# Prognostic and Predictive Values of the Autophagy Signature in Colon Cancer

**Qian Xu** ( xuqian7666@nynu.edu.cn )
Nanyang Normal University https://orcid.org/0000-0001-7061-7282
**Yugang Guo**
Nanyang Normal University
**Jintao Fang**
Nanyang Normal University
**Jiawei Zhou**
Nanyang Normal University
**Guohui Ma**
Nanyang Normal University
**You Heng**
Nanyang Normal University

**Primary research**

**Keywords:** Autophagy, Colon cancer, Signature, Nomogram

**Posted Date:** September 4th, 2020

**DOI:** https://doi.org/10.21203/rs.3.rs-63586/v1

# Abstract

**Background:** In the clinical decision-making among patients with colon cancer (COAD), making an accurate prognosis of the patients plays a central role. The effects of autophagy on the clinical outcomes of cancer, including COAD, have been widely reported in numerous studies. Here, weaim to build a novel autophagy-associated, risk-stratification scoring system to predict the overall survival(OS)of patients with COAD.

**Methods:** In this study, the candidate autophagy-related prognostic genes correlated with the survival of COAD patients from The Cancer Genome Atlas (TCGA) public RNA microarray and clinical data sets were selected as training data set. A cohort of 67 patients from TCGA and a cohort of 124 patients from GEO were used for the external validation. The autophagy-related mRNAs(ARGs) were analyzed by multivariate Cox regression analyses. Spearman correlation analysis were used to construct autophagy-related mRNAs and lncRNAs coexpression network.

**Results:** 6 autophagy-related mRNAs and 14 lncRNAs with prognostic value were extracted for constructing two novel autophagy-related RNAs signatures, respectively. Univariate and multivariate Cox regression analyses were then demonstrated that the two signature could act as independent prognostic predictor for OS. Additionally, a prognostic nomogram incorporating the clinicopathological characteristics(patient's age, tumor stage) and autophagy-related lncRNA risk score was constructed to predict the OS, which was used in the training and validation sets (5-year C-index: 0.826 and 0.895, respectively), demonstrating better discrimination ability and clinical net benefit than the risk score model. Further gene set enrichment analysis revealed that autophagy-associated lncRNAs were significantly enriched in cancer-related pathways.

**Conclusions:** The identified autophagy-related mRNAs and lncRNAs signature had important clinical implications in prognosis prediction and the user-friendly nomogram may offer an extra insight for individualized therapy of COAD.


# Background

Colon adenocarcinoma is a commonly diagnosed, malignant type of epithelial cell tumor and also one of the leading causes of mortality worldwide [1]. Despite the advances in the diagnosis and treatment strategies, the 5-year survival rate of patients with COAD remains below 50% [2], with loco-regional invasion and metastasis being the major causes of death [3, 4]. Although the histopathological classification of COAD is widely recognized, the clinicopathological and genetic factors are still unable to evaluate the survival outcomes accurately. Moreover, patients having similar risk factors might have conflicting outcomes. Consequently, a more comprehensive study is still needed in order to effectively identify the molecular biomarkers and increase the prognostic and predictive accuracy of the current assessment system.

Autophagy, which is a major catabolic process within the lysosomes, maintains the metabolic homeostasis and cell survival under metabolic pressures (e.g., energy deficiency and starvation) as well as physiological and pathological conditions (e.g., aging, apoptosis, and cancer)[5]. Its role is complex and differs among various types of cancer, such as inhibiting tumor initiation and progression in some cancers and promoting tumor survival and progression in others, making it as a potential therapeutic target for cancer [6]. In recent years, many studies have investigated the autophagic pathways in order to develop new potential targeted therapies [7–11]. In addition, autophagic drugs tend to induce cell autophagic death (type II cell death) and cause colon cancer cell death [12–15]. For instance, in the phase 1/2 clinical trials, NVP-BEZ235 was observed to induce colon cancer cell apoptosis and autophagy simultaneously [16, 17]. However, previous studies that have investigated the role of autophagy in tumorigenesis have analyzed a limited number of autophagy-related genes in either cell lines or animal models. Hence, the prognostic value of the global expression patterns of autophagy-related genes was not determined.

Long non-coding RNAs (lncRNAs), which play a role in multiple biological processes, are involved in regulating multiple ways of colon cancer progression and modulating the transcription of downstream cancer-related genes [18–20]. They are also known to be important regulatory factors in modulating autophagy and tumor progression [21]. Recent reports also stated that lncRNAs mediate the autophagy in COAD [22]. However, their potential regulatory roles in the autophagy and tumor progression in COAD need to be examined further.

Considering the autophagy-associated mRNAs (ARGs) and lncRNAs, we undertook a systematic and comprehensive biomarker discovery and validation effort to develop two risk signatures, which better predict prognosis performance than the traditional factors for COAD patients. The nomogram incorporated the autophagy-related lncRNA signature and clinical factors to predict the OS of COAD patients. In addition, we assessed the prognostic and predictive accuracy of this model in the validation set. Moreover, function and pathway enrichment analyses, as well as a co-expression network analysis of the lncRNAs and mRNAs associated with OS, were performed to provide hints concerning the roles of the prognostic genes in the development and pathophysiology of COAD.

# Materials And Methods

### LncRNA and autophagy gene screening and patient samples

A total of 232 genes from the HADb (Human Autophagy Database http://autophagy.lu/clustering/index.html) were identified as an up-to-date list of ARGs. The profiles of lncRNAs and autophagy genes and clinical information of COAD patients (colon cancer, 41 normal samples and 473 tumor samples) were acquired from the TCGA databases. A total of 124 samples from the GEO dataset (accession number: GSE72970) and 67 samples from TCGA (colon cancer primary site rectosigmoid juction) were used for the external validation. The correlation between the autophagy-related lncRNAs and ARGs was calculated using Pearson correlation coefficient. A lncRNA with a

correlation coefficient |R| > 0.3 and $P < 0.01$ was considered to be an autophagy-related lncRNA. A full mutation dataset of corresponding patients was obtained from the cBioPortal (http://www.cbioportal.org). Since our data was obtained from the TCGA and GEO databases, no ethics committee approval was required.

### Gene ontology and KEGG analysis

A functional enrichment of differentially expressed ARGs is performed using the Bohao Online Enrichment Tool (http://enrich. shbio.com/). Gene Ontology (GO) and the Kyoto Gene and Genomic Encyclopedia (KEGG) were used to assess relevant functional categories. GO and KEGG enrichment pathways with $P$ values less than 0.05 are considered to be significant categories.

### Survival analysis of autophagy-related mRNAs and lncRNAs

The clinical datasets of the COAD cohort were downloaded from the TCGA database. 426 samples with a survival duration of $t > 0$ days were retained for the survival analysis. A univariate Cox regression analysis was used in the R software to evaluate whether autophagy-related RNA was correlated with OS. RNAs with a hazard ratio (HR) < 1 were defined as a protective signature, while RNAs with a HR > 1 were defined as risky RNAs. In addition, GeneMANIA (http://www.genemania.org) was used to build protein–protein interaction (PPI) networks of the autophagy-related genes correlated with OS.

### Construction of the prognostic model of autophagy-related mRNAs and lncRNAs

Prognosis-related mRNAs and lncRNAs were constructed using multivariate Cox regression. After incorporating the expression values for each particular RNA, the risk score formula for each patient was constructed and weighted based on its estimated regression coefficients. The median risk score was used as the cutoff point according to the risk scoring formula, and the patients were divided into two groups: low-risk group and high-risk group. The survival differences between the two groups were assessed using the Kaplan–Meier curve and compared using log-rank statistical methods. The role of the risk scores in predicting patient outcomes was examined using a multivariate Cox regression analysis and stratified analysis. The prediction accuracy of the risk model was determined by a time-dependent ROC analysis. Moreover, the efficacy was also validated by the data from the GEO database.

### Univariate and multivariate cox regression analyses

To detect whether the clinical characteristics and autophagy-related RNA risk score were significantly associated with OS in COAD patients, univariate and multivariate Cox regression analyses were performed. The prognostic factors in the training set were selected with the significance threshold of log-rank $P < 0.05$. Moreover, the hazard ratio and 95% confidence intervals for each variable were calculated.

