Acta Scientific Biotechnology

Volume 5 Issue 1 February 2024

Research Article

# Gene Expression Profiling of Kidney Cancer

**Farhana Lakadia[1]*, Kshitish Acharya[2] and Kanika Mathuria[1]**

[1]*Department of Biotechnology, St. Xavier's College (Autonomous), Ahmedabad, India*

[2]*Founder of Shodhaka Life Sciences Pvt. Ltd., Bengaluru, Karnataka, India*

***Corresponding Author:** Farhana Lakadia, Department of Biotechnology, St. Xavier's College (Autonomous), Ahmedabad, India.

**Received:** December 22, 2023

**Published:** January 31, 2024



## Abstract

Kidney cancer is the seventh most prevalent cancer worldwide, with a better survival rate than some other cancers. However, due to its heterogeneity, it is classified into multiple subtypes, requiring accurate diagnosis to determine the appropriate treatment and prognosis. Gene Expression Profiling can aid in identifying the specific subtype of cancer. The focus of this project is to investigate the metabolism and molecular mechanisms of kidney tumor cells. As cancer progresses to later stages, the success rate of treatment decreases. Therefore, it is crucial to detect biomarkers that can aid in early diagnosis and later after the treatment, in prognosis. These biomarkers can also be targeted through therapies to treat the cancer. Through this project, we aim to identify potential biomarkers by subjecting a list of differentially expressed genes in tumor cells to multiple functional analysis. The bioinformatic tools used for the functional analysis are: DAVID, KEGG, REACTOME, STRING, CYTOSCAPE, JASPAR and miTARBase.

**Keywords**: Biomarkers; Differentially Expressed Genes; Kidney Cancer

## Aims and Objectives

To gain a comprehensive understanding of the molecular events, especially from the aspect of gene expression regulation, taking place in the tumor cells of Renal Carcinoma.

Identification of potential differentially expressed genes and their role in metabolism of the kidney tumor.

To identify potential biomarkers of Renal Cell Carcinoma.

## Introduction

Cancer still remains one of the most confusing and enigmatic diseases of all times. Researchers all over the world are working tirelessly to develop precise treatments that have minimal side effects. There are many types of cancer and kidney cancer is one of them. Kidney cancer is also known as Renal Cell Carcinoma.

### Statistics

According to World Cancer Research Fund International, Kidney cancer is the 13$^{th}$ most common cancer. The 5-year survival rate for patients with localized or locally advanced disease ranges from 20% to 95%, depending on the severity of the disease, whereas patients with metastatic disease have a 5-year survival rate of 0% to 10% [6]. According to the American Cancer Society, kidney cancer is the sixth most frequently occurring cancer in men and the ninth most frequently occurring cancer in women in the United States. The average age for a person to be diagnosed with kidney cancer is 64 years old, with most cases being diagnosed between the ages of 65 and 74. Kidney cancer is uncommon in individuals under the age of 45 and is more prevalent in Black and American Indian populations. According to cancer.net, an organization maintained by the American Society of Clinical Oncology (ASCO), in the year 2023, an estimated 81,800 adults (52,360 men and 29,440 women) in the United States will be diagnosed with kidney cancer. About 431,288 people worldwide were diagnosed with kidney cancer in 2020.

### Tumor heterogeneity

Previously, renal cell carcinoma (RCC) was considered a singular disease, and all patients with renal tumors underwent the same surgical procedures. Additionally, patients with advanced disease were treated with similar drugs, none of which were effective in significantly increasing survival rates, although there were occa-

sional responses. However, a greater understanding of RCC has revealed that it is comprised of various types of cancer that exhibit distinct histologies, clinical courses, and responses to therapy. Not considering the cancer heterogenicity while providing or prescribing treatments can lead to unsuccessful cure of the cancer.

Furthermore, these cancers are caused by different genes. It has been identified that at least 17 different genes can lead to RCC, and mutations in these genes can affect a cell's ability to respond to changes in oxygen, iron, nutrients, or energy, especially in cases where there are mutations in genes for tricarboxylic acid (TCA) cycle enzymes such as fumarate hydratase and succinate dehydrogenase (SDH) [6].

## Treatment

One of the factors that contributes to the mortality of patients suffering from kidney cancer is therapy resistance [7]. Any cancer can be divided into two types: benign and malignant. Benign kidney cancer is localized and does not spread in the surrounding whereas malign cancer spreads into the neighboring tissue and can later spread to a greater extent. If the kidney cancer is localized, then the most common treatment used is of the surgical method, nephrectomy. Depending upon the size of the tumor, the size of the tissue that has to be removed is determined. In case of smaller sized tumor, only a small incision is made to remove the tumor. This is termed as partial nephrectomy. In some cases where the tumor size is large, entire kidney is removed. This is termed as radial nephrectomy. However, 25-30% of kidney cancer patients have metastatic tumors. As far as metastatic Renal Cell carcinoma is concerned, 2 of the promising types of treatments are vascular endothelial growth factor (VEGF) tyrosine kinase inhibitors (TKIs) that decreases the tumor progression by inhibiting angiogenesis (formation of blood vessels)[8] and immune therapy. Immune therapy works by employing agents that block the immune checkpoint formed through the interaction between Programmed Cell Death Protein 1 (PD-1) present on tumor cells and programmed death ligand 1 (PD-L1), which is expressed on either tumor cells or myeloid cells.

