

The promising novel biomarkers and candidate small molecule drugs in kidney renal clear cell carcinoma: Evidence from bioinformatics analysis of high-throughput data

Bo Zhang^{1,2*} | Qiong Wu^{1,2*} | Ziheng Wang^{2,3*} | Ran Xu¹ | Xinyi Hu³ |
Yidan Sun⁴ | Qihong Wang² | Fei Ju² | Shiqi Ren³ | Chenlin Zhang⁵ | Lin Qin⁶ |
Qianqian Ma⁷ | You Lang Zhou²

¹Medical School of Nantong University, Nantong, P.R. China

²The Hand Surgery Research Center, Department of Hand Surgery, Affiliated Hospital of Nantong University, Nantong, P.R. China

³Department of Medicine, Nantong University Xinling college, Nantong, P.R. China

⁴Department of Oncology, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin, P.R. China

⁵Department of Spine, Chinese medicine hospital, Wuxi, P.R. China

⁶Department of Urology, The First people's Hospital of Taicang City, Taicang Affiliated Hospital of Soochow University, Suzhou, P.R. China

⁷Emergency office, Wuxi Center for disease control and prevention, Wuxi, P.R. China

Correspondence

You Lang Zhou, The Hand Surgery Research Center, Department of Hand Surgery, Affiliated Hospital of Nantong University, Nantong, P.R. China.
Email: zhouyoulang@ntu.edu.cn

Funding information

This study was funded by the National Natural Science Foundation of China (No. 81401796), Jiangsu youth medical talent project (QNRC2016702) and the Social Development Foundation of Nantong City (MS12017006-2).

Abstract

Background: Kidney renal clear cell carcinoma (KIRC) is the most common subtype of renal tumor. However, the molecular mechanisms of KIRC pathogenesis remain little known. The purpose of our study was to identify potential key genes related to the occurrence and prognosis of KIRC, which could serve as novel diagnostic and prognostic biomarkers for KIRC.

Methods: Three gene expression profiles from gene expression omnibus database were integrated to identify differential expressed genes (DEGs) using limma package. Enrichment analysis and PPI construction for these DEGs were performed by bioinformatics tools. We used Gene Expression Profiling Interactive Analysis (GEPIA) database to further analyze the expression and prognostic values of hub genes.

The GEPIA database was used to further validate the bioinformatics results. The Connectivity Map was used to identify candidate small molecules that could reverse the gene expression of KIRC.

Results: A total of 503 DEGs were obtained. The PPI network with 417 nodes and 1912 interactions was constructed. GO and KEGG pathway analysis revealed that these DEGs were most significantly enriched in excretion and valine, leucine, and isoleucine degradation, respectively. Six DEGs with high degree of connectivity (*ACAA1*, *ACADSB*, *ALDH6A1*, *AUH*, *HADH*, and *PCCA*) were selected as hub

*These authors contributed equally to this study.

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. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals, Inc.

genes, which significantly associated with worse survival of patients. Finally, we identified the top 20 most significant small molecules and pipemidic acid was the most promising small molecule to reverse the KIRC gene expression.

Conclusions: This study first uncovered six key genes in KIRC which contributed to improving our understanding of the molecular mechanisms of KIRC pathogenesis. *ACAA1*, *ACADSB*, *ALDH6A1*, *AUH*, *HADH*, and *PCCA* could serve as the promising novel biomarkers for KIRC diagnosis, prognosis, and treatment.

KEYWORDS

bioinformatics analysis, candidate small molecules, kidney renal clear cell carcinoma, novel biomarkers

1 | INTRODUCTION

Renal cell carcinoma (RCC), also known as renal cell adenocarcinoma, accounts for more than 90% of primary renal neoplasms and pelvic cancer and ranks among the top 10 malignant tumors in both men and women. The incidence and mortality of RCC have been increasing over the past two decades. (Bray et al., 2018; Escudier et al., 2016; Hsieh, Le, Cao, Cheng, & Creighton, 2018; Srigley et al., 2013) Kidney renal clear cell carcinoma (KIRC) is the most common subtype of RCC (70%–75%) and represents one of the most malignant genitourinary cancers. Nearly a quarter of patients were diagnosed with distal metastasis and advanced regional disease. Although a major breakthrough has been made in the molecular mechanisms and therapeutic strategies for KIRC, the long-term overall survival of patients has not been significantly improved, especially in patients with metastatic disease, with 5-year survival <32%. (Linehan, 2012; Shuch et al., 2015) Therefore, identifying the promising novel diagnostic and prognostic biomarkers is urgently demanded, which will contribute to enhance our understanding of KIRC initiation and progression and promote early detection of KIRC. A previous study showed that Polybromo-1 plays an important role in the proliferation and metastasis of KIRC by regulating multiple metabolic pathways including PI3K signaling pathway, glucose uptake pathway, and hypoxia response genes. (Chowdhury et al., 2016) In the past decade, with the rapid development of microarray and RNA-sequencing technology, more and more biomarkers of tumor initiation, progression, and prognosis have been identified using bioinformatics analysis. (Sun et al., 2019; Yan et al., 2019) In this study, we integrated three microarray datasets from gene expression omnibus (GEO) database (GSE781, GSE6344, and GSE100666) to identify the differentially expressed genes (DEGs) between KIRC and adjacent normal tissues, aiming to explore and determine the promising novel biomarkers associated with pathogenesis and prognosis of KIRC. Meanwhile, we revealed some candidate small molecule drugs that could reverse the gene expression of KIRC based on the CMap database. The present study uncovered six novel biomarkers that

were of great value in the diagnosis and prognosis of KIRC patients, which may contribute in revealing the molecular mechanisms of KIRC occurrence and progression. Figure 1 showed the workflow of our study.

2 | MATERIALS AND METHODS

2.1 | Data resources

We conducted a systematic search of GEO database for relevant microarray datasets in order to investigate the differential gene expression between KIRC and normal samples. GEO database (<http://www.ncbi.nlm.nih.gov/geo/>) is widely used to archive high-throughput microarray experimental data, original series, and platform records. In total, the gene expression profiles of GSE781 (Lenburg et al., 2003), GSE6344 (Gumz et al., 2007; Tun et al., 2010), and GSE100666 were downloaded from GEO database. These RNA profiles were provided on GPL96 (Affymetrix Human Genome U133A Array) and platforms GPL16951 (Phalanx Human OneArray Ver. 6 Release 1).

2.2 | Identification of DEGs

The original CEL files were downloaded and classified as KIRC and normal groups. The affy package of Bioconductor (<http://www.bioconductor.org/>) was used for data standardization and transforming raw data into expression values. The significance analysis of the empirical Bayes methods within limma package was applied to identify DEGs between KIRC samples and normal tissue sample. (Ritchie et al., 2015) The genes with the following cutoff criteria were considered as the significant DEGs: $p < 0.05$ and $|\log_2FC| > 1$.

