Identification of an autophagy-related gene expression prognostic model in endometrial carcinoma patients

Pinping Jiang
Department of Gynecology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, Jiangsu ProvinceJiangsu, China

Wei Sun
Department of Gynecology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, Jiangsu ProvinceJiangsu, China

Ningmei Shen
Department of Gynecology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, Jiangsu ProvinceJiangsu, China

Qiang Wang
Department of Hepatobiliary Surgery, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, Jiangsu Province, ChinaNanjing University Medical School Affiliated Nanjing Drum Tower Hospital, Nanjing 210029, Jiangsu, China

Shouyu Wang (✉ sywang@nju.edu.cn)
Department of Hepatobiliary Surgery, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, Jiangsu Province, ChinaNanjing University Medical School Affiliated Nanjing Drum Tower Hospital, Nanjing 210029, Jiangsu, China. Jiangsu

Shilong Fu (✉ docfusl@163.com)
Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Research

Keywords: Autophagy; TCGA; endometrial carcinoma; prognostic model; nomogram

Posted Date: May 15th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-28086/v1

Abstract

Background

Autophagy, as a lysosomal degradation pathway, has been reported to be involved in various pathologies, including cancer. However, the expression profiles of autophagy-related genes (ARGs) in endometrial cancer (EC) remain poorly understood.

Methods

In this study, we analyzed the expression of MRGs using The Cancer Genome Atlas (TCGA) data to screen differentially expressed MRGs (DE-MRGs) significantly correlated to EC patients' prognosis. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of DE-MRGs were investigated. LASSO algorithm and Cox regression analysis were performed to select MRGs closely related to EC patients' outcomes. A prognostic signature was developed and the efficacy were validated in part of and the entire TCGA EC cohort. Moreover, we developed a comprehensive nomogram including the risk model and clinical features to predict EC patients' survival probability.

Results

Ninety-four ARGs significantly dysregulated in EC samples compared with the normal control samples. Functional enrichment analysis showed these differentially expressed ARGs (DE-ARGs) were highly enriched in apoptosis, P53 signaling pathway, and various cancer development. Among the 94 DE-ARGs, we subsequently screen out four-ARGs closely related to EC patients outcomes, which are ERBB2, PTEN, TP73 and ARSA. Based on the expression and coefficiency of 4 DE-ARGs, we developed a prognostic signature and further validated its efficacy in part of and the entire TCGA EC cohort. The four ARGs signature was independent of other clinical features, and was proved to effectively distinguish high- or low-risk EC patients and predicted patients' OS accurately. Moreover, the nomogram showed the excellent consistency between the prediction and actual observation in terms of patients' 3- and 5-year survival rates.

Conclusions

It was suggested that the ARG prognostic model and the comprehensive nomogram may guide the precise outcome prediction and rational therapy in clinical practice.

Background

Endometrial carcinoma (EC) is one of the most common female reproductive malignancies, which caused nearly 90,000 deaths last year worldwide[1]. Even if early-stage endometrial cancer predicts

favorable prognosis, nearly 30% of patients were still diagnosed at late-advanced stage with regional or distant metastasis, among which, less than 20% of the individuals live longer than 5 years[2]. Currently, the major method to evaluate patients' outcomes is to combine patients' pre-operated physical condition with intraoperative observation, as well as the comprehensive pathological diagnosis after surgical operation. Nevertheless, there are several EC patients presented a high risk of cancer progression or recurrence with insensitivity to chemotherapy, which contributed to the poor prognosis[3]. Due to the limitation of clinical staging system and histological classification for prognostic prediction, it is imperative to develop novel predictive biomarkers to help clinicians evaluate patients' prognosis and guide rational therapy.

Autophagy, as a lysosomal degradation pathway, has been reported to involved in various pathologies, such as cell differentiation and death,cancer progression[4, 5]. It has been widely recognized that autophagy has two-sided effects in cancer development, it could suppress tumorigenesis at an early stage but also promote the development of already existing tumors[6]. In the progression of the cancer, autophagy suppressed the tumorigenesis in the early-stage by reducing cellular damage and chromosomal instability[7]. However, autophagy could remove the trashes s in the cells under oxidative stress which provides nutritional contents and promotes tumor growth.[8]. Therefore, it has more important clinical significance to elucidate the exact roles of ARGs in tumorigenesis and tumor progression.

In this study, we assessed the autophagy-related genes expression profile in EC patients using the Cancer Genome Atlas (TCGA) dataset. Through searching the prognosis related autophagy-related genes (ARGs), we established and validated an ARGs signature which could predict the outcomes in EC. In addition, we integrated the clinical features of patients and the ARGs signature to establish a novel nomogram model providing comprehensive therapeutic guidelines for EC treatment.

Materials And Methods

Data acquisition

From the TCGA data portal (https://tcga-data.nci.nih.gov/tcga/), we downloaded the FPKM format RNA-seq data for 548 endometrial cancer patients, which included 552 endometrial cancer and 35 adjacent non-tumor tissues. The corresponding clinical information of 541 EC patients, which contained overall survival and other clinical characteristics, were retrieved from the cBio Cancer Genomics Portal (cbioPortal, http://cbioportal.org)[9]. We searched a total of 232 autophagy-related genes from The Human Autophagy Database (HADb,http://www.autophagy.lu/index.html).

Differentially expressed ARGs and functional enrichment analysis

"Limma" R package [10] in R was used to perform the differentially-expressed ARGs (DE-ARGs) analysis between ECs and their non-tumor counterparts, with thresholds of log|Fold change (FC)|> 0.5 and P-value< 0.05. The expression level of DE-ARGs from each patient was displayed via the "ggplot" R

package. Then, we performed gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis to explore the predominant biological attributes of such differential expression ARGs via "clusterprofiler" R package[11, 12]. Adjust p-value < 0.05 was considered statistically significant. Next, the GO and KEGG enrichment annotation analysis result maps were visualized by R packages “ggplot2” and “GOplot”.

Protein-protein interaction (PPI) network construction and hub AGRs alteration analysis

The interactions between DE-ARGs were displayed in the protein-protein interaction (PPI) network. PPI network was established via the Search Tool for the Retrieval of Interacting Genes Database database(STRING, https://string-db.org/), which comprised the interaction information among certain proteins[13]. The criteria of the minimum required interaction score was set as 0.9. Then, the network result was visualized by Cytoscape software [14] and the top 15 hub genes were defined according to the ranked connectivity degrees in the string network. The alteration conditions of hub DE-ARGs in endometrial cancer were analyzed by the cbioPortal website.

Establishment of a prognostic model based on DE-ARGs

A total of 541 TCGA EC patients were randomly classified into the training cohort (n = 272) and the testing cohort (n = 269). Univariate Cox regression analyses were performed in the training cohort to find out the OS-related DE-ARGs in EC patients. The Least Absolute Shrinkage and Selection Operator (LASSO) analysis could narrow the high-dimensional regression coefficients and introduce the penalty regularization parameter λ via the cross-validation routine. To selected the key ARGs involved in the prognosis of EC patients, we further performed LASSO and multivariate Cox regression analysis via the “glmnet” R package. The risk-score signature for EC patients' prognosis prediction was to multiply each prognostic ARG expression value with the relative regression coefficient ratio calculated from the multivariate Cox regression model and finally summarize all scores. Based on the median risk score, all training cohort patients were divided into high- and low-risk groups. The outcomes of both subgroups were compared and plotted into Kaplan–Meier curves via the "survfit" R package, and the receiver operating characteristic (ROC) curve of OS prediction was also drawn to estimate the sensitivity and specificity of the prognostic model. The patients' survival status and death time of each training group, as well as the gene expression level, were also exhibited via "pheatmap" and "survival" R packages.

Evaluation of the efficacy of the prognostic DE-ARGs signature

To evaluate the efficacy of the prognostic ARGs model, we introduced the calculated model in training cohort into the testing cohort and the entire cohort. According to the cut-off risk score calculated from the training cohort, patients from the testing cohort and entire cohort were categorized as high-risk or low-risk individuals, likewise. Kaplan–Meier curve analysis, time-dependent ROC analysis, as well as patients' survival status, survival time and gene expression were also performed.

Clinical independency estimating and nomogram building

To further measure the independency of the prognostic model, we firstly screened the EC patients who contained the detailed clinical information and integrated such remaining patients' clinical characters with the ARGs expression data. The gene-expression level and clinical variance were compared between the high-risk and low-risk group patients, and the data were comprehensively displayed in the heatmap. Meanwhile, we performed Cox multivariate analysis and compared the area under the curve (AUC) values among the risk score and the clinical index to estimate the predictive efficacy of the model. The correlation between ARGs from the risk model and the clinical index was also measured. Finally, we utilized the “rms” R package to consolidated the risk score and clinical characteristics for nomogram construction.

Results

Identification of a list of differentially expressed ARGs (DE-ARGs)

After integration and analysis of the 232 autophagy-related gene expression data from 552 cancerous tissue samples and 35 non-tumor samples from TCGA EC cohort, we obtained 51 up-regulated and 43 down-regulated ARGs with the criteria of P-value <0.05 and log|Fold change (FC)|> 0.5 (Figure 2A-B). The detailed flow chart for the prediction model construction in this study was shown in Figure 1. In addition, the expression pattern of 94 differentially expressed ARGs (DE-ARGs) was shown in a box plot (Figure 2C).

Functional enrichment of the DE-ARGs

Then, we probed into the biological function of 94 DE-ARGs by performing GO and KEGG enrichment analysis. GO enrichment showed that the DE-ARGs were mainly involved in autophagy, intrinsic apoptotic signaling pathway and cellular response to oxygen levels (Figure 3A). The expression level of correlated genes in the enriched GO terms are displayed in the heatmap (Figure 3B). KEGG pathway enrichment presented that DE-ARGs were mainly involved in apoptosis, HIF-1 signaling pathway, EGFR tyrosine kinase inhibitor resistance, p53 signaling pathway, platinum drug resistance and various cancer processes (Figure 4A). The enriched KEGG pathways and related genes expression profile were presented in Figure 4B.