## Role of miRNA in cancer

miRNAs are non-coding RNA that can act as tumor promoters or suppressors depending on the microenvironment of the tumor. Under abnormal condition, miRNAs undergo dysregulation thereby immunizing cells from growth suppressors and apoptosis. They also help the cells to proliferate and become tumorous. Some miRNAs are involved in invasion and metastasis. miRNAs have been used as target for various cancer therapies [31].

## Review of Literature

A study by Linehan WM showed the importance of Gene Expression Profiling and its role in developing targeted therapy for renal cancer. His study showed that kidney cancer is a heterogenous disease with different genotypes and histology. Knowledge of the genes and their mutation can help in devising a customized and personalized treatment with reduced side effects. Also, genomic studies are required to unveil critical genes and pathways for the development of effective therapies for patients with different types of kidney cancer.

Many studies and research work has been done to identify potential biomarkers of different types of renal cancer.

In 2020, study was performed by Chen, Yeda., *et al*. on microarray datasets downloaded from Gene Expression Omnibus (GEO) database to produce a list of Differentially Expressed Genes. They used DAVID to perform functional enrichment analysis, STRING to identify the protein-protein interaction network and GEPIA to identify the survival analysis. Their study showed genes: SUCLG1, PCK2, GLDC if overexpressed, had negative effect on the tumor progression. They showed that devising targeted therapies on these genes can yield beneficial results.

Shi, Shen-Nan., *et al*. (2020) investigated the role of TIMP1 in Renal Cell Carcinoma. The data extracted from 39 normal kidney tissue and 39 kidney tumor tissues was compared to yield 310 Differentially Identified Genes. Out of which, 133 upregulated and 177 downregulated genes were subjected to protein-protein interaction analysis followed by a pathway enrichment analysis. This was extrapolated by knocking down TiMP1 in two cell lines: A498 and Caki-1 which resulted into decrease in invasion rates of the cells.

Computational approach was used by Khouja, Hamed Ishaq., *et al*. 2022 for identification of Differentially Expressed Genes. TCGA data set was utilized to extract the data at two stages (I and II) of Renal Cell Carcinoma. Their study showed that Pi3K Akt, Foxo, endocytosis, MAPK, Tight junction, cytokine cytokine receptor interaction pathway were the most altered pathways in the cancerous tissues.

Courcier, Jean., *et al*. reported the prognostic value of Carbonic Anhydrase IX (CAIX) in Renal Cell Carcinoma. It being a transmembrane protein, is highly expressed in various tumors and is a potential target for therapy. However, phase III trial are ongoing.

Role of LIF as a poor prognostic marker was reported through a study conducted by Zhong, Wenting., *et al*. (2022) using databases: TCGA, GEO, ICGC, and ArrayExpress. LIF is a pleotropic protein that is overexpressed in many tumors and has a carcinogenic effect. Its overexpression points towards a poor prognostic treatment.

A very recent study performed by Bin Wan., *et al*. 2023 identified and demonstrated the prognostic values of Differentially Methylated Genes in Renal Cell Carcinoma. Through functional analysis, they detected an upregulation of the methylated genes: TYROBP, BIRC5, BUB1B, CENPF, and MELK and their role in carcinogenesis.

## Methodology

### Data Compilation and Curation

Two data bases: SRA and ENA were used to collect the data. Keywords were designed for kidney cancer search. Target search was intended for mRNA, miRNA and fluid samples. *Homo sapiens* was chosen as a filter for organism. The data was downloaded, exported and curated. A total of 5664 hits were obtained for mRNA, 695 for miRNA and 270 for fluid. A total of 2037 sample IDs of mRNA, 570 of miRNA and 170 of fluids were analyzed and subject to further study.

The curated files were project through meta-analysis using an in-house software application. As a result, a list of Differentially Expresses Genes comprising of 305 genes was obtained. This list was used for conducting further functional analysis.

### Functional annotation of genes

David (http://www.webgestalt.org/) (Huang., *et al*. 2009) was used as a bioinformatic tool to functionally annotate the list of Differentially Expressed Genes. A list of 305 genes were given as an input. Homo sapiens was selected as an organism. The gene list was first converted Enterez Gene ID list and before submitting it to DAVID for functional annotation. Three Annotation criteria: Biological Processes, Molecular Function and Cellular Processes were analyzed.

### Protein-protein interactions and Functional analysis

The list of Differentially Expressed Genes (DEGs) was obtained by performing meta-analysis. This list was used to perform the functional analysis. STRING (https://string-db.org/cgi/input.pl) (Szklarczyk., *et al*. 2019) was used to find out the cluster formation among the sets of DEGs. Homo sapiens was set as the organism and the confidence level was set as 0.7 to achieve networks at a high confidence level. Nodes represent proteins whereas the edges represent the protein-protein associations.

The data was then extracted to Cytoscape (version 3.9.1) (Otasek., *et al*. 2019) for better visualization. A tool called network analyzer was used to study the top genes involved in protein-protein interaction. The parameters that were considered were degree of centrality (number of edges a node has), betweenness of centrality (a way of detecting the amount of influence a node has over the flow of information in a graph) and degree.

### Pathway analysis of DEGs

DAVID, KEGG (https://www.genome.jp/kegg/kegg1.html) (Kanehisa., *et al*. 2021) and REACTOME (https://reactome.org/) tools were used to perform pathway analysis. Furthermore, STRING was used to study the interaction between the intermediates of the pathways. Literature review was performed to study the function of the pathways and their potential role in tumor progression.

### Identification of miRNAs and their target genes

After performing the network analysis of the DEGs, a list of top interactive 15 genes were obtained. miRTarBase (https://mirtarbase.cuhk.edu.cn/) (Huang., *et al*. 2020) was used to identify the miRNAs of these target genes. For further analysis, only miRNAs that were validated by two or more methods were considered for further analysis.