2.3 | Functional enrichment analysis

We performed gene ontology (GO) enrichment analysis for these overlap DEGs. GO enrichment analysis is a common useful method to analyze biological process (BP), molecular functions (MF), and cellular components (CC) of genes

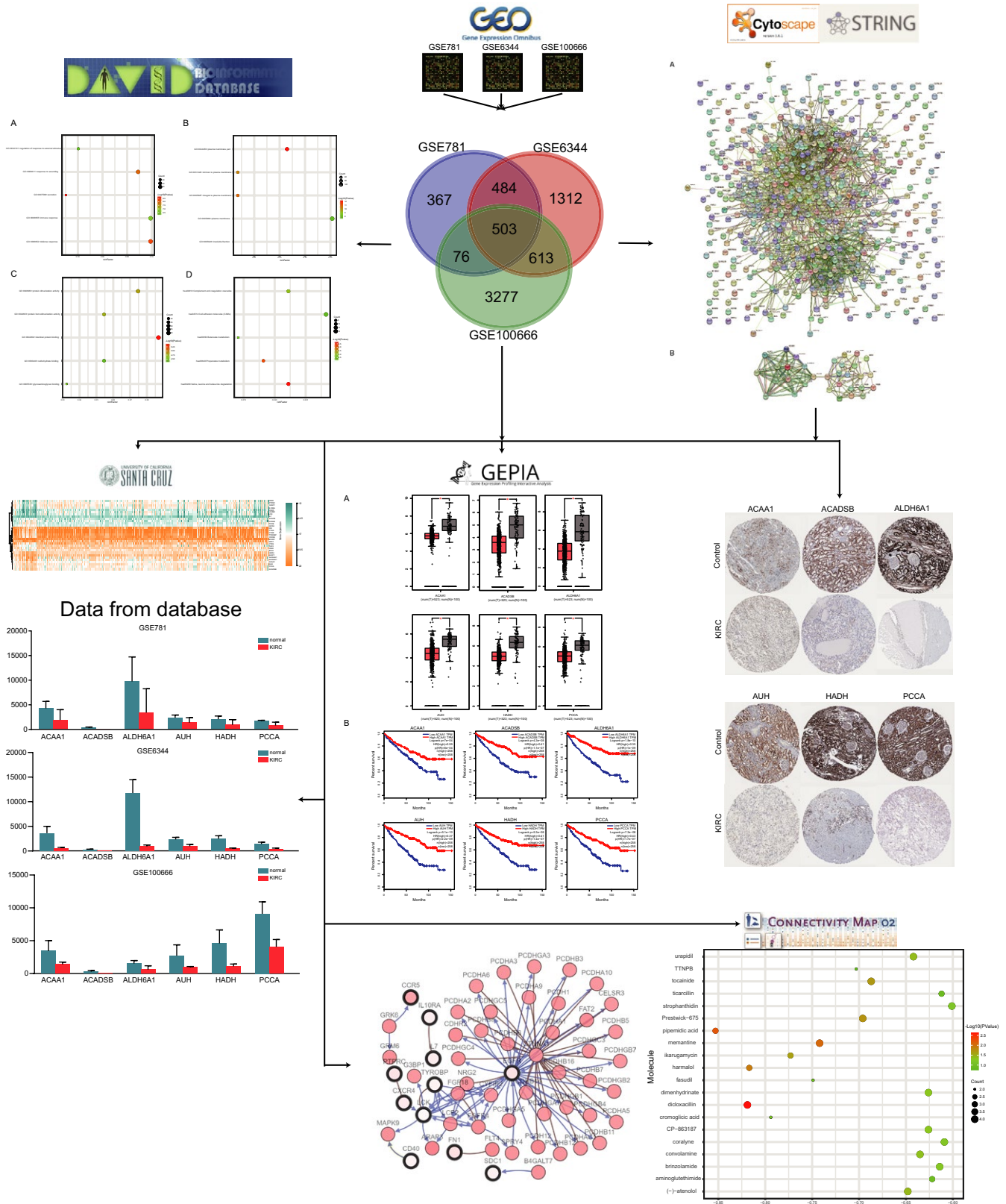


FIGURE 1 The workflow of this study for identifying key genes and pathways in KIRC

and gene products. KEGG pathway enrichment analysis was used to identify the potential functional and metabolic pathways associated with overlap DEGs. KEGG is a

collection of databases that store a large number of information about genomes, biological pathways, diseases, chemical substances, and drugs. We performed GO enrichment

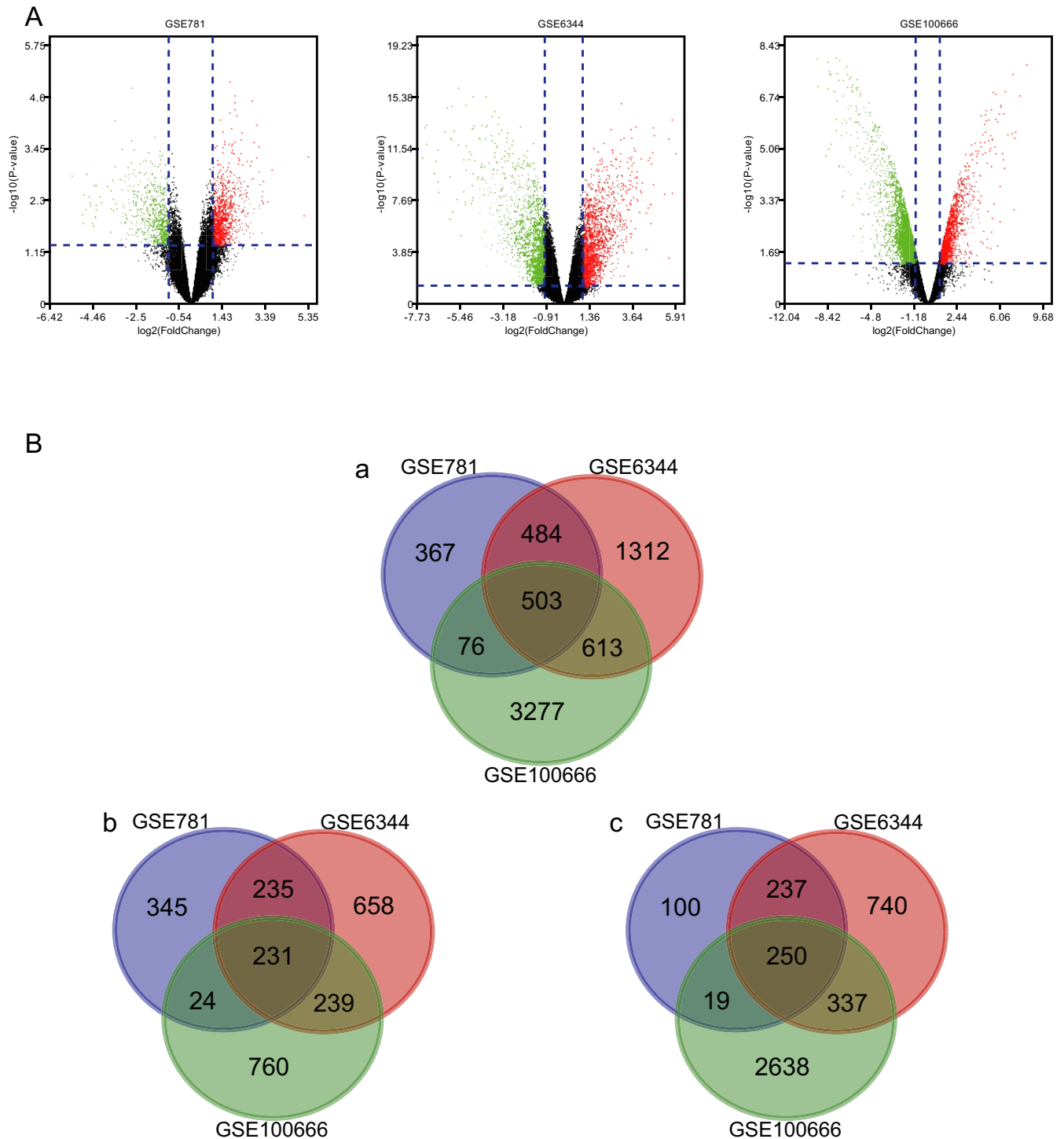


FIGURE 2 (A) Volcano plot of gene expression profile data between KIRC and normal tissues in each dataset. Red dots: significantly upregulated genes in KIRC; Green dots: significantly downregulated genes in KIRC; Black dots: nondifferentially expressed genes. $p < 0.05$ and $\log_2 \text{FC} > 1$ were considered as significant. (B) a. Venn diagram of 503 overlap DEGs from GSE781, GSE6344, and GSE100666 datasets. b. Upregulated overlap DEGs; c. Downregulated overlap DEGs

and KEGG pathway analysis by using DAVID (Database for Annotation Visualization and Integrated Discovery). DAVID is a commonly used online biological information database that provides comprehensive functional annotation and pathway interpretations (version 6.7)(<http://david.ncifcrf.gov>).

