

LETTER TO EDITOR

Four methylation-driven genes may be prognostic biomarkers in clear cell renal carcinoma

Dear Editor,

Clear cell renal carcinoma (ccRCC) is one of the most common urinary system carcinomas with a 5-year survival about 44–69%.¹ There is an urgent need to explore effective biomarkers to predict the survival of patients with ccRCC. Epigenetic mechanisms as an important part in the development of tumorigenesis mainly include heritable and reversible DNA methylation. DNA methylation could negatively regulate gene expression in the development of tumors, and be also used as biomarkers for prognostic prediction. It is important and urgent to explore the abnormality of DNA methylation-associated genes. Recently, there are more and more studies based on the Cancer Genome Atlas (TCGA) database with large amounts of genetic and clinical profiles.² In addition, MethylMix is used to identify methylation status according to the β model.³

The present study aims at understanding gene expression and methylation as well as the association with clinical information in patients with ccRCC and screening the methylation gene-based biomarkers for prognostic prediction. We obtained 524 complete clinical data and mRNA data of patients with ccRCC (539 cancer and 72 para-carcinoma tissues). Note that 2332 differential expression genes were considered as discriminatively expressed (the absolute value of fold changes is > 2 , and the value of false discovery rate is < 0.01). The methylated data were obtained from TCGA and Human Methylation 450k Illumina Infinium methylation arrays. The methylated data included 160 normal samples and 325 tumor samples. The aberrant methylation and expression of genes in ccRCC were explored by LIMMA package. Then the aberrant genes were further explored by the MethylMix package ($|\log_{2}FC| > 0$, $P < .05$, $|\text{Cor}| > 0.3$). We also explored 168 methylation-driven genes (Figure 1A,B), of which 75 genes (44.6%) were hypermethylated and the other genes were hypomethylated. The expression of genes with significant difference in tumor samples as compared with the normal were investigated, of which 51 abnormally methylated and differentially expressed genes in ccRCC (Figure 1C).

We explored the function of these genes by the Discovery 6.8 (DAVID) (<https://david.ncifcrf.gov/>) and constructed KEGG pathways using ConsensusPathDB. The results showed the function of these genes were cell periphery, platelet activation, immune response, cell-cell junction, and others. The KEGG pathways of these genes were rheumatoid arthritis, platelet activation, hepatitis C, and so on (Figure 1D,E). To analyze the function of abnormally methylated and differentially expressed genes, we calculated the relationship between the 51 genes and overall survival in the training set. The results of univariate cox analysis showed that eight genes were correlated with overall survival ($P < .05$). Of eight genes in multivariate Cox regression, four genes including Linker for Activation of T Cells (LAT), Keratin 17 (KRT17), MAGE Family Member C2 (MAGEC2), and Ribonuclease A Family Member 2 (RNASE2) were independent risk factors effecting existence ($P < .001$; Table S1). The 4-gene expression-based survival risk score was calculated as 0.991. We also compared the survival rate between patients with low and high risks according to the risk score and the Kaplan-Meier curves. The patients with high risk had shorter survival time than those with low risk ($P < .001$). The 5-year survival times of patients with low and high risks were 75.6 (95% CI = 0.67--0.825) and 50.7 (95% CI = 0.4115--0.625), respectively. The area under ROC curve (AUC) was 0.708 (Figure 1F).

We further evaluated the value of the four target genes in the prognostic significance. The 5-year survival time of patients with high risk ($n = 168$; 52.7 (95% CI = 0.446--0.624)) was shorter than those with low risk ($n = 97$; 86.2 (95% CI = 0.789--0.942), $P < .001$), respectively. The AUC was 0.732 (Figure 2A). Figure 2B demonstrated the result of the four genes in the entire group. The correlation among gene methylation and expression with patient survival using the survival package of R. KRT17, LAT, MAGEC2, and RNASE2 were significantly correlated with overall survival ($P < .05$, Figure 2C). The correlations between methylation sites and gene expressions were also analyzed. The absolute

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.

