Abstract. Clear cell renal cell carcinoma (ccRCC) originates from renal tubular epithelial cells and is the most common pathological renal cell carcinoma type with the worst prognosis. The relationship between the expression, prognosis and mechanism of ccRCC and the E2F family remains challenging. In the present study, RNA sequencing and clinical data of ccRCC from The Cancer Genome Atlas and two datasets, GSE36895 and GSE53757, from the Gene Expression Omnibus were used to identify the role of the E2F family in ccRCC. A total of 10 groups of tumor tissues and paired-normal tissues from patients with ccRCC were verified by reverse transcription-quantitative PCR. The expression, tumor grade and stage, prognosis and regulatory mechanism of the E2F family in ccRCC were analyzed. It was found that the expression levels of E2F1 to 4 and 6 to 8 were higher in ccRCC tissues than in normal tissues, whereas the expression level of E2F5 was lower in the former than in the latter. The expression levels of E2F1 to 8 were correlated with tumor stage and grade. Low expression of E2F1 to 5 and 7 to 8 was significantly associated with longer overall survival, disease-specific survival and progression-free survival times. The data revealed that the E2F family rarely has genetic mutations. The expression of E2F1, E2F2, E2F5, E2F7 and E2F8 was significantly correlated with DNA methylation, and E2F1 to E2F7 were significantly correlated with copy number and the data showed that the expression of E2Fs was significantly correlated with the cell cycle. The results of the present study suggested that E2F family genes may be potential targets for ccRCC molecular diagnosis and targeted therapy.

Introduction
Renal cell carcinoma (RCC) is a common malignant tumor in the human genitourinary system that includes numerous different pathological subtypes (1,2). The most common subtype is clear cell renal cell carcinoma (ccRCC). ccRCC originates from renal tubular epithelial cells and accounts for ~60-85% of RCCs (3,4). The prevalence of ccRCC in men is higher than that in women, and most patients are over 60 years old. It is often asymptomatic in the early stage or only has vague systemic symptoms such as fever and fatigue. The typical clinical symptoms are hematuria, pain and a palpable mass in the kidney area. It affects either kidney at an equal rate. At present, the treatment methods are limited, mainly radical nephrectomy, but relapse and metastasis easily occur after surgery and the fatality rate is high (5). The etiology of ccRCC remains unclear and possible related factors include genetics, smoking, obesity, hypertension and antihypertensive drug therapy (6). Therefore, it is very important to explore the potential biomarkers and therapeutic targets of ccRCC.

The E2F family encodes extremely important nuclear transcription factors involved in regulating the cell cycle (7,8). It was first discovered by Kovesdi et al (9) in 1986 during studies of adenovirus. There are numerous members of the E2F family. The ones that have been discovered include E2F1 to E2F8. According to the protein structure and function and transcription characteristics, they are divided into
transcription promotion factors (E2F1 to E2F3) and transcription suppressors (E2F4 to E2F8) (10). Clinical studies have found that E2F family proteins are closely related to the occurrence, development, proliferation and apoptosis of gastric, lung, liver, esophageal, prostate, bladder and ovarian cancer and other malignant tumors (7,11). In addition, E2F family proteins exhibit complex and diverse biological functions in different tumors and their expression levels are not consistent in different tumors (11). However, the expression pattern of E2F family proteins in ccRCC and their relationship with the prognosis remains unclear.

To the best of our knowledge, bioinformatics analysis has yet to be applied to explore the role of E2F family in ccRCC. In the present study, RNA sequencing (RNA-Seq) data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) were downloaded to explore the expression characteristics of E2F family proteins in ccRCC and their relationship with prognosis. Finally, the results of bioinformatics analysis we reverified with clinical samples from surgical operations. The present study may provide a new understanding of E2F family proteins in ccRCC and help to interpret the mechanisms underlying their functions.

Materials and methods

Raw data. RNA-Seq and clinical data (from 530 tumor tissues and 72 normal tissues) of ccRCC from TCGA were downloaded from https://portal.gdc.cancer.gov/. The RNA-Seq data were reported as fragments per kilobase million. The datasets GSE36895 (12) (containing 72 tumor tissues and 72 normal tissues) and GSE53757 (13) (containing 29 tumor tissues and 23 normal tissues) were downloaded from GEO (https://www.ncbi.nlm.nih.gov/geo).

