



DNA Methylation-Based Panel Predicts Survival of Patients With Clear Cell Renal Cell Carcinoma and Its Correlations With Genomic Metrics and Tumor Immune Cell Infiltration

Xiao-Ping Liu¹, Lingao Ju^{2,3}, Chen Chen^{2,3}, Tongzu Liu¹, Sheng Li^{2,3*} and Xinghuan Wang^{1*}

OPEN ACCESS

Edited by:

Massimiliano Berretta,
Centro di Riferimento Oncologico di
Aviano (IRCCS), Italy

Reviewed by:

Youlang Zhou,
Affiliated Hospital of Nantong
University, China
Anup Kumar Singh,
Department of Cancer Genetics
and Epigenetics, Beckman Research
Institute, City of Hope, United States

*Correspondence:

Sheng Li
lisheng-znyy@whu.edu.cn
Xinghuan Wang
wangxinghuan@whu.edu.cn

Specialty section:

This article was submitted to
Molecular and Cellular Oncology,
a section of the journal
*Frontiers in Cell and Developmental
Biology*

Received: 15 June 2020

Accepted: 28 September 2020

Published: 15 October 2020

Citation:

Liu X-P, Ju L, Chen C, Liu T, Li S
and Wang X (2020) DNA
Methylation-Based Panel Predicts
Survival of Patients With Clear Cell
Renal Cell Carcinoma and Its
Correlations With Genomic Metrics
and Tumor Immune Cell Infiltration.
Front. Cell Dev. Biol. 8:572628.
doi: 10.3389/fcell.2020.572628

¹ Department of Urology, Zhongnan Hospital of Wuhan University, Wuhan, China, ² Department of Biological Repositories, Zhongnan Hospital of Wuhan University, Wuhan, China, ³ Human Genetics Resource Preservation Center of Wuhan University, Wuhan, China

DNA methylation based prognostic factor for patients with clear cell renal cell carcinoma (ccRCC) remains unclear. In the present study, we identified survival-related DNA methylation sites based on the differentially methylated DNA CpG sites between normal renal tissue and ccRCC. Then, these survival-related DNA methylation sites were included into an elastic net regularized Cox proportional hazards regression (CoxPH) model to build a DNA methylation-based panel, which could stratify patients into different survival groups with excellent accuracies in the training set and test set. External validation suggested that the DNA methylation-based panel could effectively distinguish normal controls from tumor samples and classify patients into metastasis group and non-metastasis group. The nomogram containing DNA methylation-based panel was reliable in clinical settings. Higher total mutation number, SCNA level, and MATH score were associated with higher methylation risk. The innate immune, ratio between CD8⁺T cell versus Treg cell as well as Th17 cell versus Th2 cell were significantly decreased in high methylation risk group. In conclusion, we developed a DNA methylation-based panel which might be independent prognostic factor in ccRCC. Patients with higher methylation risk were associated genomic alteration and poor immune microenvironment.

Keywords: clear cell renal cell carcinoma, methylation CpG island, prognostication panel, elastic net, survival analysis

INTRODUCTION

Renal cancer, also known as renal cell carcinoma (RCC), renal adenocarcinoma, etc., accounts for about 85% of primary renal malignancies (Capitanio and Montorsi, 2016; Morris and Latif, 2017). Nearly 70–80% of RCCs are defined as clear cell renal cell carcinoma (ccRCC) (Capitanio and Montorsi, 2016). The cause of kidney cancer has not yet been clarified, and its onset may

be related to smoking, obesity, occupational exposure (such as asbestos, leather, etc.), and genetic factors (such as the absence of tumor suppressor genes) (Capitanio and Montorsi, 2016). About 30–50% of kidney cancer lacks early clinical manifestations, while nearly two-thirds of ccRCCs are detected incidentally (Gonzalez Leon and Morera Perez, 2016; Hancock and Georgiades, 2016). It has been reported that nearly 30–50% of ccRCCs will develop metastases even after undergoing radical surgery during follow-up (Fidler, 1992; Bolger, 2002; Singh et al., 2017).

Tumor stage, nuclear grade, morphology, and histology determined at the time of nephrectomy have traditionally been applied to prognosticate the progression and survival of patients with ccRCC (Golovastova et al., 2017). However, recent studies suggested that molecular heterogeneity was involved in the carcinogenesis and development of ccRCC tumors, leading to variation in cellular proliferation, metabolic activity and tumor microenvironment of ccRCC (Borys et al., 2019; Liss et al., 2019). Thus, the above traditionally clinicopathological features might be insufficient for precise tumor characterization and classification and survival prognostication (Moch et al., 2014; Zheng et al., 2018).

Currently, thanks to high-throughput sequencing technology, a variety of novel biomarkers have been introduced in clinical settings (van Vlodrop et al., 2017; Liu et al., 2018; Wu et al., 2019). Epigenetic alterations have been demonstrated to serve as key contributors to the regulation of gene expression and involved in various kinds of malignancies as well as the clinicopathology of other medical conditions (Morris and Latif, 2017; Young et al., 2017; Evelonn et al., 2019). Global DNA hypomethylation and regional hypermethylation of Cytosine-Phosphate-Guanine (CpG)-rich islands are identified in multiple kinds of malignancies, and are associated with gene expression and interact with various positive and negative feedback mechanisms (de Cubas and Rathmell, 2018). Therefore, aberrant methylation CpG sites could be selected as novel prognostication markers in human cancers including ccRCC.

In the present study, we identified a combination of 11 methylation sites associated with the survival, genomic alteration, and tumor microenvironment of patients with ccRCC, and suggested that it could be an independent prognostic factor in patients with ccRCC.

MATERIALS AND METHODS

Patient Cohort

A total of 483 renal samples in the TCGA Kidney Clear Cell Carcinoma (TCGA-KIRC) cohort obtained at the time of nephrectomy were subjected to the Illumina Infinium 450k Human DNA methylation Beadchip to obtain genome wide DNA methylation status in approximately 485,000 CpGs in renal carcinoma and adjacent normal kidney tissue samples from ccRCC patients (Cancer Genome Atlas Research Network, 2013). We obtained the beta value matrix of the TCGA-KIRC methylation profile from UCSC Xena¹. Meanwhile, we

obtained associated copy number data, mRNA expression data, somatic mutation data and clinical data of TCGA-KIRC from GDC data portal². GSE113501, measured by Illumina HumanMethylation450 BeadChip and included 115 ccRCC patients and the associated methylation profile, were downloaded for Gene Expression Omnibus (GEO)³ (Evelonn et al., 2019). We used GSE113501 as an independent validation set to verify the relationship between methylation status and disease metastasis. GSE61441, measured by Illumina HumanMethylation450 BeadChip, included 46 ccRCC tissue and 46 matched normal kidney tissues, were used to validate the relationship between the methylation status and ccRCC and normal kidney tissue (Wei et al., 2015).

Differential Methylation Analysis

R/Bioconductor package ChAMP was used to identify differentially methylated CpG sites (Morris et al., 2014). Before performing differential methylation analysis, probes met the following criteria were filtered: (1) non-CpG probes; (2) probes with SNPs; (3) probes aligning to multiple location; (4) probes located on the X, Y chromosome. Then, CpG sites with adjusted $P < 0.05$ and $\text{cut.deltaBeta} > 0.3$ were considered to be significantly differentially methylated.

Identification of Prognosis-Related CpG Sites and Development of DNA Methylation-Based Panel

The TCGA-KIRC cohort was randomly partitioned into two analyzing groups - training set and test set, in a ratio of 3:2. On the basis of the above differentially methylation CpG sites, we tried to identify prognosis-related CpG sites that were significantly (sites with Bonferroni adjusted P less than 0.01) associated with the overall survival (OS) of patients with ccRCC using univariate Cox proportional hazards regression model (CoxPH). Then, the prognosis-related CpG sites were included in an elastic net regularized CoxPH model (Zou and Hastie, 2005). The two hyperparameters α and λ were tuned through 10-fold cross-validation using the R package “glmnet” (Simon et al., 2011) and “c060” (Sill et al., 2014). CpG sites with non-zero coefficients in the elastic net regularized CoxPH model were integrated to build a lineal risk score based on the methylation levels of the CpG sites and their corresponding coefficients in the model.