### Nomogram survival rate model for independent prognostic factor

To further investigate the correlations between the independent prognostic factors and OS, we incorporated the identified independent prognostic factors with the predicted risk information in the prediction prognosis model in order to construct a nomogram that predicts the 3-, or 5-year OS using the rms package in the R software.

### Gene set enrichment analysis (GSEA)

GSEA was performed using the GSEA4.0.3 program (https://www.gsea-msigdb.org/gsea/index.jsp), in which the pathway enrichment of lncRNA was shown. In this study, we verified whether the genes that are differentially expressed between the two groups are enriched during autophagy.

## Results

### Identification of prognostic ARGs

Figure 1 shows the flow chart of our research process. After analyzing the expression profiles of 232 ARGs in the COAD tumor tissues and comparing them with the normal tissues, 16 upregulated and 20 downregulated ARGs were obtained, according to the following criteria: false discovery rate (FDR) < 0.05 and |log2 (fold change) |>1(Figs. 2A,B). Then, analysis of Kaplan-Meier was performed to determine the ability of the ARGs for OS to predict the prognosis of COAD patients, and 15 ARGs were screened. The forest map of the hazard ratio indicates that most of these genes are oncogene (HR > 1), except for SERPINA1 (Fig. 2C). Furthermore, the PPI analysis showed that these genes are closely related to autophagy, pre-autophagosomal structure, macroautophagy, autophagic vacuole assembly, cellular response to starvation, response to starvation, and cellular response to nutrient levels (Fig. 2D). Finally, given the important clinical implications of these ARGs, the genetic alterations of these genes were examined, and deep deletion and amplification were classified as two common types of mutations. A total of four genes have a mutation rate ≥ 4%, in which GRID1 is the most frequently mutated gene (6%) (Fig. 2E).

### Go and KEGG enrichment analysis of the differentially expressed ARGs

Then, the potential biological processes and pathways of the differentially expressed ARGs between COAD and non-tumor tissues were investigated. The GO term functional enrichment ($P < 0.05$) and the KEGG pathway enrichment analyses of these genes are summarized ($P < 0.01$) in Additional file 1: Fig. S1. The top 5 GO terms for biological processes include “autophagy,” “process utilizing autophagic mechanism,” “response to oxygen level,” “intrinsic apoptotic signaling pathway,” and “macroautophagy” (in decreasing order of *p* value). For the cellular components of the enrichment analysis, the top 5 terms were notably associated with “autophagosome,” “vacuolar membrane,” “autophagosome membrane,” “outer mitochondrial membrane,” and “membrane organelle outer membrane.” On the basis of molecular function, among the top 5 terms, the genes were mostly enriched in “ubiquitin protein ligase binding,” “ubiquitin-like protein ligase binding,” “protein kinase regulator activity,” “protein kinase regulator activity,” and “protein phosphatase 2A binding.” In addition, 30 KEGG pathways were considered statistically

significant ($P < 0.01$), and the top 3 significant pathways were “p53 signaling pathway,” “apoptosis,” and “human cytomegalovirus.” Overall, autophagy played an important role in the pathogenesis of COAD.

### Identification of prognostic autophagy-related LncRNA

lncRNAs dominate the upstream portion of the RNA network and function as primary effectors of the mRNAs. A gene co-expression network analysis was performed using the Pearson correlation with |R|>0.3 and $P < 0.01$ as the cutoff point. A total of 943 lncRNAs were obtained in proportion to the 36 ARGs (Additional file 5 : Table S1). Only 369 of 943 autophagy-related lncRNAs were significantly differentially expressed between patients with COAD and normal tissues (|log2FC|>1; $P < 0.05$; Figs. 3A,B; Additional file 6 : Table S2). Subsequently, 36 of 369 different expressed lncRNAs were found to be associated with OS of COAD patients, of which five lncRNAs (AP001554.1, LINC00513, SNHG16, AL137782.1, and AL590483.1) were protective genes with an HR < 1 and the remaining 31 lncRNAs were risky genes with an HR > 1 (Fig. 3C).

### Construction of the ARGs-based prognostic signature

Based on the 15 ARGs that were significantly correlated with OS in COAD patients, an autophagy-related risk signature was considered in predicting the prognosis. After multivariate Cox regression analysis, 6 ARGs which consisted of GRID1, DAPK1, RAB7A, PELP1, ULK3, and WIPI2 were identified to construct a prognostic signature for OS. Subsequently, with the expression coefficient of each independent risk gene, our six-mRNA models were formed using the following formula: prognosis score = (0.524 × expression level of UKL3) + (1.022 × expression level of GRID1) + (0.254 × expression level of DAPK1) + (0.410 × expression level of PELP1)+ (0.608 × expression level of WIP2) + (1.097 × expression level of RAB7A). Using this formula, we calculated the risk score of each patient. By using the median risk score as the threshold, we divided the patients into two groups: high-risk groups and low-risk groups (Figs. 4A,B). The heatmap of these six ARGs and the Kaplan–Meier analysis of the different OS between the two groups are displayed in Fig. 4C,D. Notably, our data showed that the high-risk group (risk scores ≥ 0.954; n = 213) had a worse prognosis (shorter OS) than the low-risk group (risk scores < 0.954; n = 213). The area under the curve (AUC) of the corresponding ROC curve for the 5-year OS is 0.728 in the training group (Fig. 4E). A similar trend was observed with a 5-year OS AUC of 0.733 in the validation group (GSE72970) from the GEO database. ($P < 0.001$; Figs. 4F,G). This indicated that the 6 mRNA-prognostic scoring system for COAD based on ARGs has a certain potential in survival prediction.

### Construction of autophagy related lncRNA-based prognostic signature

Based on the 36 differentially expressed autophagy-related lncRNAs between the tumor and normal tissues that were significantly correlated with the OS of COAD patients, multivariate Cox regression analyses were performed to select the potential prognosis-related lncRNAs. As a result, 14 lncRNAs, which consisted of CASC9, PCAT6, AP006621.2, GS1-124K5.4, MIR4435-2HG, AL354993.2, AC048344.4, AC010973.2, AL590483.1, AL137782.1, STAG3L5P-PVRIG2P-PILRB, LINC00513, SNHG16, and AP001554.1, were determined as independent prognostic indicators for OS. Furthermore, our 14-lncRNA

prognosis score was calculated using the following formula: prognosis score = (0.183 × expression level of CASC9) +(0.334 × expression level of PCAT6) + (0.348 × expression level of AP006621.2) +(0.409 × expression level of GS1-124K5.4) + (0.651 × expression level of MIR4435-2HG) +(0.838 × expression level of AL354993.2) + (0.881 × expression level of AC048344.4) +(1.145 × expression level of AC010973.2) + (− 0.882 × expression level of AL590483.1) +(− 0.875 × expression level of AL137782.1) + (− 0.688 × expression level of STAG3L5P-PVRIG2P-PILRB) +(− 0.665 × expression level of LINC00513) +(− 0.663 × expression level of SNHG16) +(− 0.573 × expression level of AP001554.1). Using this formula, we calculated the risk score of each patient. By using the median risk score as the threshold, the patients were divided into two groups: high-risk and low-risk groups. The heatmap of the 14 lncRNAs and the Kaplan–Meier analysis of the different survival durations between the two groups are shown in Fig. 5C,D. The results show that the high-risk patients (risk scores ≥ 1.009;n = 213) had shorter OS, indicating that they require more clinical attention and better clinical management, while the low-risk patients (risk scores < 1.009; n = 213) have better survival, in which a milder treatment plan is required to avoid over-treatment. The prognostic value of the risk score was also validated based on the data of the primary site rectosigmoid junction of the patient samples (n = 67) from the TCGA database. The AUC of the corresponding ROC curve of the autophagy related lncRNA-based prognostic signature prognostic index for the 5-year OS is 0.726 ($P$ < 0.001; Figs. 5F,G). As a consequence, autophagy related lncRNA-based prognostic signature also provides a robust prediction for the prognosis of COAD patients.

