In Silico Analysis for Sphingolipid Metabolism-related Genes in Human Kidney Clear Cell Carcinoma Using The Cancer Genome Atlas

Woo-Jae Park¹, Jee Young Park², Taeg Kyu Kwon²,³, Jong-Wook Park²,³, Shin Kim¹,²,³

¹Department of Biochemistry, College of Medicine, Gachon University, Incheon, Korea
²Department of Immunology, Keimyung University School of Medicine, Daegu, Korea
³Institute of Medical Science, Keimyung University, Daegu, Korea

The sphingolipid rheostat concept states that the cellular fate is largely determined by various sphingolipid metabolites and the associated signaling pathways. Aberrant regulation of the sphingolipid metabolism-related components is closely associated with cancer survival and death, including aspects like cancer development, proliferation, progression, and response to anticancer drugs. In the present study, we investigated the expression and prognostic significance of the sphingolipid metabolism-related genes in clear cell renal cell carcinoma (ccRCC), the most common pathological subtype of kidney cancer, using an RNA-sequencing dataset of The Cancer Genome Atlas Kidney Clear Cell Carcinoma (TCGA KIRC) cohort. Expression levels of various sphingolipid metabolism-related genes were significantly altered in ccRCC tissues compared with those of normal solid tissues. Notably, the expression of B4GALNT1, BNIP3, DEGS1, GAL3ST1, S1PR4, SLC26A10, SMPDL3A, and SPHK1 was significantly upregulated, whereas the expression of B4GALNT6, HPGD, LPAR1, SFTPB, ST6GALNAC5, and UGT8 was significantly downregulated in ccRCC tissues. Notably, among these significantly-altered sphingolipid metabolism-related genes, the Kaplan-Meier survival analyses showed that high expression levels of B4GALNT1, SLC26A10, and SPHK1 were associated with a poor prognosis of patients with ccRCC, whereas high expression levels of BNIP3, HPGD, and SMPDL3A were associated with a better prognosis. Taken together, our study suggests that B4GALNT1, SLC26A10, SPHK1, BNIP3, HPGD, and SMPDL3A may be novel prognostic biomarkers and targets for a therapeutic strategy to improve the treatment of ccRCC.

Keywords: Clear cell renal cell carcinoma, Sphingolipid metabolism, The Cancer Genome Atlas

Abbreviations: B4GALNT1, beta-1,4-N-acetyl-galactosaminyltransferase 1; B4GALNT6, beta-1,4-galactosyltransferase 6; BNIP3, BCL2 interacting protein 3; ccRCC, clear cell renal cell carcinoma; CERS, Ceramide synthase; CRC, Colorectal cancer; DEGS1, delta 4-desaturase, sphingolipid 1; GAL3ST1, galactose-3-O-sulfotransferase 1; HPGD, 15-hydroxyprostaglandin dehydrogenase; LPAR1, lysophosphatidic acid receptor 1; NST, normal solid tissue; PHK1, sphingosine kinase 1; S1PR4, sphingosine-1-phosphate receptor 4; SFTPB, surfactant protein B; SLC26A10, solute carrier family 26 member 10; SMPDL3A, sphingomyelin phosphodiesterase acid like 3A; ST6GALNAC5, ST6 N-acetyl-galactosaminide alpha-2,6-sialyltransferase 5; TCGA KIRC, The Cancer Genome Atlas Kidney Clear Cell Carcinoma; UGT8, UDP glycosyltransferase 8.
Introduction

Renal cell carcinoma (RCC), derived from the kidney parenchyma accounts for almost 90% of all kidney tumors [1]. According to histology of the RCC subtypes recommended by the World Health Organization, RCC is divided into three main subtypes, including clear cell renal cell carcinoma (ccRCC), papillary renal cell carcinoma, and chromophobe renal cell carcinoma, among which ccRCC is the most common subtype (80%-90%) [2]. Despite an adequate surgical resection, recurrence and metastasis occur in 25% of the patients due to the aggressiveness of ccRCC [3]. To improve this clinical limitation, there has been extensive research to find potential biomarkers that can predict the prognosis for patients with ccRCC using genomic profiling approaches such as microarray analysis, mutation analysis, and gene expression analysis [4-6]. Moreover, recent advances in biotechnology such as RNA sequencing (RNAseq) have revolutionized various fields of research, providing new insight into complex diseases such as cancer with respect to identification of the causes, establishment of targets for therapeutic strategies, and exploration of new prognostic biomarkers.