© 2020 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics

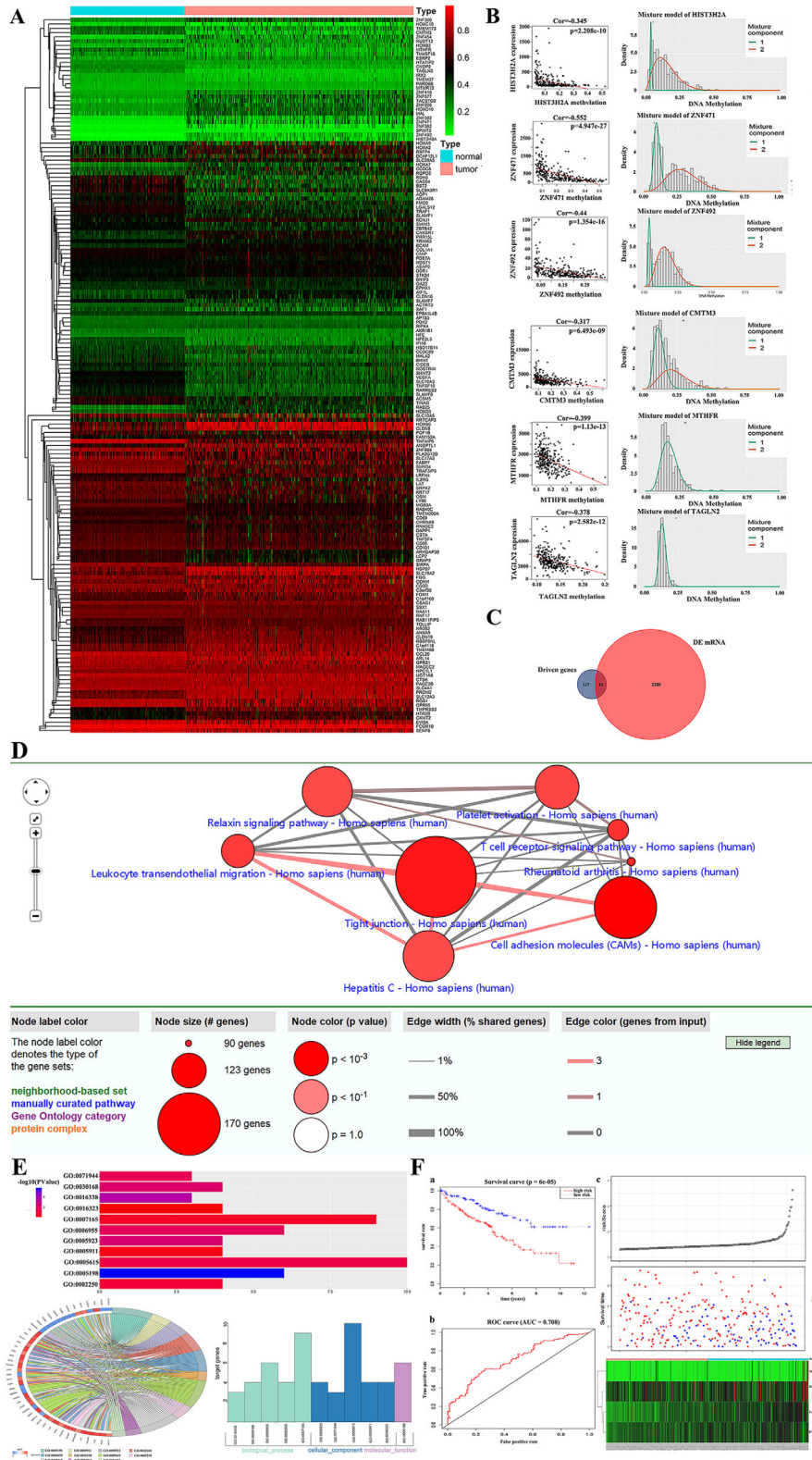


FIGURE 1 A, Heat maps of ccRCC-related aberrant methylated genes. The color from green to red shows a trend from low expression to high expression. ccRCC: clear cell renal carcinoma. B, The top hypermethylated and top hypomethylated genes of tumor samples compared with normal samples. The x-axis is the methylation degree, the y-axis is the number of methylation samples, the histogram represents the methylation distribution of cancer group, and the curve is the simulated trend curve of the methylation distribution of cancer group. The black horizontal line above the figure is the methylation level distribution of the normal sample. C, Fifty-one abnormally methylated and differentially expressed genes in ccRCC. D, The KEGG pathways of these genes were explored by ConsensusPathDB. The enriched pathways were cell adhesion molecules, Hepatitis C, leukocyte

