



OPEN ACCESS

EDITED BY

Saravanan Ganesan,
Cornell University, United States

REVIEWED BY

Ezhilarasi Chendamalai,
Washington University in St. Louis,
United States

Zeyan Li,
Shandong University, China

*CORRESPONDENCE

Guangcheng Luo,
✉ lgch@xmu.edu.cn

[†]These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION

This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 16 September 2022

ACCEPTED 27 December 2022

PUBLISHED 12 January 2023

CITATION

Fang J, Wang X, Xie J, Zhang X, Xiao Y, Li J and Luo G (2023), LGALS1 was related to the prognosis of clear cell renal cell carcinoma identified by weighted correlation gene network analysis combined with differential gene expression analysis. *Front. Genet.* 13:1046164. doi: 10.3389/fgene.2022.1046164

COPYRIGHT

© 2023 Fang, Wang, Xie, Zhang, Xiao, Li and Luo. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

LGALS1 was related to the prognosis of clear cell renal cell carcinoma identified by weighted correlation gene network analysis combined with differential gene expression analysis

Jiang Fang^{1,2†}, Xinjun Wang^{1,3†}, Jun Xie¹, Xi Zhang¹, Yiming Xiao¹, JinKun Li¹ and Guangcheng Luo^{1,3*}

¹Zhongshan Hospital Xiamen University, School of Medicine, Xiamen University, Xiamen, China, ²Affiliated Hospital of North Sichuan Medical College, Nanchong, Sichuan, China, ³The school of Clinical Medicine, Fujian Medical University, Fuzhou, China

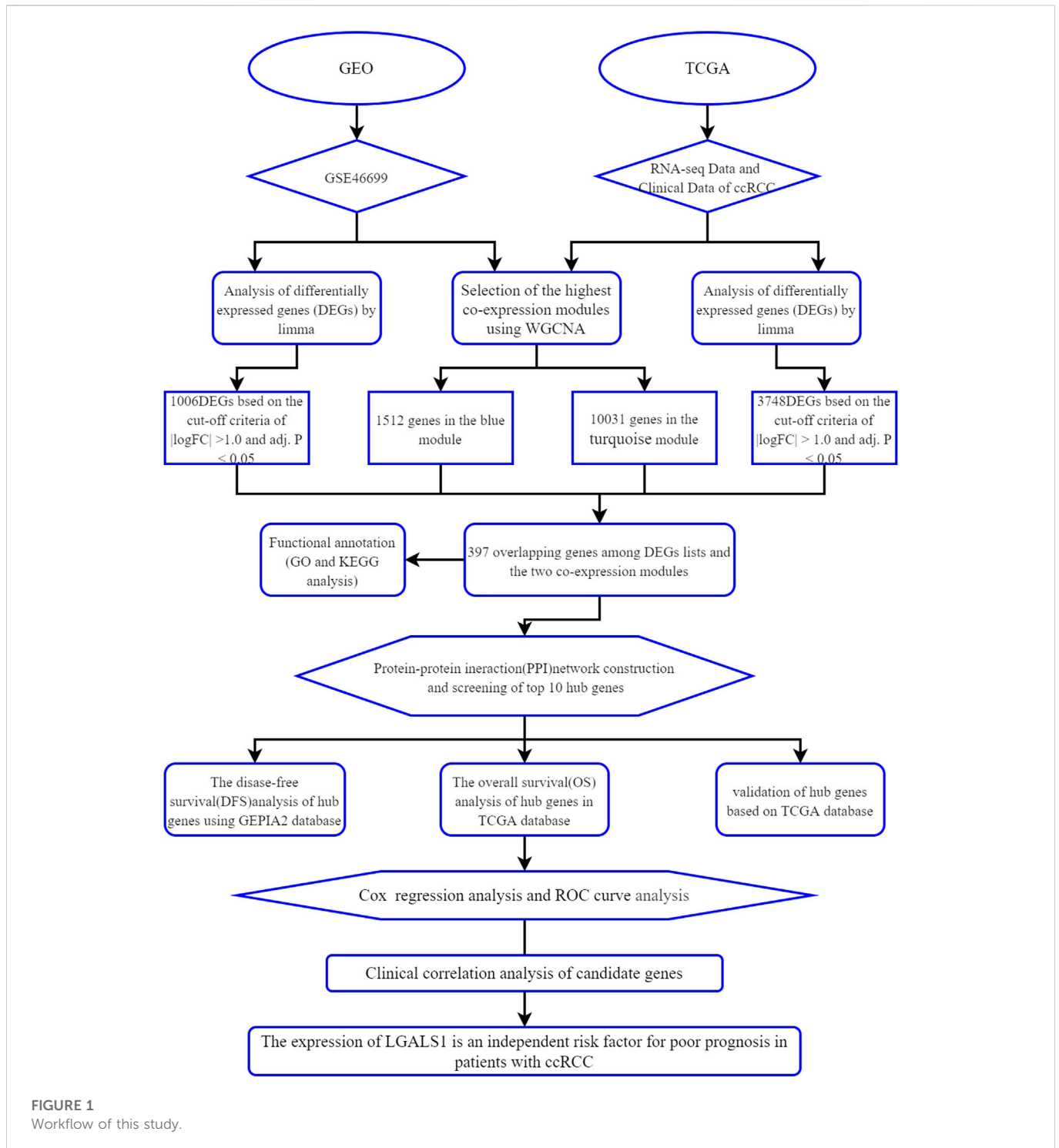
Understanding the molecular mechanism of clear cell renal cell carcinoma (ccRCC) is essential for predicting the prognosis and developing new targeted therapies. Our study is to identify hub genes related to ccRCC and to further analyze its prognostic significance. The ccRCC gene expression profiles of GSE46699 from the Gene Expression Omnibus (GEO) database and datasets from the Cancer Genome Atlas Database The Cancer Genome Atlas were used for the Weighted Gene Co-expression Network Analysis (WGCNA) and differential gene expression analysis. We screened out 397 overlapping genes from the four sets of results, and then performed Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genome (KEGG) pathways. In addition, the protein-protein interaction (PPI) network of 397 overlapping genes was mapped using the STRING database. We identified ten hub genes (KNG1, TIMP1, ALB, C3, GPC3, VCAN, P4HB, CHGB, LGALS1, EGF) using the CytoHubba plugin of Cytoscape based on the Maximal Clique Centrality (MCC) score. According to Kaplan-Meier survival analysis, higher expression of LGALS1 and TIMP1 was related to poorer overall survival (OS) in patients with ccRCC. Univariate and multivariate Cox proportional hazard analysis showed that the expression of LGALS1 was an independent risk factor for poor prognosis. Moreover, the higher the clinical grade and stage of ccRCC, the higher the expression of LGALS1. LGALS1 may play an important role in developing ccRCC and may be potential a biomarker for prognosis and treatment targets.

KEYWORDS

clear cell renal cell carcinoma, WGCNA, differential gene expression analysis, hub gene, prognosis

Introduction

Kidney cancer is one of the most common tumors, accounting for 5% and 3% of all adult malignancies in men and women, respectively, ranking sixth among men and eighth among women (Siegel et al., 2020). Renal cell carcinoma (RCC) denotes cancer originating from the renal epithelium, accounting for more than 90% of cancers in the kidney. Clear cell renal cell carcinoma (ccRCC) is the most common tumor in RCC and the cause of most cancer-related



deaths (Hsieh et al., 2017). Radiotherapy and chemotherapy are not effective for ccRCC, and surgery is the first choice for early and locally advanced ccRCC, while targeted therapies and immunotherapy are the mainstays for advanced ccRCC, including Pazopanib, Sorafenib, Tivozanib, Nivolumab, Ipilimumab, Sunitinib. (Sternberg et al., 2010; Motzer et al., 2013; Motzer et al., 2018). However, there are no available biomarkers for the prognosis and treatment of the disease at present.

Due to the recent development of high-throughput sequencing technology, genomic microarrays, and bioinformatics, bioinformatics analysis has become a new way to reveal the

pathogenesis of the disease. Weighted gene expression network analysis (WGCNA) is a systems biology method used to describe the correlation patterns of genes in microarray or RNA Sequence data. And it is an algorithm for discovering highly related gene clusters (modules) and identifying phenotypic-related modules or gene clusters (Langfelder and Horvath 2008). The hub genes and prognostic biomarkers of various cancers, including cholangiocarcinoma (Long et al., 2021), hepatocellular carcinoma (Zhang et al., 2020), and endometrial cancer (Liu et al., 2019), were identified by WGCNA.

In addition, differential gene expression analysis is another powerful analysis in transcriptomics that determines genotypic differences between two or more cell conditions to support specific hypothesis-driven research, and is one of the most common applications of RNA sequencing (RNA-seq) data and may reveal potential biomarkers of disease (McDermaid et al., 2019). Hence, the results of differential gene expression analysis and WGCNA were combined for further analysis, which can improve the recognition ability of highly related genes. The final screened genes can be used as candidate biomarkers.

In our study, differential gene expression analysis and WGCNA were performed to identify differentially co-expressed genes by using the mRNA expression data of ccRCC from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) database. GO enrichment analysis, KEGG pathway analysis, and protein-protein interaction (PPI) analysis combined with survival analysis were used to explore the development of ccRCC. Moreover, the reliability of hub genes was verified and the clinical correlation of meaningful genes was analyzed. As far as we know, this is the first time to use the data from the GEO and TCGA database for WGCNA combined with differential gene expression analysis to screen out the key genes related to ccRCC and to establish a survival model to predict the prognosis of ccRCC. The workflow of this study is displayed in Figure 1.

