CCR2 and PTPRC are regulators of Tumor Microenvironment and potential prognostic biomarkers of lung adenocarcinoma

Chengbang Wang
The First Affiliated Hospital of Guangxi Medical University

Jie Wei (✉️ 654680956@qq.com)
People's Hospital of Baise

DaLang Fang
The Affiliated Hospital of Youjiang Medical University for Nationalities

Research Article

Keywords: tumor microenvironment (TME), lung adenocarcinoma (LUAD), tumor-infiltrating immune cells (TICs)

DOI: https://doi.org/10.21203/rs.3.rs-422168/v1

License: ☇️  This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Tumor microenvironment (TME) plays an essential role in lung adenocarcinoma (LUAD) development and metastasis. With the development of TME research, it has been proved that differences in tumor-infiltrating immune cells (TICs) and gene expression profile are related to the prognosis of cancer. Our study aimed to identify key genes affecting immune state in TME of LUAD. We downloaded the RNA-seq data of LUAD cases from the TCGA database. ImmuneScore, StromalScore and ESTIMATEScore of each LUAD sample were calculated using ESTIMATE algorithm. Based on the median of different scores, LUAD samples were divided into high and low score groups. Differentially expressed genes (DEGs) between groups were obtained, and univariate Cox regression analysis and protein-protein interaction (PPI) network were used to screen shared DEGs generating in the intersection analysis. CIBORSORT algorithm was performed to calculate the relative contents of TICs for each LUAD sample, and correlation analysis between TICs and key genes was used to determine the influence of key genes to the TME. Finally, CCR2 and PTPRC, affecting the immune status of TME and the prognosis of LUAD, were acquired. Analysis based on the CIBORSORT algorithm suggested that CCR2 and PTPRC were correlated with a variety of TICs, and closely related to the clinical characteristics of the LUAD patients. Our research showed that CCR2 and PTPRC may be potential prognostic markers in LUAD, which may affect function of γδT cells and other immune cells by participating in the regulation of TME immune state.

Introduction

Lung cancer is one of the malignant tumors that threaten human health\(^1\), among which adenocarcinoma is the most common histological type of non-small cell lung cancer (NSCLC), accounting for approximately 60% of NSCLC\(^2\). Given the early symptoms of lung cancer are nonspecific, a vast majority of patients are diagnosed at advanced stages of the disease\(^3\). Although great progress has been made in the early diagnosis and treatment of lung cancer in recent years, the exploration of prognostic biomarker and novel therapeutic targets in lung cancer is still unsatisfactory\(^4,5\). Currently, immunotherapy have brought a new direction to the clinical treatment of lung cancer, and a variety of immune checkpoint inhibitors (ICIs) have shown encouraging efficacy in clinical trials\(^6,7\). However, some patients have not benefited from immunotherapy\(^8\) or developed drug resistance during medication\(^9\), and its side effects and adverse reactions are still a worrying issue. Accordingly, it is still necessary to explore and ascertain the pathogenic mechanism of lung adenocarcinoma (LUAD) to break through the current bottleneck.

Tumor microenvironment (TME) is a complex ecosystem composed of various types of cells, which can be divided into immune components dominated by immune cells and matrix components dominated by fibroblasts\(^10\). Most of the cancer characteristics are activated and sustained by varying degrees through contributions from the distinctive cells of TME. They mediate the recruitment, activation, programming and persistence of tumor cells in a variety of ways\(^11\). A growing body of studies have elucidated that tumor-infiltrating immune cells (TICs) can be used as a predictor of cancer prognosis and treatment target\(^12,13\). Nevertheless, Further studies have found that TME components have great heterogeneity.
among different tumor stages and individuals, thereby leading to the differences in prognosis\textsuperscript{14,15}. For instance, in liver cancer, tumors with higher transcriptome diversity have lower T cell cytolytic activity, which was related to the poor prognosis\textsuperscript{16}. Likewise, in colorectal cancer increased M1 macrophage levels indicated shorter overall survival (OS) of patients\textsuperscript{17}. This heterogeneity not only results in differences in patient outcome, but also increases the difficulty for clinicians to use targeted drugs\textsuperscript{18}. Therefore, the extensive search for novel and effective TME targeting sites may provide a new window of opportunity for combined immunotherapy\textsuperscript{19}.

