



Development and validation of an autophagy-related prognostic signature in esophageal cancer

Hailei Du[#], Shanshan Xie[#], Wei Guo, Jiaming Che, Lianggang Zhu, Junbiao Hang, Hecheng Li

Department of Thoracic Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Contributions: (I) Conception and design: H Li, H Du; (II) Administrative support: H Li; (III) Provision of study materials or patients: S Xie; (IV) Collection and assembly of data: W Guo; (V) Data analysis and interpretation: H Du, S Xie; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Professor Hecheng Li. Department of Thoracic Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, 197 Ruijin Er Road, Shanghai 200025, China. Email: lihecheng2000@hotmail.com.

Background: Autophagy has a dual function in cancer, and its role in carcinogenesis of the esophagus remains poorly understood. In the present study, we explored the prognostic value of autophagy in esophageal cancer (ESCA), one of the leading causes of cancer-related deaths worldwide.

Methods: Using ESCA RNA-sequencing (RNA-Seq) data from 158 primary patients with ESCA, including esophageal adenocarcinoma and esophageal squamous cell carcinoma, were downloaded from The Cancer Genome Atlas (TCGA) for this study. We obtained differentially expressed autophagy-related genes (ARGs) by the “limma” package of R. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genome (KEGG) analyses unveiled several fundamental signaling pathways associated with the differentially expressed ARGs in ESCA. Univariate Cox regression analyses were used to estimate associations between ARGs and overall survival (OS) in the TCGA ESCA cohort. A Cox proportional hazards model (iteration =1,000) with a lasso penalty was used to create the optimal multiple-gene prognostic signature utilizing an R package called “glmnet”.

Results: A prognostic signature was constructed with four ARGs (*DNAJ7B1*, *BNIP1*, *VAMP7* and *TBK1*) in the training set, which significantly divided ESCA patients into high- and low-risk groups in terms of OS [hazard ratio (HR) =1.508, 95% confidence interval (CI): 1.201–1.894, P<0.001]. In the testing set, the risk score remained an independent prognostic factor in the multivariate analyses (HR =1.572, 95% CI: 1.096–2.257, P=0.014). The area under the curve (AUC) of the receiver operating characteristic (ROC) predicting 1-year survival showed a better predictive power for the prediction model. The AUC in training and testing cohorts were 0.746 and 0.691, respectively. Therefore, the prognostic signature of the four ARGs was successfully validated in the independent cohort.

Conclusions: The prognostic signature may be an independent predictor of survival for ESCA patients. The prognostic nomogram may improve the prediction of individualized outcome. This study also highlights the importance of autophagy in the outcomes of patients with ESCA.

Keywords: Autophagy-related genes (ARGs); prognostic signature; esophageal cancer; The Cancer Genome Atlas (TCGA)

Submitted Jun 09, 2020. Accepted for publication Nov 15, 2020.

doi: 10.21037/atm-20-4541

View this article at: <http://dx.doi.org/10.21037/atm-20-4541>

Introduction

Esophageal cancer (ESCA) ranks as the sixth most common cause of cancer deaths globally (1). Despite the improved diagnostic and treatment procedures, including endoscopic resection, targeted therapy, and immunotherapy, the overall survival (OS) of ESCA patients is still poor. The 5-year survival rate for patients with advanced ESCA is less than 20% (2). Improving early diagnosis and identifying new prognostic markers and therapeutic targets are important in improving the prognosis of patients with ESCA.

Autophagy, known as type II programmed cell death, plays an important role in maintaining cell homeostasis. Abnormal autophagic levels are involved in the carcinogenesis of various organs, including ESCA (3,4). Previous studies indicated that autophagy played a dual role in cancer development, depending on stage and type of cancers (5). Studying the relationship between autophagy and tumorigenesis may facilitate the discovery of new targets for prognosis and treatment.

Various prognostic models based on gene expression data, such as long non-coding RNA (lncRNA) and microRNA (miRNA) expression, have been developed for different cancers (6,7); however, prognostic models based on autophagy-related genes (ARGs) in ESCA have not been reported. In the present study, we examined the associations between expression profiles of ARGs and clinical outcomes in ESCA patients using the database of The Cancer Genome Atlas (TCGA). We developed an ARG prognostic signature that was proved to be an independent predictor of OS in ESCA patients. Our findings provide an effective multi-dimensional biomarker strategy that predicts the prognosis of ESCA patients. We present the following article in accordance with the TRIPOD reporting checklist (available at <http://dx.doi.org/10.21037/atm-20-4541>).

Methods

The study conformed to the provisions of the Declaration of Helsinki (as revised in 2013).

Data extraction from the TCGA database

The transcriptome data and clinical information for ESCA patients were obtained from TCGA (<https://portal.gdc.cancer.gov/>). RNA-sequencing (RNA-Seq) was performed using HTSeq-FPKM (version: July 19, 2019), which is

a simple expression level normalization method. A total of 158 primary ESCA patients with gene expression data and clinical follow-up information were included in the current study. The ARGs were acquired from The Human Autophagy Database (HADb, <http://www.autophagy.lu/index.html>).

Differentially expressed ARGs and functional pathways analysis

The “Limma” package in R software was employed to identify the differentially expressed ARGs between tumor samples and non-tumor samples from patients with ESCA. The cut-offs were $|\log\text{Fold change (FC)}| > 0.5$ and P value < 0.05 . The “org.Hs.eg.db” package was utilized to extract the Entrez ID of each ARG. The “clusterProfiler”, “enrichplot” and “ggplot2” packages were applied for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEEG) analyses as well as for visualization of the enrichment terms.

Construction and validation of an ARG-related prognostic signature

The ESCA patients in the entire TCGA dataset were randomly separated into a training dataset and a testing dataset by the “caret” package of R. The clinical information related to the TCGA-ESCA cohort was downloaded, and the OS was defined as the time from diagnosis to death or the last follow-up date. ARGs that were significantly associated with OS were selected using a univariate Cox regression analysis. To identify ARGs that independently predicted OS in the training dataset, a multivariate Cox regression analysis was performed on the ARGs selected in the univariate regression analysis. A prognostic signature was developed based on these ARGs. A risk score was established based on the prognostic ARGs signature using a linear combination of the gene expression levels weighted by the regression coefficients derived from the Cox regression analysis. This model was utilized to estimate the survival of each patient in the training dataset and testing dataset. The patients in each dataset were assigned to a high-risk or a low-risk group using the median risk score as a cutoff. The predictive potential of the ARGs signature was evaluated via area under the curve (AUC) value of the receiver operating characteristic (ROC) curve at 1-year using the “survivalROC” package.

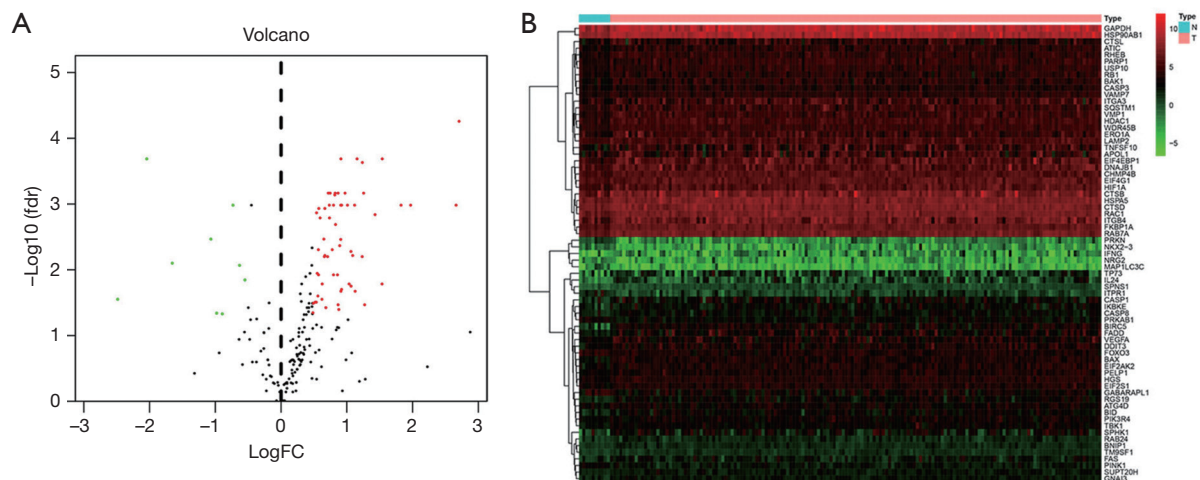


Figure 1 Differentially expressed ARGs in ESCA tumor and non-tumor samples. (A) The volcano map of 69 ARGs. The red dots indicate genes with high expression and the green dots represent genes with low expression. (B) Hierarchical clustering distribution of differentially expressed ARGs in normal and tumor samples. ARGs, autophagy-related genes; ESCA, esophageal cancer; N, normal samples; T, tumor samples; FC, fold change.

PPI network construction and module screening

The differentially expressed ARGs were submitted to the STRING database (Search Tool for the Retrieval of Interacting Genes/Proteins, <http://www.string-db.org/>, version 11.0) to identify protein-protein interaction (PPI) information. The Cytoscape 3.7.2 software was applied to further construct and visualize the PPI network. Important modules and genes were selected in the PPI network by applying the Molecular Complex Detection (MCODE) plug-in with node counts higher than 10.

Statistical analysis

A Cox regression analysis was conducted using the “survival” package. The “Limma” package was utilized to conduct the normalization and differential expression analyses. Wilcoxon rank-sum test, a non-parametric statistical hypothesis test, was used for comparing two groups. All statistical analyses were implemented using the R software (version 3.6.3). A P value <0.05 was considered to be significant.

Results

Differentially expressed ARGs in ESCA

The ESCA RNA-seq data download from the TCGA database included 159 tumor samples and 10 non-tumor samples from the patients. Among these patients, a total of

158 primary ESCA patients with gene expression data and clinical follow-up information were included in the current study. We analyzed the expression of 232 ARGs in ESCA tumor samples and non-tumor tissues using the Wilcoxon signed rank test in R. Using the criteria of $|\log_2 FC| > 0.5$ and $P < 0.05$, 69 ARGs were extracted, including 9 down-regulated genes and 60 up-regulated genes. The expression patterns of these differentially expressed ARGs were visualized using the “limma” package of R (Figure 1A,B). A box plot was generated to display the expression patterns of the 69 ARGs in ESCA and non-tumor tissues (Figure 2).

Functional enrichment analysis of the differentially expressed ARGs

The GO annotations for the differentially expressed ARGs were determined. The most enriched GO terms for biological processes were autophagy, process utilizing autophagic mechanism, and neuron death. The cellular component analysis showed enrichment of vacuolar membrane, endosome membrane, and autophagosome in the differentially expressed ARGs. In the molecular function analysis, ubiquitin-like protein ligase binding, ubiquitin protein ligase binding, and cytokine receptor binding (Figure 3A,B) were enriched in the differentially expressed ARGs. The results of the KEGG pathway analysis showed enrichment of the apoptosis, autophagy-animal, and Kaposi sarcoma-associated herpes virus infection pathways (Figure 3C,D).

