

# Identification and Validation of Two Prognostic Autophagy-related Models in Clear Cell Renal Cell Carcinoma

**Xiong-Lin Sun**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Jun-Ming Zhu**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Fei Lin**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Zhi-Bin Ke**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Xiao-Dong Li**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Hai Cai**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Fei Huang**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Xue-Yi Xue**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Qing-Shui Zheng**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Shao-Hao Chen**

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

**Ning Xu** (✉ [drxun@fjmu.edu.cn](mailto:drxun@fjmu.edu.cn))

Department of Urology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350005, China

<https://orcid.org/0000-0001-7909-7025>

---

**Primary research**

**Keywords:** Long non-coding RNAs, Gene, Prognosis, Autophagy, ccRCC

**Posted Date:** September 22nd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-78750/v1>

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

[Read Full License](#)

---

# Abstract

**Background:** Clear cell renal cell carcinoma (ccRCC) is one of the most frequent malignancies. Increasing evidence has highlighted the critical roles of autophagy-related genes and autophagy-related long non-coding RNA (lncRNA) in ccRCC development and progression. Therefore, it is necessary to identify novel biomarkers associated with autophagy in ccRCC.

**Methods:** A total of 507 ccRCC patients were included in our study and then randomly divided into a training cohort (n=255) and testing cohort (n=252). Univariate Cox regression models, Lasso regression analyses and multivariate Cox regression models were successively used for constructing gene model and lncRNA model. Receiver-operating characteristic (ROC) curve analysis, Kaplan-Meier (K-M) analysis and more functional analyses were applied for verifying the accuracy of the two models.

**Results:** The autophagy-related genes (ARGs) prognostic model was constructed based on the six ARGs (EIF4EBP1, IFNG, BID, BIRC5, CX3CL1 and RAB24) and five autophagy-related lncRNAs (AC093278.2, AC010326.3, AC099850.3, AC016773.1 and AC009549.1) and then ccRCC patients were significantly stratified into high- and low-risk groups in terms of overall survival (OS). K-M survival analyses indicated that low-risk group had a lower mortality rate than high-risk group in the six-gene prognostic risk model (training cohort:  $P=4.138e-07$ ; testing cohort:  $P=1.125e-03$ ) and the same results were obtained in the case of the five-lncRNA prognostic risk model (training cohort:  $P=4.564e-09$ ; testing cohort:  $P=2.485e-03$ ). The results of time-dependent ROC curves revealed six-gene prognostic risk model had a higher area under curve (AUC) of 0.765 than the five-lncRNA prognostic risk model at an AUC of 0.759. Therefore, the gene model is an indicator as good as the model constructed by lncRNAs. In addition, further functional analysis indicated these genes were functionally involved in regulation of endopeptidase activity, regulation of peptidase activity, autophagy, human cytomegalovirus infection, shigellosis, autophagy-animal and HIF-1 signaling pathway.

**Conclusions:** A total of six OS-related ARGs and five autophagy-related lncRNA were identified in our current study. The two autophagy-related prediction models including genes and lncRNAs are reliable prognostic and predictive biomarkers for ccRCC.

## Introduction

Clear cell renal cell carcinoma (ccRCC) is one of the most aggressive and common types in the urinary system with a steadily rising rate of 2–4% each year, which accounts for 70–85% of renal cell carcinoma (RCC)[1]. In 2020, it is reported that in the US there will be approximately 73,750 new cases and 14,840 deaths from kidney cancer[2]. Not only African Americans but also Chinese have both a higher incidence and mortality rates for ccRCC[3]. Despite the rapidly evolving of experimental technologies and therapeutic guidelines in this field, such as tyrosine kinase inhibitors (TKI)[4], optimized surgery[5] and immunotherapy[6], there is no curative treatment for advanced ccRCC. Therefore, it is warranted to develop candidate therapeutic biomarkers for improving survival of ccRCC patients.

Autophagy is a significant self-digestion process of long-lasting proteins, misfolded proteins and damaged intracellular organelles which is increasingly considered as a protective mechanism against many tumor diseases[7]. In the past few years, considerable amount of evidence suggested that autophagy played a pivotal role in the growth and progression of RCC, mainly in maintaining cellular homeostasis and providing material and energy for cell survival[8, 9].

Autophagy-related genes (ARGs) and long non-coding RNA (lncRNA) have traditionally been the hotspot of scientific research. Kong et al. revealed the expression levels of autophagy-related genes, BECN1 and ATG5, were higher in chronic lymphocytic leukemia (CLL) compared with healthy controls[10]. Liu et al. found a novel CAIF-p53-myocardin axis as a potential therapeutic target in treatment of defective autophagy-associated cardiovascular diseases[11]. To better understand the impact of ARGs and lncRNAs on ccRCC, it is necessary to construct more prediction models to improve prediction accuracy so as to help the clinical application.

In this study, the expression of ARGs and lncRNAs that related to the prognosis in ccRCC patients from The Cancer Genome Atlas (TCGA) were elected and two corresponding prognostic prediction models were built by using a series of bioinformatic tools for the first time. Receiver-operating characteristic (ROC) curve analysis, Kaplan-Meier (K-M) methods and independent prognostic analysis were used to assess the specificity and sensitivity of these models to determine the prognostic significance. Besides, further function and pathway enrichment analyses were conducted to provide hints concerning the roles of the prognostic genes in ccRCC progression. These findings demonstrated an effective multi-dimensional biomarker strategy that would be effective in monitoring autophagy alteration and predicting the prognosis in ccRCC patients.

## Methods

### ccRCC data source and patient information

RNA-seq data of 611 ccRCC tissue samples (539 primary ccRCC tissue samples, 72 normal controls) were altogether downloaded and extracted from TCGA. Among these patients, a total of 507 primary ccRCC patients with gene expression data and clinical follow-up information was involved in the present study after removal of patients with unknown clinical information and patients whose ID numbers did not match with their corresponding expression profile and clinical data. All 507 ccRCC patients were randomly divided into a training cohort (n = 255) for identifying prognostic model and a testing cohort for validating its prognostic value of the model (n = 252). A sum up of 232 ARGs were obtained from the Human Autophagy Database (HADb, [http://www. autophagy.lu/index.html](http://www.autophagy.lu/index.html)) which is an autophagy-dedicated database aiming to reserve human genes involved in autophagy. Four autophagy gene sets (go positive regulation of autophagy M15852, go regulation of autophagy M10281, go selective autophagy M24317 and kegg regulation of autophagy M6382) were extracted from the Molecular Signatures Database v4.0 (<http://www.broadinstitute.org/gsea/msigdb/index.jsp>).

# Prognostic Args And Lncrnas In Two Models

'limma' package in R software was applied to estimate differentially expressed ARGs between ccRCC and their non-tumor counterparts with cutoff of  $|\log_2 \text{Fold Change (FC)}| > 1$  and adjusted P-value  $< 0.05$ . Similarly, the autophagy-related lncRNAs in ccRCC patients were also screened using package 'limma' in R with thresholds of  $|\text{cor}| > 0.6$  and  $P < 0.001$ . Further, the univariate Cox regression models were constructed to determine prognosis-related ARGs and lncRNAs that were significantly linked to overall survival (OS) using the data in the training group ( $P < 0.001$ ). In order to identify optimal predictive ARGs and lncRNAs, Lasso regression analyses were performed by package 'glmnet' in R via the cross-validation likelihood (cvl) method for 1000 times. Multivariate Cox regression models were used to construct the OS prognostic risk models. A total of six differentially expressed ARGs and five autophagy-related lncRNAs were brought into final prognostic models. The whole process is shown in Fig. 1.

## Construction Of Prognostic Risk Model

Two prognostic risk models were respectively established based on the regression coefficients of the individual genes derived from the multivariate Cox regression model and the expression level of genes. The risk score for each patients was calculated as follows: Risk score =  $\beta_1 * \text{Exp}_1 + \beta_2 * \text{Exp}_2 + \beta_3 * \text{Exp}_3 + \dots + \beta_x * \text{Exp}_x$ .

According to the median risk score of training group as the cutoff, each set was classified into a high-risk group and a low-risk group. A lower risk score indicates a better prognosis for the ccRCC patients. To assess the efficiency of the two models, the survival analyses were conducted by K-M methods to compare the high-risk and low-risk groups by package 'survival' in R and the time-dependent ROC curves for 3-year and 5-year were plotted by package "survivalROC" in R.

## Functional Analysis

Gene set enrichment analysis (GSEA, <http://www.broadinstitute.org/gsea/msigdb/index.jsp>) was performed to identify the autophagy-related gene sets between high-risk and low-risk groups for two models. Normalized enrichment score (NES) and false discovery rate (FDR) were defined as statistical differences. Principal component analysis (PCA) of each stage of selected genes was performed by 'limma' and 'scatterplot3d' package in R. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were carried out using the DAVID (<https://david.ncifcrf.gov/>) to find the major biological attributes of these genes. Significant enrichment results were visualized for high-dimensional information using R package 'Goplot'. The Correlation circle maps were drawn using R language.

## Statistical analysis

All of the statistical tests were done with R software and Bioconductor (version 3.5.2; <https://www.r-project.org/>), SPSS software (version 24.0; Chicago, IL, USA) and GraphPad Prism 7 (San Diego, CA, USA). R 3.5.2 and Perl (64-bit, version 5.30.1.1, <http://www.perl.org>) were performed to draw plots.

The Cox regression models were used to evaluate the association between expression profiles and survival status of ccRCC patients. Univariate, LASSO and multivariate analyses were performed using the Cox proportional hazards regression model. The Kaplan–Meier curve was determined by GraphPad Prism 7. ROC curve and the corresponding area under the ROC curve (AUC) calculated from the ROC curves were used to evaluate the accuracy of the models that predicted prognosis and estimate the diagnostic value of gene expression using the package of “survivalROC” in R. Hazard ratios (HR) and 95% confidence intervals (CI) were calculated. P value < 0.05 was considered statistically significant.