The univariate analysis showed that the two autophagy-related prognostic scores were significantly associated with OS in COAD patients (Fig. 6A). After adjusting for the clinicopathological features, such as age, gender, pathology_T_stage, pathology_N_stage, pathology_M _stage, and pathologic_stage. The two prognostic signatures remained an independent prognostic indicator for OS based on the multivariate analysis (mRNA-based signatures (HR = 1.381, 95% CI = 1.208–1.578; $P$ < 0.001; Fig. 6B); LncRNA-based signatures (HR = 1.204, 95% CI = 1.134–1.278; $P$ < 0.001; Fig. 6B). Subsequently, we also determined the clinical utility of the 6-mRNA prognostic signature and 14-lncRNA prognostic signature regarding the pathology_T_stage, pathology_N_stage, pathology_M_stage, and pathologic_stage. As shown in Fig. 6C, the mRNA-based risk score tends to increase in the higher pathology_T_stage (T1–2 *vs* T3–4; $P$ = 0.034), lymph node metastasis ( N0 *vs* N1–3; $P$ < 0.001), and pathologic_stage (stage I & II *vs* stage III & IV, $P$ < 0.001). Moreover, as shown in Fig. 6D, higher 14-lncRNA-based risk scores were observed among those with higher T stage (T1–2 *vs* T3–4, $P$ = 0.002), lymph node metastasis (N0 *vs* N1–3, $P$ = 0.002), and tumor stage (stage I-II *vs* stage III-IV, $P$ = 0.001). Specifically, the identified two prognostic signatures are of clinical relevance to pathologic_stage and TNM staging of COAD. These data suggest that two prognostic signatures have the potential in diagnosis of COAD. While, according to the ROC curve(Additional file 2: Fig. S2), the prognostic model of 14 autophagy-related lncRNA(1,3,5 year AUC = 0.768, 0.76, 0.804, respectively) was superior to the 6 autophagy-related mRNA(1,3,5 year AUC = 0.694, 0.697, 0.728, respectively) and other independent prognostic factors, which was then selected for further analyses.

## Establishment of the predictive nomogram for OS based on prognostic clinical factors and the 14-lncRNA signature

To further improve the predictive accuracy of our autophagy-related lncRNA signature by examining their performance in combination with other independent prognostic factors (e.g., age, tumor stage; Figs. 6A,B), we established an easy-to-use and clinically adaptable, risk nomogram for predicting the 3- and 5-year survival probability of COAD patients (Fig. 7A). Using the nomogram, “point” was scaled to estimate the points for each variable by drawing a vertical line. Then, the “total points” were scaled to estimate the corresponding 3- and 5-year OS rates of the COAD patients. The calibration curves of nomogram for predicting the 3- and 5-year COAD patients in the training set and validation set are shown in Figs. 7B, C. The C-indices (0.826 for the training set, 0.895 for the validation set) further strengthen the good discriminative ability of our models with regard to the age and tumor stage in predicting the OS of patients with COAD.

### 14-autophagy-related lncRNAs co-expressed network

An interaction network analysis was performed on the 14 risk score-related lncRNAs and the 9 associated ARGs (BID, CAPN10, CCR2, CD46, FAS, ITPR1, TNFSF10, VEGFA, PINK1). Then, the lncRNA–mRNA risk axes were integrated into two module maps (Additional file 3: Figs. S3A,B). Positive correlations were obtained for 17 lncRNA-mRNA pairs, and negative correlations were obtained for 3 lncRNA-mRNA pairs (Additional file 3: Fig. S3C). These results suggested that the coexpressed lncRNAs-mRNA networks may play important roles in regulating autophagy and in modulating COAD patients’ survival.

### Gene set enrichment analysis of 14 lncRNAs in the model

Moreover, to gain insights into the functional roles of the 14 lncRNAs in the molecular mechanisms of COAD, GSEA enrichment analyses for the lncRNAs between the predicted high-risk and low-risk groups were performed. The results revealed that the identified 14 lncRNAs were enriched in several autophagy-related, metastasis-related, and COAD-related pathways, such as MAPK signaling pathway, VEGF signaling pathway, Hedgehog signaling pathway, Notch signaling pathway, pathway in cancer; natural killer cell-mediated cytotoxicity, GAP junction, and JAK_STAT signaling pathway (Additional file 4 : Fig. S4). Therefore, we predicted that the 14-autophagy-related lncRNAs might be potential therapeutic targets for COAD.

## Discussion

The TNM staging system is routinely used as a staging procedure for patients with COAD [1]. Because the significant heterogeneity of the prognosis and treatment response is still observed in the same stage, it is necessary to develop reliable and effective prognostic biomarkers in order to improve the clinical decisions and outcome of COAD patients [23]. In recent years, high-throughput biological technologies have been widely used to predict the prognosis and response to specific therapies for a range of tumors. Moreover, growing evidence reveals the close relationship between the ncRNAs and autophagy in the malignant progression of certain cancers, including COAD [24–26]. Targeting autophagy was found to be an alternative and novel strategy in cancer immunology, and these autophagy-related RNAs may play a

key role in cancer biology. Moreover, autophagy-related genes, such as the mRNA, miRNA, or lncRNA signature, have prognostic potential for cancer[24, 25, 27–29].

Based on the findings of this study, we performed a bioinformatics analysis of the data from the COAD and control samples obtained from the TCGA and GEO databases; two panels (14-lncRNA signature and 6-mRNA signature) relating to the prognosis for patients with COAD were identified in our study, respectively. The results of the multivariate Cox analysis indicated that these two risk scores were independent prognostic factors, which were capable to categorize the COAD patients into subgroups with significantly different risk scores. The results of the validation set revealed good reproducibility and reliability with regard to the signatures. The application of these signatures would help clinicians to select patients with a high-risk of death, thereby facilitating the development of individualized therapies for COAD patients. However, due to the heterogeneity of the disease, it is insufficient to use a single signature. Therefore, a nomogram that incorporated the patient's age, tumor stage, and lncRNA risk scores was developed in order to predict the OS of the patients with COAD. The performance of the nomogram was verified in all data sets, which guaranteed the repeatability of our model.

lncRNAs dominate the upstream portion of the RNA network and function as primary effectors of mRNAs, and several lncRNAs are identified as potential prognostic factors in COAD [20, 30]. Among these 14 lncRNAs in the model, PCAT6, CASC9, MIR4435-2HG, SNHG16, and LINC00513 have been reported to play a role in the pathogenesis and prognosis of various cancer types (31–43). For example, CASC9 has been reported to be upregulated and conferred an oncogenic function in human hepatocellular carcinoma, oral carcinoma, gastric cancer, colorectal cancer, and esophageal cancer [28, 31–33]. In colorectal cancer, CASC9 was found to be negatively correlated with the miR-193a-5p expression, and promoted tumorigenesis via the miR-193a-5p/ERBB2 axis[34]. Luo et al. [35] reported that the higher expression of the oncogenic CASC9 was associated with poor patient outcomes when interacting with CPSF3 to regulate TGF-β signaling. Consistently, it has been reported that the higher expression of PCAT6 was correlated with poorer clinical outcomes and further found that PCAT6 was able to inhibit cell apoptosis by regulating the anti-apoptotic protein ARC expression via the EZH2 in COAD [36]. Dong et al. [37] reported that the upregulation of oncogenic MIR4435-2HG promoted colorectal cancer proliferation and metastasis through partially repressing the miR-206/YAP1 axis. Moreover, the upregulated MIR4435-2HG was found to promote hepatocellular cancer cell proliferation through upregulating the miRNA-487a [38] as well as lung cancer cell progression by activating β-catenin signaling [39]. lncRNA SNHG16 was showed to be regulated by the wnt pathway in colorectal cancer [40]. Additionally, SNHG16 has also been reported to associate with many other cancers, such as bladder cancer and pediatric neuroblastoma [41, 42]. As of now, the remaining lncRNAs (AP006621.2, AL354993.2, AC048344.4, GS1-124K5.4, AC010973, AL590483.1, AL137782.1, STAG3L5P-PVRIG2P-PILRB, and AP001554.1) are rarely reported in the previous studies. Moreover, no studies have reported any association between these 14 lncRNAs and autophagy. Interestingly, in our study, an interaction network analysis was performed on the 14 risk score-related lncRNAs and their associated ARGs. For instance, AP006621.2 and STAG3L5P-PVRIG2P-PILRB were found to be associated with the expression of VEGFA in COAD. VEGFA is known to be a pivotal oncogene in various cancers [43]. PCAT6 was found to be associated with expression of FAS, and FAS is

found to be an initiator of inflammation for cancers[44]. GSEA analysis results showed that 14 lncRNAs were enriched in several autophagy-related, metastasis-related, and COAD-related pathways. Our findings might contribute additional information to future studies on elucidating the molecular mechanisms behind the roles of these lncRNAs.

