



# Comprehensive analysis of lower mitochondrial complex I expression is associated with cell metastasis of clear cell renal cell carcinoma

Futian Zhang<sup>1#</sup>, Teng Hou<sup>1#</sup>, Liang Chen<sup>1</sup>, Ming Xiong<sup>1</sup>, Menghao Zhou<sup>1</sup>, Gallina Kazobinka<sup>1,2</sup>, Jun Zhao<sup>1</sup>, Xiaomin Han<sup>1</sup>

<sup>1</sup>Department of Urology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China; <sup>2</sup>Urology Unit, La Nouvelle Polyclinique Centrale de Bujumbura, Bujumbura, Burundi

**Contributions:** (I) Conception and design: All authors; (II) Administrative support: T Hou, J Zhao, X Han; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: F Zhang, L Chen; (V) Data analysis and interpretation: F Zhang, M Xiong, M Zhou, G Kazobinka; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

**Correspondence to:** Dr. Xiaomin Han; Dr. Jun Zhao. Department of Urology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China. Email: hxmhust@126.com; 1987xh0809@hust.edu.cn.

**Background:** Clear cell renal cell carcinoma (ccRCC) is characterized by high metastasis potential. It is of great importance to explore the mechanisms underlying ccRCC metastasis and to enable development of potent therapeutics. The mitochondrial complex I (CI) had been considered to play an important role in the development of cancers, but less known in ccRCC.

**Methods:** We utilized available public databases of ccRCC, including single-cell RNA sequencing (scRNA-seq) data GSE73121 and The Cancer Genome Atlas-kidney renal clear cell carcinoma (TCGA-KIRC). Principal component analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (tSNE) analysis were evaluated the heterogeneity of metastatic renal cell carcinoma (mRCC) and primary renal cell carcinoma (pRCC). Protein-protein interaction (PPI) network identified critical gene. Gene set enrichment analysis (GSEA) and gene set variation analysis (GSVA) performed to explore the potential biologic pathways.

**Results:** Our study revealed a significant gene expression heterogeneity between pRCC and mRCC. A PPI network based on differentially expressed genes (DEGs) identified electron transport chain (ETC), especially mitochondrial CI, as the key network hub. Further analysis revealed that the role of mitochondrial CI is associated with tumor metastasis and immune responds of ccRCC. Although CI had low frequency mutations in ccRCC, CI expression is associated with the high frequency mutated genes. A prognosis model included 7 CI genes, and these had a significant effect on overall survival (OS). The area under the curve at 1, 3, and 5 years was 0.717, 0.685, and 0.728, respectively. Transcription factor analysis predicted that *PPARG* possibly is a potential transcription activator of CI genes in ccRCC.

**Conclusions:** Overall, we found that CI expression is associated with ccRCC progress. CI and *PPARG* may be potential biomarkers for metastatic ccRCC.

**Keywords:** Bioinformatics; clear cell renal cell carcinoma (ccRCC); tumor heterogeneity; single cell analysis; mitochondrial complex I (mitochondrial CI)

Submitted Feb 02, 2022. Accepted for publication Apr 12, 2022.

doi: 10.21037/tcr-22-242

**View this article at:** <https://dx.doi.org/10.21037/tcr-22-242>

## Introduction

Renal cell carcinoma (RCC) is a common malignancy of the genitourinary system. The estimated numbers of new cases and deaths of RCC were 431,288 and 179,368, respectively, in 2020 (1). The histological types including clear cell RCC (ccRCC), papillary RCC, chromophobe RCC and collecting duct RCC (2,3). ccRCC is the most common subtype of RCC, represents 70% of total cases and is characterized by high metastasis potential, resulting in about one-third of ccRCC patients having metastatic diseases at the time of diagnosis (3,4). Therefore, it is of great importance to explore the mechanisms underlying ccRCC metastasis and to enable development of potent therapeutics.

Oxidative phosphorylation (OXPHOS) is a process involving a flow of electrons through the electron transport chain (ETC), through which the mitochondrial ETC utilizes a series of electron transfer reactions to generate cellular ATP through OXPHOS (5). The mitochondrial ETC consists of protein complexes I–V, of which most of the subunits are encoded by nuclear genes (6). It has been shown that ETC activity impacts a variety of processes beyond energy balance such as reactive oxygen species (ROS) production, the redox state mitochondrial membrane potential, mitochondrial protein import, and apoptosis (7). Accumulating evidence shows that ETC plays an important role in the development of cancer, as functional electron transport provides the bioenergetic fuelling necessary to sustain tumor initiation, growth, and dissemination (8–10). For example, mitochondrial complex III is required for tumor growth (9). Specific enhancement of mitochondrial complex I (CI) activity inhibited breast cancer growth and metastasis via tumor cell  $\text{NAD}^+/\text{NADH}$  redox balance, *mTORC1* activity, and autophagy (10). However, whether mitochondrial complex is directly correlated with tumor metastasis, especially in RCC, remains largely unclear. Therefore, investigating the role of mitochondrial complex and RCC is meaningful.

In this study, we investigated the role of ETC in the ccRCC metastasis. These findings also suggested that low mitochondrial CI is associated with tumor metastasis and immune responds of ccRCC. Moreover, *PPARG* may be a potential CI transcription activator. We present the following article in accordance with the REMARK reporting checklist (available at <https://tcr.amegroups.com/>

[article/view/10.21037/tcr-22-242/rc](https://doi.org/10.21037/tcr-22-242/rc)).

## Methods

### Data preprocessing

The single-cell RNA sequencing (scRNA-seq) data were downloaded from the Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) database, accession code GEO:GSE73121 (11). The public available The Cancer Genome Atlas (TCGA) datasets were directly downloaded from the TCGA Data Portal (<https://tcga-data.nci.nih.gov/tcga/>). The scRNA-seq data were normalized and differential expression analysis by Seurat, an R package for single-cell analysis (<https://satijalab.org/seurat/>). We discard two cell samples because of a worse RNA feature and choose the top 1,500 maximum standardized variance of genes for next analysis. TCGA transcriptome data were normalized by edgeR from R package. All analyses were run in R version R 4.0.2 and R Studio version 1.3.1056.

### Principal component analysis (PCA) and t-Distributed Stochastic Neighbor Embedding (tSNE) analysis

PCA was performed on the top 1,500 maximum standardized variance screened before. Then, cluster analysis was performed in the top 20 principal components (PCs). Five PCs with the most significant difference P value were selected for the following analysis. tSNE analysis was performed after PCA based on 5 PCs. Cells were classified into two different clusters. All analyses were run in R version R 4.0.2 and R Studio version 1.3.1056.

### Enrichment analysis

Matascope, a gene annotation and analysis resource, was used for pathway and process enrichment analysis (12). Rich factor was the proportion of enriched genes in corresponding gene sets. The gene set enrichment analysis (GSEA) was performed using GSEA v4.0.2 for Windows (<http://www.gsea-msigdb.org/gsea>). The gene set variation analysis (GSVA) was performed using R package GSVA. The gene sets using in GSEA and GSVA analysis were downloaded from GSEA molecular signatures database

hallmark gene sets.

### ***Protein-protein interaction (PPI) network analysis***

PPI network based on differentially expressed genes (DEGs) was constructed by STRING (13) (Search Tool for the Retrieval of Interacting Genes) database (<https://string-db.org/>). The minimum required interaction score 0.9 was applied.

### ***Heatmap analysis and hierarchical clustering***

Gene expression data was transform into  $\log_2$ (normalized counts) for heatmap analysis. Heatmap plot was generated by ComplexHeatmap from R. Unsupervised hierarchical clustering was implemented using the Ward's minimum variance method. Two distinct subgroups were identified through the hierarchical cluster analysis. All analyses were run in R version R 4.0.2 and R Studio version 1.3.1056.

### ***Analysis of the tumor immunological microenvironment***

The ESTIMATE algorithm (<https://bioinformatics.mdanderson.org/estimate/>) was applied to calculate the tumor immune infiltration level. This algorithm can detect the stromal and immune score of each sample to predict the tumor stromal and immune infiltration and thus construct the ESTIMATE score, a microenvironment comprehensive score, to reflect tumor purity in cancer samples. Here, each sample's stromal score, immune score, and estimate score of TCGA-kidney renal clear cell carcinoma (KIRC) were calculated by the ESTIMATE algorithm from R package in R version R 4.0.2 and R Studio version 1.3.1056.