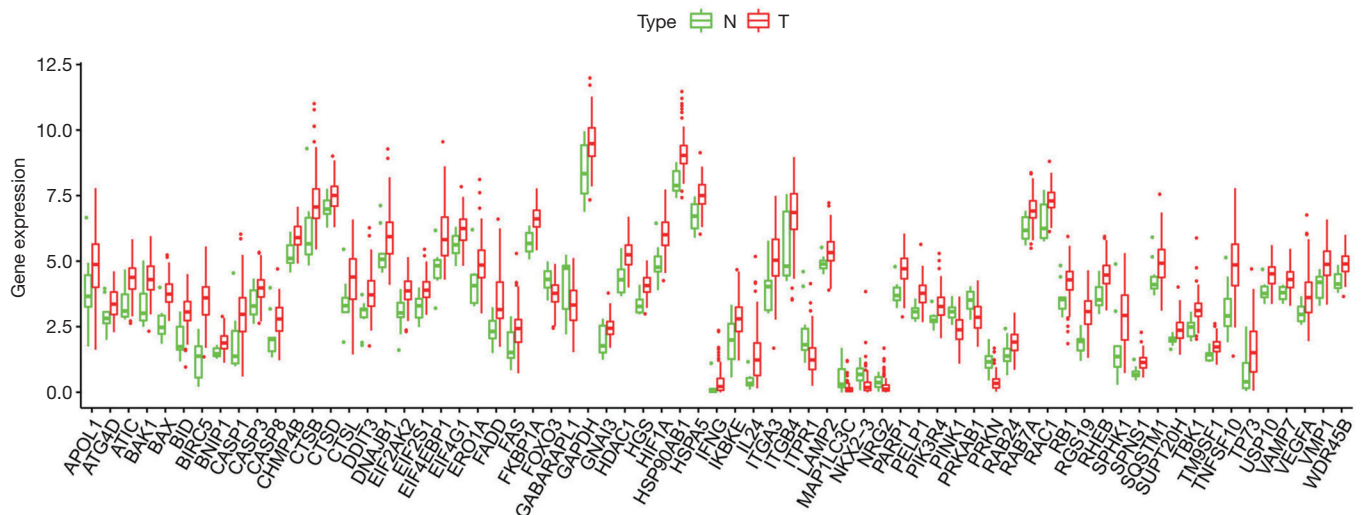


Figure 2 The expression patterns of 69 ARGs in ESCA tumor samples and paired non-tumor samples. Each red dot represents a distinct tumor sample and greens dots represent a non-tumor sample. A red bar above a gene name represents a significantly higher expression and a green bar represents a significantly lower expression. ARGs, autophagy-related genes; ESCA, esophageal cancer; N, normal samples; T, tumor samples.

More importantly, the functional analysis showed that the enriched genes were up-regulated in the most significant KEGG pathways (e.g., apoptosis and autophagy-animal).

PPIs among differentially expressed ARGs

To understand the interaction among the differentially expressed ARGs, a PPI network was constructed using STRING (Figure 4A). The PPI network was composed of three modules, including 670 nodes and 335 edges. For the ease of description, we named these modules as CASP3, VEGFA, and SQSTM1 modules. In the CASP3 module (Figure 4B), 26 edges involving 11 nodes were formed, of which the most remarkable nodes were CASP3, GAPDH, CASP8, EIF4G1, HSPA5, and DDIT3. In the VEGFA module (Figure 4C), apoptosis and necroptosis-related genes were assigned to the center of the module, including CASP1, FAS, FADD, and TNFSF10. The SQSTM1 module had the most nodes associated with autophagy and ubiquitin-like proteins (Figure 4D).

Construction of ARG prognostic signature and prediction in the training dataset

The entire group (N=158) with complete survival information and RNA-seq expression profiles was randomly separated into the training dataset (N=80, Table S1)

and testing dataset (N=78, Table S2). The summary information of ESCA patients is shown in Table 1. The detailed characteristics and population demographics of the tissue samples is shown in Table S3. Seven ARGs, including *DDIT3* (DNA damage-inducible transcript 3), *HSP90AB1* (heat shock protein 90 alpha family class B member 1), *FKBP1A* (FKBP prolyl isomerase 1A), *DNAJB1* (DnaJ heat shock protein family (Hsp40) member B1), *BNIP1* (BCL2 interacting protein 1), *VAMP7* (vesicle-associated membrane protein 7) and *TBK1* (TANK binding kinase 1) were significantly associated with OS ($P < 0.05$). These genes were selected as prognostic candidate genes using a univariate Cox regression analysis. Based on the multivariate Cox regression analysis, four genes including *DNAJB1*, *BNIP1*, *VAMP7* and *TBK1*, were selected to construct the OS prediction model. Finally, a prognostic risk score based on the prognostic signature was established as follows: $(0.61184 \times \text{expression value of } DNAJB1) + (1.59675 \times \text{expression value of } BNIP1) + (0.90984 \times \text{expression value of } VAMP7) + (0.96329 \times \text{expression value of } TBK1)$. Here, 0.61184, 1.59675, 0.90984 and 0.96329 are regression coefficients for the respective genes derived from the Cox regression analysis. Using the median value of the risk score as the cutoff, we divided the ESCA cases into high- and low-risk groups. Kaplan-Meier curves indicate that the high-risk group had a lower survival probability, compared to the low-risk group in the training dataset

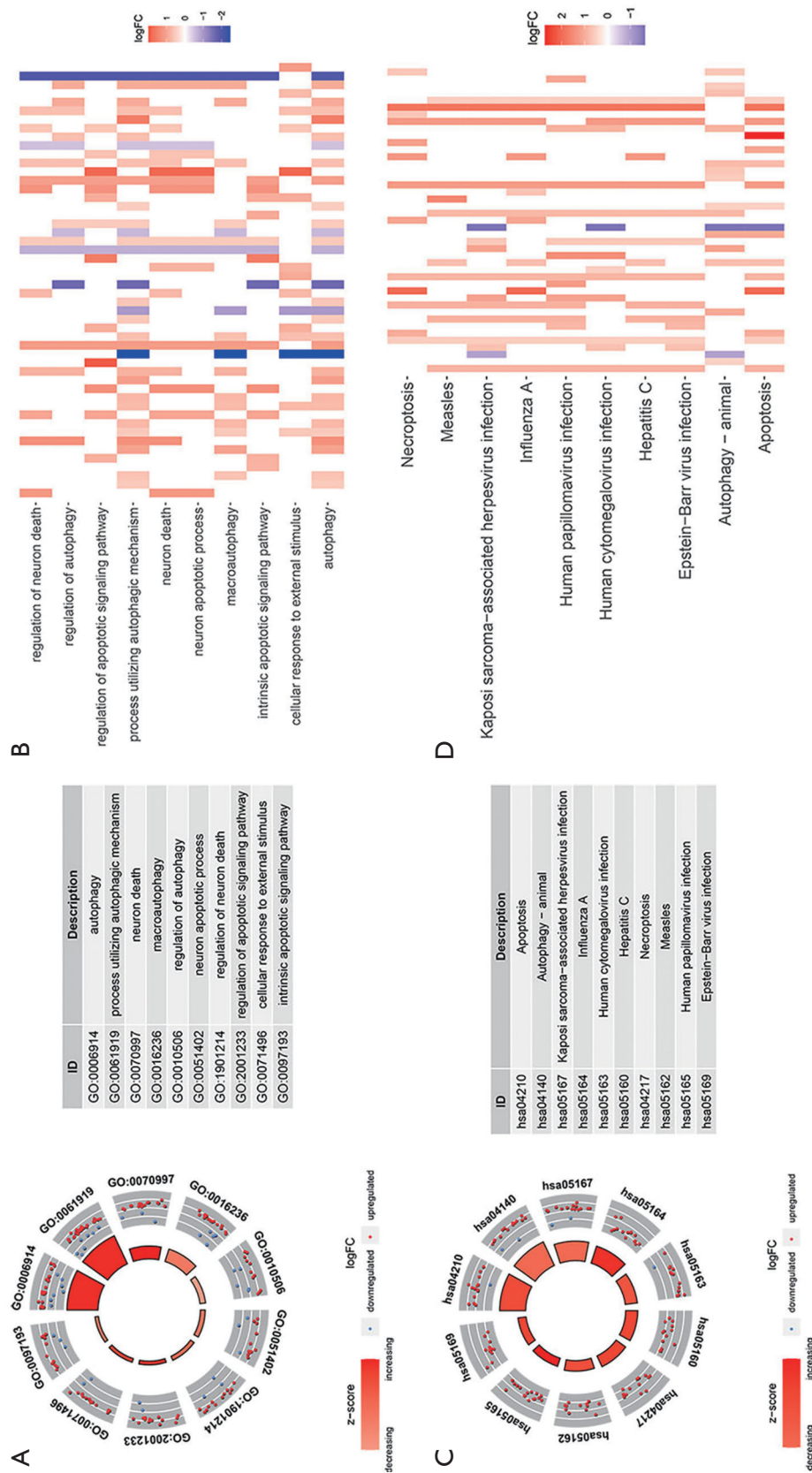


Figure 3 Functional analysis of the differentially expressed ARGs. GO (A) and KEGG (C) outer circles show a scatter plot for each term of the logFC for the assigned genes. Red circles display up-regulation, and blue ones down-regulation; (B,D) GO (B) and KEGG (D) heatmap of the relationship between ARGs and pathways. The color of each block depends on the logFC values. ARGs, autophagy-related genes; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; FC, fold change.

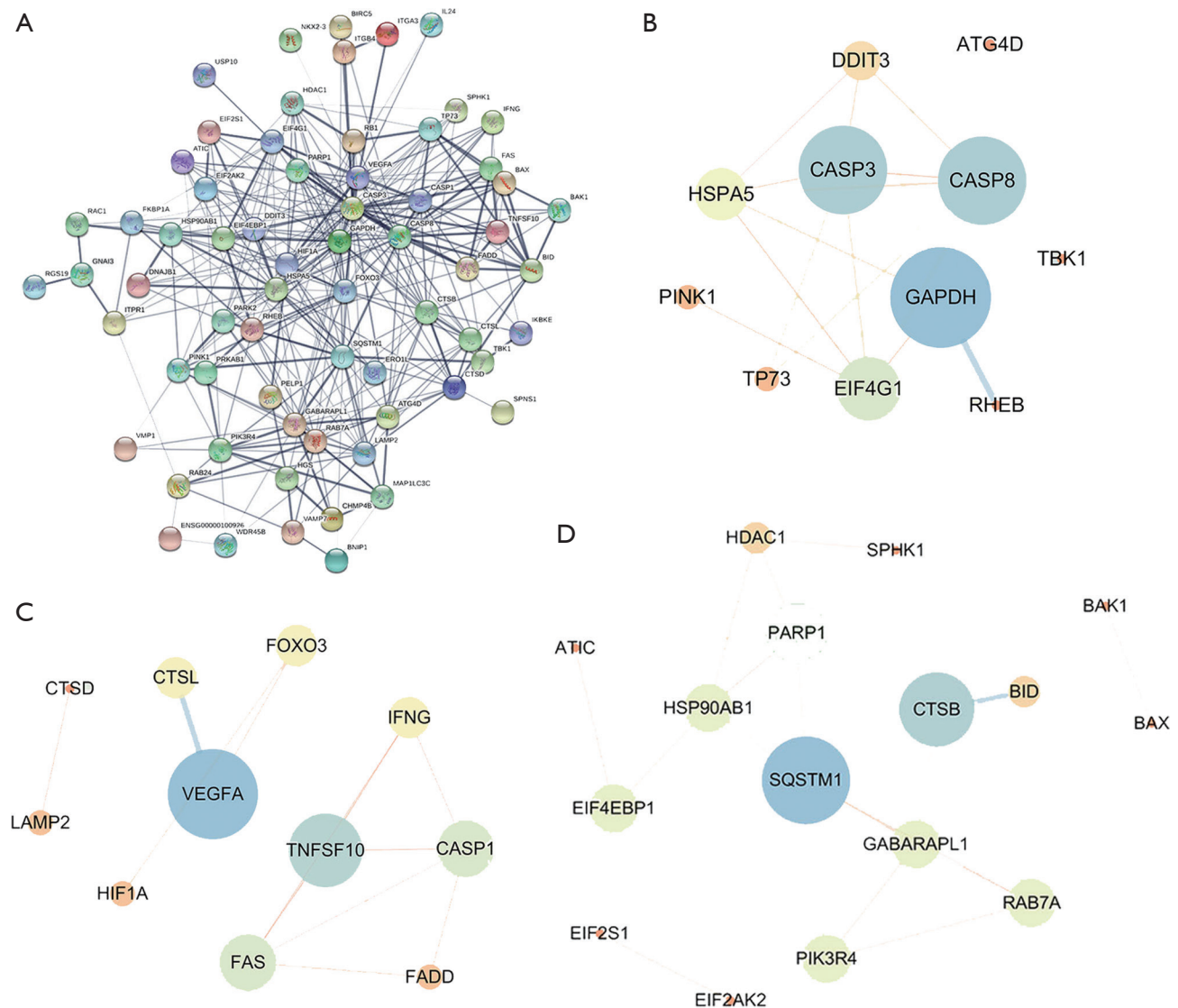


Figure 4 The three PPI networks and CASP3, VEGFA, and SQSTM1 modules. (A) PPI network; (B) CASP3 module; (C) VEGFA module; (D) SQSTM1 module. The color of a node in the PPI network indicates the logFC value of the Z score of gene expression and the size of the node represents the number of interacting proteins with the designated protein. PPI, protein-protein interaction; FC, fold change.