In this study, we used ESTIMATE algorithm to calculate the TME components of LUAD samples in the Cancer Genome Atlas (TCGA) database, and CIBERSORT algorithms were used to calculate the relative contents of various TICs, combining with univariate Cox regression analysis and protein-protein interaction (PPI) network to screen differentially expressed genes (DEGs) and finally obtained predictive biomarkers CCR2 and PTPRC. CCR2 is the receptor for monocyte chemotactic protein-1 (MCP-1), which were expressed by a variety of cell types, including monocytes, dendritic cells (DC), endothelial cells and cancer cells. The upregulation of MCP-1 was related to formation, metastasis and recurrence of multiple types of cancers\textsuperscript{20,21}. Nevertheless, some studies have indicated that the high expression of CCR2 inhibited the development of small cell lung cancer(SCLC)\textsuperscript{22}. The protein tyrosine phosphatase CD45, encoded by the PTPRC gene, which has thought to be a regulator of B cell and T cell receptor signaling, it has been reported that PTPRC expression was significantly down-regulated in acute lymphoblastic leukemia and Parkinson's disease\textsuperscript{23,24}. In patients with SCLC, its expression was positively correlated with the patients' survival\textsuperscript{25}. Here, our study revealed that the CCR2 and PTPRC might be a potential indicator for the alteration of TME status in LUAD.

Results

Survival analysis of LUAD patients in three different scores

We combined ESTIMATE algorithms and Kaplan-Meier survival analysis to profile the relationship between different scores and patients' prognosis. ImmuneScore, StromalScore and ESTIMATEScore showed a positive correlation with OS (Fig. 1). The results above suggested that the immune and matrix components in the TME could reflect the prognosis of LUAD patients, indicating the potential value of each component of TME.

Correlation of scores with clinical characteristics of LUAD patients

Clinical characteristics were downloaded from the TCGA database, and the correlation analysis between three types of scores and clinical characteristics was carried out. As shown in Fig. 2A-C, males had lower scores relative to females. Meanwhile, three kinds of scores were negatively correlated with tumor stages (Fig. 2D-F). Further analysis demonstrated that scores decreased significantly from T1 to T4 in T classification (Fig. 2G-I). However, no statistically significant changes of scores were observed in N classification(Fig. 2J-L). Likewise, M classification also showed a negative correlation with
scores (Fig. 2M-O). In a word, results above demonstrated that the immune and matrix components in the TME correlated with tumor progression and metastasis.

**DEGs Screening**
In order to analyze the changes in the gene expression profile of immune and matrix components, we divided ImmuneScore and StromalScore into high and low score groups according to the median of scores, and the DEGs of the two groups were presented in heat maps (Fig. 3A-B), thus identifying 1394 genes from ImmuneScore (including 1097 up-regulated and 297 down-regulated genes), and 1623 genes from StromalScore (including 1422 up-regulated and 201 down-regulated genes). The intersection analysis of the two scores obtained 585 up-regulated and 107 down-regulated shared genes (Fig. 3C-D). These shared DEGs may exert dual effect on the immune and matrix component of TME.

**Functional Enrichment Analysis**
Subsequently, we performed functional enrichment analysis for the shared DEGs, the results of gene ontology (GO) analysis showed that the functions of 692 shared DEGs in two groups were mainly concentrated on immune-related terms, including immune response-activating cell surface receptor signaling pathway, immune response-activating signal transduction, lymphocyte mediated immunity, and humoral immune response (Fig. 3E). Meanwhile, Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis indicated DEGs primarily enriched in cytokine-cytokine receptor interaction, chemokine signaling pathway, hematopoietic cell lineage and B cell receptor signaling pathway (Figure.3F). To sum up, DEGs obtained from the intersection analysis of ImmuneScore and StromalScore were enriched in immune function of TME, which may contain the key factors affecting immune state.

