Hypoxia-Associated Prognostic Markers and Competing Endogenous RNA Co-Expression Networks in Lung Adenocarcinoma

Lecai Xiong

Wuhan University Zhongnan Hospital

Xiao Zhou

Wuhan University Zhongnan Hospital

Yi Cai

Wuhan University Zhongnan Hospital

Peng Dai

Wuhan University Zhongnan Hospital

Jinping Zhao

Wuhan University Zhongnan Hospital

Xuefeng Zhou

Wuhan University Zhongnan Hospital

Hexiao Tang (✉ thx1245@sina.com)

Wuhan University Zhongnan Hospital <https://orcid.org/0000-0002-9675-9200>

Research

Keywords: Lung Adenocarcinoma, hypoxia, TCGA, competing endogenous RNAs theory, tumor microenvironment

Posted Date: June 16th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-556859/v1>

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

[Read Full License](#)

Abstract

Background: Lung adenocarcinoma (LUAD) is the most common form of non-small-cell lung cancer (NSCLC). Hypoxia has been found in 50 – 60% of locally advanced solid tumors and shown to be associated with poor prognosis in a variety of tumors, including NSCLC. The study focused on the hypoxia associated molecular hallmarks in LUAD.

Methods: 15 hypoxia related genes were selected to define the hypoxia status of LUAD by ConsensusClusterPlus based on the data from The Cancer Genome Atlas

(TCGA). Then, we investigate the immune status under different hypoxia status. Subsequently, we constructed prognostic models based on hypoxic-related differentially expressed genes (DEGs), identified hypoxia-related microRNAs, lncRNAs and mRNAs, and built a network based on the competing endogenous RNAs (ceRNAs) theory.

Results: Two clusters (cluster1 and cluster2) were identified with different hypoxia status, cluster 1 was defined as the hypoxia subgroup, in which all 15 hypoxia-associated genes were upregulated. The infiltration of CD4 + T cells and Tfh cells was lower, while the infiltration of regulatory T (Treg) cells, the expression of PD-1/PD-L1 and TMB were higher in cluster 1, which indicated immunosuppressive status. Based on the DEGs, a risk signature containing 7 genes was established. Furthermore, three differentially expressed microRNAs (hsa-miR-9, hsa-miR-31, hsa-miR-196b) associated prognostic under different hypoxia clusters and its related mRNAs, lncRNAs were identified, then built a ceRNAs network.

Conclusion: This study showed that hypoxia was associated with poor prognosis in LUAD and explored the potential mechanism from the perspective of gene signature and ceRNA theory.

Introduction

Lung cancer has the high morbidity and mortality, especially advanced or metastatic lung cancer, which often responds poorly to treatment. Non-small cell lung cancers (NSCLCs) account for almost 80 percent of lung cancers of which about 50 percent are lung adenocarcinoma (LUAD) (1). It is necessary to conduct more research to clarify the molecular mechanism and prognostic markers of LUAD.

Hypoxia is an inherent feature of the tumor microenvironment (TME) that arises from a skew in the balance between oxygen supply and consumption (2). Restriction of oxygen diffusion in avascular primary tumors and the higher oxygen consumption due to hyperproliferation of cancer cells leads to hypoxia. The role of hypoxia in tumor progression involves multiple cellular responses, including cancer cell survival, proliferation, epithelial-to-mesenchymal transition (EMT), invasion, angiogenesis, drug resistance, and metastasis (3). The cellular response to hypoxia is mainly mediated by the hypoxia-inducible factor (HIF) family of transcription factors, especially HIF-1 α , which regulate the expression of multiple genes involved in processes that drive the adaptation and progression of cancer cells (4). Many studies showed that the HIF-1 α protein is overexpressed in various common solid malignant tumors such

as breast cancer, colon, gastric cancer, lung cancer, and HIF-1 α expression is associated with poor survival (5) (6). Hypoxia stress also causes immunosuppression by controlling angiogenesis and favoring immune suppression and tumor resistance, even related to the part of therapy resistance (7). Therefore, it is important to provide markers for assessing the degree of hypoxia and hypoxia-related prognosis in lung cancer patients.

In recent years, non-coding RNAs (ncRNAs), including long noncoding RNA (lncRNA), microRNA(miRNA), circular RNA (circRNA), etc., are gaining prominence in biology and medicine, and numerous studies have shown that ncRNAs are involved in human cancer (8). MiRNA is defined as single-stranded ncRNAs of 20 nucleotides in length that are endogenously expressed and regulate gene expression on the posttranscriptional level. LncRNA refers to transcripts that are longer than 200 nucleotides and do not harbor protein-coding signatures, lncRNAs control the expression of genes in the nucleus by interacting with DNA, chromatin-modifying complexes, and/or various transcriptional regulators (9). The competitive endogenous RNAs (ceRNAs) hypothesis is one theory linking the function of protein-coding mRNAs to that of ncRNAs. CeRNA can bind to miRNA through miRNA response elements (MREs) to affect gene silencing caused by miRNA (10). LncRNA can impair miRNA activity through sequestration, thereby upregulating miRNA target gene expression (11). The “lncRNA-miRNA-mRNA” network has been confirmed in many human cancers (12).

In this study, we used 15 hypoxia-related gene expression signatures to characterize the different hypoxia status of lung adenocarcinoma (LUAD) samples in The Cancer Genome Atlas (TCGA) and depicted the infiltration of 24 immune cell types and tumor mutational burden (TMB) in LUAD tissues under different hypoxia conditions. Furthermore, the differentially expressed hypoxia associated miRNAs, lncRNAs mRNAs and related signaling pathways were analyzed. On this basis, a series of prognostic markers related to hypoxia were screened and a ceRNAs co-expression network in LUAD was constructed. These results have the potential to further improve the regulatory mechanisms under hypoxia in LUAD.

