An Autophagy-related Long Non-coding RNA Prognostic Signature for Laryngeal Squamous Cell Carcinoma

Yujie Shen
Fudan University Eye Ear Nose and Throat Hospital

Huiying Huang
Fudan University Eye Ear Nose and Throat Hospital

Qiang Huang
Fudan University Eye Ear Nose and Throat Hospital

liang zhou (zhoulent@126.com)
Fudan University Eye Ear Nose and Throat Hospital

Research article

Keywords: long noncoding RNA (lncRNA), autophagy, laryngeal squamous cell carcinoma (LSCC), TCGA, prognostic signature

DOI: https://doi.org/10.21203/rs.3.rs-484074/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

**Background:** Laryngeal squamous cell carcinoma (LSCC) is the second most common malignant tumor in the head and neck. Considering the role of autophagy in tumor development and drug resistance, we investigated the potential prognostic value of autophagy-related long no-coding RNAs (lncRNAs) in LSCC patients.

**Methods:** Autophagy-related lncRNAs were screened out based on the Cancer Genome Atlas (TCGA) database. Subsequently, five autophagy-related lncRNAs with prognostic value were identified through univariate and multivariate Cox regression analysis. Based on these prognosis-related lncRNAs, a risk signature was established, and the patients with LSCC were divided into the low- and high-risk groups. Finally, a nomogram based on the prognosis signature was constructed.

**Results:** The overall survival time of the high-risk group was significantly shorter than that of the low-risk one (p<0.0001). Receiver operating characteristic curve (ROC) analysis was performed to further confirm the validity of the model (ROC of training set: 0.841, ROC of test set: 0.718, ROC of entire set: 0.757). Univariate and multivariate regression analyses demonstrated that ours is an independent prognosis signature of LSCC. The nomogram integrating the five-lncRNA signature and clinical factors was then constructed for clinical application.

**Conclusion:** In summary, our signature of the five autophagy-associated lncRNAs can serve as an effective prognostic indicator for patients with LSCC.

1. Background

Laryngeal squamous cell carcinoma (LSCC), a common head and neck tumor, is showing an ascending incidence in recent years[1]. Several risk factors are etiologically involved in LSCC, among which tobacco and alcohol are the most significant[2, 3]. Although early treatment is often accompanied with favorable prognosis, statistic data have shown that at the time of diagnosis, nearly 60% of LSCC patients are already in the advanced stage (III or IV)[4]. With a strong tendency of local invasion and cervical lymph node metastasis, the advanced tumor always leads to poor prognosis[5, 6]. As the significance of molecular medicine in tumor therapy increases, the existing tumor-node-metastasis (TNM) staging system appears to be limited in assessing prognosis of LSCC. Therefore, exploring new risk signatures to evaluate the prognosis of LSCC patients is of great importance.

Autophagy, a biological process to degrade damaged organelles and recycle cellular components, can either suppress or promote tumor cells [7]. Despite its inhibitory effect in the early stage of tumor formation, autophagy contributes to tumor growth and metastasis in advanced malignancies[8, 9]. The targeted inhibition of autophagy-related pathway has been shown to exert curative effect on many tumors[10–12]. The anti-tumor effect of autophagy suppression has also been reported in LSCC, indicating its crucial function in the progression of this disease[13, 14].
Long non-coding RNAs (lncRNAs) are a class of RNAs that perform a variety of biological functions without coding any proteins. They can participate in multiple biological processes such as cell proliferation, apoptosis, and autophagy, by interacting with DNAs, RNAs, and proteins[15, 16]. Emerging evidence has implicated that lncRNAs can affect various autophagy-related genes, and thus regulate cell autophagy in many diseases[17–19]. A recent study of bladder cancer has suggested the prognostic potential of autophagy-associated lncRNAs[20]. A latest research about cervical cancer also emphasizes the diagnostic and prognostic value of lncRNAs[21]. However, the regulatory relationship between lncRNAs and cell autophagy, as well as the clinical significance of autophagy-related lncRNAs in the prognosis of LSCC remains to be further explored.