The PPI network of 94 DE-ARGs and hub gene alteration analysis

Through the STRING website, we constructed an interaction network of 94 DE-ARGs (Supplementary figure 1). There were 79 nodes and 192 edges in the network based on the interaction score criteria. Top 15 hub genes with the highest connectivity degrees in the string network were selected as follows: TP53, MAPK3, EGFR, BCL2L1, GABARAPL1, BCL2, GABARAP, GABARAPL2, BID, CASP8, ITGB1, SIRT1, CASP3, HIF1A and mTOR (Figure 5A-B). The alteration results of the hub genes showed that TP53, mTOR and BCL2L1 ranked as the most frequently altered genes. TP53 more often occurred with a missense mutation (putative driver) while mTOR presented with a missense mutation (unknown significance). BCL2L1 was frequently over-amplified among endometrial cancer patients (Figure 5C).

Identification of a four DE-ARGs based prognostic model

Next, we randomly divided the 541 TCGA EC patients into the training cohort (n = 272) and the testing cohort (n = 269). Univariate cox regression analysis was performed in the training cohort and identified 5 genes among the 94 DE-ARGs closely related to patients' outcomes, which were presented in Table 1. Subsequent lasso and multivariate cox regression analysis further picked out 4 genes significantly associated with prognosis, which are ERBB2, PTEN, TP73 and ARSA (Figure 6 and Table 2). According to the results of multivariate Cox regression analysis, we constructed the prognostic model as follows: risk score = (0.001423 × expression value of ERBB2)+ (-0.14307× expression value of PTEN) + (-0.2439× expression value of TP73) + (-0.02447 × expression value of ARSA).

According to the above signature, we calculated the risk scores of each individual in the training cohort and divided them into high-risk (n=136)and low-risk (n=136) subgroups with the mean risk score as the cut-off value. To uncover the outcome discrepancy between two subgroups, we ranked each individuals' risk score and displayed their distribution and the survival status on the dot plot (Figure 7A, C). The expression profile of 4 prognostic ARGs among the high and low-risk groups was also presented via heatmap (Figure 7B). The results showed that the high-risk group seemed to accompany more death events and patients expressed higher ERBB2 level and lower expression of TP73, PTEN and ARSA. We used these four ARGs as a signature, Kaplan-Meier curve analysis showed that the OS of the higher-risk group was significantly shorter than the low-risk group (P =1.222e-05) (Figure 7D). In addition, the ROC curve analysis revealed the area under the ROC curve (AUC) of the prognostic ARGs signature was 0.743 (Figure 7E).

Validation the efficacy of the 4 ARGs prognostic signature

To evaluate the efficacy of the prognostic signature constructed in the training cohort, we introduced the risk model into the testing cohort and the entire cohort and calculated each individuals' risk scores. As for the testing cohort, we divided the group into 142 high-risk and 127 low-risk individuals based on the training cohort' cut-off risk score. The survival time, survival status and gene expression pattern were also presented based on two subgroup patients' risk scores and the results were consistent with those in the training group (Figure 8A-C). Kaplan-Meier curve analysis also showed that high-risk subgroup patients exhibited worse outcomes compared to the low-risk subgroup in the testing cohort with shorter OS(P =2.008e-03) (Figure 8D). The area under the ROC curve (AUC) of the prognostic model was 0.711 (Figure 8E).

In accordance with the results from the training and testing group, the high-risk subgroup of the entire cohort also presented shorter survival time, worse survival status and similar gene expression profile(Figure 9A-C). Kaplan-Meier curve analysis presented that the low-risk subgroup in the entire cohort followed with longer overall survival time(P =1.034e-07) (Figure 9D). ROC curve analysis showed the AUC of the prognostic model was 0.723 (Figure 9E).

The clinical independence and correlation estimation of the risk signature

To evaluate the clinical independency of the model, we combined it with other clinical factors and then performed univariate and multivariate analysis. The results indicated that the prognostic model was able to serve as an independent prognostic indicator(P<0.001 and P=0.009, respectively) (Figure 10A-B). ROC curve analysis also showed that the AUC value of the ARGs signature was 0.705, much significantly higher than patients' age (AUC= 0.535) and weight (AUC= 0.633), the tumor grade (AUC= 0.656), histology (AUC= 0.522) and lymph node status (AUC= 0.697), and slightly lower than the clinical stage (AUC= 0.710)(Figure 10C). After removing some patients without clinical information, we integrated the risk scores, the clinical characteristics and the expression profiles of remaining EC patients to analyze the relationship between ARGs prognostic model and clinical features (Figure 10D). In addition to the significant difference of survival status among two subgroups, we observed a significant increase regarding patients' age, clinical stage, tumor grade, serous histology and lymph node metastasis in high-risk group endometrial cancer patients (Figure 10E). The correlation between each gene from the prognostic model and the clinical features were also measured. The results showed that ARSA and TP73 were significantly negatively correlated with stage, tumor grade and histology whereas ERBB2 was positively correlated with tumor grade, histology and lymph node status (Figure S2).

Nomogram building and validation

To establish a clinically applicable method for predicting the prognosis of endometrial cancer patients, we established a prognostic nomogram to predict the survival probability during 5 years based on the TCGA entire set(Figure 11A). Seven independent prognostic parameters (age, stage, grade, weight, histology, lymph node status and ARGs) were enrolled in the prediction nomogram (Figure 11A). The calibration plots show excellent consistency between the nomogram prediction and actual observation in terms of the 3- and 5-year survival rates in the TCGA cohort (Figure 11B-C).

Discussion

In this study, we thoroughly probed into the implications of ARGs in EC progress. We integrated the clinical features of patients and the ARGs signature including ERBB2, TP73, PTEN and ARSA to establish a novel nomogram model, which indicated that ARGs signature could effectively predicts the prognosis of EC patients.

By analyzing the mRNA level of EC patients from TCGA data, we obtained 94 dysregulated ARGs. Functional enrichment analysis showed that such DE-ARGs were closely associated with cancer related pathways. Among the enriched disordered signaling pathways, P53, HIF-1 signaling, and EGFR tyrosine kinase inhibitor resistance pathways have already been widely studied in endometrial cancer. The frequent mutation of TP53 anti-oncogene was proved to mediate the acceleration of EC cell proliferation. Besides, both TP53 mutation and HIF1-α/VEGF pathway activation promoted the radioresistance of endometrial carcinoma [15]. The alteration analysis results of the top 15 hub genes in our study were in accordance with these findings. Shang et.al demonstrated the therapeutic effect of targeting the NOTCH pathway could improve the effects of EGFR inhibition in EC[16]. In addition, these dysregulated genes

were also enriched in a series of the cancer process, such as lung cancer, colorectal cancer, and pancreatic cancer, which suggests that the disorder of autophagy-related genes is a common event in various cancer types. Except for TP53, mTOR and BCL2L1 from the PPI network were also exhibited frequent alternated conditions and have been widely studied in other malignant diseases. The mTOR signaling pathway could control cell growth and metabolism. Dysregulated mTOR expression in cancer cells induced cell cycle disorder and metabolism inordinance, which further impacted the tumorigenicity[17]. BCL2L1, as an antiapoptotic gene, has been observed drastically amplified in various cancer, such as gastric cancer and colorectal cancer. The genomic amplification and mutation of BCL2L1 promoted the drug resistance and survival of cancer cells [18, 19].

Recently, a variety of prognostic models based on ARGs have been established in cancer types. Wang et.al developed a three ARGs risk score in glioblastoma which integrated the expression of NRG1, ITGA3 and MAP1LC3A[20]. In pancreatic adenocarcinoma (PCa), a predictive ARGs model including KRAS, CDKN2A, TP53, and SMAD4 were constructed which precisely forecasted the outcomes of PCa patients [21]. Based on TCGA non-small cell lung cancer (NSCLC) data, Liu et.al also develop a 22-gene prognostic prediction model for NSCLC patients based on the expression profiles of autophagy-associated genes. In addition, they also validated the evaluation efficacy of the model using GEO NSCLC dataset[22]. The above studied underlined the significance of ARGs in cancer progression and proved the effectiveness of combining certain ARGs to building a gene expression-based model to estimated patients' prognosis. In the present study, the autophagy-related genes and corresponding prognostic signature in EC were studied. From uni- and multivariate Cox regression and LASSO algorithm analysis, we constructed an ARGs prognostic model based on ERBB2, PTEN, TP73, and ARSA mRNA expression level. The combination of four DE-ARGs has good robustness and reproducibility in predicting prognosis of EC patients independent from traditional clinical risk factors, with the area under the ROC curve (AUC) marked 0.743, much higher than tumor histology, grade, and patients’ age.

Besides, almost all of the genes in our risk model were reported to be associated with tumorigenesis and tumor proliferation. ERBB2, also known as HER-2, encodes a member of the epidermal growth factor (EGF) receptor family. The overexpression or mutation of this gene has been reported in numerous cancers, such as breast and ovarian cancer. Overexpression of ERBB2 in breast cancer cell lines leads to excessive phosphorylation and activation of the semaphorin receptor Plexin-B1, which promoted the invasive behavior of breast cancer and led to cancer metastasis[23]. Li et.al found increased proliferation on ERBB2/ERBB3 mutation in gallbladder carcinoma (GBC) which presented poorer prognoses. The ectopic ERBB2 mutants expression activated the PI3K/Akt signaling pathway and upregulated PD-L1 expression in GBC cells, which effectively inhibited normal T-cell-mediated cytotoxicity and led to the growth and progression of GBC [24]. Early in 2006, Morrison et.al have underlined ERBB2 as a significant oncogene role in ndometrial cancer. Their results showed that the amplification of ERBB2 was significantly correlated to high-grade and stage EC [25]. Groeneweget.al found that while trastuzumab alone did not impact uterine serous carcinoma (USC) growth, the dual therapy with trastuzumab and lapatinib significantly suppressed tumor activity in ERBB2-amplified USC xenografts models [26]. In this study, we found overexpressed ERBB2 in the high-risk group, which was in accordance with previous

