

A systematic analysis of molecular mechanisms in non-metastatic renal cancer delineates affected regulatory pathways and genes in tumor growth

Rongjun Xia, Lijuan Lin*, Shengjin Yu, Jinhui Zhang, Linlin Zheng, Lijie Zhou, Junrong Lin

School of Medicine, Eastern Liaoning University, Dandong, 118003, China

ARTICLE INFO

Original paper

Article history:

Received: July 02, 2021

Accepted: September 29, 2021

Published: December 15, 2021

Keywords:

Kidney renal cancer; Benign tumor, PPI network; Hub genes; Expression analysis

ABSTRACT

In the recent century, Kidney cancer has emerged as one of the critical renal diseases. Therefore, we analyzed gene expression profiles of non-metastatic kidney cancer to find mechanisms associated with the early-stage pathogenesis of the disease. We concentrated on the most dysregulated genes in expression to discover possible unknown proliferative molecular mechanisms and oncogenic pathways promoting kidney renal cancer growth. Survival analysis, expression profiling, and gene set over-representation analysis were conducted on the most upregulated and most down-regulated genes alongside the hub genes. Our results demonstrated that pathways engaged in the metabolism of amino acids and carbohydrates and those involved in peroxisome organization were shown to be important in developing benign tumors. Furthermore, upregulation of genes such as CXCL9 and 10 genes and CXCR4 in chemokine response pathways would bolster differentiation and engagement of immune cells in the tumor microenvironment. C3, one of the essential members of the complement system, with a high degree and betweenness centrality in the PPI network, upregulated significantly not only in our analysis but also in the validation expression profiling results and survival analysis. We also identified genes such as TYROBP, ITGB2, and EGFR to be engaged in both immunological pathways and superoxide pathways. Furthermore, we found that downregulation of Aldolase B engaged in Glycolysis and Gluconeogenesis pathways would help develop benign tumors. Finally, many top hub genes, including TYMS, PTPRC, AURKA, FN1, UBE2C, and CD53 were proposed to be engaged in the progression of non-metastatic renal tumors. This holistic interrogation calls attention to investigate further and experimentally validate the proposed molecular mechanisms.

DOI: <http://dx.doi.org/10.14715/cmb/2021.67.5.55>

Copyright: © 2021 by the C.M.B. Association. All rights reserved.



Introduction

Renal cell carcinoma is the most common type of kidney cancer in adults, and its incidence seems to be increasing in the current decades (1). This cancer usually does not reveal any symptoms in the early stages (2). It is not clear what leads to most renal cancers. Driver genetic mutations in the cells lead to an uncontrolled proliferative capacity of cells which in many cases converges into malignancy (3). Moreover, factors that can increase the risk of renal cancer include older age, obesity, hypertension, smoking, treatment for kidney failure, and family history of kidney cancer (4).

There are different therapeutic methods to cure localized renal cell carcinoma, such as surgery. This disease is typically resistant to some therapeutic ways like radiotherapy and chemotherapy. Thus, it is

required to evaluate more effective biomarkers. Previous studies have indicated that anti-VEGF and anti-mTOR antibodies have practical therapeutic effects (5). Immunological treatment includes high-dose interleukin-2, which stimulates T-cell proliferation and survival (6). Furthermore, blockage of programmed death-ligand 1 (PD-L1) has significantly mitigated RCC patients (7). However, it is still necessary to identify effective therapeutic methods and valuable biomarkers for diagnosis, prognostic, predictive, and curative usage. Systematic investigation of gene expression profiles is a practical approach to detect genes that participated in the development of kidney cancer.

In the current study, we intended to perform a systematic analysis of bulk gene expression data in order to find genes engaged in the early progression of the RCC. We investigated the known and unknown

*Corresponding authors: ldxylinlijuan@126.com

factors that control different steps of tumor progression. We aimed to find the pathogenic mechanisms controlling these steps before malignancy. This would aim us to introduce biomarkers and therapeutic genetic factors that can be targeted to prevent cancer progression. To this end, normal and benign samples from two microarray studies were downloaded from gene expression omnibus (GEO). Two datasets were constructed in R programming language and merged to recreate a larger dataset. This dataset was analyzed to exploit DE genes (differentially expressed genes) between human benign kidney renal cancer tissues versus non-cancerous ones. Afterward, protein-protein interaction (PPI) network analysis was assembled from these DEGs to identify hub genes. To identify the biological processes and pathways involved in the pathogenesis of kidney cancer, gene ontology (GO), and KEGG enrichment analysis were applied on the hub genes and most dysregulated DEGs. Expression profiling was recruited to validate recognized hub genes and DEGs. Only genes with absolute log fold change (LogFC) larger than one were considered as DEGs and utilized in the PPI network construction. The most upregulated and downregulated genes were recruited for expression profiling, survival analysis, and enrichment analysis. Consequently, a number of genes were identified, which would be candidate biomarkers for treatment, prognosis, and diagnosis of kidney renal cancer.

Materials and methods

Database Searching and Dataset Construction

Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) database was searched to detect the experiments containing high-quality transcriptomic samples appropriate to the study design. Searches were filtered for Homo sapiens, while kidney, renal, benign, neoplasm, and cancer were the keywords utilized in the search. Two microarray studies were selected. Fourteen Benign and fourteen matched normal samples were downloaded from the GSE66270 study. Afterward, twenty-four normal samples along with twenty-four matched renal cancer samples, at stage one of the disease, were downloaded from the GSE53757 study (8, 9). All ".CEL" raw data files were imported into R programming language version 4.0.0. A dataset was

constructed from each study sample, but after data preprocessing, the two datasets were merged to recreate a larger dataset.

Differentially Expressed Genes Identification

"affy" R package was used to perform absent probesets removal, background correction, and probe summarization (10). Outlier samples were identified using PCA and hierarchical clustering methods. Next, quantile normalization was implemented on datasets (11). Low variant probesets with standard deviation (SD) less than the median of all SDs were recognized and removed. In addition, the "Many to Many" problem, mapping multiple probesets to the same gene symbol, was settled down employing "nsFilter" function in "genefilter" R package (12) (13). "annotate" R package was used to map probesets to the gene symbols. Finally, using "limma" R package, linear regression was applied on variables to identify differentially expressed genes (14). Genes with absolute log fold change larger than one and Benjamini Hochberg adjusted p-value (15) less than 0.05 were selected as the DEGs. Then, the most upregulated and downregulated DEGs with absolute LogFC larger than 2 were selected for the over-representation pathway analysis.

