

# Identification of a Novel Prognostic Model for Patients with Head and Neck Squamous Cell Carcinoma Based on Autophagy-related Genes

Xuefeng Zhang

National Clinical Research Center for Oral Diseases, Sichuan University

Wanchun Yang (✉ [yangwanchunscu@163.com](mailto:yangwanchunscu@163.com))

Sichuan University

---

## Research Article

**Keywords:** TCGA, HNSCC, survival, autophagy-related genes, prognostic signature

**Posted Date:** May 27th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1663506/v1>

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

[Read Full License](#)

---

# Abstract

**Background.** Head and neck squamous cell carcinoma (HNSCC) is one of the most common cancers in the world. Experimental evidence indicates that autophagy-related genes (ARGs) are associated with the initiation and progression of HNSCC. However, it's not well defined whether ARGs have the predictive ability of prognostic significance in patients with HNSCC. In this study, we aimed to identify critical ARGs associated with the prognosis of HNSCC, and establish a prognostic signature for the risk assessment.

**Methods.** Differential expression and functional enrichment analysis were conducted to analyze the datasets from The Cancer Genome Atlas (TCGA) and the Human Autophagy Database (HADb) with R software. The prognosis-related ARGs were screened by univariate Cox regression and LASSO regression analysis. The multivariate Cox regression analysis were applied to identify the prognosis-related ARGs signature of HNSCC.

**Results.** We established the prognostic signatures based on 9 ARGs by multivariate Cox regression analysis for the overall survival (OS) analysis. The following risk scoring was based on this 9-gene prognostic signature. Kaplan-Meier curves showed that high-risk score group had worse survival than low-risk score group. ROC curves revealed that the risk signature was a predictive indicator in the outcomes of patients with HNSCC. Moreover, nomogram in the prognosis of HNSCC patients confirmed the findings that the predicted probabilities and OS rates at 1-, 3- and 5- years were relatively consistent.

**Conclusion.** In this study, we identified a novel ARG-based prediction signature, which could predict the prognosis of patients with HNSCC.

## Introduction

Originating from epithelial cells of the oral tongue, cavity, pharynx, nasal cavity, and larynx, head and neck squamous cell carcinoma (HNSCC) is one of the most common and lethal cancers, accounting for 5–10% of all human cancers [1, 2]. Despite advanced treatments by surgery, radiotherapy, and combined chemotherapy [3, 4], the 5-year survival rate for HNSCC remains ~ 60%, mainly due to the local recurrence and distant metastases [5, 6]. Numerous studies have identified the risk factors that induce HNSCC, including human papillomavirus (HPV) infection, smoking, alcohol abuse, and a variety of environmental factors and genetic mutations [7, 8]. Current studies on HNSCC are enriched in molecular targets exploration and novel treatment strategies [9]. However, to our knowledge, the prognostic prediction of HNSCC is still limited. Therefore, the pressing question remains developing specific and sensitive biomarkers for prognostic prediction.

Autophagy is a catabolic process in mammalian cells that responds to extra- and intra-cellular stress. In autophagy, cellular components are degraded in lysosomes to provide the material and energy sources [10, 11]. Emerging evidence has established that abnormalities in autophagy are closely associated with a variety of human diseases, including cancers [12]. Autophagy plays a dual role in regulating the fate of cancer cells, both promoting and suppressing tumorigenesis in different types and stages of cancer [13,

14]. Regulation of autophagy-related genes (ARGs) is thought to be the central mechanism by which autophagy exerts its function in human cancers [15, 16]. Therefore, the expression profiles of ARGs can provide the opportunity to build prognostic models for human cancers, including HNSCC.

In this study, we investigated the ARGs in HNSCC by exploring TCGA (The Cancer Genome Atlas) database, and applied interrogated bioinformatics approaches to screen genes and identified prognostic signatures associated with the survival of HNSCC patients. Our work described the prognostic values of ARGs in HNSCC, and provided novel targets for the improvement of HNSCC prognosis.

## **Materials & Methods**

### **Data acquisition**

A total of 232 autophagy-related genes (ARGs) were obtained from the Human Autophagy Database (HADb, <http://www.autophagy.lu/index.html>), an online database of a complete set of human coding genes related to autophagy. The RNA-seq data of tumor samples of HNSCC were obtained from TCGA portal (<https://tcga-data.nci.nih.gov/tcga/>), which contains 502 tumor samples and 44 adjacent non-tumor samples. The clinical characteristics and survival information of HNSCC patients were also collected.

### **Screening of differentially expressed ARGs and enrichment analysis**

Data analysis of differential expression between tumor tissue samples and their corresponding normal tissue samples downloaded from TCGA was performed using the "limma" package in R (version 4.0.3). Differential genes were obtained with a threshold of  $|\log_2(\text{FC})| > 2$  and a P value  $< 0.05$ . The differential genes were intersected with autophagy genes to obtain differentially expressed ARGs for subsequent analysis. For gene functional analysis, the Gene Ontology (GO) analysis was conducted using the "clusterProfiler" package to identify the main biological properties of differentially expressed ARGs. The results of the annotation analysis were then visualized using the "ggplot2" package for GO enrichment plots. Gene Ontology (GO) analysis consists of three main categories, namely biological processes (BP), molecular functions (MF), and cellular components (CC).