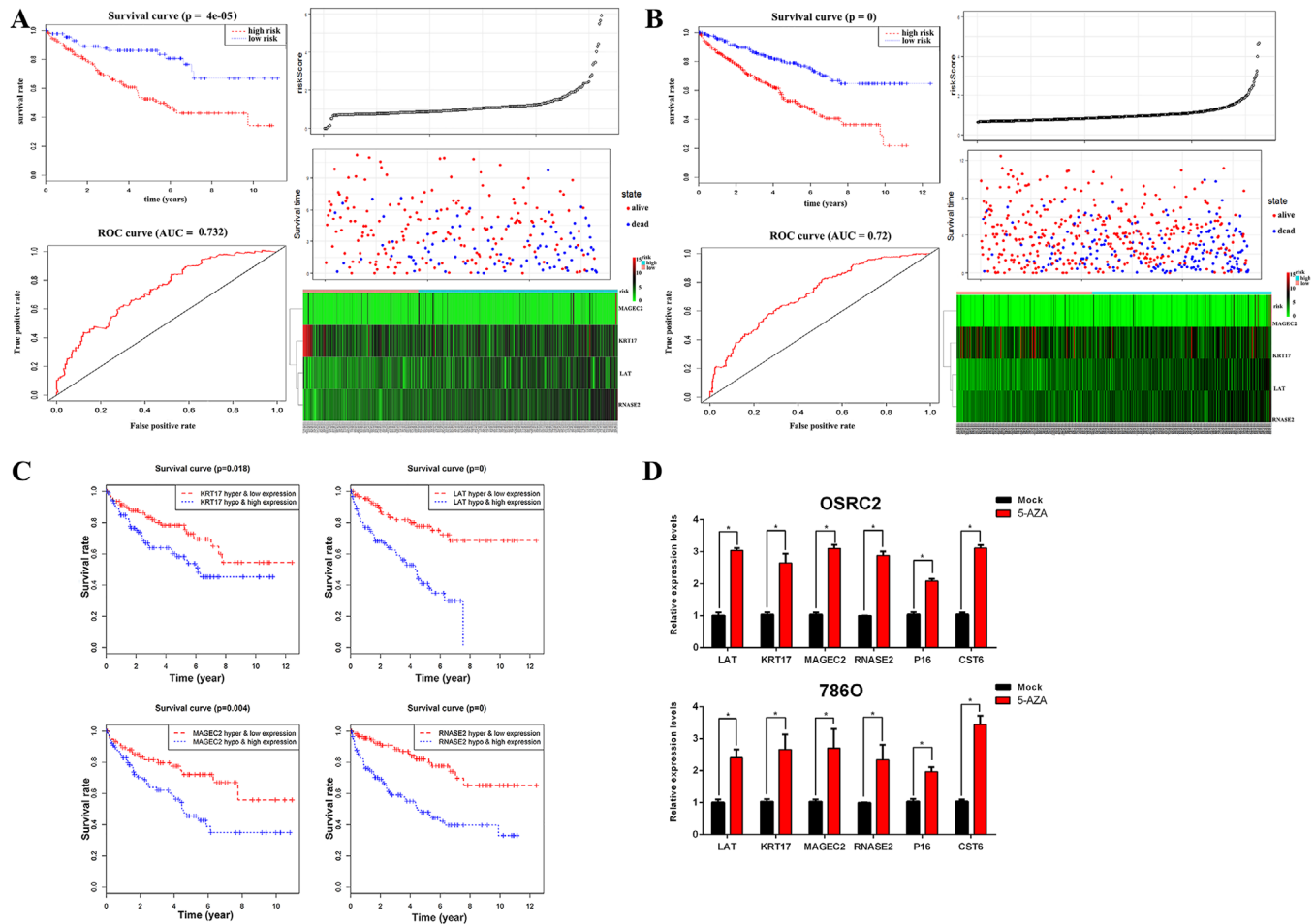


FIGURE 2 A, Identification and performance evaluation of the 4-gene signature in the testing dataset. a, Kaplan-Meier survival curve analysis for overall survival of clear cell renal cell carcinoma patients using the 4-gene signature in the testing dataset; b, ROC curve analysis of the 4-gene signature in the testing dataset; c, the distributions of the survival status and expression profiles of the four genes of patients in the testing dataset. B, Identification and performance evaluation of the 4-gene signature in the entire dataset. a, Kaplan-Meier survival curve analysis for overall survival of clear cell renal cell carcinoma patients using the 4-gene signature in the entire dataset; b, ROC curve analysis of the 4-gene signature in the entire dataset; c, the distributions of the survival status and expression profiles of the four genes of patients in the entire dataset. C, Kaplan-Meier survival curves for the combination of the four genes (LAT, KRT17, MAGEC2, and RNASE2) methylation and expression. D, The expression of four genes (LAT, KRT17, MAGEC2, and RNASE2) increased significantly in both OSRC2 and 786O cell lines following treatment with 5-Aza (8 μ M) for 72 hours. We used P16 and CST6 as positive controls in these assays. * $P < .05$

