A Novel Risk Model Based on Autophagy Pathway Related Genes for Survival Prediction in Lung Adenocarcinoma

ABCG 1  Fan Zhang
DE 1  Suzhen Xie
C 1  Zhenyu Zhang
E 1  Huanhuan Zhao
F 1  Zijun Zhao
G 2  Haiying Sun
A 1  Jiao Zheng

Corresponding Authors: Jiao Zheng, e-mail: zdd79_99@163.com, Haiying Sun, e-mail: momo426@stanford.edu
Source of support: NanFang Hospital Dean’s Funding (2016C022, 2016B009)

Background: Autophagy has a principal role in mediating tumor cell metabolism. However, the role of autophagy-pathway-related genes (APRGs) as prognostic markers remains obscure in lung adenocarcinoma (LUAD). More potential prognostic biomarkers are needed to deepen our understanding to explore the prognostic role of APRGs in LUAD.

Material/Methods: We used The Cancer Genome Atlas (TCGA) database to identify differentially expressed APRGs. Cox proportional hazard regression was used to identify prognostic APRGs, and then a risk model was constructed. The efficacy of the risk model was confirmed using a testing group. Lastly, we explored mutational signatures of prognostic of APRGs. T-tests were used to analyze all the expression patterns of genes by SPSS 19.0.

Results: Using TCGA database, 5 differently expressed APRGs were identified in LUAD patients, and functional enrichment analyze of the genes that were closely associated with the survival status in LUAD patients. Cox proportional hazard regression was facilitated to identify 9 APRGs (CCR2, LAMP1, RELA, ATG12, ATG9A, NCKAP1, ATG10, DNAJB9, and MBTPS2). Multivariate Cox proportional hazards regression analyses further identified 5 key prognostic APRGs (CCR2, LAMP1, RELA, ATG12, and MBTPS2) that were closely related to the survival status in LUAD. Then the prognostic scores based on the 5 genes as independent prognostic indicators were constructed for overall survival (OS) of LUAD patients; area under the curve (AUC) values >0.70 (all \( P < 0.05 \)). The efficacy of prognostic scores was confirmed by data from the testing group and showed significant differences between the low-risk and the high-risk groups for OS (\( P < 0.05 \)).

Conclusions: The risk model based on the construction of 5 APRGs can predict the prognosis of patients with LUAD, which may potentially predict prognostic signatures for LUAD.

MeSH Keywords: Autophagy • Carcinoma, Non-Small-Cell Lung • Prognosis

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/924710
Background

Lung carcinoma is one of the most frequent malignancies and a leading cause of cancer-related death while lung adenocarcinoma (LUAD) is considered the most common histological subtype with a sudden onset and unexpected 5-year survival rate [1,2]. Therapeutic options for LUAD vary considerably, representing a complicated obstacle to disease management [3]. Though effective targeted medicines have been rapidly invented, it is still the most lethal cancer in the world [4]. Unfortunately, very few biomarkers have been discovered to efficiently predict prognosis of LUAD patients [5]. Thus, exploring accurate markers for predicting the survival of LUAD is an urgent scientific problem.

Autophagy refers to a catabolic process that vesicles of the double-membrane structure are encapsulate damaged or denatured necrotic organelles in the cytoplasm to form autophagosomes, and then these autophagosomes fused with lysosomes are digested and degraded. These processes are regulated by specific genes that could adjust inflammation, immune response, oxidative stress, and so on [6,7]. Autophagy has been shown to have dual effects on cancer [8]. It mostly depends on the circumstance of cells whether autophagy plays a suppressed role or promoted role [9]. With the increasing research on the pathogenesis of autophagy, the study of autophagy in cancer has correspondingly increased. Autophagy-pathway-related genes (APRGs) which can be used as the prognostic signatures are required for autophagy [10,11]. However, the role of APRGs in the prognostic assessment in LUAD is still unknown. Thus, systematic studies for more accurate prognostic markers as well as novel and reliable therapeutic targets will be useful for assessment of tumor progression, metastasis, and recurrence risk of LUAD. Although it is well known that autophagy plays a crucial role in lung cancer, the role of autophagy in LUAD remains unclear and limited.

