Lactic Acid Metabolism And Transporter Related Three Genes Predict The Prognosis of Patients With Clear Cell Renal Cell Carcinoma

Tuanjie Guo  
Shanghai Jiao Tong University School of Medicine

Jian Zhang  
Shanghai Jiao Tong University School of Medicine

Tao Wang  
Shanghai Jiao Tong University School of Medicine

Zhihao Yuan  
Shanghai Jiao Tong University School of Medicine

Heting Tang  
Shanghai Jiao Tong University School of Medicine

Dongliang Zhang  
Shanghai Jiao Tong University School of Medicine

Encheng Zhang  
Shanghai Jiao Tong University School of Medicine

Siteng Chen  
Shanghai Jiao Tong University School of Medicine

Xiang Wang (✉️ seanw_hs@163.com)  
Shanghai Jiao Tong University School of Medicine

Research Article

Keywords: clear cell renal cell carcinoma, lactic acid metabolism and transporter, three genes, prognosis, cell cycle

Posted Date: December 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1109629/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Version of Record: A version of this preprint was published at Genes on March 30th, 2022. See the published version at https://doi.org/10.3390/genes13040620.
Abstract

Background: Lactic acid was previously considered a waste product of glycolysis, and has now become a key metabolite for cancer development, maintenance and metastasis. So far, numerous studies have confirmed that tumor lactic acid levels are associated with increased metastasis, tumor recurrence and poor prognosis. However, the prognostic value of lactic acid metabolism and transporter related genes in patients with clear cell renal cell carcinoma has not been explored.

Results: We selected lactic acid metabolism and transporter related twenty-one genes for LASSO cox regression analysis in the E-MTAB-1980 cohort, and finally screened three genes (PNKD, SLC16A8, SLC5A8) to construct a clinical prognostic model for patients with clear cell renal cell carcinoma. Based on the prognostic model we constructed, the over survival (hazard ratio = 4.117, 95% CI: 1.810 - 9.362, p < 0.0001) of patients in the high-risk group and the low-risk group in the training set E-MTAB-1980 cohort had significant differences, and similar results (hazard ratio = 1.909, 95% CI: 1.414 - 2.579 p < 0.0001) were also observed in the validation set TGCA cohort. Subsequently, survival prediction was carried out based on the survival prognosis model we constructed. In the E-MTAB-1980 cohort, the area under ROC curve (AUC) of 1-, 3- and 5- years was 0.71, 0.72, and 0.77, respectively. The AUC of 1-, 3- and 5- years in TCGA cohort was 0.65, 0.65, and 0.63, respectively. Using CIBERSORT algorithm to analyze the differences in immune cell infiltration in different risk groups, we found that dendritic cells and CD4+ memory cells in the high-risk group were significantly lower than those in the low-risk group, while Treg cells were higher than in the low-risk group. Finally, through gene enrichment analysis, we found that the signal pathway that is strongly related to the prognostic model is the cell cycle.

Conclusions: We constructed a prognostic model using lactic acid metabolism and transporter related genes, which can accurately predict the prognosis of patients. The poor prognosis of the high-risk group may be related to the regulation of cell cycle by lactic acid metabolism.

Background

As early as 1920, Otto Warburg had discovered that cancer cells obtain a large proportion of energy through glycolysis even when oxygen is sufficient. Cancer cells can not only obtain energy through aerobic glycolysis, but also obtain energy through glutamine metabolism[1]. Cancer cells generate a large amount of lactic acid while obtaining energy through these two ways, which creates an acidic environment for the tumor microenvironment (TME). Studies have proven that the large accumulation of lactic acid in TME not only promotes tumor growth and metastasis, but also inhibits the anti-tumor immunity of immune cells[2]. The metabolic disorder of lactic acid is closely related to the abnormal expression of the enzymes that produce and transport lactic acid[3, 4]. High expression of genes and proteins related to lactic acid metabolism and transporter often indicates high invasion and poor prognosis[5, 6].
Clear cell renal cell carcinoma (ccRCC), as a highly aggressive malignant tumor, has a variety of metabolic disorders in TME. CcRCC is usually accompanied by the reprogramming of glucose, fatty acid metabolism and the tricarboxylic acid cycle. In addition, the metabolism of some amino acids such as tryptophan, arginine and glutamine is also reprogrammed in ccRCC[7]. Studies have also shown that kidney cancer cells can promote tumor progression by decomposing fat around the kidney to produce lactic acid[8]. However, the prognostic role of lactic acid metabolism and transporter related genes in patients with clear cell renal cell carcinoma remains unclear.

