

# Identification of Autophagy-Related Risk Signatures for the Prognosis, Diagnosis, and Targeted Therapy in Cervical Cancer

**Meng Dan**

Southeast University

**Hua Jin**

Nantong Tumor Hospital

**Xing Zhang**

Southeast University

**Wenjing Yan**

Southeast University

**Qianqian Xia**

Southeast University

**Siyuan Shen**

Southeast University

**Shuqian Xie**

Southeast University

**Mengjing Cui**

Southeast University

**Bong Ding**

Southeast University Zhongda Hospital

**Yun Gu**

Nanjing Maternity and Child Health Care Hospital

**Shizhi Wang** (✉ [101011631@seu.edu.cn](mailto:101011631@seu.edu.cn))

Southeast University <https://orcid.org/0000-0002-6083-6753>

---

## Research Article

**Keywords:** cervical cancer, autophagy, mRNA, biomarkers, qPCR

**Posted Date:** April 9th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-398453/v1>

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

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Cancer Cell International on July 8th, 2021.

See the published version at <https://doi.org/10.1186/s12935-021-02073-w>.

# Abstract

## Background

To rummage autophagy-related prognostic, diagnostic, and therapeutic biomarkers in cervical cancer (CC).

## Methods

The RNA-sequence and clinical information were from the TCGA and GTEx databases. We operated Cox regression to determine signatures related to overall survival (OS) and recurrence-free survival (RFS) respectively. The diagnostic and therapeutic effectiveness of prognostic biomarkers were further explored.

## Results

We identified nine (*VAMP7*, *MTMR14*, *ATG4D*, *KLHL24*, *TP73*, *NAMPT*, *CD46*, *HGS*, *ATG4C*) and three risk signatures (*SERPINA1*, *HSPB8*, *SUPT20H*) with prognostic values for OS and RFS respectively. Six risk signatures (*ATG4C*, *ATG4D*, *CD46*, *TP73*, *SERPINA1*, *HSPB8*) were selected for qPCR. We screened five prognostic signatures (*ATG4C*, *CD46*, *HSPB8*, *MTMR14*, *NAMPT*) with diagnostic function through the GEO database. Correlation between our models and treatment targets certificated the prognostic score provided a reference for precision medicine.

## Conclusions

We constructed OS and RFS prognostic models in CC. Autophagy-related risk signatures might serve as diagnostic and therapeutic biomarkers.

## Background

Cervical cancer (CC) is one of the most common gynecological malignancies in women[1]. It is estimated that 600,000 women are diagnosed as CC, and more than 340,000 people die from CC each year. The continuous infection of high risk human papillomavirus (HR-HPV) is a necessary condition for CC, but its specific mechanism and co-factors have not yet been clarified[2]. Studies have shown that more than 70% of CC patients are diagnosed as advanced, and 15–61% of patients have lymph node or distant metastasis and recurrence, then the survival rate of these patients is significantly reduced[3]. Although the use of HPV vaccine and the progress of surgery, radiotherapy and chemotherapy have reduced the incidence and mortality of CC, late metastasis, recurrence and drug resistance are still huge challenges for its clinical treatment[4]. Limited by the economic level, the burden of disease is heavier in areas with scarce medical resources[5]. To reduce the burden and provide better treatment to patients, it is necessary

to make accurate judgments on the prognosis of patients in order to achieve the purpose of formulating reasonable treatment plans and avoiding overtreatment or undertreatment. In recent years, autophagy has been discovered to play a momentous role in CC[6, 7].

Autophagy is an evolutionarily conserved important process in eukaryotes, which is responsible for the turnover of intracellular substances[8–10]. It has not been fully clarified whether autophagy plays a positive or negative role in physiological or pathological processes including cancer. Studies have shown that autophagy dysregulation plays a vital role in the occurrence and development of CC. Hu *et al.* found that autophagy was undermined in CC cells with gradually reduced positive expression rate of autophagy-related proteins Beclin1 and LC3 in the normal cervical group, CIN (cervical intraepithelial neoplasia) III group, and CC group[11]. Li *et al.* discovered that high-risk HPV infection could promote the development of CC by inhibiting autophagy of the host[12]. Zhang *et al.* reported that microtubule-associated protein 7 (MAP7) promoted the migration, invasion and progression of CC by regulating the autophagy pathway[13].

RNA sequence (RNA-seq) is a technique for observing the differences of cells at the gene level, thereby analyzing the biological behavior of different cells. At present, through the construction and verification of prediction models, a large number of studies have applied RNA-seq results to the diagnosis, treatment, and prognosis of CC. Wu *et al.* forecasted two lncRNA signatures as potential prognostic biomarker in CC based on the TCGA database[14]. Ding *et al.* screened immune-related genes' (IRGs) prognostic implications in cervical cancer and endometrial cancer[15]. Wang *et al.* identified *DAPK1* promoter hypermethylation as a biomarker for intraepithelial lesion and cervical cancer[16]. However, the literature on the prognostic, diagnostic and therapeutic value of autophagy-related genes in CC is scarce.

In the present study, we acquired the transcriptome profile and clinical information of CC patients from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx). According to the Human Autophagy-dedicated Database(HADb), we searched for autophagy-related genes (ARGs) in CC. After Cox regression analysis, ARGs related to overall survival (OS) and recurrence-free survival (RFS) were determined and their respective prognostic models were established. Kaplan-Meier curve, area under the curve (AUC) and nomogram all indicated that these models had an excellent predictive performance. Endometrial cancer (UCEC) and head and neck cancer (HNSCC) were used as external verification data to verify the effectiveness of the models. The functions of ARGs included in the prognostic models were further analyzed by Gene Set Enrichment Analysis (GSEA). Finally, the diagnostic value of risk signatures was detected based on the Gene Expression Omnibus (GEO) database. We found that the risk score was significantly correlated with clinical treatment targets. In this paper, we screened out valuable biomarkers for CC by constructing prognostic models containing ARGs and studied their diagnostic value and treatment orientation for clinical.

## Methods

### Data gathered and pre-neatening of the training set

The RNA-seq transcriptome data (HTSeq-Counts) composed of 306 CC with clinical information and 3 normal samples was downloaded from the TCGA data portal (<https://tcga-data.nci.nih.gov/tcga/>). Five normal cervix samples from the GTEx portal (<http://gtexportal.org>) were added to expand the number of normal samples. For subsequent difference analysis, we normalized the transcriptome data from these two datasets and integrated data by R program (<http://www.r-project.org/webcite>) v4.0.2 using the limma package (<http://www.bioconductor.org/>). HADb (<http://www.autophagy.lu/>) is the first Human Autophagy- appropriated Database. We obtained a list of 232 human autophagy genes, as shown in Additional file 1. Subsequently 217 autophagy-related genes (ARGs) expressed in CC were shown in Additional file 2. We matched the autophagy-related gene expression profiles with the OS and RFS according to the patient's ID numbers respectively, and those that did not match were excluded.

### **Identification and functional annotation of the differently expressed autophagy-related genes (DEARGs)**

DEARGs were appraised by the R program using the edgeR package (Bioconductor-edgeR) with cut-off values as  $FDR < 0.05$  and  $|\log_2FC| > 1$ . To investigate the biomedical molecular mechanism of DEARGs, we carried out Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment under R environment using clusterProfiler package (Bioconductor-clusterProfiler). Then the STRING database (<https://string-db.org/>) was exploited to construct protein-protein interaction (PPI) network of DEARGs, and we imported the data into the software of Cytoscape 3.7.2 to visualize the interaction of PPI network.

### **Construction of the ARGs-based risk signatures**

Firstly, univariate Cox regression analysis was performed to screen ARGs related to OS and RFS, and  $P < 0.05$  was considered significant. Subsequently, to prevent overfitting of the models, we conducted Least absolute shrinkage and selection operator (Lasso) regression by using the R package "glmnet" (<https://CRAN.R-project.org/package=glmnet>). Finally, multivariate Cox regression analysis was carried out to evaluate the risk signatures of the predictive model and obtain their regression coefficients. We calculated the patient's risk score based on the expression level of the prognostic genes and the risk coefficient. The risk score was used to predict the prognosis of CC patients, and then patients were divided into high- and low-risk groups according to the median risk score ( $P < 0.05$ ). The Kaplan–Meier (K-M) survival curve was plotted to evaluate the statistical differences in OS and RFS between the two groups respectively. In addition, we generated receiver operating characteristic (ROC) curves to determine the accuracy of the prognostic models.

Univariate and multivariate Cox regression analyses were conducted to investigate whether autophagy-related risk biomarkers could be regarded as independent predictors of OS and RFS in CC patients from the TCGA data set. Clinical parameters such as age, stage, grade, T, race and BMI were considered for OS. The clinical information related to RFS was incomplete, we only included age, stage, and grade for independent prognostic analysis. The nomograms were generated to visualize the risk score and the survival probability of the patient calculated according to the prognostic genes in the OS and RFS model

using the “rms” and “survival” packages, thus realizing the combination of bioinformatics research and clinical practice. The calibration curve of the prediction model was used to evaluate the difference between the predicted probability and the actual probability.

### **Validation of the prognostic model in the testing set**

To verify the interpretability and prediction accuracy of the model, we introduced two other cancer types, they were endometrial cancer (UCEC) and head and neck cancer (HNSCC), both of which were directly or indirectly related to CC. HPV infection is an important risk factor for head and neck cancer[17, 18]. Endometrial cancer is also a common uterine malignant tumor. It has been reported that CC may cause endometrial cancer after radiotherapy[19].

The mRNA expression data and clinical OS data of patients with UCEC were from UCSC Xena (<http://xena.ucsc.edu/>). GSE117973, a global gene expression analysis of 77 primary head and neck cancer cases consisting of clinical progression-free survival (PFS) information, was downloaded from the comprehensive Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>). Similarly, we calculated the patient's risk score based on the risk signatures selected in the training set and divided samples into high- and low-risk groups, the K-M survival curves and ROC curves were plotted to verify the validity of the models in the testing set.

