Quantitative Biology 2022, 10(1): 79–93
https://doi.org/10.15302/J-QB-021-0279

RESEARCH ARTICLE

DNA methylation profiling reveals new potential subtype-specific gene markers for early-stage renal cell carcinoma in Caucasian population

Alvaro Filbert Liko[1], Edward Ciputra[1], Nathaniel Alvin Sanjaya[1], Priskila Cherisca Thenaka[1], David Agustriawan[2,*]

[1] Department of Biomedicine, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta 13210, Indonesia
[2] Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta 13210, Indonesia
* Correspondence: david.agustriawan@i3l.ac.id

Received December 17, 2020; Revised July 1, 2021; Accepted July 2, 2021

***Background*: Renal cell carcinoma (RCC) is among the top adult cancers worldwide, with a challenging management due to lack of early diagnosis, therapy resistance, and diverse molecular background. Aberrant DNA methylation has been associated with RCC development due to transcription deregulation. We discovered potential DNA methylation-based biomarkers for stage I RCC in Caucasian population from The Cancer Genome Atlas (TCGA) database.**
***Methods*: Patients' clinical, methylation beta-value, and mRNA expression data were retrieved. Differential methylation and expression analysis were conducted to obtain differentially methylated CpG-gene pairs. Inversely correlated CpG-gene pairs between their expression and methylation levels were selected using Pearson's correlation test and then screened for any recorded somatic mutations. Their biomarker capacities were analyzed using the Kaplan-Meier and receiver operating characteristic analysis, followed by protein network and functional enrichment analysis.**
***Results*: We obtained differentially methylated CpGs in clear cell (KIRC) and papillary RCC (KIRP) but not chromophobe RCC (KICH). Six inversely correlated CpG-gene pairs with no reported cancer-associated mutations were selected. Prognostic values were found in *ATXN1* and *RFTN1* for KIRC, along with *GRAMD1B* and *TM4SF19* for KIRP, while diagnostic values were found in *VIM* and *RFTN1* for KIRC, along with *TNFAIP6* and *TM4SF19* for KIRP. Both subtypes showed enrichment of immune and metabolism-related pathways.**
***Conclusion*: We discovered novel potential DNA methylation-based prognostic and diagnostic markers for early-stage RCC in Caucasian population. Validation by wet laboratory analysis and adjustments for confounding variables might be needed, considering our study limitation to specific race.**

Keywords: biomarker; chromophobe renal cell carcinoma; clear cell renal cell carcinoma; DNA methylation; papillary renal cell carcinoma; TCGA

Author summary: Renal cell carcinoma (RCC) is strongly associated with epigenetic aberrations, including DNA methylation. The lack of early diagnosis, however, leads to unexplored epigenetic landscape in stage I RCC. Herein, we identified novel, methylation-driven, and subtype-specific gene markers of three early-stage RCC subtypes (KIRC, KIRP, KICH) with potential prognostic and diagnostic values, from TCGA database. Strikingly, stage I KIRC displays total hypomethylation, while KICH displays insignificant methylation aberration. Importantly, the identified genes (*ATXN1*, *RFTN1*, *VIM*, *HLA-B*, *TM4SF19*, *TNFAIP6*, *GRAMD1B*) correlate with immune infiltration and altered metabolism as the main hallmarks of RCC, which may contribute to better understanding of early-stage RCC.

© The Author(s) 2021. Published by Higher Education Press

INTRODUCTION

Renal cell carcinoma (RCC) is a disease where malignant cells formed in the renal tubules and responsible for more than 90% of kidney cancer. There are more than 140,000 annual RCC related-deaths around the world based on the latest data from the World Health Organization. RCC also denotes the sixth and tenth most commonly diagnosed cancer in men and women, respectively [1]. As a "silent" cancer, RCC is often diagnosed at advanced stages, with inherent resistance to several lines of treatment [2]. It has several subtypes, which include clear cell RCC (KIRC, ~75%), papillary RCC (KIRP, 15%–20%), and chromophobe RCC (KICH, ~5%) [3,4]. Different RCC subtypes bear their unique molecular profiles that contribute to the complexity of RCC pathogenesis and thus management. Therefore, it needs to be further characterized based on the subtypes in order to develop a better diagnosis and management of RCC [5].

With our current understanding, the molecular etiology of RCC is still not fully understood. However, familial RCC, which accounts for around 3% of cases, is known to be caused by mutation in von Hippel-Lindau (*VHL*) tumor suppressor gene, which is responsible for the regulation of hypoxia inducible factor-alpha (HIF-α) signaling pathway [6]. Otherwise, most RCC cases are sporadic and associated with epigenetic alteration instead of genetic mutation [2]. Epigenetic alterations, such as DNA methylation, histone modification, or RNA silencing, have been extensively studied in cancer, including RCC [7]. For instance, the chromatin modifying genes, *PBRM1* and *SEDT2*, are mutated in 69.3%, 53%, and 14.9% of KIRC, KIRP, and KICH, respectively [8]. Interestingly, these mutations are associated with methylome remodeling in RCC, which are supported by numerous studies. For instance, Chen *et al.* provides a link between aberrant DNA methylation and protein expression within each RCC subtype [9]. There have also been efforts in utilizing DNA methylation for novel classification or proposing prognostic markers [2,10]. However, most studies did not correlate their results with the histological subtypes as the major determinant for patient management [11]. Not to mention, none of them focused on early-stage RCC [12,13].

In this study, we identified novel methylation-driven gene markers by differential methylation analysis of three RCC subtypes (KIRC, KIRP, and KICH) in their early stage (stage I) based on the American Joint Committee on Cancer (AJCC) classification in Caucasian population using the TCGA database, while elucidating its clinical relevance by analyzing their prognostic and diagnostic values, as well as exploring their protein-protein interaction and pathway enrichment related to each RCC subtype. We focused on stage I RCC considering its clinical manifestations that are unclear to specific subtypes [14]. Meanwhile, Caucasian population was chosen due to the data availability in the TCGA. So far, there have been no significant differences in overall incidence and mortality rates of RCC between populations [15]. This study is also limited to the observation of the direct effect of DNA methylation towards gene expression, meaning that the observed genes were limited to those having an inverse correlation with their methylation pattern.

RESULTS

Differential methylation analysis of stage I RCC tissue samples

In this study, we retrieved clinical data, DNA methylation beta-value data, and selected mRNA expression data of 120 KIRC stage I patients, 118 KIRP stage I patients, and 27 KICH stage I patients. From these patient data, methylation beta-values of 485,577 CpG sites in 64 normal tissue samples, 120 KIRC stage I samples, 118 KIRP stage I samples, and 20 KICH stage I samples were retrieved. Using Fisher's exact test, we obtained 74 DMRs in KIRC stage I samples, in which all of them are hypomethylated (Fig. 1A). Meanwhile, 520 DMRs were obtained in KIRP stage I samples, in which 378 of them are hypomethylated and 142 of them are hypermethylated (Fig. 1B). The complete list of the differentially methylated CpGs is available in the Supplementary Materials. Surprisingly, we did not obtain any DMRs between KICH Stage I and normal samples. We noticed some overlapping CpGs that are hypomethylated in both KIRC and KIRP, in which *ATXN1* becomes the most significant. Therefore, we conducted further analysis on *ATXN1* expression in both RCC subtypes.

Differential expression analysis of selected DMRs

After obtaining the DMRs, we continued with the differential expression analysis of the selected genes that correspond to the CpG sites. The mRNA expression data were retrieved from 36 normal tissue samples, 120 KIRC stage I samples, and 118 KIRP stage I samples. After sorting, we also compared the change in expression level with the methylation value. We only used the genes that have an inverse relationship with the methylation value of its corresponding CpG sites, meaning an overexpressed gene should be hypomethylated, and vice versa. From the matched genes, we selected the top 5 overexpressed and underexpressed genes to be compared between the diseased patients and normal individuals. For KIRC, from the 74 DMRs that correspond to 39

© The Author(s) 2021. Published by Higher Education Press

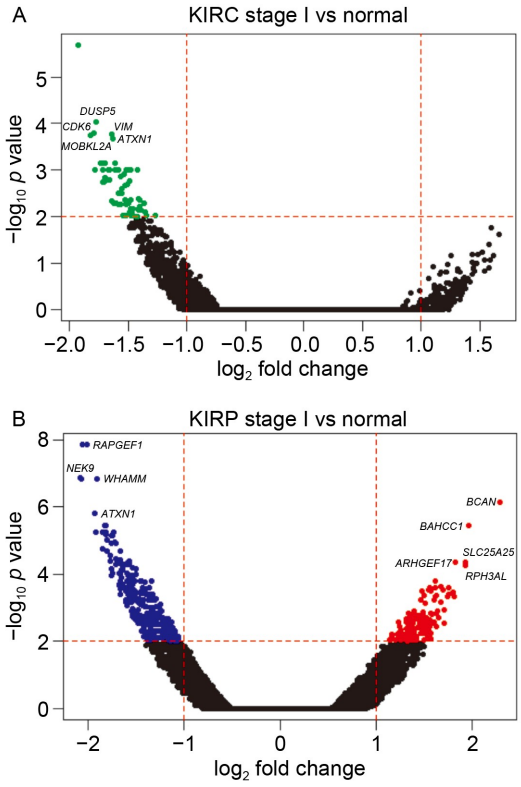

Figure 1. Volcano plots of differentially methylated CpG regions (DMRs) in stage I RCC. (A) 74 hypomethylated CpGs in KIRC with their corresponding genes (top 5) are shown in green dots. (B) 142 hypomethylated and 378 hypermethylated CpGs in KIRP with their corresponding genes (top 5) are shown in blue and red dots, respectively. Black dots indicate CpGs that are not differentially methylated.

genes, 22 differentially expressed genes (DEGs) were obtained, in which only 8 genes were overexpressed, which follows our criteria. Despite there being downregulated genes in KIRC Stage I samples, we did not include such genes since they do not have a negative correlation with their CpG methylation beta-values. We compared the expression level of the top 5 upregulated genes between KIRC Stage I and normal samples (Fig. 2A), with *VIM* as the most notable gene expression changes.

