Effects of autophagy-associated genes on the prognosis for lung adenocarcinoma

Chongxiang Chen¹2#, Siliang Chen³#, Xiaochun Hu³, Jiaojiao Wang⁴, Tianmeng Wen⁵, Juan Fu⁶, Huan Li¹

¹Department of Intensive Care Unit, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China; ²Guangzhou Institute of Respiratory Diseases, State Key Laboratory of Respiratory Disease, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China; ³Department of Hematology, Sun Yat-Sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China; ⁴Department of Tuberculosis, Fuzhou Pulmonary Hospital of Fujian, Teaching Hospital of Fujian Medical University, Fuzhou 350008, China; ⁵School of Public Health, Sun Yat-sen University, Guangzhou 510080, China; ⁶Department of Ultrasound and Electrocardiogram, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China

Contributions: (I) Conception and design: H Li, J Fu; (II) Administrative support: H Li, J Fu; (III) Provision of study materials or patients: C Chen, S Chen; (IV) Collection and assembly of data: C Chen, S Chen; (V) Data analysis and interpretation: X Hu, J Wang, T Wen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

These authors contributed equally to this work.

Correspondence to: Huan Li. Department of Intensive Care Unit, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, 651 Dongfeng Dong Road, Guangzhou 510060, China. Email: lihuan@sysucc.org.cn; Juan Fu. Department of Ultrasound and Electrocardiogram, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, 651 Dongfeng Dong Road, Guangzhou 510060, China. Email: fujuan@sysucc.org.cn.

Background: Several studies show that autophagy plays an important part in the biological processes of lung adenocarcinoma. Therefore, this work aimed to establish one scoring system on the basis of the expression profiles of differentially expressed autophagy-related genes (DEARGs) in patients with lung adenocarcinoma.

Methods: The Cancer Genome Atlas (TCGA) was applied to retrieve lung adenocarcinoma data. The overall survival (OS)-associated DEARGs were selected for the DEARG scoring scale. Moreover, the online database Kaplan-Meier Plotter (www.Kmplot.com) was employed to verify the accuracy of our results.

Results: The expression patterns of DEARG were detected in lung adenocarcinoma as well as normal lung tissues. A gene set related to autophagy was identified, along with 9 genes that showed marked significance in predicting the lung adenocarcinoma prognosis. According to the cox regression results, DEARGs (including ITGB4, BIRC5, ERO1A, and NLRC4) were applied to calculate the DEARGs risk score. Patients with lower DEARG risk scores were associated with better OS. Moreover, based on analysis with the receiver operating characteristic (ROC) curve, DEARGs accurately distinguished the healthy tissues from lung adenocarcinoma tissues [area under the curve (AUC) value of >0.6].

Conclusions: A scoring system is constructed based on the primary DEARGs, which accurately predicts the outcomes of lung adenocarcinoma.

Keywords: Lung adenocarcinoma; autophagy; scoring system

Submitted Nov 26, 2019. Accepted for publication Feb 04, 2020.
doi: 10.21037/tcr.2020.02.07
View this article at: http://dx.doi.org/10.21037/tcr.2020.02.07
Introduction

Nowadays, there are over 200,000 lung cancer cases among cancer diseases in 2019 worldwide (1). As the leading cause of death among various cancers (2), the total lung cancer deaths exceed the sum of those of the following three most prevalent cancers, namely, colon, breast and pancreatic cancers (1). The prognosis for lung cancer is relatively poor due to its hidden symptoms at the early disease stage (3). Thus, more attention should be paid to develop the early screening approaches and the novel treatments for lung cancer (2).

As a catabolic process, autophagy maintains the cellular homeostasis through the lysosomes-recycling intracellular components (4,5). As verified by previous studies, autophagy plays various roles, including the neutral, tumor-suppressive, or tumor promoting activities (6-8). At the early stage of cancer, protein and organelles in tumor cells are degraded by autophagy, which relieves the chromosomal instability and cell injury, thus suppressing tumor development (9-11). Nonetheless, autophagy facilitates the survival of cancer cells in the presence of stress conditions when cancer is formed, therefore promoting tumor development (12,13). Besides, those autophagy effects on cancer are examined through the immune system, microenvironmental stress, and nutrient availability (14).