## Results

### 1. Six ARGs and five autophagy-related lncRNAs were screened out from the training cohort

There were 507 ccRCC patients with raw expression data and corresponding clinical follow-up information extracted from TCGA project in the present study which were randomly divided into a training dataset (n = 255) and a testing dataset (n = 252). A total of 222 ARGs were integrated from HADB database. Considered as the criteria of FDR < 0.05 and  $|\log_2(\text{FC})| > 1$ , 45 differentially expressed genes were finally obtained between tumor and non-tumor tissues including 36 up-regulated and 9 down-regulated ARGs (Fig. 2). In addition, we identified 467 autophagy-related lncRNAs by co-expression networks (corFilter = 0.6, P < 0.001). Univariate Cox regression analyses were utilized to single out the prognostic ARGs and lncRNAs in the training group, resulting in 12 ARGs (P < 0.001) and 79 autophagy-related lncRNAs (P < 0.001) associated with the ccRCC patients' overall survival. Lasso regression analyses of those screened genes (Fig. 3A, B,  $\log(\lambda.\text{min}) = -2.9642$ ) and lncRNAs (Fig. 3C, D,  $\log(\lambda.\text{min}) = -2.4841$ ) above were performed for further multivariate Cox regression model to avoid overfitting the prognostic problems. Six ARGs (EIF4EBP1, IFNG, BID, BIRC5, CX3CL1 and RAB24, Fig. 4A) and five autophagy-related lncRNAs (AC093278.2, AC010326.3, AC099850.3, AC016773.1 and AC009549.1, Fig. 4B) were finally identified to retain their prognostic significance (Table 1, Table 2). There were five ARGs (EIF4EBP1, IFNG, BID, BIRC5, and RAB24) and four autophagy-related lncRNAs (AC010326.3, AC099850.3, AC016773.1 and AC009549.1) associated with shorter survival which were regarded as prognostic risky factors whereas the remaining one ARG (CX3CL1) and one autophagy-related lncRNA (AC093278.2) with negative coefficient of multivariate Cox regression analyses may be protective factors due to the close association between their high expression and longer patients' survival.

Table 1

Six ARGs and five autophagy-related lncRNAs were selected by multivariate Cox regression

Type	Gene name	Coefficient	HR	HR.95L	HR.95H	P value
gene	EIF4EBP1	0.006	1.006	0.999	1.013	0.114
	IFNG	0.119	1.126	1.002	1.266	0.047
	BID	0.113	1.119	1.013	1.238	0.027
	BIRC5	0.093	1.097	1.016	1.184	0.018
	CX3CL1	-0.012	0.988	0.974	1.002	0.100
	RAB24	0.049	1.051	0.999	1.104	0.054
lncRNA	AC093278.2	-0.123	0.884	0.801	0.976	0.014
	AC010326.3	0.101	1.107	0.966	1.268	0.143
	AC099850.3	0.101	1.107	1.051	1.165	0.000
	AC016773.1	0.340	1.406	0.924	2.137	0.111
	AC009549.1	0.079	1.083	0.991	1.183	0.079
HR, hazard ratio; lncRNAs, long non-coding RNAs						

Table 2

List of the six ARGs and five autophagy-related lncRNAs associated with prognosis in ccRCC patient

ID	Gene name	Protein name	Location	Positive correlation	Negative correlation
1978	EIF4EBP1	Eukaryotic translation initiation factor 4E-binding protein 1	Chr8	IGFBP4 VKORC1 PPP1R14BP3	RBL2 TSTD2 INIP
3458	IFNG	Interferon gamma	Chr12	-	-
637	BID	BH3-interacting domain death agonist	Chr22	TMSB10 ARPC3 KIAA0930 PDCD5 RGS19	WLS DCAF11 L2HGDH DYNLL2 TSPYL1
332	BIRC5	Baculoviral IAP repeat-containing protein 5	Chr17	-	-
6376	CX3CL1	Fractalkine	Chr16	-	-
53917	RAB24	Ras-related protein Rab-24	Chr5	FAM193B APBB3 ZNF692 SPPL2B ANKZF1	DPH3 EIF4G2 CLTC TM9SF2 YWHAB
ENSG00000261269	AC093278.2	-	Chr5	LDB2 GIMAP8	PPIF ATP6V0E2
ENSG00000269867	AC010326.3	-	Chr19	-	-
ENSG00000265303	AC099850.3	-	Chr17	-	-
ENSG00000207009	AC016773.1	-	Chr4	TEPSIN GSDMB CHKB NOL12	PCYOX1 CHMP3 SERINC1 CCDC47
ENSG00000270607	AC009549.1	-	Chr11	-	-
Chr, Chromosome; lncRNAs, long non-coding RNAs					

## 2. Construction of a six-gene prognostic risk model and a five-lncRNA prognostic risk model in the training cohort

Two prognostic risk models were respectively constructed by integrating the expression values of the genes and lncRNAs selected above and corresponding estimated regression coefficient derived from multivariate Cox regression analysis. The formula of risk score for each patient was calculated as follows (six-gene prognostic risk model):  $\text{risk score} = (0.0057 \times \text{expression level of EIF4EBP1}) + (0.1189 \times \text{expression level of IFNG}) + (0.1128 \times \text{expression level of BID}) + (0.0925 \times \text{expression level of BIRC5}) + (-0.012 \times \text{expression level of CX3CL1}) + (0.0493 \times \text{expression level of RAB24})$ . The five-lncRNA prognostic risk model were built as follows:  $\text{risk score} = (-0.1233 \times \text{expression level of AC093278.2}) + (0.1014 \times \text{expression level of AC010326.3}) + (0.1012 \times \text{expression level of AC099850.3}) + (0.3404 \times \text{expression level of AC016773.1}) + (0.0794 \times \text{expression level of AC009549.1})$ . All the 507 ccRCC patients in the training dataset and the testing dataset for two models were divided into high-risk and low-risk groups according to the median values of the OS-related prediction model, respectively.

## 3. A comparison was made of the predictive ability between the two prognostic risk model for training dataset and testing dataset

To compare the performance of prognostic risk model in predicting the clinical outcome of ccRCC patients, the K-M survival curves were plotted to analyze the different survival time between the high-risk and low-risk groups. K-M survival analyses indicated that low-risk group (5-year survival rate 83.1%) had a lower mortality rate than high-risk group (5-year survival rate 52.9%) in the six-gene prognostic risk model (Fig. 5A,  $P = 4.138e-07$ ). Patients in the high-risk group had significantly shorter survival (5-year survival rate 49.4%) than those in the low-risk group (5-year survival rate 85.9%) in the five-lncRNA prognostic risk model (Fig. 6A,  $P = 4.546e-09$ ). ROC curves analyses were carried out to show prognostic effectiveness for predicting the 3-year and 5-year survival. The 3- and 5-year of AUC for training dataset in the six-gene prognostic risk model were respectively 0.708 and 0.73.8 (Fig. 5C, D), while those in the five-lncRNA prognostic risk model were 0.740 and 0.757 (Fig. 6C, D). The preliminary analyses demonstrated the five-lncRNA model had a better competitive performance than the another in the training dataset.

With the increase in the risk scores, Fig. 5G and Fig. 6G showed the distribution of the risk score, the number of patients in two risk groups and prognostic expression profiles in the training dataset. The expression levels of ARGs (EIF4EBP1, IFNG, BID, BIRC5, and RAB24) and lncRNAs (AC010326.3, AC099850.3, AC016773.1 and AC009549.1) were up-regulated and the expression levels of remaining ARG (CX3CL1) and lncRNA (AC093278.2) were down-regulated.

To further validate the prognostic power of the two models, patients in the testing dataset were marked as a high-risk group and a low-risk group by the same cutoff derived from the training dataset. In consistent with the findings in the training dataset, K-M survival analysis (Fig. 5B, 6B) and ROC curves analysis (Fig. 5E, F; 6E, F) compared with the training dataset were summarized in Table 3. The heatmap of expression profiles, the risk scores and distribution and patients' survival status in the TCGA database

were also performed in testing dataset (Fig. 5H, 6H). The partial analyses demonstrated the six-gene model had a better competitive performance than the latter in the testing dataset.

Table 3  
Predictive performances of two prognosis prediction models based on six ARGs and five autophagy-related lncRNAs

Type	Gene model			lncRNA model		
	K-M analysis P value	AUC (3 year)	AUC (5 year)	K-M analysis P value	AUC (3 year)	AUC (5 year)
Training dataset	4.138e-07	0.708	0.738	4.564e-09	0.74	0.757
Testing dataset	1.125e-03	0.684	0.701	2.485e-03	0.67	0.693

K-M, Kaplan-Meier; lncRNAs, long non-coding RNAs; AUC, area under curve

#### 4. Comparison of the relationships between clinical parameters and two prognosis-related prediction model

Age, gender, grade, tumor stage, T-staging and M-staging were combined with models to evaluate whether the prognostic power of the two models were independent of other clinical characters.

The univariate Cox regression analyses demonstrated six-gene prognostic risk model (HR = 1.019, 95% CI = 1.010–1.029, Fig. 7A) and five-lncRNA prognostic risk model (HR = 1.022, 95% CI = 1.011–1.032, Fig. 7B) were all available independent prognostic indicator ( $P < 0.001$ ). Six-gene prognostic risk model ( $P = 0.002$ , HR = 1.014, 95% CI = 1.005–1.023, Fig. 7C) had similar prognostic efficiency with another model ( $P = 0.002$ , HR = 1.016, 95% CI = 1.006–1.027, Fig. 7D) after adjustment for the same clinical variables by multivariate Cox regression analyses. In addition, time-dependent ROC curves were plotted and calculated by clinical characters. The results revealed six-gene prognostic risk model (Fig. 7E) had a higher AUC of 0.765 than the five-lncRNA prognostic risk model (Fig. 7F) at an AUC of 0.759. All results above demonstrated a similar competitive performance of the two prognostic risk model for survival prediction.