### **Construction of prognostic model**

To establish the prognostic model using ARGs in HNSCC, the univariate Cox regression analysis was performed to screen for ARGs significantly associated with overall survival (OS) under a threshold of  $P < 0.05$ . Next, LASSO Cox regression analysis and multivariate Cox regression analysis were performed to establish the ARGs prognostic signature. The prognostic value of the ARGs prognostic signature was assessed by risk scores. The Kaplan-Meier survival curve was plotted to assess the efficacy of the risk signature, comparing the high-risk score group and low-risk score group based on the predictive characteristics. The predictive value of the prognostic signature was then assessed using the areas under

the curve (AUC) of the receiver-operator characteristic (ROC) curve by the "timeROC" package in R software. A nomogram was created with the "rms" package and the "survival" package to predict the OS for HNSCC patients. The calibration curves of the nomogram were plotted to evaluate the gap between the nomogram and actual incidence.

## Statistical analysis

All statistical analysis in this work were performed by the R software (version 4.0.3). Results were considered to be of statistical significance with the p value < 0.05.

## Results

### Identification of differentially expressed ARGs in HNSCC.

As a first step toward understanding the prognostic value of autophagy in HNSCC, we screened the differentially expressed genes related to autophagy in HNSCC. As presented by the Venn diagram, among 10,342 differentiated expression genes in HNSCC, there were 104 autophagy-related genes (ARGs), including 69 up-regulated and 35 down-regulated genes (Fig. 1A).

To further figure out the biological functions of these 104 genes, we performed Gene Ontology (GO) enrichment analysis. Results showed that the differentially expressed ARGs were mainly enriched in biological processes (BP) of macroautophagy/autophagy (Fig. 1B). In cellular components (CC), the top enriched terms were autophagosome assembly (Fig. 1C). In molecular functions (MF), the top ranked function terms were cysteine-type endopeptidase activities and ubiquitin-like protein ligase bindings (Fig. 1D).

### Construction of the prognostic models using ARGs in HNSCC.

We assessed the prognostic value of above 104 ARGs using univariate Cox regression, and found 21 ARGs to be significantly associated with the prognosis of HNSCC patients. Based on these 21 ARGs, we used LASSO Cox regression analysis ("lars" package and "glmnet" package in R software) to screen out 14 prognosis-related genes. Next, we employed the multivariate Cox regression analysis to finally identify 9 most prognostically valuable ARGs, including ATG4A, RAF1, BAG3, RAB11A, MAP2K7, STK11, ATIC, FADD and USP10. The risk scores were then calculated based on the gene expression levels and the associated Cox regression coefficients, as risk score =  $\text{expr MAP2K7} \times -0.1020 + \text{expr STK11} \times -0.0592 + \text{expr FADD} \times 0.0109 + \text{expr USP10} \times 0.0248 + \text{expr RAB11A} \times 0.0149 + \text{expr RAF1} \times -0.0496 + \text{expr ATIC} \times 0.0328 + \text{expr ATG4A} \times 0.0989 + \text{expr BAG3} \times 0.0087$  (Fig. 2A-B).

Then we classified HNSCC patients into high- and low-risk score groups based on the median of the OS-related prediction signature. The Kaplan-Meier survival curves clearly showed that the low-risk group had lower mortality and better prognosis than that in the high-risk score group ( $p < 0.001$ ) (Fig. 2C). Additionally, the ROC curves for the overall survival (OS) were applied to reveal the predictive performance

of the 9-gene based risk signature. The AUC values of ROC for 1-, 3-, and 5- years were 0.685, 0.78, and 0.749, respectively, indicating an excellent predictive value of the signature (Fig. 2D).

### **Identification of the ARGs prognostic signature in HNSCC.**

To further evaluate the ARGs prognostic signature, we established the risk factor association map by the "ggplot2" package and the "pheatmap" package in R software, and analyzed the distribution of risk scores, survival time and correlated ARG expression for HNSCC patients. As shown in the association map, the risk scores of patients increased in parallel with decreased survival time and increased numbers of death (Fig. 3A-B).

Next, we analyzed the expression profiles of the 9 ARGs including ATG4A, RAF1, BAG3, RAB11A, MAP2K7, STK11, ATIC, FADD and USP10. It's noted that MAP2K7, STK11 and RAF1 were highly expressed in the group of low-risk scores, whereas genes including RAB11A, FADD, BAG3 and USP10 were in the group of high-risk scores (Fig. 3C).

To further confirm the effect of the 9 ARGs on the prognosis of HNSCC, we performed Kaplan-Meier logarithm test to assess the OS difference between the different ARGs expressions. Results showed that low expression levels of MAP2K7, STK11 and RAF1 were tightly correlated with the poor survival time in HNSCC patients, whereas low expression levels of RAB11A, FADD, BAG3 and USP10 were closely associated with better prognosis (Fig. 4). All the survival analysis results were in accordance with the gene expression patterns with different risk-scores.

### **Establishment of the nomogram prognostic model in HNSCC.**

To provide better prognosis for HNSCC patients based on our risk score model, we conducted a nomogram to predict the OS of 1-, 3-, and 5-years using five prognostic factors, including gender, stage, grade, age, and risk scores (Fig. 5). The calibration curves showed that the predicted calibration curves were close to the standard curves, suggesting that the nomogram could be reliable in the prediction of OS in HNSCC patients.

## **Discussion**

HNSCC represents the 6th most prevalent cancer worldwide, with malignant behaviors and low survival rates [17]. One of the major reasons for the poor prognosis of HNSCC was due to the fact that ~ 25% of patients developed tumor recurrence within 5 years after primary diagnosis [18]. However, clinical diagnosis of HNSCC is often made at an intermediate or even late stage, which inevitably leads to a poor prognosis. Therefore, there is an urgent need to identify reliable and effective biomarkers in the diagnosis, treatment, and prognostic assessment of HNSCC.