Given the above, some specific autophagy-related lncRNAs may be important in evaluating the prognosis of patients with LSCC, or even provide potential targets for gene therapy. In this study, autophagy-related lncRNAs were identified using The Cancer Genome Atlas (TCGA) database, and a novel prognosis signature for LSCC was constructed based on autophagy-related lncRNAs, aiming to explore new indicators for the prognosis of LSCC patients.

2 Materials And Methods

2.1 Data collection and processing

The available RNA-seq data and associated clinical information of 111 LSCC tumor samples and 12 matched normal tissues were downloaded from the TCGA database (https://tcga-data.nci.nih.gov/tcga/) [22].

2.2 Dataset processing and identification of differentially expressed lncRNAs

The differentially expressed lncRNAs between LSCC and normal tissues were all normalized by the “limma” package in R software (version: x64 3.6.1)[23], with screening criteria of | log (FC) | ≥ 0.5 and adj.p < 0.05. Pearson correlation was applied to calculate the correlation between the lncRNAs and autophagy-related mRNAs. A lncRNA with a correlation coefficient > 0.3 and P < 0.001 was considered to be autophagy-related [24].

2.3 Construction of the prognostic signature and risk score calculation

We randomly divided 111 samples from TCGA into the training set and the test set with a ratio of 1:1 using the “caret” package in R software. Through the univariate Cox regression analysis, we firstly screened lncRNAs with p < 0.05 in the training set. Multivariate Cox regression analysis was subsequently used to select autophagy-related lncRNAs with prognostic value and then the prognostic signature was constructed. The lncRNAs with a p value < 0.05 were included to establish the risk score. We calculated the risk score for each LSCC patient using the following formula: Risk score = \( \beta_{\text{gene1}} \times \exp_{\text{gene1}} + \beta_{\text{gene2}} \times \exp_{\text{gene2}} \).
\[ \exp_{\text{gene}_2} + \ldots + \beta_{\text{gene}_n} \times \exp_{\text{gene}_n}. \] Patients with LSCC were divided into the high- and low-risk groups according to their median risk score. The co-expression network between the autophagy-related mRNAs and the IncRNAs included in the prognosis signature was constructed and visualized using Cytoscape v3.7.1 software[25].

### 2.4 Analysis of the independence of the prognostic signature

Kaplan-Meier survival analysis was used to compare the overall survival rate between the high- and low-risk groups. The time-dependent receiver operating characteristic (ROC) analyses were performed using “survivalROC” packages to evaluate the sensitivity and independence of the signature. These analyses were conducted in training set, test set and entire set. The independence of IncRNA signature for the overall survival (OS) of LSCC patients was further explored by univariate and multivariate analyses in the entire set.

### 2.5 Establishment and assessment of nomogram

Using the “rms” package of R software, a novel nomogram incorporating the IncRNA signature and other clinical factors was established using univariate and multivariate analyses. Calibration curves were conducted to evaluate whether the predicted survival by the nomogram was consistent with the actual survival[26]. The predictive performance of the prognostic model was also evaluated using area under the curve (AUC) in the ROC analysis[27].

### 2.6 Statistical analysis

Based on the critical conditions of \(|\log(FC)| \geq 0.5\) and \(\text{adj.} \ p < 0.05\), differently expressed IncRNAs were identified using the “edgeR” package in R software. Univariate and multivariate analyses were conducted to evaluate survival. OS was analyzed using the Kaplan–Meier method and the log-rank test was performed to assess the statistically significant differences between the high-risk and low-risk groups. To explore the predictive accuracy of the five-IncRNA signature and the prognostic nomogram, time-dependent ROC analysis was performed by the “survivalROC” package in R software, and calibration curve by “rms” package. All data were processed and analyzed using Perl software (version 5.30.1) and R software (version: x64 3.6.1). A \(p\) value \(< 0.05\) was considered statistically significant.