Methods

Study Cohort

We followed the methods of Zhang et al. 2020(13). The LUAD gene expression profile and miRNA mature strand expression RNAseq illuminaHiseq data of TCGA were retrieved from UCSC Xena (TCGA-LUAD) (14), the phenotype of LUAD samples was also gained. The gene expression data included 59 normal samples and 515 tumor samples, among them, 565 samples have complete clinical data. Exchanging the accession number to the ID of miRNA was performed by miRbase database (15) and miRBaseVersions.db R package. This study complied with the publication guidelines of TCGA, and ethics approval and informed consent were not required.

Classification of hypoxia status

Fifteen hypoxia related gene expression signatures were selected for our analysis according to the published studies: *ACOT7*, *ADM*, *ALDOA*, *CDKN3*, *ENO1*, *LDHA*, *MIF*, *MRPS17*, *NDRG1*, *P4HA1*, *PGAM1*, *SLC2A1*, *TPI1*, *TUBB6*, *VEGFA*, which have been shown to perform the hypoxia status (16). Spearman's rank correlation was performed to assess the correlation among these genes by "corrplot" package, and the PPI network was built using the STRING database. Two different hypoxia status groups (cluster1 and cluster2) among 515 TCGA-LUAD tumor samples were selected by using ConsensusClusterPlus package with 50 iterations, resample rate of 0.8. Principal component analysis (PCA) was analyzed and visualized by "limma" package and "ggplot2" package. The differential expression of these genes between tumor samples and normal samples, cluster1 and cluster2 were analyzed by "limma" package with a cut-Off P < 0.05, then visualized by heatmap and vioplot.

Immune cell infiltration and tumor mutational burden (TMB) analysis

The data of 24 immune cell types and infiltration acquired from the ImmuCellAI

Database (17). The mutation data of LUAD patients were obtained from the TCGA Xena Hub mentioned earlier. The TMB score for each sample was calculated by the formula as follows (18): $TMB = (\text{total mutation} / \text{total covered bases}) \times 10^6$. The relationship of immune cells/TMB score and hypoxia status was analyzed by "limma" package with a cutoff of $|\log_2(\text{fold-change})| (|\log FC|) > 1$ and adjusted P value < 0.05(19).

Identification of differentially expressed genes (DEGs) related to hypoxia

DEGs between cluster1 and cluster2 were identified by "limma" package with a cutoff of $|\log_2(\text{fold-change})| > 1$ and adjusted P value < 0.01. Then the PPI network of DEGs was constructed via STRING database (20), and the crucial sub-network was selected by the MCODE APP in Cytoscape 3.7.2 according to the rules as follows: degree cut-off = 10, node score cut-off = 0.2, Max depth = 100, and k-Score = 2 (21). Gene Ontology (GO) and KEGG pathway enrichment analysis of sub-network were performed in DAVID database (22), and visualized by drawing bubble charts.

Identification of DEGs and construction of prognostic models

Randomly divide 506 samples with complete clinical data into the training cohort and testing cohort. Univariate Cox regression analysis was performed on the training cohort, and then LASSO Cox regression (23) was employed to select the powerful independent prognostic markers with P < 0.05 for OS in LUAD and construct prognostic models. The risk score was calculated by the following formula:

$$RS = \sum_{i=1}^n \text{Coe}f(i) X(i)$$

n represents the gene number in the module, Coef (i) is the coefficient of each gene; X(i) means the mRNA expression level of each gene. When Coef (i) is less than 0, it means that the corresponding gene plays a protective effect on the patient. When Coef (i) is greater than 0, the gene represents the opposite trend for survival. All LUAD samples in the training cohort were divided into high and low risk groups. Then, the prognostic values of risk scores in the two groups were analyzed by Kaplan Meier method, sensitivity and specificity assessments were estimated using the receiver operating characteristic (ROC) curves. And The relationship between risk score, immune cell infiltration, TMB score, and clinical characteristics were analyzed in all samples. The prognostic model was verified in a testing cohort and all sample cohort.

Differentially expression of miRNA

The differentially expressed miRNAs between cluster1 and cluster2 were analyzed by “limma” package with adjusted P value < 0.05 and $|\log FC| \geq 1$. The prognostic values of differentially expressed miRNAs in LUAD were assessed with Kaplan-Meier Plotter (<https://kmplot.com/analysis/>), and the samples were divided into two groups by the best cut-off value by the tool automatically and calculated via the Kaplan-Meier analysis and Logrank-P test for the 120 months’ OS. We selected the miRNA related to prognosis, and identified the target genes via the mirDIP database (24). Target genes supported by four or more databases were regarded as candidate genes for further analysis. Find the intersection genes of candidate genes and hypoxia-related differential genes through the Venn diagram. The correlations between intersection genes and corresponding miRNAs were calculated via the Starbase database (24) based on the data from TCGA-LUAD, and the prognostic values of selected candidate genes were analyzed by Kaplan-Meier Plotter database, and miRNA-regulated genes related to the prognosis of hypoxia are identified.

Identify target lncRNAs of candidate miRNAs

The target lncRNAs of candidate miRNAs were predicted via StarBase database. lncRNAs that are negatively correlated with miRNA (P value < 0.05, correlation coefficient < - 0.1) and positively correlated with target genes (P value < 0.05, correlation coefficient > 0.1) are selected (24), and constructed the miRNAs-lncRNAs network via Cytoscape 3.7.2.