#### 5. Comparison of the different status between high-risk and low-risk in two prognostic risk model

In order to further explore these differences in statuses, PCA were performed for all genes extracted from TCGA, ARGs and autophagy-related lncRNAs and six genes and five lncRNAs selected finally under the analyses of two models, respectively. Figure 8 indicated the divider line between the red dots and the green dots became increasingly obvious with the small number of candidate genes selected. Four typical autophagy-related gene sets were employed for gene functional enrichment analysis to identify the low-risk group is enrich in autophagy ( $P < 0.01$ ) except for 'kegg regulation of autophagy' in the lncRNA model

( $P = 0.4$ ). The higher the normalized enrichment score (NES) score is obtained, the greater the enrichment of the gene set. The results showed that the two gene sets (go positive regulation of autophagy and kegg regulation of autophagy) had a greater enrichment in the low-risk group for six-gene prognostic risk model while another two gene sets (go selective autophagy and go regulation of autophagy) indicated a better enrichment in the five-lncRNA prognostic risk model (Fig. 9).

## 6. Further functional analysis of the six-gene prognostic risk model and five-lncRNA prognostic risk model

To explore the primary mechanism of the gene model, GO enrichment and KEGG analysis were performed according to the differentially expressed ARGs. The ARGs were found significantly related to top GO biological processes according to the results of DAVID, including regulation of endopeptidase activity, regulation of peptidase activity, autophagy and process utilizing autophagic mechanism (Fig. 10,  $P < 0.05$ ). The heatmap of the relationship between ARGs and GO enrichment analysis was also displayed. Moreover, 5 KEGG signaling pathways functionally involved ARGs, consisting of human cytomegalovirus infection, shigellosis, autophagy-animal, HIF-1 signaling pathway and Kaposi sarcoma-associated herpesvirus infection (Fig. 11,  $P < 0.05$ ). In addition, the expression correlation between prognostic ARGs and lncRNAs and all genes extracted from ccRCC in TCGA were examined. Correlation circle maps were plotted based on genes which had both positive correlation and negative correlation genes (Fig. 12, Table 2). Color intensity is proportional to the correlation coefficients. These results suggested that the two models might affect autophagy-related biological processes and pathways involved in ccRCC.

## Discussion

Although there is a growing interest to use new first-line treatment options as non-operative interventions or therapy protocols for ccRCC, including TKI cabozantinib, the immunotherapy combination of nivolumab /ipilimumab, and the combination of atezolizumab /bevacizumab, the individual prognosis of ccRCC is commonly less than satisfactory[4, 6]. A number of lines of evidence suggested autophagy played an important role in cell survival, development, and cancer progression. Liu et al. found ccRCC was associated with mono-allelic loss and/or mutation of the critical autophagy-related gene ATG7, and autophagy had an anticancer role in ccRCC tumorigenesis[12]. Messai et al. suggested EPAS1 or the autophagy sensor ITPR1 may serve as a potential therapeutic target in the future because of the improvement of NK cell responses in patients with RCC[13]. lncRNAs are a type of functional non-coding transcripts that are longer than 200 nt. Recent studies have revealed lncRNAs played crucial roles in various biological processes and regulated gene expression via diverse mechanisms in eukaryotes[14]. Genes could produce a polypeptide chain or functional RNA which consist a series of nucleotide sequences and support the basic structure and performance of life. Therefore, this is definitely an area where additional research input is needed in investigating the relevance of the expression levels of ARGs or autophagy-related lncRNAs for ccRCC.

In our research, six ARGs (EIF4EBP1, IFNG, BID, BIRC5, CX3CL1 and RAB24) and five autophagy-related lncRNAs (AC093278.2, AC010326.3, AC099850.3, AC016773.1 and AC009549.1) were considered closely

associated with patients' survival in ccRCC. IF4EBP1 acts as a significant effector in mTOR signaling pathway and the protein it encodes competitively binds to eukaryotic translation initiation factor 4E (EIF4E) to inhibit EIF4E complex assembly and thus blocking translation[15, 16]. IF4EBP1 is involved in the occurrence and progression of a variety of malignancies and is associated with poor survival. Cha et al. identified EIF4EBP1 were anomaly up-regulated in hepatocellular carcinoma, and the protein overexpression indicated a short survival and poor prognosis[17]. An autophagy-related prognostic signature including EIF4EBP1 also has correlation with distant metastasis-free survival for breast cancer[18]. Additionally, IF4EBP1 was found correlated with poor prognosis of ccRCC and was a risk factor for ccRCC survival in the model we constructed. Interferon gamma (IFNG) has been known as the regulator for both tumor immune surveillance and tumorigenesis. Dong et al. showed that immunity-related GTPase family member 1 was important for IFNG-mediated regulation of tumor cell growth in melanoma[19]. Also, specific IFNG was close association with autophagy and immune responses against *Mycobacterium tuberculosis*[20, 21]. Overexpression of the BH3-interacting domain death agonist (BID) protein sensitizes certain cancer cell lines to apoptosis induced by anticancer agents so that BID fused with TAT cell penetrating peptide and doxorubicin may be considered as a potential therapeutic method for advanced prostate cancer treatment[22]. However, the mechanisms underlying the resistance of tumor cell to BID have yet been fully understood. BIRC5 is a well-known cancer therapeutic target which has remained a central goal of survivin studies in the cancer field. It has been used in the field of gastric cancer[23], breast cancer[24], Leukemic[25] and so on. Besides, some previous studies had found the expression levels of BIRC5 was higher in RCC tissues than in adjacent normal tissues and was potential therapeutic targets for personalized medicine[26, 27]. CX3CL1 is a multifunctional inflammatory chemokine with a single receptor CX3CR1. Yao et al. revealed the expression of CX3CR1 was associated with the process of cellular migration in vitro and tumor metastasis of ccRCC in vivo by clinical and molecular cellular evidence[28]. RAB24 is an atypical RAB protein that promotes the process of late autophagy for long time. RAB24 could terminate the autophagic and endocytic pathways under nutrient-rich conditions[29]. Taken together, these findings suggest the six genes might be important regulators in ccRCC progression and might function via the involvement of the above processes and pathways. AC099850.3 has once been used to construct the lncRNA-miRNA-mRNA ceRNA network in squamous cell carcinoma of the tongue (SCCT) in order to benefit the diagnosis and treatment of SCCT[30]. Similarly, it was reported that AC016773.1 could clarify the pathogenesis of ccRCC and might provide potential therapeutic targets in laryngeal squamous cell carcinoma (LSCC)[31, 32]. The rest of three lncRNAs above are either poorly investigated or have not been reported, which means our findings suggested further research for them is imperative.

In the current study, the expression levels of ARGs and autophagy-related lncRNAs were extracted from TCGA by performing a genome-wide analysis to identify some molecular biomarkers associated with autophagy for detecting the prognosis of ccRCC patients. In order to improve its accuracy and credibility of the models, all patients were randomly divided into two equal portions, training cohort (n = 255) and testing cohort (n = 252). A total of 12 ARGs and 79 autophagy-related lncRNAs were found in the univariate Cox regression analysis. six ARGs (EIF4EBP1, IFNG, BID, BIRC5, CX3CL1 and RAB24) and five

autophagy-related lncRNAs (AC093278.2, AC010326.3, AC099850.3, AC016773.1 and AC009549.1) were used for constructing two OS-related prediction models via further Lasso regression analyses and multivariate Cox regression, respectively. The two prediction models were then classified ccRCC patients into the high-risk group and low-risk group with significantly different overall survival. K-M curve analysis demonstrated the competitive performance of the two models for predicting 5-year survival in the training dataset and testing dataset. Next, the prognostic power of the prediction models was further validated in ROC curves and independent prognostic analyses (age, gender, grade, T-staging and M-staging), demonstrated good reliability and reproducibility of the two models and even independent prognostic indicators in predicting ccRCC patients' survival. Finally, GO and KEGG analyses suggested that these genes were functionally associated with regulation of endopeptidase activity, regulation of peptidase activity, autophagy, human cytomegalovirus infection, shigellosis, autophagy-animal and HIF-1 signaling pathway. Consequently, further functional characterization of the six ARGs and five autophagy-related lncRNAs would be beneficial to unraveling the underlying mechanism of ccRCC.

To the best of our knowledge, this is the first study that compares gene and lncRNA models for predicting outcome of ccRCC patients. The two models were capable to stratify ccRCC patients into two risk groups with significantly different OS, showing greater prognostic power and having a chance to be routinely used as candidate biomarkers in the clinical work. However, our results indicated the six-gene prognostic risk model is superior to the lncRNA one through a series of verification methods in a way. The main probable reason is that the gene encode proteins directly involved in biological process whereas the lncRNA's main function is to regulate the coding potential of the genome. It is also the reason why the cancer genes were first discovered and described in 1991 but lncRNAs have just been studied for a few years[33]. In any case, all the two models are particular excellent so that we suggest ccRCC patients can determinate the five ARGs and six lncRNAs to improve the accuracy and operability of prognosis evaluation.

However, there were some deficiencies needed to be corrected in the future. Firstly, the data only came from TCGA instead of combination of information from multiple databases. Secondly, the deeper mechanisms why ARGs modulate the initiation and progression of ccRCC have not been elucidated. Therefore, additional work is needed to make up for these shortcomings and assess the efficiency of the two models.

In brief, a total of six OS-related ARGs (EIF4EBP1, IFNG, BID, BIRC5, CX3CL1 and RAB24) and five autophagy-related lncRNAs (AC093278.2, AC010326.3, AC099850.3, AC016773.1 and AC009549.1) were identified in our current study. The two models can be used as independent prognostic biomarkers in stratifying risk subgroups in terms of survival for patients with ccRCC. The gene model had similarly super performance for risk stratification compared to the lncRNA-related model. Moreover, further research of these pivotal genes and lncRNAs may guide the tailored therapy of ccRCC.

## Declarations

## Acknowledgments

None

## Funding

This study was supported by Natural Science Foundation of Fujian Province (Grant number: 2018J01172), Startup Fund for scientific research, Fujian Medical University (Grant number: 2017XQ1063), Fujian educational research programs for young and middle-aged teachers (Grant number: JT180187).