**Intersection Analysis of PPI Network and Univariate Cox Regression Analysis**
We then constructed PPI network using STRING database to further screen the 692 shared genes. The top 30 genes sorted by the number of nodes were shown in bar plot (Fig. 4A-B). Univariate Cox regression analysis of the 692 shared genes were performed showing that 22 genes were related to the OS of LUAD (Figure. 4C). Having performed the intersection analysis of PPI network and univariate Cox regression analysis, we ultimately obtained two overlapping genes CCR2 and PTPRC (Fig. 4D).

**Integrated analysis of CCR2 and PTPRC**
Next, we performed integrated analysis of these two genes. As shown in Fig. 5A-B, the expression of CCR2 and PTPRC in tumors were significantly lower than that of normal samples. Survival analysis showed that patients in the high expression group of the two key genes had longer OS than those in the low expression group (Fig. 5C-D). In summary, there was a positive correlation between the expression of PTPRC and CCR2 and the prognosis of LUAD patients. We further analyzed the correlation between the key genes, CCR2 and PTPRC, and clinical characteristics (Fig. 6A-J). The expression of both key genes was significantly higher in males relative to females. Further analysis showed that the two key genes expression was negatively correlated with tumor stage of LUAD, and its expression decreased as the T and N classification increased. However, there was no significant correlation between expression of the two genes and M classification. In conclusion, we can conclude that these two genes may affect the
prognosis by affecting the immune state in TME based on the previous analysis, and they could be used as reliable biomarkers of LUAD prognosis.

**Correlation analysis between key genes and TICs**

The CIBERSORT algorithm was performed on 535 tumor samples to calculate the relative content of 22 types of TICs in each sample, the relative content of TICs in tumor samples and correlations analysis between TICs were shown in Fig. 7a-b. The correlation analysis between the high and low expression groups of 2 key genes showed that 15 TICs were closely related to CCR2, including 10 positively related TIC subtypes and 5 negatively related TIC subtypes (Figure. 8). 13 types of TICs were correlated with PTPRC, including 7 positively related TIC subtypes and 6 negatively related TIC subtypes (Figure. 9). We subsequently divided LUAD patients into high and low groups based on median content of each TIC subtype and found that patients with a higher proportion of γδT cells and plasma cells had a better prognosis, while higher proportion of resting NK cells led to diametrically opposite results (Figure. 10). Taken together, these two genes were key factor affecting immune state of LUAD.

**Discussion**

Lung cancer is the most common cause of cancer-related deaths in the world’s population with a 5-year survival rate ranges from 4–17% depending on tumor stage and regional variation. According to statistics, there were 2.1 million lung cancer cases globally leading to 1.761 million deaths in 2018, ranking first in the incidence of malignant tumors. Although many advances have been made in understanding the mechanism of this disease, more than 1 million people die from lung cancer each year worldwide. Currently, the role of TME in tumorigenesis and cancer progression has become clearer, making it possible to predict treatment responsiveness and survival prognosis. Notably, the immune components in TME could mediate anti-tumor effect, and studies have indicated the correlation between tumor immune microenvironment and the efficacy of immunotherapy.

In our study, key genes, CCR2 and PTPRC, were obtained by analyzing the matrix and immune components of TME in LUAD. CCR2 was a chemokine receptor that regulated the immune response by inducing the recruitment of macrophages and monocytes to the site of inflammation. Intriguingly, numerous studies have elucidated that CCR2 was correlated with the occurrence and development of many diseases. High expression of CCR2 may lead to the poor prognosis of various cancers, and inhibition of CCR2 can enhance the therapeutic effect of PD-1 on tumor suppression. Paradoxically, the study by Zheng et al. proved that CCL2 was significantly down-regulated in SCLC and led to the proliferation of cancer cells. Our study indicated that the expression of CCR2 in LUAD was significantly lower compared with the corresponding normal tissue, and patients with high CCR2 expression had longer OS. Further analysis of the relationship between the expression of CCR2 and tumor stage shed bright light on the results that with the progress of T classification, the expression of CCR2 decreased. In a word, the above results demonstrated that CCR2 may play an anti-tumor effect in LUAD. In addition, the other key gene, PTPRC, encoded an evolutionarily highly conserved protein tyrosine phosphatase CD45,
which was only expressed on nucleated cells of the hematopoietic system and was considered to be a regulator of B and T cell receptor signaling\textsuperscript{24,37,38}. In recent years, the role of CD45 in cancer and TME remained controversial. Chen et al. indicated that the accumulation of CD45 + CD71 + erythroid cells in liver cancer may play an immunosuppressive effect in TME\textsuperscript{39}. Studies have confirmed that a higher proportion of CD45 + cells was closely related to the poor prognosis of NSCLC patients. However, in other types of tumors, high expression of CD45 may be related to the good prognosis of the tumor, indicating that CD45 may play a completely distinctive role in different tumors\textsuperscript{24,25}. Our study revealed that the expression of PTPRC in LUAD was significantly down-regulated compared with normal tissue, and patients with high PTPRC expression had better prognosis. Further analysis of the relationship between the expression of PTPRC and clinic characteristics indicated that with the progress of tumor stage, the expression of PTPRC decreased significantly.