[hazards ratio (HR) =1.508, 95% confidence interval (CI): 1.201–1.894, $P=1.346e-05$; *Figure 5A*]. The ROC curve for the ARGs signature model is shown in *Figure 5B*, and the AUC for 1-year survival was 0.746. *Figure 5C,D,E* show the distribution of patient risk scores, survival time, and the expression of risk genes in the training dataset.

Construction of a nomogram based on the four ARGs

To develop a quantitative method for the prediction

of ESCA prognosis, we used the prognostic signature derived from the four ARGs to construct a nomogram. Based on the multivariate Cox analysis results, we drew a horizontal line to determine the points for each autophagy prognosis gene and calculated the total score of each patient by adding the points for all variables. The points were normalized to a 0 to 100 distribution. By calculating the total score, we could estimate the 1-, 2-, and 3-year survival of each ESCA patient (*Figure 6A*). Additionally, we evaluated the relationship between different clinical

Table 1 The summary information of patients with ESCA

Characteristic	Training dataset, n (%)	Testing dataset, n (%)	P
Age, years			0.3451
≤65	53 (66.25)	45 (57.69)	
>65	27 (33.75)	33 (42.31)	
Gender			0.6052
Male	70 (87.50)	65 (83.33)	
Female	10 (12.50)	13 (16.67)	
Grade			0.7484
G1–2	39 (48.75)	42 (53.85)	
G3	22 (27.50)	21 (26.92)	
Unknown	19 (23.75)	15 (19.23)	
T			0.6579
T1–2	34 (42.50)	30 (38.46)	
T3–4	40 (50.00)	39 (50.00)	
Unknown	6 (7.50)	9 (11.54)	
N			0.5182
N0	33 (41.25)	32 (41.03)	
N1–3	41 (51.25)	36 (46.15)	
Unknown	6 (7.50)	10 (12.82)	
M			0.1649
M0	65 (81.25)	54 (69.23)	
M1	4 (5.00)	4 (5.13)	
Unknown	11 (13.75)	20 (25.64)	
Stage			0.8538
I–II	43 (53.75)	41 (52.56)	
III–IV	29 (36.25)	27 (34.62)	
Unknown	8 (10.00)	10 (12.82)	

ESCA, esophageal cancer.

characteristics and the prognosis of ESCA by performing a Cox regression analysis. We found that tumor stage, distant metastasis, lymph node involvement, and risk score were associated with OS of ESCA patients ($P < 0.01$; *Figure 6B*). Importantly, as shown in *Figure 6C*, risk score remained as an independent risk predictor.

Validation of the ARG prognostic signature

To examine the robustness of the four-ARGs signature, the

prognostic value of the ARGs signature was validated in the testing dataset. Based on the expression value of the four prognostic ARGs, the prognostic risk score was calculated for patients in the testing dataset. Each patient was marked as a high- or low-risk based on the median risk score obtained from the training dataset. The Kaplan-Meier survival curves for the high- and low-risk groups in the testing dataset were significantly different ($P = 6.876 \times 10^{-3}$; *Figure 7A*). The AUC for the ROC curve of the OS-related predictive signatures in the testing dataset was 0.691 (*Figure 7B*).

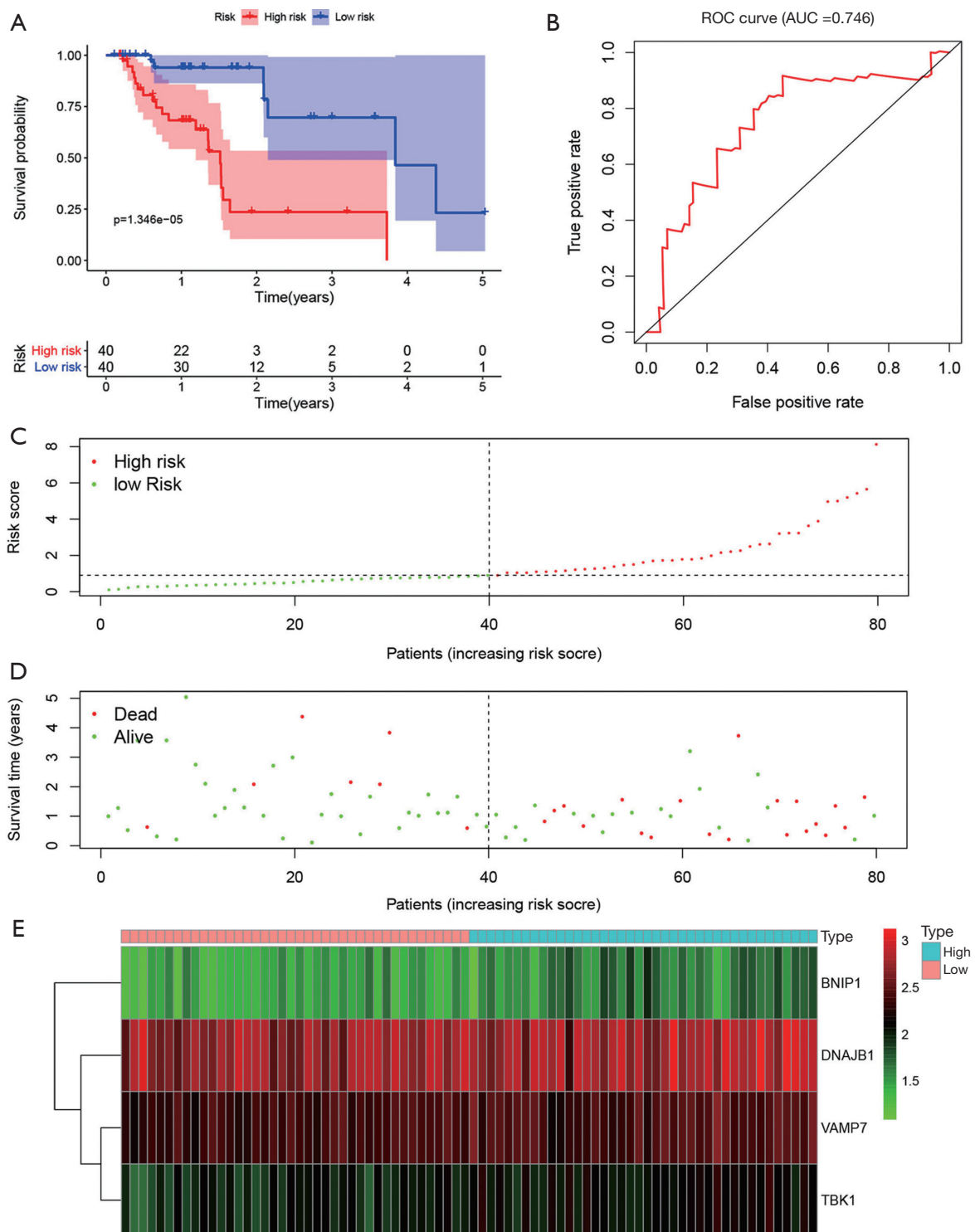


Figure 5 The ARGs prognostic signature for ESCA patients in the training dataset. (A) Kaplan-Meier curves plot show that patients in the high-risk group had significantly worse OS compared with those in the low-risk group; (B) ROC curve for the OS-related prognostic signature; (C) the number of patients in different risk groups; (D) scatterplots of ESCA patients with different survival status; (E) expression of risk genes in ESCA patients with different risks (low, pink; high, blue). ARGs, autophagy-related genes; ESCA, esophageal cancer; ROC, receiver operating characteristic; OS, overall survival. AUC, area under the curve.

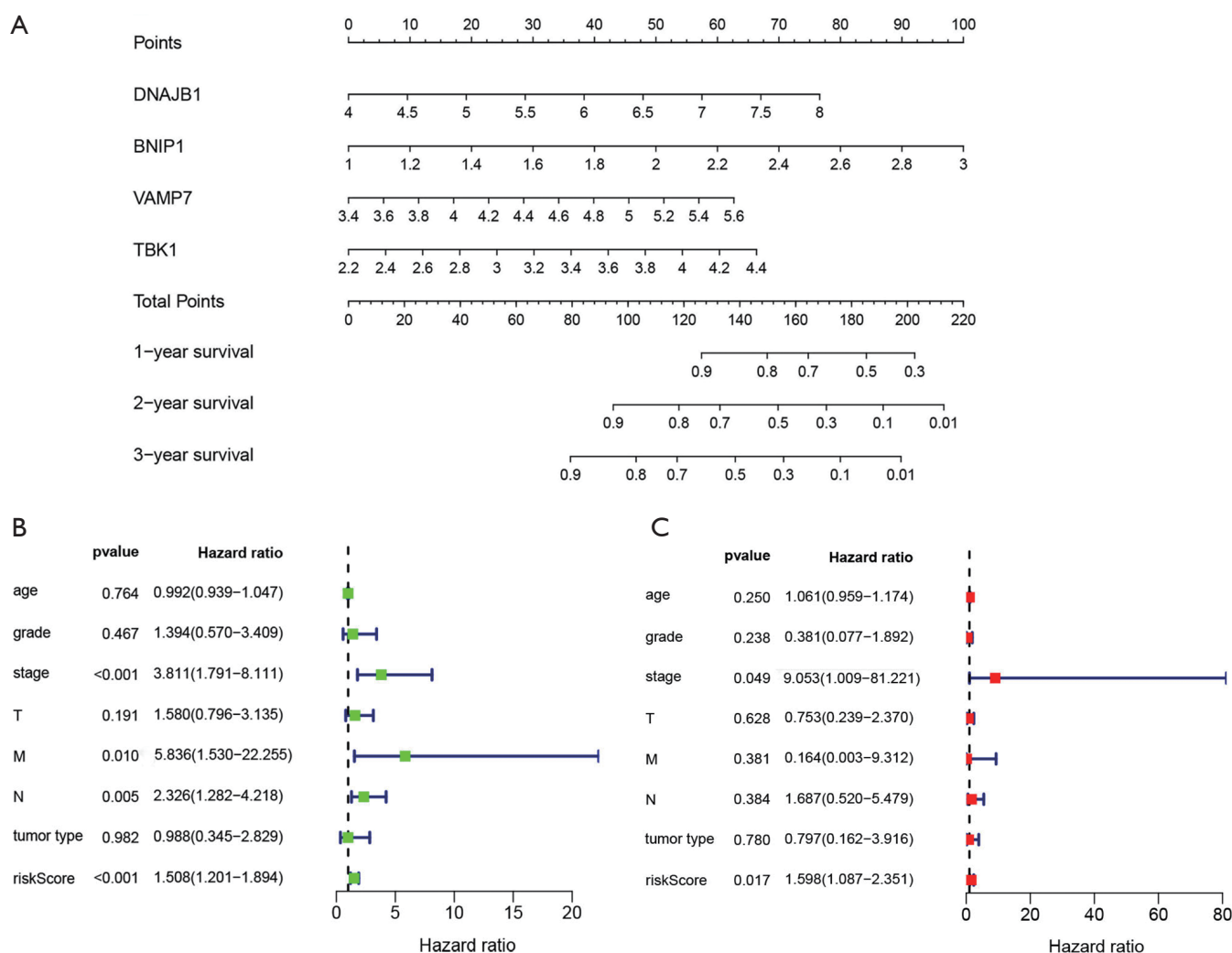


Figure 6 The nomogram and Cox regression analysis for the relationship between different clinical characteristics and the prognosis of ESCA. (A) The nomogram for predicting OS developed in training dataset; (B) univariate Cox regression analysis results; (C) multivariate Cox regression analysis results. ESCA, esophageal cancer; OS, overall survival.