### Transcription factor analysis

The main goal of this analysis was to detect the number of Transcription Factors Encoding genes. To do that, Tansfac was used. The search was performed by selecting the tool for 'Factor'. The organism was selected as *Homo sapiens*. After obtaining the list of all the Transcription Factors Encoding Genes, JASPAR Core was used for further analysis.

The information about the name of the Transcription factor, the family and sequence logo were collected. The functional annotation of the genes was done by performing literature review.

### Integrated analysis by network building and visualization

Data obtained from all the above-mentioned steps was collected to do the final visualization using cystoscope. The organism was selected a *Homo sapiens*. The list of the Differentially expressed genes was given as the input. The network was analyzed using F1 network analyzation. Protein-protein interaction, interaction of transcription factors and proteins involved in the pathways were studied using this tool.

## Results

### Functional Annotation of genes

To functionally annotate the genes, DAVID was used. Three criteria of annotation: The genes were annotated based on their functions in biological processes (BP), Cellular processes (CC), Molecular functions (MF).

The most significant biological processes that the genes of the Differentially Expressed Gene list took part in were signal transduction, immune response, cell adhesion, inflammatory response, cell-cell signaling. The most significant Cellular components that the genes of the Differentially Expressed Gene list were plasma membrane, extracellular exosomes (which has shown to promote tumor) [23]. And lastly, the greatest number of Molecular Functions that the genes of the differentially Expressed Gene list were involved in are ATP binding, Calcium ion binding, protein homodimerization. The genes of the list show differential expression in renal disease and cancer as well.

Table 1 shows enlists the different Differentially Expressed Genes and their role in numerous processes.

### Protein-protein interactions and Functional analysis

A total of 268 nodes were identified and amongst them, a total of 106 edges were identified. Some of the protein-protein interactions took place to give rise to various biological processes like chemokine-mediated signaling pathway, synapse assembly, humoral response, cell chemotaxis, leukocyte migration. Protein-protein interaction for molecular function included Heparin binding, calcium binding, chemokine activity and so on. Many proteins took

| Annotation type | Types | Gene name |
|---|---|---|
| Disease | Renal | HMGCS2, CXCL13, CXCL8, C1QTNF7, CD27, LRP2, ANK3, APOL1, AQP2, ALOX15B, CDH9, CFH, CCNA2, CDKN1A, CYP8B1, EPHX1, ERBB4, FGG, FMO2, FRZB, LGALS2, GJA4, GPC5, HSP90B1, ID3, IL12RB2, KDR, KNG1, LIPG, HLA-DPB2, MMP1, OLR1, PLAT, PDGFRB, PKLR |
| | Cancer | HMGCS2, PFKFB2, ABCA6, CCL20, CCR6, CXCL1, CXCL14, CXCL5, CXCL6, CXCL8, C1QTNF7, CD27, DIRAS3, EHF, EYA2, FBXW7, FCGR2B, GATA3-AS1, GC, LRP2, NLRP11, RALYL, RUNX1, RUNX3, SEC61G, ACO1, ADH1C, ANGPTL1, AIFM1, ALOX15B, AHRR, CREB5, CDH11, CHL1, CCN3, CBLN4, COL11A1, CFH, CUBN, CCNA2, CDKN1A, DTL, EPHX1, ERBB4, FMO2, FRZB, GJA4, GPC5, GNRH1, IGK, INHBA, ID3, IGFBP2, IFI27, IFITM1, IRF1, IL1RL1, IL12RB2, IL17RB, IL7R, DIO1, KDR, KIF23, HLA-DRB4, MRC1, MMP1, MT1X, MT-ATP6, MAD2L1, NELL1, NCF2, NNMT, PLAT, PDGFRB, KCNMB2, KCNK10, PRLR, RMST |
| Biological Processes | Inflammation | CCL20, CXCL1, CXCL13, CXCL5, CXCL6, CXCL8, CEBPB, FCGR2B, NOX4, ITGAL, IL1RL1, KNG1, LYZ, OLR1 |
| | Cell Adhesion | ADAM12, FREM2, CDH11, CDH16, CHL1, CCN3, CLDN19, COLFA1, IBSP, ITGAL, IL32, LRRN2, LY9, OLFM4, PCDHB16, PCDHB2, PCDHB7 |
| | Immune Response | CCL20, CCR6, CXCL13, CXCL14, CXCL5, CXCL6, CXCL8, CEPBP, FCGR2B, GPR183, LIF, CTSW, IGLC1, IL1RL1, IL32, IL7R, HLA-DRB4, PIK3R6 |
| Cellular Components | Exosome Formation | AEBP1, ATAD2, FREM2, GC, LCK, LRP2, RAB17, AQP2 |
| | Plasma Membrane | ADAM12, ABCA6, ABCA8, CCR6, CLEC5A, CD27, CD3G, CD69, DIRAS3, FRMD6, FREM2, FCAMR, FCGR2B, GPR183, LCK, LRP2, NOX4, RAB17, RIMBP2 |
| Molecular Process | ATP Binding | PFKFB2, ABCA6, ABCA8, ATAD2, LCK, NEK5, NLRC3, NLRP11, NRK, RUNX1, RUNX2, ACSL5, ERBB4, KDR, PANK1, PALM3, PNCK, PKLR |

**Table 1:** Functional Annotation of Genes using DAVID.

part in plasma membrane stability and vesicle formation. Table 3 enlists clusters and its Differentially Expresses Genes.

