrepant autophagy status divided

by the signature that caused the different OS between patients in high- and low-risk groups. All the findings have indirectly proved the autophagy property of our signature.

With that, the mutant information was explored and represented in Fig. 7. With gene sets analysis, we got the top 3 mutant genes in both groups, TP53, PIK3CA and TTN, which occupied more than 70% mutant in total and all are proved to have close relationship with BC progression[Duffy, Synnott and Crown 2018, Martinez-Saez, et al. 2020, Feng, et al. 2020]. The research not only demonstrated the mutant details, but also the frequencies of the gene sets, making it easier to understand the distinct gene profiles in high- and low-risk groups.

Distinguished from other similar research, we extended our study by comparing the abundance of multiple immune cells in high- and low-risk groups. We have found several distinct cell types, most of which were immune activator, observed increased in high-risk groups. With the research, we could draw a brief conclusion that the BC patients with high-risk scores were likely to have a more abundant immune cell types and higher immune scores, which may be caused by higher burden of tumor induced activation of the immune system.

Ultimately, verification is necessary. So, we investigated the expression of target lncRNAs between tumor tissues and normal tissues and found that 8 out of the 12 autophagy-related lncRNAs showed significant difference. From the research we also found the risk factors for prognosis might be the protector for tumor occurrence.

There are two articles about constructing autophagy-related lncRNA prognostic model, one of which was from Li, etc.[Li, et al. 2021], whose study is relatively brief, only involving the construction and exploration of prognostic model, pathway analysis and primary verification. The other is reported by zhang, etc. [Zhang, et al. 2020]. Although the research consists more experiments, the study did not further detect the survival differences between high- and low-risk groups within various clinical subgroups, or carry out preliminary verification of the obtained signature in the end. So, the experiment is also incomplete. Besides, by comparison, it is found that the AUC of our signature is 0.882, which has a better prediction efficiency than that of 0.788. Surprisingly, compared with the 9 autophagy-related lncRNA obtained in the paper, we only shared two jointly screened lncRNA, MAPT-AS1 and ST7-AS1, which were protective factors for BC patients in both of the prognostic models. So, the two lncRNAs might play crucial roles in the prognosis of BC.

This study is aimed at the risk stratification of patients in line with autophagy-related lncRNA expression level to predict their survival prognosis. This paper not only is comprehensive, but also explored the immune status of patients in high- and low-risk groups, which may be of positive significance for future studies on the association between autophagy and immunity. However, there are still some improvable research in this paper. For example, our study did not investigate the difference among molecular subtyping of BC. In addition, there is a lack of multidimensional proof for the signature verification in this paper, and the few of data on male BC patients may lead to insufficient representativeness of some studies, so more data are needed to further verify the results of these experiments.

Not only lncRNA, there are plenty of non-coding RNAs that could play essential regulatory roles in autophagy process in tumors including but not limited to breast cancer [Liang, Ling, Mehrpour, Saw, Liu, Tan, Tian, Zhong, Lin, Luo, Lin, Li, Zhou, Hamai, Codogno, Li, Song and Gong 2020, Huang, et al. 2019, Li, et al. 2018]. And in addition to tumor prognosis, they may also contribute to cancer screening, treatment and evaluation. Therefore, there is a lot of research space and potential clinical application value for this study.

Conclusion

Our study acquired an independent autophagy-related lncRNA signature, which had universality among various clinical features. The signature involves 12 lncRNA and plays significant roles in predicting prognosis of breast cancer. The results might have a positive effect for further research.

Abbreviations

breast cancer (BC), chaperone-mediated autophagy (CMA), Long non-coding RNA (lncRNA), non-coding RNAs (ncRNA), The Cancer Genome Atlas (TCGA), fragments per kilobase of transcript per million (FPKM), tumor depth (T), lymph node (N) phase, metastasis (M) phase, Human Autophagy Database (HADb), receiver operating characteristic (ROC), Kyoto Encyclopedia Genes and Genomes (KEGG), Gene Ontology (GO), Principal component analysis (PCA), Gene set enrichment analysis (GSEA), overall survival (OS); area under the ROC curve (AUC), extracellular matrix (ECM), single nucleotide polymorphism (SNP)

Declarations

Funding

This study was supported by the National Natural Science Foundation of China (No. 82003239), Hunan Province Natural Science Foundation (Youth Foundation Project) (NO.2019JJ50945), and the Science Foundation of Xiangya Hospital for Young Scholar (NO. 2018Q012).

Competing interests

The authors declare that they have no competing interests

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the TCGA and HADb repository, [<https://portal.gdc.cancer.gov/>

<http://www.autophagy.lu/clustering/index.html>]

Authors' contributions

LZZ designed this study and directed the research group in all aspects, including planning, execution, and analysis of the study. WWX drafted the manuscript. LN, ZQ and SL collected the data. LZZ provided the statistical software, performed the data analysis, arranged the Figures and Tables. LZZ and SLF revised the manuscript. All authors have read and approved the final version of the manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Acknowledgements

Not applicable.

References

1. Y.S. Sun, et al., Risk Factors and Preventions of Breast Cancer, *Int J Biol Sci*, 13 (2017) 1387-1397<https://doi.org/10.7150/ijbs.21635>
2. M. Gage, et al., Translational advances regarding hereditary breast cancer syndromes, *J Surg Oncol*, 105 (2012) 444-451<https://doi.org/10.1002/jso.21856>
3. C. Costa, et al., PTEN Loss Mediates Clinical Cross-Resistance to CDK4/6 and PI3Kalpha Inhibitors in Breast Cancer, *Cancer Discov*, 10 (2020) 72-85<https://doi.org/10.1158/2159-8290.CD-18-0830>
4. M.J. Duffy, et al., Mutant p53 in breast cancer: potential as a therapeutic target and biomarker, *Breast Cancer Res Treat*, 170 (2018) 213-219<https://doi.org/10.1007/s10549-018-4753-7>
5. D. Dustin, et al., ESR1 mutations in breast cancer, *Cancer*, 125 (2019) 3714-3728<https://doi.org/10.1002/cncr.32345>
6. G. Xiao, et al., Gain-of-Function Mutant p53 R273H Interacts with Replicating DNA and PARP1 in Breast Cancer, *Cancer Res*, 80 (2020) 394-405<https://doi.org/10.1158/0008-5472.CAN-19-1036>
7. G. Corso, et al., Prognosis and outcome in CDH1-mutant lobular breast cancer, *Eur J Cancer Prev*, 27 (2018) 237-238<https://doi.org/10.1097/CEJ.0000000000000405>
8. C. Gao, et al., SNP mutation-related genes in breast cancer for monitoring and prognosis of patients: A study based on the TCGA database, *Cancer Med*, 8 (2019) 2303-2312<https://doi.org/10.1002/cam4.2065>
9. D.J. Klionsky, Autophagy revisited: a conversation with Christian de Duve, *Autophagy*, 4 (2008) 740-743<https://doi.org/10.4161/auto.6398>
10. N. Mizushima, M. Komatsu, Autophagy: renovation of cells and tissues, *Cell*, 147 (2011) 728-741<https://doi.org/10.1016/j.cell.2011.10.026>