## Author Contributions

Conceptualization, Xiong-Lin Sun and Ning Xu; Data curation, Zhi-Bin Ke, Xiao-Dong Li and Hai Cai; Software, Fei Huang; Supervision, Qing-Shui Zheng and Xue-Yi Xue; Writing – original draft, Xiong-Lin Sun, Jun-Ming Zhu, Fei Lin and Zhi-Bin Ke; Writing – review & editing, Shao-Hao Chen and Ning Xu.

## Conflict of interests

All authors declare no conflict of interests.

## Data Availability Statement

All data generated or analyzed during the present study was downloaded from TCGA and HADb database.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Not applicable.

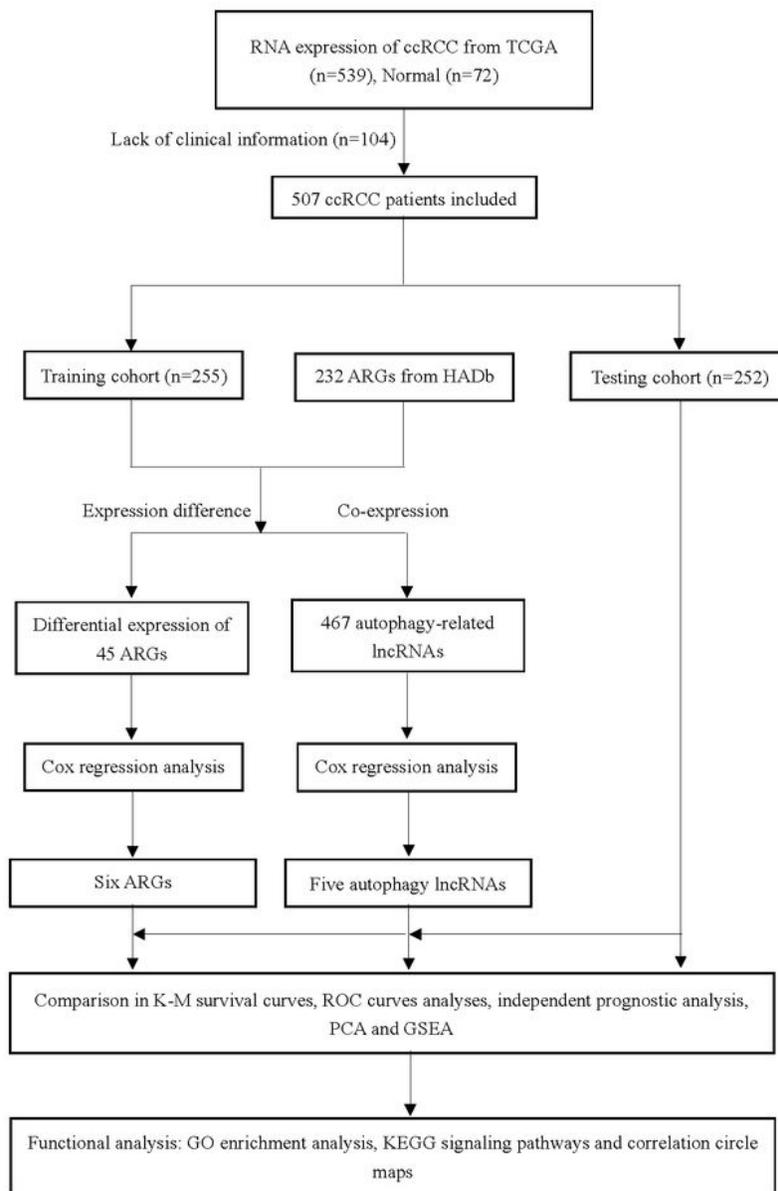
## References

1. Cairns P. Renal cell carcinoma. *Cancer biomarkers: section A of Disease markers*. 2010;9(1–6):461–73.
2. Siegel RL, et al. *Cancer statistics, 2020*. *CA: a cancer journal for clinicians*. 2020;70(1):7–30.
3. Chen WQ, et al. [Report of Cancer Incidence and Mortality in China, 2014]. *Zhonghua zhong liu za zhi [Chinese journal of oncology]*. 2018;40(1):5–13.
4. Dalbeni A, et al. Effects of Antiangiogenic Drugs on Microcirculation and Macrocirculation in Patients with Advanced-Stage Renal Cancer. *Cancers*. 2018;11(1).

5. Alyaev YG, et al. [Computer-assisted laparoscopic surgery for renal carcinoma]. *Urologiia (Moscow, Russia: 1999)*. 2018(3):30–38.
6. Rini BI, et al. The society for immunotherapy of cancer consensus statement on immunotherapy for the treatment of advanced renal cell carcinoma (RCC). *J Immunother Cancer*. 2019;7(1):354.
7. Levine B, Kroemer G. Autophagy in the pathogenesis of disease. *Cell*. 2008;132(1):27–42.
8. Singla M, Bhattacharyya S. Autophagy as a potential therapeutic target during epithelial to mesenchymal transition in renal cell carcinoma: An in vitro study. *Biomedicine pharmacotherapy = Biomedecine pharmacotherapie*. 2017;94:332–40.
9. Haas NB, et al. Autophagy Inhibition to Augment mTOR Inhibition: a Phase I/II Trial of Everolimus and Hydroxychloroquine in Patients with Previously Treated Renal Cell Carcinoma. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2019;25(7):2080–87.
10. Kong YL, et al. Expression of autophagy related genes in chronic lymphocytic leukemia is associated with disease course. *Leukemia research*. 2018;66:8–14.
11. Liu CY, et al. LncRNA CAIF inhibits autophagy and attenuates myocardial infarction by blocking p53-mediated myocardin transcription. *Nature communications*. 2018;9(1):29.
12. Liu XD, et al. Autophagy mediates HIF2 $\alpha$  degradation and suppresses renal tumorigenesis. *Oncogene*. 2015;34(19):2450–60.
13. Messai Y, et al. The autophagy sensor ITPR1 protects renal carcinoma cells from NK-mediated killing. *Autophagy*. 2015:0.
14. Zhu J, et al. Function of lncRNAs and approaches to lncRNA-protein interactions. *Science China Life sciences*. 2013;56(10):876–85.
15. Harris TE, Lawrence JC. Jr. TOR signaling. *Science's STKE: signal transduction knowledge environment*. 2003;2003(212):re15.
16. Gingras AC, et al. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annual review of biochemistry*. 1999;68:913–63.
17. Cha YL, et al. EIF4EBP1 overexpression is associated with poor survival and disease progression in patients with hepatocellular carcinoma. *PloS one*. 2015;10(2):e0117493.
18. Gu Y, et al. Autophagy-related prognostic signature for breast cancer. *Molecular carcinogenesis*. 2016;55(3):292–9.
19. Dong H, et al. IFN $\gamma$ -induced Irgm1 promotes tumorigenesis of melanoma via dual regulation of apoptosis and Bif-1-dependent autophagy. *Oncogene*. 2015;34(42):5363–71.
20. Rovetta AI, et al. IFN $\gamma$ -mediated immune responses enhance autophagy against *Mycobacterium tuberculosis* antigens in patients with active tuberculosis. *Autophagy*. 2014;10(12):2109–21.
21. Wu S, et al. Genetic variants in IFN $\gamma$  and IFNGR1 and tuberculosis susceptibility. *Cytokine*. 2019;123:154775.
22. Orzechowska EJ, et al. Synergy of BID with doxorubicin in the killing of cancer cells. *Oncol Rep*. 2015;33(5):2143–50.

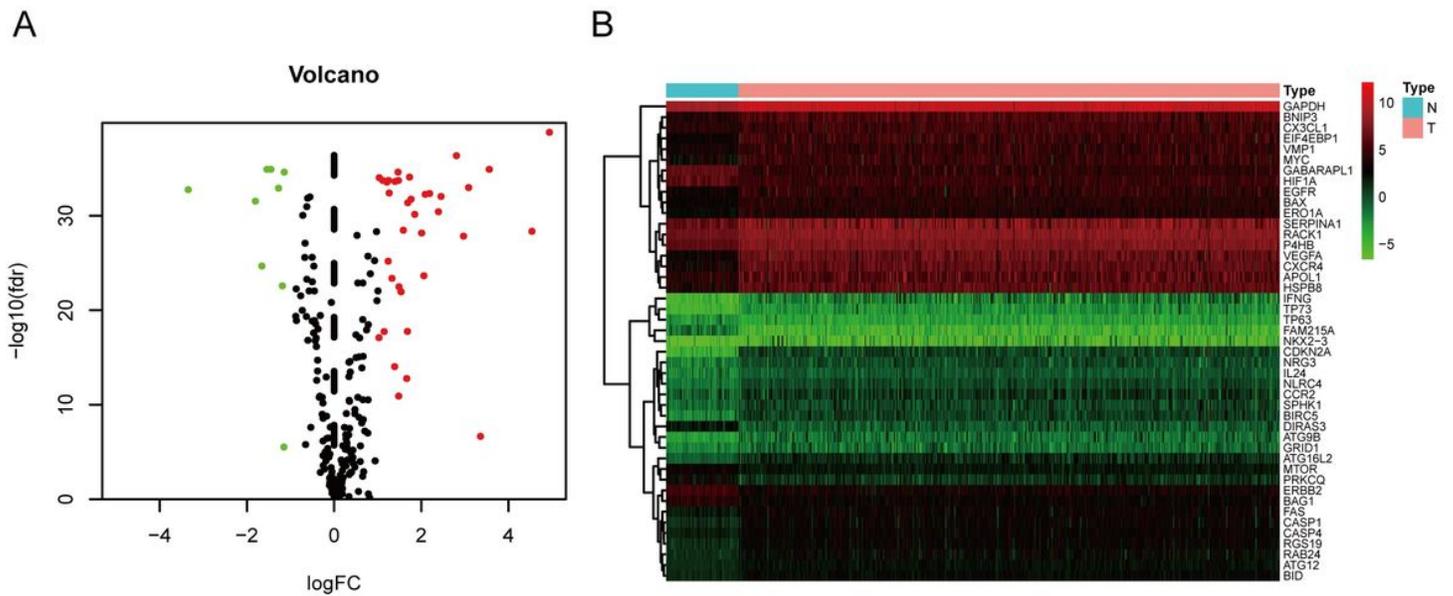
23. Zou J, et al. Dysregulation of miR-195-5p/-218-5p/BIRC5 axis predicts a poor prognosis in patients with gastric cancer. *J Biol Regul Homeost Agents*. 2019;33(5):1377–85.
24. Sušac I, et al. Polymorphisms in Survivin (BIRC5 Gene) Are Associated with Age of Onset in Breast Cancer Patients. *Journal of oncology*. 2019;2019:3483192.
25. Narimani M, et al. Knockout Of BIRC5 Gene By CRISPR/Cas9 Induces Apoptosis And Inhibits Cell Proliferation In Leukemic Cell Lines, HL60 And KG1. *Blood and lymphatic cancer: targets and therapy*. 2019;9:53–61.
26. Zhang H, et al. MALAT1 accelerates the development and progression of renal cell carcinoma by decreasing the expression of miR-203 and promoting the expression of BIRC5. *Cell proliferation*. 2019;52(5):e12640.
27. Robinson RL, et al. Comparative STAT3-Regulated Gene Expression Profile in Renal Cell Carcinoma Subtypes. *Frontiers in oncology*. 2019;9:72.
28. Yao X, et al. Expression of CX3CR1 associates with cellular migration, metastasis, and prognosis in human clear cell renal cell carcinoma. *Urol Oncol*. 2014;32(2):162–70.
29. Ylä-Anttila P, et al. RAB24 facilitates clearance of autophagic compartments during basal conditions. *Autophagy*. 2015;11(10):1833–48.
30. Zhou RS, et al. Integrated analysis of lncRNA-miRNA-mRNA ceRNA network in squamous cell carcinoma of tongue. *BMC Cancer*. 2019;19(1):779.
31. Wang J, et al. Construction and comprehensive analysis of dysregulated long non-coding RNA-associated competing endogenous RNA network in clear cell renal cell carcinoma. *Journal of cellular biochemistry*. 2018.
32. Kong X, et al. Comprehensive analysis of differentially expressed profiles of lncRNAs, mRNAs, and miRNAs in laryngeal squamous cell carcinoma in order to construct a ceRNA network and identify potential biomarkers. *Journal of cellular biochemistry*. 2019;120(10):17963–74.
33. Chen X, et al. Expression and prognostic relevance of MAGE-A3 and MAGE-C2 in non-small cell lung cancer. *Oncology letters*. 2017;13(3):1609–18.

