A Five Autophagy-Related Long Non-Coding RNA Prognostic Model for Patients with Lung Adenocarcinoma

Boxuan Liu
Xi'an Jiaotong University Second Affiliated Hospital

Yun Zhao
Xi'an Jiaotong University Medical College First Affiliated Hospital

Shuanying Yang (yangshuanying@xjtu.edu.cn)
Xi'an Jiaotong University Second Affiliated Hospital  https://orcid.org/0000-0002-2132-9361

Primary research

Keywords: lung adenocarcinoma, autophagy, long non-coding RNA, prognostic model, The Cancer Genome Atlas

DOI: https://doi.org/10.21203/rs.rs-543033/v1

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

Background: Lung adenocarcinoma is the most occurred pathological type among non-small cell lung cancer. Although huge progress has been made in terms of early diagnosis, precision treatment in recent years, the overall 5-year survival rate of a patient remains low. In our study, we try to construct an autophagy-related lncRNA prognostic signature that may guide clinical practice.

Methods: The mRNA and lncRNA expression matrix of lung adenocarcinoma patients were retrieved from TCGA database. Next, we constructed a co-expression network of lncRNAs and autophagy-related genes. Lasso regression and multivariate Cox regression were then applied to establish a prognostic risk model. Subsequently, a risk score was generated to differentiate high and low risk group and a ROC curve and Nomogram to visualize the predictive ability of current signature. Finally, gene ontology and pathway enrichment analysis were executed via GSEA.

Results: A total of 1,703 autophagy-related lncRNAs were screened and five autophagy-related lncRNAs (LINC01137, AL691432.2, LINC01116, AL606489.1 and HLA-DQB1-AS1) were finally included in our signature. Judging from univariate (HR=1.075, 95% CI: 1.046–1.104) and multivariate (HR =1.088, 95%CI = 1.057 − 1.120) Cox regression analysis, the risk score is an independent factor for LUAD patients. Further, the AUC value based on the risk score for 1-year, 3-year, 5-year, was 0.735, 0.672 and 0.662 respectively. Finally, the lncRNAs included in our signature were primarily enriched in autophagy process, metabolism, p53 pathway and JAK/STAT pathway.

Conclusions: Overall, our study indicated that the prognostic model we generated had certain predictability for LUAD patients’ prognosis.

Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide. Newly diagnosed lung cancer patients in the US has exceeded 2 million a year. Non-small cell lung cancer (NSCLC) accounts for nearly 80% of all new lung cancer cases including lung adenocarcinoma and lung squamous carcinoma. For lung adenocarcinoma, it is the most common pathological type of NSCLC. Despite the advancement of treatment strategy, the prognosis of lung adenocarcinoma (LUAD) patients has limited improvement in 5 year survival rate. Therefore, it is of great need for us to establish novel prognostic signature to adjust each patients’ treatment.

Autophagy is a protein degradation process with multiple steps in eukaryotes through autophagosomes and lysosomes and the whole process makes a vital part in maintaining homeostasis. Studies have shown autophagy plays a critical part in various human diseases such as heart disease, tumorigenesis and tumor progression as well as neurology malfunction. In recent studies, multiple findings indicate autophagy involves in tumor occurrence, maintenance as well as progression. Further analysis shows the process of autophagy can be classified into one of the tumor suppressor mechanisms and enhance chemotherapy response. In past years, numerous researchers have put efforts to find new treatment strategies for LUAD patients through autophagy related pathways.

Non-coding RNA (ncRNA) is a class of RNAs with no protein-coding ability and a Majority of human genomes are transcribed into this kind of RNAs. Long ncRNAs (LncRNAs) are the most extensive studied with a length over 200 nucleotides and involve in tumorigenesis in various cancer types and play a role in cell cycle, apoptosis and chemoresistence. In recent years, several studies have found that the expression of certain lncRNA may also serve as biomarkers that can help physicians to predict a patients’ prognosis. Kumar et al. found the high expression of p53 in LUAD patients predicted a poor prognosis. Zeng et al. established a five-lncRNA signature which can serve as an
independent survival predictor. Up till now, there were no autophagy-related lncRNA risk models had been established to guide daily clinical practice.