Meanwhile, for KIRP, 192 genes were differentially expressed from 259 genes obtained from the 520 DMRs. Unlike in KIRC, 5 overexpressed and underexpressed genes were able to fulfill our criteria in KIRP. We compared the expression level of those 10 genes between cancer and normal tissue samples (Fig. 2B). *TNFAIP6* has the most notable increase of expression in cancer samples, while *CASR* has the most notable decrease of expression. The selected genes and their corresponding CpG annotations within the genes are shown in Table 1, by searching the Epigenome Wide Association Studies (EWAS) Atlas database using its probe ID.

Correlation tests

From those selected genes, we confirmed the association of each gene expression with their CpG methylation value using Pearson's correlation test (Table 1). The tests showed an inverse correlation in 4 genes for KIRC, with *VIM* having the strongest correlation. Meanwhile, 8 genes in KIRP showed a negative correlation, with *TNFAIP6* having the strongest correlation. Interestingly, there is no significant correlation between *ATXN1* methylation and gene expression in KIRP, while there is a moderately significant correlation in KIRC. This confirms a preliminary validation of DNA methylation-mediated silencing or activation of the selected genes.

Somatic mutation validation

Prior to the investigation of the biomarker capacity of the selected genes, the catalogue of somatic mutations in cancer (COSMIC) database was used to validate the mutation rate and role of the selected genes in cancer progression. The results are summarized in Table 2, which indicates no recorded association between all the selected genes with any cancer types. The results validate the novelty of the selected genes, particularly as a potential marker in RCC. There are still recorded occurrences of somatic mutations according to the database. However, since it does not contribute to cancer development, it might not affect the expression of the genes found in our analysis, thus signifying the direct effect of methylation aberration towards gene expression changes.

Overall survival analysis

In order to select the potential candidate prognostic markers for KIRC and KIRP stage I patients, we conducted OS analysis using the Kaplan-Meier curve while also calculating the hazard ratio and their 95% confidence intervals of each gene. We found that high *ATXN1*, as well as high RFTN1 expressions, are associated with lower overall survival in KIRC (Fig. 3A, B), whereas high *TM4SF19*, as well as high *GRAMD1B* expressions, are associated with lower overall survival in KIRP (Fig. 3D, E). Interestingly, there is no significant association of *VIM* and *TNFAIP6* expressions on the overall survival in KIRC and KIRP, respectively, although they showed the strongest correlation (Fig. 3C, F). Furthermore, one-way ANOVA tests showed that *ATXN1* and *VIM* have the highest upregulation in KIRC compared to KIRP and KICH, while *TNFAIP6* and *TM4SF19* have the highest upregulation in KIRP compared to KIRC and KICH

© The Author(s) 2021. Published by Higher Education Press

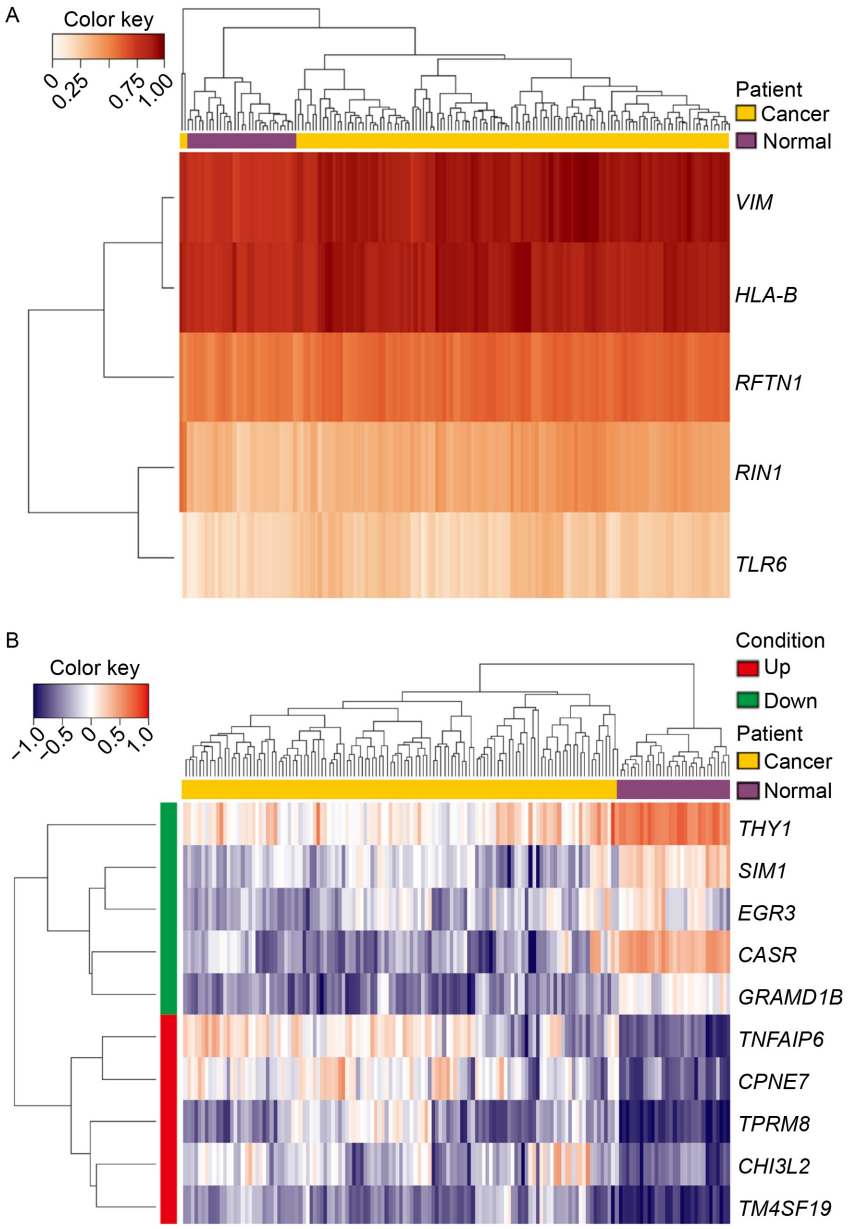

Figure 2. Heatmaps of selected differentially expressed genes in stage I RCC. The genes were selected from the differential methylation analysis results. (A) Top 5 upregulated genes in KIRC were selected. (B) Top 5 upregulated and downregulated genes in KIRP were selected.

(Supplementary Fig. S1). This supports our preliminary selection of RCC-specific prognostic markers.

Receiver operating characteristic (ROC) analysis

The diagnostic values of the selected genes were also investigated using ROC analysis. We analyzed all selected genes with significant methylation-expression correlation from each subtype. The area-under-curve (AUC) cut-off value of 0.8 was used as a threshold for a potential diagnostic biomarker. In addition, we performed ROC analysis for *ATXN1* in KIRC since it shows potential prognostic value. For KIRC, only *RFTN1* and *VIM* expressions may be used to differentiate between KIRC and healthy individuals (Fig. 3G), as well as between other RCC subtypes (Fig. 3I), whereas *ATXN1* has the least ability to distinguish between healthy and KIRC patients (AUC = 0.78). For KIRP, *TM4SF19* and *TNFAIP6* expressions may be able to differentiate between KIRP and healthy individuals (Fig. 3H). However, when compared to other RCC subtypes, their AUC values drop below the cutoff value, highlighting its low accuracy in distinguishing between RCC subtypes (Fig. 3J).

© The Author(s) 2021. Published by Higher Education Press

Table 1 Selected CpG-gene pairs and their correlation test results

Sub-type	CpG	Gene symbol	Annota-tion	*p*-value	r coefficient
KIRC	cg26063719	*VIM*	Body	<0.0001	–0.58
	cg25954539	*HLA-B*	Body	<0.0001	–0.51
	cg08995609	*RIN1*	Body	0.0008	–0.30
	cg14665413	*TLR6*	Promoter	0.08	–0.15
	cg10502118	*RFTN1*	N/A	<0.0001	–0.56
	cg22274117	*ATXN1*	Body	0.002	–0.26
KIRP	cg01942558	*TNFAIP6*	Promoter	<0.0001	–0.75
	cg10225951	*CPNE7*	Body	<0.0001	–0.55
	cg02952295	*TRPM8*	Body	<0.0001	–0.38
	cg21090033	*TM4SF19*	Body	0.028	–0.20
	cg26366091	*CHI3L2*	Body	<0.0001	–0.61
	cg14394550	*EGR3*	Body	0.07	–0.16
	cg02955631	*THY1*	Body	<0.0001	–0.64
	cg06504526	*SIM1*	Body	0.52	–0.10
	cg17637327	*GRAMD1B*	Body	<0.0001	–0.45
	cg07629454	*CASR*	Body	0.004	–0.25

N/A: not available

Protein-protein interaction and functional enrichment analysis

Consequently, we conducted protein-protein interaction studies of *ATXN1*, *HLA-B*, *RFTN1*, and *VIM* for KIRC as well as *TM4SF19*, *GRAMD1B*, and *TNFAIP6* for

Table 2 Cancer-associated somatic mutation validation of selected genes using COSMIC v91 database

