



## A Two-DNA Methylation Signature to Improve Prognosis Prediction of Clear Cell Renal Cell Carcinoma

Shanping Shi<sup>1</sup>, Shazhou Ye<sup>1</sup>, Xiaoyue Wu<sup>1</sup>, Mingjun Xu<sup>1</sup>, Renjie Zhuo<sup>2</sup>, Qi Liao<sup>2</sup>, and Yang Xi<sup>1</sup>

<sup>1</sup>Diabetes Center, Zhejiang Provincial Key Laboratory of Pathophysiology, Institute of Biochemistry and Molecular Biology, Medical School of Ningbo University, Ningbo, Zhejiang;

<sup>2</sup>Department of Preventative Medicine, Medical School of Ningbo University, Ningbo, Zhejiang, China.

**Purpose:** Effective biomarkers and models are needed to improve the prognostic prospects of clear cell renal cell carcinoma (ccRCC). The purpose of this work was to identify DNA methylation biomarkers and to evaluate the utility of DNA methylation analysis for ccRCC prognosis.

**Materials and Methods:** An overview of genome-wide methylation of ccRCC tissues derived from The Cancer Genome Atlas (TCGA) database was download for analysis. DNA methylation signatures were identified using Cox regression methods. The potential clinical significance of methylation biomarkers acting as a novel prognostic markers was analyzed using the Kaplan-Meier method and receiver operating characteristic (ROC) curves.

**Results:** This study analyzed data for 215 patients with information on 23171 DNA methylation sites and identified a two-DNA methylation signature (cg18034859, cg24199834) with the help of a step-wise multivariable Cox regression model. The area under the curve of ROCs for the two-DNA methylation signature was 0.819. The study samples were stratified into low- and high-risk classifications based on an optimal threshold, and the two groups showed markedly different survival rates. Moreover, the two-DNA methylation marker was suitable for patients of varying ages, sex, stages (I and IV), and histologic grade (G2).

**Conclusion:** The two-DNA methylation signature was deemed to be a potential novel prognostic biomarker of use in increasing the accuracy of predicting overall survival of ccRCC patients.

**Key Words:** Carcinoma, renal cell, biomarkers, DNA methylation, prognosis

### INTRODUCTION

Clear cell renal cell carcinoma (ccRCC) is the most potentially lethal subtype of renal cell carcinoma (RCC) and accounts for 94% of metastatic RCCs.<sup>1-3</sup> Early diagnosis of ccRCC is difficult, and by the time it is identified, it typically has already metastasized in 30-35% of patients.<sup>4</sup> In early stage patients (stages I

and II), most ccRCCs are treated by resection via surgery, with survival rates over 5 years greater than 70%, while as for patients in stages III (regional spread) and IV (metastatic), which have a poor prognosis, survival rates at 5 years are only 50% and 10%, respectively.<sup>5-7</sup> Therefore, predictive methods with high specificity and sensitivity may improve the prognosis of patients.

Important targets in cancer research, cancer initiation and proliferation are regulated by epigenetic modifications.<sup>8</sup> As one of the basic mechanism of epigenetic modification, DNA methylation plays as a pivotal role in the processes of chromatin structure regulation and gene regulation.<sup>9</sup> Abnormal DNA methylation was reported to show a significant relationship with the occurrence and prognosis of various cancers.<sup>10</sup> Dubrowskaja, et al.<sup>11</sup> demonstrated that neurofilament heavy (*NEFH*)-specific hypermethylation in RCC is relevant to overall survival rates among patients undergoing targeted therapy. Mir-874, which is downregulated in breast cancer, is correlat-

**Received:** June 10, 2019 **Revised:** August 30, 2019

**Accepted:** September 3, 2019

**Corresponding author:** Yang Xi, PhD, Diabetes Center, Zhejiang Provincial Key Laboratory of Pathophysiology, Institute of Biochemistry and Molecular Biology, Medical School of Ningbo University, 360 Dacheng Rd, Ningbo, Zhejiang 315211, China.

Tel: 86-574-87600754, Fax: 86-574-87608638, E-mail: xiyang@nbu.edu.cn

•The authors have no potential conflicts of interest to disclose.

© Copyright: Yonsei University College of Medicine 2019

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ed with the prognosis of breast cancer.<sup>12</sup> Zhang, et al.<sup>13</sup> demonstrated that decreased of mir-874 expression is mediated by DNA methylation of the promoter region and that DNA methylation is directly related to the prognosis of breast cancer. Furthermore, DNA methylation of specific genes has already been used as biomarkers in early diagnosis and prognosis,<sup>14,15</sup> such as aldehyde dehydrogenase 1 family member A3 (*ALDH1A3*) promoter hypermethylation in glioma CpG island methylator phenotype-primary glioblastoma<sup>16</sup> and hypomethylation of cyclin dependent kinase inhibitor 2A (*CDKN2A*) in colon cancer.<sup>17</sup> Moreover, for patients with stage I non-small cell lung cancer, a shorter relapse-free survival rate has been found to be associated with higher methylation levels of CpG sites within five genes (*HIST1H4F*, *PCDHGB6*, *NPBWR1*, *ALX1*, and *HOXA9*).<sup>18</sup>

It was reported that ccRCC typically shows features of epigenetic alterations and methylated CpGs accompanying progress of disease,<sup>19,20</sup> although a specific prognostic model has not been proposed. This study analyzed whole-genome methylation profiles for ccRCC in The Cancer Genome Atlas (TCGA) database to identify potential biomarkers for predicting survival times for ccRCC patients.

