Construction of a protein-based model for prognosis prediction of kidney renal clear cell carcinoma: an investigation based on functional proteomics data

Huiming Jiang (jianghm0228@163.com)
Meizhou People's Hospital

Meizhou People's Hospital

Haibin Chen
Shantou University Medical College

Pei Wan
Meizhou People's Hospital

NanHui Chen
Meizhou People's Hospital

Research article

**Keywords:** kidney renal clear cell carcinoma, proteomics, risk score, prognosis

**DOI:** https://doi.org/10.21203/rs.3.rs-56345/v1

**License:** © This work is licensed under a Creative Commons Attribution 4.0 International License.

Read Full License
Abstract

Background

Several studies have shown prognostic value of gene-based models at mRNA level in kidney renal clear cell carcinoma (KIRC). However, protein-based models for prognosis prediction of KIRC are rarely reported, then we conduct this study.

Methods

Proteomics and clinical data of KIRC were acquired from The Cancer Proteome Atlas (TCPA) and The Cancer Genome Atlas (TCGA), respectively. Prognosis-associated proteins were screened by univariate Cox regression analysis and log-rank test. Patients were grouped into training set and testing set. The protein-based prognostic model was constructed by lasso Cox regression analysis in training set and validated in test set and whole set.

Results

A five-protein model (ACC1, IGFBP2, MIG6, PEA15 and RAD51) was constructed for KIRC prognosis prediction. It could well classify KIRC patients into high-risk and low-risk group with significantly different survival. The results was validated in the testing set and whole set. Age, AJCC stage and risk score based on the five-protein model were identified as independent prognostic parameters and they were used to construct a nomogram. The calibration plot showed the nomogram had good agreement between predicted and actual outcomes. Time-dependent ROC curves revealed the nomogram performed best in predicting the 1-year, 3-year and 5-year overall survival (OS) compared with other independent prognostic parameters. DCA demonstrated the nomogram had obviously clinical net benefit. Furthermore, several cancer-related biological signaling pathways were enriched in the functional enrichment analysis.

Conclusion

Our study developed an effective protein-based model to predict the OS of KIRC, which may help clinicians to offer individual treatment.

Background

Kidney cancer is one of the most lethal urological diseases globally. Kidney renal clear cell carcinoma (KIRC) is the highest mortality, invasion, and metastasis subtype, which accounts for about 70–75% of kidney cancer [1, 2]. With the increasing incidence and mortality in recent years, KIRC remains a big challenge worldwide [3]. It was reported that the newly diagnosed cases in the USA had increased to 73,820 in 2019 and resulted in nearly 14,770 deaths [4]. Approximately 15% RCC patients have
progressed into distant metastasis at clinical diagnosis, and the 5-year survival rates are less than 10% in patients with metastasis [5]. Thus, accurate prognostic models are urgently needed to distinguish the high-risk patients with poor survival. Conventional parameters such as TNM stage are not able to provide enough accurate predictions because of the great heterogeneity of KIRC and the predictive ability is far from satisfying [6].

In recent years, with widespread use of high-throughput sequencing technology to screen biomarkers of cancer, several studies have shown the great potential of gene-based model at mRNA level in predicting the survival of KIRC [7–9]. Since lots of gene regulation happens post-transcriptionally, functional proteomics still represents a powerful approach to reveal the pathophysiology of cancer. However, protein-based models for prognosis prediction of KIRC are rarely reported. Our study aimed to identify a prognostic model base on proteins for KIRC using data from The Cancer Proteome Atlas (TCPA).

TCPA is a comprehensive, cutting-edge, protein-centric pan-cancer analytic platform which is freely available at http://tcpaportal.org. Using the reverse-phase protein arrays (RPPAs) platform, a technology that can quantitatively assess lots of protein markers in thousands of samples in a sensitive, cost-effective and high-throughput manner, approximately 8,000 tumor samples across 32 cancer types through The Cancer Genome Atlas (TCGA) project [10, 11]. TCPA has become a vital bioinformatics resource because of the advantage of quantitative protein expression and survival data from TCGA patient.

**Methods**

**Data extraction**

We downloaded the protein expression data of 445 KIRC patients from TCPA database (http://tcpaportal.org) and the clinical information of 537 KIRC patients from TCGA database (https://portal.gdc.cancer.gov/), respectively. Next, we matched the patients’ clinical information and protein expression data and removed the patients with a follow-up period less than one month. Finally, 433 patients with protein expression data and survival information were identified.