Characterizing the Prognostication Ability of the DNA Methylation-Based Panel

Univariate and multivariable CoxPH model were performed to analyze the relation between the DNA methylation-based panel and the OS of patients with ccRCC in clinical settings. For the CoxPH model, the DNA methylation-based panel, age, and tumor stage were continuous variables, while hemoglobin, serum calcium, and gender were taken as categorized variable. Time-dependent receiver operating characteristic curve (ROC)

¹<https://xenabrowser.net/>

²<https://portal.gdc.cancer.gov>

³<https://www.ncbi.nlm.nih.gov/geo/>

analysis were performed using the R package “survivalROC” (Kamarudin et al., 2017), and the 1-, 3-, 5-, 7-, and 10-year Area under the ROC Curves (AUCs) were visualized using the R packages “ggplot2.” Moreover, the prediction ability of the DNA methylation-based panel was also validated in two independent validation cohorts (GSE61441, Wei et al., 2015; and GSE113501, Evelonn et al., 2019). The risk scores of patients in the two cohorts were calculated based on the methylation levels of the CpG sites in the DNA methylation-based panel and their responding coefficients in the elastic net penalized CoxPH model.

Clinical Application of the DNA Methylation-Based Panel

To determine the clinical utility of the methylation-based panel, the DNA methylation-based panel and other clinical variables (age, hemoglobin level, serum calcium level, tumor stage, and gender) were incorporated into a multivariable survival model, and then developed a nomogram incorporating the DNA methylation-based panel. The nomogram was also calibrated at 3-, and 5- years. Moreover, decision curve analysis (DCA) (Vickers and Elkin, 2006) was performed to determine the clinical benefit of the DNA methylation containing panel. Finally, we compared the prognostic performance of the DNA methylation-based panel with those of several currently available biomarkers in the whole TCGA-KIRC cohort using C-index.

Correlation Analysis Between the CpG Site Methylation and the Corresponding Gene Expression

The fragments per kilobase of exon model per million reads mapped (FPKM) normalized gene expression profile of the TCGA-KIRC cohort was used to identify the relationship between the CpG site methylation status and the expression of genes regulated by the corresponding sites. To this end, Spearman’s correlation analysis was performed and visualized using the R package “ggpubr.” Meanwhile, the role of specific gene on the OS of patients with ccRCC was also characterized.

Analyzing Associations Between the DNA Methylation-Based Panel and Genomic Metrics of Patients With ccRCC

To investigate the mechanisms that the DNA methylation-based panel affected the survival of ccRCC patients, we performed Spearman’s rank correlation analysis to characterize the relationship between the methylation risk of ccRCC patients and genomic metrics of ccRCC patients including total numbers of mutations, clonal heterogeneity measured by the mutant-allele tumor heterogeneity (MATH) (Mroz and Rocco, 2013), and somatic copy number alteration (SCNAs) (Mroz and Rocco, 2013). R package “maftool” (Mayakonda et al., 2018) was used to calculate the total mutation number and MATH of each ccRCC on the basis of VarScan2 preprocessed mutation data. The SCNA levels of ccRCC patients in the TCGA-KIRC cohort were calculated as previously introduced (Davoli et al., 2017).

Analyzing the Associations Between the DNA Methylation-Based Panel and Tumor Microenvironment (TME) of Patients With ccRCC

Next, we tried to study the relationship between the methylation status of the multi-CpG site and tumor microenvironment. We used gene signatures of 24 types of immune cells introduced by Bindea et al. (2013) to calculate the specific immune infiltration using the single cell gene set enrichment analysis (ssGSEA) method implemented in the R package GSVA (Hanzelmann et al., 2013). As introduced by Senbabaoglu et al. (2016), we used the mean of the standardized values for CD8⁺ T cell, T helper cell, T cell, T central and effector memory cell, Th1 cell, Th2 cell, Th17 cell, and Treg cells as T cell infiltration score (TIS), while the mean standardized value of the 24 types of immune cells was treated as overall immune infiltration score (OIIS). Moreover, we analyzed the correlation between the methylation risk and four inhibitory checkpoint molecules (PDCD1, PDCD1LG2, CD274, and CTLA-4). Spearman’s rank correlation analysis was used to characterize the correlations between specific immune cell composition (or checkpoint molecules) and the methylation risk of the multi-CpG site.

Detection of the Methylation Levels of the Multi-CpG Sites in Clinical Settings

ccRCC samples and corresponding normal renal tissues were collected from patients who received surgery during hospitalization at department of urology, Zhongnan hospital of Wuhan university, which was approved by the Ethics Committee at Zhongnan Hospital of Wuhan University, the informed consent was available from all participants as well. The procedure of bisulfite sequencing PCR (BSP) and methylation specific PCR (MSP) and data analysis were the same with our previous study (Ge et al., 2019).

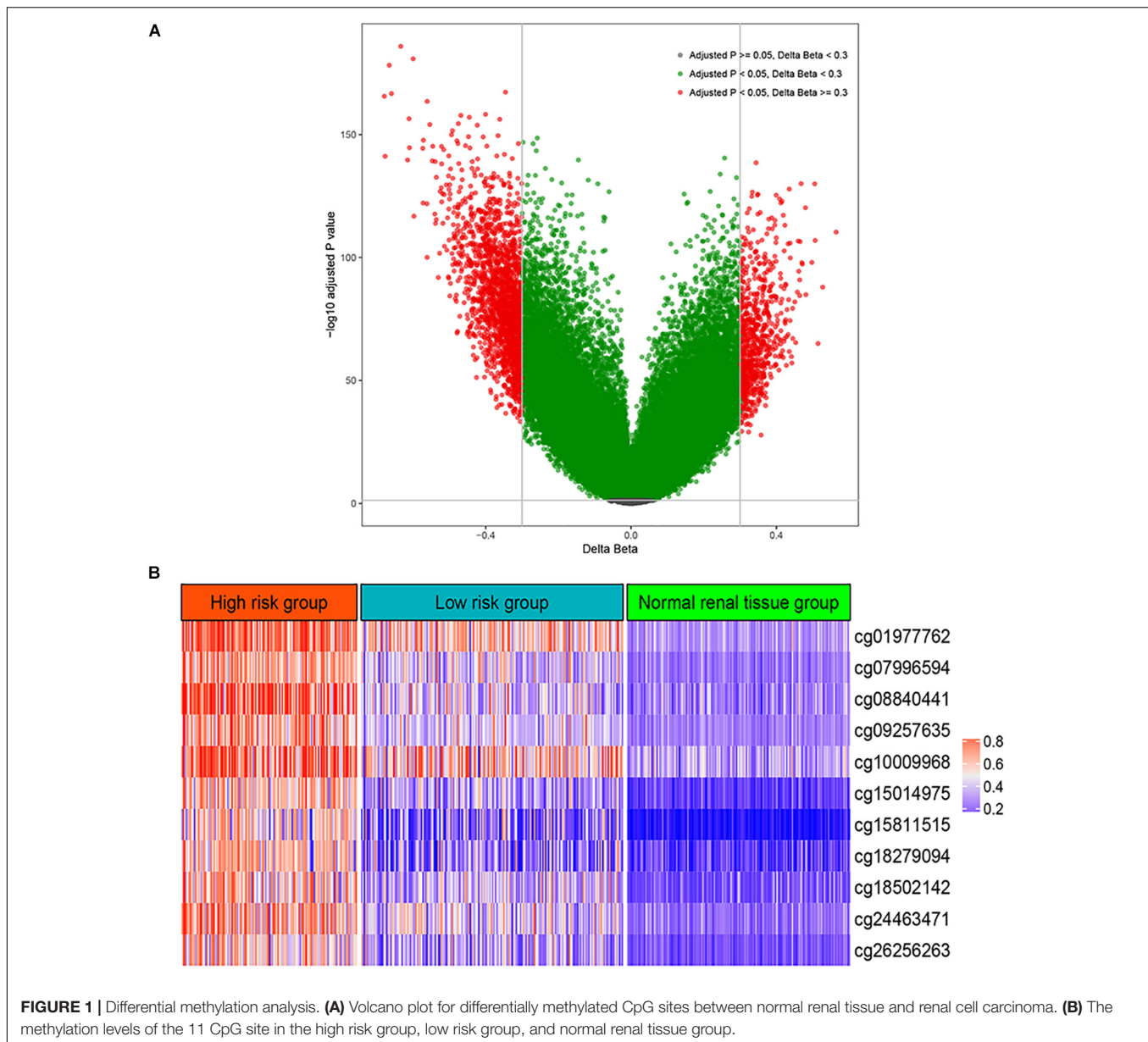
RESULTS

Differentially Methylated CpG Sites Between ccRCC and Adjacent Normal Renal Tissue

A total of 483 renal tissues were subjected to DNA methylation profiling, including 160 pairs of ccRCC sample and adjacent normal renal samples. Thus, we calculated differentially methylated CpG sites based on the paired tumors and normal samples. As shown in **Figure 1A**, a total of 2,628 CpG sites were found to be differentially methylated between ccRCC and normal renal tissue at adjusted $P < 0.05$ and $\text{cut.deltaBeta} > 0.3$.