## MATERIALS AND METHODS

### DNA methylation information for ccRCC tissues in the TCGA dataset

The DNA methylation information of ccRCC (platform: Illumina Human Methylation 27; downloaded in March 27, 2019; depending on GRCh38, the genomic coordinates of the CpGs were gained) and relevant clinical data were obtained from the TCGA database (<https://cancergenome.nih.gov/>). No further approval from the Ethics Committee was requested due to the data being obtained from the TCGA dataset. We studied how DNA methylation levels impacted survival in ccRCC using dates reported in survival records of the patients.  $\beta$  values were used to express DNA methylation levels, and  $M/(M+U)$  was computed. In  $M/(M+U)$ , M is defined as the signal from methylated beads, while U is defined as the signal from unmethylated beads at targeted CpG sites. Ultimately, 215 samples comprised 96 ccRCCs of grade 2, 82 of grade 3, and 28 of grade 4<sup>21</sup> and 113 of stage I, 27 of stage II, 49 of stage III, and 25 of stage IV.<sup>22</sup> In total, 23171 DNA methylation sites were selected in this work, and samples with relevant clinical records were obtained from the TCGA dataset. According to the TCGA series number, these 215 samples were assigned into a training dataset (60%) and a validation dataset (40%): one was used for discriminating and designing prognostic markers, while the other was applied to prove the precision of the biomarkers in survival prediction.

### Prognostic identification and choosing CpG sites

Detection of methylation signatures was performed by means of the computing environment R with Survival package (R version 3.5.1; Revolution Analytics, Mountain View, CA, USA). First, methylation markers ( $p < 0.05$ ) found to be closely related to patient survival were defined as candidate markers through univariate Cox proportional hazard analysis. Then, robust likelihood-based survival analysis modeling and multivariable Cox regression analysis was applied to further screen for CpG sites.<sup>23</sup>

### Establishment and validation of risk scores

We established risk scores for prognosis-related CpG sites by means of estimation of the regression coefficients in the multivariable Cox regression analysis of the training set.  $p$  values were characterized using the Wilcoxon rank-sum test, followed by false discovery rate correction. Then, we ranked the samples according to the risk score. We drew cross-validated time-dependent receiver operating characteristic (ROC) curves to verify the prediction accuracy of this model and to develop an effective optimal threshold. Based on this threshold, the samples were classified into high- and low-risk groups. Thereafter, cumulative survival time was calculated by a nonparametric statistic (Kaplan–Meier estimator), with a log-rank test (Mantel–Cox). Also, differences between the two groups in overall survival time were compared. Furthermore, the risk score system was evaluated by fitting in the checking set and the complete set. Also, by using different regrouping methods, we predicted the performance of the DNA methylation signature.

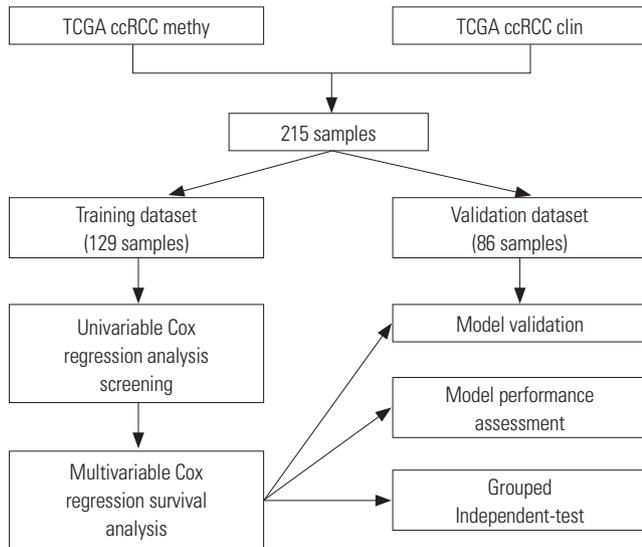
## RESULTS

### Collection of data from TCGA

The study was performed as indicated in Fig. 1. In total, 23171 DNA methylation sites were collected for 215 patient samples from the TCGA DNA methylation and clinical dataset. The median age of the patients was 60 years (range, 33–86 years), and the median survival time was 2386 days. The patients were randomly assigned to a training set ( $n=129$ , used to identify key CpG sites) and to a testing set ( $n=86$ , used to verify the methylation signature) based on status (alive/dead). The clinical data, including age, sex, ethnicity, stage, and status, are listed in the Table 1.

### Screening and identification of DNA methylation markers in the training dataset

Methylation levels were used as variables in the training dataset to identify DNA methylation markers related with ccRCC, and based thereon, univariate Cox proportional hazard regression analysis and robust likelihood-based survival analysis were performed. The first 19 significant prognosis-related CpG sites were chosen ( $p < 0.001$ , Supplementary Table 1, only on-



**Fig. 1.** Flow-chart of the study. The order of analyses to develop the risk score model and to validate the efficiency of the signature to predict prognostic outcomes. TCGA, The Cancer Genome Atlas; ccRCC, clear cell renal cell carcinoma.

line). Next, the 19 CpG sites of prognostic significance were put into multivariable Cox regression survival analysis to further screen CpG sites in order to identify an optimum prognostic model for predicting ccRCC. Finally, two methylation sites (cg18034859, cg24199834) in the hazard ratio (HR) model were identified; their risk coefficients are listed in Table 2. cg24199834 was indicated as a risk factor with a HR of 1.03, while the methylation level of cg24199834 was inversely proportional to the risk of death, with an HR of 0.78, indicating that cg18034859 was a protective factor. Accordingly, we developed the following risk score formula: Risk score =  $-0.2455 \times \beta$  value of cg18034859 +  $0.0304 \times \beta$  value of cg24199834. The formula demonstrated that the hypermethylation levels of cg24199834 are associated with higher risk, while the hypomethylation levels of cg18034859 are related to lower risk.