### **Detection of the diagnostic values of risk indicators GEO database**

The diagnostic values of the risk indicators for OS and RFS were assessed using ROC curve analysis from consolidated data sets from GSE63514 (24 normal samples and 28 CC), GSE75132 (21 normal samples and 20 CC), GSE6791 (8 normal samples and 20 CC). We combined the data and removed the batch effect for further analysis

### **Gene Set Enrichment Analysis of the 12 risk signatures**

GSEA 4.1.0 was used to demonstrate the role of each gene in CC. We chose C2.CP: KEGG.V7.0.symbols.gmt file as the reference gene set file, the values of  $P < 0.05$  and  $FDR < 0.25$  after performing 1,000 permutations were considered a significant difference.

### **Correlation between risk score and clinical therapeutic targets**

Here we studied the relationship between risk score and the common targeted therapeutic targets in CC by Pearson's correlation analysis. The therapy targets are listed as follow: AKT serine/threonine kinase 1 (*AKT1*, *AKT*), BCL2 apoptosis regulator (*BCL2*, *Bcl-2*, *PPP1R50*), mechanistic target of rapamycin kinase (*MTOR*, *FRAP*), tumor protein p53 (*TP53*, *P53*), vascular endothelial growth factor A (*VEGFA*, *VEGF*).

### **Expression of the 12 risk signatures at the protein level**

The Human Protein Atlas database(<https://www.proteinatlas.org/>) is a free data resource that contains more than 26,000 kinds of antibodies. The immunohistochemical results of 12 risk signatures in normal

and pathological tissues were found in this database.

## **RNA expression detection of selected risk signatures in cervical cancer and normal tissue samples**

To detect certain risk signatures(ATG4C, ATG4D, CD46, TP73, SERPINA1, HSPB8) expression in human samples by qPCR, a total of seven fresh CC tissues and paired adjacent non-tumor tissues were acquired from patients between September 2019 and October 2020 at Zhongda Hospital. RNA extraction and qRT-PCR can be described in the previous literature published by our research group[20]. The primer sequences involved in this study are shown in Additional1 file 3. All samples were stored at - 80 °C before total RNA extraction.

## **Statistical Analysis**

Statistical analyses were performed with the use of R 4.0.2 (www.r-project.org), SPSS 26, and GraphPad Prism 8.0.2. All statistical tests were two-sided, and  $P < 0.05$  were considered statistically significant.

# **Results**

## **Identification and functional annotation of DEARGs**

The flow diagram of the integral analysis is described in Figure 1. We obtained 217 autophagy genes expressed in 306 CC patients and 8 normal controls. Among these expressed autophagy genes, there were 53 differentially expressed genes (Figure 2a), including 21 down-regulated and 32 up-regulated ( $FDR < 0.05$ ,  $|\log_2FC| > 1$ ). The box plot revealed the expression pattern of differential genes in tumor and normal samples (Supplementary Figure 1a), all DEARGs were represented in Additional1 file 4. To rummage the signal pathways of DEARGs, we performed GO and KEGG enrichment analysis (Figure 2b,c). In the term of biological process (BP), the enriched pathways such as autophagy, macroautophagy, cellular response to external stimulus, intrinsic apoptotic signaling pathway and extrinsic apoptotic signaling pathway closely related to cell proliferation and migration. For the cellular component (CC) part, autophagosome, autophagosome membrane and vacuolar membrane were involved in autophagy. The results of molecular function (MF) of the genes showed these genes participated in many binding reactions. KEGG analysis showed 53 DEARGs were included in critical pathways associated with cancer development, such as apoptosis, autophagy – animal, platinum drug resistance and p53 signaling pathway. The detailed results of significant top 8 GO annotation and top 10 KEGG pathway enrichment were shown in Additional1 file 5. The PPI network showed the connection between the various genes (Figure 2d). The hub genes might be identified by the number of connections, CASP3 had the most connected nodes with other genes (Supplementary Figure 1b).

## **Construction and verification of autophagy prognostic model in the training set**

We first performed univariate Cox regression analysis to find autophagy gene signatures related to OS and RFS respectively. Among the 35 genes related to the OS, 13 genes were protective with  $HR < 1$ , and 22

genes were dangerous with HR > 1 (Figure 3a). Among the 9 genes related to the RFS, 3 genes were protective with HR < 1, and 6 genes were dangerous with HR > 1 (Figure 3b). Lasso regression was adopted to further screen variables to ensure the stability of our model, and finally 20 candidate genes related to OS and 8 candidate genes related to RFS were obtained (Figure 3c-f). Multivariate Cox regression analysis was utilized to determine the genes that constructed the predictive model and their regression coefficients. For OS and RFS models, 9 autophagy-related gene signatures (*VAMP7*, *MTMR14*, *ATG4D*, *KLHL24*, *TP73*, *NAMPT*, *CD46*, *HGS*, *ATG4C*) and 3 autophagy-related gene signatures (*SERPINA1*, *HSPB8*, *SUPT20H*) were gotten. The coefficient of each gene was shown in Table 1. Moreover, we could find that at the level of mRNA or protein, not all risk signatures related to survival were differential genes (Supplementary Figure 2a-l). The results from the K-M analysis of single-gene illustrated 8 genes were significantly related to survival, including five protective genes like *ATG4D*, *KLHL24*, *TP73*, *HSPB8*, *SERPINA1* ( $P = 0.025, 0.009, 0.030, 0.050, 0.014$  respectively) and three dangerous genes like *CD46*, *HGS*, *SUPT20H* ( $P = 9.609e-04, 0.007, 0.014$  respectively), which was consistent with the Cox regression analysis (Supplementary Figure 3a-l). Simultaneously, it could be speculated that these risk genes might have a synergistic effect, not just individual genes. According to the expression of model genes and regression coefficients to calculate the patient's risk score, the formula was as follows:

$$\text{risk score} = \sum_{i=1}^n \text{Coef}_i * x_i$$

Among them, coef represents the coefficient and  $x_i$  represents the relative expression value of each risk signature.

The median risk score was applied as a cut-off value to divide patients into high- and low-risk groups. The OS and RFS survival analysis respectively between the two groups were significantly different (Figure 4a, c,  $P = 2.636e-08$  for OS,  $P = 2.104e-03$  for RFS) lower risk score generally presages a better survival prognosis.

The AUC of ROC analyses at 1, 3, and 5 years was 0.783, 0.830, and 0.824 for OS and 0.682, 0.793, and 0.843 for RFS (Figure 4b, d.). Additionally, the risk score, the number of survivals and the expression of model genes between the high- and low-risk group had a difference. The expression of protective genes was higher in the low-risk group, and the expression of dangerous genes was higher in the high-risk group. The low-risk groups had fewer deaths than the high-risk groups (Figure 4e-j).

### **OS and RFS prognostic risk models were independent predictive indicators in the CC patients of the TCGA database**

We further analyzed whether the risk score could assume independent predictive indicators for CC patients. The clinical information was given in Table 2. Univariate Cox regression analysis showed that age ( $P = 0.030$ , HR = 1.022), stage ( $P < 0.001$ , HR = 1.811), T ( $P = 0.017$ , HR = 1.397), BMI ( $P = 0.036$ , HR = 0.954) and the risk score ( $P < 0.001$ , HR = 2.973) correlated with OS of CC patients (Figure 5a). Differently, only risk score ( $P < 0.001$ , HR = 2.867) was associated with RFS of CC patients (Figure 5c). Next, we

performed multivariate Cox regression analysis using the above clinical parameters and risk score. The results showed that the prognostic risk model could serve as an independent prognostic indicator to predict OS and RFS in patients respectively ( $P < 0.001$ , HR = 2.759 for OS,  $P < 0.001$ , HR = 2.785 for RFS) (Figure 5b,d). Furthermore, multivariate Cox analysis showed that stage significantly correlated with OS ( $P = 0.009$ , HR = 1.436). These results demonstrated that both prognostic models could be independently used to predict OS and RFS in CC patients. The nomogram calculated the survival prediction value of the individual at 1, 3, and 5 years based on the total score and the probability of the outcome event (Figure 5e, f). The calibration curve showed that the predicted risk was consistent with the actual risk (Supplementary Figure 4a-f).

### **Verification of autophagy-related predictive signatures in the testing group**

The effect of the prediction model is likely to be different due to changes in scenarios and populations. To verify the validity of the prognostic model and improve its generalizability, we conducted external verifications on the OS and RFS prognosis model. Since other appropriate gene expression data and clinical information about CC patients were not available, we chose UCEC and HNSCC as external verification data. We downloaded the transcriptome data of UCEC patients and clinical data including OS from the UCSC Xena database, then calculated the risk score based on the expression of nine OS model genes and regression coefficients in the UCEC patients from our testing group. Furthermore, GSE117973 containing mRNA expression information and PFS clinical data was used for validating the prognostic model of RFS and risk score was calculated. In the OS and RFS validation sets, patients were divided into high- and low-risk groups based on the calculated median risk score respectively. K-M survival curves showed that in UCEC patients and HNSCC patients, the higher risk score was associated with adverse outcomes (Supplementary Figure 4g, i). The AUC also proved that the risk signature had good accuracy with 0.571, 0.635, and 0.669 at 3, 5 and 7 years respectively for the OS of UCEC patients. For the PFS of HNSCC patients, the AUC was 0.443, 0.571, 0.635 at 1, 3 and 5 years respectively (Supplementary Figure 4h, j).

### **Diagnostic value of risk signatures in GEO consolidated data sets in CC**

We identified the diagnostic application of risk signatures in the combined data using ROC curve analysis. The AUC for *ATG4C*, *ATG4D*, *CD46*, *HSPB8*, *MTMR14* and *SUPT20H* were 0.704, 0.640, 0.697, 0.627, 0.743 and 0.705 respectively (95% CI = 0.612-0.796, 0.542-0.737, 0.603-0.791, 0.528-0.726, 0.654-0.832, 0.614-0.796), (Figure 6a-f). This demonstrated that risk signatures were potential diagnostic markers. Diagnostic values of other risk signatures without statistical significance were shown in Supplementary Figure 5a-f.

