Identification of methylation-driven genes related to prognosis in clear-cell renal cell carcinoma

Jia Wang1,2,3* | Qiujing Zhang1,3* | Qingqing Zhu1,3 | Chengxiang Liu4 | Xueli Nan5 | Fuxia Wang6 | Lihua Fang7 | Jie Liu3 | Chao Xie3 | Shuai Fu3 | Bao Song8

1School of Medicine and Life Sciences, University of Jinan-Shandong Academy of Medical Sciences, Jinan, China
2Department of Oncology, Zibo Maternal and Child Health Hospital, Zibo, China
3Department of Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China
4Department of Oncology, Jinan Jigang Hospital, Jinan, China
5Department of Oncology, Wu Di People Hospital, Binzhou, China
6Department of Oncology, YunCheng Country People’s Hospital, YunCheng, China
7Department of Oncology, Chang Qing District People’s Hospital, Jinan, China
8Basic Laboratory, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China

Correspondence
Bao Song, No. 440 Jiyan Road, Hualiyin, Jinan, Shandong 250117, China.
Email: sbwzzlj_sdch@163.com

Abstract
With the participation of the existing treatment methods, the prognosis of advanced clear-cell renal cell carcinoma (ccRCC) is poor. More evidence indicates the presence of methylation in ccRCC cancer cells, but there is a lack of studies on methylation-driven genes in ccRCC. We analyzed the open data of ccRCC in The Cancer Genome Atlas database to obtain ccRCC-related methylation-driven genes, and then carried out pathway enrichment, survival, and joint survival analyses. More important, we deeply explored the correlation between differential methylation sites and the expression of these driving genes. Finally, we screened 29 methylation-driven genes via MethylMix, of which six were significantly associated with the survival of ccRCC patients. This study demonstrated that the effect of hypermethylation or hypomethylation on prognosis is different, and the level of methylation of key methylation sites is associated with gene expression. We identified methylation-driven genes independently predicting prognosis in ccRCC, which offers theoretical support in bioinformatics for the study of methylation in ccRCC and a new perspective for the epigenetic study of ccRCC.

KEYWORDS
clear-cell renal cell carcinoma, methylation-driven genes, MethylMix, prognosis, the Cancer Genome Atlas (RRID:SCR_014514)

1 | BACKGROUND

Renal cell carcinoma ranks among the top 10 in the world in the diagnosis rate of both sexes (5% for men and 2% for women; Capitanio et al., 2019), and shows an increasing trend year by year. Clear-cell renal cell carcinoma (ccRCC) is the most common subtype (70–80%) (Greef & Eisen, 2016). About 25% of the patients have lost the opportunity of operating at the time of diagnosis (Lalani et al., 2019), and approximately only 10% of them survive 5 years later (van den Heuvel et al., 2019). And the group of patients who had an occasion of surgery had a 20% chance of postoperative recurrence (Motzer, 2016). Later, the discovery that the Von Hippel Lindau (VHL) gene is present in more than 90% of ccRCC (Nargund, Osmanbeyoglu, Cheng, & Hsieh, 2017) enriched the treatment of renal clear-cell carcinoma. However, the response rate of the existing treatment is only about 27% recently, and at least 75% of the patients who received traditional treatment make progress 2 years later (Motzer,

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. Journal of Cellular Physiology Published by Wiley Periodicals, Inc.
In spite of immunotherapy improving the overall response rate, the incidence of serious adverse events also increases (Rini et al., 2019). With the clinical efficacy of immunotherapy, the exploration of the mode of combined immunotherapy was followed, but at present, there were no reliable biomarkers for early screening and prognosis judgment of ccRCC, which limits the progress of treatment of renal clear-cell carcinoma. Although early studies have discovered that the prognostic markers of ccRCC may be mutated genes such as VHL (Gulati et al., 2014). The heterogeneity of ccRCC itself requires more effective biomarkers to evaluate the prognosis. For this reason, there is an urgent need to ascertain valuable molecular targets in the in-depth study of ccRCC.

Advances in research techniques have led us to a deeper understanding of the adverse diseases in our own professional field. For example, we recognize that the high expression of CD36 (cluster of differentiation 36) transcriptional group (Xu, Qu, Wang, Zhang, & Ye, 2019) and P21 (RAC1) activated kinase 1 (PAK1) protein (Qu, Liu, Bai, Xu, & Guo, 2019) is a sign of poor prognosis in ccRCC. However, BCL2 associated athanogene 1 and NOP56 ribonucleoprotein are different from CD36 transcriptional group and PAK1 protein, their high expression indicates a better prognosis (Giridhar et al., 2017). The role of gene methylation in the development of cancer has been gradually perceived and recognized. Researchers have successfully identified methylated genes or methylation-driven genes that forecast the prognosis of esophageal squamous cell carcinoma (ESCC; Roy et al., 2019), lung squamous cell carcinoma (Gao et al., 2019), lung adenocarcinoma (C. Su et al., 2019), and hepatocellular carcinoma (G.-X. Li et al., 2019). Moreover, the function of epigenetic abnormalities in ccRCC has been confirmed. Bioinformatics analysis of microarray data has broad prospects and clinical significance. Therefore, the era of clinical application of methylated and methylation-driven genes is bound to come. However, there is short of previous evidence for the exploration of methylation-driven genes in ccRCC.

The opening of a large database such as The Cancer Genome Atlas (TCGA; Tomczak, Czerwińska, & Wiznerowicz, 2015) makes it possible to meet the deeper and more accurate needs of scientists for disease exploration. It contains genetic data such as human-methylated genes needed by the researchers, as well as a variety of clinical prognosis information, supplemented by bioinformatics technology so that we continue to have a new understanding of the occurrence and development of cancer and promote the rapid development of the discipline field. Our team collected the opened big data of TCGA and analyzed the data using the MethylMix software package (Gevaert, 2015) developed by Gevaert et al. to filter out the differentially methylated genes connected with the prognosis of ccRCC. Then, the gene enrichment pathway was visualized, and the survival and joint survival analysis were carried out to ascertain the relationship between the methylation-driven genes and the survival of the patient. The purpose of this study was to evaluate the relationship between methylation-driven genes, gene loci, and messenger RNA (mRNA) data for the sake of understanding the cancer mechanism involved in methylation ulteriorly and providing valuable medical evidence for the treatment and prognosis of ccRCC.