### 3. Results

#### 3.1 Establishment of a co-expression network for autophagy-related IncRNAs.

The research design is illustrated in Fig. 1. A total of 123 samples were extracted from the TCGA database, among which 111 were laryngeal cancer tissues and 12 were matched normal tissues. Differentially expressed IncRNAs (DEIncRNAs) were screened according to the screening criteria \((|\log(FC)| \geq 0.5, \text{adj.} \ p < 0.05)\). Totally, 1683 downregulated IncRNAs and 890 upregulated IncRNAs were
illustrated using a volcano map and the top 20 up-regulated and down-regulated DElncRNAs were shown in a heatmap (Fig. 2). Finally, we extracted 372 autophagy-related IncRNAs based on the filtering criteria of correlation score < 0.3 and \( p \) value < 0.001.

### 3.2 Construction of a prognostic model based on autophagy-related IncRNAs in LSCC patients.

The 111 tumor samples in TCGA database were randomly divided into a training set (\( n = 56 \)) and a test set (\( n = 55 \)). Further assessment combined with survival time was carried out in the training set by univariate Cox regression analysis, and seven autophagy-related IncRNAs were identified (Table 1). Then five IncRNAs (NCK1-DT, PTOV1-AS2, AC012640.2, AC023310.4 and AL513318.2) were screened out using multivariate Cox regression analysis for further analysis (Table 2). The risk score formula of the prognosis signature was as follows:

\[
\text{Risk score} = (-0.814837805 \exp \text{NCK1-DT}) + (-1.558931361 \exp \text{PTOV1-AS2}) + (0.979350461 \exp \text{AC012640.2}) + (0.751261346 \exp \text{AC023310.4}) + (0.394548801 \exp \text{AL513318.2})
\]

Subsequently, we verified the connection between these five IncRNAs and the prognosis of LSCC patients through a survival curve (\( p < 0.05 \)). AL513318.2, AC012640.2 and AC023310.4 were confirmed to be unfavorable prognostic factors for LSCC while the other two IncRNAs were exactly opposite. (Fig. 3A).

According to the risk coefficients acquired from the multivariate Cox regression analysis, a Sankey diagram was also produced to reveal the relationship among these autophagy-associated genes, the extracted IncRNAs and their corresponding risk types. The results indicated that AC012640.2, AC023310.4 and AL513318.2 were risk factors for the prognosis of LSCC while NCK1-DT and PTOV1-AS2 were protective factors (Fig. 3B).

The relationships between these prognostic IncRNAs we identified in signature and the associated autophagy genes are shown in Fig. 4.

### 3.3 Validation of the risk model of LSCC

Based on the risk score, patients with LSCC were divided into the high- and low-risk groups according to the median risk score. Significant differences between the two risk levels were displayed by Kaplan-Meier survival curve in all three (training, test, entire) sets, suggesting a strong association between high risk and poor prognosis (Fig. 5). Moreover, compared with the previously published IncRNA signature[28], the area under the receiver operating curve (AUC) of all three sets were larger (training set = 0.841; test set = 0.718; entire set = 0.757), further confirming the better effectiveness of our signature in the prognosis of LSCC(Fig. 5). The distribution of the risk scores and survival time are shown in Fig. 6. The heatmap in Fig. 6 displays the expression distributions of three sets for the five autophagy-related IncRNAs, with the color change (from green to red) indicating the expression levels rising from low to high.
The above results demonstrated the good performance of the risk score based on our signature as a prognostic predictor to evaluate the OS of patients with LSCC.