| Name | Betweenness Centrality | Closeness Centrality | Degree |
|---|---|---|---|
| CXCL8 | 0.232562 | 0.355102 | 27 |
| IL7R | 0.068471 | 0.307965 | 24 |
| FCGR2B | 0.050483 | 0.312388 | 16 |
| CXCL1 | 0.031045 | 0.306878 | 16 |
| PLEK | 0.163661 | 0.315789 | 15 |
| CCL20 | 0.011942 | 0.296422 | 15 |
| CD69 | 0.019386 | 0.309059 | 15 |
| GPR29 | 0.006624 | 0.287129 | 15 |
| IL33 | 0.016912 | 0.287603 | 15 |
| LCK | 0.039501 | 0.303665 | 14 |
| LIF | 0.076017 | 0.311828 | 13 |
| CXCL13 | 0.003658 | 0.269767 | 12 |
| BGN | 0.047583 | 0.260479 | 11 |
| CD27 | 0.048147 | 0.285714 | 11 |
| KDR | 0.111726 | 0.314647 | 11 |

**Table 2:** Top 15 interactive genes amongst the list of the Differentially Expressed Genes.

| Cluster Number | Type of interaction | Genes involved |
|---|---|---|
| 1 | Cytokine-Chemokine signaling pathway | CXCL8, CXCL13, CXCL1, CCL20, CXCL5, GPR29, CXCL6 |
| 2 | IL-17 signaling pathway | CEBPB, CXCL8, CCL20, MMP1, NOV, IL33, CXCL1, CXCL6, CXCL5 |
| 3 | Mitotic Cell cycle, spindle formation | KIF23, ATAD2, CCNA2, KIF4A, MAD2L1 |
| 4 | T cell differentiation | IL7R, CD3G, LCK |
| 5 | PPAR signaling pathway | HMGCS2, EHHADH, FABP6, FABP3, ACSL5 |
| 6 | Type I interferon signaling pathway | IRF1, IFI27, IFITM1 |

**Table 3**: Protein-Protein interaction amongst the list of the Differentially Expressed Genes.

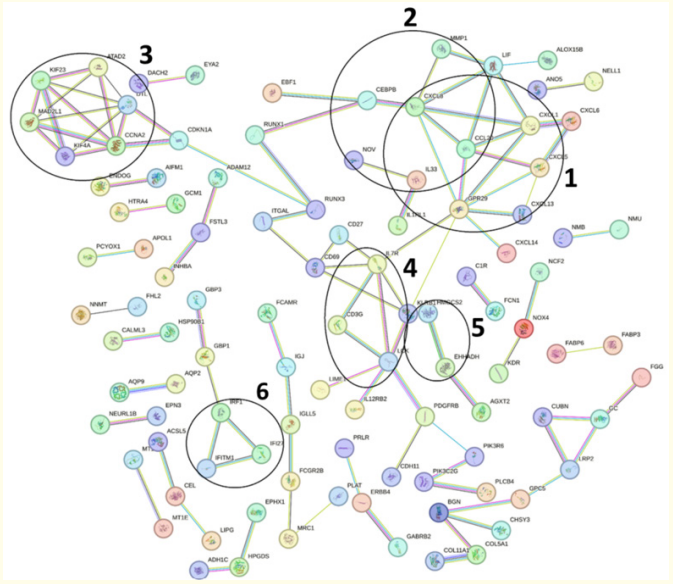



**Figure 1:** Protein-Protein interaction using STRING. 268 nodes showing protein/gene products and 106 edges were obtained. Key: 1: Cytokine-Chemokine signaling pathway, 2: IL-17 signaling pathway, 3: Cell cycle, 4: T cell differentiation, 5: PPAR signaling pathway, 6: Type I interferon signaling pathway.

Cluster 1 was enriched with cytokine-chemokine pathway as its main function. Along with that, it was involved in cell killing, humoral immune response, cell chemotaxis, response to stress, leukocyte migration and cell migration. Cluster 2 was linked with IL signaling pathway. Cluster 3 is mainly associated with Cell division and cell cycle process. Along with that, proteins also take part in chromosome and microtubule organization. Cluster 4 was found to be linked with T cell differentiation. Cluster 5 is linked with PPAR signaling pathway whereas cluster 6 is linked with Type 1 interferon signaling pathway.

## Pathway analysis of DEGs

The aim of this analysis was to understand the pathways that Differentially Expressed Genes take part in. This analysis was done using various tools: Cytoscape, DAVID, Kegg, Reactome and STRING.

Majority of the genes are involved in the cytokine-chemokine pathway, PI3K-Akt Pathway, IL17 Pathway, TNF Signaling Pathway, cGMP-PKG Signaling Pathway.

| No. | Pathway | Genes involved |
|---|---|---|
| 1 | Cytokine-chemokine Pathway | CCL20, CCR6, CXCL1, CXCL13, CXCL14, CXCL5, CXCL6, CXCL8, CD27, LIF, INHBA, IL1RL1, IL12RB2, IL17RB, IL20RB, IL32, IL33, IL7R, PRLR |
| 2 | PPAR Pathway | HMGCS2, ACSL5, CYP8B1, EHHADH, FABP3, FABP6, FABP7, MMP1, OLR1 |
| 3 | IL17 Pathway | CCL20, CXCL1, CXCL5, CXCL6, CXCL8, CEBPB, HSP90B1, IL17RB, MMP1 |
| 4 | TNF Signaling Pathway | CCL20, CXCL1, CXCL5, CXCL6, CEBPB, LIF, CREB5, IRF1 |
| 5 | cGMP-PKG Signaling Pathway | CREB5, CALML3, CNGA1, KNG1, MYLK4, PIK3R6, PLCB4, KCNMB2 |
| 6 | Regulation of PI3K/AKT Signaling | LCK, ERBB4, IL1RL1, IL33, PDGFRB |