The sphingolipid rheostat is a biochemical concept that involves regulation of the interconversion of sphingolipid metabolites, including ceramide, sphingosine, and sphingosine-1-phosphate (S1P), and their related signaling pathways, which play an important role in cell fate determination [7]. Accumulating evidences indicate that the dysregulated sphingolipid metabolism-related genes have clinicopathological relevance in various human cancers [8,9]. For example, a high expression of sphingosine kinase 1 (SPHK1) was significantly associated with a worse outcome in survival analysis of patients with breast cancer [10], and a high expression of ceramide synthase (CERS) 5 was significantly associated with a poor prognosis in colorectal cancer (CRC) patients [11]. However, research is scarce on the expression and prognostic potential of various sphingolipid metabolism-related genes in ccRCC. Thus, in the present study, we investigated the mRNA expression levels of the dysregulated sphingolipid metabolism-related genes in ccRCC using gene expression RNAseq data from The Cancer Genome Atlas Kidney Clear Cell Carcinoma (TCGA KIRC) cohort and their prognostic relevance.

Materials and Methods

Gene expression databases and cluster analysis

The gene expression RNAseq dataset (level 3, dataset ID: TCGA.KIRC.sampleMap/HiSeqV2) and clinical characteris-tic datasets (TCGA.KIRC.sampleMap/KIRC_clinicalMatrix and survival/KIRC_survival.txt) for the TCGA KIRC cohort were downloaded from the UCSC Xena database (https://xena.ucsc.edu). The RNAseq dataset provides gene-level transcript estimates with log2(x+1)-transformed counts normalized using the RNA-Seq by Expectation-Maximization (RSEM) normalization package. The RNAseq dataset of TCGA KIRC comprises 606 samples, including 533 primary tumor tissues, one additional new primary tumor tissue, and 72 normal solid tissues (NST). The mRNA expression levels of the sphingolipid metabolism-related genes [12] were sorted from the TCGA KIRC data. We calculated the |log2Fold-Change (FC)| to screen for genes with more than 2-fold differences in expression levels between groups. This study met the publication guidelines provided by TCGA (http://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/ using-tcga/citing-tcga). Cluster analysis was performed using the Cluster 3.0 software [13] to classify the samples into statistically similar groups, and the resulting heatmaps were visualized in TreeView [14] (version 1.6). Expression levels in the heatmaps have been quantile normalized [15], median-centered, and scaled for visualization.

Kaplan-Meier (KM) survival analysis with log rank test

Prior to a survival analysis, we calculated the median gene expression values of the selected sphingolipid metabolism-related genes from TCGA KIRC and determined the values as the cutoff values for survival analysis. Based on the median value of each gene, ccRCC samples of TCGA KIRC was divided into 2 groups. The survival of the two groups was analyzed using the KM method, and statistical significance was assessed using the log-rank test.

Statistical analysis

Statistical analysis was performed with the SPSS software (version 25.0; IBM SPSS, Armonk, NY, USA). The Kolmogorov-Smirnov test was used for testing normality of the data. Differences between groups were then statistically analyzed using the Student’s t-test. The association between inter-individual mRNA levels of the sphingolipid metabolism-related genes was assessed using the Pearson’s correlation coefficients for continuous variables. In general, \( p \ < \ 0.05 \) was considered to denote significance in all statistical analyses performed in the study.
Results

Altered expression levels of the sphingolipid metabolism-related genes in TCGA KIRC cohort.

