Original Article

Hypoxia-related prognostic model in bladder urothelial reflects immune cell infiltration

Luanfeng Li\textsuperscript{1,2,3,4,5}, Wensi Liu\textsuperscript{1,2,3,4}, Haichao Tang\textsuperscript{1,2,3,4}, Xiangyi Wang\textsuperscript{1,2,3,4}, Xinli Liu\textsuperscript{6}, Zhaojin Yu\textsuperscript{1,2,3,4}, Yanan Gao\textsuperscript{1,2,3,4}, Xiaobin Wang\textsuperscript{7}, Minjie Wei\textsuperscript{1,2,3,4,5}

\textsuperscript{1}Department of Pharmacology, School of Pharmacy, China Medical University, Shenyang 110122, Liaoning, China; \textsuperscript{2}Liaoning Key Laboratory of Molecular Targeted Anti-Tumor Drug Development and Evaluation, Shenyang 110122, Liaoning, China; \textsuperscript{3}Liaoning Cancer immune peptide drug Engineering Technology Research Center, Shenyang 110122, Liaoning, China; \textsuperscript{4}Key Laboratory of Precision Diagnosis and Treatment of Gastrointestinal Tumors (China Medical University), Ministry of Education, Shenyang 110122, Liaoning, China; \textsuperscript{5}Shenyang Kangwei Medical Laboratory Analysis Co. LTD, Shenyang, Liaoning, China; \textsuperscript{6}Medical Oncology Department of Gastrointestinal Cancer, Liaoning Cancer Hospital & Institute, Cancer Hospital of China Medical University, Shenyang 110042, Liaoning, China; \textsuperscript{7}Center of Reproductive Medicine, Shengjing Hospital of China Medical University, Shenyang 117004, Liaoning, China

Received July 5, 2021; Accepted September 9, 2021; Epub October 15, 2021; Published October 30, 2021

Abstract: Hypoxia is a common feature of tumor microenvironment (TME). This study aims to establish the genetic features related to hypoxia in Bladder urothelial carcinoma (BLCA) and investigate the potential correlation with hypoxia in the TME and immune cells. We established a BLCA outcome model using the hypoxia-related genes from The Cancer Genome Atlas using regression analysis and verified the model using the Gene Expression Omnibus GSE32894 cohort. We measured the effect of each gene in the hypoxia-related risk model using the Human Protein Atlas website. The predictive abilities were compared using the area under the receiver operating characteristic curves. Gene Set Enrichment Analysis was utilized for indicating enrichment pathways. We analyzed immune cell infiltration between risk groups using the CIBERSORT method. The indicators related to immune status between the two groups were also analyzed. The findings indicated that the high-risk group had better outcomes than the low-risk group in the training and validation sets. Each gene in the model affected the survival of BLCA patients. Our hypoxia-related risk model had better performance compared to other hypoxia-related markers (HIF-1α and GLUT-1). The high-risk group was enriched in immune-related pathways. The expression of chemokines and immune cell markers differed significantly between risk groups. Immune checkpoints were more highly expressed in the high-risk group. These findings suggest that the hypoxia-related risk model predicts patients’ outcomes and immune status in BLCA risk groups. Our findings may contribute to the treatment of BLCA.

Keywords: Hypoxia, predictive biomarker, immune infiltrates, BLCA

Introduction

Bladder urothelial carcinoma (BLCA) is a common urinary malignant tumor [1] that can be diagnosed using invasive cystoscopy. Although this method has been routinely used in clinical diagnosis, it is expensive and does not predict outcomes [2-4]. Hypoxic regions in BLCA impair cellular biological functions due to insufficient oxygenation, allowing immune escape by inhibiting immune cells in the microenvironment, thereby interfering with the treatment of solid tumors [5-7]. The high expression level of the hypoxia-related marker in BLCA is associated with poor outcomes, as in other solid tumors [8]. Similarly, many hypoxia-related markers have been found in the core hypoxic regions of non-muscle-invasive and muscle-invasive bladder tumors [9, 10]. BLCA is more susceptible to advanced progression and distant metastasis under hypoxic conditions [11]. For these reasons, it is critical to identify hypoxia-related biomarkers to diagnose and treat BLCA.

Bioinformatics analysis is used to mine potential hub genes and related-biological processes in various diseases. Zhang et al. found that hypoxia participates in molecular mechanisms
of biological processes involved in BLCA progression using Gene Set Enrichment Analysis (GSEA) [12]. However, no hypoxia risk model has been established for BLCA.

In this study, the gene expression matrix and data from BLCA patients using the Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov/) and the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) were used to identify hypoxia-related biomarkers and establish a model that predicts outcome in BLCA. We also explored the correlation between the model and immunity, which could guide individual clinical treatment decisions, even immunotherapy, and provide a scientific basis for developing BLCA therapies.

