Iran. J. Biotechnol. October 2024;22(4): e3817 DOI: 10.30498/ijb.2024.421319.3817

Research Article

Bioinformatic Approaches for the Identification of Novel Tumor Suppressor Genes and Cancer Pathways in Renal Clear Cell Carcinoma

Hassan Dastsooz [1,2,3#], Elham Mohammadisoleimani[#4], Hamed Haghi-Aminjan[5], Zahra Firoozi[4], Yaser Mansoori[4, 6*]

[1]Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy.
[2]Candiolo, C/o IRCCS, IIGM-Italian Institute for Genomic Medicine, Turin, Italy.
[3]Candiolo Cancer (IT), FPO-IRCCS, Candiolo Cancer Institute, Turin, Italy.
[4]Department of Medical Genetics, Fasa University of Medical Sciences, Fasa, Iran.
[5]Pharmaceutical Sciences Research Center, Ardabil University of Medical Sciences, Ardabil, Iran.
[6]Noncommunicable Diseases Research Center, Fasa University of Medical Sciences, Fasa, Iran.

[*]*Corresponding author*: Yaser Mansoor, Noncommunicable Diseases Research Center, Fasa University of Medical Sciences, Fasa, Iran. Tel: +98-07153350994, Fax: +98-07153350996, E-mail: fums.mansoori@gmail.com

[#]These authors contributed equally to this work.

Received: 2023/10/21 ; **Accepted:** 2024/05/15

Background: Clear cell renal cell carcinoma (ccRCC, KIRC) is the most prevalent subtype of RCC, and even with different available therapies, the average progression-free survival is worse. Therefore, the identification of new molecular targets could be helpful for its therapeutic purposes.

Materials and Methods: We used the Cancer Genome Atlas to perform bioinformatic analyses for genes with possible tumor suppressor roles in KIRC.

Objective: This research aims to identify new prognostic biomarkers and potential therapeutic targets for this type of cancer.

Results: We identified 14 down-regulated genes in KIRC that had not previously been studied or poorly studied, with the majority of them impacted by increased promoter methylation. Eight genes showed shorter overall survival and worse prognosis, indicating their function as tumor suppressors, and six genes revealed good prognosis. From the 8 genes, *C7ORF41* and *CTXN3* revealed only downregulation in most cancers, proposing them as highly potential tumor suppressors. Among these 8 genes, the function of *CTXN3* in cancers is unknown. Moreover, we identified the *CWH43* gene as the major signature of KIRC. In addition, we found different genes as signatures of KIRC tumor stages and grades.

Conclusions: Our results may shed light on identifying KIRC pathogenesis and developing effective therapeutic targets for renal cancers, mainly KIRC.

Keywords: Tumor suppressor; Methylation; Prognostic biomarkers; Renal clear cell carcinoma; The Cancer Genome Atlas (TCGA)

1. Background

Renal cell carcinoma (RCC) is a heterogeneous group of tumors among the most prevalent forms of human malignancies (1). Several RCC subtypes have been described, and the most common RCC subtypes include clear cell, papillary, and chromophobe RCCs (1, 2). Clear cell RCC (ccRCC, KIRC) is the most prevalent subtype of RCC, accounting for approximately 80% of its cases. Around thirty percent of KIRC cases lead to metastasis, leading to a worse survival rate (1, 3-5).

Copyright *© 2024 The Author(s); Published by National Institute of Genetic Engineering and Biotechnology. This is an open access article, distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses /by-nc/4.0/) which permits others to copy and redistribute material just in noncommercial usages, provided the original work is properly cited.*

So far, several contributing factors to KIRC have been identified, including genetic and environmental risk factors. Von Hippel–Lindau (VHL) gene and the protein polybromo-1 (PBRM-1) gene are among the main genes involved in RCC pathogenesis (6). However, with the development of advanced molecular genetics techniques, other involved genes with over- or under-mRNA and protein expression levels have been reported in this cancer. Their altered expressions mainly result from duplication, deletions, methylation changes, fusion proteins, and among others, the progression and invasion of KIRC. Despite substantial advances in diagnosis and the 5-year survival rate, the prognosis remains poor (7). RCC is a heterogeneous illness with several subgroups, each with unique clinical and genetic features. Clustering algorithms and machine learning, for example, may help categorize RCC into subtypes based on molecular markers. This data may be used to personalize therapy and predict outcomes for patients (8, 9). Moreover, with the emergence of the database of cancer (the Cancer Genome Atlas, TCGA), several genes have been reported to be involved in cancers (10-12). The TCGA offers detailed molecular characterization of numerous cancer types, including genomic, transcriptomic, epigenomic, and proteomic data (13). The TCGA data collection is especially focused on cancer-related molecular data and offers standardized and curated datasets for various cancer types. Other databases, for example, GEO datasets, cover a broader variety of biological study disciplines and provide a diversified selection of datasets from different experimental methods and sample kinds (14).

2. Objectives

In this study, we showed that several downregulated genes can be considered tumor suppressor genes in KIRC and applied as useful signatures of KIRC grades, stages, subtypes, and prognostic prediction. These signatures may be used for diagnostic purposes and the development of personalized medicine since our study revealed their correlations with cancer progression and invasion.

3. Materials and Methods

3.1. Top 250 Down-Regulated Genes In KIRC

Firstly, we investigated top under-expressed genes in TCGA-KIRC using the UALCAN webserver (15) (http://ualcan.path.uab.edu) to provide genes with significant transcripts per million (TPM) values ($P <$ .001) and median TPM value (= >10), giving a gene list based on mean (tumor) / mean (normal). UALCAN is a robust, easy-to-use, and online tool that has an abundance of information, including simple access to publicly accessible cancer OMICS data, patient survival statistics for genes, epigenetic control of gene expression evaluation, and pan-cancer gene expression analysis. It enables users to find biomarkers or do in silico validation of putative candidate genes.

We looked for the differential expression of 250 under-expressed genes between normal and tumor TCGA-KIRC samples as a heatmap. Next, from these top 250 under-expressed genes, we systematically searched in article databases for genes with unknown functions in KIRC. The up-regulated genes are mostly factors that affect the proliferation of cells, while the down-regulated genes are proteins that are exposed to or released from the cell surface (16). Downregulated expression data was utilized in this investigation since using upregulated data would increase the amount of work. Then, for our genes with a possible novel role involved in KIRC, we performed different analyses using the UALCAN webserver to show the deferential expression as box and whisker plots in TCGA-KIRC and its subtypes, grades, and stages. Moreover, we also confirmed their deferential expression using another web server, GSCALite (17) (http://bioinfo.life.hust.edu.cn/web/GSCALite/). Regarding this webserver, differential expression of TCGA mRNA in KIRC is in normalized RNA-Seq by Expectation Maximization (RSEM) values based on matched tumor/normal data from TCGA. Mean (tumor) / mean (normal) is applied to show the fold change (FC). The genes with fold change (FC) > 2 and false discovery rate (FDR) < 0.05 are considered to provide results as images.

3.2. Cancer Prognosis Analysis

Using UALCAN, Kaplan-Meier plots were produced to illustrate overall survival with the use of TCGA-KIRC patient survival data. In this analysis, samples are divided into high expression and lower expression, with TPM above and below the upper quartile, respectively. The statistically significant difference between the survival curves of these groups compared by log-rank is shown as P-value.

