A promising prognostic signature for lung adenocarcinoma (LUAD) patients basing on 6 hypoxia-related genes

Jie Luo, PhD* , Xiaotian Du, MM

Abstract

Background: Hypoxia signaling plays a critical role in the development of lung adenocarcinoma (LUAD). We herein aimed to explore the prognostic value of hypoxia-related genes and construct the hypoxia-related prognostic signature for LUAD patients.

Methods: A total of 26 hypoxia-related genes were collected. Five hundred thirteen and 246 LUAD samples were obtained from the Cancer Genome Atlas and Gene Expression Omnibus databases, respectively. Univariate Cox regression and LASSO Cox regression analyses were conducted to screen the hypoxia-related genes associated with the prognosis of LUAD patients, which would be used for constructing prognosis predictive model for LUAD patients. Multivariate Cox regression analysis was done to determine the independent prognostic factors. The Nomogram model was constructed to predict the prognosis of LUAD patients.

Results: Based on 26 hypoxia-related genes, LUAD samples could be divided into 4 clusters with different prognoses. Among which, 6 genes were included to construct the Risk Score and the LUAD patients with higher Risk Score had worse prognosis. Besides, the Nomogram based on all the independent risk factors could relatively reliably predict the survival probability. And 9 types of immune cells' infiltration was significantly differential between high and low risk LUAD patients.

Conclusion: The Risk Score model based on the 6 crucial hypoxia-related genes could relatively reliably predict the prognosis of LUAD patients.

Abbreviations: EGFR = epidermal growth factor receptor, LUAD = lung adenocarcinoma, OS = overall survival, ROC = receiver operating characteristic, TCGA = the Cancer Genome Atlas, TNM = tumor, lymph node, metastasis.

Keywords: hypoxia-related genes, LASSO Cox regression, LUAD, overall survival, prognosis

1. Introduction

Lung cancer has consistently been one of the most malignant tumors and is classified into small-cell lung carcinomas and non-small cell lung carcinomas.[1] Lung adenocarcinoma (LUAD) is the most prevalent subtype of non-small cell lung carcinomas,[2] resulting from various factors.[3–5] Despite the advancements in cancer therapy, such as immunotherapy and non-invasive surgical resection, the 5-year overall survival (OS) of LUAD patients is still approximately 17.4%.[6,7] Recently, the clinical results of molecularly targeted therapies for LUAD patients are encouraging,[8,9] while drug resistance still blocks curing LUAD patients.[10,11] The prognosis of LUAD is still a major clinical challenge, which is remarkably different among the individual patients with LUAD.[12] Therefore, it is necessary to construct the prognostic signatures closely associated with patient survival, which might benefit and assist the selection and development of the more practical therapeutic strategies of LUAD.

Hypoxia is described as a mismatch of cellular oxygen demand and supply, which is a feature of solid cancers as a result of high proliferative rates, increased metabolism, and poor vascularization.[13] Hypoxia response is mainly mediated by the stabilization of the hypoxia-related genes, such as hypoxia inducible factor proteins, which could transcriptionally activate over 300 genes.[14] It has been evidenced that multiple cancers have been marked with heterogeneous hypoxia and hypoxia serves as an adverse prognostic feature.[15] Moreover, hypoxia is correlated with increasing metastasis and resistance to chemotherapy and radiotherapy, thereby leading to poor survival.[16] In some cancers, combining hypoxia-targeting treatment with

radiotherapy has been indicated to improve local control of tumors and OS.[17,18] Hypoxia gene signatures are successfully derived for multiple tumor sites, including cervical, head and neck, bladder, and soft tissue sarcoma cancers.[19] Besides, these hypoxia signatures are not only independent prognostic factors, but also predictors of benefiting from hypoxia-modifying therapy in head and neck and bladder cancers.[19] Previous investigation in prostate cancer shows the prognostic significance for signatures derived in other tumor types or associated with the hypoxia marker pimonidazole.[20] Importantly, hypoxia-related genes may serve as the prognostic biomarkers for tumors.[21] And recent reports have revealed the prognostic hypoxia-related signature in early-stage LUAD patients.[22,23] However, hypoxia related gene-based prognostic signature has not been explored in various stages' LUAD patients as far as we know.

In this study, we are interested in investigating the prognostic value of 26 hypoxia-related genes in LUAD patients, via integrating the bioinformatics analyses of LUAD mRNA data obtained from the public databases. Our findings would provide more alternatives for clinical strategies of LUAD patients.

2. Materials and methods

2.1. Data collection

The mRNA expression profiling data of 513 LUAD samples with complete survival information was downloaded from the Cancer Genome Atlas database (TCGA, https://tcga-data.nci.nih.gov/tcga/). After excluding patients with incomplete survival information, 500 patients were used for the following analysis. The detailed clinical information of the samples has been provided in Table 1. The mRNA expression profiling data with complete survival information of GSE31210 dataset contained 246 LUAD patients.[24] The expression value of the samples was analyzed using the Affymetrix Human Genome U133 Plus 2.0 Array platform.

2.2. Hypoxia-related genes

Our study used the 26 hypoxia-related genes which were previously reported in the effective hypoxic modification therapy for laryngeal cancer,[17] and the information of the 26 hypoxia-related genes was provided in Supplemental Digital Content (see Table S1, http://links.lww.com/MD2/A755. 26 Hypoxia-related genes).

2.3. Cluster analysis

The consensus cluster analysis was performed using *ConsensusClusterPlus* package of *R* based on the mRNA expression value of the 26 hypoxia-related genes.[25] And the principal component analysis was conducted on the samples as well.

2.4. LASSO Cox regression analysis

The correlation of hypoxia-related gene expression with OS of LUAD patients was evaluated by univariate Cox regression, in which the threshold was $P < .05$. The LASSO Cox regression analysis was performed to further identify the genes associated with the prognosis of LUAD using the *glmnet R* package.[26] The

Table 1

Clinicopathological characteristics of LUAD patients from TCGA database.

Characteristics		Patients (N=500)	
		No.	%
Gender	Female	270	54
	Male	230	46
Age	≤66 (median)	249	49.8
	>66 (median)	241	48.2
	Unknown	10	2
Race	White	188	37.6
	Asian	4	0.8
	Black or African American	37	7.4
	American Indian or Alaska native	1	0.2
	Unknown	19	3.8
Radiation therapy	No	361	72.2
	Yes	58	11.6
	Unknown	81	16.2
EGFR mutation	No	433	86.6
	Yes	63	12.6
	Unknown	4	0.8
Vital status	Alive	318	63.6
	Dead	182	36.4

EGFR = epidermal growth factor receptor, LUAD = lung adenocarcinoma, TCGA = the Cancer Genome Atlas.

Risk Score was calculated based on the selected genes using the following formula:

$$\text{Risk score} = \sum_{i=1}^{n} \text{Coef}_i * X_i$$

The $Coef_i$ was the risk coefficient of each factor calculated by the LASSO-Cox model, and X_i was the mRNA expression value of the genes. And the samples were separated into low-risk and high-risk sets based on the cutoff value.

2.5. Survival analysis

The OS was analyzed by the *survival R* package and *survminer R* package based on the Kaplan-Meier method.[27] The difference significance was evaluated by the two-sided log-rank test. The time-dependent receiver operating characteristic (ROC) curve was generated by the *survivalROC R* package.[28]

2.6. Analysis of immune infiltration

The immune infiltration difference of 22 immune cells in the samples was analyzed by using CIBERSORT software combined with the LM22 feature matrix.[29] The sum of the proportions of all estimated immune cells in each sample is equal to 1.

2.7. Nomogram construction

A concise nomogram of predicting the survival of LUAD patients was established using the *rms R* package based on all independent prognostic factors identified by multivariate Cox regression analysis.[30] The calibration curve of the nomogram was obtained, and the relationship between the predicted probability of nomogram and the actual incidence rate was determined. The $P < .05$ was considered statistically significant.