Materials and methods

Download and preprocess data from the GEO and TCGA databases

The Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) and the Cancer Genome Atlas (TCGA; <https://portal.gdc.cancer.gov/>) are publicly available cancer databases. Download the microarray data of ccRCC from the GEO database and the RNA-seq data and clinical information of ccRCC from the TCGA database. Firstly, we downloaded the normalized gene expression matrix files and platform annotations of the GSE46699 dataset from the GEO database. GSE46699 was composed of 66 tumor tissues and 64 normal tissues from the ccRCC samples. Then, we used the platform annotation file to convert the probe into gene symbols and removed the repeated probes of the same gene by determining the median expression value of all the corresponding probes. In consequence, a total of 21654 genes were chosen for further analysis.

The RNA sequencing (RNA-Seq) expression data of 611 samples and the corresponding clinical information of 537 cases were downloaded from the TCGA database. Among the 611 samples, 72 cases were normal and 539 cases were ccRCC. The *edgeR* R package tutorial recommends that genes with low expression do not need to be further analyzed (Robinson et al., 2010). Therefore, we used CPM (count per million) to correct the data and analyze the genes with CPM >1. We used the *rpkm* function in the *edgeR* R package to select the data by dividing the gene count by the gene length, then the *rpkm* values

TABLE 1 Clinical data clinical correlation analysis of survival related genes of TCGA-KIRC.

		Sample number	
Gender	Female		191
	Male		346
Grade	G1		14
	G2		230
	G3		207
	G4		78
	GX		5
	unknow		3
Stage	Stage I		269
	Stage II		57
	Stage III		125
	Stage IV		83
	unknow		3

were output. Finally, 14684 genes with RPKM values were obtained for the next analysis.

Identify key co-expression modules by WGCNA

In our research, the R package termed WGCNA was used to construct a gene co-expression network based on the gene expression data profiles of GSE46699 and TCGA-KIRC (Langfelder and Horvath 2008). WGCNA was used to find the modules with the most significant differences between normal and tumor samples and to extract genes from the modules. An adjacency matrix was constructed to describe the correlation strength between nodes. The formula of the adjacency matrix was as follows:

$$s_{ij} = |\text{cor}(x_i, x_j)|_{a_{ij}} = S_{ij}^{\beta}$$

where i and j represented two different genes, and X_i and X_j were their expression values, respectively. s_{ij} represented Pearson's correlation coefficient, and a_{ij} represented the strength of the correlation between two genes. β was a soft threshold, which was the Pearson correlation coefficient β of each pair of genes. Pearson correlation matrix was transformed into an adjacency matrix (scale-free network) by a β -power operation. In this study, we chose the soft power $\beta = 16$ and 2. Then, we converted the adjacency matrix into a topological overlap matrix (TOM) and the corresponding degree of dissimilarity (1-TOM). Subsequently, the hierarchical clustering dendrogram of the 1-TOM matrix was constructed, and the similar gene expression was divided into different gene co-

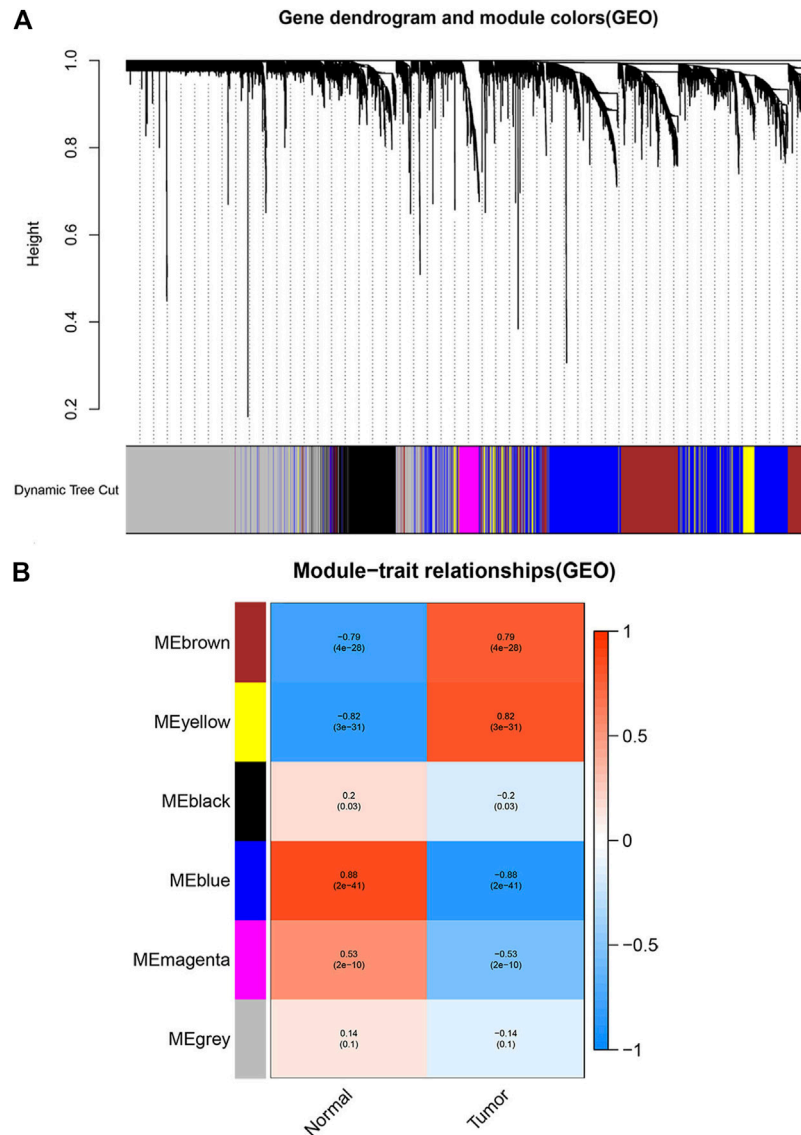


FIGURE 2

Screen the modules correlated with the clinical information from the GSE46699 dataset. **(A)** The hierarchical clustering of genes based on the 1-TOM matrix is used to sort the clustering tree of co-expression network modules. Mark each module with a different color. **(B)** Module-trait relationship in GSE46699. Each row corresponds to a color module, and each column corresponds to the tumor and normal. Each cell consists of the corresponding p -value and correlation.

expression modules, each module contained at least 50 genes. According to the previous research (Li et al., 2019), we calculated the module-trait correlation between the modules and clinical trait information to further determine the functional modules in the co-expression network. The method of dynamic tree cutting was used to identify modules from the hierarchical clustering tree, and the module eigengene (ME) of each module was calculated. ME represented the overall expression level of the module. Finally, the module with a high correlation coefficient was considered as a candidate module related to clinical traits and was chosen for follow-up analysis.

Screening the differentially expressed genes (DEGs)

In this study, the *limma* R package was used to screen the differentially expressed genes (DEGs) between ccRCC and normal kidney tissue in the data downloaded from the GEO and TCGA, respectively (Ritchie et al., 2015), using the Benjamini-Hochberg method to adjust the p -value to control the false discovery rate (FDR). The adj. $p < 0.05$ and $|\log FC| > 1.0$ were selected as cut-off criteria for DEGs. Volcano plots were drawn using the *ggplot2* R package in R version 4.03.

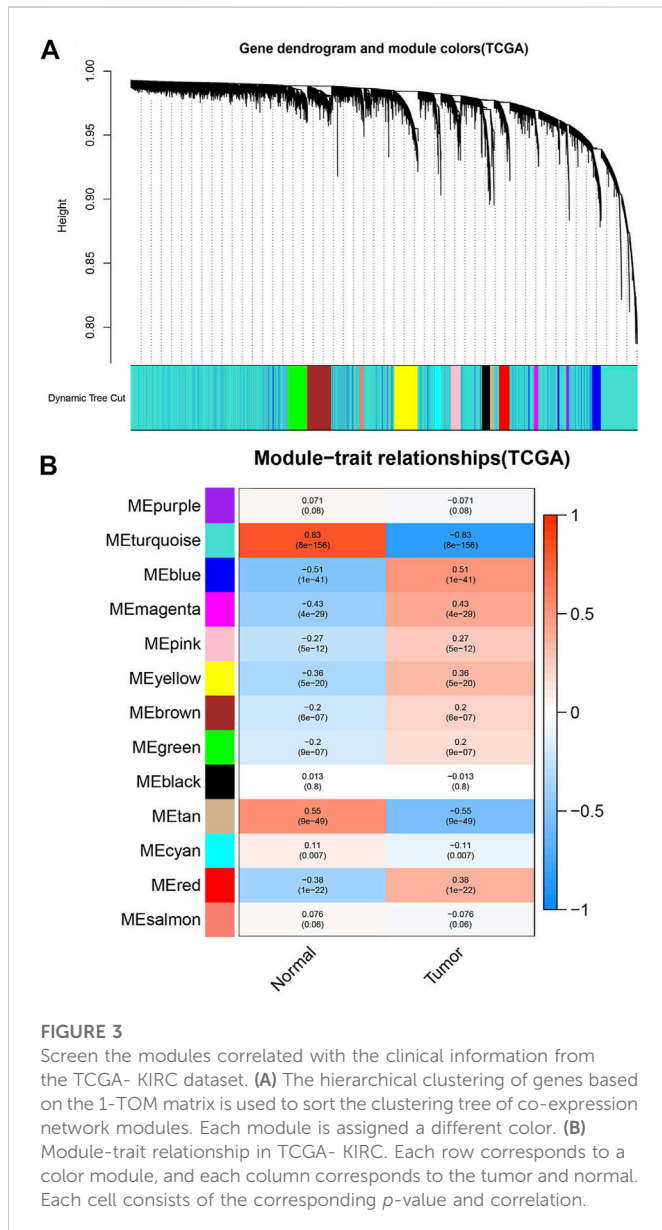


FIGURE 3
Screen the modules correlated with the clinical information from the TCGA- KIRC dataset. **(A)** The hierarchical clustering of genes based on the 1-TOM matrix is used to sort the clustering tree of co-expression network modules. Each module is assigned a different color. **(B)** Module-trait relationship in TCGA- KIRC. Each row corresponds to a color module, and each column corresponds to the tumor and normal. Each cell consists of the corresponding p -value and correlation.