Development of a Combination of Multi-CpG Site (DNA Methylation-Based Panel) Relating With the Survivals of Patients With ccRCC

Firstly, a total of 56 CpG sites were shown to be significantly (Bonferroni adjusted $P < 0.01$) correlated with the overall survival (OS) of patients with ccRCC using univariate CoxPH

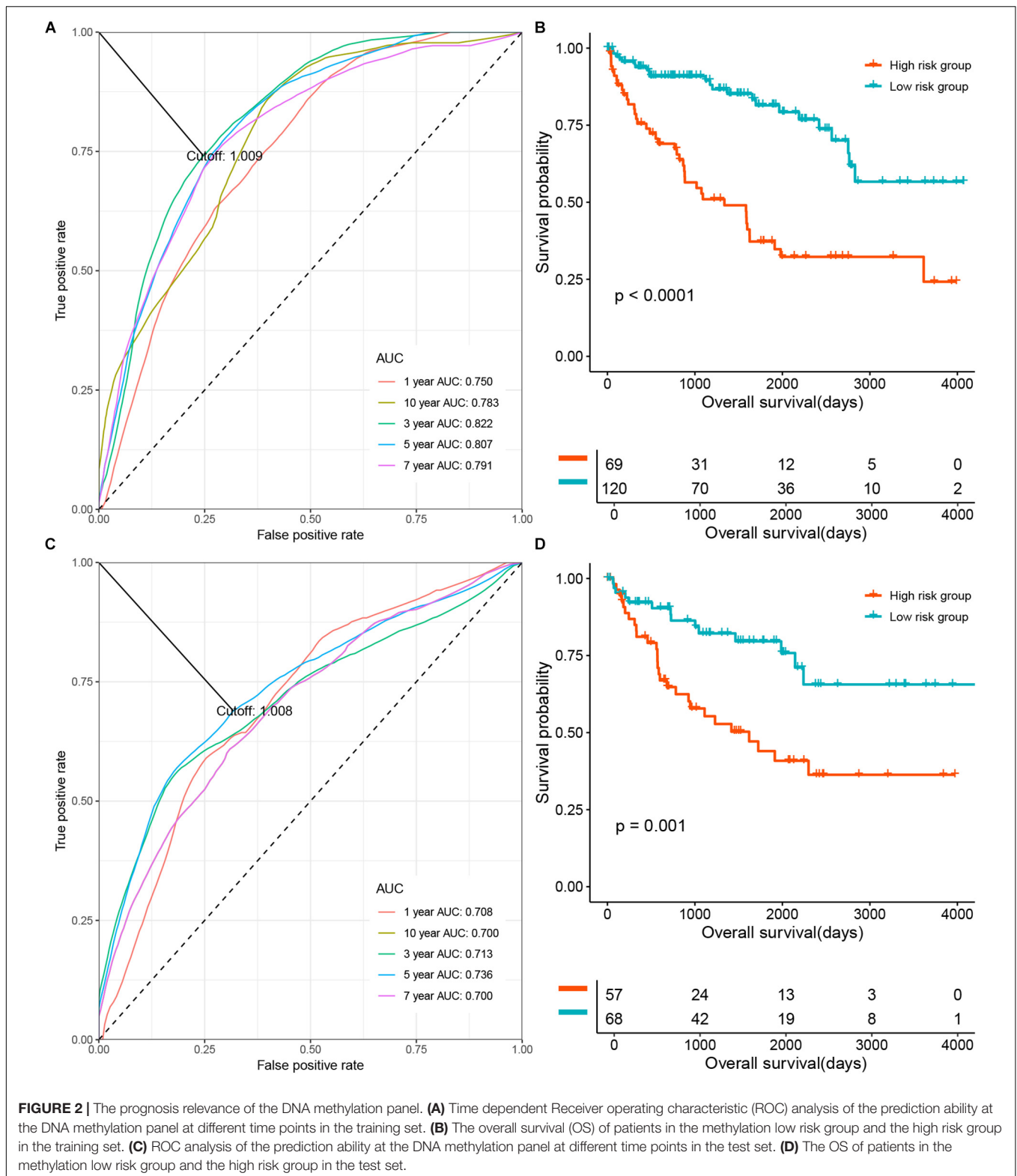


model (**Supplementary Table S1**). Then, these 56 CpG sites were included into an elastic net penalized CoxPH model. After 10-fold cross-validation, optimal α (0.0129) and λ (10.6629) were utilized to fit a final elastic net penalized CoxPH model, resulting in 11 CpG sites whose coefficients were not zero (**Supplementary Figure S1** and **Supplementary Table S1**). Therefore, we built a DNA methylation-based panel based on the coefficients and methylation levels of these 11 CpG sites (**Figure 1B** and **Supplementary Figure S2**).

Prognostication Relevance of the DNA Methylation-Based Panel

In order to determine the prognostication relevance of the DNA methylation-based panel, patients were categorized into two

risk groups according to the optimal cutoff of the risk score calculated based on the coefficients and methylation levels of these 11 CpG sites. Then, we performed survival analysis on the OS of ccRCC patients. As shown in **Figure 2A**, in the training set, time-dependent receiver operating characteristic (ROC) analysis suggested that the area under the ROC curves (AUCs) at 1-, 3-, 5-, 7-, and 10-year were 0.750, 0.783, 0.822, 0.807, and 0.791, respectively. Kaplan-Meier (KM) curve suggested patients in the low-risk group demonstrated significantly longer OS compared with those in the high-risk group ($P < 0.0001$) (**Figure 2B**). Meanwhile, time-dependent ROC analysis also suggested that the DNA methylation-based panel showed high sensitivity and specificity in predicting the OS of ccRCC patients in the test set, with 1-, 3-, 5-, 7- and 10-year AUCs of 0.708, 0.700, 0.713, 0.736,



and 0.700, respectively (Figure 2C). KM survival analysis suggested that the DNA methylation-based panel could classified patients into significant risk groups ($P = 0.001$) (Figure 2D).

Moreover, multivariable CoxPH model suggested the DNA methylation-based panel was an independent prognostication factor after adjusting other clinical variables including age,

hemoglobin level, serum calcium level, tumor stage, and gender (Supplementary Tables S2, S3).