2.8. Statistical analysis

The difference of immune cell infiltration in the samples was analyzed by the Wilcoxon rank-sum test with the threshold of $P < .05$. The statistical analysis was done in R software (version v3.5.2).

3. Results

3.1. The 26 hypoxia-related genes can effectively distinguish the LUAD samples with different prognosis

To evaluate the correlation of 26 hypoxia-related genes with the prognosis of LUAD patients, 500 LUAD samples with complete survival information from TCGA database were used to the consensus cluster analysis, and the cluster number k = 4 was selected according to the cumulative distribution function of the cluster (Fig. 1A and B). The consensus matrix diagram (Fig. 1C) and the heat map (Fig. 1D) showed that the consensus clustering effect was effective and could clearly distinguish 4 clusters. The principal component analysis presented a significant difference among the 4 clusters (Fig. 1E). Kaplan-Meier survival analysis suggested that 4 clusters of patients demonstrated remarkable difference of OS, and the prognosis of the Cluster2 samples was the worst (Fig. 1F).

3.2. The Risk Score model based on the identified 6 hypoxia-related genes can reliably predict the prognosis of LUAD patients

To explore the correlation of hypoxia-related genes with the prognosis of LUAD, we further constructed the prognostic model based on the hypoxia-related genes. The LUAD samples from TCGA database were subjected to univariate Cox regression analysis, in which the expression of the hypoxia-related genes served as the continuous variable. Our data revealed that 10 genes were significantly associated with the OS of the LUAD samples ($P < .05$) (Fig. 2A), including *ALDOA, ANGPTL4, BNC1, CDKN3, FAM83B, KRT17, LDHA, SLC16A1, SLC2A1,* and *TPI1*. Moreover, the hazard ratio of these 10 genes were greater than 1, which indicated that high expressions of these genes were potentially related to poor prognosis of LUAD patients. These significant genes were entered into LASSO COX regression analysis to construct the prognostic model for LAUD patients, and the model achieved the best performance with 6 genes, including *ANGPTL4, BNC1, CDKN3, FAM83B, LDHA,* and *SLC16A1* (Fig. 2B and C). The Risk Score of the signature for OS was identified: Risk Score = 0.0767 * Express Value of *ANGPTL4* + 0.0307 * Express Value of *BNC1* + 0.1034 * Express Value of *CDKN3* + 0.0386 * Express Value of *FAM83B*
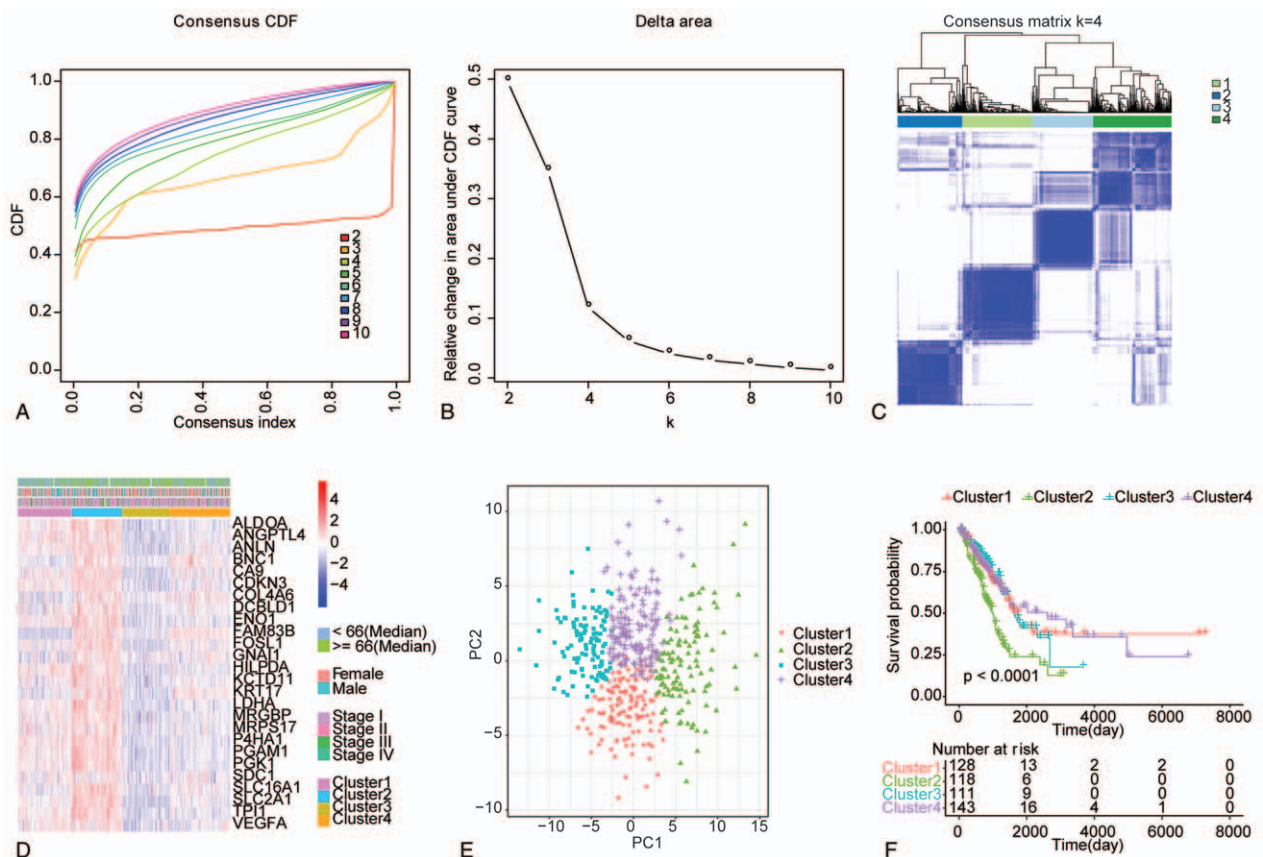