Intersecting the genes in the module of interest

The overlapping genes, obtained by intersecting the DEGs with the co-expression genes extracted from the co-expression network, were used to determine potential prognostic genes. Use the *VennDiagram* R package in R software to map the genes we obtained into a Venn diagram (Chen and Boutros 2011).

Functional annotation and pathway enrichment analysis

The *clusterProfiler* R package was used to visualize gene ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway of key genes (Yu et al., 2012). The cut-off criterion was set at adjusted $p < 0.05$.

PPI network construction and hub genes identification

The key genes were imported into the *STRING* online database (<https://string-db.org/>) (Szklarczyk et al., 2019). Genes with a confidence score ≥ 0.9 were chosen to build a protein-protein interaction (PPI) network model and visualized by Cytoscape (v3.8.2) (Shannon et al., 2003). A Cytoscape plugin, cytoHubba, can be used to rank nodes in a network by their network features. Among the eleven methods to find hub nodes, Maximal Clique Centrality (MCC) was the most effective one. The MCC of each node was calculated by CytoHubba (Chin et al., 2014), and the genes with the top 10 MCC scores were identified as hub genes for this study.

Hub genes validation

To verify the reliability of the hub genes, the differentially expressed genes (DEGs) of hub genes obtained from the TCGA database were tested by Wilcoxon test in R software using the *limma* R package (Ritchie et al., 2015), and the expression level of each hub gene between normal tissues and tumor tissues was visualized as a boxplot. Set the cut-off criterion of DEGs to $p < 0.05$.

Verification of hub genes with survival analysis

To validate whether the survival of ccRCC patients was affected by hub genes, the data obtained from the TCGA, including the clinical data and gene expression data of 530 ccRCC samples, were used to seek the relationship between hub genes and Overall survival (OS) in patients by Kaplan-Meier univariate survival analysis with the survival R package in R software. In addition, using the online tool GEPIA2 (<http://gepia2.cancer-pku.cn/#index>) to determine the relationship between hub genes expressed in ccRCC patients and Disease-free survival (DFS) (Tang et al., 2019). The log-rank $p < 0.05$ of the survival-related hub genes was considered statistically significant.

Construction of the prognostic risk model

Cox regression analysis was used to construct the prognostic risk model of ccRCC, and the *survival* package in R was used for univariate and multivariate Cox proportional hazard regression analysis. The Cox analysis signature included age, gender, grade as well stages. The sensitivity and specificity of the Receiver Operation characteristic (ROC) curve were used to evaluate the prognostic value of the signature. All analyses were performed using R.

Correlation between hub genes with the clinicopathological characters

Clinicopathological data of ccRCC patients, including gender, Grade, and tumor stage were collected from the TCGA database. We conducted a correlation analysis between hub genes and clinical traits. The analysis was conducted by chi-square test under R.

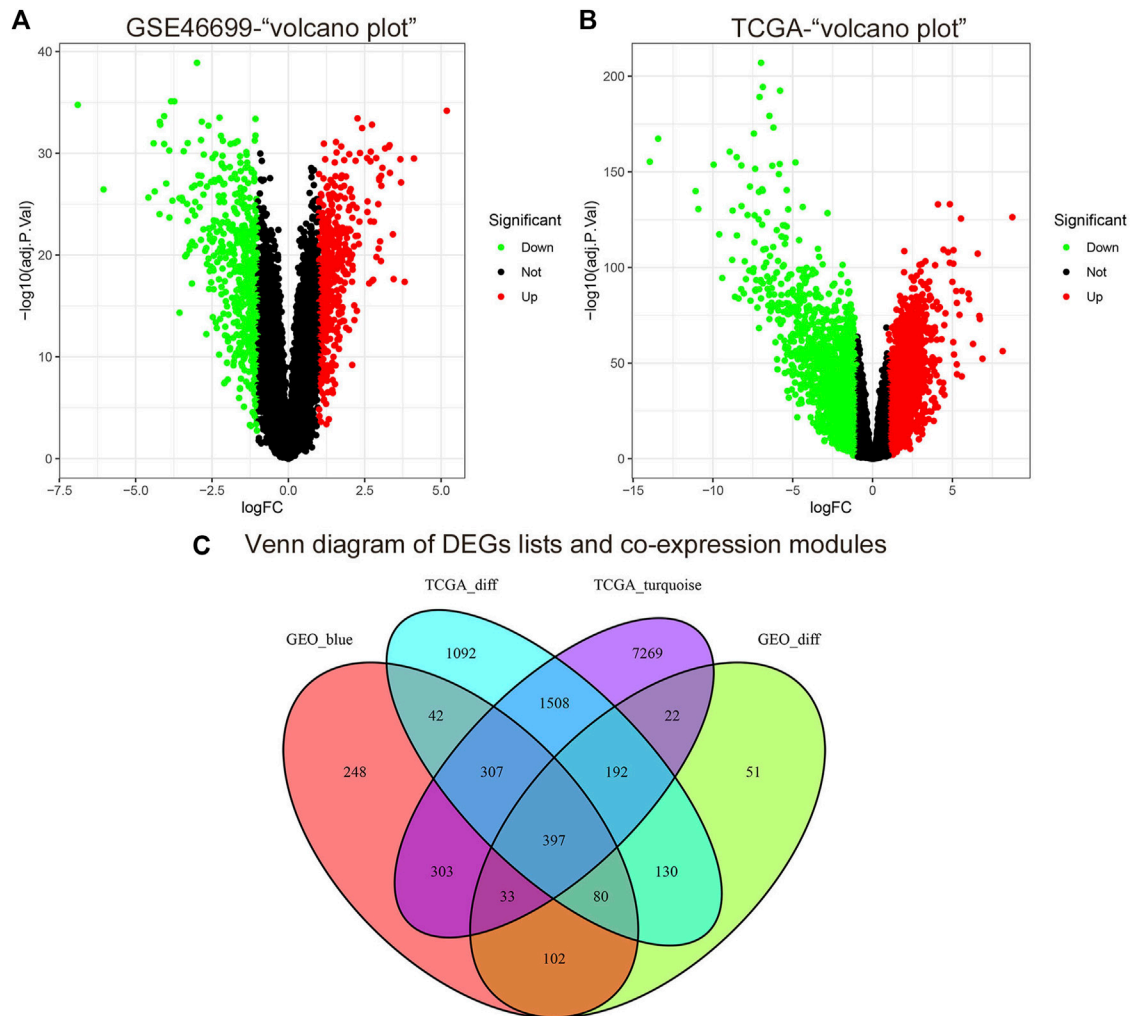


FIGURE 4

The cut-off criteria for screening differentially expressed genes in GSE46699 and TCGA datasets of ccRCC were $|\log FC| > 1.0$ and $\text{adj. } p < 0.05$. **(A)** Volcano map of differentially expressed genes in the GSE46699 data set. **(B)** Volcano map of differentially expressed genes in the TCGA dataset. **(C)** Venn diagram of gene crossover between differentially expressed genes and co-expression module. In total, 397 overlapping genes in the intersection of differentially expressed genes and two co-expression modules.

environment. And the *ggplot2* R package was used to draw boxplots. The criterion for statistical significance was set as $p < 0.05$.