Several studies report that certain scoring systems can be developed, including the immune score, stromal score or DNA damage-repair, to assess the biological processes of tumors (15-17), which reminds us that a scoring system of autophagy may also be built in various tumors (18,19). Autophagy inhibition and stimulation are reported as the treatments for autophagy, and they are also the crucial parts in patient assessment (14,20-22). Furthermore, numerous studies also demonstrate that, the activation of autophagy is remarkably related to chemo-resistance, cancer dormancy and stem cell survival (23).

Previous study has reported the relationship of autophagy with the non-small cell lung cancer (NSCLC) (3). Nonetheless, few studies use the large-scale expression patterns for assessing the autophagy effect on the prognosis for lung adenocarcinoma. This study aimed to build a novel scoring system based on the screened crucial differentially expressed autophagy-related genes (DEARGs), which might serve as a perspective tool in assessing autophagy.

Methods

Gene set related to autophagy

Genes involved in the autophagy process were identified using the human autophagy database (Human Autophagy Database, http://autophag.lu/clustering/index.html).

Patient samples

Gene expression profiles and clinical data of lung adenocarcinoma patients were retrieved from The Cancer Genome Atlas (TCGA) database, and a total of 594 specimens were collected based on TCGA microarray as the cohort.

Processes

Principal component analysis (PCA) was carried out using the R program, to investigate those different expression profiles of genes in the enrolled specimens. The receiver-operating characteristics (ROC) curves of R package were plotted using the survival ROC of R package to assess the survival specificity and sensitivity. Then, the values of area under the ROC curve (AUC) were determined based on those plotted ROC curves. In addition, the online database Kaplan-Meier Plotter (www.Kmplot.com) was applied in verifying the accuracy of our results.

Analyses of functional enrichments

For the differentiated expressed genes (DEGs), the R 3.6.1 was utilized to analyze their functional enrichment for confirming the three categories of Gene Ontology (GO), including biological processes (BP), molecular functions (MF), and cellular components (CC).

Statistical analyses

Multivariate and univariate Cox regression analyses were used for assessing the associations among DEARGs, the risk score value (established by DEARGs) and other clinical baseline characteristics for prognosis prediction. According to the risk score, patients with lung adenocarcinoma were classified as low or high-risk group based on the median risk score. Survival analyses were regarded as the crucial results; and patient prognosis [overall survival (OS)] was compared among various ranges of score using log-rank test and Kaplan-Meier analysis.

The time-dependent ROC curves were plotted to calculate the predictive ability of our risk model. Bioinformatics analysis was carried out using the R software (version 3.6.1). P value of <0.05 was considered as statistically significant.
Differences in clinical baseline characteristics between low-risk and high-risk groups were examined through $\chi^2$ test or Student’s $t$-test.

**Results**

**DEARGs within lung adenocarcinoma tissues**

Expression of 232 ARGs was analyzed in 535 lung adenocarcinoma and 59 normal lung tissue samples using Wilcoxon signed-rank test, and 30 differentially expressed ARGs (DEARGs) were found. The results showed that 18 of the included DEARGs were up-regulated, and 12 were down-regulated ($|\log_{2}FC| > 1$, FDR < 0.05, Figure 1).

**Prognostic factors of DEARGs in patients with lung adenocarcinoma**

A total of 31 DEARGs expression within lung adenocarcinoma cases was analyzed through univariate Cox regression analysis for identifying the related DEARGs for prognosis prediction (Figure 2). According to our results, 9 genes (ITGB4, DLC1, HSPB8, GAPDH, NLRC4, BIRC5, ERO1A, DAPK2, and ATIC) were related to the survival of lung adenocarcinoma (Figure 2B). Multivariate Cox proportional hazard regression analysis was also carried out to identify the best model in predicting the patient prognosis. Based on the results, NLRC4 (HR 0.63, 95% CI: 0.43–0.90), ITGB4 (HR 1.14, 95% CI: 1.02–1.28), ERO1A (HR 1.39, 95% CI: 1.14–1.70) and BIRC5 (HR 1.14, 95% CI: 1.01–1.30) were significant predictors of survival.
0.98–1.31) were found to be the risk factors that independently predicted survival. The risk score was calculated according to the following formula, risk score = (0.1331 × ITGB4 expression) − (0.4698 × NLRC4 expression) + (0.1270 × BIRC5 expression) + (0.3314 × ERO1A expression). The high-risk DEARGs were related to poor prognosis; on the contrary, the low-risk ones were markedly associated with superior prognosis.