### ***Construction of risk model***

Forty-six CI genes expression were obtained. Univariate Cox regression analysis was used to screen the genes that were significantly related to ccRCC prognosis. Multivariate Cox regression was performed to further reduce the number of genes and obtained the risk scoring model related to ccRCC prognosis. Then, risk model was constructed based on 7 CI genes. Survival analysis, receiver operating characteristic (ROC) curves, univariate and multivariate Cox regression was also evaluated in the next. All analyses were run in R version R 4.0.2 and R Studio version 1.3.1056 with the following packages: survival, glmnet, survivalROC, gplot, and forestplot.

### ***Prediction of CI genes transcription factor***

DAVID bioinformatics resources (<https://david.ncifcrf.gov/home.jsp>) was used to identify potential transcription factor binding site in 5'-untranslated region (5'-UTR) of CI genes. In the next, correlation analysis was performed between transcription factors and CI genes. Then, we combined the results of DAVID bioinformatics resources prediction and correlation analysis to build a regulatory network using cytoscape (<https://cytoscape.org/>). Among that, positive correlation coefficient transcription factor was considered to be transcription activator, and negative was transcription repressor. UALCAN database (<http://ualcan.path.uab.edu/>) was used to identify the expression of transcription activator.

### ***Statistical analyses***

Statistical analyses in this study were performed using GraphPad Prism 8 or R. Survival curves were calculated with the Kaplan-Meier method and compared using the log-rank test. In general,  $P < 0.05$  was considered significant in this study.

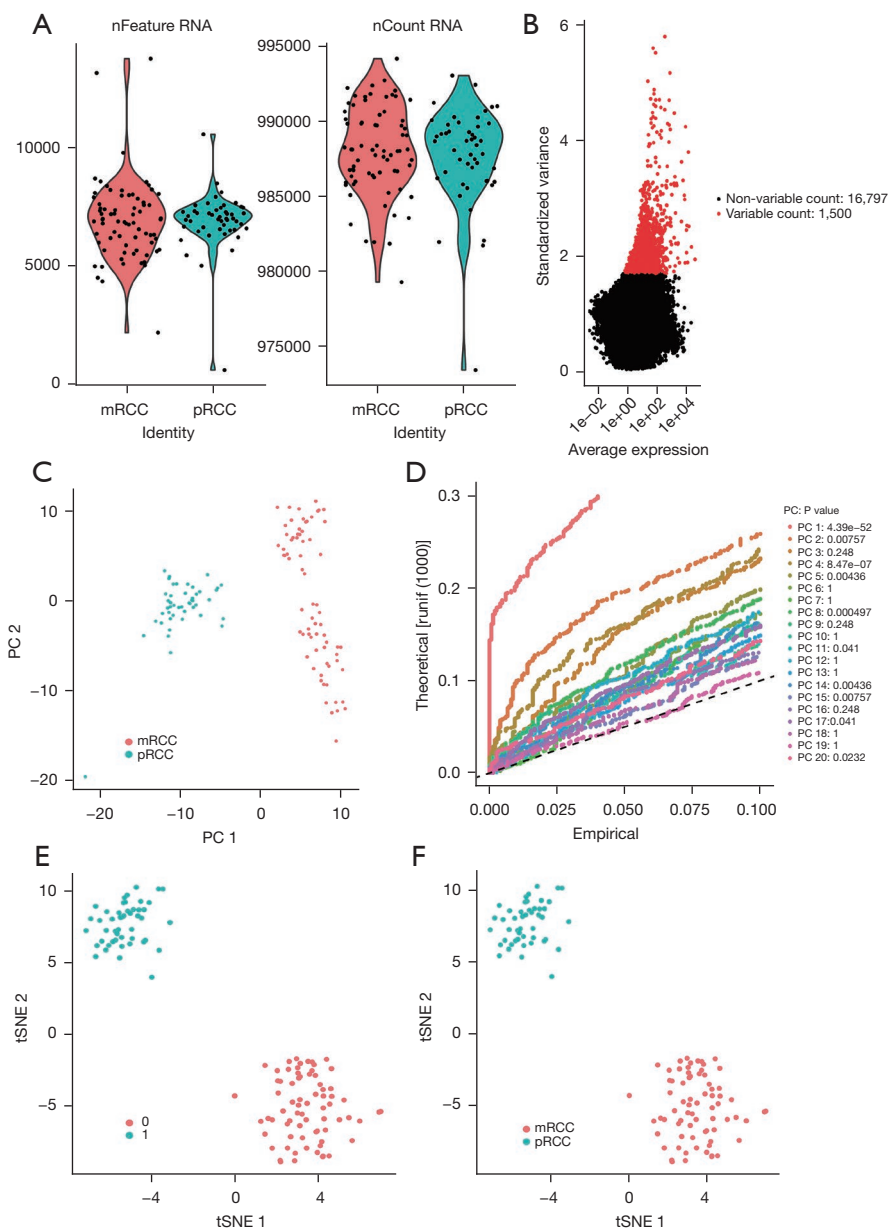
### ***Ethical statement***

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

## **Results**

### ***Tumor cell heterogeneity between metastatic renal cell carcinoma (mRCC) and primary renal cell carcinoma (pRCC)***

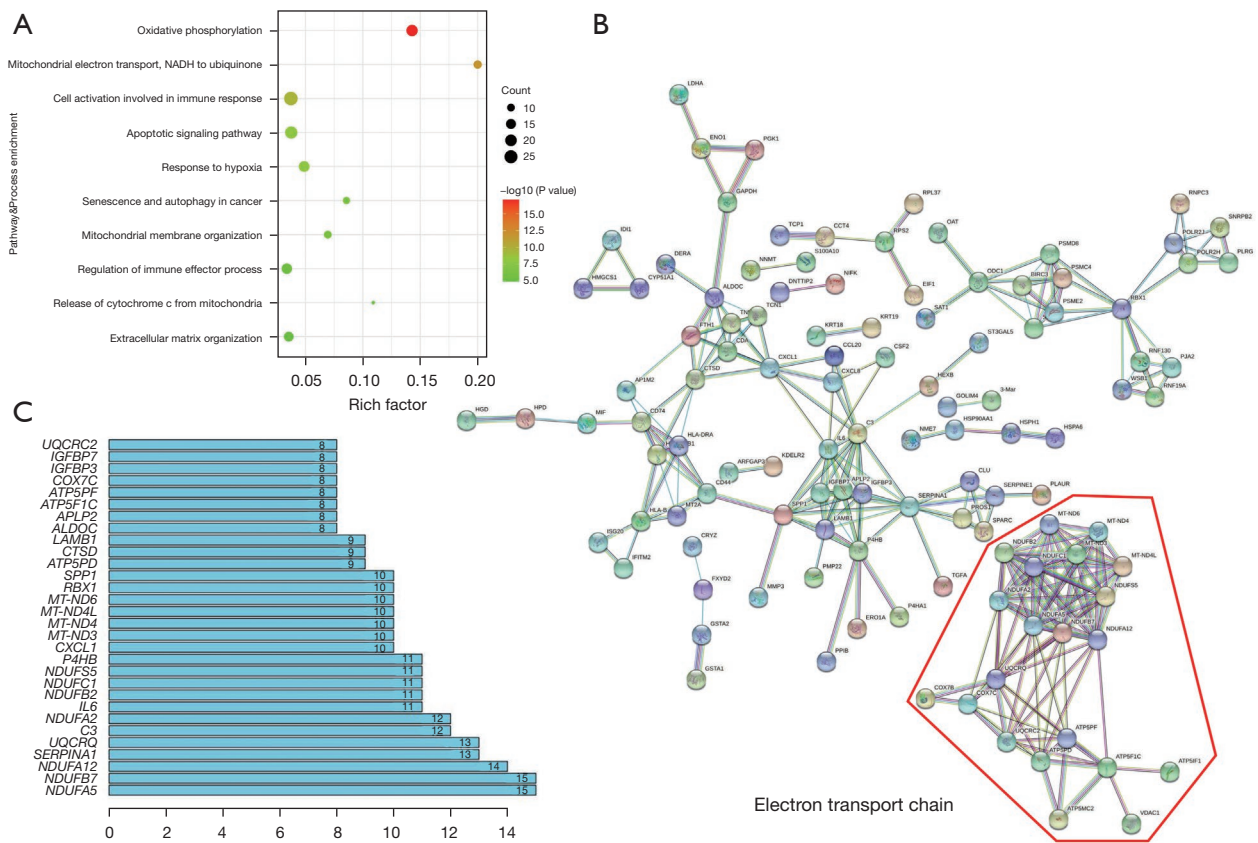
Prior to data analysis, rigorous quality control of the scRNA-seq was conducted. The numbers of RNA feature and RNA count of each cell were shown in *Figure 1A*. All samples have a similar RNA counts, while two samples were discarded because of inadequate RNA quality. Following normalization, we chose the top 1,500 maximum variance of genes for further identification of significant regulatory genes (*Figure 1B*). In order to determine tumor cell heterogeneity between mRCC and pRCC, the 1,500 variant genes screened before were analyzed using PCA. As showed in the cell distribution in the top 2 PCs (*Figure 1C*), discreteness was obviously observed, indicating that PCA analysis could achieve the purpose of dimensionality reduction and heterogeneity exists between pRCC and



