A Valuable Plasma Membrane Protein Prognostic Model in Clear Cell Renal Cell Carcinoma

Haosheng Liu  
Guangdong Provincial People's Hospital  
https://orcid.org/0000-0002-0799-3880

Jianxiong Fang  
Guangdong Provincial People's Hospital

Tianqi Liu  
Guangdong Provincial People's Hospital

Zhenhui Zhang  
Guangdong Provincial People's Hospital

Chao Zhao  
Guangdong Provincial People's Hospital

Tianming Peng  
Guangdong Provincial People's Hospital

Yong Li  
Guangdong Cardiovascular Institute

Xiaoyong Pu  
Guangdong Provincial People's Hospital

Jiumin Liu  
jljumin@163.com  
Guangdong Provincial People's Hospital  
https://orcid.org/0000-0002-3058-9405

Research article

Keywords: clear cell renal cell carcinoma, plasma membrane proteins, prognostic, bioinformatic

DOI: https://doi.org/10.21203/rs.3.rs-74217/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background: Renal cell carcinoma (RCC) accounts for 3% of human cancers, and clear cell renal cell carcinoma (ccRCC) is the most common pathological type of RCC. Cell surface proteins have been shown to play an important role in the occurrence and progression of various cancers. In this study, we focused on plasma membrane proteins (PMPs), to explore their potential value in ccRCC.

Methods: The PMPs expression profiles and ccRCC patients’ clinical information were downloaded from TCGA database. Through a series of bioinformatic methods, we established a plasma membrane proteins prognostic model and verify its value in multiple ways.

Results: Multivariate cox regression analysis and area under receiver operating characteristic curve indicated that this model was an effective independent predictor of ccRCC clinical outcomes. It has good prognostic value in different groups of clinical features. Combined with other two clinical characteristics, a nomogram was constructed to predict patient survival at 1, 3, and 5 years.

Conclusions: Our study is the first to explore the prognostic value of plasma membrane proteins in clear cell renal cell carcinoma. We hope our work could provide a new viewpoint for ccRCC prognosis and draw people’s attention to plasma membrane proteins in clear cell renal cell carcinoma.

1 Background

Renal cell carcinoma represents around 3% of all cancers, with the highest incidence occurring in Western countries [1]. Generally, during the last two decades until recently, there has been an annual increase of about 2% in incidence both worldwide and in Europe leading to approximately 99,200 new RCC cases and 39,100 kidney cancer-related deaths within the European Union in 2018 [2]. With the development of human society, the problems brought by tumor are more and more prominent. Therefore, an in-depth study of valuable prognostic tools for clinical decisions is vital currently. Proteomic analysis of tumor tissue samples and recognition of potential protein biomarkers in serum or plasma is an evolving field. Direct analysis of proteins has several advantages over indirect analysis, such as transcriptome analysis, although it requires more tissue and takes more time.

In recent years, cell surface proteins have come into focus because they are readily available and have the potential to become new drug targets. Plasma membrane proteins (PMP) account for about 50% of the cell membrane weight and their functions are complex and varied [3]. Researches show that plasma membrane proteins mediate or initiate phenotypic changes associated with malignant transformation, such as cell proliferation, adhesion, and migration [4–5]. HER-2, a highly expressed receptor protein exist in many types of cancer, can promote the proliferation and invasion of tumor cells when activated [6]. Some PMPs are differentially expressed between tumor and normal tissue, which may be potential therapeutic targets or biomarkers. After analyzed VHL-associated changes in plasma membrane proteins, Aggelis V et al identified 19 differentially expressed proteins, which found to be potential biomarkers for ccRCC [7]. These researches show that the PMP disorder may be closely related to the occurrence and
development of cancer. However, as far as we know, large-scale gene expression signature has rarely been used to investigate the association between PMP and ccRCC. A more comprehensive understanding of the effects of PMP on tumors could help in the clinical diagnosis of renal cancer and even provide a new, precise direction for treatment.