In this study, we used lactic acid metabolism and transporter related genes to construct a prognostic model, which is closely related to the survival and prognosis of patients. We found that the infiltration of immune cells in two risk groups is different. Finally, we also found that the underlying mechanism of this model to predict prognosis is related to cell cycle signaling pathways.

Materials And Methods

Patient cohorts and data sources.

In this study, we included two cohorts, which are the training set E-MTAB-1980 (https://www.ebi.ac.uk/) cohort and the validation set The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) cohort[9]. Only patients with complete gene expression matrix and clinical information will be included in the study, otherwise they will be shaved off. In the E-MTAB-1980 cohort, a total of 101 ccRCC patients' mRNA sequencing data and corresponding clinical information were extracted from the E-MTAB-1980 database, their gene expression data has been standardized by the uploader. A total of 526 patients in the TCGA cohort were included in the study. Their mRNA sequencing data and clinical information were downloaded through the TCGA database, and all gene expression data were standardized through log₂(tpm+1). The statistical description of the clinical information of the patients in the two cohorts is detailed in Table 1. Simultaneously, twenty-one lactic acid metabolism and transporter related genes were obtained from GSEA (http://www.gsea-msigdb.org), and all genes can be found in the supplementary material S1.

Table 1: Basic clinical characteristics of patients in the TCGA cohort and E-MTAB-1980 cohort.
<table>
<thead>
<tr>
<th></th>
<th>TCGA Cohort (526)</th>
<th>E-MTAB-1980 Cohort 101</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 65</td>
<td>194 (36.9%)</td>
<td>44 (43.6%)</td>
</tr>
<tr>
<td>&lt; 65</td>
<td>332 (63.1%)</td>
<td>57 (56.4%)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>343 (65.2%)</td>
<td>77 (23.8%)</td>
</tr>
<tr>
<td>Female</td>
<td>183 (34.8%)</td>
<td>24 (23.8%)</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>13 (2.5%)</td>
<td>13 (12.9%)</td>
</tr>
<tr>
<td>G2</td>
<td>224 (42.6%)</td>
<td>59 (58.4%)</td>
</tr>
<tr>
<td>G3</td>
<td>204 (38.8%)</td>
<td>22 (21.8%)</td>
</tr>
<tr>
<td>G4</td>
<td>74 (14.1%)</td>
<td>5 (5.0%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>8 (1.5%)</td>
<td>2 (2.0%)</td>
</tr>
<tr>
<td><strong>Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>261 (49.6%)</td>
<td>67 (66.3%)</td>
</tr>
<tr>
<td>II</td>
<td>57 (10.8%)</td>
<td>11 (10.9%)</td>
</tr>
<tr>
<td>III</td>
<td>125 (23.8%)</td>
<td>14 (13.9%)</td>
</tr>
<tr>
<td>IV</td>
<td>83 (15.8%)</td>
<td>13 (12.9%)</td>
</tr>
<tr>
<td><strong>T stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>267 (50.8%)</td>
<td>68 (67.3%)</td>
</tr>
<tr>
<td>T2</td>
<td>69 (13.1%)</td>
<td>11 (10.9%)</td>
</tr>
<tr>
<td>T3</td>
<td>179 (34.0%)</td>
<td>21 (20.85)</td>
</tr>
<tr>
<td>T4</td>
<td>11 (2.1%)</td>
<td>1 (1.0%)</td>
</tr>
<tr>
<td><strong>N stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2</td>
<td>\</td>
<td>4 (4%)</td>
</tr>
<tr>
<td>N1</td>
<td>16 (3.1%)</td>
<td>3 (3.0%)</td>
</tr>
<tr>
<td>N0</td>
<td>238 (45.2%)</td>
<td>94 (93.1%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>272 (51.7%)</td>
<td>/</td>
</tr>
<tr>
<td><strong>M stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M1 78(14.8%)</td>
<td>M0 418(79.5%)</td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Survival</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dead</td>
<td>171(32.5%)</td>
<td>23(22.8%)</td>
</tr>
<tr>
<td>Living</td>
<td>355(67.5%)</td>
<td>78(77.25%)</td>
</tr>
</tbody>
</table>

Establishment and verification of prognostic model.