### **Single-gene Gene Set Enrichment for the risk signatures**

Single-gene GSEA of the risk signatures has shown the potential roles of the genes in CC. The visualized results were filtered with  $P < 0.05$  and FDR  $< 0.25$ , as shown in Supplementary Figure 6a-k. These genes were enriched in six different signal pathways, namely KEGG\_REGULATION\_OF\_AUTOPHAGY (*ATG4C*,

*ATG4D*, *KLHL24*, *TP73*), KEGG\_MTOR\_SIGNALING\_PATHWAY (*HGS*, *VAMP7*), KEGG\_NEUROTROPHIN\_SIGNALING\_PATHWAY (*KLHL24*), KEGG\_PROSTATE\_CANCER (*SERPINA1*, *SUPT20H*), KEGG\_PROSTATE\_CANCER (*SUPT20H*) and KEGG\_FOCAL\_

ADHESION (*SUPT20H*). Details were shown in Additional1 file 6.

### **Risk score could be applied as a sensitive indicator of target therapy**

Pearson correlation analysis manifested our OS-related risk score was significantly related to the mRNA expression level of *AKT1* (cor = 0.1565,  $P = 0.0075$ ), *BCL2* (cor = -0.2589,  $P < 0.0001$ ), *MTOR* (cor = 0.2013,  $P = 0.0006$ ), *TP53* (cor = 0.2013,  $P = 0.0006$ ), *VEGFA* (cor = -0.3215,  $P < 0.0001$ ) and RFS-related risk score was correlated with the mRNA expression level of *AKT1* (cor = 0.2630,  $P = 0.0002$ ), *TP53* (cor = 0.1744,  $P = 0.0130$ ), *VEGFA* (cor = 0.2093,  $P = 0.0028$ ), as shown in Figure 6g-n. These results pointed out that patients with higher OS-related risk score and RFS-related risk score might respond better to therapies targeting *AKT1* and *VEGFA*, patients with lower OS-related risk score might respond better to therapies targeting *BCL2*. Besides, patients with higher OS-related risk score might be more sensitive to drugs targeting *MTOR*. Interestingly, patients with higher OS-related risk score and RFS-related risk score had opposite sensitivity to *TP53*, which was worth exploring further.

### **RNA expression in CC and normal samples**

Finally, among the 12 autophagy biomarkers related to the survival of cervical cancer, based on literature search, we selected 6 genes (*ATG4C*, *ATG4D*, *CD46*, *TP73*, *SERPINA1* and *HSPB8*) to verify their RNA expression in CC and normal cervical tissue samples. The results showed that except for *SERPINA1*, the other five genes were down-regulated in cervical cancer (Figure 7a-f, all  $P \leq 0.05$ ). This result shows that *ATG4C*, *ATG4D*, *CD46*, *TP73* and *HSPB8* may play a protective role in the progression of cervical cancer. Interestingly, *ATG4D*, *CD46*, *TP73* and *HSPB8* were negatively correlated with the risk score, which is consistent with the experimental results. *ATG4C* is positively correlated with the risk score, which is contrary to the above qPCR results.

## **Discussion**

CC is the most universal gynecological malignancy. In economically underdeveloped countries, the incidence and mortality of CC remain high all year round. 2018 International Federation of Obstetrics and Gynecology (FIGO) CC patient staging system proposed for the first time that imaging and pathological examination results should be included in the staging standard, which could reduce unnecessary surgery and the risk of complications in patients [21], but it had obvious limitations for patients with lymph node metastasis. Recently prognostic models of different cancer patients including CC have been constructed based on bioinformatics technology [14, 22–24]. It has been reported autophagy has a significant impact on the occurrence and development of tumors, autophagy-related genes can serve as prognostic markers [25–27]. However, few studies have reported the predictive relationship of CC transcriptome expression data on survival at the level of autophagy.

Inspired by the reported cellular and molecular effects of autophagy in CC, here we adopted autophagy-related genes expressed in CC rather than differentially expressed autophagy-related genes to construct prognostic models. After Cox regression analysis, we finally found nine autophagy-related risk signatures (*VAMP7*, *MTMR14*, *ATG4D*, *KLHL24*, *TP73*, *NAMPT*, *CD46*, *HGS*, *ATG4C*) were related to OS, and three autophagy-related risk signatures (*SERPINA1*, *HSPB8*, *SUPT20H*) were related to RFS. K-M survival analysis and the AUC illustrated our models had a great predictive performance. Additionally, the risk scores of the OS and RFS model could be considered as independent predictive indicators. Finally, the nomogram and calibration chart showed that the risk signatures could accurately assess the survival of CC patients. Vesicle-Associated Membrane Protein 7 (*VAMP7*) is a member of the synaptobrevin family which was found to be highly expressed and associated with a shorter survival period in esophageal cancer which could act as a dangerous gene[28]. However, we found out that *VAMP7* and risk score were negatively correlated contrary to previous conclusions. Myotubularin-related protein 14 (*MTMR14*) is a member of the myotubularin (MTM)-related protein family. Knockdown of *MTMR14* promoted cell apoptosis and inhibited migration in liver cancer cells[29]. However, Liu *et al.* put forward an opposite conclusion, they believed that the deficiency of *MTMR14* could promote autophagy and proliferation of mouse embryonic fibroblasts[30], which is consistent with our results. Autophagy Related 4D Cysteine Peptidase (*ATG4D*) and its paralog gene *ATG4C* played an important role as autophagy regulators that linked mitochondrial dysfunction with apoptosis in cancers such as breast carcinoma, human uterine fibroids, colorectal cancer and glioma[31–34]. Here we found that *ATG4C* might act as a risk factor, while *ATG4D* did the opposite. *KLHL24* has been uncovered as a prognostic predictor for acute myeloid leukemia[35]. Bockers suggested that *KLHL24* associated with *ESR1* could act as the upstream regulator to promote breast tumor growth[36]. In CC, high expression level of *KLHL24* can predict a better prognosis. *TP73* is a member of the *TP53* family whose expression has been observed altered in most human cancers and associated with the prognosis. In our present study, *TP73* played a protective role in the prognosis of CC. Nicotinamide phosphoribosyltransferase (*NAMPT*) possesses various functions in human cells, and altered *NAMPT* expression is associated with human carcinogenesis[37–39]. Currently, inhibition of *NAMPT* as a therapeutic strategy in cancer has attracted more and more attention. Our research once again confirmed the dangerous role of *NAMPT*. *CD46* has been found to be a prognostic indicator and highly expressed *CD46* was associated with poor prognosis[40, 41], which strongly supported our results. Hepatocyte Growth Factor-Regulated Tyrosine Kinase Substrate (*HGS*) is one of the master regulators whose expression gradually increased with CRC canceration and was an independent poor prognostic factor[42]. We also found that overexpressed *HGS* was associated with poor prognosis in CC. Serpin Family A Member 1 (*SERPINA1*) is a protein with a highly conserved structure and the most important protease inhibitor in humans whose expression is increased in malignant tumors, and it has been proven to be a predictor of poor prognosis for high-grade gliomas (HGGs), cutaneous squamous cell carcinoma (SCC) and non-small-cell lung cancer (NSCLC)[43–45]. Contrary to the above conclusion, *SERPINA1* played an inhibitory role in the poor prognosis of CC, which suggested that different cancers had different mechanisms. Heat Shock Protein Family B (Small) Member 8 (*HSPB8*) belongs to the small heat shock protein(*HSP20*)family. It has been reported that high *HSPB8* levels have a worse prognosis than low *HSPB8* expression[46, 47]. Veysiére *et al.* proposed that the mutation of *SPT20* Homolog, SAGA

Complex Component (*SUPT20H*) was significantly related to Rheumatoid Arthritis (RA)[48]). Additionally, *SUPT20H* could serve as an independent prognostic biomarker in gliomas and multiple myeloma (MM) [49, 50].

Next, we selected six genes(*ATG4C*, *ATG4D*, *CD46*, *TP73*, *SERPINA1* and *HSPB8*) from the above twelve risk biomarkers for qPCR verification according to literature retrieval. The experimental results confirm that *ATG4D*, *CD46*, *TP73* and *HSPB8* are decreased in cervical cancer, and the prognosis model shows a negative correlation with the risk score, it could be inferred that they had a protective effect on the occurrence and prognosis of cervical cancer. *ATG4C* expression is decreased in cervical cancer, but it is a risk factor in the prognosis model, it is worthy of further study.

The gene expression information and clinical survival information of patients with endometrial cancer (TCGA-UCEC) were validated to determine the predictive value of the OS model. Similarly, head and neck cancer (GSE117973) was downloaded to identify the predictive value of the RFS model. The results of the external validation set showed that our risk score was significantly related to the survival of patients, and a high risk score could indicate a poor prognosis for patients. Although the type of cancer in the testing set was different from the training set, studies have shown radiotherapy for CC can increase the incidence of endometrial cancer[19]. One of the important pathogenic factors for head and neck cancer is HPV infection[51]. So we believed that the prediction models for CC patients could be applied for endometrial cancer and head and neck cancer. These results certificated that the models had wide applicability.

Bioinformatics functional analysis showed that 53 DE-ARGs were mainly enriched in autophagy and apoptosis pathways. Based on the results, we hypothesized that autophagy played an important role in the occurrence and development of CC. Single-gene GSEA revealed the risk signatures included in our computational prognostic models could participate in autophagy, cell adhesion to affect the cell cycle, cell proliferation and migration.

Finally, we further explored whether model genes could be regarded as diagnostic markers. Interestingly, ROC curve analysis revealed that the patient might be accurately distinguished from the normal samples according to the expression of the risk signatures. In the present study, prognostic markers and diagnostic signatures were combined in CC for the first time, which explained that the incidence and prognosis of cancer were continuous and indivisible. Additionally, OS-related risk score and RFS-related risk score could prompt the patient's sensitivity to the target drug to guide reasonable clinical treatment.

## Conclusion

In summary, we have constructed predictive models to predict the OS and RFS of CC patients based on ARGs expressed in CC and verified the effectiveness of the models in external validation sets. Furthermore, these model genes are expected to be diagnostic and treatment indicators, which may require a larger sample size to confirm this finding.

# Declarations

## Ethics approval and consent to participate

All the subjects involved have been informed of the purpose and significance of this study, signed the informed consent and provided the necessary biological samples for testing and analysis. The study has been approved by the ethics committee of Southeast University.

## Consent for publication

All the subjects involved have agreed to publish relevant research results.