Subtype	Gene symbol	Mutation occurrence	Relation to cancer
KIRC	*VIM*	0.983%	Absence
	HLA-B	1.77%	Absence
	RFTN1	3.061%	Absence
	ATXN1	5.917%	Absence
KIRP	*TNFAIP6*	0.89%	Absence
	TM4SF19	0.575%	Absence
	GRAMD1B	4.39%	Absence

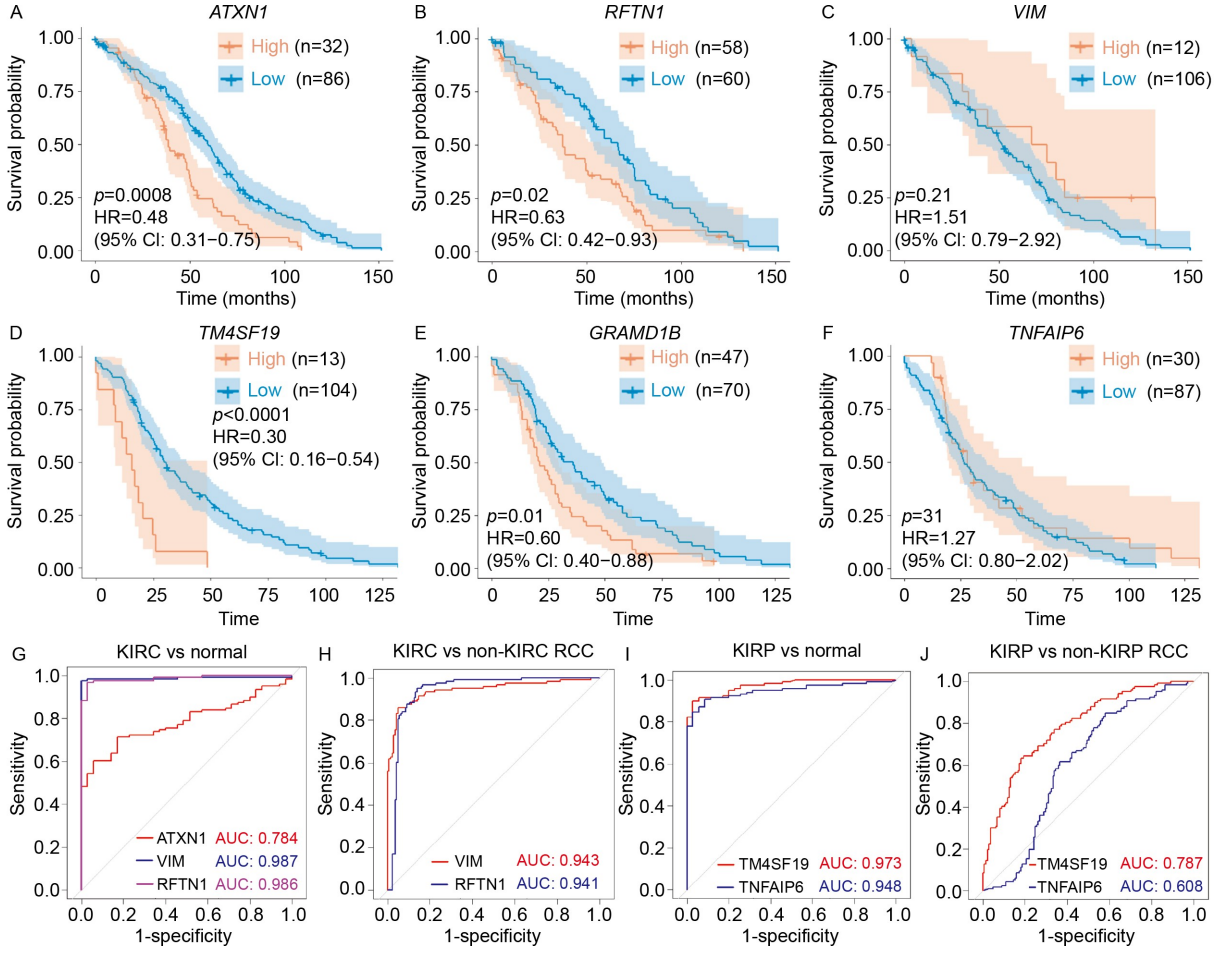

Figure 3. Biomarker analysis of selected genes in stage I RCC. (A–F) Kaplan-Meier OS analysis of selected genes in KIRC (A–C) and KIRP (D–F). The *x*-axis is the time (in months) whereas the *y*-axis is the survival probability obtained from the vital status data. The HR values are for the low gene expression groups. (G –J) ROC analysis of selected genes in KIRC (G–H) and KIRP (I–J). AUC value of 0.8 was chosen as the cut-off value.

© The Author(s) 2021. Published by Higher Education Press

KIRP by using the STRING web tool. The interactome networks were created independently in different RCC subtypes. The multiple protein analysis *of ATXN1*, *HLA-B*, *RFTN1*, and *VIM* results in a total of 34 nodes and 173 interactions (Fig. 4A). We found that *RFTN1* nodes are not connected to the general network and considered as a singleton. Therefore, we constructed a single network analysis for *RFTN1* which results in a total of 24 nodes and 49 interactions (Fig. 4B). Functional enrichment analysis of these significant nodule interactions were then performed. We showed that these interactions are highly associated with immune response, protein expression, and machinery, cell apoptosis, autophagy, organic acid metabolism, pathways involved in specific diseases, graft-versus-host disease, and influenza (Fig. 5A).

Meanwhile, for KIRP, the multiple protein analysis of *TM4SF19*, *GRAMD1B*, and *TNFAIP6* results in a total of 33 nodes and 465 interactions (Fig. 4C), with all genes forming singletons. Therefore, we performed single protein network analysis for each other singletons. *TM4SF19* interaction network results in a total of 27 nodes and 55 interactions (Fig. 4D), while *GRAMD1B* interaction network results in a total of 20 nodes and 32 interactions (Fig. 4E). Functional enrichment analysis of the significant interactions revealed that immune signaling, inflammation, metabolism-related pathways, and metabolite transport are enriched (Fig. 5B). In general, immune-associated and metabolic pathways are enriched in both KIRC and KIRP. The term names for each GO process or KEGG pathway are provided in Supplementary Tables S1 and S2.

DISCUSSION

DNA methylation role at early stage of different RCC subtype

Genome-wide DNA methylation profiling using Illumina Human Methylation 450k data from the TCGA database revealed distinct DMRs in stage I KIRC and KIRP, but not KICH. This is consistent with a previous study reporting no hypermethylated CpGs in KICH, suggesting the minimal influence of methylation towards its progression [9,16]. Interestingly, all DMRs in stage I KIRC are hypomethylated, suggesting that early-stage KIRC may favor gene activation rather than silencing. This is in accordance with a study reporting that KIRC with larger tumor sizes and higher grades are associated with hypermethylation [17]. Differential expression analysis and Pearson's correlation tests showed some genes with positive correlation between their

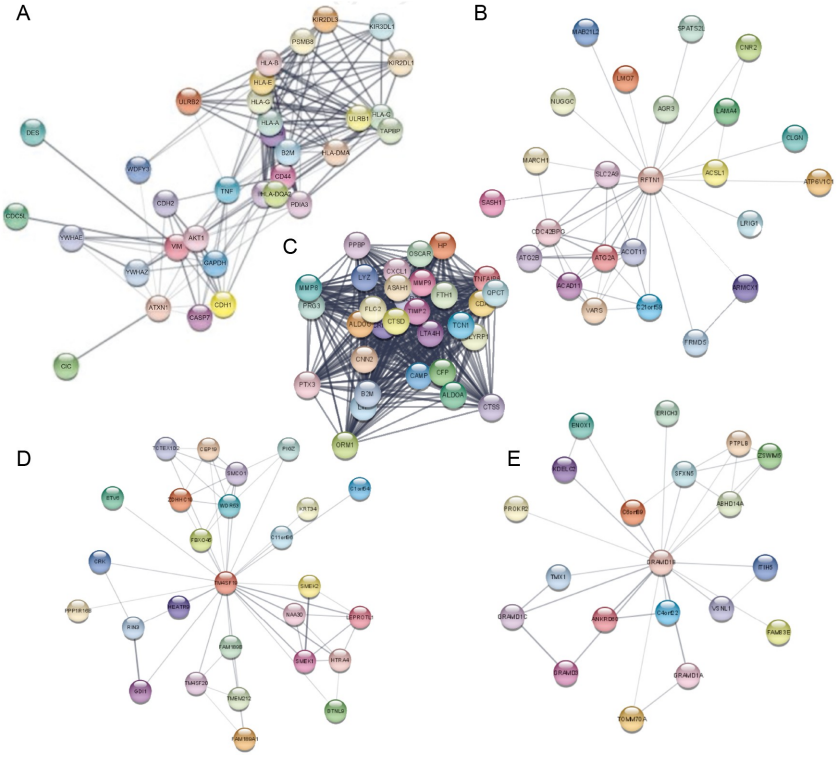

Figure 4. Protein-protein interactions of selected gene markers in stage I RCC. Single protein networks of singletons were constructed. (A) Protein networks of *ATXN1*, *VIM*, and *HLA-B* in KIRC. (B) Single protein network of *RFTN1* in KIRC. (C) Single protein network of *TNFAIP6* in KIRP. (D) Single protein network of *TM4SF19* in KIRP. (E) Single protein network of *GRAMD1B* in KIRP. The significance of the interaction is indicated by the line thickness.