## 2 | MATERIALS AND METHODS

### 2.1 | Data source and data methods

We downloaded the data needed for this study from the open website of TCGA (https://portal.gdc.cancer.gov/, RRID:SCR_014514) using the kidney renal clear-cell carcinoma (KIRC) cohort, including methylation, mRNA expression, and clinical information in patients with ccRCC. Methylation data of 485 specimens (160 normal samples and 325 cancer samples involved) and mRNA expression data of 611 samples (comprising 72 normal specimens and 539 ccRCC specimens) were obtained.

DNA methylation data in patients with ccRCC were assayed using the Illumina Methylation 450k Bead chip (Walker et al., 2015), a technical platform for a large-scale study of DNA methylation. And then, we exploit R-based LIMMA (https://bioconductor.org/packages/release/bioc/html/limma.html, RRID:SCR_010943) software package (Ritchie et al., 2015) and EdgeR (https://bioconductor.org/packages/release/bioc/html/edgeR.html, RRID:SCR_012802) package (Robinson, McCarthy, & Smyth, 2010) to standardize the downloaded data to acquire differentially methylated genes ($p = .05$, logFold Change ($\log FC$) = 0.5) and expressed genes (FC = 5, adjusted $p = .01$). LIMMA and EdgeR are R software packages used to analyze differential genes and differentially expressed data, respectively. The operation process is as follows: the data of ccRCC were introduced into R platform, and the differentially methylated genes and differentially expressed genes were obtained and mapped after LIMMA and EdgeR packages treatment, filtration, and standardization. It is worth noting that differentially methylated genes and differentially expressed genes are genes with different degrees of methylation and different degrees of expression in normal and tumor tissues, respectively. However, the methylation-driven genes are genes with different degree of methylation and expression between normal and tumor tissues. Therefore, the MethylMix algorithm was utilized for calculating the relationship between gene methylation level and gene expression advanced through R language with $|\log FC| \geq 0$, adjusted $p < .05$, and correlation coefficient (Cor) $< -0.3$ as screening conditions. MethylMix (https://bioconductor.riken.jp/packages/3.1/bioc/html/MethylMix.html) is an algorithm based on the $\beta$-mixed model to identify methylation states and compares them with the normal DNA methylation state. It implemented the differential methylation value, that is, the difference between the tumor methylation state and the normal methylation state, to identify disease-related methylation-driven genes. And the operation of MethylMix package requires the input of three specific data sets (the methylation data of the tumor samples, the methylation data of the normal tissue samples, and the corresponding gene expression data of the tumor samples). As a result, we submit the data to MethylMix according to the requirements of the algorithm, and finally screened out differentially methylation-driven genes for survival analysis. All the data we download is open to the TCGA platform, so we do not need the approval of the local ethics committee.
2.2 Path enrichment analysis of methylation-driven genes

To further comprehend the biological functions of these methylation-driven genes, we applied ConensusPathDB (http://cpdb.molgen.mpg.de/; Herwig, Hardt, Lienhard, & Kamburov, 2016) to visualize the analysis of pathway of methylation-driven genes, and took the cut-off value of $p$ value = 0.05 as the criterion. ConensusPathDB integrates the interaction networks of *Homo sapiens*, containing signal transduction, gene regulation, and drug-target interaction, and biochemical metabolism either, which currently integrates 32 public databases. It can interact with the genetic information we collected to avoid redundancy perfectly, so it has become one of the favorite tools for visual methylation-driven gene pathway enrichment analysis in the research process.

2.3 Survival analysis and joint survival analysis

To explore the effect and significance of methylation-driven genes on the prognosis of patients with ccRCC, we take the overall survival (OS) of

### TABLE 1 Twenty-nine methylation-driven genes in ccRCC

| Gene       | normal Mean | Tumor mean | log FC | adjust P | cor   |
|------------|-------------|------------|--------|----------|-------|
| NNMT       | 0.490801045 | 0.386640567 | −0.34415 | 1.78E−4 | −0.319852 |
| MOB3A      | 0.683159551 | 0.557690028 | −0.29276 | 2.89E−6 | −0.429547 |
| BCAM       | 0.412729523 | 0.515318674 | 0.320268 | 1.51E−8 | −0.478336 |
| PRR15L     | 0.401467303 | 0.519562987 | 0.372016 | 4.53E−5 | −0.470521 |
| C11orf21   | 0.660522453 | 0.533692924 | −0.3076 | 1.86E−3 | −0.340058 |
| HOXB6      | 0.624464597 | 0.866363149 | 0.472352 | 7.77E−5 | −0.57175 |
| ZNF300P1   | 0.226865565 | 0.341288606 | 0.589155 | 1.43E−5 | −0.436561 |
| HHLA2      | 0.410567575 | 0.304387165 | −0.43171 | 8.02E−5 | −0.494496 |
| BST2       | 0.535399548 | 0.37064583 | −0.35037 | 4.00E−5 | −0.426027 |
| NR0B2      | 0.578674355 | 0.686782026 | 0.247101 | 5.73E−5 | −0.431367 |
| MTHFR      | 0.298172025 | 0.182097665 | −0.71143 | 7.85E−5 | −0.3702 |
| ACSM5      | 0.570660517 | 0.451334871 | −0.33843 | 2.05E−4 | −0.477774 |
| AC009506.1 | 0.082717108 | 0.167930097 | 1.021603 | 1.13E−4 | −0.399622 |
| MAL        | 0.201505706 | 0.313709011 | 0.638606 | 3.03E−4 | −0.345857 |
| GRAP2      | 0.601094479 | 0.489241065 | −0.29705 | 1.16E−4 | −0.342031 |
| ZNF418     | 0.185958338 | 0.260853528 | 0.488261 | 2.18E−4 | −0.509882 |
| HOXA7      | 0.266592109 | 0.364415225 | 0.450949 | 6.13E−3 | −0.41498 |
| C1orf116   | 0.699315024 | 0.761773712 | 0.12342 | 5.73E−7 | −0.530969 |
| KRTCAP3    | 0.443374147 | 0.578722251 | 0.384346 | 4.80E−6 | −0.577152 |
| CD69       | 0.629105907 | 0.553558638 | −0.18457 | 2.19E−5 | −0.360412 |
| FMO2       | 0.516856494 | 0.438365926 | −0.23763 | 5.60E−2 | −0.48796 |
| EVI2A      | 0.883065324 | 0.809783928 | −0.12498 | 5.21E−8 | −0.485002 |
| ZNF888     | 0.776904328 | 0.67156532 | −0.21021 | 5.46E−14 | −0.408986 |
| RIPK4      | 0.316498368 | 0.334545342 | 0.008004 | 2.16E−13 | −0.481455 |
| UGT1A8     | 0.846779043 | 0.813971416 | −0.05701 | 4.47E−13 | −0.413703 |
| PCGEM1     | 0.827941383 | 0.72499483 | −0.19156 | 3.29E−7 | −0.330849 |
| ATRACTR3   | 0.328281826 | 0.369746258 | 0.171881 | 1.57E−6 | −0.431851 |
| FABP7      | 0.723408963 | 0.643630606 | −0.16858 | 1.29E−5 | −0.624072 |
| SXX1       | 0.678468784 | 0.683103154 | 0.009878 | 0.0004706 | −0.471696 |