**Evaluation of the predictive performance of the two-DNA methylation signature**

After screening, 0.54 was identified as the optimal threshold cut-off point, exhibiting the best performance (Fig. 2A). The

**Table 1.** Clinicopathological Characteristics of ccRCC Patients from TCGA

Characteristics	Groups	Patients		
		Total (n=215)	Training dataset (n=129)	Validation dataset (n=86)
Age (yr)	Median	60	59	60
	Range	33–86	33–83	33–86
	≤60	106	66 (62.26)	40 (37.74)
	>60	108	62 (57.41)	46 (42.59)
	Unknown	1	1 (100.00)	0 (0.00)
Stage	I	109	65 (59.63)	44 (40.37)
	II	30	16 (53.33)	14 (46.67)
	III	51	31 (60.78)	20 (39.22)
	IV	24	17 (70.83)	7 (29.17)
	Unknown	1	0 (0.00)	1 (100.00)
Histologic	G2	94	53 (56.38)	41 (43.62)
	G3	89	57 (64.04)	32 (35.96)
	G4	25	15 (60.00)	10 (40.00)
	Others	7	4 (57.14)	3 (42.86)
Gender	Male	145	83 (57.24)	62 (42.76)
	Female	70	46 (65.71)	24 (34.29)
Status	Alive	146	86 (58.90)	60 (41.10)
	Dead	69	43 (62.32)	26 (37.68)

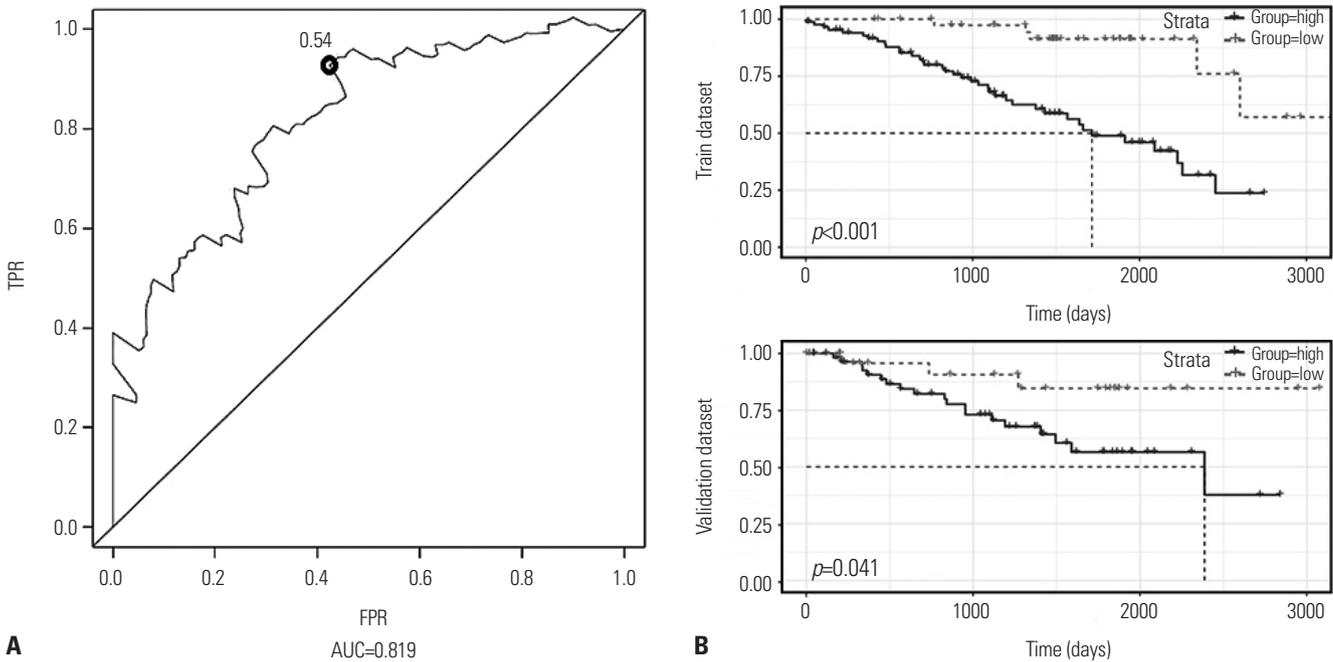
ccRCC, clear cell renal cell carcinoma; TCGA, The Cancer Genome Atlas. Values are presented as n (%) unless otherwise indicated.

**Table 2.** Multivariable Survival Analysis

CpG sites	Gene	Position	Coef*	Exp (coef)*	p value <sup>†</sup>	95% CI*
cg18034859	MYLK2	chr20:29870349	-0.2455	0.7823	<0.001	0.6992–0.8752
cg24199834	POU4F2	chr4:147779576	0.0304	1.0309	<0.001	1.0159–1.0461

CI, confidence interval.