Figure 1. Cluster analysis of LUAD samples based on hypoxia-related genes. (A) The cumulative distribution function (CDF) graph of consensus cluster analysis with k 2-10 was shown. (B) Relative change of the area under the CDF curve with k 2-10 was presented. The x-axis was the number of k, which was the number of clusters, and the y-axis was the relative change of the area under the CDF curve. (C) The consensus matrix diagram with k = 4 was shown. (D) Heatmap of the expression levels of 26 hypoxia-related genes in 4 clusters was shown. Behavioral genes were listed as samples. Red indicated high expression and blue indicated low expression. (E) The principal component analysis (PCA) was performed. The dots with different colors represented samples in different groups. The distance of the dots revealed the similarity of the hypoxia-related gene expression. (F) The Kaplan-Meier survival curve was shown. The x-axis was time and the y-axis was survival rate. Different colors represented different groups. The P value was calculated based on the log-rank test. LUAD = lung adenocarcinoma.
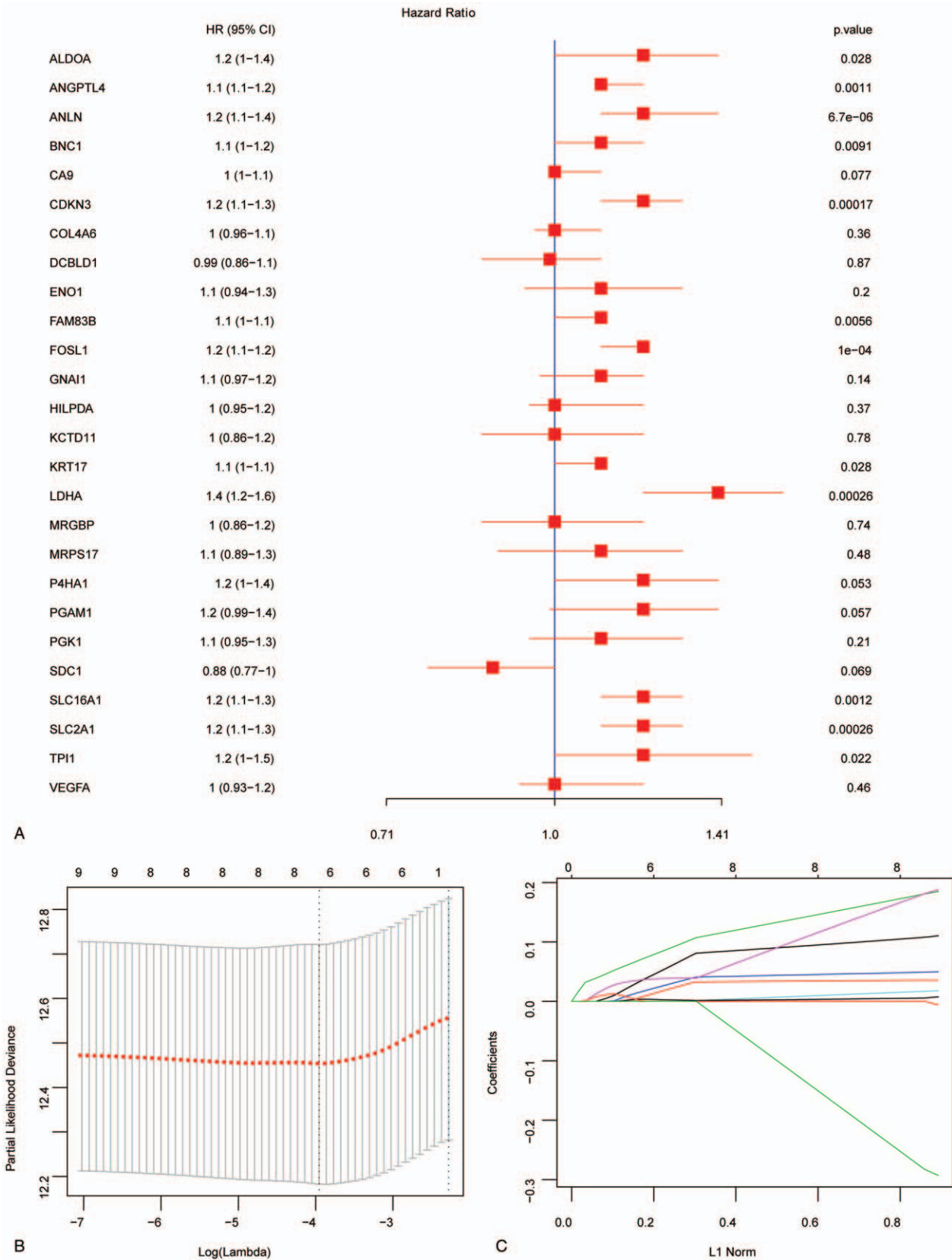
Figure 2. Construction of prognostic model of LUAD patients. (A) A total of 10 hypoxia-related genes were identified by Cox regression analysis in LUAD samples, in which the expression of the identified genes served as the continuous variable. The result was shown in the forest map, in which HR represented hazard ratio and 95% CI represented 95% confidence interval. (B) The optimal lambda value was selected in the LASSO COX regression analysis. (C) Coefficient spectrum of LASSO Cox regression model was shown. LUAD = lung adenocarcinoma.

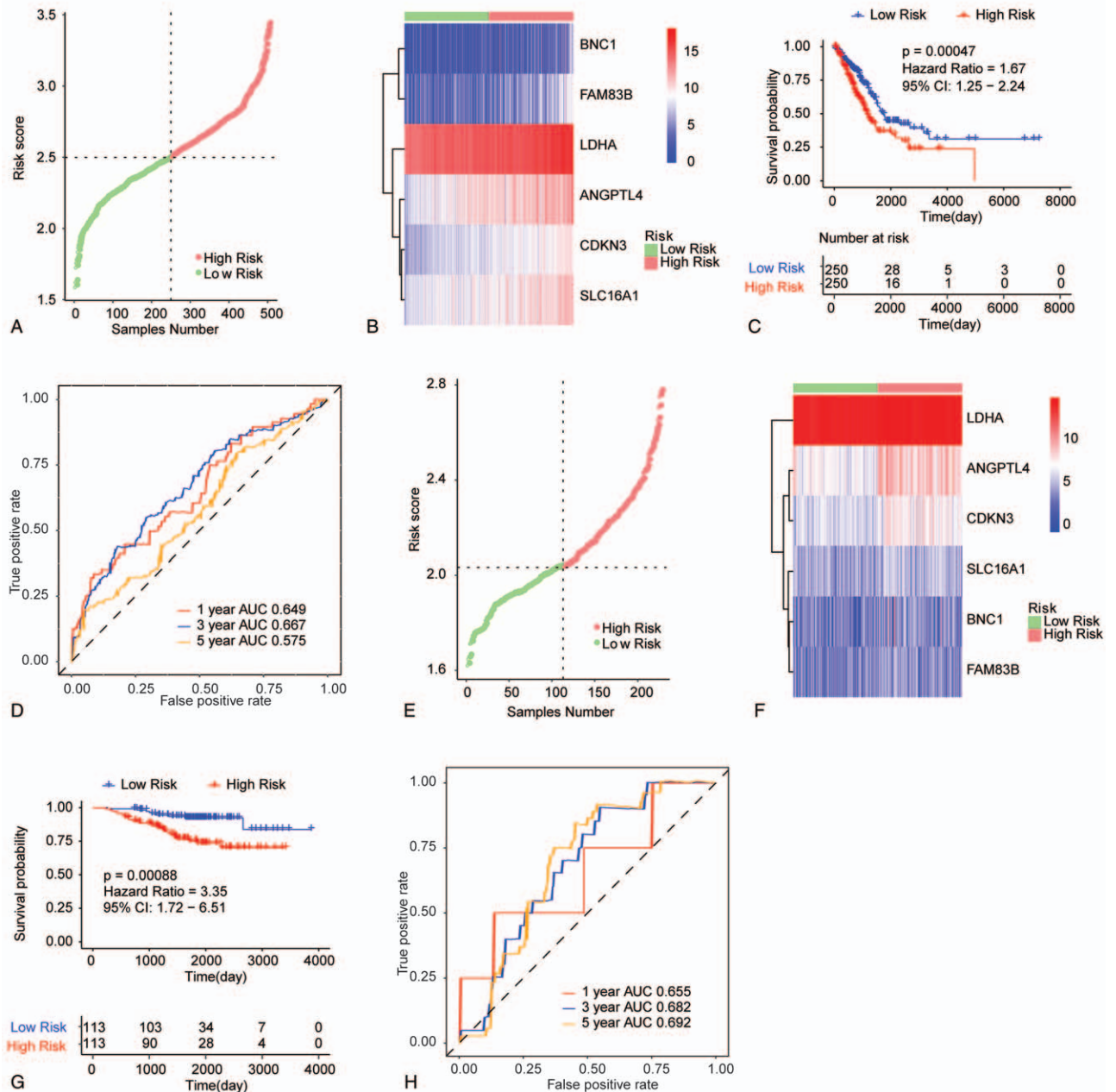**Figure 3.** The Risk Score model based on the identified 6 hypoxia-related genes can reliably predict the prognosis of LUAD patients. (A) The distribution of Risk Scores of the LUAD samples in TCGA dataset was shown. The dot represented the samples, the red dot represented the samples with the higher Risk Score, and the green dot represented the samples with the lower Risk Score. (B) The clustering heatmap of the expression levels of 6 hypoxia-related genes in TCGA dataset was shown. Behavioral genes were listed as the samples. Red indicated high expression and blue indicated low expression. (C) Kaplan-Meier survival curve of TCGA dataset was shown. The x-axis was time and the y-axis was survival rate. Different colors represented different groups. (D) The time-dependent ROC analysis was performed in TCGA dataset. The x-axis was the false positive rate (False Positive), the y-axis was the true positive rate (True Positive), and the accuracy of the prediction was evaluated by the AUC value. (E) The distribution of Risk Scores of the LUAD samples in the GEO dataset was shown. (F) The clustering heatmap of the expression levels of 6 hypoxia-related genes in the GEO dataset was shown. (G) Kaplan-Meier survival curve of GEO dataset was shown. (H) The time-dependent ROC analysis was performed in the GEO dataset. AUC = area under the ROC curve, GEO = Gene Expression Omnibus, LUAD = lung adenocarcinoma, ROC = receiver operating characteristic, TCGA = the Cancer Genome Atlas.