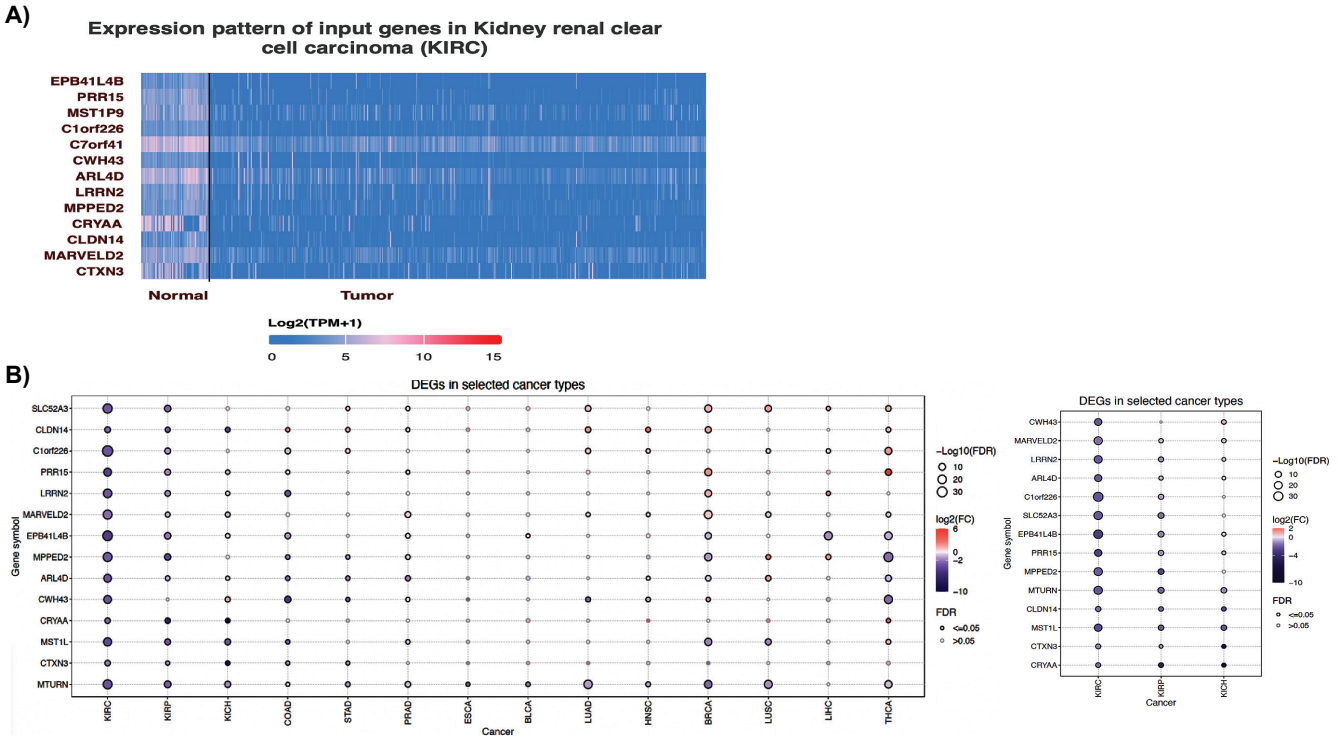

Figure 1. Expression profile of 21 genes with uncharacterized function in KIRC in cancers using UALCAN and GSCALite webservers. A) Deferential expression of these genes in KIRC as a heatmap scored based on the statistically significant under-expression. **B)** Differential expressions of the genes as bubble plot in renal cancer subtypes including KIRC, KIRP and KICH and other cancers as follow: COAD, BRCA, LUSC, THCA, LUAD, LIHC, PRAD, higher color intensity shows that the significant value is more different. The point size indicates statistical significance, indicating the bigger size representing more significant value.

3.3. Methylation Analysis

To analyze the methylation data of TCGA-KIRC for our genes, we used the GSCALite web server. For differential methylation, tumor-normal pairs are applied, and a student T-test is used to show the methylation difference between these paired samples, and the p-value is adjusted by FDR (FDR <= 0.05 is considered as significant). Moreover, the correlation between methylation and clinical overall survival data of TCGA-KIRC was also performed using GSCALite. In addition, by integrating DNA methylation and mRNA expression data of TCGA-KIRC data, methylation correlation to mRNA expression for our genes was also performed based on the Pearson correlation coefficient. The P-value is adjusted by FDR, and genes with FDR<=0.05 are considered for the final result.

3.4. Cancer Pathway Analysis

Using GSCALite, we used TCPA reverse phase protein array (RPPA) data from TCGA-KIRC samples to search for famous cancer-related pathways, which include the tuberous sclerosis complex (TSC)/mammalian target of rapamycin (mTOR), receptor tyrosine kinase (RTK), phosphoinositide 3-kinases (PI3K)/AKT protein kinase (AKT), estrogen receptor (ER), androgen receptor (AR), epithelial-mesenchymal transition (EMT), DNA damage response, cell cycle, and apoptosis pathways. This analysis uses median-centered RBN RPPA data and gives the difference in pathway activity score (PAS) between proteins with positive roles and those with negative functions in a specific pathway. If the PAS from a gene in the group with high expression is higher than its PAS in the group with low expression, it indicates the activation role of the gene on a particular pathway, but the opposite condition represents its inhibition role (p-value is adjusted by FDR and FDR<=0.05 is considered as significant).

3.5. Copy Number Variation (CNV) Analysis

Using GSCALite, we studied CNV percentage (heterozygous and homozygous CNV) and the correlation between CNV and mRNA expression (mRNA RSEM data) of our identified genes in TCGA-KIRC (these data are merged by TCGA barcode to show the final result based on Person correlation coefficient and P-value adjusted by FDR).

3.6. Drug Sensitivity

In the current study, we also investigated the expression correlation (based on the Spearman correlation) between our identified genes and drug sensitivity using GSCALite with data on Genomics of Drug Sensitivity in Cancer (GDSC). The positive correlation indicates drug resistance of the high expression of gene and the negative correlation represents drug sensitivity of the low expression of a gene.

4. Results

4.1. Novel Under-Expressed Genes In KIRC

Using TCGA-KIRC data, we looked for the top 250 down-regulated genes in KIRC compared to matched TCGA normal tissues, and we found that fourteen of these genes have not been fully studied or only mentioned as a list of identified genes in KIRC (18) (**Fig. 1 and Table 1**). Moreover, regarding the *CTXN3* gene, there are no functional studies in cancers.

As seen in **Figure 1A**, we showed the differential expression of these 14 genes as a heatmap ranked based on the statistical significance of their downregulation (more details of statistical values are given in **Table 1**). Moreover, **Figure 1B** shows the differential expressions as a bubble plot, which represents the overall expression of these genes in three subtypes of renal cancers (KIRC; kidney renal papillary cell carcinoma, KIRP; and kidney chromophobe, KICH) and other common cancers retained in final results based on parameters mentioned in method section for differential expression analysis. These cancers include colon adenocarcinoma (COAD), breast invasive carcinoma (BRCA), thyroid carcinoma (THCA), lung squamous cell carcinoma (LUSC), lung adenocarcinoma (LUAD), head and neck squamous cell carcinoma (HNSC), prostate adenocarcinoma (PRAD), stomach adenocarcinoma (STAD), liver hepatocellular carcinoma (LIHC), bladder urothelial carcinoma

Table 1. Differential expression of our 14 identified genes with uncharacterized function in KIRC.

Gene	Normal-vs-Primary (mRNA Expression)
Proline rich 15 ***(PRR15)***	Nm (20.492 TPM)-vs-Tm (0.562 TPM), Decreased: $p< 10^{-16}$
Cortexin 3 ***(CTXN3)***	Nm (19.224 TPM)-vs-Tm (0.068 TPM), Decreased: $p< 10^{-10}$
Maturin, neural progenitor differentiation regulator homolog (MTURN, C7ORF41)	Nm (90.963 TPM)-vs-Tm (9.088 TPM), Decreased: $p< 10^{-12}$
Chromosome 1 open reading frame 226 ***(C1ORF226)***	Nm (11.403 TPM)-vs-Tm (1.031 TPM), Decreased: $p< 10^{-12}$
Macrophage stimulating 1 like ***(MST1L, also known as MST1P9)***	Nm (25.141 TPM)-vs-Tm (0.939 TPM), Decreased: $p< 10^{-16}$
Metallophosphoesterase domain containing 2 ***(MPPED2)***	Nm (14.393 TPM)-vs-Tm (1.505 TPM), Decreased: $p< 10^{-12}$
MARVEL domain containing 2 ***(MARVELD2)***	Nm (27.384 TPM)-vs-Tm (3.657 TPM), Decreased: $p< 10^{-12}$
Cell wall biogenesis 43 C-terminal homolog ***(CWH43)***	Nm (12.143 TPM)-vs-Tm (0.034 TPM), Decreased: $p< 10^{-12}$
Erythrocyte membrane protein band 4.1 like 4B ***(EPB41L4B)*** *	Nm (12.36 TPM)-vs-Tm (0.096TPM), Decreased: $p< 10^{-12}$
Crystallin alpha A ***(CRYAA)***	Nm (42.994TPM)-vs-Tm (0.133TPM), Decreased: $p< 10^{-15}$
Solute carrier family 52 member 3 (***C20ORF54*** also known as ***SLC52A3***)	Nm (15.181TPM)-vs-Tm (0.766TPM), Decreased: $p< 10^{-12}$
ADP ribosylation factor like GTPase 4D ***(ARL4D)***	Nm (45.856TPM)-vs-Tm (3.656TPM), Decreased: $p< 10^{-12}$
Leucine rich repeat neuronal 2 ***(LRRN2)***	Nm (17.405TPM)-vs-Tm (0.666TPM), Decreased: $p< 10^{-12}$
Claudin 14 ***(CLDN14)***	Nm (10.351TPM)-vs-Tm (0.889TPM), Decreased: $p< 10^{-8}$

* EPB41L4B was recently reported by Bing Zhang *et al* to be correlated with worse prognosis in ccRCC patients based on the GEPIA and ULCAN websites. Nm: normal median, Tm: tumor media, TPM; transcript per million.