\*In multivariable Cox regression analysis; <sup>†</sup>p value was calculated through the Wilcoxon rank-sum test followed by false discovery rate adjustment for multiple corrections.



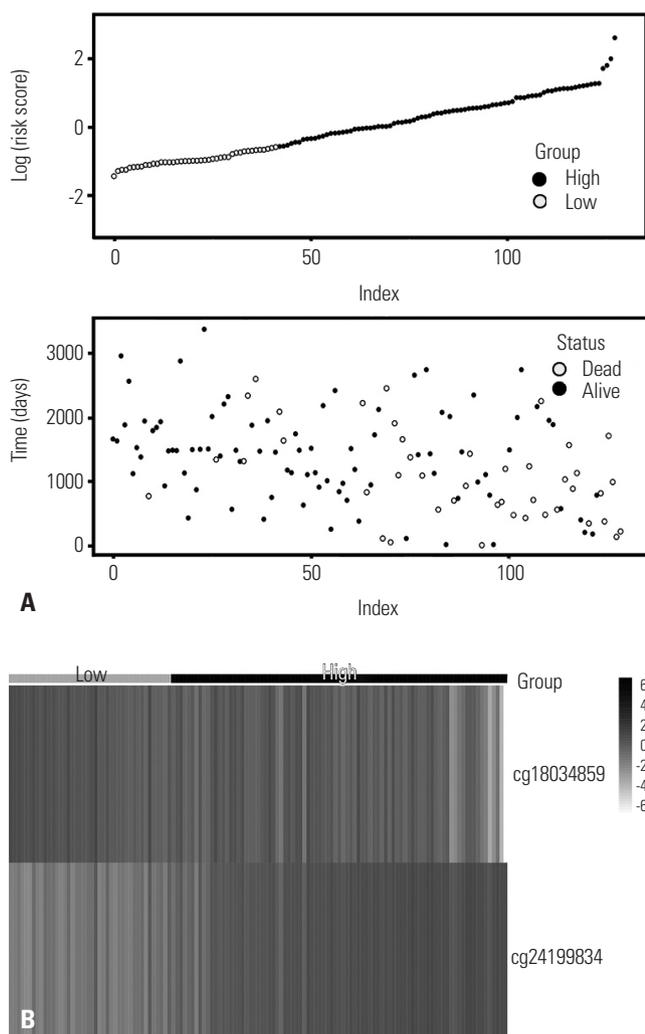
**Fig. 2.** Evaluation of the predictive performance of the two-DNA methylation signature. (A) Receiver operating characteristic analysis of the sensitivity and specificity for survival time by the two-DNA methylation signature in the training dataset. The black dot represents the optimal validation dataset using the two-DNA methylation signature. (B) The Kaplan-Meier analysis was used to visualize the survival probability for the low-risk versus high-risk groups of patients based on the optimal cut-off point. Rows represent survival time (days), and columns represent survival rate. TPR, true positive rate; FPR, false positive rate; AUC, area under the curve of receiver operating characteristics.

area under the curve of ROCs (AUC) for the two-DNA methylation signature was 0.819, indicating a favorable prognostic value in prediction of patient survival. Then, the training dataset was divided into high-risk (n=87) and low-risk groups (n=42). Kaplan-Meier survival analysis was applied to compare the survival time of the two groups. The results indicated that patients in the high-risk group had shorter survival time than the low-risk group. A significantly worse prognosis was embodied in the high-risk group ( $p < 0.001$ , Fig. 2B). To further validate the prognostic ability of the two-DNA methylation signature, we divided the validation dataset into high-risk (n=58) and low-risk groups (n=28). Similar to the above, the high-risk group had significantly shorter survival time in Kaplan-Meier survival analysis (Fig. 2B). Fig. 3A shows the risk score distribution and the scatter plot of patient states, with higher risk corresponding to more deaths. Heat mapping depicts the methylation levels of cg18034859 and cg24199834 between the two groups for the entire dataset. Significantly higher methylation levels for cg24199834 were confirmed, while the opposite was true for cg18034859 (Fig. 3B).

**Predictive performance of the two-DNA methylation signature**

Studies have reported that age,<sup>24</sup> sex,<sup>25</sup> pathological stage,<sup>26</sup> and histologic grade etc.<sup>27</sup> are associated with prognostic survival. Accordingly, all patient samples were regrouped according to clinicopathological characteristics, and the two-DNA methylation signature was validated separately in the different

groups. First, the effect of the two-DNA methylation signature among the different age groups of patients (based on survival status, “alive” used the final follow-up time; otherwise the time of death):  $\leq 60$  (n=106, 49.53%) and  $> 60$  (n=108, 50.47%) years were analyzed (Fig. 4A). From Kaplan-Meier curves, we found that the overall survival of patients in different risk groups differed between the two age cohorts ( $p < 0.01$ ), and AUC values were 0.860 in patients younger than 60 years old and 0.666 in patients older than 60 years old (Fig. 4B), exhibiting the independent relationship between the two-DNA methylation signature and age. Ricketts, et al.<sup>25</sup> found that the spread of ccRCC among the sexes is distinctly uneven. Sex may affect mutation spectra and hold implications for prognosis. A summary analysis of age, sex, pathological stage, and histologic grade is listed in Table 3. The two risk groups both in male and female patients ( $p < 0.01$ ) reflected obviously distinct survival times, and the AUCs for male and female sex were 0.772 and 0.832, respectively. Taking the sample number into account, we verified the predictive value of the two-DNA methylation signature in G2 (n=96) and G3 (n=82), with correlated AUC values of 0.877 and 0.628. About one quarter of the patients showed distant metastasis at initial diagnosis, and our results showed that overall survival was significantly different between the risk groups in pathological stage IV patients ( $p = 0.01$ ). The AUCs in stages I, II, III, and IV cohorts were 0.714, 0.686, 0.721, and 0.577, respectively. From the above results, the two-DNA methylation signature was proven to be an independent applicable value for predicting patient survival.