## Availability of data and material

Further information and requests should be directed to the Lead Contact, Shizhi Wang ([shizhiwang2009@seu.edu.cn](mailto:shizhiwang2009@seu.edu.cn)). The datasets supporting the conclusions of this article are available at the TCGA data portal (<https://tcga-data.nci.nih.gov/tcga/>), the GTEx portal (<http://gtexportal.org>) and the comprehensive Gene Expression Omnibus (GEO;<https://www.ncbi.nlm.nih.gov/geo/>)

## Competing interests

There is no conflict of interest.

## Funding

This work was supported by Natural Science Foundation of China (81872684), the Fundamental Research Funds for the Central Universities, Southeast University "Zhongying Young Scholars" Project, the Six Talent Peaks Project in Jiangsu Province (wsw-201), the "SIX ONE" Talent Research Project for the High-level Health Personnel of Jiangsu Province (LGY2018037), the Fifth Scientific Research Project of Nantong ("226 Project"), the Research Project from Nantong Commission of Health (MB2020018), National Natural Science Foundation of China (81802105), Jiangsu Provincial Key Research and Development Program (BE2017619), and Nanjing Technological Development Program (201605049).

## Authors' contributions

D.M. conceived the study, performed data processing and wrote the article. X.Z. and W.Y. downloaded datasets of CC. S.W. critically revised the article for research content and administrative support. The final manuscript was read and approved by all authors.

## Acknowledgements

Not applicable.

# References

1. Cohen PA, Jhingran A, Oaknin A, Denny L. Cervical cancer. *Lancet (London, England)* 393(10167), 169-182 (2019).10.1016/s0140-6736(18)32470-x
2. Jin J. HPV Infection and Cancer. *Jama* 319(10), 1058 (2018).10.1001/jama.2018.0687
3. 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: a cancer journal for clinicians* 68(6), 394-424 (2018).10.3322/caac.21492
4. Bjurberg M, Beskow C, Kannisto P, Lindahl G. [Cervical cancer is a clinical challenge]. *Lakartidningen* 112 (2015).
5. Pimple SA, Mishra GA. Global strategies for cervical cancer prevention and screening. *Minerva ginecologica* 71(4), 313-320 (2019).10.23736/s0026-4784.19.04397-1
6. Mattoscio D, Medda A, Chiocca S. Human Papilloma Virus and Autophagy. *International journal of molecular sciences* 19(6), (2018).10.3390/ijms19061775
7. Chen H, Deng Q, Wang W, Tao H, Gao Y. Identification of an autophagy-related gene signature for survival prediction in patients with cervical cancer. *Journal of ovarian research* 13(1), 131 (2020).10.1186/s13048-020-00730-8
8. Xilouri M, Stefanis L. Chaperone mediated autophagy in aging: Starve to prosper. *Ageing research reviews* 32 13-21 (2016).10.1016/j.arr.2016.07.001
9. Li WW, Li J, Bao JK. Microautophagy: lesser-known self-eating. *Cellular and molecular life sciences : CMLS* 69(7), 1125-1136 (2012).10.1007/s00018-011-0865-5
10. Feng Y, He D, Yao Z, Klionsky DJ. The machinery of macroautophagy. *Cell research* 24(1), 24-41 (2014).10.1038/cr.2013.168
11. Hu YF, Lei X, Zhang HY *et al.* Expressions and clinical significance of autophagy-related markers Beclin1, LC3, and EGFR in human cervical squamous cell carcinoma. *OncoTargets and therapy* 8 2243-2249 (2015).10.2147/ott.S86844
12. Li X, Gong Z, Zhang L *et al.* Autophagy knocked down by high-risk HPV infection and uterine cervical carcinogenesis. *International journal of clinical and experimental medicine* 8(7), 10304-10314 (2015).
13. Zhang L, Liu X, Song L, Zhai H, Chang C. MAP7 promotes migration and invasion and progression of human cervical cancer through modulating the autophagy. *Cancer cell international* 20 17 (2020).10.1186/s12935-020-1095-4
14. Wu W, Sui J, Liu T *et al.* Integrated analysis of two-lncRNA signature as a potential prognostic biomarker in cervical cancer: a study based on public database. *PeerJ* 7 e6761 (2019).10.7717/peerj.6761
15. Ding H, Fan GL, Yi YX, Zhang W, Xiong XX, Mahgoub OK. Prognostic Implications of Immune-Related Genes' (IRGs) Signature Models in Cervical Cancer and Endometrial Cancer. *Frontiers in genetics* 11 725 (2020).10.3389/fgene.2020.00725
16. Wang XB, Cui NH, Liu XN *et al.* Identification of DAPK1 Promoter Hypermethylation as a Biomarker for Intra-Epithelial Lesion and Cervical Cancer: A Meta-Analysis of Published Studies, TCGA, and GEO

- Datasets. *Frontiers in genetics* 9 258 (2018).10.3389/fgene.2018.00258
17. Rieke DT, Keilholz U. Systemic Treatment in HPV-Induced Recurrent or Metastatic HNSCC. *Recent results in cancer research. Fortschritte der Krebsforschung. Progres dans les recherches sur le cancer* 206 149-160 (2017).10.1007/978-3-319-43580-0\_11
  18. Brianti P, De Flammineis E, Mercuri SR. Review of HPV-related diseases and cancers. *The new microbiologica* 40(2), 80-85 (2017).
  19. Lopes J, Horta M, Cunha TM. Endometrial cancer after radiation therapy for cervical carcinoma: A radiological approach. *European journal of radiology* 105 283-288 (2018).10.1016/j.ejrad.2018.06.008
  20. Panina Y, Germond A, Masui S, Watanabe TM. Validation of Common Housekeeping Genes as Reference for qPCR Gene Expression Analysis During iPS Reprogramming Process. *Scientific reports* 8(1), 8716 (2018).10.1038/s41598-018-26707-8
  21. Yan DD, Tang Q, Chen JH, Tu YQ, Lv XJ. Prognostic value of the 2018 FIGO staging system for cervical cancer patients with surgical risk factors. *Cancer management and research* 11 5473-5480 (2019).10.2147/cmar.S203059
  22. Liang B, Li Y, Wang T. A three miRNAs signature predicts survival in cervical cancer using bioinformatics analysis. *Scientific reports* 7(1), 5624 (2017).10.1038/s41598-017-06032-2
  23. Dai F, Chen G, Wang Y *et al.* Identification of candidate biomarkers correlated with the diagnosis and prognosis of cervical cancer via integrated bioinformatics analysis. *OncoTargets and therapy* 12 4517-4532 (2019).10.2147/ott.S199615
  24. Gao C, Zhou C, Zhuang J *et al.* MicroRNA expression in cervical cancer: Novel diagnostic and prognostic biomarkers. *Journal of cellular biochemistry* 119(8), 7080-7090 (2018).10.1002/jcb.27029
  25. An Y, Bi F, You Y, Liu X, Yang Q. Development of a Novel Autophagy-related Prognostic Signature for Serous Ovarian Cancer. *Journal of Cancer* 9(21), 4058-4071 (2018).10.7150/jca.25587
  26. Qiu J, Sun M, Wang Y, Chen B. Identification and validation of an individualized autophagy-clinical prognostic index in gastric cancer patients. *Cancer cell international* 20 178 (2020).10.1186/s12935-020-01267-y
  27. Wang SS, Chen G, Li SH *et al.* Identification and validation of an individualized autophagy-clinical prognostic index in bladder cancer patients. *OncoTargets and therapy* 12 3695-3712 (2019).10.2147/ott.S197676
  28. Chen C, Chen S, Cao H *et al.* Prognostic significance of autophagy-related genes within esophageal carcinoma. *BMC cancer* 20(1), 797 (2020).10.1186/s12885-020-07303-4
  29. Li Z, Rong L, Lian H, Cheng J, Wu X, Li X. Knockdown MTMR14 promotes cell apoptosis and inhibits migration in liver cancer cells. *Gene* 691 106-113 (2019).10.1016/j.gene.2018.11.099
  30. Liu J, Lv Y, Liu QH, Qu CK, Shen J. Deficiency of MTMR14 promotes autophagy and proliferation of mouse embryonic fibroblasts. *Molecular and cellular biochemistry* 392(1-2), 31-37 (2014).10.1007/s11010-014-2015-5

31. Zhang XM, Li HJ, Wang D, Li C, Chen Q, Li CL. [Expressions and Clinical Significance of Autophagy-related Genes ATG2B, ATG4D, ATG9B in Breast Carcinoma]. *Sichuan da xue xue bao. Yi xue ban = Journal of Sichuan University. Medical science edition* 47(2), 184-188 (2016).
32. Andaloussi AE, Habib S, Soylemes G *et al.* Defective expression of ATG4D abrogates autophagy and promotes growth in human uterine fibroids. *Cell death discovery* 3 17041 (2017).10.1038/cddiscovery.2017.41
33. Gil J, Ramsey D, Pawlowski P *et al.* The Influence of Tumor Microenvironment on ATG4D Gene Expression in Colorectal Cancer Patients. *Medical oncology (Northwood, London, England)* 35(12), 159 (2018).10.1007/s12032-018-1220-6
34. Fu Z, Luo W, Wang J *et al.* Malat1 activates autophagy and promotes cell proliferation by sponging miR-101 and upregulating STMN1, RAB5A and ATG4D expression in glioma. *Biochemical and biophysical research communications* 492(3), 480-486 (2017).10.1016/j.bbrc.2017.08.070
35. Chen XX, Li ZP, Zhu JH, Xia HT, Zhou H. Systematic Analysis of Autophagy-Related Signature Uncovers Prognostic Predictor for Acute Myeloid Leukemia. *DNA and cell biology* 39(9), 1595-1605 (2020).10.1089/dna.2020.5667
36. Böckers M, Paul NW, Efferth T. Organophosphate ester tri-o-cresyl phosphate interacts with estrogen receptor  $\alpha$  in MCF-7 breast cancer cells promoting cancer growth. *Toxicology and applied pharmacology* 395 114977 (2020).10.1016/j.taap.2020.114977
37. Zhou SJ, Bi TQ, Qin CX, Yang XQ, Pang K. Expression of NAMPT is associated with breast invasive ductal carcinoma development and prognosis. *Oncology letters* 15(5), 6648-6654 (2018).10.3892/ol.2018.8164
38. Lucena-Cacace A, Otero-Albiol D, Jiménez-García MP, Muñoz-Galvan S, Carnero A. NAMPT Is a Potent Oncogene in Colon Cancer Progression that Modulates Cancer Stem Cell Properties and Resistance to Therapy through Sirt1 and PARP. *Clinical cancer research : an official journal of the American Association for Cancer Research* 24(5), 1202-1215 (2018).10.1158/1078-0432.Ccr-17-2575
39. Guo Q, Han N, Shi L *et al.* NAMPT: A potential prognostic and therapeutic biomarker in patients with glioblastoma. *Oncology reports* 42(3), 963-972 (2019).10.3892/or.2019.7227
40. Surowiak P, Materna V, Maciejczyk A *et al.* CD46 expression is indicative of shorter revival-free survival for ovarian cancer patients. *Anticancer research* 26(6c), 4943-4948 (2006).
41. Maciejczyk A, Szelachowska J, Szynglarewicz B *et al.* CD46 Expression is an unfavorable prognostic factor in breast cancer cases. *Applied immunohistochemistry & molecular morphology : AIMM* 19(6), 540-546 (2011).10.1097/PAI.0b013e31821a0be9
42. Sun Y, Zheng W, Guo Z *et al.* A novel TP53 pathway influences the HGS-mediated exosome formation in colorectal cancer. *Scientific reports* 6 28083 (2016).10.1038/srep28083
43. Ookawa S, Wanibuchi M, Kataoka-Sasaki Y *et al.* Digital Polymerase Chain Reaction Quantification of SERPINA1 Predicts Prognosis in High-Grade Glioma. *World neurosurgery* 111 e783-e789 (2018).10.1016/j.wneu.2017.12.166