Note: ACSM5, acyl-CoA synthetase medium chain family member 5; ATRACTR3, actin-related protein T3; BCAM, basal cell adhesion molecule; BST2, bone marrow stromal cell antigen 2; C11orf21, chromosome 11 open reading frame 21; C1orf116, chromosome 1 open reading frame 11; ccRCC, clear-cell renal cell carcinoma; CD69, cluster of differentiation 69; EVI2A, ecotropic viral integration site 2A; FABP7, fatty acid binding protein 7; FMO2, flavin containing monoxygenase 2; GRAP2, GRB2 related adaptor protein 2; HOXA7, HERV-H LTR-associating protein 2; HOXA5, homeobox A5; HOXB6, homeobox B6; KRTCAP3, keratinocyte associated protein 3; MAL, Mal T cell differentiation protein; MOB3A, MOB kinase activator 3A; MTHFR, methylenetetrahydrofolate reductase; NNMT, nicotinamide N-methyltransferase; NR0B2, nuclear receptor subfamily 0 Group B member 3; PCGEM1, prostate-specific transcript; PRR15L, proline-rich protein 15; RIPK4, receptor-interacting serine/threonine kinase protein 4; SSX1, SSX family member 1; UGT1A8, UDP-glucuronosyltransferase 1-8; ZNF300P1, zinc finger protein 300 pseudogene 1; ZNF418, zinc finger protein 418; ZNF888, zinc finger protein 888.
patients, that is, from the time when the patient is confirmed to have the disease and to the time of death (for any reason), as the survival time and use it for final prognosis evaluation, survival analysis, and joint survival analysis. Survival analysis was used to show the relationship between the hypermethylated/hypomethylated level of genes and overall survival, while combined survival analysis was used to analyze the effect of methylated level combined with gene expression on survival. Therefore, we used Kaplan-Meier curve (Schultz, Peterson, & Breslau, 2002) to analyze and evaluate the relationship between methylation-driven genes and survival rate, and verified this correlation by Log-rank test (Koletsi & Pandis, 2017) on the survival R software package platform (Singh & Mukhopadhyay, 2011), so as to screen the possibility of independent prognosis of methylation-driven genes. And the setting of $p < .05$ was viewed as statistically significant. Then, the survival R package was further used to obtain the joint survival curve in the joint survival analysis to analyze the relationship between gene methylation level and expression and prognosis of ccRCC. It is worth noting that, based on the downloaded clinical prognostic information and the information of the related sites of the methylation-driven genes, we combined the prognostic key genes obtained from the survival analysis and the combined analysis to make sure the correlation between the expression of methylation-driven genes and the methylated sites of key genes through the Kaplan-Meier curve drawn by the survival R package (set $p < .05$, $|\text{Cor}| > 0.5$ as the screening condition).

**FIGURE 1** Thermal map of methylation-driven genes associated with ccRCC. (The color change from green to red in the heat map illustrates the trend from low to high methylation. $|\log \text{FC}| \geq 0$, adjusted $p < .05$ and Cor $< -0.3$). ccRCC, clear-cell renal cell carcinoma; FC, fold change
3 | RESULTS

3.1 | Acquisition of methylation-driven genes

To study the methylation-driven genes associated in patients with ccRCC, we downloaded methylated data from 160 normal specimens and 325 methylated cancer samples, as well as mRNA expression data from 72 normal specimens and 539 ccRCC samples from the TCGA database. First of all, we used LIMMA software package and EdgeR package to extract abnormal methylation data and gene expression data, and obtained 105 differentially methylated genes and 257 differentially expressed genes, respectively. Second, the data were integrated and grouped, and 29 methylation-driven genes were screened via using MethylMix R package (Table 1). Afterward the connection between gene methylation and gene expression was visualized on the R platform (Figure 1). Six of the most representative of these genes (chromosome 11 open reading frame 21 [C11orf21], ecotropic viral integration site 2A [EVI2A], proline-rich protein 15 [PRR15L], receptor-interacting serine/threonine kinase protein 4 [RIPK4], SSX family member 1 [SSX1], and zinc finger protein 418 [ZNF418]) are shown in Figures 2 and 3 and the rest in Figures S1 and S2.

3.2 | Path enrichment analysis of methylation-driven genes

We utilized ConsensusPathDB online software for path enrichment analysis to further study the mechanism of methylation-driven genes in the occurrence and development of ccRCC. The results of pathway enrichment analysis showed that the methylation-driven genes were mainly enriched in the universal transcriptional pathway, RNA polymerase II transcription and gene expression, and were significantly related to the enriched events (Figure 4).