### 3.4 The risk model was closely related to clinicopathological features of LSCC

Univariate and multivariate Cox regression analyses demonstrated the risk signature was an independent predictive factor of LSCC patients (Fig. 7, Table 3). Based on the prognosis signature, a nomogram integrating the signature and other clinicopathological features was established to predict the 1-, 3- and 5-year OS of LSCC patients (Fig. 8). Calibration curves further confirmed the consistency between the nomogram-predicted OS and the actual OS in the entire set (Fig. 8). Moreover, combined with ROC analysis, AUCs of the nomogram were 0.814, 0.841 and 0.766 for 1-, 3-, and 5-year OS, demonstrating a much better predictive accuracy of the nomogram compared with that of a single clinical factor (Fig. 8).

### 4. Discussion

LSCC consists of heterogeneous histological subtypes and extensive changes may occur in the clinical course [1, 29]. With the increasingly diversified and individualized treatment strategies of LSCC, the existing TNM staging system seems to have certain limitations, making it necessary to explore reliable and novel prognostic biomarkers[30]. With the help of high-throughput biological technology, IncRNAs have been found closely associated with transcription, posttranscription and various cell functions such as cell apoptosis and autophagy, which indicates its potential value as a prognostic indicator of tumors[31]. For instance, IncRNAs have been identified as effective prognostic biomarkers in colorectal cancer and kidney renal clear cell carcinoma [32, 33]. Whereas few studies have investigated the role of IncRNAs in the prognosis of LSCC. Therefore, it is significant to establish a IncRNA signature to predict the prognosis of patients with LSCC.

Autophagy, a highly conserved regulatory mechanism for eukaryotic cells, has been considered to participate in the development and progression of tumors[34]. Though the autophagy-related cell death exerts anti-tumor effects at the early stages, the genomic instability and necrosis-induced inflammation caused by autophagy can promote tumorigenesis and tumor growth[35]. Additionally, autophagy plays an important role in maintaining tumor growth and metabolism in a microenvironment of low oxygen and nutrient deficiency, as well as enhancing the resistance of tumor cells to chemoradiotherapy and promoting tumor recurrence[36]. Recently, growing evidence has indicated the potential relationship between autophagy and LSCC. Autophagy suppression was reported to enhance DNA damage and cell death of LSCC cells[13]. LncRNA H19 were found to regulate the autophagy-related drug resistance in LSCC by targeting miR-107[37]. Through inhibiting autophagy, circPARD3 was demonstrated to drive malignant progression and chemoresistance of LSCC[38].

Considering the importance of both autophagy and IncRNAs in LSCC, we used the TCGA dataset to explore the prognostic value of autophagy-related IncRNAs in this study. Firstly, we identified differential expressed autophagy-associated IncRNAs. Subsequently, five autophagy-related IncRNAs that enabled
the classification of high- and low risk LSCC patients were screened out. It was found that LSCC patients in the high-risk group had a shorter survival when compared with those in the low-risk group. Additionally, the five-IncRNA signature also exhibited a high prediction accuracy when calculating AUC in ROC analysis. Combined with the univariate and multivariate Cox regression analyses, we finally validated the five-IncRNA signature as an independent prognostic indicator for LSCC.

Among these five selected IncRNAs, AC012640.2, AC023310.4 and AL513318.2 were risk-associated genes, while NCK1-DT, PTOV1-AS2 were protective factors. At present, the specific role of IncRNAs in oncogenesis is still controversial. For example, NCK1-DT, also named as NCK1-AS1, is considered as a risk factor in the prognostic prediction of cervical cancer, which is contrary to the result in our research[39]. Hence, the role of IncRNAs in tumor genesis and development needs further research.

There are some limitations in our study. Firstly, the performance of this prognostic signature could be more reliable if it was validated using other independent external datasets with long-term follow up. In addition, due to the relative limited amount of data available for LSCC in TCGA, the number of samples from LSCC patients (n = 111) was obviously larger than those matched (n = 12), which may have biased our results to some extent. Moreover, further investigations, including immunohistochemistry, quantitative real-time PCR and other biochemical experiments are needed to confirm our findings. Meanwhile, the actual role of some identified autophagy-related IncRNAs still remains to be further studied.