Reverse transcription-quantitative (RT-q) PCR. A total of 10 groups of tumor tissues and paired-normal tissues were obtained from 10 patients who underwent radical resection at The Affiliated Suqian First People's Hospital of Nanjing Medical University between January 2020 and August 2021. The patients were diagnosed as ccRCC by imaging and pathological examination, and did not receive chemotherapy or radiotherapy before operation. All experimental procedures were approved (approval no. 2018-SL-0026) by the Ethics Committee of The Affiliated Suqian First People's Hospital of Nanjing Medical University (Suqian, China). Written informed consent was provided by all patients prior to the study. All patients (age range, 55-68 years; median age, 62 years; seven men and three women) were diagnosed with ccRCC by laboratory examination and imaging examination. The clinical information of the patients is provided in Table SI. The mRNA expression of the E2Fs was examined using RT-qPCR. Total RNA was extracted using TRIzol® reagent (Invitrogen; Thermo Fisher Scientific, Inc.) from the tissues. Total RNA was converted into cDNA, and quantitated using a Fastking One Step Reverse Transcription and Fluorescence Quantitative Kit (Tiangen Biotech Co., Ltd.) as a fluorophore according to the manufacturer's protocol. The primers were designed by online tool 'primerBank' (https://pga.mgh.harvard.edu/primerbank). β-actin was used as the internal reference gene. The primer sequences of E2Fs are presented in Table SII. Expression levels of mRNAs relative to β-actin were determined using the 2−ΔΔCq method (14).

Definitions of clinical survival and recurrence types. Raw counts of RNA-sequencing data (level 3) of ccRCC from TCGA were downloaded from https://portal.gdc.cancer.gov/. A total of 3 types of clinical survival and recurrence outcomes were selected in the present study: overall survival (OS), disease-specific survival (DSS), progression-free survival (PFS) and disease-free survival (DFS) (15). Survival analysis with the log-rank test was used to compare the survival difference between the normal and tumor groups. The hazard ratio (HR) was used to indicate the risk difference between the two groups.

Mutation and methylation analysis. The data were downloaded from TCGA (https://portal.gdc.cancer.gov/) and cBioPortal (https://www.cbioportal.org/datasets), copy number and mutation analysis were performed using cBioPortal (https://www.cbioportal.org/) (15). MeV software (http://projects/mev-tm4) was used to create heatmaps. Pearson's correlation analysis and mapping were performed using R 4.1.1 (https://www.r-project.org/). P<0.05 indicates a significant correlation.

Enrichment analysis. Gene enrichment analysis was performed to determine the correlation between E2F expression and the cell cycle. The data were downloaded from TCGA. Gene set enrichment analysis (https://www.gsea-msigdb.org/gsea) was used to perform gene enrichment analysis in three datasets, REACTOME, (Kyoto Encyclopedia of Genes and genomes (KEGG) and Gene Ontology (GO). A normalized enrichment score (NES)>0 means E2F expression is positively correlated with the cell cycle, and a false discovery rate (FDR)<0.05 was considered to indicate a statistically significant difference.

Statistical analysis. The Mann-Whitney U test was performed using GraphPad Prism 9 software (GraphPad Software, Inc.) and R 4.1.1 was used to determine E2F expression between normal tissues and tumor tissues. Kaplan-Meier survival curves and log-rank tests were used to evaluate the effect of E2F expression on survival. The Kruskal-Wallis test was used to detect the differences between groups in normal tissue + stage/grade and stage/grade. P<0.05 was considered to indicate a statistically significant difference.

Results

E2F expression in patients with ccRCC. RNA-Seq data from 530 tumors and 72 normal tissue samples from the TCGA dataset were analyzed (Fig. 1A). In data of TCGA, compared with normal tissues, E2F1, E2F2, E2F3, E2F4, E2F6, E2F7 and E2F8 were overexpressed in cancer tissues, and E2F5 was expressed at low levels in cancer tissues (Fig. 1B). In dataset GSE36985, the expression levels of E2F1, E2F2, E2F6, E2F7 and E2F8 in tumor tissues were higher than those in normal tissues, and the expression of E2F5 was lower than that in normal tissues. There was no statistically significant difference between the two groups of E2F3 and E2F4 (Fig. 1C). In dataset GSE53757, the expression of E2F2, E2F6, E2F7 and E2F8 in tumor tissues was higher than that in normal tissues,
and the expression of E2F4 and E2F5 was lower than that in normal tissues. There was no statistically significant difference between the two groups of E2F1 and E2F3 (Fig. 1D). In addition, the RT-qPCR verification results of clinical specimens collected from ccRCC surgery revealed that compared with normal tissues, E2F5 expression in cancer tissues was low, and the E2F1, E2F, E2F4, E2F6, E2F7 and E2F8 were overexpressed (Fig. 1E).

E2F family expression and pathological status of patients with ccRCC. The relationship of E2F family expression with stage and grade was analyzed, respectively in TCGA data. The expression of E2Fs is significantly different in the staging and grading of patients with ccRCC. The higher the pathological stage was, the higher the expression of E2F1, E2F2, E2F3, E2F4, E2F6, E2F7 and E2F8 and the lower the expression of E2F5 (Fig. 2A). The results of pathological grading were similar to the results of staging (Fig. 2B).