Currently, clinical and fundamental studies have demonstrated a variety of risk factors and gene mutation/pathways in the initiation and progression of HNSCC [19]. But the prognostic determinant of HNSCC is still limited. In this study, we developed a risk score model using 9 autophagy-related genes to

predict the prognosis of HNSCC patients. The risk scores based on prognostic models could be applied clinically to provide convenient and better prognostic monitoring.

At the molecular level, experimental evidence has highlighted the importance of autophagy in HNSCC. Autophagy is an essential catabolic process in cells, which degrades dysfunctional/impaired organelles and macromolecules, and thus maintains the cellular homeostasis [20]. There are accumulating evidence suggesting the role of autophagy-mediated cell survival or progression in HNSCC [21]. In the early stages of tumor formation, autophagy is considered to be anti-cancer and protect normal cells from transforming into malignant tumor cells. When tumor progressed, the autophagy in tumor cells may protect tumor cells from stress-induced death, and contribute to radio- and chemo-resistance. Therefore, autophagy itself is considered as a double-edged sword in the development of HNSCC. Here, we focused on the differentially expressed ARGs in HNSCC, and established a prognostic model using 9 selected ARGs. We noted that several reports have identified prognostic models in HNSCC using immune-related genes [22] and pyroptosis-related genes [23]. Furthermore, another pilot studies also demonstrated the ARG signature in the prediction of prognosis of patients with HNSCC [24–27]. Our work supplements current understandings of the prognostic signature of HNSCC, and reveals that autophagy/ARGs may play a role in the development of HNSCC.

It's noted that most of our identified ARGs (ATG4A, RAF1, BAG3, RAB11A, MAP2K7, STK11, ATIC, FADD and USP10) had been reported to be involved in HNSCC by either bioinformatic or experimental studies. For example, mutations of STK11 gene produced functionally inactive proteins, and contributed to the development of HNSCC [28]. FADD, an adaptor molecule for death receptor-mediated apoptosis, regulates cell proliferation and has a role in predicting long-term outcomes in patients with HNSCC [29, 30]. Our study provides integrative evidence to demonstrate that the 9 ARGs identified had abilities to predict the prognosis of HNSCC. Of note, the 9 genes were divided into two subgroups, one is negative associated with risk scores (e.g., MAP2K7, STK11 and RAF1) and the other is positive associated with risk scores (e.g., RAB11A, FADD, BAG3 and USP10). This notion was consistent with the OS curves of these ARGs in HNSCC.

There are some limitations in this study. We have not provided experimental evidence to further verify our results. Moreover, future prospective studies with large-sized cohorts and studies are needed to confirm the predictive effect and clinical value of our model.

## Conclusion

In conclusion, we performed comprehensive bioinformatics analysis to investigate the role and prognostic value of ARGs for HNSCC. We identified a risk signature based on 9 ARGs that can classify HNSCC patients into subgroups with different prognosis. Our study provides a novel insight into the autophagic status of HNSCC, and affords the potential application of this signature in clinical predictions.

# Declarations

## Ethical declarations

All gene expression datasets were collected from TCGA database. This study was approved by the Ethical Committee in the West China Hospital of Stomatology of Sichuan University and conducted according to the principles expressed in the Declaration of Helsinki.

## Authors' contribution

Study design: Xuefeng Zhang and Wanchun Yang. Data retrieve: Xuefeng Zhang. Statistical analysis: Xuefeng Zhang. Result interpretation: Xuefeng Zhang and Wanchun Yang. Writing: Xuefeng Zhang and Wanchun Yang. Funding acquisition: Xuefeng Zhang.

## Conflict of interest

The authors have declared no conflicts of interest.

## Availability of Data and Materials

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

## Funding

This work was supported by The National Natural Science Foundation of China (Grant Nos. 82002877). This fund contributed to the design of the study, collection, analysis, and interpretation of data, and writing the manuscript.