In this study, we try to clarify the possible role of PMP played in ccRCC and explore their potential value in prognosis as well as targeted therapy. The PMP expression profiles and patients’ clinical information of ccRCC were downloaded from TCGA database. Then, we identified differentially expressed PMPs through computational methods. A number of bioinformatic analysis were used to study underlying regulatory mechanisms. What’s more, a valuable prognostic model was built to predict ccRCC patient's overall survival, as well as provide a new viewpoint for precise therapy of ccRCC. To verify the diagnosis effectiveness of the model, we obtained ccRCC data from GEO and ArrayExpress database. The external validation results show that this PMP based model has certain prognostic value in ccRCC. Our work reveals the prognostic value of PMPs in ccRCC for the first time. There is no doubt that more experiments are needed to further explore the underlying mechanism.

2 Materials And Methods

2.1 Data acquisition

Level 3 RNA sequencing and clinical data from 539 KIRC and 72 paracancerous samples were downloaded from The Cancer Genome Atlas (TCGA) database (https://tcga-data.nci.nih.gov/tcga/). Based on the requirement to the data integrality, patients that met the following criteria were excluded from subsequent analysis: (1) patients with survival time less than 30 days, (2) insufficient information of stage, grade, age and gender. Finally, 482 tumor samples which were from different individuals and 68 paracancerous samples were selected from the training set in this study.

Meanwhile, three microarray datasets were downloaded from GEO and ArrayExpress database (GSE29609, GSE22541, E-MTAB-3267) which includes 116 KIRC patients with corresponding clinical information for external validation as testing set. The “sva” R package was used to eliminate the batch effect.

The plasma membrane protein list was obtained from The Human Protein Atlas database (http://www.proteinatlas.org/). Since the data were downloaded from public database, the ethical approval is not required.

2.2 Differential gene analysis

To identify the differentially expressed genes (DEGs) and differentially expressed PMPs (DEPMPs), “limma” R package was used to normalized expression matrix, then make comparisons between tumor and para cancerous tissues. DEGs and DEPMPs were identified using the threshold of a log |fold change|
> 1 and an p value < 0.05. We extracted DEPMPs from all DEGs and used GO and KEGG pathway enrichment analysis to investigate DEPMPs’ molecular function.

### 2.3 Survival related DEPMPs’ molecular characteristics

We screened survival related DEPMPs through univariate cox proportional hazards regression analysis. In order to explore the clinical values of those survival related DEPMPs comprehensively, some public databases were used. The protein–protein interaction (PPI) networks was constructed by submitting gene list to the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database ([https://string-db.org/](https://string-db.org/)). Since transcription factors play an important role in the initiation of gene expression, we also built a transcription factor regulation network through Cistrome Cancer database ([http://cistrome.org/CistromeCancer/](http://cistrome.org/CistromeCancer/)), which contains over 23,000 ChIP-seq and chromatin accessibility profiles from human and mouse genomes. It provides 318 TFs’ binding information [8]. The interaction between those DEPMPs and corresponding TFs was constructed by cytoscape software.

### 2.4 Plasma membrane protein prognostic model (PMPPM)

Then, we used these survival related PMPs to construct a prognostic model according to multivariate cox proportional hazards regression analysis. After multiplied the expression level of PMPs involved in the model by their Cox regression coefficients, we obtained each patients’ Risk Score. The median of Risk Score was regarded as a cutoff to divided patients into high risk and low risk groups, survival analysis was assessed by Kaplan–Meier (K–M) methods. In addition, the predictive value of this model was evaluated by areas under the curve (AUC) of the receiver-operator characteristic (ROC) curve using package “survivalROC“ in R and multivariate cox proportional hazards regression analysis.

We further explored the relationship between the PMPPM and clinical characteristics included age, gender, grade, TNM stage, T stage and metastasis. The characteristics of each PMP involved in this model were found in UALCAN database ([http://ualcan.path.uab.edu/index.html](http://ualcan.path.uab.edu/index.html)). Besides, we obtained copy number variations information from Cbioportal database ([http://www.cbioportal.org/](http://www.cbioportal.org/))[9]. Finally, a nomogram was established to investigate patients 1-year, 2-year, and 3-year overall survival.