## Figures



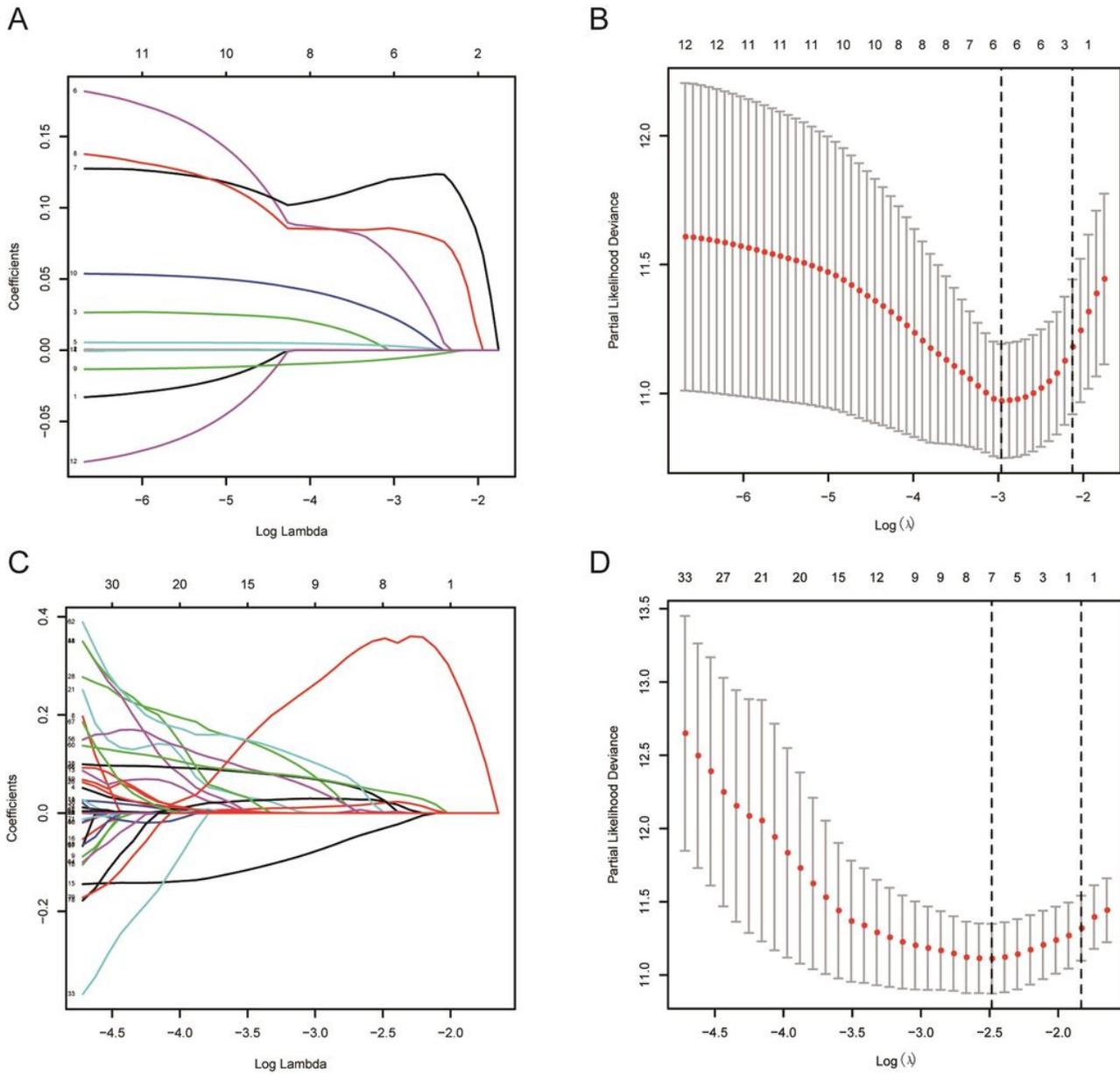
**Figure 1**

A total of six differentially expressed ARGs and five autophagy-related lncRNAs were brought into final prognostic models.



**Figure 2**

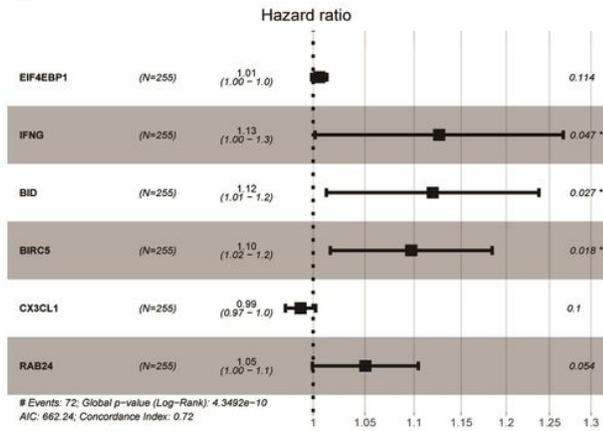
Considered as the criteria of  $\text{FDR} < 0.05$  and  $|\log_2(\text{FC})| > 1$ , 45 differentially expressed genes were finally obtained between tumor and non-tumor tissues including 36 up-regulated and 9 down-regulated ARGs