**Materials and methods**

**Data processing and analysis**

See Figure 1 for the diagram describing our process. RNA-seq clinical and transcriptome data were obtained for the training and validation sets from TCGA and GEO. Cohorts with no complete transcriptome data, overall survival, and survival status were excluded. Statistical analysis of data from 433 BLCA samples obtained from TCGA and 224 samples from the GSE32894 data set (the corresponding probe is GPL6947-13512). The “limma” package in R software (R version3.6.2., https://www.r-project.org/) was used for data normalization.

**Construction of protein-protein interaction (PPI) network**

We built aPPI network using STRING (https://string-db.org). The top 50 genes ranked the highest connected nodes of hypoxia genes were selected with the help of the R software (R version3.6.2.).

**Generation of hypoxia-related risk model**

We performed univariate and multivariate regression analysis to mine genes significantly associated with prognosis with the help of the “survival” package in R software. The risk score was calculated based on the formula below. Patients were divided into two groups, high risk group and low-risk group depending on the mid-value.

\[
\text{Risk score} = \sum_{i=1}^{n} (\text{Exp}i \times \text{Coei})
\]

**Kaplan-meier survival analysis**

We plotted overall survival (OS) for both groups using Kaplan-Meier analysis in R (‘survival’ package). We plotted receiver operating characteristic curves (ROC) for determining the capacity of hypoxia-related risk models to predict OS. We used the “survivorROC” package in R software. We also performed survival analysis according to hypoxia marker models proposed in the literature using the Sangerbox website (http://sangerbox.com) to compare the predictive capacity of various models. ROC curves were drawn. \(P<0.05\) indicates statistically significant differences.

**Correlation of genes in hypoxia-related risk model**

We used Spearman correlation analysis to calculate correlations of gene expression using this model. Heatmaps and boxplots were utilized to visualize the differential expressions of these genes in different stages of T.

The expression of genes in the hypoxia-related risk model was obtained from the normal and pathological tissues in the Human Protein Atlas (HPA, proteinatlas.org). The impact of gene expression on BLCA survival was determined from the HPA.

**Gene set enrichment analysis**

We performed enrichment analysis on samples from the high-risk group using the HALLMARK gene set. We considered a false discovery rate (FDR) of <0.25, and \(P<0.05\) was deemed to be significant. The top 20 significantly enriched pathways were selected according to the FDR-q values.

**The proportion of infiltrating immune cells**

We used CIBERSORT to calculate the fractions of 22 types of infiltrating immune cells in the two groups from TCGA and GEO. We acquired immune-related genes from the Tracking Tumor Immunophenotype platform (http://biocc.hrbmu.edu.cn/TIP/index.jsp). The expression of immune-related genes is shown as box-and-whisker plots. The expression of phenotypic markers of immune cells in both risk groups was visualized using bar plots. Scatter plots were drawn, and Pearson coefficients were calculated to determine correlations between the
expression of immune checkpoints and hypoxia-related risk scores.

**Statistical analysis**

Kaplan-Meier method was applied to assess OS and the log-rank test was used for the difference analysis. All data analyses were conducted with R software. Data with $P$ less than 0.05 was considered to be statistically significant.

**Results**

*Screening hypoxia-related genes and constructing a prognostic risk prediction model in BLCA*

All hypoxia-related genes were derived from the HALLMARK gene set, and the PPI values among them were calculated using PPI network analysis. The top 50 hypoxia-related hub genes were selected according to adjacent nodes (Figure 2A). A prognostic model was established with seven hypoxia-related genes (EGFR, CAV1, VEGFA, FBP1, GAPDH, SDC4, BGN) using univariate and multivariable Cox regression analysis (Figure 2B, 2C). The risk score formula was as follows:

$$\text{Risk score} = 0.256 \times \text{EGFR} - 0.139 \times \text{CAV1} - 0.216 \times \text{VEGFA} - 0.110 \times \text{FBP1} + 0.225 \times \text{GAPDH} - 0.166 \times \text{SDC4} + 0.166 \times \text{BGN}.$$

We validated the risk score formula using cohort GSE32894. Each patient’s risk score was computed in TCGA training set and the GEO validation set. Patients were assigned to risk
Biomarker based on hypoxia genes for BLCA patients

Figure 2. Identification of the candidate hypoxia-related genes in the TCGA cohort. A. Protein-protein interaction (PPI) network in TCGA (left). The top 50 genes were selected based on the number of nodes (right). B. Univariate Cox regression analysis was used to identify candidate hypoxia-related genes. C. Multivariate Cox regression analysis was used to identify candidate hypoxia-related genes. D. Kaplan-Meier survival analysis for bladder cancer patients in TCGA (left) and GEO (right) databases, stratified according to risk scores (high vs. low). E. Receiver operating characteristic curve analysis of the prognostic accuracy of the hypoxia-related risk model. F. Patient risk scores in TCGA and GEO databases. G. Survival in the high- and low-risk patient groups in TCGA and GEO databases. PPI, Protein-protein interaction; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus. P values were obtained from independent-samples t-test. P<0.05 was considered to statistically significant. Survival analysis was conducted using the Kaplan-Meier method, and differences between cohorts were assessed using the log-rank test.