### 2.5 Statistical analysis

Heatmap of DEGs and DEPMPs were plotted using “pheatmap“ R package with zero-mean normalization. Two groups of boxplots were analyzed using Wilcoxon-test. “clusterProfiler” R package was used to build GO and KEGG pathway enrichment analysis. We calculated area under the ROC curve through the “survival ROC” R package. For Kaplan–Meier curves, p-values and hazard ratio (HR) with 95% confidence interval (CI) were generated by log-rank tests and univariate Cox proportional hazards regression. All analytical methods above and R packages were performed using R software version 3.6.1 (The R Foundation for Statistical Computing, 2019). All statistical tests were two-sided. P-value <0.05 was considered as statistically significant.
3 Results

3.1 Identify for DEPMPs

The patients’ clinical information is shown in Table 1. A total of 7369 DEGs were screened, which included 5467 up regulated and 1902 down regulated genes (Fig 1A). 98 up regulated and 61 down regulated DEPMPs were found in these DEGs (Fig 1B). GO enrichment analysis indicate that these DEPMPs were mainly enriched in actin filament organization. For CC, they were enriched in membrane region. Besides, they were mostly enriched in cell adhesion molecule binding in MF categories (Fig 2A). KEGG pathway enrichment analysis show that DEPMPs were mainly enriched in actin filament organization as well as regulation of actin filament organization (Fig 2B).

3.2 Characteristics of survival related PMPs

45 survival related DEPMPs (HR>1) and 56 survival related DEPMPs (HR<1) were screened out by univariate cox regression analysis. PPI network shows that those proteins interact with each other in ccRCC (Fig S1). In addition, 318 transcription factors (TFs) expression profiles were examined and 60 of them were differentially expressed in ccRCC and normal tissues (Fig S2A). Then we established a network with 101 survival related DEPMPs and these 60 TFs. The cut-off values are correlation scores > 0.4 and P-value < 0.01. The regulation diagram clearly illustrates the regulation relationship between TFs and these PMPs (Fig S2B).

3.3 Construct and analyze PMP prognostic model

We constructed a prognostic model according to multivariate Cox regression analysis' results (Table 2). ccRCC patients were divided into two groups with different clinical outcomes (Fig 3). The formula was as follows: [ULBP2 Expression level * (0.13216)] + [EPB41L4A Expression level * (-0.02912)] + [MPP5 Expression level * (-0.07191)] + [CASP4 Expression level * (0.04215)] + [ARHGEF12 Expression level * (-0.01793)] + [KNTC1 Expression level * (0.12334)] + [UBE2S Expression level * (0.03137)] + [ZNF598 Expression level * (0.06594)] + [CYFIP2 Expression level * (-0.01219)]. According to discrete clinical outcomes, the PMPPM may be a significant tool for differentiating ccRCC patients (Fig 4A). The AUC value was 0.758, indicate that the model has certain potential in survival monitoring (Fig 4B). After adjusting for age, gender, tumor grade, tumor stage, tumor size and distant metastasis status and other parameters, the multivariate cox regression analysis shows the PMPPM was an independent predict factor (Fig 4C, 4D). The Risk Score was significantly higher in advanced grade patients, advanced stage patients, distant metastasis patients (Fig S3). As to the nine gene themselves, we have shown here a box plot of the differences in the expression of these genes in tumor versus normal tissue (Fig S4A). Besides, we identified that, in mutation, Amplification was the most common type and CYFIP2 is the most frequently mutated gene. (Fig S4B). In addition, we get their protein and pan-cancer mRNA expression levels and survival analysis from UALCAN database as a supplement (Fig S5-7).