**Table 4:** Pathway Analysis of the Differentially Expressed Genes.

Role of cytokine and chemokine signaling pathway is well established in cancer progression. The TNF networking is mainly associated with inflammatory response. TNF alpha, secreted my tumor cells, is a key intermediate cytokine that has been found to be associated with cancer progression, growth and metastasis [28]. Moreover, it is an established prognosis marker and target for renal cell carcinoma [28]. PI3K/AKT Signaling is linked to cell cycle, cell growth, survival and proliferation, tumor formation, metastasis and drug resistance [29].

cGMP-PKG Signaling Pathway plays an important role in inhibiting apoptosis and mitosis arrest. PDE5, an intermediate is a target for any inhibition cancer therapies [30].

Table 4 enlists the pathways and the respective list of genes involved in them.

## Identification of miRNAs and their target genes

The miRNAs of top 15 most interactive genes found from the list of the Differentially Expressed Genes and the Transcription Factor Encoding Genes were found out. The list of miRNAs with their target genes are given in table 5.

## Transcription factor analysis

Out of the list of 305 Differentially Expressed Genes, 8 were found to be playing role as the Transcription Factor Encoding

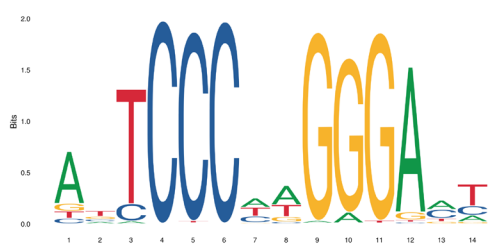
| Target Gene | miRNA |
|---|---|
| CXCL8 | hsa-miR-183-5p |
| IL7R | hsa-miR-3924, hsa-miR-4775, hsa-miR-3682-3p, hsa-miR-190b, hsa-miR-190a-5p, hsa-miR-32-3p, hsa-miR-4474-3p, hsa-miR-6842-3p, hsa-miR-377-3p, hsa-miR-365b-5p, hsa-miR-365a-5p, hsa-miR-8052, hsa-miR-3199, hsa-miR-5787, hsa-miR-4505, hsa-miR-4430, hsa-miR-3652, hsa-miR-6131, hsa-miR-1228-3p, hsa-miR-889-5p, hsa-miR-942-3p, hsa-miR-26b-5p |
| FCGR2B | hsa-miR-18a-5p |
| CXCL1 | hsa-miR-302e, hsa-miR-27b-5p |
| PLEK | hsa-miR-3923, hsa-miR-1250-3p, hsa-miR-153-5p, hsa-miR-5696, hsa-miR-579-3p, hsa-miR-664b-3p, hsa-miR-520a-5p, hsa-miR-525-5p, hsa-miR-183-3p, hsa-miR-4452, hsa-miR-3180-5p, hsa-miR-6875-3p, hsa-miR-4659a-3p, hsa-miR-4659b-3p |
| CCL20 | hsa-miR-21-5p |
| CD69 | hsa-miR-92a-3p |

| | |
|---|---|
| CCR6 | hsa-miR-518a-3p, hsa-miR-150-5p, hsa-miR-127-5p, hsa-miR-7843-3p, hsa-miR-4749-3p, hsa-miR-6806-3p, hsa-miR-3928-5p, hsa-miR-4635, hsa-miR-4539, hsa-miR-106a-3p, hsa-miR-6879-3p, hsa-miR-4430, hsa-miR-3652, hsa-miR-564, hsa-miR-122-5p, hsa-miR-504-3p, hsa-miR-3135b, hsa-miR-4638-5p, hsa-miR-1307-3p, hsa-miR-500b-3p, hsa-miR-183-5p, hsa-miR-454-3p, hsa-miR-4295, hsa-miR-3666, hsa-miR-301a-3p, hsa-miR-301b-3p, hsa-miR-130b-3p, hsa-miR-130a-3p, hsa-miR-5197-5p, hsa-miR-640, hsa-miR-5579-3p, hsa-miR-1245a, hsa-miR-8079, hsa-miR-129-5p, hsa-miR-8055, hsa-miR-4309, hsa-miR-3198, hsa-miR-4294, hsa-miR-1289, hsa-miR-8082, hsa-miR-4534, hsa-miR-1273f, hsa-miR-4708-5p, hsa-miR-6088, hsa-miR-4770, hsa-miR-143-3p, hsa-miR-3124-3p, hsa-miR-6514-5p, hsa-miR-150-3p, hsa-miR-518a-5p |
| IL33 | hsa-miR-524-3p |
| LCK | hsa-miR-335-5p |
| LIF | hsa-miR-330-5p, hsa-miR-199a-5p, hsa-miR-346, hsa-miR-26a-5p, hsa-miR-125a-5p, hsa-miR-125b-5p, hsa-miR-199b-5p, hsa-miR-3180-5p, hsa-miR-4279, hsa-miR-4645-5p, hsa-miR-4650-5p, hsa-miR-4673, hsa-miR-4711-3p, hsa-miR-4757-5p, hsa-miR-5193, hsa-miR-660-3p, hsa-miR-6733-3p, hsa-miR-6744-3p, hsa-miR-6749-3p, hsa-miR-6845-3p, hsa-miR-223-3p, hsa-miR-26b-5p, hsa-miR-335-5p |
| CXCL13 | hsa-miR-26b-5p |
| BGN | hsa-miR-4311, hsa-miR-6828-5p, hsa-miR-8485, hsa-miR-206, hsa-miR-3125, hsa-miR-3916, hsa-miR-4476, hsa-miR-4533, hsa-miR-6758-5p, hsa-miR-6856-5p, hsa-miR-6859-5p, hsa-miR-6876-5p |
| CD27 | hsa-miR-335-5p |
| KDR | hsa-miR-338-3p, hsa-miR-16-5p, hsa-miR-17-3p, hsa-miR-200b-3p, hsa-miR-15b-5p, hsa-miR-200c-3p, hsa-miR-19b-1-5p, hsa-miR-335-5p, hsa-miR-106b-5p, hsa-miR-548aa, hsa-miR-4799-5p, hsa-miR-548ap-3p, hsa-miR-548as-3p, hsa-miR-548at-3p, hsa-miR-548t-3p, hsa-miR-548ay-3p, hsa-miR-34a-5p, hsa-miR-155-5p, hsa-miR-1236-3p |

