Research Article

Data Mining of Prognostic Microenvironment-Related Genes in Clear Cell Renal Cell Carcinoma: A Study with TCGA Database

Bin Chen, Wei Chen, Jing Jin, Xueping Wang, Yifang Cao, and Yi He

Urology Department, Jiaxing First Hospital, First Affiliated Hospital of Jiaxing University, Jiaxing, Zhenjiang 314001, China

Correspondence should be addressed to Yifang Cao; caoyifang3301@sohu.com and Yi He; 84748480@qq.com

Received 18 July 2019; Revised 8 September 2019; Accepted 27 September 2019; Published 30 October 2019

Academic Editor: Marcello Ciaccio

Copyright © 2019 Bin Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Clear cell renal cell carcinoma (ccRCC) is one of the most prevalent kidney malignancies. The tumor microenvironment (TME) is highly related to the oncogenesis, progress, and prognosis of ccRCC. The aim of this study was to infer the level of infiltrating stromal and immune cells and assess the prognostic value of them. The gene expression profile was obtained from TCGA and used for calculating the stromal and immune scores. Based on a cut-off value, patients were divided into low- and high-stromal/immune score groups. Survival analysis was performed to evaluate the prognostic value of stromal and immune scores. Moreover, differentially expressed genes (DEGs) that are highly related to TME were determined and applied for functional enrichment analysis and protein-protein interaction (PPI) network. The Kaplan-Meier plot demonstrated that patients with high-immune scores and stromal scores had poorer clinical outcome. In addition, a total of 89 DEGs were identified and mainly involved in 5 pathways. The top 5 degree genes were extracted from the PPI network; among them, IL10 and XCR1 were highly associated with prognosis of ccRCC. The results of the present study demonstrated that ESTIMATE algorithm-based stromal and immune scores may be a credible indicator of cancer prognosis and IL10 along with XCR1 may be a potential key regulator for the TME of ccRCC.

1. Introduction

Renal cell carcinoma (RCC) is the most prevalent kidney malignant tumor globally [1], and it is estimated that over 350,000 cases are diagnosed with RCC each year [2]. Clear cell renal cell carcinoma (ccRCC) is the most common and invasive form in all RCC and comprises about 70–80% of all RCC cases [3, 4].

In recent years, the tumor microenvironment (TME) has already attracted a huge amount of interest from researchers. TME is a complicated system which consists of an extracellular matrix, stromal cells (like fibroblasts, occasionally adipocytes, and mesenchymal stromal cells), and immune cells (such as B and T lymphocytes, macrophages, and natural killer cells) [5]. Immune and stromal cells are the most important component of nontumor cells in TME and have shown a potential value for diagnosis and prognosis prediction of cancers [6, 7]. Previous studies have found that the extent of stromal cells could provide a prognostic factor for patients with solid cancers [8]. In addition, it also has been reported that activated CD8+ T cell density in TME is associated with favorable clinical outcomes of ccRCC [9, 10]. Nevertheless, several immune cells have the opposite effect. For example, the recruitment of CD4+ T cells in TME could promote RCC proliferation through modulating TGFβ1/YBX1/HIF2α signals [11]. Moreover, regulatory T cells (Tregs) can also inhibit tumor immune responses by releasing immunosuppressive cytokines [12]. Macrophages have been reported to have crucial function in both promoting and blocking cancer growth. Macrophage M2 presenting CD163 and CD204 is highly associated with poor prognosis of RCC [13], whereas Hutterer et al. found that tumor-associated macrophages could independently reduce the risk of death in RCC [14].

Immunohistochemistry (IHC) and flow cytometry are the most commonly used technology for determining immune and stromal cells in TME by detecting marker proteins [15]. However, due to the restriction of the channel of markers, conventional technology could not evaluate diverse immune cells simultaneously [16]. As an alternative, algorithms based on a large scale of gene expression profile have
been applied for predicting tumor purity in many researches [17]. The ESTIMATE (Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression data) algorithm developed by Yoshihara et al. is a new tool for inferring the level of infiltrating stromal and immune cells by calculating stromal and immune scores [17]. Subsequent researches on glioblastoma [18], colon cancer [19], and breast cancer [20] have been investigated with the ESTIMATE algorithm and shown good effectiveness of this algorithm. Nevertheless, research on ccRCC using the ESTIMATE algorithm has not been reported in detail.