3.4 Validation of PMP prognostic model
After using R package, sva, to eliminate the batch effect, we used the RNA-sequencing data from GEO and ArrayExpress database to validate the PMP prognostic model. The risk score of every patient in testing set was calculated out as above, and the patients were divided into high-risk and low-risk groups based on the median risk score of training set. It turned out that the high-risk group also had visibly worse prognosis than the low-risk one (Fig 5A). Besides, the AUC of the ROC for risk score was 0.741, meaning it performed well in assessing and predicting the prognosis of patients with ccRCC (Fig 5B). In different groups of clinical characteristics, the low-risk group also had visibly better prognosis than the high-risk one (Fig 6,7). Taken together, we believe that the PMP prognostic model we constructed had certain efficiency and credibly clinical application value.

3.5 Predictive nomogram

All independent prognostic factors identified by multivariate Cox regression analysis included the PMPPM were used to establish a nomogram to predict patients’ overall survival at 1, 3, and 5 years (Fig8).

4 Discussion

Protein is the material basis of life, the basic organic matter that constitutes cells, and the main undertaker of life activities. Membrane proteins play an important role in many life activities of organisms, such as cell proliferation and differentiation, energy conversion, signal transduction and material transport. It is estimated that about 60% of drug targets are membrane proteins. Abnormal membrane protein expression causes a variety of diseases including cancer. In recent years, research on the structure and function of membrane proteins has become a hot topic [10]. However, the exact mechanism or the role of plasma membrane proteins in renal cell carcinoma still unclear. In this study, we downloaded a large number of ccRCC data from TCGA, which helped to obtain a comprehensive analysis of plasma membrane proteins in ccRCC patients. After compared gene expression between ccRCC and normal patients’ tissues, we identified 159 DEPMPs. GO and KEGG pathway enrichment analysis shows that these PMPs were mainly enriched in actin lament organization. Studies have found that actin lament organization plays a role in a variety of tumors such as prostate cancer, head and neck cancers and melanomas [11]. However, there is no reports about actin lament organization in renal cancer, further experimental exploration is needed in the future.

101 survival related PMPs were screened out by univariate Cox regression analysis. With the help of online websites, we could learn more about the molecular characteristics and internal or external relationships of those survival related PMPs. First, the protein-protein network shows that those PMPs interact closely with each other. Then, the TF-PMP network shows that transcription factors FOXM1, NCAPG, CENPA, MYBL2, EOMES, IRF4, IKZF1 and BATF are closely related to these PMPs. Some previously studies have shown those TFs were connected with occurrence and progression of ccRCC [12–18]. Base on the above findings, we have good reason to believe that those PMPs play a significant role in ccRCC as a whole.
In order to explore whether these survival related PMPs have prognostic value in ccRCC, we constructed a prognostic model according to multivariate cox regression analysis. Survival and ROC analysis indicated that the prognostic model shows considerable value of prognostic prediction. The positive results were also confirmed by external data. What's more, we did a comprehensive analysis of the relationship between the model and clinical parameters. The risk score was higher for advanced grade, stage, and distant metastasis patients. In addition, we also found these nine PMPs involved in this model were closely related to tumor grade, stage and distant metastasis respectively.

With the help of an online database, we explored these PMPs further. We found that mutations are common in these genes and CYFIP2 was the most frequently mutated gene. CYFIP2, cytoplasmic FMR1-interacting protein 2, was reported to be a candidate p53 target gene. CYFIP2-induced apoptosis is part of a coordinated p53-dependent response in cancer cells [20]. Nevertheless, studies on CYFIP2 in kidney cancer are rare. The survival analysis of CYFIP2 in different subgroups include tumor grade, race, and gender show that CYFIP2 was closely related to the overall survival of ccRCC patients. Due to the high frequency of CYFIP2 mutation, we think it is necessary to pay more attention to its certain mechanism in ccRCC. Current research shows that the remaining eight PMPs are all closely related to cancer too. ULBP2 expression was reported to increase in ovarian cancer cells and high expression of ULBP2 is an indicator of poor prognosis in ovarian cancer [21]. EPB41L4A is a target gene for the Wnt/β-catenin pathway. High expression of EPB41L4A indicates good survival in multiple myeloma [22]. Gene fusions are frequent early genomic rearrangements in prostate cancer. The unique in-frame MPP5–FAM71D fusion product is important for proliferation of prostate cancer cells [23]. CASP4, a member of the inflammatory caspases, has been shown to promote the proliferation of many kinds of cancer cells [24]. Xie et al found a novel SNP of ARHGEF12 that may involve ARHGEF12-RhoA-p38 signaling in erythroid regeneration in ALL patients after chemotherapy [25]. Recent advances have demonstrated that kinetochore-associated proteins are upregulated and serve significant roles in the carcinogenesis of numerous types of cancer. KNTC1 may have an essential role in mediating cell viability and apoptosis in human ESCC cells and may serve as a novel therapeutic target for esophageal squamous cell carcinoma [26]. Ubiquitin-conjugating enzyme E2S (UBE2S) knockdown suppressed the malignant characteristics of breast cancer cells, such as migration, invasion, and anchorage-independent growth [27]. In this study, we found that they were associated with overall survival in patients with ccRCC. However, up to now, existing research cannot fully explain the certain mechanism of those genes. Thus, more large sample prospective studies and basic experiment further defined the relation between kidney cancer and plasma membrane protein is needed.