© The Author(s) 2021. Published by Higher Education Press

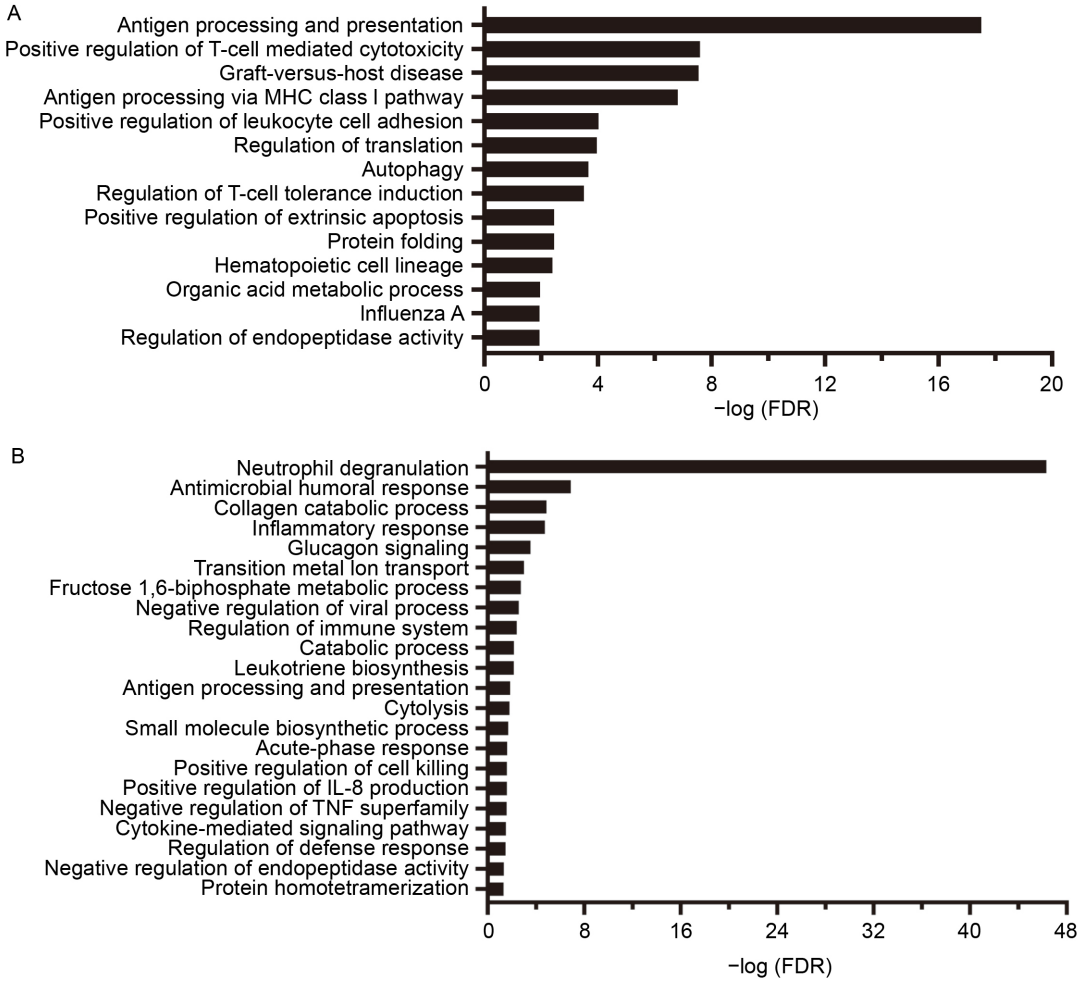

Figure 5. Functional enrichment analysis of the most significant protein interactions in stage I RCC. FDR value of 0.05 was used as the cutoff point. (A) Pathways that are associated with the selected genes in KIRC. (B) Pathways that are associated with the selected genes in KIRP.

methylation and expression, which is contrary to the accepted association. These contradictions are not uncommon, since other factors may also contribute to gene expression changes [18]. Therefore, we only selected genes with negative correlation.

CpG site annotations were obtained from the EWAS Atlas database, which features extensive knowledge regarding the association of epigenomics and biological features [19]. Interestingly, we found that most CpG sites are intragenic, since the majority of CpG islands are located at the transcription starting site (TSS) [20]. Unlike TSS CpG sites that mainly regulate transcription of the corresponding exons, intragenic sites can promote gene transcription of introns and non-corresponding exons, as well as regulate histone methylation, allowing a more delicate regulation [21]. The latter is more probable since intragenic CpG islands between tissues are more diverse than TSS sites [20]. Not to mention, an important validation is that based on COSMIC v91 that uses human genome version 38 (GRCh38), none of the genes have reported cancer-associated mutations, thus supporting our findings in which their expression changes are indeed methylation-driven [22].

OS and ROC analysis showed some genes with both prognostic and diagnostic values, whereas some only serve as either a prognostic or diagnostic marker. This illustrates the ambiguity of the word ‘biomarker’ itself, since not all diagnostic biomarkers can serve as a prognostic biomarker and vice versa [23]. In addition, none of the selected genes have been reported in recent reviews regarding methylation markers in RCC [12,13]. Regardless, subsequent confirmation of these findings is required, such as gene expression and methylation studies using actual RCC cell lines, animal models, or patient specimens.

© The Author(s) 2021. Published by Higher Education Press

Immune response modulation as potential phenotype of early-stage KIRC

Protein-protein interaction studies and functional enrichment analysis showed a strong association with immune response, metabolic pathways, as well as antigen processing in KIRC. These findings are supported by the known hallmarks of KIRC, notably increased immune signature, glycolysis rate, and cytoplasmic lipid droplets [24,25]. ATXN1, HLA-B, and VIM create a single interaction network, while RTXN1 forms a singleton.

ATXN1 (ataxin-1) has a main role in spinocerebellar ataxia type 1 (SCA1), a neurological disorder [26]. Meanwhile, *VIM* encodes for a mesenchymal marker, vimentin, an established immunohistochemical marker of KIRC [27]. The low and indirect interaction between ataxin-1 and vimentin may be attributed to co-expression, thus might explain why *VIM* expression is not associated with OS in KIRC although being highly overexpressed. Their interaction has been confirmed by a former co-immunoprecipitation study [28]. We also highlight the interaction of ataxin-1 and caspase-7 (encoded by *CASP7*), one of the executors of apoptosis, along with other functions, such as cell differentiation [29]. Its activation in renal epithelium is known to be p53-dependent, inferring that its loss-of-function in KIRC is likely to be due to p53 depletion and thus associated with worse prognosis [30–32]. In rat pheochromocytoma, *ATXN1* has also been reportedly upregulated in hypoxic conditions, which is prominent in KIRC [33]. Paradoxically, ataxin-1 is shown to inhibit epithelial-mesenchymal transition (EMT), while in another study, it promotes cervical cancer proliferation [34,35].

RFTN1 (raftlin) is one of the key constituents of lipid rafts, essential for protein trafficking required for cancer progression [36]. It is also involved in the innate immune response and angiogenesis regulation [37,38]. Since lipid rafts play a crucial role in glucose transporter stability, raftlin may be involved in this prominent phenotype of KIRC [39,40]. Moreover, raftlin interacts significantly with CDC42BPG, ATG2A, and ATG2B. *CDC42BPG* encodes for CDC42-binding kinase, which is overexpressed in RCC [41–43]. However, based on the Human Protein Atlas, its overexpression seems to be associated with better prognosis in RCC [44,45]. Both ATG2A and ATG2B are related to autophagy, in which its paradoxical role in cancer is often pronounced [46]. Comprehensive expression analysis records that they are minimally expressed in normal adult tissues, including the kidney [47].

Variation of phenotypes in early-stage KIRP

Despite the three genes (*TNFAIP6, TM4SF19*, *GRAMD1B*) being differentially expressed in KIRP, they did not have any pathway association, thus interaction networks were constructed individually. The neutrophil degranulation pathway is generally enhanced in KIRP, which is consistent with the previous findings [48,49]. This highlights the similar pathway enrichment in both KIRC and KIRP, which is the immune signature.

TNFAIP6 (tumor necrosis factor alpha-induced protein 6) is a secreted protein that responds towards TNF-α signaling and is associated with cell migration. Interestingly, we found that TNFAIP6 interacts significantly with MMP-9, which is closely related to neutrophil degranulation and increased cell invasion. Importantly, overexpression of TNFAIP6 has been associated with worse prognosis in urothelial, colorectal, gastric, and non-small cell lung cancers [50–54]. A study also highlights that downregulation of TNFAIP6 drives EMT in normal renal epithelium, suggesting its possible downregulation in late-stage KIRP[55,56].

TM4SF19 encodes for a transmembrane protein related to immune evasion due to its overexpression by cancer-associated fibroblasts [57]. Although its role in KIRP has not been elucidated, its other family members have been recently pronounced as potential cancer targets [58]. It has also been associated with resistance to afatinib in non-small cell lung cancer [59]. TM4SF19 is found to interact significantly with LEPROTL1, SMEK1, and HTRA4. Serum leptin and adiponectin levels, which are partially regulated by LEPROTL1 (leptin receptor overlapping transcript-like 1), have been positively associated with the risk of RCC [60,61]. SMEK1 (suppressor of MEK1) is involved in various cellular processes, such as DNA repair, apoptosis, microtubule organization, and cell cycle arrest [62]. HTRA4 (high temperature requirement protease A 4), a specific marker for preeclampsia, is reportedly overexpressed in glioblastoma multiforme and breast cancer, as well as associated with the regulation of junctional proteins and microtubules [63–65]. Based on the current understanding of TM4SF19 and its interactions, TM4SF19 may have broad functions in KIRP progression.