3.3 | Prognostic evaluation and survival analysis

The prognostic value of 29 methylation-driven genes was construed by survival R software package, and nine genes were detected to be independent prognostic indicators of ccRCC (Figure 5). Among them, the
methylation levels of EVI2A, C11orf21, SSX1, bone marrow stromal cell antigen 2 (BST2), and MOB kinase activator 3A (MOB3A) were significantly positively correlated with the OS of patients with ccRCC, while the other four genes (PRR15L, ZNF418, homeobox B6 [HOXB6], RIPK4) were inversely connected with survival. Besides, for the sake of studying the correlation between methylation-driven genes, gene expression and the prognosis of patients, we coalesced these data for analysis, with $p < .05$ as statistically significant; and found that the gene expression and methylation levels of 11 genes (SSX1, PRR15L, methylenetetrahydrofolate reductase [MTHFR], ZNF418, EVI2A, RIPK4, flavin containing monoxygenase 2 [FMO2], C11orf21, HERV-H LTR-associating protein 2 [HHLA2], Mal T cell differentiation protein [MAL], nicotinamide N-methyltransferase [NNMT]) were significantly correlated with the prognosis of patients (Figure 6). The hypermethylation/low expression levels of SSX1, EVI2A, C11orf21, and NNMT were significantly higher than those of the patients with hypomethylation/high expression levels. On the contrary, the high methylated/low expression of the other seven genes represented poor survival.

On the basis of the above, by taking the intersection of survival analysis and joint survival analysis, we locked six prognostic genes C11orf21, EVI2A, PRR15L, RIPK4, SSX1, and ZNF418 as independent prognostic factors or biomarkers, and explored the correlation between the expression of driving genes and the corresponding methylated sites, our team found that not all methylation sites were associated with the expression of driving genes. We gained sites significantly related to C11orf21 and EVI2A, PRR15L, RIPK4, SSX1, and ZNF418, respectively. The details are as follows: zero site in C11orf21, one site in EVI2A...
(cg23352695), three sites in PRR15L (cg03496533, cg18052778, cg15738800), one site in RIPK4 (cg05306310), one site in SSX1 (cg17158811), and five sites in ZNF418 (cg15060012, cg11998703, cg21444693, cg26356061, cg18673377; Table 2). We ascertained that there was no significant correlation between the expression of C11orf21 methylation-driven gene and the level of methylation at all its corresponding sites (Figure S3), while the other five genes were negatively related to the level of methylation at the corresponding sites (Figure 7).

4 | DISCUSSION

Nearly, 80% to 90% of renal cell carcinoma is clear-cell carcinoma of the kidney (Ljungberg et al., 2015), which is one of the deadliest kidney masses in the world, with an estimated 140,000 deaths from ccRCC each year (Y. Huang et al., 2019). There is growing evidence that intratumoral heterogeneity, prevalent in ccRCCs, may affect therapeutic effect to a large extent (Crusz et al., 2016). Studies have shown that the changes in gene expression caused by methylation are connected with the development and regulation of a variety of human malignant tumors. Otherwise, unlike genetic changes, the epigenetic changes reason from methylation can be reversed (Zhao et al., 2018). Hence, it has gradually attracted extensive attention of researchers because it is noninvasive for diagnosis and detection of cancer (Luttmer et al., 2016); but the early screening of ccRCC markers are lacking recently, coupled with the limited accuracy, so the new prognostic biomarkers are in sole need of further exploration. High-throughput sequencing technology gives us a higher ability to detect promising biomarkers correlated with the prognosis and treatment of ccRCC. In recent years, the view that epigenetic changes such as methylation are participated in the progression of ccRCC has been confirmed (Shenoy et al., 2015), and the relationship between gene methylation and ccRCC has been explored more and more. It has been found that the methylation of promoter regions of ion transport mechanism is affiliated with the risk of ccRCC (Deckers et al., 2017), respectively. The high/low methylation of DNA is associated with the invasiveness of the disease and can significantly affect the survival of patients (X. Su et al., 2017). Moreover, van Vlodrop et al. (2017) demonstrate that the four gene promoters composed of Gremlin 1, DAN Family BMP Antagonist, neuralized E3 ubiquitin protein ligase 1, Ladinin 1, and neurofilament heavy can be used as prognostic markers of ccRCC. However, there is no report on the study of abnormal methylation driving genes in ccRCC. And so far, only one methylation-driven genes, Cytoshein 1 Interacting protein, has been reported to have an effect on ccRCC in previous studies, and the paper concludes that hypermethylation of this driver gene may be a feature of good
prognosis in KIRC (Gevaert, Tibshirani, & Plevritis, 2015). Thus, it is pressed for searching new prognosis-related driving genes in ccRCC for more accurate individualized treatment and good prognosis evaluation.

In this study, we detected TCGA data to explore the relationship between methylation-driven genes and the prognosis of ccRCC to determine effective prognostic biomarkers. First, Limma and EdgeR packages were carried out to process the data, and then 29 methylation-driven genes were screened based on MethylMix algorithm. Then, we made use of pathway analysis to further understand their biological functions. The results revealed that methylated genes were mainly attached oneself to universal transcriptional pathway, RNA polymerase II transcription and gene expression (transcription). It was a hint that the interaction and possible synergistic effect of methylation-driven genes on the gene function level of renal clear-cell carcinoma cell proliferation. Then, nine genes (EVI2A, C11orf21, SSX1, BST2, MOB3A, PRR15L, ZNF418, HOXB6, and RIPK4) were won as independent prognostic indicators by using survival R software package. In addition, 11 genes (SSX1, PRR15L, MTHFR, ZNF418, EVI2A, RIPK4, FMO2, C11orf21, HHLA2, MAL, and NNMT) were discovered to be significantly correlated with the prognosis of patients ($p < .05$) in joined survival

**FIGURE 5** Kaplan-Meier survival curves of nine driving genes. BST2, bone marrow stromal cell antigen 2; C11orf21, chromosome 11 open reading frame 21; EVI2A, ecotropic viral integration site 2A; HOXB6, homeobox B6; MOB3A, MOB kinase activator 3A; PRR15L, proline-rich protein 15; RIPK4, receptor-interacting serine/threonine kinase protein 4; SSX1, SSX family member 1; ZNF418, zinc finger protein 418
analysis so as to ulteriorly study the relationship between the expression of methylation-driven genes and the prognosis of patients. In conclusion, we speculated that hypermethylation of methylation-driven genes affects the metabolic level of cells by inhibiting the expression of genes, which resulted in a decline in the survival of patients with ccRCC. Through the intersection of survival analysis and joint survival analysis, we identified six specific methylation-driven genes (PRR15L, C11orf21, ZNF418, EVI2A, RIPK4, and SSX1) to explore the relationship between their methylation levels and methylation sites.