**Figure 1** Tumor cell heterogeneity between mRCC and pRCC. (A) The feature and count of each cell sample from scRNA-seq. (B) The standardized variance of each gene. Red dots represent genes with top 1,500 maximum variance. (C) PCA plot based on scRNA-seq. (D) Theoretical, empirical score and P value of each PCA component. (E) tSNE analysis plot based on 5 PCs. (F) The results of cell types. mRCC, metastatic renal cell carcinoma; pRCC, primary renal cell carcinoma; scRNA-seq, single-cell RNA sequencing; tSNE, t-Distributed Stochastic Neighbor Embedding; PCA, principal component analysis.

mRCC. Twenty PCs also were obtained and p value of each PC was calculated (Figure 1D). Then, cluster analysis was performed in these 20 PCs, and eventually five PCs with the most significant difference were identified, including PC1, PC4, PC5, PC8, and PC14. Subsequently, cells in these

5 PCs were classified into 2 clusters via tSNE analysis (Figure 1E). Cell origins were traced, showing that cluster 1 exactly belongs to pRCC and cluster 0 exactly belongs to mRCC (Figure 1F). The gene expression heterogeneity between pRCC and mRCC was shown in the Figure 1F.



**Figure 2** ETC genes have differential expression between mRCC and pRCC. (A) Pathway and process enrichment based on DEGs. (B) The PPI network based on DEGs. (C) Genes with the most number of nodes. ETC, electron transport chain; mRCC, metastatic renal cell carcinoma; pRCC, primary renal cell carcinoma; DEGs, differentially expressed genes; PPI, protein-protein interaction.

**ETC genes have differential expression between mRCC and pRCC**

Differential expression analysis was performed between mRCC and pRCC. The results showed that 254 genes were expressed significantly. To investigate the biological function of DEGs, we performed a pathway and process enrichment analysis (Figure 2A). Pathway and process enrichment analysis of DEGs revealed significantly enriched in OXPHOS, mitochondrial electron transport, and response to hypoxia. In the meanwhile, mitochondrial membrane organization and release of cytochrome c from mitochondria also enriched. It means that mitochondrion and mitochondrial function, such as OXPHOS and ETC, probably played an important role in mRCC. Of cause, we also observed further enrichment in other cancer associated pathway and process, such as cell activation involved in

immune response, apoptotic signaling pathway, senescence and autophagy in Cancer, regulation of immune effector process, and extracellular matrix organization.

We utilized STRING database to construct PPI network for identifying the function and correlation of DEG’s proteins (Figure 2B). We found significant PPIs among these proteins, which revealed that those proteins might affect cell biological behavior via mutual adjustment. Several interactive groups were obvious and ETC proteins were the most significant interaction groups. In addition, quantifying the number of nodes showed that ETC genes, especially ubiquinone oxidoreductase core genes, also named mitochondrial CI, possessed more node numbers (Figure 2C). Thus, ETC genes had the potential to function on the metastasis of ccRCC cell. Besides that, many immunity-related genes with close association were also

found in the PPI network.

### ***Underexpressed ETC genes were associated with tumor metastasis process and higher immune responds***

We compared the expression of ETC genes between pRCC and mRCC [ $\log_2$ fold change (FC) mRCC/pRCC; *Figure 3A*]. Four mitochondrial encoding genes overexpressed in mRCC. By contrast, all nuclear encoded genes are underexpressed. Then, gene expression of each sample was displayed by the heatmap of hierarchical clustering (*Figure 3B*). An obviously expressed differential between pRCC and mRCC was shown.

The GSVA based on scRNA-seq was used for determining the variation of pathway in mRCC compared to pRCC (14). The hallmark gene sets which have significant difference in GSVA enrichment scores ( $|\log_2FCs| > 0.25$ ; P value  $< 0.05$ ) are sorted in decreasing order of  $\log_2FCs$ , with the most up-regulated pathway appearing at the top and the most down-regulated pathway appearing at the bottom (*Figure 3C*). Meanwhile, the heatmap showing each scRNA-seq sample's GSVA enrichment score of hallmark gene sets that were significantly up- and down-regulated (*Figure 3D*). The GSVA results revealed that the MYC TARGETS V1, MYC TARGETS V2, and TNFA SIGNALING VIA NFKB pathways have a higher GSVA enrichment score in mRCC cell sample. These pathways control several genes involved in cell growth, migration, and proliferation in ccRCC as previously reported. INTERFERON\_GAMMA\_RESPONSE and INTERFERON ALPHA RESPONSE also enriched in mRCC cell sample, it's similar to what has been previously reported (15). On the contrast, pRCC cell had higher GSVA enrichment score in GLYCOLYSIS, OXIDATIVE PHOSPHORYLATION, HYPOXIA, FATTY ACID METABOLISM, and HEME METABOLISM process. These highly enriched process means that mitochondrial function disorder and high immune responds is involved in mRCC.

In order to verify the relationship between ETC genes and tumor metastasis, we further validated using data from TCGA-KIRC patients. Most of ETC genes were downregulated in tumor compared with normal tissues (*Figure S1A*). The heatmap showing ETC genes expression in each tumor samples based on TCGA-KIRC data (*Figure S1B*). Hierarchical cluster was used, according to the expression levels of the 21 genes, to classify patients into two subgroups. Group A has a higher genes expression and group B has a lower genes expression. But the top 5 mutated