(Ashburner et al., 2000; Dennis et al., 2003; Gene Ontology Consortium, 2006; Huang da, Sherman, & Lempicki, 2009; Papanicolaou, Gebauer-Jung, Blaxter, Owen McMillan, & Jiggins, 2008) $p < 0.05$ was considered statistically significant.

2.4 | Protein–protein interaction (PPI) network construction and module analysis

The Search Tool for the Retrieval of Interacting Genes database (STRING, <https://string-db.org/>) was used to construct the PPI network.(Szkarczyk et al., 2015) The interactions with a combined score >0.4 were considered significant. The PPI network was visualized through the Cytoscape software. (Smoot, Ono, Ruschinski, Wang, & Ideker, 2011) Subsequently, Molecular Complex Detection (MCODE) plug-in of the Cytoscape software was applied to screen significant modules from the PPI network. The degree cutoff = 10, node score cutoff = 0.2, k-core = 2, and max. depth = 100 were used as selection criterion.(Bandettini et al., 2012) We also performed Go and KEGG pathway enrichment analysis for the most significant modules. In addition, the Networks Gene Oncology tool (BiNGO) plugin

in Cytoscape was used to perform and visualize the biological process analysis of the hub genes.(Maere, Heymans, & Kuiper, 2005)

2.5 | Analysis and validation of hub genes

The hierarchical clustering of module genes was constructed by UCSC Cancer Genomics Browser (<http://genome-cancer.ucsc.edu>). A network of hub genes and their coexpression genes was established by cBioPortal online platform (<http://www.cbioportal.org>). The Gene Expression Profiling Interactive Analysis (GEPIA) was used to further confirm the reliability of the expression level of hub genes and their prognostic value.(Cerami et al., 2012; Gao et al., 2013; Tang et al., 2017) The hazard ratio (HR) with 95% confidence intervals (CI) of overall survival was calculated. The Kaplan–Meier curve and boxplot were visualized to present

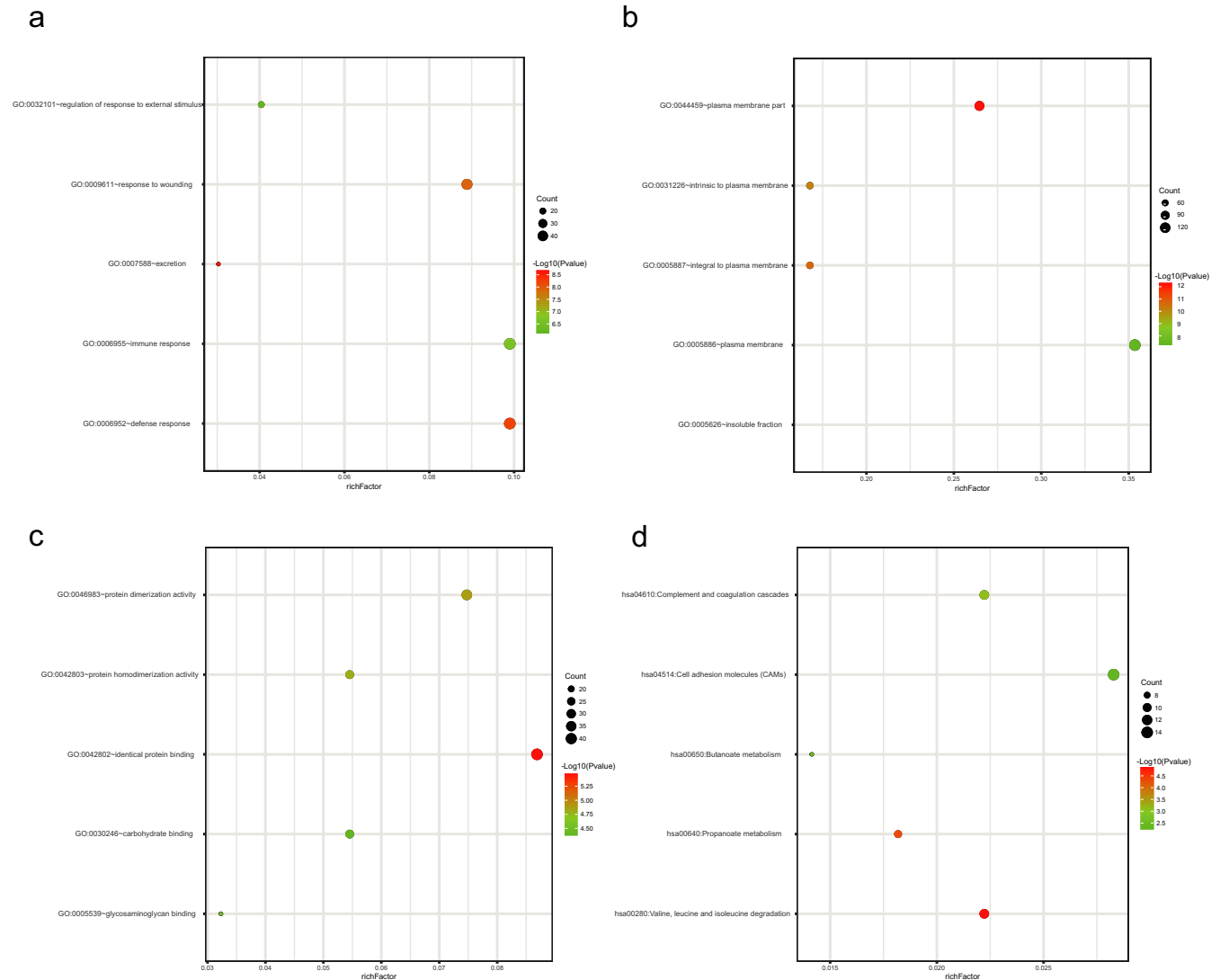


FIGURE 3 Functional and signaling pathway analysis of the overlapped DEGs in KIRC. (a) Biological processes (b) Cellular components (c) Molecular function (d) KEGG pathway

these relationships between gene expression and patients' prognosis. The human protein atlas (HPA, www.proteinatlas.org) database was used to determine the protein expression of the hub genes between KIRC and normal tissues based on immunohistochemistry.

2.6 | Identification of small molecules

We compared the overlap DEGs with those in the Connectivity Map database (CMap, <http://www.broadinstitute.org/cmap/>) to predict potential small active molecular that could reverse the current biological status of KIRC.(Lamb et al., 2006) First, the overlap DEGs were divided into upregulated and downregulated groups. Then different expression significant probesets from each group were selected for gene set enrichment analysis. Finally, the enrichment scores ranging from -1 to $+1$ were calculated. A positive connectivity value (closer to $+1$) indicated that the corresponding small molecules could induce the gene expression of KIRC, whereas a negative connectivity value (closer to -1) indicated greater similarity among the genes and the small molecules that could reverse the KIRC cell status.

3 | RESULTS

3.1 | Identification of DEGs in KIRC

After the standardization of GSE781, GSE6344, and GSE100666 datasets, a total of 503 overlap DEGs expressed in KIRC tissues were identified with the Limma package. The volcano plot of DEGs of KIRC in each dataset was presented in Figure 2A. The Venn diagrams showed the 503 overlap DEGs among the three datasets (Figure 2Ba), consisting of 231 upregulated genes (Figure 2Bb) and 250 downregulated genes (Figure 2Bc).