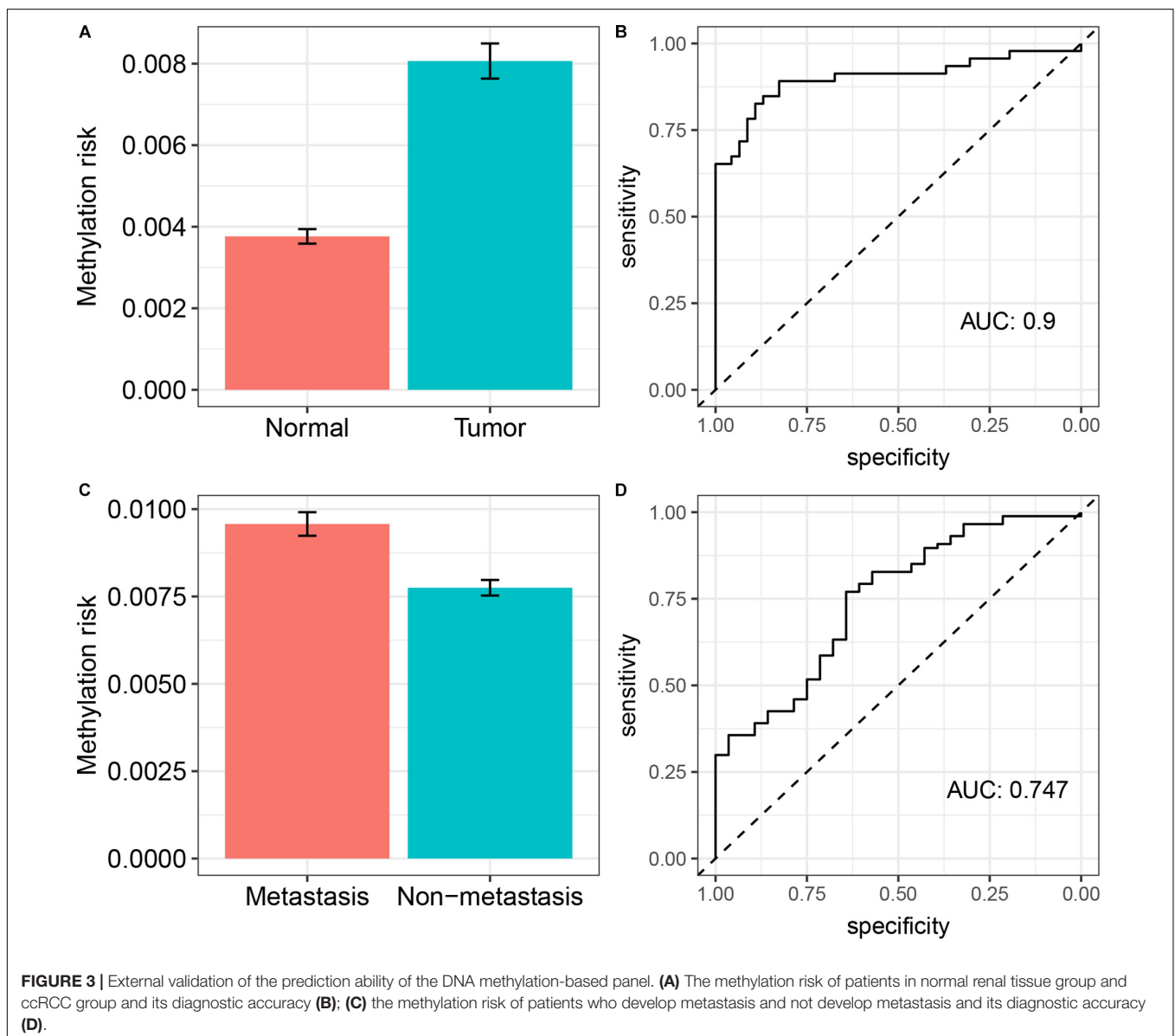
External Validation of the Prediction Ability of the DNA Methylation-Based Panel

Next, we tried to validate the prediction ability of the DNA methylation-based panel in two independent validation cohorts (GSE61441, Wei et al., 2015; GSE113501, Evelonn et al., 2019). We calculated the risk scores of the 46 paired ccRCC tumor and adjacent normal kidney tissues (GSE61441) as well as the risk scores of the ccRCC samples in GSE113501 on the basis of the methylation levels of the CpG sites in the DNA methylation-based panel and their responding coefficients in the elastic net penalized CoxPH model. As shown in Figures 3A,B,

the risk scores of tumor samples (ccRCC) were significantly higher (Wilcoxon test $P < 0.0001$) compared with those of adjacent normal tissues, which achieved an AUC of 0.9. Meanwhile, the DNA methylation-based panel could also significantly predict the metastasis of ccRCC (Wilcoxon test $P < 0.0001$), with an AUC of 0.747 (Figures 3C,D).

DNA Methylation-Based Panel Containing Nomogram and Its Clinical Application

To determine the clinical utility of the methylation-based panel, we incorporated the DNA methylation-based panel and other clinical variables (age, hemoglobin level, serum calcium level, tumor stage, and gender) into a multivariable survival model, and then developed a nomogram incorporating the DNA