In the present retrospective study, we applied the ESTIMATE algorithm for the analysis of gene expression profiles of ccRCC which were collected from The Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov) to infer stromal and immune scores for the first time. The association of stromal and immune scores with clinicopathological parameter as well as clinical prognosis of ccRCC patients was also investigated.

2. Materials and Methods

2.1. Data Profile. The gene expression profiles of ccRCC were downloaded from TCGA (Apr 2019) and then were subjected to background correction and normalization with Perl 5.0 (http://www.perl.org/). Meanwhile, relevant clinical characteristics of cancer cases were also collected. Patients with follow-up time <30 days or lacking pathologic diagnosis would be removed.

2.2. ESTIMATE Algorithm. As described previously [17], with the ESTIMATE package in R (version 3.5.2, https://www.r-project.org), stromal and immune scores of each sample were calculated. The optimal cut-off values were evaluated with the online tool: Cutoff Finder (http://molpath.charite.de/cutoff/assign.jsp) [21]. Based on the cut-off values, patients were divided into low- and high-stromal/immune score groups. Group comparisons of stromal/immune scores between different clinical indexes were performed by the t-test with SPSS 20.0. P value < 0.05 was considered statistically significantly different.

2.3. Expression Analysis of Differentially Expressed Genes (DEGs). The Bioconductor package, “edgeR” (http://www.bioconductor.org/packages/release/bioc/html/edgeR.html) was utilized to identify DEGs between low- and high-stromal/immune score groups. The overlapping DEGs would be used for further analysis.

2.4. Functional Enrichment Analysis and Protein-Protein Interaction (PPI) Network. All overlapping DEGs were utilized for Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis with P value (adjusted P value) <0.05 as the threshold. In addition, the PPI network of overlapping DEGs would be obtained from STRING (https://string-db.org) [22] with confidence >0.7 as a cut-off criterion. Then, the data of the PPI network was reconstructed by Cytoscape (version 3.6, https://cytoscape.org), and the top five degree DEGs were selected as the most important targets. The relationship between hub genes and tumor-infiltrating immune cells (B cell, CD4 T cell, CD8 T cell, neutrophil, macrophage, and dendritic cell) was deciphered using TIMER (Tumor IMmune Estimation Resource, https://cistrome.shinyapps.io/timer/).

2.5. Survival Analysis. The Kaplan-Meier plots were conducted to illustrate survival difference between the low- and high-stromal/immune score groups with overall survival of ccRCC patients. Univariable and multivariable Cox regression model was used to determine independent prognostic factors. The P value < 0.05 was set as the cut-off value.

2.6. Expression Profile of Immunomodulators. In recent years, immune checkpoint inhibitors have been approved for the therapy of various cancers, including renal cell carcinoma [23]. In this study, several key immunomodulators (LAG-3, TIM-3, CTLA-4, IFN-γ, ICOS, ICAM-1, TIGIT, PD-1, PDL-1, NKG2A, and VISTA) were quantified in both normal kidney tissues and ccRCC tissues. The difference of expression of immunomodulators between normal and ccRCC samples as well as low- and high-stromal/immune score groups was compared by the t-test.

3. Results

3.1. Patient Characteristics. After removing cases with follow-up time <30 days, a total of 587 samples (72 normal kidney tissue samples and 515 ccRCC tissue samples) were collected from TCGA. The detailed demographic and baseline characteristics of 515 ccRCC patients are described in Table 1. All patients were diagnosed with ccRCC pathologically.

3.2. Evaluation of Immune Scores and Stromal Scores. Based on the cut-off values for stromal scores (841) or immune scores (1780), patients were assigned to low- and high-stromal/immune score groups. As shown in Figure 1(a), immune scores in Fuhrman grade 3/4 were significantly increased compared with those in Fuhrman grade 1/2 (P < 0.001). In addition, immune scores in stage iii/iv were also significantly increased.