It should to be noted that some limitations exist in this study. First, at present, the molecular mechanisms behind key PMPs are still covered, our study need to be validated by more experiment. Second, other researches may draw different results due to different experimental variations and statistical methods. Beyond these limitations, this study focused on potential molecular mechanisms and clinical significance of PMP, we hope this prognostic model could inspire medical scientists on ccRCC prognosis and precise therapy.
Conclusion

Clear cell renal cell carcinoma (ccRCC) is the most common pathological type of renal cell carcinoma. Many researches indicated that plasma membrane proteins take a nonnegligible part in the occurrence and progression of tumor. This work focused on potential molecular mechanisms and clinical significance of plasma membrane proteins. For the first time, we systematically and comprehensively analyzed the plasma membrane protein expression profile data of clear cell renal cell carcinoma and constructed a valuable prognosis model. We hope that this model could serve as a beneficial complement of ccRCC diagnosis and treatment.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Both RNA profiles data and clinical characteristics of ccRCC are publicly available and in open access platforms

Competing interests

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

Funding

This work was supported by Guangdong Natural Science Foundation (2019A1515012019), Science and technology program of Guangdong Province (2017ZC0275).

Authors’ contributions

The study was designed by LJ and PX. LH and FJ participated in data collection. LH, FJ, LT, ZZ, ZC, PT and LY participated in statistical analysis. LH and FJ wrote the draft, LJ and PX corrected the manuscript. All authors read and approved the final manuscript.

Jianxiong Fang and Liu Haosheng contributed equally to this paper, was the co-first author

Acknowledgements
References

1. J Ferlay, M Colombet, I Soerjomataram, T Dyba, G Randi, M Bettio, et al. Cancer incidence and
mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. Eur J Cancer,
2018. 103: 356-387. Doi: 10.1016/j.ejca.2018.07.005.

2. Levi F, Ferlay J, Galeone C, Lucchini F, Negri E, Boyle P, et al. The changing pattern of kidney cancer
incidence and mortality in Europe. BJU Int, 2008;101(8):949-58. Doi: 10.1111/j.1464-
410X.2008.07451.x

3. Leth-Larsen R, Lund RR, Ditzel HJ. Plasma Membrane Proteomics and Its Application in Clinical
Cancer Biomarker Discovery. Mol Cell Proteomics.2010;9(7) :1369-1382. Doi: 10.1074/mcp.R900006-
MCP200

4. Harvey S, Zhang Y, Landry F, Miller C, Smith, J. W. Insights into a plasma membrane signature.
Physiol. Genomics 2001;5(3):129-136. Doi: 10.1152/physiolgenomics.2001.5.3.129.

5. Gschwind A, Fischer O M, Ullrich A. The discovery of receptor tyrosine kinases: targets for cancer
therapy. Nat. Rev. Cancer 2004;4(5):361–370. Doi: 10.1038/nrc1360.

6. Alberto Ocaña, Atanasio Pandiella. Targeting HER receptors in cancer. Curr Pharm Des.
2013;19(5):808-817. Doi: 10.2174/138161213804547303.