Among the six genes, the high methylated level of PRR15L, ZNF418, and RIPK4 prognosticated low survival rate, and the form of high methylated/low expression represented poor prognosis. PRR15L is a protein-coding gene located on chromosome 17 and expressed in renal tissue. According to statistics, there are deficient studies on PRR15L gene at present, and this gene is reported for the first time as an index to evaluate the prognosis of ccRCC by us. ZNF418 gene encodes a zinc finger protein, which plays a negative regulatory role in mitogen-activated protein kinase signaling pathway (Y. Li et al., 2008). Its low expression is concerned with the occurrence and development and poor prognosis of gastric cancer (Hui et al., 2018), and can be used as a biomarker for the diagnosis of ESCC (Pu et al., 2017). The overexpression of RIPK4 in receptor interaction is involved in the growth of some tumor cells (X. Huang et al., 2013). And there are some studies that showed that RIPK4 can facilitate the differentiation of epidermis through phosphorylation and participate in the carcinogenesis of skin (Lee et al., 2017) and encouraged the occurrence of nasopharyngeal carcinoma (Gong, Luo, Yang, Jiang, & Liu, 2018). Furthermore, protein kinase RIPK4 promotes the invasiveness of bladder urothelial carcinoma (Liu et al., 2018) and pancreatic cancer (Qi et al., 2018) through nuclear factor-κB pathway and RAF1/MEK/ERK pathway, respectively. Equally important, it is closely correlated with bone metastasis of breast cancer (Zhang, He, & Zhang, 2019) and lymph node metastasis of cervical cancer (Azimohammadi et al., 2017). We made the conclusion that patients with high methylation level of PRR15L, ZNF418, and RIPK4 had a poor prognosis compared with patients with low methylation level of methylation cleared. Moreover, combined survival analysis further displayed that their high methylated/low expression was a feature of low survival rate.

FIGURE 6 Results of combined survival analysis of 11 genes. Regard $p < .05$ as statistically significant, C11orf21, chromosome 11 open reading frame 21; EVI2A, ecotropic viral integration site 2A; FMO2, flavin containing monoxygenase 2; HHLA2, HERV–H LTR-associating protein 2; MAL, Mal T cell differentiation protein; MTHFR, methylenetetrahydrofolate reductase; NNMT, nicotinamide N-methyltransferase; PRR15L, proline-rich protein 15; RIPK4, receptor-interacting serine/threonine kinase protein 4; SSX1, SSX family member 1; ZNF418, zinc finger protein 418.
The lower the methylation level of the other three genes EVI2A, C11orf21, and SSX1, the worse the survival, and the combined survival analysis uncovered that the patients with low methylated/high expression had worse prognosis than those with high methylated/low expression. EVI2A is a protein-coding gene that can be employed to diagnose diseases (Lo, Shen, Baumgarner, Cramer, & Losse, 2012). In addition, some studies have revealed that it is also a specific tumor suppressor factor of lymphocytes (X.-W. Li et al., 2014), and the upregulation of EVI2A gene expression may increase the malignant risk of malignant peripheral schwannoma (MPNST; Pasmant et al., 2011). In our study, patients with hypomethylation/high expression of EVI2A gene had a low survival rate, which was consistent with the study of MPNST. C11orf21 located on chromosome 11p15.5, in the meta-analysis of chronic lymphocytic leukemia (CLL), S. I. Berndt et al found that C11orf21 is related to the arises of CLL (Berndt et al., 2013), but it needs to be further verified. The expression of SSX1 gene is related to the poor prognosis of patients with colon cancer (Hilal, Novikov, Novikov, & Karaulov, 2017). Moreover, Panigrahi et al discovered that specific chromosome translocation t (X: 18; p11; q11) is a feature of synovial sarcoma.

### TABLE 2  Correlation between methylation sites and gene expression

| Gene symbol | Methylation site | Correlation | p Value   |
|-------------|------------------|-------------|-----------|
| EVI2A       | cg23352695       | −0.577      | 5.43E−32  |
| PRR15L      | cg03496533       | −0.673      | 9.05E−47  |
|             | cg18052778       | −0.664      | 2.95E−45  |
|             | cg15738800       | −0.501      | 2.74E−23  |
| RIPK4       | cg05306310       | −0.502      | 1.8E−23   |
| SSX1        | cg17158811       | −0.563      | 2.99E−30  |
| ZNF418      | cg15060012       | −0.62       | 4.48E−38  |
|             | cg11998703       | −0.603      | 1.69E−35  |
|             | cg21444693       | −0.577      | 5.26E−32  |
|             | cg26356061       | −0.556      | 2.21E−29  |
|             | cg18673377       | −0.534      | 8.9E−27   |

Note: EVI2A, ecotropic viral integration site 2A; PRR15L, proline-rich protein 15; RIPK4, receptor-interacting serine/threonine kinase protein 4; SSX1, SSX family member 1; ZNF418, zinc finger protein 418.