Therefore, we applied mRNA and lncRNA expression profile from TCGA database to generate a risk model to predict LUAD patients’ prognosis.

**Methods**

**2.1. Raw data acquisition and initial analysis**

The mRNA and lncRNA expression profile and its clinical data of LUAD patients were downloaded from The Cancer Genome Atlas (TCGA) database ([https://portal.gdc.cancer.gov/](https://portal.gdc.cancer.gov/)). The Practical Extraction and Report Language (Perl) script was then applied to extract survival time, age, gender, tumor stage and TNM stage to merge into a single file. After initial screening the clinical file, we exclude patients with a short follow up time (<30 days) and the final patients we obtained for subsequent analysis were 454.

**2.2. Screening Autophagy-Related lncRNAs**

A list of autophagy related genes(Supplementary File 1) were obtained from the Human Autophagy Database ([HADb, http://autophagy.lu/clustering/index.html](http://autophagy.lu/clustering/index.html)). Then, these autophagy-related gene matrix were filtered out and went through a log2 transformation. Pearson correlation test was exploited to filter out the most correlated lncRNAs with a cut-off value of correlation coefficient was set as $|R| > 0.3$ and $P < 0.001$.

**2.3. Identification of Autophagy-Related lncRNAs signature for lung adenocarcinoma**

After initial screening of autophagy-related lncRNAs, Cox regression was used to determine if the lncRNA was significantly correlated with patients’ prognosis. Then, the least absolute shrinkage and selection operator (Lasso) regression were adopted to construct a prognostic risk score base on the following formula:

$$\text{risk score} = \text{coef(lncRNA1)} \times \text{expr(lncRNA1)} + \text{coef(lncRNA2)} \times \text{expr(lncRNA2)} + \ldots + \text{coef(lncRNA}_n) \times \text{expr(lncRNA}_n).$$

$\text{coef (lncRNA}_n)$ was defined as the coefficient of lncRNAs.

$\text{expr (lncRNA}_n)$ was defined as the expression of lncRNAs.

According to the median risk score, all included samples were then allocated into a high-risk group and a low-risk group.

**2.4. Independent survival analysis of the signature for LUAD patients**

To test the credibility of the prognostic model we constructed, we applied both univariate and multivariate Cox regression to assess patients’ survival with clinicopathological factors and risk score. The receiver operating characteristic (ROC) curves were generated using the survivalROC R package. Further, we drew a Nomogram to visualize a patient’s survival probability under our prognostic signature and the index of concordance ($C$-index) was also calculated to show the accuracy of the signature.
2.5. Functional Analysis

Gene Set Enrichment Analysis (GSEA) (http://software.broadinstitute.org/gsea/index.jsp) is a software that calculates whether a set of genes exhibits significant differences between two groups. We conducted GSEA analysis using the risk score as the phenotype to get significantly up and down regulated GO terms and KEGG pathways.

2.6. Statistical Analysis

All statistical analyses were performed using R Studio (version 1.1.453). The limma R package was applied to differentiate the survival-related IncRNAs and the co-expression network of the IncRNAs-mRNAs was established and visualized via Cytoscape and Sankey diagram. The Kaplan-Meier survival analysis was visualized using survival R package and the p-value was calculated via log-rank tests. The univariate and multivariate Cox regression analysis were applied to assess the relationship between risk score and clinopathological parameters via t-test. The Nomogram was drawn by applying the rms R package. The ROC curve was used to visualize the credibility of the the signature and an AUC value over 0.60 was considered had certain credibility. Two-tailed $p < .05$ was considered statistically significant.