7. Aggelis V, Craven RA, Peng, J, Harnden P, Cairns DA, Maher ER, et al. Proteomic identification of
differentially expressed plasma membrane proteins in renal cell carcinoma by stable isotope
labelling of a von Hippel-Lindau transfectant cell line model. Proteomics.2009;9(8) :2118-30.
doi:10.1002/pmic.200800756.

8. Mei S, Meyer C, Zheng R, Qin Q, Wu Q, Jiang P, et al. Cistrome Cancer: A Web Resource for Integrative
Gene Regulation Modeling in Cancer. Cancer Res. 2017;77(21): e19-22. Doi: 10.1158/0008-
5472.CAN-17-0327

9. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal:
an open platform for exploring multidimensional cancer genomics data. Cancer Discov.
2012;2(5):401–04. Doi: 10.1158/2159-8290.CD-12-0095

10. Qian H, Wu X, Du X, Yao X, Zhao X, Lee J, et al. Structural Basis of Low-pH-Dependent Lysosomal
Cholesterol Egress by NPC1 and NPC2. Cell.2020;182(1) :98-111.e18. Doi:
10.1016/j.cell.2020.05.020

11. Michelle Peckham. How myosin organization of the actin cytoskeleton contributes to the cancer
phenotype. Biochem Soc Trans.2016;44(4) :1026-34. Doi: 10.1042/BST20160034.

12. Yu Y, Lu W, Zhou X, Huang H, Shen S, Guo L. MicroRNA-132 suppresses migration and invasion of
renal carcinoma cells. J Clin Lab Anal. 2019;34: e22969. Doi:10.1002/jcla.22969.

13. Yamada Y, Arai T, Kojima S, Sugawara S, Kato M, Okato A, et al. Regulation of antitumor miR-144-5p
targets oncogenes: Direct regulation of syndecan-3 and its clinical significance. Cancer
14. Sun SS, Fu Y, Lin JY. Upregulation of MYBL2 independently predicts a poorer prognosis in patients with clear cell renal cell carcinoma. Oncol Lett. 2020;19(4):2765-2772. Doi: 10.3892/ol.2020.11408.

15. Dielmann A, Letsch A, Nonnenmacher A, Miller K, Keilholz U, Busse A. Favorable prognostic influence of T-box transcription factor Eomesodermin in metastatic renal cell cancer patients. Cancer Immunol Immunother. 2016;65(2):181-192. Doi:10.1007/s00262-015-1786-1.

16. Dannenmann SR, Thielicke J, Stöckli M, Matter C, von Boehmer L, Cecconi V, et al. Tumor-associated macrophages subvert T-cell function and correlate with reduced survival in clear cell renal cell carcinoma. Oncoimmunology. 2013;2(3): e23562. Doi: 10.4161/onci.23562.

17. Li XH, Yang CZ, Wang J. Network spatio-temporal analysis predicts disease stage-related genes and pathways in renal cell carcinoma. Genet Mol Res. 2016;15(2). Doi: 10.4238/gmr.15028061.

18. Wang B, Ran Z, Liu M, Ou Y. Prognostic Significance of Potential Immune Checkpoint Member HHLA2 in Human Tumors: A Comprehensive Analysis. Front Immunol. 2019; 10:1573. Doi: 10.3389/fimmu.2019.01573.

19. Saller E, Tom E, Brunori M, Otter M, Estreicher A, Mack DH, et al. Increased apoptosis induction by 121F mutant p53. EMBO J. 1999; 18:4424–4437. Doi: 10.1093/emboj/18.16.4424.

20. RS Jackson, Cho YJ, Stein S, Liang P. CYFIP2, a direct p53 target, is leptomycin-B sensitive. Cell Cycle. 2007; 6:95–103. Doi: 10.4161/cc.6.1.3665.

21. Li K, Mandai M, Hamanishi J, Matsumura N, Suzuki A, Yagi H, et al. Clinical significance of the NKG2D ligands, MICA/B and ULBP2 in ovarian cancer: high expression of ULBP2 is an indicator of poor prognosis. Cancer Immunol Immunother. 2009; 58(5):641-52. Doi: 10.1007/s00262-008-0585-3.