11. N. Mizushima, et al., Autophagosome formation in mammalian cells, *Cell Struct Funct*, 27 (2002) 421-429<https://doi.org/10.1247/csf.27.421>
12. J. Lee, et al., Autophagy, mitochondria and oxidative stress: cross-talk and redox signalling, *Biochem J*, 441 (2012) 523-540<https://doi.org/10.1042/BJ20111451>
13. F. Reggiori, D.J. Klionsky, Autophagy in the eukaryotic cell, *Eukaryot Cell*, 1 (2002) 11-21<https://doi.org/10.1128/ec.01.1.11-21.2002>
14. L. Poillet-Perez, E. White, Role of tumor and host autophagy in cancer metabolism, *Genes Dev*, 33 (2019) 610-619<https://doi.org/10.1101/gad.325514.119>
15. B. Cui, et al., Autophagy and the Immune Response, *Adv Exp Med Biol*, 1206 (2019) 595-634https://doi.org/10.1007/978-981-15-0602-4_27
16. D.J. Wu, I.E. Adamopoulos, Autophagy and autoimmunity, *Clin Immunol*, 176 (2017) 55-62<https://doi.org/10.1016/j.clim.2017.01.007>
17. P. Boya, et al., Autophagy in stem cells: repair, remodelling and metabolic reprogramming, *Development*, 145 (2018) <https://doi.org/10.1242/dev.146506>
18. L. Ou, et al., The mechanisms of graphene-based materials-induced programmed cell death: a review of apoptosis, autophagy, and programmed necrosis, *Int J Nanomedicine*, 12 (2017) 6633-6646<https://doi.org/10.2147/IJN.S140526>
19. C. He, et al., Exercise-induced BCL2-regulated autophagy is required for muscle glucose homeostasis, *Nature*, 481 (2012) 511-515<https://doi.org/10.1038/nature10758>
20. E.E. Mowers, et al., Functions of autophagy in the tumor microenvironment and cancer metastasis, *FEBS J*, 285 (2018) 1751-1766<https://doi.org/10.1111/febs.14388>
21. J.M.M. Levy, et al., Targeting autophagy in cancer, *Nat Rev Cancer*, 17 (2017) 528-542<https://doi.org/10.1038/nrc.2017.53>
22. L. Zhao, et al., LncRNA SNHG14/miR-5590-3p/ZEB1 positive feedback loop promoted diffuse large B cell lymphoma progression and immune evasion through regulating PD-1/PD-L1 checkpoint, *Cell Death Dis*, 10 (2019) 731<https://doi.org/10.1038/s41419-019-1886-5>
23. W. Zhao, et al., LncRNA HOTAIR influences cell growth, migration, invasion, and apoptosis via the miR-20a-5p/HMGA2 axis in breast cancer, *Cancer Med*, 7 (2018) 842-855<https://doi.org/10.1002/cam4.1353>
24. G.H. Wei, X. Wang, LncRNA MEG3 inhibit proliferation and metastasis of gastric cancer via p53 signaling pathway, *Eur Rev Med Pharmacol Sci*, 21 (2017) 3850-3856
25. C. Zhou, et al., LncRNA PVT1 promotes gemcitabine resistance of pancreatic cancer via activating Wnt/beta-catenin and autophagy pathway through modulating the miR-619-5p/Pygo2 and miR-619-5p/ATG14 axes, *Mol Cancer*, 19 (2020) 118<https://doi.org/10.1186/s12943-020-01237-y>
26. C.Y. Liu, et al., LncRNA CAIF inhibits autophagy and attenuates myocardial infarction by blocking p53-mediated myocardin transcription, *Nat Commun*, 9 (2018) 29<https://doi.org/10.1038/s41467-017-02280-y>

27. Z. Sun, et al., The expression of lncRNA-MALAT1 in breast cancer patients and its influences on prognosis, *Cell Mol Biol (Noisy-le-grand)*, 66 (2020) 72-78
28. Y. Yu, et al., lncRNA PCNAP1 predicts poor prognosis in breast cancer and promotes cancer metastasis via miR3405pdependent upregulation of SOX4, *Oncol Rep*, 44 (2020) 1511-1523<https://doi.org/10.3892/or.2020.7699>
29. J. Gu, et al., Effect of the lncRNA GAS5-MiR-23a-ATG3 Axis in Regulating Autophagy in Patients with Breast Cancer, *Cell Physiol Biochem*, 48 (2018) 194-207<https://doi.org/10.1159/000491718>
30. G. Liang, et al., Autophagy-associated circRNA circCDYL augments autophagy and promotes breast cancer progression, *Mol Cancer*, 19 (2020) 65<https://doi.org/10.1186/s12943-020-01152-2>
31. C. Liu, et al., FSIP1 regulates autophagy in breast cancer, *Proc Natl Acad Sci U S A*, 115 (2018) 13075-13080<https://doi.org/10.1073/pnas.1809681115>
32. O. Martinez-Saez, et al., Frequency and spectrum of PIK3CA somatic mutations in breast cancer, *Breast Cancer Res*, 22 (2020) 45<https://doi.org/10.1186/s13058-020-01284-9>
33. H. Feng, et al., lncRNA TTN-AS1 Regulates miR-524-5p and RRM2 to Promote Breast Cancer Progression, *Onco Targets Ther*, 13 (2020) 4799-4811<https://doi.org/10.2147/OTT.S243482>
34. X. Li, et al., A novel autophagy-related lncRNA prognostic risk model for breast cancer, *J Cell Mol Med*, 25 (2021) 4-14<https://doi.org/10.1111/jcmm.15980>
35. R. Zhang, et al., Identification and Validation of an Autophagy-Related lncRNA Signature for Patients With Breast Cancer, *Front Oncol*, 10 (2020) 597569<https://doi.org/10.3389/fonc.2020.597569>
36. T. Huang, et al., MIR93 (microRNA -93) regulates tumorigenicity and therapy response of glioblastoma by targeting autophagy, *Autophagy*, 15 (2019) 1100-1111<https://doi.org/10.1080/15548627.2019.1569947>
37. H. Li, et al., miR-519a enhances chemosensitivity and promotes autophagy in glioblastoma by targeting STAT3/Bcl2 signaling pathway, *J Hematol Oncol*, 11 (2018) 70<https://doi.org/10.1186/s13045-018-0618-0>