Result

3.1. Establishment of a Co-expression Network

A total of 14,142 lncRNAs were separated form the patients’ expression profiles we downloaded from TCGA database. In searching of HADb, we found 257 autophagy related genes(Supplementary File1) of which 210 genes were identified in our TCGA-LUAD expression matrix (Table S1). Next, we constructed a co-expression network between the autophagy related gene we identified and the correlated lncRNAs with a cut-off value of $|R^2| >0.3$ and $P < 0.001$. Finally, a total of 1,703 autophagy-related lncRNAs (ARlncRNAs) were filtered out for further analysis.

3.2. Development of Prognostic Risk Model from Autophagy-related IncRNA

Based on our univariate Cox regression result, 74 autophagy-related IncRNAs were found to have a significant correlation with patients’ survival in lung adenocarcinoma. 57 of the 74 ARlncRNAs were favorable factors (HR<1) while 17 ARlncRNAs showed a harmful results (Table S2). The Lasso regression found 5 autophagy-related IncRNAs (Figure1a) and subsequent multivariate Cox regression included the 5 ARlncRNAs (LINC01137, AL691432.2, LINC01116, AL606489.1 and HLA-DQB1-AS1) into our risk score model (Figure1b-c, Table1). Besides, the Kaplan-Meier survival analysis was conducted to evaluate the prognostic value of each of the the ARlncRNAs. All of the five ARlncRNAs had significant prognostic value in predicting patients’ overall survival among which AL691432.2 and HLA-DQB1-AS1 were considered protective factor while the other three ARlncRNAs were risk factors (Figure 2). Moreover, we had established a co-expression network and Sanky diagram to show the correlation between significant ARlncRNAs and autophagy related genes (Figure 3). Meanwhile, the network also revealed that AL606489.1 had a significant correlation with BIRC6 and WDFY3 from the Pearson correlation test (Supplementary file 2). Finally, according to the results of the multivariate Cox regression analysis, a risk score model was generated based on the following formula: Risk Score= (0.0538 × LINC01137 expression) + (-0.0853 × AL691432.2 expression) + (0.0718 × LINC01116 expression) + (0.221 × AL606489.1 expression) + (-0.0562 × HLA-DQB1-AS1 expression).
3.3. The Prognostic Influence of the Established Signature

According to our risk score formula, the risk score of each patient was calculated and assigned to high/low risk group. The median risk score in our analysis was 0.9939 and 223 patients were assigned to low risk group while 222 patients was assigned to high risk group. The Kaplan-Meier survival analysis showed that the risk score was significantly correlated with patients’ overall survival in lung adenocarcinoma with a p-value <0.001 (Figure 4e). For the time-dependent ROC curve, the AUC value for 1-year, 3-year, 5-year, 7-year was 0.735, 0.672,0.662 and 0.732 respectively which indicating a reliability of this signature combination to predict patients’ survival (Figure 4a-d). Next, the risk score distribution along with the survival time between high and low risk group showed in Figure 5a implies a poorer survival probability in high risk group. For this prognostic signature, high expression of AL691432.2 and HLA-DQB1-AS1 were found in low risk group while the other three were associated with high risk group (Figure 5b). Furthermore, we found a small number of deaths in the low risk group compared to high risk group.

3.4. Clinical Value of the Autophagy-Related lncRNA Signature

Univariate and multivariate Cox regression analysis were conducted on this 5- ARlncRNAs signature in patients’ cohort to evaluate whether this signature can be an independent factor of other relevant information such as gender, age and TNM stage. In our univariate Cox regression analysis, the risk score and tumor stage were independent prognostic indicators with a HR of 1.075 (95% CI: 1.046–1.104, P < 0.001) and 1.666 (95% CI: 1.409–1.969, P < 0.001) respectively (Table 2, Figure 6a). In the multivariate Cox regression analysis, the overall survival was used as a dependent variable and other clinical factors (age, gender, stage and TNM) were regarded as covariates. The multivariate Cox regression analysis indicated that risk score still was an independent factor in our analysis with a HR of 1.088 (95% CI = 1.057 – 1.120, P < 0.001), Table 3, Figure 6b). Further, according to the nomogram we drew from our analysis, risk score and tumor stage were the most two contributors to 1-year, 3-year and 5-year overall survival of patients with lung adenocarcinoma (Figure 6c). Besides, to evaluate the reliability of this scoring system, the AUC value under the ROC curve we drew from our results indicated risk score (0.668) as well as tumor stage (0.733) had certain credibility in predicting a patient’s prognosis based on the 5-ARlncRNAs signature (Figure 6d). The C-index of our model was 0.726 (95% CI: 0.671-0.781). A detailed clinical stratification on risk score showed this 5-ARlncRNAs signature had a close relationship with tumor stage especially the tumor size as well as the lymph nodes metastasis (Table 4).