groups according to the median value. Based on Kaplan-Meier curves (Figure 2D), the low-risk group had better outcomes than the high-risk group. The prognostic accuracy of the risk
Biomarker based on hypoxia genes for BLCA patients

score model for OS with time-dependent ROC curve showed that the area under the curve (AUC) in the training set were 0.639, 0.684, and 0.695, for 1-, 3-, and 5-year survival, respectively; those of the validation set were 0.563, 0.749, and 0.707, respectively. These findings suggest that the model predicted the outcome (Figure 2E). Each patient’s risk score was plotted in the training and validation sets (Figure 2F), and we found the patients in the low-risk group had longer OS than the patients in the high-risk group (Figure 2G). These findings suggest that the hypoxia-related risk model could predict survival from BLCA.

Relationship among genes in hypoxia-related risk model and their survival analysis

To determine the contribution of each gene to the risk model, we performed heatmap analysis. The risk groups differed in terms of hypoxia-related gene expression (Figure 3A). Spearman correlation analysis revealed that hypoxia-related genes were less correlated with each other, indicating these genes might be representative (Figure 3B). According to HPA, expression levels of CAV1, FBP1, SDC4, and VEGFA were lower in BLCA than normal tissues, suggesting these genes were protective. Expression levels of EGFR, BGN, and GAPDH were higher in BLCA than in normal tissues, suggesting that these genes were risk factors (Figure 3C). These findings were consistent with our multivariate regression model. Genes in the hypoxia-related risk model were associated with survival from BLCA (Figure 3D). These findings might increase understanding of the role of these genes in BLCA.

Validation of the ability of the hypoxia-related risk model to predict pathological parameters in BLCA

To confirm the ability of the hypoxia-related risk model to predict outcomes, we used univariate and multivariate Cox regression analysis of the risk scores and clinicopathological parameters, including age, gender, and T staging, N staging, and grade. We found that the risk score, T staging, N staging, and age were independent outcome predictors in TCGA training set. The risk score and T staging were independent prognostic factors (Figure 4A, 4B). To determine the relationship between T staging and the expressions of the gene in the hypoxia-related model, it has been suggested that expression levels of VEGFA, FBP1, CAV1, and BGN in different T staging exhibited significant differences, indicating the hypoxia-related model is closely related to tumor progression (Figure 4C, 4D). To sum up, these findings suggest that our risk model predicted outcome in BLCA.

Verification of hypoxia-related prognostic markers in BLCA

To further illustrate the superiority of our model as a prognostic marker, we compared the potential of the risk model to other hypoxia-related markers. The glucose transporter 1 (GLUT-1) predicts the outcome in BLCA. Hypoxia-inducible factor-1 (HIF-1α), a hub biomarker for hypoxia, was selected for reference and comparison. All metrics established in BLCA patients were from the Sangerbox website. We found that low-risk patients had longer OS in our hypoxia-related risk model by dividing the high and low-risk groups using optimal cutoff values. Further comparison of AUC values from time-dependent ROC curves revealed that, in the training set of our hypoxia-related risk model, the 1-, 3-, 5-year AUC values of the nomogram were 0.64, 0.68, and 0.68, respectively (Figure 5A), which were significantly higher than GLUT-1 (Figure 5B, 1-, 3-, 5-year AUC values: 0.52, 0.56, 0.54) and HIF-1α (Figure 5C, 1-, 3-, 5-year AUC values: 0.53, 0.54, 0.51). Likewise, the validation dataset confirmed the above results. Our findings suggest that our risk model has more reliable predictive capabilities than previously employed markers.

Validation of ability of our risk model to predict the tumor immune microenvironment

Hypoxia influences the tumor microenvironment (TME), which in turn modulates immune status. To elucidate the inherent associations between the hypoxia-related outcome model and immune status and provide a basis for subsequent immunotherapy, GSEA was used for functional enrichment in high-risk BLCA patients. We found that a series of immune-related pathways were enriched, including JAK-STAT3 signaling, NF-κB signaling, IFN-γ signaling, and inflammatory responses (Figure 6A; Table 1). We determined associations between hypoxia-related genes and infiltration of immune cells in the BLCA TME and infiltration of immune cells in both groups in the TCGA and
Biomarker based on hypoxia genes for BLCA patients

Figure 3. Gene expression and correlation analysis in hypoxia-related risk model. A. Heat maps of expression levels of the genes in the hypoxia-related risk model for the high- and low-risk groups in TCGA (left) and GEO databases (right). B. Correlations among the genes in the hypoxia-related risk model based on TCGA (left) and GEO (right) databases. Positive and negative correlations are indicated in red and green, respectively. C. Validation of the protein expression in the hypoxia-related risk model on the HPA website. D. Survival of the genes for BLCA patients in the...
hypoxia-related model on the HPA website using Kaplan-Meier survival analysis. HPA, Human Protein Atlas; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; BLCA, Bladder Urothelial Carcinoma. P values were obtained from independent-samples t-test. P<0.05 was considered to statistically significant. Survival analysis was conducted using the Kaplan-Meier method, and differences between cohorts were assessed using the log-rank test.