The heatmap revealed that various sphingolipid metabolism-related genes were dysregulated in ccRCC compared with NST of patients with ccRCC (Fig. 1). The Student’s t-test confirmed that various sphingolipid metabolism-related genes were significantly dysregulated in ccRCC tissues compared with NST ($p < 0.05$; Fig. 2A). Specifically, beta-1,4-N-acetyl-galactosaminyltransferase 1 ($B4GALNT1$), the BCL2 interacting protein 3 ($BNIP3$), delta 4-desaturase, sphingolipid 1 ($DEGS1$), galactose-3-O-sulfotransferase 1 ($GAL3ST1$), the sphingosine-1-phosphate receptor 4 ($S1PR4$),

---

**Fig. 1.** Heatmap presenting the relative mRNA expression levels of the sphingolipid metabolism-related genes in the TCGA KIRC cohort. The data are presented in matrix format in which each row represents an individual gene and each column represents a single tissue. Each cell in the matrix represents the expression level of a gene feature in an individual tissue. The red and green colors in cells reflect relative high and low expression levels, respectively, as indicated in the scale bar. The samples were ordered from NSTs to ccRCC samples according to the standardized expression level of each gene as indicated. TCGA KIRC, The Cancer Genome Atlas kidney clear cell carcinoma; ccRCC, clear cell renal cell carcinoma; NST, normal solid tissue.
Fig. 2. (A) Heatmap showing significantly dysregulated mRNA expression of the sphingolipid metabolism-related genes in ccRCC tissues compared with the NST of TCGA KIRC cohort. The data are presented in matrix format in which each row represents an individual gene and each column represents a single tissue. Each cell in the matrix represents the expression level of a gene feature in an individual tissue. The red and green colors in cells reflect relative high and low expression levels, respectively, as indicated in the scale bar. The samples were ordered from NSTs to ccRCC samples according to the standardized expression level of each gene as indicated. *p<0.05 (NST versus ccRCC). (B) Relative mRNA levels of various sphingolipid metabolism-related genes in TCGA KIRC samples. *Student t-test, p<0.05, \(|\log_{2}\text{FoldChange}| \geq 1.0\). TCGA KIRC, The Cancer Genome Atlas kidney clear cell carcinoma; ccRCC, clear cell renal cell carcinoma; NST, normal solid tissue.
the solute carrier family 26 member 10 (SLC26A10), sphingomyelin phosphodiesterase acid like 3A (SMPDL3A), and SPHK1 were significantly upregulated in ccRCC, whereas beta-1,4-galactosyltransferase 6 (B4GALT6), 15-hydroxyprostaglandin dehydrogenase (HPGD), the lysophosphatidic acid receptor 1 (LPAR1), the surfactant protein B (SFTPB), ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 5 (ST6GALNAC5), and UDP glycosyltransferase 8 (UGT8) were significantly downregulated in ccRCC compared to their respective expression in NST (the Student’s t-test, \( p < 0.05, |\log_{2}FC| \geq 1.0; \text{Fig. 2B} \)). These 14 dysregulated sphingolipid metabolism-related genes may therefore play an important role in the ccRCC pathophysiology.

Inter-individual correlations between sphingolipid metabolism-related genes showing altered mRNA expression levels in ccRCC.

Based on the Pearson’s correlation coefficient, we found 69 significant correlations between the mRNA expression levels of the 14 significantly dysregulated sphingolipid metabolism-related genes in ccRCC sorted from TCGA KIRC cohort (Table 1).

Survival analysis of the significantly dysregulated sphingolipid metabolism-related genes in ccRCC.

Survival data were available for a total of 605 ccRCC patients for the KM survival analysis. The KM plot and log-rank test indicated that higher mRNA expression levels of B4GALNT1, SLC26A10, and SPHK1 were associated with an unfavorable overall survival in ccRCC patients (Fig. 3A-3C; \( p = 0.013, p < 0.001, \) and \( p < 0.001, \) respectively). However, higher mRNA expression levels of BNIP3, HPGD, and SMPDL3A were associated with a favorable overall survival in ccRCC patients (Fig. 3D-3F; \( p = 0.005, p < 0.001, \) and \( p = 0.009, \) respectively).