**Fig. 3.** Risk score analysis of the training set. The distribution of two-DNA methylation based risk core, patient survival, and methylation levels of two CpGs were analyzed in the training set (n=129). (A) Two-DNA methylation signature risk score distribution. 'Time' means 'survival time'. (B) Heat-map of the DNA methylation profiles. Rows represent CpG sites, and columns represent patients.

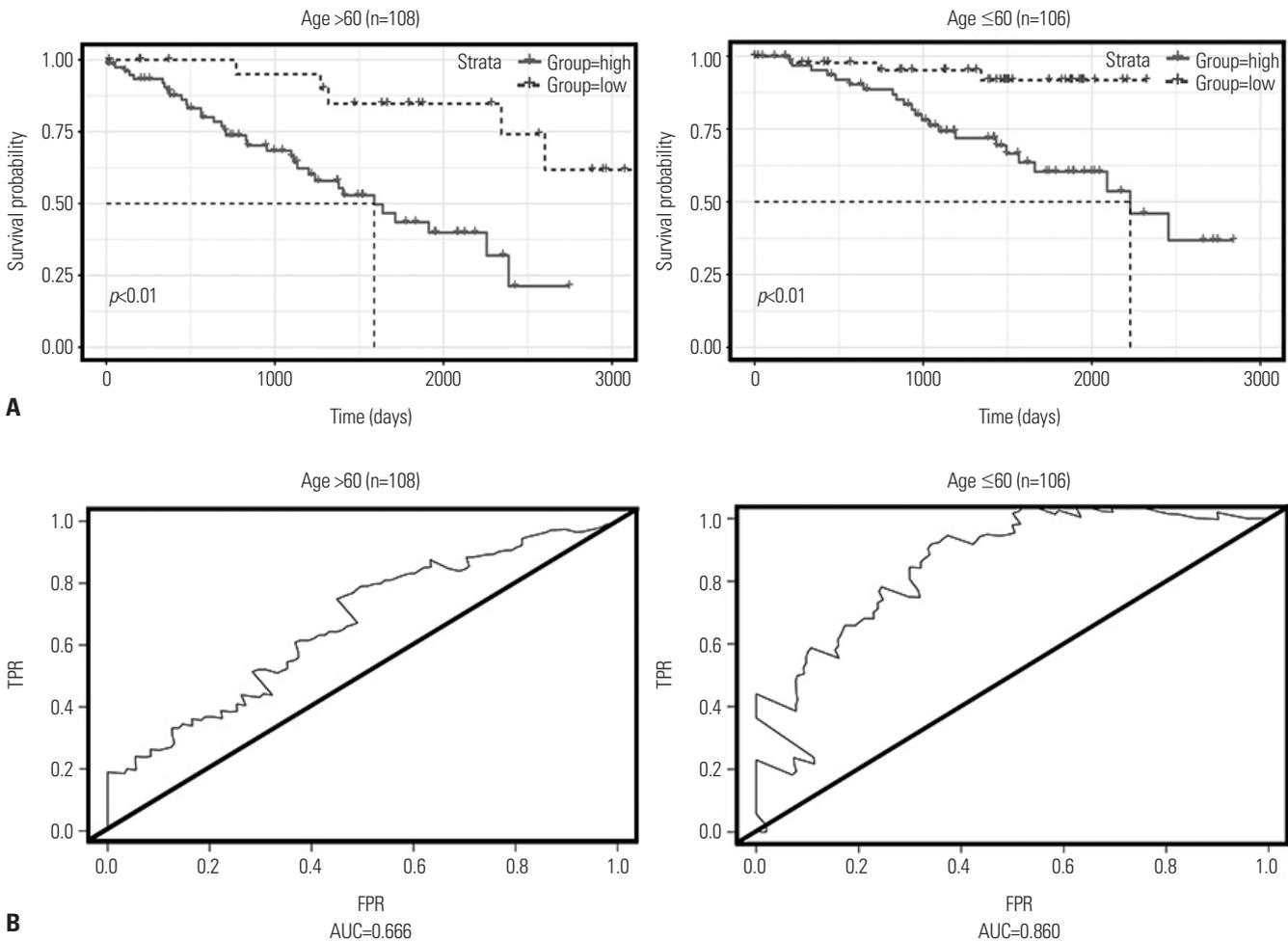
## DISCUSSION

A lack of effective and reliable prognostic biomarkers or models remain the focus of ccRCC research. Studies have demonstrated that molecular markers, including DNA methylation, can be used to predict the prognosis of ccRCC.<sup>28-31</sup> Compared to using individual DNA methylation factors as predictors, combinations of DNA methylation could achieve higher sensitivity and specificity.<sup>32</sup> In the present study, we generated a risk score model based on a two-DNA methylation signature significantly related to ccRCC prognosis in Cox regression and ROC analysis. Median survival time was used to divide high- and low- risks groups, which has more clinical practicability than median risk score. Calculating patient risk scores could be helpful to predict the probability of patients reaching median survival time. Time-dependent ROC curves showed that

our two-DNA methylation signature has stable and good performance in prognosis prediction. The model could distinguish different risk groups with totally varied survival time, which was verified in training and validation datasets. Independent grouping tests showed that, the two-DNA methylation signature possessed independent predictive value from age and sex, and was more suitable to forecast survival time for patients in stage I and G2.