3.2 | Enrichment analyses

To better understand the biological functions of overlap DEGs among the three datasets, we performed GO function and KEGG pathway enrichment analysis using DAVID. For biological processes, GO analysis results indicated that upregulated and downregulated DEGs were significantly enriched in excretion, defense response, response to wounding, immune response, and regulation of response to external

Category	Term	<i>p</i> value
GOTERM_BP_FAT	GO:0007588~excretion	2.32E-09
GOTERM_BP_FAT	GO:0006952~defense response	6.27E-09
GOTERM_BP_FAT	GO:0009611~response to wounding	1.32E-08
GOTERM_BP_FAT	GO:0006955~immune response	2.22E-07
GOTERM_BP_FAT	GO:0032101~regulation of response to external stimulus	6.55E-07
GOTERM_CC_FAT	GO:0044459~plasma membrane part	9.91E-13
GOTERM_CC_FAT	GO:0005887~integral to plasma membrane	3.18E-11
GOTERM_CC_FAT	GO:0031226~intrinsic to plasma membrane	9.96E-11
GOTERM_CC_FAT	GO:0005886~plasma membrane	5.00E-08
GOTERM_CC_FAT	GO:0005626~insoluble fraction	2.16E-07
GOTERM_MF_FAT	GO:0042802~identical protein binding	3.31E-06
GOTERM_MF_FAT	GO:0,046,983 ~ protein dimerization activity	1.33E-05
GOTERM_MF_FAT	GO:0042803~protein homodimerization activity	1.54E-05
GOTERM_MF_FAT	GO:0005539~glycosaminoglycan binding	2.85E-05
GOTERM_MF_FAT	GO:0030246~carbohydrate binding	4.20E-05
KEGG_PATHWAY	hsa00280:Valine, leucine and isoleucine degradation	1.49E-05
KEGG_PATHWAY	hsa00640:Propanoate metabolism	5.26E-05
KEGG_PATHWAY	hsa04610:Complement and coagulation cascades	7.77E-04
KEGG_PATHWAY	hsa00650:Butanoate metabolism	0.003257
KEGG_PATHWAY	hsa04514:Cell adhesion molecules (CAMs)	0.005074

TABLE 1 Functional and pathway enrichment analysis of the overlap DEGs

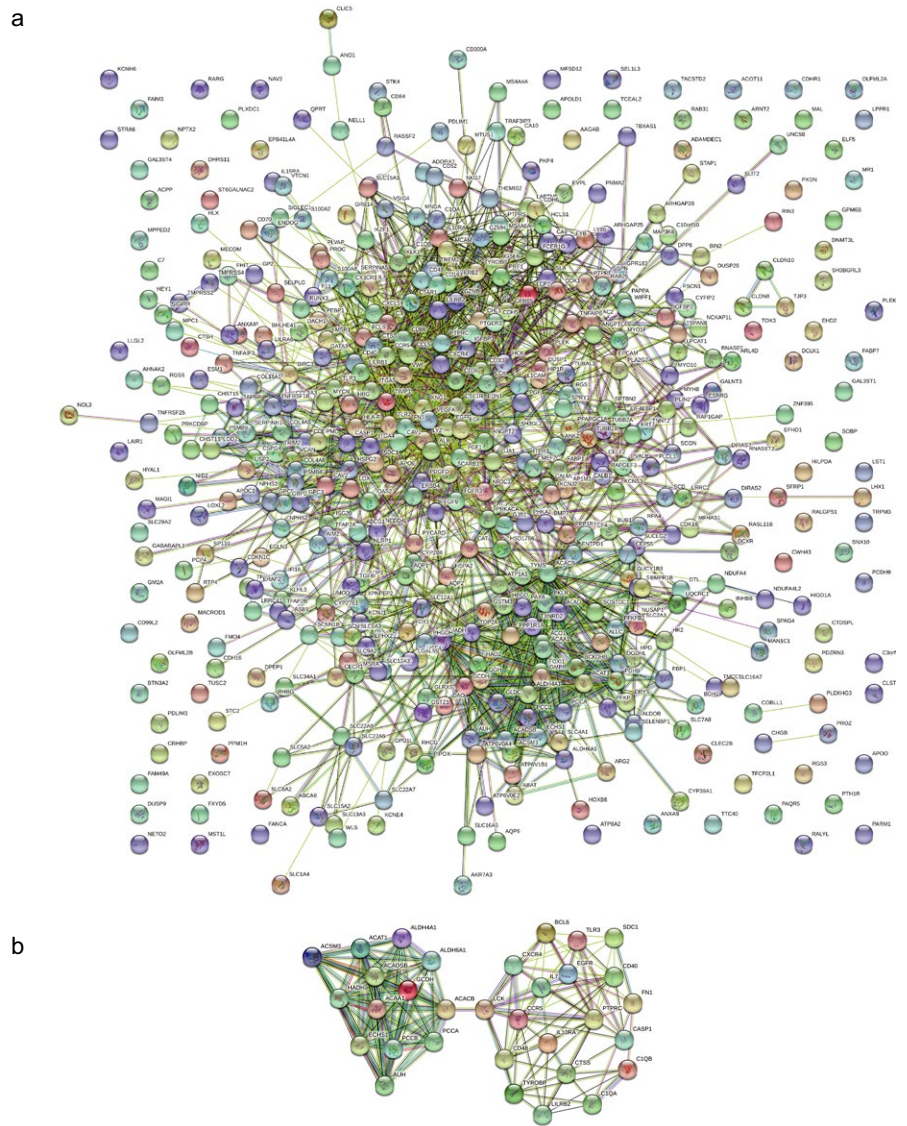


FIGURE 4 Protein–protein interaction networks construction and module analysis

TABLE 2 Functional and pathway enrichment analysis of genes in the most significant modules

ID	Pathway description	Observed gene count	False discovery rate
hsa280	Valine, leucine and isoleucine degradation	9	3.19E–15
hsa640	Propanoate metabolism	6	1.13E–09
hsa71	Fatty acid degradation	6	4.99E–09
hsa1212	Fatty acid metabolism	6	7.49E–09
hsa1100	Metabolic pathways	13	4.30E–07
hsa650	Butanoate metabolism	4	3.44E–06
hsa380	Tryptophan metabolism	4	1.60E–05
hsa1200	Carbon metabolism	5	1.81E–05
hsa310	Lysine degradation	4	3.44E–05
hsa4060	Cytokine–cytokine receptor interaction	6	7.84E–05

homodimerization activity, glycosaminoglycan binding, and carbohydrate binding. Additionally, the results of KEGG pathway analysis revealed that these DEGs were mainly enriched in valine, leucine, and isoleucine degradation, propanoate metabolism, complement and coagulation cascades, butanoate metabolism, and cell adhesion molecules (CAMs). (Figure 3; Table 1).

3.3 | PPI network construction and module analysis

The PPI network among the DEGs was established by using the STRING database, with 417 nodes and 1912 interactions. (Figure 4a) Figure 4b presented the most significant module extracted from this PPI network by MCODE. Using KEGG pathway enrichment analysis, the genes in this module were mainly enriched in valine, leucine and isoleucine degradation, propanoate metabolism, and fatty acid degradation. (Table 2) Hierarchical clustering indicated that the module genes could basically differentiate the KIRC tissues from the noncancerous tissues. (Figure 5a) The biological process analysis indicated that the module genes were significantly related to cell cycle arrest, negative regulation of cell cycle, and cell cycle arrest in response to pheromone. (Figure 5b) *ACAA1*, *ACADSB*, *ALDH6A1*, *AUH*, *HADH*, and *PCCA* with high degree of connectivity were selected as hub genes.