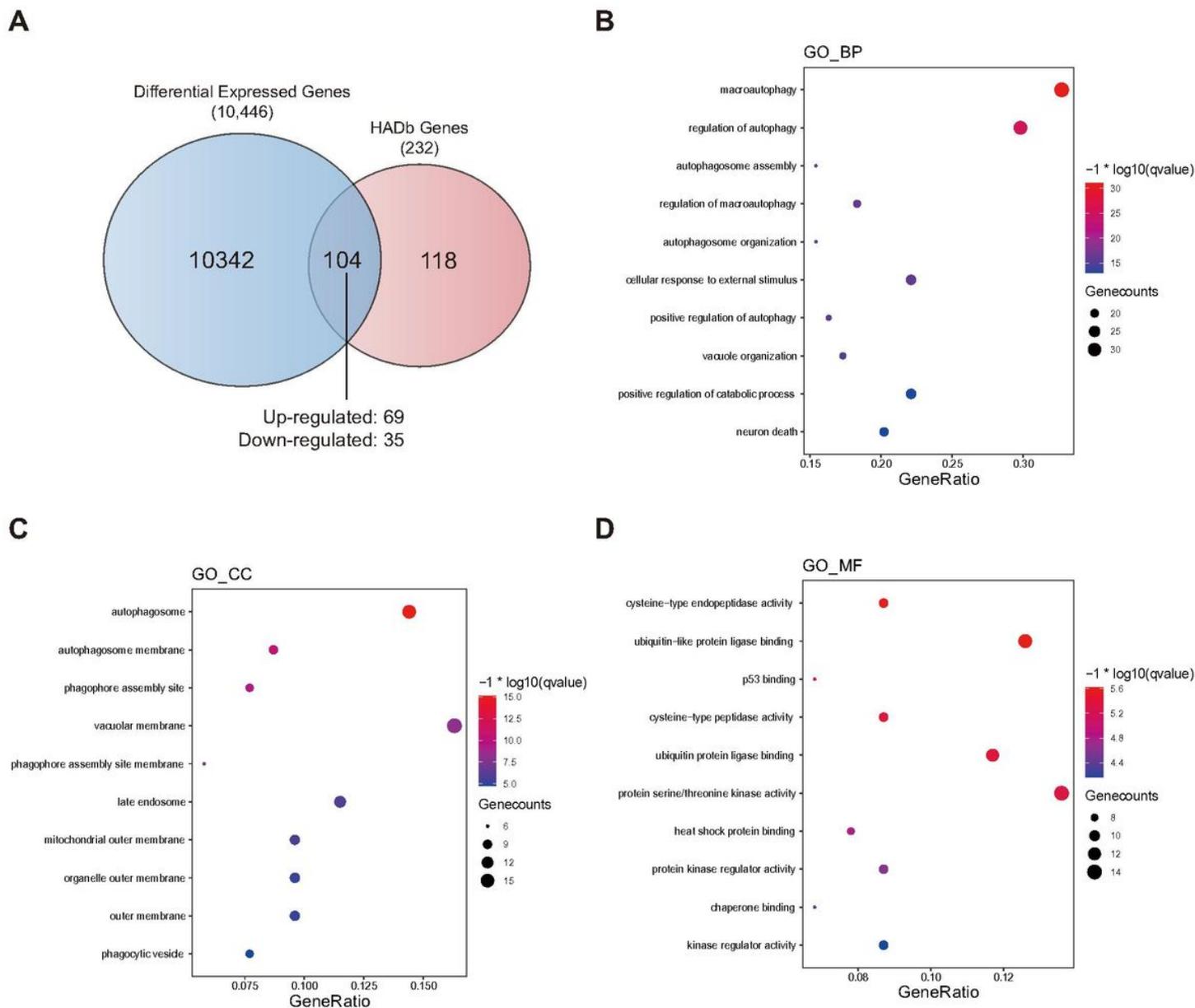
# References

1. Johnson DE, Burtneß B, Leemans CR, Lui VWY, Bauman JE, Grandis JR: **Head and neck squamous cell carcinoma**. Nat Rev Dis Primers 2020, **6**(1):92.
2. Marur S, Forastiere AA: **Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment**. Mayo Clin Proc 2016, **91**(3):386–396.
3. Gougis P, Moreau Bachelard C, Kamal M, Gan HK, Borcoman E, Torossian N, Bieche I, Le Tourneau C: **Clinical Development of Molecular Targeted Therapy in Head and Neck Squamous Cell Carcinoma**. JNCI Cancer Spectr 2019, **3**(4):pkz055.
4. Polverini PJ, D'Silva NJ, Lei YL: **Precision Therapy of Head and Neck Squamous Cell Carcinoma**. J Dent Res 2018, **97**(6):614–621.
5. Citron F, Armenia J, Franchin G, Polesel J, Talamini R, D'Andrea S, Sulfaro S, Croce CM, Klement W, Otasek D *et al*: **An Integrated Approach Identifies Mediators of Local Recurrence in Head and Neck Squamous Carcinoma**. Clin Cancer Res 2017, **23**(14):3769–3780.

6. Duprez F, Berwouts D, De Neve W, Bonte K, Boterberg T, Deron P, Huvenne W, Rottey S, Mareel M: **Distant metastases in head and neck cancer.** *Head Neck* 2017, **39**(9):1733–1743.
7. Gingerich MA, Smith JD, Michmerhuizen NL, Ludwig M, Devenport S, Matovina C, Brenner C, Chinn SB: **Comprehensive review of genetic factors contributing to head and neck squamous cell carcinoma development in low-risk, nontraditional patients.** *Head Neck* 2018, **40**(5):943–954.
8. Benson E, Li R, Eisele D, Fakhry C: **The clinical impact of HPV tumor status upon head and neck squamous cell carcinomas.** *Oral Oncol* 2014, **50**(6):565–574.
9. Leemans CR, Braakhuis BJ, Brakenhoff RH: **The molecular biology of head and neck cancer.** *Nat Rev Cancer* 2011, **11**(1):9–22.
10. Klionsky DJ, Petroni G, Amaravadi RK, Baehrecke EH, Ballabio A, Boya P, Bravo-San Pedro JM, Cadwell K, Cecconi F, Choi AMK *et al*: **Autophagy in major human diseases.** *EMBO J* 2021, **40**(19):e108863.
11. Ktistakis NT, Tooze SA: **Digesting the Expanding Mechanisms of Autophagy.** *Trends Cell Biol* 2016, **26**(8):624–635.
12. Levy JMM, Towers CG, Thorburn A: **Targeting autophagy in cancer.** *Nat Rev Cancer* 2017, **17**(9):528–542.
13. Miller DR, Thorburn A: **Autophagy and organelle homeostasis in cancer.** *Dev Cell* 2021, **56**(7):906–918.
14. Guo JY, Xia B, White E: **Autophagy-mediated tumor promotion.** *Cell* 2013, **155**(6):1216–1219.
15. Levine B, Kroemer G: **Biological Functions of Autophagy Genes: A Disease Perspective.** *Cell* 2019, **176**(1–2):11–42.
16. Li X, He S, Ma B: **Autophagy and autophagy-related proteins in cancer.** *Mol Cancer* 2020, **19**(1):12.
17. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.** *CA Cancer J Clin* 2018, **68**(6):394–424.
18. Chuang SC, Scelo G, Tonita JM, Tamaro S, Jonasson JG, Kliewer EV, Hemminki K, Weiderpass E, Pukkala E, Tracey E *et al*: **Risk of second primary cancer among patients with head and neck cancers: A pooled analysis of 13 cancer registries.** *Int J Cancer* 2008, **123**(10):2390–2396.
19. Blise KE, Sivagnanam S, Banik GL, Coussens LM, Goecks J: **Single-cell spatial architectures associated with clinical outcome in head and neck squamous cell carcinoma.** *NPJ Precis Oncol* 2022, **6**(1):10.
20. Mizushima N: **A brief history of autophagy from cell biology to physiology and disease.** *Nat Cell Biol* 2018, **20**(5):521–527.
21. Sannigrahi MK, Singh V, Sharma R, Panda NK, Khullar M: **Role of autophagy in head and neck cancer and therapeutic resistance.** *Oral Dis* 2015, **21**(3):283–291.
22. Jiang P, Li Y, Xu Z, He S: **A signature of 17 immune-related gene pairs predicts prognosis and immune status in HNSCC patients.** *Transl Oncol* 2021, **14**(1):100924.