Compared to other acquainted prognostic markers, it is found that the two-DNA methylation signature possesses noticeable higher sensitivity and specificity in predicting ccRCC prognosis. Zhan, et al.<sup>28</sup> reported a five-gene prognostic model to predict prognosis in ccRCC, with an AUC of 0.783. The six-lncRNA signature identified in a previous study<sup>33</sup> was described as a prognostic biomarker of ccRCC, with an AUC of 0.696. In this study, low- and high-risk groups could be well distinguished by way of the two-DNA methylation signature. Also, the two-DNA methylation signature performed well in associated log-rank tests, confirming its independent predictive value for patient survival.

Exploring another database (<http://genome.ucsc.edu/>), we found cg18034859 and cg2419834 are located in the promoter regions of myosin light chain kinase 2 (*MYLK2*) and POU class 4 homeobox 2 (*POU4F2*) genes, respectively, indicating the potential influence of its methylation on gene expression. *POU4F2* is a member of the POU-domain transcription factor family, which is involved in determining the fate of retinal ganglion cells (RGCs) and initiating RGC differentiation.<sup>34</sup> The methylation levels of *POU4F2* were quantified by quantitative methylation-specific PCR using urine sediment DNA from another study, and the researchers generated a two-DNA methylation signature [combination of *POU4F2* and protocadherin 17 (*PCDH17*)] for bladder cancer. The model showed high sensitivity and specificity of 90.00% and 93.96%, respectively, reflecting the ability of detecting bladder cancer effectively among pathologically varied sample groups.<sup>35</sup> Moreover, urinary levels of *POU4F2* hypermethylation were considered to be having a bearing on non-muscle invasive bladder cancer recurrence, which is in good association with our results for a high DNA methylation level of cg2419834 at *POU4F2*.<sup>36</sup> In addition, gene expression levels of *POU4F2* in ccRCC were not detectable based on Gene Expression Profiling Interactive Analysis (GEPIA, <http://gepia.cancer-pku.cn/index.html>) and TCGA database (data not shown). Meanwhile, *MYLK2* encodes a calcium/calmodulin-dependent serine/threonine kinase and is a risk factor for the development of acute lung injuries.<sup>37</sup> Soung, et al.<sup>38</sup> reported that *MYLK2* does not play a significant role in cancer pathogenesis because the kinase domain of *MYLK2* is rarely mutated in common human carcinomas. The gene expression data of *MYLK2* were downloaded from TCGA, while the clinical data of ccRCC were the same. We merged and ranked samples according to the expression levels of *MYLK2* and divided them into high-expression



**Fig. 4.** Stratified analysis. Kaplan-Meier and receiver operating characteristic (ROC) analyses of patients with clear cell renal cell carcinoma of different age. Grouping was based on age at initial diagnosis:  $\leq 60$  ( $n=106$ , 49.53%),  $>60$  ( $n=108$ , 50.47%). (A) Kaplan-Meier analysis with two-sided log-rank test was performed to estimate the differences in survival time between the low-risk and high-risk patients. (B) ROC curves of the two-DNA methylation signature were used to demonstrate the sensitivity and specificity in predicting the survival time of patients. TPR, true positive rate; FPR, false positive rate; AUC, area under the curve of receiver operating characteristics.

**Table 3.** Results of Kaplan-Meier and ROC Analyses Based on the Different Regrouping Methods

Regrouping factors	Group	Sample size	p value	AUC
Age (yr)	$\leq 60$	106	0.001	0.860
	$>60$	108	0.003	0.666
Stage	I	113	0.006	0.714
	II	27	0.290	0.686
	III	49	0.100	0.721
	IV	25	0.012	0.577
Histologic grade*	G2	96	$<0.001$	0.877
	G3	82	0.145	0.628
Gender	Female	76	$<0.001$	0.832
	Male	139	0.005	0.772

ROC, receiver operating characteristic; AUC, area under the curve of ROCs. \*G1 and GX (the degree of differentiation is unknown) were excluded in this study as these tumors may exhibit different biological behaviors. G4 was also excluded due to too few samples.

( $n=89$ ) and low-expression groups ( $n=89$ ). Kaplan-Meier survival analysis methods exhibited a difference in survival time between high-expression and low-expression groups: the survival time of patients in the high-expression group was significantly shorter ( $p=0.028$ ) (Supplementary Fig. 1, only online). The results indicated that *MYLK2* may be an independent prognostic marker of survival for patients with ccRCC.

In conclusion, we discovered that the survival time of patients is greatly affected by a two-DNA methylation signature, which was detected from genome-whole analysis of DNA methylation data for 215 samples with ccRCC, and the significance for patients of varying ages, sex, stages, and histologic grades was confirmed. The two-DNA methylation signature could play an important role as a prognostic marker for forecasting the survival time of patients with ccRCC. In the future, research may want to focus on a functional mechanism for the two-DNA methylation marker and its potential association in carcinogenesis.