We further calculated relative content of TICs of LUAD samples using CIBERSORT algorithm, the results found that patients with high proportion of \( \gamma \delta \) T cells presented a better prognosis. \( \gamma \delta \) T lymphocytes are a subset of T lymphocytes, which can directly inhibit tumor cells through cytotoxicity\textsuperscript{40,41}. It has been confirmed that \( \gamma \delta \) T cells participated in the anti-tumor effect of lung cancer\textsuperscript{42} and prostate cancer\textsuperscript{43}. Previous studies have confirmed that \( \gamma \delta \) T cells have the property of suppressing and inhibiting a variety of tumor cell lines\textsuperscript{44,45}. However, some \( \gamma \delta \) T lymphocyte subtypes were unexpected promoters for tumorigenesis and cancer development as their functions were affected by the immunosuppressive signal of TME\textsuperscript{46}. Our research found that the proportion of \( \gamma \delta \) T cells was significantly correlated with the expression of CCR2 and PTPRC (\( \rho < 0.05 \)). This may suggest that CCR2 and PTPRC were the key factors that drive the anti-tumor effect of \( \gamma \delta \) T cells in the TME. In summary, CCR2 and PTPRC may affect the function of immune cells, such as \( \gamma \delta \) T cells, by participating in the regulation of the immune activity of TME and exert an impact on the prognosis of LUAD. CCR2 and PTPRC can be used as biomarkers to predict the immune response of TME and provide new therapeutic targets for LUAD. Notwithstanding, our study has some limitations. First, we still need to validate the findings of the study through in \textit{vivo} and in \textit{vitro} experiments. Secondly, we didn’t clarify the way in which key genes affect tumor immunity and exert anti-tumor effects.

**Materials And Methods**

**Extraction of Data**

Through the TCGA database (https://portal.gdc.cancer.gov/), we obtained RNA-seq data and clinical characteristics of 594 LUAD samples (535 tumor samples and 59 healthy samples).

**Analyzing of ImmuneScore, StromalScore, and ESTIMATEScore**

We analyzed the immune and matrix components in the TME using R language version 4.0.3 loaded with estimate package, presented by ImmuneScore, StromalScore, and ESTIMATEScore. The higher the score reflects the higher corresponding component in TME. ESTIMATEScore represents the sum of immune and matrix components.
**Survival Analysis**

The survival and survminer R package were used for survival analysis. Kaplan–Meier analysis was used to perform the survival curve as well, $p < 0.05$ was considered significant.

**Analysis of Correlation Between Clinical Characteristics and Scores**

We analyzed the relationship between clinical characteristics and different scores of LUAD patients downloaded from the TCGA via R language, and used the Wilcoxon rank sum test or the Kruskal-Wallis rank sum test as a test of significance.

**Screening of DEGs**

LUAD samples were marked as high or low score group based on the median of ImmuneScore and StromalScore. We used the package "limma" to analyze the DEGs between the high and low score groups. DEGs with $|\log_{2}FC| > 1$ and false discovery rate (FDR) < 0.05 were statistically significant. We used the R language with package pheatmap to construct the heatmaps of DEGs.

**Enrichment Function Analysis**

To explore the functions of the DEGs, we used clusterProfiler, enrichplot, and ggplot2 R package for the analysis. Terms with $p$-value and $q$-value < 0.05 were considered significant.