Figures

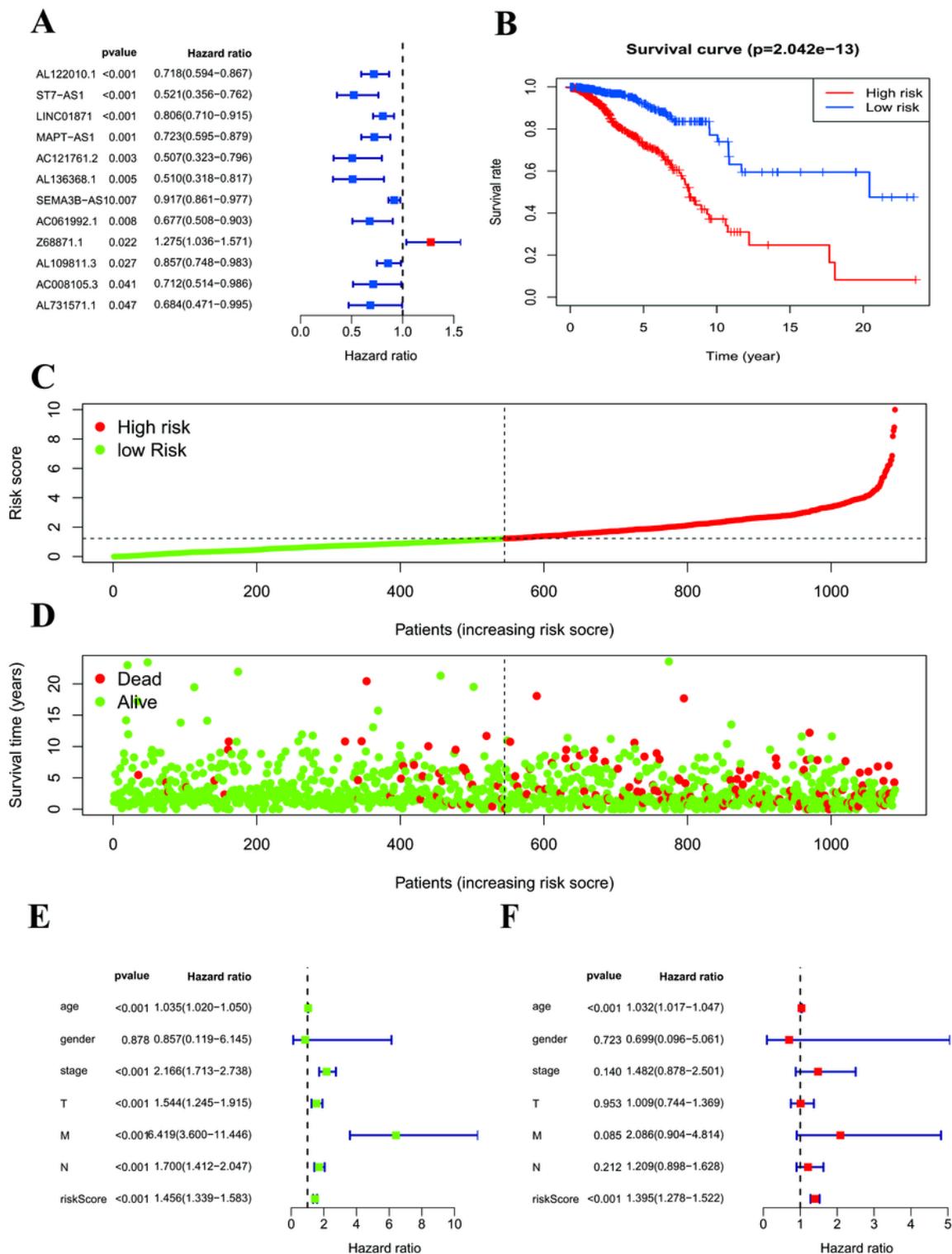


Figure 1

The construction and validation of the prognostic signature. (a) Univariate Cox regression analysis showed 12 autophagy-related lncRNAs associated with the OS of the breast cancer. (b) The high risk group presented a worse OS than that of low-risk group. (c) Risk scores of BC patients were presented based on the high- and low-risk groups. (d) The scatterplot of relationship between risk scores and

survival time/survival outcomes. Forest plot of several factors, including age, gender, stage ,T, N, M phase and risk score by univariate Cox regression (e) and multivariate Cox regression (f).