### Figure 4

Independent prognostic value of risk score and genes in hypoxia-related model. A. Single-factor prognostic analysis included age, gender, TNM stage, and the risk score of BLCA patients in TCGA and GEO databases. B. Multifactor prognostic analysis included clinicopathological parameters and the risk score of BLCA patients in TCGA and GEO databases. C. Heat maps showing the expression levels of genes in the hypoxia-related risk model in TCGA and GEO databases for different T stages. D. Comparisons of the expression levels of hypoxia-related genes in different T stages from TCGA and GEO databases. TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; TNM, Tumor-node-metastasis. P values were obtained from independent-samples t-test. P<0.05 was considered to statistically significant. *P<0.05, **P<0.01, ***P<0.001.
Biomarker based on hypoxia genes for BLCA patients

Figure 5. Comparison and verification of predictive ability of hypoxia models. A-C. Kaplan-Meier survival analysis for BLCA patients in TCGA and GEO databases, stratified according to risk scores or mRNA expressions. Receiver operating characteristic curve analysis of the prognostic accuracy of the models. TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus. P values were obtained from independent-samples t-test. P<0.05 was considered statistically significant. Survival analysis was conducted using the Kaplan-Meier method, and differences between cohorts were assessed using the log-rank test.

GEO databases (Figure 6B). The correlation between the hypoxia-related risk model and immune infiltrating cells was analyzed by the bubble-plot. A higher risk score correlated with a higher proportion of immune cells such as M0 and M2 macrophages. By contrast, the proportion of immune cells such as plasma cells, naive B cells, and T cells was higher in the low-risk group than in the high-risk group. The risk groups had significant differences in the proportion of immune cells as visually displayed in a histogram (P<0.05) (Figure 6C, 6D). The above results suggest that our risk model appeared to act as a cue on immune status.

Validation of the predictive value of hypoxia-related risk model in immune phenotypes

Further comparison of relative expression of immune cell marker genes in the risk groups revealed that the expression of phenotype-related marker genes of plasma cells in the low-risk group was significantly higher than that of the high-risk group. However, the proportion of M2 macrophage phenotype-related marker genes was significantly higher in the high-risk group (Figure 7A). We analyzed the expression of M0-/M2-related chemokines. It was found that chemokines were abundantly enriched in the high-risk group. In particular, the chemokines produced after polarization to M2, such as CCL18, showed significantly high expression in the high-risk group, possibly explaining the high content of M2 macrophages in the high-risk group (Figure 7B). Using the "Tracking Tumor Immune Phenotype" online platform (http://biocc.hrbmu.edu.cn/TIP/index.jsp), a series of immune regulation-related genes were further screened. Using a heat map, genes related to negative immune regulation were...
Biomarker based on hypoxia genes for BLCA patients

A

Immune-associated pathway in high-risk group (TCGA)

Enrichment plot:
- HALLMARK_INFLAMMATORY_RESPONSE
- HALLMARK_IL6_JAK_STAT1_SIGNALLING
- HALLMARK_TH2_SIGNALING_VIA_NFKB
- HALLMARK_INTERFERON_GAMMA_RESPONSE

B

Immune cell Percent in TCGA (%)

- B cells naive
- B cells memory
- Plasma cells
- T cells CD8
- T cells CD4 naive
- T cells CD4 memory resting
- T cells CD4 memory activated
- T cells follicular helper
- T cells regulatory (Tregs)
- T cells gamma delta
- NK cells resting
- NK cells activated
- Monocytes
- Macrophages M0
- Macrophages M1
- Macrophages M2
- Dendritic cells resting
- Dendritic cells activated
- Mast cells resting
- Mast cells activated
- Eosinophils
- Neutrophils
Biomarker based on hypoxia genes for BLCA patients

Figure 6. Pathway enrichment analysis and tumor-infiltrating immune cell analysis in BLCA patients. A. Enriched gene sets in the HALLMARK collection in high-risk group of the training set. B. Heat map of immune cell infiltration in high- and low-risk group from TCGA or GEO databases. C. Bubble chart of the correlation between the patient risk score and the proportion of immune infiltrating cells in the TCGA database. D. Immune infiltrating cells are significantly associated with hypoxia-related risk scores in TCGA database ($P<0.05$). GSEA, Gene Set Enrichment Analysis; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus. $P$ values were obtained from independent-samples t-test. $P<0.05$ was considered statistically significant.
Biomarker based on hypoxia genes for BLCA patients