In this research, we searched and downloaded APRGs and analyzed RNA sequence data as well as clinicopathological characteristics of LUAD patients from The Cancer Genome Atlas (TCGA), then we performed a functional analysis of differentially expressed genes. To identify prognostic APRGs, Cox proportional hazard regression model was established to study the association between the expression of these obtained genes and the survival of LUAD patients. The prognostic score was used as an assessment for the patient’s risk score (Figure 1). These findings are expected to effectively monitor autophagy and predict the prognosis of patients with LUAD.

Material and Methods

Data mining

In our analysis, 234 APRGs were downloaded from the Human Autophagy Database (HADb, http://www.autophagy.lu/index.html). The gene expression data and the corresponding clinical data in LUAD were obtained from TCGA.

Functional of differentially expressed APRGs

To estimate the underlying mechanism of differentially expressed APRGs, the downloaded data was analyzed by the R package edgeR program, with a value of |log2(Fold Change)| >1 and \(P < 0.01\). Then we performed the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways to analyze the functional annotation of differently expressed APRGs by clusterProfiler package of R software.

Figure 1. The flowchart of identification of LUAD survival-related autophagy prognostic signature. Abbreviations: LUAD, lung adenocarcinoma.
Identification of prognostic biomarkers for LUAD

To identify the prognostic genes and validate the efficacy of risk score, we studied 458 patients with LUAD for whom we could obtain entire overall survival (OS) information (Table 1). These patients were randomly divided into 2 groups, a training group (n=230) and a testing group (n=228). The data from the training group was used to construct the Cox regression model, while data from the testing group was used to confirm the steadiness of model. Lasso regression analysis was used to get a more refined model. Next, these prognostic genes were analyzed by univariate Cox regression analysis with the expression data standardized into [log2 (data+1)]. Ultimately, the prognosis related APRGs with a P value <0.05 by univariate analysis were chosen for multivariate Cox analysis. Patients of the training group were divided into a high-risk group and a low-risk group according to their median risk score. The accuracy of the risk score was evaluated using the survival receiver operator characteristic (ROC).

Construction of an individualized prognostic index

To further analyzed the relationship between the risk score and clinical parameters, multivariate Cox regression analysis was performed according to gender, age, pathological stage, and the risk score obtained from the prognostic model as covariates.

Mutational signatures of prognostic APRGs in LUAD

To analyze the mutational frequency of the prognostic genes that are related to OS in LUAD patients, a total of 564 samples from TCGA dataset were included in this research.

Statistical analysis

All the data in our study were analyzed by SPSS 19.0 (IBM SPSS) and R 3.2.2. Multivariate and univariate Cox analyses analyzed the independent prognostic value by using survival package of R. T-tests were used to analyze the expression patterns of genes in LUAD and the clinicopathological parameters. A P-value <0.05 was viewed as statistically significant.

Results

Identification of differentially expressed APRGs in LUAD

By searching the data of LUAD in TCGA, 551 cases of the RNAseq data (54 normal issues and 497 tumor issues) as well as prognosis data of 486 LUAD patients were obtained. Totally 234 genes related to autophagy were extracted via HADb. By using edgeR software, 5 differentially expressed APRGs between tumor tissues and normal tissues were identified for further study (P<0.001, Figure 2). 4 genes were remarkably higher expressed than in tumor issues while 1 gene was remarkably lower expressed (P<0.001).

Functional annotation of differentially expressed APRGs

To explore the biological function and the pathways of the differentially expressed APRGs, we further performed GO and KEGG analysis of all the differentially expressed APRGs in LUAD. The most enriched GO item for biological processes (BP) was “positive regulation of tyrosine phosphorylation of STAT protein”; changes in molecular function (MF) were mainly enriched in “ubiquitin-protein ligase binding” (Figure 3). However, KEGG pathway enrichment analysis showed there was no statistically significant pathway.