$+0.039$ * Express Value of *LDHA* $+0.0017$ * Express Value of *SLC16A1*. Then all the LUAD samples from TCGA cohort and GSE31210 of Gene Expression Omnibus cohort were separated into high-risk group and low-risk group based on the median of the Risk Score (Fig. 3A and E). Besides, the expression levels of the hypoxia-related genes in the model were significantly different between high-risk group and low-risk group (Fig. 3B and F). The

survival analysis demonstrated that the OS of LAUD patients in low-risk group was longer than that in high-risk group (Fig. 3C and G). Moreover, ROC analysis showed that the 1-year, 3-year, and 5-year area under the ROC curve of patients with LUAD from TCGA database were 0.649, 0.667, and 0.575, respectively (Fig. 3D). The 1-year, 3-year, and 5-year area under the ROC curve of LUAD patients from the GSE31210 dataset were 0.655,

0.682, and 0.692, respectively (Fig. 3H), indicating that the Risk Score model can effectively predict the prognosis of patients with LUAD in both datasets. Together these data suggested that the Risk Score model based on the identified 6 hypoxia-related genes could relatively reliably predict the prognosis of LUAD patients.

3.3. Immunosuppressive microenvironment may contribute to the poor prognosis of LUAD patients

Then, the immune infiltration difference of 22 immune cells was analyzed and presented using CIBERSORT software combined with the LM22 feature matrix. The infiltration ratios of immune cells were different in the LUAD samples from TCGA database (Fig. 4A). The correlation between the infiltration ratios of different types of immune cells was relatively weak (Fig. 4B). The significant difference of 9 types of immune cells' infiltration was identified in the low-risk group compared with the high-risk group (Fig. 4C), including B cells naive, B cells memory, T cells

CD4 memory resting, T cells CD4 memory activated, NK cells resting, Macrophages M0, Macrophages M1, Mast cells resting, and Neutrophils, implying that the differential immune infiltration might affect the prognosis of LUAD.

The expressions of immune checkpoints have become biomarkers for choosing immunotherapy of LUAD patients.[31–33] Therefore, we analyzed the correlation between the Risk Score and the key immune checkpoints, including CTLA4, PDL1, PDL2, TIM3, LAG3, and TIGIT. Our data showed that the Risk Score was significantly correlated with these immune checkpoints (Fig. 5A) (see Table S2, Supplemental Digital Content, http://links.lww.com/MD2/A756. Analysis of correlations between Risk Score and immune checkpoints). Meanwhile, the expressions of these immune checkpoints were remarkably higher in the high-risk set compared with that in low-risk set ($P < .05$) (Fig. 5B), indicating that immunosuppressive microenvironment may contribute to the poor prognosis of LUAD patients.
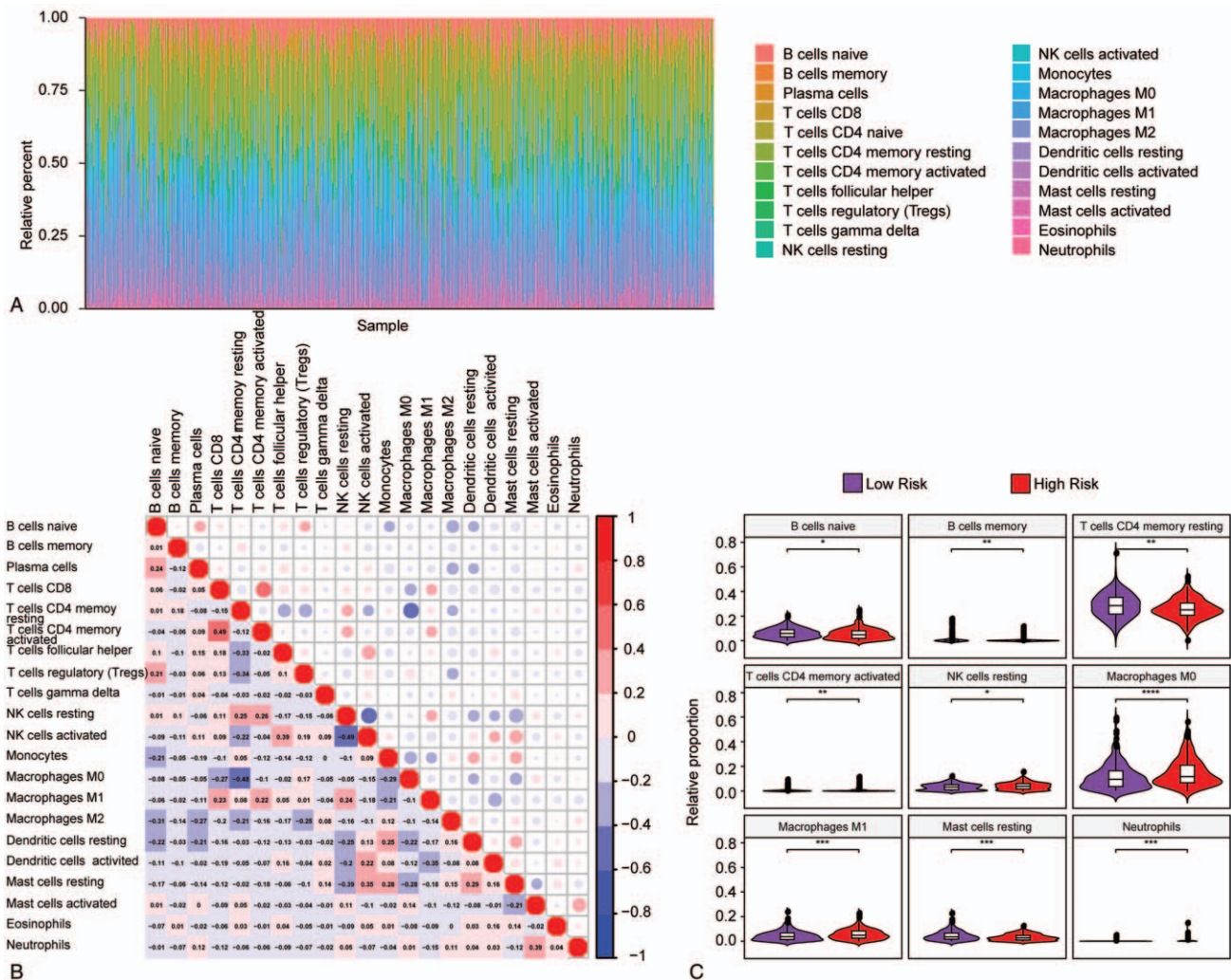