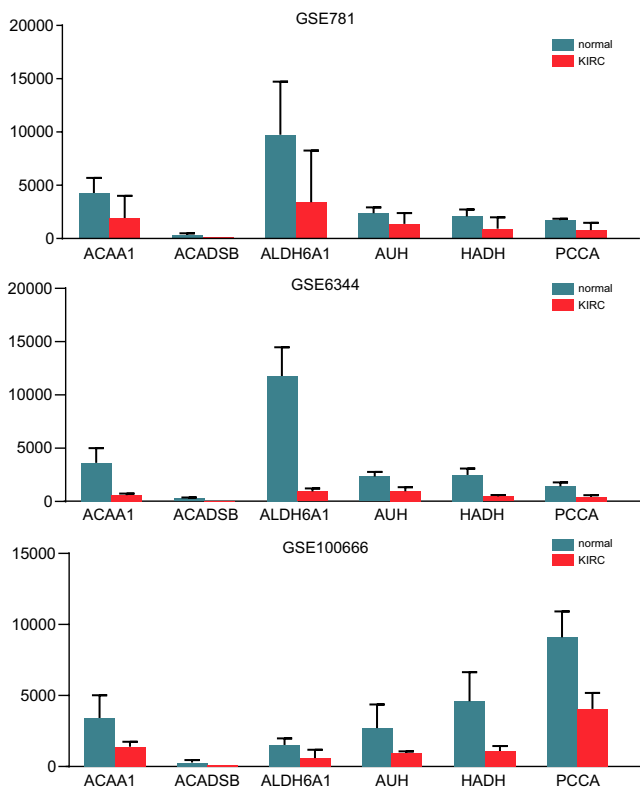


FIGURE 6 The expression level of hub genes between KIRC and normal tissues in three datasets

The expression level of *ACAA1*, *ACADSB*, *ALDH6A1*, *AUH*, *HADH*, and *PCCA* in KIRC tissues was significantly downregulated among each dataset. (Figure 6). Table S1 showed the full names and function roles of these hub genes.

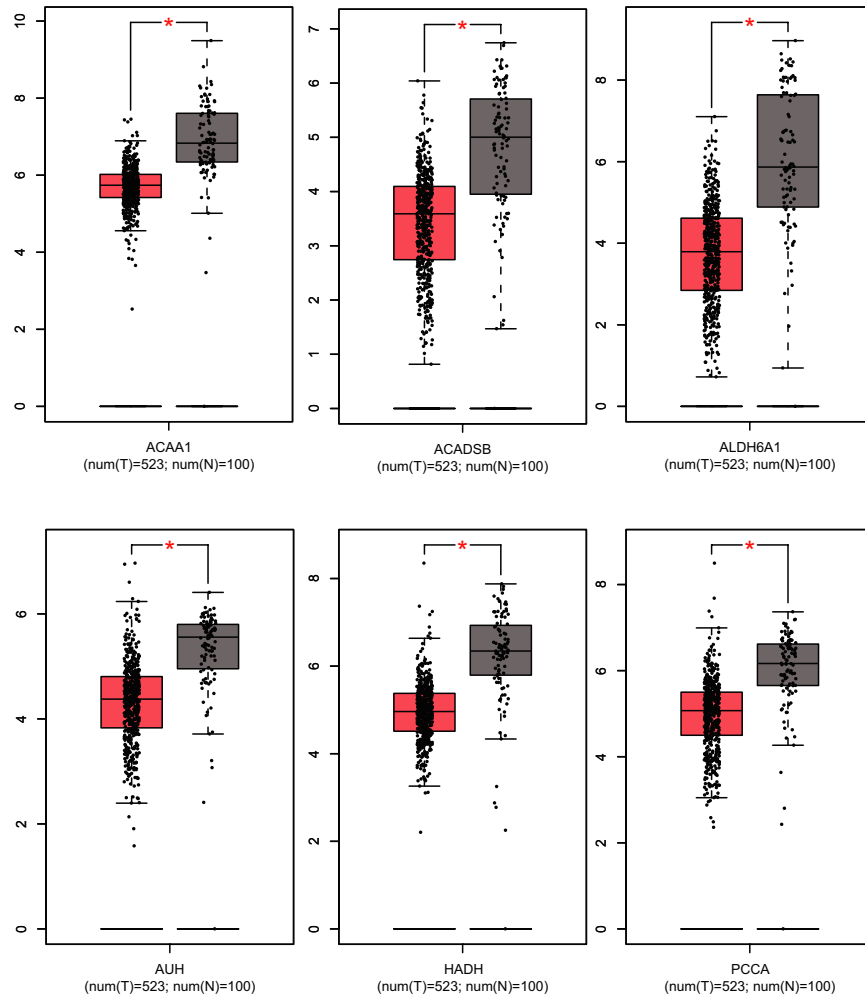
3.4 | Analysis and validation of hub genes

The hub gene expression and their prognostic correlation were further analyzed by GEPIA database. The mining of GEPIA database further confirmed that these hub genes were significantly downregulated in KIRC tissues compared to normal tissues. (Figure 7a) Based on the prognostic analysis of 516 KIRC patients in GEPIA database, we found that downregulation of *ACAA1*, *ACADSB*, *ALDH6A1*, *AUH*, *HADH*, and *PCCA* were correlated with significantly worse overall survival of patients. (Figure 7b) The expression level of these hub genes could represent the important prognostic biomarkers for predicting the survival of KIRC patients. Due to the gene expression not always consistent with its protein content, (Maier, Guell, & Serrano, 2009) we used HPA database to further analyze the protein level of hub genes. The immunohistochemical staining results indicated significantly higher positivity for *ACAA1*, *ACADSB*, *ALDH6A1*, *AUH*, *HADH*, and *PCCA* in adjacent normal tissues than in cancer tissues. (Figure 8) A network of the module genes and their coexpression genes was constructed using cBioPortal online platform. (Figure 9a) To further explore the molecular mechanism of the hub genes in KIRC, we predicted potential transcription factors and constructed a regulatory network of lncRNA, miRNA, and mRNA which could be involved in the expression of hub genes by GGBI analysis. (Gene-Cloud Biotechnology Information, Figure S1 & S2).

3.5 | Identification of related active small molecules

In order to screen and identify potential therapeutic drugs for KIRC, we uploaded these DEGs which were divided into upregulated and downregulated groups into the CMap database and then matched them with small molecule treatment. The top 20 most significant small molecules with their enrichment scores were shown in Table 3 and Figure 9b. The small molecules of pipemidic acid (enrichment score = -0.853) and dicloxacillin (enrichment score = -0.819) were related to highly significant negative scores, which could imitate the status of normal cells. These potential small molecule drugs could reverse the gene expression induced by KIRC, thus providing new directions and molecular mechanisms for the future development of new targeted drugs for treating KIRC. However, further studies were still required to confirm the value of these candidate small molecules in treating KIRC.