Table 1: The baseline characteristics of all patients.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Subtype</th>
<th>No.</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>&lt;65</td>
<td>250</td>
<td>48.5</td>
</tr>
<tr>
<td></td>
<td>≥65</td>
<td>265</td>
<td>51.6</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>341</td>
<td>66.2</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>174</td>
<td>33.8</td>
</tr>
<tr>
<td>TNM staging</td>
<td>Stage i/i</td>
<td>315</td>
<td>61.2</td>
</tr>
<tr>
<td></td>
<td>Stage iii/iv</td>
<td>200</td>
<td>38.8</td>
</tr>
<tr>
<td>Fuhrman grade</td>
<td>G 1/2</td>
<td>235</td>
<td>45.6</td>
</tr>
<tr>
<td></td>
<td>G 3/4</td>
<td>280</td>
<td>54.4</td>
</tr>
<tr>
<td>Immune score</td>
<td>Low</td>
<td>296</td>
<td>57.5</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>219</td>
<td>42.5</td>
</tr>
<tr>
<td>Stromal score</td>
<td>Low</td>
<td>317</td>
<td>61.6</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>198</td>
<td>38.4</td>
</tr>
<tr>
<td>Survival status</td>
<td>Alive</td>
<td>354</td>
<td>68.7</td>
</tr>
<tr>
<td></td>
<td>Dead</td>
<td>161</td>
<td>31.3</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>515</td>
<td>100</td>
</tr>
</tbody>
</table>


increased (Figure 1(b), \( P < 0.001 \)). However, for stromal scores, there was no statistical difference between different Fuhrman grades (\( P = 0.738 \)) and TNM staging (\( P = 0.083 \)). When compared with patients without lymph node metastasis, patients with 1, 2, and 3 lymph node metastases had higher immune scores (Figure 1(c), \( P = 0.012 \), \( P = 0.032 \), and \( P = 0.016 \)). Similar results were observed in patients with distant metastasis (Figure 1(d), \( P = 0.001 \)). Nevertheless, no difference of immune scores was observed between patients with or without lymph node metastasis (\( P = 0.529 \)) and patients with or without distant metastasis (\( P = 0.685 \)).

The associations of stromal/immune scores and corresponding overall survival were analyzed by the Kaplan-Meier plot and evaluated with the log-rank test. The Kaplan-Meier plot demonstrated that high-immune scores as well as stromal scores were negatively correlated with favorable outcome of ccRCC patients (Figure 2). In addition, 5-year survival rates of low- and high-immune score groups were 69.3\% and 52.2\%, respectively (HR = 1.659, 95\% CI [1.204, 2.248]). And, for low- and high-stromal score groups, the rates were 65.7\% and 54.6\%, respectively (HR = 1.409, 95\% CI [1.024, 1.938]). In addition, the result of multivariable Cox regression indicated that both stromal and immune scores were independent prognostic factors (Table 2).

3.3. Functional Enrichment Analysis and PPI Network. A total of 89 overlapping DEGs were identified between low- and high-stromal/immune score groups, including 42
Figure 2: Prognostic values for overall survival: (a) immune scores and (b) stromal scores.
upregulated and 47 downregulated overlapping genes (Figures 3(a) and 3(b)). To better understand the role of DEGs, KEGG and GO analyses were carried out. KEGG analysis revealed 89 overlapping DEGs which were mainly involved in 5 pathways, such as cytokine-cytokine receptor interaction, NF kappa B signaling pathway, and primary immunodeficiency (Figure 3(d)). In addition, 84 GO terms (67 terms of biological process, 1 term of cellular component, and 16 terms of molecular function) were enriched (Figure 3(e), Supplemental Table 1).

For exploring the interplay among 89 overlapping DEGs, a PPI network was built using the STRING tool with confidence >0.7 as the cut-off criterion. 24 nodes (21 upregulated and 3 downregulated DEGs) along with 39 edges consisted of the PPI network (Figure 3(c)). Moreover, in the PPI network, CD79a, CD19, CCL19, IL10, and XCR1 were the remarkable nodes as they had the most connections with other nodes. Furthermore, survival analysis revealed a significant correlation between the expression of IL10 (0.030) and XCR1 (P = 0.012) and prognosis of ccRCC (Supplemental Figure 1). In addition, the result showed that both IL10 and XCR1 were closely related to the infiltration of tumor immune cells (Supplemental Figure 2).

3.4. Expression Profile of Immunomodulators. As shown in Figure 4(a), all of 11 immunomodulators (LAG-3, TIM-3, CTLA-4, IFN-γ, ICOS, ICAM-1, TIGIT, PD-1, PDL-1, NKG2A, and VISTA) were significantly increased in ccRCC samples compared with normal kidney samples. Moreover, in the high-stromal score group, the expression of 11 immunomodulators was all upregulated (Figure 4(b)). In the high-immune score group, similar phenomenon was observed except for PD-1 and PDL-1 (Figure 4(c)). Then, we further explored the prognostic value of 11 immunomodulators. Survival analysis suggested that patients with low expression of LAG3 or CTLA-4 had a longer overall survival than those with high expression of LAG3 (HR = 1.448, 95% CI [1.063, 1.972], Figure 4(d)) or CTLA-4 (HR = 1.513, 95% CI [1.11, 2.062], Figure 4(e)).