Network Construction

STRING database was used to generate the Interactions from all DEGs. Interactions were emerged based on five sources of evidence namely Experiments, Databases, Co-expression, Gene fusion and Co-occurrence. Using "igraph" package in R software (16), the giant component of the network was extracted from the whole network. Next, different network descriptive and centralities were computed using the same package to first determine that it is a scale-free biological network, secondly find essential hub genes.

Enrichment Analysis

Enrichment analysis was performed using the ClueGO Cytoscape plugin (17). Over-representation analysis was performed using Fisher's exact test. Enriched terms for biological processes were obtained from the GO repository. For pathway enrichment analysis, information in KEGG (18) databases were

used. P-values were adjusted using the Benjamini-Hochberg method, and the cutoff was set at 0.05.

Expression Profiling and Survival Analysis

Genes were given to the GEPIA2 web-server to validate identified DEGs based on Datasets in TCGA and GTEx genomic databases (19-21). Expression profiles were compared between tumor and normal samples in Kidney Renal Clear Cell Carcinoma (KIRC) datasets. To create boxplots, LogFC and p-value cutoff were set at 0.5 and 0.05 respectively. TPM normalized data were log2 transformed. For survival plots, Overall Survival was selected, and the median was chosen for group cutoff with a 95% confidence interval. All KIRC datasets with monthly time-scale were selected to obtain the survival results.

Results and discussion

Data Integration and Preprocessing

A large expression matrix encompassing 38 samples in the case group, 38 samples in the control group, and 54675 probe sets at the row was assembled for downstream analysis. Figure 1 presents the PCA plot employing "prcomp" function in R to recognize the outlier samples. GSM1300068, GSM1300066 samples in the cancer group and GSM1300067 and GSM1300097 samples in the normal group lied at a distance from other group members; therefore, they were considered as the outliers.

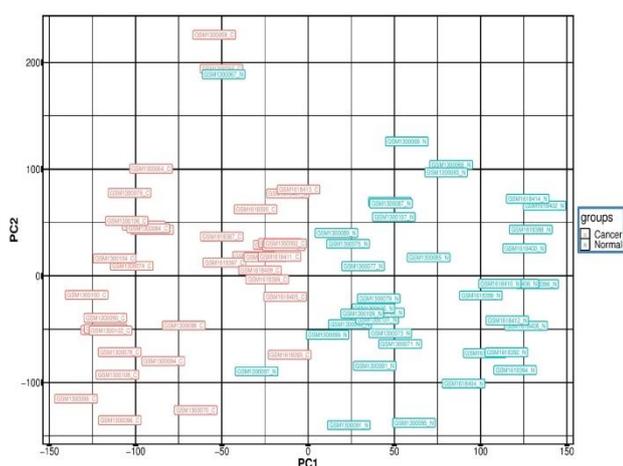


Figure 1. The PCA plot. It illustrates the first eigenvector (PC1) and second eigenvector (PC2) for cancer and normal samples. Data were scaled before plotting

To ensure the outlier samples, boxplots were created before data normalization in Figure 2. Normal

boxplots in 2A were not able to depict the batch differences. However, the RLE plot in 2B was more influential in describing the batch effects. Thus, samples whose median was either less than the negative one or larger than the positive one were considered as the outliers. GSM1300069 sample in the normal group was added to the biased samples. Other trespassing samples were already recognized nicely by the PCA plane.

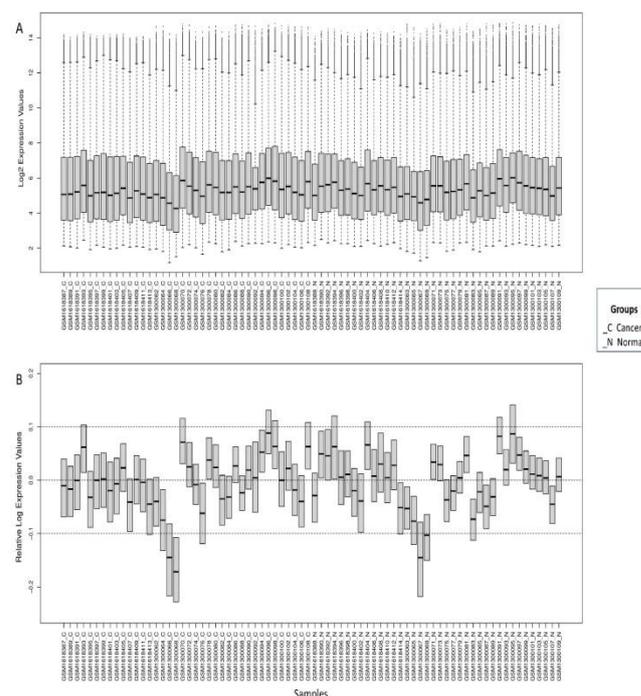


Figure 2. Boxplots. A presents the non-normalized log2 expression value of genes in all samples, while B exhibits the IQR for relative log expression value of genes (RLE) before normalization in all samples. Samples with RLE less than either 0.1 or -0.1 were considered as the outliers.

In the next step, data were normalized using the quantile method. Standard deviation (SD) for all probesets was computed. The median of all SDs, 0.38, was used as the cutoff to remove low-variant genes. From multiple probesets mapping to the same gene symbol, one of them was selected based on a larger IQR. Therefore, the final number of probesets was 7378, each representing a gene. Using the limma R package, differentially expressed genes were recognized. 693 Genes with absolute log fold change larger than one and adjusted p-value less than 0.05 were selected as the final DEGs including 331 downregulated and 362 upregulated genes.

Protein-Protein Interaction Network

Final DEGs were imported into the STRING database, and interactions were built based on the five sources of evidence mentioned in the methods. A network with 401 nodes and 1762 interactions emerged. Then the giant component of this network with 351 nodes and 1720 edges was extracted for further analysis, illustrated in Figure 3.

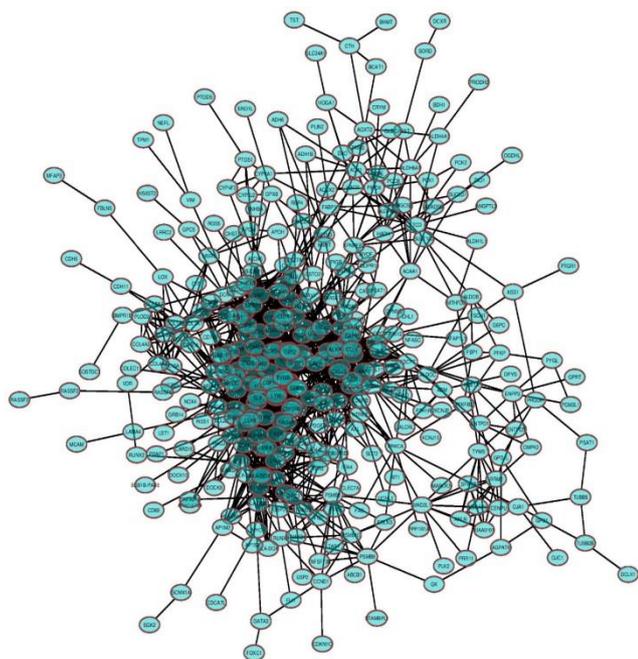


Figure 3. The giant component PPI Network. Interactions were attained from STRING interactome database. Interaction information from five sources of evidence was recruited to insert an edge between two nodes (see Methods).