This study has several limitations. The neoadjuvant treatment information were not collected across the public cohorts. Another limitation is the potential reporting bias because all samples were from the retrospective collection. Further prospective studies are required to validate the results.

In conclusion, the study provided a novel 6-ARGs and a 14-lncRNA prognostic signatures which were developed and validated in order to categorize the COAD patients into subgroups with different prognostic risks, respectively. Furthermore, we established a risk nomogram which incorporates our lncRNA risk score and key clinicopathological features, offering an easy-to-deploy tool for identifying high-risk COAD patients and predicting the prognosis of patients suffering from this lethal malignancy.

## Declarations

# Authors' contributions

QX and YG-G conceived the study. GH-M collected the database.YG-G, FJ-T, GH-M and JW-Z analyzed and interpreted these data. QX and YH wrote the manuscripts. The final draft was read and approved by all authors.

# Funding

This work was supported by the Natural Science Foundation of China (81701185), High School Key Research Project of Henan Province (18A320003), the key scientific and technological project of Henan Province (202102310406).

# Data availability statement

The datasets analyzed for this study can be found in the Human Autophagy Database (http://www.autophagy.lu/), cBioPortal (http://www.cbioportal.org), The Cancer Genome Atlas (https://portal.gdc.cancer.gov/), and Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/).

# Ethics approval and consent to participate

No ethical approval nor informed consent was required in this study due to the public-availability of the data used.

# Consent for publication

Not applicable.

# Competing interests

The authors declare that they have no competing interests.

# Acknowledgements

Not applicable.

## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018, 68(6):394-424.
2. Yang SY, Sales KM, Fuller B, Seifalian AM, Winslet MC: Apoptosis and colorectal cancer: implications for therapy. Trends Mol Med 2009, 15(5):225-233.
3. Zheng X, He K, Zhang L, Yu J: Crizotinib induces PUMA-dependent apoptosis in colon cancer cells. Mol Cancer Ther 2013, 12(5):777-786.
4. Brenner H, Kloor M, Pox CP: Colorectal cancer. Lancet 2014, 383(9927):1490-1502.
5. Mizushima N, Levine B, Cuervo AM, Klionsky DJ: Autophagy fights disease through cellular self-digestion. Nature 2008, 451(7182):1069-1075.
6. Levine B: Unraveling the role of autophagy in cancer. Autophagy 2006, 2(2):65-66.
7. Amaravadi RK, Lippincott-Schwartz J, Yin XM, Weiss WA, Takebe N, Timmer W, DiPaola RS, Lotze MT, White E: Principles and current strategies for targeting autophagy for cancer treatment. Clin Cancer Res 2011, 17(4):654-666.
8. Janku F, McConkey DJ, Hong DS, Kurzrock R: Autophagy as a target for anticancer therapy. Nature Reviews Clinical Oncology 2011, 8(9):528-539.
9. Sakitani K, Hirata Y, Hikiba Y, Hayakawa Y, Ihara S, Suzuki H, Suzuki N, Serizawa T, Kinoshita H, Sakamoto K *et al*: Inhibition of autophagy exerts anti-colon cancer effects via apoptosis induced by p53 activation and ER stress. BMC Cancer 2015, 15:795.
10. Lampada A, O'Prey J, Szabadkai G, Ryan KM, Hochhauser D, Salomoni P: mTORC1-independent autophagy regulates receptor tyrosine kinase phosphorylation in colorectal cancer cells via an mTORC2-mediated mechanism. Cell Death Differ 2017, 24(6):1045-1062.
11. Izumi M, Nakamura S: Chloroplast Protein Turnover: The Influence of Extraplastidic Processes, Including Autophagy. Int J Mol Sci 2018, 19(3).

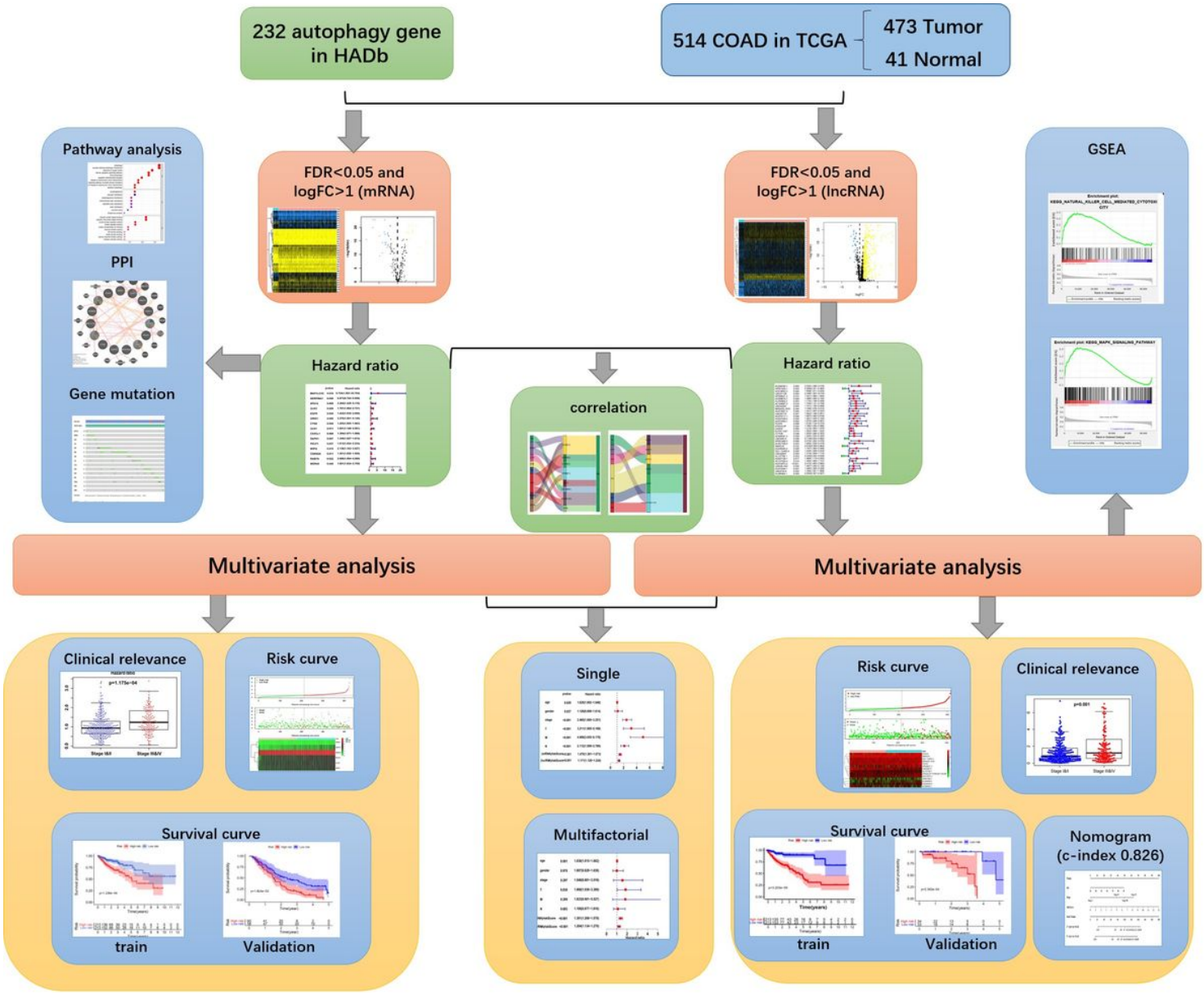
12. Zeng X, Kinsella TJ: A novel role for DNA mismatch repair and the autophagic processing of chemotherapy drugs in human tumor cells. Autophagy 2007, 3(4):368-370.
13. Ren SX, Shen J, Cheng AS, Lu L, Chan RL, Li ZJ, Wang XJ, Wong CC, Zhang L, Ng SS *et al*: FK-16 derived from the anticancer peptide LL-37 induces caspase-independent apoptosis and autophagic cell death in colon cancer cells. PLoS One 2013, 8(5):e63641.
14. Fu W, Li X, Lu X, Zhang L, Li R, Zhang N, Liu S, Yang X, Wang Y, Zhao Y *et al*: A novel acridine derivative, LS-1-10 inhibits autophagic degradation and triggers apoptosis in colon cancer cells. Cell Death Dis 2017, 8(10):e3086.
15. Luan Y, Li Y, Zhu L, Zheng S, Mao D, Chen Z, Cao Y: Codonopis bulleynana Forest ex Diels inhibits autophagy and induces apoptosis of colon cancer cells by activating the NF-kappaB signaling pathway. Int J Mol Med 2018, 41(3):1305-1314.
16. Maira SM, Stauffer F, Brueggen J, Furet P, Schnell C, Fritsch C, Brachmann S, Chene P, De Pover A, Schoemaker K *et al*: Identification and characterization of NVP-BEZ235, a new orally available dual phosphatidylinositol 3-kinase/mammalian target of rapamycin inhibitor with potent in vivo antitumor activity. Mol Cancer Ther 2008, 7(7):1851-1863.
17. Liu J, Long S, Wang H, Liu N, Zhang C, Zhang L, Zhang Y: Blocking AMPK/ULK1-dependent autophagy promoted apoptosis and suppressed colon cancer growth. Cancer Cell Int 2019, 19:336.
18. Huang JZ, Chen M, Chen, Gao XC, Zhu S, Huang H, Hu M, Zhu H, Yan GR: A Peptide Encoded by a Putative lncRNA HOXB-AS3 Suppresses Colon Cancer Growth. Mol Cell 2017, 68(1):171-184 e176.
19. Mercer TR, Dinger ME, Mattick JS: Long non-coding RNAs: insights into functions. Nature Reviews Genetics 2009, 10(3):155-159.
20. Yue B, Qiu S, Zhao S, Liu C, Zhang D, Yu F, Peng Z, Yan D: LncRNA-ATB mediated E-cadherin repression promotes the progression of colon cancer and predicts poor prognosis. J Gastroenterol Hepatol 2016, 31(3):595-603.
21. Wu Q, Meng WY, Jie Y, Zhao H: LncRNA MALAT1 induces colon cancer development by regulating miR-129-5p/HMGB1 axis. J Cell Physiol 2018, 233(9):6750-6757.
22. Li Y, Li C, Li D, Yang L, Jin J, Zhang B: lncRNA KCNQ1OT1 enhances the chemoresistance of oxaliplatin in colon cancer by targeting the miR-34a/ATG4B pathway. Onco Targets Ther 2019, 12:2649-2660.
23. Gasch C, Bauernhofer T, Pichler M, Langer-Freitag S, Reeh M, Seifert AM, Mauermann O, Izbicki JR, Pantel K, Riethdorf S: Heterogeneity of epidermal growth factor receptor status and mutations of KRAS/PIK3CA in circulating tumor cells of patients with colorectal cancer. Clin Chem 2013, 59(1):252-260.
24. Giannopoulou E, Antonacopoulou A, Matsouka P, Kalofonos HP: Autophagy: novel action of panitumumab in colon cancer. Anticancer Res 2009, 29(12):5077-5082.
25. Hu Q, Wang YB, Zeng P, Yan GQ, Xin L, Hu XY: Expression of long non-coding RNA (lncRNA) H19 in immunodeficient mice induced with human colon cancer cells. Eur Rev Med Pharmacol Sci 2016, 20(23):4880-4884.