**Figure 3**

Lasso regression analyses of those screened genes

A



B

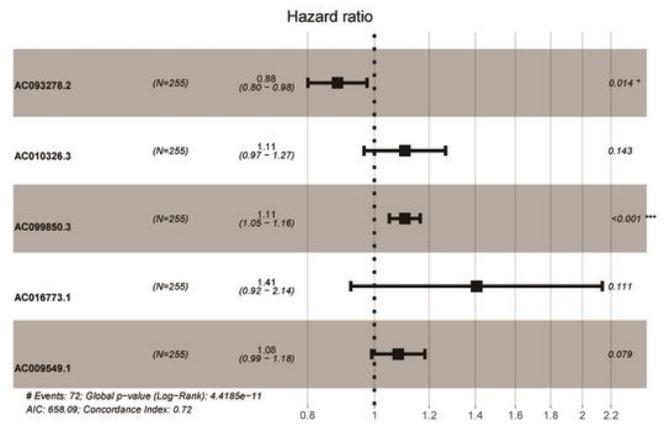
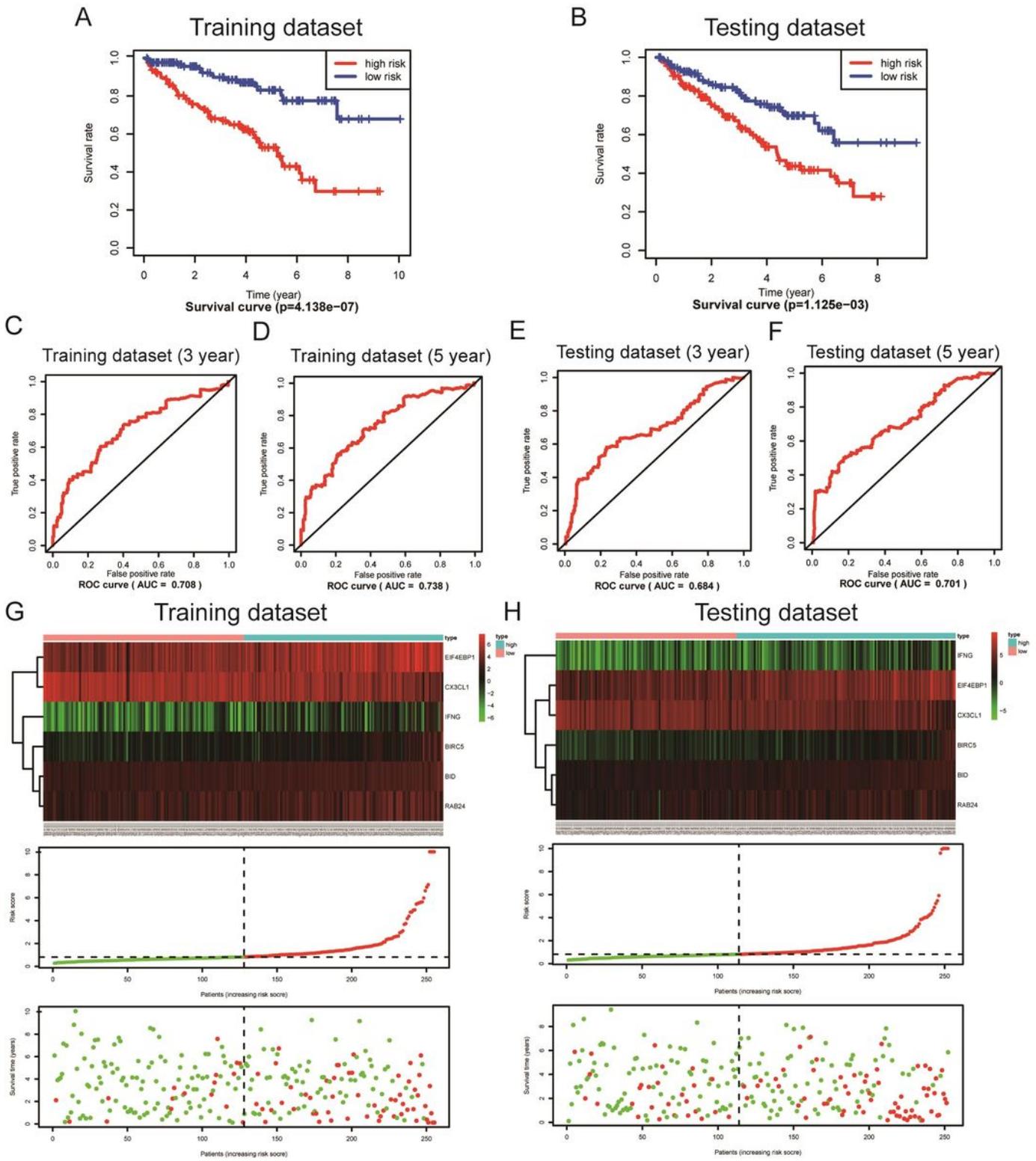


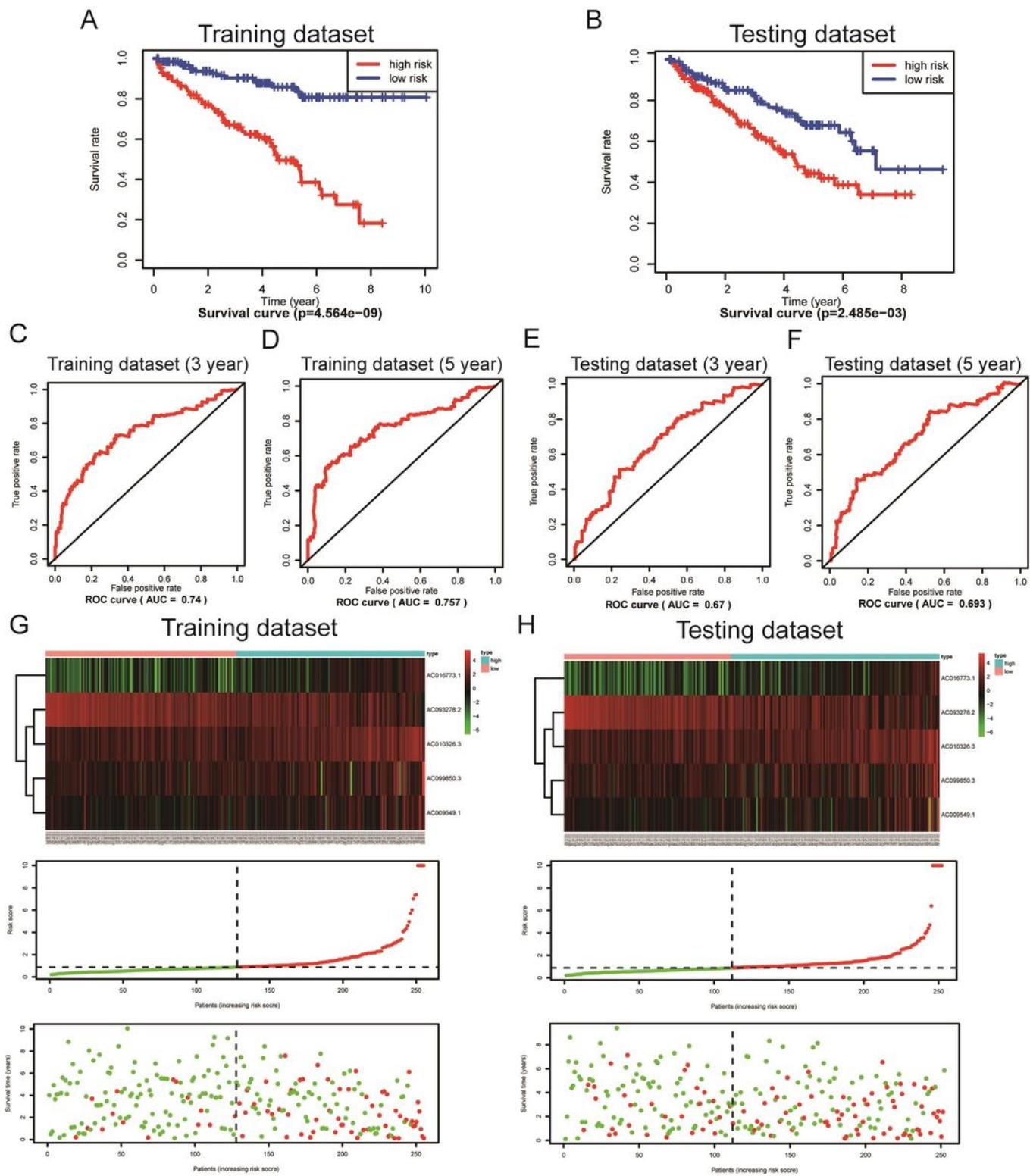
Figure 4

multivariate Cox regression model to avoid overfitting the prognostic problems. Six ARGs