Figure 4. Immune cell infiltration in LUAD patients of high and low risk groups. (A) The immune infiltration difference of 22 immune cells in LUAD patients was analyzed using CIBERSORT software combined with the LM22 feature matrix. (B) The correlation matrix of 22 immune cell infiltration was shown. Red represented a positive correlation and blue represented a negative correlation. The darker the color, the greater the correlation. (C) The infiltration difference of 9 types of immune cells was analyzed in the low-risk set compared with the high-risk set of the samples. The difference of immune cell infiltration in the samples was analyzed by the Wilcoxon rank-sum test with the threshold of $P < .05$, $^{*}P < .05$, $^{**}P < .01$, $^{***}P < .001$, $^{****}P < .0001$. LUAD = lung adenocarcinoma.
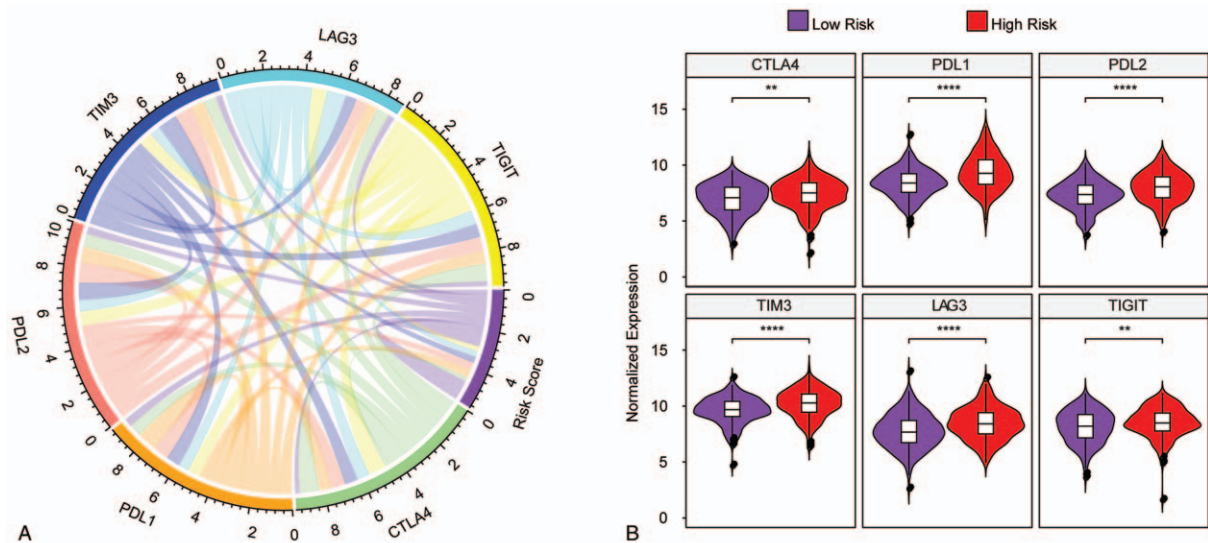
Figure 5. Expression of immune checkpoints in LUAD patients of high and low risk groups. (A) The correlation of the Risk Score with the expression of the 6 key immune checkpoints, including CTLA4, PDL1, PDL2, TIM3, LAG3, and TIGIT, was presented in a Chord diagram. (B) The expression of the immune checkpoints was analyzed in the low-risk set compared with the high-risk set of LUAD patients. LUAD = lung adenocarcinoma.
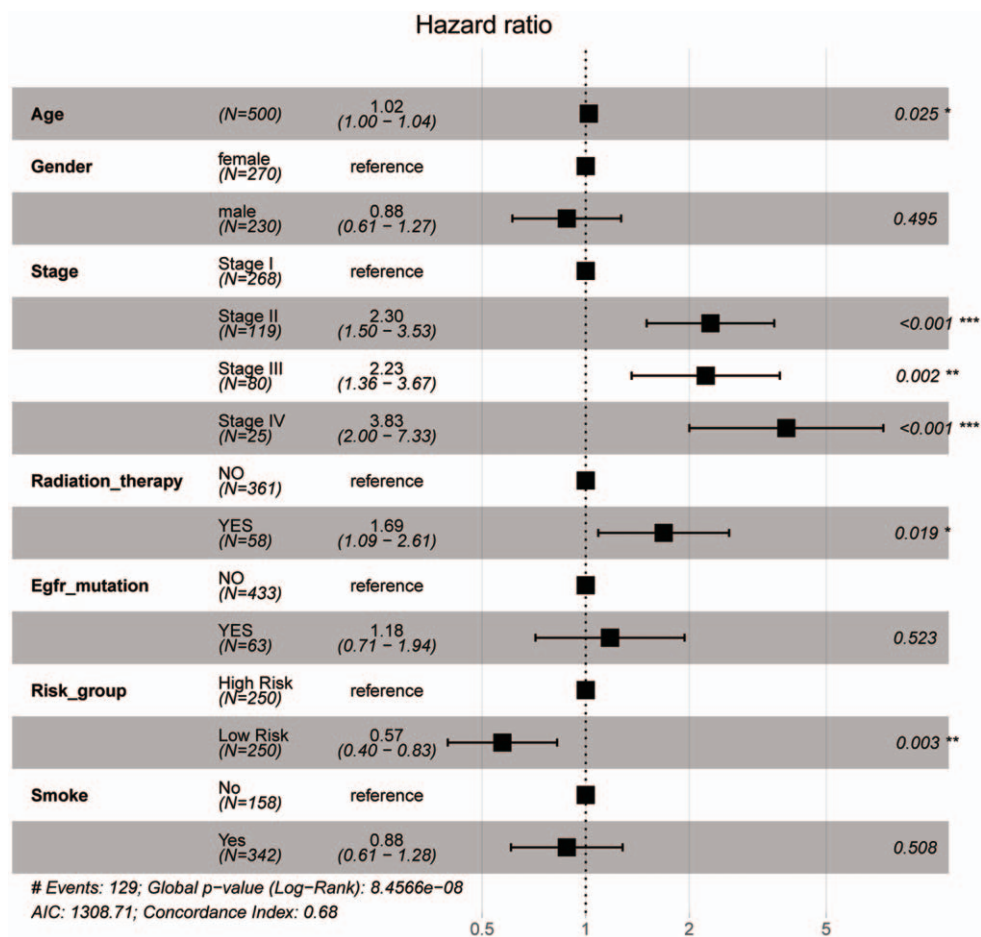
Figure 6. The Risk Score can be used as an independent risk factor to predict the prognosis of LUAD patients. The multivariate Cox regression analysis, containing age, gender, TNM stage, radiation therapy, EGFR mutation, and the Risk Score, was performed in the LUAD samples from TCGA database. Compared with the reference samples, the samples with hazard ratio larger than 1 represented a higher risk of death, and the samples with hazard ratio less than 1 represented a lower risk of death. EGFR = epidermal growth factor receptor, LUAD = lung adenocarcinoma, TCGA = the Cancer Genome Atlas, TNM = tumor, lymph node, metastasis.

3.4. The Risk Score can be used as an independent risk factor to predict the prognosis of LUAD patients

To further evaluate whether the Risk Score was an independent signature, we performed the multivariate Cox regression analysis, containing age, gender, tumor, lymph node, metastasis (TNM) stage, radiation therapy, epidermal growth factor receptor (EGFR) mutation, smoking status, and the Risk Score. Our data confirmed that the age, TNM stage, radiation therapy, and Risk Score were closely associated with the OS of the patients, which were independent prognostic factors for LUAD patients. Those LUAD patients with high-Risk Score had a high risk of death and Risk Score was a poor prognostic factor (hazard ratio = 4.64; 95% CI [2.24–9.62]; $P < .001$) (Fig. 6).

In order to further explore the prognostic value of Risk Score in LUAD patients with different clinicopathological factors, including gender, TNM stage, and EGFR mutation, we regrouped LUAD samples according to these factors and performed Kaplan-Meier survival analysis. We found that the OS of patients in the high-risk sets was significantly lower than that in low-risk sets of patients in female samples (Fig. 7A), male samples (Fig. 7B), Stage I + Stage II (Fig. 7C), Stage III + Stage IV samples (Fig. 7D), EGFR wild-type samples (Fig. 7E), and EGFR mutation samples (Fig. 7F). Together these data indicated that the Risk Score can be used as an independent risk factor to predict the prognosis of LUAD patients.