a



b

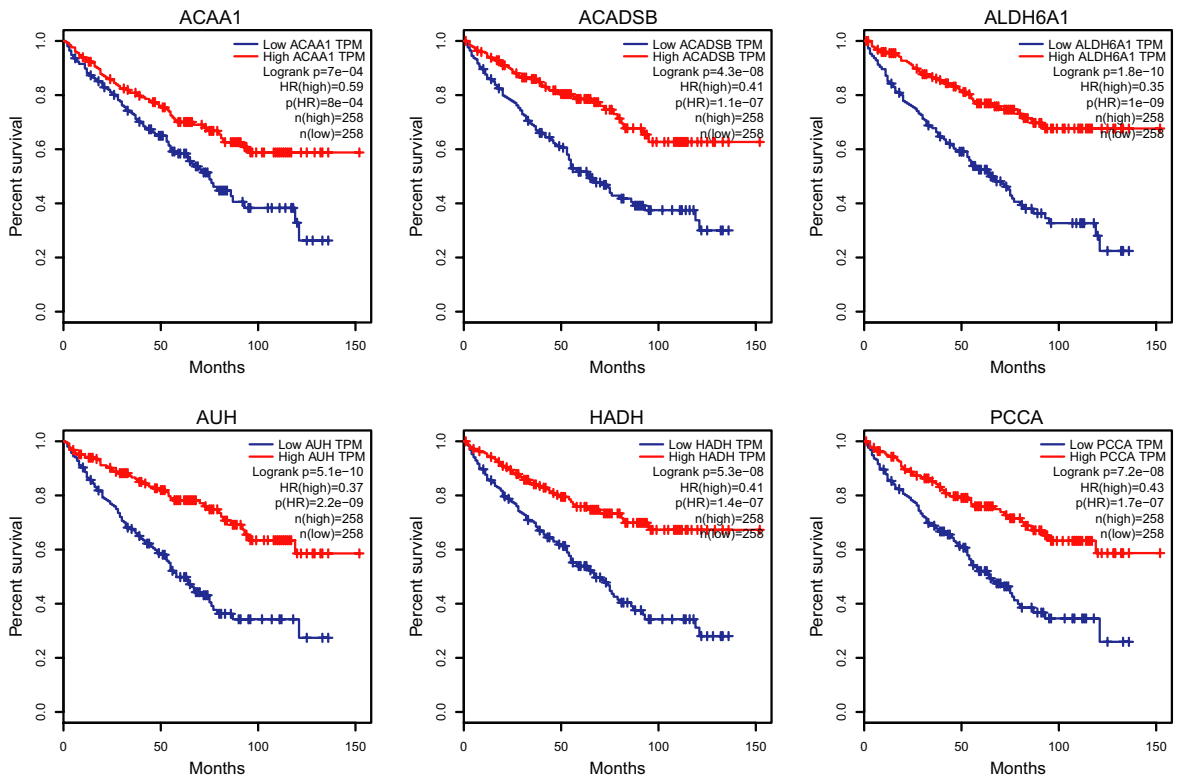


FIGURE 7 (a) The expression of hub genes between KIRC tissues and normal tissues. (b) The prognostic value of hub genes

4 | DISCUSSION

KIRC is the most lethal urological disease and the incidence has increased over the past few decades. KIRC is also a heterogeneous histological disease and the mechanisms underlying its occurrence and progression are complex. (Van Poppel et al., 2011; Xue et al., 2012) Therefore, identifying the key genes related to the pathogenesis and prognosis of KIRC is urgently demanded. In the present study, we integrated gene expression profiles of GSE781, GSE6344, and GSE100666 from GEO database to screen the differentially expressed genes between KIRC and adjacent normal tissues

by bioinformatics techniques. The purpose of our study was to find the novel diagnostic and prognostic biomarkers for KIRC. Additionally, we also identified the candidate small molecule drugs that could reverse the tumoral status of KIRC in order to provide new directions for the future development of new drugs for treating KIRC.

A total of 503 overlap DEGs were found among the three datasets, including 231 significantly upregulated genes and 250 downregulated genes. We analyzed the biological process, molecular function, and cellular component of these overlap DEGs using GO enrichment analysis in order to explore their role in the occurrence and development of KIRC. Excretion, defense response, and response to wounding were

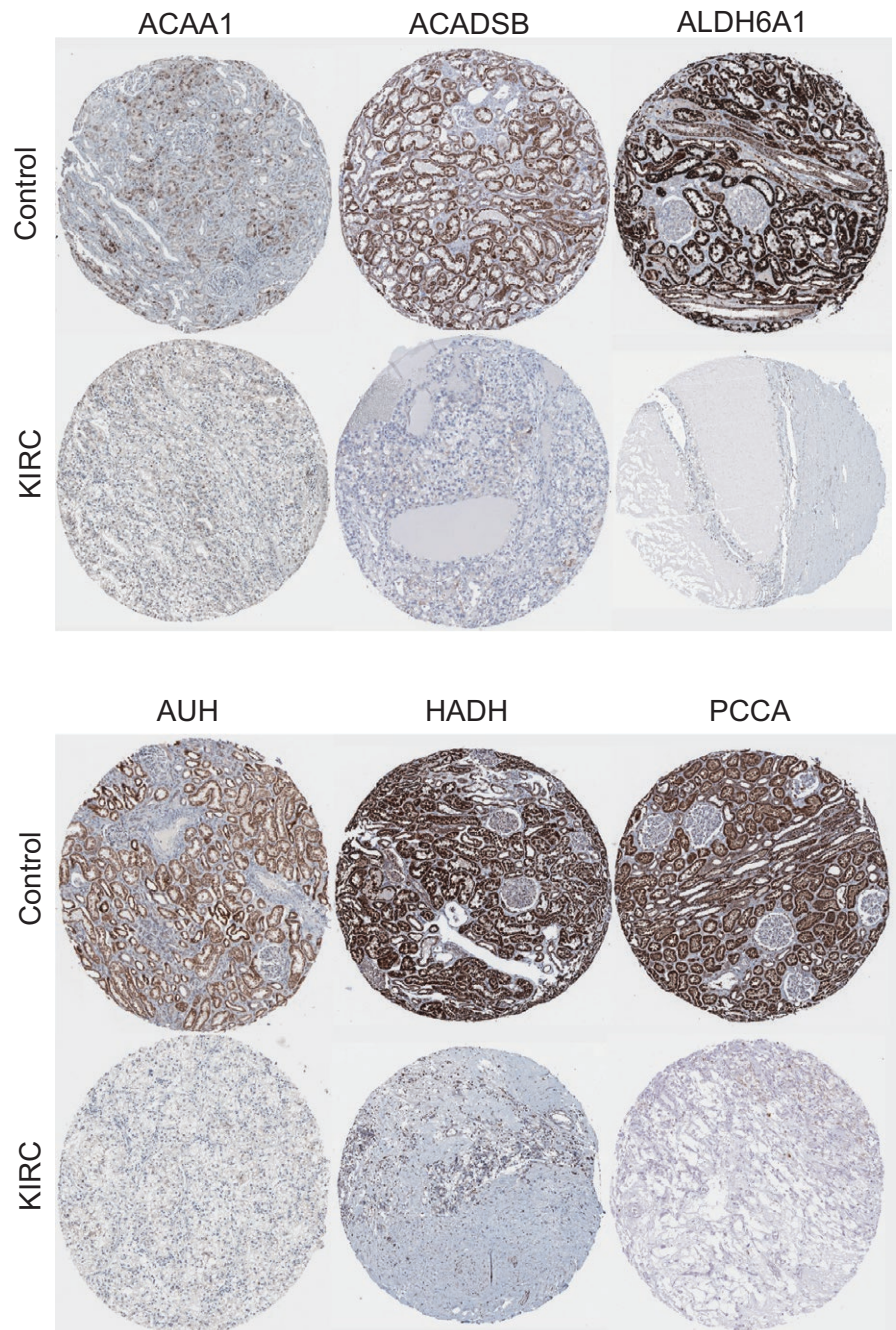


FIGURE 8 Representative immunohistochemistry staining results reveal the protein level expression of hub genes in KIRC and normal tissues

the top three significantly major functions among the biological process. Molecular function enriched for the DEGs were mainly within identical protein binding, protein dimerization activity, and protein homodimerization activity. Changes in cell component were mainly associated with plasma membrane part, integral to plasma membrane, and intrinsic to plasma membrane. Furthermore, the results of KEGG pathway suggested these DEGs were significantly enriched in valine, leucine and isoleucine degradation, propanoate metabolism, complement and coagulation cascades, butanoate metabolism, and cell adhesion molecules. The cell adhesion molecules are widely involved in the intercellular and cellular–extracellular matrix interactions of tumors. Some adhesion molecules play an important role in tumor recurrence, metastasis, and invasion. (Okegawa, Pong, Li, & Hsieh, 2004) Based on the construction of PPI network and module analysis,

ACAA1, *ACADSB*, *ALDH6A1*, *AUH*, *HADH*, and *PCCA* with high degree of connectivity were selected as hub genes. They were significantly downregulated in KIRC tissues compared with normal tissues. We used GEPIA database to further validate the results of bioinformatics analysis. Similarly, the mining of GEPIA database also confirmed that these hub genes exhibited significant differences in expression between KIRC and normal tissues. More importantly, these hub genes have important value in predicting the survival of KIRC patients. The low expression level of *ACAA1*, *ACADSB*, *ALDH6A1*, *AUH*, *HADH*, and *PCCA* was significantly correlated with the prognosis of patients. The immunohistochemical staining results revealed that the expression level of hub gene was consistent with their protein expression level, which further validated the accuracy of our findings. The present study first revealed six key genes with the diagnostic, prognostic, and