## ACKNOWLEDGEMENTS

This work was supported by the Zhejiang Provincial Natural Science Foundation of China (grant number LY17C060002).

## AUTHOR CONTRIBUTIONS

**Conceptualization:** Shanping Shi and Yang Xi. **Data curation:** Shanping Shi, Renjie Zhuo, Qi Liao, and Shazhou Ye. **Formal analysis:** Shanping Shi and Xiaoyue Wu. **Funding acquisition:** Yang Xi. **Investigation:** Shanping Shi, Xiaoyue Wu, and Mingjun Xu. **Methodology:** Shanping Shi, Shazhou Ye, Xiaoyue Wu, and Mingjun Xu. **Project administration:** Yang Xi. **Resources:** Shanping Shi, Shazhou Ye, and Mingjun Xu. **Software:** Shanping Shi and Shazhou Ye. **Supervision:** Yang Xi. **Validation:** Xiaoyue Wu, Shazhou Ye, and Renjie Zhuo. **Visualization:** Shanping Shi and Qi Liao. **Writing—original draft:** Shanping Shi and Yang Xi. **Writing—review & editing:** Shazhou Ye, Xiaoyue Wu, and Mingjun Xu.

## ORCID iDs

Shanping Shi <https://orcid.org/0000-0002-0672-4419>  
 Shazhou Ye <https://orcid.org/0000-0001-5181-5902>  
 Xiaoyue Wu <https://orcid.org/0000-0003-3784-0751>  
 Mingjun Xu <https://orcid.org/0000-0001-5924-7129>  
 Renjie Zhuo <https://orcid.org/0000-0001-7944-8550>  
 Qi Liao <https://orcid.org/0000-0001-8769-9325>  
 Yang Xi <https://orcid.org/0000-0001-8630-5945>

## REFERENCES

- Thiesen HJ, Steinbeck F, Maruschke M, Koczan D, Ziemens B, Hakenberg OW. Stratification of clear cell renal cell carcinoma (ccRCC) genomes by gene-directed copy number alteration (CNA) analysis. *PLoS One* 2017;12:e0176659.
- Ljungberg B, Bensalah K, Canfield S, Dabestani S, Hofmann F, Hora M, et al. EAU guidelines on renal cell carcinoma: 2014 update. *Eur Urol* 2015;67:913-24.
- Young JR, Coy H, Douek M, Lo P, Sayre J, Pantuck AJ, et al. Clear cell renal cell carcinoma: identifying the gain of chromosome 12 on multiphasic MDCT. *Abdom Radiol (NY)* 2017;42:236-41.
- Mehdi A, Riazalhosseini Y. Epigenome aberrations: emerging driving factors of the clear cell renal cell carcinoma. *Int J Mol Sci* 2017;18. pii: E1774.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin* 2017;67:7-30.
- Drucker BJ. Renal cell carcinoma: current status and future prospects. *Cancer Treat Rev* 2005;31:536-45.
- Scélo G, Brennan P. The epidemiology of bladder and kidney cancer. *Nat Clin Pract Urol* 2007;4:205-17.
- Pan Y, Liu G, Zhou F, Su B, Li Y. DNA methylation profiles in cancer diagnosis and therapeutics. *Clin Exp Med* 2018;18:1-14.
- van der Kooi EL, de Greef JC, Wohlgenuth M, Frants RR, van Asseldonk RJ, Blom HJ, et al. No effect of folic acid and methionine supplementation on D4Z4 methylation in patients with facioscapulo-humeral muscular dystrophy. *Neuromuscul Disord* 2006;16:766-9.
- Hao X, Luo H, Krawczyk M, Wei W, Wang W, Wang J, et al. DNA methylation markers for diagnosis and prognosis of common cancers. *Proc Natl Acad Sci U S A* 2017;114:7414-9.
- Dubrowskaja N, Gebauer K, Peters I, Hennenlotter J, Abbas M, Scherer R, et al. Neurofilament Heavy polypeptide CpG island methylation associates with prognosis of renal cell carcinoma and prediction of antivasular endothelial growth factor therapy response. *Cancer Med* 2014;3:300-9.
- Wang L, Gao W, Hu F, Xu Z, Wang F. MicroRNA-874 inhibits cell proliferation and induces apoptosis in human breast cancer by targeting CDK9. *FEBS Lett* 2014;588:4527-35.
- Zhang L, Yan DL, Yang F, Wang DD, Chen X, Wu JZ, et al. DNA methylation mediated silencing of microRNA-874 is a promising diagnosis and prognostic marker in breast cancer. *Oncotarget* 2017;8:45496-505.
- Li W, Zhang X, Lu X, You L, Song Y, Luo Z, et al. 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Res* 2017;27:1243-57.
- Jeschke J, Bizet M, Desmedt C, Calonne E, Dedeurwaerder S, Gaudraud S, et al. DNA methylation-based immune response signature improves patient diagnosis in multiple cancers. *J Clin Invest* 2017;127:3090-102.
- Zhang W, Yan W, You G, Bao Z, Wang Y, Liu Y, et al. Genome-wide DNA methylation profiling identifies ALDH1A3 promoter methylation as a prognostic predictor in G-CIMP- primary glioblastoma. *Cancer Lett* 2013;328:120-5.
- Shukla S, Pia Patric IR, Thinagararjan S, Srinivasan S, Mondal B, Hegde AS, et al. A DNA methylation prognostic signature of glioblastoma: identification of NPTX2-PTEN-NF-κB nexus. *Cancer Res* 2013;73:6563-73.
- Sandoval J, Mendez-Gonzalez J, Nadal E, Chen G, Carmona FJ, Sayols S, et al. A prognostic DNA methylation signature for stage I non-small-cell lung cancer. *J Clin Oncol* 2013;31:4140-7.
- Evelönn EA, Landfors M, Haider Z, Köhn L, Ljungberg B, Roos G, et al. DNA methylation associates with survival in non-metastatic clear cell renal cell carcinoma. *BMC Cancer* 2019;19:65.
- Peters I, Dubrowskaja N, Hennenlotter J, Antonopoulos WI, Von Klot CAJ, Tezval H, et al. DNA methylation of neural EGFL like 1 (NELL1) is associated with advanced disease and the metastatic state of renal cell cancer patients. *Oncol Rep* 2018;40:3861-8.
- Dagher J, Delahunt B, Rioux-Leclercq N, Egevad L, Srigley JR, Coughlin G, et al. Clear cell renal cell carcinoma: validation of World Health Organization/International Society of Urological Pathology grading. *Histopathology* 2017;71:918-25.
- Delahunt B, Chevillat C, Martignoni G, Humphrey PA, Magi-Galluzzi C, McKenney J, et al. The International Society of Urological Pathology (ISUP) grading system for renal cell carcinoma and other prognostic parameters. *Am J Surg Pathol* 2013;37:1490-504.
- Cho HJ, Yu A, Kim S, Kang J, Hong SM. Robust likelihood-based survival modeling with microarray data. *J Stat Softw* 2009;29:1-16.
- Jung EJ, Lee HJ, Kwak C, Ku JH, Moon KC. Young age is independent prognostic factor for cancer-specific survival of low-stage clear cell renal cell carcinoma. *Urology* 2009;73:137-41.
- Ricketts CJ, Linehan WM. Gender specific mutation incidence and survival associations in Clear Cell Renal Cell Carcinoma (ccRCC). *PLoS One* 2015;10:e0140257.
- Ebru T, Fulya OP, Hakan A, Vuolat YC, Necdet S, Nuray C, et al. Analysis of various potential prognostic markers and survival data in clear cell renal cell carcinoma. *Int Braz J Urol* 2017;43:440-54.
- Chevillat JC, Blute ML, Zincke H, Lohse CM, Weaver AL. Stage pT1 conventional (clear cell) renal cell carcinoma: pathological features associated with cancer specific survival. *J Urol* 2001;166:453-6.
- Zhan Y, Guo W, Zhang Y, Wang Q, Xu XJ, Zhu L. A five-gene signature predicts prognosis in patients with kidney renal clear cell carcinoma. *Computational and Mathematical Methods in Medicine* 2015;2015. Article ID 842784.