3.5. The nomogram model can reliably predict the prognosis of LUAD patients

Next, we constructed a nomogram model based on all independent prognostic factors, including the age, TNM stage, Radiation

therapy, and Risk Score (Fig. 8A). The results for predicting 1-year, 3-year, and 5-year of OS probability showed that the nomogram-predicted survival probability closely matched with the actual survival probability (Fig. 8B–D), suggesting that the nomogram model can reliably predict the prognosis of LUAD patients.

4. Discussion

Lung cancer serves as the most prevailing malignancy type and leads to the most cancer-related deaths.[34] LUAD with poor prognosis is one of the major subtypes of lung cancer.[35] Increasing evidence has demonstrated that the hypoxia signaling is involved in modulating the development of LUAD, and hypoxia-related genes may serve as the prognostic biomarkers for LUAD patients. It has been reported that hypoxia-stimulated expression of *GBE1* contributes to LUAD progression by regulating metabolic reprogramming.[36] Hypoxia-inducible factor-1α is correlated with the genetic aberrations in LUAD.[37] The previous study also showed the identification and validation of the hypoxia-related gene signature to predict OS of patients with early-stage LUAD.[22] Genome-wide analysis reveals the roles of hypoxia-related DNA methylation-driven genes in the progression of LUAD.[38] Hypoxia-induced cell stemness causes drug resistance and poor prognosis of LUAD.[39] In this study, consensus cluster analysis based on 26 hypoxia-related genes showed that 4 clusters of the LUAD samples presented significantly different OS, implying that these hypoxia-related genes may contribute to the different prognosis of LUAD patients. Furthermore, the Risk Score based on the 6 key hypoxia-related genes could relatively reliably predict the prognosis of LUAD patients.
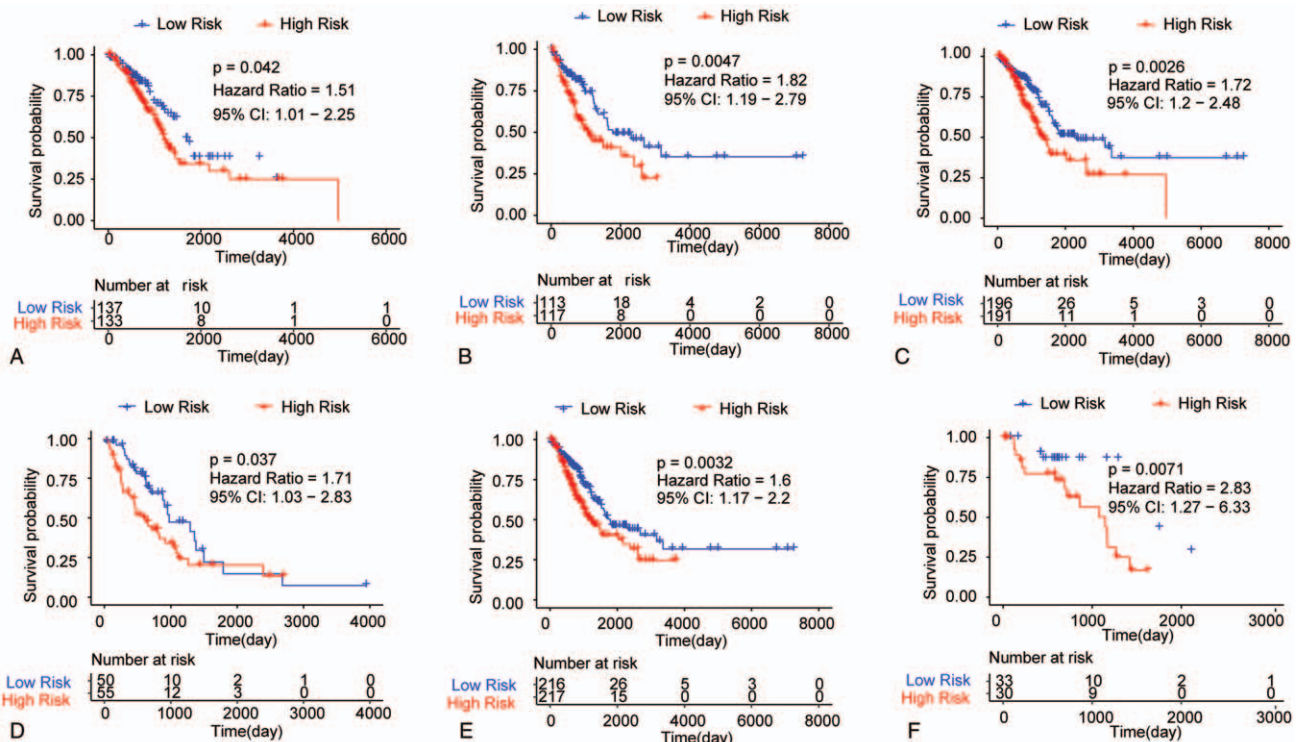
Figure 7. Survival curves of LUAD samples with different clinicopathological factors. (A–F) The Kaplan-Meier OS analysis was conducted in high-risk and low-risk sets of the LUAD patients with different clinicopathological factors, including Gender, TNM Stage, and EGFR mutation. The x-axis was time and the y-axis was survival rate. Different colors represented different groups. The P value was calculated based on the log-rank test. EGFR = epidermal growth factor receptor, LUAD = lung adenocarcinoma, OS = overall survival, TNM = tumor, lymph node, metastasis.

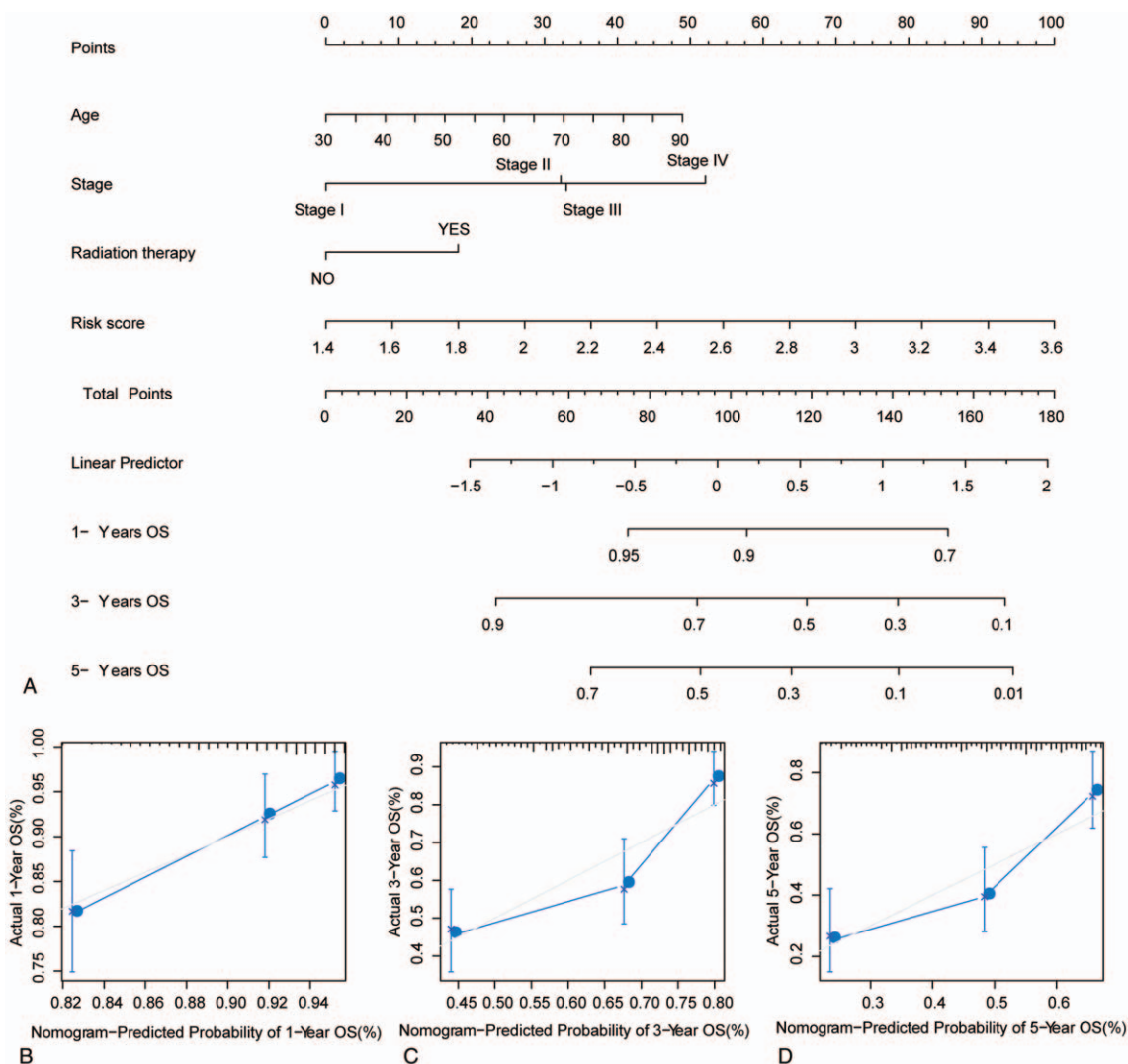**Figure 8.** The nomogram model can reliably predict the prognosis of LUAD patients. (A) The nomogram model consisted of age, TNM stage, radiation therapy, and Risk Score was constructed to predict 1-year, 3-year, and 5-year OS of LUAD patients. (B–D) The calibration curves of the nomogram for the estimation of survival rates at 1-year, 3-year, and 5-year was shown, respectively. The x-axis represented the predicted survival rate of the nomogram, and the y-axis represented the actual survival rate. LUAD = lung adenocarcinoma, OS = overall survival, TNM = tumor, lymph node, metastasis.