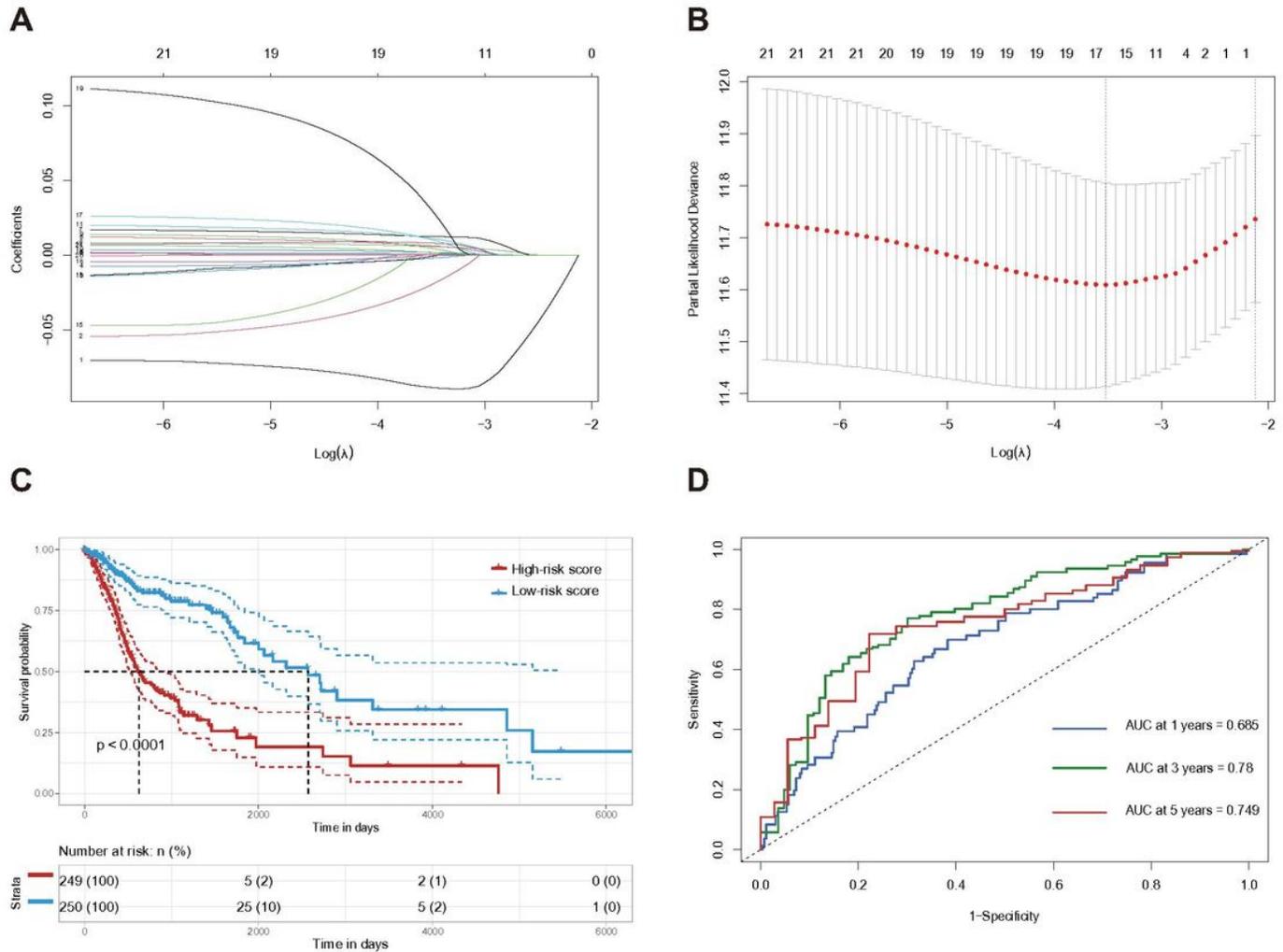
23. Shen Y, Li X, Wang D, Zhang L, Li X, Xia T, Shang X, Yang X, Su L, Fan X: **Novel prognostic model established for patients with head and neck squamous cell carcinoma based on pyroptosis-related genes.** *Transl Oncol* 2021, **14**(12):101233.
24. Li Q, Wang J, Meng X, Chen W, Feng J, Mao J: **Identification of autophagy-related gene and lncRNA signatures in the prognosis of HNSCC.** *Oral Dis* 2021.
25. Ren Z, Zhang L, Ding W, Luo Y, Shi Z, Shrestha B, Kan X, Zhang Z, Ding J, He H *et al*: **Development and validation of a novel survival model for head and neck squamous cell carcinoma based on autophagy-related genes.** *Genomics* 2021, **113**(1 Pt 2):1166–1175.
26. Feng H, Zhong L, Yang X, Wan Q, Pei X, Wang J: **Development and validation of prognostic index based on autophagy-related genes in patient with head and neck squamous cell carcinoma.** *Cell Death Discov* 2020, **6**:59.
27. Li C, Wu ZH, Yuan K: **Autophagy-Related Signature for Head and Neck Squamous Cell Carcinoma.** *Dis Markers* 2020, **2020**:8899337.
28. Qiu W, Schonleben F, Thaker HM, Goggins M, Su GH: **A novel mutation of STK11/LKB1 gene leads to the loss of cell growth inhibition in head and neck squamous cell carcinoma.** *Oncogene* 2006, **25**(20):2937–2942.
29. Pattje WJ, Melchers LJ, Slagter-Menkema L, Mastik MF, Schrijvers ML, Gibcus JH, Kluin PM, Hoegen-Chouvalova O, van der Laan BF, Roodenburg JL *et al*: **FADD expression is associated with regional and distant metastasis in squamous cell carcinoma of the head and neck.** *Histopathology* 2013, **63**(2):263–270.
30. Rasamny JJ, Allak A, Krook KA, Jo VY, Policarpio-Nicolas ML, Sumner HM, Moskaluk CA, Frierson HF, Jr., Jameson MJ: **Cyclin D1 and FADD as biomarkers in head and neck squamous cell carcinoma.** *Otolaryngol Head Neck Surg* 2012, **146**(6):923–931.