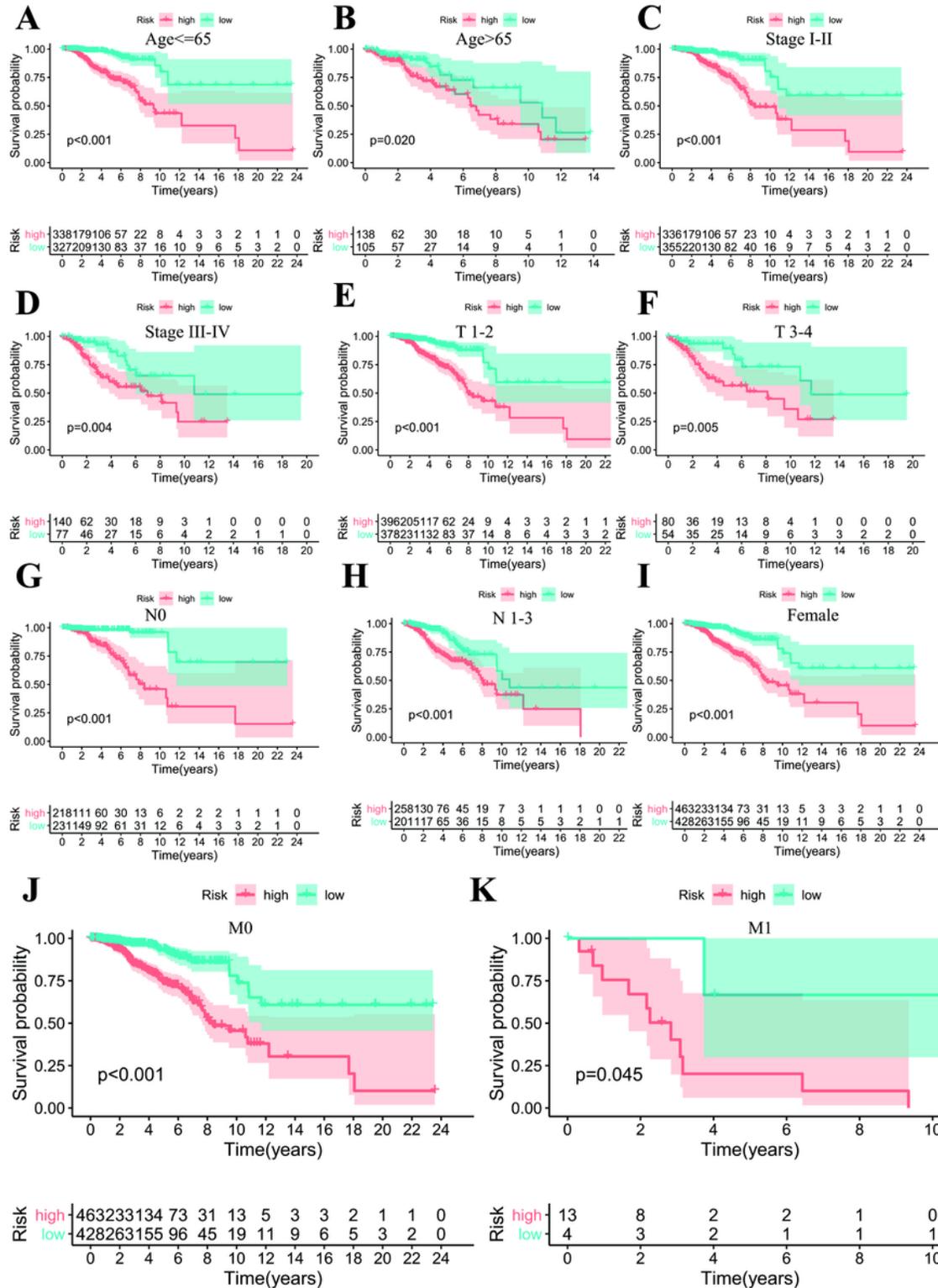


Figure 2

Comparison of survival time between high- and low-risk groups within the clinical subgroups. The survival difference between high- and low-risk groups within the subgroups of Age ≤ 65 (a) and Age > 65

(b), Stage I&II (c) and Stage III&IV (d), T1-2 (e), and T3-4 (f) phase, N0 (g) and N1-3 (h) phase, gender (i) as well as M0 (j) and M1(k) phase.