22. Zhang W, Lai R, He X, Liu X, Zhang Y, Yang Z, et al. Clinical prognostic implications of EPB41L4A expression in multiple myeloma. J Cancer. 2020; 11(3):619-629. Doi: 10.7150/jca.33805.

23. Teles Alves I, Hartjes T, McClellan E, Hiltemann S, Böttcher R, Dits N, et al. Next-generation sequencing reveals novel rare fusion events with functional implication in prostate cancer. Oncogene. 2015;34(5):568-77. Doi: 10.1038/onc.2013.591

24. Papoff G, Presutti D, Lalli C, Bolasco G, Santini S, Manelfi C, et al. CASP4 gene silencing in epithelial cancer cells leads to impairment of cell migration, cell-matrix adhesion and tissue invasion. Sci Rep. 2018; 8(1):17705. Doi: 10.1038/s41598-018-35792-8.

25. Xie Y, Gao L, Xu C, Chu L, Gao L, Wu R, et al. ARHGEF12 regulates erythropoiesis and is involved in erythroid regeneration after chemotherapy in acute lymphoblastic leukemia patients. Haematologica. 2020; 105(4): 925-936. Doi: 10.3324/haematol.2018.210286

26. Liu CT, Min L, Wang YJ, Li P, Wu YD, Zhang ST, et al. shRNA-mediated knockdown of KNTC1 suppresses cell viability and induces apoptosis in esophageal squamous cell carcinoma. Int J Oncol. 2019; 54(3):1053-1060. Doi: 10.3892/ijo.2019.4672.

27. Ayesha AK, Hyodo T, Asano E, Sato N, Mansour MA, Ito S, et al. UBE2S is associated with malignant characteristics of breast cancer cells. Tumour Biol. 2016; 37(1):763-72. Doi: 10.1007/s13277-015-
### Tables

**Table 1**

Patients’ clinical information

| Items               | Group | No. of patients (%) |
|---------------------|-------|---------------------|
| Age (years)         | <65   | 304 (63.1%)         |
|                     | ≥ 65  | 178 (36.9%)         |
| Gender              | Male  | 317 (65.8%)         |
|                     | Female| 165 (34.2%)         |
| Grade               | grade 1&2 | 213 (44.2%)     |
|                     | grade 3&4 | 269 (55.8%)     |
| AJCC-stage          | stage I&II | 285 (59.1%)    |
|                     | stage III&IV | 197 (40.9%)   |
| T                   | T 1&2 | 302 (62.7%)         |
|                     | T 3&4 | 180 (37.3%)         |
| M                   | M0    | 404 (83.8%)         |
|                     | M1    | 78 (16.2%)          |

**Table 2**

Result of multivariate cox proportional hazards regression analysis

| Gene         | Coef | HR    | HR.95L | HR.95H | p-value    |
|--------------|------|-------|--------|--------|------------|
| ULBP2        | 0.13216 | 1.03064 | 1.02059 | 1.04079 | 1.58E-09  |
| EPB41L4A     | -0.02912 | 0.98098 | 0.97342 | 0.98861 | 1.17E-06  |
| MPP5         | -0.07191 | 0.97475 | 0.96333 | 0.98631 | 2.11E-05  |
| CASP4        | 0.04215  | 1.01328 | 1.00920 | 1.01738 | 1.48E-10  |
| ARHGEF12     | -0.01793 | 0.98821 | 0.97939 | 0.99711 | 9.55E-3   |
| KNTC1        | 0.12334  | 1.03338 | 1.02285 | 1.04401 | 3.31E-10  |
| UBE2S        | 0.03137  | 1.00647 | 1.00474 | 1.00820 | 2.12E-13  |
| ZNF598       | 0.06594  | 1.02362 | 1.01650 | 1.03079 | 5.57E-11  |
| CYFIP2       | -0.01289 | 1.01328 | 1.00920 | 1.01738 | 1.48E-10  |