3.5. Functional Analysis

GSEA was then applied to find the relevant gene ontology terms and KEGG pathways that involved between the high and low risk group. For GO terms, we found a total of 147 upregulated in high risk group and 324 downregulated in high risk group with a P<0.001. In our analysis, the most concentrated biological process including cell metabolism, cell division as well as T cell selection (Table S3). For KEGG analysis, a total of 20 pathways were enriched among which 14 were upregulated in high risk group. The enriched pathways were mostly related with p53 signaling pathway, sugar metabolism, protein export, DNA replication and JAK-STAT signaling pathway (Table S3). Moreover, most of the GO terms and KEGG pathways enriched in our analysis were found to be closely related to the occurrence and development of lung adenocarcinoma, indicating that the five IncRNA may play a role in lung cancer related functions.

Discussion
With the development of novel treatment options in recent years, the overall survival time of lung cancer patients improves greatly. However, recurrence and metastasis remain the major cause of patients’ mortality. Autophagy, a process that degrades and recycles cellular component to maintain homeostasis, has been extensively studied to play as a double-edged sword in cancer development\textsuperscript{11,12}. Researchers also put a lot emphasis on lncRNA for it played crucial role in tumorigenesis in various cancers and may serve as biomarkers in cancer diagnosis and prognosis\textsuperscript{13}. Thus, this led us to find if there were potential specific autophagy-related lncRNAs signature that can be used to predict patients’ prognosis.

In our study, we had established a risk model of 5 autophagy-related lncRNAs as an independent tool to predict a patients’ survival based on Lasso regression and Cox regression (LINC01137, AL691432.2, LINC01116, AL606489.1 and HLA-DQB1-AS1). So far, among the 5 autophagy-related lncRNAs we have included in our risk model, only LINC01116 and HLA-DQB1-AS1 had been found to be related with cancer. HLA-DQB1-AS1 has been found to be a protective immune related lncRNA in LUAD patients which is consistent with our study that it has a low expression in high risk group\textsuperscript{14}. For LINC01116, numerous studies have identified that it plays a role in tumorigenesis, promotes proliferation, inhibits proliferation and even contributes to drug resistance in various cancers such as glioma, breast cancer, gastric cancer, ovarian cancer and lung cancer\textsuperscript{15-21}. In lung cancer, Zeng \textit{et al} found LINC01116 overexpressed in LUAD patients and this contributes to tumor proliferation and metastasis\textsuperscript{21}. In another study, LINC01116 was found to be involved in miR-744-5p/SCN1B axis which resulted in the exacerbation of lung squamous carcinoma\textsuperscript{22}. Additionally, LINC01116 contributes to gefitinib resistance by regulating the IFI44 expression\textsuperscript{19}. Moreover, LINC01116 has also been found in cisplatin resistance in lung adenocarcinoma via the EMT process\textsuperscript{23} which is consistent with our result that it serves as a risk factor in patients’ prognosis. LINC01137 was found to be an immune-related biomarker in diagnosis of psoriasis and also an indicator for chemical stress response\textsuperscript{24,25} and the upregulation of it may result in poor overall survival.