Results

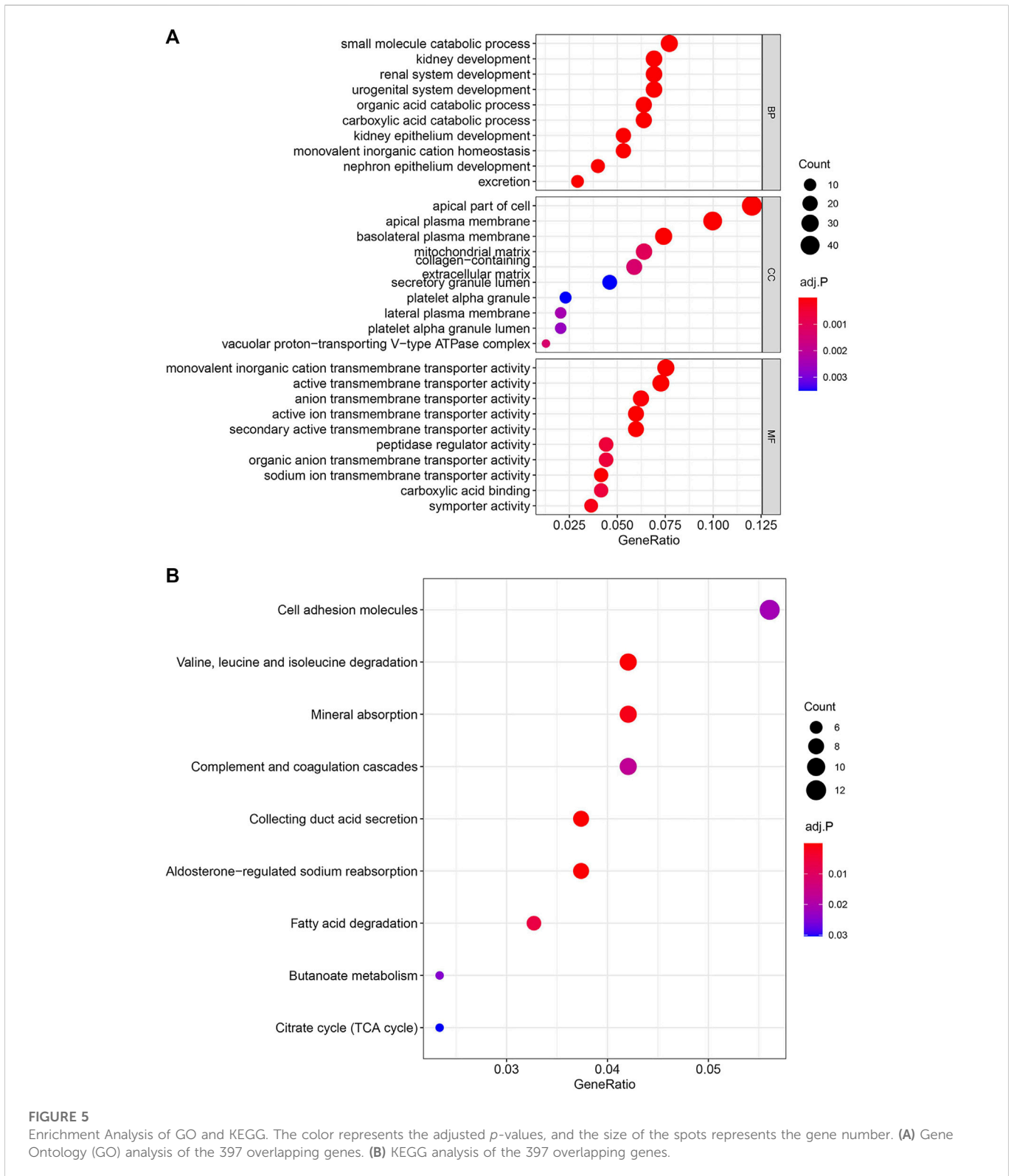
Construction of weighted gene Co-expression and division of modules

The *WGCNA* R package was used to construct the Weighted gene co-expression networks from the GSE46699 and TCGA- KIRC datasets for finding the functional clusters of ccRCC patients. The screened genes in the GSE46699 dataset were divided into five co-expression modules (Figure 2A) and the screened genes in TCGA- KIRC were divided into 13 co-expression modules (Figure 3A) (excluding the grey module, the genes of this gene set were not clustered into any module), and each module was assigned a color. After that, to assess the correlation between two clinical traits (normal and tumor) and each module, we drew the

module-traits relationships heatmap. Figure 2B; Figure 3B showed the module-trait relationships. The blue module of the GSE46699 dataset and the turquoise module of TCGA- KIRC have the highest correlation with normal tissues (blue module: $r = .88$, $p = 2e-41$; turquoise module: $r = .83$, $p = 8e-156$), the blue module of the GSE46699 and the turquoise module of TCGA-KIRC included 1,512 and 10,031 co-expression genes, respectively.

DEGs screening and identification of common genes between co-expression modules and DEGs

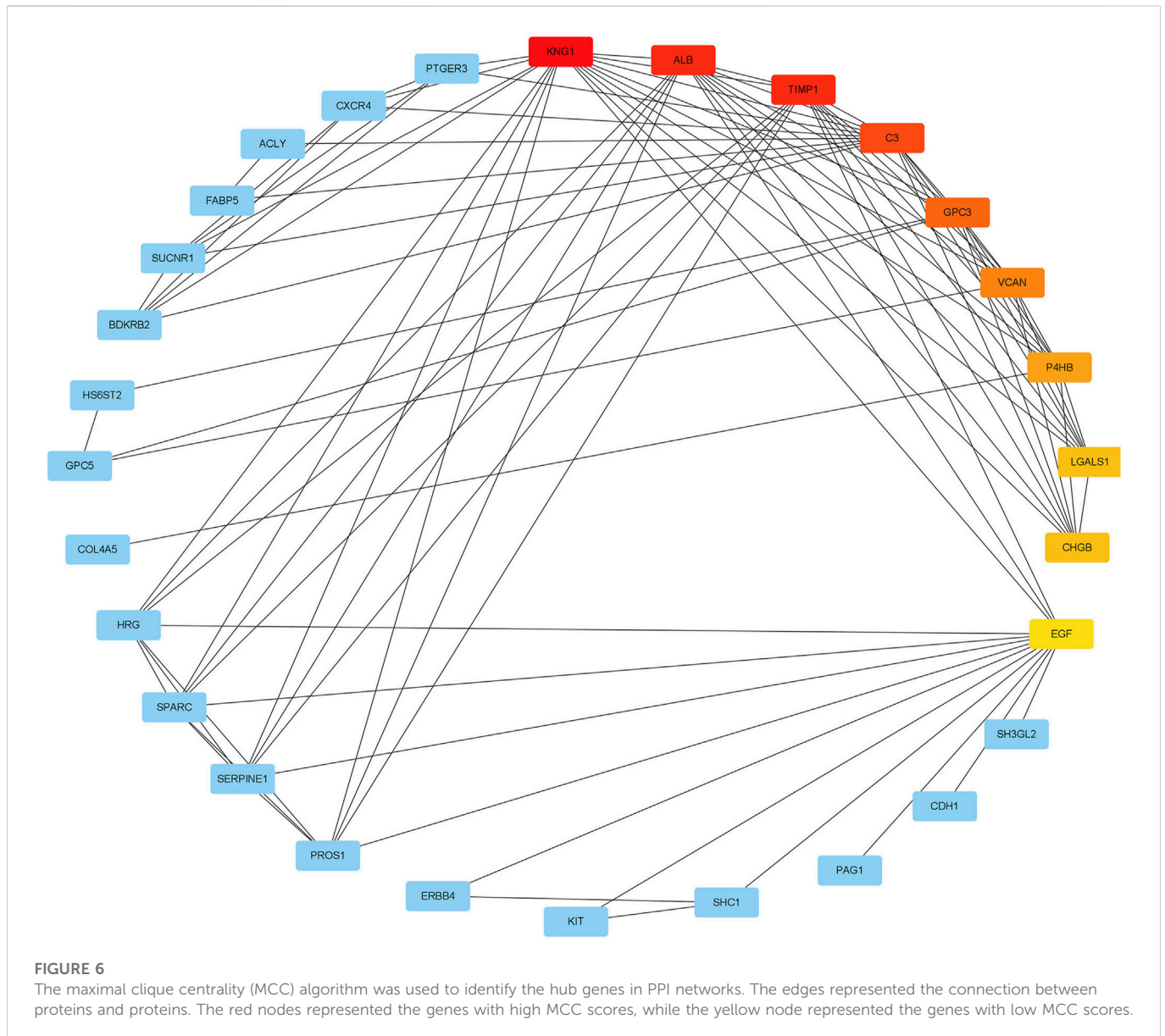
According to the cut-off criteria of $|\log FC| \geq 1.0$ and $\text{adj. } p < .05$, we used the *limma* R package to find that 1,007 DEGs in the GSE46699 dataset (Figure 4A) and 3,748 DEGs in TCGA-KIRC (Figure 4B) were abnormally regulated in tumor tissues. Then, we got a Venn diagram using the *VennDiagram* R package (Figure 4C). In total, 397 overlapping genes were extracted for further study.



Enrichment analysis of GO and KEGG

To explore the biological processes and signal pathways of the main enrichment of these 397 genes, the *clusterProfiler* R package was used for GO and KEGG analysis. The GO enrichment analysis includes three parts: molecular function (MF), biological process (BP), and cell component (CC). GO

analysis revealed that the most significant molecular function (MF), biological process (BP), and cell component (CC) were monovalent inorganic cation transmembrane transporter activity, small molecule catabolic process, and apical part of cell, respectively (Figure 5A). And KEGG analysis showed the signaling pathway of key genes was mostly related to the Cell adhesion molecules (Figure 5B).



Construction of PPI network and identification of hub genes

We used the *STRING* database to construct a PPI network between overlapping genes. Then, the MCC algorithm of the CytoHubba plugin was used to select the hub genes from the PPI network, which was plotted in Figure 6. According to the MCC scores, the genes with the top ten highest scores were determined as hub genes, including LGALS1, TIMP1, KNG1, ALB, C3, GPC3, VCAN, P4HB, CHGB, and EGF.