The distribution of risk scores, survival status, and expression for the four prognostic ARGs in the testing dataset are shown in *Figure 7C,D,E*, respectively.

Independence of the prognostic value of the ARGs signature derived from other clinical variables

We evaluated the independence of the prognostic value of the ARGs signature based on other clinical variables. Univariate and multivariate Cox regression analyses were performed using age, grade, stage (T, N, M), tumor type (adenocarcinoma, squamous cell carcinoma), and the

prognostic risk score model as covariables. Multivariate Cox analysis revealed that the four-ARG signature was significantly associated with OS in each dataset (training: *Figure 6B,C*; testing: *Figure 8A,B*). Thus, the ARGs signature is a potential prognostic marker for ESCA. The AUCs of the ROC curves were 0.709 (risk score), 0.594 (age), 0.583 (grade), 0.655 (stage), 0.611 (T), 0.676 (lymph node status), and 0.509 (distant metastasis) (*Figure 9*).

Discussion

ESCA is a fatal malignant tumor, causing approximately

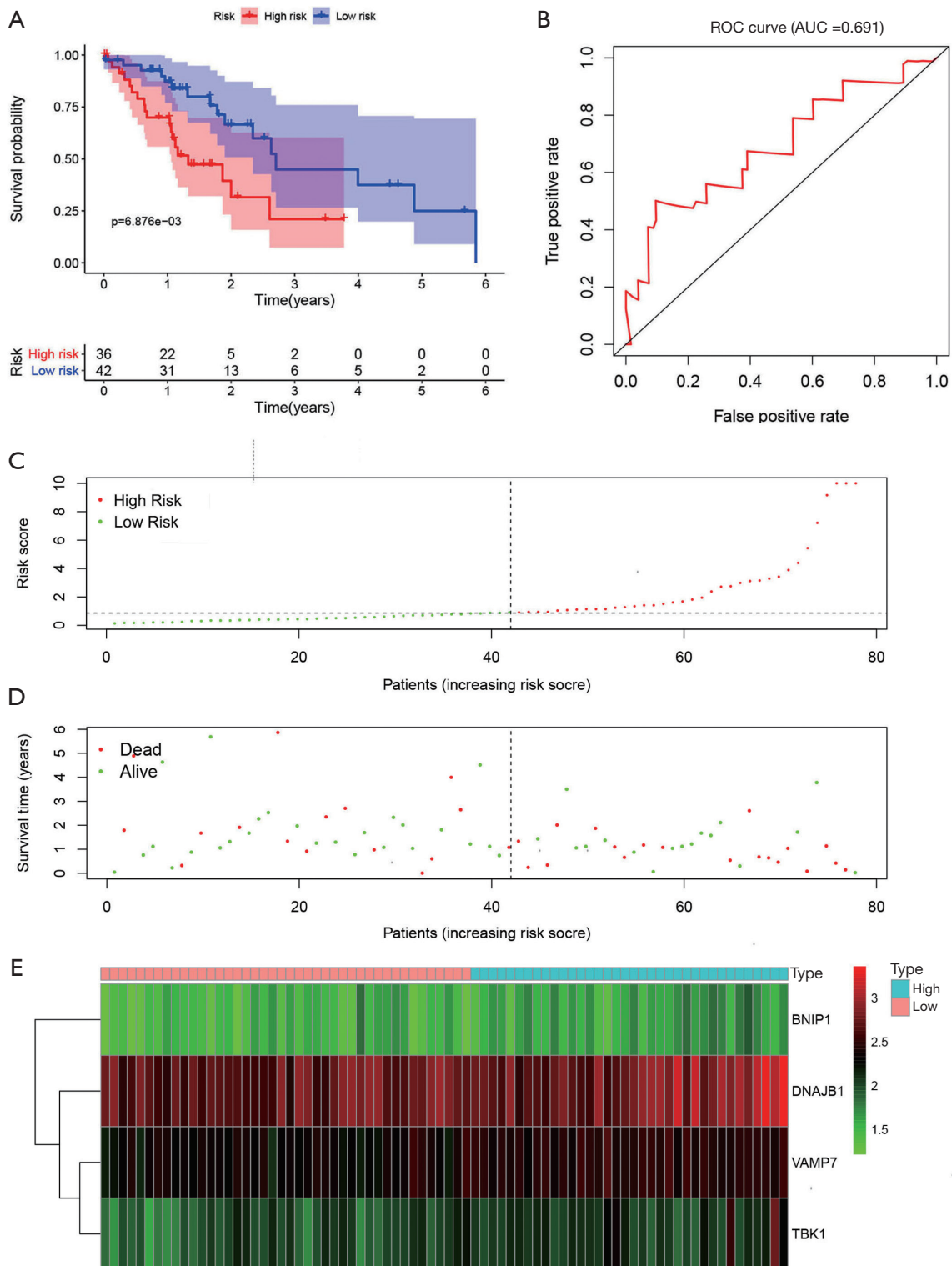


Figure 7 Validation of the predictive 4 ARGs prognostic signature in the testing dataset. (A) Kaplan-Meier curves results; (B) the AUC of ROC curves for predicting 1-year survival; (C) the number of patients in different risk groups; (D) scatterplots of ESCA patients with different survival status; (E) expression of risk genes in ESCA patients with different risks (low, pink; high, blue). ARGs, autophagy-related genes; AUC, area under the curve; ROC, receiver operating characteristic; ESCA, esophageal cancer.

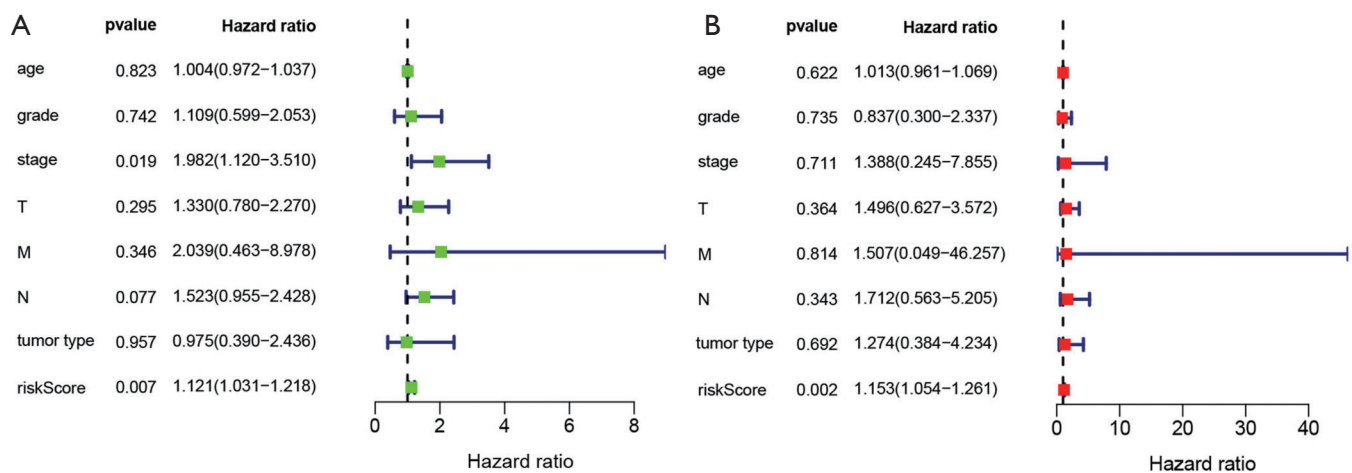


Figure 8 Cox regression analysis of the relationship between different clinical characteristics and the prognosis of ESCA. (A) Univariate Cox regression analysis results in the testing dataset; (B) multivariate Cox regression analysis results in the testing dataset. ESCA, esophageal cancer.

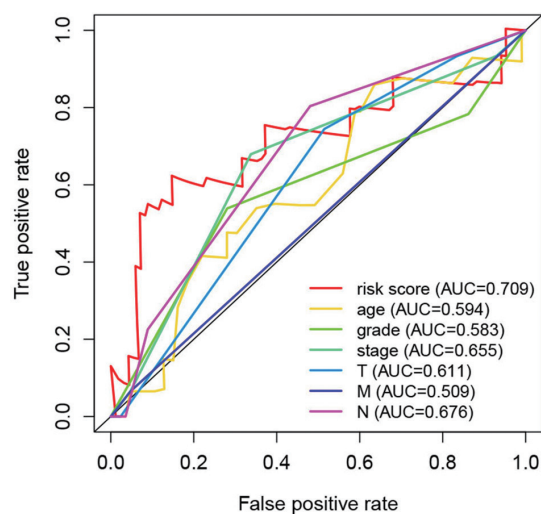


Figure 9 Comparison of risk score and clinical features in predicting the accuracy of patient survival prognosis. AUC, area under the curve.

500,000 deaths worldwide in 2018 (8), and the incidence is still growing. Unfortunately, there are no effective molecular biomarkers for monitoring ESCA prognosis. In recent years, the focus of ESCA research has turned to tumor autophagy (9). However, most of the current research on autophagy in ESCA only focuses on a single gene involved in autophagy (10,11). A comprehensive analysis of ARGs has not been conducted to explore the clinical significance of autophagy in ESCA. In the present study, we

analyzed the RNA-seq expression profiles and corresponding clinical information of ESCA patients downloaded from TCGA. To obtain the genes essential for ESCA from the perspective of autophagy, we first screened 69 ARGs that were differentially expressed between ESCA tumor and non-tumor tissues. A functional analysis showed that the genes enriched in the most significant KEGG pathway (e.g., apoptosis and autophagy-animal) were upregulated. Based on these results, we speculated that autophagy is activated during the initiation of ESCA. In addition, analysis of the PPI network for ESCA indicated that the main nodules were related to autophagy, apoptosis, and ubiquitin-like protein process. Subsequently, patients were randomly divided into a training dataset and a testing dataset. Univariate Cox regression and multivariate Cox regression analyses were performed in the training dataset. An autophagy-related prognostic signature was constructed based on the four key ARGs. This prognostic signature was then validated in the testing dataset. Kaplan-Meier curves showed the high-risk group had worse OS compared to the low-risk group in the two datasets. Univariate Cox regression and multivariate Cox regression analyses suggested that the four-ARG signature was as an independent risk predictor for patient survival, independent of other clinical factors.

These four ARGs play roles in the progression of various tumors. *DNAJ7B1* is associated with a variety of cellular processes, including the proteasome pathway (12), endoplasmic reticulum stress (13), and viral infection (14). Lately, more attention has been paid to the functions of

DNAJB1 in the progression of cancer. *DNAJB1* inhibits mitogen-inducible gene 6 (MIG6) stabilization and increases lung cancer cell proliferation (15). On the other hand, *DNAJB1* acts as an autophagy-related protein involved in the development of tumors. *DNAJB1* also targets programmed cell death 5 (PDCD5) to suppress p53-mediated apoptosis and enhance the proliferation of cancer cells (16). The up-regulation of *DNAJB1* correlates with poor prognosis for cholangiocarcinoma patients (17). Bnip1, a member of the BH3-only protein family, influences apoptosis, proliferation, invasion, and migration in cervical cancer cells through the regulation of mammalian target of rapamycin (mTOR) expression (18). Although *TBK1* was thought to only play a role in the immune response (19), a recent study showed that *TKB1* promotes the oncogenic phenotype in cancer and regulates autophagy (20). As for *VAMP7*, it was shown to be involved in multiple vesicular transport events, including autophagy (21). However, little is known about the functional mechanism of *VAMP7* in ESCA.