value of $Cor > 0.2$ was the cut off value. The expression of LAT, MAGEC2, and RNASE2 were associated with the methylation level of multiple sites (Figure S10). To explore the correlations between methylation and expression, OSRC2 and 786O cells of ccRCC were treated by 5-Aza (8 μ M) to

explore the relationship between the blocking of promoter in methylation and expression. The expression of LAT, KRT17, MAGEC2, and RNASE2 genes increased significantly in both cells after the treatment, as compared with P16 and CST6 as controls (Figure 2D).

transendothelial migration, platelet activation, relaxin signaling, rheumatoid arthritis, T-cell receptor signaling, and tight junction. E, The function of these genes was explored by DAVID. GO:0071944-cell periphery, GO:0030168-platelet activation, GO:0016338-calcium independent cell-cell adhesion, GO:0016323-basolateral plasma membrane, GO:0007165-signal transduction, GO:0006955-immune response, GO:0005923-bicellular tight junction, GO:0005911-cell-cell junction, GO:0005615-extracellular space, GO:0005198-structural molecule activity, and GO:0002250-adaptive immune response. F, Identification and performance evaluation of the four-gene signature in the training dataset. a, Kaplan-Meier survival curve analysis for overall survival of clear cell renal cell carcinoma patients using the 4-gene signature in the training dataset; b, ROC curve analysis of the 4-gene signature in the training dataset; c, the distributions of the survival status and expression profiles of the four genes of patients in the training dataset

In conclusion, our study explored the methylation-driven and differently expressed genes in ccRCC and found four genes with potentials of prognostic biomarkers for ccRCC patients.

ACKNOWLEDGMENTS

This work was supported by Haiyan Foundation of Harbin Medical University Cancer Hospital (No. JJZD2018-04 to Q.Y.X.), Hei Long Jiang Postdoctoral (LBH -Z18224 to X.Y.W.), China Postdoctoral Science Foundation (2018M641861 to X.Y.W.), Fundamental Research Funds for the Provincial Universities in Heilongjiang (2017LCZX94 to X.Y.W.), and National Natural Science Fund of China (No. 81903120 to H.Y. and 81902317 to X.Y.W.).

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

All the data and materials are available upon request from the authors.

AUTHOR CONTRIBUTIONS

H. Yin and X.Y. Wang performed experiments. H. Yin and H.Y. Zhang wrote the manuscript. Q.Y. Xu designed the study. All authors approved the final manuscript.

Hang Yin¹
Haiyang Zhang²
Xiaoyuan Wang³
Qingyong Xu¹

¹Department of Radiation Oncology, Harbin Medical University Cancer Hospital, Harbin, China

²Key Laboratory of Hepatosplenic Surgery, Ministry of Education, The First Affiliated Hospital of Harbin Medical University, Harbin, China

³Department of Oncology, Harbin Medical University Cancer Hospital, Harbin, China

Correspondence

Xiaoyuan Wang and Qingyong Xu, Department of Oncology, Harbin Medical University Cancer Hospital, Harbin, Heilongjiang Province, China
Email: 346650985@qq.com, xqingyonghmu@163.com

REFERENCES

1. Motzer RJ, Tannir NM, McDermott DF, et al. Nivolumab plus ipilimumab versus sunitinib in advanced renal-cell carcinoma. *N Engl J Med.* 2018; 378: 1277-1290.
2. Hutter C, Zenklusen JC. The cancer genome atlas: creating lasting value beyond its data. *Cell.* 2018; 173: 283-285.
3. Cedoz PL, Prunello M, Brennan K, Gevaert O. MethylMix 2.0: an R package for identifying DNA methylation genes. *Bioinformatics.* 2018; 34: 3044-3046.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.