The transitivity of this network was 43%, its diameter was nine, including TST, CTH, AGXT2, HADH, PSMB8, GBP2, HLA-DRA, AP1M2, SCNN1A, and SGK2, but the edge density was 3%. The degree distribution of the network presented in Figure 4 depicts the scale-free properties of the biological network (22).

The ninth decile of all betweenness and degree centralities was used as the cutoff to select the hub genes. Eighteen hub genes were recognized presented in Table 1. ITGB2 gene had both the maximum betweenness and degree centralities, 4471 and 60 respectively. C3AR1 and CTSS genes had high values for the two centralities as well.

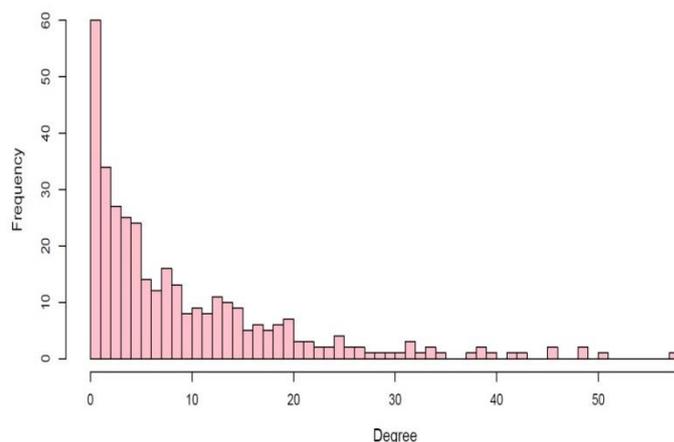


Figure 4. Degree distribution of the network. The network was assembled from a small number of nodes with high degrees and a large number of nodes with small degrees.

Table 1. Degree and Betweenness centralities for hub genes. DEGs are sorted based on the highest degree. The status column depicts the expression status of the genes. Up means, genes were upregulated in the cancer group compared to the normal group but down means gene was downregulated in the cancer group compared to the normal group.

DE Genes	Degree	Betweenness	Status
ITGB2	60	4471	Up
TYROBP	58	2496	Up
C3AR1	51	4089	Up
LCP2	49	1357	Up
PTPRC	46	1700	Up
CD53	46	1636	Up
CTSS	42	4412	Up
HCK	40	1706	Up
MNDA	39	1793	Up
KNG1	35	2201	Down
C3	34	3072	Up
EGFR	29	4457	Up
SRGN	27	2140	Up
EGF	27	2462	Down
FN1	26	2013	Up
IRF8	26	2390	Up
GNG2	24	2049	Up
CD44	24	3856	Up

Gene Expression Profiles and Survival Analysis

Figure 5 illustrates the expression profiling and survival analysis results for the most dysregulated DEGs in TCGA and GTEx databases. Besides, Betweenness centrality was evaluated for feature selection because these genes would be involved in the master routes of the PPI network. They would be

part of signaling pathways that significantly impact the pathogenesis of the disease. Downregulated and upregulated hub genes having betweenness centrality larger than 3000 were selected from Table 1 (Panel C). Five genes with the largest LogFCs (Panel A) and seven genes with the smallest LogFCs (Panel B) were taken as well. Downregulation of the two central genes, KNG1 and EGF, was unquestionable in the boxplots. KNG1 was among the most downregulated genes having high betweenness centrality. Among the significant genes in panel C, EGF witnessed a perceptible downregulation. Other central genes were upregulated in renal cancer. All the boxplot results conferred to the expression of DEGs, validating their expression in kidney renal benign tumors. Kappa Miller algorithm was used for survival analysis. Significant results (Log-rank < 0.05) for survival analysis did not confer to our analysis and expression profiling boxplots except for ALDOB and C3 genes.

Gene Set Enrichment Analysis

Most upregulated genes with LogFC larger than two, most downregulated genes with LogFC less than negative two, and hub genes in Table 1 were employed to perform the over-representation analysis. Hub genes with absolute LogFC larger than two that are present in Figure 6 were TYROBP, LCP2, PTPRC, CTSS, MNDA, KNG1, C3 and EGF. Most upregulated genes, most downregulated genes and hub genes were visualized in three separate clusters with different colors. Notably, common DEGs between the clusters were kept only in one of the clusters. Biological Process terms in GO repository and signaling pathway terms in KEGG database with Benjamini-Hochberg adjusted p-value less than 0.5 went under enrichment analysis. Nodes' size exhibits the significance of the terms. The same terms were also illustrated in Figure 7 in the form of bar plots. "Response to chemokine" group contained the most number of terms for most upregulated genes (7A). "Protein targeting to peroxisome" group encompassed the most number of terms for most downregulated genes (7B) and the "positive regulation of superoxide anion generation" group contained the most number of terms for hub genes (7C). Furthermore, a great number of downregulated genes were among the enriched sets involved in catabolic and metabolic processes, which highlights the importance of

metabolic processes in the emergence of progressive kidney tumors.

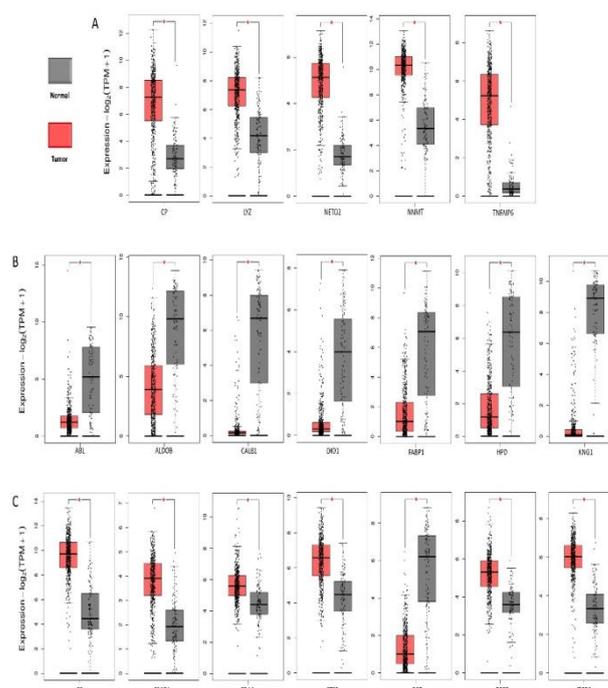


Figure 5. Gene expression profile boxplots for kidney renal carcinoma in TCGA and GTEx repositories. Red boxes are tumor gene expression values and gray boxes are normal gene expression values. Data were log₂ transformed and normalized using the TPM method. Panel A exhibits the expression values for the most upregulated genes with LogFC larger than four. Panel B depicts the most downregulated genes with LogFC less than negative four, and panel C displays the same plots for the highly central genes.