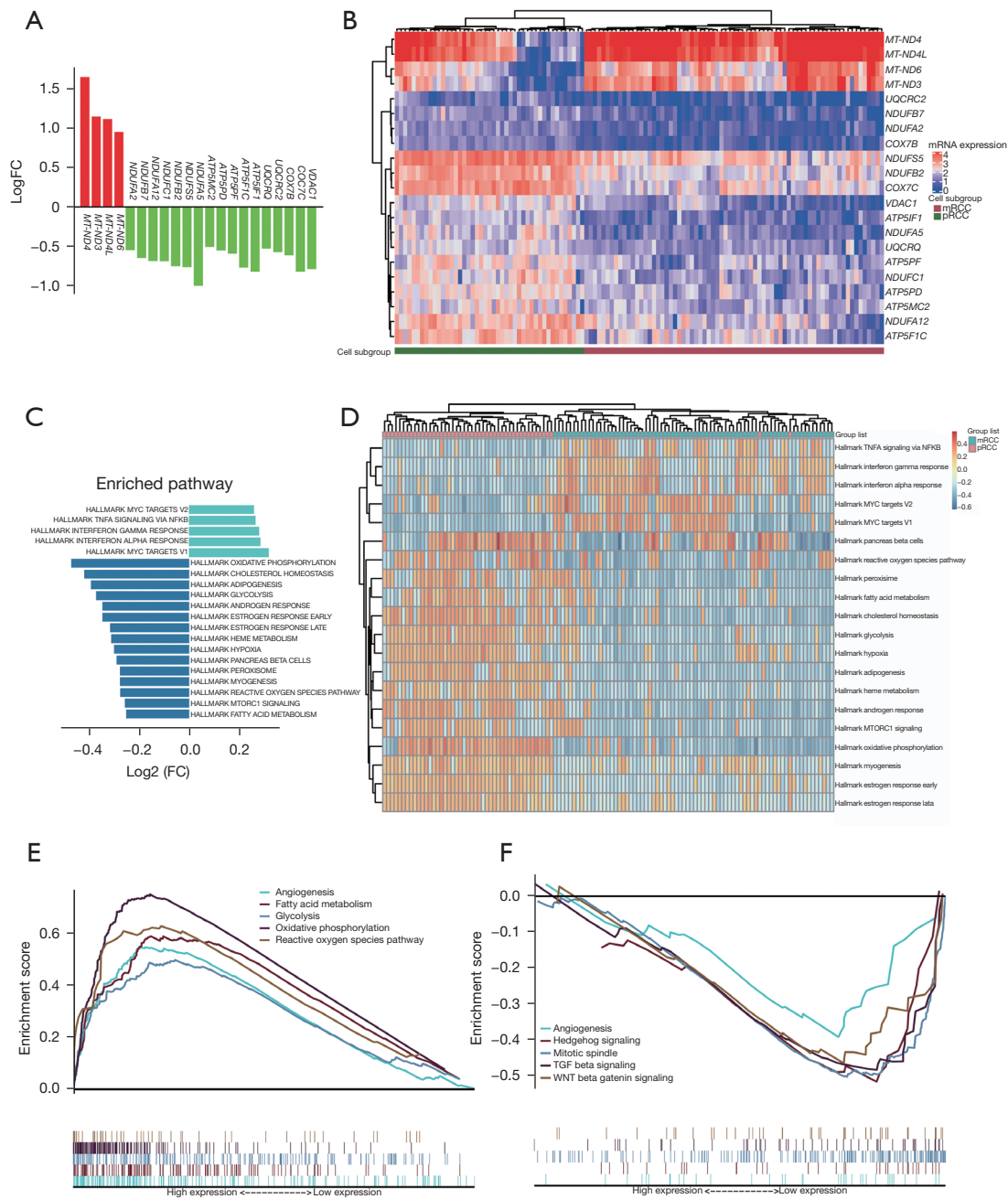
genes and clinical features in ccRCC do not show clear differences in two groups. GSEA enrichment results showed that cell proliferation, angiogenesis, and metastasis pathway was closely correlated with lower ETC genes expression (*Figure 3E*). Such as ANGIOGENESIS, MITOTIC SPINDLE, HEDGEHOG SIGNALING, TGF BETA SIGNALING, and WNT BETA CATENIN SIGNALING pathway were enriched in group B. On the contrast, GLYCOLYSIS, OXIDATIVE PHOSPHORYLATION, REACTIVE OXYGEN SPECIES PATHWAY, ADIPOGENESIS, and FATTY ACID METABOLISM were enriched in group A (*Figure 3F*). This result based on TCGA data were similar with GSVA results of scRNA-seq.

### ***Underexpressed CI genes were associated with tumor metastasis process and higher immune responds***

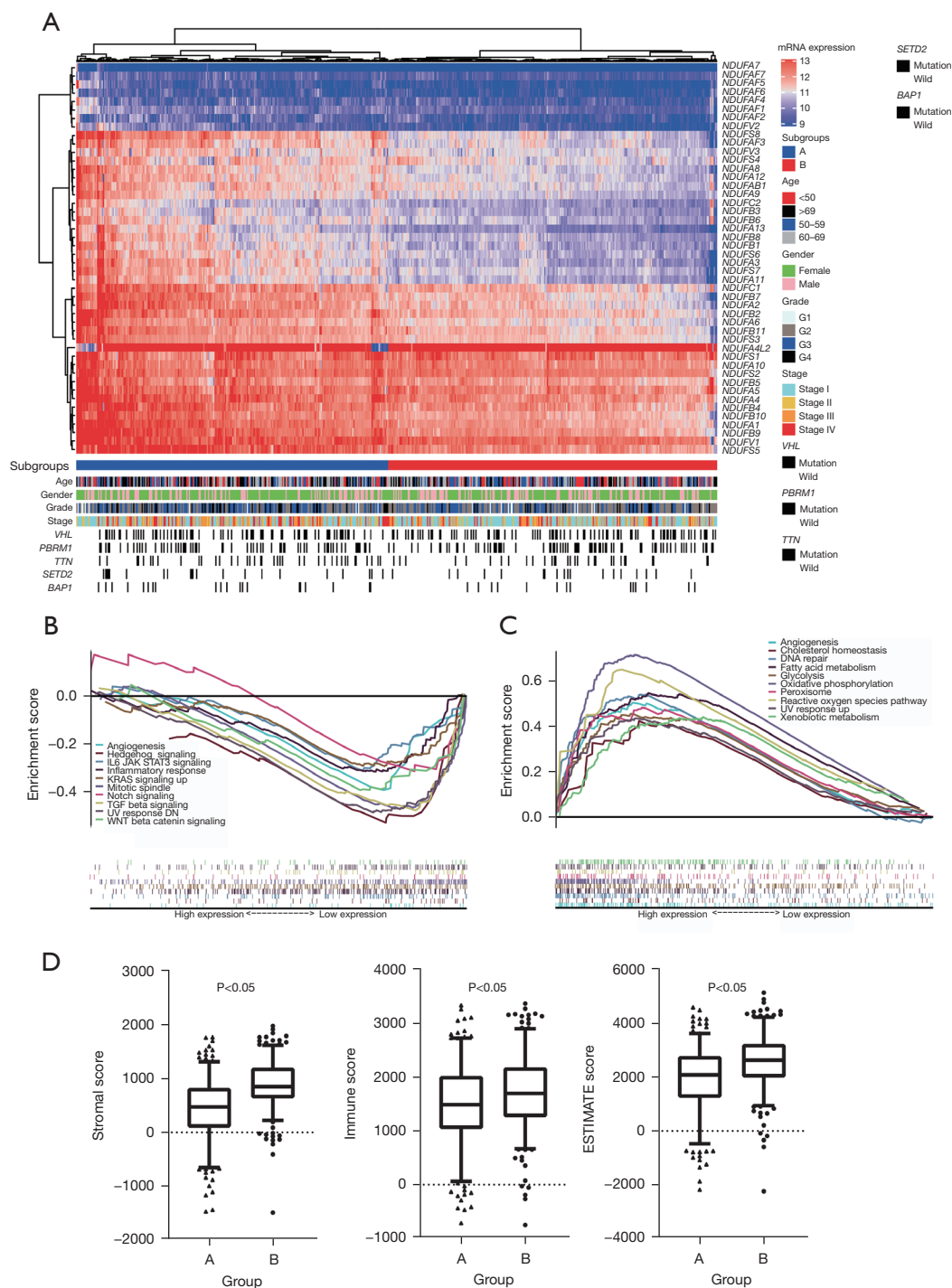
CI is the gatekeeper of the ETC, crucial for respiration in many aerobic organisms and catalyzes the first step of NADH oxidation (16). Also because ubiquinone oxidoreductase core genes possessed more node numbers and closely related to each other, we next focus on CI genes for further analysis.

So, we utilized hierarchical cluster classifying patients into two subgroups to verify CI genes expression difference in TCGA-KIRC tumor samples (*Figure 4A*). Group A has a higher CI genes expression and group B has a lower CI genes expression. Top 5 mutated genes and clinical features in ccRCC do not show clear differences in two groups. A GSEA analysis based on the two groups showed that cell proliferation, angiogenesis, and metastasis pathway enriched in lower CI genes expression (*Figure 4B*). Such as ANGIOGENESIS, HEDGEHOG\_SIGNALING, KRAS\_SIGNALING\_UP, MITOTIC\_SPINDLE, NOTCH\_SIGNALING, TGF\_BETA\_SIGNALING, and WNT\_BETA\_CATENIN\_SIGNALING. By contrast, ADIPOGENESIS, FATTY\_ACID\_METABOLISM, GLYCOLYSIS, OXIDATIVE\_PHOSPHORYLATION, PEROXISOME, and REACTIVE\_OXYGEN\_SPECIES\_PATHWAY enriched in higher CI genes expression (*Figure 4C*).