GRAMD1B (GRAM protein domain 1b) is involved in lipid and protein signaling, along with many unknown functions. Multiple studies suggest a multifaceted role of GRAMD1B in the JAK/STAT signaling pathway, with negative and positive feedbacks in breast and gastric cancers, respectively [66,67]. Its overexpression is also associated with chemoresistance in ovarian cancer and *TP53* mutation in non-small cell lung cancer [68,69]. We showed that in KIRP, high expression of *GRAMD1B* correlates with lower OS although being downregulated. Such paradoxical results suggest that GRAMD1B may have unique and dynamic roles in KIRP.

© The Author(s) 2021. Published by Higher Education Press

Despite all the results obtained, this study can be subjected to further validations. The proposed conformational wet-lab study is shown in Fig. 6. Moreover, further adjustment towards confounding variables may be needed, such as patient age and comorbidity [70]. However, so far, there is no significant correlation between age and gene expression across RCC subtypes [71]. Not to mention, the patient age obtained in this study is also normally distributed (Supplementary Fig. S2). A recent study by Wong *et al.* suggested that RCC patients with renal stones, chronic kidney diseases, and cerebrovascular disease, lead to a higher risk of mortality, followed by significant changes in immune metagenes [72]. Unfortunately, comorbidities cannot be adjusted since due to lack of data for specific subtype. We have also minimized the effect of confounding variables by limiting the data scope to 'Stage I' and 'Caucasian population'.

CONCLUSIONS

Overall, this study unravels potential DNA methylation-based subtype-specific biomarkers in KIRC and KIRP. In stage I KIRC, *VIM, ATXN1,* and *RFTN1* gene expression are inversely correlated with their methylation values and have prognostic and/or diagnostic values. These genes are associated with immune signature and metabolic pathways which are known to be altered in KIRC. Meanwhile, *GRAMD1B*, *TNFAIP6* and *TM4SF19* gene expression are inversely correlated with their methylation values and have prognostic and/or diagnostic values in stage I KIRP. These genes are responsible for neutrophil degranulation, which correlates with tumor progression and invasiveness, and have also been reported in KIRP. Surprisingly, no significant methylation alterations were identified in KICH. Nevertheless, wet-lab experiments are still needed to further confirm our findings, as a step ahead towards a better molecular characterization and early diagnosis of RCC.

MATERIALS AND METHODS

The pipeline of our study is described in Fig. 7. The methods consist of data retrieval, differential methylation analysis, differential gene expression analysis, somatic mutation validation, correlation tests, OS analysis, ROC analysis, protein-protein interaction studies, and functional enrichment analysis.

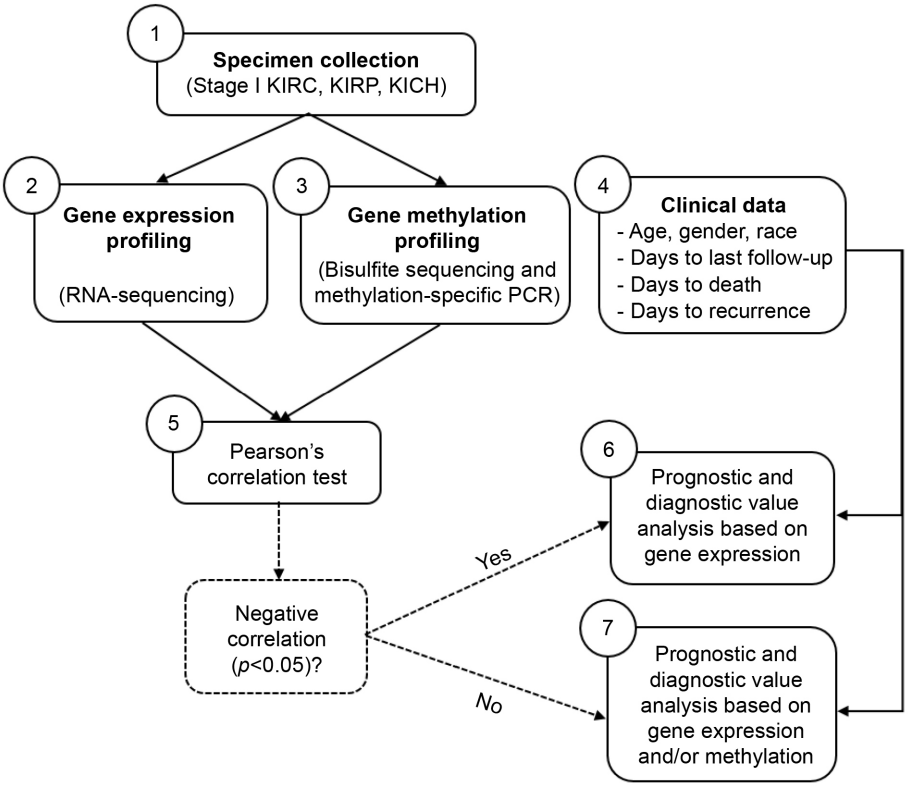

Figure 6. Suggested pipeline for conformational wet-lab study. If the negative correlation is significant, it can be assumed that the gene expression itself may be used as a prognostic or diagnostic marker, whereas if insignificant, both gene expression and methylation will be analyzed for any prognostic or diagnostic value.

© The Author(s) 2021. Published by Higher Education Press

Data retrieval

Clinical data, methylation beta-value data, and mRNA expression data (normalized counts) of KIRC, KIRP, and KICH Stage I patients were retrieved from The Cancer Genome Atlas (TCGA) database by using the TCGA Assembler [73] and TCGAbiolinks packages [74] from April to May 2020. Only patients with complete data (clinical data, methylation beta-value data, and gene expression data) were selected. The patients were filtered based on AJCC pathologic stage, ethnicity, and type of array used. Illumina Human Methylation Array 450k probes were chosen for the methylation beta-value data, whereas normalized counts were chosen for the mRNA expression data.

Differential methylation analysis

Differential methylation analysis was conducted in R 3.6.3 software by using the 'edgeR' package [75]. Normal tissue methylation data from KIRC and KIRP Stage I patients were merged and compared with data from each RCC subtype, since no normal tissue samples were available for KICH. The beta-values were analyzed

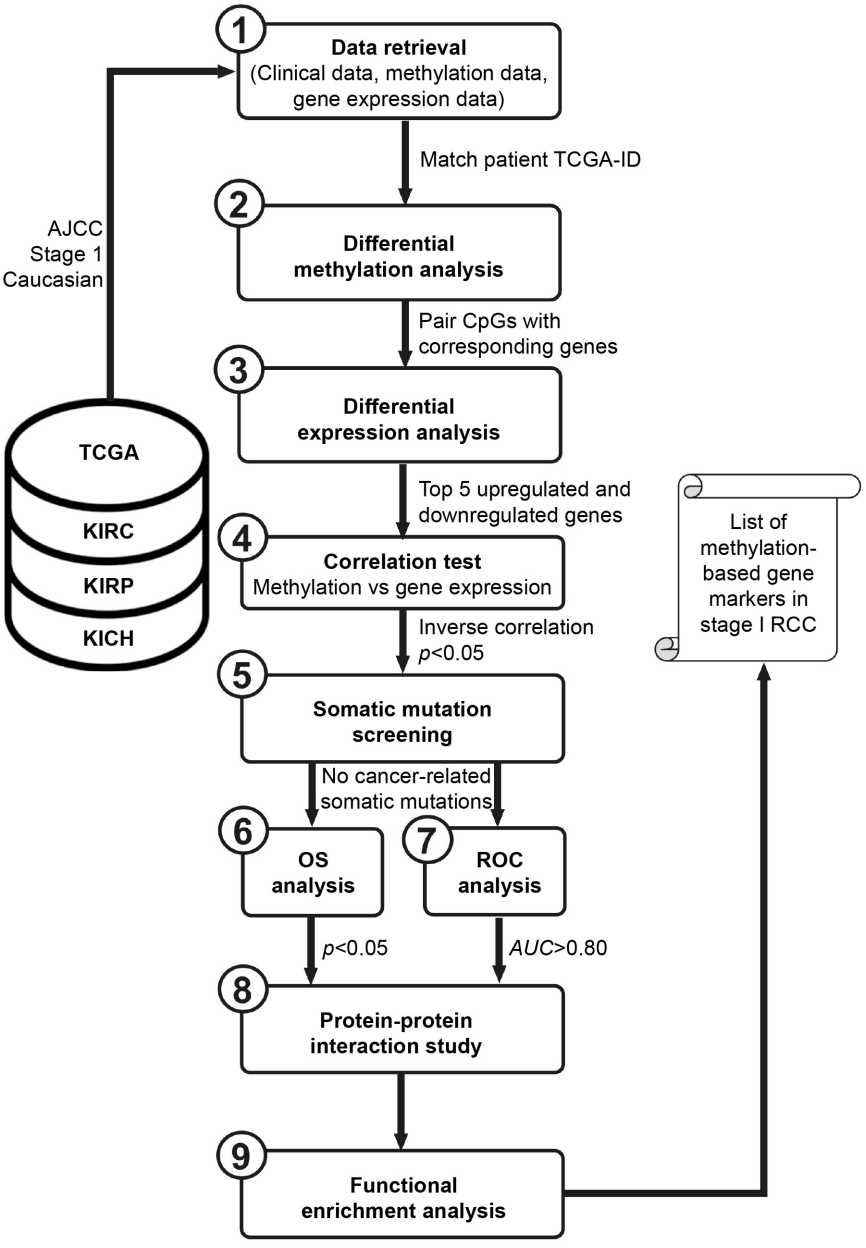

Figure 7. Pipeline of the study. The numbers indicate the step-by-step methods that were conducted.