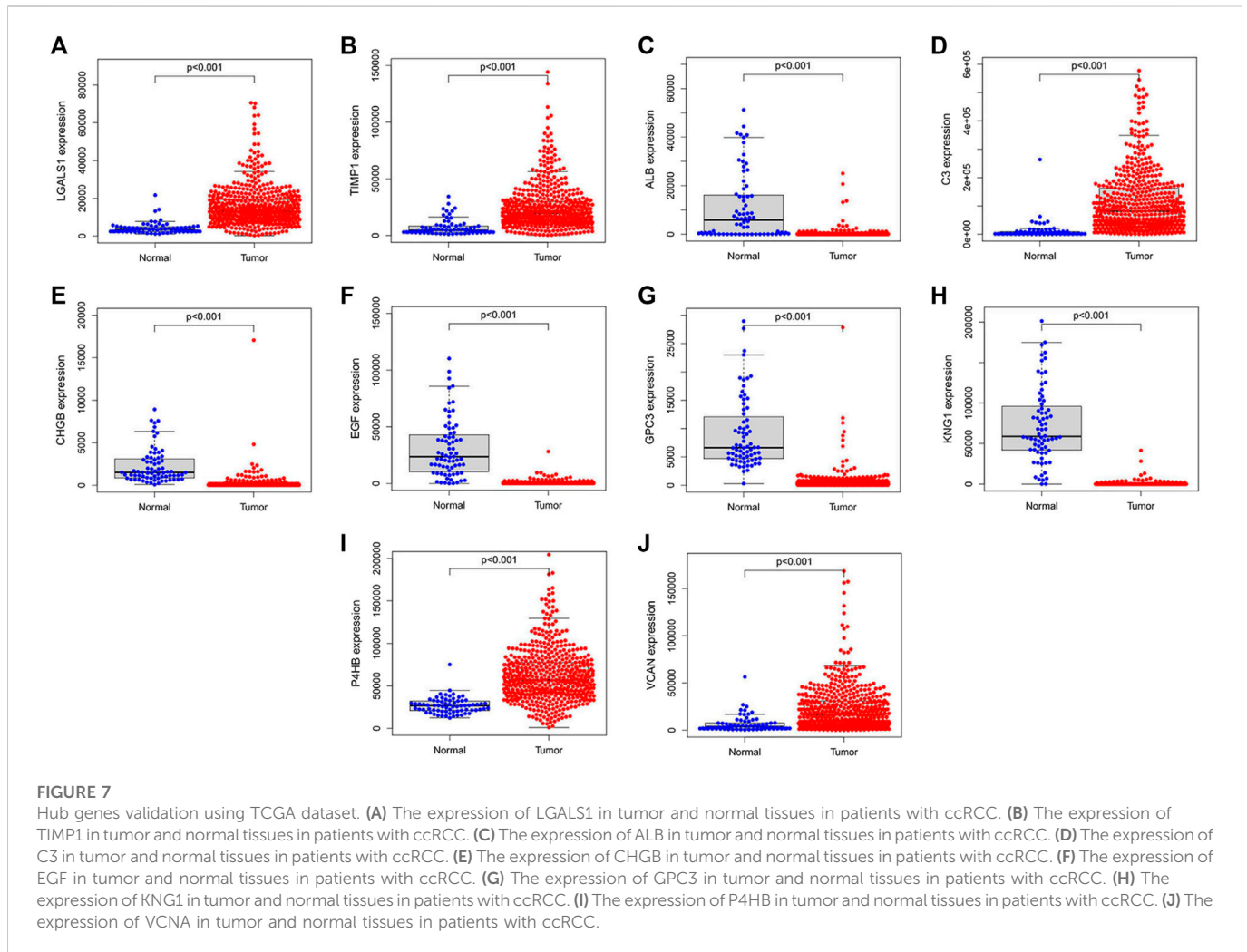
Validation and survival analysis of hub genes

We validated the expression levels of hub genes among the patients of the TCGA-KIRC dataset, which was displayed in Figure 7. We found that ten hub genes in ccRCC up-regulated significantly more than that in normal tissues. Then, we used the

survival R package and GEPIA2 database to do OS and DFS analysis on ten hub genes through the Kaplan-Meier curve, to explore prognostic values of hub genes in ccRCC patients. As shown in Figure 8, the results illustrated that LGALS1 and TIMP1 were significantly associated with the overall survival of the ccRCC patients ($p < 0.05$), and higher expression of LGALS1 and TIMP1 was related to poorer overall survival (OS) in patients with ccRCC. And Disease-Free survival in patients with ccRCC was associated with the expression levels of LGALS1, TIMP1, C3, CHGB, GPC3, P4HB, and VCAN ($p < .05$) (Figure 9), high levels of LGALS1, TIMP1, C3, CHGB, GPC3, P4HB, and VCAN were correlated with low Disease-Free survival in patients with ccRCC.

Verification of prognostic model

Univariate regression analysis showed that the expression of LGALS1 and TIMP1, age, grade, and stage were related to the



prognosis of ccRCC (Figures 10D, F). Multivariate regression analysis showed that the expression of LGALS1, age, grade and stage were independent risk factors correlated with the prognosis of ccRCC patients (Figures 10C, E). To evaluate the prognostic validity of the risk score, the ROC curve showed that the AUCs of 1-year, 3-year, and 5-year survival expressed by LGALS1 were .606, .587 and .601, respectively, while the expression of TIMP1 were .654, .569, and .585, respectively, indicating that the risk model was significantly effective and applicable.

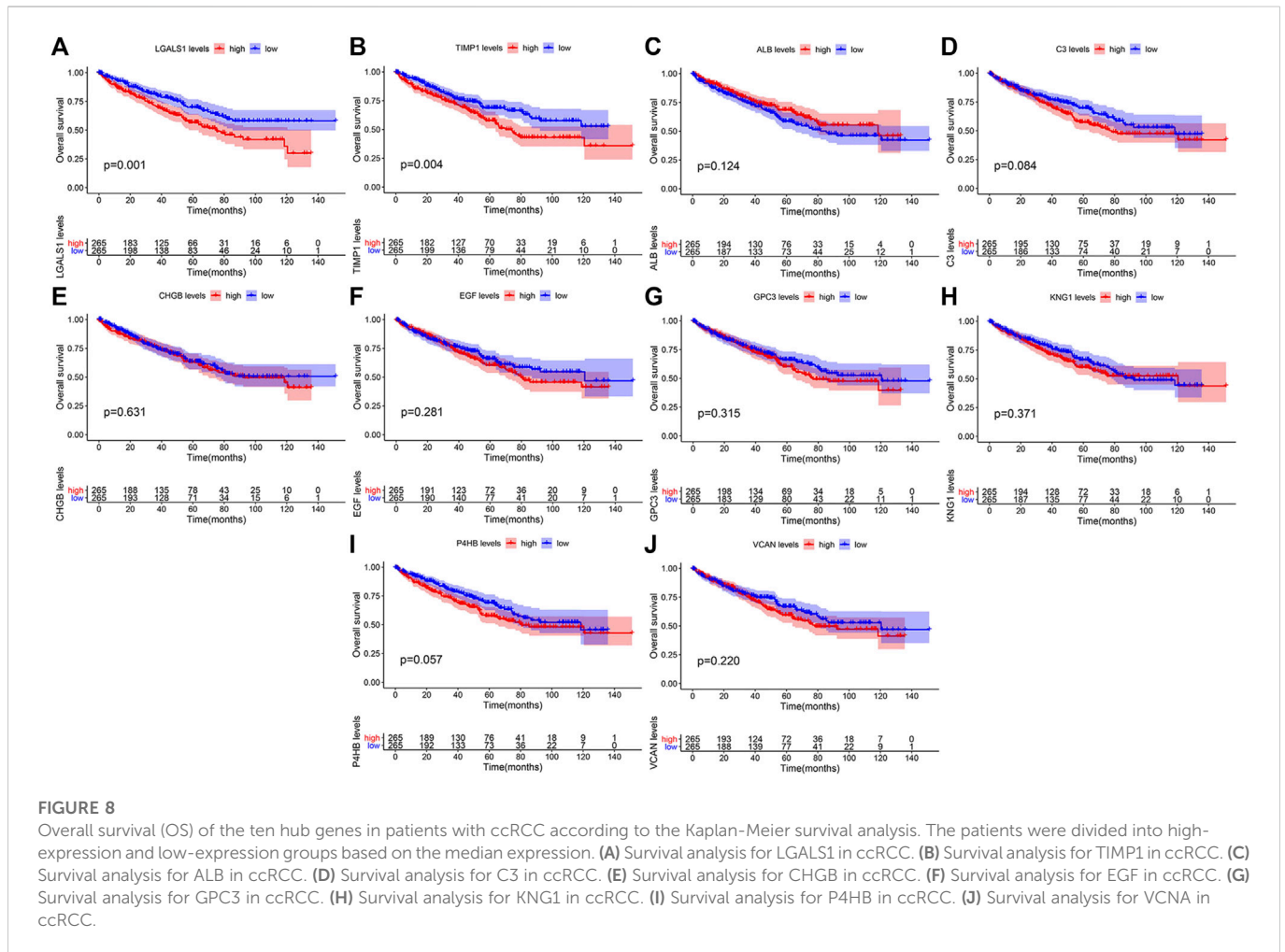
Clinical correlation analysis of survival-related genes

We analyzed the correlation between the expression of LGALS1 and clinical data downloaded from the TCGA database (Table 1). As shown in Figure 11, there were significant differences in the expression of LGALS1 from Grade 2 to Grade 3 ($p = 0.05$) and Grade 2 to Grade 4 ($p = 0.00074$). In addition, there were significant differences in the expression of LGALS1 from Stage I to Stage III ($p = 0.025$) and Stage I to Stage IV ($p = 0.0045$). Moreover, there are differences in the expression of LGALS1 ($p = 0.023$) between males

and females in ccRCC patients. These results showed that the higher the clinical grade and stage of ccRCC, the higher the expression of LGALS1. And LGALS1 was highly expressed in male patients with ccRCC.