**PPI Network Construction**

The STRING public database (https://string-db.org/) was used to generate the PPI network, and results were reconstructed with Cytoscape version 3.8.1. The nodes with the confidence of interactive association more than 0.7 were selected to generate the PPI network.

**Univariate COX Regression Analysis**

R language combined with “survival” package was used in univariate Cox regression analysis. DEGs with a value of $p < 0.05$ were selected for following analysis.

**TICs Analysis**

We used the CIBERSORT algorithm on 535 tumor samples to calculate the relative contents of 22 types of TICs in each LUAD sample, samples with $p < 0.05$ were used for subsequent analysis.

**Conclusion**

Through bioinformatics to evaluate data in TCGA, our research showed that CCR2 and PTPRC may be potential prognostic markers in LUAD, which may affect the immune function of $\gamma\delta T$ cells and other immune cells by participating in the regulation of TME immune activity.
Declarations

Acknowledgements

This study was funded by the National Natural Science Foundation of China (grant number 81770759).

Author contributions

C.B.W participated in design, writing and modification of all the paper. D.L.F took part in the statistical analysis and prepared figures 1-6. J.W. provided important advices for this paper.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

References


**Figures**
Figure 1

Correlation of different scores with the survival time of LUAD patients. (A) Kaplan–Meier survival analysis for LUAD patients grouped by median of ImmuneScore (B) Kaplan–Meier survival analysis for LUAD patients grouped by median of StromalScore (C) Kaplan–Meier survival analysis for LUAD patients grouped by median of ESTIMATEScore.

Figure 2

Correlation analysis between scores and clinical characteristics. (A-C) Correlation analysis between scores and gender. (D-F) Correlation analysis between scores and tumor stage. (G-I) Correlation analysis
between scores and T classification. (J-L) Correlation analysis between scores and N classification. (M-O) Correlation analysis between scores and M classification.

Figure 3

Heatmaps, Venn plots, and functional enrichment analyses for DEGs. (A) Heatmap for DEGs generated in ImmuneScore. (B) Heatmap for DEGs generated in StromalScore. (C-D) Venn plots showing shared up-regulated and down-regulated DEGs in ImmuneScore and StromalScore. (E-F) GO and KEGG enrichment analyses for shared DEGs in ImmuneScore and StromalScore.
Figure 4

PPI network and univariate Cox regression analysis. (A) Construction of PPI network based on STRING database. (B) The top 30 genes ordered by the number of nodes. (C) Forest plot of univariate Cox regression analysis showing the genes affecting the prognosis of LUAD. (D) Venn plot showing the shared genes between the top 30 genes in PPI and significant factors in univariate COX regression analysis.
Figure 5

Integrated analysis of CCR2 and PTPRC. (A-B) Different expression of CCR2 and PTPRC in the normal and cancer tissue. (C-D) Kaplan–Meier survival analysis of LUAD between different CCR2 and PTPRC expression level and prognosis.
Figure 6

The correlation analysis between CCR2 and PTPRC expression and clinic characteristics. (A-E) Correlation analysis between CCR2 expression level and different clinic characteristics. (F-J) Correlation analysis between PTPRC expression level and different clinic characteristics.
Figure 7

TICs analysis in tumor samples and correlation analysis of different TICs subtypes. (A) the relative content of 22 kinds of TICs in each sample based on CIBERSORT algorithm. (B) correlation analysis of 22 kinds of TICs.
Figure 8

Correlation analysis between the relative contents of TICs and CCR2 expression (A) Violin plot showing the difference in the contents of 22 immune cells in tumor tissues of high and low CCR2 expression groups. (B) Scatter plot showing the correlation analysis between the proportions of 15 types of TICs and the expression of CCR2.
Figure 9

Correlation analysis between the relative contents of TICs and PTPRC expression (A) Violin plot showing the difference in the contents of 22 immune cells in tumor tissues of high and low PTPRC expression groups. (B) Scatter plot showing the correlation analysis between the proportions of 13 types of TICs and the expression of PTPRC.
Figure 10

Three key immune cells affecting patients' survival.