First of all, we extracted the mRNA expression of twenty-one lactic acid metabolism and transporter related genes and the clinical information of the patients in the E-MTAB-1980 data set. Subsequently, based on the glmnet package in the R language environment, the least absolute contraction and selection operator (LASSO)-cox regression analysis was carried out in order to screen out the genes most relevant to the prognosis. We set lambda to 1000 when calculating the coefficients to ensure the robustness of the model. The formula for risk score is as follows: \( \text{risk score} = \sum_{i=1}^{n}(\alpha_i \times \text{exp}_i) \). In this formula, \( n \) represents the number of prognostic genes, \( i \) represents the prognostic genes, \( \alpha \) represents the coefficient of the \( i \) gene, and \( \text{exp} \) represents the mRNA expression value of the \( i \) gene. According to the median of risk score, patients were divided into high-risk groups and low-risk groups. Subsequently, we used the same method to validate this prognostic model in the TCGA cohort.

Univariate Cox regression analysis and survival analysis.

Cox regression analysis is implemented through the dplyr and survival packages. Time-dependent ROC curve and area under the curve (AUC) were derived from R software using timeROC package. The survival analysis in this study is implemented through survival and survminer packages. False discovery rate (FDR) is 0.05 as the statistical boundary.

Immune cell infiltration in tumor microenvironment.

The CIBERSORT algorithm is used to evaluate the difference in immune cell infiltration between the high-risk group and the low-risk group of ccRCC patients, and explore the immunological mechanism of the model to predict prognosis. According to the instructions in the article[10], the LM22 file is used as the signature of immune cells, and the permutations are set as 1000. All immune cell proportion analysis processes are implemented in the R environment.

Identification of differentially expressed genes

We use the edger package to normalize the gene count value, and then use the limma package to select genes with \( |\log_{2}\text{FC}| \geq 1 \) and FDR < 0.05 in the TCGA database.
Functional enrichment analysis.

In order to explore the potential mechanism by which our prognostic model can predict the prognosis of ccRCC patients, we performed a weighted gene co-expression network analysis (WGCNA) analysis using the previously identified differentially expressed genes in cancer and adjacent tissues. Whereafter, select the gene module most relevant to the risk score, extract the genes from the gene module and perform the Kyoto Encyclopedia of Genes and Genomes (KEGG) approach and genetic ontology (GO) analysis in Metascape[11].

Statistical analysis.

In this study, the Two-sided Student’s t-test and one-way ANOVA were used to detect statistically significant differences between groups. Kaplan–Meier curve analysis was carried out by using log-rank test to compare OS. Cox regression analysis were performed to explore whether risk score could act as a prognostic factor for ccRCC. All statistical analysis were performed in R (4.0.0) and GraphPad Prism 8.

Results

Identification of three genes lactic acid metabolism and transporter related as prognostic models of ccRCC patient.

Given that lactic acid metabolism is disordered in the TME, we first compared whether there are significant differences in the expression of lactic acid metabolism and transporter related genes in clear renal cell carcinoma tissues and normal adjacent kidney tissues. Among all the twenty-one lactic acid metabolism and transporter related genes, twenty genes have significant expression differences. Compared with normal tissues, nine genes are up-regulated and eleven genes are down-regulated in cancer tissues (Figure 1A), which is consistent with lactic acid metabolism disorder[12]. In order to make the prognostic model more robust, we used a small sample of E-MTAB-1980 cohort as a training set to construct a prognostic model, and a total of three genes (PNKD, SLC16A8, SLC5A8) were screened to construct the model (Figure 1B), their respective weights can be found in Table 2. Subsequently, we used the aforementioned formula to calculate the risk score of each patient in the E-MTAB-1980 cohort, and divided them into a high-risk group and a low-risk group (high score means high risk) based on the median score for survival analysis. There is a significant difference in OS (hazard ratio = 4.117, 95% CI: 1.810 - 9.362, p < 0.0001) between the two groups of patients, and patients in the high-risk group have a lower survival rate (Figure 1C). Homoplastically, in the TCGA cohort, it can be observed that the OS (hazard ratio = 1.909, 95% CI: 1.414 - 2.579 p < 0.0001) of ccRCC patients has similar result (Figure 1D). Finally, we integrated the selected genes and the clinical information of the patients into the heamap for analysis, and the correlation between the risk score and gene expression and clinical pathological information can be clearly observed (Figure 1E).