The signature based on our 5 autophagy-related lncRNAs has a AUC value of 0.735, 0.672 and 0.662 for 1-year, 3-year and 5-year respectively which indicating a certain reliability in predicting patients’ survival. Multivariate Cox analyses also reveals that both the risk score based on the 5 autophagy-related lncRNAs as well as the tumor stage can be used as independent prognostic indicators. Taken together, we believe this signature has the potential to predict a patients’ survival with certain credibility.

Next, we conducted functional enrichment analysis of the 5 autophagy-related lncRNA we included in our prognostic signature. The GSEA results showed that the most significantly enriched biological process were autophagy-related metabolism and immune related T cell selection. For KEGG pathway, p53 signaling pathway as well as JAK/STAT signaling pathways were involved in lung adenocarcinoma. P53 is a well known signaling pathway that has been studied in various cancers and it plays an important role in regulating autophagy. Studies showed that either blocking p53 expression with molecule agent or applying p53-knockout cell line will result in enhanced autophagy\textsuperscript{26}. Further, inhibition of p53 degradation may prevent autophagy suggesting an apoptotic role of autophagy\textsuperscript{27}. JAK/STAT is another signaling pathway that has been well studied in tumorigenesis. In lung cancer, nearly 55% of patients and most of lung cancer cell lines were found to have a much higher expression of STAT3 and this overexpression of STAT3 had a close relationship with the occurrence of lung cancer\textsuperscript{28}. Further, a study conducted by Sun \textit{et al} suggested that the activation of JAK1/STAT1 had a positive correlation with a patients’ TNM stage especially with nodal metastasis\textsuperscript{29}. Even if there are no relevant studies indicate a potential relationship between the lncRNAs and the signaling pathways, we will continue to study this potential mechanism in our future work.

Several limitations still remain in our work. First of all, the patients included in our study is only a small fraction so that the accuracy of the result may have certain deviation. Secondly, we may need a prospective study to test the availability of the established prognostic model in our study. Thirdly, more basic biological experiments are required to detect the
expression level of current lncRNAs in tumor samples as well as to reveal the potential molecular mechanism that involves.

**Conclusion**

In conclusion, our study provided a comprehensive analysis of autophagy-related IncRNA in lung adenocarcinoma and constructed a coexpression network which indicates a potential mechanism underlying the autophagy-related IncRNA. The five autophagy-related IncRNAs prognostic signature significantly correlated with a patient's survival and the generated risk score was considered an independent factor for lung adenocarcinoma patients. In summary, the five autophagy-related IncRNAs and the generated prognostic model might be a potential biomarkers and therapeutic targets for the patients with lung adenocarcinoma.

**Declarations**

Ethics approval and consent to participate

The original experimental data we used in the study was generated in TCGA database which did not need to be approved by Human Ethics Committee.

Consent for publication

Not applicable.

Authors’ contributions

BL analyzed the data and wrote the manuscript. ZY was responsible for downloading the data. SY was responsible for the conception and guidance of the study.

Competing Interest

The authors declare that they have no competing interests.

Availability of data and materials

The data included in our study can be retrieved from TCGA database.

Funding

The present study was funded by the National Natural Science Foundation of China (grant no. 81672300).

Acknowledgements

We appreciated Xia Meng for kindly help us to download the necessary data for database.

**References**


**Tables**

Table 1: Coefficient and Survival analysis of included IncRNAs based on TCGA-LUAD data.

<table>
<thead>
<tr>
<th>IncRNA</th>
<th>Coefficient</th>
<th>HR</th>
<th>HR.95L</th>
<th>HR.95H</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINC01137</td>
<td>0.053750024</td>
<td>1.05522079</td>
<td>1.021371157</td>
<td>1.090192245</td>
</tr>
<tr>
<td>AL691432.2</td>
<td>-0.085254979</td>
<td>0.918278113</td>
<td>0.837259143</td>
<td>1.007137037</td>
</tr>
<tr>
<td>LINC01116</td>
<td>0.071843654</td>
<td>1.074487339</td>
<td>1.036366241</td>
<td>1.114010662</td>
</tr>
<tr>
<td>AL606489.1</td>
<td>0.221059368</td>
<td>1.247397484</td>
<td>1.10313008</td>
<td>1.410532186</td>
</tr>
<tr>
<td>HLA-DQB1-AS1</td>
<td>-0.05618464</td>
<td>0.945364568</td>
<td>0.904206835</td>
<td>0.98839572</td>
</tr>
</tbody>
</table>