**Table 5**: Lists of genes and their miRNA.

Genes. These 8 genes are as follows: CEBPB, EBF1, ID3, IRF1, EMX2, RUNX1, RUNX3, GCM1. Out of these, only GCM1 is a non prognostic biomarker.

Detailed information about each gene with its function is given in the table 6.

| Gene Name | TF Name | Family | Function | Sequence Logo |
|---|---|---|---|---|
| RUNX1 | RUNX1 | Runt-related factors | RUNX1 has been linked to skin cancer in mice<br><br>Can act both as an oncogenic or a tumor-suppressive factor [9] |  |
| | AML1 | | AML with RUNX1-RUNX1T1 fusion is a prognosis marker [10]<br><br>Mutation is directly involved in acute myeloid leukemia [11] | |
| | AML1 deltaN | | Modulator of AML1 | |
| EBF1 | EBF1 | Early B-Cell Factor-related factors | Key pioneer transcription factor of B-cell specification and commitment<br><br>Upstream Transcription Factor of PNO1<br><br>Potential Biomarker |  |

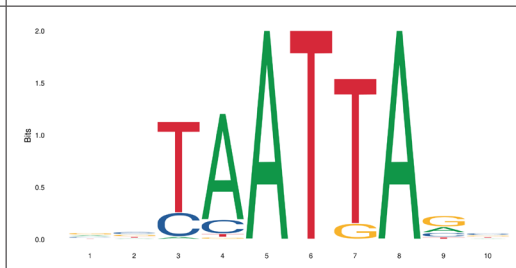
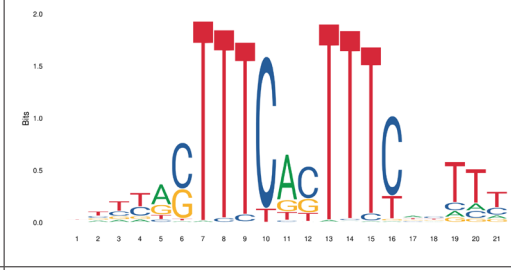
| | | | | |
|---|---|---|---|---|
| GCM1 | GCMa | GCM | Non-Prognostic |  |
| EMX2 | EMX2 | Homeo domain factors | Prognostic marker in Renal Cancer |  |
| IRF1 | IRF1 | Interferon-regulatory factors | Prognostic marker in Renal Cancer |  |
| CEBPB | NF-IL6-2 | CEBP-related | tumor formation, growth, metastasis and treatment resistance in a variety of cancers |  |
| ID3 | HEIR-1 | | Transcriptional regulator and prognostic marker in many cancer types | |

**Table 6**

## Integrated analysis by network building and visualization

A network was obtained and visualized using Cytoscape. Different clusters were further analyzed to obtain comprehensive insight into the molecular interaction among the proteins. Transcription factors were visualized and their interaction with different proteins were studied. Transcription factor CEBPB was found to be interacting with CXCL8, CCL20, FABP3 and LYZ. Cluster 1 and cluster 2 showed Calcium signaling pathway and Calcium Ion Binding properties respectively. In cluster 1, CALML3 inhibits CNGA1. Cluster 3 was found to be associated with Cytokine Chemokine Signaling pathway whereas Cluster 4 collectively showed Adaptive Immune Response, Cluster 5 is linked with IL signaling pathway. The F1 network of integrated analysis is given below:

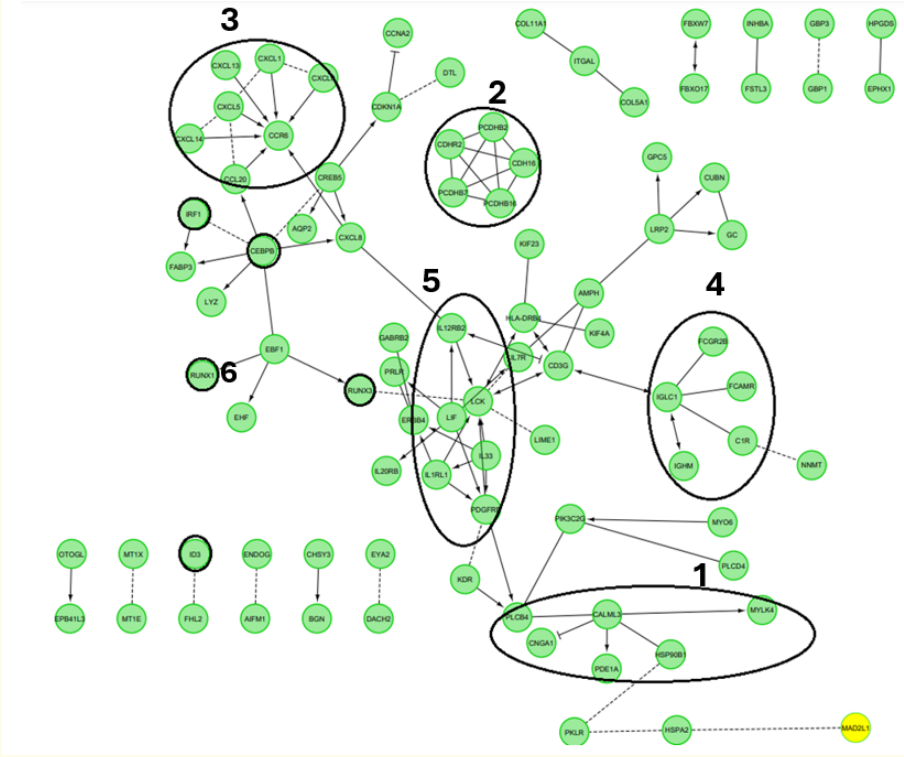



**Figure 2:** Integrated Network Analysis using Cytoscape. Key: 1: Calcium signaling pathway, 2: Calcium Ion Binding, 3: Cytokine Chemokine Signaling pathway, 4: Adaptive Immune Response, 5: IL signaling pathway, 6: Trasncription Factors.

## Discussion

Renal Cell Carcinoma is a complex disease and is associated with accumulation of many gene products that contribute in cell-cell adhesion, exosome formation, Leukocyte migration and much more. It is linked to mutation in genes, upregulation and down regulation of many genes, and miRNA alteration all of which are collectively linked to carcinogenesis. In present study, we collected data from about 3000 sample IDs which was then subjected to sorting and curating. This data was collected from two data bases: SRA and ENA. Different kidney cancer types at different stages were considered while collecting the data. This data was subjected to meta-analysis which gave us a list of 305 Differentially Expressed Genes in Renal cell Carcinoma. The genes were functionally annotated using DAVID to understand their role in the tumor microenvironment. This Gene Expression Profiling is a crucial step in studying the metabolic, molecular and cellular processes taking place in the tumor cells. We found a list of genes that contributed to both: cancer and kidney disorders. This list includes HMGCS2, CXCL8, C1QTNF7, CD27, LRP2, ALOX15B, CFH, CCNA2, CDKN1A, EPHX1, ERBB4, FMO2, FRZB, GJA4, GPC5, ID3, IL12RB2, KDR, MMP1, PLAT, PDGFRB, PKLR genes**.** Any alteration in pathway can destabilize the homeostasis of the cell and tissue environment. Various pathways that these genes were a part of are Cytokine, chemokine pathway, PPAR pathway, cGMP-PKG Signaling Pathway, TNF signaling pathway. Alterations in these pathways have shown role in carcinogenesis. From the list of 305 genes, we found the top most interactive and differentially expressed genes: CXCL8, IL7R, FCGR2B, CXCL1, PLEK, CCL20, CD69, GPR29, IL33, LCK, LIF, CXCL13, BGN, CD27, KDR. miRNAs have shown to have effect on cancer progression. They are often treated as the targets of cancer therapies. We have found the miRNAs for these top interactive target genes. Through subsequent research, we can explore their role in carcinogenesis.

The top most interactive gene was found to be CXCL8, an intermediate of many pathways. It plays a role in immune response, defense response, chemokine cytokine networking, chemotaxis, cell adhesion, inflammatory response, signal transduction and many other events involved in carcinogenesis

## Conclusion

Through this project, we were able to identify and functionally annotate differentially expressed genes in control versus cancer samples. This enabled us to gain insight into the gene expression regulation, biological processes, and pathways in the cells of kidney cancers, which is an important step in identifying potential biomarkers. After performing the functional analysis, we concluded that CXCL8 is highly expressed and is an intermediate of many biological processes, metabolic pathways and tumor-related processes. Therefore, its potential role as a biomarker for this cancer type should be explored through subsequent research.

## Acknowledgement

This project's success is owed to the individuals who have provided me with the opportunity, guidance, support, and strength throughout the entire process. I would like to express my sincerest gratitude to them.

Firstly, I would like to thank the Head of Department of Biotechnology at St. Xavier's College, Ahmedabad, Dr. Sudeshna Menon, and the professors for their unwavering support throughout the course.

I am also grateful to Dr. Kshitish Acharya, the Founder and Scientific Advisor of Shodhaka Life Sciences, whose insights helped me grasp the true essence of this project. Additionally, I extend my heartfelt gratitude to my mentor, Ms. Kanika Mathuria, who was always available to guide me throughout the entire process. I would like to extend my gratitude to the team at Shodhaka Life Sciences for organizing training sessions that helped me acquire the necessary skills.

Last but not the least, I would like to thank my family and friends for their constant support and presence during this extensive course.

## Bibliography

1. Hausman DM. "What Is Cancer?". *Perspectives in Biology and Medicine* 62.4 (2019): 778-784.

2. Owens B. "Kidney cancer". *Nature* 537.7620 (2016): S97.

3. Wu Y., *et al*. "Epigenetic and transcriptomic characterization reveals progression markers and essential pathways in clear cell renal cell carcinoma". *Nature Communications* 14.1 (2023): 1-25.