(BLCA), and esophageal carcinoma (ESCA). Most of the genes showed significant under-expression in KIRC, KIRP, KICH, and COAD (order in significant values). It is worth mentioning that we found *MTURN* (also known as *C7ORF41*) with only under-expression in most cancers, indicating its potential role as a tumor suppressor.

In addition, a comparison between three types of kidney cancers showed that lower expression of most of our identified genes can be observed in its three types, but *CWH43*, *EPB41L4B*, and *LRRN2*, showed over-expression in the KICH subtype; *MPPED2*, *C1ORF226*, and *C20ORF54* did not show changes in KICH subtype; and *CWH43* did not have altered expression in KIRP subtype. Therefore, the *CWH43* gene can be considered as the major signature of KIRC since it is only under-expressed in this renal subtype. In addition, *MARVELD2*, and *CRYAA* only showed altered expression (down-regulation) in the three subtypes of renal cancer, indicating their possible roles in the kidney and signature of renal cancers.

We next evaluated the expression levels of these genes based on the TCGA-KIRC molecular and histological subtypes, grades, and pathological stages. Regarding pathological stages, our data showed statistically significant downregulation of all these genes across all different stages compared to matched normal samples. However, under-expression for *PRR15*, *CTXN3, CLDN14, LRRN2, C20ORF54,* and *CRYAA* was only significant when pathological stages were compared to matched normal samples but not between each stage separately, indicating their important role in the initiation of tumorigenesis in KIRC. While all these genes can be considered as signatures of KIRC stage 1, some of them can be specifically used for certain stages. For instance, we identified the following genes as signatures of different stagesthat can help distinguish the early and the advanced KIRC: *MST1P9* for stage 2; *MARVELD2* and *C7ORF41* for stage 3; and *MST1P9* for stage 4 (**Table 2 and Deta file 1**).

Regarding KIRC grades, our data revealed all grades showed significant under-expression for all these genes compared to matched normal samples. While all these genes can be considered as signatures of KIRC grade 1, some genes were specific for given grades as follows: *PRR15*, *EPB41L4B,* and *C20ORF54* for grade 2 and *C7ORF41* for grade 4. Moreover, when each tumor grade is compared with each other, there is no significant under-expression between them for the following genes: *CTXN3*, *C1ORF226*, *CWH43*, *ARL4D*, *LRRN2*, and *CLDN14* (**Table 2 and Deta File 2**).

Regarding ccRCC subtypes, our data showed that approximately all these genes had significant under-expression in both ccA and ccB KIRC compared to matched normal cases. However, we identified significant downregulated differences for some genes when ccA and ccB are compared to each other (**Table 2**). These genes as signatures of ccA subtype (with lower expression) included *C1ORF226, PRR15*, *CLDN14*, *LRRN2*, *ARL4D*, *C20ORF54 (SLC52A3)*, and *EPB41L4B*. Signatures of ccB subtype (with lower expression) were *CWH43*, *MPPED2,* and *CTXN3* (**Table 2**).

4.2. Cancer Prognosis

The OS time between the higher expression level and lower-expression level of all the 14 identified genes with uncharacterized function in TCGA-KIRC were compared, and our data showed a shorter OS with a worse prognosis in cases with lower expression levels of 8 genes compared to their higher expression levels (**Fig. 2**, p-value < 0.05, a shorter OS with worse prognosis). These genes include *PRR15, CTXN3, C7ORF41, C1ORF226, MST1P9, MPPED2, MARVELD2,* and *CWH43.* The worst OS time can be seen for *CWH43* and *C7ORF41* ($p < 0.0001$). However, a longer OS with good prognosis was observed in cases with lower expression levels of 6 genes compared to their higher expression levels (**Fig. 3**, p-value < 0.05, a longer OS with good prognosis). These genes included *EPB41L4B, CRYAA, C20ORF54, ARL4D, LRRN2,* and *CLDN14.*

4.3. Methylation Analysis

Our data revealed that in TCGA-KIRC, most of the 14 genes are affected by upper promoter methylation compared to matched normal samples, mainly for *MPPED2*, *EPB41L4B*, and *ARL4D (***Fig. 4A**). However, *CTXN3* showed lower promoter methylation (**Fig. 4A**). Some genes, which include *C7ORF41*, *C1ORF226*, *C20ORF54*, and *MST1P9* (*MST1L*) were not retained in the final results of methylation difference analysis based on the parameters explained in the method section.

Table 2. Expression levels of 14 identified genes based on TCGA-KIRC molecular and histological subtypes, grades, and pathological stages.

Gene	ccRCC Subtypes	Tumor Grade		Tumor stages	
PRR15	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-16}$ ccA-vs-ccB: $p< 10^{-8}$	N-vs-G1: $p< 10^{-12}$ N-vs-G2: $p< 10^{-16}$ N-vs-G3: $p< 10^{-16}$ N-vs-G4: $p< 10^{-16}$	G1-vs-G2:0.0077 G1-vs-G3: p=0.91 G1-vs-G4: p=0.63 G2-vs-G3: p=0.25 G2-vs-G4: p=0.99 G3-vs-G4: p=0.44	N-vs-S1: $p < 10^{-16}$ N-vs-S2: $p < 10^{-16}$ N-vs-S3: $p < 10^{-16}$ N-vs-S4: $p < 10^{-12}$	S1-vs-S2: p=0.83 S1-vs-S3: p=0.16 S1-vs-S4: p=0.26 S2-vs-S3: p=0.48 S2-vs-S4: p= 0.49 S3-vs-S4: p= 0.91 Stage Signature:1
CTXN3	N-vs-ccA: $p< 10^{-9}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: $p< 10^{-5}$	N-vs-G1: $p< 10^{-10}$ N-vs-G2: $p< 10^{-11}$ N-vs-G3: $p< 10^{-10}$ N-vs-G4: $p< 10^{-10}$	G1-vs-G2: p=0.7 G1-vs-G3: p=0.54 G1-vs-G4: p=0.64 G2-vs-G3: p=0.2 G2-vs-G4: p=0.82 G3-vs-G4: p=0.25	N-vs-S1: $p< 10^{-10}$ N-vs-S2: $p< 10^{-10}$ N-vs-S3: $p< 10^{-9}$ N-vs-S4: $p< 10^{-10}$	S1-vs-S2: p=0.8 S1-vs-S3: p=0.5 S1-vs-S4: p=0.83 S2-vs-S3: p=0.43 S2-vs-S4: p=0.99 S3-vs-S4: p=0.46 Stage Signature:1
C7ORF41 (MTURN)	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: p=0.27	N-vs-G1: $p< 10^{-12}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-12}$	G1-vs-G2: p=0.81 G1-vs-G3: p=0.16 G1-vs-G4: p=0.038 G2-vs-G3: $p< 10^{-6}$ G2-vs-G4: $p< 10^{-12}$ G3-vs-G4: $p< 10^{-4}$	N-vs-S1: $p< 10^{-12}$ N-vs-S2: $p< 10^{-12}$ N-vs-S3: $p< 10^{-12}$ N-vs-S4: $p< 10^{-12}$	S1-vs-S2: p=0.57 S1-vs-S3: $p< 10^{-9}$ S1-vs-S4: $p< 10^{-9}$ S2-vs-S3: p=0.0037 S2-vs-S4: $p< 10^{-4}$ S3-vs-S4: p=0.36 Stage Signature:1 and 3
C1ORF226	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: p=0.0016	N-vs-G1: $p< 10^{-12}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-12}$	G1-vs-G2: p=0.29 G1-vs-G3: p=0.89 G1-vs-G4: p=0.92 G2-vs-G3: p=0.1 G2-vs-G4: p=0.41 G3-vs-G4: p=0.83	N-vs-S1: $p< 10^{-12}$ N-vs-S2: $p< 10^{-12}$ N-vs-S3: $p< 10^{-12}$ N-vs-S4: $p< 10^{-12}$	S1-vs-S2: p=0.25 S1-vs-S3: p=0.168 S1-vs-S4: p=0.01 S2-vs-S3: p=0.097 S2-vs-S4: p=0.037 S3-vs-S4: p=0.35 Stage Signature:1
MST1P9 (MST1L)	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: p=0.053	N-vs-G1: $p< 10^{-11}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-12}$	G1-vs-G2: p=0.72 G1-vs-G3: p=0.58 G1-vs-G4: p=0.3 G2-vs-G3: p=0.015 G2-vs-G4: $p< 10^{-4}$ G3-vs-G4: p=0.065	N-vs-S1: $p< 10^{-12}$ N-vs-S2: $p< 10^{-12}$ N-vs-S3: $p< 10^{-16}$ N-vs-S4: $p< 10^{-12}$	S1-vs-S2: p=0.0027 S1-vs-S3: p=0.19 S1-vs-S4: $p< 10^{-9}$ S2-vs-S3: p=0.12 S2-vs-S4: p=0.19 S3-vs-S4: p=0.0013 Stage Signature: 1, 2 and 4
MPPED2	N-vs-ccA: $p< 10^{-16}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: p=0.013	N-vs-G1: $p< 10^{-12}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-12}$	G1-vs-G2: p=0.8 G1-vs-G3: p=0.12 G1-vs-G4: p=0.056 G2-vs-G3: p=0.018 G2-vs-G4: $p< 10^{-4}$ G3-vs-G4: p=0.16	N-vs-S1: $p< 10^{-12}$ N-vs-S2: $p< 10^{-12}$ N-vs-S3: $p< 10^{-12}$ N-vs-S4: $p< 10^{-12}$	S1-vs-S2: p=0.65 S1-vs-S3: $p< 10^{-4}$ S1-vs-S4: $p< 10^{-5}$ S2-vs-S3: p=0.1 S2-vs-S4: p=0.62 S3-vs-S4: p=0.36 Stage Signature:1
MARVELD2	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: p=0.16	N-vs-G1: $p< 10^{-12}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-12}$	G1-vs-G2: p=0.051 G1-vs-G3: p=0.43 G1-vs-G4: p=0.43 G2-vs-G3: p=0.05 G2-vs-G4: $p< 10^{-7}$ G3-vs-G4: $p< 10^{-4}$	N-vs-S1: $p< 10^{-12}$ N-vs-S2: $p< 10^{-12}$ N-vs-S3: $p< 10^{-12}$ N-vs-S4: $p< 10^{-12}$	S1-vs-S2: p=0.46 S1-vs-S3: $p< 10^{-4}$ S1-vs-S4: $p< 10^{-8}$ S2-vs-S3: p=0.021 S2-vs-S4: p=0.0026 S3-vs-S4: p=0.89 Stage Signature: 1 and 3