We also find immune regulatory pathway, such as IL6\_JAK\_STAT3\_SIGNALING and INFLAMMATORY\_RESPONSE, also be found enriched with lower CI genes expression. In order to explore the affection of CI genes on immune microenvironment, we used ESTIMATE of R to estimate stromal and immune cells in KIRC tumor sample, then compared the StromalScore, ImmuneScore



**Figure 3** Underexpressed ETC genes were associated with tumor metastasis process and higher immune responds. (A) Gene expression of ETC genes. Positive logFC represents higher expression and negative logFC represents lower expression in mRCC. (B) A heatmap showing gene expression of each sample. The cell types were shown below. (C) The differential enriched pathway based on GSEA result. (D) The GSEA score of each sample in differential enriched pathway based on GSEA result. (E,F) The GSEA results of TCGA sample based on ETC gene expression. FC, fold change; mRCC, metastatic renal cell carcinoma; pRCC, primary renal cell carcinoma; ETC, electron transport chain; GSEA, gene set variation analysis; GSEA, gene set enrichment analysis; TCGA, The Cancer Genome Atlas.



**Figure 4** Underexpressed mitochondrial CI genes were associated with tumor metastasis process and higher immune responds. (A) A heatmap showing CI gene expression based on TCGA sample. All sample are divided into two subgroups by hierarchical clustering analysis. The corresponding gene mutations and clinical features for each sample are shown below. (B,C) The GSEA results of TCGA sample based on CI gene expression. (D) The StromalScore, ImmuneScore and ESTIMATEScore of two groups clustered by hierarchical clustering analysis. CI, complex I; TCGA, The Cancer Genome Atlas; GSEA, gene set enrichment analysis.



and ESTIMATEScore of groups A and B (Figure 4D). The higher StromalScore, ImmuneScore and ESTIMATEScore in group A which had a lower CI genes expression sample. A lower CI gene expression could be associated with a strong immune response and immune microenvironment.

#### **Multiple mutated genes in ccRCC were associated with CI gene expression**

Renal oncocytomas as benign tumors are characterized by CI mutated (17). But CI mutated and what effects of high-frequency mutated genes for CI expressions are not clear in ccRCC. Next, we identified the five most frequently mutated genes and CI gene mutations in ccRCC based on TCGA-KIRC data (Figure 5A). *VHL*, *PBRM1*, *TTN*, *SETD2*, and *BAP1* were the most frequent gene mutations, which are same as previous researches (2,3,18), but few CI gene mutations could find in ccRCC. It means that the CI gene mutations didn't play an important role in the development of ccRCC, so differences of CI genes expression may be a principal mechanism to affect cell migration. Next, we calculated  $\log_2FC$  (wild/mutation) of each gene to verify whether *VHL*, *PBRM1*, *TTN*, *SETD2*, and *BAP1* mutations can affect CI genes expression (Figure 5B). Besides *NDUFA4L2* had an obvious overexpression in *VHL* and *PBRM1* mutated samples compared to wild samples, most of CI genes had a higher expression in wild samples compared to *VHL* and *PBRM1* mutated samples. By contrast, most of CI genes had a higher expression in *SETD2* and *BAP1* mutated samples compared to wild sample. In addition, there was no significant CI genes expression difference between *TTN* wild samples and mutated sample.

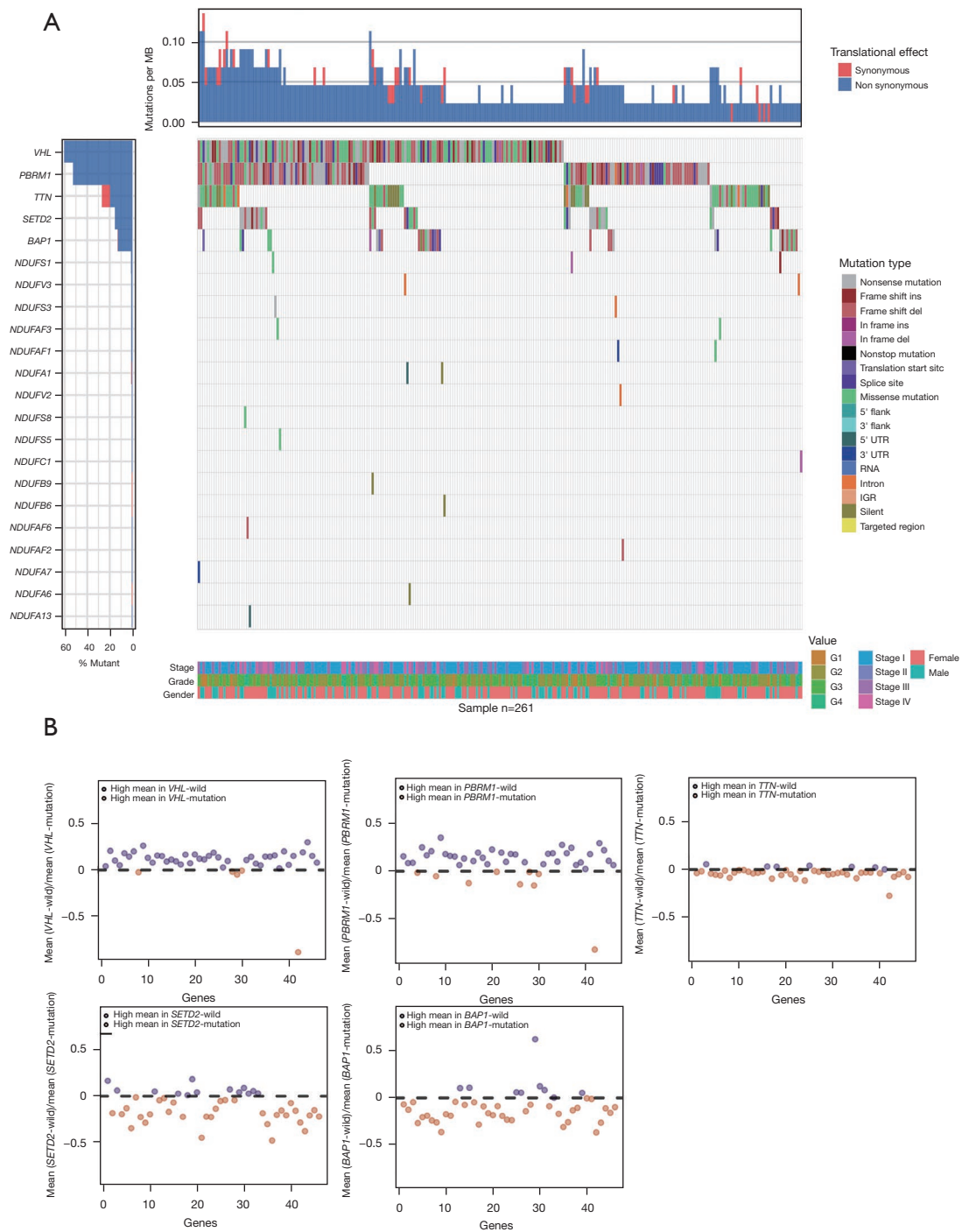
#### **A risk model based on CI expression can predict patient's prognosis**

To explore whether the patient's survival prognosis can be predicted by CI genes, univariate Cox regression analysis was conducted at first (Figure 6A). Univariate Cox regression analysis revealed 10 CI genes had significant P value ( $P < 0.05$ ) in overall survival (OS). *NDUFS1*, *NDUFS4*, *NDUFB6*, *NDUFA5*, and *NDUFA10* were designated as low-risk genes, while *NDUFAF3*, *NDUFA58*, *NDUFAF6*, *NDUFS6*, and *NDUFA3* were identified as high-risk genes. Then multivariate Cox proportional hazard regression analysis was conducted on the 10 CI genes to further screen, 7 CI genes were selected for a model construction.