Prognostic value of E2Fs in patients with ccRCC. To evaluate the clinical significance of the expression of E2Fs in the survival of patients with ccRCC, E2F expression was assessed in 530
ccRCC clinical samples (≤4,000 days of follow-up) from TCGA for OS, DSS and PFS. Low expression of E2F1, E2F2, E2F3, E2F4, E2F5, E2F7 and E2F8 were significantly associated with a longer OS time (HR=1.989, P<0.0001; HR=2.111, P<0.0001; HR=1.779, P=0.000; HR=1.941, P<0.0001; HR=2.194, P<0.0001; HR=2.656, P<0.0001; HR=1.616, P=0.001), a longer DSS time (HR=3.042, P<0.0001; HR=3.121, P<0.0001; HR=2.058, P=0.005; HR=2.433, P<0.0001; HR=2.355, P<0.0001; HR=3.746, P<0.0001; HR=2.537, P<0.0001) and a longer PFS time (HR=2.030, P<0.0001; HR=1.976, P<0.0001; HR=1.406, P=0.041; HR=1.621, P=0.004; HR=1.477, P=0.033; HR=2.579, P<0.0001; HR=1.908, P<0.0001) in patients with ccRCC, respectively (Fig. 3). However, low expression of E2F6 was not associated with OS (HR=1.319; P=0.068), DSS (HR=1.314; P=0.156) or PFS (HR=0.717; P=0.077) in patients with ccRCC (Fig. 3). However, it was found that the expression of E2F1-E2F8 was not significantly associated with DFS. These results suggested that the mRNA expression levels of E2F1, E2F2, E2F3, E2F4, E2F5, E2F7 and E2F8 may be useful for the prediction of survival of patients with ccRCC.

**Mechanism of E2F imbalance in patients with ccRCC.** The mutation rates of E2F1-E2F8 were 0, 0.2, 0, 0.2, 1, 0, 0 and 0.4%, respectively (Fig. 4A). The expression of E2F1, E2F2, E2F5 and E2F8 was significantly negatively correlated with DNA methylation ($r=-0.25$, $P<0.0001$; $r=-0.35$, $P<0.0001$;
Figure 3. Expression of E2Fs was associated with a favorable prognosis in patients with ccRCC. Kaplan-Meier plots depicting the OS, DSS and PFS of patients with high and low expression of E2Fs. All patients with cervical squamous cell carcinoma were included in the TCGA database. OS, overall survival; DSS, disease-specific survival; PFS, progression-free survival; HR, hazard ratio.
Figure 4. Mechanism of E2Fs dysregulation in patients with ccRCC. (A) Mutation rate of the E2F family in patients with ccRCC. (B) Heatmap displaying E2F expression levels and DNA methylation levels in ccRCC tissues and normal tissues. (C) Association between E2F expression levels and DNA methylation levels in ccRCC tissues. (D) Association between E2F expression level and copy number in ccRCC tissues. ccRCC, clear cell renal cell carcinoma.
recent years. They are atypical family factors and can work in concert with hypoxia-inducible factors to jointly regulate the transcription of vascular endothelial growth factor. E2Fs have been reported to have roles in a variety of cancer types, such as breast (18), liver (19) and gastric cancer (20), since they can regulate numerous cellular functions related to cell cycle progression. Although certain E2F family members have been confirmed to play promising roles in ccRCC, the distinct roles of E2Fs in the development, progression and metastasis of ccRCC remain to be elucidated. In the present study, the expression, mutation and prognostic values of different E2Fs in patients with ccRCC were analyzed.

E2F1 is the most frequently studied gene in the E2F family. E2F1 is an important transcription-promoting factor located on human chromosome 20q11, ~11 kb in size, mainly comprised of six introns and seven exons, and it can encode proteins with a size of more than 400 amino acids (10,21). E2F1 has a very obvious tissue specificity. It can form a heterodimer with retinoblastoma and bind to the corresponding DNA sequence, thereby enhancing or inhibiting the activity of E2F1. The transcription factor E2F1 can regulate biological processes such as the cell cycle, cell proliferation, cell apoptosis and cell differentiation (10). In the present study, E2F1 was highly expressed in cancer tissues and its high expression was closely related to a worse tumor grade and staging and a poor prognosis. In ccRCC, the expression of E2F1 is also significantly positively correlated with the cell cycle.