CWH43	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-7}$ c c A - v s - c c B : p=0.018	N-vs-G1: $p< 10^{-10}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-14}$	G1-vs-G2: p=0.85 G1-vs-G3: p=0.63 G1-vs-G4: p=0.61 G2-vs-G3: p=0.68 G2-vs-G4: p=0.67 G3-vs-G4: p=0.94	N-vs-S1: $p < 10^{-12}$ N-vs-S2: $p < 10^{-13}$ N-vs-S3: $p < 10^{-12}$ N-vs-S4: $p < 10^{-16}$	S1-vs-S2: p=0.8 S1-vs-S3: p=0.019 S1-vs-S4: p=0.37 S2-vs-S3: p=0.073 S2-vs-S4: p=0.55 S3-vs-S4: p=0.32 Stage Signature:1
EPB41L4B	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: $p< 10^{-5}$	N-vs-G1: $p< 10^{-12}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-12}$	G1-vs-G2: p=0.025 G1-vs-G3: p=0.035 G1-vs-G4: p=0.096 G2-vs-G3: p=0.73 G2-vs-G4: p=0.9 G3-vs-G4: p=0.88	N-vs-S1: $p < 10^{-12}$ N-vs-S2: $p< 10^{-12}$ N-vs-S3: $p< 10^{-12}$ N-vs-S4: $p< 10^{-12}$	S1-vs-S2: p=0.28 S1-vs-S3: p 0.026 S1-vs-S4: p=0.85 S2-vs-S3: p=0.1 S2-vs-S4: p= 0.27 S3-vs-S4: p= 0.23 Stage Signature:1
CRYAA	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-6}$ ccA-vs-ccB: p=0.11	N-vs-G1: $p< 10^{-12}$ N-vs-G2: $p< 10^{-11}$ N-vs-G3: $p< 10^{-8}$ N-vs-G4: $p< 10^{-11}$	G1-vs-G2:0.09 G1-vs-G3: p=0.2 G1-vs-G4: p=0.013 G2-vs-G3: p=0.6 G2-vs-G4: p=0.58 G3-vs-G4: p=0.42	N-vs-S1: $p < 10^{-12}$ N-vs-S2: $p < 10^{-5}$ N-vs-S3: $p < 10^{-5}$ N-vs-S4: $p < 10^{-12}$	S1-vs-S2: p=0.37 S1-vs-S3: p=0.28 S1-vs-S4: p=0.77 S2-vs-S3: p=0.88 S2-vs-S4: p= 0.39 S3-vs-S4: p= 0.3 Stage Signature:1
C20ORF54 ***(SLC52A3)***	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: $p< 10^{-8}$	N-vs-G1: $p< 10^{-12}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-12}$	G1-vs-G2: p=0.021 G1-vs-G3: p=0.19 G1-vs-G4: p=0.01 G2-vs-G3: p=0.15 G2-vs-G4: p=0.34 G3-vs-G4: p=0.061	N-vs-S1: $p< 10^{-12}$ N-vs-S2: $p< 10^{-12}$ N-vs-S3: $p< 10^{-12}$ N-vs-S4: $p < 10^{-12}$	S1-vs-S2: p=0.47 S1-vs-S3: p=0.81 S1-vs-S4: p=0.48 S2-vs-S3: p=0.43 S2-vs-S4: p=0.68 S3-vs-S4: p=0.42 Stage Signature:1
ARL4D	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-14}$ ccA-vs-ccB: $p< 10^{-8}$	N-vs-G1: $p< 10^{-4}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-16}$	G1-vs-G2: p=0.35 G1-vs-G3: p=0.26 G1-vs-G4: p=0.28 G2-vs-G3: p=0.86 G2-vs-G4: p=0.26 G3-vs-G4: p=0.67	N-vs-S1: $p< 10^{-15}$ N-vs-S2: $p < 10^{-12}$ N-vs-S3: $p< 10^{-12}$ N-vs-S4: $p< 10^{-16}$	S1-vs-S2: p=0.084 S1-vs-S3: p=0.001 S1-vs-S4: p=0.03 S2-vs-S3: p=0.33 S2-vs-S4: p=0.8 S3-vs-S4: p=0.44 Stage Signature:1
LRRN2	N-vs-ccA: $p< 10^{-12}$ N-vs-ccB: $p< 10^{-12}$ ccA-vs-ccB: $p< 10^{-8}$	N-vs-G1: $p< 10^{-8}$ N-vs-G2: $p< 10^{-12}$ N-vs-G3: $p< 10^{-12}$ N-vs-G4: $p< 10^{-12}$	G1-vs-G2: p=0.79 G1-vs-G3: p=0.89 G1-vs-G4: p=0.8 G2-vs-G3: p=0.79 G2-vs-G4: p=0.98 G3-vs-G4: p=0.82	N-vs-S1: $p< 10^{-12}$ N-vs-S2: $p< 10^{-12}$ N-vs-S3: $p< 10^{-12}$ N-vs-S4: $p< 10^{-12}$	S1-vs-S2: p=0.8 S1-vs-S3: p=0.84 S1-vs-S4: p=0.34 S2-vs-S3: p=0.96 S2-vs-S4: p=0.62 S3-vs-S4: p=0. 6 Stage Signature:1
CLDN14	N-vs-ccA: $p< 10^{-9}$ N-vs-ccB: $p< 10^{-8}$ ccA-vs-ccB: p=0.013	N-vs-G1: $p< 10^{-8}$ N-vs-G2: $p< 10^{-8}$ N-vs-G3: $p< 10^{-8}$ N-vs-G4: $p< 10^{-8}$	G1-vs-G2: p=0.78 G1-vs-G3: p=0.45 G1-vs-G4: p=0.57 G2-vs-G3: p=0.54 G2-vs-G4: p=0.7 G3-vs-G4: p=0.84	N-vs-S1: $p< 10^{-8}$ N-vs-S2: $p< 10^{-9}$ N-vs-S3: $p< 10^{-8}$ N-vs-S4: $p< 10^{-7}$	S1-vs-S2: p=0.23 S1-vs-S3: p=0.94 S1-vs-S4: p=0.26 S2-vs-S3: p=0.06 S2-vs-S4: p=0.097 S3-vs-S4: p=0.23 Stage Signature:1