26. Kourkoumpetis T, Chen L, Ittmann M, Graham DY, El-Serag HB, Jiao L: Altered Expression of Autophagy-Related Genes in Human Colon Cancer. Gastroenterology 2017, 152(5):S1029.
27. Eissa S, Matboli M, Awad N, Kotb Y: Identification and validation of a novel autophagy gene expression signature for human bladder cancer patients. Tumour Biol 2017, 39(4):1010428317698360.
28. Zhang H, Lu X, Wang N, Wang J, Wang L: Autophagy-related gene expression is an independent prognostic indicator of glioma. Oncotarget 2017, 8(37):60987-61000.
29. Wang SS, Chen G, Li SH, Pang JS, Cai KT, Yan HB, Huang ZG, He RQ: Identification and validation of an individualized autophagy-clinical prognostic index in bladder cancer patients. Onco Targets Ther 2019, 12:3695-3712.
30. Bo H, Fan L, Li J, Liu Z, Zhang S, Shi L, Guo C, Li X, Liao Q, Zhang W *et al*: High Expression of lncRNA AFAP1-AS1 Promotes the Progression of Colon Cancer and Predicts Poor Prognosis. J Cancer 2018, 9(24):4677-4683.
31. Pan Z, Mao W, Bao Y, Zhang M, Su X, Xu X: The long noncoding RNA CASC9 regulates migration and invasion in esophageal cancer. Cancer Med 2016, 5(9):2442-2447.
32. Noh JH, Gorospe M: AKTions by Cytoplasmic lncRNA CASC9 Promote Hepatocellular Carcinoma Survival. Hepatology 2018, 68(5):1675-1677.
33. Fang J, Chen W, Meng XL: LncRNA CASC9 Suppressed the Apoptosis of Gastric Cancer Cells through Regulating BMI1. Pathol Oncol Res 2020, 26(1):475-482.
34. Ding Y, Li X, Zhang Y, Zhang J: Long Non-Coding RNA Cancer Susceptibility 9 (CASC9) Up-Regulates the Expression of ERBB2 by Inhibiting miR-193a-5p in Colorectal Cancer. Cancer Management and Research 2020, Volume 12:1281-1292.
35. Luo K, Geng J, Zhang Q, Xu Y, Wu J: LncRNA CASC9 interacts with CPSF3 to regulate TGF-β signaling in colorectal cancer. Journal of Experimental & Clinical Cancer Research 2019, 38(1).
36. Huang W, Su G, Huang X, Zou A, Wu J, Yang Y, Zhu Y, Liang S, Li D, Ma F *et al*: Long noncoding RNA PCAT6 inhibits colon cancer cell apoptosis by regulating anti-apoptotic protein ARC expression via EZH2. Cell Cycle 2019, 18(1):69-83.
37. Dong X, Yang Z, Yang H, Li D, Qiu X: Long Non-coding RNA MIR4435-2HG Promotes Colorectal Cancer Proliferation and Metastasis Through miR-206/YAP1 Axis. Frontiers in Oncology 2020, 10:160.
38. Kong Q, Liang C, Jin Y, Pan Y, Tong D, Kong Q, Zhou J: The lncRNA MIR4435-2HG is upregulated in hepatocellular carcinoma and promotes cancer cell proliferation by upregulating miRNA-487a. Cell Mol Biol Lett 2019, 24:26.
39. Qian H, Chen L, Huang J, Wang X, Ma S, Cui F, Luo L, Ling L, Luo K, Zheng G: The lncRNA MIR4435-2HG promotes lung cancer progression by activating beta-catenin signalling. J Mol Med (Berl) 2018, 96(8):753-764.
40. He X, Ma J, Zhang M, Cui J, Yang H: Long Non-Coding RNA SNHG16 Activates USP22 Expression to Promote Colorectal Cancer Progression by Sponging miR-132-3p. Onco Targets Ther 2020, 13:4283-

4294.