Kidney cancer is one of the most common cancers. Therefore, we analyzed gene expression profiles of Benign kidney cancer to find mechanisms associated with the pathogenesis of the disease. We concentrated on the most dysregulated genes and the hub genes in the analyses in order to recognize key important affected pathways and biological processes triggering the progression of cancer before malignancy.

Renal carcinoma is a complex disease and it appears with different histopathological symptoms, genetic changes, and resistant responses to the types of therapies. Therefore, reprogramming of cancer metabolism is a crucial way to identify therapeutic and diagnostic approaches in renal carcinoma. In addition, based on recent studies, cancer cells use

amino acids as the common energy source and their metabolism is increased in many tumors (23).

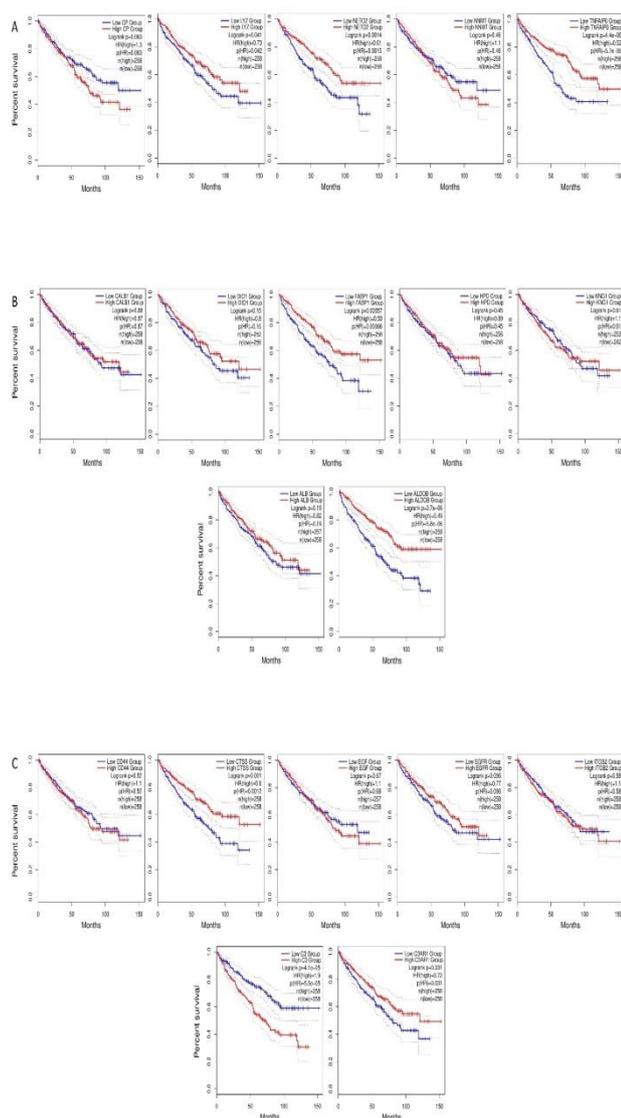


Figure 6. Survival analysis plots. patients having a low expression of genes are in blue while those having high expression of the genes are in Red. The time scale of the survival rate is based on months. Panel A exhibits the survival plots for the most upregulated genes with LogFC larger than four. Panel B depicts survival analysis results for the most down-regulated genes with LogFC less than negative four. Panel C shows the same plots for highly central genes.

As a result, the metabolism of amino acids is increased in different cancers (24). In our study, the enrichment of the "Glycine, serine and threonine metabolism" pathway was occurred together with "Peroxisome organization" pathways (Figure 8B) which highlights the role of amino acids metabolism in peroxisomes to promote tumor growth (25, 26). PIPOX, DAO are the two genes connecting the two

pathways together (Figure 7). PIPOX is a peroxisomal enzyme catalyzing oxidative demethylation of sarcosine to yield glycine in mammals (27). DAO is a D-amino acid oxidase that brings about the oxidation of d-amino acids to hydrogen peroxide (28). The downregulation of these two genes has to be further investigated in renal cancer since the DAO system has previously been reported as an efficient anti-cancer therapy (29, 30).

"Response to chemokine" were other important pathways enriched by the most upregulated genes (Figure 8A). CXCL9 and 10 genes and CXCR4 are engaged in the most number of chemokine pathways. The two chemokine ligands are expressed in the majority of cancers. Chemokines are proteins that induce chemotaxis, promote differentiation of immune cells, and cause tissue extravasation. Some reports are claiming the paradox roles for chemokine-receptor axis where it suppresses tumor growth by differentiation and activation of immune cells while there is some evidence showing their anti-tumor activity by activation of tumor proteases such as Cathepsin B (CTSB) (31-33). Its upregulation in benign renal tumors should be further investigated. C3 gene is an important member of the complement system against pathogens by enhancing immune responses (34). In the Classical pathway, this protein is cleaved to form C3a and C3b. C3b binds to the surface of cells is now protected from factor H-mediated inactivation. In the alternative pathway, the surface-bound C3b then binds protein factor B to form C3bB. it is cleaved again to form C3bBb. The C3bBb complex is stabilized by binding oligomers of factor P (properdin). It acts enzymatically to cleave much more C3, and more C3b proteins are attached to the same surface as C3bBbP. many abnormal cells such as cancer cells are accumulated on the surface by C3b proteins. Consequently, the C3bBbP complex enzyme bind again to another C3b to form C3bBbC3bP. This enzyme acts on C5 protein (C5 convertase). It cleaves C5 to C5a and C5b. The C5b then recruits and assembles C6, C7, C8 and multiple C9 molecules to assemble the Membrane Attack Complex (MAC). This creates a hole or pore in the membrane that can kill the tumor cells (34, 35). However, it is worth investigating why renal benign tumors increase expression of the C3 gene which might increment the rate of MAC formation on the surface of tumor cells.