**Identification of prognosis-related proteins**

Expression data of 233 proteins were included in each patient. Univariate Cox regression analysis and log-rank test were used to screen the prognostic proteins.

**Construction of the protein-based prognostic model**

433 patients were randomly grouped into a training set and testing set. The training set was applied to construct the Cox regression hazards model, while the testing set and the whole set was applied to validate the performance of the model. The potential risk proteins were identified by Lasso regression analysis and then they were used to construct a protein-based prognostic model. Patients’ risk score was calculated by the following computational formula: Risk score = Coefficient (protein 1) x Expression value
of protein 1 + Coefficient (protein 2) x Expression value of protein 2 + ⋯ + Coefficient (protein i) x Expression value of (protein i). “Protein i” mean the ith selected protein, and “coefficient (protein i)” mean the corresponding regression coefficient [12]. The optimal cut-off value of risk score was investigated by the R package “survminer” [13]. Patients were separated into high-risk and low-risk group base on the cut-off value. The higher risk score means worse survival for KIRC.

**Validation of the protein-based prognostic model**

The risk score of each patient in the testing set and whole set was calculated with the same prognostic model. Then the ROC analysis, Kaplan-Meier (KM) survival analysis was performed to validate the model's predictive power. Besides, the protein expression level of the prognostic proteins was explored in the Human Protein Atlas online database (http://www.proteinatlas.org).

**Independence prognostic value of the protein-based model**

The independent prognostic value of protein-based model was assessed by the methods of Univariate Cox regression analysis, multivariate Cox regression analysis and stratified analysis. The association between the risk score and clinicopathological features was also explored.

**Constructing and estimating a predictive nomogram**

All of the independent prognostic parameters identified were used to construct a composite nomogram. The calibration plot of the nomogram was applied to check the consistency between predictive and actual outcome of KIRC. Time-dependent ROC curves were plotted to compare the predictive accuracy and decision curve analysis (DCA) was used to compare the clinical net benefit among nomogram and other independent prognostic parameters.

**Functional enrichment analysis**

We performed the correlation analysis to identified the co-expressed proteins with the prognostic proteins in the model. The correlation coefficients > 0.4 and P value < 0.001 was considered significantly. Then gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis were conducted on Metascape (http://metascape.org/gp/index.html#/main/step1).

**Statistical analysis**

All preprocessing processes and statistical analysis were realized by R software v3.6.1 (R Foundation for Statistical Computing, Vienna, Austria), SPSS v25 (SPSS Inc.) and Strawberry Perl v5.30.1.1. Chi-square or Fisher's exact test was used for qualitative variables and t-test for quantitative variables. p < 0.05 was considered statistically significant except where noted.

**Results**
Identification of prognosis-related proteins

The study was conducted as shown in the flowchart (Fig. 1). We identified 77 proteins significantly related to overall survival (OS) of KIRC including 35 proteins with HR > 1 and 42 proteins with HR < 1 (Fig. 2A). It means the high expression of proteins was associated with poor survival of patients when HR > 1.

Construction of a five-protein prognostic model in the training set

Totally 433 patients with a follow-up period longer than one month were randomly divided into training set (n = 216) and testing set (n = 217). There was not significant difference in clinical characteristic between two groups (Table 1). Lasso Cox regression analysis were performed in the training set (Fig. 2B) and five risk proteins were successfully identified (Table 2). The proteins were acetyl-CoA carboxylase 1 (ACC1), insulin-like growth factor binding protein 2 (IGFBP2), mitogen-inducible gene 6 (MIG6), proliferation and apoptosis adaptor protein 15 (PEA15) and RAD51 recombinase (RAD51). Finally, a protein-based prognostic model was established by the five proteins. The risk score = 0.438 * Expression of ACC1 + 0.251 * Expression of IGFBP2 − 0.229 * Expression of MIG6 + 0.283 * Expression of PEA15 + 0.348 * Expression of RAD51. According to the optimal cut-off value of risk score, patients were divided into high-risk and low-risk group (Fig. 3A). More dead patients corresponded to the higher risk score in the training set (Fig. 3D). Heatmap of the five prognostic proteins presented an obviously different expression between the high-risk and low-risk group (Fig. 3G). The KM curve show that patients in high-risk group suffered a significantly poorer OS than patients in low-risk group (p < 0.001) (Fig. 4A). The area under the ROC (AUC) for 1-year, 3-year and 5-year OS were 0.789, 0.753, and 0.785, respectively. (Fig. 4D). Taking together, the results indicated that our five-protein model had a well performance in survival prediction for KIRC.
Table 1  
Clinical features of KIRC patients in training set, testing set and whole set.