## Figures



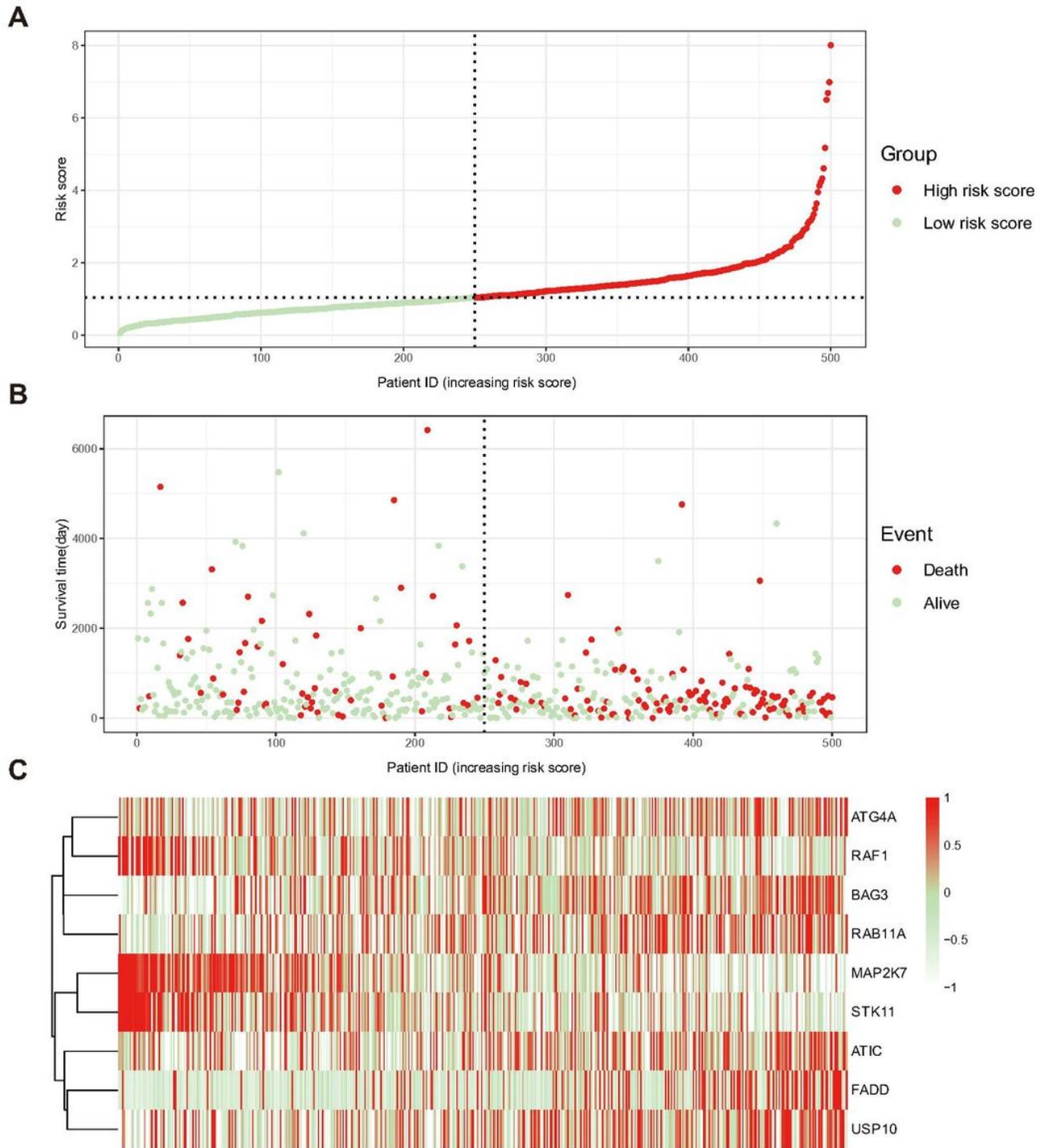
**Figure 1**

**Identification of differentially expressed ARGs in HNSCC. (A)** Venn diagram showing the differentially expressed autophagy-related genes (ARGs) in HNSCC from TCGA dataset. **(B-D)** Results of GO analysis showing altered ARGs profiles in TCGA dataset, including biological process (B), cellular component (C) and molecular function (D).



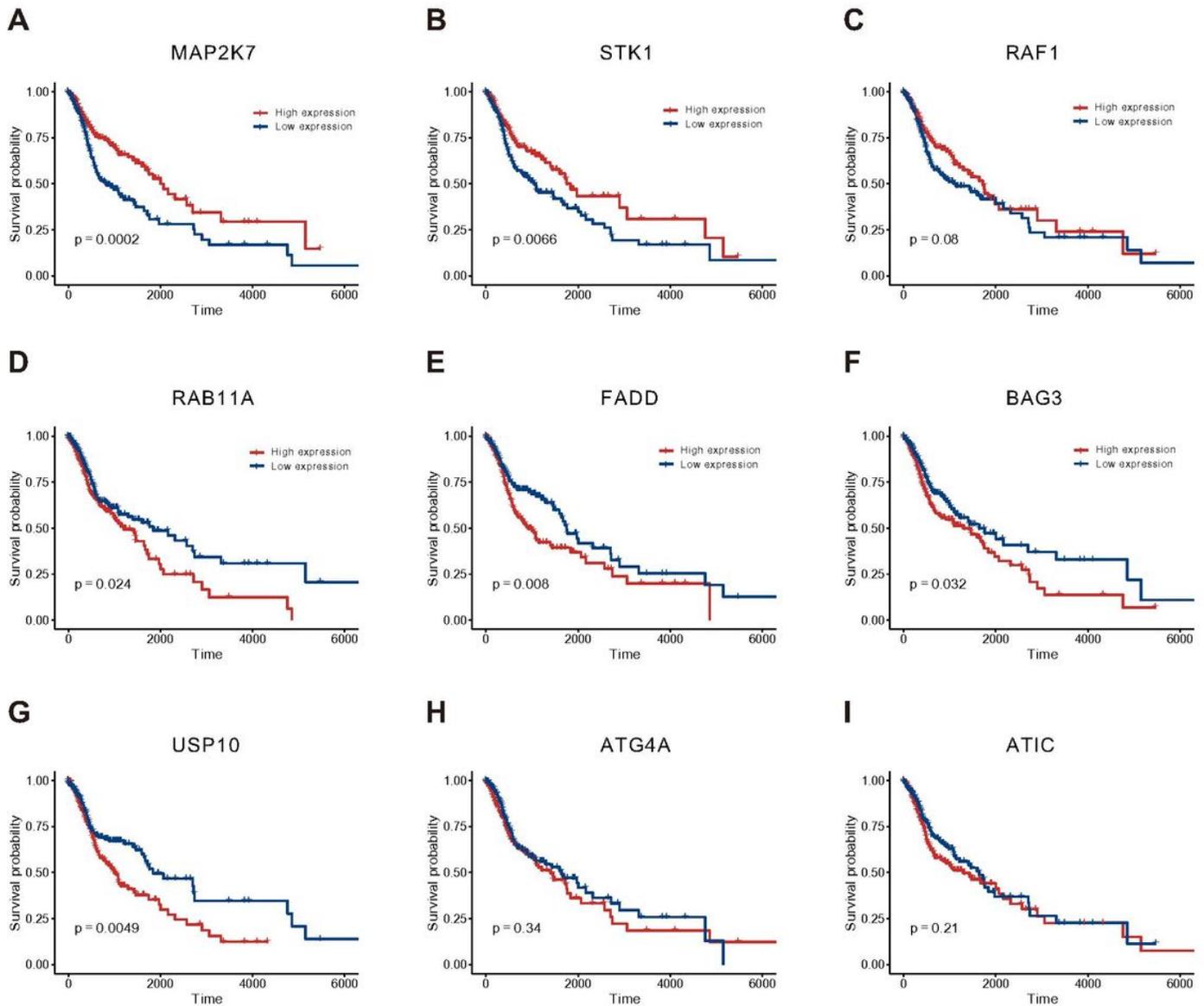
**Figure 2**

**Construction of the risk score based ARGs signature for OS in HNSCC patients.** (A) The LASSO regression analyses of ARGs-related genes. (B) Cross-validation for tuning the parameter selection in the LASSO regression. (C) Kaplan–Meier analysis of HNSCC patients was stratified by median risk in TCGA. High risk scores are associated with poor survival. (D) ROC curves of risk score in prognosis prediction of HNSCC.