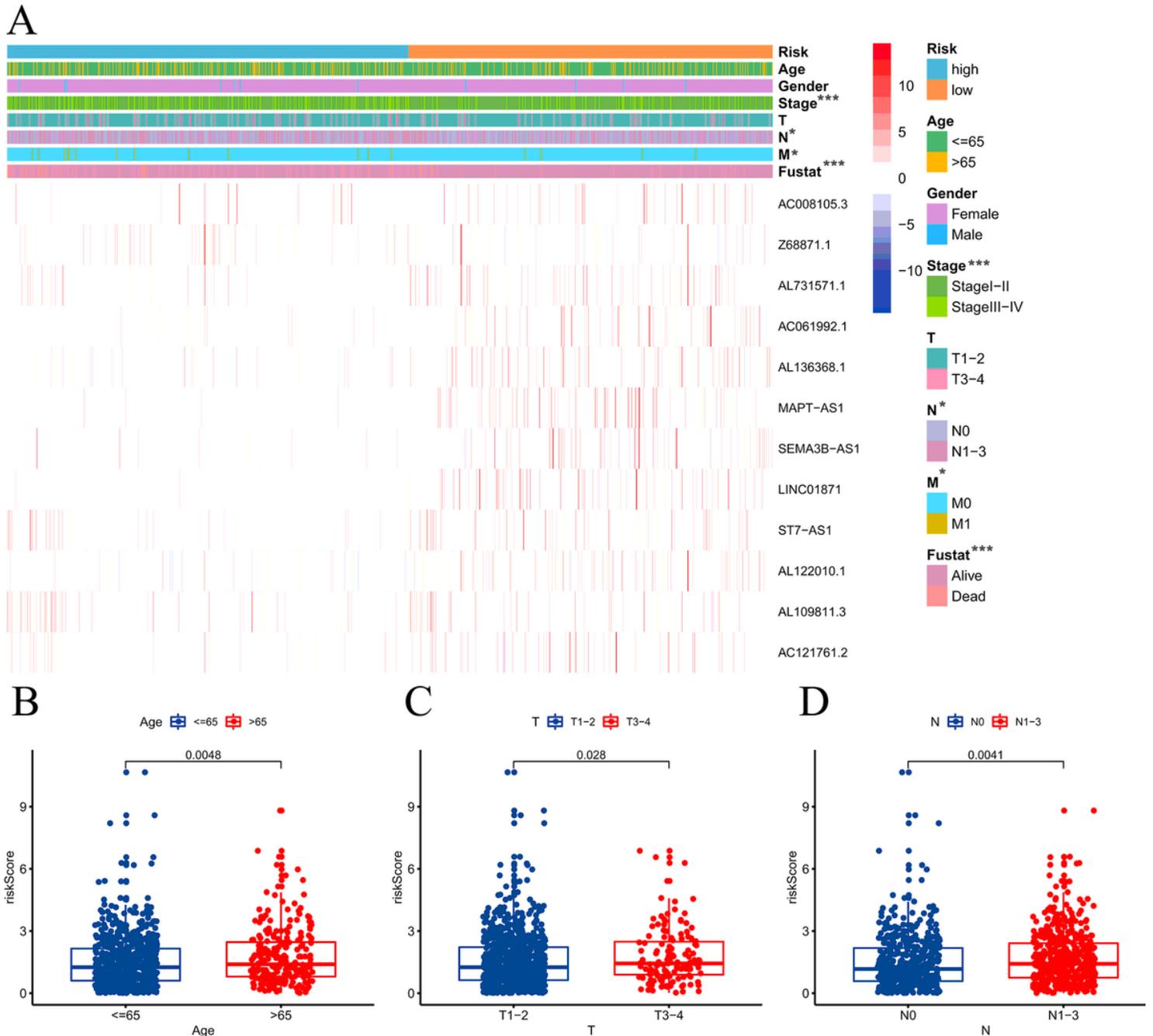


Figure 3

Some clinical factors are correlated with risk score. (a) Heat map of the clinical features and expression of the 12 lncRNAs in high- and low-risk groups. (b) Comparison of risk scores of different ages. (c) Comparison of risk scores of different T phases. (d) Comparison of risk scores of different N phases. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

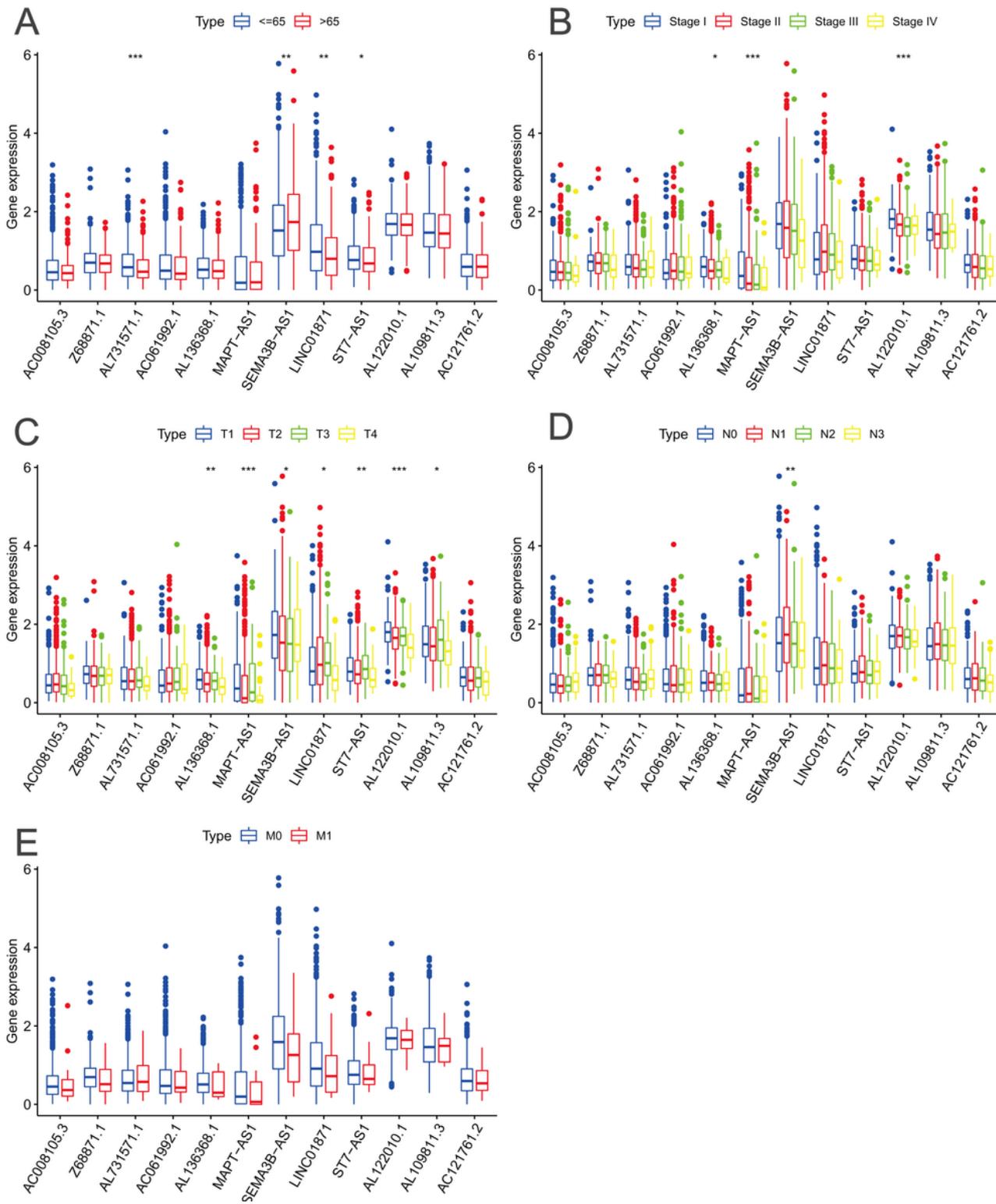


Figure 4

The expression level of the 12 autophagy-related lncRNAs changed with clinical factors. a-e represents age, stage, T, N and M phase, respectively. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