Some prognostic signatures have been derived from ARGs in several tumor types, including glioblastoma (22), hepatocellular carcinoma (23), and lung cancer (24). Our results indicated that an autophagy-related prognostic signature based on the four ARGs can be employed for the prognostic prediction of survival in ESCA patients. This prognostic tool may be the basis for establishing personalized treatment procedures, based on patient risk, for improved therapy of ESCA. We explored and validated the prognostic value of ARGs in ESCA. However, our research has several limitations. (I) Due to the relatively small number of ESCA tumor and non-tumor samples in the TCGA database, the present study inevitably has the problem of period. (II) Although we have verified the model in the testing dataset, further investigations *in vitro*, *in vivo*, and in independent ESCA cohorts are needed to confirm the accuracy of the model. (III) Furthermore, other potential prognostic variables associated with OS in ESCA, such as lymphovascular invasion (LVI) and neutrophil-to-lymphocyte ratio (NLR) need to be explored. (IV) Finally, the functional mechanisms of the four ARGs should be further investigated in ESCA.

Conclusions

In summary, we constructed and verified a novel ARG prognostic signature that could improve the individualized outcome prediction for patients with ESCA. For those ESCA patients with high risk scores, a more comprehensive

neoadjuvant/adjuvant therapy or closer follow-up are indicated for an improved prognosis.

Acknowledgments

The results of this study are based on the data from TCGA (<https://www.cancer.gov/tcga>). We thank all the authors who provided the data for this study. Meanwhile, we would also like to thank Prof Qingyi Wei from Duke Cancer Institute for his scientific editing.

Funding: This study was supported by Shanghai Municipal Commission of Health and Family Planning Outstanding Academic Leaders Training Program (2017BR055).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <http://dx.doi.org/10.21037/atm-20-4541>

Peer Review File: Available at <http://dx.doi.org/10.21037/atm-20-4541>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm-20-4541>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study conformed to the provisions of the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Domper Arnal MJ, Ferrández Arenas Á, Lanás Arbeloa Á.

- Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and Eastern countries. *World J Gastroenterol* 2015;21:7933-43.
2. Torre LA, Siegel RL, Ward EM, et al. Global Cancer Incidence and Mortality Rates and Trends--An Update. *Cancer Epidemiol Biomarkers Prev* 2016;25:16-27.
 3. Poillet-Perez L, White E. Role of tumor and host autophagy in cancer metabolism. *Genes Dev* 2019;33:610-9.
 4. Saxena R, Klochkova A, Murray MG, et al. Roles for Autophagy in Esophageal Carcinogenesis: Implications for Improving Patient Outcomes. *Cancers (Basel)* 2019;11:1697.
 5. Yazdani HO, Huang H, Tsung A. Autophagy: Dual Response in the Development of Hepatocellular Carcinoma. *Cells* 2019;8:91.
 6. Gu JX, Zhang X, Miao RC, et al. Six-long non-coding RNA signature predicts recurrence-free survival in hepatocellular carcinoma. *World J Gastroenterol* 2019;25:220-32.
 7. Siriwardhana C, Khadka VS, Chen JJ, et al. Development of a miRNA-seq based prognostic signature in lung adenocarcinoma. *BMC Cancer* 2019;19:34.
 8. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394-424.
 9. Liu H, Zhao J, Fu R, et al. The ginsenoside Rk3 exerts anti-esophageal cancer activity in vitro and in vivo by mediating apoptosis and autophagy through regulation of the PI3K/Akt/mTOR pathway. *PLoS One* 2019;14:e0216759.
 10. Ma Z, Chen C, Tang P, et al. BNIP3 induces apoptosis and protective autophagy under hypoxia in esophageal squamous cell carcinoma cell lines: BNIP3 regulates cell death. *Dis Esophagus* 2017;30:1-8.
 11. Wu J, Zhang D, Li J, et al. MACC1 induces autophagy to regulate proliferation, apoptosis, migration and invasion of squamous cell carcinoma. *Oncol Rep* 2017;38:2369-77.
 12. Yamazaki S, Uchiumi A, Katagata Y. Hsp40 regulates the amount of keratin proteins via ubiquitin-proteasome pathway in cultured human cells. *Int J Mol Med* 2012;29:165-8.
 13. Lenna S, Farina AG, Martyanov V, et al. Increased expression of endoplasmic reticulum stress and unfolded protein response genes in peripheral blood mononuclear cells from patients with limited cutaneous systemic sclerosis and pulmonary arterial hypertension. *Arthritis Rheum* 2013;65:1357-66.
 14. Batra J, Tripathi S, Kumar A, et al. Human Heat shock protein 40 (Hsp40/DnaJB1) promotes influenza A virus replication by assisting nuclear import of viral ribonucleoproteins. *Sci Rep* 2016;6:19063.
 15. Park SY, Choi HK, Seo JS, et al. DNAJB1 negatively regulates MIG6 to promote epidermal growth factor receptor signaling. *Biochim Biophys Acta* 2015;1853:2722-30.
 16. Cui X, Choi HK, Choi YS, et al. DNAJB1 destabilizes PDCD5 to suppress p53-mediated apoptosis. *Cancer Lett* 2015;357:307-15.
 17. Ren H, Luo M, Chen J, et al. Identification of TPD52 and DNAJB1 as two novel bile biomarkers for cholangiocarcinoma by iTRAQ-based quantitative proteomics analysis. *Oncol Rep* 2019;42:2622-34.
 18. Li FH, Xiang L, Ran L, et al. BNIP1 inhibits cell proliferation, migration and invasion, and promotes apoptosis by mTOR in cervical cancer cells. *Eur Rev Med Pharmacol Sci* 2019;23:1397-407.
 19. Xiao Y, Zou Q, Xie X, et al. The kinase TBK1 functions in dendritic cells to regulate T cell homeostasis, autoimmunity, and antitumor immunity. *J Exp Med* 2017;214:1493-507.
 20. Durand JK, Zhang Q, Baldwin AS. Roles for the IKK-Related Kinases TBK1 and IKKε in Cancer. *Cells* 2018;7:139.
 21. Fader CM, Aguilera MO, Colombo MI. ATP is released from autophagic vesicles to the extracellular space in a VAMP7-dependent manner. *Autophagy* 2012;8:1741-56.
 22. Wang QW, Liu HJ, Zhao Z, et al. Prognostic Correlation of Autophagy-Related Gene Expression-Based Risk Signature in Patients with Glioblastoma. *Onco Targets Ther* 2020;13:95-107.
 23. Mao D, Zhang Z, Zhao X, et al. Autophagy-related genes prognosis signature as potential predictive markers for immunotherapy in hepatocellular carcinoma. *PeerJ* 2020;8:e8383.
 24. Liu Y, Wu L, Ao H, et al. Prognostic implications of autophagy-associated gene signatures in non-small cell lung cancer. *Aging(Albany NY)* 2019;11:11440-62.
- Cite this article as:** Du H, Xie S, Guo W, Che J, Zhu L, Hang J, Li H. Development and validation of an autophagy-related prognostic signature in esophageal cancer. *Ann Transl Med* 2021;9(4):317. doi: 10.21037/atm-20-4541