| Characteristic               | Whole set | Training set | Testing set | P-value |
|-----------------------------|-----------|--------------|-------------|---------|
| Total                       | n = 433   | n = 216      | n = 217     |         |
| Dead (%)                    | 280 (64.7)| 135 (62.5)   | 145 (66.8)  | 0.367   |
| Median age, years (range)   | 60 (26–88)| 59 (26–88)   | 61 (29–88)  | 0.211   |
| Age ≥ 60 years (%)          | 228 (52.7)| 107 (49.5)   | 121 (55.8)  |         |
| Female (%)                  | 141 (32.6)| 63 (29.2)    | 78 (35.9)   | 0.151   |
| Tumor grade                 |           |              |             | 0.183   |
| G1                          | 8 (1.8)   | 1 (0.5)      | 7 (3.2)     |         |
| G2                          | 181 (41.8)| 89 (41.2)    | 92 (42.4)   |         |
| G3                          | 171 (39.5)| 89 (41.2)    | 82 (37.8)   |         |
| G4                          | 71 (16.4) | 36 (16.7)    | 35 (16.1)   |         |
| Unknow                      | 2 (0.4)   | 1 (0.5)      | 1 (0.5)     |         |
| AJCC stage                  |           |              |             | 0.754   |
| Stage I                     | 210 (48.5)| 104 (48.1)   | 106 (48.8)  |         |
| Stage II                    | 43 (9.9)  | 24 (11.1)    | 19 (8.8)    |         |
| Stage III                   | 103 (23.8)| 48 (22.2)    | 55 (25.3)   |         |
| Stage IV                    | 75 (17.3) | 39 (18.1)    | 36 (16.6)   |         |
| Unknow                      | 2 (0.5)   | 1 (0.5)      | 1 (0.5)     |         |

Abbreviations: KIRC, kidney renal clear cell carcinoma; AJCC, American Joint Committee on Cancer.
Table 2
Proteins contained in the prognostic model.

| Protein | Description                  | Gene symbol | Coefficient | HR  | 95% CI      | P-value |
|---------|------------------------------|-------------|-------------|-----|-------------|---------|
| ACC1    | acetyl-CoA carboxylase 1     | ACACA       | 0.438       | 2.9 | 1.9–4.4     | 2.60E-07|
| IGFBP2  | insulin-like growth factor binding protein 2 | IGFBP2 | 0.251 | 1.7 | 1.4–2.2     | 1.50E-06|
| MIG6    | mitogen-inducible gene 6    | ERRFI1      | -0.229      | 0.35| 0.21–0.58   | 5.00E-05|
| PEA15   | proliferation and apoptosis adaptor protein 15 | PEA15 | 0.283 | 3.2 | 1.6–6.1     | 0.00055 |
| RAD51   | RAD51 recombinase            | RAD51       | 0.348       | 5.1 | 2.4–11      | 2.90E-05|

Validation of the five-protein prognostic model in the testing set and whole set

Validation was conducted in the testing set and whole set with the same method above. Patients were also divided into high-risk and low-risk group with the same cut-off value (Fig. 3B-C) and More death events occurred in the high-risk group (Fig. 3E-F). Different expression of five proteins was also demonstrated between high-risk and low-risk group (Fig. 3H-I). Meanwhile, the KM curve show the OS was significantly poorer in the high-risk group than that in the low-risk group (both p < 0.001) (Fig. 4B-C). The AUC were 0.730, 0.705, 0.725 and 0.763, 0.732, 0.755 for 1-year, 3-year and 5-year OS in the testing set and whole set, respectively (Fig. 4E-F). All of the results were consistent with that in the training set which suggested the reliability of our model.

Moreover, the expression of the five proteins was also validated in the Human Protein Atlas online database. The immunohistochemistry (IHC) staining show a higher expression of ACC1, PEA15 and RAD51 in tumor tissue than that in normal tissue (Fig. 2C). However, data of IGFBP2 and MIG6 was not found.