![Figure 7](image-url) The correlation between methylated sites and matched driving gene expression of EVI2A (a), PRR15L (b–d), RIPK4 (e), SSX1 (f) and ZNF418 (g–k), (p < .05, |Cor| > 0.5), EVI2A, ecotropic viral integration site 2A; PRR15L, proline-rich protein 15; RIPK4, receptor-interacting serine/threonine kinase protein 4; SSX1, SSX family member 1; ZNF418, zinc finger protein 418.
markers of ccRCC. It is interesting to note that the expression levels of PRR15L, ZNF418, and RIPK4, which can be used as prognostic methylation to analyze the data we got from TCGA, and the outcomes was 29 To sum up, we conducted R soft packages and MethylMix technology or less except PRR15L, whereas, their research in ccRCC is still in the clear in this study. And we believe that the study of methylation sites caused by hypermethylation. Otherwise, none of the methylation sites had a statistically positive or negative relationship with the C11orf21 driving gene, which may be due to the limitations of our research. Therefore, further experiments are needed to verify the related sites found in this study and to find the sites that are not clear in this study. And we believe that the study of methylation sites of driving genes will play an important role in our further study of the mechanism.

In a word, the development of bioinformatics provides more and more evidence for the important role of DNA methylation in tumor. The six effective genes obtained in this study have been studied more or less except PRR15L, whereas, their research in ccRCC is still in the blank period. Our study discussed the role of PRR15L, C11orf21, ZNF418, EVI2A, RIPK4, and SSX1 in ccRCC for the first time. We used R software to analyze the methylation-driven genes of ccRCC in TCGA data. After that, for the sake of verifying the relationship between the degree of gene methylation and the prognosis of patients, we utilized survival analysis. In addition, to further ensure the effect of its practical application, we united the level of gene methylation, gene expression and patient prognosis survival information for joint survival analysis, and finally carried out a more accurate site analysis. This study has found potential methylation-driven genes that can be used as biomarkers of ccRCC, and the exploration of related sites has further refined our research, giving us a new understanding of ccRCC from the perspective of bioinformatics. However, it is undeniable that these ccRCC-related differential driving genes need further experimental verification on the basis of a rigorous attitude.

5 | CONCLUSION

To sum up, we conducted R soft packages and MethylMix technology to analyze the data we got from TCGA, and the outcomes was 29 methylation-driven genes which have a close relevance with ccRCC. Afterwards, we explored the connection of driving gene methylation, gene expression and the prognosis of patients via survival analysis and joint survival analysis on the basis of available data, and identified the methylation-driven genes, EVI2A, C11orf21, SSX1, PRR15L, ZNF418, and RIPK4, which can be used as prognostic markers of ccRCC. It is interesting to note that the expression levels of the five methylation-driven genes except C11orf21 are negatively correlated with the degree of methylation of the sites. Therefore, the development and evolution of ccRCC may be due in part to the methylation of these driving genes, and the genes we obtained in this study can be used as prognostic markers for ccRCC. In conclusion, the results of our study provide new ideas for the diagnosis, therapy, and prognosis of ccRCC, so that the methylation-driven genes contribute to ccRCC discovered, and open the way for the future phase of methylation-driven genes into clinical application.

ACKNOWLEDGMENTS

We sincerely thank the Cancer Genome Atlas database for its data sharing.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS

J.W. and Q.J.Z. contributed to the design of study, J.W., Q.Q.Z. and C.X.L. performed data analysis, X.L.N. and F.X.W. collected research data, L.H.F. and J.L. provided analytical tools; Q.J.Z. and B.S. involved in drafting the manuscript, C.X. and S.F. revising the manuscript critically. All the authors were fully involved in the writing of the paper and gave final approval to the version we submitted.

DATA ACCESSIBILITY

The data that support the findings of this study are openly available in TCGA at https://portal.gdc.cancer.gov/ (RRID:SCR_014514)

ORCID

Bao Song http://orcid.org/0000-0003-0846-1621

REFERENCES

Azimohammadi, S., Azimohammadi, S., Safari, A., Kaghazian, M., Sadrkhanlo, M., Behnod, V., ... Seifoleslami, M. (2017). High-level expression of RIPK4 and EZH2 contributes to lymph node metastasis and predicts favorable prognosis in patients with cervical cancer. Oncology Research, 25(4), 495−501. https://doi.org/10.3727/096504016X14749735594687
Berndt, S. I., Skibola, C. F., Joseph, V., Camp, N. J., Nieters, A., Wang, Z., ... Slager, S. L. (2013). Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. Nature Genetics, 45(8), 868−876. https://doi.org/10.1038/ng.2652
Capitanio, U., Bensalah, K., Bex, A., Boorjian, S. A., Bray, F., Coleman, J., ... Russo, P. (2019). Epidemiology of renal cell carcinoma. European Urology, 75(1), 74−84. https://doi.org/10.1016/j.euro.2018.08.03
Cruz, S. M., Tang, Y. Z., Sarker, S.-J., Prevoow, W., Kiyani, I., Beltran, L., ... Gerlinger, M. (2016). Heterogeneous response and progression patterns reveal phenotypic heterogeneity of tyrosine kinase inhibitor response in metastatic renal cell carcinoma. BMC Medicine, 14(1), 185. https://doi.org/10.1186/s12916-016-0729-9
Deckers, I. A., van Engelond, M., van den Brandt, P. A., Van Neste, L., Soetekouw, P. M., Aarts, M. J., ... & Schouten, I. L. (2017). Promoter CpG island methylation in ion transport mechanisms and associated dietary intakes jointly influence the risk of clear-cell renal cell cancer. *International Journal of Epidemiology, 46*(2), 622–631. https://doi.org/10.1093/ije/dyw266

Gao, C., Zhuang, J., Zhou, C., Ma, K., Zhao, M., Liu, C., ... & Sun, C. (2019). Prognostic value of aberrantly expressed methylation gene profiles in lung squamous cell carcinoma: A study based on The Cancer Genome Atlas. *Journal of Cellular Physiology, 234*(5), 6519–6528. https://doi.org/10.1002/jcp.27389

Gevaert, O. (2015). Methyl Mix: An R package for identifying DNA methylation-driven genes. *Bioinformatics*, 31(11), 1839–1841. https://doi.org/10.1093/bioinformatics/btv020

Gevaert, O., Tibshirani, R., & Plevritis, S. K. (2015). Pancancer analysis of DNA methylation-driven genes using MethylMix. *Genome Biology*, 16, 17. https://doi.org/10.1186/s13059-014-0579-8