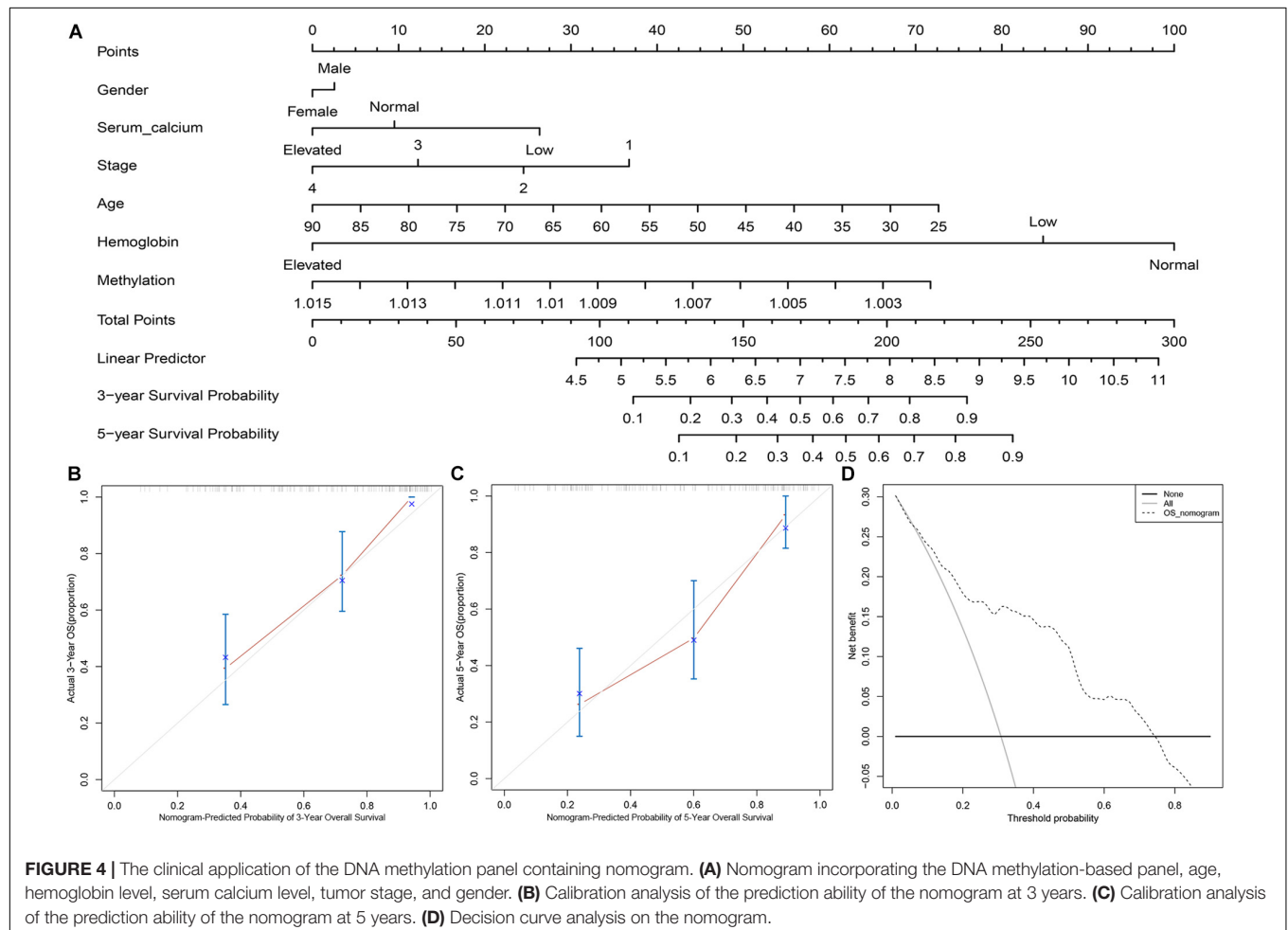


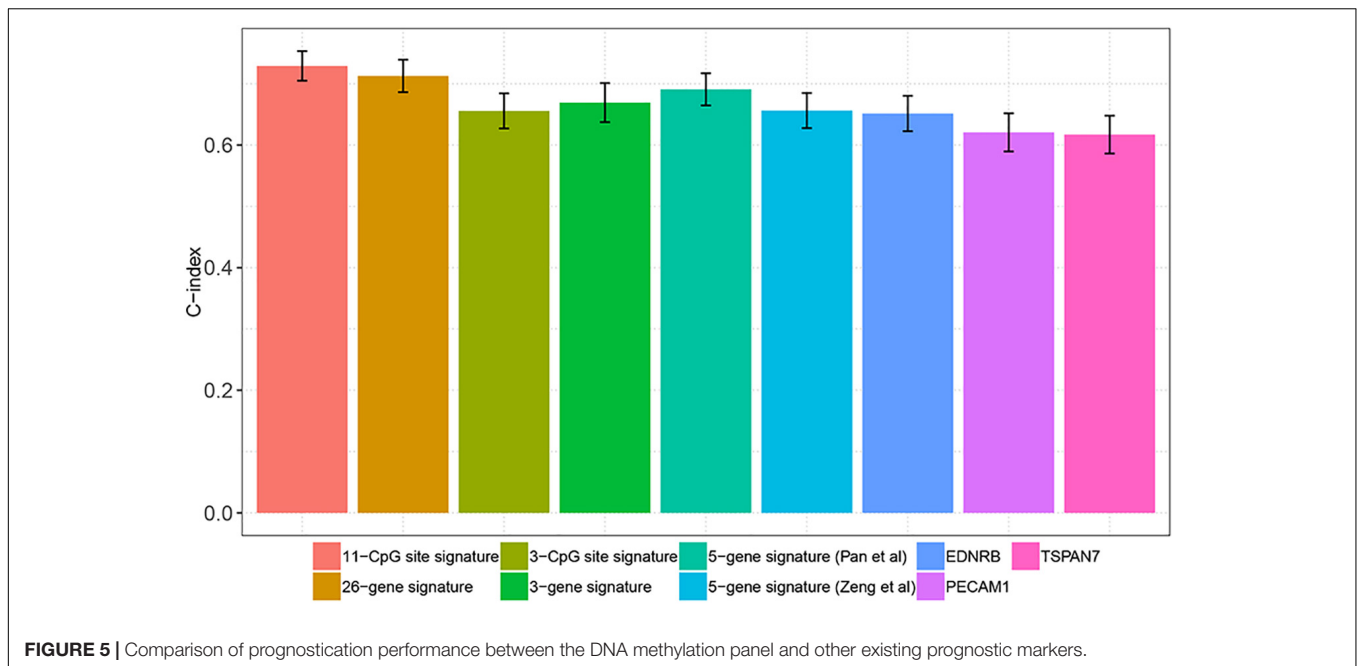
methylation-based panel (**Figure 4A**). Clinically, the nomogram could be used in such a way that the score (according to the “Points” line) of each variable is calculated, and then the score of each variable is summed to obtain a total score (according to the “Total points” line), and then a physician could estimate the 3-year survival probability and 5-year survival probability of patients with ccRCC based on the total score of each patient. The C-indexes were 0.7447 for internal validation with bootstrap method and 0.788 for external validation, respectively. Meanwhile, calibration analysis suggested a good agreement between the predicted and observed 3-, and 5- year OS probabilities (**Figures 4B,C**). Moreover, DCA curve suggested that the DNA methylation panel containing nomogram showed better prediction ability for the OS within a threshold probability of 1–75% (**Figure 4D**).

Comparison of Prognostication Performance Between the DNA Methylation-Based Panel and Other Existing Prognostic Markers

Several studies have reported that a variety of molecular markers can be used to predict the survival of patients with ccRCC.

Wuttig et al. (2012) suggested that PECAM1, EDNRB, and TSPAN7 could predict the survival of ccRCC patients. Kosari et al. (2005) classified ccRCC patients into aggressive and non-aggressive groups using a 36-gene signature and confirmed that this 35-gene signature predicts patient survival. Pan et al. (2019) introduced a 5-gene signature including OTX1, MATN4, PI3, ERVV-2, and NFE4 that was associated with the progression and prognosis of ccRCC patients. Kang et al. (2019) identified a 3- CpG site based promoter methylation signature associated with aggressive tumor phenotype and progression free survival in patients with ccRCC after surgical treatment. Zeng et al. (2019) indicated that the combination of IDUA, NDST1, SAP30L, CRYBA4, and SI was a prognostic signature (5-gene signature) in ccRCC. Chen et al. (2019) applied a least absolute shrinkage and selection operator (LASSO) penalized CoxPH model to identify a 3-gene signature that showed significantly prognostic ability in ccRCC. In the present study, we tried to evaluate and compare the prognostication performances of the above prognostic biomarkers and our DNA methylation-based panel based on C-index. As shown in **Figure 5**, the C-index of our DNA methylation-based panel was significantly higher compared with the 3-CpG site signature (0.73 versus 0.66, $P = 0.0295$), 3-gene signature (0.73 versus 0.67, $P = 0.024$),





two 5-gene signatures (0.73 versus 0.66, $P = 0.0114$ and 0.73 versus 0.69, $P = 0.0543$), EDNRB (0.73 versus 0.65, $P = 0.0015$), TSPAN7 (0.73 versus 0.62, $P = 0.0002$), and PECAM1 (0.73 versus 0.62, $P < 0.0001$). At the same time, the predictive performance of our DNA methylation-based panel was comparable with the 35-gene signature (0.73 versus 0.71, $P = 0.2777$).

Associations Between Methylation and Gene Expression and the Impact of the Associated Gene Expression on the Survival of Patients

As an epigenetic modification, DNA methylation has been demonstrated to play an important role in the regulation of gene expression (Spainhour et al., 2019). Thus, we tried to identify the correlations between the methylation levels of the 11 CpG sites and the expression levels of their corresponding genes. Unlike conventional understanding, the methylation levels in most sites are positively correlated with gene expressions. CpG site cg10009968 methylation level was positively correlated with the expression of CARD11 ($R = 0.37$, $P < 0.0001$); Methylation levels of cg07996594 ($R = 0.43$, $P < 0.0001$), cg15014975 ($R = 0.39$, $P < 0.0001$), cg24463471 ($R = 0.43$, $P < 0.0001$), and cg26256263 ($R = 0.4$, $P < 0.0001$) were positively correlated with the expression levels of RUNX3; The methylation level of CpG site cg01977762 was positively correlated with the expression level of UHRF1 ($R = 0.32$, $P < 0.0001$) and cg08840441 methylation was positively correlated with GMIP expression ($R = 0.45$, $P < 0.0001$); Meanwhile, there was a trend of positive correlation between the methylation level of CpG site cg15811515 ($R = 0.075$, $P = 0.19$) and the expression level of CSDAP1; Moreover,

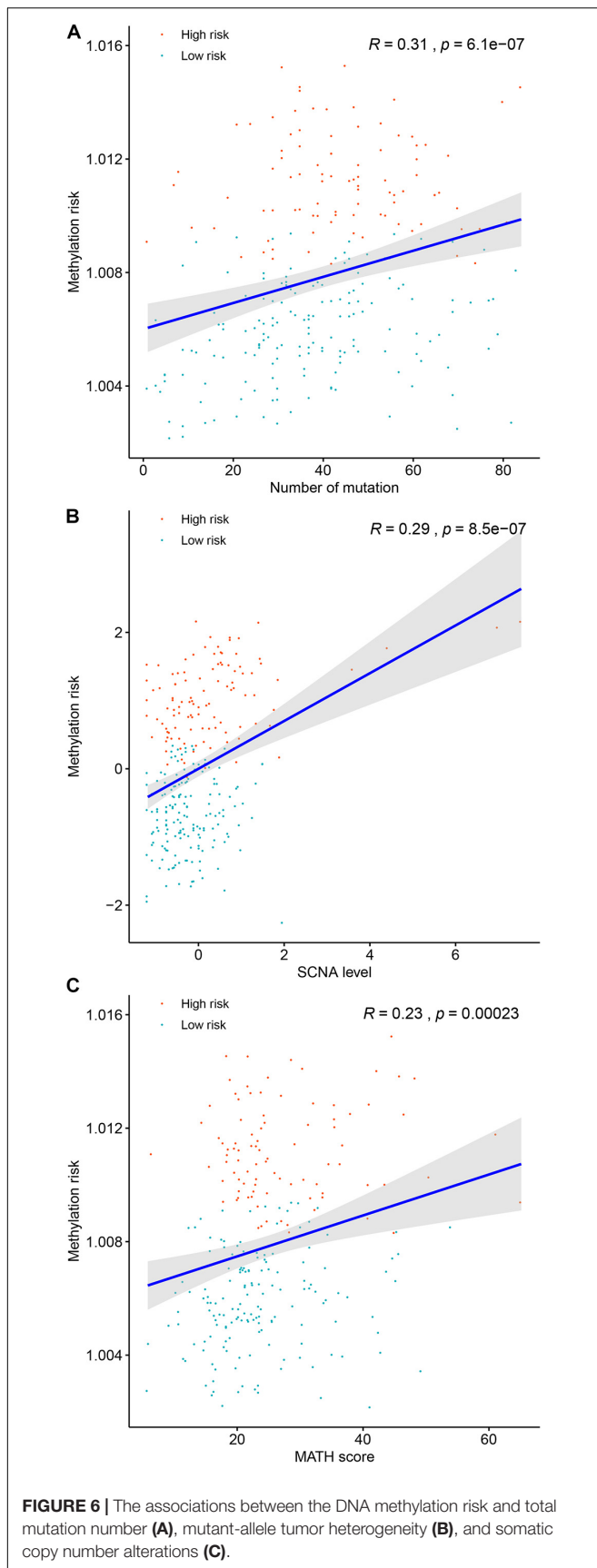
the expression of genes related with the 11 CpG sites could significantly stratify patients into different survival groups (Supplementary Figure S3).