| Table 1. Gene set enrichment analyses of high-risk group in TCGA BLCA samples |
|---------------------------------|----------------|----------------|----------------|-----------------|----------------|
| Name                            | ES             | NES            | NOM P-value    | FDR q-value     | FWER P-value    | Rank at MAX   |
| Hallmark_epithelial_mesenchymal_transition | 0.785456       | 2.3276258      | 0              | 0               | 0               | 6237          |
| Hallmark_inflammatory_response   | 0.68982        | 2.2883618      | 0              | 9.34E-04        | 0.002           | 7066          |
| Hallmark_apical_junction        | 0.582675       | 2.279953       | 0              | 0.001264694     | 0.03            | 5698          |
| Hallmark_complement             | 0.601588       | 2.2350826      | 0.00210084     | 0.001415409     | 0.03            | 6921          |
| Hallmark_allograft_rejection    | 0.687183       | 2.1634717      | 0.004175365    | 0.003710467     | 0.009           | 7171          |
| Hallmark_coagulation            | 0.59552        | 2.1140666      | 0              | 0.00490311      | 0.015           | 8266          |
| Hallmark_angiogenesis           | 0.677615       | 2.0700555      | 0              | 0.006818818     | 0.024           | 7500          |
| Hallmark_IL2STAT5_signaling     | 0.514615       | 2.0774574      | 0.001996008    | 0.007468107     | 0.023           | 7692          |
| Hallmark_hypoxia                | 0.526527       | 2.0419757      | 0.002123142    | 0.007637098     | 0.032           | 5688          |
| Hallmark_KRAS_signaling_up      | 0.501232       | 2.0119126      | 0.004024145    | 0.008965819     | 0.044           | 8132          |
| Hallmark_MTORC1_signaling       | 0.561707       | 1.9666486      | 0.01247013     | 0.013026817     | 0.064           | 5578          |
| Hallmark_apoptosis              | 0.4807         | 1.9077333      | 0.006396588    | 0.019406212     | 0.095           | 6722          |
| Hallmark_IL6_JAK_STAT3_signaling| 0.587044       | 1.8920499      | 0.006263048    | 0.020746794     | 0.108           | 7090          |
| Hallmark_myogenesis             | 0.508341       | 1.8633606      | 0.022633744    | 0.024279423     | 0.13            | 7591          |
| Hallmark_hedgehog_signaling     | 0.568482       | 1.8527048      | 0.006060606    | 0.02462458      | 0.136           | 5864          |
| Hallmark_TNFa_signaling_VIA_NFKB | 0.53393        | 1.719862       | 0.04526749     | 0.029674731     | 0.186           | 7365          |
| Hallmark_apical_surface         | 0.488358       | 1.8005064      | 0.01026694     | 0.029942015     | 0.182           | 4284          |
| Hallmark_UV_response_DN         | 0.47509        | 1.8080297      | 0.006012024    | 0.030107137     | 0.175           | 3806          |
| Hallmark_interferon gamma_response | 0.636303      | 1.8092804      | 0.040339705    | 0.03190229      | 0.175           | 7092          |
| Hallmark_glycolysis             | 0.441299       | 1.7679787      | 0.017021276    | 0.033044443     | 0.2             | 5690          |
| Hallmark_reactive_oxygen_species_pathway | 0.518838      | 1.68988       | 0.032986962    | 0.049855        | 0.271           | 7019          |
| Hallmark_G2M_checkpoint         | 0.5881         | 1.6712279      | 0.06198347     | 0.052687917     | 0.293           | 8571          |
| Hallmark_mitotic_spindle        | 0.461691       | 1.6442664      | 0.055009823    | 0.05769695      | 0.308           | 8399          |
| Hallmark_interferon alpha_response | 0.65164       | 1.6323211      | 0.085365854    | 0.0579011       | 0.318           | 6801          |
| Hallmark_unfolded_protein_response | 0.460538      | 1.5964074      | 0.0751073      | 0.066218555     | 0.355           | 7780          |
| Hallmark_HEME_metabolism        | 0.362973       | 1.5404751      | 0.03941909     | 0.072797954     | 0.425           | 7646          |
| Hallmark_TGF_BETA_signaling     | 0.454284       | 1.5445051      | 0.06407767     | 0.07373564      | 0.419           | 5753          |
| Hallmark_E2F_targets            | 0.58466        | 1.56686       | 0.115226336    | 0.07388786      | 0.391           | 8609          |
| Hallmark_UV_response_UP         | 0.38287        | 1.5473924      | 0.039748956    | 0.07478651      | 0.412           | 8183          |
| Hallmark_MYC_targets_V1         | 0.516636       | 1.5521855      | 0.11943322    | 0.07618481     | 0.407           | 8344          |
| Hallmark_WNT_beta_catenin_signaling | 0.46144       | 1.4933735      | 0.083333336    | 0.0864872      | 0.481           | 8778          |
| Hallmark_P53_pathway            | 0.366206       | 1.4944457      | 0.074         | 0.08871294      | 0.479           | 7284          |
| Hallmark_androgen_response       | 0.393805       | 1.4745541      | 0.06326531     | 0.089081176     | 0.503           | 6377          |
| Hallmark_estrogen_response_late  | 0.358566       | 1.479784       | 0.059548255    | 0.08985082      | 0.5             | 6625          |
| Hallmark_notch_signaling         | 0.406555       | 1.4611244      | 0.055226825    | 0.091368124     | 0.517           | 4567          |
| Hallmark_KRAS_signaling_DN      | 0.324951       | 1.392642       | 0.049603175    | 0.115419306     | 0.592           | 10049         |
## Biomarker based on hypoxia genes for BLCA patients