4. Discussion

cCRCC is one of the most prevalent kidney malignancies and accounts for approximately 3% of adult cancer [24, 25]. Previous studies have shown that the tumor microenvironment (TME) was highly related to the oncogenesis, progress, and prognosis of cCRCC [26]. TME is the place where the immune system and tumor interplay, reflecting the plasticity of both the tumor and immune system [27]. Tumor development is highly dependent on TME, and any alterations of the composition of TME may influence the evolution of malignancies [27]. Understanding the change may help the development of therapeutic strategies. Stromal cells and immune cells are the important components of TME, which play a critical role in the development of cancers [6]. The ESTIMATE algorithm, a tool based on a large scale of gene expression profile, has been used for the investigation of glioblastoma, colon cancer, and breast cancer and shown good precision and practicality [18–20]. However, it has not been applied for the research on cCRCC.

In the current study, we attempted to infer the level of infiltrating stromal and immune cells in cCRCC by calculating stromal and immune scores with the ESTIMATE algorithm. Survival curves according to Kaplan-Meier showed that low-immune scores as well as stromal scores predicted a favorable prognosis in cCRCC patients. In addition, evidence indicating a significant increase in immune scores was witnessed in patients in Fuhrman grade 3/4 and stage iii/iv, with lymph node metastasis as well as with distant metastasis, whereas for stromal scores, no differences were observed between different Fuhrman grades and TNM staging, with or without lymph node metastasis or distant metastasis.

For exploring the potential mechanism of the change of TME, we identified tumor microenvironment-related genes which would be further utilized for functional enrichment analysis and constructing the PPI network. A total of 89 overlapping differentially expressed genes (DEGs) (42 upregulated and 47 downregulated genes) were determined. KEGG analysis demonstrated that 5 pathways were enriched by 89 overlapping DEGs, including cytokine-cytokine receptor interaction, NF kappa B signaling pathway, and primary immunodeficiency. The NF kappa B signaling pathway is involved in immunity, inflammation, and cell survival [28]. Lua et al. found that the NF kappa B signaling pathway could affect the prognosis of RCC by decreasing the local inflammatory infiltrate and regulating TME [29]. Additionally, Morais et al. reported that inhibition of the NF kappa B signaling pathway attenuated the progression of RCC [30]. Meanwhile, the PPI network was constructed to predict the interaction relationship among 89 overlapping DEGs, and the top 5 degree genes (CD79a, CD19, CCL19, IL10, and

### Table 2: Results of univariate and multivariable Cox regression analysis.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>HR Univariate analysis 95% CI</th>
<th>P</th>
<th>HR Multivariate analysis 95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (&lt;65/≥65)</td>
<td>1.690 (1.232, 2.318)</td>
<td>0.001</td>
<td>1.512 (1.096, 2.087)</td>
<td>0.012</td>
</tr>
<tr>
<td>Sex (female/male)</td>
<td>0.937 (0.680, 1.292)</td>
<td>0.691</td>
<td>0.921 (0.665, 1.275)</td>
<td>0.619</td>
</tr>
<tr>
<td>Fuhrman grade (1 and 2/3 and 4)</td>
<td>2.742 (1.923, 3.911)</td>
<td>&lt;0.001</td>
<td>1.771 (1.218, 2.574)</td>
<td>0.003</td>
</tr>
<tr>
<td>Stage (I and ii/iii and iv)</td>
<td>4.361 (3.162, 6.085)</td>
<td>&lt;0.001</td>
<td>3.410 (2.394, 4.857)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Stromal scores (low/high)</td>
<td>1.411 (1.134, 1.925)</td>
<td>0.030</td>
<td>1.236 (1.110, 1.551)</td>
<td>0.045</td>
</tr>
<tr>
<td>Immune scores (low/high)</td>
<td>1.647 (1.207, 2.248)</td>
<td>0.002</td>
<td>1.321 (1.197, 1.671)</td>
<td>0.029</td>
</tr>
</tbody>
</table>

