Construction an autophagy-clinical prognostic index in oral cancer

Jirui Jiang
Tongji University

Zhanglong Zhen
Tongji University

Junhao Yin
Shanghai Jiao Tong University School of Medicine

Xiaoli Zeng
Tongji University

ZhengNan Shan
Tongji University

Shengjiao Li (✉ 07824@tongji.edu.cn)
Tongji University

Jia Li
Tongji University

Research Article

Keywords: autophagy-related genes, oral cancer, prognostic index, overall survival

Posted Date: August 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1912380/v1

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

Read Full License
Abstract

**Background:** Autophagy, is a metabolic pathway that occurs in eukaryotic cells and regulated by autophagy-related genes (ARGs). The occurrence and development of many diseases are caused by abnormal autophagy. The purpose of this article is to explore the relationship between autophagy and prognosis of oral cancer, hoping to provide a new way for early diagnosis and guide doctors to make subsequent treatment decisions.

**Methods:** Download the RNA seq and clinical features of 305 oral cancer and 30 non-tumor patients from The Cancer Genome Atlas (TCGA) dataset. Filtered out differential expression autophagy-related genes (ARGs), and gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyzed these ARGs. Cox regression analysis filtered out the prognostic ARGs and constructed a risk score models for overall survival (OS). Divided patients into high-risk and low-risk groups based on median risk score. Kaplan-Meier analyzed the overall survival (OS). Next, receiver operating characteristic (ROC) curve verified the predictive accuracy of the model. Furthermore, we performed stratification analyses to explore the relationship between the prognostic signature and clinicopathological variables. Lastly, we used another date set to verify the model. All data was processed by R (version 3.6.0) and perl (version 5.18.4).

**Results:** The K-M plot showed the overall survival rate of the high-risk group was lower than the low-risk group's ($P=2.216e-10$). And Cox regression analysis suggested that the autophagy prognostic index was an independent prognostic factor. Furthermore, the ARGs prognostic model was confirmed in dataset of GSE65858.

**Conclusion:** This study constructed an autophagy-related signature of oral cancer, which can foresee the prognosis of patients. It will open up new prospects for fight against oral cancer.

Introduction

Head and neck squamous cell carcinomas (HNSCC) are a kind of common malignant tumor, which most are oral cavity and oropharynx. Smoking, excessive drinking, poor eating habits, and infection are risk factors for HNSCC. Chewing betel nut is also one of the reasons for the high incidence of oral cancer in individual countries and regions[1]. Oral cancer patients are predominantly male, but female patients have also increased in recent years[2, 3]. Oral squamous cell carcinoma (OSCC) accounts for more than 90% of oral cancers[4], and its treatment is mainly surgery, with radiotherapy and chemotherapy as adjuvant therapy. The first operation is often the key to cure. If the excision is not complete, patients are at high risk for locoregional recurrence and subsequent development of a new primary cancer[5, 6]. When some advanced tumors are removed, it often causes dysfunction and facial deformity, causing difficulty in swallowing and dysphonia. For patients with larger defects, prostheses or artificial prostheses are required after surgery[7]. If the tumor can be detected early and treated early, it will save a lot of time and money. Therefore, The prognosis of tumors is greatly influenced by early detection and early treatment.
Eukaryotic cells through autophagy degrade useless organelles, proteins and pathogens, which occur in lysosomes. Autophagy is a response of cells to pressure, and regulated by autophagy-related genes [6–7]. By degrading damaged cellular components and metabolites in lysosomes, it can help cells to carry out metabolism and renewal of organelles. Studies have shown that autophagy participate in all kinds of physiological and pathological processes such as cell homeostasis, aging, immunity, tumorigenesis and neurodegenerative diseases [8]. At different stages of the cancer, autophagy may have opposite roles, such as acting as a tumor suppressor or initiator. Autophagy may regulate tumorigenesis and development according to most experts, although it is controversial [9]. Therefore, this study used the public database TCGA to explore the relationship between autophagy and oral cancer, hoping to provide a way for early diagnosis and early treatment for patients with oral cancer.