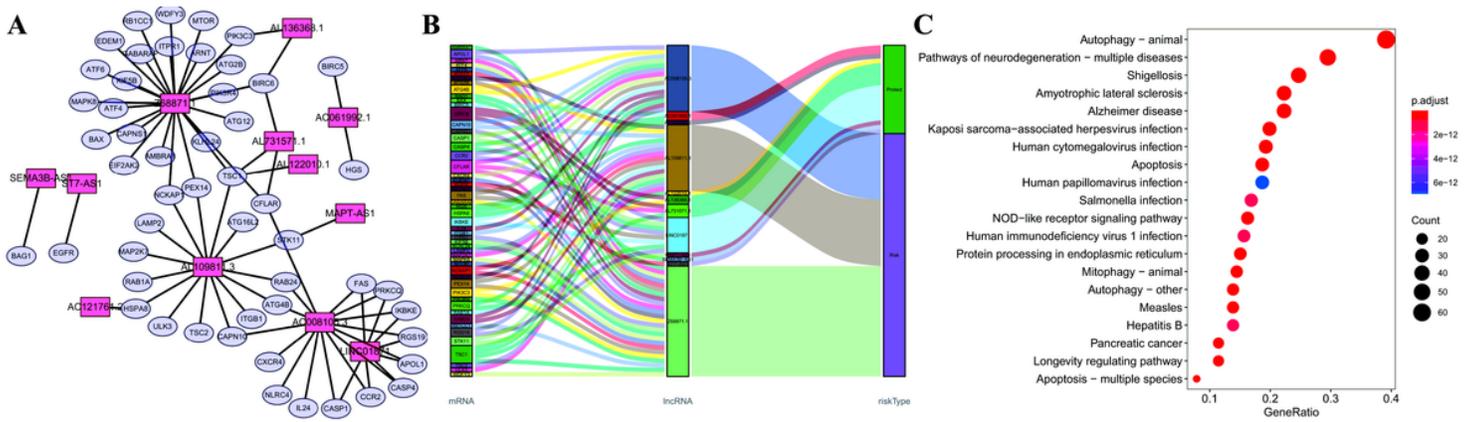


Figure 5

Co-expression and function analysis of the signature. (a) The network of the 12 autophagy-related lncRNAs and their regulatory mRNAs. (b) The relationship among mRNAs, lncRNAs and risk factors presented by a Sankey diagram. (c) KEGG enrichment analysis showed the functions of the co-expressed mRNAs

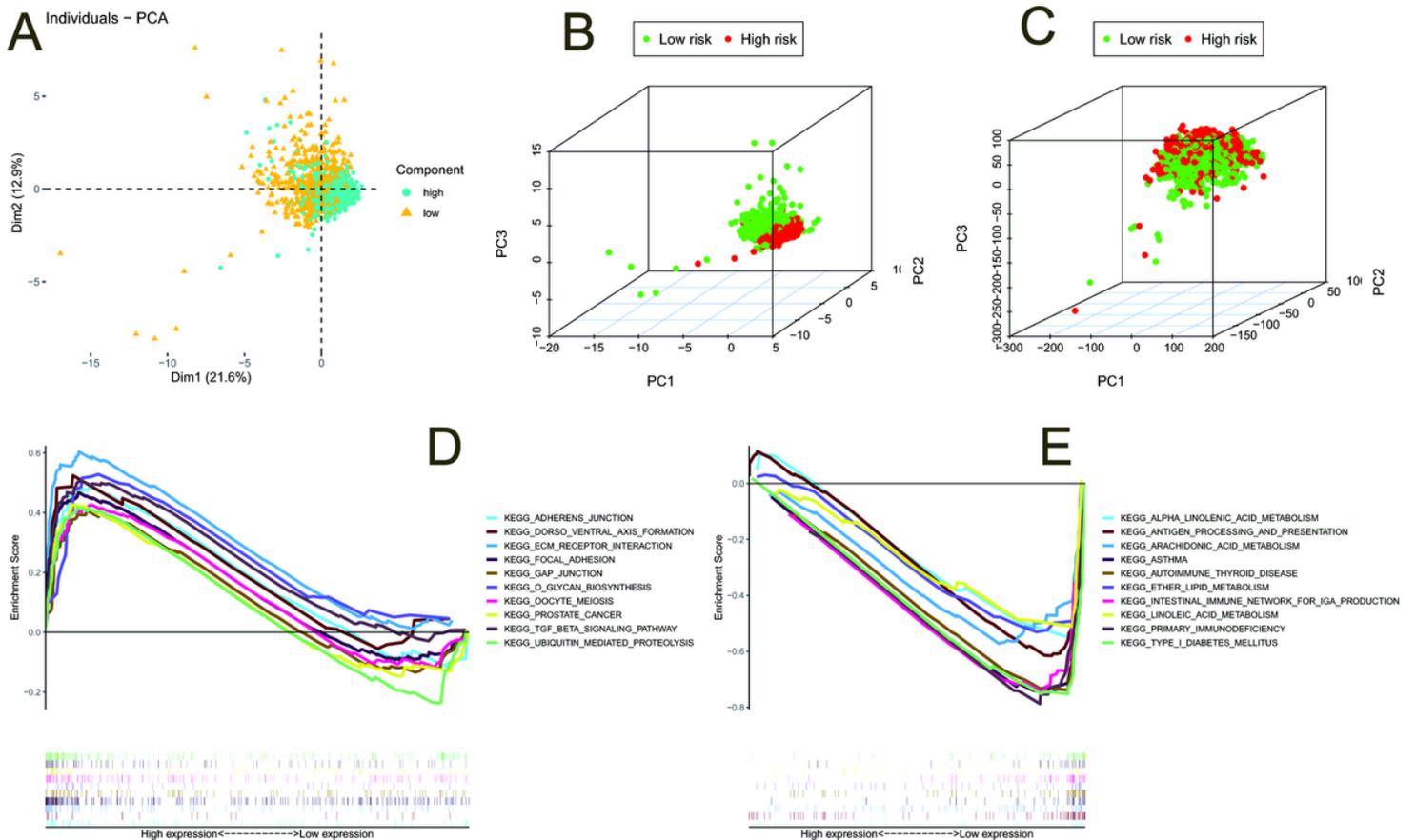


Figure 6

PCA and GSEA analysis showed the essence of the genes in high- and low-risk groups. According to PCA, there were separations of autophagy-related lncRNAs sets between the high- (a) and low-risk (b) groups. (c) There was no significant difference of the genome-wide expression profiles between the two groups

by PCA. The GSEA analysis showed top 10 gene sets enriched in high-risk group (d) and low-risk group (e).