© The Author(s) 2021. Published by Higher Education Press

using Fisher's exact test to determine the differentially methylated CpGs between the normal and RCC samples. Log_2 fold change of less than −1 or more than 1 and false discovery rate (FDR) of less than 0.01 were used as the cut-off points. Volcano plots of the differentially methylated regions (DMRs) of the CpG sites were constructed in R 3.6.3 software.

Differential gene expression analysis

Differential gene expression analysis was conducted in R 3.6.3 software by using the 'edgeR' package [75]. Normal tissue expression data from KIRC and KIRP stage I patients were merged and compared its expression with data from each RCC subtype. Only selected genes that correspond to the DMRs were subjected to this analysis, by using Fisher's exact test. DEG with the FDR < 0.01, and sorted the list based on the log fold change value (from the highest to lowest) and *p*-value (from the lowest to highest). Heatmaps of the differentially expressed genes were constructed in R 3.6.3 software. The selected CpG-gene pairs were screened for the CpG site annotation using the epigenome wide association studies (EWAS) Atlas by using the CpG probe ID [19].

Correlational tests

Pearson's correlation tests were conducted between methylation beta-values and relative gene expression (in log_{10}) of selected genes after the differential expression analysis in R 3.6.3 software. The CpG-gene data pairs were determined by matching the patient TCGA ID. Inverse correlations with $p < 0.05$ were considered as statistically significant.

Somatic mutation validation

The selected genes were screened for any recorded somatic mutations using the Catalogue of Somatic Mutations in Cancer (COSMIC) database by using the gene symbol [76]. The percentage of reported somatic mutations and its association with cancer progression were retrieved.

Overall survival (OS) analysis

In order to evaluate the prognostic values of the selected genes, OS analysis using Kaplan-Meier curves was conducted in R 3.6.3 software. Patients' clinical data (vital status, time of diagnosis, time of last follow-up or time of death) and mRNA expression data were matched using the patient TCGA ID. The cut-off level for low and high expressions, as well as the 95% hazard ratio (HR) confidence intervals (CI) were determined by using the 'survival' and 'survminer' packages, as demonstrated by previous study [77] (Supplementary Tables S3 and S4). 95% CI of HR of < 1 and $p < 0.05$ were considered as statistically significant.

Receiver operating characteristic (ROC) analysis

ROC analysis was performed in two stages and was conducted in R 3.6.3 software by using the 'pROC' package [78]. Primary ROC analysis was conducted to select genes with diagnostic value in discriminating RCC and normal kidney. The mRNA expression values of each gene and the status (cancer or normal) were plotted into the ROC curve to determine its specificity and sensitivity values. AUC value of 0.8 were used as a threshold for acceptable diagnostic value [79]. The chosen genes were then subjected to the secondary ROC analysis to validate their diagnostic value as subtype-specific marker. The mRNA expression values of each gene and the corresponding RCC subtype (*e.g.*, KIRC or non-KIRC) were plotted into the ROC curve to determine its specificity and sensitivity values. Genes with AUC value of more than 0.8 were selected as potential subtype-specific diagnostic biomarkers.

Protein-protein interaction and functional enrichment analysis

The selected gene markers were subjected for protein-protein interaction analysis by using the Search Tool For The Retrieval Interacting Genes (STRING) v11.0 [80]. Multiple protein analyses were performed first to identify possible interaction between selected genes. If there are disconnected proteins in the network, the genes were subjected to single protein analysis. The network edges are set to show the confidence interval by the thickness of the line. The confidence ranges are low (< 0.4), medium (0.4−0.7), and high confidence (> 0.7). The maximum number of interactors are limited to no more than 30 interactors in the first shell. The interaction map was visualized by using Cytoscape 3.8.0. The most significant nodule interaction is then extracted by using the MCODE application on Cytoscape 3.8.0. The selection is based on the degree cutoff of 10, node score cutoff of 0.2, k-core of 2, and maximum depth of 100. It was followed by functional enrichment analysis for gene ontology (GO) process and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway. The results of the functional enrichment analysis were filtered for redundant terms with 0.5 redundant cutoff. The interaction network images and functional enrichment table were exported as the result of protein-protein interaction analysis, as previously described [81].

© The Author(s) 2021. Published by Higher Education Press

Statistical analysis

Other statistical analyses were conducted in R 3.6.3 software, which include normal Q-Q plots for the RCC patient age distribution and one-way ANOVA of the selected gene expressions among different RCC subtypes and normal tissue samples, followed by Tukey post-hoc test. $p < 0.05$ was considered as statistically significant.

ABBREVIATIONS

AJCC	American Joint Committee on cancer
ATG2A	Autophagy-related 2A
ATG2B	Autophagy-related 2B
ATXN1	Ataxin-1
AUC	Area under the curve
CASR	Calcium sensing receptor
CDC42	Cell division control protein 42 homolog
CDC42BPG	CDC42 binding protein kinase gamma
CHI3L2	Chitinase 3-like 2
CI	Confidence interval
COSMIC	Catalogue of somatic mutations in cancer
CPNE7	Copine 7
CRC	Colorectal cancer
DNA	Deoxyribonucleic acid
DMRs	Differentially methylated regions
DEGs	Differentially expressed genes
ECM	Extracellular matrix
EGR3	Early growth response 3
EWAS	Epigenome wide association study
FDR	False discovery rate
Glut-4	Glucose transporter 4
GO	Gene ontology
GRAM	Glucosyltransferases, rab-like GTPase activators and myotubularins
GRAMD1B	GRAM domain containing 1B
HIF-α	Hypoxia inducible factor-alpha
HLA-B	Human leukocyte antigen B
HR	Hazard ratio
HTRA4	HtrA serine peptidase 4
ICGC	International cancer genome consortium
JAK/STAT	Janus kinase/signal transducers and activators of transcription
KEGG	Kyoto Encyclopedia of Genes and Genomes
KICH	Chromophobe renal cell carcinoma
KIRC	Clear cell renal cell carcinoma
KIRP	Papillary renal cell carcinoma
LEPROTL1	Leptin receptor overlapping transcript like 1
MEK1	Mitogen-activated protein 2 kinase 1
MMP-9	Matrix metalloproteinase 9
mRNA	Messenger ribonucleic acid
OS	Overall survival
PI3K	Phosphoinositide 3-kinase
PTX3	Pentraxin 3
RAC1	Rac family small GTPase 1
RCC	Renal cell carcinoma
RFTN1	Raftlin-1
RIN1	Ras and Rab interactor 1
ROC	Receiver operating characteristic
SCA1	Spinocerebellar ataxia type 1
SIM1	Single-minded homolog 1
SMEK1	Suppressor of MEK1
STRING	Search tool for the retrieval of interacting genes/proteins
TCGA	The Cancer Genome Atlas
TLR6	Toll-like receptor 6
TM4SF19	Transmembrane 4 L six family member 19
TNFAIP6	Tumor necrosis factor-alpha induced protein 6
TP53	Tumor suppressor protein 53
TRPM8	Transient receptor potential cation channel subfamily melastatin member 8
TSS	Transcription start site
VHL	Von Hippel-Lindau
VIM	Vimentin

SUPPLEMENTARY MATERIALS

The supplementary materials can be found online with this article at https://doi.org/10.15302/J-QB-021-0279.

ACKNOWLEDGEMENTS

This research was funded by an internal grant from Indonesia International Institute for Life Sciences (No. 001/SK/WRII-IBSII/I/2020). Unconditional in-kind support was provided by our colleague Stefanus Bernard for helping us with the bioinformatics coding.

COMPLIANCE WITH ETHICS GUIDELINES

The authors Alvaro Filbert Liko, Edward Ciputra, Nathaniel Alvin Sanjaya, Priskila Cherisca Thenaka and David Agustriawan declare no competing financial interests. All procedures performed in studies were in accordance with the ethical standards of the institution or practice at which the studies were conducted, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

OPEN ACCESS

This article is licensed by the CC By under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you