Discussion

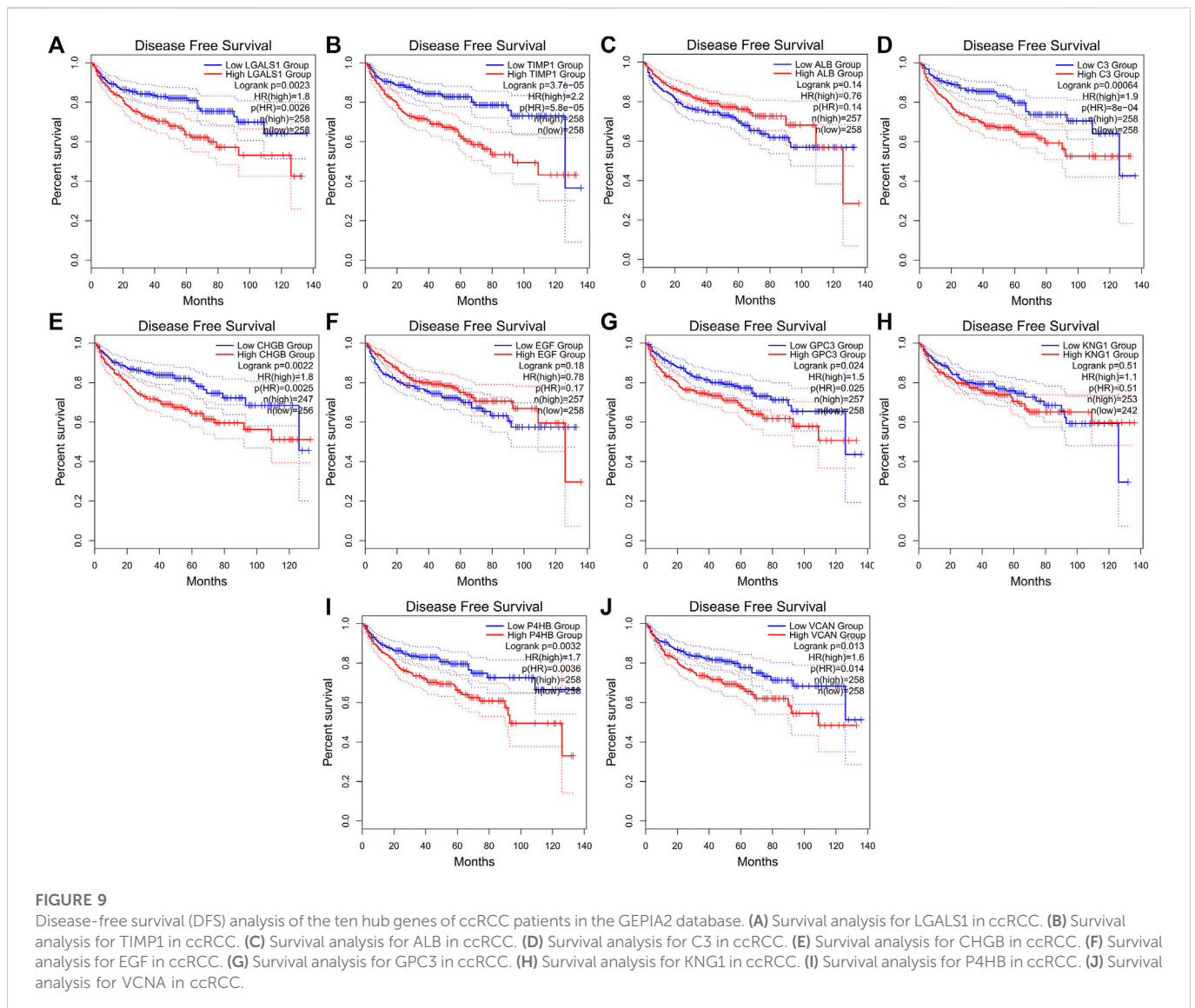
For kidney cancer, about one-third of patients were diagnosed with regional or distant metastases. The 5-Year Relative Survival of patients with distant metastasis was only 13.9%. Radical nephrectomy is the standard treatment for localized primary kidney cancer, but approximately one-quarter of these have relapses in distant sites (Choueiri and Motzer 2017). With the development of bioinformatics, although some biomarkers related to ccRCC have been found in recent years, such as AOX1 (Xiong et al., 2021), circCHST15 (Gui et al., 2021), DDX39 (Bao et al., 2021) and PPAR α (Luo et al., 2019), the risk of death in patients with ccRCC is still high, so it is necessary to find more reliable markers and obtain more treatments to reduce the risk of death in patients with ccRCC. We identified 397 differentially expressed genes with consistent expression trends in the GSE46699 and TCGA-KIRC



databases by WGCNA and Differential gene expression analysis. Furthermore, we analyzed the KEGG pathway and GO enrichment analysis of the 397 genes, and found that they play a significant role in many biological processes. Based on the MCC scores of the CytoHubba plugin in Cytoscape, the top 10 hub genes correlated with ccRCC were picked out, including, KNG1, TIMP1, ALB, C3, GPC3, VCAN, P4HB, CHGB, LGALS1, and EGF. And all ten hub genes up-regulated significantly in ccRCC compared with the normal tissues. Among ten hub genes, the high expression of LGALS1 and TIMP1 were correlated significantly with the poor OS, and the high expression of LGALS1, TIMP1, C3, CHGB, GPC3, P4HB, and VCAN were correlated with the low DFS of ccRCC significantly. Univariate and multivariate Cox proportional hazard analysis showed that LGALS1 expression was an independent risk factor for poor prognosis. Finally, the correlation between the expression of LGALS1 and clinicopathological features was analyzed.

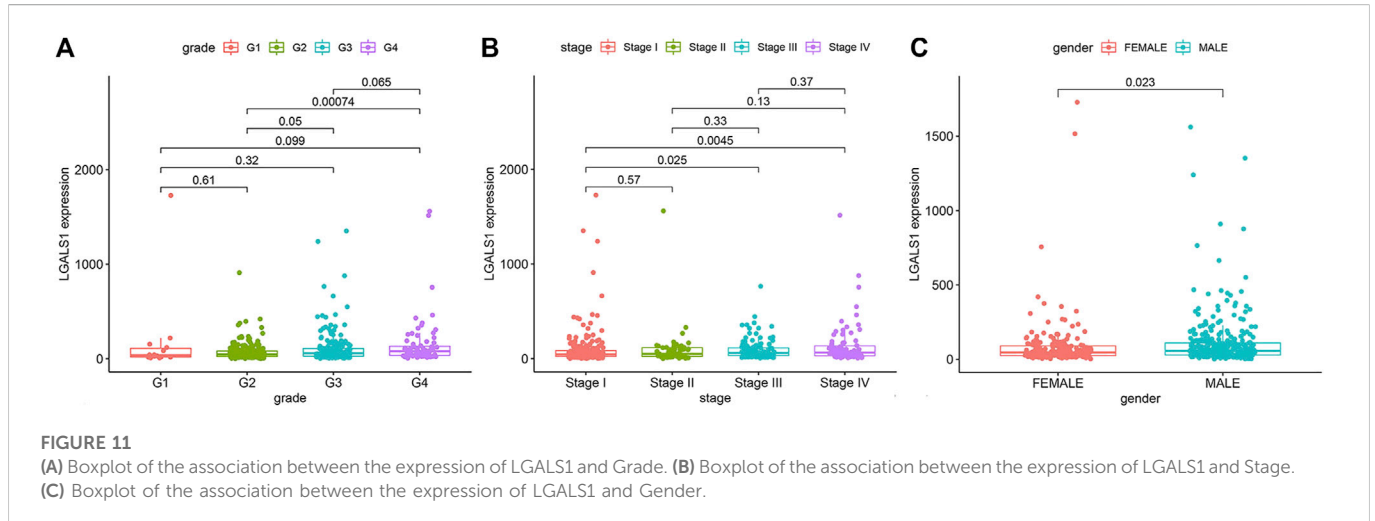
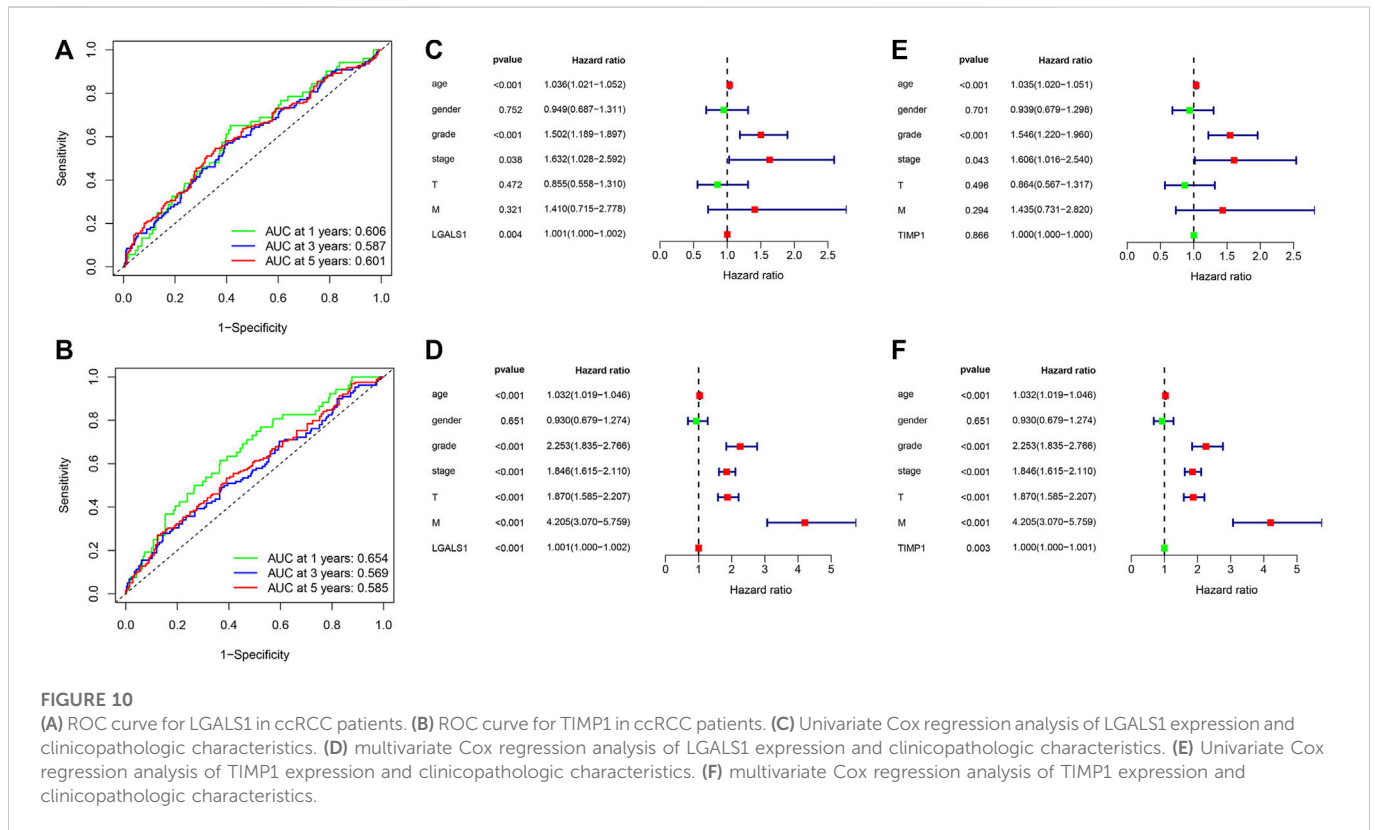
In the Genome database, the gene that encodes galectin-1 is named LGALS1 (Barondes et al., 1994). Galectin-1 is a homodimer composed of subunits of approximately 130 amino acids, abundant in skeletal, smooth, and cardiac muscle, motor and sensory neurons, thymus, kidney, and placenta. And the carbohydrate recognition domain (CRD) is responsible for β -galactoside binding (Barondes et al., 1994). Galectins play important roles in metastasis, angiogenesis,