Construction of the ceRNA network and related PPI network

The proteins related to microRNA-regulated genes were predicted in STRING database, and built a network of proteins, miRNA, and lncRNA through Cytoscape 3.7.2(21).

Statistical Analysis

All statistical calculations were performed in R software, SPSS or online bioinformatic databases and tools as mentioned. Wilcox Test was used to compare mRNA expression, infiltration score of immune cells and risk score. Chi-square test was used to compare clinical and pathological parameters and other categorical variables. Differentially expressed miRNAs and mRNAs were calculated by “limma” R package. The Kaplan-Meier curve and Log-rank P test, Univariable COX and LASSO Cox regression were

used to analyze the survival outcomes. ROC curve was utilized to assess diagnostic effect. The visualization of the data was done by R 3.6.3, Graphpad 8.0 and Cytoscape 3.7.2.

Results

Consensus clustering identified two clusters of LUAD with different hypoxia status

Fifteen hypoxia related gene expression signatures were selected: *ACOT7*, *ADM*, *ALDOA*, *CDKN3*, *ENO1*, *LDHA*, *MIF*, *MRPS17*, *NDRG1*, *P4HA1*, *PGAM1*, *SLC2A1*, *TPI1*, *TUBB6*, *VEGFA*. These highly enriched for hypoxia-regulated pathways (16). Firstly, we compared the expression differences of these genes in normal tissues and tumor tissues, the results showed 10 of them were highly expressed in LUAD (Fig. 1A, B). The interrelationships and correlation were analyzed (Fig. 1C), and the PPI network included 12 genes.

According to the expression similarity of these 15 hypoxia related genes, 515 tumor samples were clustered into different hypoxia status by the Consensus Clustering Method. In the CDF curve of the consistent matrix, although the middle section of the CDF curve was flat when $K = 2$ and $K = 3$ (Fig. 1E), the interference between subgroups could be reduced to minimal when $K = 2$ (Fig. 1F-I). Thus, two subgroups marked cluster 1 ($n = 271$) and cluster 2 ($n = 244$) were identified. And the result of PCA suggested that there was a significant distinction between the two clusters (Fig. 1J). To show the hypoxia state of the two clusters more intuitively, we drew a heatmap and a violin map based on the expression of these 15 genes in the two clusters. The result showed that all of these 15 hypoxia related genes were upregulated in cluster 1 (Fig. 1K-L), and we could define cluster 1 as “hypoxia subgroup” compared with cluster 2.

Clinicopathological characteristics, immune cell infiltration, immune checkpoints, and TMB score of different hypoxia status in LUAD

After excluding samples with incomplete clinical data, the associations between hypoxia status and clinicopathological characteristics were analyzed through chi-square test in 475 samples. The results showed that the hypoxia status in LUAD was significantly associated with age, tumor TNM stage, T stage and N stage (Fig. 2A, Table 1). The data of 24 immune cell infiltration and infiltration scores in LUAD were obtained from ImmuCellAI database. The differences of immune cell infiltration and infiltration scores were shown in Fig. 2B-C. Compared with cluster 2, CD4+ T, Follicular helper T cells (Tfh) that promote tumor immunity were lower in cluster 1 (both $P < 0.01$), while the nTreg cells and iTreg cells that inhibit tumor immunity were higher in cluster 1 ($P = 0.049, 0.000$; Fig. 2C-F). The results showed that hypoxia is not conducive to the tumor immune process. Gene mutations are an important cause of tumorigenesis and development, and TMB is used to predict immune checkpoint blockade therapy efficacy (25). Hence, we evaluated the difference of TMB score in cluster 1 and cluster 2. The TMB score for each sample was calculated according to methods, the result showed that TMB score was higher in cluster 1 compared

with cluster 2 (Fig. 2G). Furthermore, the expression of PDCD1 (PD-1) and CD274 (PD-L1), the most common immune checkpoints, were higher in cluster 1 (Fig. 2H-I).

Table 1
Clinicopathological characteristics of cluster 1 and cluster 2 in LUAD patients.

Characteristic	Total	Cluster 1 (n = 250)		Cluster 2 (n = 225)		P Value
		No. of patients	%	No. of patients	%	
Gender						
Female	253	124	49.01	129	50.99	0.092
Male	222	126	56.76	96	43.24	
Age						
≤ 65	227	139	61.23	88	38.77	< 0.001
>65	248	111	44.76	137	55.24	
TNM Stage						
Stage I-II	374	185	49.47	189	50.53	0.002
Stage III-IV	101	65	64.36	36	35.64	
T Stage						
T1-T2	413	213	51.57	200	48.42	0.004
T3-T4	62	37	59.68	25	40.32	
M Stage						
M0	317	169	53.31	148	46.69	0.418
M1	22	14	63.63	8	36.36	
MX	136	67	49.26	69	50.74	
N Stage						
N0-N1	404	205	50.74	199	49.26	0.001
N2-N3	71	45	63.38	26	36.62	

Identification of hypoxia-related differentially expressed genes (DEGs) and enrichment analysis

Up- and down- regulated DEGs ($|\log_2(\text{fold-change})| > 1$ and adjusted P value < 0.01) were selected according to cluster. A total of 980 differentially expressed genes were selected, including 440 upregulated and 540 downregulated (Fig. 3A). Due to the large number of DEGs, a crucial sub-network

with 119 nodes and 6153 edges was selected by the MCODE APP in Cytoscape 3.7.2, and constructed the PPI network on STRING database (Fig. 3B). The top 10 items of each GO (biological processes, molecular functions and cellular components) and KEGG pathway analysis were shown as bubble diagrams (Fig. 3C-F).