Independent prognostic value of the five-protein model in the whole set

Univariate Cox regression and multivariate Cox regression analysis were performed to detect whether the risk score calculated by our five-protein model had independent prognostic value from other clinicopathological parameters including age, gender, grade and American Joint Committee on Cancer (AJCC) stage. Univariate Cox regression analysis indicated that age, grade, AJCC stage and risk score were related to the OS of KIRC (Fig. 5A). However, only age, AJCC stage and risk score were confirmed as independent prognostic factors by multivariate Cox regression analysis (Fig. 5B). Moreover, stratified analysis revealed that patients in high-risk group had significantly poorer OS than patients in low-risk group according to the age, gender, grade and AJCC stage (all p < 0.001) (Fig. 6A-D). Furthermore, the risk
score was also significantly increasing with the grade (p < 0.001), AJCC stage (p < 0.001) and age (p = 0.033) progress, but not with gender (p = 0.41) (Fig. 6E-H).

**Constructing and estimating the protein-based nomogram in the whole set**

In order to construct a more reliable predictive model for clinical practice, a composite nomogram was established including three independent prognostic factors (age, AJCC stage and risk score) (Fig. 5C). The calibration plot indicated that the 1-year, 3-year and the 5-year OS predicted by nomogram was in good agreement with the actual outcome (Fig. 7A-C). The AUC of nomogram was 0.855, 0.817 and 0.799 for 1-year, 3-year and 5-year OS, respectively, which were larger than that of age, AJCC stage and risk score (Fig. 7D-F). The DCA demonstrated the nomogram could improve clinical net benefit especially for 5-year OS (Fig. 7G-I). Taking together, the results suggested that our protein-based nomogram could perform best in predicting the OS for KIRC.

**Functional enrichment analysis**

Correlation analysis show that most of the co-expressed proteins were associated with RAD51 (Fig. 8A). Several cancer-promoting proteins were positive related to the five proteins in the model such as TIGAR and EEF2 (Fig. 8B). Apoptotic signaling pathway, p53 signaling pathway and PI3K-Akt signaling pathway were enriched by the functional enrichment analysis of the co-expressed proteins (Fig. 8C-D).

**Discussion**

It is still a big challenge globally for the treatment of kidney cancer due to the high incidence and mortality. TNM stage is the most important prognostic predictor for cancers including KIRC, but the clinical outcomes can differ greatly even among patients with the same stage because of the heterogeneity of KIRC [14]. Therefore, novel biomarkers are necessary for the establishment of more accurate prognostic models. Recently, more and more proteins have been reported to be an independent prognostic factors for KIRC, but none study has focused on protein-based models for prognosis prediction of KIRC [15–17].

In our study, we firstly screened the prognosis-related proteins with the data form TCPA. Next, a five-protein model (including ACC1, IGFBP2, MIG6, PEA15 and RAD51) was developed and assessed for its prognostic value in KIRC. The risk score calculated by the model was an independent prognostic factor for KIRC and it could efficiently divide patients into high-risk and low-risk group whose OS were significantly different. The consistent results were also obtained in the subgroup analysis based on age, gender, grade and AJCC stage. Furthermore, the higher risk score was significantly associated with higher grade, AJCC stage and age. Additionally, in order to further improve the predictive power of our five-protein model, a nomogram integrating risk score and other independent prognostic parameters (age and AJCC stage) was constructed. The AUC of the constructed nomogram were 0.855, 0.817 and 0.799 for 1-year, 3-year and 5-year OS, respectively, larger than that of any parameters contained in nomogram. An AUC > 0.60 was regarded as acceptable for predictions, and an AUC > 0.75 was deemed to have excellent
predictive value [18, 19]. Thus, our nomogram had a perfect predictive power. Additionally, the calibration plot of the nomogram revealed satisfied agreement between the predicted and actual outcome. Finally, the DCA indicated the nomogram could improve clinical net benefit for KIRC. Collectively, these findings suggested our five-protein model could effectively distinguish high-risk patients with poor survival and the protein-based nomogram could accurately predict the OS for KIRC which might help clinicians make best clinical decisions.