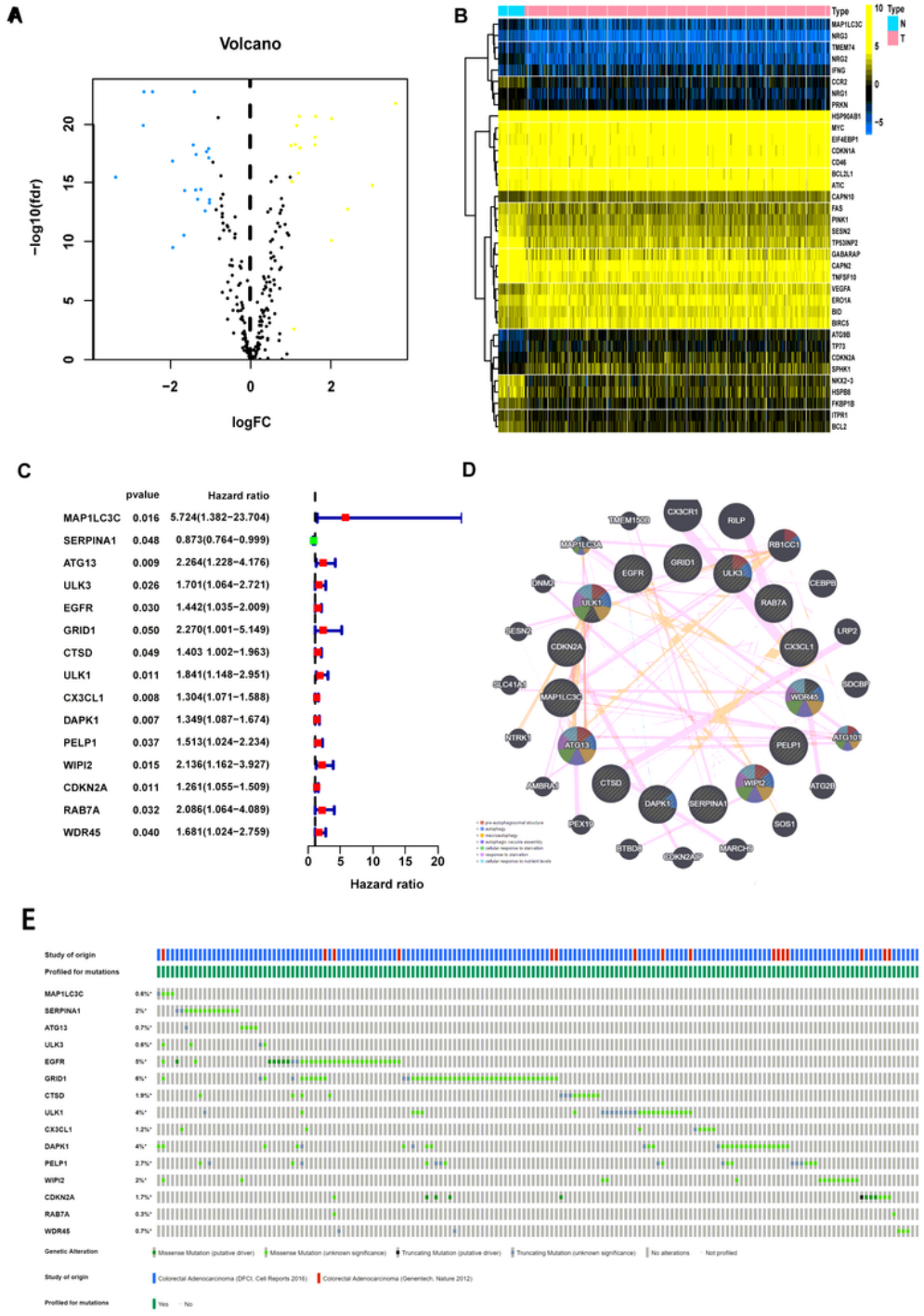
41. Deng D, Yang S, Wang X: Long non-coding RNA SNHG16 regulates cell behaviors through miR-542-3p/HNF4α axis via RAS/RAF/MEK/ERK signaling pathway in pediatric neuroblastoma cells. Bioscience Reports 2020, 40(5).
42. Feng F, Chen A, Huang J, Xia Q, Chen Y, Jin X: Long noncoding RNA SNHG16 contributes to the development of bladder cancer via regulating miR-98/STAT3/Wnt/beta-catenin pathway axis. J Cell Biochem 2018, 119(11):9408-9418.
43. Xue Z, Cui C, Liao Z, Xia S, Zhang P, Qin J, Guo Q, Chen S, Fu Q, Yin Z: Identification of LncRNA Linc00513 containing lupus-associated genetic variants as a novel regulator of interferon signaling pathway. Frontiers in immunology 2018, 9:2967.
44. Cullen SP, Martin SJ: Fas and TRAIL 'death receptors' as initiators of inflammation: Implications for cancer. In: Seminars in cell & developmental biology: 2015: Elsevier; 2015: 26-34.

# Figures




## Figure 1

The flowchart of our research process.




## Figure 2

Prognostic ARGs screening. (A) The volcano plot of the differentially expressed ARGs. The yellow dots indicate a high expression and the blue dots a low expression. (B) The heat map shows 15 prognostic ARGs between COAD and paired non-tumor samples, with yellow dots indicating significantly upregulated mRNAs, blue dots indicating downregulated mRNAs, and black dots indicating no differences in the mRNAs. “N” indicates non-tumor tissues; “T” indicates tumor tissues. (C)Risk ratio forest plot showed the

prognostic value of the 15 genes. (D)The protein–protein interactions among the ARGs that are associated with the prognosis of patients with COAD.The different colors of the network nodes indicate the biological functions of the set of enrichment genes. (E) The mutation frequency of the prognosis-related ARGs.

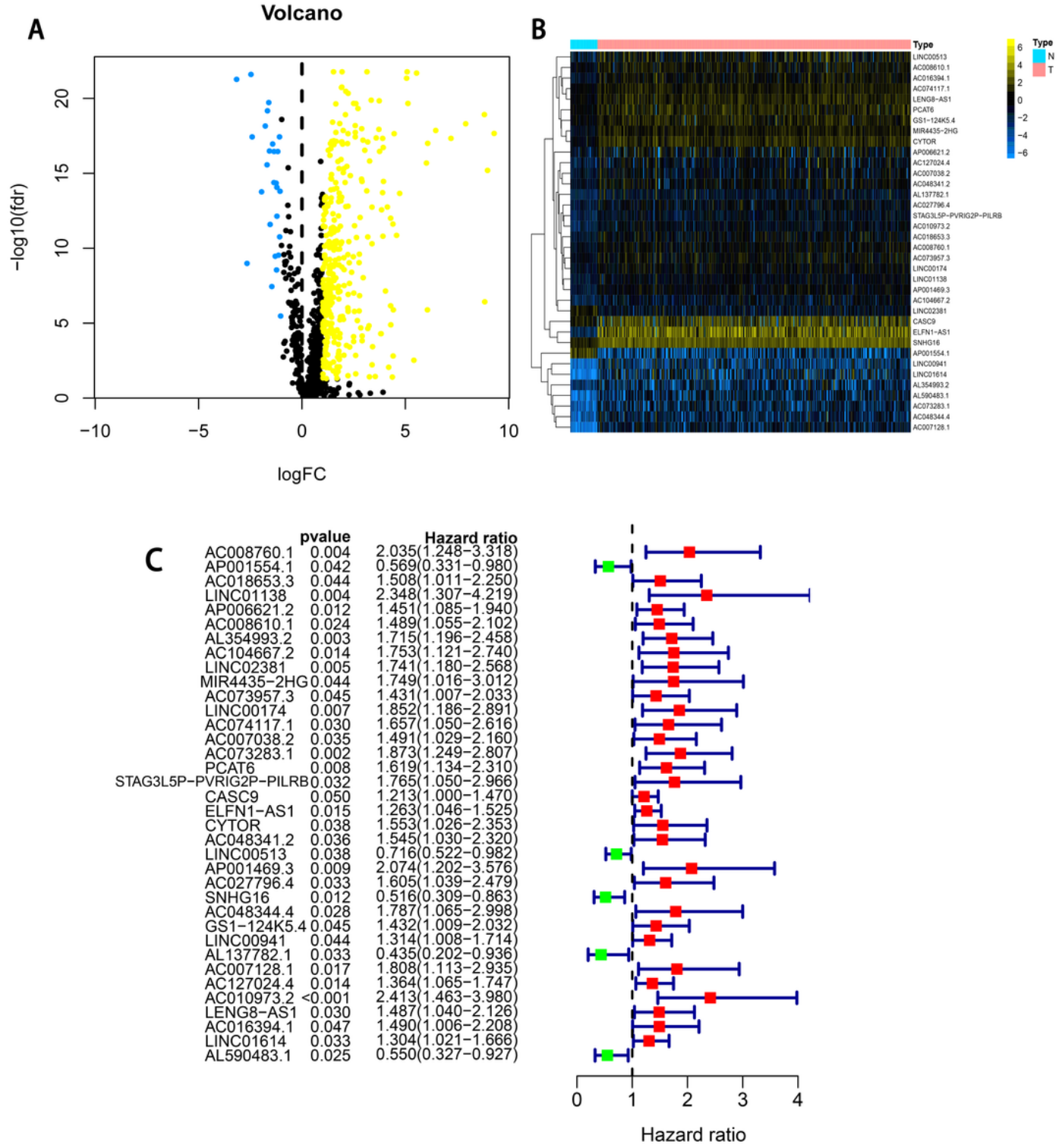



**Figure 3**

Prognostic lncRNAs screening. (A) The volcano plot of the differentially expressed autophagy-related lncRNAs. The yellow dots indicates the high expression and the blue indicates the low expression. (B) The heat map show 36 autophagy-related lncRNAs that associated with prognostic of patients with COAD. The yellow dots indicates significantly upregulated lncRNAs, blue dots representing downregulated lncRNAs, and black dots represented no differences lncRNAs.(C) Risk ratio forest plot showed the prognostic value of the 36 autophagy-related lncRNAs.