Discussion

Since the sphingolipid rheostat was proposed as an import-

Table 1. Pearson’s correlation analysis between inter-individual components of sphingolipid metabolism-related genes

<table>
<thead>
<tr>
<th>Samples</th>
<th>Correlations between components</th>
<th>Pearson’s correlation coefficient value</th>
<th>( p )-Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCGA KIRC tissues from TCGA KIRC cohort (n = 533)</td>
<td>B4GALNT1 and B4GALT6</td>
<td>-0.271</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALNT1 and BNIP3</td>
<td>-0.129</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>B4GALNT1 and GAL3ST1</td>
<td>-0.267</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALNT1 and SLC26A10</td>
<td>0.768</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALNT1 and SMPDL3A</td>
<td>-0.244</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALNT1 and SPHK1</td>
<td>0.181</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALNT1 and ST6GALNAC5</td>
<td>0.087</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>B4GALNT1 and UGT8</td>
<td>-0.110</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and DEGS1</td>
<td>-0.141</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and GAL3ST1</td>
<td>-0.127</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and HPGD</td>
<td>0.158</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and S1PR4</td>
<td>-0.232</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and SFTPB</td>
<td>0.352</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and SLC26A10</td>
<td>-0.346</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and SMPDL3A</td>
<td>0.120</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and SPHK1</td>
<td>-0.255</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and ST6GALNAC5</td>
<td>-0.183</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>B4GALT6 and UGT8</td>
<td>0.347</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>BNIP3 and DEGS1</td>
<td>0.123</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>BNIP3 and GAL3ST1</td>
<td>0.410</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>BNIP3 and HPGD</td>
<td>-0.106</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>BNIP3 and LPAR1</td>
<td>-0.201</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>BNIP3 and SFTPB</td>
<td>-0.353</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>BNIP3 and SLC26A10</td>
<td>-0.104</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>BNIP3 and SMPDL3A</td>
<td>0.419</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>BNIP3 and SPHK1</td>
<td>-0.320</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

*Pearson’s correlation coefficient analysis.
In determining the cell fate, many studies have focused on the contribution of sphingolipid metabolism to the mechanism of cancer development, as well as for the identification of new targets for a cancer therapeutic strategy and prognostic factors [16]. However, the expression and potential prognostic markers of the sphingolipid metabolism-related genes remain to be investigated for many cancer types, including ccRCC. In this study, using the RNAseq data of TCGA KIRC cohort, we evaluated the dysregulation of sphingolipid metabolism-related genes in ccRCC and identified potential prognostic markers.

We identified eight sphingolipid metabolism-related genes (B4GALNT1, BNIP3, DEGS1, GAL3ST1, S1PR4, SLC26A10, SMPDL3A, and SPHK1) that were significantly upregulated in ccRCC in TCGA-KIRC cohort, and the KM survival analysis showed that higher expression levels of B4GALNT1, SLC26A10, and SPHK1 were significantly correlated with a worse prognosis. These results were found to be consistent with recent studies showing that worse overall survival was significantly associated with upregulated SPHK1 expression in non-small cell lung cancer [17] and that B4GALNT1 expression was upregulated in ccRCC [18]. Although SLC26A10 expression was found to be significantly upregulated in colorectal carcinoma compared to adenoma, the prognostic significance of SLC26A10 was not evident [19]. SPHK1 is a major enzyme of the S1P metabolism pathway, and plays an important role in tumorigenesis [20], cancer angiogenesis [21], intercellular communication of the tumor environment [22], and formation of the prometastatic environment [23]. B4GALNT1 is a key enzyme in generation of the gangliosides GM2/GD2, and has also been shown to be involved in tumorigenesis [24]. Thus, our results suggest that SPHK1 and B4GALNT1 may play crucial roles as oncogenes in ccRCC.

Similarly, we hypothesized that BNIP3 and SMPDL3A also functioned as oncogenes in ccRCC, and that their higher expression levels may contribute to poor prognoses. Noteworthy, BNIP3 and SMPDL3A were significantly upregulated in ccRCC, but higher expression levels of these genes were sig-

Fig. 3. Survival analyses of the sphingolipid metabolism-related genes in ccRCC samples of TCGA KIRC cohort. Kaplan-Meier plot of overall survival in subjects with high versus low (A) B4GALNT1, (B) SLC26A10, (C) SPHK1, (D) BNIP3, (E) HPGD, and (F) SMPDL3A mRNA expression. TCGA KIRC, The Cancer Genome Atlas kidney clear cell carcinoma; ccRCC, clear cell renal cell carcinoma.
Sphingolipid Metabolism-related Genes have Prognostic Significances in TCGA KIRC