4. Linehan WM., *et al*. "The Metabolic Basis of Kidney Cancer". *Cancer Discovery* 9.8 (2019): 1006-1021.

5. Wu Y., *et al*. "Epigenetic and transcriptomic characterization reveals progression markers and essential pathways in clear cell renal cell carcinoma". *Nature Communications* 14.1 (2023): 1681.

6. Lane BR., *et al*. "Cancer of the kidney". In: DeVita VT, Lawrence TS, Rosenberg SA, editors. Cancer Principles and Practice of Oncology. Philadelphia, PA: Wolters Kluwer (2015): 865-884.

7. Li F., *et al*. "Kidney cancer biomarkers and targets for therapeutics: survivin (BIRC5 XIAP, MCL-1, HIF1α, HIF2α, NRF2, MDM2, MDM4, p53, KRAS and AKT in renal cell carcinoma". *Journal of Experimental and Clinical Cancer Research: CR* 40.1 (2021): 254.

8. Rini BI. "Vascular endothelial growth factor-targeted therapy in renal cell carcinoma: current status and future directions". *Clinical Cancer Research* 13.4 (2007): 1098-1106.

9. Blyth K., *et al*. "The RUNX genes: gain or loss of function in cancer". *Nature Reviews Cancer* 5 (2005): 376-387.

10. Höllein A., *et al*. "Molecular characterization of AML with RUNX1-RUNX1T1 at diagnosis and relapse reveals net loss of co-mutations. *HemaSphere* 3.1 (2019): e178.

11. Yuan Y., *et al*. "AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations". *Proceedings of the National Academy of Sciences of the United States of America* 98.18 (2001): 10398-10403.

12. Zhang Y W., *et al*. "A novel transcript encoding an N-terminally truncated AML1/PEBP2 alphaB protein interferes with transactivation and blocks granulocytic differentiation of 32Dcl3 myeloid cells". *Molecular and Cellular Biology* 17.7 (1997): 4133-4145.

13. Tanaka T., *et al*. "An acute myeloid leukemia gene, AML1, regulates hemopoietic myeloid cell differentiation and transcriptional activation antagonistically by two alternative spliced forms". *The EMBO journal,* 14.2 (1995): 341-350.

14. Shen Z., *et al*. "Transcription Factor EBF1 Over-Expression Suppresses Tumor Growth in vivo and in vitro via Modulation of the PNO1/p53 Pathway in Colorectal Cancer". *Frontiers in Oncology* 10 (2020): 1035.

15. The Human Protein Atlas (2023).

16. Okamoto J., *et al*. "EMX2 is epigenetically silenced and suppresses growth in human lung cancer". *Oncogene* 29.44 (2010): 5969-5975.

17. Chen F., *et al*. "IRF1 suppresses Ki-67 promoter activity through interfering with Sp1 activation". *Tumor Biology* 33 (2012): 2217-2225.

18. Tolomeo M and Grimaudo S. "The "Janus" Role of C/EBPs Family Members in Cancer Progression". *International Journal of Molecular Sciences* 21.12 (2020): 4308.

19. Chen F., *et al*. "Upregulation of Id3 inhibits cell proliferation and induces apoptosis in A549/DDP human lung cancer cells in vitro". *Molecular Medicine Reports* 14 (2016): 313-318.

20. Chen F., *et al*. "Upregulation of Id3 inhibits cell proliferation and induces apoptosis in A549/DDP human lung cancer cells in vitro". *Molecular Medicine Reports* 14 (2016): 313-318.

21. Wang T., *et al*. "Identification and immunoprofiling of key prognostic genes in the tumor microenvironment of hepatocellular carcinoma". *Bioengineered* 12.1 (2021): 1555-1575.

22. Chen Y., *et al*. "Identifying the novel key genes in renal cell carcinoma by bioinformatics analysis and cell experiments". *Cancer Cell International* 20 (2020): 331.

23. Grange C., *et al*. "Extracellular vesicles and carried miRNAs in the progression of renal cell carcinoma". *International Journal of Molecular Sciences* 20 (2019): 1832.

24. Wan B., *et al*. "Identification of Differentially Methylated Genes Associated with Clear Cell Renal Cell Carcinoma and Their Prognostic Values". *Journal of Environmental and Public Health* (2023): 8405945.

25. Shi S N., *et al*. "Identification of potential novel differentially-expressed genes and their role in invasion and migration in renal cell carcinoma". *Aging* 12.10 (2020): 9205-9223.

26. Courcier J., *et al*. "Carbonic Anhydrase IX in Renal Cell Carcinoma, Implications for Disease Management". *International Journal of Molecular Sciences* 21.19 (2020): 7146.

27. Zhong W., *et al*. "Elevated expression of LIF predicts a poor prognosis and promotes cell migration and invasion of clear cell renal cell carcinoma". *Frontiers in Oncology* 12 (2022): 934128.

28. Balkwill F. "TNF-α in promotion and progression of cancer". *Cancer and Metastasis Reviews* 25 (2006): 409-416.

29. Rascio F., *et al*. "The Pathogenic Role of PI3K/AKT Pathway in Cancer Onset and Drug Resistance: An Updated Review". *Cancers* 13.16 (2021): 3949.

30. Ren Y., *et al*. "Essential role of the cGMP/PKG signaling pathway in regulating the proliferation and survival of human renal carcinoma cells". *International Journal of Molecular Medicine* 34.5 (2014): 1430-1438.

31. Peng Y and Croce C. "The role of MicroRNAs in human cancer". *Signal Transduction and Targeted Therapy* 1 (2016): 15004.