HR: hazard ratio; 95% CI: 95% confidence interval.
Figure 3: Analysis of DEGs. (a) 42 upregulated overlapping DEGs and (b) 47 downregulated overlapping DEGs. (c) Protein-protein interaction (PPI) network with confidence >0.7. Red and green nodes represent upregulated and downregulated genes, respectively. (d) The enriched pathways with $P < 0.05$. (e) The enriched GO terms with $P < 0.05$ and gene count >5.
Figure 4: Continued.
XCR1) were extracted for further analysis. Among the top 5 degree genes, 2 DEGs (IL10 and XCR1) were highly relative to clinical outcome of ccRCC patients as well as the infiltration of tumor immune cells. IL10, a cytokine produced by monocytes and lymphocytes primarily, has pleiotropic effects on regulating immune response by stimulating B cells and inhibiting macrophages and helper T cells [31]. A previous study showed that the RCC patients with high expression of IL10 had a lower incidence of distant metastasis [32]. However, another research suggested that IL10-producing B cells...
were higher in advanced-stage RCC and could decrease the proportion of T cells in TME [33]. XCR1, the only receptor of the chemokine XCL1, is expressed in dendritic cells and has an implicated function in dendritic cell-mediated cytotoxic immune response [34]. Besides, the overexpression of XCR1 can promote the growth, migration, and invasion in breast cancer and non-small-cell lung cancer [35, 36]. All data suggested that IL10 and XCR1 may be potential key regulators for the TME of ccRCC and novel markers for the prognosis of ccRCC.

Recently, immune checkpoint inhibitors have evolved treatment strategies in oncology and have been approved for the therapy of various cancers including RCC [23]. In this study, we determined 11 immunomodulators (LAG-3, TIM-3, CTLA-4, IFN-γ, ICOS, ICAM-1, TIGIT, NKG2A, and VISTA) that were upregulated in both high-immune score group and high-stromal score group. In addition, PD-1 and PDL-1 were also increased in the high-stromal score group. Kaplan-Meier analysis demonstrated that longer overall survival among patients with low expression of LAG-3 and CTLA-4 was observed. CTLA-4, a homolog of CD28, is expressed by T cells and could inhibit the T cell immune response by diminishing costimulatory signal [37]. LAG-3 is expressed in activated CD4+ and CD8+ T cells and participates in helper T cell response [38]. Several researches have shown that LAG-3 expression was related to the metastasis and prognosis of various cancers such as breast cancer, lung cancer, and ovarian cancer [39–41].

There are a few limitations to be addressed in this study. Firstly, due to the fact that all patients were gathered from TCGA database, the potential of selection bias could not be excluded and it was not possible to collect all information of patients, such as the organ distributions with metastasis and information on whether to use anti-inflammatory drugs or not. Secondly, there was no experimental research conducted to examine the functions of IL10 and XCR1 in ccRCC. Thus, further investigation both in vitro and in vivo is demanded to testify the discovery of this research.

In summary, we applied the ESTIMATE algorithm to calculate stromal and immune scores which were highly associated with the clinical outcome of ccRCC. In addition, we identified 89 microenvironment-related genes, and data from the PPI network and survival analysis revealed that IL10 and XCR1 may be potential key regulators for the TME of ccRCC and could be useful for outlining the prognosis of ccRCC patients. However, more experimental research both in vitro and in vivo is needed to examine the finding of this research.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there is no conflict of interests.

Authors’ Contributions

YH and YF conceived and designed the study. WC and BC performed the majority of the experiments and wrote the paper. JJ and XW analyzed the experimental data. YH and YF reviewed and edited the manuscript. All authors approved the manuscript and agree to be accountable for all aspects of the research in ensuring that the accuracy or integrity of any part of the work is appropriately investigated and resolved. Bin Chen and Wei Chen contributed equally to this work.

Acknowledgments

This work was supported by the Jiaxing Science and Technology Project (2018AD32091 and 2018AD32083) and the Key Subject of Jiaxing Medicine (2019-zc-07).

Supplementary Materials

Supplemental Table 1 Gene Ontology analysis of 89 overlapping DEGs. Supplemental Figure 1: prognostic values for overall survival. (A) IL10. (B) XCR1. Supplemental Figure 2: the relationship between hub genes and tumor-infiltrating immune cells (B cell, CD4 T cell, CD8 T cell, neutrophil, macrophage, and dendritic cell). (Supplementary Materials)

References