Table S1 The training dataset

ID	OS (years)	fustat	DDIT3	HSP90AB1	FKBP1A	DNAJB1	BNIP1	VAMP7	TBK1
TCGA-2H-A9GL	0.493150685	1	4.827826371	8.970044921	6.956794355	5.997463221	2.423222977	4.27904124	3.622804475
TCGA-L5-A88T	1.901369863	0	2.654444248	7.955963111	6.197298878	6.91372478	1.660706871	3.806357056	2.478899863
TCGA-R6-A6L4	1.35890411	1	3.577958072	8.693904024	6.759510071	6.797365824	1.495175289	4.885442095	2.883341829
TCGA-IG-A8O2	0.389041096	1	4.451781482	9.654673975	6.980076726	5.932668546	1.917065832	4.990701686	3.201977401
TCGA-Z6-AAPN	0.221917808	0	3.941938251	9.574178918	7.092371314	7.173117437	2.398539452	4.678727168	2.949928936
TCGA-L5-A4OX	0.619178082	1	3.580676144	9.194472719	6.111061786	7.745737509	1.841659193	4.667248388	3.476281875
TCGA-JY-A6FH	3.564383562	0	2.789262465	9.076230385	5.76938272	5.347621369	1.789629104	4.18832021	2.461993929
TCGA-LN-A49S	1.095890411	0	2.955916838	8.879125251	6.641958809	6.564232886	1.939746005	4.326831408	2.87028442
TCGA-LN-A8HZ	1.02739726	0	3.938977329	8.971609939	6.798874825	6.30004225	1.734502833	3.980882617	2.736373463
TCGA-JY-A6FA	3.728767123	1	4.505311703	9.429841054	5.419283026	6.02997536	1.845603656	4.826518998	3.544737601
TCGA-IG-A6QS	0.830136986	1	3.552475444	9.301597477	6.781877679	6.513609176	1.846896833	4.600574843	2.711346004
TCGA-IG-A3YA	1.731506849	0	2.892253638	8.610190806	6.430352639	5.500102693	1.839777587	4.586666118	3.028583613
TCGA-V5-AASV	1.279452055	0	3.368847116	8.882714571	5.8094785	6.34992536	1.520640491	3.501940976	2.251194326
TCGA-L5-A4OP	0.597260274	0	2.538235936	10.12866716	7.770724011	5.517561999	2.171751899	4.255537744	2.738880871
TCGA-L5-A4ON	1.528767123	1	4.21978652	10.99759727	6.778897162	7.258559439	1.684689495	4.848191257	3.370370485
TCGA-IG-A3QL	1.673972603	0	2.977132257	9.136798398	7.050914117	6.345453506	1.770393952	4.2982558	2.792020305
TCGA-LN-A7HZ	1.098630137	0	3.71496641	9.255750917	7.052177845	5.406374419	1.91411251	4.364283164	3.181030176
TCGA-JY-A93E	2.101369863	0	3.880533706	8.92312065	6.816484512	5.117682992	1.442653734	4.660006735	3.077585342
TCGA-L5-A4OH	2.717808219	0	4.216075322	8.693056873	7.42608189	4.713457179	1.74080368	4.450796875	3.290687586
TCGA-LN-A9FP	1.002739726	0	3.616239682	8.134282179	6.242406294	4.576906398	1.319206921	3.879451846	2.900247219
TCGA-L5-A8NF	0.221917808	1	5.056688962	9.095747396	6.165624859	5.481255387	2.295275779	4.62205033	3.315560425
TCGA-L5-A4OO	0.276712329	0	3.393407621	8.071163403	6.628313984	5.440777754	1.926128777	3.888138061	3.855940959
TCGA-LN-A4A1	1.049315068	0	4.014711459	9.178792808	6.298019481	6.068160232	2.014054581	4.3005572	2.759815018
TCGA-L5-A893	0.252054795	0	2.627126539	8.219530646	6.312579577	5.544569391	2.036303516	4.01531994	2.693455053
TCGA-ZR-A9CJ	1.643835616	1	4.807556885	8.741279896	7.071307895	6.620831237	2.339571554	4.401319178	3.705768945
TCGA-IG-A97I	1.01369863	0	4.67336081	9.581857367	6.704082083	6.769031163	2.211186366	3.640251878	2.987497872
TCGA-KH-A6WC	0.523287671	0	1.766230043	7.415025101	6.185635656	7.152983455	1.297307649	3.897640258	2.250549619
TCGA-R6-A8WC	0.191780822	0	4.168735089	9.359903044	6.247705587	5.514609916	1.672084486	4.626546021	3.558500308
TCGA-IC-A6RE	0.64109589	0	3.997665185	8.799810302	6.847100763	5.021003475	2.198824792	4.617500625	2.983758777
TCGA-L5-A8NW	3.84109589	1	2.849987454	8.095240349	6.153231183	6.542483415	1.47412917	4.256373639	3.239190422
TCGA-LN-A7HY	1.002739726	0	3.591496778	9.064763027	6.498016846	5.328998089	2.690389225	4.214981566	2.889217396
TCGA-2H-A9GF	2.147945205	1	4.913540497	9.382232338	6.339090874	4.587228757	1.940633976	4.779305953	3.099426001
TCGA-L5-A8NS	1.117808219	0	4.113392127	8.311202916	5.957089037	6.043918389	1.760357224	4.64710782	3.42288422
TCGA-VR-AA4G	1.254794521	0	3.511055565	8.917980658	6.39234765	6.991039922	1.971422156	4.06764474	3.15938328
TCGA-R6-A6KZ	0.421917808	1	3.996410441	9.457982253	6.504448625	5.09183479	2.546459889	4.212763864	3.203744767
TCGA-2H-A9GI	1.191780822	1	4.040165634	8.939686417	6.22424318	4.687106761	1.949822734	4.691879086	3.619335686
TCGA-VR-A8EU	1.526027397	1	2.969442472	9.097479119	6.508355438	6.479622003	1.827541838	4.931424937	2.937903093
TCGA-L5-A891	0.312328767	0	4.356784005	9.478581157	6.965957398	5.23219833	2.020889025	3.848248354	2.520401127
TCGA-L5-A88W	2.093150685	1	4.294123715	8.725477034	6.805765795	6.505030263	2.183388055	3.982509022	2.320542731
TCGA-JY-A93C	1.931506849	0	3.474274059	8.271334533	6.700533182	5.03683049	2.366208837	4.529735511	3.383182933
TCGA-XP-A8T6	2.090410959	1	3.082085497	8.717037178	6.373392081	6.338339726	1.583222042	4.325619553	2.538504889
TCGA-LN-A49U	1.279452055	0	3.149268147	9.117042784	5.946308375	5.437497753	1.658801924	4.054448252	3.157955781
TCGA-LN-A7HX	1.019178082	0	3.335858945	9.329326125	6.807888105	5.814202525	1.625439768	4.525267092	3.217087215
TCGA-LN-A4A3	1.063013699	0	3.147803578	9.626747392	7.529423526	6.659413704	2.125942322	4.033962975	2.915110501
TCGA-L5-A8NG	2.997260274	0	3.937109888	9.912088897	6.509205097	4.764028147	1.583497903	4.645371954	3.404192496
TCGA-LN-A4A4	1.049315068	0	3.284619252	9.023827526	6.129183055	5.86452616	1.76466802	3.787445002	3.390306829
TCGA-Z6-A8JE	0.175342466	0	3.452088137	8.85596644	6.938312401	5.822454931	2.208212604	4.328239867	3.643980434
TCGA-LN-A49P	1.02739726	0	4.300710204	8.820363756	6.71365183	6.408940915	2.491396748	5.067056548	3.335804505
TCGA-L5-A4OW	0.594520548	1	3.315709191	9.487419522	6.739900899	5.312552864	1.569227961	4.798654571	3.455154047
TCGA-VR-A8EP	1.663013699	0	4.291865789	8.842561232	5.777570684	6.385815475	1.494902998	4.380511283	3.288302448
TCGA-2H-A9GO	1.353424658	1	5.303238266	9.09383242	6.606879975	5.371851205	2.612079741	4.96444438	3.385815889
TCGA-2H-A9GQ	0.350684932	1	5.253921591	9.971030473	7.348717533	5.992868791	2.096826877	4.534667995	4.245250731
TCGA-R6-A6DN	0.665753425	1	3.221582481	9.666495293	6.439795154	5.786362689	2.279582725	3.500058848	3.620005697
TCGA-L5-A8NV	4.380821918	1	5.293069891	9.406931805	6.323264133	5.027103613	1.943288614	4.248269923	3.092103055
TCGA-L5-A4OJ	1.750684932	0	3.987531888	8.743005939	6.822117	4.757674553	1.974020425	4.361457473	3.291834854
TCGA-JY-A6FB	5.032876712	0	3.568377532	11.4707031	6.57772776	5.999144273	1.774877281	3.550287512	2.930583694
TCGA-IG-A3YB	0.219178082	0	4.240392187	8.355256823	6.358778829	5.240705483	2.103535574	3.837628438	2.566320779
TCGA-L5-A88Z	0.616438356	0	5.440921976	9.092438709	6.699494721	7.323753752	1.863527382	4.286592876	3.158915734
TCGA-VR-A8Q7	3.57260274	0	2.797230769	9.729918193	6.649087422	4.652210223	1.292390897	4.621832588	3.425215492
TCGA-L5-A4OG	0.391780822	0	5.214517296	8.566661441	6.914852699	6.349799953	1.716211299	3.906968499	3.196442434
TCGA-V5-A7RE	1.369863014	0	3.749967547	9.395055728	6.332777561	6.246039765	1.919768727	4.354996524	2.971719564
TCGA-L5-A4OU	2.416438356	0	4.414071255	8.947668147	6.779148417	6.760296518	1.98306522	4.431677798	3.377227279
TCGA-2H-A9GK	0.635616438	1	3.733458471	10.79812992	6.921188751	5.053892006	1.700155146	3.938247207	3.034048929
TCGA-R6-A6XG	3.2	0	3.200781949	9.646613789	6.592511007	5.010280806	2.893132293	3.801546703	3.186744737
TCGA-Q9-A6FW	0.652054795	0	3.762331805	8.431608335	6.30194407	6.992034209	1.730347366	3.984077982	2.971544474
TCGA-IG-A4P3	1.553424658	1	3.667198421	8.806021746	6.533274905	6.323714668	1.802678377	4.044156274	3.717245401
TCGA-LN-A49M	1.054794521	0	2.81780476	9.109260514	6.182299197	6.049905603	1.128059093	5.34711143	3.286546451
TCGA-L5-A8NK	1.128767123	0	4.107169199	8.464565741	6.360561598	6.639509208	1.443001776	4.603428314	2.921035179
TCGA-M9-A5M8	2.75890411	0	2.982214012	8.95078566	6.526655413	5.33779992	1.444355386	4.172800311	3.352434956
TCGA-L5-A8NN	0.457534247	0	6.270365228	9.245535356	6.94509146	4.10191812	2.593186423	4.033171885	3.702761697
TCGA-VR-A8EZ	1.515068493	1	5.213103831	9.04673796	7.424273804	6.151930033	2.470724267	4.399376885	3.208545976
TCGA-Z6-A9VB	0.109589041	0	2.888983021	10.01699488	7.014909812	6.399442856	1.593443653	3.981180456	3.150315633
TCGA-LN-A9FR	1.021917808	0	3.400333807	9.483771851	6.201351567	6.00345877	1.549179729	4.374977551	2.656929915
TCGA-2H-A9GN	0.745205479	1	3.559974384	9.887415854	6.067007605	7.534549021	2.142309021	4.254821037	3.200961944
TCGA-LN-A7HW	1	0	3.419043301	9.213812388	5.709825756	6.108118772	1.618081219	4.676163017	2.761581017
TCGA-LN-A49O	1.117808219	0	3.306929762	8.617228313	6.16147053	6.636413928	1.632633497	4.382095179	2.853812854
TCGA-IC-A6RF	1.306849315	0	2.837697894	8.379577263	6.228612392	6.185030883	1.665081949	3.748654817	3.015292947
TCGA-LN-A4A8	1.293150685	0	3.513035884	9.057370391	7.479867551	5.402529141	1.481332528	5.463765647	4.103988526
TCGA-V5-A7RC	0.284931507	1	4.422648792	11.20445866	7.122608533	6.269965957	2.389198657	3.915365055	3.062476642
TCGA-LN-A5U5	0.37260274	1	3.978027183	9.233727515	6.745542568	6.417232445	2.230542431	4.127413722	3.689774389

fustat: 1 = detach; 0 = alive. OS, overall survival.