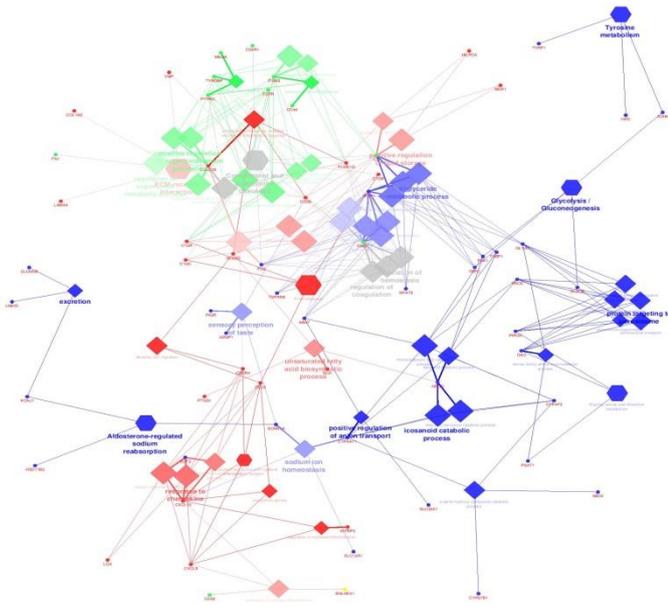


Figure 7. The enrichment graph. The size of the nodes depicts the significance of terms. Biological Process terms are in the shape of a diamond and KEGG pathways are rectangular. Green terms were obtained from hub genes, Blue terms were obtained from most downregulated genes and red terms were attained from most upregulated genes. Gray terms are non-specific terms (terms achieved by enrichment of genes in more than one set).

Functional Analysis of hub genes gave rise to the enrichment of superoxide generation pathways (Figure 8C). TYROBP, ITGB2, and EGFR are the genes engaged in these pathways. Superoxide or O_2^- is a member of ROS small molecules whose moderate rise in cancer cells accelerates cellular proliferation and tumor growth (36). Therefore, renal benign tumors might bolster these pathways to increment the rate of superoxide thereby enhancing cell proliferation.

TYROBP encodes transmembrane signaling polypeptide which caused activate motif in the cytoplasmic domain and it is considered as an immune receptor. TYROBP acts a crucial role in inflammatory responses, for example, it is related to the killer cell immunoglobulin-like receptor family or it activates signaling pathways like Jun NH2-terminal kinase in microglia and causes activate inflammatory responses (37, 38). Findings in the present study showed that TYROBP as a top hub gene (degree 41) was correlated with renal cancer and there are several possible explanations for this result. For example, it was reported that upregulated TYROBP is related to developing this cancer in these patients and plays

crucial functions in the pathogenic inflammatory process in kidney cancer (37, 39, 40).

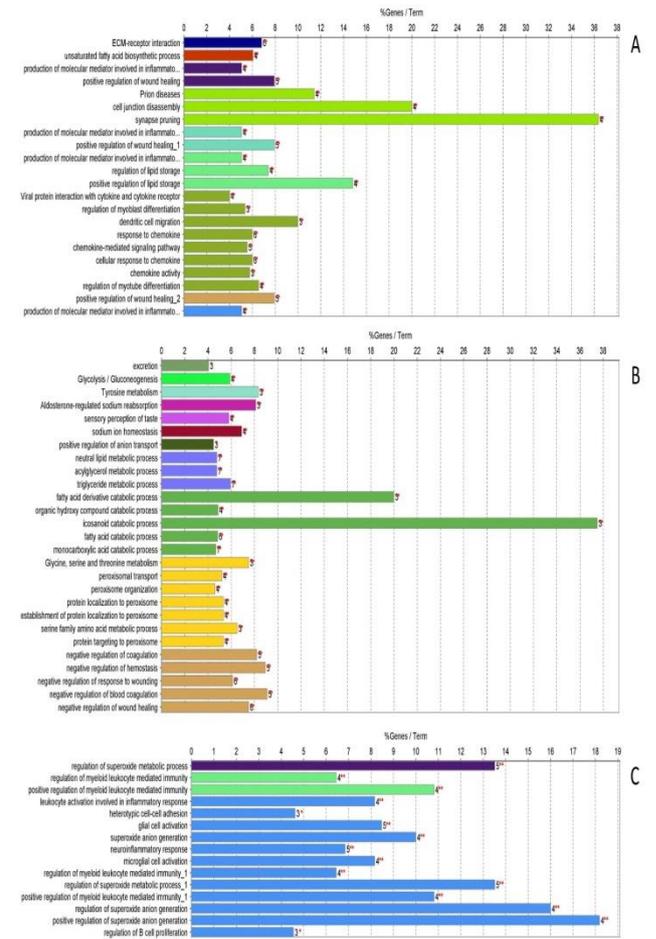


Figure 8. Enrichment bar plots. A shows the results for most upregulated genes, B exhibits the results of most downregulated genes and C illustrates the results for hub genes. Numbers next to each bar present the number of engaged genes in that term and the length of bar plots depicts the percentage of a total number of genes in that term.

Furthermore, it was proposed that the evaluation of the gene expression might be used as an indicator to screen prognosis in this tumorigenic disease.

ITGB2 encodes integrin $\beta 2$ and this is a usual subunit for many receptors. It plays a role as the cell's mechanical anchor to the ECM (41, 42). The expression of the gene is associated with deficiency of leukocyte adhesion and it is involved in the expression of components of receptors at the surface of leukocytes (42). Previous studies indicated that ITGB2 was overexpressed in several cancers, such as colorectal cancer, liver metastasis, glioblastoma, and kidney cancer (42-45). Moreover, ITGB2 correlated negatively with the estimated glomerular filtration rate

in patients with chronic kidney disease and nephrotic syndrome (45, 46). It was observed that this gene was related to survival in cases after surgery and it resulted that low expression of this gene caused increased survival in patients with colorectal cancer (43). These results are consistent in good agreement with other DEG analysis which has shown that ITGB2 is upregulated; however, its obtained degree is varied compared with other data analysis, also current research was proved ITGB2 was associated with this cancer.

EGFR is one of the known genes engaged in many oncogenic pathways including Ras/MAPK, PI3K/AKT, and phospholipase C (PLC)/protein kinase C (PKC) signaling cascade. Our Analysis depicted its up-regulation which was further supported by expression profiling in Figure 5C. However, its ligand called EGF exhibited an expression reduction in both our Analysis and the expression profiling. Therefore, EGF downregulation in kidney renal cancer should be more integrated. KNG1 is the most downregulated and central gene in our Analysis which presented a highly significant downregulation in Figure 5B as well. This gene is worth more investigation. However, the survival results for EGF and KNG1 didn't show significant (Log-rank > 0.05).