<table>
<thead>
<tr>
<th>Hallmark</th>
<th>Value</th>
<th>p-value</th>
<th>fold change</th>
<th>FDR</th>
<th>Count</th>
<th>tags</th>
<th>list</th>
<th>signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hallmark_PI3K_AKT_MTOR_signaling</td>
<td>0.35112</td>
<td>1.37708</td>
<td>0.106</td>
<td>0.612</td>
<td>9645</td>
<td>34%</td>
<td>17%</td>
<td>41%</td>
</tr>
<tr>
<td>Hallmark_estrogen_response_early</td>
<td>0.32310</td>
<td>1.32962</td>
<td>0.13389</td>
<td>0.14196</td>
<td>4827</td>
<td>28%</td>
<td>9%</td>
<td>31%</td>
</tr>
<tr>
<td>Hallmark_cholesterol_homeostasis</td>
<td>0.34580</td>
<td>1.25490</td>
<td>0.21327</td>
<td>0.18466</td>
<td>4339</td>
<td>23%</td>
<td>8%</td>
<td>25%</td>
</tr>
<tr>
<td>Hallmark_spermatogenesis</td>
<td>0.31486</td>
<td>1.17982</td>
<td>0.24267</td>
<td>0.23170</td>
<td>11722</td>
<td>37%</td>
<td>21%</td>
<td>47%</td>
</tr>
<tr>
<td>Hallmark_cholesterol_homeostasis</td>
<td>0.34580</td>
<td>1.25490</td>
<td>0.21327</td>
<td>0.18466</td>
<td>4339</td>
<td>23%</td>
<td>8%</td>
<td>25%</td>
</tr>
<tr>
<td>HALLMARK_MYC_targets_V2</td>
<td>0.40920</td>
<td>1.13440</td>
<td>0.34959</td>
<td>0.26181</td>
<td>8609</td>
<td>41%</td>
<td>15%</td>
<td>49%</td>
</tr>
<tr>
<td>Hallmark_proteasome_beta_cells</td>
<td>0.40770</td>
<td>1.10020</td>
<td>0.35523</td>
<td>0.28389</td>
<td>16104</td>
<td>55%</td>
<td>28%</td>
<td>77%</td>
</tr>
<tr>
<td>Hallmark_protein_secretion</td>
<td>0.30152</td>
<td>1.02800</td>
<td>0.43153</td>
<td>0.34483</td>
<td>6377</td>
<td>27%</td>
<td>11%</td>
<td>30%</td>
</tr>
<tr>
<td>Hallmark_DNA_repair</td>
<td>0.26497</td>
<td>0.93889</td>
<td>0.48140</td>
<td>0.43317</td>
<td>5443</td>
<td>23%</td>
<td>10%</td>
<td>26%</td>
</tr>
</tbody>
</table>
Biomarker based on hypoxia genes for BLCA patients

Figure 7. Relationships between immunophenotypes and hypoxia-related risk model. A. The expression of marker genes on the surface of tumor-infiltrating immune cells in high- and low-risk groups. B. The expression of chemokines of tumor-infiltrating immune cells in high- and low-risk groups. C. Heat maps showing the expression levels of negative regulatory immune genes in high- and low-risk groups. D, E. Scatter plots showed the correlations between hypoxia-related risk scores and the expression of immune checkpoints from TCGA and GEO databases. Pearson coefficients were used to assess the correlation between the two factors. The box plots showed the expression of the immune checkpoints in high- and low-risk groups. TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus. P values were obtained from independent-samples t-test. P<0.05 was considered statistically significant. *P<0.05, **P<0.01, ***P<0.001.
highly expressed in the high-risk group (Figure 7C). We compared expression levels of immune checkpoints between risk groups and found that PD1 and CD70 positively correlated with the risk score. Expression levels PD1 and CD70 were significantly higher in the high-risk group (Figure 7D, 7E). We verified these findings using data from GEO. In summary, these findings suggest that our hypoxia-related risk model provides a basis for predicting immune features and using immunotherapy to treat BLCA.