Cancer-related proteins play a vital role in the survival of patients with cancer, all the five proteins in our risk model were reported to be involved in cancer. ACC1, IGFBP2, PEA15 and RAD51 are negative prognostic proteins with HR > 1, while MIG6 is a positive prognostic protein with HR < 1. ACC1 is a kind of lipid metabolism enzyme and converts from acetyl CoA to malonyl CoA which is the rate limiting step in de novo fatty acid synthesis (FASyn) [20]. Lots of cancers demonstrate an increase in FASyn and decrease of FASyn inhibits the tumor growth [21, 22]. It is observed an overexpression of ACC1 in several kinks of cancers, including breast cancer, lung cancer and prostate cancer [23, 24]. Higher ACC1 level is significantly related to poor survival in patients with lung cancer [25]. However, the role of ACC1 in KIRC is still rarely reported. IGFBP2 belongs to IGFBP family and functions as a carrier of IGF-I and IGF-II [26]. Several studies have reported that IGFBP2 can promote tumorogenesis, tumor angiogenesis and metastasis [27, 28]. Overexpression of IGFBP2 is found to be associated with metastasis and poor survival in lung cancer [29]. Furthermore, IGFBP2 is reported to play important role in renal cell carcinoma metastasis [30]. Thus, IGFBP2 may be a potential target for anticancer therapy. MIG6 is also named epidermal growth factor receptor (EGFR) feedback inhibitor 1, functioning as a negative regulator of EGFR signaling pathways [31]. Downregulation of MIG6 is found in a variety of cancers including breast cancer, lung cancer, papillary thyroid cancers and endometrial carcinoma and is closely related to poor prognosis [32–35]. Another study shows that MIG6 may suppress the epithelial proliferation through the inhibition of AKT activation [36]. Functional enrichment analysis in our study also revealed that PI3K-Akt signaling pathway was involved in KIRC, but the actual functions of MIG6 in KIRC remain unknown. PEA15 is a member of the anti-apoptotic family and it functions through ERK1/2 binding and death-effector domain [37]. It is reported that PEA15 promotes liver metastasis of colorectal cancer by ERK/MAPK signaling pathway [38]. Meanwhile, PEA15 can contribute to the AKT-regulated cisplatin resistance and is associated with the survival in gastric cancer [39]. Whether PEA15 plays a similar role in KIRC remains unknown. RAD51, the central homologous recombination protein, is important for the cellular response to DNA damage [40]. It participates in several cellular processes including apoptosis which is the top 1 of GO enrichment analysis in our study [41]. Recent studies revealed that RAD51 was overexpression in many cancers including pancreatic adenocarcinoma, breast cancer and lung cancer [42–44]. Evidence also indicated that RAD51 was associated with tumor resistance to chemotherapy and survival [43, 45]. Yet, the relationship between RAD51 and KIRC remain unexplored.

As the five-protein model established in our study is based on protein expression level without any post-transcriptional regulation, it may be more cost-effective and reliable in clinical practice. Although our protein-based model shows a good performance in the survival prediction for KIRC, it is worth noting that several limitations still existed. Firstly, the five-protein model was generated base on the data from TCPA,
in which most patients were Asian or White. Secondly, our prognostic model was not validated in external cohort. Therefore, it is necessary to validate our model in multiple centers with different populations. Finally, since most of proteins contained in the model are rarely explored in KIRC, rigorous basic experiments are indispensable.

In conclusion, our study developed an effective protein-based model to predict the OS of KIRC, which may help clinicians to offer individual treatment.

**Declarations**

**Ethics approval and consent to participate**

Not required.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Availability of data and materials**

All data generated or analyzed during this study are included in this published article or are available from the stated public repository.

**Funding**

The work was supported by the Peiyu Project of Meizhou People's Hospital (Grant/Award No. PY-C2020004).

**Authors' contributions**

JH and CN designed the study. JH wrote the manuscript and CH revised it. WP, CH and CN analyzed the data. All authors agree with the results and conclusions of this manuscript.

**Acknowledgements**
We are thankful to The Cancer Proteome Atlas (TCPA) and Human Protein Atlas for the source of data used in our study.