In our study, the original RNA seq and clinical data about oral cancer patients were obtained from the The Cancer Genome Atlas (TCGA) database. Download suitable clinical cases based on primary site and tumor type. Screening out the differentially expressed autophagy-related genes, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were performed on these genes, a risk prognosis model was built by univariate analysis and multivariate analysis. Calculated the risk score for each patient, and the patients were divided into high-risk or low-risk group according to the median score. For the high-risk group and the low-risk group, the survival curve shows that the two groups of patients are statistically different, and the area under the ROC curve (AUC) also suggests that the model can better predict the prognosis of patients. In order to verify this model's accuracy, we downloaded another dataset GSE65858 from the Gene Expression Omnibus (GEO) database. Although this dataset is about head and neck squamous cell carcinoma, the results still show that the model can better predict the prognosis of patients. Perhaps our research can provide new ideas for predicting the prognosis of oral cancer patients.

Materials And Methods

Download patients information and datasets

First, download 257 autophagy related genes (ARGs) from the Human Autophagy Database (HADb, http://www.autophagy.lu/index.html), which is a public repository containing information about the human genes described so far as involved in autophagy. According to primary site of cancer, download the original RNA seq data set and clinical features of the oral cancer from TCGA, included 305 OSCC tissue samples and 30 non-tumor samples (https://portal.gdc.cancer.gov/). The validation dataset GSE65858 is downloaded from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/). Use R (version 3.6.0) and perl (version 5.18.4) software to process data.

Enrichment Analysis of Differentially Expressed ARGs

We obtained differentially expressed autophagy-related genes by screening criteria: log2 fold Change |log FC| > 1, false discovery rate FDR < 0.05. To further explore potential biological functions, we used the R
package “clusterprofiler” for GO analysis, including biological process (BP), cell components (CC), molecular function (MF). At the same time, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed. Use the “GOplot package” for visualization.

**Identification of prognosis-related genes and Construction the model**

After combining survival time and survival status with differential genes, the autophagy genes associated with prognosis were identified by univariate analysis using the “survival” package. According to the HR value, the gene was judged as a high-risk gene or a low-risk gene. After multivariate Cox regression analysis, the model was optimized, and five autophagy genes were finally determined to construct the prognostic risk model. The risk score is calculated as follows: \( \text{risk score} = \sum_{k=1}^{n} \text{coef}_k \cdot x_k \), coef is the coefficient, \( n \) is the number of selected autophagy genes.

**Statistical Analysis**

The risk score of each patient was calculated according to the expression level and coefficient. Risk scores were used to predict patient outcomes. Patients were divided into high- and low-risk groups based on the median risk score, with higher score representing greater risk. In order to verify the accuracy of the autophagy-related gene prognosis model, plotted the survival curves of the two groups. And analyze the relationship about prognostic signature for OS and clinicopathological variables.

**Model validation**

Downloaded the GEO dataset GSE65858, which contains transcriptome data and clinical samples of 270 head and neck squamous cell carcinomas. Based on the previously constructed autophagy-related gene prognostic model, verify the accuracy of model.

**Results**

**Differentially expressed ARGs**

RNA expression data and clinical information of 305 OSCCs and 30 normal tissues were downloaded from the TCGA database. Extracted the expression values of 257 autophagy gene. Through the screening conditions of FDR < 0.05 and |log2(Fold Change)|>1, 24 up-regulated and 11 down-regulated autophagy-related genes were screened. Heatmaps and volcano plots are shown in Figs. 1a–b. Boxplots for the 35 differentially expressed genes are shown in Figs. 1c. There are 11 down-regulated genes and 24 up-regulated genes.