Associations Between Genomic Metrics and the Methylation Risk of ccRCC Patients

Genomic abnormalities have been found in a variety of human tumors and global genomic aberration was found in the progression of human malignancy and closely related to their progression. Thus, we analyzed the correlations between common genomic abnormalities (total mutation load, SCNA and MATH and the DNA methylation-based panel associated methylation risk of ccRCC patients). As shown in Figure 6, the methylation risk of ccRCC patients was significantly correlated with total mutation number ($R = 0.31$, $P < 0.0001$, Figure 6A), SCNA ($R = 0.29$, $P < 0.0001$, Figure 6B), and MATH ($R = 0.23$, $P < 0.0001$, Figure 6C). These findings indicated that patients with higher methylation risk were associated with greater clonal heterogeneity.

Association Between Immune Cell Infiltration and the Methylation Risk of ccRCC Patients

Tumors are not just masses of monoclonal malignant cells, it also encompasses a variety of immune cells and stromal cells, which together constitute a complex tumor microenvironment. Tumor-infiltrating immune cells have been reported to play a significant role in tumor progression and management. Therefore, determination of the relationship between the methylation risk and the tumor-infiltrating immune cells could help clarify the mechanisms underlying the survival impact



of the methylation of the 11 CpG sites and help design immunogenic effects of anticancer therapies. ssGSEA-based immune cell infiltration level was calculated as mentioned above. As shown in **Figure 7A** and **Supplementary Figure S4**, Spearman's rank correlation analysis revealed weak positive associations between the methylation risk and T cell infiltration score (TIS) ($R = 0.14$, $P = 0.011$) and overall immune infiltration score (OIS) ($R = 0.15$, $P = 0.011$). While significantly negative correlations were found between the methylation risk and innate immune cells of NK cells ($R = -0.32$, $P < 0.0001$, **Figure 7B**) and mast cells ($R = -0.27$, $P < 0.0001$, **Figure 7C**).

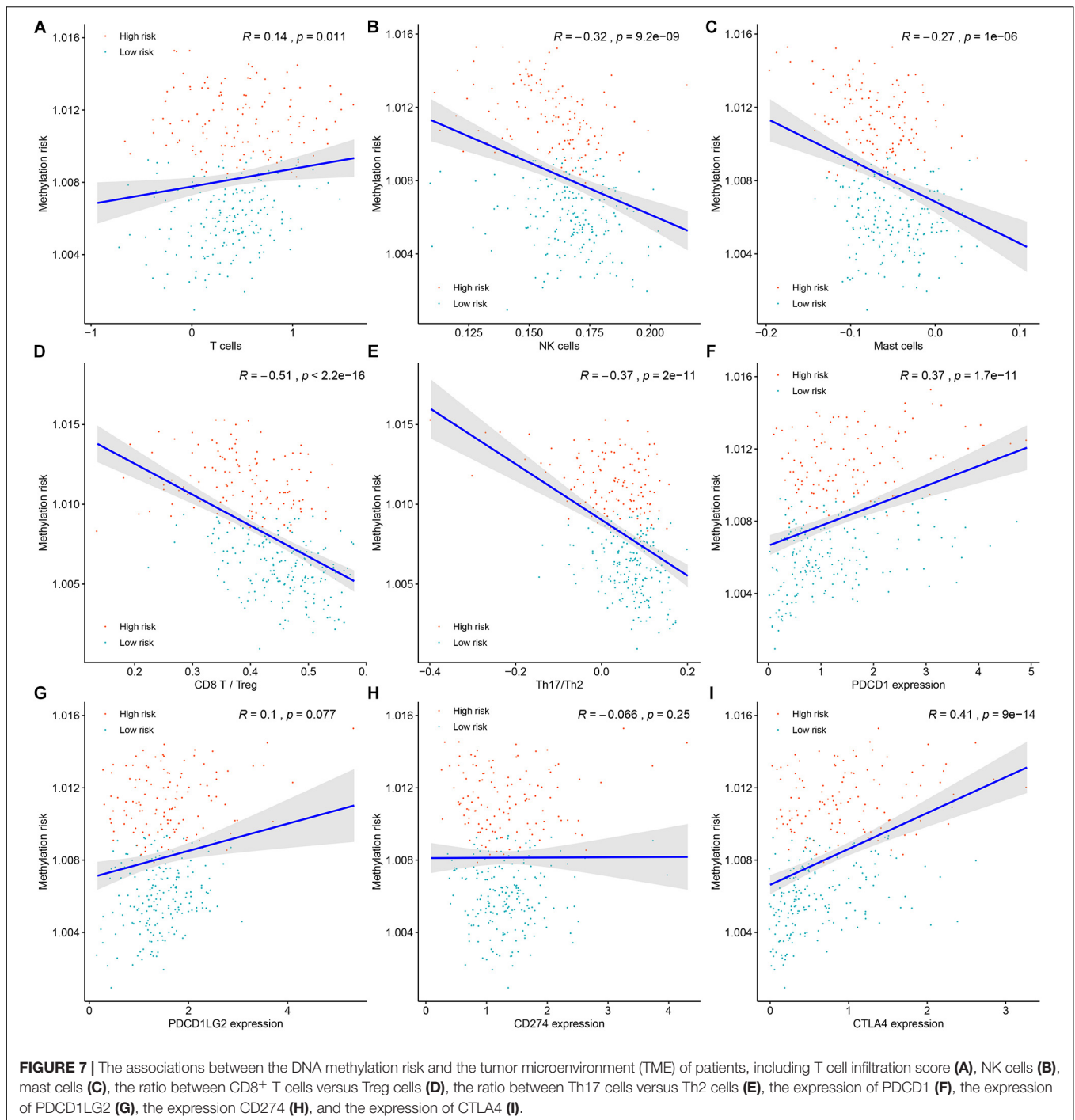
The ratio between protumorigenic immune cells versus antitumorigenic cells is more likely to determine whether the net effect of these cells is tumor promotion versus inhibition compared with the absolute count of certain immune cell types. As shown in **Figures 7D,E**, the ratio between CD8⁺ T cells versus Treg cells ($R = -0.51$, $P < 0.0001$) as well as the ratio between Th17 cells versus Th2 cells ($R = -0.31$, $P < 0.0001$) was negatively correlated with the methylation risk of ccRCC patients. Next, we tried to determine the associations between the expression of immunotherapeutic targets PD-1 (PDCD1), PD-L1 (CD274), PDCD1LG2 and CTLA-4 (CTLA4) (**Figures 7F-I**) and the methylation risk of ccRCC patients, and Spearman's correlation analysis revealed that the expressions of PDCD1 ($R = 0.37$, $P < 0.0001$, **Figure 7F**), PDCD1LG2 ($R = 0.1$, $P = 0.077$, **Figure 7G**), and CTLA4 ($R = 0.41$, $P < 0.0001$, **Figure 7I**) were positively correlated with the methylation risk of patients with ccRCC.

The Methylation Levels of the CpG Sites in ccRCC Samples and Its Normal Controls

We investigated the methylation levels of our the 11 CpG sites in 5 ccRCCs and 5 normal controls. As shown in **Supplementary Figure S5**, the methylation levels of cg15014975, cg07996594, cg26256263, cg18502142, cg15811515, cg09257635, cg10009968, cg08840441, and cg01977762 was obviously higher in ccRCCs compared with those in normal renal tissues (no significantly differences regarding the methylation levels of the others two CpG sites (cg24463471 and cg18279094) between the two groups), which was in accordance with the result of differentially methylation analysis above. Therefore, these results indicated that the 11 CpG sites containing DNA methylation panel would be reliable in clinical settings.