© The Author(s) 2021. Published by Higher Education Press

give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

REFERENCES

1. Capitanio, U., Bensalah, K., Bex, A., Boorjian, S. A., Bray, F., Coleman, J., Gore, J. L., Sun, M., Wood, C. and Russo, P. (2019) Epidemiology of renal cell carcinoma. Eur. Urol., 75, 74−84
2. Lasseigne, B. N. and Brooks, J. D. (2018) The role of DNA methylation in renal cell carcinoma. Mol. Diagn. Ther., 22, 431−442
3. Tabibu, S., Vinod, P. K. and Jawahar, C. V. (2019) Pan-renal cell carcinoma classification and survival prediction from histopathology images using deep learning. Sci. Rep., 9, 10509
4. Shuch, B., Amin, A., Armstrong, A. J., Eble, J. N., Ficarra, V., Lopez-Beltran, A., Martignoni, G., Rini, B. I. and Kutikov, A. (2015) Understanding pathologic variants of renal cell carcinoma: distilling therapeutic opportunities from biologic complexity. Eur. Urol., 67, 85−97
5. Evelönn, E. A., Landfors, M., Haider, Z., Köhn, L., Ljungberg, B., Roos, G. and Degerman, S. (2019) DNA methylation associates with survival in non-metastatic clear cell renal cell carcinoma. BMC Cancer, 19, 65
6. Maher, E. R. (2018) Hereditary renal cell carcinoma syndromes: diagnosis, surveillance and management. World J. Urol., 36, 1891−1898
7. Sharma, S., Kelly, T. K. and Jones, P. A. (2010) Epigenetics in cancer. Carcinogenesis, 31, 27−36
8. Ricketts, C. J., De Cubas, A. A., Fan, H., Smith, C. C., Lang, M., Reznik, E., Bowlby, R., Gibb, E. A., Akbani, R., Beroukhim, R., *et al*. (2018) The Cancer Genome Atlas comprehensive molecular characterization of renal cell carcinoma. Cell Rep., 23, 313−326.e5
9. Chen, F., Zhang, Y., Şenbabaoğlu, Y., Ciriello, G., Yang, L., Reznik, E., Shuch, B., Micevic, G., De Velasco, G., Shinbrot, E., *et al*. (2016) Multilevel genomics-based taxonomy of renal cell carcinoma. Cell Rep., 14, 2476−2489
10. Chen, W., Zhuang, J., Wang, P. P., Jiang, J., Lin, C., Zeng, P., Liang, Y., Zhang, X., Dai, Y. and Diao, H. (2019) DNA methylation-based classification and identification of renal cell carcinoma prognosis-subgroups. Cancer Cell Int., 19, 185
11. Nguyen, D. P., Vertosick, E. A., Corradi, R. B., Vilaseca, A., Benfante, N. E., Touijer, K. A., Sjoberg, D. D. and Russo, P. (2016) Histological subtype of renal cell carcinoma significantly affects survival in the era of partial nephrectomy. Urol. Oncol., 34, 259.e1−259.e8
12. Peters, I., Merseburger, A., Tezval, H., Lafos, M., Tabrizi, P., Mazdak, M., Wolters, M., Kuczyk, M. A., Serth, J. and von Klot, C.-A. J. (2020) The prognostic value of DNA methylation markers in renal cell cancer: A systematic review. Kidney Cancer, 4, 3−13
13. Lommen, K., Vaes, N., Aarts, M. J., van Roermund, J. G., Schouten, L. J., Oosterwijk, E., Melotte, V., Tjan-Heijnen, V. C., van Engeland, M. and Smits, K. M. (2021) Diagnostic DNA methylation biomarkers for renal cell carcinoma: A systematic review. Eur. Urol. Oncol., 4, 215−226
14. Cairns, P. (2010) Renal cell carcinoma. Cancer Biomark., 9, 461−473
15. Lipworth, L., McLaughlin, J. K., Tarone, R. E. and Blot, W. J. (2011) Renal cancer paradox: higher incidence but not higher mortality among African-Americans. Eur. J. Cancer Prev., 20, 331−333
16. Malouf, G. G., Su, X., Zhang, J., Creighton, C. J., Ho, T. H., Lu, Y., Raynal, N. J. M., Karam, J. A., Tamboli, P., Allanick, F., *et al*. (2016) DNA methylation signature reveals cell ontogeny of renal cell carcinomas. Clin. Cancer Res., 22, 6236−6246
17. Evelönn, E. A., Degerman, S., Köhn, L., Landfors, M., Ljungberg, B. and Roos, G. (2016) DNA methylation status defines clinicopathological parameters including survival for patients with clear cell renal cell carcinoma (ccRCC). Tumour Biol., 37, 10219−10228
18. Spainhour, J. C., Lim, H. S., Yi, S. V. and Qiu, P. (2019) Correlation patterns between DNA methylation and gene expression in The Cancer Genome Atlas. Cancer Inform., 18, 1176935119828776
19. Li, M., Zou, D., Li, Z., Gao, R., Sang, J., Zhang, Y., Li, R., Xia, L., Zhang, T., Niu, G., *et al*. (2019) EWAS Atlas: A curated knowledgebase of epigenome-wide association studies. Nucleic Acids Res., 47, D983−D988
20. Jeziorska, D. M., Murray, R. J. S., De Gobbi, M., Gaentzsch, R., Garrick, D., Ayyub, H., Chen, T., Li, E., Telenius, J., Lynch, M., *et al*. (2017) DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. Proc. Natl. Acad. Sci. USA, 114, E7526−E7535
21. Deaton, A. M. and Bird, A. (2011) CpG islands and the regulation of transcription. Genes Dev., 25, 1010−1022
22. Jia, P. and Zhao, Z. (2017) Impacts of somatic mutations on gene expression: an association perspective. Brief Bioinform., 18, 413−425
23. Strimbu, K. and Tavel, J. A. (2010) What are biomarkers? Curr. Opin. HIV AIDS, 5, 463−466
24. Zhang, S., Zhang, E., Long, J., Hu, Z., Peng, J., Liu, L., Tang, F., Li, L., Ouyang, Y. and Zeng, Z. (2019) Immune infiltration in renal cell carcinoma. Cancer Sci., 110, 1564−1572
25. Muglia, V. F. and Prando, A. (2015) Renal cell carcinoma: histological classification and correlation with imaging findings. Radiol. Bras., 48, 166−174
26. Sánchez, I., Balagué, E. and Matilla-Dueñas, A. (2016) Ataxin-1 regulates the cerebellar bioenergetics proteome through the GSK3β-mTOR pathway which is altered in Spinocerebellar ataxia type 1 (SCA1). Hum. Mol. Genet., 25, 4021−4040
27. Williams, A. A., Higgins, J. P., Zhao, H., Ljunberg, B. and Brooks, J. D. (2009) CD 9 and vimentin distinguish clear cell

from chromophobe renal cell carcinoma. BMC Clin. Pathol., 9, 9
28. Chen, H. K., Fernandez-Funez, P., Acevedo, S. F., Lam, Y. C., Kaytor, M. D., Fernandez, M. H., Aitken, A., Skoulakis, E. M. C., Orr, H. T., Botas, J., *et al*. (2003) Interaction of Akt-phosphorylated ataxin-1 with 14-3-3 mediates neurodegeneration in spinocerebellar ataxia type 1. Cell, 113, 457−468
29. Svandova, E., Lesot, H., Vanden Berghe, T., Tucker, A. S., Sharpe, P. T., Vandenabeele, P. and Matalova, E. (2014) Non-apoptotic functions of caspase-7 during osteogenesis. Cell Death Dis., 5, e1366
30. Vilella-Arias, S. A., Rocha, R. M., da Costa, W. H., Zequi, S. C., Guimarães, G. C., Verjovski-Almeida, S., Soares, F. A. and Reis, E. M. (2013) Loss of caspase 7 expression is associated with poor prognosis in renal cell carcinoma clear cell subtype. Urology, 82, 974.e1−974.e7
31. Yang, C., Kaushal, V., Haun, R. S., Seth, R., Shah, S. V. and Kaushal, G. P. (2008) Transcriptional activation of caspase-6 and -7 genes by cisplatin-induced p53 and its functional significance in cisplatin nephrotoxicity. Cell Death Differ., 15, 530−544
32. Kang, J. H., Lee, J. S., Hong, D., Lee, S. H., Kim, N., Lee, W. K., Sung, T. W., Gong, Y. D. and Kim, S. Y. (2016) Renal cell carcinoma escapes death by p53 depletion through transglutaminase 2-chaperoned autophagy. Cell Death Dis., 7, e2163
33. Ke, W., Lu, Z. N., Deng, X. R., Li, W. L., Du, M., Yang, H. and Liu, Y. M. (2017) MicroRNA-18a targets ATXN1 to alleviate injury induced by permanent middle cerebral artery occlusion in mice. Int. J. Clin. Exp. Med., 10, 965−971
34. Kang, A. R., An, H. T., Ko, J. and Kang, S. (2017) Ataxin-1 regulates epithelial-mesenchymal transition of cervical cancer cells. Oncotarget, 8, 18248−18259
35. Kang, A. R., An, H. T., Ko, J., Choi, E. J. and Kang, S. (2017) Ataxin-1 is involved in tumorigenesis of cervical cancer cells via the EGFR-RAS-MAPK signaling pathway. Oncotarget, 8, 94606−94618
36. Staubach, S. and Hanisch, F.-G. (2011) Lipid rafts: signaling and sorting platforms of cells and their roles in cancer. Expert Rev. Proteomics, 8, 263−277
37. Watanabe, A., Tatematsu, M., Saeki, K., Shibata, S., Shime, H., Yoshimura, A., Obuse, C., Seya, T. and Matsumoto, M. (2011) Raftlin is involved in the nucleocapture complex to induce poly(I: C)-mediated TLR3 activation. J. Biol. Chem., 286, 10702−10711
38. Bayliss, A. L., Sundararaman, A., Granet, C. and Mellor, H. (2020) Raftlin is recruited by neuropilin-1 to the activated VEGFR2 complex to control proangiogenic signaling. Angiogenesis, 23, 371−383
39. Wu, G., Wang, Q., Xu, Y., Li, J., Zhang, H., Qi, G. and Xia, Q. (2019) Targeting the transcription factor receptor LXR to treat clear cell renal cell carcinoma: agonist or inverse agonist? Cell Death Dis., 10, 416
40. Yuan, T., Hong, S., Yao, Y. and Liao, K. (2007) Glut-4 is translocated to both caveolae and non-caveolar lipid rafts, but is partially internalized through caveolae in insulin-stimulated adipocytes. Cell Res., 17, 772−782
41. Unbekandt, M. and Olson, M. F. (2014) The actin-myosin regulatory MRCK kinases: regulation, biological functions and associations with human cancer. J. Mol. Med. (Berl.), 92, 217−225
42. Arias-Romero, L. E. and Chernoff, J. (2013) Targeting Cdc42 in cancer. Expert Opin. Ther. Targets, 17, 1263−1273
43. Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., *et al*. (2010) Towards a knowledge-based Human Protein Atlas. Nat. Biotechnol., 28, 1248−1250
44. Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., *et al*. (2017) A pathology atlas of the human cancer transcriptome. Science, 357, eaan2507
45. Human Protein Atlas. Expression of CDC42BPG in cancer. Retrieved in 30 June 2020. Data available from the website of The Human Protein Atlas
46. Velikkakath, A. K., Nishimura, T., Oita, E., Ishihara, N. and Mizushima, N. (2012) Mammalian Atg2 proteins are essential for autophagosome formation and important for regulation of size and distribution of lipid droplets. Mol. Biol. Cell, 23, 896−909
47. Kusama, Y., Sato, K., Kimura, N., Mitamura, J., Ohdaira, H. and Yoshida, K. (2009) Comprehensive analysis of expression pattern and promoter regulation of human autophagy-related genes. Apoptosis, 14, 1165−1175
48. Yu, B., Zhang, J., Sun, Z., Cao, P., Zheng, X., Gao, Z., Cao, H., Zhang, F. and Wang, W. (2021) Interferon-inducible protein 16 may be a biomarker and prognostic factor in renal cell carcinoma by bioinformatics analysis. Medicine (Baltimore), 100, e24257
49. Wen, L., Guo, L., Zhang, W., Li, Y., Jiang, W., Di, X., Ma, J., Feng, L., Zhang, K. and Shou, J. (2019) Cooperation between the inflammation and coagulation systems promotes the survival of circulating tumor cells in renal cell carcinoma patients. Front. Oncol., 9, 504
50. Guo, F. and Yuan, Y. (2020) Tumor necrosis factor alpha-induced proteins in malignant tumors: progress and prospects. OncoTargets Ther., 13, 3303−3318
51. Chan, T. C., Li, C. F., Ke, H. L., Wei, Y. C., Shiue, Y. L., Li, C. C., Yeh, H. C., Lee, H. Y., Huang, S. K., Wu, W. J., *et al*. (2019) High TNFAIP6 level is associated with poor prognosis of urothelial carcinomas. Urol. Oncol., 37, 293.e11−293.e24
52. Offenberg, H., Brünner, N., Mansilla, F., Ørntoft Torben, F. and Birkenkamp-Demtroder, K. (2008) TIMP-1 expression in human colorectal cancer is associated with TGF-B1, LOXL2, INHBA1, TNF-AIP6 and TIMP-2 transcript profiles. Mol. Oncol., 2, 233−240
53. Zhang, X., Xue, J., Yang, H., Zhou, T. and Zu, G. (2021) TNFAIP6 promotes invasion and metastasis of gastric cancer and indicates poor prognosis of patients. Tissue Cell, 68, 101455
54. Shin, S.-B., Jang, H.-R., Xu, R., Won, J.-Y. and Yim, H. (2020) Active PLK1-driven metastasis is amplified by TGF-β signaling that forms a positive feedback loop in non-small cell lung cancer. Oncogene, 39, 767−785
55. Bommaya, G., Meran, S., Krupa, A., Phillips, A. O. and Stead-