44. Farshchian M, Kivisaari A, Ala-Aho R *et al.* Serpin peptidase inhibitor clade A member 1 (SerpinA1) is a novel biomarker for progression of cutaneous squamous cell carcinoma. *The American journal of pathology* 179(3), 1110-1119 (2011).10.1016/j.ajpath.2011.05.012
45. Ercetin E, Richtmann S, Delgado BM *et al.* Clinical Significance of SERPINA1 Gene and Its Encoded Alpha1-antitrypsin Protein in NSCLC. *Cancers* 11(9), (2019).10.3390/cancers11091306
46. Shen J, Li M, Min L. HSPB8 promotes cancer cell growth by activating the ERK-CREB pathway and is indicative of a poor prognosis in gastric cancer patients. *Oncology reports* 39(6), 2978-2986 (2018).10.3892/or.2018.6376
47. Yang W, Niu L, Zhao X *et al.* Development and validation of a survival model based on autophagy-associated genes for predicting prognosis of hepatocellular carcinoma. *American journal of translational research* 12(10), 6705-6722 (2020).
48. Veyssiere M, Perea J, Michou L *et al.* A novel nonsense variant in SUPT20H gene associated with Rheumatoid Arthritis identified by Whole Exome Sequencing of multiplex families. *PloS one* 14(3), e0213387 (2019).10.1371/journal.pone.0213387
49. Xu Y, Li R, Li X *et al.* An Autophagy-Related Gene Signature Associated With Clinical Prognosis and Immune Microenvironment in Gliomas. *Frontiers in oncology* 10 571189 (2020).10.3389/fonc.2020.571189
50. Zhu FX, Wang XT, Zeng HQ, Yin ZH, Ye ZZ. A predicted risk score based on the expression of 16 autophagy-related genes for multiple myeloma survival. *Oncology letters* 18(5), 5310-5324 (2019).10.3892/ol.2019.10881
51. Syrjänen S, Rautava J, Syrjänen K. HPV in Head and Neck Cancer-30 Years of History. *Recent results in cancer research. Fortschritte der Krebsforschung. Progres dans les recherches sur le cancer* 206 3-25 (2017).10.1007/978-3-319-43580-0\_1

## Tables

**Table 1** Multivariate Cox regression analyses of OS and RFS in CC patients

OS			
Gene	Coefficient	HR(95%CI)	Pvalue
<i>VAMP7</i>	-0.99787	0.36866 (0.21117-0.64362)	0.00045
<i>CD46</i>	0.71823	2.05081(1.29362-3.25118)	0.00225
<i>NAMPT</i>	0.41726	1.51780(1.13852-2.02344)	0.00445
<i>ATG4C</i>	1.02923	2.79890(1.35789-5.76911)	0.00529
<i>HGS</i>	0.75585	2.12943(1.20152-3.77393)	0.00963
<i>TP73</i>	-0.20373	0.81569(0.69885-0.95205)	0.00979
<i>MTMR14</i>	-0.64426	0.52505(0.31495-0.87530)	0.01348
<i>KLHL24</i>	-0.28758	0.75008(0.56843-0.98978)	0.04209
<i>ATG4D</i>	-0.36031	0.69746(0.49070-0.99134)	0.04460
RFS			
Gene	Coefficient	HR(95%CI)	Pvalue
<i>SERPINA1</i>	-0.34683	0.70693(0.55833-0.89507)	0.00397
<i>SUPT20H</i>	1.15581	3.17661(1.47404-6.84571)	0.00317
<i>HSPB8</i>	-0.24263	0.78457(0.63195-0.97403)	0.02792

**Table 2** Clinicopathological features of cervical cancer patients

<b>OS</b>			
<b>Clinical parameters</b>	<b>Variable</b>	<b>Total(253)</b>	<b>Percentages</b>
Age	≤50	154	60.7%
	≥50	99	39.3%
Pathological stage	Stage I	131	51.8%
	Stage II	60	23.7%
	Stage III	37	14.6%
	Stage IV	20	8.0%
	Unknown	5	2.0%
Histological grade	G1	15	5.9%
	G2	117	46.2%
	G3	99	39.1%
	G4	1	0.4%
	GX	21	8.3%
T	T1	115	45.5%
	T2	63	24.9%
	T3	15	5.9%
	T4	10	4.0%
	TX	50	19.8%
Race	American Indian or Alaska Native	6	2.4%
	Asian	18	7.1%
	Black or African American	23	9.1%
	Native Hawaiian or other Pacific	1	0.4%
	White	182	71.9%
	Unknown	23	9.1%
BMI	≤24	80	31.6%
	≥24	173	68.4%
<b>RFS</b>			
<b>Clinical parameters</b>	<b>Variable</b>	<b>Total(208)</b>	<b>Percentages</b>

Age	≤50	133	63.9%
	>50	75	36.1%
Pathological stage	Stage I	113	54.3%
	Stage II	50	24.0%
	Stage III	34	16.3%
	Stage IV	11	5.3%
Histological grade	G1	17	8.2%
	G2	101	48.6%
	G3	89	42.8%
	G4	1	0.5%

## Supplementary Information

### Supplementary Figure 1 The box plot of 53 DEARGs and barplot of PPI network

**a.** The expression pattern of DEARGs in CC and normal samples. **b.** The number of nodes in the top 30 autophagy-related genes of the PPI network.

### Supplementary Figure 2 Immunohistochemistry (IHC) results showing protein levels of autophagy-related genes in CC and normal tissues

**a.** IHC results of *ATG4C* in CC (staining:Medium; intensity:Moderate; quantity:>75%; location: Cytoplasmic membranous) and in normal tissue (staining:Low;intensity:Weak; quantity:75%-25%; location:Cytoplasmic membranous). **b.** IHC results of *ATG4D* in CC (staining:Not detected; intensity:Negative; quantity:None; location:None) and in normal tissue(staining: Low;intensity:Weak; quantity:75%-25%; location:Cytoplasmic membranous). **c.** IHC results of *CD46* in CC(staining:Medium; intensity:Moderate; quantity:>75%; location:Cytoplasmic membranous) and in normal tissue(staining:Low; intensity:Weak; quantity:>75%; location:

Cytoplasmic membranous). **d.** IHC results of *HGS* in CC(staining:Not detected; intensity:Negative; quantity:None; location:None) and in normal tissue(staining:Not detected; intensity:Negative; quantity:None; location:None). **e.** IHC results of *HSPB8* in CC(staining: Medium; intensity:Moderate; quantity:>75%; location:Cytoplasmic membranous) and in normal tissue(staining:High; intensity:Strong; quantity:>75%; location:Cytoplasmic membranous). **f.** IHC results of *KLHL24* in CC(staining:Not detected; intensity:Negative; quantity:None; location:None) and in normal tissue(staining:Not detected; intensity:Negative; quantity:None; location:None). **g.** IHC results of *MTMR14* in CC(staining:Not detected; intensity:Negative; quantity:None; location: None) and in normal tissue(staining:Not detected; intensity:Negative; quantity:None; location: None). **h.** IHC results of *NAMPT* in CC(staining:Not detected;

intensity:Negative; quantity:None; location:None) and in normal tissue(staining:Not detected; intensity:Negative; quantity:None; location:None). **i.** IHC results of *SERPINA1* in CC(staining:Not detected; intensity:Negative; quantity:None; location:None) and in normal tissue(staining:Not detected; intensity:Negative; quantity:None; location:None). **j.** IHC results of *SUPT20H* in CC(staining:Medium; intensity:Moderate; quantity:>75%; location:Cytoplasmic membranous) and in normal tissue(staining:Medium; intensity:Moderate; quantity:>75%; location:Cytoplasmic membranous). **k.** IHC results of *TP73* in CC(staining:Not detected; intensity:Negative; quantity:None; location:None) and in normal tissue(staining:Medium;intensity:Strong; quantity:<25% ;location:Nuclear). **l.** IHC results of *VAMP7* in CC(staining:Medium; intensity:Moderate; quantity:>75%; location: Cytoplasmic membranous) and in normal tissue(staining:Medium; intensity:Moderate; quantity: 75%-25%; location:Cytoplasmic membranous).

### **Supplementary Figure 3 Kaplan–Meier survival curves for CC based on the prognostic risk genes**

**a-l** shows the results of *ATG4C*, *ATG4D*, *CD46*, *HGS*, *KLHL24*, *MTMR14*, *NAMPT*, *TP73*, *VAMP7*, *HSPB8*, *SERPINA1* and *SUPT20H* respectively.