References

1. Hakimi AA, Voss MH, Kuo F, Sanchez A, Liu M, Nixon BG, Vuong L, Ostrovnaya I, Chen YB, Reuter V, et al. Transcriptomic Profiling of the Tumor Microenvironment Reveals Distinct Subgroups of Clear Cell Renal Cell Cancer: Data from a Randomized Phase III Trial. Cancer discovery. 2019;9(4):510–25.
2. Shuch B, Amin A, Armstrong AJ, Eble JN, Ficarra V, Lopez-Beltran A, Martignoni G, Rini BI, Kutikov A. Understanding pathologic variants of renal cell carcinoma: distilling therapeutic opportunities from biologic complexity. European urology. 2015;67(1):85–97.
3. Gray RE, Harris GT. Renal Cell Carcinoma: Diagnosis and Management. Am Family Phys. 2019;99(3):179–84.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. Cancer J Clin. 2019;69(1):7–34.
5. Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JHR, O’Brien T, Martincorena I, Tarpey P, Angelopoulos N, Yates LR, et al. Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. Cell. 2018;173(3):611–23 e617.
6. Lam JS, Klatte T, Kim HL, Patard JJ, Breda A, Zisman A, Pantuck AJ, Figlin RA: Prognostic factors and selection for clinical studies of patients with kidney cancer. Critical reviews in oncology/hematology 2008, 65(3):235–262.
7. Chen L, Luo Y, Wang G, Qian K, Qian G, Wu CL, Dan HC, Wang X, Xiao Y. Prognostic value of a gene signature in clear cell renal cell carcinoma. Journal of cellular physiology. 2019;234(7):10324–35.
8. Wu J, Jin S, Gu W, Wan F, Zhang H, Shi G, Qu Y, Ye D. Construction and Validation of a 9-Gene Signature for Predicting Prognosis in Stage III Clear Cell Renal Cell Carcinoma. Frontiers in oncology. 2019;9:152.
9. Zeng MH, Qiu JG, Xu Y, Zhang XH. IDUA, NDST1, SAP30L, CRYBA4, and SI as novel prognostic signatures clear cell renal cell carcinoma. Journal of cellular physiology 2019.
10. Chen MM, Li J, Wang Y, Akbani R, Lu Y, Mills GB, Liang H. TCPA v3.0: An Integrative Platform to Explore the Pan-Cancer Analysis of Functional Proteomic Data. Molecular cellular proteomics: MCP. 2019;18(8 suppl 1):15-s25.
11. Li J, Lu Y, Akbani R, Ju Z, Roebuck PL, Liu W, Yang JY, Broom BM, Verhaak RG, Kane DW, et al. TCPA: a resource for cancer functional proteomics data. Nature methods. 2013;10(11):1046–7.
12. Tibshirani R. The lasso method for variable selection in the Cox model. Statistics in medicine. 1997;16(4):385–95.
13. Diboun I, Wernisch L, Orengo CA, Koltzenburg M. Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. BMC Genomics. 2006;7:252.
14. Martinez-Salamanca JI, Huang WC, Millan I, Bertini R, Bianco FJ, Carballido JA, Ciancio G, Hernandez C, Herranz F, Haferkamp A, et al. Prognostic impact of the 2009 UICC/AJCC TNM staging system for
renal cell carcinoma with venous extension. European urology. 2011;59(1):120–7.

15. Hao H, Wang Z, Ren S, Shen H, Xian H, Ge W, Wang W. Reduced GRAMD1C expression correlates to poor prognosis and immune infiltrates in kidney renal clear cell carcinoma. PeerJ. 2019;7:e8205.

16. Niu H, Li F, Wang Q, Ye Z, Chen Q, Lin Y. High expression level of MMP9 is associated with poor prognosis in patients with clear cell renal carcinoma. PeerJ. 2018;6:e5050.

17. Wang K, Ruan H, Song Z, Cao Q, Bao L, Liu D, Xu T, Xiao H, Wang C, Cheng G, et al: **PLIN3 is up-regulated and correlates with poor prognosis in clear cell renal cell carcinoma.** Urologic oncology 2018, 36(7):343.e349-343.e319.

18. Cho SH, Pak K, Jeong DC, Han ME, Oh SO, Kim YH. The AP2M1 gene expression is a promising biomarker for predicting survival of patients with hepatocellular carcinoma. Journal of cellular biochemistry. 2019;120(3):4140–6.

19. Han ME, Kim JY, Kim GH, Park SY, Kim YH, Oh SO. SAC3D1: a novel prognostic marker in hepatocellular carcinoma. Scientific reports. 2018;8(1):15608.

20. Szutowicz A, Kwiatkowski J, Angielski S. Lipogenetic and glycolytic enzyme activities in carcinoma and nonmalignant diseases of the human breast. British journal of cancer. 1979;39(6):681–7.

21. Currie E, Schulze A, Zechner R, Walther TC, Farese RV Jr. Cellular fatty acid metabolism and cancer. Cell Metabol. 2013;18(2):153–61.

22. Svensson RU, Shaw RJ. Lipid Synthesis Is a Metabolic Liability of Non-Small Cell Lung Cancer. Cold Spring Harb Symp Quant Biol. 2016;81:93–103.

23. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, Lapuk A, Neve RM, Qian Z, Ryder T, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiology. Cancer cell. 2006;10(6):529–41.

24. Swinnen JV, Vanderhoydonc F, Elgamal AA, Eelen M, Vercaeren I, Joniau S, Van Poppel H, Baert L, Goossens K, Heyns W, et al. Selective activation of the fatty acid synthesis pathway in human prostate cancer. International journal of cancer. 2000;88(2):176–9.