researches. PTEN as a tumor suppressor gene is closely involved in the maintenance of genomic stability. Function loss of PTEN is one of the most common genetic aberrations in endometrioid endometrial carcinomas. Mutter et.al found that nearly 83% and 55% PTEN mutation rate happened in endometrioid endometrial adenocarcinomas and precancers respectively while no PTEN mutation was observed in normal endometria [27]. Likewise, the germline PTEN mutations were found to be a critical cause of early-onset EC patients [28]. Interestingly, based on the discovery that the function loss of PTEN induced defects in the repair of DNA double-strand breaks by homologous recombination, Dedes et.al showed that PTEN-deficient cells exhibited significantly greater sensitivity to poly(adenosine diphosphate ribose) polymerase inhibitors (PARP-i) than wild-type PTEN EC cell lines [29]. Here, we showed frequent downregulation of PTEN happened in high-risk group patients, which uncovered the significant role of PTEN in EC prognosis and proved the efficiency of our risk model. As for TP73, Horvilleur et.al showed that p73-α, an isoform protein encoded by TP73, significantly inhibited MYCN gene and protein expression in neuroblastoma (NB), which was closely related to NB malignancy and neuronal differentiation propensity [30]. However, the opposite effect of TP73 was also observed in other human cancers. The imbalance between transcriptionally active (TAp73) or inactive (ΔNp73) isoforms was reported to contribute to tumorigenesis and resistance to chemotherapy. In acute promyelocytic leukemia (APL), Lucena-Araujo et.al presented a higher ΔNp73/TAp73 RNA expression associated with worse outcomes and a higher risk of relapse in patients [31]. While the biological role of TP73 in endometrial cancer has yet not been studied, our risk model initially categorized TP53 as a protective gene. Arylsulfatase A (ARSA) and its variants were reported closely related to metachromatic leucodystrophy (MLD) [32]. Yet, the correlation between ARSA and cancer development was rarely studied. Laidler et.al assessed the serum level of ARSA in 96 healthy volunteers and 368 patients with different cancer types. While the serum level of ARSA was unable to be a valuable tumor indicator in breast and gastrointestinal tract cancer, they observed higher expression in EC patients than the normal group[33]. However, the expression of ARSA was on serum level and only nine EC subjects were enrolled in the study, which partly explained the discrepancy result presented in our study for the downregulation of ARSA observed in high-risk group EC tissue samples.

Nomograms have been widely used in clinical practice for guiding treatment. To our knowledge, the nomogram in our study is the first to incorporate an autophagy-related signature with age, stage, tumor grade and lymph node status that effectively predicting the survival of EC patients. The high consistency of calibration plots also suggested that the excellent predictive performance of the nomogram in clinical use. This visual scoring system could provide a novel and precise method to help both gynecologists and patients in performing individualized survival estimation and facilitating the selection of better treatment options.

Conclusion

In summary, we identified 94 commonly dysregulated autophagy-related genes in EC. The most enriched biological pathways regarding DE-ARGs were apoptosis and autophagy as well as various cancer development processes. The top 15 hub genes in PPI network were also selected and analyzed. We also

performed cox and lasso regression to establish and validate a robust prognostic model combining four dysregulated ARGs. Based on the risk model, we further distinguished patients with an elevated risk of mortality independent of other clinical features. In addition, we constructed a synthetic nomogram that combined clinical characters and the risk model and demonstrated its efficacy in predicting EC patients' prognosis. The 4 ARGs model and nomogram may guide the outcome prediction and rational therapy for doctors in clinical practice.

Abbreviations

EC
endometrial cancer
ARGs
autophagy-related genes
GEO
Gene Expression Omnibus
TCGA
The Cancer Genome Atlas
GO
gene ontology
KEGG
Kyoto Encyclopedia of Genes and Genomes
PPI
protein–protein interaction
STRING
Search Tool for the Retrieval of Interacting Genes Database
ROC
Receiver operating characteristic
AUC
area under the curve
LASSO
Least Absolute Shrinkage and Selection Operator
OS
overall survival
ERBB2
erb-b2 receptor tyrosine kinase 2
ERBB3
erb-b2 receptor tyrosine kinase 3
PTEN
phosphatase and tensin homolog
TP73

tumor protein p73
ARSA
arylsulfatase A
HADb
The Human Autophagy Database
FC
fold change
TP53
tumor protein p53
MAPK3
mitogen-activated protein kinase 3
EGFR
epidermal growth factor receptor
BCL2L1
BCL2 like 1
GABARAPL1
GABA type A receptor associated protein like 1
BCL2
BCL2 apoptosis regulator
GABARAP
GABA type A receptor-associated protein
GABARAPL2
GABA type A receptor associated protein like 2
BID
BH3 interacting domain death agonist
CASP8
caspase 8
CASP3
caspase 3
ITGB1
integrin subunit beta 1
SIRT1
sirtuin 1
HIF1A
hypoxia inducible factor 1 subunit alpha
mTOR
mechanistic target of rapamycin kinase
VEGF
vascular endothelial growth factor
NRG1

neuregulin 1
MAP1LC3A
microtubule associated protein 1 light chain 3 alpha
PCa
pancreatic adenocarcinoma
NSCLC
non-small cell lung cancer
GBC
gallbladder carcinoma
USC
uterine serous carcinoma
PARP-i
poly adenosine diphosphate ribose polymerase inhibitor
NB
neuroblastoma
APL
acute promyelocytic leukemia
MLD
metachromatic leucodystrophy

Declarations

Acknowledgements

None.

Author Contributions

Shilong Fu, Shouyu Wang and Pinping Jiang designed the project. Wei Sun and Ningmei Shen contributed on data analysis and prepared the main manuscript. Qiang Wang revised and submitted the manuscript. All authors reviewed the manuscript.

Fundings

This study was supported by the National Natural Science Foundation of China (81773383, 81370078, 81903085); and the Science Foundation for Distinguished Young Scholars of Jiangsu Province (BK20170047); and the Fundamental Research Funds for the Central Universities (021414380439); and the Project funded by China Postdoctoral Science Foundation (2019M651808).

Availability of data and materials

The expression data were deposited in the TCGA database and the clinical information was retrieved from cBioPortal website. Besides, please contact the author for data and materials requests.

Ethics approval and consent to participate

Not required.

Consent for publication

That the article is original, has not already been published in a journal, and is not currently under consideration by another journal.

Competing interests

All financial and non-financial competing interests must be declared in this section.

References

1. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2019.** *CA Cancer J Clin* 2019, **69:**7-34.
2. Koh WJ, Abu-Rustum NR, Bean S, Bradley K, Campos SM, Cho KR, Chon HS, Chu C, Clark R, Cohn D, et al: **Cervical Cancer, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology.** *J Natl Compr Canc Netw* 2019, **17:**64-84.
3. Brooks RA, Fleming GF, Lastra RR, Lee NK, Moroney JW, Son CH, Tatebe K, Veneris JL: **Current recommendations and recent progress in endometrial cancer.** *CA Cancer J Clin* 2019, **69:**258-279.
4. Amaravadi RK, Kimmelman AC, Debnath J: **Targeting Autophagy in Cancer: Recent Advances and Future Directions.** *Cancer Discov* 2019, **9:**1167-1181.
5. Rabinowitz JD, White E: **Autophagy and metabolism.** *Science* 2010, **330:**1344-1348.
6. White E: **The role for autophagy in cancer.** *J Clin Invest* 2015, **125:**42-46.
7. White E, Karp C, Strohecker AM, Guo Y, Mathew R: **Role of autophagy in suppression of inflammation and cancer.** *Curr Opin Cell Biol* 2010, **22:**212-217.
8. Guo JY, Chen HY, Mathew R, Fan J, Strohecker AM, Karsli-Uzunbas G, Kamphorst JJ, Chen G, Lemons JM, Karantza V, et al: **Activated Ras requires autophagy to maintain oxidative metabolism and tumorigenesis.** *Genes Dev* 2011, **25:**460-470.
9. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al: **The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.** *Cancer Discov* 2012, **2:**401-404.
10. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: **limma powers differential expression analyses for RNA-sequencing and microarray studies.** *Nucleic Acids Res* 2015, **43:**e47.
11. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25:**25-29.
12. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K: **KEGG: new perspectives on genomes, pathways, diseases and drugs.** *Nucleic Acids Res* 2017, **45:**D353-D361.

13. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al: **STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets.** *Nucleic Acids Res* 2019, **47:**D607-D613.
14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13:**2498-2504.
15. Miyasaka A, Oda K, Ikeda Y, Sone K, Fukuda T, Inaba K, Makii C, Enomoto A, Hosoya N, Tanikawa M, et al: **PI3K/mTOR pathway inhibition overcomes radioresistance via suppression of the HIF1-alpha/VEGF pathway in endometrial cancer.** *Gynecol Oncol* 2015, **138:**174-180.
16. Shang C, Lang B, Meng LR: **Blocking NOTCH pathway can enhance the effect of EGFR inhibitor through targeting CD133+ endometrial cancer cells.** *Cancer Biol Ther* 2018, **19:**113-119.
17. Mossmann D, Park S, Hall MN: **mTOR signalling and cellular metabolism are mutual determinants in cancer.** *Nat Rev Cancer* 2018, **18:**744-757.
18. Park H, Cho SY, Kim H, Na D, Han JY, Chae J, Park C, Park OK, Min S, Kang J, et al: **Genomic alterations in BCL2L1 and DLC1 contribute to drug sensitivity in gastric cancer.** *Proc Natl Acad Sci U S A* 2015, **112:**12492-12497.
19. Cho SY, Han JY, Na D, Kang W, Lee A, Kim J, Lee J, Min S, Kang J, Chae J, et al: **A Novel Combination Treatment Targeting BCL-XL and MCL1 for KRAS/BRAF-mutated and BCL2L1-amplified Colorectal Cancers.** *Mol Cancer Ther* 2017, **16:**2178-2190.
20. Wang Z, Gao L, Guo X, Feng C, Lian W, Deng K, Xing B: **Development and validation of a nomogram with an autophagy-related gene signature for predicting survival in patients with glioblastoma.** *Aging (Albany NY)* 2019, **11:**12246-12269.
21. Yue P, Zhu C, Gao Y, Li Y, Wang Q, Zhang K, Gao S, Shi Y, Wu Y, Wang B, et al: **Development of an autophagy-related signature in pancreatic adenocarcinoma.** *Biomed Pharmacother* 2020, **126:**110080.
22. Liu Y, Wu L, Ao H, Zhao M, Leng X, Liu M, Ma J, Zhu J: **Prognostic implications of autophagy-associated gene signatures in non-small cell lung cancer.** *Aging (Albany NY)* 2019, **11:**11440-11462.
23. Worzfeld T, Swiercz JM, Looso M, Straub BK, Sivaraj KK, Offermanns S: **ErbB-2 signals through Plexin-B1 to promote breast cancer metastasis.** *J Clin Invest* 2012, **122:**1296-1305.
24. Li M, Liu F, Zhang F, Zhou W, Jiang X, Yang Y, Qu K, Wang Y, Ma Q, Wang T, et al: **Genomic ERBB2/ERBB3 mutations promote PD-L1-mediated immune escape in gallbladder cancer: a whole-exome sequencing analysis.** *Gut* 2019, **68:**1024-1033.
25. Morrison C, Zanagnolo V, Ramirez N, Cohn DE, Kelbick N, Copeland L, Maxwell GL, Fowler JM: **HER-2 is an independent prognostic factor in endometrial cancer: association with outcome in a large cohort of surgically staged patients.** *J Clin Oncol* 2006, **24:**2376-2385.
26. Groeneweg JW, Hernandez SF, Byron VF, DiGloria CM, Lopez H, Scialabba V, Kim M, Zhang L, Borger DR, Tambouret R, et al: **Dual HER2 targeting impedes growth of HER2 gene-amplified uterine serous**

carcinoma xenografts. *Clin Cancer Res* 2014, **20:**6517-6528.