Giridhar, K. V., Sosa, C. P., Hillman, D. W., Sanhueza, C., Dalpiaz, C. L., Costello, B. A., ... Kohli, M. (2017). Whole blood mRNA expression-based prognosis of metastatic renal cell carcinoma. *International Journal of Molecular Sciences, 18*(11), 2326. https://doi.org/10.3390/ijms18112326

Gong, Y., Luo, X., Yang, J., Jiang, Q., & Liu, Z. (2018). RIPK4 promoted the tumorigenicity of nasopharyngeal carcinoma cells. *Biomedicine & Pharmacotherapy, 108*, 1–6. https://doi.org/10.1016/j.biopha.2018.08.147

Greef, B., & Eisen, T. (2016). Medical treatment of renal cancer: New horizons. *British Journal of Cancer, 115*(5), 505–516. https://doi.org/10.1038/bjc.2016.230

Gulati, S., Martinez, P., Aarts, M. J., Schouten, I. L., Soetekouw, P. M., ... & Schouten, I. L. (2017). Promoter CpG island methylation in ion transport mechanisms and associated dietary intakes jointly influence the risk of clear-cell renal cell cancer. *International Journal of Epidemiology, 46*(2), 622–631. https://doi.org/10.1093/ije/dyw266

Huang, Y., Wang, J., Jia, P., Li, X., Pei, G., Wang, C., ... & Pandis, N. (2017). Survival analysis, part 2: Kaplan-Meier method and the log-rank test. *American Journal of Orthodontics and Dentofacial Orthopedics*, 152(4), 569–571. https://doi.org/10.1016/j.ajo.2017.07.008

Lalani, A.-K. A., Mc Gregor, B. A., Albíges, L., Choueiri, T. K., Motzer, R., Powles, T., ... & Bex, A. (2019). Systemic treatment of metastatic clear renal cell carcinoma in 2018: Current paradigms, use of immunotherapy, and future directions. *European Urology, 75*(1), 100–110. https://doi.org/10.1016/j.euro.2018.10.010

Lee, P., Jiang, S., Li, Y., Yue, J., Gou, X., Chen, S.-Y., ... & Wu, X. (2017). Phosphorylation of Pkp1 by RIPK4 regulates epithelial differentiation and skin tumorigenesis. *The EMBO Journal, 36*(13), 1963–1980. https://doi.org/10.15252/embj.201695679

Li, G.-X., Ding, Z.-Y., Wang, Y.-W., Liu, T.-T., Chen, W.-X., Wu, J.-J., ... & Zhang, B.-X. (2019). Integrative analysis of DNA methylation and gene expression identify a six epigenetic driver signature for predicting prognosis in hepatocellular carcinoma. *Journal of Cellular Physiology, 234*(7), 11942–11950. https://doi.org/10.1002/jcp.27882

Li, X.-W., Rees, J. S., Xue, P., Zhang, H., Hamala, S. W., Sanderson, B., ... & Jackson, A. P. (2014). New insights into the DT40 B cell receptor cluster using a proteomic proximity labeling assay. *The Journal of Biological Chemistry, 289*(21), 14434–14447. https://doi.org/10.1074/jbc.M113529578

Li, Y., Yang, D., Bai, Y., Mo, X., Huang, W., Yuan, W., ... & Wu, X. (2008). ZNF418, a novel human KRAB/C2H2 zinc finger protein, suppresses MAPK signaling pathway. *Molecular and Cellular Biochemistry, 310*(1–2), 141–151. https://doi.org/10.1007/s11007-006-9764-4

Liu, J.-Y., Zeng, Q.-H., Cao, P.-G., Xie, D., Chen, X., Yang, F., ... & Cao, K. (2018). RIPK4 promotes bladder urothelial carcinoma cell aggressiveness by upregulating VEGF-A through the NF-κB pathway. *British Journal of Cancer, 118*(12), 1617–1627. https://doi.org/10.1038/s41416-018-0116-8

Ljungberg, B., Bensalah, K., Canfield, S., Dabestani, S., Hofmann, F., Hora, M., ... & Bex, A. (2015). EUA guidelines on renal cell carcinoma: 2014 update. *European Urology, 67*(5), 913–924. https://doi.org/10.1016/j.euro.2015.01.005

Lo, C.-L., Shen, F., Baumgarner, K., Cramer, M. J., & Lossie, A. C. (2012). Identification of 12956/SvEvTac-specific polymorphisms on mouse chromosome 11. *DNA and Cell Biology, 31*(3), 402–414. https://doi.org/10.1089/dna.2011.1353

Luttmer, R., De Strooper, L. M. A., Dijkstra, M. G., Berkhoff, J., Snijders, P. J. F., Steenbergen, R. D. M., ... & Meijer, C. J. L. M. (2016). FAM194A methylation analysis in self-samples compared with cervical scrapes for detecting cervical (pre) cancer in HPV-positive women. *British Journal of Cancer, 115*(5), 579–587. https://doi.org/10.1038/bjc.2016.200

Motzer, R. J. (2016). Perspective: What next for treatment? *Nature, 537*(7620), 511–511. https://doi.org/10.1038/5375111a

Nargund, A. M., Osmanbeyoglu, H. U., Cheng, E. H., & Hsieh, J. J. (2017). SWI/SNF tumor suppressor gene PBRM1/BAF180 in human clear cell kidney cancer. *Molecular & Cellular Oncology, 4*(4), https://doi.org/10.1080/23723556.2017.1342747

Panigrahi, M. K., Pradhan, G., Sahoo, N., Mishra, P., Patra, S., & Mohapatra, P. R. (2018). Primary pulmonary synovial sarcoma: A reappraisal. *Journal of Cancer Research and Therapeutics, 14*(3), 481–489. https://doi.org/10.4103/0973-1482.204883

Pasmant, E., Masliah-Planchon, J., Lévy, P., Laurendeau, I., Ortonne, N., Parfait, B., ... & Bieche, I. (2011). Identification of genes potentially involved in the increased risk of malignancy in NF1-microdepleted patients. *Molecular Medicine, 17*(1–2), 79–87. https://doi.org/10.2119/molmed.2010.00079