25. Li EQ, Zhao W, Zhang C, Qin LZ, Liu SJ, Feng ZQ, Wen X, Chen CP. Synthesis and anti-cancer activity of ND-646 and its derivatives as acetyl-CoA carboxylase 1 inhibitors. European journal of pharmaceutical sciences: official journal of the European Federation for Pharmaceutical Sciences. 2019;137:105010.

26. Fisher MC, Meyer C, Garber G, Dealy CN. Role of IGFBP2, IGF-I and IGF-II in regulating long bone growth. Bone. 2005;37(6):741–50.

27. Tombolan L, Orso F, Guzzardo V, Casara S, Zin A, Bonora M, Romualdi C, Giorgi C, Bisogno G, Alaggio R, et al. High IGFBP2 expression correlates with tumor severity in pediatric rhabdomyosarcoma. Am J Pathol. 2011;179(5):2611–24.

28. Hu Q, Huang L, Kuang X, Zhang H, Ling G, Chen X, Li K, Deng Z, Zhou J. Is insulin-like growth factor binding protein 2 associated with metastasis in lung cancer? Clin Exp Metastasis. 2014;31(5):535–41.
29. Tang D, Yao R, Zhao D, Zhou L, Wu Y, Yang Y, Sun Y, Lu L, Gao W. Trichostatin A reverses the chemo-resistance of lung cancer with high IGFBP2 expression through enhancing autophagy. Scientific reports. 2018;8(1):3917.

30. Katayama H, Tamai K, Shibuya R, Nakamura M, Mochizuki M, Yamaguchi K, Kawamura S, Tochigi T, Sato I, Okanishi T, et al. Long non-coding RNA HOTAIR promotes cell migration by upregulating insulin growth factor-binding protein 2 in renal cell carcinoma. Scientific reports. 2017;7(1):12016.

31. Hackel PO, Gishizky M, Ullrich A. Mig-6 is a negative regulator of the epidermal growth factor receptor signal. Biological chemistry. 2001;382(12):1649–62.

32. Anastasi S, Sala G, Huiping C, Caprini E, Russo G, Iacovelli S, Lucini F, Ingvarsson S, Segatto O. Loss of RALT/MIG-6 expression in ERBB2-amplified breast carcinomas enhances ErbB-2 oncogenic potency and favors resistance to Herceptin. Oncogene. 2005;24(28):4540–8.

33. Ruan DT, Warren RS, Moalem J, Chung KW, Griffin AC, Shen W, Duh QY, Nakakura E, Donner DB, Khanafshar E, et al. Mitogen-inducible gene-6 expression correlates with survival and is an independent predictor of recurrence in BRAF(V600E) positive papillary thyroid cancers. Surgery. 2008;144(6):908–13. discussion 913 – 904.

34. Kim TH, Franco HL, Jung SY, Qin J, Broaddus RR, Lydon JP, Jeong JW. The synergistic effect of Mig-6 and Pten ablation on endometrial cancer development and progression. Oncogene. 2010;29(26):3770–80.

35. Li Z, Qu L, Zhong H, Xu K, Qiu X, Wang E. Low expression of Mig-6 is associated with poor survival outcome in NSCLC and inhibits cell apoptosis via ERK-mediated upregulation of Bcl-2. Oncol Rep. 2014;31(4):1707–14.

36. Yoo JY, Kang HB, Broaddus RR, Risinger JI, Choi KC, Kim TH. MIG-6 suppresses endometrial epithelial cell proliferation by inhibiting phospho-AKT. BMC Cancer. 2018;18(1):605.

37. Mace PD, Wallez Y, Egger MF, Dobaczewska MK, Robinson H, Pasquale EB, Riedl SJ. Structure of ERK2 bound to PEA-15 reveals a mechanism for rapid release of activated MAPK. Nature communications. 2013;4:1681.

38. Tang B, Liang W, Liao Y, Li Z, Wang Y, Yan C. PEA15 promotes liver metastasis of colorectal cancer by upregulating the ERK/MAPK signaling pathway. Oncol Rep. 2019;41(1):43–56.

39. Jiang X, Zhang C, Li W, Jiang D, Wei Z, Lv M, Xie X, Sun X. PEA15 contributes to the clinicopathology and AKT-regulated cisplatin resistance in gastric cancer. Oncol Rep. 2019;41(3):1949–59.