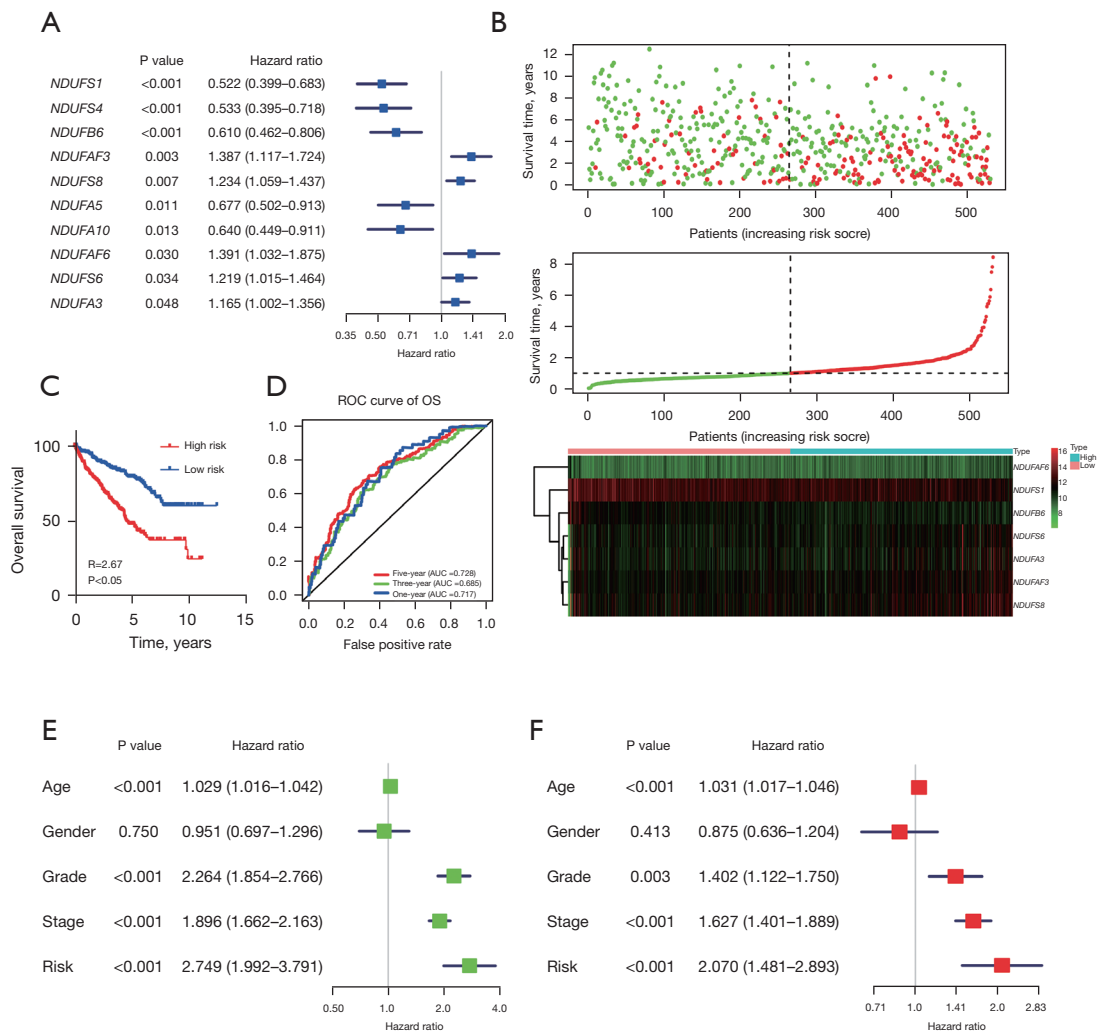
According to the 7 CI genes expression, the risk model was constructed as following formula: risk score =  $(-0.39 \times NDUFS1) + (-0.81 \times NDUFB6) + (0.43 \times NDUFAF3) + (0.71 \times NDUFS8) + (0.30 \times NDUFAF6) + (0.41 \times NDUFS6) + (-0.80 \times NDUFA3)$ . After that, patient's survival status and 7 CI genes expression were arranged according to the risk score (Figure 6B), revealing that the proportion of patients with poor prognosis was increased with risk score increased. *NDUFAF6*, *NDUFS6*, *NDUFAF3*, and *NDUFS8* increased as the risk score increased, while *NDUFS1*, *NDUFB6*, and *NDUFA3* decreased as the risk score increased. Then, patients were subdivided into high- and low-risk groups based on the cut-off of the median risk score. The Kaplan-Meier OS plot showed significant differences between high- and low-risk groups (Figure 6C). The ROC curves for OS showed that the 1, 3, and 5 years AUC of the risk score model were 0.717, 0.685, and 0.728, respectively (Figure 6D). Univariate Cox regression analyses revealed the 7 CI genes risk model could be a prediction maker for prognosis (Figure 6E). Multivariate Cox regression analyses revealed the 7 CI genes risk model could be independent prognostic factors (Figure 6F).

#### **PPARG could be a powerful CI genes transcription activator**

To verify the potential transcription factor of CI genes, we utilized DAVID bioinformatics resources to identify potential transcription factor binding site in CI genes' 5'-UTR. 11 transcription factors were predicted that more than half of CI genes had corresponding binding sites (Figure 7A). A complex regulatory network was shown in Figure 7B based on the prediction of DAVID bioinformatics resources. For simplifying this network and identifying transcription factor accurately, a correlation analysis between transcription factors and each CI genes was conducted (Figure 7C). We screened the correlation coefficient results by  $|cor| > 0.3$ , positive correlation coefficient transcription factor was considered to be transcription activator, and negative was transcription repressor. Then, we combined the results of regulatory network and correlation analysis, rebuilding the regulatory network (Figure 7D). *PPARG* and *MECOM* were transcription factors, while *FOXJ2*, *MEF2A*, *POU2F1*, *ZEB1*, and *RUNX1* were transcription repressor. Significantly, *PPARG* showed powerful regulation ability and could regulate 22 CI genes expression. Next, we assessed *PPARG* expression in different stage and grade using UALCAN database (Figure 7E, 7F) (19). *PPARG*



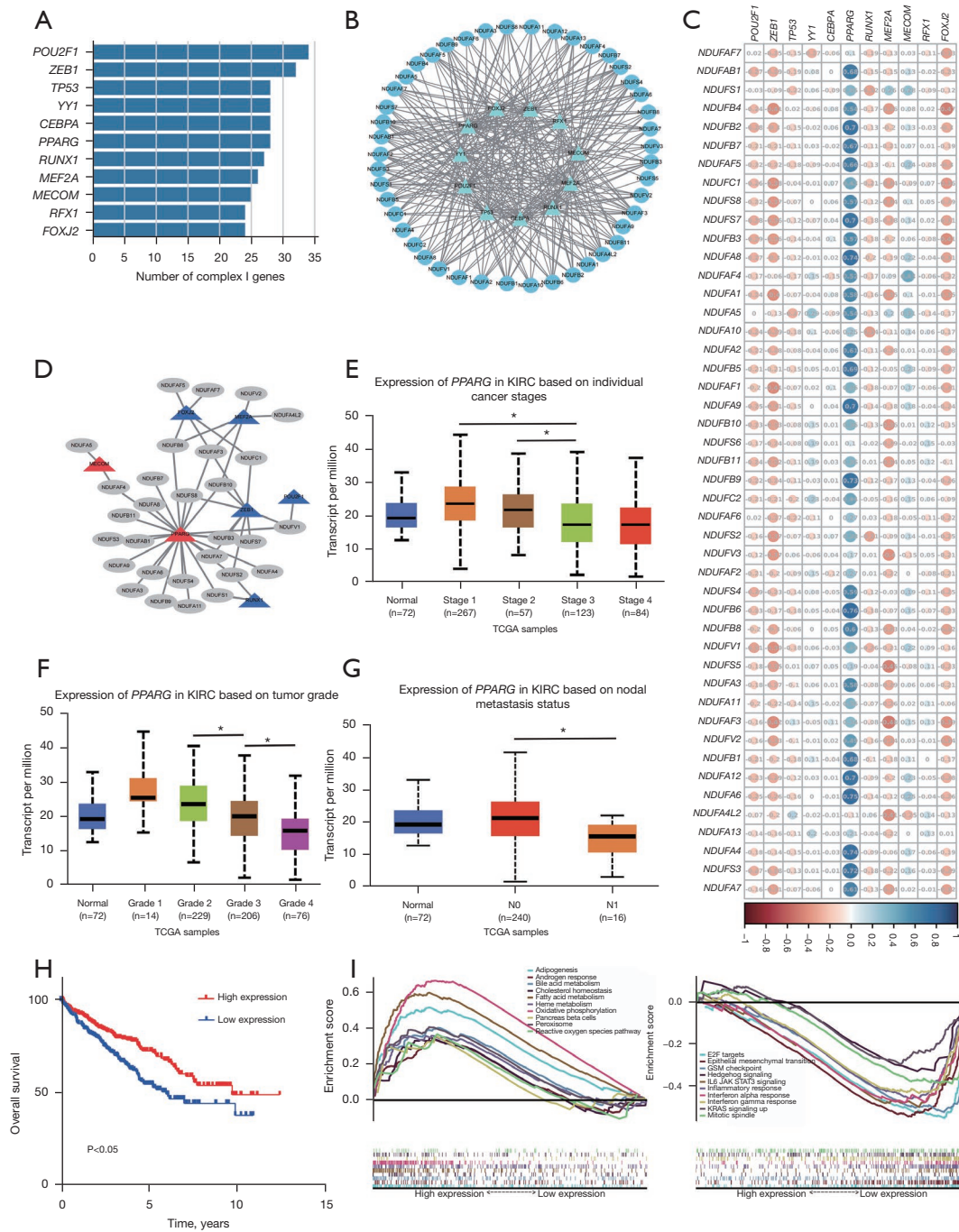
**Figure 5** Multiple mutated genes in ccRCC was associated with CI gene expression. (A) A waterfall plot showing the high-frequency mutated genes and CI gene mutation. clinical features for each sample are shown below. (B) CI gene expression based on multiple mutated genes. Upper represents high expression in wild sample and lower represents high expression in mutated sample. MB, megabase; ccRCC, clear cell renal cell carcinoma; CI, complex I.