Isoforms of aldolases (ALDOB, ALDOB and ALDOC) are abundant in the human body and play roles in glycolysis, fructolysis, and the synthesis of glyceraldehyde and ATP (47). ALDOB catalyzes the conversion of fructose-1,6-bisphosphate into dihydroxyacetone phosphate and glyceraldehyde-3-phosphate (48). It is overexpressed in colorectal adenocarcinoma promoting epithelial to mesenchymal transition (EMT) while under-expressed in gastric cancer. This gene not only was suppressed in our Analysis but also its suppression was supported by expression profiling and survival analysis (47). Its expression should also be checked in metastatic renal cancer.

Lysozyme (LYZ) is a naturally occurring enzyme found in bodily secretions such as tears, saliva, and milk. It functions as an antimicrobial agent by cleaving the peptidoglycan component of bacterial cell walls, which leads to cell death (49). Lysozyme (LYZ) gene has been previously reported in human breast cancer (50). An old study introduced it as an anti-cancer agent (51). A more recent study has

reported its anti-tumor impacts from hen egg white (52). Nonetheless, there is not enough evidence connecting this enzyme to different cancers. Therefore, its upregulation in kidney renal cancer should be further investigated. Contrary, survival analysis showed that patients expressing a higher amount of this gene are more likely to survive which calls for further investigation.

NETO2 is a recent hotspot in cancer pathogenesis highlighted mostly in multiple recent studies (53, 54). This transmembrane protein activates PI3K/AKT and NF- κ B and Snail by their phosphorylation resulting in the invasion of gastric cancer cells (55). In Colorectal cancer it enhances activation of EMT-related kinases and transcription factors (56). Its role in RCC and tumor growth (non-metastatic role) must be investigated. Furthermore, the survival results in Figure 6A are contrary to our analysis and expression profiling. TNFAIP6 is another novel gene that has been investigated recently in gastric cancer. (57). CTSS had a high value for betweenness and degree centralities. It is a hub gene with LogFC larger than three verified in Figure 5. However, the survival rate for patients expressing the three mentioned genes is higher in Figure 6C.

Previous papers confirmed that overexpression of the TYMS gene could increase the risk of tumor development in some cancers like breast cancer, colorectal cancer, lung cancer, pancreatic carcinoma, gastric carcinoma, thyroid carcinoma, and lymphoma cancer (58, 59). Our data analysis also proposed that TYMS was upregulated in the cases with renal tumors. PTPRC is another gene, which encodes one member of the PTP family. These molecules organize different cellular processes associated with tumorigenesis, cell differentiation, and processes of proliferation and oncogenesis (5). The gene regulates antigen receptor signaling in immune cells including T cells and B cells (44, 58). The results concur with other researches which has shown that the PTPRC gene was expressed highly in kidney tumors (46, 60). Interestingly, several studies reported that PTPRC was identified as a hub gene with a high degree being used as a biomarker to screen renal cancer status (44, 61). Another hub gene, AURKA, can increase the expression of anti-apoptotic molecules in tumorigenic tissues. This gene encodes auroa kinase A, which affects the spindle, centrosome, and chromosome

segregation therefore it controls G2 and mitosis phases in the cell cycle (62). In addition, it was indicated that overexpression of AURKA can cause genomic instability and produce cells with multinuclear, overexpressed in neuroblastoma, liver cancer, and kidney carcinoma (63). FN1 (fibronectine 1) is a famous gene belongs to ligand glycoprotein family, expressed in different types of cells. FN1 plays a role in a various cellular mechanisms such as differentiation, migration, adhesion of cells, and carcinogenesis (64). Several studies showed that the expression of this gene is correlated positively with gastrointestinal cancers including gastric cancer and colorectal carcinoma (65, 66). FN1 was upregulated in tumor samples compared with normal samples in our analysis. Moreover, FN1 was overexpressed in breast tissues and led to metastasis and it was known as a hub gene. It was reported that FN1 was related to drug resistance in tumorigenic cells (64). SRGN is a member of the intracellular proteoglycan family and a large number of cells can produce it. It is apparent that stimulation of IL- β leads to increase SRGN expression in chondrocytes and endothelial cells (67). Also, SRGN is involved in inflammatory processes. It was found the expression of this gene increased in cancers and affected cell responses to some cytokines including IL-1 β and TNF α (67, 68). The CD53 gene encodes a number of proteins in the surface of immune cells like NK cell, monocyte, B cell, etc. and can organize immune responses (69). In addition, this gene control production of TNF α . CD53 plays a potential role in the migration of leukocytes to lymphoid organs (70). Deficiency of this gene is related to some infections in different organs, for example, intestinal and pulmonary infections. Also, it was observed that immune cells, which belong to hematopoietic stem cells, and hematological tumors expressed the tetraspanin CD53 (70, 71). Based on the immune role of CD53, the expression of CD53 was expressed highly in cases with renal cancer.

In summary, we applied accurate methods including network analysis, survival analysis, expression profiling and gene set enrichment analysis to analyze the expression of candidate biomarker genes between metastatic samples compared to normal samples. However, the main limitations need to be addressed that is designing in vitro and in vivo experiments to confirm these findings. Expression of

these genes particularly C3 which conferred to both expression profiling and survival results, as well as the hub genes with high degrees, can be monitored during the progression of cancer from early stages of the disease to high-grade stages of the disease. This aims to understand better the pathogenesis mechanism of these genes and develop a more accurate combinatory therapeutic target. Moreover, our findings showed that the major metabolic pathways and the enzymes of peroxisomes are related to the development of this disease, which required to be further investigated.

Funding Resources

No funding is applicable.

Acknowledgements

None.

Interest conflict

The authors declare no conflict of interest.