**A higher DEARG score was related to the worse OS**

Table 1 presents the baseline features of all the included cases. As suggested by our log-rank test results obtained using the Kaplan-Meier survival curve, compared with low-risk cases, the high-risk cases showed poor prognosis (Figure 2A). Therefore, the ROC curves were plotted to determine the effect of the as-constructed signature on patient OS. Based on our results, the AUC value for those prognostic models to predict OS was 0.735. To better predict the prognosis for lung adenocarcinoma patients, baseline characteristics (including age, gender, Tumor, Node, and Metastasis stage in all patients) were used in ROC curve analysis. Then, their distributions were ranked and analyzed, which revealed that stage (AUC 0.714), node stage (AUC 0.674), tumor stage (AUC 0.653), gender (AUC 0.562), age (AUC 0.508), and metastasis stage (AUC 0.502) were followed by the risk score (Figure 2C).

As shown in the figure, high-risk cases in OS group had...
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total (n=522)</th>
<th>Percentages (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;65</td>
<td>262</td>
<td>50.19</td>
</tr>
<tr>
<td>≤65</td>
<td>241</td>
<td>46.17</td>
</tr>
<tr>
<td>Unknown</td>
<td>19</td>
<td>3.64</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>280</td>
<td>53.64</td>
</tr>
<tr>
<td>Male</td>
<td>242</td>
<td>46.36</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage I</td>
<td>279</td>
<td>53.45</td>
</tr>
<tr>
<td>Stage II</td>
<td>124</td>
<td>23.75</td>
</tr>
<tr>
<td>Stage III</td>
<td>85</td>
<td>16.28</td>
</tr>
<tr>
<td>Stage IV</td>
<td>26</td>
<td>4.98</td>
</tr>
<tr>
<td>Unknown</td>
<td>8</td>
<td>1.53</td>
</tr>
<tr>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>172</td>
<td>32.95</td>
</tr>
<tr>
<td>T2</td>
<td>281</td>
<td>53.83</td>
</tr>
<tr>
<td>T3</td>
<td>47</td>
<td>9.00</td>
</tr>
<tr>
<td>T4</td>
<td>19</td>
<td>3.64</td>
</tr>
<tr>
<td>Unknown</td>
<td>3</td>
<td>0.57</td>
</tr>
<tr>
<td>M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0</td>
<td>353</td>
<td>67.62</td>
</tr>
<tr>
<td>M1</td>
<td>25</td>
<td>4.79</td>
</tr>
<tr>
<td>Unknown</td>
<td>144</td>
<td>27.59</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>335</td>
<td>64.18</td>
</tr>
<tr>
<td>N1</td>
<td>98</td>
<td>18.77</td>
</tr>
<tr>
<td>N2</td>
<td>75</td>
<td>14.37</td>
</tr>
<tr>
<td>N3</td>
<td>2</td>
<td>0.38</td>
</tr>
<tr>
<td>Unknown</td>
<td>12</td>
<td>2.30</td>
</tr>
<tr>
<td>Survival rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survival (n=1)</td>
<td>167</td>
<td>31.99</td>
</tr>
<tr>
<td>Death (n=0)</td>
<td>355</td>
<td>68.01</td>
</tr>
</tbody>
</table>
up-regulated BIRC5, ITGB4, and ERO1A, whereas down-regulated NLRC4. Figure 3 displays the distributions of risk gene expression, survival status and risk score among these cases. Therefore, functional enrichment analysis was carried out on these genes. According to those 10 most significant GO terms with regard to MF, CC and BP categories; the above-mentioned DEGs were potentially associated with immune cell regulation, ECM, cytokine activation and binding of receptors (Figure 4).

**Multivariate Cox regression analysis on risk score as well as baseline clinical features**

The results showed that the tumor stage (HR 1.861, 95% CI: 1.159–2.990), and risk score (HR 2.052, 95% CI: 1.621–
Figure 4 GO enrichment analysis. (A) GO functional enrichment analysis: Barplot; (B) GO functional enrichment analysis: bubble plot; (C) GO circle plot; (D) GO bubble plot; (E) GO heat plot.
Figure 5 Univariate and multivariate cox regression analyses for OS among lung adenocarcinoma cases. (A,B) Univariate and multivariate cox regression analyses, respectively. OS, overall survival.