Identification of prognostic APRGs in the training group

The univariate Cox analysis and Lasso method initially identified a total of 9 APRGs (CCR2, LAMP1, RELA, ATG12, ATG9A, NCKAP1, ATG10, DNAJB9, and MBTPS2) that had a close relationship between gene expression and OS in the training group (Figure 4A, 4B). The significantly different variables were identified in the multivariate analyses. After that, CCR2, LAMP1, RELA, ATG12, and MBTPS2 were selected as independent prognostic indicators for OS of LUAD. These optimal GO item for biological processes (BP) was “positive regulation of tyrosine phosphorylation of STAT protein”; changes in molecular function (MF) were mainly enriched in “ubiquitin-protein ligase binding” (Figure 3). However, KEGG pathway enrichment analysis showed there was no statistically significant pathway.
that patients with low-risk prognosis had significantly better survival than patients with high-risk prognosis ($P=4.67\times10^{-5}$, Figure 4C). Besides, the number of LUAD patients in the training group was ranked by the risk scores (Figure 5A, 5C, 5E). Our analyses showed that the 5-year OS rate for the high-risk group was 23.1% (95% CI=12.6–42.1%), whereas the low-risk group was 56.9% (95% CI=41.8%–77.4%). The AUC value for the ROC curve was 0.711 for 5 years, which implied that the 5 APRGs prognosis risk assessment model was credible and effective (Figure 4E).

Validation of prognostic APRGs in the testing group

To validate the prediction accuracy of the prognostic risk model, the testing group served as the validation set, by calculating the risk scores of each patient (Figure 5B, 5D, 5F). We also divided the testing group patients into a high-risk group (n=114) and a low-risk group (n=114) in the OS model. Kaplan-Meier survival curve analysis showed significant differences between the high-risk group and the low-risk group for OS ($P=4.79\times10^{-4}$) (Figure 4D). Our analysis showed that 5-year OS rates for the high-risk group and low-risk group were 18.0% (95% CI=9.21–35.3%) and 56.9% (95% CI=41.8–77.4%) respectively, which indicated that patients in the low-risk group also had better OS than patients in the high-risk group. The AUC value for the OS model was 0.703 at 5 years (Figure 4F).

Prognostic risk models are independently associated with OS of LUAD patients

We also analyzed the relationship between the risk model and key clinical parameters such as age, gender, stage, and pathological factors (Table 3). In univariate analysis, we found that T-stage (hazard ratio [HR]=1.632, 95% CI: 1.315–2.024, $P<0.001$), N-stage (HR=1.790, 95% CI: 1.459–2.196, $P<0.001$), tumor stage (HR=1.654 95% CI: 1.401–1.951, $P<0.001$), risk score (HR=1.380, 95% CI: 1.283–1.484, $P<0.001$) were suggested as independent prognostic factors (Figure 6A). Moreover, multivariate analysis suggested that tumor stage (HR=1.930 95% CI: 1.190–3.130, $P<0.01$) and risk score (HR=1.318, 95% CI:...
Figure 3. The heat plot (A) and bubble plot (B) of enriched Gene Ontology (GO) terms. The change in color from blue to red represents the increase in the adjusted P-value, and the length of the bar indicates the number of gene enrichment terms.
1.218–1.425, \( P<0.001 \) were selected as the significant risk factors for LUAD patients with complementary value (Figure 6B). These results demonstrated that the prognostic model can be independently used to predict OS for LUAD patients. We then assessed the accuracy of the risk score model using the ROC curve analysis, the area under ROC curve (AUC) values were 0.705 and 0.723, for risk score and stage respectively.

Mutational analysis of 5 prognostic APRGs involved in the LUAD network

The mutational analyses on the 5 prognostic genes revealed more than 0.9% genetic alterations in LUAD. Gene LAMP1 had the highest rate of alteration (2.6%) followed by MBTPS2 (1.7%), ATG12 (0.9%), and CCR2 (0.9%). RELA had the lowest rate of alteration (0.4%).

Discussion

LUAD is one of the leading causes of malignant lung tumors and involved high morbidity and mortality rates in 2019 [12].
Figure 4. Prognostic index of 5 APRGs. (A) Lasso regression of 5 APRGs using the OS model. (B) Lasso co-efficient profiles of 5 APRGs by optimal lambda. Kaplan-Meier plot represents that patients in the high-risk group (red line) had significantly shorter overall survival time than those in the low-risk group (blue line). (C) Left: training group (D) Right: testing group. Time-dependent ROC curve analysis shows AUC values for OS in LUAD patients. (E) Left: training group. (F) Right: testing group.

APRGs – autophagy-pathway-related genes; ROC – receiver-operator characteristic; AUC – area under the curve; OS – overall survival. LUAD – lung adenocarcinoma.
inefficient prognosis of LUAD patients leads to drug resistance and metastasis. It is clear that exploring the prognosis of LUAD may affect outcomes of patients with LUAD and will significantly benefit the diagnosis, treatment, and prognosis assessment.