### **Supplementary Figure 4 Calibration curve of nomogram diagram and exploration of risk signature in the testing set**

**a-c.** Calibration plots for 1-year , 3-year and 5-year survival for the nomogram of OS model. **d-f.** Calibration plots for 1-year , 3-year and 5-year survival for the nomogram of RFS model. **g.** K-M curve of the high- and low-risk UCEC patients in the testing group for OS model. **h.** The 3-year, 5-year and 7-year ROC curves in the testing group of UCEC patients for OS. **i.** K-M curve of the high- and low-risk HNSCC patients in the testing group for the RFS model. **j.** The 1-year, 3-year and 5-year ROC curves in the testing group of HNSCC patients for RFS.

### **Supplementary Figure 5 AUC of risk signature without diagnostic value**

**a-f.** AUC of risk signature without diagnostic value (*HGS*, *KLHL24*, *NAMPT*, *SERPINA1*, *TP73*, *NAMPT*).

### **Supplementary Figure 6 Single gene function enrichment analysis and target sensitivity detection**

**a-k.** Single gene function enrichment analysis of 12 prognostic risk signatures.

Additional file 1 232 autophagy-related genes from HADb

Additional file 2 217 autophagy-related genes expressed in CC

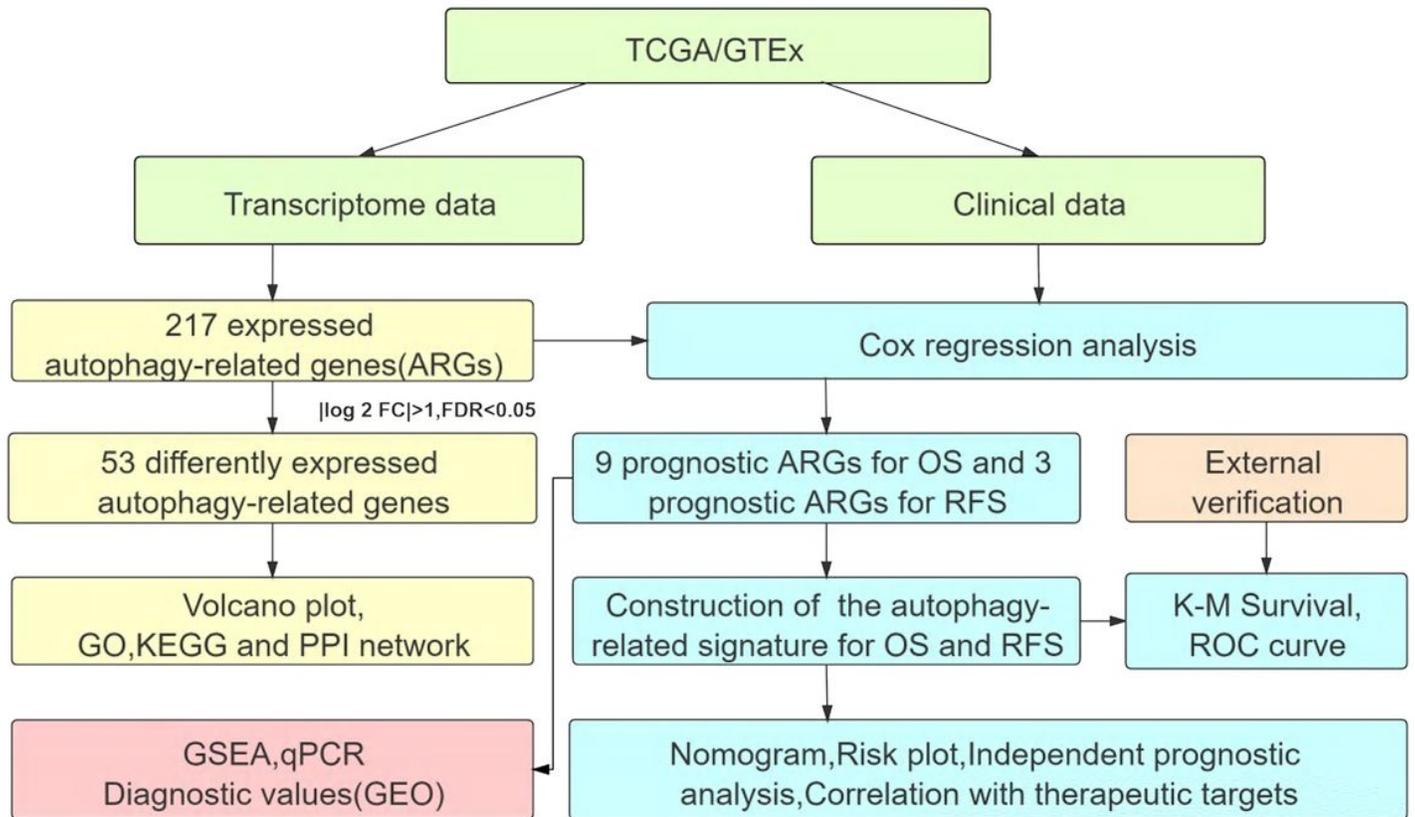
Additional file 3 Primer sequence for qPCR

Additional file 4 53 differently expressed ARGs (DEARGs) in CC

Additional file 5 Significant top 8 GO annotation and top 10 KEGG pathway enrichment

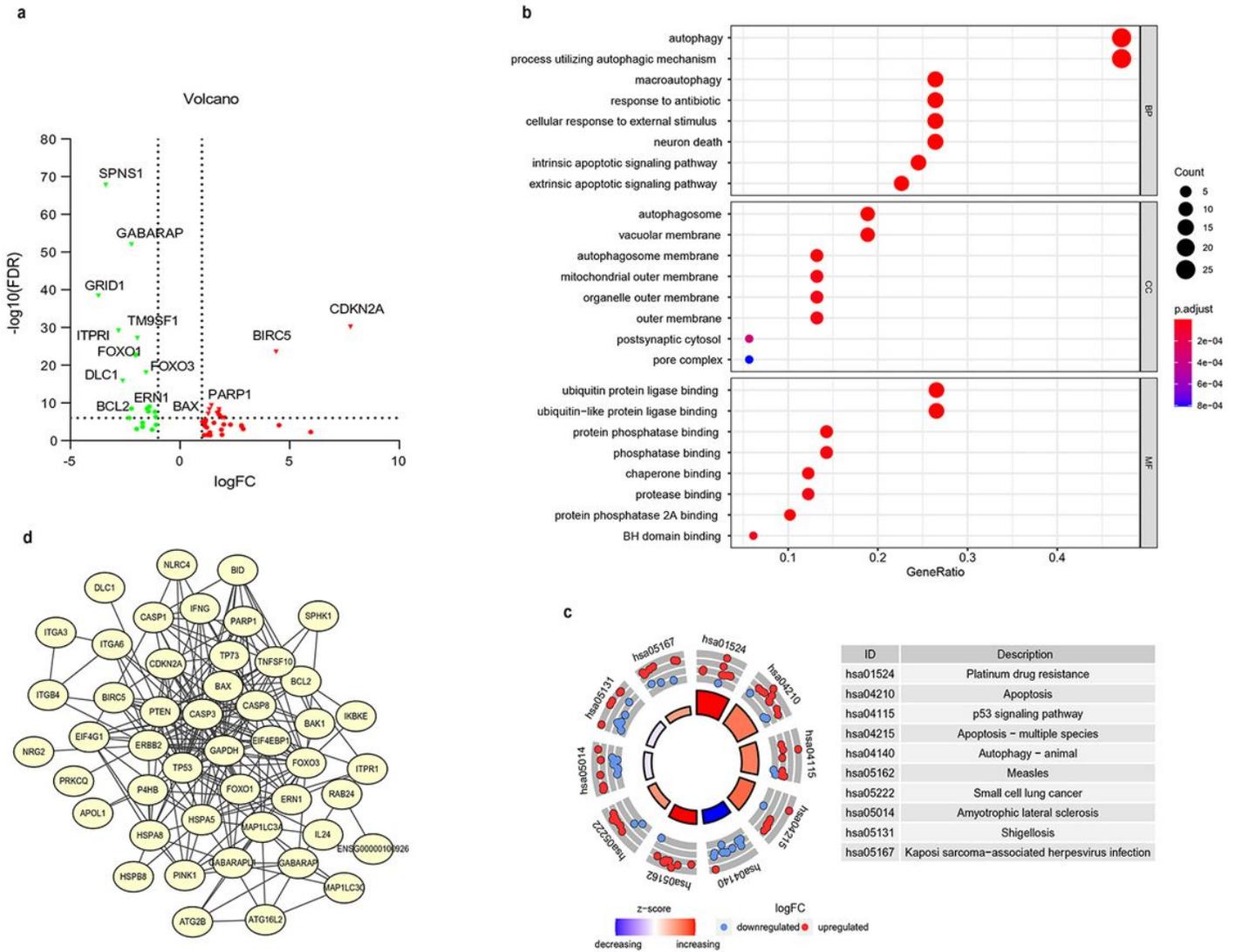
Additional file 6 Detailed results of GSEA enrichment analysis