29. Liang B, Zhao J, Wang X. A three-microRNA signature as a diagnostic and prognostic marker in clear cell renal cancer: an In Silico analysis. *PLoS One* 2017;12:e0180660.
30. Song J, Peng J, Zhu C, Bai G, Liu Y, Zhu J, et al. Identification and validation of two novel prognostic lncRNAs in kidney renal clear cell carcinoma. *Cell Physiol Biochem* 2018;48:2549-62.
31. Xu D, Zhang S, Zhang S, Liu H, Li P, Yu L, et al. NOD2 maybe a biomarker for the survival of kidney cancer patients. *Oncotarget* 2017; 8:101489-99.
32. Dai W, Teodoridis JM, Zeller C, Graham J, Hersey J, Flanagan JM, et al. Systematic CpG islands methylation profiling of genes in the wnt pathway in epithelial ovarian cancer identifies biomarkers of progression-free survival. *Clin Cancer Res* 2011;17:4052-62.
33. Zuo S, Wang L, Wen Y, Dai G. Identification of a universal 6-lncRNA prognostic signature for three pathologic subtypes of renal cell carcinoma. *J Cell Biochem* 2018 Oct 30 [Epub]. Available at: <https://doi.org/10.1002/jcb.28012>.
34. Wu F, Kaczynski TJ, Sethuramanujam S, Li R, Jain V, Slaughter M, et al. Two transcription factors, Pou4f2 and Isl1, are sufficient to specify the retinal ganglion cell fate. *Proc Natl Acad Sci U S A* 2015; 112:E1559-68.
35. Maskell LJ, Mahadeo AV, Budhram-Mahadeo VS. POU4F2/Brn-3b transcription factor is associated with survival and drug resistance in human ovarian cancer cells. *Oncotarget* 2018;9:36770-9.
36. Reinert T, Borre M, Christiansen A, Hermann GG, Ørntoft TF, Dyrskjød L. Diagnosis of bladder cancer recurrence based on urinary levels of EOMES, HOXA9, POU4F2, TWIST1, VIM, and ZNF154 hypermethylation. *PLoS One* 2012;7:e46297.
37. Christie JD, Ma SF, Aplenc R, Li M, Lanken PN, Shah CV, et al. Variation in the myosin light chain kinase gene is associated with development of acute lung injury after major trauma. *Crit Care Med* 2008;36:2794-800.
38. Soung YH, Lee JW, Kim SY, Nam SW, Park WS, Lee JY, et al. Mutational analysis of the kinase domain of MYLK2 gene in common human cancers. *Pathol Res Pract* 2006;202:137-40.