DISCUSSION

In the present study, a total of 2,628 of 361,127 CpG sites were differentially methylated between normal renal tissue and ccRCC and the global picture of DNA methylation patients in the high risk group, low risk group, and normal renal tissue group was shown in **Supplementary Figure S6**. Furthermore, we randomly divided patients in the TCGA-KIRC cohort into a training set and a test set, and used the elastic net regularized



CoxPH model in the training set to find 11 methylation sites closely related to the overall survival of the patient. At the same time, we used the univariate and multivariable CoxPH models in the training set and test set to verify the predictive performance of the DNA methylation-based panel on the OS of ccRCC patients.

In clinical applications, we constructed a nomogram containing the DNA methylation-based panel and other clinical features such as patient age, gender, and pathological stage,

etc., and confirmed the reliability of its clinical use using internal and external validation methods. External validation results indicated that the DNA methylation-based panel was able to significantly distinguish between normal renal tissue and ccRCC as well as metastasis and non-metastasis ccRCCs. Finally, in predicting the OS of ccRCC patients, we confirmed that our DNA methylation-based panel was superior to other existing predictive molecular markers. Taken together, the above results indicated that the DNA

methylation-based panel exhibited strong predictive value and might serve as an independent prognostic factor in patients with ccRCC.

Of the 11 CpG sites found in this study, 8 CpG sites (cg10009968, cg15811515, cg18279094, cg07996594, cg15014975, cg24463471, cg26256263, and cg01977762) are located in the promoter regions of the corresponding genes, and there are 3 CpG sites (cg09257635, cg08840441, and cg18502142) located in the gene body region of the corresponding genes. It is generally believed that methylation modifications of the promoter region can down-regulate the expression of the corresponding genes. However, unlike traditional understanding, our correlation analysis indicated that the methylation modifications of some promoter regions were positively correlated with the expression of the corresponding genes. This was consistent with the findings of Spainhour et al. (2019), indicating that a large number of positive correlations (about 30%) between methylation and gene expression in the 33 cancer types of TCGA database. Thus, there is an amount of more to be learned regarding the role of DNA methylation beyond the commonly accepted silencing role.

Associations between genomic metrics and prognosis of patients with malignance have been reported previously. Thus, in the present study, we tried to identify the associations between commonly accepted genomic metrics (total mutation load, SCNA, and MATH) and the methylation risk of ccRCC patients. As expected, patients with higher methylation risk were associated with higher total mutation load, which was consistent with previous observations in other tumors. SCNA, also known as aneuploidy, was widespread in multiple cancers and had been recognized as a driver event in carcinogenesis (Davoli et al., 2017). It was widely accepted that higher SCNA levels were correlated with high proliferation of cancer cells and poorer survival of cancer patients (Davoli et al., 2017). Our correlation analysis suggested that higher SCNA levels of ccRCC patients were associated with higher methylation risk, which were associated with poor survival of ccRCC patients. As another genomic alteration, MATH, a measure of intratumor heterogeneity, utilized the broadness of the distribution of mutant allele frequencies and had been used to assess the clonal and genetic heterogeneity of tumors. Higher MATH score corresponded to higher heterogeneity which fosters tumor evolution (McGranahan and Swanton, 2017). In the present study, we showed that patients with higher MATH score were significantly associated higher methylation risk, which ultimately resulted in poor survivals of patients.

Tumor immune cell infiltration represented an important component of tumor microenvironment and was used to classified multiple tumors into different immune subtypes associated different heterogeneity and survival (Thorsson et al., 2018). In the present study, we investigated the associations between the methylation risk and the infiltrating immune cells of ccRCC patients. Higher innate immune activity (measured by NK cell score and mast cell score) was associated lower methylation risk, unexpectedly, we observed

weak positive correlations between TIS ($R = 0.14$, $P = 0.011$) and OIIS ($R = 0.15$, $P = 0.011$) and the methylation risk of ccRCC patients. Given that functions of tumor-infiltrating immune cells are different. For example, CD8⁺ T cells are always associated with cytolytic activity, while regulatory T cells (Tregs) and tumor associated macrophages (TAMs) have been demonstrated to be correlated with protumor functions (Bindea et al., 2013). Thus, we investigated the correlation between the ratio of protumorigenic immune cells versus antitumorigenic cells and the methylation risk of ccRCC patients. As a result, ratio between CD8⁺ T cells (antitumorigenic) versus Treg cells (protumorigenic) was significantly correlated with lower methylation risk of ccRCC patients. Meanwhile, we found the ratio between Th17 cells (antitumorigenic) versus Th2 (protumorigenic) cells was significantly decreased in high versus low methylation risk patients. PD-1, also known as PDCD1, is a member of the CD28 superfamily that negatively regulates T cell activation in immune responses and plays a crucial role in immune tolerance by delivering negative signals upon interaction with its two ligands, PD-L1(CD274) or PD-L2 (PDCD1LG2) (Jin et al., 2011). CTLA-4 (CTLA4) is another immune receptor that negatively regulated T cells activation in immune response (Egen et al., 2002). Through correlation analysis, we observed that higher methylation risk was correlated with higher expression of these immune inhabitation molecules (PD-1, PDCD1LG2, CTLA-4). This might explain why higher TIS and OIIS was associated with higher methylation risk, i.e., the increased expression of PD-1, PDCD1LG2, and CTLA-4 mediates immune tolerance.

Taken together, we developed a DNA methylation-based panel which might be an independent prognostic factor in ccRCC. High levels of total mutation number, SCNA level, and MATH score were associated with higher methylation risk. The innate immune, and ratios between CD8⁺T cell versus Treg cell as well as Th17 cell versus Th2 cell were significantly decreased in high methylation risk group.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: TCGA-KIRC methylation profile was downloaded from UCSC Xena (<https://xenabrowser.net/>); copy number data, mRNA expression data, somatic mutation data and clinical data of TCGA-KIRC were downloaded from GDC data portal (<https://portal.gdc.cancer.gov>). GSE11501 were downloaded from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE113501>; GSE61441 were downloaded from <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61441>.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee at Zhongnan Hospital of Wuhan University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XW and SL designed the study. X-PL collected the data, performed the statistical analysis, conducted the methylation detection, and wrote the manuscript. LJ participated in the methylation detection. CC and TL reviewed the manuscript. All the authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank the patients and investigators who participated in TCGA and GEO for providing data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2020.572628/full#supplementary-material>