The development of LUAD is complicated and multiple gene contributors participate in the progression of LUAD.[40] In the present study, we have firstly identified 10 hypoxia-related genes associated with the OS of the LUAD samples. Moreover, LASSO COX regression analysis was conducted to optimize the genes and 6 genes were more closely associated with the prognosis of LUAD patients, including *ANGPTL4, BNC1, CDKN3, FAM83B, LDHA,* and *SLC16A1*, based on which, the Risk Score was constructed. Abnormally expressed genes could be served as the potential prognostic biomarkers of LUAD. For example, isoform specific gene expression analysis reveals the prognostic value of *KRAS* in LUAD patients.[41] Elevated *MACC1* expression is able to predict the poor prognosis in LUAD.[42] Moreover, previous studies have identified that some of the hypoxia-related genes participate in the progression of LUAD. It has been reported that *ALDOA* promotes LUAD metastasis by interacting with γ-actin.[43] Besides, *ANGPTL4* is closely associated with the metastasis and OS of LUAD patients.[44] *BNC1* is aberrantly expressed in the LUAD samples.[45] The overexpression of *CDKN3* and *LDHA* is

correlated with poor survival of LUAD.[46,47] *FAM83B* and *SLC16A1* have been reported as biomarkers for the prognosis of LUAD.[48,49] *KRT17* contributes to LUAD progression by promoting cell proliferation and invasion.[50] *SLC2A1* inhibits the metabolic reprogramming to repress LUAD progression.[51] All the above evidence could support our findings directly or indirectly. Additionally, the OS of low-risk LUAD patients was better than that of high-risk LUAD patients. And the nomogram model based on the independent risk factors, including age, TNM stage, radiation therapy, and Risk Score, showed that the predicted survival closely matched with the best prediction performance. Collectively, our findings were consistent with the previous reports about the correlation of these genes with LUAD, further suggesting that our Risk Score and nomogram model could reliably predict the prognosis of LUAD patients.

Immune microenvironment plays a critical role in the development of LUAD.[52] Meanwhile, the immune checkpoints are involved in the modulation of LUAD progression.[53] Moreover, immune checkpoints, such as PDL1, are closely correlated with

hypoxia signaling in LUAD.[54] Our immune infiltration analysis revealed that 9 types of immune cells' infiltration was significantly differential between high and low risk LUAD patients, comprising B cells naive, B cells memory, T cells CD4 memory resting, T cells CD4 memory activated, NK cells resting, Macrophages M0, Macrophages M1, Mast cells resting, and Neutrophils. The Risk Score was remarkably correlated with the immune checkpoints, including CTLA4, PDL1, PDL2, TIM3, LAG3, and TIGIT. Our data further imply the potential correlation of immune infiltration and immune checkpoints with the hypoxia-related genes in LUAD patients. These data suggested that the immunosuppressive microenvironment may contribute to the poor prognosis of LUAD patients.

However, considering the complex pathopoiesis and heterogeneity LUAD, there were still some limitations in our present study. Firstly, our prognostic signature for LUAD patients based on hypoxia-related genes was only validated in 1 independent dataset, additional validation datasets would improve the reliability of the signature. Besides, our study had initially explored the prognostic value of hypoxia-related genes in LUAD patients, more details of the deepening mechanism involving these hypoxia-related genes in LUAD should be further investigated.

5. Conclusion

In conclusion, the 26 hypoxia-related genes can effectively distinguish the LUAD samples with differential prognoses. The Risk Score model based on the 6 hypoxia-related genes, including *ANGPTL4, BNC1, CDKN3, FAM83B, LDHA,* and *SLC16A1*, and the nomogram model based on the independent prognostic signature, could relatively reliably predict the prognosis of LUAD patients. Our findings provide valuable alternative predictive models which might contribute to improve the prognosis of LUAD patients.

Author contributions

Conceptualization: Jie Luo.
Data curation: Jie Luo, Xiaotian Du.
Funding acquisition: Jie Luo.
Resources: Jie Luo.
Software: Jie Luo, Xiaotian Du.
Supervision: Jie Luo, Xiaotian Du.
Validation: Jie Luo, Xiaotian Du.
Visualization: Jie Luo, Xiaotian Du.
Writing – original draft: Jie Luo, Xiaotian Du.
Writing – review & editing: Jie Luo, Xiaotian Du.

References

[1] Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C. Progress and prospects of early detection in lung cancer. Open Biol 2017;7:170070.

[2] Devarakonda S, Morgensztern D, Govindan R. Genomic alterations in lung adenocarcinoma. Lancet Oncol 2015;16:e342–51.

[3] Zhang C, Zhang G, Sun N, et al. Comprehensive molecular analyses of a TNF family-based signature with regard to prognosis, immune features, and biomarkers for immunotherapy in lung adenocarcinoma. EBioMedicine 2020;59:102959.

[4] Nicoleau S, Wojciak-Stothard B. Beyond thrombosis: the role of platelets in pulmonary hypertension. SciMed J 2020;2:243–71.

[5] Kosvyra A, Maramis C, Chouvarda I. Developing an integrated genomic profile for cancer patients with the use of NGS data. Emerg Sci J 2019;3:157–67.

[6] Denisenko TV, Budkevich IN, Zhivotovsky B. Cell death-based treatment of lung adenocarcinoma. Cell Death Dis 2018;9:117.

[7] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018;68:7–30.

[8] Kobayashi Y, Mitsudomi T. Not all epidermal growth factor receptor mutations in lung cancer are created equal: perspectives for individualized treatment strategy. Cancer Sci 2016;107:1179–86.

[9] Karachaliou N, Santarpia M, Gonzalez Cao M, et al. Anaplastic lymphoma kinase inhibitors in phase I and phase II clinical trials for non-small cell lung cancer. Expert Opin Investig Drugs 2017;26:713–22.

[10] Sacco JJ, Al-Akhrass H, Wilson CM. Challenges and strategies in precision medicine for non-small-cell lung cancer. Curr Pharm Des 2016;22:4374–85.