 © The Author(s) 2021. Published by Higher Education Press

man, R. (2011) Tumour necrosis factor-stimulated gene (TSG)-6 controls epithelial-mesenchymal transition of proximal tubular epithelial cells. Int. J. Biochem. Cell Biol., 43, 1739−1746

56. Pasquier, J., Abu-Kaoud, N., Al Thani, H. and Rafii, A. (2015) Epithelial to mesenchymal transition in a clinical perspective. J. Oncol., 2015, 792182
57. Peng, Q., Zhao, L., Hou, Y., Sun, Y., Wang, L., Luo, H., Peng, H. and Liu, M. (2013) Biological characteristics and genetic heterogeneity between carcinoma-associated fibroblasts and their paired normal fibroblasts in human breast cancer. PLoS One, 8, e60321
58. Fu, F., Yang, X., Zheng, M., Zhao, Q., Zhang, K., Li, Z., Zhang, H. and Zhang, S. (2020) Role of transmembrane 4 L six family 1 in the development and progression of cancer. Front. Mol. Biosci., 7, 202
59. Chung, C. T., Yeh, K. C., Lee, C. H., Chen, Y. Y., Ho, P. J., Chang, K. Y., Chen, C. H., Lai, Y. K. and Chen, C. T. (2020) Molecular profiling of afatinib-resistant non-small cell lung cancer cells *in vivo* derived from mice. Pharmacol. Res., 161, 105183
60. Scherer, A., Günther, O. P., Balshaw, R. F., Hollander, Z., Wilson-McManus, J., Ng, R., McMaster, W. R., McManus, B. M. and Keown, P. A. (2013) Alteration of human blood cell transcriptome in uremia. BMC Med. Genomics, 28, 23
61. Liao, L. M., Schwartz, K., Pollak, M., Graubard, B. I., Li, Z., Ruterbusch, J., Rothman, N., Davis, F., Wacholder, S., Colt, J., *et al*. (2013) Serum leptin and adiponectin levels and risk of renal cell carcinoma. Obesity (Silver Spring), 21, 1478−1485
62. Yeo, X. Y. (2018) Investigation and validation of the functional role of sMEK1 acetylation in DNA damage repair and HDACi-mediated radiosensitisation. Master's thesis. University of Oxford
63. Wenta, T., Rychlowski, M., Jarzab, M. and Lipinska, B. (2019) HtrA4 protease promotes chemotherapeutic-dependent cancer cell death. Cells, 8, 1112
64. Singh, H., Zhao, M., Chen, Q., Wang, Y., Li, Y., Kaitu'u-Lino, T. J., Tong, S. and Nie, G. (2015) Human HtrA4 expression is restricted to the placenta, is significantly up-regulated in early-onset preeclampsia, and high levels of HtrA4 cause endothelial dysfunction. J. Clin. Endocrinol. Metab., 100, E936−E945
65. Ding, L., Li, L. M., Hu, B., Wang, J. L., Lu, Y. B., Zhang, R. Y., He, X., Shi, C., Wu, L. M., Wu, C. M., *et al*. (2020) TM4SF19 aggravates LPS-induced attenuation of vascular endothelial cell adherens junctions by suppressing VE-cadherin expression. Biochem. Biophys. Res. Commun., 533, 1204−1211
66. Khanna, P., Chua, P. J., Wong, B. S. E., Yin, C., Thike, A. A., Wan, W. K., Tan, P. H. and Baeg, G. H. (2017) GRAM domain-containing protein 1B (GRAMD1B), a novel component of the JAK/STAT signaling pathway, functions in gastric carcinogenesis. Oncotarget, 8, 115370−115383
67. Khanna, P., Lee, J. S., Sereemaspun, A., Lee, H. and Baeg, G. H. (2018) GRAMD1B regulates cell migration in breast cancer cells through JAK/STAT and Akt signaling. Sci. Rep., 8, 9511
68. Wu, S. Y., Yang, X., Gharpure, K. M., Hatakeyama, H., Egli, M., McGuire, M. H., Nagaraja, A. S., Miyake, T. M., Rupaimoole, R., Pecot, C. V., *et al*. (2014) 2′-OMe-phosphorodithioate-modified siRNAs show increased loading into the RISC complex and enhanced anti-tumour activity. Nat. Commun., 5, 3459
69. Lv, T., Lv, H., Fei, J., Xie, Y., Lian, D., Hu, J., Tang, L., Shi, X., Wang, J., Zhang, S., *et al*. (2020) p53-R273H promotes cancer cell migration via upregulation of neuraminidase-1. J. Cancer, 11, 6874−6882
70. Hajian-Tilaki, K. (2013) Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J Intern Med, 4, 627−635
71. Jung, M., Ramankulov, A., Roigas, J., Johannsen, M., Ringsdorf, M., Kristiansen, G. and Jung, K. (2007) In search of suitable reference genes for gene expression studies of human renal cell carcinoma by real-time PCR. BMC Mol. Biol., 8, 47
72. Wong, C. S., Chen, T. T., Chang, W. P., Wong, H. S., Wu, M. Y., Adikusuma, W., Lin, Y. F. and Chang, W. C. (2020) Prognostic effect of comorbid disease and immune gene expression on mortality in kidney cancer-a population based study. Cancers (Basel), 12, 1654
73. Wei, L., Jin, Z., Yang, S., Xu, Y., Zhu, Y. and Ji, Y. (2018) TCGA-assembler 2: software pipeline for retrieval and processing of TCGA/CPTAC data. Bioinformatics, 34, 1615−1617
74. Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T. S., Malta, T. M., Pagnotta, S. M., Castiglioni, I., *et al*. (2016) TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res., 44, e71
75. Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26, 139−140
76. Sondka, Z., Bamford, S., Cole, C. G., Ward, S. A., Dunham, I. and Forbes, S. A. (2018) The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat. Rev. Cancer, 18, 696−705
77. In, J. and Lee, D. K. (2019) Survival analysis: part II − applied clinical data analysis. Korean J. Anesthesiol., 72, 441−457
78. Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C. and Müller, M. (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics, 12, 77
79. Yang, S. and Berdine, G. (2017) The receiver operating characteristic (ROC) curve. Chronicles, 5, 34−36
80. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., *et al*. (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res., 47, D607−D613
81. Zhang, B., Wu, Q., Wang, Z., Xu, R., Hu, X., Sun, Y., Wang, Q., Ju, F., Ren, S., Zhang, C., *et al*. (2019) The promising novel biomarkers and candidate small molecule drugs in kidney renal clear cell carcinoma: Evidence from bioinformatics analysis of high-throughput data. Mol. Genet. Genomic Med., 7, e607

© The Author(s) 2021. Published by Higher Education Press