27. Mutter GL, Lin MC, Fitzgerald JT, Kum JB, Baak JP, Lees JA, Weng LP, Eng C: **Altered PTEN expression as a diagnostic marker for the earliest endometrial precancers.** *J Natl Cancer Inst* 2000, **92:**924-930.
28. Blumenthal GM, Dennis PA: **Germline PTEN mutations as a cause of early-onset endometrial cancer.** *J Clin Oncol* 2008, **26:**2234; author reply 2234.
29. Dedes KJ, Wetterskog D, Mendes-Pereira AM, Natrajan R, Lambros MB, Geyer FC, Vatcheva R, Savage K, Mackay A, Lord CJ, et al: **PTEN deficiency in endometrioid endometrial adenocarcinomas predicts sensitivity to PARP inhibitors.** *Sci Transl Med* 2010, **2:**53ra75.
30. Horvilleur E, Bauer M, Goldschneider D, Mergui X, de la Motte A, Benard J, Douc-Rasy S, Cappellen D: **p73alpha isoforms drive opposite transcriptional and post-transcriptional regulation of MYCN expression in neuroblastoma cells.** *Nucleic Acids Res* 2008, **36:**4222-4232.
31. Lucena-Araujo AR, Kim HT, Thome C, Jacomo RH, Melo RA, Bittencourt R, Pasquini R, Pagnano K, Gloria AB, Chauffaille Mde L, et al: **High DeltaNp73/TAp73 ratio is associated with poor prognosis in acute promyelocytic leukemia.** *Blood* 2015, **126:**2302-2306.
32. Hettiarachchi D, Dissanayake VHW: **Three novel variants in the arylsulfatase A (ARSA) gene in patients with metachromatic leukodystrophy (MLD).** *BMC Res Notes* 2019, **12:**726.
33. Laidler P, Kowalski D, Silberring J: **Arylsulfatase A in serum from patients with cancer of various organs.** *Clin Chim Acta* 1991, **204:**69-77.

Tables

Table1 Univariate Cox regression identified 5 ARGs correlated to endometrial cancer patients' OS

Gene ID	HR	HR.95L	HR.95H	P-value
CDKN2A	1.026531627	1.009917387	1.043419189	0.00165897
ERBB2	1.001501287	1.00049562	1.002507965	0.003426745
TP73	0.755473706	0.607575446	0.939373908	0.011649278
ARSA	0.971274216	0.94540475	0.997851557	0.034335042
PTEN	0.902900646	0.816729251	0.998163804	0.045947383

Abbreviation: ARGs: Autophagy related genes; HR: Hazard Ratio, OS: overall survival time.

Table2 Multivariate cox regression selected 4 ARGs correlated to endometrial cancer patients' OS

Gene ID	HR	HR.95L	HR.95H	P-value
ERBB2	1.001423915	1.000390213	1.002458685	0.00692645
PTEN	0.866697329	0.778078128	0.965409814	0.009332535
TP73	0.783565307	0.628160539	0.977416683	0.030580492
ARSA	0.975830053	0.950211196	1.002139627	0.071466256

Abbreviation: ARGs: Autophagy related genes; HR: Hazard Ratio, OS: overall survival time.

Supplemental Figure Legends

Supplementary Figure 1. The protein-protein network of 94 DE-ARGs.

Supplementary Figure 2. Clinical characteristics of ERBB2, ARSA, and TP73.

Figures

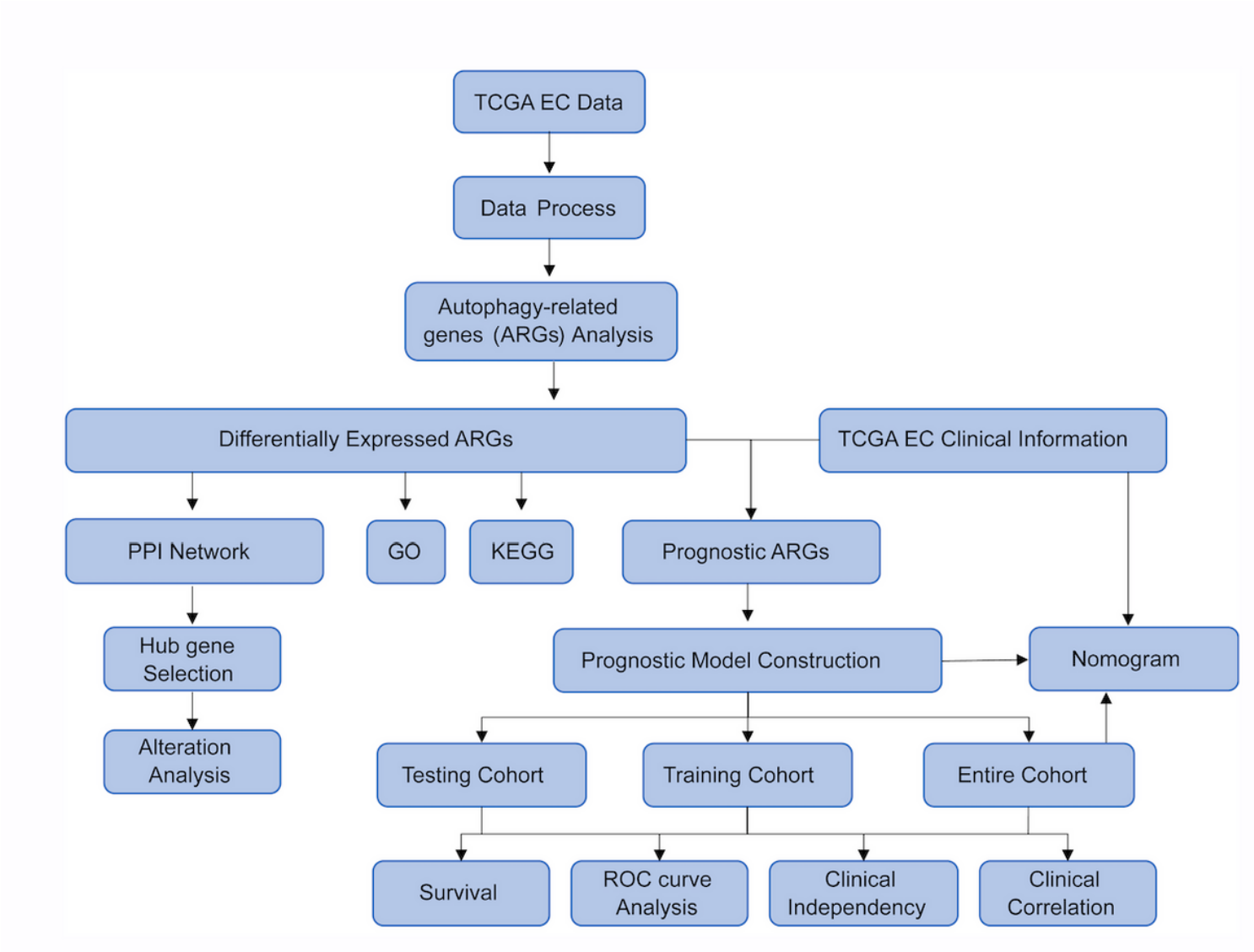


Figure 1

The flow chart of the analysis procedure in identifying autophagy-related prognostic signature.