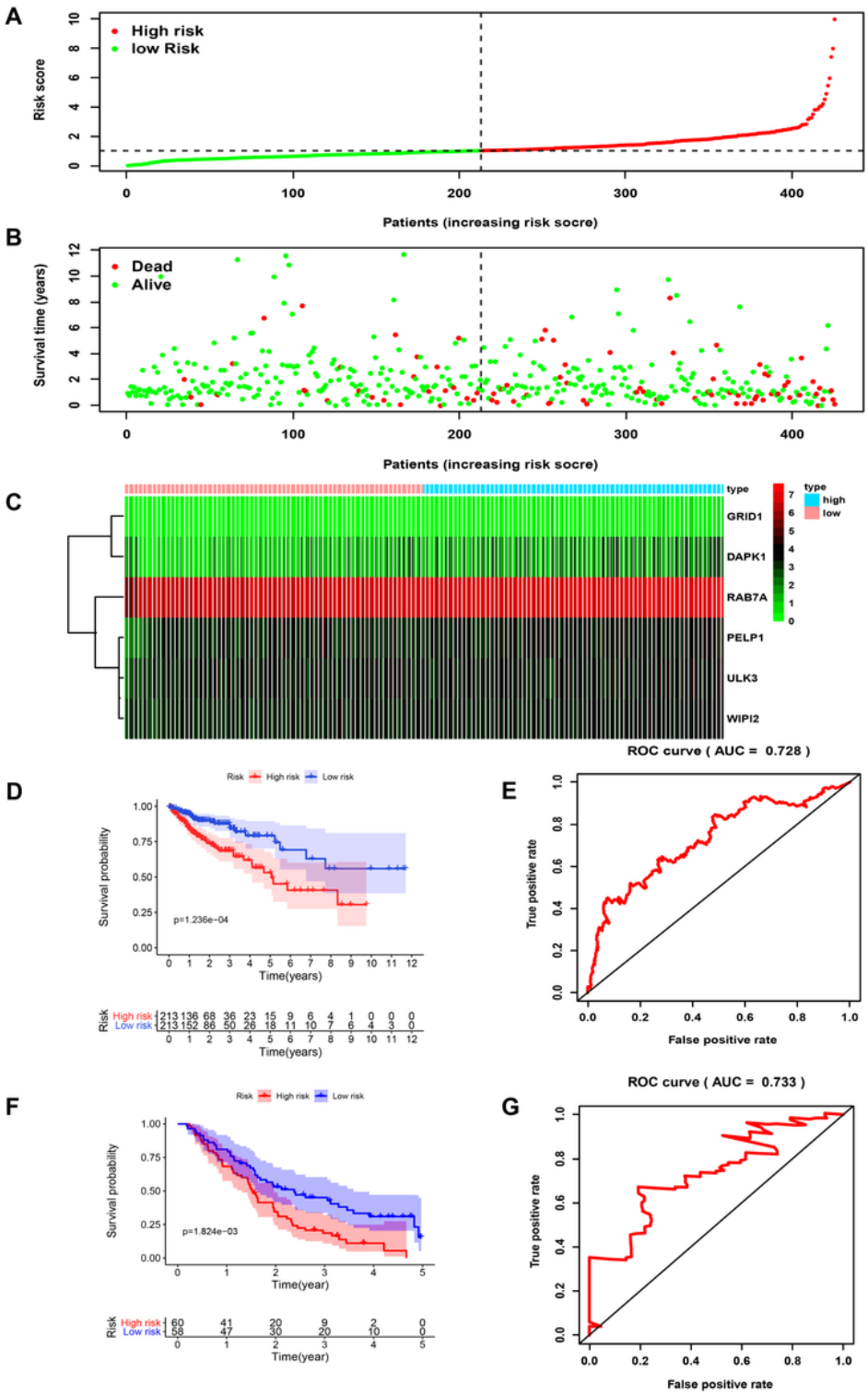
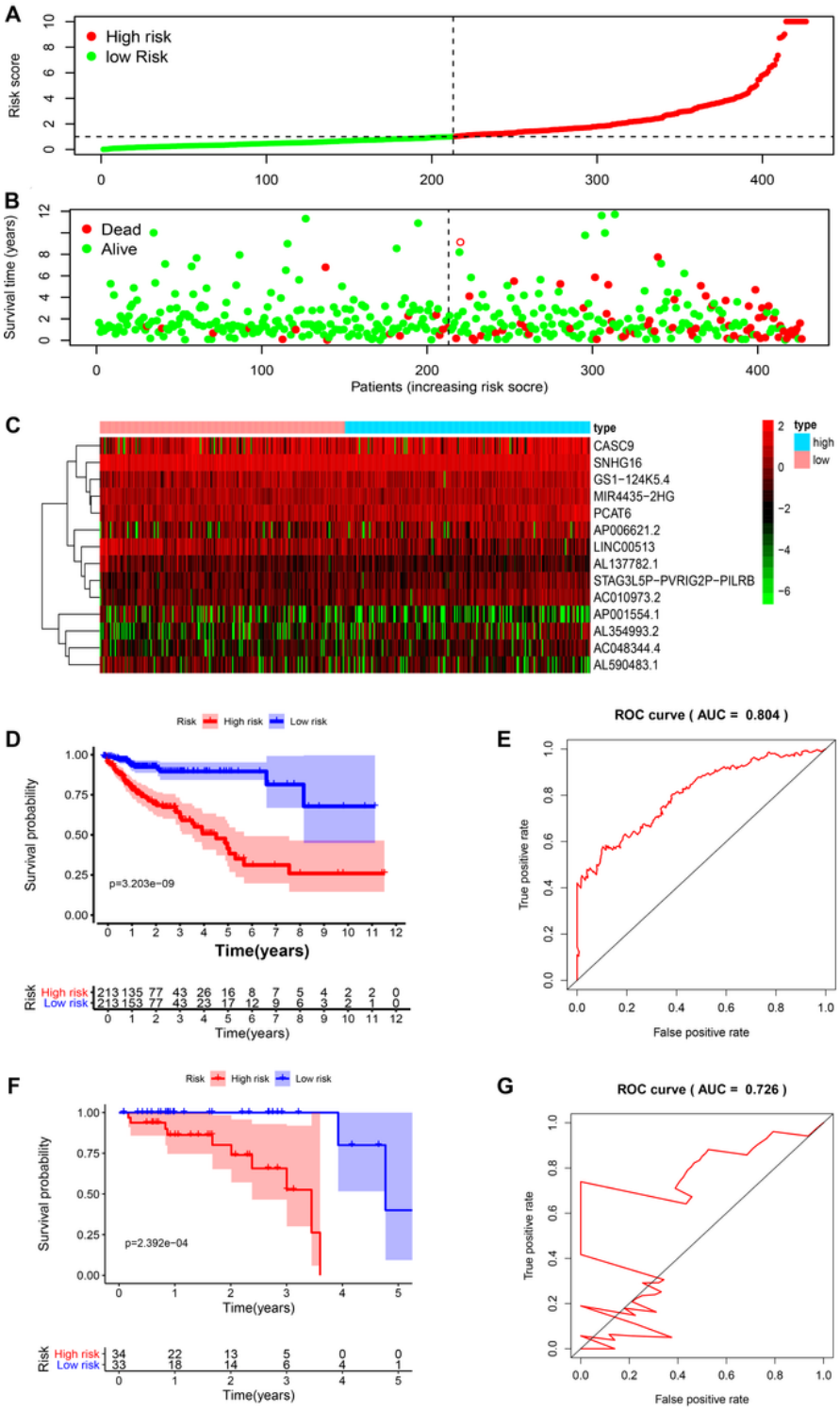