Table 2: Clinicopathological characteristics and risk scores under univariate Cox regression.
Table 3: Clinicopathological characteristics and risk scores under multivariate Cox regression.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>z</th>
<th>HR</th>
<th>HR.95L</th>
<th>HR.95H</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>-0.0004919</td>
<td>0.00990325</td>
<td>-0.0496714</td>
<td>0.99950821</td>
<td>0.98029482</td>
<td>1.01909818</td>
<td>0.96038425</td>
</tr>
<tr>
<td>gender</td>
<td>0.01585295</td>
<td>0.18971575</td>
<td>0.08356158</td>
<td>1.01597927</td>
<td>0.70048447</td>
<td>1.4735714</td>
<td>0.93340501</td>
</tr>
<tr>
<td>stage</td>
<td>0.51022051</td>
<td>0.08543962</td>
<td>5.97170846</td>
<td>1.66565845</td>
<td>1.40883332</td>
<td>1.96930185</td>
<td>2.35E-09</td>
</tr>
<tr>
<td>T</td>
<td>0.47317741</td>
<td>0.11238298</td>
<td>4.2104012</td>
<td>1.60508612</td>
<td>1.28776844</td>
<td>2.00059372</td>
<td>2.55E-05</td>
</tr>
<tr>
<td>M</td>
<td>0.59479213</td>
<td>0.3069535</td>
<td>1.93772713</td>
<td>1.81265412</td>
<td>0.99319756</td>
<td>3.3082189</td>
<td>0.05265652</td>
</tr>
<tr>
<td>N</td>
<td>0.59090872</td>
<td>0.10551225</td>
<td>5.60038014</td>
<td>1.80562848</td>
<td>1.46830475</td>
<td>2.2204479</td>
<td>2.14E-08</td>
</tr>
<tr>
<td>riskScore</td>
<td>0.07193084</td>
<td>0.01355209</td>
<td>5.30773206</td>
<td>1.07458103</td>
<td>1.04641417</td>
<td>1.10350606</td>
<td>1.11E-07</td>
</tr>
</tbody>
</table>