N: normal; G: grade, S: stage; p: p value

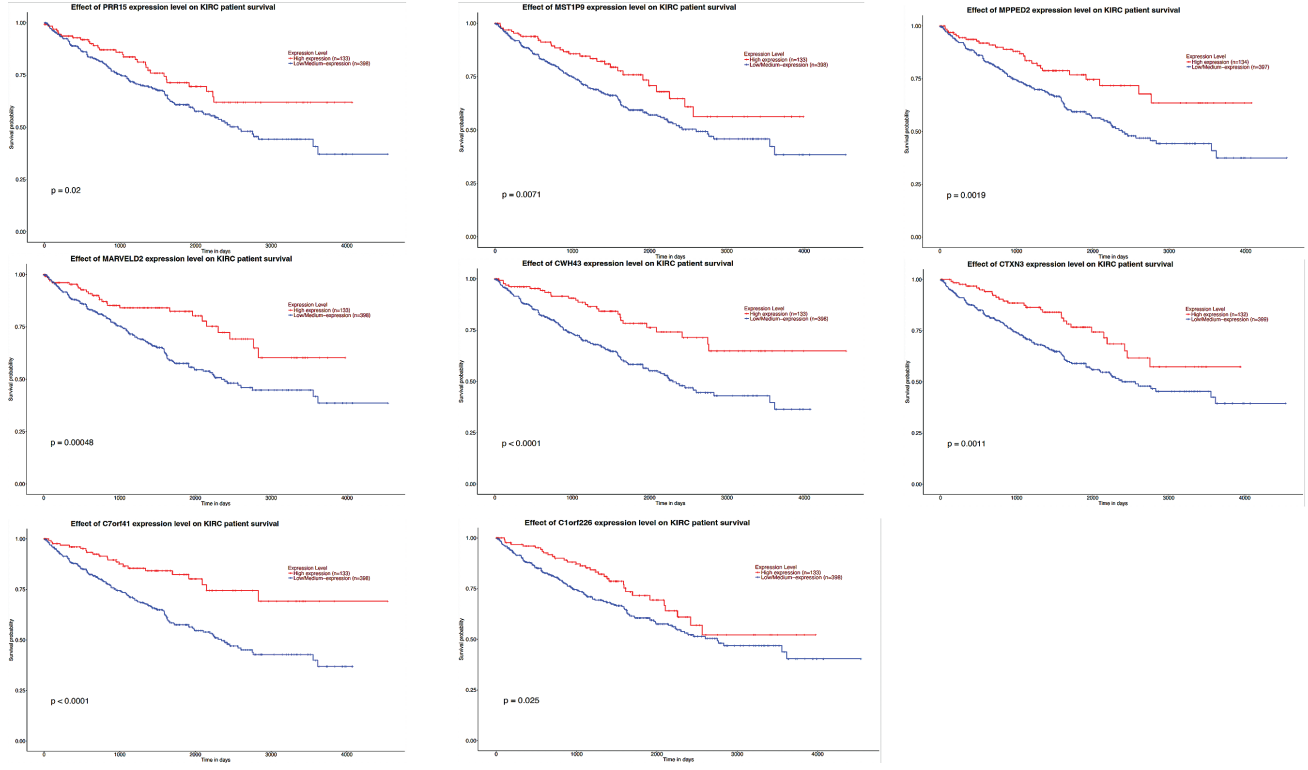

Figure 2. Shorter OS with worse prognosis in cases with lower expression levels of 8 genes compared to their higher expression using UALCAN webserver.

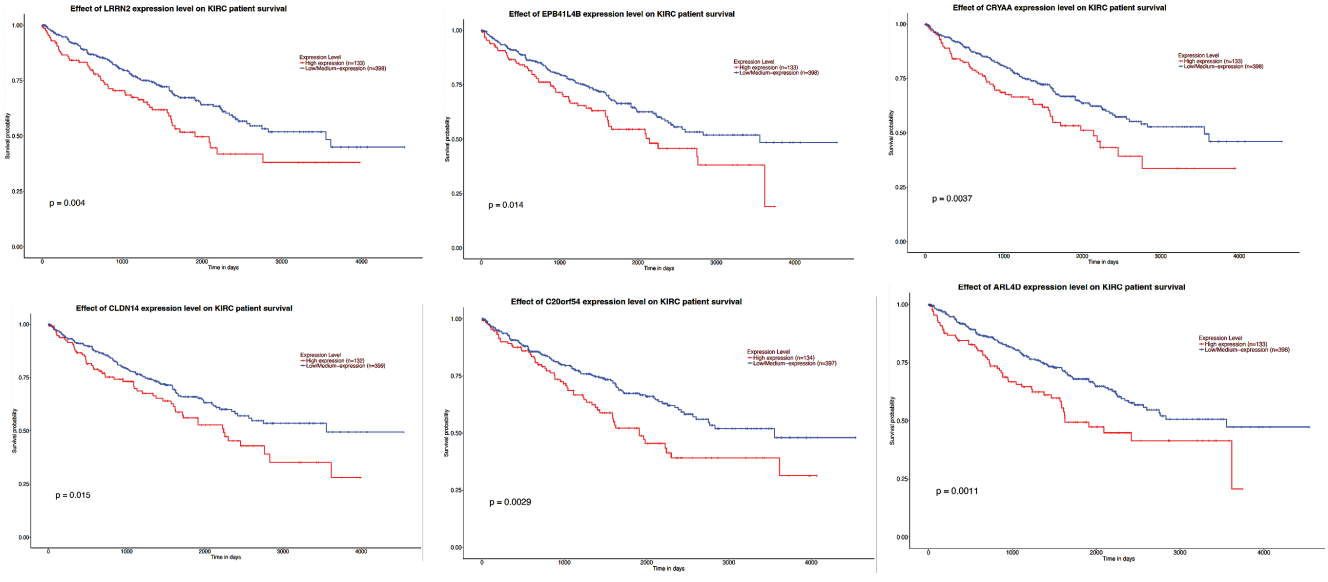

Figure 3. Longer OS with good prognosis in cases with lower expression levels of 6 genes compared to their higher expression levels using UALCAN webserver.