Construction of prognostic models based on hypoxia-related DEGs

Of all 515 tumor samples, 505 of them contained overall survival (OS) data. The entire set ($n = 505$) were randomly split into training ($n = 253$) and testing set ($n = 252$) at 1:1 ratio. The gene expression profiles and the corresponding survival information of LUAD patients in train sets were screened step by step through the univariate Cox regression analysis and LASSO Cox regression analysis. 210 genes were identified through the univariate Cox regression analysis (Table S1), and 7 genes were identified as key prognostic hypoxic-related genes finally (Fig. 4A-C). The training set risk score for OS = $(0.475429 \times \text{expression level of } CKS2) + (-0.11256 \times \text{expression level of } ELF5) + (-0.15491 \times \text{expression level of } FAM184A) + (-0.05916 \times \text{expression level of } GLB1L3) + (-0.17306 \times \text{expression level of } GNMT) + (-0.16417 \times \text{expression level of } IRX5) + (-0.1664 \times \text{expression level of } RIC3)$.

Samples in the training set were divided into high- or low-risk groups according to the median risk score. The results of Kaplan-Meier analysis showed that the high-risk group had a significantly worse prognosis compared with low-risk group (Fig. 4D). ROC curves based on train set were plotted to assess the sensitivity and specificity of the prediction, the AUC value was 0.809 (Fig. 4E). The results of survival analysis and the AUC of ROC curve in testing and entire sets were similar with the above (Figure S1A, B, E, F). Since risk score has such a distinctive distinction, can it become an independent predictor? Univariate and multivariate analyses were performed in train, test and entire sets to test whether the risk signature could be an independent predictor, all results showed that risk score might be independent risk factors in LUAD patients (all $P < 0.05$; Fig. 4F, G, Figure S1C, D, G, H).

The expression levels of the 7 genes in all TCGA-LUAD samples ($n = 475$) were visualized in the vioplot and heatmap (Fig. 4H, I), and all TCGA-LUAD samples were classified into high-risk and low-risk groups. The results showed that there were

significant differences associated with risk score in terms of cluster ($P < 0.0001$), TNM stage ($P = 0.0022$), T stage ($P < 0.0001$), N stage ($P = 0.0016$) (Fig. 4J-M). The results showed that cluster1 (Relative hypoxia subgroup) had a higher risk score than cluster2, and the higher the tumor stage, the higher the risk score.

Identification of prognosis-related differentially expressed miRNAs under the hypoxic status

Compared with cluster 2, 11 upregulated miRNAs and 1 downregulated miRNAs were identified (Fig. 5A, Table S2). The prognostic value of these 12 miRNAs was evaluated via the Kaplan-Meier Plotter database. There were 3 upregulated miRNAs associated with OS (hsa-miR-196b, hsa-miR-31, hsa-miR-9).

Survival analysis showed high expression of hsa-miR-196b, hsa-miR-31, hsa-miR-9 in LUAD were associated with worse OS (Fig. 5B-D; HR = 1.68, 1.86, 1.49 respectively; all $P < 0.01$).

Identification of candidate target DEGs regulated by prognosis-related miRNA

Target genes of miRNA (hsa-miR-9, hsa-miR-31, hsa-miR-196b) were obtained via the mirDIP database, and those supported by four or more databases were regarded as candidate target genes for further analysis (Figure S2). Identify the intersection of candidate target genes and the top 50% of downregulated genes through the Venn diagram (Fig. 5E). Three target genes (*PCSK2*, *FREM2*, *ALPL*) in hsa-miR-9, two (*ATP1A2*, *PRSS12*) in hsa-miR-196b, and one (*CEBPA*) in hsa-miR-31 were identified. Except for *PRSS12*, other target genes were related to prognosis, we defined them as candidate target DEGs (*PCSK2*, *FREM2*, *ALPL*, *ATP1A2*, *CEBPA*). All candidate target DEGs were downregulated in the relative hypoxia group (cluster 1), and low expression was associated with poor survival (HR = 0.71, 0.59, 0.74, 0.67, 0.55 respectively, all $P < 0.05$, Fig. 5F-J).

Hypoxia-related competitive endogenous RNA (ceRNA) regulation network

Target lncRNA of hsa-miR-196b, hsa-miR-31, and hsa-miR-9 were predicted via the Starbase database, those that were negatively related to the candidate miRNAs (Fig. 5K) and positively related to the candidate target DEGs are selected (Table 2). Finally, 6 candidate lncRNAs were selected (*NEAT1*, *AC000123.3*, *ARHGAP27P1-BPTFP1-KPNA2P3*, *AC020978.7*, *AC021078.1*, *MIR497HG*). The local protein network between candidate target DEGs (*PCSK2*, *FREM2*, *ALPL*, *ATP1A2*, *CEBPA*) was constructed, then added candidate miRNAs (hsa-miR-196b, hsa-miR-31, hsa-miR-9) and candidate lncRNAs (*NEAT1*, *AC000123.3*, *ARHGAP27P1-BPTFP1-KPNA2P3*, *AC020978.7*, *AC021078.1*, *MIR497HG*) (Fig. 5L). In this network, loss of candidate lncRNAs leads to an increase in candidate miRNAs, which suppress the expression of *PCSK2*, *FREM2*, *ALPL*, *ATP1A2*, *CEBPA*, leading to a worse survival in LUAD.