Keimyung Med J

Significantly correlated with better prognoses in TCGA-KIRC cohort. BNIP3 is a well-known tumor suppressor gene that regulates cell death and mitophagy [25], and its higher expression level has previously been correlated with a good prognosis [26], providing further support that BNIP3 plays an important role as a tumor suppressor gene in the ccRCC pathophysiology. However, further investigation is needed to clarify the reason for the increase in BNIP3 expression levels in ccRCC tissues compared with those in NST. SMPDL3A functions as a liver X receptor target gene; however, little is known about its potential roles in cancer. Conversely, the other six dysregulated sphingolipid metabolism-related genes (B4GALT6, HPGD, LPAR1, SFTPB, ST6GALNAC5, and UGT8) were significantly downregulated in ccRCC compared with those in NST. Notably, higher expression levels of HPGD were significantly correlated with a better prognosis in TCGA-KIRC cohort. HPGD is responsible for regulation of prostaglandin metabolism [27] and also acts as a tumor suppressor gene [28]. A decreased level of HPGD was correlated with a poor prognosis in patients with triple-negative breast cancer [29], which was found to be in line with our results, suggesting an important role in the pathophysiology of ccRCC. Moreover, in the present study we provide the first evidence for the significantly-altered regulation of SLC26A10, B4GALT6, and SMPDL3A and their prognostic significance in ccRCC. However, previous studies have demonstrated conflicting results in finding a poor prognosis associated with higher expression of BNIP3 in patients with non-small cell lung cancer [30] and uveal melanoma [31], and higher expression of HPGD in patients with breast cancer [32]. Therefore, more detailed analysis will be needed to determine the regulatory mechanisms of sphingolipid metabolism-related genes and their association with the pathophysiology of ccRCC.

Significant correlations between the sphingolipid metabolism-related genes have been reported in various CRC cohorts, including CERS2 and CERS4, CERS4 and CERS5, and CERS5 and CERS6 [33]. Inter-individual correlations between the significantly dysregulated sphingolipid metabolism-related genes of ccRCC revealed 69 significant correlations; however, no genes were correlated with these six CERS genes as none of these genes showed significantly-altered expression levels in ccRCC compared with those in NST. Therefore, cancer signaling pathways altered in ccRCC may be related to the correlation between the sphingolipid metabolism-related genes other than the CERS genes.

In conclusion, we identified the significantly-altered sphingolipid metabolism-related genes and their prognostic significance in patients with ccRCC using TCGA-KIRC cohort. Notably, our study suggests that B4GALNT1, SLC26A10, SPHK1, BNIP3, HPGD, and SMPDL3A may be potential prognostic biomarkers and promising therapeutic candidates for ccRCC.

Acknowledgements

We would like to thank all the members of our research group for their enthusiastic participation in this study.

Funding

The present study was supported by the National Research Foundation of Korea (grant funded by the Korea Government Ministry of Science, ICT and Future Planning; grant nos. 2014R1A5A2010008, NRF-2016R1D1A1B04930619, and NRF-2019R111A3A01063114).

Availability of data and materials

The datasets analyzed during the present study are available from The Cancer Genome Atlas (https://www.cancer.gov/tcga) and the UCSC Xena (https://xena.ucsc.edu). The datasets generated in the present study are available from the corresponding author upon reasonable request.

Authors contributions

WJP and SK contributed to the conception and design of the study, analysis of the data, interpretation of results, and the writing of the manuscript. JYP and SK contributed to the acquisition of data and the writing of the manuscript. TKK, JWP, and SK reviewed the manuscript. SK edited the manuscript. All authors read and approved the manuscript and agree to be accountable for all aspects of the research in ensuring that the accuracy and integrity of any part of the work are appropriately investigated and resolved.

Ethics approval

Not applicable.

Patient consent

Not applicable.
Conflict of interest

All authors declare no conflicts-of-interest related to this article.

References

26. Chourasia AH, Tracy K, Frankenerger C, Boland ML, Sharifi...