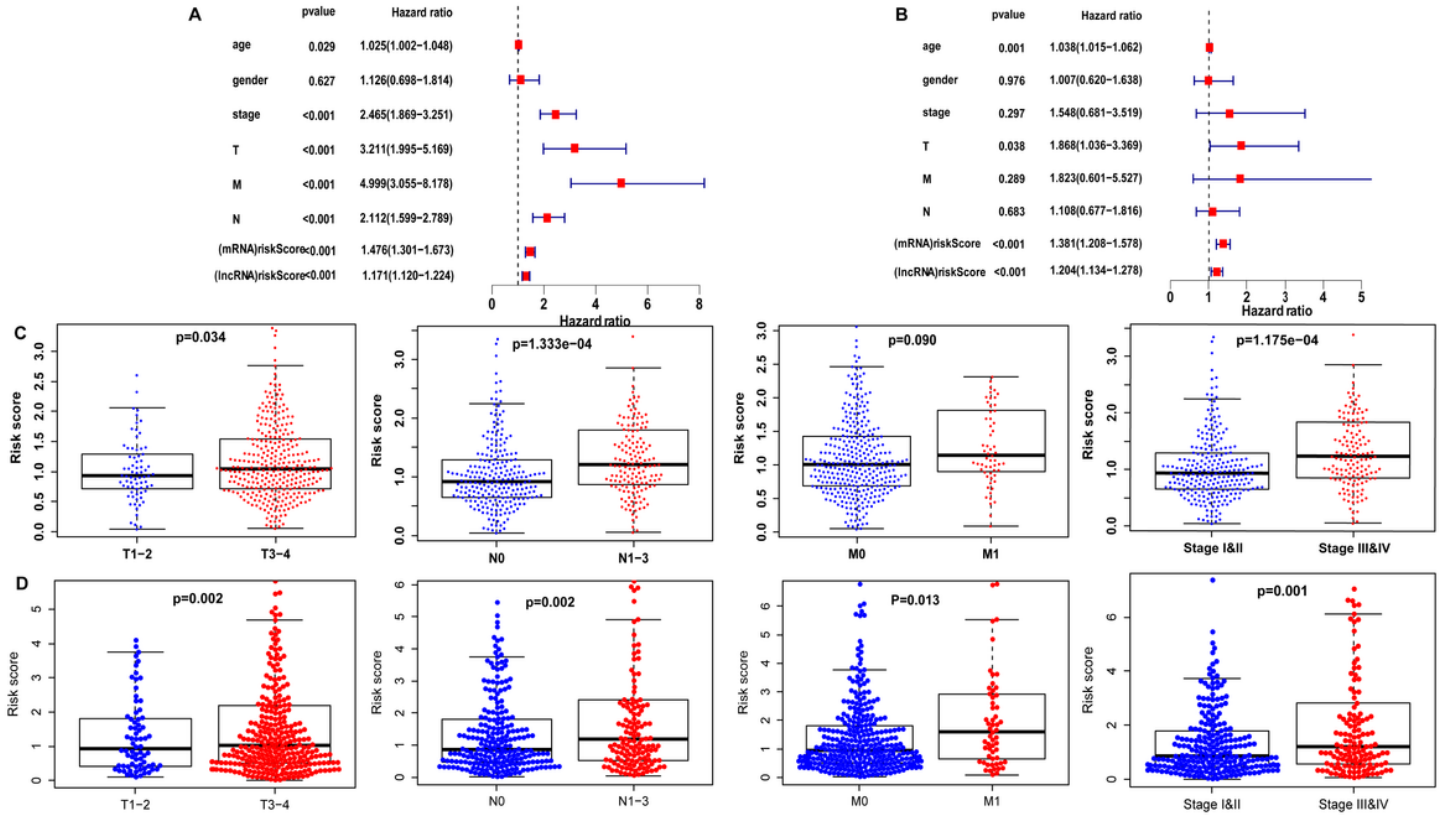

Figure 4

Development of a risk score system based on ARGs. (A) Distribution of risk score. The risk score increased from blue to red. (B) The survival status and duration of COAD cases. Blue and red scatter indicate alive and dead, respectively. (C) The heatmap shows the expression distributions of the 6 ARGs in the low-risk and high-risk groups, with the color changing from blue to red, indicating an increasing trend from low levels to high levels. The Kaplan–Meier survival curves (high vs. low-risk) in the training set (D) and the validation set (F). ROC curves of the two risk score system to evaluate their capability in predicting the OS of the COAD patients in the training set (E) and the validation set (G).



## Figure 5

Development of a risk score system based on autophagy-related lncRNAs. (A) Distribution of risk score. The risk score increased from green to red. (B) The survival status and duration of each patients. Blue and red scatter represent alive and dead, respectively. (C) The heatmap shows the expression distributions of the 8 lncRNAs in the low-risk and high-risk groups, with the color changing from blue to red, indicating an increasing trend from low levels to high levels. The Kaplan–Meier survival curves (high vs. low-risk) in the training set (D) and the validation set (F). Using the ROC curves of the risk score system to evaluate their capability in predicting the OS of the COAD patients in the training set (E) and the validation set (G).




## Figure 6

Clinical utility of prognostic signature for COAD. (A,B) Forest plot of univariate Cox regression and multivariate Cox regression analyses of risk factors associated with OS. (C,D) The clinical utility of the 6-mRNA prognostic signatures and 14-lncRNA prognostic signatures regarding the pathology_T_stage, pathology_N_stage, pathology_M_stage, and pathologic_stage.

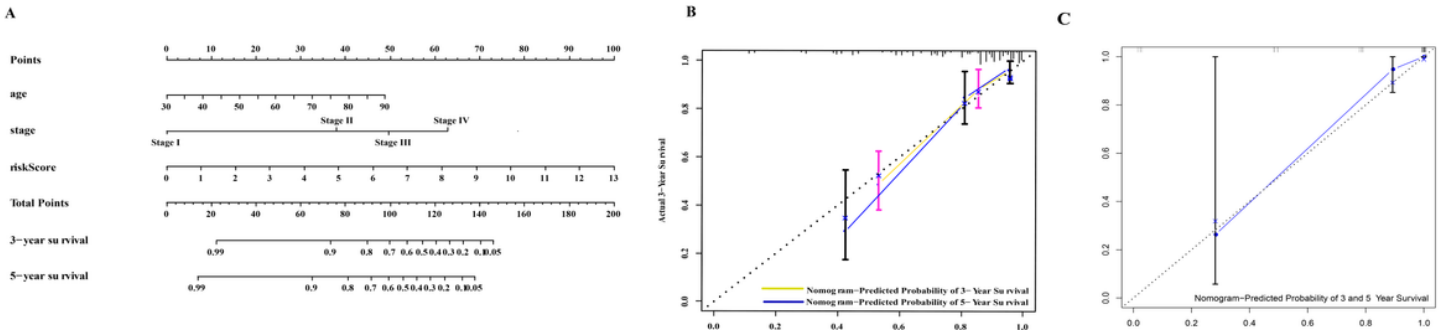



**Figure 7**

Nomograms predicting survival and calibration curves in COAD patients. (A)The nomogram used to predict the 3- and 5-year OS based on the 8-lncRNA risk score, age, and tumor stage. (B,C) Calibration curves of nomogram for predicting the 3- and 5-year OS of the COAD patients in the training set (B) and in the validation set (C). The dotted line represents a perfect fit, the solid lines represent the predictive performance, and the error bars reflect 95% CIs of the nomograms.

# Supplementary Files

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

- SupplenmentaryFigureS2.tif
- SupplementaryTableS2.docx
- SupplementaryTableS1.docx
- SupplementaryFigureS4.tif
- SupplementaryFigureS3.tif
- SupplementaryFigureS1.tif