Table 2
The correlation coefficient of lncRNA-microRNA and lncRNA-mRNA in TCGA-LUAD.
ARHGAP27P1*=ARHGAP27P1-BPTFP1-KPNA2P3

LncRNA	MicroRNA	Correlation coefficient	P value	mRNA	Correlation coefficient	P value
NEAT1	hsa-miR-196b-5p	-0.171	p < 0.001	<i>ATP1A2</i>	0.204	p < 0.001
AC000123.3	hsa-miR-196b-5p	-0.142	p < 0.001	<i>ATP1A2</i>	0.22	p < 0.001
ARHGAP27P1*	hsa-miR-196b-5p	-0.185	p < 0.001	<i>ATP1A2</i>	0.277	p < 0.001
AC020978.7	hsa-miR-9-5p	-0.194	p < 0.001	<i>PCSK2</i>	0.137	0.0016
NEAT1	hsa-miR-9-5p	-0.087	0.0486	<i>FREM2</i>	0.213	p < 0.001
AC021078.1	hsa-miR-9-5p	-0.097	0.0276	<i>FREM2</i>	0.118	0.0066
AC020978.7	hsa-miR-9-5p	-0.194	p < 0.001	<i>FREM2</i>	0.321	p < 0.001
AC020978.7	hsa-miR-9-5p	-0.194	p < 0.001	<i>ALPL</i>	0.182	p < 0.001
MIR497HG	hsa-miR-31-5p	-0.173	p < 0.001	<i>CEBPA</i>	0.206	p < 0.001

Discussion

Tumors are composed of cancer cells and TME. TME contains tumor-infiltrating immune cells, cancer-associated fibroblasts (CAFs), endothelial cells, the extracellular matrix, and a wide range of metabolites and cytokines (26). As tumors grow, preexisting blood vessels fail to sufficiently perfuse the tumor and oxygen levels drop, which lead to a hypoxic and acidic environment (27). Hypoxia is rapidly sensed by the O2/prolyl hydroxylase (PHD)/Von Hippel Lindau (VHL) axis, which induces the stabilization of HIF-1 α or HIF-2 α and the subsequent activation of a gene signature that orchestrates the cellular adaptation to hypoxia (28). Furthermore, an acidic environment caused by hypoxia impacts the metabolic and functional reprogramming of cancer cells and tumor-associated stromal cells (29). Therefore, in connection with our research direction, the effect of hypoxia on lung adenocarcinoma is a topic worthy of further discussion.

According to published studies (16, 30), 15-gene expression signature (*ACOT7*, *ADM*, *ALDOA*, *CDKN3*, *ENO1*, *LDHA*, *MIF*, *MRPS17*, *NDRG1*, *P4HA1*, *PGAM1*, *SLC2A1*, *TPI1*, *TUBB6*, and *VEGFA*) that has been shown to perform the best when classifying hypoxia status were selected. These 15 genes make up a common hypoxia signature which will be upregulated and are consistently co-expressed with previously

validated hypoxia-regulated genes under hypoxic conditions in various cancers (31). In this study, TCGA-LUAD samples were clustered into different hypoxia status according to the 15 genes expression. All of these 15 genes were upregulated in cluster 1 (Fig. 1K-L), which were defined as “hypoxic subgroup”. In the analysis of the relationship between hypoxia status of lung adenocarcinoma and clinicopathological characteristics, we found that hypoxia is associated with worse TNM staging, which suggests that hypoxia is associated with poor prognosis (Fig. 2A, Table 1).

Numerous studies indicate that the hypoxia of TME is an important reason for promoting tumor immunosuppression and resistance to immunotherapy (32). Tumor hypoxia region can recruit immunosuppressive cells such as myeloid derived suppressor cells (MDSC), tumor-associated macrophages (TAM), tumor-associated neutrophils (TANs) and Tregs, and negatively affect the activation of CD8 + T cells and CD4 + T cells (33). Hypoxia cancer cells, via HIF-1 α , secrete the chemokine CCL28 that recruits CXCR10 + Tregs into tumors (34). TGF- β is a cytokine highly abundant in hypoxic regions of the tumor (35), which leads to the arise of TANs. TGF- β induces the production of Foxp3 and ROR γ t in CD4 + T cells, which induces the differentiation of Tregs and enhances immunosuppression (36). In this study, the infiltration of 24 immune cell types were compared in cluster1 and cluster2. The results showed that the infiltration of CD4 + T cells and Tfh cells was lower, while the infiltration of nTreg cells and iTreg cells was higher in cluster 1 (Fig. 2), which indicates that there is an immunosuppressive state in cluster1. A study in breast cancer showed that hypoxic increased TMB by driving genome instability and altering DNA damage repair pathways (37). The same phenomenon was observed in this study. The TMB score, the PD-1 and PD-L1 expression of cluster1 was significantly higher than that of cluster 2 (Fig. 2H), which indicated immunosuppressive status.