Table 2: Selected genes and associated weights in the prognosis model.
Genes  |  Weights  
--- | ---  
PNKD  |  0.43603432  
SLC16A8  |  -0.05299190  
SLC5A8  |  -0.07168776  

**Risk score is strongly correlated with the patient’s clinicopathological information.**

We want to know whether there is a difference in risk scores in different subgroups of clinicopathology information. In the TCGA cohort, we found that the risk score gradually increased as the Grade increased (Figure 2A). Similarly, the patient’s risk score will also gradually increase with the increase of Stage (Figure 2B). The T stage of renal cell carcinoma mainly represents the size of the tumor, and the risk score also increases as the tumor increases (Figure 2C). In addition, the most important thing is that when the risk score increases, the risk of tumor metastasis also gradually increases (Figure 2D, E). Taken together, the higher the lactic acid metabolism risk score, the more likely the tumor is to develop into a serious clinical pathological state. It can be explained that our prognostic model is strongly related to clinicopathological information.

**The constructed prognostic model can be used as a prognostic factor.**

Considering that the risk score can significantly distinguish the survival rate and clinicopathological information of ccRCC patients, we want to know whether the risk score can be used as a prognostic factor for ccRCC patients. Therefore, we performed univariate Cox regression analysis in two independent cohorts, and the hazard ratio of the risk score in the E-MTAB-1980 cohort was 6.406, \( p < 0.0001 \). Simultaneously, we also got analogous conclusion in the TCGA cohort (Table.3). In summary, we can conclude that the risk score of lactic acid metabolism and transport can be used as a prognostic factor for ccRCC patients.

**Table 3: Univariate Cox regression analysis in the TCGA cohort and E-MTAB-1980 cohort.**

| Cohort | TCGA cohort | E-MTAB-1980 cohort |  
|---|---|---|---  
| | HR (95% CI) | p-Value | HR (95% CI) | p-Value  
| Age  | 1.030(1.016-1.043) | <0.001 | 1.044(1.002-1.087) | 0.040  
| Gender  | 0.952(0.693-1.306) | 0.759 | 2.265(0.673-7.626) | 0.187  
| Grade  | 2.333(1.896-2.871) | <0.001 | 2.982(1.671-5.320) | <0.001  
| Stage  | 1.907(1.666-2.182) | <0.001 | 2.289(1.654-3.167) | <0.001  
| Risk score  | 3.323(2.072-5.327) | <0.001 | 6.406(2.571-15.960) | <0.001  

In order to explore whether our prognostic model can predict the survival time of patients, we performed ROC curve analysis on patients in two independent cohorts. In the E-MTAB-1980 cohort, patients’ 1-, 3- and 5-year AUC reached 0.71, 0.72, and 0.77, respectively (Figure 3A).
Simultaneously, in the TCGA cohort, the 1-, 3- and 5-year AUC of the patients were 0.65, 0.65, and 0.63, respectively (Figure 3B). Therefore, the prognostic model we established can be used as a reference factor for predicting survival of ccRCC patients.

**More Treg cells in the tumor microenvironment of patients with high score.**

Lactic acid metabolism can affect the distribution and phenotype of immune cells[13]. In order to explore whether there are significant statistical differences in the proportion of immune cells in patients with different risk groups, we used the CIBERSORT algorithm for analysis. In the TME of ccRCC patients, we found that the top three immune cells with the largest proportion are M2 type macrophages, CD4+ T memory cells, and CD8+ T cells (Figure 4A). So as to show the proportion of immune cells in each patient more intuitively, we visualized the proportion of immune cells in each patient (Figure 4B). To further study the relationship between our risk model and immune cell infiltration, we analyzed the differences in immune infiltration of ccRCC patients between high-risk and low-risk groups (Figure 4C). We found that the tumors of patients in the high-risk group had more Treg cells and macrophages, and fewer dendritic cells and CD8+ cells than those in the low-risk group. This may be related to the worse the prognosis of ccRCC patients with higher CD8+ T cell infiltration[14]. In addition, it may also be related to the ability of lactic acid to supply Treg cell metabolism and proliferation[15]. In short, for the first time, we established a prognostic model based on lactic acid metabolism and transporter related genes to analyze the infiltration of immune cells in the TME.