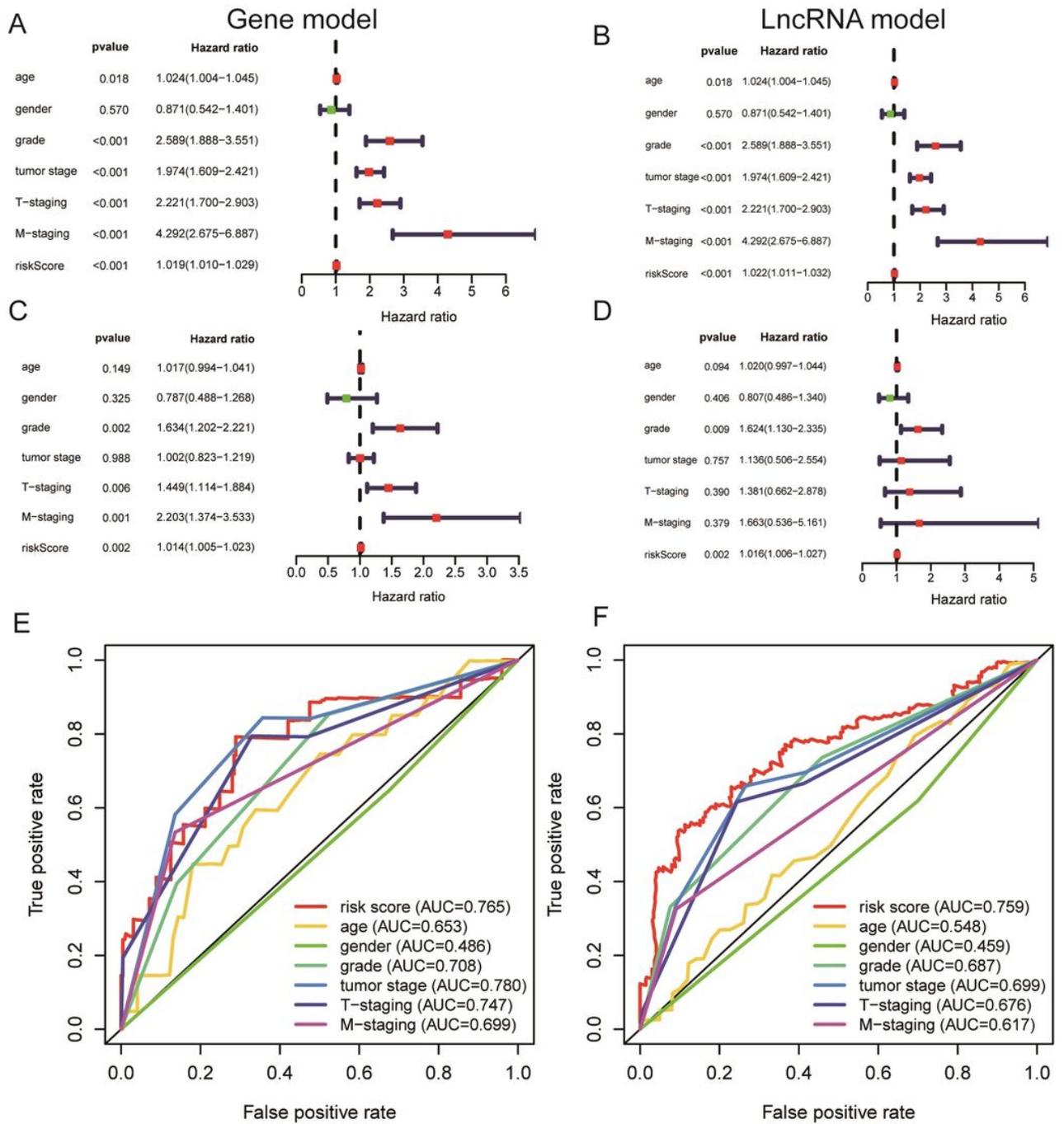
**Figure 5**

lower mortality rate than high-risk group



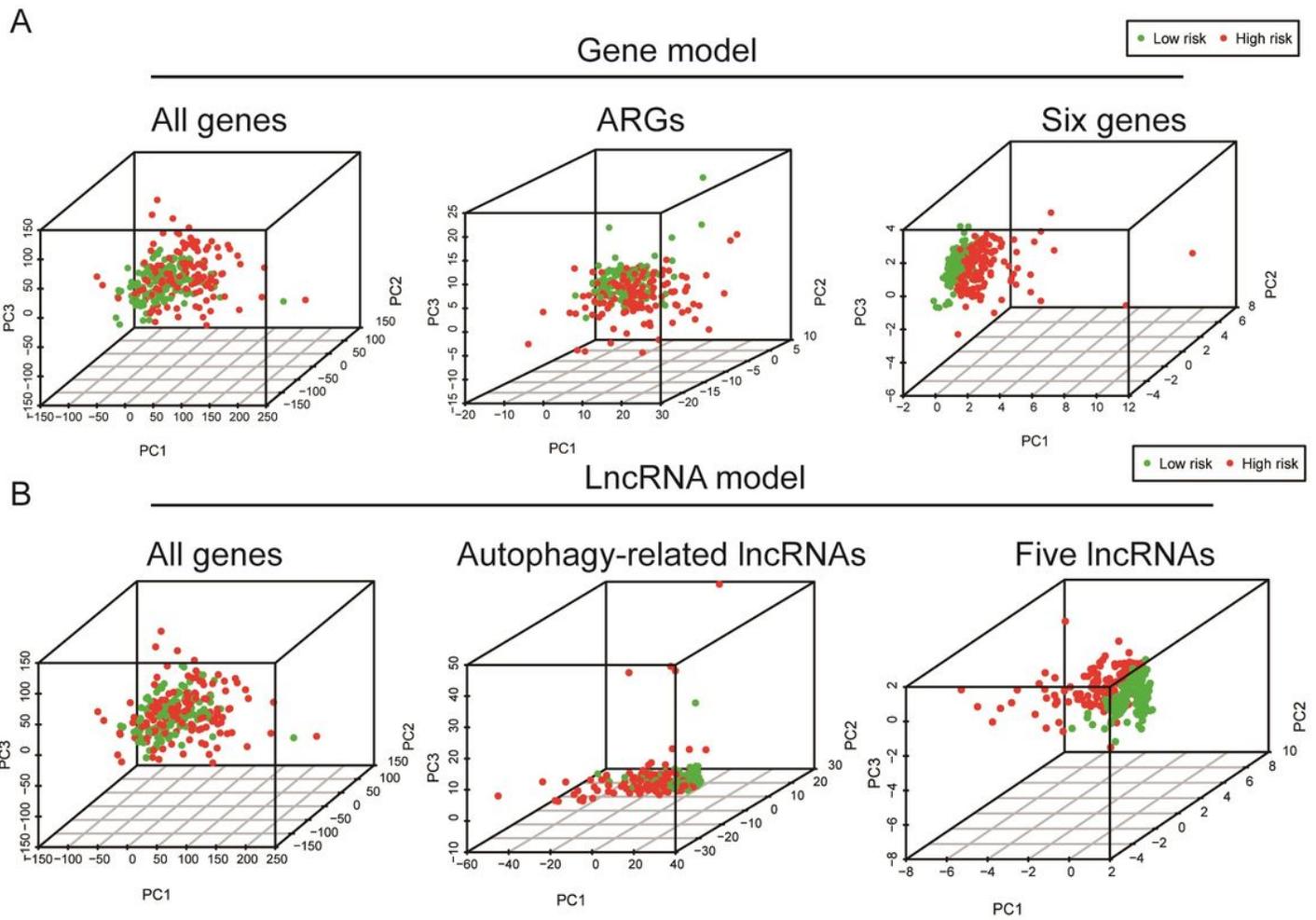
**Figure 6**

The heatmap of expression profiles, the risk scores and distribution and patients' survival status in the TCGA database were also performed in testing dataset



**Figure 7**

The univariate Cox regression analyses demonstrated six-gene prognostic risk model

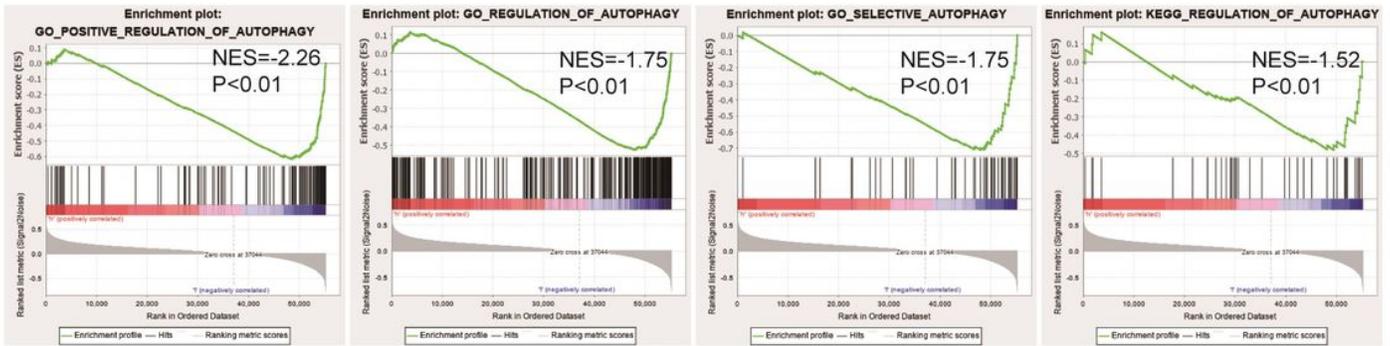


**Figure 8**

PCA were performed for all genes extracted from TCGA, ARGs and autophagy-related lncRNAs and six genes and five lncRNAs selected finally under the analyses of two models, respectively

A

## Gene model



B

## LncRNA model

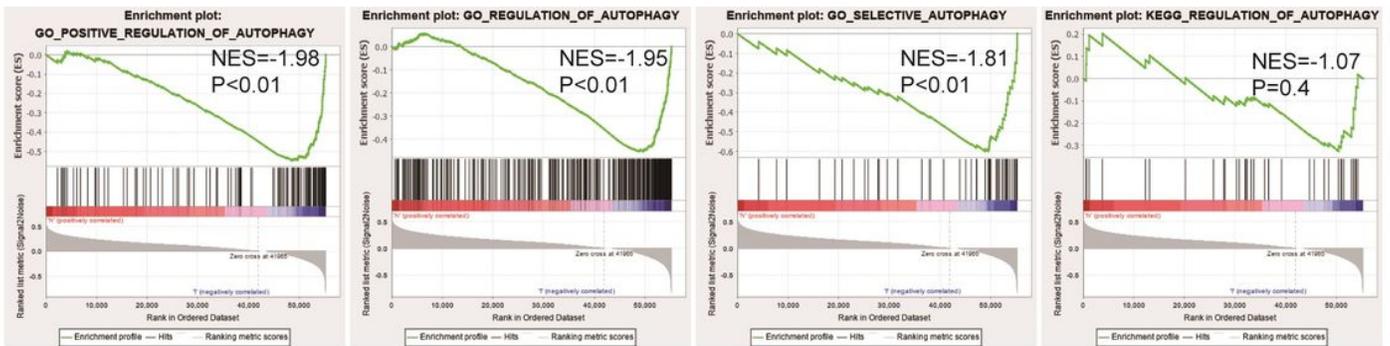
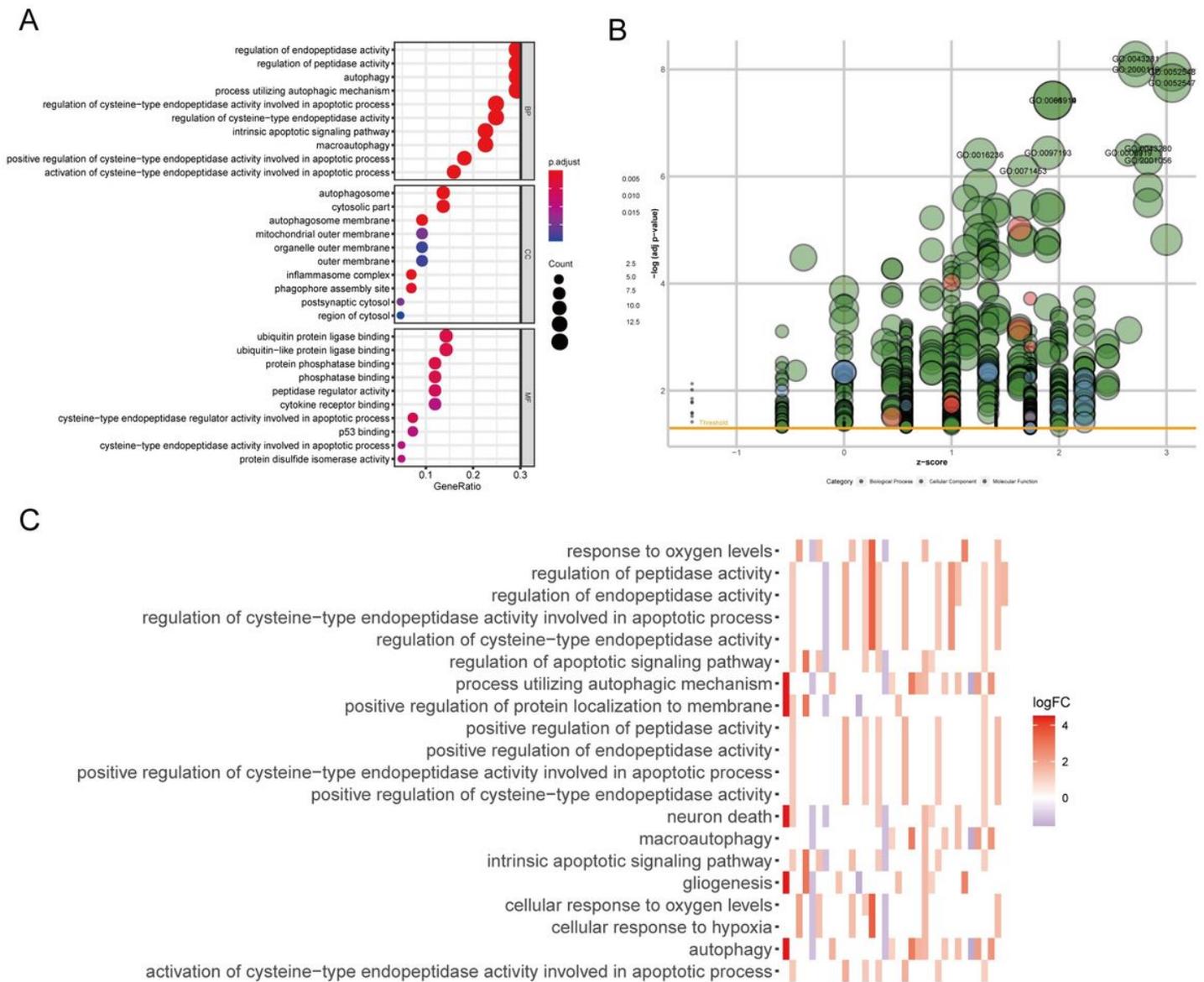