Under hypoxia conditions, tumor cells activate multiple adaptive pathways to promote the evolution of a more aggressive tumor phenotype including the activation of DNA damage repair proteins, altered metabolism, and decreased proliferation (38). In the study, the differentially expressed genes between cluster 1 and cluster 2 were identified and performed GO/KEGG analysis. The results showed the differentially expressed genes related to metabolism, such as ATP binding, related to cell cycle and proliferation, such as cell division, cell proliferation, G1/S/G2/M transition of the mitotic cell cycle, etc., related to DNA damage repair, such as DNA replication, DNA repair, etc., related immune regulation, such as FoxO signaling pathway (Fig. 3). HIF stable expression under hypoxia, and promotes angiogenesis through VEGF-A, glycolysis, and pH control through CA-IX (39). There is extensive evidence showing down regulation of numerous proteins involved in homologous recombination, mismatch repair, base excision repair, and nucleotide excision repair under hypoxic conditions (40). FoxO signaling pathway is a pivotal regulator of Tregs cell function, which promotes immune suppression (41). Furthermore, there is a large amount of DGEs enriched in microtubule-based movement, protein kinase activity, p53 signaling pathway, microRNAs in cancer, etc., which all related to tumor invasion and metastasis. These underlying mechanisms together lead to the negative impact of hypoxia on tumor prognosis. To better predict the prognosis of patients, we constructed a risk signature containing 7 genes by univariate Cox and LASSO Cox regression analysis, which showed a well-prediction ability (Fig. 4).

MiRNAs and lncRNAs are identified as key regulators of gene expression in various biological and pathological processes (42). We identified potential ncRNA regulatory pathways involving miRNA, lncRNA and mRNA based on ceRNA theory, and built a PPI network which might promote the development of LUAD (Fig. 5). Researches showed hsa-miR-196b was potential biomarker in LUAD (43), *ATP1A2* mutation was found in pulmonary carcinoid tumors and involved in multiple biological processes, such as cellular metabolism, immune regulation, etc, and *NEAT1* function as a competing endogenous lncRNA in multiple tumors (44). According to the results of this research, *NEAT1* as a sponge of hsa-miR-196b alleviates its repression on *ATP1A2*, and regulates multiple biological in LUAD. Similarly, hsa-miR-31 and its predicting target lncRNA (*MIR497HG*), mRNA (*CEBPA*) all involved in multiple tumors (45) (46), there is also a competing endogenous between them in LUAD according to our results. The levels of hsa-miR-9 correlate with tumor grade and metastatic status (47), and its upregulation led to enhanced NSCLC cell invasion and adhesion via the regulation of multiple pathways (48). According to this study, samples in TCGA-LUAD with high expression of hsa-miR-9 have worse survival, and various lncRNAs (AC020978.7, *NEAT1*, AC021078.1) and mRNAs (*PCSK2*, *FREM2*, *ALPL*) targeted are all related to survival.

In conclusion, this study explored the role and potential mechanism of hypoxia in LUAD from the perspective of gene signature and ceRNA theory in silico analyses, the results showed hypoxia promoted tumor progression and immunosuppressive status through multiple pathways, and the regulatory effect of ceRNA theory on LUAD had also been observed. However, it is undeniable that our research still has some limitations. All our results are based on silico analyses of TCGA and some website analysis. More functional experiments are needed to verify the results of this research, which will be the focus of our next exploration.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

This study met the publication guidelines stated by TCGA (<https://cancergenome.nih.gov/publications/publicationguidelines>).

Competing interests

The authors declare that they have no competing interests

Funding

This project was supported by the Program of Excellent Doctoral (Postdoctoral) of Zhongnan Hospital of Wuhan University [Grant No. ZNYB2019002], Independent research project of Wuhan University [2042020kf0134], and Natural Science Foundation of Hubei Province [2020CFB703].

Author Contributions

LCX and HXT contributed to the conception of the study. LCX, XZ, JPZ contributed to experimental technology and experimental design. LCX, XZ, YC, PD performed the data analyses. LCX, XZ, YC, PD wrote the manuscript. HXT and JPZ supervised the study. All authors contributed to the article and approved the submitted version.

Acknowledgments

Thank you very much for the technical support provided by Mr. Pengju Gong.

References

1. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *International journal of cancer*. 2019;144(8):1941–53.
2. Abou Khouzam R, Goutham HV, Zaarour RF, Chamseddine AN, Francis A, Buart S, et al. Integrating tumor hypoxic stress in novel and more adaptable strategies for cancer immunotherapy. *Semin Cancer Biol*. 2020;65:140–54.
3. Schito L, Semenza GL. Hypoxia-Inducible Factors: Master Regulators of Cancer Progression. *Trends in cancer*. 2016;2(12):758–70.
4. Wigerup C, Pålman S, Bexell D. Therapeutic targeting of hypoxia and hypoxia-inducible factors in cancer. *Pharmacol Ther*. 2016;164:152–69.
5. Bos R, Zhong H, Hanrahan CF, Mommers EC, Semenza GL, Pinedo HM, et al. Levels of hypoxia-inducible factor-1 alpha during breast carcinogenesis. *J Natl Cancer Inst*. 2001;93(4):309–14.
6. Zhong H, De Marzo AM, Laughner E, Lim M, Hilton DA, Zagzag D, et al. Overexpression of hypoxia-inducible factor 1alpha in common human cancers and their metastases. *Cancer research*. 1999;59(22):5830–5.
7. Jing X, Yang F, Shao C, Wei K, Xie M, Shen H, et al. Role of hypoxia in cancer therapy by regulating the tumor microenvironment. *Mol Cancer*. 2019;18(1):157-.
8. Esteller M. Non-coding RNAs in human disease. *Nature reviews Genetics*. 2011;12(12):861–74.
9. Novikova IV, Hennelly SP, Tung CS, Sanbonmatsu KY. Rise of the RNA machines: exploring the structure of long non-coding RNAs. *J Mol Biol*. 2013;425(19):3731–46.
10. Qi X, Zhang DH, Wu N, Xiao JH, Wang X, Ma W. ceRNA in cancer: possible functions and clinical implications. *J Med Genet*. 2015;52(10):710–8.

11. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nature reviews Genetics*. 2016;17(5):272–83.
12. Anastasiadou E, Jacob LS, Slack FJ. Non-coding RNA networks in cancer. *Nature reviews Cancer*. 2018;18(1):5–18.
13. Gong PJ, Shao YC, Huang SR, Zeng YF, Yuan XN, Xu JJ, et al. Hypoxia-Associated Prognostic Markers and Competing Endogenous RNA Co-Expression Networks in Breast Cancer. *Front Oncol*. 2020;10:579868.
14. Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, et al. The UCSC Genome Browser database: 2014 update. *Nucleic Acids Res*. 2014;42(Database issue):D764-D70.
15. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res*. 2019;47(D1):D155-D62.
16. Ye Y, Hu Q, Chen H, Liang K, Yuan Y, Xiang Y, et al. Characterization of Hypoxia-associated Molecular Features to Aid Hypoxia-Targeted Therapy. *Nat Metab*. 2019;1(4):431–44.
17. Miao YR, Zhang Q, Lei Q, Luo M, Xie GY, Wang H, et al. ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy. *Advanced science (Weinheim, Baden-Wurttemberg, Germany)*. 2020;7(7):1902880.
18. Liu J, Nie S, Wu Z, Jiang Y, Wan Y, Li S, et al. Exploration of a novel prognostic risk signatures and immune checkpoint molecules in endometrial carcinoma microenvironment. *Genomics*. 2020;112(5):3117–34.
19. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
20. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017;45(D1):D362-d8.
21. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*. 2011;27(3):431–2.
22. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*. 2003;4(5):P3.
23. Ma C, Luo H, Cao J, Zheng X, Zhang J, Zhang Y, et al. Identification of a Novel Tumor Microenvironment-Associated Eight-Gene Signature for Prognosis Prediction in Lung Adenocarcinoma. *Front Mol Biosci*. 2020;7:571641-.
24. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2014;42(Database issue):D92-7.
25. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Annals of oncology: official journal of the European Society for Medical Oncology*. 2019;30(1):44–56.

26. Maman S, Witz IP. A history of exploring cancer in context. *Nature reviews Cancer*. 2018;18(6):359–76.
27. Casazza A, Di Conza G, Wenes M, Finisguerra V, Deschoemaeker S, Mazzone M. Tumor stroma: a complexity dictated by the hypoxic tumor microenvironment. *Oncogene*. 2014;33(14):1743–54.
28. Palazon A, Goldrath AW, Nizet V, Johnson RS. HIF transcription factors, inflammation, and immunity. *Immunity*. 2014;41(4):518–28.
29. Chang CH, Qiu J, O'Sullivan D, Buck MD, Noguchi T, Curtis JD, et al. Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. *Cell*. 2015;162(6):1229–41.
30. Haider S, McIntyre A, van Stiphout RG, Winchester LM, Wigfield S, Harris AL, et al. Genomic alterations underlie a pan-cancer metabolic shift associated with tumour hypoxia. *Genome Biol*. 2016;17(1):140.
31. Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br J Cancer*. 2010;102(2):428–35.
32. Riera-Domingo C, Audigé A, Granja S, Cheng WC, Ho PC, Baltazar F, et al. Immunity, Hypoxia, and Metabolism-the Ménage à Trois of Cancer: Implications for Immunotherapy. *Physiological reviews*. 2020;100(1):1–102.
33. Multhoff G, Vaupel P. Hypoxia Compromises Anti-Cancer Immune Responses. *Adv Exp Med Biol*. 2020;1232:131–43.
34. Facciabene A, Peng X, Hagemann IS, Balint K, Barchetti A, Wang LP, et al. Tumour hypoxia promotes tolerance and angiogenesis via CCL28 and T(reg) cells. *Nature*. 2011;475(7355):226–30.
35. Fridlender ZG, Sun J, Kim S, Kapoor V, Cheng G, Ling L, et al. Polarization of tumor-associated neutrophil phenotype by TGF-beta: "N1" versus "N2" TAN. *Cancer cell*. 2009;16(3):183–94.
36. Zhou L, Lopes JE, Chong MM, Ivanov II, Min R, Victora GD, et al. TGF-beta-induced Foxp3 inhibits T(H)17 cell differentiation by antagonizing RORgamma function. *Nature*. 2008;453(7192):236–40.
37. Hassan Venkatesh G, Bravo P, Shaaban Moustafa Elsayed W, Amirtharaj F, Wojtas B, Abou Khouzam R, et al. Hypoxia increases mutational load of breast cancer cells through frameshift mutations. *Oncoimmunology*. 2020;9(1):1750750.
38. Samanta D, Semenza GL. Metabolic adaptation of cancer and immune cells mediated by hypoxia-inducible factors. *Biochimica et biophysica acta Reviews on cancer*. 2018;1870(1):15–22.
39. Loboda A, Jozkowicz A, Dulak J. HIF-1 and HIF-2 transcription factors—similar but not identical. *Mol Cells*. 2010;29(5):435–42.
40. Scanlon SE, Glazer PM. Multifaceted control of DNA repair pathways by the hypoxic tumor microenvironment. *DNA Repair*. 2015;32:180–9.
41. Ouyang W, Liao W, Luo CT, Yin N, Huse M, Kim MV, et al. Novel Foxo1-dependent transcriptional programs control T(reg) cell function. *Nature*. 2012;491(7425):554–9.
42. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature*. 2014;505(7483):344–52.