**Figure 6** A risk model based on CI expression can predict patient's prognosis. (A) The result of univariate Cox regression analyses. (B) A combined plot showing the relationship among risk score, survival time, and expression of 7 genes. (C) Kaplan-Meier OS curves for patients with high/low risk score. (D) Time-dependent ROC analysis showing the 1-, 3- and 5-year AUC. (E) Univariate Cox regression analyses. (F) Multivariate Cox regression analyses. ROC, receiver operating characteristic; OS, overall survival; AUC, area under the curve; CI, complex I.

had no significant difference between normal and tumor sample, but showed a lower expression in high stage and grade than low stage and grade. Besides that, *PPARG* expression was lower in patients with node metastasis (Figure 7G). The Kaplan-Meier OS plot showed significant differences based on *PPARG* expression (Figure 7H), and higher *PPARG* expression had a better prognosis. GSEA analysis showed *PPARG* high expression is associated with *ADIPOGENESIS*, *BILE\_ACID\_METABOLISM*, *FATTY\_ACID\_METABOLISM*, *HEME\_METABOLISM*, *OXIDATIVE\_PHOSPHORYLATION*, and *REACTIVE\_OXYGEN\_SPECIES\_PATHWAY*,

while low expression is associated with *E2F\_TARGETS*, *EPITHELIAL\_MESENCHYMAL\_TRANSITION*, *G2M\_CHECKPOINT*, *HEDGEHOG\_SIGNALING*, *KRAS\_SIGNALING\_UP*, and *MITOTIC\_SPINDLE* (Figure 7I). This GSEA result was similar with that based on CI genes expression. We also find negative correlations between *PPARG* expression and *StromalScore*, *ImmuneScore* and *ESTIMATEScore* (Figure S1C). The correlation coefficients are  $-0.33$ ,  $-0.32$ , and  $-0.38$ , respectively. It reveals that *PPARG* could regulate immune microenvironment by their ability to transcription activated CI gene expressions.



**Figure 7** *PPARG* could be a powerful CI genes transcription activator. (A) The number of CI genes binding with each transcription factor. (B) A network showing the connect between CI gene and transcription factor. (C) The result of correlation analysis. (D) A network showing the connect between CI gene and transcription factor after screening. Red triangle represents transcription activator and blue represents transcription inhibitor. (E) The expression of *PPARG* in KIRC based on individual cancer stages. (F) The expression of *PPARG* in KIRC based on tumor grade (G) The expression of *PPARG* in KIRC based on nodal metastasis status. (H) Kaplan-Meier OS curves for patients based on *PPARG* expression. (I) The GSEA result based on *PPARG* expression. \*,  $P < 0.05$ . KIRC, kidney renal clear cell carcinoma; TCGA, The Cancer Genome Atlas; CI, complex I; OS, overall survival; GSEA, gene set enrichment analysis.

## Discussion

Previous studies have shown that low mitochondrial respiratory chain content correlates with tumor aggressiveness in RCC (15,20). Besides, mitochondrial function is tightly associated with the activity of the respiratory chain complexes. But the functions of specific proteins in mitochondrial respiratory electron transfer chain contents have not been clarified. Here, we explored the heterogeneity between mRCC and pRCC and found that ETC, especially CI, was decreased in mRCC, indicating low CI have a significant role in ccRCC metastasis, which is accomplished by activation of multiple signaling pathways. During the metastatic cascade, cancer cells tightly interact with the immune system and they influence each other (21). The ccRCC is characterized by robust immune cell infiltration and response to immunotherapy (22). Additionally, the cancer cell intrinsic inflammation was involved in ccRCC metastasis (23). In support of these results, we showed that CI gene expression was associated with immune responds and immune microenvironment, implying that a lower CI genes expression indirectly promote cancer cell metastasis through the regulation of immune responses.

Emerging evidence has shown that CI encoding genes play an important role in the dysfunction of OXPHOS system (24). Mutations of CI related genes may not only contribute to the carcinogenesis of renal oncocytoma, but also could be used as potential diagnostic markers to distinguish renal oncocytoma from chromophobe RCC (25). However, there are few reports on how *VHL* and *PBRM1* mutation influence CI in ccRCC (2,3,18). According to previous reports, loss of *pVHL* function causes hypoxia-inducible factor (HIF)-mediated glycolysis enhancement and inhibition of mitochondrial function in RCC cells (26-28). Our data showed that CI has low mutation frequency, implying that the differences in CI gene expression, rather than mutations, may influence ccRCC metastasis. Meanwhile, the expression of most CI encoding genes was higher in *VHL* and *PBRM1* mutated sample, and lower in *SETD2* and *BAP1* mutated samples. The results suggest that *VHL*, *PBRM1*, *SETD2*, and *BAP1* mutation may influence mitochondrial function.

It has been demonstrated that HIF and hypoxia-related pathways play critical roles in the development and progress of RCC (26,29-31). Small molecule inhibitors that target HIF are promising therapeutical agents for RCC. It has recently been reported that low oxygen levels

stabilized HIF-1 $\alpha$  and increased CI levels. Meanwhile, knockdown of HIF-1 $\alpha$  abrogated the effect of hypoxia on CI levels (31). Here, we showed that *PPARG* can function as a transcription activator for CI genes. Moreover, lower *PPARG* expression is associated with advanced tumor stage, pathological grade, lymph node metastasis, worse prognosis, and higher potential of metastasis. Thus, *PPARG* has potential for use as a therapeutic target to enhance the immunotherapy treatment's efficacy in RCC.

In conclusion, our study revealed that CI plays an important role in ccRCC progress. CI expression is associated with the high frequency mutated genes. We constructed a 7 CI genes model to predict the prognosis of ccRCC. *PPARG* could be a potential transcription activator for CI. CI and *PPARG* may be potential biomarkers for metastatic ccRCC.

## Acknowledgments

*Funding:* None.

## Footnote

*Reporting Checklist:* The authors have completed the REMARK reporting checklist. Available at <https://tcr.amegroups.com/article/view/10.21037/tcr-22-242/rc>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <https://tcr.amegroups.com/article/view/10.21037/tcr-22-242/coif>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