40. Benson FE, Baumann P, West SC. Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. Nature. 1998;391(6665):401–4.

41. Hine CM, Seluanov A, Gorbunova V. Use of the Rad51 promoter for targeted anti-cancer therapy. Proc Natl Acad Sci USA. 2008;105(52):20810–5.

42. Maacke H, Jost K, Opitz S, Miska S, Yuan Y, Hasselbach L, Luttges J, Kalthoff H, Sturzbecher HW. DNA repair and recombination factor Rad51 is over-expressed in human pancreatic adenocarcinoma. Oncogene. 2000;19(23):2791–5.
43. Qiao GB, Wu YL, Yang XN, Zhong WZ, Xie D, Guan XY, Fischer D, Kolberg HC, Kruger S, Stuerzbecher HW. High-level expression of Rad51 is an independent prognostic marker of survival in non-small-cell lung cancer patients. British journal of cancer. 2005;93(1):137–43.

44. Le Scodan R, Cizeron-Clairac G, Fourme E, Meseure D, Vacher S, Spyратos F, de la Lande B, Cvitkovic F, Lidereau R, Bieche I. DNA repair gene expression and risk of locoregional relapse in breast cancer patients. Int J Radiat Oncol Biol Phys. 2010;78(2):328–36.

45. Xu ZY, Loignon M, Han FY, Panasci L, Aloyz R. Xrcc3 induces cisplatin resistance by stimulation of Rad51-related recombinational repair, S-phase checkpoint activation, and reduced apoptosis. J Pharmacol Exp Ther. 2005;314(2):495–505.

**Figures**

![Study flow diagram](image)

**Figure 1**

Study flow.
Figure 2

Identification of 5 significantly prognostic proteins in KIRC. (a) Volcano plot of the significantly prognostic proteins in KIRC. Log2 HR is displayed on the x-axis and the -log10 (p value) on the y-axis. P value < 0.05 was considered statistically significant. Proteins with P value < 0.0001 were labeled. (b) LASSO coefficient profiles and deviance profiles of the 77 prognostic proteins. (c) The representative protein expression of the five proteins (data of IGFBP2 and MIG6 was not found) in KIRC and normal kidney tissue. Data were from the Human Protein Atlas online database (http://www.proteinatlas.org). KIRC, kidney renal clear cell carcinoma; HR, hazard ratio.
Figure 3

Risk score analyses of five-protein model. Risk score distribution of patients in the training set (a), testing set (b) and whole set (c). Survival status of the patients in the training set (d), testing set (e) and whole set (f). Patients in high-risk and low-risk group were on the right and left side of dotted line, respectively. Heatmap of the five prognostic proteins in the training set (g), testing set (h) and whole set (i).
Figure 4

Prognostic analyses of five-protein model. Kaplan-Meier survival analysis in the training set (a), testing set (b) and whole set (c). Time-dependent ROC analysis in the training set (d), testing set (e) and whole set (f). ROC, receiver operating characteristic.
Figure 5

Identifying the independent prognostic parameters and construction of protein-based nomogram. (a) Forrest plot of univariate Cox regression analysis. (b) Forrest plot of multivariate Cox regression analysis. (c) Nomogram integrated protein-based risk score, age and AJCC stage. KIRC, kidney renal clear cell carcinoma; AJCC, American Joint Committee on Cancer.
Figure 6

Stratified survival analysis of the five-protein model and the association between the risk score and clinicopathological features. Stratified survival analysis according to stage (a), grade (b), age (c) and gender (d). The association between the risk score and stage (e), grade (f), age (g) and gender (h).
Figure 7

Performance of protein-based nomogram in predicting OS for KIRC and comparison of the predictive power among age, stage, risk score and protein-based nomogram. The calibration plot of the nomogram for consistency check between predictive OS and actual outcome in 1-year, 3-year and 5-year OS (a-c). The time-dependent ROC curves of the age, stage, risk score and nomogram in predicting 1-year, 3-year and 5-year OS (d-f). The decision curve analysis of the nomogram compared for 1-year, 3-year and 5-year OS (g-i). OS, overall survival; ROC, receiver operating characteristic.
**Figure 8**

Correlation analysis and functional enrichment analysis of the five proteins in the prognostic model. (a) The sankey diagram shows the correlation between the five proteins and other proteins with correlation coefficients > 0.4 and P value < 0.001. (b) Correlation between the expression of RAD51 and TIGAR, ACC1 and EEF2. (c) Top 20 of GO enrichment analysis. (d) Top 10 of KEGG enrichment analysis. GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.