**WGCNA reveals that high risk score is related to cell cycle.**

Finally, in order to explore the potential mechanism of our model for predicting prognosis, we performed a weighted gene co-expression network analysis in the TCGA cohort based on the previously identified differentially expressed genes. We analyzed 2936 differentially expressed genes in the TCGA cohort through WGCNA. According to the suggestion of pickSoftThreshold, the soft threshold power of the β value is set to 14 (Figure 5A). Subsequently, all genes related to ccRCC are hierarchically clustered into 10 gene modules (Figure 5B). The correlation analysis shows that the green model (MEgreen) has the highest correlation with the risk score (Figure 5C). Then we performed a functional enrichment analysis of the genes in the green module, and we found that the most relevant signal pathway for risk score is the cell cycle (Figure 5D). Ultimately, we performed protein-protein interaction analysis (Figure 5E) on the genes in the first three signal pathways of Figure D.

**Discussion**

Clear cell renal cell carcinoma is the most common malignant tumor in kidney tumors. The prognosis of patients with tumors that can be completely removed by surgery is good. At present, the most commonly used factors that can be used as a prognostic reference for patients are mainly TNM stage and Fuhrman grade system, but there is no other reference factors can guide the prognosis[16]. Therefore, it is urgent to develop a tool that can predict the prognosis of ccRCC patients.
We build a prognostic model based on the small sample E-MTAB-1980 cohort as the training set, and then validate the model in the large sample TCGA cohort, which makes our model more convincing. The constructed model can clearly separate the survival curves of patients in the high-risk and low-risk groups in the two cohorts, and the differences are statistically significant, which fully demonstrates that the model is powerful and convincing. In addition, the patient's Stage, Grade, T stage, lymph node metastasis and distant metastasis are all associated with the risk score, and the score will increase as the disease progresses. Nevertheless, the model we built can also accurately predict the survival time of patients, which fully affirms the important value of our model and the possibility of clinical use.

Excessive lactic acid production by cancer cells is related to the abnormal expression of lactic acid metabolism enzymes[13]. The effect of lactic acid on immune cells has gradually become transparent. Above all, it can induce M2-like polarization of tumor-associated macrophages. The presence of lactic acid alone can induce the expression of VEGF and ARG1 in macrophages, thereby promoting tumor growth and metastasis[17]. This is consistent with our previous conclusion that M2 macrophages account for a large proportion in the tumor microenvironment of ccRCC patients. Lactic acid can inhibit the differentiation of monocytes into dendritic cells. Higher lactic acid levels in the tumor microenvironment may hinder the formation and accumulation of dendritic cells[13, 18, 19]. This is consistent with our analysis that the high-risk group had a lower percentage of dendritic cells. In addition, the lactic acid in the tumor microenvironment can make CD8+ cells highly express PD-1 and reduce the secretion ability of cytokines[20]. More importantly, Treg cells can use lactic acid as a source of nutrients to promote their own proliferation and differentiation[15]. Interestingly, we also observed a higher proportion of Treg cells in patients in the high-risk group. It can be explained that the poor prognosis of patients in the high-risk group is related to the suppression of anti-tumor immune function.

The accumulation of lactic acid in TME is not only due to the high expression of lactic acid metabolizing enzymes, but the lactic acid transporter also plays an indispensable role in it. When the lactic acid transporter is highly expressed, it can significantly increase the invasion and metastasis ability of tumor cells, and inhibiting the lactic acid transporter can reduce the proliferation and metastasis ability of cancer cells[5, 21]. The latest research has found that the use of small molecule targeted inhibitors of lactic acid transporters can reverse the immunosuppressive microenvironment of solid tumors and improve the safety and anti-tumor efficacy of tumor immunotherapy[22–24]. Blocking lactic acid metabolism and transport is a very promising treatment for cancer.