REFERENCES

- Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39, 782–795. doi: 10.1016/j.immuni.2013.10.003
- Bolger, G. (2002). Nephrectomy for metastatic renal-cell cancer. *N. Engl. J. Med.* 346, 1095–1096. doi: 10.1056/NEJM200204043461417
- Borys, A. M., Seweryn, M., Golabek, T., Belch, L., Klimkowska, A., Toton-Zuranska, J., et al. (2019). Patterns of gene expression characterize T1 and T3 clear cell renal cell carcinoma subtypes. *PLoS One* 14:e0216793. doi: 10.1371/journal.pone.0216793
- Cancer Genome Atlas Research Network (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 499, 43–49. doi: 10.1038/nature12222
- Capitani, U., and Montorsi, F. (2016). Renal cancer. *Lancet* 387, 894–906. doi: 10.1016/S0140-6736(15)00046-X
- Chen, L., Luo, Y., Wang, G., Qian, K., Qian, G., Wu, C. L., et al. (2019). Prognostic value of a gene signature in clear cell renal cell carcinoma. *J. Cell. Physiol.* 234, 10324–10335. doi: 10.1002/jcp.27700
- Davoli, T., Uno, H., Wooten, E. C., and Elledge, S. J. (2017). Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* 355:eaaf8399. doi: 10.1126/science.aaf8399
- de Cubas, A. A., and Rathmell, W. K. (2018). Epigenetic modifiers: activities in renal cell carcinoma. *Nat. Rev. Urol.* 15, 599–614. doi: 10.1038/s41585-018-0052-7
- Egen, J. G., Kuhns, M. S., and Allison, J. P. (2002). CTLA-4: new insights into its biological function and use in tumor immunotherapy. *Nat. Immunol.* 3, 611–618. doi: 10.1038/ni0702-611
- Evelonn, E. A., Landfors, M., Haider, Z., Kohn, L., Ljungberg, B., Roos, G., et al. (2019). DNA methylation associates with survival in non-metastatic clear cell renal cell carcinoma. *BMC Cancer* 19:65. doi: 10.1186/s12885-019-5291-3
- Fidler, I. J. (1992). The biology of renal cancer metastasis. *Semin. Urol.* 10, 3–11.
- Ge, Q., Lu, M., Ju, L., Qian, K., Wang, G., Wu, C. L., et al. (2019). miR-4324-RACGAP1-STAT3-ESR1 feedback loop inhibits proliferation and metastasis of bladder cancer. *Int. J. Cancer* 144, 3043–3055. doi: 10.1002/ijc.32036
- Golovastova, M. O., Korolev, D. O., Tsoy, L. V., Varshavsky, V. A., Xu, W. H., Vinarov, A. Z., et al. (2017). Biomarkers of renal tumors: the current state and clinical perspectives. *Curr. Urol. Rep.* 18:3. doi: 10.1007/s11934-017-0655-1
- Gonzalez Leon, T., and Morera Perez, M. (2016). Renal cancer in the elderly. *Curr. Urol. Rep.* 17:6. doi: 10.1007/s11934-015-0562-2
- Hancock, S. B., and Georgiades, C. S. (2016). Kidney cancer. *Cancer J.* 22, 387–392. doi: 10.1097/PPO.0000000000000225
- Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 14:7. doi: 10.1186/1471-2105-14-7
- Jin, H. T., Ahmed, R., and Okazaki, T. (2011). Role of PD-1 in regulating T-cell immunity. *Curr. Top. Microbiol. Immunol.* 350, 17–37. doi: 10.1007/82_2010_116
- Kamarudin, A. N., Cox, T., and Kolamunnage-Dona, R. (2017). Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med. Res. Methodol.* 17:53. doi: 10.1186/s12874-017-0332-6
- Kang, H. W., Park, H., Seo, S. P., Byun, Y. J., Piao, X. M., Kim, S. M., et al. (2019). Methylation signature for prediction of progression free survival in surgically treated clear cell renal cell carcinoma. *J. Korean Med. Sci.* 34:e144. doi: 10.3346/jkms.2019.34.e144
- Kosari, F., Parker, A. S., Kube, D. M., Lohse, C. M., Leibovich, B. C., Blute, M. L., et al. (2005). Clear cell renal cell carcinoma: gene expression analyses identify a potential signature for tumor aggressiveness. *Clin. Cancer Res.* 11, 5128–5139. doi: 10.1158/1078-0432.CCR-05-0073
- Liss, M. A., Chen, Y., Rodriguez, R., Pruthi, D., Johnson-Pais, T., Wang, H., et al. (2019). Immunogenic heterogeneity of renal cell carcinoma with venous tumor thrombus. *Urology* 124, 168–173. doi: 10.1016/j.urol.2018.09.018
- Liu, X. P., Yin, X. H., Meng, X. Y., Yan, X. H., Wang, F., and He, L. (2018). Development and validation of a 9-gene prognostic signature in patients with multiple myeloma. *Front. Oncol.* 8:615. doi: 10.3389/fonc.2018.00615
- Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28, 1747–1756. doi: 10.1101/gr.239244.118
- McGrath, N., and Swanton, C. (2017). Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* 168, 613–628. doi: 10.1016/j.cell.2017.01.018
- Moch, H., Srigley, J., Delahunt, B., Montironi, R., Egevad, L., and Tan, P. H. (2014). Biomarkers in renal cancer. *Virchows Arch.* 464, 359–365. doi: 10.1007/s00428-014-1546-1
- Morris, M. R., and Latif, F. (2017). The epigenetic landscape of renal cancer. *Nat. Rev. Nephrol.* 13, 47–60. doi: 10.1038/nrneph.2016.168

- Morris, T. J., Butcher, L. M., Feber, A., Teschendorff, A. E., Chakravarthy, A. R., Wojdacz, T. K., et al. (2014). ChAMP: 450k Chip analysis methylation pipeline. *Bioinformatics* 30, 428–430. doi: 10.1093/bioinformatics/btt684
- Mroz, E. A., and Rocco, J. W. (2013). MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol.* 49, 211–215. doi: 10.1016/j.oraloncology.2012.09.007
- Pan, Q., Wang, L., Zhang, H., Liang, C., and Li, B. (2019). Identification of a 5-gene signature predicting progression and prognosis of clear cell renal cell carcinoma. *Med. Sci. Monit.* 25, 4401–4413. doi: 10.12659/MSM.917399
- Senbabaoglu, Y., Gejman, R. S., Winer, A. G., Liu, M., Van Allen, E. M., de Velasco, G., et al. (2016). Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* 17:231. doi: 10.1186/s13059-016-1092-z
- Sill, M., Hielscher, T., Becker, N., and Zucknick, M. (2014). c060: extended inference with lasso and elastic-net regularized cox and generalized linear models. *J. Stat. Softw.* 62, 1–22. doi: 10.1080/03610926.2020.1821892
- Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011). Regularization paths for Cox's proportional hazards model via coordinate descent. *J. Stat. Softw.* 39, 1–13.
- Singh, A. K., Winslow, T. B., Kermany, M. H., Goritz, V., Heit, L., Miller, A., et al. (2017). A pilot study of stereotactic body radiation therapy combined with cytoreductive nephrectomy for metastatic renal cell carcinoma. *Clin Cancer Res.* 23, 5055–5065. doi: 10.1158/1078-0432.CCR-16-2946
- Spainhour, J. C., Lim, H. S., Yi, S. V., and Qiu, P. (2019). Correlation patterns between DNA methylation and gene expression in the cancer genome atlas. *Cancer Inform.* 18:1176935119828776. doi: 10.1177/1176935119828776
- Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48, 812–830.e14. doi: 10.1016/j.immuni.2018.03.023
- van Vlodrop, I. J. H., Joosten, S. C., De Meyer, T., Smits, K. M., Van Neste, L., Melotte, V., et al. (2017). A four-gene promoter methylation marker panel consisting of GREM1, NEURL, LAD1, and NEFH predicts survival of clear cell renal cell cancer patients. *Clin. Cancer Res.* 23, 2006–2018. doi: 10.1158/1078-0432.CCR-16-1236
- Vickers, A. J., and Elkin, E. B. (2006). Decision curve analysis: a novel method for evaluating prediction models. *Med. Decis. Making* 26, 565–574. doi: 10.1177/0272989X06295361
- Wei, J. H., Haddad, A., Wu, K. J., Zhao, H. W., Kapur, P., Zhang, Z. L., et al. (2015). A CpG-methylation-based assay to predict survival in clear cell renal cell carcinoma. *Nat. Commun.* 6:8699. doi: 10.1038/ncomms9699
- Wu, J., Jin, S., Gu, W., Wan, F., Zhang, H., Shi, G., et al. (2019). Construction and validation of a 9-gene signature for predicting prognosis in Stage III clear cell renal cell carcinoma. *Front. Oncol.* 9:152. doi: 10.3389/fonc.2019.00152
- Wuttig, D., Zastrow, S., Fussel, S., Toma, M. I., Meinhardt, M., Kalman, K., et al. (2012). CD31, EDNRB and TSPAN7 are promising prognostic markers in clear-cell renal cell carcinoma revealed by genome-wide expression analyses of primary tumors and metastases. *Int. J. Cancer* 131, E693–E704. doi: 10.1002/ijc.27419
- Young, J. R., Coy, H., Douek, M., Lo, P., Sayre, J., Pantuck, A. J., et al. (2017). Clear cell renal cell carcinoma: identifying the gain of chromosome 12 on multiphasic MDCT. *Abdom Radiol. (NY)* 42, 236–241. doi: 10.1007/s00261-016-0868-7
- Zeng, M. H., Qiu, J. G., Xu, Y., and Zhang, X. H. (2019). IDUA, NDST1, SAP30L, CRYBA4, and SI as novel prognostic signatures clear cell renal cell carcinoma. *J. Cell. Physiol.* 234, 16320–16327. doi: 10.1002/jcp.28297
- Zheng, Z., Cai, Y., Chen, H., Chen, Z., Zhu, D., Zhong, Q., et al. (2018). CXCL13/CXCR5 axis predicts poor prognosis and promotes progression through PI3K/AKT/mTOR pathway in clear cell renal cell carcinoma. *Front. Oncol.* 8:682. doi: 10.3389/fonc.2018.00682
- Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net (vol B 67, pg 301, 2005). *J. R. Stat. Soc. Series B Stat. Methodol.* 67:768. doi: 10.1111/j.1467-9868.2005.00527.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Liu, Ju, Chen, Liu, Li and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.