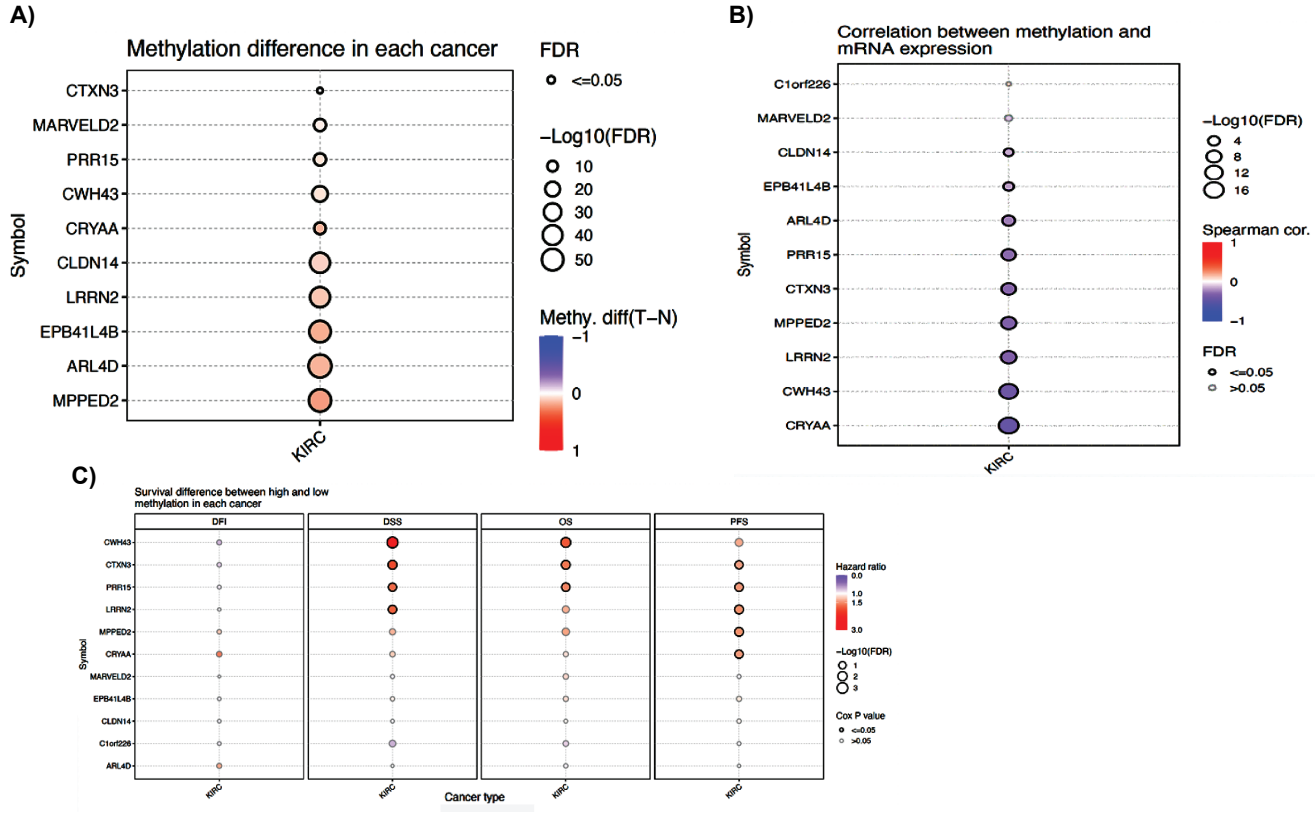

Figure 4. Promoter methylation status in our identified genes with uncharacterized function in KIRC using GSCALite webserver. A) Methylation difference between KIRC and matched normal samples for the genes. The blue and red point shows decreased and increased promoter methylation, respectively, the higher color intensity illustrates that the significant value is more different. The point size shows a statistically significant value, indicating the bigger size represents a more significant value. **B)** Person correlation between methylation and mRNA expression of our identified genes. The blue point shows a negative correlation which represents that the increased promoter methylation of these genes leads to the gene under-expression. The red point indicates a positive correlation, representing that increased promoter methylation leads to gene over-expression. The higher color intensity illustrates the higher significant correlation. The point size shows a statistically significant value, indicating the bigger size represents a more significant value. **C)** Overall survival difference between hypermethylation and hypomethylation. Only genes with log-rank p-value significant (<=0.05) are retained in the result. The red point shows a low worse for the gene with a high methylation group and the blue point represents the opposite. The point size indicates statistical significance, indicating the bigger size represents a more significant value.

Regarding the correlation between methylation and gene expression, 9 out of these 14 identified genes in our study showed a negative correlation, mainly for *CRYAA* (FDR: 2.26e-17, spearman correlation: -0.45) (**Fig. 4B**). *C7ORF41*, *C20ORF54*, and *MST1P9* did not retain in final results based on parameters used in this analysis explained in the method section. We also analyzed survival differences between KIRC patients with hyper- and hypomethylation. Our data showed significantly worse survival for patients with high methylation in certain genes, particularly CWH43 and CTXN3 (**Fig. 4C**).

4.4. Cancer Pathway Analysis

As shown in **Figure 5** and **Table 3**, we found different proportions of activation or inhibition roles of our studied genes in cancer pathways. **Figure 5B** and **Table 3** show that *MTURN* had only activation role (not inhibition) on particular pathways, including hormone AR, Hormone ER, and RTK pathways. Regarding inhibition on cancer

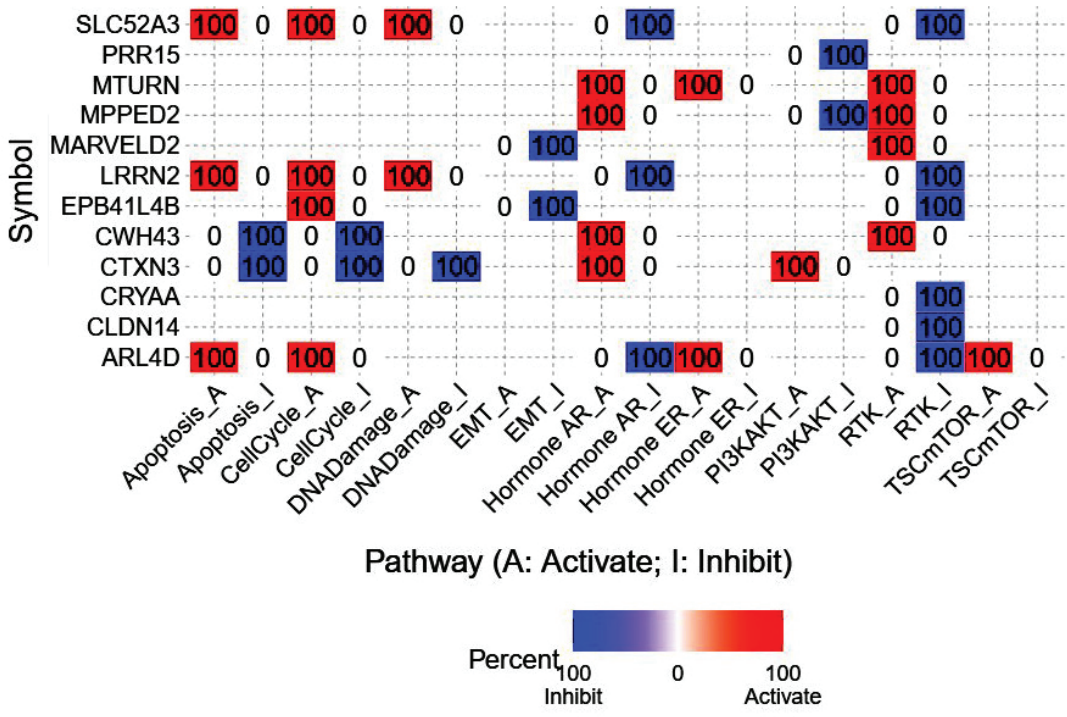

Figure 5. Role of the 14 genes in cancer pathways using GSCALite webserver. A) it shows heatmap percentage. It only illustrates genes with inhibition or activation roles in = > 5 cancer types. Pathway_A: activation of the pathway; Pathway_I: inhibition of the pathway.

pathways, some genes only showed inhibition roles as follows: *PRR15* on PI3KAKT pathway, *CRYAA*, and *CLDN14* on RTK pathway. However, some genes were involved in both inhibition and activation of different pathways (**Fig. 5B and Table 3**).

Notably, 12 genes were involved in different pathways, mainly the hormone RTK pathway (mainly inhibition: 6 genes versus 4 genes for activation, **Fig. 5 and Table 3**). Other main pathways were Hormone AR (4 genes for activation versus 3 for inhibition), cell cycle (4 genes for activation versus 2 for inhibition), apoptosis (3 genes for activation versus 2 for inhibition), DNA damage response (2 genes for activation vs. 1 for inhibition), PI3KAKT (2 genes for inhibition vs. 1 for activation), EMT only for inhibition, TSCmTOR andHormone ER only for activation. Moreover, the only gene that was involved in TSCmTOR pathway was *ARL4D*. Furthermore, among these genes, some of them were only involved in one pathway, including *PRR15* in PI3KAKT pathway and *CRYAA* and *CLDN14* in RTK pathway.

4.5.CNV Analysis

In our study, we found several genes with heterozygous amplification and deletion in TCGA-KIRC (Figure 6A). The frequency of amplification was higher for *CTXN3, PRR15, MTURN (C7orf41),* and *MARVELD2*. However, for heterozygous deletion, the frequency was higher for *EPB41L4B* (**Fig. 6A**). Regarding homozygosity of deletion and amplification, only *CTXN3* showed a much more homozygous amplification percentage, 12.69%, (**Fig. 6B**). Moreover, the Pearson correlation between CNV and mRNA expression (RSEM) revealed no significant negative correlation to mRNA expression for these genes (**Fig. 6C**). However, some genes showed a positive correlation (**Fig. 6C**).

4.6. Drug Sensitivity

In our study, we also investigated gene expression correlation with drug resistance using GDSC drug data. Our results found a negative correlation for some genes, which revealed that their low gene expression was sensitive to some drugs (**Fig. 7**, blue points).

Table 3. Role of our proposed genes in cancer pathways.