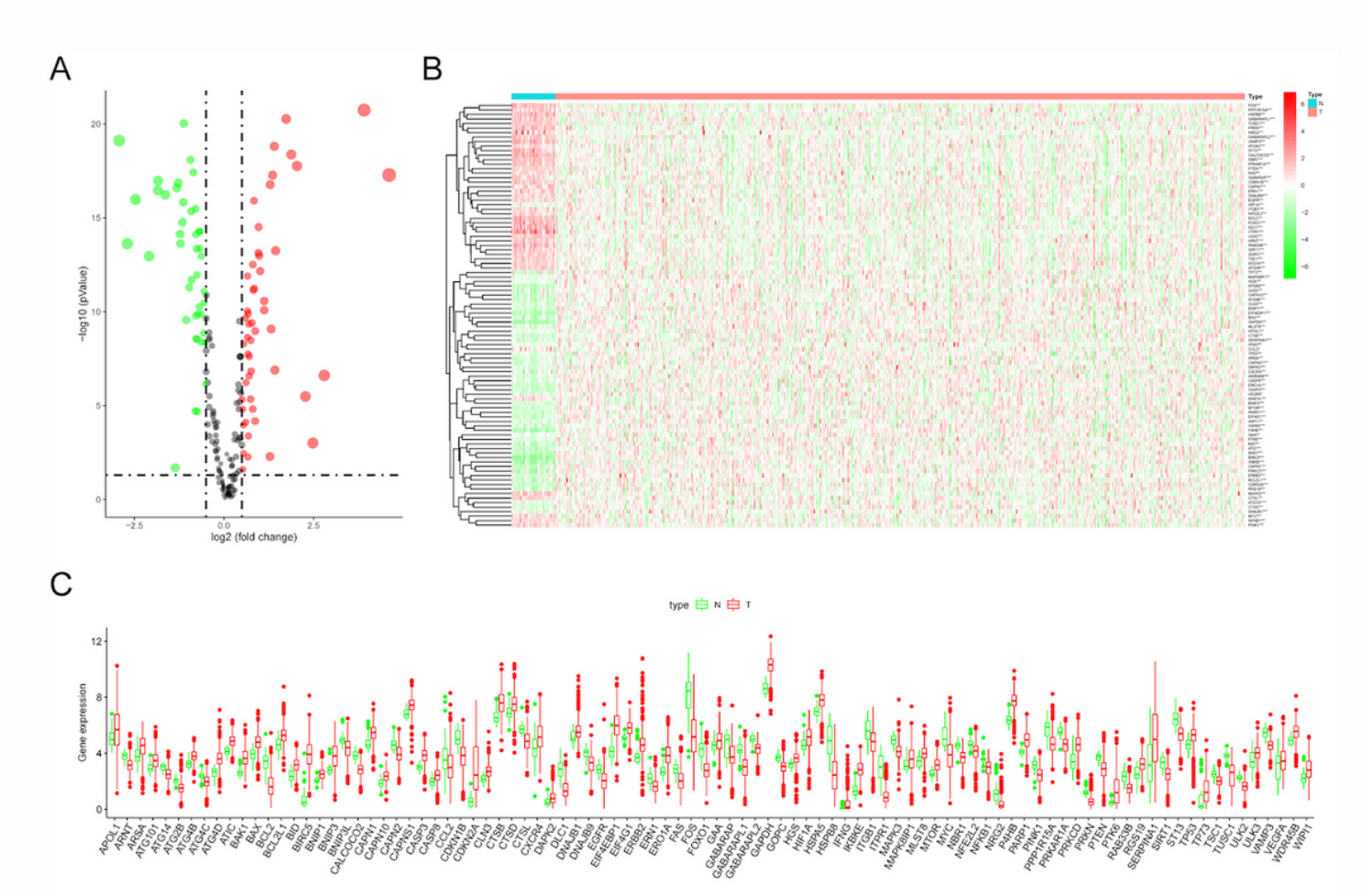


Figure 2

The expression profiles of 94 differentially expressed autophagy-related genes (DE-ARGs) between TCGA endometrial cancer (EC) and normal tissues. (A) Volcano plot of DE-ARGs in EC and normal samples of the TCGA dataset. The vertical axis indicates the -log(P-value), and the horizontal axis indicates the log2 (fold change [FC]). The red dots and the green dots represent up- and down-regulated genes,respectively (P-value <0.05 and |log2(FC)|>0.5). (B) Heat map of the 94 DE-ARGs in the entire TCGA EC cohort. Red and green indicate higher expression and lower expression, respectively. (C) Box plot of the expression of the DEGs between cancerous and normal tissues.

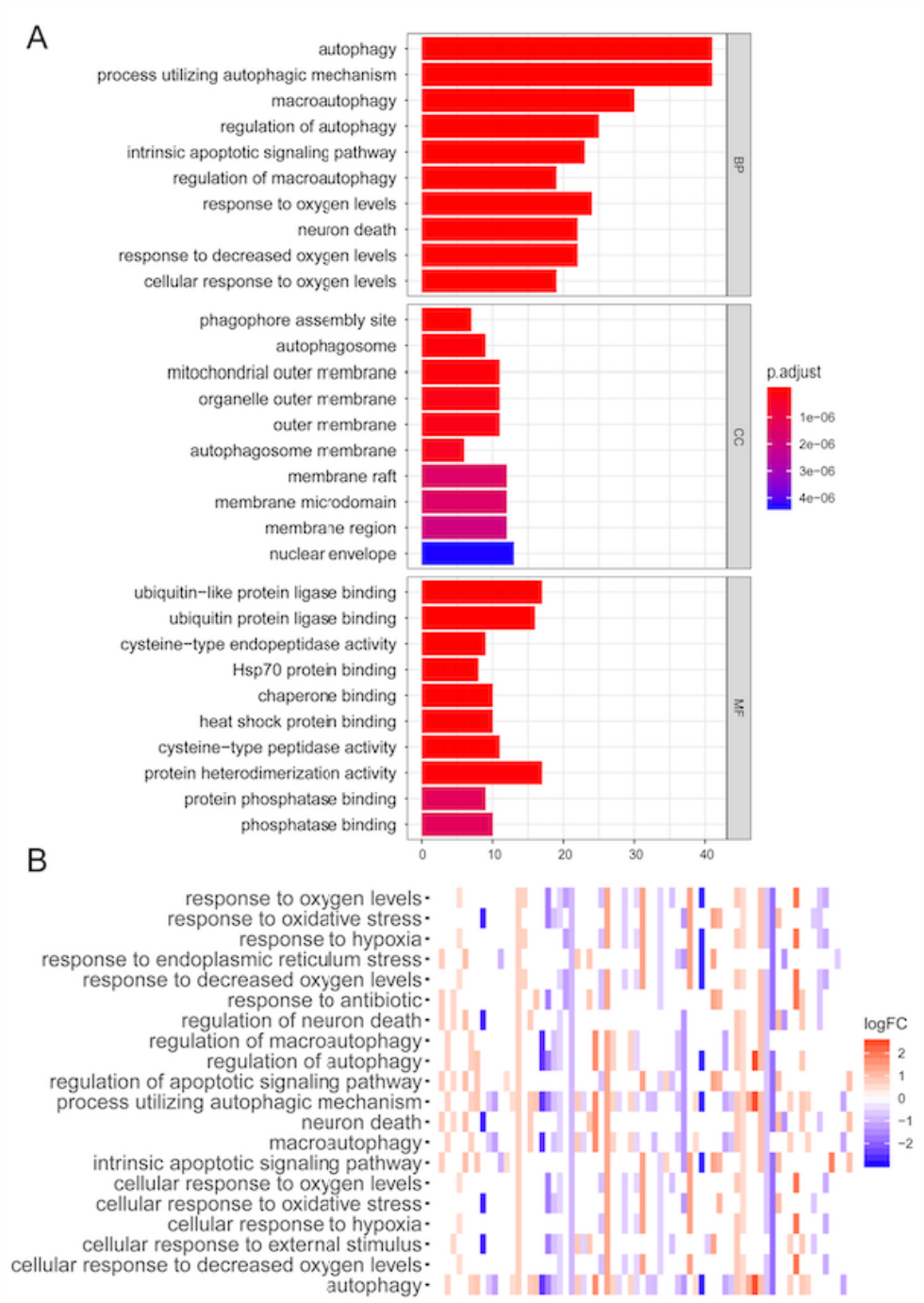
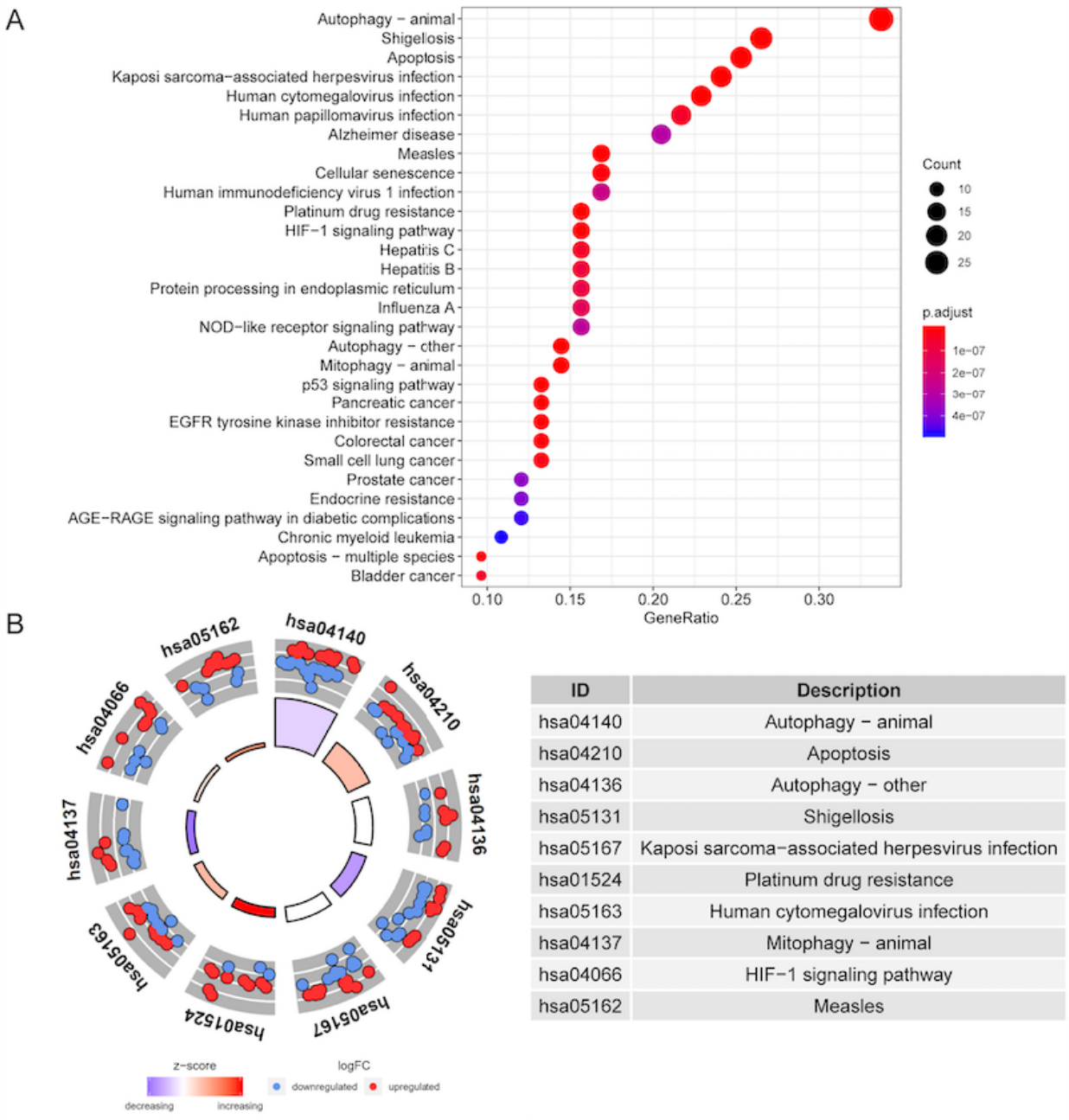


Figure 3

Gene ontology (GO) functional enrichment of differentially expressed ARGs. (A) GO analysis shows the biological processes, cellular component and molecular functions involved in differential genes. (B) Heatmap of the expression of DEGs in the enriched GO items.