Discussion

As a common characteristic in microenvironments of solid tumors [13], intratumoral hypoxia is associated with poor outcomes. Hypoxic gene signatures have been used to predict the outcome of various tumors, including head and neck cancers, breast tumors, and carcinoma of the lung. The combination of biomarkers in a predictive model improves the predictive value over individual biomarkers. We established a prognostic model using seven hypoxia-related genes (EGFR, VEGFA, CAV1, BGN, FBP1, SDC4, GAPDH) based on the TCGA database to predict prognosis, survival, and immune status in BLCA. Meanwhile, it was validated in the GEO dataset.

The high-risk group had worse OS according to the hypoxia-related risk model. Oxygen is essential for energy metabolism. Under hypoxic conditions, TME is affected by the regulation of several energy metabolism pathways. TME also affects the metabolism of immune cells, tumor progression, and treatment resistance [14]. Therefore, hypoxic areas can be regarded as metabolic areas within the tumor, which fine-tunes tumor-associated immune responses [15]. Considering that hypoxia is a critical node in tumor progression, we selected hypoxia as the starting point for the predictive outcome and immune status of BLCA patients.

Using Spearman analysis and HPA, we found that each hypoxia-related gene in the risk model was representative. Epidermal growth factor receptor (EGFR) is a receptor tyrosine kinase [16] frequently overexpressed or mutated in several tumors and promotes BLCA progression by means of the VEGF receptor (R)2/nuclear factor-kB signaling pathway [17]. Vascular endothelial growth factor-a (VEGFA) is the first hypoxia-induced angiogenesis factor that promotes proliferation, migration, and formation of the endothelial matrix [18]. VEGFA is a pro-angiogenic factor; however, Pan et al. found that high levels of secreted VEGFA were induced by hypoxia in a mouse tumor model, causing tumor vascular regression and inhibiting tumor growth [19]. Similarly, syndecan (SDC) is highly expressed in almost all malignant tumors [20]. SDC4 silencing reversed the phenotypic transformation of hypoxia-resistant endothelial cells; SDC4 might be an attractive target for tumor therapy [21]. Biglycan (BGN), first identified in bone tissue, is highly expressed in pancreatic cancer, colorectal cancer, and intrahepatic cholangiocarcinoma [22]. Zhao et al. found that BGN can be used as a promising prognostic biomarker and therapeutic target for BLCA [23]. BGN not only triggers pro-inflammatory Toll-like receptors and inflammasomes-signaling, but also stimulates the production of pro-inflammatory cytokines (eg. TNF-α, IL-1β, IL-6), which are key mediators of inflammation in tumor development [24]. Fructose-1,6-biphosphatase (FBP1) is the rate-limiting enzyme in gluconeogenesis [25]. Nutrients are acquired by enhancement of tumor glycolysis under hypoxia [26]. FBP1, as a negative regulator of glycolysis, is frequently down-regulated in many types of tumors and could increase glycolytic capacity to contribute to intratumoral hypoxia [25]. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a glycolytic enzyme that is upregulated in response to hypoxic stress in endothelial cells, and its overexpression is associated with upregulation of HIF-1α protein [27, 28]. These findings suggest that each hypoxia gene in the hypoxia-related risk model could be crucial in cancer progression.

CAV1 has been shown to promote the invasion and migration of tumor cells in some cancers. By reviewing the literature, it was found that CAV1 could inhibit the proliferation and metastasis of CRC and pancreatic cancer cells [29, 30]. Similarly, the loss of CAV1 is a marker of hypoxia and oxidative stress [31], which is inextricably linked with tumor progression. Therefore, CAV1 might inhibit tumorigenesis, especially under hypoxic conditions. However, there is no clear demonstration of the mechanism of CAV1. We found that the effect of CAV1 on survival in univariate Cox regression analysis and HPA was inconsistent with the result of the multivariate Cox regression analysis. Similar
results were reported by Chen et al. [32]. Multicollinearity analysis was used to diagnose and confirm that the multivariate COX regression model was established reasonably. It may be the case that multivariate Cox regression differs from the original univariate Cox regression because of various factors, especially the influence of hypoxia conditions on the dependent variables.

The predictive ability of our risk model was compared using GLUT-1 and HIF-1α, critical regulators of molecular responses to hypoxia [33] involved in tumor cell biological processes [34]. HIF-1α participates in the acute hypoxic response and regulates adaptation to hypoxic conditions [35, 36]. HIF-1α is highly expressed in hypoxic conditions [37]. GLUT-1, a biomarker for tumor hypoxia, is up-regulated in various tumors. Boström et al. demonstrated that GLUT-1 independently predicted poor outcomes [38]. We verified this in TCGA; however, we found the opposite results in the GEO database, possibly because of the different sources of patient samples between the two databases. This phenomenon has been previously reported [39]. To test the sensitivity and specificity of our model, GLUT-1 and HIF-1α were compared with the hypoxia-related risk model. The 1-, 3-, 5-year AUC values showed that our model had a stronger predictive ability than the two markers and could be used to predict outcomes in BLCA.