[11] Lin JJ, Riely GJ, Shaw AT. Targeting ALK: precision medicine takes on drug resistance. Cancer Discov 2017;7:137–55.

[12] Calvayrac O, Pradines A, Pons E, Mazieres J, Guibert N. Molecular biomarkers for lung adenocarcinoma. Eur Respir J 2017;49:1601734.

[13] Harris AL. Hypoxia–a key regulatory factor in tumour growth. Nat Rev Cancer 2002;2:38–47.

[14] Semenza GL. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. Annu Rev Pathol 2014;9:47–71.

[15] Milosevic M, Warde P, Menard C, et al. Tumor hypoxia predicts biochemical failure following radiotherapy for clinically localized prostate cancer. Clin Cancer Res 2012;18:2108–14.

[16] 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;102:428–35.

[17] Eustace A, Mani N, Span PN, et al. A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in laryngeal cancer but not bladder cancer. Clin Cancer Res 2013;19:4879–88.

[18] Overgaard J. Hypoxic modification of radiotherapy in squamous cell carcinoma of the head and neck–a systematic review and meta-analysis. Radiother Oncol 2011;100:22–32.

[19] Betts GN, Eustace A, Patiar S, et al. Prospective technical validation and assessment of intra-tumour heterogeneity of a low density array hypoxia gene profile in head and neck squamous cell carcinoma. Eur J Cancer 2013;49:156–65.

[20] Ragnum HB, Vlatkovic L, Lie AK, et al. The tumour hypoxia marker pimonidazole reflects a transcriptional programme associated with aggressive prostate cancer. Br J Cancer 2015;112:382–90.

[21] Fjeldbo CS, Hompland T, Hillestad T, et al. Combining imaging- and gene-based hypoxia biomarkers in cervical cancer improves prediction of chemoradiotherapy failure independent of intratumour heterogeneity. EBioMedicine 2020;57:102841.

[22] Sun J, Zhao T, Zhao D, et al. Development and validation of a hypoxia-related gene signature to predict overall survival in early-stage lung adenocarcinoma patients. Ther Adv Med Oncol 2020;12: 1758835920937904.

[23] Shi R, Bao X, Unger K, et al. Identification and validation of hypoxia-derived gene signatures to predict clinical outcomes and therapeutic responses in stage I lung adenocarcinoma patients. Theranostics 2021;11: 5061–76.

[24] Yamauchi M, Yamaguchi R, Nakata A, et al. Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma. PLoS One 2012;7:e43923.

[25] Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010;26: 1572–3.

[26] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1–22.

[27] Li M, Spakowicz D, Burkart J, et al. Change in neutrophil to lymphocyte ratio during immunotherapy treatment is a non-linear predictor of patient outcomes in advanced cancers. J Cancer Res Clin Oncol 2019;145:2541–6.

[28] Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics 2000;56: 337–44.

[29] Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12: 453–7.

[30] Li W, Liu J, Zhao H. Identification of a nomogram based on long non-coding RNA to improve prognosis prediction of esophageal squamous cell carcinoma. Aging (Albany NY) 2020;12:1512–26.

[31] Mezquita L, Auclin E, Ferrara R, et al. Association of the lung immune prognostic index with immune checkpoint inhibitor outcomes in patients with advanced non-small cell lung cancer. JAMA Oncol 2018;4:351–7.

[32] McGranahan N, Furness AJ, Rosenthal R, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 2016;351:1463–9.

[33] Dong ZY, Zhong WZ, Zhang XC, et al. Potential predictive value of TP53 and KRAS mutation status for response to PD-1 blockade immunotherapy in lung adenocarcinoma. Clin Cancer Res 2017;23:3012–24.

[34] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68: 394–424.

[35] Cancer Genome Atlas Research NetworkComprehensive molecular profiling of lung adenocarcinoma. Nature 2014;511:543–50.

[36] Li L, Yang L, Fan Z, et al. Hypoxia-induced GBE1 expression promotes tumor progression through metabolic reprogramming in lung adenocarcinoma. Signal Transduct Target Ther 2020;5:54.

[37] Putra AC, Tanimoto K, Arifin M, Hiyama K. Hypoxia-inducible factor-1alpha polymorphisms are associated with genetic aberrations in lung cancer. Respirology 2011;16:796–802.

[38] Li H, Tong L, Tao H, Liu Z. Genome-wide analysis of the hypoxia-related DNA methylation-driven genes in lung adenocarcinoma progression. Biosci Rep 2020;40:BSR20194200.

[39] Zhao M, Zhang Y, Zhang H, et al. Hypoxia-induced cell stemness leads to drug resistance and poor prognosis in lung adenocarcinoma. Lung Cancer 2015;87:98–106.

[40] Liu Q, Li A, Tian Y, et al. The expression profile and clinic significance of the SIX family in non-small cell lung cancer. J Hematol Oncol 2016;9:119.

[41] Yang IS, Kim S. Isoform specific gene expression analysis of KRAS in the prognosis of lung adenocarcinoma patients. BMC Bioinformatics 2018;19(Suppl 1):40.

[42] Guo T, Zhao S, Li Z, Li F, Li J, Gu C. Elevated MACC1 expression predicts poor prognosis in small invasive lung adenocarcinoma. Cancer Biomark 2018;22:301–10.

[43] Chang YC, Chiou J, Yang YF, et al. Therapeutic targeting of aldolase A interactions inhibits lung cancer metastasis and prolongs survival. Cancer Res 2019;79:4754–66.

[44] Zhang L, Zhang Z, Yu Z. Identification of a novel glycolysis-related gene signature for predicting metastasis and survival in patients with lung adenocarcinoma. J Transl Med 2019;17:423.

[45] Yuan F, Lu L, Zou Q. Analysis of gene expression profiles of lung cancer subtypes with machine learning algorithms. Biochim Biophys Acta Mol Basis Dis 2020;1866:165822.

[46] Fan C, Chen L, Huang Q, et al. Overexpression of major CDKN3 transcripts is associated with poor survival in lung adenocarcinoma. Br J Cancer 2015;113:1735–43.

[47] Yu C, Hou L, Cui H, et al. LDHA upregulation independently predicts poor survival in lung adenocarcinoma, but not in lung squamous cell carcinoma. Future Oncol 2018;14:2483–92.

[48] Okabe N, Ezaki J, Yamaura T, et al. FAM83B is a novel biomarker for diagnosis and prognosis of lung squamous cell carcinoma. Int J Oncol 2015;46:999–1006.

[49] Stewart PA, Parapatics K, Welsh EA, et al. A pilot proteogenomic study with data integration identifies MCT1 and GLUT1 as prognostic markers in lung adenocarcinoma. PLoS One 2015;10:e0142162.

[50] Liu J, Liu L, Cao L, Wen Q. Keratin 17 promotes lung adenocarcinoma progression by enhancing cell proliferation and invasion. Med Sci Monit 2018;24:4782–90.

[51] Wang Y, Shi S, Ding Y, et al. Metabolic reprogramming induced by inhibition of SLC2A1 suppresses tumor progression in lung adenocarcinoma. Int J Clin Exp Pathol 2017;10:10759–69.

[52] Lavin Y, Kobayashi S, Leader A, et al. Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. Cell 2017;169:750–65.e17.

[53] Skoulidis F, Byers LA, Diao L, et al. Co-occurring genomic alterations define major subsets of KRAS-mutant lung adenocarcinoma with distinct biology, immune profiles, and therapeutic vulnerabilities. Cancer Discov 2015;5:860–77.

[54] Koh J, Jang JY, Keam B, et al. EML4-ALK enhances programmed cell death-ligand 1 expression in pulmonary adenocarcinoma via hypoxia-inducible factor (HIF)-1alpha and STAT3. Oncoimmunology 2016;5: e1108514.