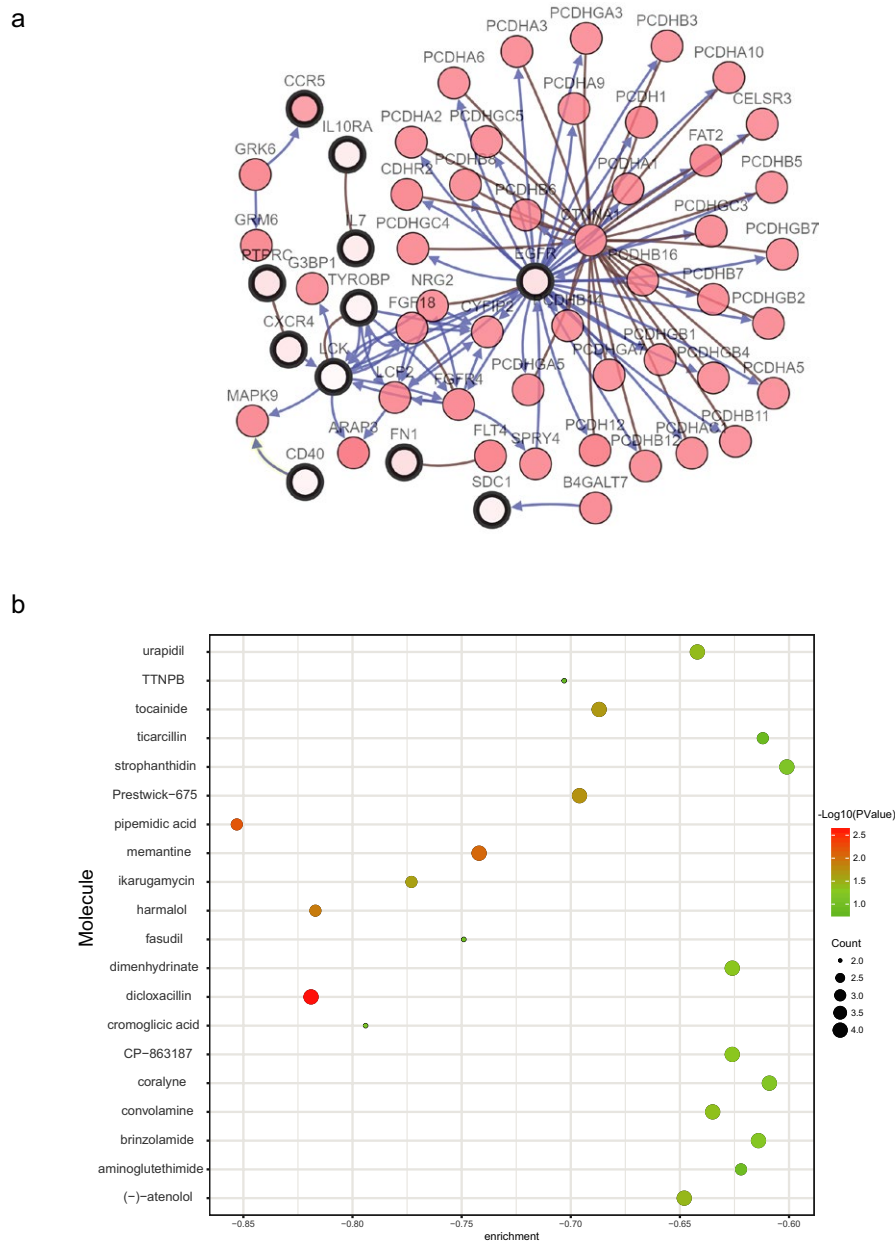


FIGURE 9 (a) The network of module genes and their coexpression genes constructed by cBioPortal. Nodes with thick outline: hub genes; Nodes with thin outline: coexpression genes. (b) Pop plot of top 20 identified small molecules that could reverse the gene expression of KIRC

therapeutic value in KIRC. No studies have reported the role of these six key genes in the initiation and progression of KIRC. In order to further enhance the understanding of these hub genes in the pathogenesis of KIRC, we predicted the corresponding potential transcription factors and establish a regulatory network of lncRNA–miRNA–mRNA for each hub gene. The construction of these regulatory networks will contribute in revealing the potential mechanism of these hub genes in KIRC initiation and progression. *CENPK*, a member of the centromeric protein family, was markedly upregulated in ovarian cancer tissues and associated with poor prognosis in patients. *ALDH6A1* and *ALDH6A1* were significantly upregulated in metastatic prostate cancer and could be used to predict the survival of patients. (Cho et al., 2018) A previous study showed that *HADH* knockdown can promote tumor cell migration and invasion by activating Akt signaling pathway. (Shen et al., 2017) Evidence indicated that *ACADSB* was also markedly downregulated in HCC cell lines. (Nwosu et al., 2018) However, no studies have reported the effects of hub genes on the initiation and progression of KIRC. The above studies indicated these hub genes could play an important role in the occurrence and development of KIRC.

In addition, we used the CMap database to determine a set of potential small molecules which may imitate the gene expression of normal cells. The small molecules with higher

negative enrichment scores meant that they are more likely to reverse KIRC gene expression. This work will contribute to developing new target therapeutic drugs for KIRC. The efficacy and safety of pipemidic acid, the most significant small molecule (enrichment score = -0.853), has not been investigated as yet in KIRC. Meanwhile, the relationship between dicloxacillin (enrichment score = -0.819) and KIRC remains little known. Therefore, further studies were urgently demanded to reveal the huge potential of these listed small molecules in treating KIRC.

In conclusion, through a comprehensive bioinformatics analysis for the microarray datasets of KIRC, we first uncovered six novel key genes associated with KIRC initiation and progression. These hub genes could serve as promising novel biomarkers for the diagnosis, prognosis, and treatment of KIRC. A series of detailed analysis were further confirmed their vital roles on the pathogenesis of KIRC. In addition, this study identified a group of candidate small molecule drugs, which could provide new directions for the development of new targeted drugs for KIRC. Taken together, this work provided new insights for KIRC to integrate multiple biomarkers in clinical practices.

CONFLICTS OF INTERESTS

None.

REFERENCES

- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, 25(1), 25–29. <https://doi.org/10.1038/75556>
- Bandettini, W. P., Kellman, P., Mancini, C., Booker, O. J., Vasu, S., Leung, S. W., ... Arai, A. E. (2012). MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: A clinical validation study. *Journal of Cardiovascular Magnetic Resonance*, 14, 83. <https://doi.org/10.1186/1532-429x-14-83>
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6), 394–424. <https://doi.org/10.3322/caac.21492>
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., ... Schultz, N. (2012). The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discovery*, 2(5), 401–404. <https://doi.org/10.1158/2159-8290.cd-12-0095>
- Cho, S. Y., Kang, S., Kim, D. S., Na, H. J., Kim, Y. J., Choi, Y. D., & Cho, N. H. (2018). HSP27, ALDH6A1 and prohibitin act as a trio-biomarker to predict survival in late metastatic prostate cancer. *Anticancer Research*, 38(11), 6551–6560. <https://doi.org/10.21873/anticancer.13021>

TABLE 3 List of the 20 most significant small molecule drugs that can reverse the tumoral status of KIRC

CMap name	Enrichment	p
Pipemidic acid	−0.853	0.00637
Dicloxacillin	−0.819	0.00203
Harmalol	−0.817	0.01216
Cromoglicic acid	−0.794	0.08303
Ikarugamycin	−0.773	0.024
Fasudil	−0.749	0.12454
Memantine	−0.742	0.00871
TTNPB	−0.703	0.17597
Prestwick−675	−0.696	0.01814
Tocainide	−0.687	0.02085
(−)-Atenolol	−0.648	0.03648
Urapidil	−0.642	0.03957
Convolamine	−0.635	0.04323
Dimenhydrinate	−0.626	0.04812
CP−863187	−0.626	0.04846
Aminoglutethimide	−0.622	0.11319
Brinzolamide	−0.614	0.05602
Ticarcillin	−0.612	0.12627
Coralyne	−0.609	0.05954
Strophanthidin	−0.601	0.06595