E2F2 and E2F3 are involved in regulation of the cell cycle and are highly expressed in breast cancer tissues (18). At present, little is known about the expression and role of E2F2 in ccRCC. In the present study, it was revealed that the expression of E2F2 in ccRCC tissues is higher than that in normal tissues, and this expression is significantly related to the tumor stage and grade of patients with ccRCC. In addition, in all patients with ccRCC, high expression of E2F2 was significantly associated with a poor OS, DSS and PFS, which appeared to be consistent with the role of E2F2 as an oncogene. Notably, high expression of E2F3 in cancer tissues was observed in TCGA, but this phenomenon was not identified in the two GEO datasets or the clinical sample validation.

The present study showed that E2F4 may have carcinogenic effects and the high expression of E2F4 is related to poor prognostic factors of ccRCC, such as high TNM stage and high grade. Furthermore, it has been reported that the expression of E2F4 in prostate (22) and breast cancer (18) is higher than that in normal tissues. Although E2F4 is traditionally classified as a cell cycle inhibitor, its pro-proliferation and anti-apoptotic activities have been confirmed in various human cell lines. Notably, gene enrichment analysis in the present study revealed that the positive correlation between E2F4 and the cell cycle was not significant, and the possible relationship of E2F4 with the early stages of carcinogenesis needs to be clarified.

E2F5 has different expression levels in different types of tumors. Studies have reported that E2F5 is overexpressed in glioblastoma (23) and prostate cancer (24). However, E2F5 was downregulated in MCF7 human breast cancer cells, significantly impairing cell proliferation, migration and invasion in vitro and increasing cell cycle arrest in the G0/G1 phase (25). In the present study, it was proved that the expression of E2F5 in ccRCC tissues is lower than that in normal tissues. According to previous studies, the expression of E2F6
Figure 5. Association between E2F expression and the cell cycle. Gene set enrichment analysis showing cell cycle signature enrichment analysis in three datasets: REACTOME, KEGG and GO. KEGG, Kyoto Encyclopedia of Genes and genomes; GO, Gene Ontology.
in breast cancer tissues is lower than that in normal tissues, but this expression was not correlated with tumor stage in patients with breast cancer (18,26). In the present study, it was identified that the expression of E2F6 in ccRCC tissues was higher than that in normal tissues, and the expression level was significantly positively correlated with the copy number but not significantly correlated with the prognosis of patients with ccRCC. Studies have found that E2F7 and E2F8 are unique inhibitory genes that have a vital inhibitory effect on cell proliferation (27-29). However, the present study demonstrated that E2F7 and E2F8 are highly expressed in ccRCC tissues, suggesting that they may play different biological roles in different cell types.

The change of DNA methylation status is an important factor in tumorigenesis. This change includes the decrease of the overall methylation level of the genome and the abnormal increase of the local methylation level of CpG island, resulting in the instability of the genome and the low expression of tumor suppressor genes (30). It was revealed that E2F1, E2F2 and E2F8 expression may be increased due to hypomethylation in tumors, and E2F4, E2F6 and E2F7 expression may be increased due to an increased copy number in tumors. In addition, the present study showed that the expression of E2F family in ccRCC was significantly positively correlated with cell cycle.

In conclusion, except for the E2F3 and E2F5, all E2F family members are highly expressed in ccRCC tissues, and their expression is closely related to the pathological status and survival prognosis of patients with ccRCC, but different members of the family have differences in different tissue samples. The present findings suggested that E2F family genes may be potential targets for molecular diagnosis and targeted therapy of ccRCC, transcriptional E2F1-5, 7, and 8 were potential prognostic markers for the improvement of ccRCC survival and prognostic accuracy. Targeted therapy against single or combined E2F family proteins may improve the therapeutic effect and patient outcomes. It is expected that the findings of the present study will contribute to available knowledge, improve treatment designs, and enhance the accuracy of prognosis for patients with ccRCC.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Medical Research Project of Jiangsu Provincial Health Commission (grant no. Z2020074), the Six Talent Peaks Project in Jiangsu (grant no. WSN-342), the Jiangsu Provincial Medical Youth Talent Project (grant no. QNRC2016480), Suqian Guiding Science and Technology Project (grant no. Z2018174) and the Science and Technology Project of Suqian (grant no. S201720).

Availability of data and materials

The datasets generated and analyzed during the current study are available in TCGA (portal.gdc.cancer.gov) and GEO (ncbi.nlm.nih.gov/geo/).

Authors' contributions

ZL, YS and HG conceptualized and designed the research and performed bioinformatics analysis. YW and XYZ performed the experiments. JS, XY, XCZ, HL and XJY analyzed and interpreted the data. ZL and JS drafted and edited the manuscript. XY and XCZ supervised the project. ZL, YS and HG confirm the authenticity of all the raw data. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

The present study was approved (approval no. 2018-SL-0026) by the Ethics Committee of The Affiliated Suqian First People's Hospital of Nanjing Medical University (Suqian, China). Written informed consent was provided by all patients prior to the study.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.