43. Liu Y, Xie D, He Z, Zheng L. Integrated analysis reveals five potential ceRNA biomarkers in human lung adenocarcinoma. *PeerJ*. 2019;7:e6694.
44. Yu M, Zhang X, Li H, Zhang P, Dong W. MicroRNA-588 is downregulated and may have prognostic and functional roles in human breast cancer. *Medical science monitor: international medical journal of experimental clinical research*. 2017;23:5690–6.
45. Nagy Á, Lánczky A, Menyhárt O, Győrffy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci Rep*. 2018;8(1):9227.
46. Wu YL, Zhou C, Lu S, Qin S, Pan H, Wu G, et al. Erlotinib versus gemcitabine/cisplatin in Chinese patients with EGFR mutation-positive advanced non-small-cell lung cancer: Crossover extension and post-hoc analysis of the ENSURE study. *Lung Cancer*. 2019;130:18–24.
47. Maniadakis N, Fragoulakis V, Pallis AG, Simou E, Georgoulas V. Economic evaluation of docetaxel-gemcitabine versus vinorelbine-cisplatin combination as front-line treatment of patients with advanced/metastatic nonsmall-cell lung cancer in Greece: A cost-minimization analysis. *Ann Oncol*. 2010;21(7):1462–7.
48. Han SH, Kim HJ, Gwak JM, Kim M, Chung YR, Park SY. MicroRNA-222 Expression as a Predictive Marker for Tumor Progression in Hormone Receptor-Positive Breast Cancer. *Journal of breast cancer*. 2017;20(1):35–44.

Figures

Figure 1

Consensus Clustering identified two clusters of LUAD with different hypoxia status. (A-B) The heatmap and violin plot of the 15 hypoxia related gene expression signature in TCGA-LUAD tumor and normal samples (Wilcox Test). (C) Spearman correlation analysis of the 15 hypoxia related gene expression signature. (D) The PPI network of the 15 hypoxia related gene expression signature. (E) The CDF value of consensus index. (F) Relative change in area under CDF curve for k = 2–9. (G) The tracking plot for k = 2

to k = 9. (H, I) Consensus matrix for k=2 and k=3. (J) Principal component analysis of the total RNA expression profile. (K-L) The heatmap and violin plot of the 15 hypoxia related gene expression signature in cluster1 and cluster2 (Wilcox Test). *P < 0.05, **P < 0.01, and ***P < 0.001.

Figure 2

Immune cell infiltration and TMB score of different hypoxia status in LUAD. (A)The heatmap of clinicopathological characteristics in cluster 1 and cluster 2. (B) The violin plot of infiltration score and 24

immune cell types in cluster1 and cluster2 (Wilcox Test). (C-F) The infiltration of CD4+ T cell, Thf cell, iTreg cell, and nTreg cell in cluster1 and cluster2 (Wilcox Test). (G) The TMB score of cluster 1 and cluster 2. (H-I) The expression of PDCD1 (PD-1) and CD274 (PD-L1) in cluster 1 and cluster 2.*P < 0.05, **P < 0.01, and ***P < 0.001.

Figure 3

Identification hypoxic-related differentially expressed genes (DEGs) and enrichment analysis. (A) Volcano plot for DEGs in cluster1 and cluster2. Red and green dots represent up-regulated and down-regulated DEGs in cluster1 relative to cluster2, respectively ($P < 0.01$, $|\log FC| > 1$). (B) PPI network of a crucial sub-network with 119 nodes and 6153 edges among DEGs. (C-E) The top 10 items of GO analysis: biological processes, molecular functions and cellular components of the DEGs in the sub-network. (F) Pathway enrichment analysis of the DEGs in the sub-network.

Figure 4

Construction of prognostic models based on hypoxic-related DEGs. (A-B) LASSO Cox regression was conducted to construct the most powerful prognostic markers. (C) Identify 7 powerful prognostic markers and the coefficients by multivariate Cox regression via LASSO in train set (Wilcox Test). (D) Kaplan-Meier overall survival (OS) curves for patients in high- and low-risk group in train set ($P < 0.001$). (E) ROC curve and AUC value for risk score in train set. (F-G) Univariate and Multivariate Cox regression analysis of the associated between clinicopathological features (including risk score) in train set. (H) The heatmap shows the expression of the 7 powerful prognostic markers and the distribution of clinicopathological characteristics in high-risk group and low-risk group of all TCGA-LUAD samples (Chi-square Test). (I) The relationship between Risk score and cluster status, TNM stage, T stage and N stage (Wilcox Test). * $P < 0.05$, ** $P < 0.01$, and *** $P < 0.001$.

Figure 5

Identification prognosis-related differentially expressed microRNAs and construction of the ceRNA regulation network under hypoxic status. (A) Volcano plot for differentially expressed microRNAs in cluster1 and cluster2. Red and green dots represent up-regulated and down-regulated in cluster1 relative to cluster2, respectively ($P < 0.01$, $|\log FC| > 1$). (B-D) The overall survival curves of hsa-miR-196b, hsa-miR-31, hsa-miR-9 in TCGA-LUAD. (E) Venn diagrams showing the intersection between predicted target genes

of hsa-miR-196b/hsa-miR-31/hsa-miR-9 and DEGs. (F-L) The overall survival curves of PCSK2, FREM2, ALPL, ATP1A2, CEBPA in TCGA-LUAD. (K) The microRNA-lncRNA networks of the target lncRNAs of hsa-miR-196b, hsa-miR-31 and hsa-miR-9. (L) The ceRNA regulation network based on differentially expressed microRNAs and DEGs under hypoxic status.

Supplementary Files

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

- [SupplementaryFigures.docx](#)
- [SupplementaryTables.docx](#)