To better understand the characteristics of high-risk patients in our risk model, we employed the GSEA function and found that the high-risk group displayed significant enrichment in immune-related pathways, including JAK-STAT3, NF-κB, and IFN-γ. In TME, IL-6/JAK/STAT3 signaling drives proliferation, invasion, and metastasis, while strongly inhibiting anti-tumor immune responses [40]. NF-κB is an essential factor that could be used in tumor immunosurveillance [41]. Hypoxia-sensitive pathways are thought to be critical regulators of immune cell function [42]. For example, therapeutic strategies for HIF in the immune system might be beneficial for anti-tumor immune responses [43]. Immune cell infiltration in tumor cells is closely related to tumor outcomes. Therefore, we focused on immune-related functions. While explaining the characteristics of our risk model, we also analyzed the predictive performance of the model in the immune direction.

By measuring the proportions of immune-infiltrating cells using CIBERSORT, we found that higher risk scores correlated with more significant contents of M0 and M2 macrophages. Naïve B cells also showed differences in each risk group. Immune cells in the TME correlated with survival from tumors. Xue et al. found that M2 macrophages are the most common cells that infiltrate the microenvironment in BLCA [44]. This finding was consistent with our results. M2 macrophages promote the polarization of TAM to M2 under hypoxic conditions and are involved in promoting angiogenesis, cell proliferation, and immunosuppression of tumors [37, 45]. Naïve B cells significantly positively correlated with better survival. They mediate anti-tumor effects by secreting IFN-γ and enhancing T cell activation [46]. Plasma cells exert anti-tumor immunity by participating in synergistic interactions among lymphocyte subpopulations [47]. The greater proportion of plasma cells in the low-risk group might explain their better survival and outcome.

We measured expression levels of genes and chemokines related to immune cells. The transcription factor BCL6 enhances the function of high-affinity antibody-secreting plasma cells [48]. SDC1 (CD138) is also used as a marker for plasma cells [49]. M2 macrophages expressed high levels of CD163, MS4A4A, VSIG4, and chemokines (CCL17 and CCL18) [50]. Similarly, high levels were observed in the high-risk group. We confirmed a relationship between our risk model and immune negative regulatory genes. Immune negative regulatory genes were expressed highly in the high-risk group, suggesting that our model predicted the immune status of TME.

Immune checkpoint inhibitors, including anti-PD-1 monoclonal antibody and anti-CD70 antibody, have shown great potential to control tumors through immune activation [51]. CD70 can be transiently expressed on B cells [52] and is believed to play a role in tumor proliferation and evasion of immune surveillance [53]. We found that these immune checkpoints showed markedly higher expression in the high-risk group. The findings indicate that our risk model might serve as a proxy for a patient’s
immune status and may inform decision-making for the treatment of BLCA.

This study has certain limitations. The outcome data were derived from public databases, and patient sample volumes were limited. More real clinical data are needed to validate our findings.

In conclusion, the hypoxia-related risk model reliably predicted outcome and immune status in BLCA. The model might be more helpful for cancer treatment and immunotherapy options than traditional treatment regimens. The model provides prospects for clinical applications as a biomarker and paves the way for research developments in BLCA.

Acknowledgements

This work was supported by the NSFC joint fund for regional innovation and development (No. U20A20413), National Natural Science Foundation of China (NSFC, No. 82073884), General project of Natural Science Foundation of Liaoning Province (No. 2020-M S-141), General project of scientific research of Shenyang High Level Talent Innovation and Entrepreneurship Team (2019-SYRCCY-B-01), Major Special S&T Projects in Liaoning Province (2019JH1/10300005), and Shenyang S&T Projects (20-204-4-22). The authors would like to acknowledge the Key Laboratory of Precision Diagnosis and Treatment of Gastrointestinal Tumors (China Medical University), Ministry of Education for providing the space and equipment for conducting the study.

Disclosure of conflict of interest

None.

Address correspondence to: Minjie Wei, Department of Pharmacology, School of Pharmacy, China Medical University, No. 77, Puhe Road, Shenyang North New Area, Shenyang 110122, Liaoning, China. Tel: +86-24-31939448; Fax: +86-24-31939448; E-mail: njwei@cmu.edu.cn; Xiaobin Wang, Center of Reproductive Medicine, Shengjing Hospital of China Medical University, 7 Mulan Road, Benxi Economic Development Zone, Shenyang 117004, China. Tel: +86-15365751222; E-mail: 1657438092@qq.com

References

nein 3 is a hypoxia-upregulated oncogene enhancing cell invasion and tumorigenesis in human bladder carcinoma cells. Int J Mol Sci 2019; 20: 980.


[33] Tirpe AA, Gulei D, Corteia SM, Crivii C and Ber-indan-Neagoe I. Hypoxia: overview on hypoxia-
mediated mechanisms with a focus on the role of HIF genes. Int J Mol Sci 2019; 20: 6140.