Pathway	Genes that only inhibit a pathway	Genes that only activate a pathway
Apoptosis	*CTXN3* and *CWH43*	*ARL4D, LRRN2,* and *C20ORF54 (SLC52A3)*
Cell cycle	*CTXN3* and *CWH43*	*ARL4D, LRRN2, EPB41L4B,* and *C20ORF54 (SLC52A3)*
DNA damage response	*CTXN3*	*LRRN2* and *C20ORF54 (SLC52A3)*
EMT	*MARVELD2* and *EPB41L4B*	
Hormone AR	*ARL4D, LRRN2, C20ORF54 (SLC52A3)*	*CTXN3, CWH43, MPPED2* and *MTURN*
Hormone ER		*ARL4D* and *MTURN*
PI3K/AKT	*MPPED2, PRR15*	*CTXN3*
RTK	*ARL4D, CLDN14, CRYAA, EPB41L4B, LRRN2, C20ORF54 (SLC52A3)*	*MARVELD2, CWH43, MPPED2, MTURN*
TSC/mTOR		*ARL4D*

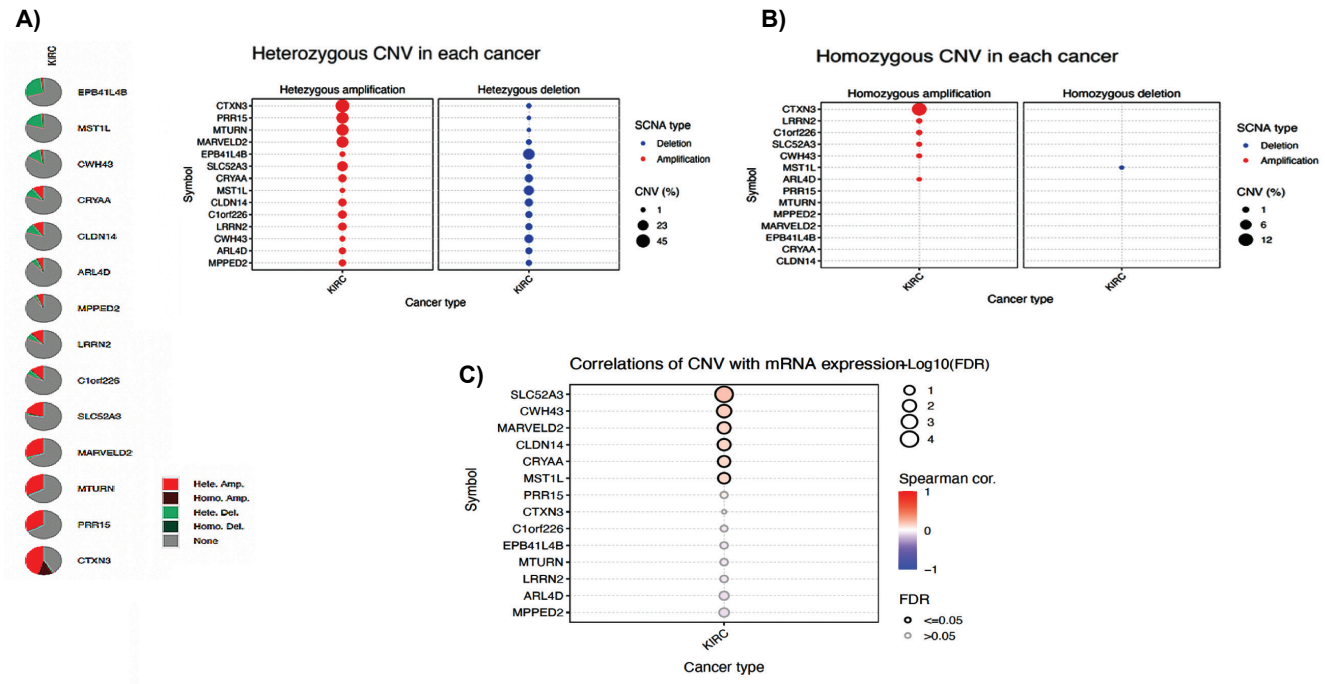

Figure 6. CNV analysis in our 14 identified genes with uncharacterized function in KIRC using GSCALite webserver. A) CNV Pie chart shows the total percentage of CNVs in the 14 genes in TCGA-KIRC patients. Hete Amp: heterozygous amplification; Hete Del: heterozygous deletion; Homo Amp: homozygous amplification; Homo Del: homozygous deletion; None: no CNV. **B)** Heterozygous and homozygous CNV profile shows the proportion of heterozygous and homozygous amplification and deletion in each gene in TCGA-KIRC samples. Results only give genes affecting more than 5% CNV in cancers. **C)** Correlation between CNV and gene expression. Genes with a significant correlation between mRNA expression and CNV percentage (FDR<=0.05) are given. The blue bubble shows a negative correlation which represents that the gene with a highly frequent CNV leads to the gene under-expression. The red bubble indicates a positive correlation which shows that the gene with a highly frequent CNV is affected by the gene over-expression. The higher color intensity represents the higher significant correlation. The point size shows a statistically significant value, indicating the bigger size represents more significant value.

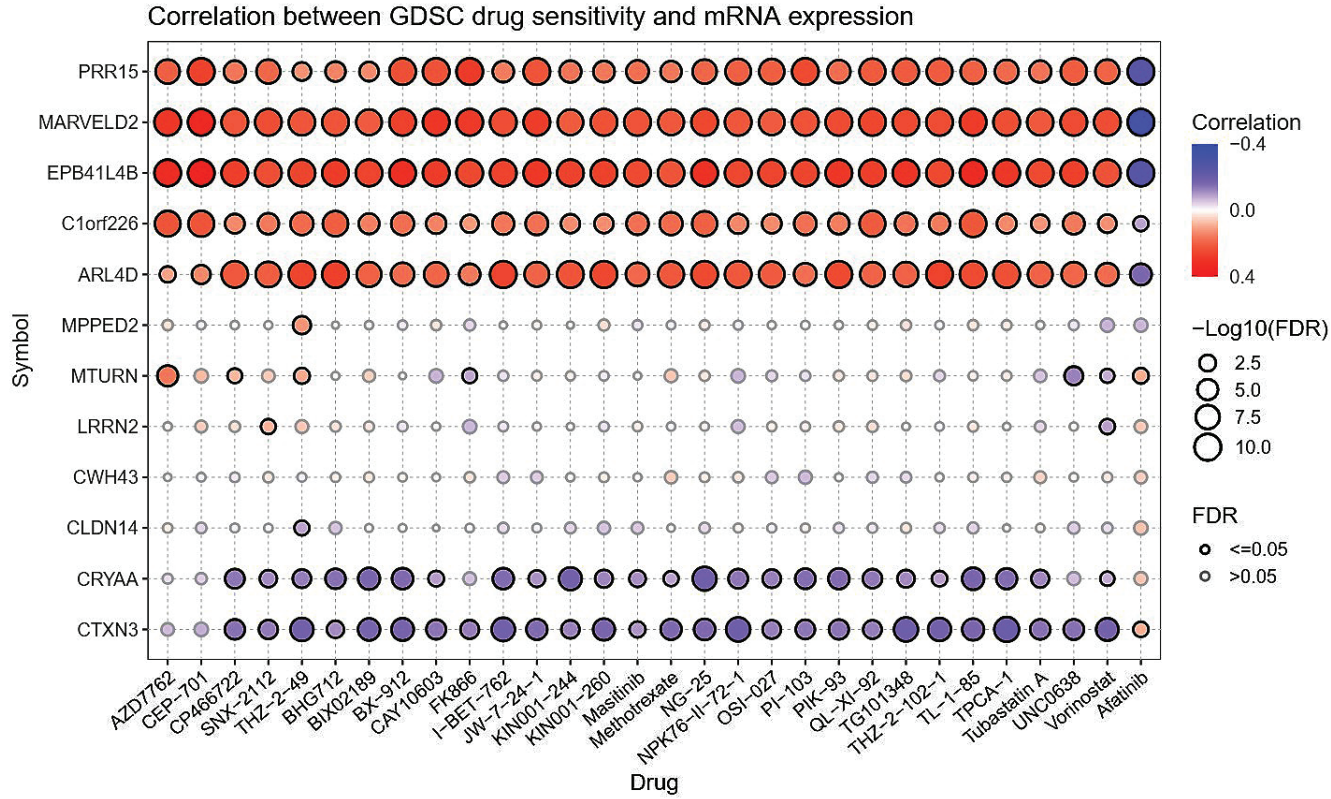

Figure 7. Drug sensitivity analysis from GDSC using GSCALite webserver. The Spearman correlation is considered between gene expression and drug sensitivity. The positive correlation (red color) represents drug resistance for genes with high expression, negative correlation (blue points) shows the opposite. The higher color intensity represents the higher significant correlation. The point size shows a statistically significant value, indicating the bigger size represents a more significant value.

As an example, regarding *EPB41L4B, MARVELD2, PRR15, C1ORF226,* and *ARL4D,* they showed the most significant sensitivity for Afatinib. As seen in Figure 7, drug sensitivity to low expression of *MARVELD2* and *PRR15* can be applied to three subtypes of renal cancer since we also identified decreased expression of these genes in these subtypes as mentioned in previous sections.

5. Discussion

In this study, we demonstrated a comprehensive analysis of 14 uncharacterized down-regulated genes in KIRC concerning the matched normal counterparts. We found that these genes showed notably low expression in KIRC, and also most of them had downregulation in KICH, KIRP, and COAD. It should be noted that *C7ORF41* was previously reported generally in RCC but not specifically in KIRC (19). Our study revealed 8 genes with worse OS and 6 genes with good prognosis in KIRC. We showed that these 8 genes can be considered tumor suppressor genes in KIRC since they showed lower expression across different pathological stages and tumor grades in this cancer and also revealed shorter survival time and worse prognosis. We found that *C7ORF41* and *CTXN3* could function only as tumor suppressors since they only showed downregulation in the cancers.