References

1. Capitanio U, Bensalah K, Bex A et al. Epidemiology of renal cell carcinoma. *Eur Urol* 2019; 75(1): 74-84.
2. Zhou Y, Funston G, Lyratzopoulos G, Walter FM. Improving the timely detection of bladder and kidney cancer in primary care. *Adv Ther* 2019; 36(7): 1778-1785.
3. Macheret M, Halazonetis TD. DNA replication stress as a hallmark of cancer. *Ann Rev Pathol: Mech Dis* 2015; 10: 425-448.
4. Chow W-H, Dong LM, Devesa SS. Epidemiology and risk factors for kidney cancer. *Nat Rev Urol* 2010; 7(5): 245.
5. Yuan L, Zeng G, Chen L et al. Identification of key genes and pathways in human clear cell renal cell carcinoma (ccRCC) by co-expression analysis. *Int J Biol Sci* 2018; 14(3): 266.
6. Drake CG, Stein MN. The immunobiology of kidney cancer. *J Clin Oncol* 2018; 36(36): 3547-3552.
7. Motzer RJ, Escudier B, McDermott DF et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med* 2015; 373(19): 1803-1813.
8. Wotschovsky Z, Gummlich L, Liep J et al. Integrated microRNA and mRNA signature associated

with the transition from the locally confined to the metastasized clear cell renal cell carcinoma exemplified by miR-146-5p. *PloS one* 2016; 11(2): e0148746.

9. Von Roemeling CA, Radisky DC, Marlow LA et al. Neuronal pentraxin 2 supports clear cell renal cell carcinoma by activating the AMPA-selective glutamate receptor-4. *Cancer Res* 2014; 74(17): 4796-4810.

10. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinforma* 2004; 20(3): 307-315.

11. Do JH, Choi D-K. Normalization of microarray data: single-labeled and dual-labeled arrays. *Mol Cell (Springer Science & Business Media BV)* 2006; 22(3).

12. Ramasamy A, Mondry A, Holmes CC, Altman DG. Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med* 2008; 5(9): e184.

13. Gentleman R, Carey V, Huber W, Hahne F. Genefilter: Methods for filtering genes from microarray experiments. R package version 2011; 1(0).

14. Smyth GK. Limma: linear models for microarray data. *Bioinformatics and computational biology solutions using R and Bioconductor*: Springer; 2005: 397-420.

15. Ferreira J, Zwinderman A. On the Benjamini–Hochberg method. *Ann Stat* 2006; 34(4): 1827-1849.

16. Luke DA. *A user's guide to network analysis in R*: Springer; 2015.

17. Bindea G, Mlecnik B, Hackl H et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinforma* 2009; 25(8): 1091-1093.

18. Pico AR, Kelder T, Van Iersel MP, Hanspers K, Conklin BR, Evelo C. WikiPathways: pathway editing for the people. *PLoS Biol* 2008; 6(7): e184.

19. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nuc Acids Res* 2019; 47(W1): W556-W560.

20. Weinstein JN, Collisson EA, Mills GB et al. The cancer genome atlas pan-cancer analysis project. *Nat Genetics* 2013; 45(10): 1113.

21. Lonsdale J, Thomas J, Salvatore M et al. The genotype-tissue expression (GTEx) project. *Nat Genetics* 2013; 45(6): 580-585.

22. Broido AD, Clauset A. Scale-free networks are rare. *Nature communications* 2019; 10(1): 1-10.

23. Li Z, Zhang H. Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. *Cell Mol Life Sci* 2016; 73(2): 377-392.

24. Ulaner GA, Schuster DM. Amino acid metabolism as a target for breast cancer imaging. *PET Clinc* 2018; 13(3): 437.

25. Wanders RJ. Metabolic functions of peroxisomes in health and disease. *Biochimie* 2014; 98: 36-44.

26. Wanders RJ, Waterham HR, Ferdinandusse S. Metabolic interplay between peroxisomes and other subcellular organelles including mitochondria and the endoplasmic reticulum. *Frontiers Cell Dev Biol* 2016; 3: 83.

27. Goyer A, Johnson TL, Olsen LJ et al. Characterization and metabolic function of a peroxisomal sarcosine and pipicolate oxidase from Arabidopsis. *J Biol Chem* 2004; 279(17): 16947-16953.

28. Huang J-L, Chen X-L, Guo C, Wang Y-X. Contributions of spinal D-amino acid oxidase to bone cancer pain. *Amino Acids* 2012; 43(5): 1905-1918.

29. Rosini E, Pollegioni L, Ghisla S, Orru R, Molla G. Optimization of d-amino acid oxidase for low substrate concentrations—towards a cancer enzyme therapy. *FEBS J* 2009; 276(17): 4921-4932.

30. Bava A, Gornati R, Cappellini F, Caldinelli L, Pollegioni L, Bernardini G. D-amino acid oxidase–nanoparticle system: a potential novel approach for cancer enzymatic therapy. *Nanomed* 2013; 8(11): 1797-1806.

31. Tokunaga R, Zhang W, Naseem M et al. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation—a target for novel cancer therapy. *Cancer Treat Rev* 2018; 63: 40-47.

32. Nazari A, Ahmadi Z, Hassanshahi G et al. Effective treatments for bladder Cancer affecting CXCL9/CXCL10/CXCL11/CXCR3 Axis: a review. *Oman Med J* 2020; 35(2): e103.

33. Bronger H, Karge A, Dreyer T et al. Induction of cathepsin B by the CXCR3 chemokines CXCL9 and CXCL10 in human breast cancer cells. *Oncol Lett* 2017; 13(6): 4224-4230.