Table S2 The testing dataset

ID	OS (year)	fustat	DDIT3	HSP90AB1	FKBP1A	DNAJB1	BNIP1	VAMP7	TBK1
TCGA-L5-A8NU	5.846575342	1	4.199072732	8.288675123	6.6482317	4.917927349	2.067159753	3.951623099	2.957914203
TCGA-R6-A6XQ	0.528767123	1	4.262723291	8.763269808	7.191064151	6.029689414	2.022167509	5.11092653	3.187968653
TCGA-LN-A7HV	0.876712329	0	3.714558755	9.229722977	6.992561119	5.93186551	1.884917095	3.870938724	2.344967932
TCGA-S8-A6BV	1.668493151	0	4.252639904	9.878413834	7.114791332	6.796336167	1.749740791	4.743578875	3.146978326
TCGA-JY-A6FE	0.306849315	1	3.357585784	8.901993662	7.242033661	4.632596275	2.080845577	4.342373402	2.247818776
TCGA-LN-A8I0	1.115068493	0	4.407540384	8.2827391	6.697584798	6.340447974	1.399529877	3.395398475	2.905674565
TCGA-VR-AA4D	3.775342466	0	3.911616809	8.725487259	6.67663554	6.062617812	2.835956605	4.509011064	3.391991199
TCGA-V5-A7RB	0.44109589	1	4.025737561	8.912499277	6.585528647	5.99116762	2.607598527	4.37065463	3.172022757
TCGA-L5-A8NL	1.101369863	0	4.017193463	8.809586766	6.373227522	5.729176593	1.997210415	4.46409707	3.529764032
TCGA-IG-A51D	1.419178082	0	2.634722431	9.57510486	6.983084809	5.447945774	2.182052476	4.134896121	3.110427224
TCGA-L5-A4OE	2	1	3.735684634	9.151996021	6.383866695	6.734384469	1.500829071	4.189267927	3.466948118
TCGA-LN-A49W	1.104109589	0	3.037053582	8.578329511	6.397038357	5.3137079	1.936554355	4.524871888	3.42034922
TCGA-L5-A8NR	0.726027397	0	4.903733894	8.333759143	6.144571373	5.771239601	1.747783952	4.211092428	3.49389963
TCGA-L5-A8NT	2.260273973	0	4.225954213	8.487307064	6.333878299	5.296853744	1.331835093	4.344080327	3.53500916
TCGA-VR-A8EW	0.676712329	1	5.104419696	8.894695614	7.417003522	8.10798439	2.112540677	4.260871749	2.662580825
TCGA-LN-A8I1	1.098630137	0	5.154648418	9.216335298	7.668157427	6.201263339	2.162233243	3.589864137	3.101433252
TCGA-2H-A9GR	2.704109589	1	3.883979699	8.969787721	6.421706138	4.763447636	1.948598123	4.482552066	2.965698154
TCGA-IG-A3I8	1.304109589	0	3.70945497	8.794595542	6.995016358	5.739048514	1.402706353	4.085941269	3.188732282
TCGA-IG-A3YC	1.676712329	0	3.669009752	8.866825162	7.063760697	5.626661657	1.819306443	4.080683243	3.116686965
TCGA-R6-A8W5	1.315068493	1	3.118854218	9.719083478	6.499683116	4.886554049	1.820185091	4.481888533	2.919943542
TCGA-L5-A43J	0	1	3.445697044	8.501755198	6.940694154	5.238488814	2.096769541	4.213642012	2.991917969
TCGA-IG-A50L	0.043835616	0	3.584465888	8.551682118	6.666337638	7.036350572	1.63692525	4.419190985	3.158279338
TCGA-IG-A5S3	1.950684932	0	3.892735871	8.966416047	6.780180156	4.899205606	2.195981867	3.302831252	3.420485636
TCGA-L5-A88S	1.290410959	0	3.370688941	9.410639812	7.260132009	6.310800888	1.953198551	4.132890188	2.296048843
TCGA-LN-A9FO	0.010958904	0	5.429068463	9.640637009	6.904477748	9.284367722	2.289380314	4.985081838	3.772338668
TCGA-VR-A8ER	1.035616438	1	4.781741584	9.410236362	7.420269324	7.31629511	2.367340869	4.300685314	2.929696708
TCGA-V5-AASX	0.747945205	0	2.422795434	9.506860946	6.581242016	5.133474341	1.322727127	4.49800827	2.704992774
TCGA-VR-A8ET	0.128767123	1	4.08527769	8.624788565	6.422412847	7.559366392	1.792264305	4.33282887	5.878286469
TCGA-IG-A4QS	0.323287671	1	4.58431254	9.282349912	7.261394475	5.075551884	2.08510576	4.58230369	3.087835518
TCGA-IG-A7DP	1.238356164	0	3.611938387	8.38343819	6.797638183	4.417029522	2.088171794	4.12158076	3.234861336
TCGA-L7-A56G	0.904109589	1	2.513000752	8.925938613	6.420978845	6.41819872	1.571130205	3.944328111	2.904218945
TCGA-2H-A9GG	1.671232877	1	4.497432965	8.201927442	5.875552937	4.786805728	1.848834396	4.152295837	2.898033396
TCGA-L5-A4OM	3.994520548	1	3.335541046	9.147650697	6.236166234	5.932902503	1.347142195	4.919567732	3.22502176
TCGA-L5-A8NJ	1.37260274	0	3.761004367	8.715073112	6.300068066	5.584229546	2.332801294	4.179718841	2.937732888
TCGA-Z6-A8JD	0.284931507	0	4.68027259	9.169492128	6.143965379	8.193684529	2.223139235	4.096713046	2.528141947
TCGA-LN-A4A5	1.865753425	1	3.107633474	8.82846276	6.521882435	5.00270162	1.867416935	4.760363612	3.518740324
TCGA-LN-A9FQ	1.071232877	0	2.716728443	9.031629256	6.613404498	6.777659049	1.481729309	3.957791119	3.139896412
TCGA-L5-A8NI	1.123287671	1	4.661955912	9.797306968	5.946920541	7.606077955	2.379150707	4.59615874	3.332433115
TCGA-IG-A625	1.068493151	1	2.914472104	9.176324019	6.648331694	5.536953371	1.349198237	4.926625716	3.639322474
TCGA-RE-A7BO	0.583561644	1	4.66079003	9.031425248	7.279989369	5.073597901	2.30303196	4.00341964	2.989790254
TCGA-2H-A9GJ	4.879452055	1	3.679721045	9.027329356	6.047096858	4.578984589	1.703839139	3.91168583	2.968118592
TCGA-IG-A5B8	0.065753425	1	3.749838864	9.139522447	7.119343726	7.193117258	2.491604547	4.551715336	2.907715581
TCGA-VR-A8EY	2.309589041	0	4.796416575	9.367308593	7.049451073	6.036503554	2.538913067	3.537499844	2.306701748
TCGA-L5-A88V	0.216438356	0	4.710229516	9.300249589	6.955575519	4.970090701	1.642495953	4.401958066	2.514128574
TCGA-R6-A8WG	1.057534247	1	3.210161368	8.868445989	6.025075187	5.013508588	1.512261855	5.380992185	3.799012941
TCGA-LN-A49Y	1.038356164	0	4.637795031	8.837986565	6.022209564	5.39889309	2.381195781	3.979254633	3.115990925
TCGA-JY-A93F	2.002739726	0	2.455902934	8.534422729	6.691155855	6.791182428	1.816134775	3.796639282	2.850046682
TCGA-L5-A8NM	0.646575342	1	3.312395369	9.035965025	6.511355443	5.462754185	2.147323633	4.158947716	3.441338326
TCGA-LN-A4MQ	1.02739726	0	3.443075223	8.703734038	6.107973618	5.367102283	1.877346129	4.31767743	4.041360398
TCGA-L5-A4OS	3.482191781	0	4.046177587	9.525793858	7.167907826	4.661001612	2.08254932	4.629197478	3.447606834
TCGA-VR-A8EX	2.342465753	1	3.424756988	8.407481497	6.567587428	6.820775068	1.687879029	3.860167259	2.665785229
TCGA-L7-A6VZ	0.863013699	0	4.161861891	9.970807649	6.54125571	5.192935842	1.925077318	4.981173148	3.271131758
TCGA-LN-A49X	1.052054795	0	2.354254108	9.293276901	7.110052068	4.908414239	1.753744007	4.281279103	2.934784069
TCGA-L5-A4OI	1.665753425	0	3.188258167	8.507171848	6.417985304	5.226446104	1.926141403	3.756809814	3.067349426
TCGA-R6-A6Y0	4.495890411	0	3.22279158	7.67095998	6.42898067	6.870658161	1.786322851	3.905822606	2.969706219
TCGA-2H-A9GM	1.161643836	1	4.813047988	10.67710562	7.125514727	4.993211955	2.24371386	4.656578743	3.218133592
TCGA-L5-A8NQ	1.780821918	1	2.765224697	8.657968511	6.560763079	6.485464891	1.622038016	3.611511407	2.137826976
TCGA-LN-A5U6	1.02739726	0	3.468280932	8.533115941	7.323001602	6.856476575	2.038238096	3.719260119	2.526706612
TCGA-IG-A97H	1.208219178	0	3.486624282	8.88423708	7.14636909	5.993441365	2.358774329	4.385617402	2.912549275
TCGA-L5-A43E	2.520547945	0	4.160424903	9.23270608	6.722800702	4.870307997	1.640842081	4.239040532	3.405097203
TCGA-VR-A8EQ	1.901369863	1	3.686833289	9.23707704	6.147631106	5.413886153	1.789904909	4.118580665	2.83122322
TCGA-VR-AA7I	1.326027397	1	3.280846772	8.816638743	6.847654977	5.897763429	1.607926327	4.706273596	3.195489341
TCGA-V5-AASW	0.77260274	0	4.049273535	8.400253991	6.68180624	4.830623457	1.823517451	4.264423083	3.411452415
TCGA-R6-A8W8	0.24109589	1	3.13275466	10.46286827	6.639949569	5.897222755	1.873171027	4.59328026	2.895618273
TCGA-L5-A88Y	0.030136986	0	3.32235132	9.937172226	7.203890227	5.907790797	1.379479014	3.578287844	2.582447134
TCGA-JY-A93D	2.630136986	1	3.884041052	8.60030544	6.494039846	5.621471695	1.570234608	4.506845637	3.448919711
TCGA-XP-A8T8	1.197260274	0	3.109314754	9.381449941	6.166142591	6.421782593	1.943799456	4.26680078	2.639791302
TCGA-L5-A8NE	4.624657534	0	3.661454328	9.023681177	5.594459274	5.559113347	1.809469599	4.166524251	2.026629404
TCGA-JY-A939	1.808219178	0	4.130080813	10.67164367	6.164916377	5.370657071	2.106418342	4.07342438	3.086228894
TCGA-LN-A4A9	0.961643836	1	3.139927673	9.15324294	5.925800322	6.562840599	1.508080741	3.717970278	3.399542591
TCGA-JY-A6FD	5.668493151	0	3.046376681	8.65970019	5.755862595	5.984754298	1.831882509	3.82515571	2.541863453
TCGA-2H-A9GH	2.605479452	1	3.590918139	8.77867643	6.813091097	5.267095657	2.533694157	5.016083973	3.049253841
TCGA-L5-A4OT	0.408219178	1	5.971828173	9.666420227	6.417926411	8.914611178	1.991722431	4.507515126	3.519003044
TCGA-R6-A6DQ	0.632876712	1	5.1257573	9.77314102	7.086522662	5.431159926	2.31208339	4.684368246	3.675572075
TCGA-VR-A8EO	1.567123288	0	4.249816699	9.857720703	6.676206966	6.922418771	2.065550033	4.7855304	2.714891197
TCGA-S8-A6BW	1.698630137	0	4.431719894	8.456499807	5.84652942	6.221475761	1.827325579	4.205609902	4.737176753
TCGA-LN-A5U7	2.104109589	0	2.810750119	8.868932103	6.848535312	7.001446478	1.989445684	4.46760555	3.219720086
TCGA-L5-A8NH	1.076712329	1	3.776290368	9.900583976	6.821666852	6.250500049	1.846099317	4.51917955	3.068628165

fustat: 1 = detach; 0 = alive. OS, overall survival.