Autophagy serves a pivotal role in maintaining cellular homeostasis and diminishes tumorigenesis [13]. The relationship between autophagic responses and tumor has been widely studied. Previous research has shown that APRGs can mediate the death of tumor cells [14]. Autophagy not only promotes proliferation to favor of tumor cells, but also induces apoptosis of cells. Recent research has demonstrated that ectopic expression of KISS1 downregulates the 2 key modulators of autophagy, ATG5 and ATG7. This finding implicates autophagy is involved in the mechanism of breast cancer metastatic transformation [15]. Current clinical practice evaluation of the outcome of patients who suffer from LUAD mainly depends on TMN stage, which is not precise enough. The prognostic effects of APRGs in LUAD remains to be elucidated [16,17].

During the past decades, deep mining of omics big data became an important assessment for the diagnosis, treatment, and prognosis of cancer. TCGA program provides an easy access to obtain high-quality tumor genomic data with clinical information. Herein, with the assistance of TCGA, we first mined the public data of patients with LUAD and found that 5 APRGs were significant differentially expressed in tumor samples. Then we performed bioinformatics analyses using multiple bioinformatics tools. This showed that the differently expressed APRGs tended to be an important process involved in the evolution of LUAD, providing clues to the underlying mechanisms of LUAD. By performing univariate analyses and multivariate analyses, a list of 5 core prognostic APRGs (CCR2, LAMP1, RELA, ATG12, and MBTPS2) was used to facilitate our risk model, showing their effect on the outcomes of patients with LUAD. Finally, we investigated the correlation of the expression of APRGs and clinic features. All the results demonstrated our risk score may well predict prognosis and provide novel messages for clinical application in LUAD patients. CCR2

Figure 5. Prognosis of high-risk and low-risk LUAD patients. Risk score distribution of low-risk (green) and high-risk (red) in LUAD patients in (A) training group and (B) testing group. Scatter plot of survival status of LUAD patients in (C) training group and (D) testing group. Red dots (dead); green dots (alive). Expression of risk genes in the high-risk (blue) and low-risk (pink) of the OS model in (E) training group and (F) testing group. LUAD – lung adenocarcinoma; OS – overall survival.
studies in cancer research have mostly focused on regulating immune cell recruitment, which can drive fundamental mechanisms such as growth, angiogenesis, and progression of a variety of tumors including lung cancer [18,19]. LAMP1 is an important regulator of maturation of autophagosomes and phagosomes. Directly targeting LAMP1 can modulate autophagy by accelerating autophagosome formation and degradation. Moreover, LAMP1 is involved in inhibition of autophagy-mediated proliferation and metastasis of pancreatic ductal adenocarcinoma [20]. ATG12, a key member of ATG family, has been shown to mediate its functions to conjugate to ATG5 and ATG3 to promote mitochondrial fusion and mitochondrial apoptosis [21]. Sun et al. [22] also found that targeting ATG12 can exert influence on the promotion of autophagy and regulation of EPI sensitivity in breast cancer cells. RELA, a protein subunit of NF-xB, mainly exerts its effects on post-translational modifications, including acetylation, phosphorylation, ubiquitination, and methylation, and has been proven to be implicated in numerous types of tumorigenesis [23]. Previous research has shown that overexpression of RELA can inhibit proliferation and apoptosis rates of lung cancer cells [24]. Interestingly, our results indicated that high expression of RELA can inhibit proliferation and apoptosis rates of lung cancer cells [24]. Research has shown that overexpression of RELA can inhibit proliferation and apoptosis rates of lung cancer cells [24].

In summary, we identified 5 key prognostic APRGs that could provide a more accurate estimate of OS in LUAD patients. These findings may also provide a multidimensional biomarker strategy that can more effectively and comprehensively message the role of autophagy in LUAD and may hold great promise for enhancing diagnostic accuracy and predicting treatment response. The main limitation of our findings was that our study was conducted and used data already collected from patients in public databases. Nevertheless, further prospective experimental studies are needed to validate prospective clinical trials and the mechanisms of APRGs modulating the initiation and progression of LUAD.

Conclusions

Based on TCGA, we identified 5 prognostic APRGs and developed a risk model to better predict the OS prognosis of LUAD patients,
and determined the independent prognosis of LUAD patients, and which could provide accurate novel prognosis predictor for high-risk population screening and provide another promising predictive biomarker for clinical practice in LUAD treatment.

References:


Conflict of interest

None.