34. Dunkelberger JR, Song W-C. Complement and its role in innate and adaptive immune responses. *Cell Res* 2010; 20(1): 34-50.
35. Bayly-Jones C, Bubeck D, Dunstone MA. The mystery behind membrane insertion: a review of the complement membrane attack complex. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2017; 372(1726): 20160221.
36. Piran M, Piran M, Sepahi N, Rahimi A. Systems Biomedicine of Colorectal Cancer Reveals Potential Targets for CRC Treatment. *bioRxiv* 2020.
37. Wang Y, Wei H, Song L, Xu L, Bao J, Liu J. Gene Expression Microarray Data Meta-Analysis Identifies Candidate Genes and Molecular Mechanism Associated with Clear Cell Renal Cell Carcinoma. *Cell J (Yakhteh)* 2020; 22(3).
38. Chen L, Yuan L, Wang Y et al. Co-expression network analysis identified FCER1G in association with progression and prognosis in human clear cell renal cell carcinoma. *Int J Biol Sci* 2017; 13(11): 1361.
39. Yang J-F, Shi S-N, Xu W-H et al. Screening, identification and validation of CCND1 and PECAM1/CD31 for predicting prognosis in renal cell carcinoma patients. *Aging (Albany NY)* 2019; 11(24): 12057.
40. Meng M, Zhang W, Tang Q et al. Bioinformatics analyses on the immune status of renal transplant patients, a systemic research of renal transplantation. *BMC Med Genomics* 2020; 13(1): 24.
41. De Fraipont F, El Atifi M, Cherradi N et al. Gene expression profiling of human adrenocortical tumors using complementary deoxyribonucleic acid microarrays identifies several candidate genes as markers of malignancy. *J Clin Endocrinol Metab* 2005; 90(3): 1819-1829.
42. Di Jia SL, Li D, Xue H, Yang D, Liu Y. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. *Aging (Albany NY)* 2018; 10(4): 592.
43. Cavalieri D, Dolara P, Mini E et al. Analysis of gene expression profiles reveals novel correlations with the clinical course of colorectal cancer. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics* 2007; 16(11): 535-548.
44. Zhang F, Wu P, Wang Y et al. Identification of significant genes with prognostic influence in clear cell renal cell carcinoma via bioinformatics analysis. *Translational Androl Urol* 2020; 9(2): 452.
45. Schuetz AN, Yin-Goen Q, Amin MB et al. Molecular classification of renal tumors by gene expression profiling. *J Mol Diag* 2005; 7(2): 206-218.
46. Song E, Song W, Ren M et al. Identification of potential crucial genes associated with carcinogenesis of clear cell renal cell carcinoma. *J Cell Biochem* 2018; 119(7): 5163-5174.
47. Chang Y-C, Yang Y-C, Tien C-P, Yang C-J, Hsiao M. Roles of aldolase family genes in human cancers and diseases. *Trends in Endocrinology & Metabolism* 2018; 29(8): 549-559.
48. He J, Jin Y, Chen Y et al. Downregulation of ALDOB is associated with poor prognosis of patients with gastric cancer. *OncoTargets and Ther* 2016; 9: 6099.
49. Oliver W, Wells J. Lysozyme as an alternative to growth promoting antibiotics in swine production. *J Animal Sci Biotechnol* 2015; 6(1): 1-7.
50. Serra C, Vizoso F, Alonso L et al. Expression and prognostic significance of lysozyme in male breast cancer. *Breast Cancer Res* 2002; 4(6): 1-8.
51. Sava G, Benetti A, Ceschia V, Pacor S. Lysozyme and cancer: role of exogenous lysozyme as anticancer agent. *Anticancer Res* 1989; 9(3): 583-591.
52. Mahanta S, Paul S, Srivastava A, Pastor A, Kundu B, Chaudhuri TK. Stable self-assembled nanostructured hen egg white lysozyme exhibits strong anti-proliferative activity against breast cancer cells. *Colloids and Surfaces B: Biointerfaces* 2015; 130: 237-245.
53. Xu J-C, Chen T-y, Liao L-t et al. NETO2 promotes esophageal cancer progression by inducing proliferation and metastasis via PI3K/AKT and ERK pathway. *Int J Biol Sci* 2021; 17(1): 259.
54. Fedorova MS, Snezhkina AV, Lipatova AV et al. NETO2 Is Deregulated in Breast, Prostate, and Colorectal Cancer and Participates in Cellular Signaling. *Frontiers Genetic* 2020; 11: 1587.
55. Liu J-y, Jiang L, He T et al. NETO2 promotes invasion and metastasis of gastric cancer cells via activation of PI3K/Akt/NF- κ B/Snail axis and predicts outcome of the patients. *Cell Death Dis* 2019; 10(3): 1-14.
56. Fedorova MS, Snezhkina AV, Pudova EA et al. Upregulation of NETO2 gene in colorectal cancer. *BMC Genetic* 2017; 18(1): 45-52.

57. Zhang X, Xue J, Yang H, Zhou T, Zu G. TNFAIP6 promotes invasion and metastasis of gastric cancer and indicates poor prognosis of patients. *Tissue Cell* 2021; 68: 101455.
58. Zhang S, Yan L, Cui C et al. Identification of TYMS as a promoting factor of retroperitoneal liposarcoma progression: Bioinformatics analysis and biological evidence. *Oncol Rep* 2020; 44(2): 565-576.
59. Hossain A, Asa TA, Quinn J, Rahman M, Huq F, Moni MA. Network-based genetic profiling, and therapeutic target identification of thyroid cancer. *bioRxiv* 2018: 480632.
60. Wang J, Yuan L, Liu X et al. Bioinformatics and functional analyses of key genes and pathways in human clear cell renal cell carcinoma. *Oncol Lett* 2018; 15(6): 9133-9141.
61. Ruan H, Li S, Tong J, Cao Q, Song Z, Wang K. The screening of pivotal gene expression signatures and biomarkers in renal carcinoma. *J Cancer* 2019; 10(25): 6384.
62. Yan G, Liu Z. Identification of differentially expressed genes in hepatocellular carcinoma by integrated bioinformatic analysis. *BioRxiv* 2019: 570846.
63. Zhang J, Qin X, Liu X et al. Identification of abnormally expressed genes and their roles in invasion and migration of hepatocellular carcinoma. *Aging* 2020; 12(10.18632).
64. Liu X, Ma Y, Yang W, Wu X, Jiang L, Chen X. Identification of therapeutic targets for breast cancer using biological informatics methods. *Molecular medicine reports* 2015; 12(2): 1789-1795.
65. Xiang L, Xie G, Ou J, Wei X, Pan F, Liang H. The extra domain A of fibronectin increases VEGF-C expression in colorectal carcinoma involving the PI3K/AKT signaling pathway. *PloS one* 2012; 7(4): e35378.
66. Wang W, Chin-Sheng H, Kuo L-J et al. NNK enhances cell migration through $\alpha 7$ -nicotinic acetylcholine receptor accompanied by increased of fibronectin expression in gastric cancer. *Ann Surgical oncol* 2012; 19(3): 580-588.
67. Scuruchi M, D'Ascola A, Avenoso A, Mandraffino G, Campo S, Campo GM. Serglycin as part of IL-1 β induced inflammation in human chondrocytes. *Arch Biochem Biophysics* 2019; 669: 80-86.
68. Rudd E. *Genetic and cellular studies of familial hemophagocytic lymphohistiocytosis*: Institutionen för kvinnors och barns hälsa/Department of Women's and ...; 2007.
69. Khuder SA, Al-Hashimi I, Mutgi AB, Altorok N. Identification of potential genomic biomarkers for Sjögren's syndrome using data pooling of gene expression microarrays. *Rheumatol Int* 2015; 35(5): 829-836.
70. Beckwith KA, Byrd JC, Muthusamy N. Tetraspanins as therapeutic targets in hematological malignancy: a concise review. *Frontiers Physiol* 2015; 6: 91.
71. Yeung L, Anderson JM, Wee JL et al. Leukocyte Tetraspanin CD53 Restrains $\alpha 3$ Integrin Mobilization and Facilitates Cytoskeletal Remodeling and Transmigration in Mice. *J Immunol* 2020; 205(2):521-32.