Table S3 The characteristics and population demographics of the tissue samples

ID	Age	Gender	Grade	Stage	T	M	N	Tumor Type	Race	Reflux History	Alcohol History
TCGA-L5-A4OG	79	Female	G1	I	T1	M0	N0	Adeno	White	No	Yes
TCGA-L5-A4OJ	70	Female	G3	I	T1	M0	N0	Adeno	White	Yes	Yes
TCGA-L5-A893	71	Female	Unknown	I	T1	M0	N0	Adeno	White	Yes	Yes
TCGA-2H-A9GJ	57	Male	G2	I	T1	M0	N0	Adeno	Unknown	Yes	No
TCGA-JY-A6FB	77	Male	G2	I	T1	M0	N0	Adeno	White	Unknown	Yes
TCGA-L5-A4OH	71	Male	Unknown	I	T1	M0	N0	Adeno	White	No	Yes
TCGA-L5-A43E	74	Male	Unknown	I	T1	M0	N0	Adeno	White	Yes	Yes
TCGA-IC-A6RF	69	Female	Unknown	I	T1	M0	N0	SCC	White	Yes	Yes
TCGA-KH-A6WC	82	Male	Unknown	I	T1	M0	N0	SCC	White	No	Yes
TCGA-V5-A7RE	45	Male	G2	I	T1	M0	N0	Adeno	White	Yes	Yes
TCGA-JY-A93F	58	Female	G1	I	T2	M0	N0	SCC	White	No	Yes
TCGA-LN-A811	67	Female	G3	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A9FP	60	Female	G3	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A49M	62	Male	G1	II	T2	M0	N0	SCC	Asian	No	No
TCGA-LN-A4A9	58	Male	G1	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A5U7	46	Male	G1	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A7HV	58	Male	G1	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A9FO	42	Male	G2	II	T2	M0	N0	SCC	Asian	No	No
TCGA-IG-A97I	58	Male	G2	II	T2	M0	N0	SCC	White	No	No
TCGA-LN-A4A5	49	Male	G2	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A4A8	52	Male	G2	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A7HW	59	Male	G2	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A7HX	72	Male	G2	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A7HZ	49	Male	G2	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A8I0	52	Male	G2	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-IG-A3QL	54	Male	G2	II	T2	M0	N0	SCC	White	Unknown	Yes
TCGA-VR-A8ET	64	Male	G2	II	T2	M0	N0	SCC	White	Unknown	Yes
TCGA-L5-A8NU	84	Male	G3	II	T2	M0	N0	Adeno	White	Yes	No
TCGA-LN-A8HZ	56	Male	G3	II	T2	M0	N0	SCC	Asian	No	Yes
TCGA-L5-A8NQ	71	Male	G3	II	T2	M0	N0	SCC	White	Yes	Yes
TCGA-IG-A5B8	72	Male	G1	I	T3	M0	N0	SCC	White	No	Yes
TCGA-VR-A8EY	44	Female	G1	II	T3	M0	N0	SCC	White	No	No
TCGA-L5-A8NK	84	Female	G1	II	T3	M0	N0	SCC	White	Yes	Yes
TCGA-JY-A6FD	51	Female	G1	II	T3	M0	N0	SCC	Asian	Unknown	Yes
TCGA-IG-A3I8	51	Female	G2	II	T3	M0	N0	SCC	White	Unknown	No
TCGA-Z6-AAPN	57	Male	G1	II	T3	M0	N0	SCC	White	Unknown	Unknown
TCGA-LN-A49U	62	Male	G1	II	T3	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A49S	59	Male	G2	II	T3	M0	N0	SCC	Asian	No	No
TCGA-LN-A49X	44	Male	G2	II	T3	M0	N0	SCC	Asian	No	No
TCGA-2H-A9GR	80	Male	G2	II	T3	M0	N0	Adeno	Unknown	Unknown	No
TCGA-LN-A49P	71	Male	G2	II	T3	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A49Y	77	Male	G2	II	T3	M0	N0	SCC	Asian	No	Yes
TCGA-M9-A5M8	58	Male	G2	II	T3	M0	N0	Adeno	Unknown	No	Yes
TCGA-L5-A8NV	75	Male	G3	II	T3	M0	N0	Adeno	White	Yes	No
TCGA-LN-A49O	47	Male	G3	II	T3	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A4A1	60	Male	G3	II	T3	M0	N0	SCC	Asian	No	Yes
TCGA-LN-A9FO	62	Male	G3	II	T3	M0	N0	SCC	Asian	No	Yes
TCGA-VR-A8EO	49	Male	G3	II	T3	M0	N0	SCC	White	Yes	Yes
TCGA-L5-A4OU	81	Male	Unknown	II	T3	M0	N0	Adeno	White	Yes	Yes
TCGA-IG-A5S3	69	Female	G2	II	T3	M0	N0	SCC	White	Unknown	Yes
TCGA-Z6-A8JD	53	Male	G2	II	T3	M0	N0	SCC	White	Unknown	No
TCGA-L5-A8NT	69	Male	Unknown	II	T3	M0	N0	Adeno	White	No	Yes
TCGA-L5-A8NS	76	Male	Unknown	II	T3	M0	N0	Adeno	White	Yes	Yes
TCGA-VR-AA7I	70	Male	G1	III	T4	M0	N0	SCC	White	No	Yes
TCGA-IG-A3YA	53	Male	G2	III	T4	M0	N0	SCC	White	Unknown	Yes
TCGA-L5-A8NH	54	Male	Unknown	IV	T1	M1	N0	Adeno	White	Yes	Yes
TCGA-L5-A8NF	57	Male	Unknown	IV	T1	M1	N0	Adeno	White	No	Yes
TCGA-L5-A4OP	67	Female	G2	I	T1	Unknown	N0	Adeno	White	Yes	Yes
TCGA-L5-A8BY	76	Male	Unknown	Unknown	T1	Unknown	N0	Adeno	White	No	Yes
TCGA-S8-A6BW	51	Male	G1	I	T2	Unknown	N0	SCC	White	Yes	Yes
TCGA-L5-A88S	84	Male	Unknown	I	T3	Unknown	N0	SCC	White	Yes	No
TCGA-L5-A88W	67	Male	Unknown	II	T3	Unknown	N0	SCC	White	No	Yes
TCGA-L5-A43J	90	Male	G3	II	T3	Unknown	N0	SCC	White	No	Yes
TCGA-V5-AASV	67	Male	G3	II	T3	Unknown	N0	SCC	Black	Yes	Yes
TCGA-L5-A4OM	54	Female	Unknown	I	T1	Unknown	N0	SCC	White	No	Yes
TCGA-L5-A88Z	70	Female	Unknown	II	T1	M0	N1	SCC	White	Yes	Yes
TCGA-RE-A7BO	72	Female	G3	II	T1	M0	N1	Adeno	White	No	No
TCGA-XP-A8T8	49	Male	G1	II	T1	M0	N1	SCC	Black	No	Yes
TCGA-2H-A9GH	44	Male	G2	II	T1	M0	N1	Adeno	Unknown	No	No
TCGA-VR-AA4D	53	Male	G2	II	T1	M0	N1	Adeno	White	Yes	No
TCGA-2H-A9GM	53	Male	G2	II	T1	M0	N1	Adeno	Unknown	Unknown	No
TCGA-IG-A51D	63	Male	G2	II	T1	M0	N1	SCC	White	Unknown	No
TCGA-IC-A6RE	59	Male	G2	II	T1	M0	N1	Adeno	White	Yes	Yes
TCGA-JY-A939	77	Male	G3	II	T1	M0	N1	Adeno	White	No	Yes
TCGA-L5-A4OS	86	Female	G3	II	T2	M0	N1	Adeno	White	Yes	Yes
TCGA-L5-A8NM	84	Female	Unknown	II	T2	M0	N1	Adeno	White	No	Yes
TCGA-L5-A4OW	56	Female	Unknown	II	T2	M0	N1	Adeno	Unknown	Yes	Yes
TCGA-IG-A6QS	54	Male	G1	II	T2	M0	N1	SCC	White	No	Yes
TCGA-LN-A5U6	54	Male	G2	II	T2	M0	N1	SCC	Asian	No	Yes
TCGA-LN-A9FR	70	Male	G2	II	T2	M0	N1	SCC	Asian	No	Yes
TCGA-XP-A8T6	54	Male	G2	II	T2	M0	N1	SCC	Black	No	Yes
TCGA-JY-A6FH	53	Male	G2	II	T2	M0	N1	Adeno	White	No	Yes
TCGA-IG-A4P3	48	Male	G2	II	T2	M0	N1	SCC	White	No	Yes
TCGA-JY-A6FA	51	Male	G2	II	T2	M0	N1	SCC	Asian	Yes	Yes
TCGA-L5-A88T	86	Male	Unknown	II	T2	M0	N1	Adeno	White	No	Yes
TCGA-L5-A8NE	77	Male	Unknown	II	T2	M0	N1	Adeno	White	No	Yes
TCGA-L5-A4OX	60	Male	Unknown	II	T2	M0	N1	Adeno	White	Yes	Yes
TCGA-L5-A8NW	55	Male	Unknown	Unknown	T2	M0	N1	Adeno	White	Yes	Yes
TCGA-L5-A8NR	81	Female	Unknown	III	T3	M0	N1	Adeno	Unknown	Yes	No
TCGA-LN-A4A3	61	Male	G2	III	T3	M0	N1	SCC	Asian	No	No
TCGA-2H-A9GG	66	Male	G2	III	T3	M0	N1	Adeno	Unknown	Unknown	No
TCGA-2H-A9GI	68	Male	G2	III	T3	M0	N1	Adeno	Unknown	Unknown	No
TCGA-LN-A7HY	50	Male	G2	III	T3	M0	N1	SCC	Asian	No	Yes
TCGA-VR-A8EQ	73	Male	G2	III	T3	M0	N1	Adeno	White	No	Yes
TCGA-LN-A4MQ	46	Male	G3	III	T3	M0	N1	SCC	Asian	No	No
TCGA-2H-A9GL	74	Male	G3	III	T3	M0	N1	Adeno	Unknown	No	No
TCGA-L5-A8NI	79	Male	G3	III	T3	M0	N1	Adeno	White	No	No
TCGA-2H-A9GK	43	Male	G3	III	T3	M0	N1	Adeno	Unknown	Yes	No
TCGA-2H-A9GF	67	Male	G3	III	T3	M0	N1	Adeno	Unknown	Unknown	No
TCGA-2H-A9GN	70	Male	G3	III	T3	M0	N1	Adeno	Unknown	Unknown	No
TCGA-2H-A9GQ	80	Male	G3	III	T3	M0	N1	Adeno	Unknown	Unknown	No
TCGA-LN-A49W	73	Male	G3	III	T3	M0	N1	SCC	Asian	No	Yes
TCGA-JY-A6FE	49	Male	G3	III	T3	M0	N1	SCC	White	Unknown	Yes
TCGA-L5-A8NG	77	Male	Unknown	III	T3	M0	N1	Adeno	White	Yes	No
TCGA-LN-A4A4	36	Male	Unknown	III	T3	M0	N1	SCC	Asian	No	Yes
TCGA-L5-A8NJ	77	Male	Unknown	III	T3	M0	N1	Adeno	White	No	Yes
TCGA-L5-A8NL	56	Male	Unknown	III	T3	M0	N1	Adeno	White	Yes	Yes
TCGA-L5-A8NN	81	Male	Unknown	III	T3	M0	N1	Adeno	White	Yes	Yes
TCGA-IG-A50L	58	Male	G2	III	T3	M0	N1	SCC	Asian	No	No
TCGA-IG-A3YB	61	Male	G2	III	T3	M0	N1	SCC	White	Unknown	No
TCGA-Z6-A9VB	53	Male	G2	III	T3	M0	N1	SCC	White	Unknown	Unknown
TCGA-IG-A3YC	62	Male	G2	III	T3	M0	N1	SCC	White	Unknown	Yes
TCGA-Z6-A8JE	57	Male	G3	III	T3	M0	N1	SCC	White	Unknown	No
TCGA-VR-A8Q7	60	Male	G3	III	T3	M0	N1	SCC	White	No	Yes
TCGA-JY-A93E	61	Male	G3	III	T3	M0	N1	Adeno	White	Yes	Yes
TCGA-L5-A891	51	Male	Unknown	Unknown	T3	M0	N1	Adeno	White	No	Yes
TCGA-VR-A8ER	54	Male	Unknown	III	T4	M0	N1	SCC	Unknown	No	Yes
TCGA-VR-A8EU	51	Male	G2	IV	T1	M1	N1	SCC	White	Yes	Yes
TCGA-LN-A5U5	57	Male	G2	IV	T3	M1	N1	SCC	Asian	No	Yes
TCGA-VR-A8EX	63	Male	G2	IV	T1	M1	N1	SCC	Unknown	No	No
TCGA-L5-A4OT	77	Male	Unknown	IV	T3	M1	N1	Adeno	White	Yes	Yes
TCGA-R6-A8W5	60	Male	G2	IV	T3	M1	N1	Adeno	White	Yes	No
TCGA-2H-A9GO	58	Male	G3	IV	T3	M1	N1	Adeno	Unknown	Unknown	No
TCGA-L5-A88V	60	Male	Unknown	III	T3	Unknown	N1	Adeno	White	Yes	Yes
TCGA-S8-A6BV	76	Male	G2	III	T3	Unknown	N1	Adeno	White	No	Yes
TCGA-V5-A7RB	59	Male	Unknown	Unknown	Unknown	Unknown	N1	Adeno	Unknown	Yes	Yes
TCGA-L5-A4ON	65	Male	Unknown	II	T1	Unknown	N1	Adeno	White	Yes	Yes
TCGA-VR-AA4G	51	Female	G2	III	T2	M0	N2	SCC	White	No	No
TCGA-Q9-A6FW	61	Male	G2	III	T3	M0	N2	Adeno	White	No	No
TCGA-IG-A625	60	Male	G2	III	T3	M0	N2	SCC	Asian	No	Yes
TCGA-VR-A8EW	57	Male	G2	III	T3	M0	N2	SCC	Asian	No	Yes
TCGA-IG-A4QS	71	Male	G3	III	T3	M0	N2	Adeno	White	No	No
TCGA-VR-A8EP	51	Male	G3	III	T3	M0	N2	SCC	White	No	No
TCGA-JY-A93C	47	Male	G3	III	T3	M0	N2	Adeno	Asian	No	Yes
TCGA-IG-A8O2	62	Male	G3	III	T3	M0	N2	SCC	Asian	No	Yes
TCGA-L5-A4OE	81	Male	G2	III	T3	Unknown	N2	Adeno	White	Unknown	Yes
TCGA-JY-A93D	51	Male	G3	III	T2	M0	N3	Adeno	White	Yes	Yes
TCGA-VR-A8EZ	47	Male	G2	III	T3	M0	N3	SCC	White	No	No
TCGA-L5-A4OO	75	Male	G2	III	T3	M0	N3	Adeno	White	Yes	Yes
TCGA-L7-A6VZ	62	Male	G2	III	T3	Unknown	N3	Adeno	White	Yes	Unknown
TCGA-L5-A4OI	79	Male	G2	III	T3	Unknown	N3	Adeno	White	No	Yes
TCGA-ZR-A9CJ	65	Male	G3	III	T3	Unknown	N3	Adeno	White	Unknown	No
TCGA-IG-A97H	36	Male	G3	II	T3	M0	Unknown	SCC	Asian	No	Yes
TCGA-IG-A7DP	50	Female	G2	III	T4	Unknown	Unknown	Adeno	White	No	No
TCGA-R6-A6DQ	74	Female									