Our data proposed that*MARVELD2* and *CRYAA* with possible roles in the kidney could be considered as signatures of renal cancers due to their down-regulation only in the three subtypes of renal cancer. However, it is worth mentioning that our data proposed the *CWH43* gene as the major signature of KIRC due to its only under-expression in this renal subtype. This data will be very helpful in distinguishing the subtypes of renal cancer where it is necessary to consider specific

treatments.

Regarding pathological KIRC stages and grades, our data revealed different signatures for the stages and grades (**Table 2** and **Deta File 1, 2**). These data are very useful for the diagnosis of this cancer and applying possible specific treatments. We found some gene signatures for the early stage II (*MST1P9*) and advanced stages of KIRC [stage III (*MARVELD2* and *C7ORF41*) and IV *(MST1P9*)], which were found to have discerning power between different KIRC stages. These genes with down-regulated expression may help to find the main biological changes during tumorigenesis in KIRC and serve as biomarkers for KIRC stages. Moreover, our data revealed two genes as both grade and stage signatures of KIRC, which included *C7ORF41*, representing the involvement of the same pathways in KIRC differentiation and progression.

From 8 new possible tumor suppressor genes in KIRC, *CTXN3* has not been functionally and not fully studied in cancers yet, and our data can propose it as a tumor suppressor gene.

Previous research indicates an altered expression of cortexin 3 (*CTXN3*), either alone or in combination with modifications to DISC1, could subtly interfere with GABAergic neurotransmission and/or metabolism of amyloid precursor protein (APP) in the developing brain, which could expose the affected individual to a higher risk of schizophrenia later in life (20). The function of *CTXN3* (cortexin 3) has not been exactly reported, but it was predicted that it is a highly conserved protein in vertebrates with a single membrane-spanning domain that may involve extracellular or intracellular signaling pathways in the kidney or brain (21). Our data extracted from STRING predicted its possible indirect involvement in the Hippo/SWH (Sav/Wts/Hpo) signaling pathway, which has an important role in tumor suppression **(Additional File 3, panel 1**). Moreover, our data extracted from GTEx identified its highest expression in normal kidneys compared to other tissues **(Deta file 3, panel 2**), and our results representing its under-expression and worse prognosis in KIRC propose its function as a tumor suppressor.

Regarding *MST1P9* (*MST1L, p*utative macrophage stimulating 1-like protein), it is a pseudogene with possible serine-type endopeptidase activity (22), and with our data showing its downregulated expression, shorter OS, and worse prognosis, it can be proposed that this gene be a tumor suppressor. Information about other genes is summarized in **Additional File 4**. The present research offers new insight into the molecular mechanisms that led to RCC development, although the possible involvement of these genes requires further investigation.

Acknowledgment

The data that support the findings of this study are available from the corresponding author upon reasonable request. TCGA data extracted from UALCAN (http://ualcan.path.uab.edu) and GSCALite (http://bioinfo.life.hust.edu.cn/web/GSCALite/) was used for our analysis in the current study.

Authors' contributions

HD and EM designed the experiments, performed experiments, and collected data; HD, EM, HHA, ZF, and YM discussed the results and strategy. YM Supervised, directed, and managed the study. All authors approved the version to be published.

Conflict of interest

None. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The raw data supporting the conclusions of this article are available from the corresponding author upon reasonable request.

References

1. Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, *et al*. Renal cell carcinoma. *Nat Rev Dis Primers*. 2017;**3**:17009. doi: 10.1038/nrdp.2017.9
2. Moch H. [The WHO/ISUP grading system for renal carcinoma]. Pathologe. 2016;**37**(4):355-360. doi:10.1007/s00292-016-0171-y
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin*. 2015;**65**(1):5-29. doi:10.3322/caac.21254
4. Alchahin AM, Mei S, Tsea I, Hirz T, Kfoury Y, Dahl D, *et al*. A transcriptional metastatic signature predicts survival in clear cell renal cell carcinoma. *Nat Commun*. 2022;**13**(1):5747. doi:10.1038/s41467-022-33375-w
5. Dudani S, de Velasco G, Wells JC, Gan CL, Donskov F, Porta C, *et al*. Evaluation of Clear Cell, Papillary, and Chromophobe Renal Cell Carcinoma Metastasis Sites and Association With Survival. *JAMA Netw Open*. 2021;**4**(1):e2021869. doi:10.1001/jamanetworkopen.2020.21869
6. Motzer RJ, Jonasch E, Agarwal N, Bhayani S, Bro WP, Chang SS, *et al*. Kidney Cancer, Version 2.2017, NCCN Clinical

Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2017;**15**(6):804-834. doi:10.6004/jnccn.2017.0100
7. Chen YY, Hu HH, Wang YN, Liu JR, Liu HJ, Liu JL, *et al*. Metabolomics in renal cell carcinoma: From biomarker identification to pathomechanism insights. *Arch Biochem Biophys*. 2020;**695**:108623. doi:10.1016/j.abb.2020.108623
8. Li F, Jin Y, Pei X, Guo P, Dong K, Wang H, *et al*. Bioinformatics analysis and verification of gene targets for renal clear cell carcinoma. *Comput Biolog Chem*. 2021;**92**:107453. doi:10.1016/j.compbiolchem.2021.107453
9. Somsuan K, Aluksanasuwan S. Bioinformatic analyses reveal the prognostic significance and potential role of ankyrin 3 (ANK3) in kidney renal clear cell carcinoma. *Genomics Inform*. 2023;**21**(2):e22. doi:10.5808/gi.23013
10. Xiao W, Wang X, Wang T, Xing J. Overexpression of BMP1 reflects poor prognosis in clear cell renal cell carcinoma. *Cancer Gene Ther*. 2019. doi:10.1038/s41417-019-0107-9
11. Batai K, Imler E, Pangilinan J, Bell R, Lwin A, Price E, *et al*. Whole-transcriptome sequencing identified gene expression signatures associated with aggressive clear cell renal cell carcinoma. *Genes Cancer*. 2018;**9**(5-6):247-256. doi:10.18632/genesandcancer.183
12. Han G, Zhao W, Song X, Kwok-Shing Ng P, Karam JA, Jonasch E, *et al*. Unique protein expression signatures of survival time in kidney renal clear cell carcinoma through a pan-cancer screening. *BMC Genomics*. 2017;**18**(Suppl6):678. doi:10.1186/s12864-017-4026-6
13. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, *et al*. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;**45**(10):1113-1120. doi:10.1038/ng.2764
14. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, *et al*. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res*. 2013;**41**(Database issue):D991-D995. doi:10.1093/nar/gks1193
15. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, *et al*. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia*. 2017;**19**(8):649-658. doi:10.1016/j.neo.2017.05.002
16. Danielsson F, Skogs M, Huss M, Rexhepaj E, O'Hurley G, Klevebring D, *et al*. Majority of differentially expressed genes are down-regulated during malignant transformation in a four-stage model. *Proc Natl Acad Sci U S A*. 2013;**110**(17):6853-6858. doi:10.1073/pnas.1216436110
17. Liu CJ, Hu FF, Xia MX, Han L, Zhang Q, Guo AY. GSCALite: a web server for gene set cancer analysis. *Bioinformatics*. 2018;**34**(21):3771-3772. doi:10.1093/bioinformatics/bty411
18. 18. Wu P, Xiang T, Wang J, Lv R, Wu G. TYROBP is a potential prognostic biomarker of clear cell renal cell carcinoma. *FEBS Open Bio*. 2020;**10**(12):2588-2604. doi:10.1002/2211-5463.12993
19. Deng J, Kong W, Mou X, Wang S, Zeng W. Identifying novel candidate biomarkers of RCC based on WGCNA analysis. *Per Med*. 2018;**15**(5):381-394. doi:10.2217/pme-2017-0091
20. Šerý O, Lochman J, Povová J, Janout V, Plesník J, Balcar VJ. Association between 5q23.2-located polymorphism of CTXN3 gene (Cortexin 3) and schizophrenia in European-Caucasian males; implications for the aetiology of schizophrenia. *Behavioral and Brain Functions*. 2015;**11**(1):10. doi:10.1186/s12993-015-0057-9
21. Wang HT, Chang JW, Guo Z, Li BG. In silico-initiated cloning and molecular characterization of cortexin 3, a novel human gene specifically expressed in the kidney and brain, and well conserved in vertebrates. *Int J Mol Med*. 2007;**20**(4):501-510. doi:10.3892/ijmm.20.4.501
22. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, *et al*. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics*. 2016;**54**:1-30-1-1- 3. doi:10.1002/cpbi.5