- Chowdhury, B., Porter, E. G., Stewart, J. C., Ferreira, C. R., Schipma, M. J., & Dykhuizen, E. C. (2016). PBRM1 regulates the expression of genes involved in metabolism and cell adhesion in renal clear cell carcinoma. *PLoS ONE*, *11*(4), e0153718. <https://doi.org/10.1371/journal.pone.0153718>
- Dennis, G. Jr, Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for annotation, visualization, and integrated discovery. *Genome Biology*, *4*(5), P3.
- Escudier, B., Porta, C., Schmidinger, M., Rioux-Leclercq, N., Bex, A., Khoo, V., ... Horwich, A. (2016). Renal cell carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*, *27*(suppl 5), v58–v68. <https://doi.org/10.1093/annonc/mdw328>
- Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., ... Schultz, N. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signalling*, *6*(269), p1. <https://doi.org/10.1126/scisignal.2004088>
- Gene Ontology Consortium. (2006). The Gene Ontology (GO) project in 2006. *Nucleic Acids Research*, *34*(Database issue), D322–326. <https://doi.org/10.1093/nar/gkj021>
- Gumz, M. L., Zou, H., Kreinest, P. A., Childs, A. C., Belmonte, L. S., LeGrand, S. N., ... Copland, J. A. (2007). Secreted frizzled-related protein 1 loss contributes to tumor phenotype of clear cell renal cell carcinoma. *Clinical Cancer Research*, *13*(16), 4740–4749. <https://doi.org/10.1158/1078-0432.ccr-07-0143>
- Hsieh, J. J., Le, V., Cao, D., Cheng, E. H., & Creighton, C. J. (2018). Genomic classifications of renal cell carcinoma: A critical step towards the future application of personalized kidney cancer care with pan-omics precision. *The Journal of Pathology*, *244*(5), 525–537. <https://doi.org/10.1002/path.5022>
- Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. <https://doi.org/10.1038/nprot.2008.211>
- Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., ... Golub, T. R. (2006). The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science*, *313*(5795), 1929–1935. <https://doi.org/10.1126/science.1132939>
- Lenburg, M. E., Liou, L. S., Gerry, N. P., Frampton, G. M., Cohen, H. T., & Christman, M. F. (2003). Previously unidentified changes in renal cell carcinoma gene expression identified by parametric analysis of microarray data. *BMC Cancer*, *3*, 31. <https://doi.org/10.1186/1471-2407-3-31>
- Linehan, W. M. (2012). Genetic basis of kidney cancer: Role of genomics for the development of disease-based therapeutics. *Genome Research*, *22*(11), 2089–2100. <https://doi.org/10.1101/gr.131110.111>
- Maere, S., Heymans, K., & Kuiper, M. (2005). BiNGO: A Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, *21*(16), 3448–3449. <https://doi.org/10.1093/bioinformatics/bti551>
- Maier, T., Guell, M., & Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Letters*, *583*(24), 3966–3973. <https://doi.org/10.1016/j.febslet.2009.10.036>
- Nwosu, Z. C., Battello, N., Rothley, M., Pioronska, W., Sitek, B., Ebert, M. P., ... Dooley, S. (2018). Liver cancer cell lines distinctly mimic the metabolic gene expression pattern of the corresponding human tumours. *J Exp Clin Cancer Res*, *37*(1), 211. <https://doi.org/10.1186/s13046-018-0872-6>
- Okegawa, T., Pong, R. C., Li, Y., & Hsieh, J. T. (2004). The role of cell adhesion molecule in cancer progression and its application in cancer therapy. *Acta Biochimica Polonica*, *51*(2), 445–457. 035001445
- Papanicolaou, A., Gebauer-Jung, S., Blaxter, M. L., Owen McMillan, W., & Jiggins, C. D. (2008). ButterflyBase: A platform for lepidopteran genomics. *Nucleic Acids Research*, *36*(Database issue), D582–587. <https://doi.org/10.1093/nar/gkm853>
- 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>
- Shen, C., Song, Y. H., Xie, Y., Wang, X., Wang, Y., Wang, C., ... Zhou, J. (2017). Downregulation of HADH promotes gastric cancer progression via Akt signaling pathway. *Oncotarget*, *8*(44), 76279–76289. <https://doi.org/10.18632/oncotarget.19348>
- Shuch, B., Amin, A., Armstrong, A. J., Eble, J. N., Ficarra, V., Lopez-Beltran, A., ... Kutikov, A. (2015). Understanding pathologic variants of renal cell carcinoma: Distilling therapeutic opportunities from biologic complexity. *European Urology*, *67*(1), 85–97. <https://doi.org/10.1016/j.eururo.2014.04.029>
- Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., & Ideker, T. (2011). Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics*, *27*(3), 431–432. <https://doi.org/10.1093/bioinformatics/btq675>
- Srigley, J. R., Delahunt, B., Eble, J. N., Egevad, L., Epstein, J. I., Grignon, D., ... Argani, P. (2013). The International Society of Urological Pathology (ISUP) vancouver classification of renal neoplasia. *American Journal of Surgical Pathology*, *37*(10), 1469–1489. <https://doi.org/10.1097/PAS.0b013e318299f2d1>
- Sun, G., Li, Y., Peng, Y., Lu, D., Zhang, F., Cui, X., ... Li, Z. (2019). Identification of differentially expressed genes and biological characteristics of colorectal cancer by integrated bioinformatics analysis. *Journal of Cellular Physiology*, <https://doi.org/10.1002/jcp.28163>
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., ... von Mering, C. (2015). STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, *43*(Database, issue), D447–452. <https://doi.org/10.1093/nar/gku1003>
- Tang, Z., Li, C., Kang, B., Gao, G., Li, C., & Zhang, Z. (2017). GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Research*, *45*(W1), W98–w102. <https://doi.org/10.1093/nar/gkx247>
- Tun, H. W., Marlow, L. A., von Roemeling, C. A., Cooper, S. J., Kreinest, P., Wu, K., ... Copland, J. A. (2010). Pathway signature and cellular differentiation in clear cell renal cell carcinoma. *PLoS ONE*, *5*(5), e10696. <https://doi.org/10.1371/journal.pone.0010696>
- Van Poppel, H., Becker, F., Cadeddu, J. A., Gill, I. S., Janetschek, G., Jewett, M. A., ... Zhu, G. (2011). Treatment of localised renal cell carcinoma. *European Urology*, *60*(4), 662–672. <https://doi.org/10.1016/j.eururo.2011.06.040>
- Xue, Y. J., Xiao, R. H., Long, D. Z., Zou, X. F., Wang, X. N., Zhang, G. X., ... Liu, M. (2012). Overexpression of FoxM1 is associated with tumor progression in patients with clear cell renal cell carcinoma. *Journal of Translational Medicine*, *10*, 200. <https://doi.org/10.1186/1479-5876-10-200>

Yan, X., Wan, H., Hao, X., Lan, T., Li, W., Xu, L., ... Wu, H. (2019). Importance of gene expression signatures in pancreatic cancer prognosis and the establishment of a prediction model. *Cancer Manag Res*, *11*, 273–283. <https://doi.org/10.2147/cmar.s185205>

SUPPORTING INFORMATION

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

How to cite this article: Zhang B, Wu Q, Wang Z, et al. The promising novel biomarkers and candidate small molecule drugs in kidney renal clear cell carcinoma: Evidence from bioinformatics analysis of high-throughput data. *Mol Genet Genomic Med*. 2019;7:e607. <https://doi.org/10.1002/mgg3.607>