tumor immunity, proliferation, and apoptosis (Barondes et al., 1994; Méndez-Huergo et al., 2017; Wdowiak et al., 2018; Huang et al., 2021). Galectin-1 can influence the proliferation of CD8⁺T cells and the immunosuppressive capacity of CD8⁺CD122⁺PD-1⁺Tregs, lower Galectin-1 expression result in reduced tumor growth (Cagnoni et al., 2021). Galectin-1 can interact with oncogenic RAS protein on the cell surface, affect the proliferation of tumor cells through the RAS pathway (Paz et al., 2001; Shih et al., 2019), and promote tumor progression and chemotherapy resistance by up-regulating p38 MAPK, ERK, and cyclooxygenase-2 (Chung et al., 2012). It can also promote tumor invasion and metastasis by activating the FAK/PI3K/AKT/mTOR pathway (Su et al., 2020). In addition, Galectin-1 can affect tumorigenesis by activating the Hh signal pathway (Martínez-Bosch et al., 2014). Galectin-1 is a multifunctional target during Pancreatic ductal adenocarcinoma progression (Orozco et al., 2018). Up-regulation of LGALS1 expression can not only promote the occurrence and development of non-small cell lung cancer cells (Sun et al., 2022), but also promote the proliferation and cell cycle progression of esophageal squamous cell carcinoma cells (Cui et al., 2022). The high expression of Galectin-1 is associated with the migration and invasion of gastric cancer cells and poor prognosis of patients with prostate cancer (Shih et al., 2018) and hepatocellular carcinoma (Tsai



et al., 2022). LGALS1 affects the occurrence and development of tumors in many ways, and its high expression is related to the poor prognosis of tumors, which is consistent with our result in ccRCC. Olena Masui et al. verified the abnormal protein expression of Galectin-1 in metastatic and primary renal cell carcinoma by Western blot and immunohistochemical analysis (Masui et al., 2013). N M A White et al. showed that Galectin-1 is a downstream effector molecule of miR-22 and participates in the HIF/mTOR signal axis in renal cell carcinoma (White et al., 2014). In addition, some of their findings were completely consistent with our study, including LGALS1 mRNA expression in ccRCC is significantly higher than normal kidney tissue of the same patient, patients with high LGALS1 expression are correlated with poor overall survival and disease-free survival, and compared with low-grade (grade I or II) tumors, high expression levels of LGALS1 are correlated with high-

grade tumors (grade III or IV) (White et al., 2014). Galectin-1 has emerged as a therapeutic target and reliable biomarker for a variety of tumors, which can be used to describe the clinical response and prognosis of patients. It is also expected to become a therapeutic target and prognostic marker for ccRCC. Some galectin-1 inhibitors have been developed for the treatment of tumors. For example, the Galectin-1 inhibitor showed significant anti-cancer effects both *in vitro* and *in vivo* in thyroid cancer lines expressing Gal-1 (Gheysen et al., 2021), and the Galectin-1 inhibitor had been used to treat B-cell precursor acute lymphoblastic leukemia (Paz et al., 2018) and head and neck squamous cell carcinomas (Koonce et al., 2017). Moreover, the Galectin-1 inhibitor combined with sorafenib was used in the treatment of liver cancer (Leung et al., 2019). The new strategy of targeting Galectin-1 in the treatment of ccRCC needs more research and is expected to improve the prognosis of ccRCC patients.



Nevertheless, there were some limitations in this study. Even if we have performed a lot of bioinformatics analysis to determine the potential diagnostic genes of ccRCC, the number of ccRCC samples obtainable in the public database is still limited, which may lead to potential errors/biases. In addition, the molecular mechanisms of the effects of LGALS1 on the prognosis and survival of patients

with ccRCC need more experiments and verification on a larger scale.

In conclusion, we identified a hub gene, LGALS1, by WGCNA and differential gene expression analysis in patients with ccRCC, which can serve as a potential biomarker for predicting prognosis and treatment targets. Our research provided new insights

for exploring the prognosis and therapeutic targets of ccRCC patients.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

Designed the study: JF, GL, and XW. Carried out the research and analyzed the data: JF. Contributed materials/analysis tools: JX, JL, and YX. Contributed to the writing of the manuscript: JF, XW, JX, JL, YX, and XZ. Revised the manuscript: GL and XW. In addition, thanks for the GEO and TCGA databases.

Funding

This work was supported by Project of Science and Technology Bureau of Xiamen (3502Z20184033 and 3502Z20194022), Natural Science Foundation of Fujian Province (2020J011213), and the Young/Middle-aged Talent Cultivation Project funded by Fujian Provincial

References

- Bao, Y., Jiang, A., Dong, K., Gan, X., Gong, W., Wu, Z., et al. (2021). DDX39 as a predictor of clinical prognosis and immune checkpoint therapy efficacy in patients with clear cell renal cell carcinoma. *Int. J. Biol. Sci.* 17 (12), 3158–3172. doi:10.7150/ijbs.62553
- Barondes, S. H., Castronovo, V., Cooper, D. N., Cummings, R. D., Drickamer, K., Feizi, T., et al. (1994). Galectins: A family of animal beta-galactoside-binding lectins. *Cell* 76 (4), 597–598. doi:10.1016/0092-8674(94)90498-7
- Barondes, S. H., Cooper, D. N., Gitt, M. A., and Leffler, H. (1994). Galectins. Structure and function of a large family of animal lectins. *J. Biol. Chem.* 269 (33), 20807–20810. doi:10.1016/s0021-9258(17)31891-4
- Cagnoni, A. J., Giribaldi, M. L., Blidner, A. G., Cutine, A. M., Gatto, S. G., Morales, R. M., et al. (2021). Galectin-1 fosters an immunosuppressive microenvironment in colorectal cancer by reprogramming CD8(+) regulatory T cells. *Proc. Natl. Acad. Sci. U. S. A.* 118 (21), e2102950118. doi:10.1073/pnas.2102950118
- Chen, H., and Boutros, P. C. (2011). VennDiagram: A package for the generation of highly-customizable Venn and euler diagrams in R. *BMC Bioinforma.* 12, 35. doi:10.1186/1471-2105-12-35
- Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8, S11. doi:10.1186/1752-0509-8-S4-S11
- Chouei, T. K., and Motzer, R. J. (2017). Systemic therapy for metastatic renal-cell carcinoma. *N. Engl. J. Med.* 376 (4), 354–366. doi:10.1056/NEJMra1601333
- Chung, L. Y., Tang, S. J., Sun, G. H., Chou, T. Y., Yeh, T. S., Yu, S. L., et al. (2012). Galectin-1 promotes lung cancer progression and chemoresistance by upregulating p38 MAPK, ERK, and cyclooxygenase-2. *Clin. Cancer Res.* 18 (15), 4037–4047. doi:10.1158/1078-0432.CCR-11-3348
- Cui, Y., Yan, M., Wu, W., Lv, P., Wang, J., Huo, Y., et al. (2022). ESCCAL-1 promotes cell-cycle progression by interacting with and stabilizing galectin-1 in esophageal squamous cell carcinoma. *NPJ Precis. Oncol.* 6 (1), 12. doi:10.1038/s41698-022-00255-x
- Gheysen, L., Soumoy, L., Trelcat, A., Verset, L., Journe, F., and Saussez, S. (2021). New treatment strategy targeting galectin-1 against thyroid cancer. *Cells* 10 (5), 1112. doi:10.3390/cells10051112
- Gui, C. P., Liao, B., Luo, C. G., Chen, Y. H., Tan, L., Tang, Y. M., et al. (2021). circCHST15 is a novel prognostic biomarker that promotes clear cell renal cell carcinoma cell proliferation and metastasis through the miR-125a-5p/EIF4EBP1 axis. *Mol. Cancer* 20 (1), 169. doi:10.1186/s12943-021-01449-w