## Figures



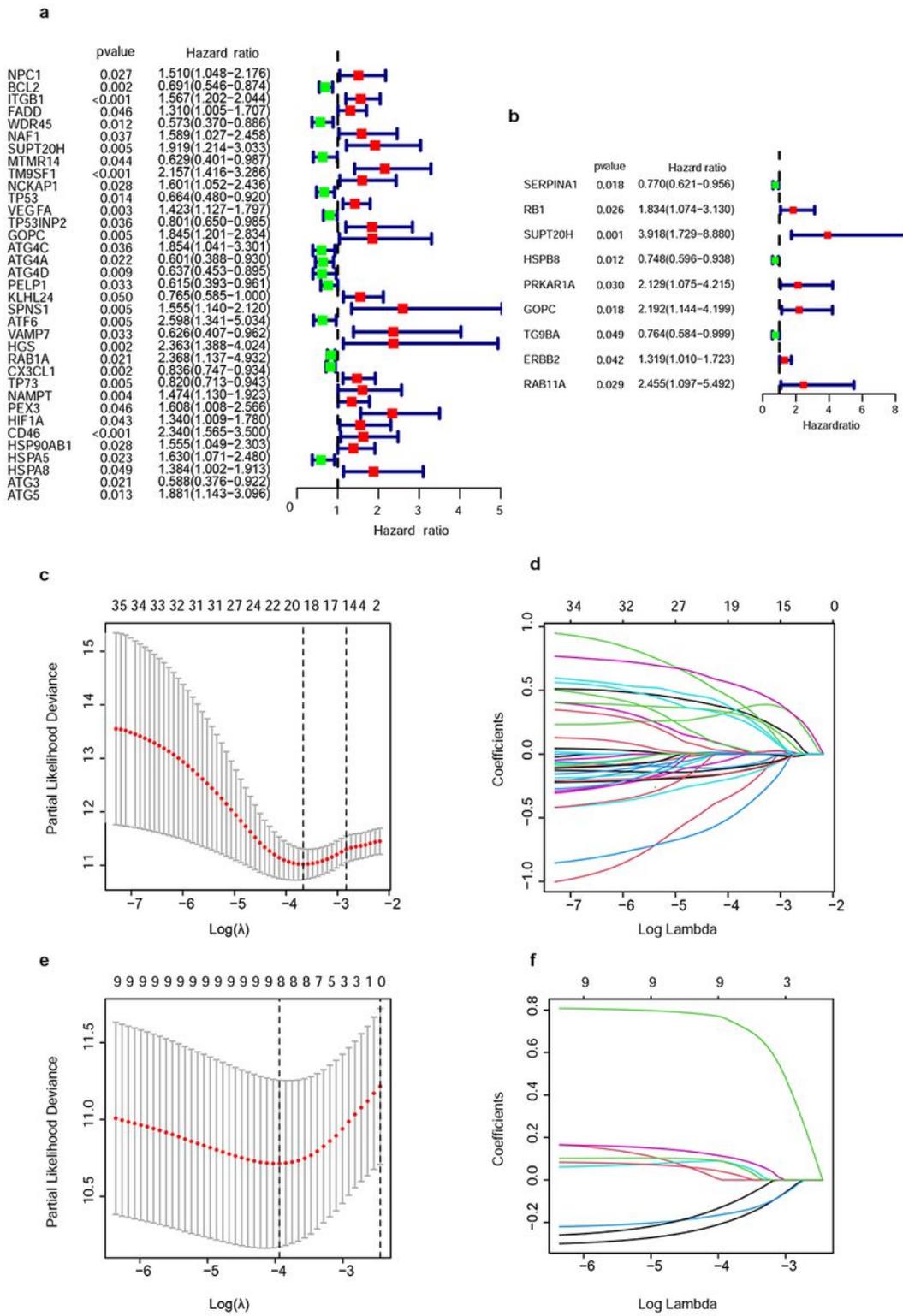
**Figure 1**

The flow diagram of the integral analysis



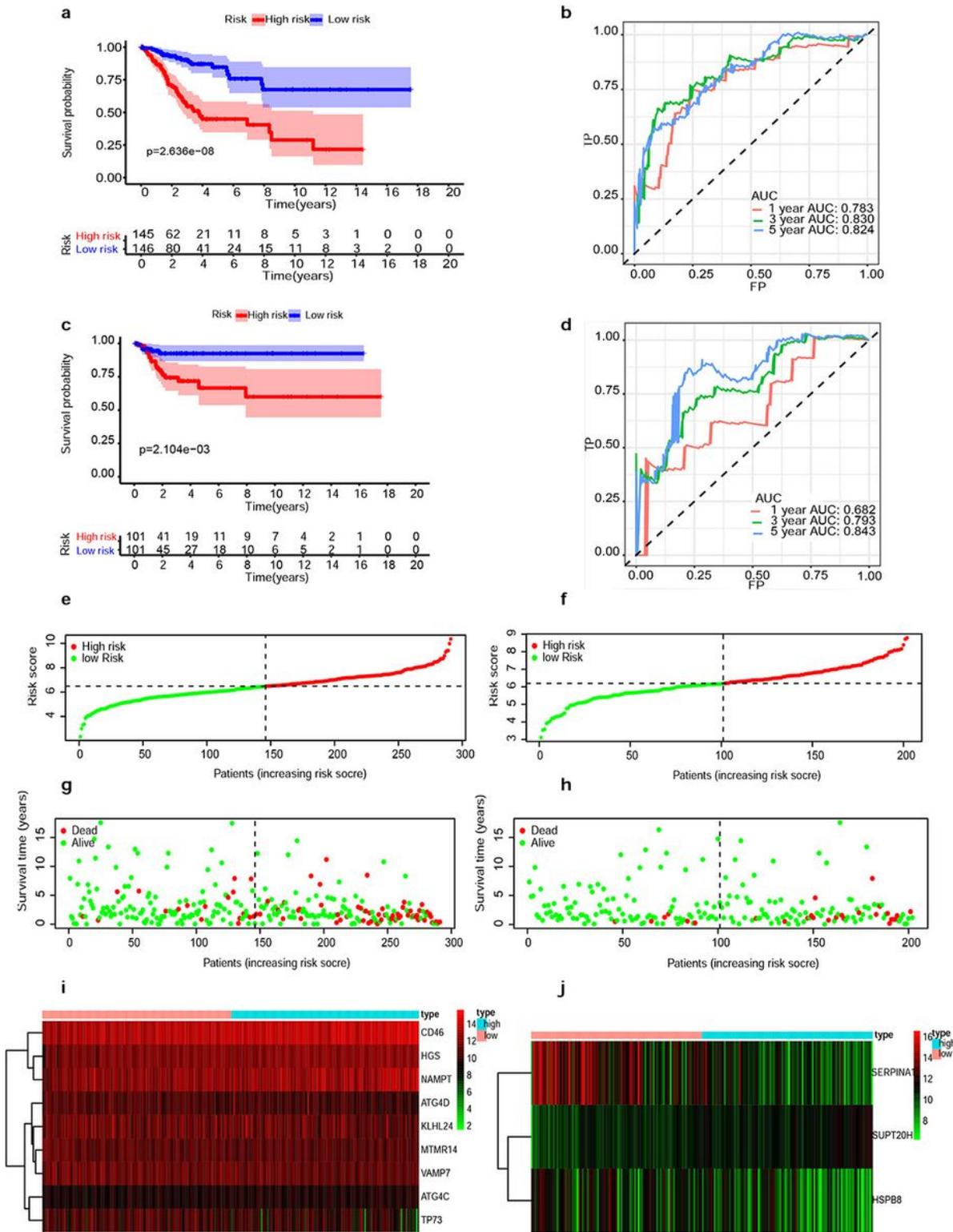
**Figure 2**

Identification and functional annotation of differently expressed autophagy-related genes (DEARGs) in CC patients a. Volcano plot for DEARGs in the tumor and normal samples, in the volcano map, we used the ordinate “6” as the boundary for gene annotation. b. Top 8 GO analysis of 53 DEARGs from the three aspects of BP, CC and MF. c. Top 10 significant KEGG signal pathways. d. PPI network of 53 DEARGs.



**Figure 3**

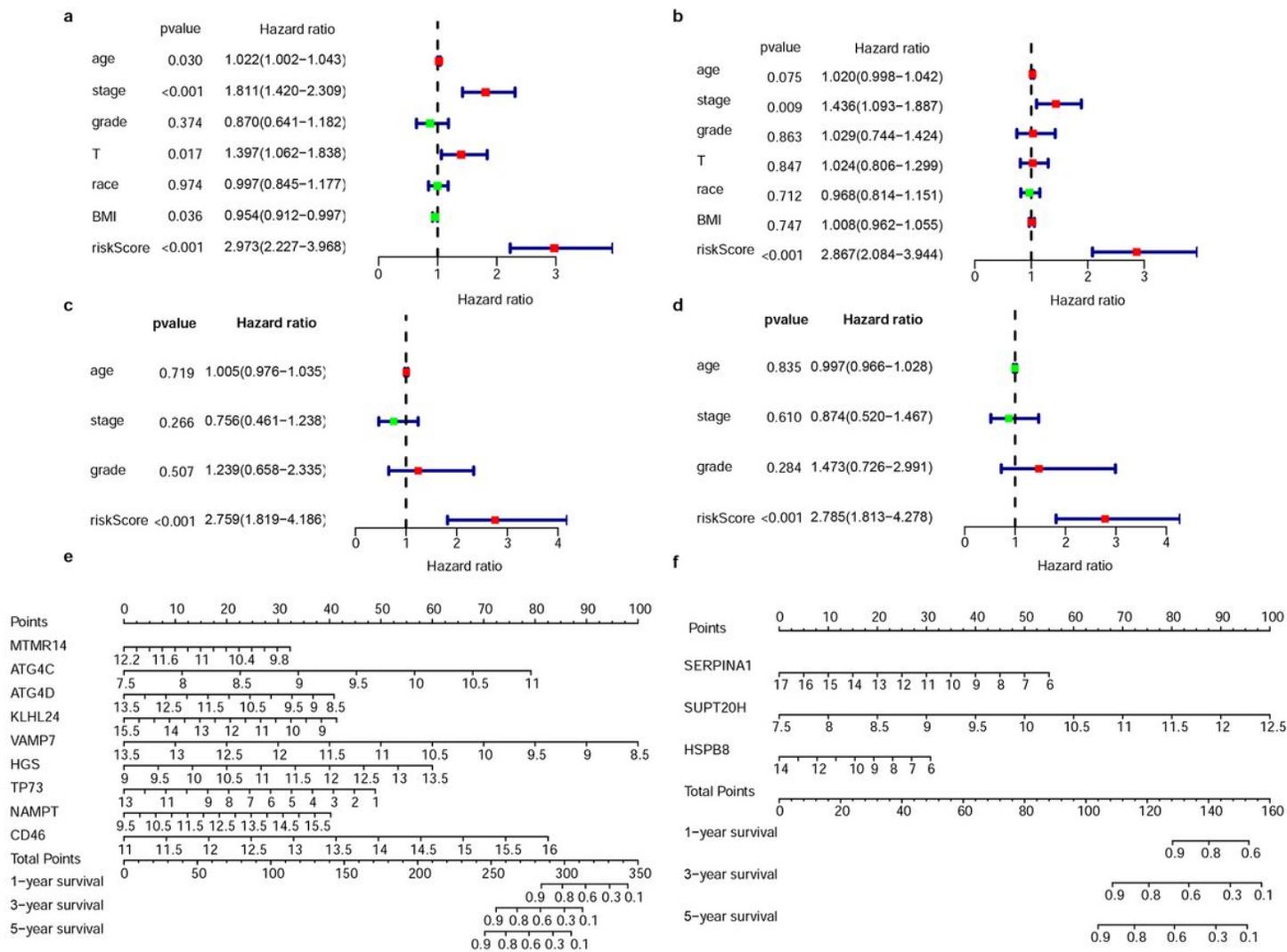
Construction of OS and RFS risk prognostic models a. Univariate Cox regression analysis of the risk signatures for OS. b. Univariate Cox regression analysis of the risk signatures for RFS. c, d. Lasso Cox regression model for OS. e, f. Lasso Cox regression model for RFS.