Gene enrichment analysis found that the constructed prognostic model is strongly related to the cell cycle signaling pathway, and the Grade and Stage also increase with the increase of the risk score. The evaluation criterion of Fuhrman grade system mainly depends on the cell division. T stage is mainly judged based on the size of the tumor, which happens to depend on the cell cycle. This is reflected in the accumulation of lactic acid in the TME to promote tumor proliferation. Based on the cell cycle signaling pathway, we also showed the protein-protein interaction network, suggesting that the risk score is related to the interaction of these proteins.
The three genes screened by our prognostic model have almost no researchers reporting their role in tumorigenesis and development. Among them, the PNKD gene has only been reported in Tourette Disorder or Tic disorder[25]. The SLC16A8 gene has not been found to be associated with the disease. The biological behavior of SLC5A8 in cervical cancer has been reported, which can alleviate the progression of cervical cancer by regulating the Wnt signaling pathway[26]. These three genes have greater research space in the future, and it will also be the direction and focus of our research.

We innovatively used lactic acid metabolism and transporter related genes to construct a prognostic model for ccRCC patients, and carefully analyzed the role of lactic acid-related genes in the prognosis of cancer patients. Although our model is excellent, it is inevitable that there are certain shortcomings. For example, the cohorts we use are all public databases, without our own cohort as verification. In addition, there are still some unavoidable deviations in our data analysis process. Even though, the model we established still has very important clinical significance.

Conclusions

We used lactic acid metabolism and transporter related genes to construct a clinical prognosis model for ccRCC patients. The model can predict the prognosis of patients well, and the mechanism for predicting the prognosis of this model is related to the cell cycle.

List Of Abbreviations

AUC: area under ROC curve; TME: tumor microenvironment; ccRCC: clear cell renal cell carcinoma; TCGA: The Cancer Genome Atlas; LASSO: least absolute contraction and selection operator; FDR: false discovery rate; WGCNA: weighted gene co-expression network analysis.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

In this study, we included two public datasets, which are the E-MTAB-1980 (https://www.ebi.ac.uk/) and The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/). The datasets used and analyzed in this study are also available from the corresponding author upon reasonable request.

Competing interests
The authors declare no conflict of interest.

**Funding**

This work was supported by the National Natural Science Foundation of China (81570607).

**Authors’ contributions**

XW and SC performed study concept and design; TG, TW, and JZ performed development of methodology and writing, review and revision of the paper; HT, DZ, EZ and ZY provided acquisition, analysis and interpretation of data, and statistical analysis; All authors read and approved the final manuscript.

**Acknowledgements**

None.

**References**


Figures

Figure 1
Establishment and verification of prognostic model. (A) Differential mRNA expressions of lactate metabolism and transport related genes between clear cell renal cell carcinoma and normal renal tissue. *: P < 0.05; **: P < 0.01; ***: P < 0.001; ns: no significance. (B) The tenfold cross-validated error and coefficients at varying levels of penalization plotted against the log (lambda) sequence for the least absolute shrinkage and selection operator analysis. (C) Kaplan-Meier survival analysis of overall survival stratified by risk score for ccRCC patients in the E-MTAB-1980 cohort. (D) Kaplan-Meier survival analysis to validate the prognostic model in the TCGA cohort. (E) Heatmap illustrated the expression of the selected genes and the distribution of clinicopathologic factors in the TCGA cohort. ccRCC, clear cell renal cell carcinoma; TCGA, the cancer genome atlas.

Figure 2

Relationship between risk score and different clinicopathological information groups based on TCGA cohort. (A) The relationship between risk score and Grade. (B) The relationship between risk score and Stage. (C) The relationship between risk score and T stage. (D) The relationship between risk score and lymph node metastasis. (E) The relationship between risk score and distant metastasis.
Figure 3

Survival prediction based on prognostic model. (A) ROC curve of the E-MTAB-1980 cohort. (B) ROC curve of the TCGA cohort.
Figure 4

Distribution of immune cells in tumor microenvironment in the TCGA cohort. (A) Each type of immune cell accounts for information in the tumor microenvironment. (B) Information on the proportion of immune cells in each patient in the TCGA cohort. (C) The difference in the proportion of immune cells between the high-risk group and the low-wind group. *: P < 0.05; **: P < 0.01; ***: P < 0.001; ****: P < 0.0001; ns: no significance.
Figure 5

Gene weight co-expression network analysis and gene function enrichment analysis. (A) Soft power estimation in ccRCC for WGCNA. (B) Gene dendrogram with different colors showing the modules identified by WGCNA. (C) The relationship between gene modules and clinical characteristic. (D) Potentially enriched pathways of the co-expressed genes in green module. (E) The protein-protein interaction network of the top three signaling pathways.
Supplementary Files

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

- supplementarymaterial.docx