**Functional enrichment of the differentially expressed ARGs**

In order to understand the biological function of these genes, GO function and KEGG pathway enrichment analysis of 35 differentially expressed ARGs. The pictures (Fig. 2) were visualized to display the top 30 GO...
enrichment analysis and pathway enrichment analysis. The biological processes of differential genes are mainly concentrated in regulation of apoptotic signaling pathway, neuron death, etc. The cell components of differential genes are mainly concentrated in integrin complex, protein complex involved in cell adhesion etc. The molecular function of differential genes are mainly concentrated in receptor ligand activity, cell adhesion molecule binding etc. KEGG results showed that the main pathways were mainly enriched in viral infection, apoptosis and other pathways.

Identification Of Prognostic Args

Univariate analysis was carried out to screen autophagy-related genes (ATGs) significantly associated with overall survival (OS) in the dataset (Fig. 3). Then, the multivariable regression analysis was performed to optimize and establish the prognosis model of oral cancer ATGs. 5 genes (BIRC5, BECN1, VEGFA, GAPDH, SPHK1) were finally used to build a prediction model (Table 1). The risk score is calculated as follows: 

$$risk\_score = \sum_{k=1}^{n} coef_k \times X_k$$

where coef is the coefficient, nt, and X is the expression level of each selected autophagy gene. We divided patients into high risk and low risk groups according to the value of the median risk score.

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Full name</th>
<th>Coefficient</th>
<th>HR</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRC5</td>
<td>Baculoviral IAP repeat containing 5</td>
<td>0.756974082</td>
<td>2.131815752</td>
<td>0.002960485</td>
</tr>
<tr>
<td>BECN1</td>
<td>Beclin 1</td>
<td>0.640582536</td>
<td>1.897585969</td>
<td>0.006356973</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
<td>0.424285798</td>
<td>1.528498374</td>
<td>0.005222257</td>
</tr>
<tr>
<td>SPHK1</td>
<td>Sphingosine kinase 1</td>
<td>0.258881825</td>
<td>0.771914239</td>
<td>0.017139479</td>
</tr>
<tr>
<td>VEGFA</td>
<td>Vascular endothelial growth factor A</td>
<td>0.227814298</td>
<td>1.25585209</td>
<td>0.022838053</td>
</tr>
</tbody>
</table>

Autophagy prognostic index is an independent prognostic factor

Calculated each patient’s risk score. Divide patients into high-risk and low-risk groups based on median risk score (Fig. 4a). As the risk values increases, the mortality increases (Fig. 4b). Besides, we can draw the conclusion that the overall survival rate of the high-risk group was lower than the low-risk group’s (P = 2.216e - 10) from the K-M plot (Fig. 4c). There is a clear difference about the five-year survival rate, the high-risk group and the low-risk group were 27.2% and 60.4% (supplementary 1).

To verify whether the autophagy-related signature and clinicopathological variables is an independent prognostic factor for overall survival (OS) of oral cancer patients, we performed Cox regression analysis. If the P value < 0.05 in both univariate and multivariate cox regression analysis, indicate that this factor
could predict prognosis significantly. Figure 5a–5b showed that risk score can be used as an independent criterion for predicting patient prognosis compared to other clinical variables. The ROC curve also further confirmed the accuracy and applicability of prognostic index for predicting the prognosis of patients with oral cancer (Fig. 5c).

And then we analyzed the relationship between the prognostic signature and clinicopathological variables. The result (Fig. 6) shows the risk score of male was higher than female ($P = 0.012$), the stage III–IV was higher than stage I–II ($P = 2.122e-04$), and T3–4 was higher than that of T1–2 ($P = 0.020$), while the M and N were not significant (Figu 14).

**External data for the model verify**

To verify that this model is useful on other datasets, we validated it on another dataset. We downloaded the dataset GSE65858 from the GEO database, which contains transcriptome data and clinical samples of 270 head and neck squamous cell carcinomas. We use the model to validate against this database. The results (Figure 7) showed that there was statistically significant difference in survival between the two groups, and the AUC value of the ROC curve was 0.602. Although this is a head and neck squamous cell carcinomas dataset, our results are well-established in this dataset.