Health and Family Planning Commission and Xiamen Health and Family planning Commission (2021GGB028).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1046164/full#supplementary-material>

- Hsieh, J. J., Purdue, M. P., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., et al. (2017). Renal cell carcinoma. *Nat. Rev. Dis. Prim.* 3, 17009. doi:10.1038/nrdp.2017.9
- Huang, Y., Wang, H. C., Zhao, J., Wu, M. H., and Shih, T. C. (2021). Immunosuppressive roles of galectin-1 in the tumor microenvironment. *Biomolecules* 11 (10), 1398. doi:10.3390/biom11101398
- Koonce, N. A., Griffin, R. J., and Dings, R. P. M. (2017). Galectin-1 inhibitor OTX008 induces tumor vessel normalization and tumor growth inhibition in human head and neck squamous cell carcinoma models. *Int. J. Mol. Sci.* 18 (12), 2671. doi:10.3390/ijms18122671
- Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559
- Leung, Z., Ko, F. C. F., Tey, S. K., Kwong, E. M. L., Mao, X., Liu, B. H. M., et al. (2019). Galectin-1 promotes hepatocellular carcinoma and the combined therapeutic effect of OTX008 galectin-1 inhibitor and sorafenib in tumor cells. *J. Exp. Clin. Cancer Res.* 38 (1), 423. doi:10.1186/s13046-019-1402-x
- Li, W., Lu, J., Ma, Z., Zhao, J., and Liu, J. (2019). An integrated model based on a six-gene signature predicts overall survival in patients with hepatocellular carcinoma. *Front. Genet.* 10, 1323. doi:10.3389/fgene.2019.01323
- Liu, J., Zhou, S., Li, S., Jiang, Y., Wan, Y., Ma, X., et al. (2019). Eleven genes associated with progression and prognosis of endometrial cancer (EC) identified by comprehensive bioinformatics analysis. *Cancer Cell Int.* 19, 136. doi:10.1186/s12935-019-0859-1
- Long, J., Huang, S., Bai, Y., Mao, J., Wang, A., Lin, Y., et al. (2021). Transcriptional landscape of cholangiocarcinoma revealed by weighted gene coexpression network analysis. *Brief. Bioinform* 22 (4), bbaa224. doi:10.1093/bib/bbaa224
- Luo, Y., Chen, L., Wang, G., Qian, G., Liu, X., Xiao, Y., et al. (2019). PPAR α gene is a diagnostic and prognostic biomarker in clear cell renal cell carcinoma by integrated bioinformatics analysis. *J. Cancer* 10 (10), 2319–2331. doi:10.7150/jca.29178
- Martínez-Bosch, N., Fernández-Barrena, M. G., Moreno, M., Ortiz-Zapater, E., Munné-Collado, J., Iglesias, M., et al. (2014). Galectin-1 drives pancreatic carcinogenesis through stroma remodeling and Hedgehog signaling activation. *Cancer Res.* 74 (13), 3512–3524. doi:10.1158/0008-5472.CAN-13-3013
- Masui, O., White, N. M., DeSouza, L. V., Krakovska, O., Matta, A., Metias, S., et al. (2013). Quantitative proteomic analysis in metastatic renal cell carcinoma reveals a unique set of proteins with potential prognostic significance. *Mol. Cell Proteomics* 12 (1), 132–144. doi:10.1074/mcp.M112.020701

- McDermid, A., Monier, B., Zhao, J., Liu, B., and Ma, Q. (2019). Interpretation of differential gene expression results of RNA-seq data: Review and integration. *Brief. Bioinform* 20 (6), 2044–2054. doi:10.1093/bib/bby067
- Méndez-Huergo, S. P., Blidner, A. G., and Rabinovich, G. A. (2017). Galectins: Emerging regulatory checkpoints linking tumor immunity and angiogenesis. *Curr. Opin. Immunol.* 45, 8–15. doi:10.1016/j.coi.2016.12.003
- Motzer, R. J., Nosov, D., Eisen, T., Bondarenko, I., Lesovoy, V., Lipatov, O., et al. (2013). Tivozanib versus sorafenib as initial targeted therapy for patients with metastatic renal cell carcinoma: Results from a phase III trial. *J. Clin. Oncol.* 31 (30), 3791–3799. doi:10.1200/JCO.2012.47.4940
- Motzer, R. J., Tannir, N. M., McDermott, D. F., Aren Frontera, O., Melichar, B., Choueiri, T. K., et al. (2018). Nivolumab plus Ipilimumab versus Sunitinib in advanced renal-cell carcinoma. *N. Engl. J. Med.* 378 (14), 1277–1290. doi:10.1056/NEJMoa1712126
- Orozco, C. A., Martinez-Bosch, N., Guerrero, P. E., Vinaixa, J., Dalotto-Moreno, T., Iglesias, M., et al. (2018). Targeting galectin-1 inhibits pancreatic cancer progression by modulating tumor-stroma crosstalk. *Proc. Natl. Acad. Sci. U. S. A.* 115 (16), E3769–e3778. doi:10.1073/pnas.1722434115
- Paz, A., Haklai, R., Elad-Sfadia, G., Ballan, E., and Kloog, Y. (2001). Galectin-1 binds oncogenic H-Ras to mediate Ras membrane anchorage and cell transformation. *Oncogene* 20 (51), 7486–7493. doi:10.1038/sj.onc.1204950
- Paz, H., Joo, E. J., Chou, C. H., Fei, F., Mayo, K. H., Abdel-Azim, H., et al. (2018). Treatment of B-cell precursor acute lymphoblastic leukemia with the Galectin-1 inhibitor PTX008. *J. Exp. Clin. Cancer Res.* 37 (1), 67. doi:10.1186/s13046-018-0721-7
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26 (1), 139–140. doi:10.1093/bioinformatics/btp616
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303
- Shih, T. C., Fan, Y., Kiss, S., Li, X., Deng, X. N., Liu, R., et al. (2019). Galectin-1 inhibition induces cell apoptosis through dual suppression of CXCR4 and Ras pathways in human malignant peripheral nerve sheath tumors. *Neuro Oncol.* 21 (11), 1389–1400. doi:10.1093/neuonc/noz093
- Shih, T. C., Liu, R., Wu, C. T., Li, X., Xiao, W., Deng, X., et al. (2018). Targeting galectin-1 impairs castration-resistant prostate cancer progression and invasion. *Clin. Cancer Res.* 24 (17), 4319–4331. doi:10.1158/1078-0432.CCR-18-0157
- Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics. *CA Cancer J. Clin.* 70 (1), 7–30. doi:10.3322/caac.21590
- Sternberg, C. N., Davis, I. D., Mardiak, J., Szczylak, C., Lee, E., Wagstaff, J., et al. (2010). Pazopanib in locally advanced or metastatic renal cell carcinoma: Results of a randomized phase III trial. *J. Clin. Oncol.* 28 (6), 1061–1068. doi:10.1200/JCO.2009.23.9764
- Su, Y. L., Luo, H. L., Huang, C. C., Liu, T. T., Huang, E. Y., Sung, M. T., et al. (2020). Galectin-1 overexpression activates the FAK/PI3K/AKT/mTOR pathway and is correlated with upper urinary urothelial carcinoma progression and survival. *Cells* 9 (4), 806. doi:10.3390/cells9040806
- Sun, H., Zhang, H., Yan, Y., Li, Y., Che, G., Zhou, C., et al. (2022). NCAPG promotes the oncogenesis and progression of non-small cell lung cancer cells through upregulating LGALS1 expression. *Mol. Cancer* 21 (1), 55. doi:10.1186/s12943-022-01533-9
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47 (D1), D607–D613. doi:10.1093/nar/gky1131
- Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* 47 (W1), W556–W560. doi:10.1093/nar/gkz430
- Tsai, Y. T., Li, C. Y., Huang, Y. H., Chang, T. S., Lin, C. Y., Chuang, C. H., et al. (2022). Galectin-1 orchestrates an inflammatory tumor-stroma crosstalk in hepatoma by enhancing TNFR1 protein stability and signaling in carcinoma-associated fibroblasts. *Oncogene* 41 (21), 3011–3023. doi:10.1038/s41388-022-02309-7
- Wdowiak, K., Francuz, T., Gallego-Colon, E., Ruiz-Agamez, N., Kubiczko, M., Grochoła, I., et al. (2018). Galectin targeted therapy in oncology: Current knowledge and perspectives. *Int. J. Mol. Sci.* 19 (1), 210. doi:10.3390/ijms19010210
- White, N. M., Masui, O., Newsted, D., Scorilas, A., Romaschin, A. D., Bjarnason, G. A., et al. (2014). Galectin-1 has potential prognostic significance and is implicated in clear cell renal cell carcinoma progression through the HIF/mTOR signaling axis. *Br. J. Cancer* 110 (5), 1250–1259. doi:10.1038/bjc.2013.828
- Xiong, L., Feng, Y., Hu, W., Tan, J., Li, S., and Wang, H. (2021). Expression of AOX1 predicts prognosis of clear cell renal cell carcinoma. *Front. Genet.* 12, 683173. doi:10.3389/fgene.2021.683173
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287. doi:10.1089/omi.2011.0118
- Zhang, Q., Wang, J., Liu, M., Zhu, Q., Li, Q., Xie, C., et al. (2020). Weighted correlation gene network analysis reveals a new stemness index-related survival model for prognostic prediction in hepatocellular carcinoma. *Aging (Albany NY)* 12 (13), 13502–13517. doi:10.18632/aging.103454