Table 4: The relationship of Clinicopathological characteristics and risk score.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>z</th>
<th>HR</th>
<th>HR.95L</th>
<th>HR.95H</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>0.00640912</td>
<td>0.00995471</td>
<td>0.64382751</td>
<td>1.0064297</td>
<td>0.98698369</td>
<td>1.02625884</td>
<td>0.51968729</td>
</tr>
<tr>
<td>gender</td>
<td>-0.121044</td>
<td>0.19536558</td>
<td>-0.6195771</td>
<td>0.88599495</td>
<td>0.60413746</td>
<td>1.29935173</td>
<td>0.53553628</td>
</tr>
<tr>
<td>stage</td>
<td>0.66984495</td>
<td>0.24266002</td>
<td>2.76042563</td>
<td>1.95393433</td>
<td>1.21438775</td>
<td>3.1438553</td>
<td>5.77E-03</td>
</tr>
<tr>
<td>T</td>
<td>0.10686601</td>
<td>0.12429339</td>
<td>0.85978835</td>
<td>1.11278514</td>
<td>0.87219299</td>
<td>1.41974399</td>
<td>3.90E-01</td>
</tr>
<tr>
<td>M</td>
<td>-0.8767001</td>
<td>0.62004428</td>
<td>-1.4139314</td>
<td>0.41615393</td>
<td>0.12344439</td>
<td>1.4029321</td>
<td>0.15738205</td>
</tr>
<tr>
<td>N</td>
<td>0.01858194</td>
<td>0.20797052</td>
<td>0.0893489</td>
<td>1.01875566</td>
<td>0.67771205</td>
<td>1.53142191</td>
<td>9.29E-01</td>
</tr>
<tr>
<td>riskScore</td>
<td>0.05757551</td>
<td>0.0144736</td>
<td>3.9779673</td>
<td>1.05926526</td>
<td>1.02963851</td>
<td>1.08974448</td>
<td>6.95E-05</td>
</tr>
<tr>
<td>Clinical</td>
<td>n</td>
<td>Mean</td>
<td>SD</td>
<td>t</td>
<td>P</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;60</td>
<td>203</td>
<td>1.184</td>
<td>0.744</td>
<td>-0.4800803</td>
<td>0.632</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤60</td>
<td>102</td>
<td>1.234</td>
<td>0.927</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>158</td>
<td>1.131</td>
<td>0.714</td>
<td>-1.5470808</td>
<td>0.123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>147</td>
<td>1.275</td>
<td>0.896</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I-II</td>
<td>232</td>
<td>1.137</td>
<td>0.796</td>
<td>-2.433708</td>
<td>0.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III-IV</td>
<td>73</td>
<td>1.403</td>
<td>0.823</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>266</td>
<td>1.158</td>
<td>0.789</td>
<td>-2.2319123</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-4</td>
<td>39</td>
<td>1.494</td>
<td>0.892</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>285</td>
<td>1.188</td>
<td>0.804</td>
<td>-0.9871843</td>
<td>0.335</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>1.386</td>
<td>0.874</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>195</td>
<td>1.096</td>
<td>0.657</td>
<td>-2.7228457</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-3</td>
<td>110</td>
<td>1.386</td>
<td>1.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TCGA: The Cancer Genome Atlas database, LUAD: lung adenocarcinoma

**Figures**
**Figure 1**

Selection of IncRNA using Lasso regression. (a) Lasso coefficient of the five included IncRNAs. (b) Profiles of Lasso coefficients. (c) Univariate analysis of included IncRNAs from the samples.
Figure 2

The Kaplan-Meier survival analysis of five included IncRNAs. LINC01137, AL691432.2 and LINC01116 were independent unfavorable factors. AL691432.2 and HLA-DQB1-AS1 were independent protective factors for lung adenocarcinoma.
Figure 3

The co-expression network of autophagy-related IncRNA-mRNA and Sankey diagram. (a) mRNA – Autophagy-related IncRNAs – risk type relationship showed in Sankey diagram. (b) The co-expression network visualized using Cytoscape 3.7.2 software.
Figure 4

The prognostic indicators of the five autophagy-related IncRNAs signature. (a) 1-year survival ROC curve for LUAD patients. (b) 3-year survival ROC curve for LUAD patients. (c) 5-year survival ROC curve for LUAD patients. (d) 7-year survival ROC curve for LUAD patients. (e) Kaplan-Meier survival curve of the high-risk and low-risk groups for LUAD patients.
Figure 5

The analysis of the risk score from the generated risk model. (a) Expression profiles of lncRNAs in different groups, (b) the risk curve of each sample in high and low risk group, (c) the survival plot of each sample based on the risk score.
Figure 6

The evaluation of the constructed signature's prognostic credibility in LUAD patients. (a,b) The univariate and multivariate Cox regression analysis of risk score and clinicopathological characteristics. (d) The integrated ROC curves of risk score and clinical features, (c) The nomogram of 1-year, 3-year or 5-year survival predictibility based on risk score, age and TNM stage.
Figure 7

The functional enrichment analysis based on autophagy-related IncRNAs. (a,b) Upregulated gene ontology terms in high risk group and low risk group, (b) Upregulated KEGG pathways in high risk group and low risk group.

Supplementary Files

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

- GSEAanalysisresult.docx
- SupplementaryFile1AutophagyrelatedgenefromHADb.pdf
- SupplementaryFile2correlationanalysisofIncRNAandmRNA.docx
- UnvariateresultsofARIncRNAsbasedonTCGALUADdata.docx