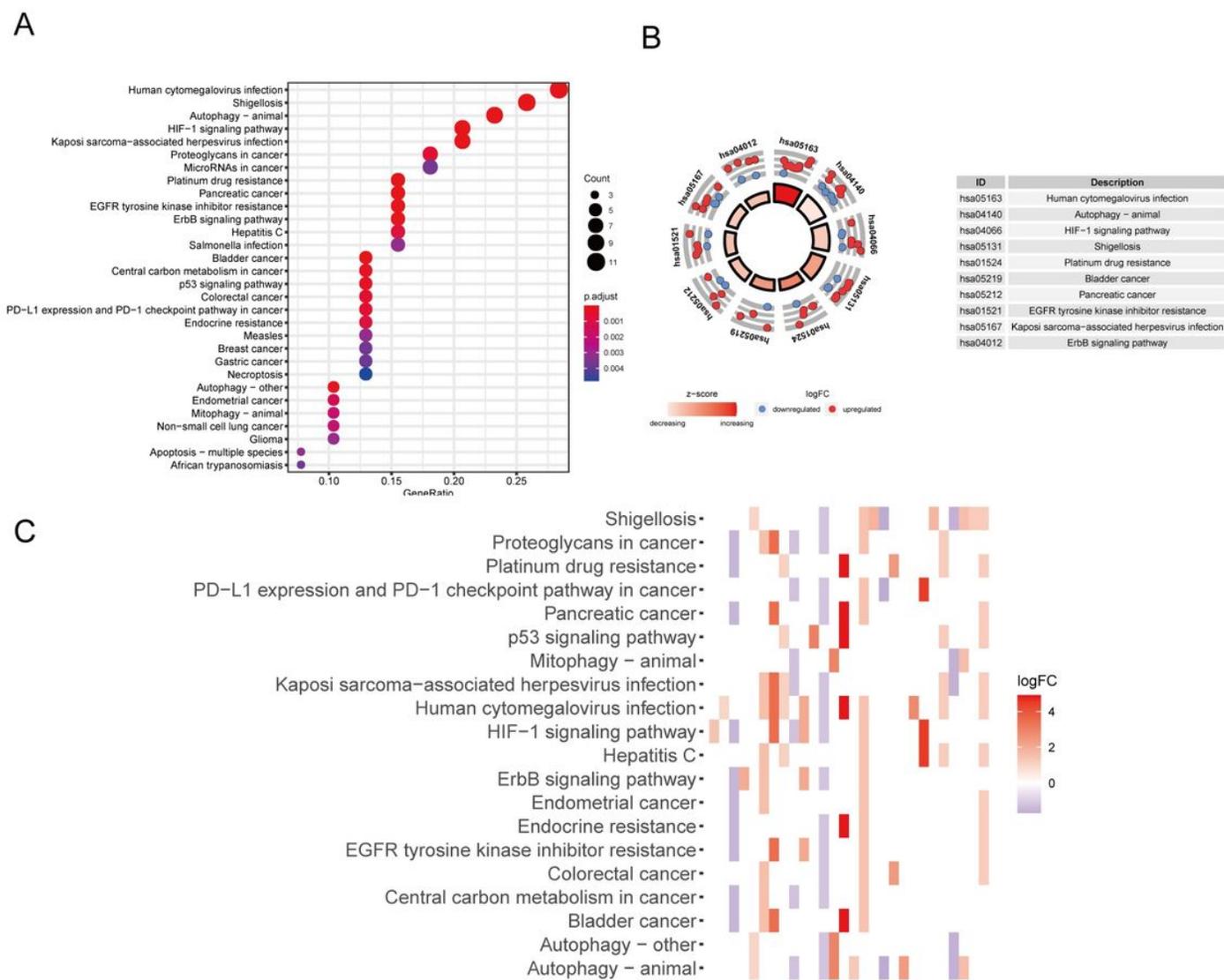
Figure 9

The results showed that the two gene sets (go positive regulation of autophagy and kegg regulation of autophagy) had a greater enrichment in the low-risk group for six-gene prognostic risk model while another two gene sets (go selective autophagy and go regulation of autophagy) indicated a better enrichment in the five-lncRNA prognostic risk model



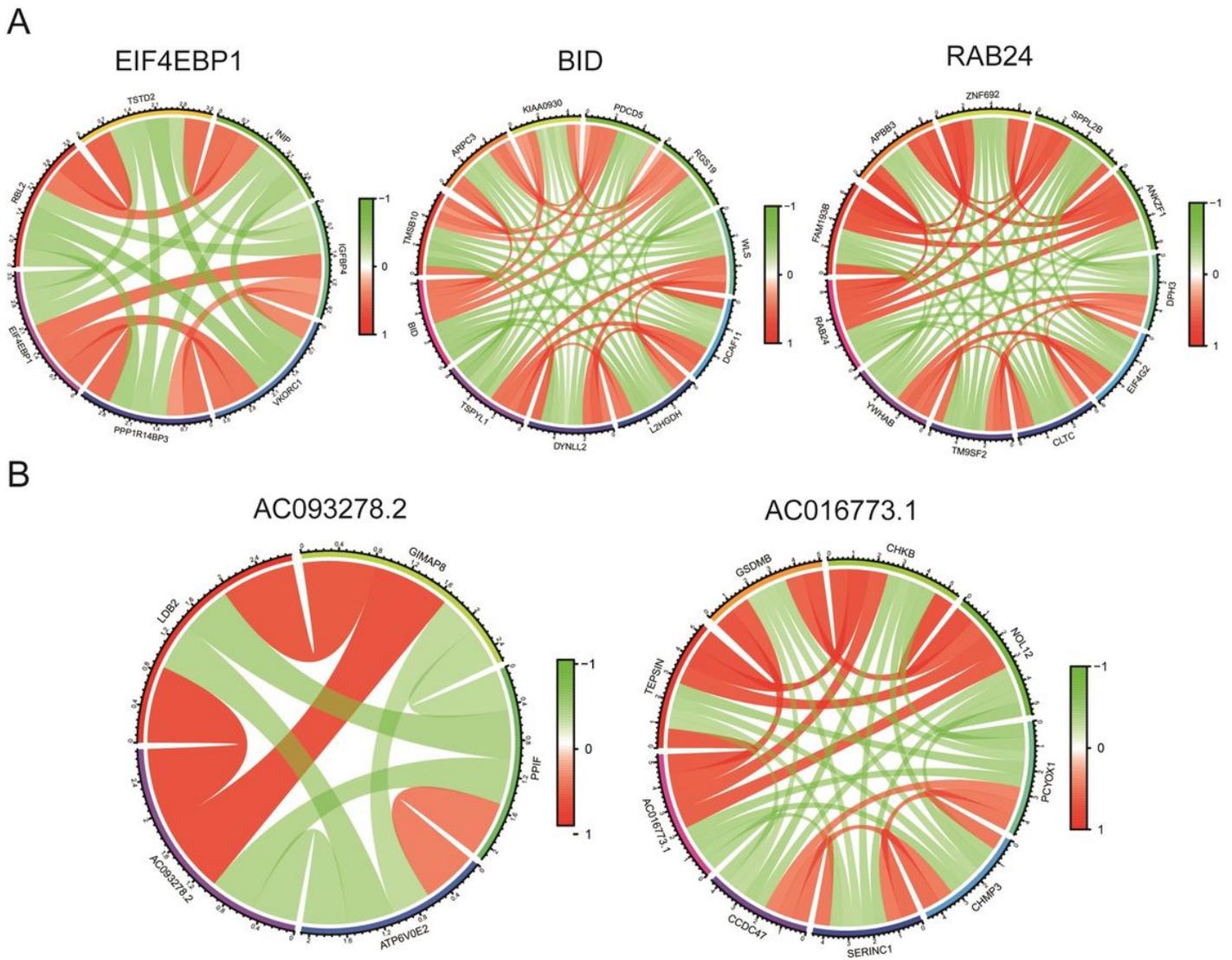
**Figure 10**

The ARGs were found significantly related to top GO biological processes according to the results of DAVID, including regulation of endopeptidase activity, regulation of peptidase activity, autophagy and process utilizing autophagic mechanism



**Figure 11**

The heatmap of the relationship between ARGs and GO enrichment analysis was also displayed. Moreover, 5 KEGG signaling pathways functionally involved ARGs, consisting of human cytomegalovirus infection, shigellosis, autophagy-animal, HIF-1 signaling pathway and Kaposi sarcoma-associated herpesvirus infection



**Figure 12**

Correlation circle maps were plotted based on genes which had both positive correlation and negative correlation genes