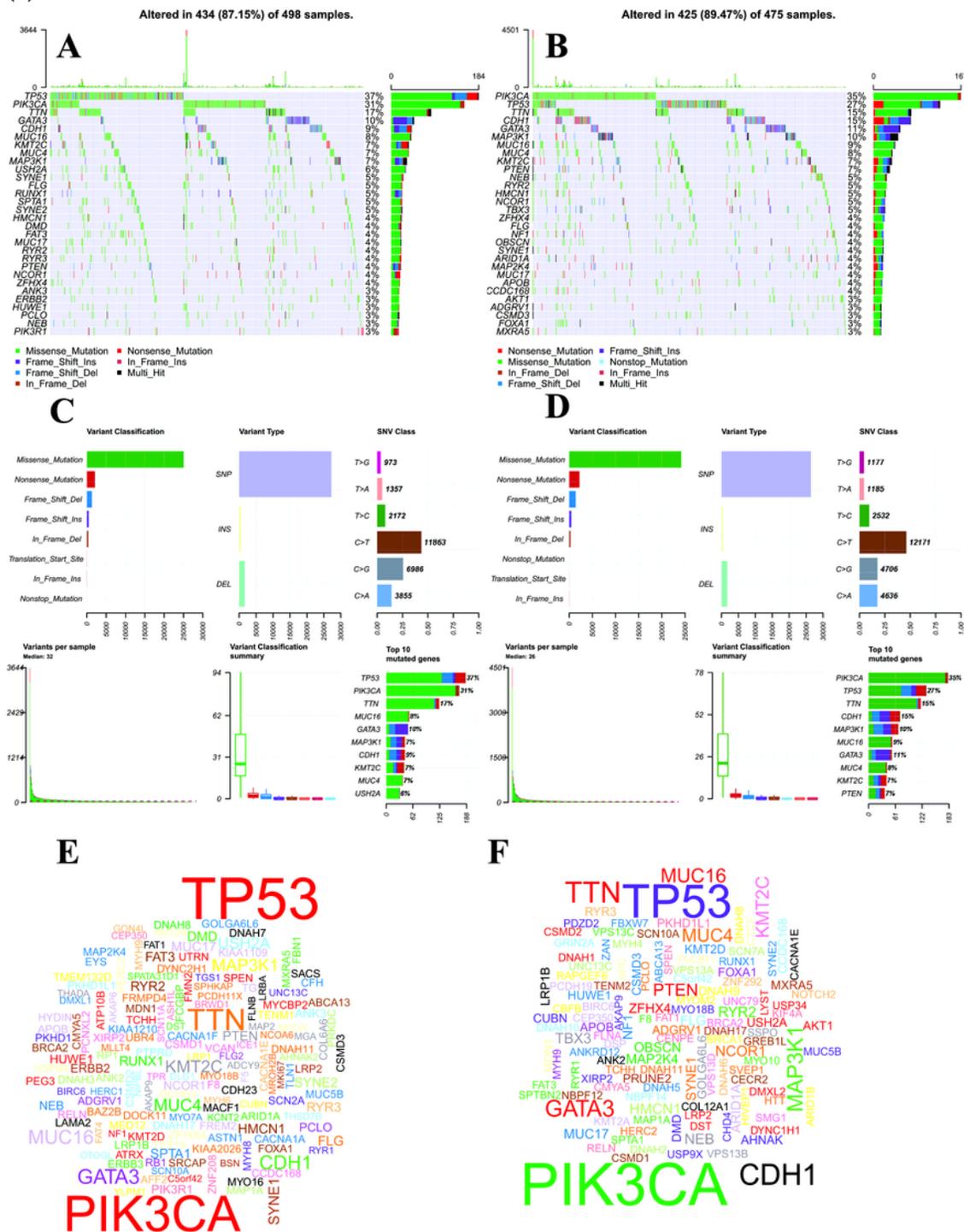
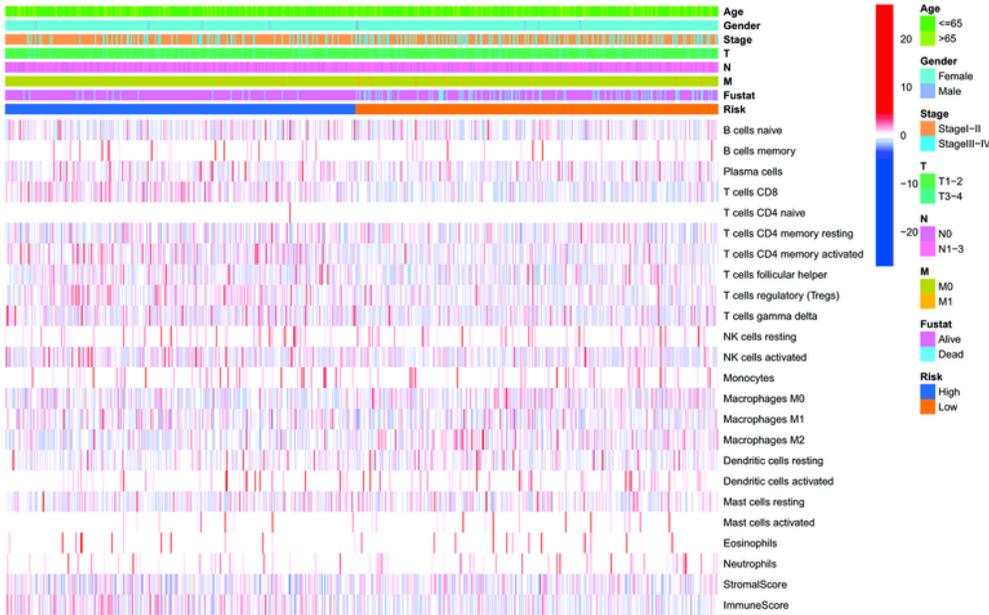


Figure 7

Comparison of mutation profiles of high- and low-risk groups. Waterfall plots show various mutation types in high- (a) and low-risk groups (b). The variant classification and type, the SNV summary of the variant per sample and variant classification, and the top 10 mutated genes in high- (c) and low-risk

groups (d). Gene cloud revealed the mutated frequencies by the size of the gene in high- (e) and low-risk groups (f).