ID	Description
hsa04140	Autophagy − animal
hsa04210	Apoptosis
hsa04136	Autophagy − other
hsa05131	Shigellosis
hsa05167	Kaposi sarcoma−associated herpesvirus infection
hsa01524	Platinum drug resistance
hsa05163	Human cytomegalovirus infection
hsa04137	Mitophagy − animal
hsa04066	HIF−1 signaling pathway
hsa05162	Measles

Figure 4

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment of differentially expressed ARGs. (A) KEGG analysis shows significantly enriched pathways of DE-ARGs. The node color changes gradually from red to blue in descending order according to the adjusted P values. The size of the node represents the number of counts. (B) Circle plot of the enriched DEGs in the KEGG items.

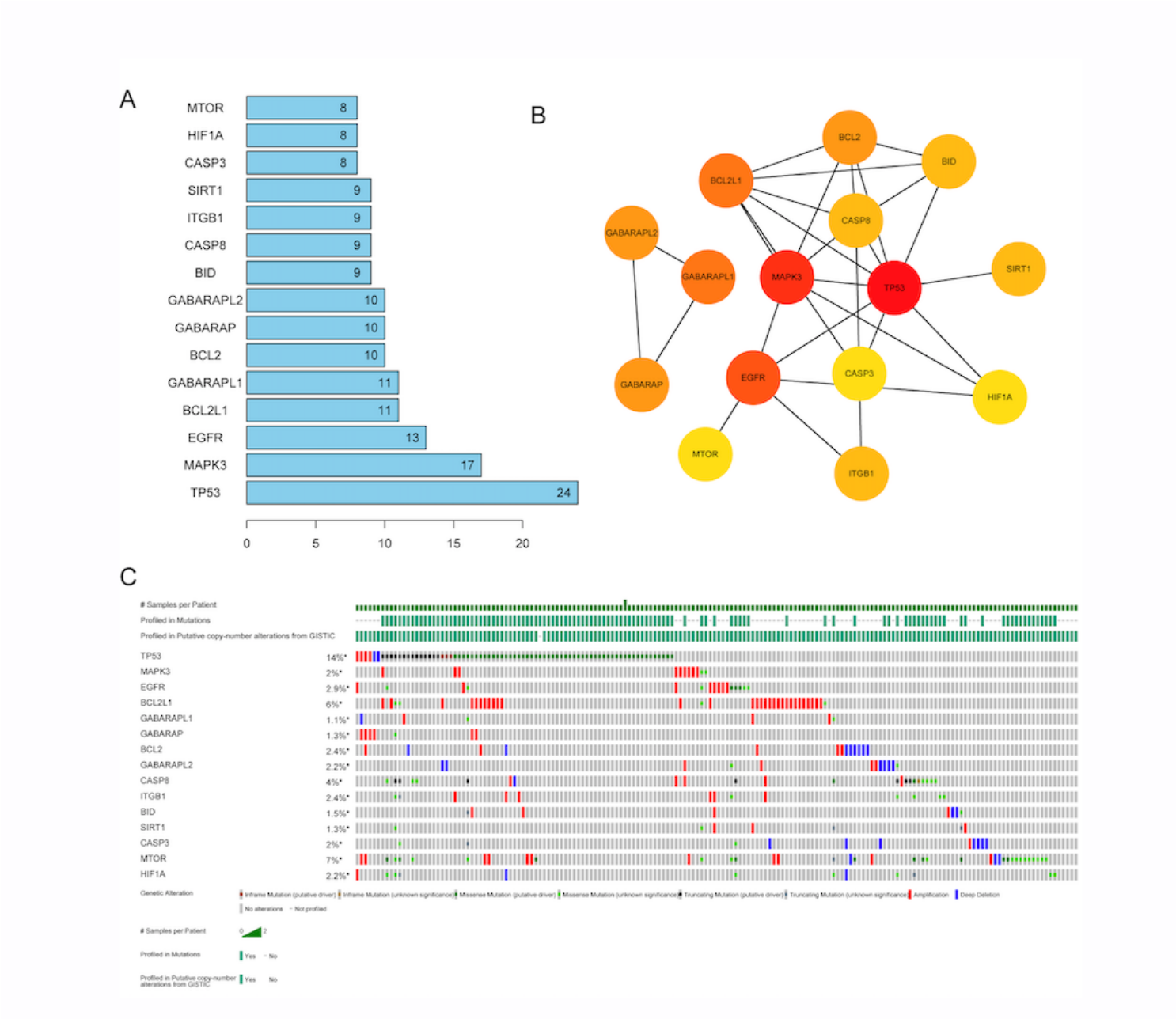


Figure 5

Protein-protein network of hub DE-ARGs and alteration analysis. (A) The highest degree of hub genes was ranked; (B) The interaction network of the top 15 hub genes; (C) The gene mutation overview of 15 hub genes in TCGA EC patients.

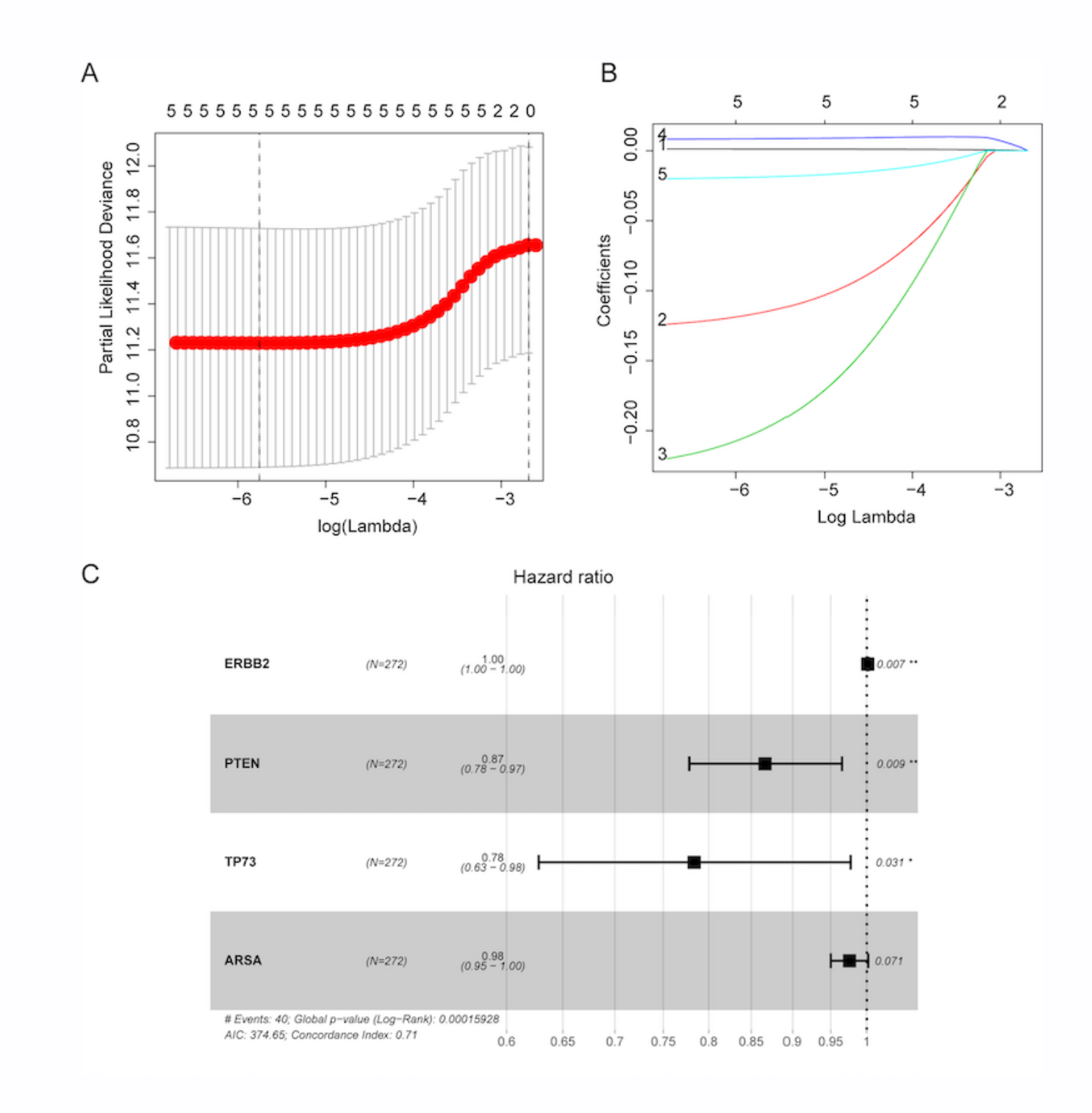


Figure 6

Identification of prognosis related ARGs using LASSO and cox regression analysis. (A) Plots of the cross-validation error rates. Each dot represents a lambda value along with error bars to give a confidence interval for the cross-validated error rate; (B) LASSO coefficient profiles of the ARGs associated with the overall survival of endometrial cancer; (C) multivariate cox regression identified 4 prognostic ARGs in the training cohort.