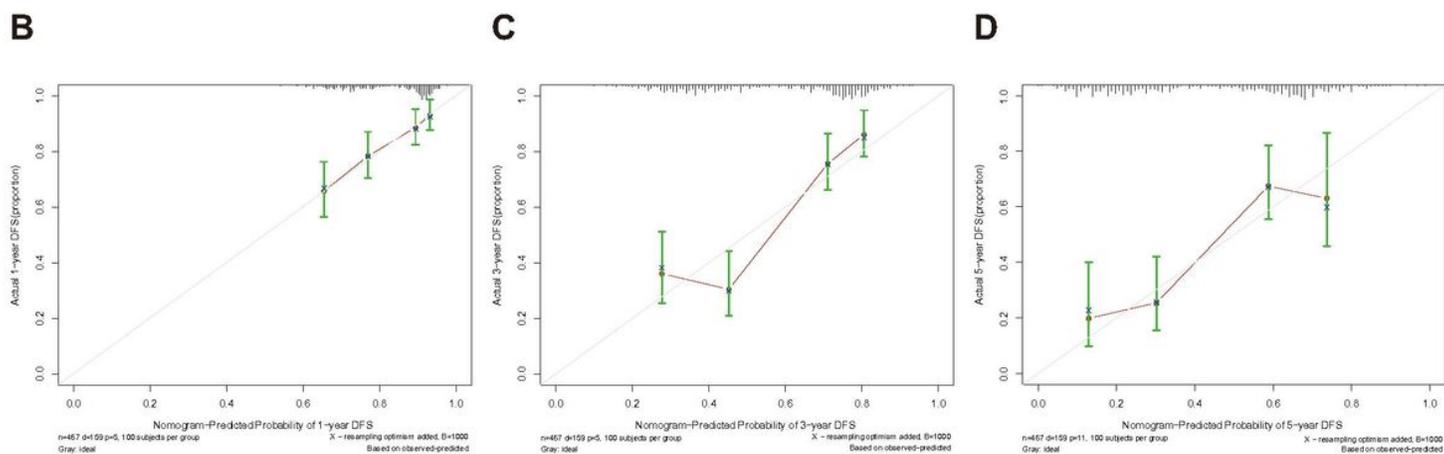
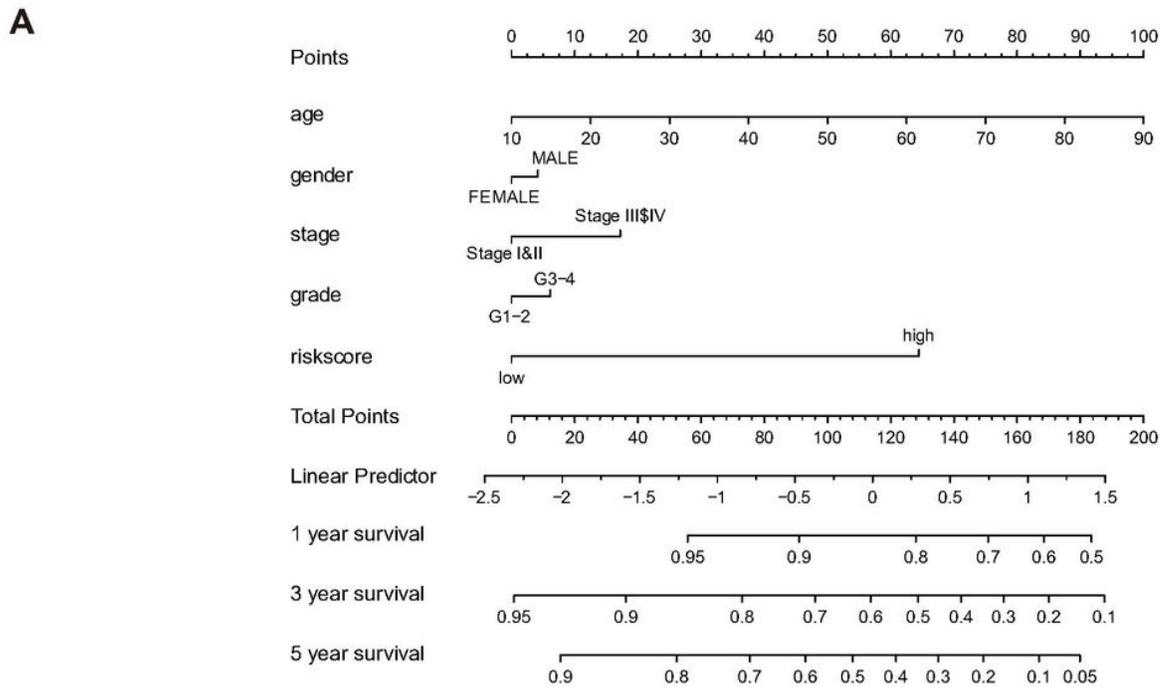
**Figure 3**

**Identification of risk score based ARGs signature of patients with HNSCC. (A-B)** Risk plot of each point sorted based on risk score, representing one patient. Green and red points represent patients with low- and high-risk, respectively. **(C)** Distribution of risk score and ARG expression of HNSCC patients in TCGA.



**Figure 4**

**The correlation of 9 ARGs with the survival time in patients with HNSCC. (A-I)** Kaplan–Meier survival analysis of screened 9 ARGs with HNSCC using the information from TCGA dataset. Patients are divided into low and high groups by median each gene expression level.



**Figure 5**

**Construction a nomogram to predict the patient survival in HNSCC. (A)** The nomogram using age, gender, stage, grade, and risk score to predict the OS of patients with HNSCC. **(B-D)** The calibration plot to evaluate the nomogram. Y-axis, actual survival. X-axis, predicted survival of 1-year (B), 3- year (C), and 5- year (D), respectively.