5. Conclusions

In this study, we identified differentially expressed autophagy-related IncRNAs, and subsequently constructed a five-IncRNA signature that could predict the survival outcomes of LSCC patients. Moreover, by combining the signature with other clinicopathological features, a prognostic nomogram was established to validate its independence and accuracy. Our study provides potential prognostic biomarkers and therapeutic targets for LSCC, as well as better understanding of the value of IncRNAs in carcinomas.

Abbreviations

LSCC
laryngeal squamous cell carcinoma
IncRNAs
long no-coding RNAs
DEIncRNAs
differentially expressed IncRNAs
TCGA
The Cancer Genome Atlas
ROC
Receiver operating characteristic
AUC
Area under curve

TNM
tumor-node-metastasis

OS
overall survival

Declarations

Ethics approval and consent to participate: Not applicable. All data in this study are publicly available and no permission was required to perform this study.

Consent for publication: Not Applicable.

Availability of data and materials: The datasets generated and/or analyzed during the current study are available in the TCGA (http://www.cbioportal.org) database.

Competing interests: The authors declare that they have no competing interests.

Funding: The present study was supported by grants from the National Natural Science Foundation of China (81972529).

Authors’ contributions: YJS conceived this work and HYH wrote the paper. YJS collected and preprocessed the data from TCGA. QH performed the analysis and HYH prepared the figures and tables. QH helped to interpret the results. LZ revised the manuscript and supervised the entire study. All authors have read and approved the manuscript. YJS, HYH and QH contributed equally to this study.

Acknowledgements: Not Applicable.

Author information:

YJS, HYH and QH contributed equally to this study.

Department of Otorhinolaryngology Head and Neck Surgery, Eye, Ear, Nose, and Throat Hospital, Fudan University, Shanghai, 200031, P.R. China.

Yujie Shen, Huiying Huang, Qiang Huang & Liang Zhou

References


Tables

Due to technical limitations, table 1-3 is only available as a download in the Supplemental Files section.

Figures
Figure 1

Flowchart for the analysis of autophagy-related long non-coding RNA prognostic signature for laryngeal squamous cell carcinoma.
Figure 2

Heatmap and corresponding volcano plot of DElncRNAs. A. Volcano plot of DElncRNAs. Each dot represents a gene, with red means up-regulated while green means down-regulated. B. Heatmap of the top 20 up-regulated and down-regulated DElncRNAs in 111 laryngeal cancer tissue and 12 normal tissue samples.

Figure 3

Kaplan–Meier survival curve for the five selected autophagy-associated IncRNA for LSCC in TCGA database and their Sankey diagram. A. The Kaplan–Meier survival analysis indicated that a high expression of AL513318.2, AC012640.2 and AC023310.4 always related to worse prognosis for patients with LSCC, while NCK1-DT and PTOV1-AS2 were considered to be protective factors. B. Sankey diagram displays the relationships among autophagy-related genes, autophagy-related DElncRNAs, and risk types.
Figure 4

Co-expression network of autophagy genes and lncRNAs we identified in signature. Red nodes represent autophagy-related lncRNAs, and the blue nodes represent autophagy genes.
Figure 5

Kaplan–Meier survival curves and ROC curves to evaluate the effectiveness of five-lncRNA signature. There were significant differences in OS between the high- and low-risk groups in all three sets. The areas under the ROC of three sets all exceeded 0.7 (A Training set, B Test set, C Entire set).
Figure 6

The risk score, survival status and expression of five selected autophagy-related lncRNAs for patients in high- and low-risk groups (A Training set, B Test set, C Entire set).

Figure 7

Figure 8

Nomogram construction and assessments in LSCC. A. Nomograms to predict OS of patients with LSCC. B. Calibration curves and ROC analysis for nomogram at the 1-tear survival time C. Calibration curves and ROC analysis for nomogram at the 3-tear survival time D. Calibration curves and ROC analysis for nomogram at the 5-tear survival time

Supplementary Files

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

- table3.xlsx
- table1.xlsx
- table2.xlsx