**Discussion**

Oral cancer is a highly heterogeneous disease, and the outcomes of treatment is often unsatisfactory. The locoregional recurrence is the most reason of failure. Clinically, clinical stage at presentation and presence of metastatic lymph nodes are used to predict the survival. Besides, these indicators such as tumor size, histological grades, cell types, age, TNM stage can reveal prognosis [8]. Sometimes although patients with the same TNM stage they have very different survival times in terms of prognosis. However, when radical resection is required for some advanced malignant tumors, dysfunction and severe facial deformities are often caused [9]. Therefore, the early diagnosis of oral cancer is very important. In particular, the early occurrence of deep tumors greatly increases the difficulty of diagnosis. At present, the early clinical misdiagnosis rate can be seen at any time, and the misdiagnosis rate is close to 30%. Clinically, oral and maxillofacial tumors are easily misdiagnosed as gingivitis, traumatic ulcers, maxillary sinusitis, osteomyelitis, and tuberculosis [10]. Therefore, errors in early diagnosis will delay the treatment, resulting in later deformities and loss of language and chewing functions. Appropriate risk stratification, avoiding a one-size-fits-all approach, is necessary to develop an individualized treatment plan for patients.

Autophagy is important in synthesis and degradation of cellular components in intracellular. Many studies has explored the relationship between oral cancer and autophagy. Altered autophagy in fibroblasts enhances OSCC cell migration. Autophagy can affect the phenotype of myofibroblasts and cancer-associated fibroblasts. Increased numbers of senescence markers and autophagosomes in cancer-associated fibroblasts in OSCC [11]. In the occurrence of oral cancer, the PIK3CA gene is mutated,
which activates PI3K and promotes autophagy[12]. In the early stage of tumorigenesis, autophagy is inhibited, but as the tumor volume increases, the cells in the outer layer of the tumor near the blood vessels can continue to proliferate. The cancer cells in the center lack nutrients, but autophagy can help them survive[13]. The treatment of oral cancer also involves autophagy. The proliferation and survival of OSCC cells could be inhibit by curcumin, and the anti-cancer function via both autophagy and apoptosis[14]. RANKL-RANK pathway is a classical mechanism when bone is invade by OSCC tumor. LC3-II expressed increased in OSCC cells when treated with RANKL, indicating that RANKL participated autophagy in OSCC cells[15].

In our study, we screened 5 key genes associated with prognosis by bioinformatics. The five key genes been reported in varying degrees to be related to autophagy in cancer and bone metabolism. BECN1 plays a central role in autophagy that triggers a cascade of proteins involved in autophagosome formation. BECN1 plays an important role in the occurrence and progression of autophagy and interacts apoptosis and autophagy in OSCC cells. Treatment of OSCC tumor cells with RANKL increased the expression of BECN 1 and LC3, which are autophagy genes and associated with cancer cell lifespan[16]. Cytoplasmic expression of BECN 1 correlates with tumor grade and lymph node metastasis. Expression level of BECN 1 could reflect the overall survival and disease recurrence of OSCC[17]. BIRC5 is one of the inhibitors in apoptosis protein family and suppresses autophagy in normal tissues. Lung adenocarcinoma and acute myelocytic leukemia (AML) patients with poor prognosis always express high BIRC5[18-19]. The discovery of BIRC5 is more than two decades, and the use of BIRC5 as a cancer therapeutic target is well-known. Although many studies have explore the BIRC5’s functional role in cancers, clinical application of targeted drugs remains a challenge[20].

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is often seen as a housekeeping gene. Many studies suggest that GAPDH is involved directly in tumor progression, invasiveness, and metastases and serves as a new therapeutic target[21]. Some studies suggested GAPDH could regulate cell death. GAPDH may be regarded as a valuable target for regulating autophagy in cancer therapy via modulating of nuclear GAPDH[22].