Associations of risk signature and DEARGs with clinical features (including gender, age, stage, T, N, M stage)

As demonstrated by our results, BIRC5, ERO1A, NLRC4 and risk score were related to the survival status and stage (Figure 6).

The Kaplan-Meier Plot dataset was used for result verification

The results showed that these four genes were related to the poor OS and progression-free survival (PFS); however, difference in the post-progression survival (PPS) was not statistically significant (Figure 7).

Discussion

NSCLC is generally classified as three principle histological subtypes, including lung adenocarcinoma (40–70%), large-cell carcinoma (10–15%), and squamous-cell carcinoma (20–30%) (24). This study aimed to find out a more helpful approach to predict the prognosis for lung adenocarcinoma.

Under physiological situation, autophagy functions as an intracellular quality control system to maintain the homeostasis; besides, it is mainly responsible for removing the damaged proteins (25). Previously, Mathew et al. (26)
Figure 6 Clinical correlations among DEARGs included in the risk score, risk score, and baseline clinical characteristics. DEARGs, differentially expressed autophagy-related genes.
Figure 7 Verification of DEARGs included in the risk score by Kaplan Meier Plotter (www.Kmplot.com). DEARGs, differentially expressed autophagy-related genes.
indicated that, the loss of autophagy within pancreas and liver resulted in oxidative stress, inflammation as well as chronic tissue damage, which revealed the benefits of autophagy in tumor initiation. Autophagy is the new therapeutic approach, and its process in tumor has long been proposed. Nevertheless, according to the results by Sui et al. (27), autophagy probably affected patient responses to chemoradiotherapy, which worked as an assisting tool.

Firstly, this study found the different expression of autophagy between lung adenocarcinoma and normal lung tissues, which might explain the role of autophagy. However, no consensus is reached for the controversial functions of autophagy so far (28); therefore, it may be helpful to develop a useful autophagy-related risk score to assist the clinical workers in better applying the autophagy-targeted therapy for lung adenocarcinoma patients. Obviously, compared with other classical clinical indicators, the risk score calculated by the DEARGs score was superior.

In previous studies, Liu et al. indicated that the processes of autophagy and apoptosis were crucial in the pathophysiology of NSCLC (3). First of all, plenty of evidence demonstrates that autophagy assists in the migration and invasion of lung cancer cells by toll-like receptors (TLR4)-triggered by TL3 (29). Moreover, several studies report that various substances are beneficial for lung cancer patients, such as Marsdenia tenacissima (30), Juglanin (31), SBI0206965 (32), di-limonene (33), CK1α (34) and miR-138 (35). Lastly, some drugs not only play certain roles in the autophagy, but also target some other antigens [including CD47 (36) and ULK1(37)], thus rendering them with the synergistic antioxidant effects.

Among the independent risk genes included in this study, we are especially interested in BIRC5 and ERO1A. Among them, the high BIRC5 level indicates cancer development and dismal prognosis for lung adenocarcinoma (38). Knocking down BIRC5 induces the apoptosis of cancer cells, which is thereby considered to be an underlying target of AML treatment (39). For ERO1A, Takei et al. (40) indicated that it was over-expressed in various cancer types, and its expression acted as a novel factor related to poor cancer prognosis. Additionally, Tanaka et al. (41) also showed that, ERO1A played an important part in the formation of disulfide bond under the combined action of the programmed cell death-1 (PDI). Moreover, ERO1A regulates MHC class I expression on cell surface and redox status, thus resulting in changes in CD8+ T cells susceptibility (42).

Conclusions

All in all, DEARGs are used to successfully establish a novel scoring system in this study, which accurately assesses the outcomes for patients with lung adenocarcinoma. The included DEARGs are confirmed by the online dataset called Kaplan-Meier Plotter (www.Kmplot.com). According to our results, DEARGs are regarded as the treatment targets and prognostic biomarkers within lung adenocarcinoma. Moreover, our results should be verified in additional studies.

Acknowledgments

Funding: None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The data comes from database, so ethical reviews are exempt.

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

References

4. Bouhamdani N, Comeau D, Cormier K, et al. STF-