A



B

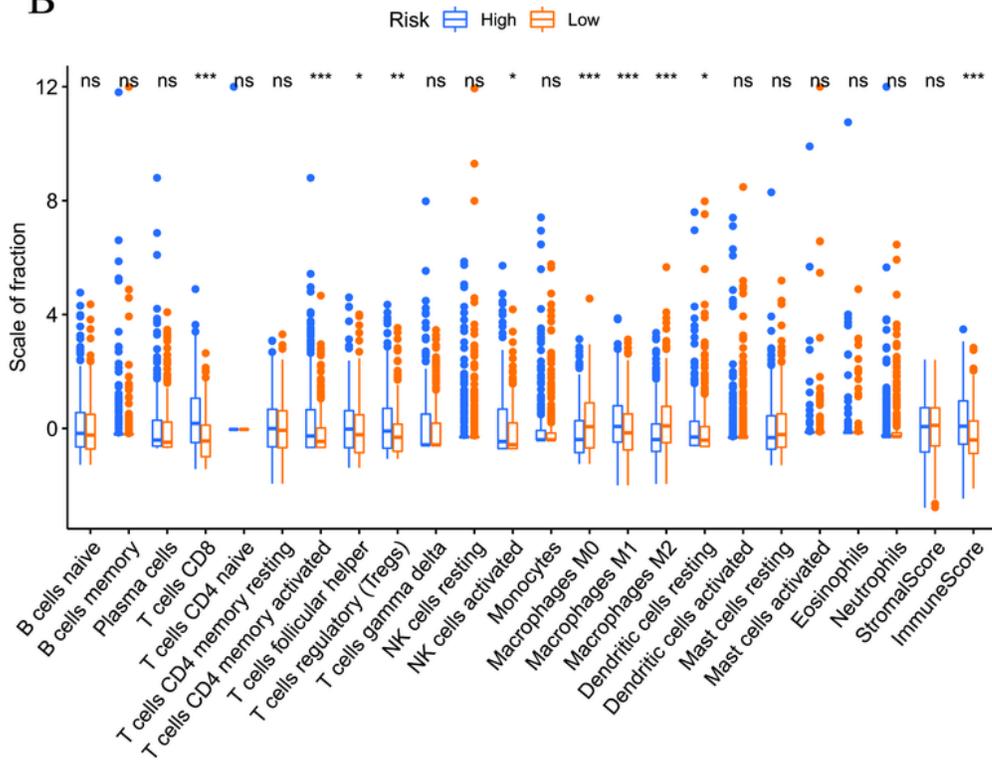


Figure 8

Immune status difference between high- and low-risk groups. (a) Heat map and (b) box plot presented the distinction of various immune cells, respectively. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

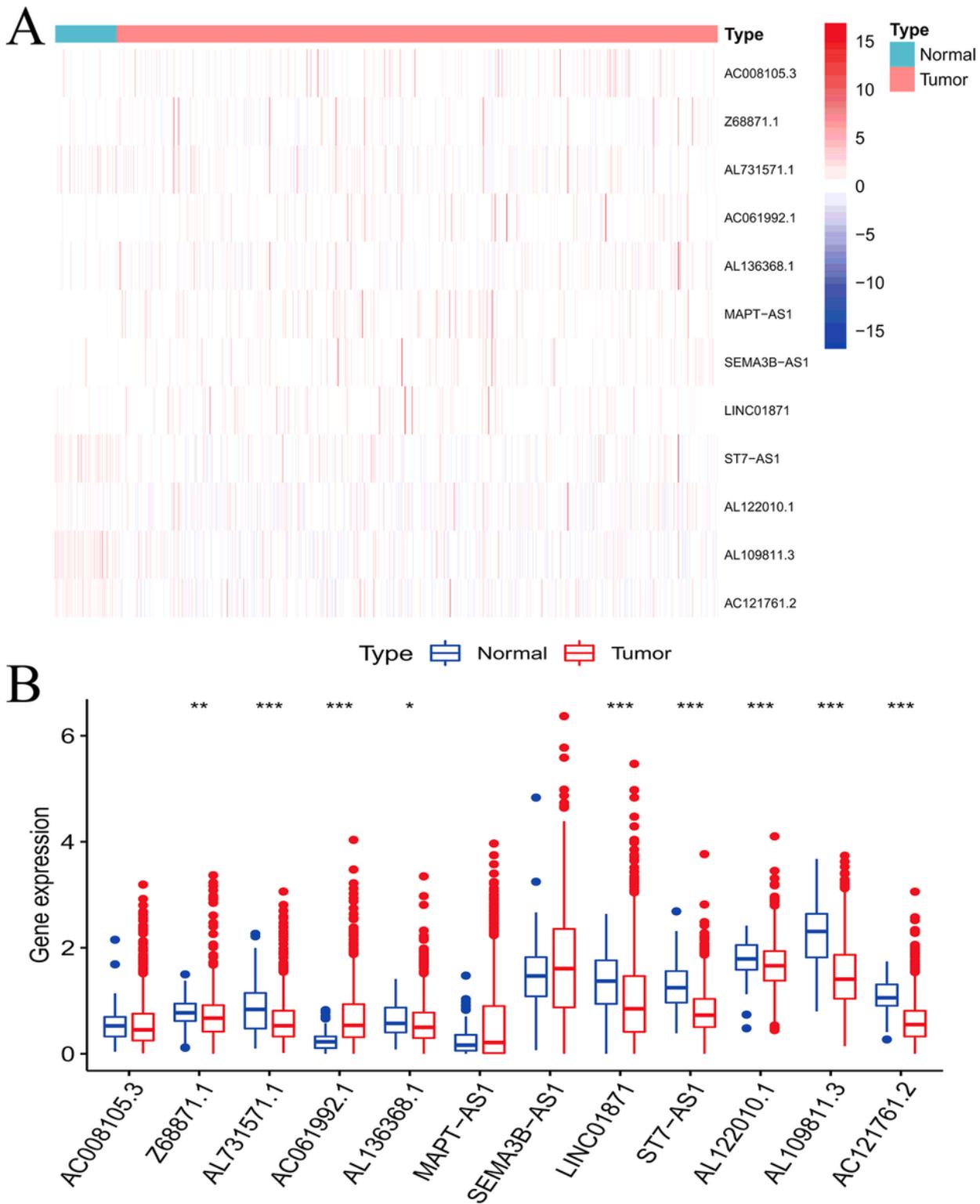


Figure 9

Further verification of the discrepant expression of the 12 autophagy-related lincRNAs between tumor and normal tissues. The heat map(a) and box plot(b) demonstrated the different expression level between 1109 BC tissues and 113 normal tissues in the TCGA data set. * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementarymaterialFigureS1.tif](#)
- [SupplementarymaterialFigureS2.tif](#)
- [SupplementarymaterialTableS1S3.xls](#)