SPHK1 is widely participated in cell growth, proliferation, and antiapoptosis[23]. SPHK1–S1P (Sphingosine 1-phosphate) could regulate osteoclastogenesis and play a role in communication between osteoclasts and osteoblasts. In the process of rank mediated transformation of BMMs into osteoclasts, SPHK1 expression level and activity increased. SPHK1 overexpression weakened osteoclastogenesis via modulation of p38 and ERK activities, low expression enhanced, conversely[24]. VEGFA can induce tumor angiogenesis and cancer progression[25]. VEGFA was highly expressed in ovarian cancer compared with normal ovarian in epithelial cells. The VEGFA high expression was correlated with chemotherapy resistance and which was mediated by autophagy [26]. In our study, we found that high expression of BIRC5, BECN1, VEGFA, GAPDH and SPHK1 were associated with worse patient prognosis, which is the same as most reported results in literature.
There are some limitations in our study. Oral cancer is a part of HNSC, so the validation set may cause deviation in the results. In addition, this result needs more clinical data to be verified, and the specific function of the included genes in the occurrence and development of tumors need to be further explored. Furthermore, the location of oral cancer has a significant impact on prognosis, and individual differences affect the prognosis.

Inhibiting of autophagy has emerged as a potential anti-cancer therapeutic approach, which deserves further research and exploration. Therefore, exploring this mechanism will open up new prospects for fight against oral cancer. The greatest value of tumor prognostic models is to foresee the later disease progression of patients, and it is expected that this model can guide doctors to make subsequent treatment decisions.

Declarations

Ethical Approval

TCGA and GEO belong to public databases. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open source data, so there are no ethical issues and other conflicts of interest.

Competing interests

I declare that the authors have no competing interests as defined by BMC, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Authors' contributions

Jirui Jiang: Conceptualization; formulation or evolution of overarching research goals and aims
ZhengNan Shan: Design of methodology; creation of models
Junhao Yin: Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components
Xiaoli Zeng: Conducting a research and investigation process
Zhanglong Zhen: Application of statistical
Shengjiao Li: presentation of the published work
Jia Li: specifically visualization; data presentation

Funding
This work was supported by the Shanghai Municipal Natural Science Foundation (No. 22ZR1467200 and No. 19140904800). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

The sample data comes from TCGA database (The sample data comes from TCGA database and GEO database) and GEO database (https://www.ncbi.nlm.nih.gov/geo/).

All data analysed during this study are included in this article.

References

5. Gharat SA, Momin M, Bhavsar C. Oral Squamous Cell Carcinoma: Current Treatment [3]
13. Cancer Genome Atlas Network Comprehensive genomic characterization of head and neck squamous cell carcinomas
Figures

Figure 1
Differentially expressed ARGs. a: Heatmap of the expression levels of 35 differentially expressed ARGs in TCGA. N normal; T Tumor; Red upregulation; Green downregulation. b: Differentially expressed autophagy-related genes (ARGs) between oral cancer and normal oral tissues. Red indicates high expression and green low expression. Black shows those genes showed no difference between OSCC and normal oral tissues. c: Expression patterns of 35 autophagy-related genes (ARGs) in oral cancer types and matched non-tumor samples.

Figure 2

Figure 3
Expression profile and prognostic value of ARGs.

Figure 4
Characteristics of prognostic gene signatures. a: Distribution of risk score and patient survival time, and status of oral cancer. b: Distribution of risk score and patient survival time, and status of OSCC. c: Kaplan–Meier analysis of TCGA oral cancer patients was stratified by median risk.

Figure 5
Prognostic indicators based on ARGs show good predictive performance. a: A forest plot of univariate Cox regression analysis in OSCC. b: A forest plot of multivariate Cox regression analysis in oral cancer. c: Survival-dependent receiver operating characteristic (ROC) curves validate the prognostic significance of ARGs-based prognostic indicators.
Figure 6

Clinicopathological significance of the prognostic index of oral cancer. P values were at different gender, different tumor stage, size.

Figure 7

External date verified the value of prognostic model. a. Kaplan–Meier analysis of HNSC patients was stratified by risk score. b. The AUC value of the ROC curve was 0.602.

Supplementary Files

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

- S.txt