**Figure 4**

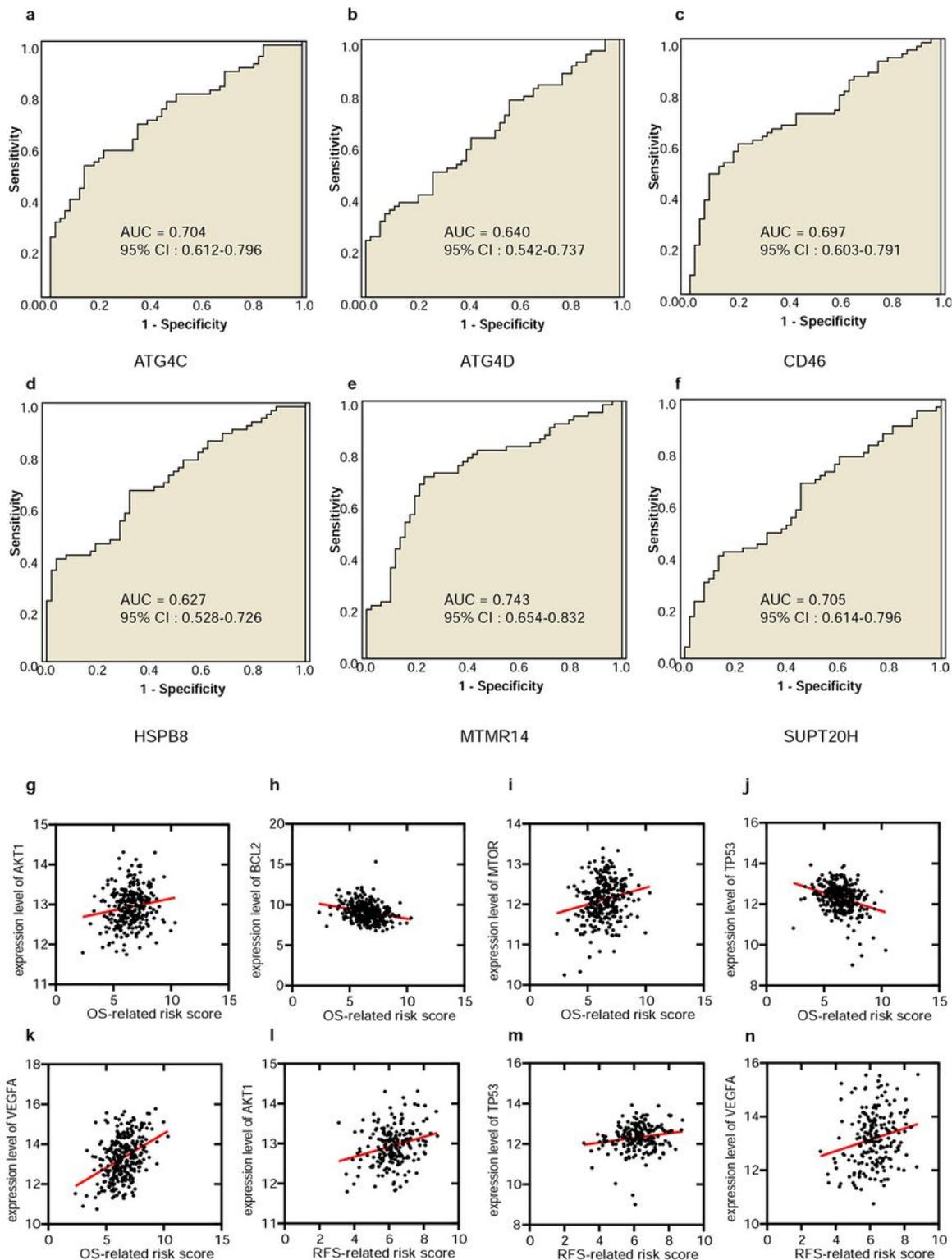
Characteristics of prognostic gene signatures a, c. K-M curves for OS (a) and RFS (c) in the high- and low-risk groups when stratified by the autophagy-related signatures. b, d. ROC curve of risk score at 1,3,5 years for OS (b) and RFS (d) respectively. e, f. Distribution of OS-related risk score and RFS-related risk score, the black dotted line is the optimal cut-off value for dividing patients into low- and high-risk groups.

g, h. Distribution of patient survival time, and status. i, j. Heatmap of autophagy-related gene expression profiles in the prognostic signature of CC.



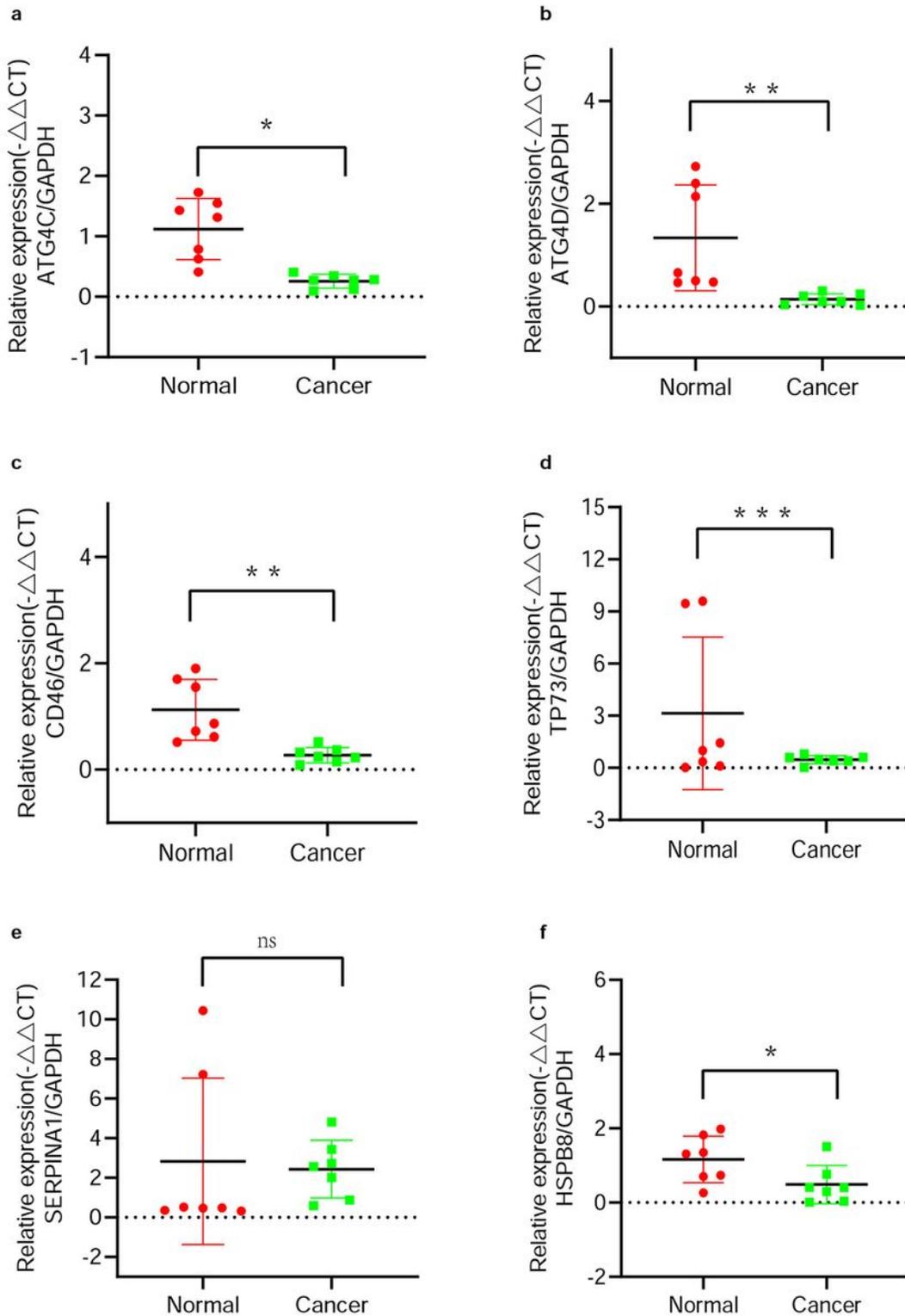
**Figure 5**

Independent prognostic analysis and nomogram diagram a. Univariate Cox regression analysis. Forest plot of associations between risk factors and the survival for OS. b. Multiple Cox regression analysis. The autophagy-associated gene signature is an independent predictor of CC patients for OS. c. Univariate Cox regression analysis. Forest plot of associations between risk factors and the survival for RFS. d. Multiple Cox regression analysis. The autophagy-associated gene signature is an independent predictor of CC patients for RFS. e. A nomogram of the CC cohort (training set) used to predict the OS. f. A nomogram of the CC cohort (training set) used to predict the RFS.



**Figure 6**

Discrimination of diagnostic value and therapeutic target sensitivity based on risk signature a-f. AUC of risk signature with diagnostic value (ATG4C, ATG4D, CD46, HSPB8, MTMR14, SUPT20H). g-k. Correlation of the OS-related risk score and expression level of targets of precise treatment in CC. l-n. Correlation of the RFS-related risk score and expression level of targets of precise treatment in CC.



**Figure 7**

RNA expression of ATG4C, ATG4D, CD46, TP73, SERPINA1 and HSPB8 in CC and normal samples a-f. qPCR determined the RNA expression of ATG4C, ATG4D, CD46, TP73, SERPINA1 and HSPB8 in CC and normal samples. Quantitative normalization of the gene was performed in each sample using GAPDH expression as an internal control.

## Supplementary Files

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

- [Additionalfile16.xlsx](#)
- [SupplementaryFigure16.pdf](#)