Pu, W., Wang, C., Chen, S., Zhao, D., Zhou, Y., Ma, Y., ... & Wang, M. (2017). Targeted bisulfite sequencing identified a panel of DNA methylation-based biomarkers for esophageal squamous cell carcinoma (ESCC). *Clinical Epigenetics, 9*, 129. https://doi.org/10.1186/s13377-017-0430-7

Qi, Z.-H., Hu, X.-X., Zhang, S.-R., Xu, J.-Z., Li, S., Gao, H.-L., ... & Liu, L. (2018). RIPK4/PEBP1 axis promotes pancreatic cancer cell migration and invasion by activating RAF1/MEK/ERK signaling. *International Journal of Oncology, 52*(4), 1105–1116. https://doi.org/10.3892/ijo.2018.4269

Qu, Y., Liu, L., Bai, Q., Xu, J., & Guo, J. (2019). Tumor-infiltrating regulatory T lymphocytes orchestrate oncogenic PAK1-confined immune evasion in clear-cell renal carcinoma. *European Urology Supplements, 18*, e86–e87. https://doi.org/10.1016/j.eursups.2018.09.006

Rini, B. I., Pilmack, E. R., Stus, V., Gafanov, R., Hawkins, R., Nosov, D., ... & KEYNOTE-426 Investigators (2019). Pembrolizumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *The New England Journal of Medicine, 380*(12), 1116–1127. https://doi.org/10.1056/NEJMoa1816714

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., ... & Smyth, G. K. (2015). limma powers differential expression analyses for
RNA-sequencing and microarray studies. Nucleic Acids Research, 43(7), e47. https://doi.org/10.1093/nar/gkv007

Robinson, M. D., Mc Carthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Roy, R., Kandimalla, R., Sonohara, F., Koike, M., Kodera, Y., Takahashi, N., ... Goel, A. (2019). A comprehensive methylation signature identifies lymph node metastasis in esophageal squamous cell carcinoma. International Journal of Cancer, 144(5), 1160–1169. https://doi.org/10.1002/ijc.31755

Schultz, L. R., Peterson, E. L., & Breslau, N. (2002). Graphing survival curve estimates for time-dependent covariates. International Journal of Methods in Psychiatric Research, 11(2), 68–74

Shenoy, N., Vallumsetti, N., Zou, Y., Galeas, J. N., Shrivastava, M., Hu, C., ... Verma, A. (2015). Role of DNA methylation in renal cell carcinoma. Journal of Hematology & Oncology, 8, 88. https://doi.org/10.1186/s13045-015-0180-y

Singh, R., & Mukhopadhyay, K. (2011). Survival analysis in clinical trials: Basics and must know areas. Perspectives in Clinical Research, 2(4), 145–148. https://doi.org/10.4103/2229-3485.86872

Su, C., Shi, K., Cheng, X., Han, Y., Li, Y., Yu, D., & Liu, Z. (2019). Methylation of CLEC14A is associated with its expression and lung adenocarcinoma progression. Journal of Cellular Physiology, 234(3), 2954–2962. https://doi.org/10.1002/jcp.27112

Su, X., Zhang, J., Mouawad, R., Compérat, E., Roupérè, M., Allanic, F., ... Malouf, G. G. (2017). NSD1 inactivation and SETD2 mutation drive a convergence toward loss of function of H3K36 writers in clear cell renal cell carcinomas. Cancer Research, 77(18), 4835–4845. https://doi.org/10.1158/0008-5472.CAN-17-0143

Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. Contemporary Oncology, 19(1A), A68–A77. https://doi.org/10.5114/co.2014.47136

van den Heuvel, C. N. A. M., van Ewijk, A., Zeelen, C., de Bitter, T., Huygen, M., Mulders, P., ... Leenders, W. P. J. (2019). Molecular profiling of druggable targets in clear cell renal cell carcinoma through targeted RNA sequencing. Frontiers in Oncology, 9, 117. https://doi.org/10.3389/fonc.2019.00117

van Vlodrop, I. J. H., Joosten, S. C., De Meyer, T., Smits, K. M., Van Neste, L., Melotte, V., ... van Engeland, M. (2017). A four-gene promoter methylation marker panel consisting of GREM1, NEURL, LAD1, and NEFH predicts survival of clear cell renal cell cancer patients. Clinical Cancer Research, 23(8), 2006–2018. https://doi.org/10.1158/1078-0432.CCR-16-1236

Walker, D. L., Bhagwate, A. V., Baheti, S., Smalley, R. L., Hilker, C. A., Sun, Z., ... Cunningham, J. M. (2015). DNA methylation profiling: Comparison of genome-wide sequencing methods and the infinium human methylation 450 bead chip. Epigenomics, 7(8), 1287–1302. https://doi.org/10.2217/EPI.15.64

Xu, W.-H., Qu, Y.-Y., Wang, J., Zhang, H.-L., & Ye, D. (2019). Elevated CD36 expression correlates with increased visceral adipose tissue and predicts poor prognosis in ccRCC patients. Journal of Clinical Oncology, 37(7_suppl), https://doi.org/10.1200/JCO.2019.37.7_suppl.571. 571–571

Zhang, Y., He, W., & Zhang, S. (2019). Seeking for correlative genes and signaling pathways with bone metastasis from breast cancer by integrated analysis. Frontiers in Oncology, 9, 138. https://doi.org/10.3389/fonc.2019.00138

Zhao, L., Duan, Y.-T., Lu, P., Zhang, Z.-J., Zheng, X.-K., Wang, J.-L., & Feng, W.-S. (2018). Epigenetic targets and their inhibitors in cancer therapy. Current Topics in Medicinal Chemistry, 18(28), 2395–2419. https://doi.org/10.2174/1568026619666181224095449.

https://portal.gdc.cancer.gov/

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Wang J, Zhang Q, Zhu Q, et al. Identification of methylation-driven genes related to prognosis in clear-cell renal cell carcinoma. J Cell Physiol. 2020:235:1296–1308. https://doi.org/10.1002/jcp.29046