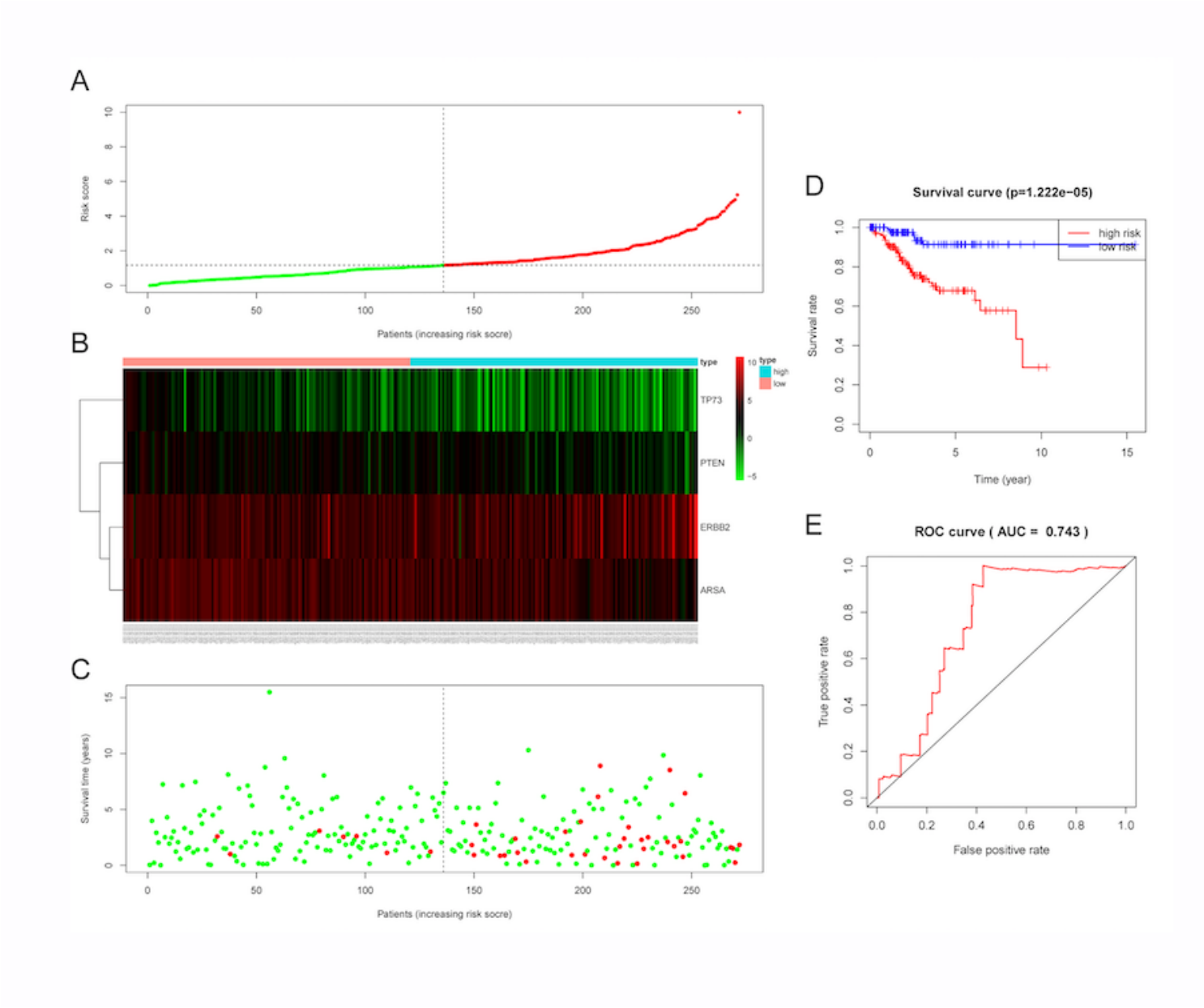


Figure 7

Prognostic analysis of the model in the TCGA training cohort. (A) The risk score, (B) expression heatmap, (C)survival status, (D) Kaplan-Meier survival and (E) time-dependent ROC curves of the prognostic model for the TCGA EC training cohort.

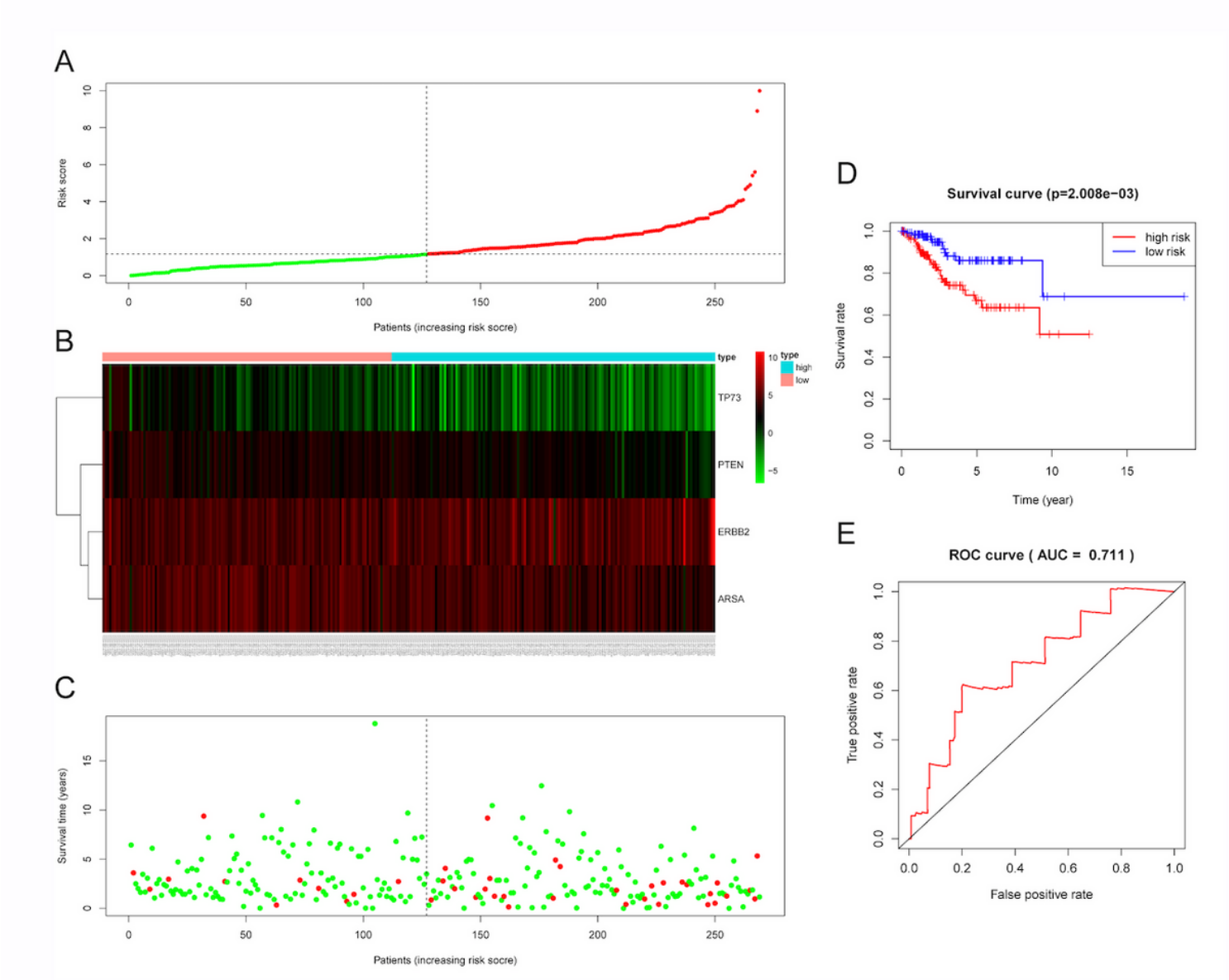


Figure 8

Validation of the efficacy of the risk signature in the TCGA testing cohort. (A) The risk score, (B) expression heatmap, (C)survival status, (D) Kaplan-Meier survival and (E) time-dependent ROC curves of the prognostic model for the TCGA EC testing cohort.