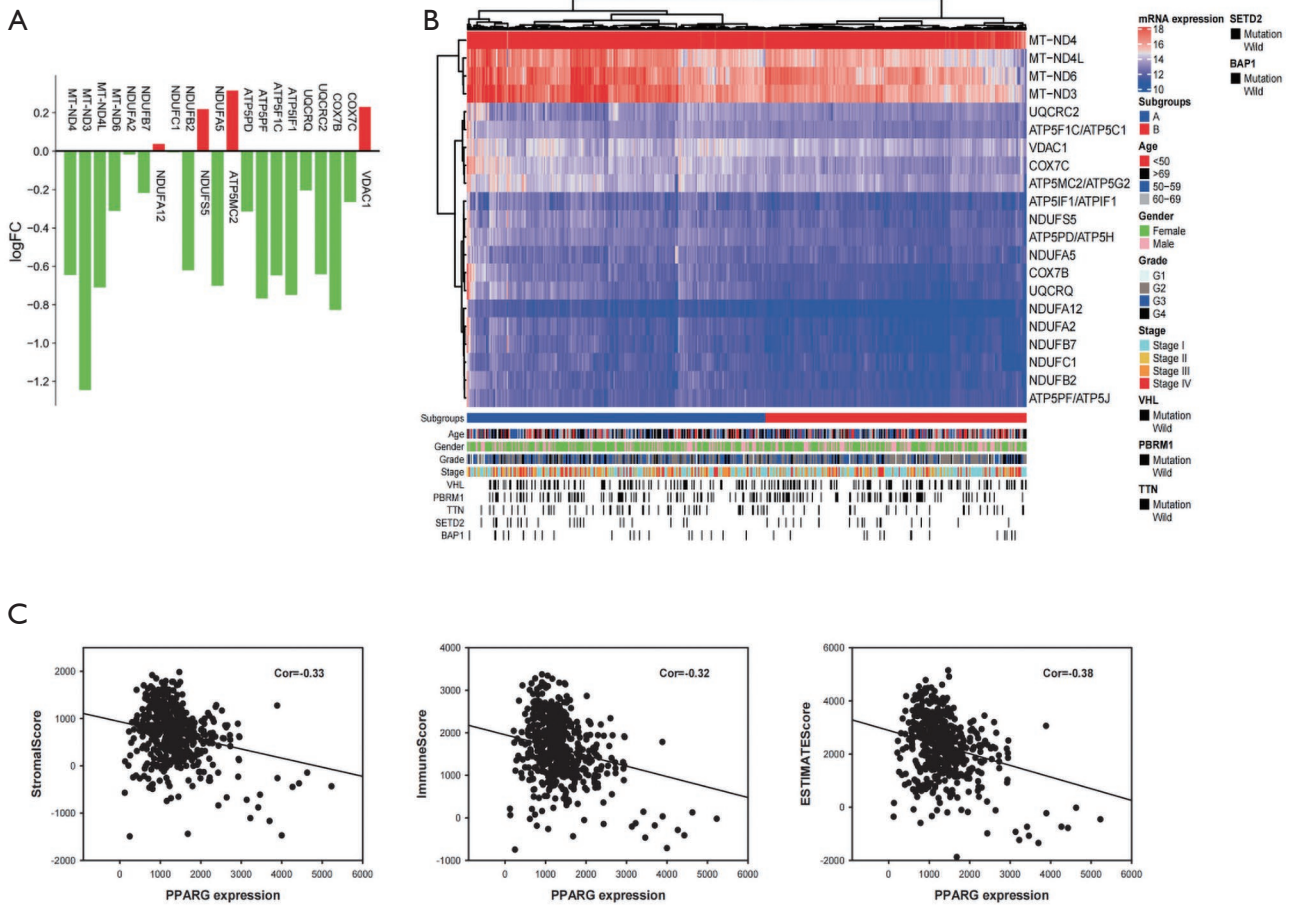
*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

- Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71:209-49.
- Hsieh JJ, Purdue MP, Signoretti S, et al. Renal cell carcinoma. *Nat Rev Dis Primers* 2017;3:17009.
- Rini BI, Campbell SC, Escudier B. Renal cell carcinoma. *Lancet* 2009;373:1119-32.
- Wiechno P, Kucharz J, Sadowska M, et al. Contemporary treatment of metastatic renal cell carcinoma. *Med Oncol* 2018;35:156.
- Nolfi-Donagan D, Braganza A, Shiva S. Mitochondrial electron transport chain: Oxidative phosphorylation, oxidant production, and methods of measurement. *Redox Biol* 2020;37:101674.
- Mai N, Chrzanowska-Lightowlers ZM, Lightowlers RN. The process of mammalian mitochondrial protein synthesis. *Cell Tissue Res* 2017;367:5-20.
- Birsoy K, Wang T, Chen WW, et al. An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. *Cell* 2015;162:540-51.
- Raimondi V, Ciccarese F, Ciminale V. Oncogenic pathways and the electron transport chain: a dangerROS liaison. *Br J Cancer* 2020;122:168-81.
- Martínez-Reyes I, Cardona LR, Kong H, et al. Mitochondrial ubiquinol oxidation is necessary for tumour growth. *Nature* 2020;585:288-92.
- Santidrian AF, Matsuno-Yagi A, Ritland M, et al. Mitochondrial complex I activity and NAD<sup>+</sup>/NADH balance regulate breast cancer progression. *J Clin Invest* 2013;123:1068-81.
- Kim KT, Lee HW, Lee HO, et al. Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. *Genome Biol* 2016;17:80.
- Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019;10:1523.
- Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 2021;49:D605-12.
- Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
- Chakraborty S, Balan M, Sabarwal A, et al. Metabolic reprogramming in renal cancer: Events of a metabolic disease. *Biochim Biophys Acta Rev Cancer* 2021;1876:188559.
- Hirst J. Mitochondrial complex I. *Annu Rev Biochem* 2013;82:551-75.
- Gopal RK, Calvo SE, Shih AR, et al. Early loss of mitochondrial complex I and rewiring of glutathione metabolism in renal oncocytoma. *Proc Natl Acad Sci U S A* 2018;115:E6283-90.
- D'Avella C, Abbosh P, Pal SK, et al. Mutations in renal cell carcinoma. *Urol Oncol* 2020;38:763-73.
- Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* 2017;19:649-58.
- Simonnet H, Alazard N, Pfeiffer K, et al. Low mitochondrial respiratory chain content correlates with tumor aggressiveness in renal cell carcinoma. *Carcinogenesis* 2002;23:759-68.
- Blomberg OS, Spagnuolo L, de Visser KE. Immune regulation of metastasis: mechanistic insights and therapeutic opportunities. *Dis Model Mech* 2018;11:dmm036236.
- Mier JW. The tumor microenvironment in renal cell cancer. *Curr Opin Oncol* 2019;31:194-9.
- Nishida J, Momoi Y, Miyakuni K, et al. Epigenetic remodelling shapes inflammatory renal cancer and neutrophil-dependent metastasis. *Nat Cell Biol* 2020;22:465-75.
- Distelmaier F, Koopman WJ, van den Heuvel LP, et al. Mitochondrial complex I deficiency: from organelle dysfunction to clinical disease. *Brain* 2009;132:833-42.
- Mayr JA, Meierhofer D, Zimmermann F, et al. Loss of complex I due to mitochondrial DNA mutations in renal oncocytoma. *Clin Cancer Res* 2008;14:2270-5.
- Schödel J, Grampp S, Maher ER, et al. Hypoxia, Hypoxia-inducible Transcription Factors, and Renal Cancer. *Eur Urol* 2016;69:646-57.
- Briston T, Stephen JM, Thomas LW, et al. VHL-Mediated Regulation of CHCHD4 and Mitochondrial Function. *Front Oncol* 2018;8:388.
- Zhang H, Gao P, Fukuda R, et al. HIF-1 inhibits mitochondrial biogenesis and cellular respiration in VHL-deficient renal cell carcinoma by repression of C-MYC activity. *Cancer Cell* 2007;11:407-20.
- Cho H, Du X, Rizzi JP, et al. On-target efficacy of a HIF-2 $\alpha$  antagonist in preclinical kidney cancer models. *Nature*

- 2016;539:107-11.
30. Masoud GN, Li W. HIF-1 $\alpha$  pathway: role, regulation and intervention for cancer therapy. *Acta Pharm Sin B* 2015;5:378-89.
31. Saldana-Caboverde A, Nissanka N, Garcia S, et al. Hypoxia Promotes Mitochondrial Complex I Abundance via HIF-1 $\alpha$  in Complex III and Complex IV Efficient Cells. *Cells* 2020;9:2197.

**Cite this article as:** Zhang F, Hou T, Chen L, Xiong M, Zhou M, Kazobinka G, Zhao J, Han X. Comprehensive analysis of lower mitochondrial complex I expression is associated with cell metastasis of clear cell renal cell carcinoma. *Transl Cancer Res* 2022;11(6):1488-1502. doi: 10.21037/tcr-22-242



**Figure S1** ETC genes were downregulated in tumor. (A) Gene expression of ETC genes based on TCGA-KIRC sample. Positive logFC represents higher expression and negative logFC represents lower expression in tumor sample. (B) A heatmap showing ETC gene expression based on TCGA sample. All sample are divided into two subgroups by hierarchical clustering analysis. The corresponding gene mutations and clinical features for each sample are shown below. (C) The correlation between PPARG expression and StromaScore, ImmuneScore and ESTIMATEScore. FC, fold change; ETC, electron transport chain; TCGA, The Cancer Genome Atlas; KIRC, kidney renal clear cell carcinoma.