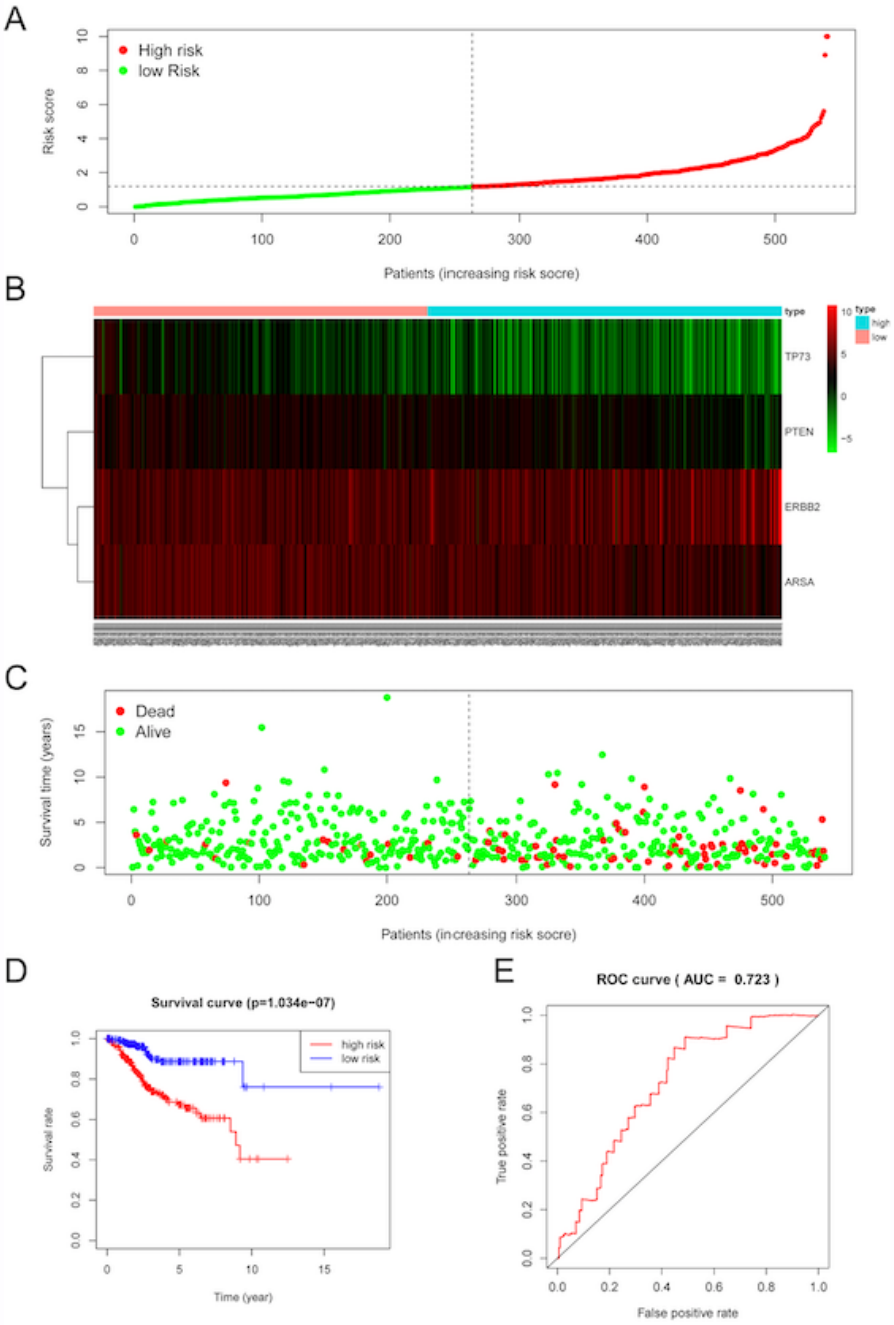


Figure 9

Estimation of the efficacy of the risk signature in TCGA entire EC cohort. (A) The risk score, (B) expression heatmap, (C)survival status, (D) Kaplan-Meier survival and (E) time-dependent ROC curves of the prognostic model for the TCGA EC entire cohort.

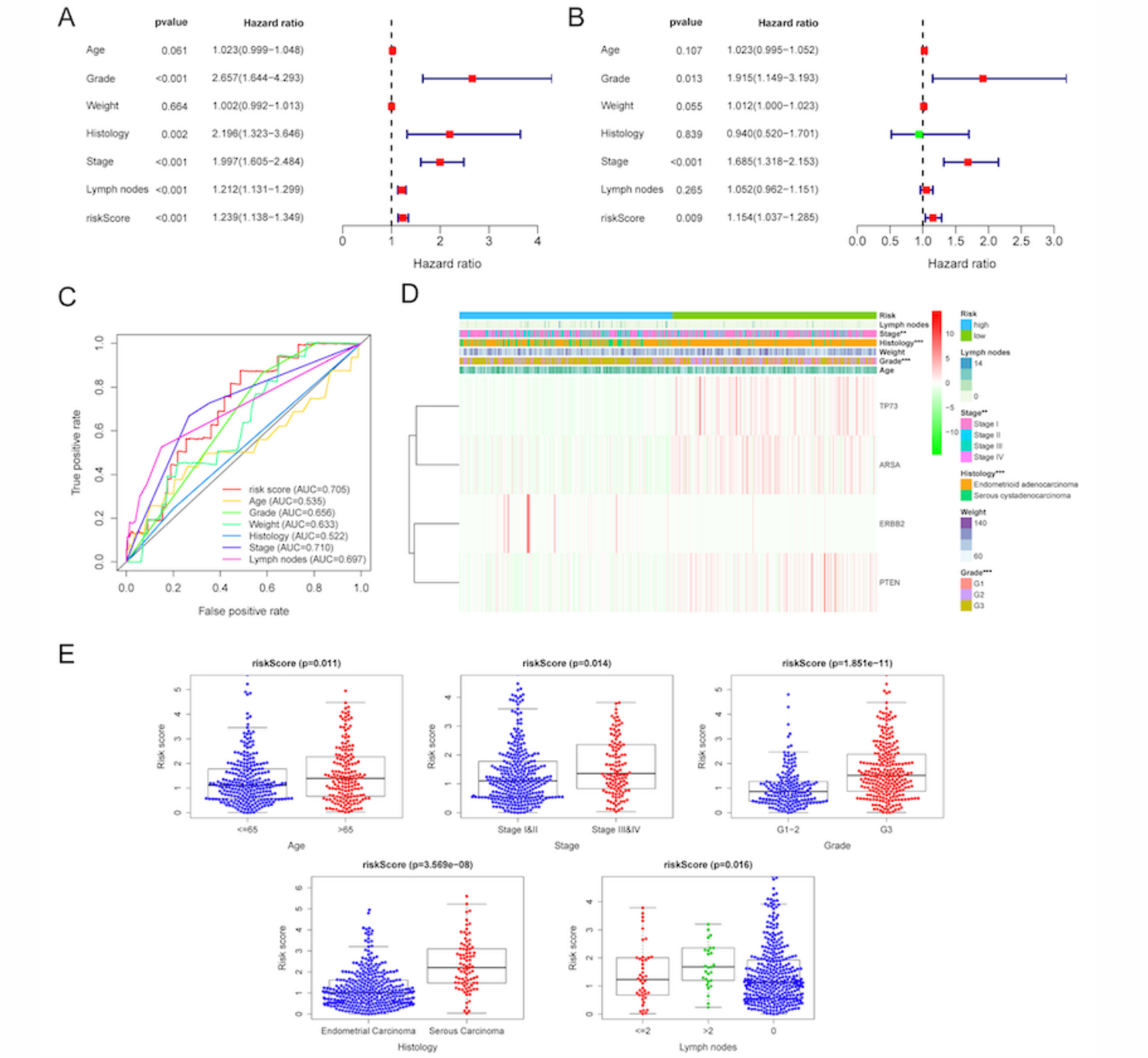


Figure 10

Clinical characteristics of the prognostic ARGs signature. Univariate (A) and multivariate (B) regression analysis, as well as time-dependent ROC curve analysis (C) of the prognostic value between the risk model and EC patients' OS status when compared to or combined with clinical factors; (D) Heat map showing the expression of 4 ARGs in the risk model and the clinicopathological features of patients with EC; (E) Clinicopathological significance of the prognostic signature of endometrial cancer.

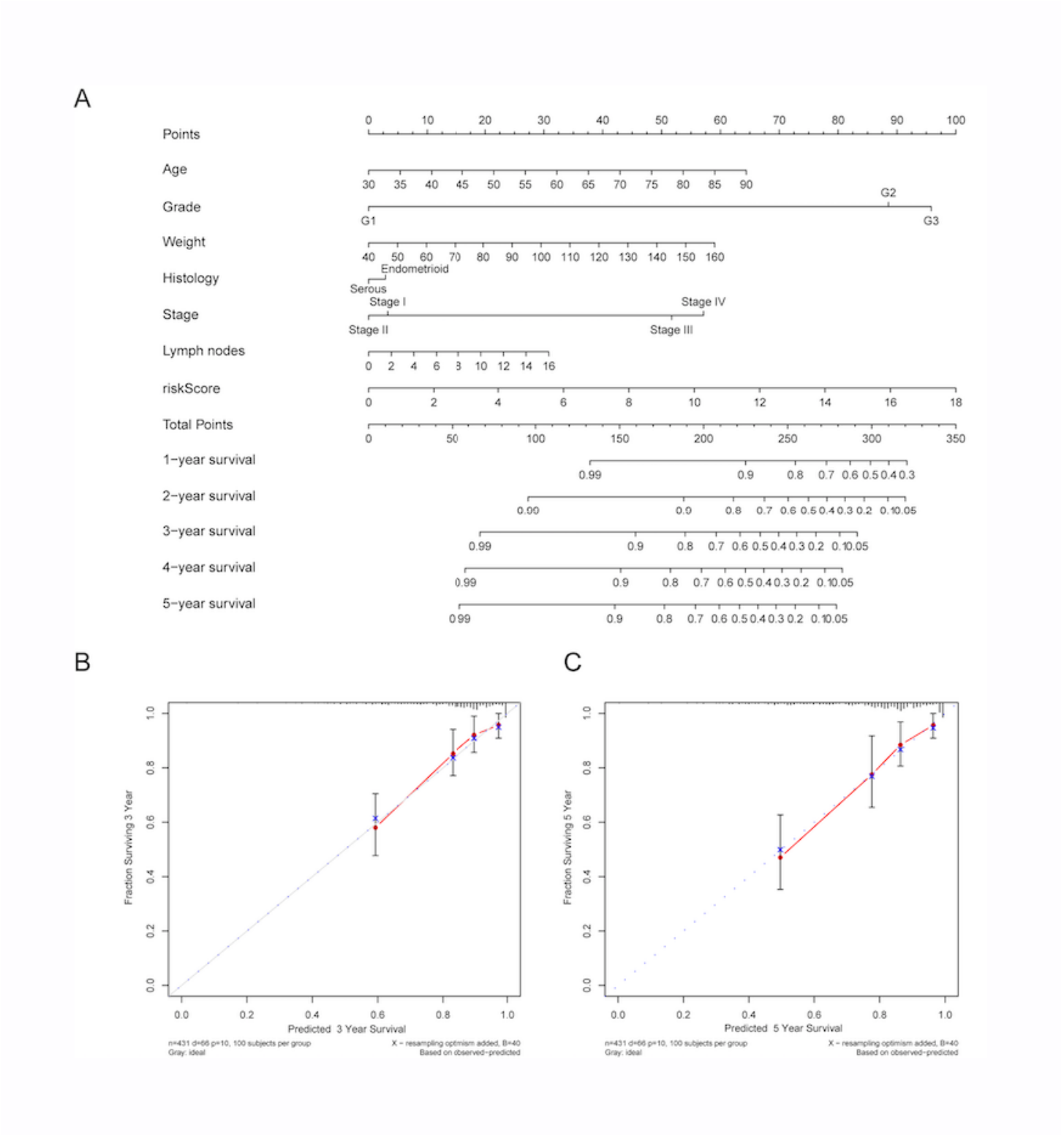


Figure 11

Nomogram for predicting the 1-,2-, 3-,4- and 5-year survival probability of patients with EC. (A) Prognostic nomogram for EC patients; (B) Calibration curves for the nomogram at 3-, and 5-year.

Supplementary Files

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

- FigureS1.png
- FigureS2.tif