SETD2 deficiency promotes the transition from PKD to ccRCC by dysregulation of sphingomyelin metabolism

Yu Zhang¹*, Changwei Liu¹², Tianbao Ye³, Lu Chen⁴, Xiaoqun Yang⁵, Luting Zhou⁵, Huimin An⁴.

1. School of Biomedical Engineering, Shanghai Jiao Tong University, 1954 Huashan Road, Shanghai 200030, China.
2. State Key Laboratory of Oncogenes and Related Genes, Renji-Med X Clinical Stem Cell Research Center, Ren Ji Hospital, School of Medicine and School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China.
3. Department of Cardiology, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China.
4. Department of Urology, Ruijin Hospital, Shanghai Jiao tong University School of Medicine, 197 Ruijin Second Road, Shanghai, 200025, People's Republic of China.
5. Department of Pathology, Ruijin Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai, China.

*Corresponding authors’ e-mail address: yuzhang_sjtu@163.com

Abstract:

Clear cell renal cell carcinoma (ccRCC) is a malignant tumor with dysregulated lipid metabolism, and is highly incidental in polycystic kidney disease (PKD) patients. However, SETD2 is an important tumor suppressor gene in ccRCC generation, yet the underlying mechanisms, especially the effects of lipid metabolism, remain largely unexplored. Here, we revealed extensive and large-scale metabolic reprogramming events in a SETD2-deficient ccRCC mouse model by performing a multi-omics study comprising transcriptomics, proteomics, metabolomics, lipidomics and metabolic mass spectrometry imaging approaches. Our data unveiled dramatic alteration of fatty acid biosynthesis, glycerolipid metabolism, glycerophospholipid metabolism, tricarboxylic acid cycle (TCA), carbohydrate digestion and absorption, protein digestion and absorption and biosynthesis of amino acids, which eventually led to dysregulated sphingomyelin metabolism-related metabolic pathways. Clinically, we discovered that TCA and amino acid metabolism are positively associated with SETD2, while glycolysis, protein catabolic and lipids biosynthesis are negatively associated with SETD2. Mass spectrometry imaging of clinical ccRCC tissues revealed that SETD2 mutation is associated with upregulated sphingomyelin biosynthesis in human ccRCC. Our study provides a comprehensive resource of biological data to support future investigations of SETD2-deficient ccRCC, facilitating the development of metabolically targeted therapeutic modalities.
Introduction

Polycystic kidney disease (PKD) is the most common hereditary cystic renal disorder derived from excessive renal epithelial cell proliferation and is the common cause of end-stage renal disease\(^1\). PKD is usually characterized by the formation of abnormal fluid secretion and extracellular matrix deposition in cystic epithelial cells, which are mostly located in the distal regions of kidneys\(^3\). Renal cell carcinoma (RCC), a parenchymal cancer of the kidney derived from renal tubular epithelial cells, is one of the ten most common cancers world-wide\(^4\). RCC is known to occur more frequently in patients with PKD than in the general population\(^6\). However, little is known about alterations in signaling pathways during the transition from PKD to RCC.

Renal cell carcinoma is estimated to account for 2.2% of all cancer burdens globally, with more than 400,000 new diagnoses globally in 2020\(^9\). Clear cell renal cell carcinoma (ccRCC) is the most common form of RCC, accounting for 75% of all cases and is the predominant form of cancer-associated deaths\(^5\). To understand the potential molecular changes driving ccRCC carcinogenesis, various high-throughput genomic studies have been performed to unveil the discriminating features of ccRCC\(^10\). The deletion of chromosome 3p leads to the inactivation of various tumor suppressor genes (VHL, PBRM1, BAP1 and SETD2), which has been considered to be the earliest driver event in ccRCC\(^11\). The array of tumor suppressor genes, which are mutated or inactivated in tumorigenesis, are reported to be involved in the alterations and perturbations of various metabolic events, including citrate cycle (TCA cycle), fatty acid biosynthesis, glycolysis, amino acid metabolism, carbohydrate absorption and modulation of pathways important for redox balance\(^14\). The typical phenotypic feature of ccRCC is the clear cell morphology, which is associated with excessive accumulation of lipids in the cytoplasm\(^17\). The specific excess accumulation of lipids during the progression from PKD to ccRCC needs to be further explored and clarified.

SET-domain-containing 2 (SETD2) is an epigenetic regulator promoting RNA polymerase II (Pol II)-associated histone methyltransferase, which is vital to distinct gene expression\(^20\) and protein expression regulation\(^22,23\). SETD2 participates in diverse chromatin biological processes including transcriptional regulation\(^24\), DNA damage repair, cross-talk of histone modification, alternating splicing, non-histone targets methylation and embryonic development\(^25\). In addition, SETD2 is one of the most frequently mutated chromatin-regulated genes and has been significantly associated with the progression of various cancers, including ccRCC\(^20\). We have established a mouse model of SETD2-deficiency-induced ccRCC and demonstrated that SETD2-deficiency can promote the progression of ccRCC and stronger lipid accumulation in ccRCC\(^19\). Previous studies have characterized the mechanisms underlying SETD2-deficiency-induced ccRCC from a genetic perspective\(^19\). However, the mechanisms at the proteomics and metabolomics levels remain to be elucidated. Our previous studies also showed that correcting the abnormal lipid metabolism associated with SETD2 deficiency in liver and HCC could be exploited as potential
preventive and therapeutic strategies\textsuperscript{27}. The histone methyltransferase gene SETD2 is an epigenetic modifier with tumor suppressor functionality and metabolism regulation whose deficiency has recently been associated with ccRCC development\textsuperscript{28-33}. It is still critical to expand our understanding of SETD2 deficiency-induced ccRCC pathogenesis to facilitate the development of new therapies. However, little is known from \textit{in vivo} studies about the impact of SETD2 inactivation on metabolic alterations in ccRCC and the consequential pathogenesis of ccRCC development.

Multi-omics analysis contributes to a comprehensive understanding of the pathogenesis of complex diseases such as cancer\textsuperscript{34}, central nervous system disease\textsuperscript{35}, diabetes\textsuperscript{36}, microbiome\textsuperscript{37} and viral infection\textsuperscript{38} and the discovery of potential therapeutic targets. In addition, proteins are the direct performers of biological functions, the recent studies have adopted proteomics and transcriptomics to analyze clinical ccRCC\textsuperscript{12,39}. These studies demonstrated that ccRCC exhibits an extensive metabolic dysregulation including oxidative phosphorylation-related metabolism, glycolysis, TCA cycle, amino acid metabolism and lipid metabolism etc.\textsuperscript{12,39}. As metabolites are the direct executors of metabolomics reaction and metabolic mass spectrometry imaging can identify metabolites \textit{in situ}, metabolic reprogramming in ccRCC needs to be comprehensively studied at the proteomics, transcriptomics, metabolomics and mass spectrometry imaging levels. We also have focused on RNA-sequencing studies in cultured ccRCC cells and/or fresh samples from ccRCC mouse models and revealed that multilevel regulation of Wnt/\(\beta\)-Catenin signaling pathway by SETD2 suppresses the transition from PKD to ccRCC, which provides a potential therapeutic strategy for high-risk patients\textsuperscript{19}. However, there has been no attempt to perform comprehensive analysis of SETD2-deficiency-induced ccRCC in mouse models and clinical ccRCC samples with SETD2-deficiency by incorporating multi-omics techniques.

Identifying SETD2 deficiency-induced metabolic reprogramming in ccRCC not only contributes to a better understanding of SETD2-ccRCC interactions but also help to develop potential strategies for inhibiting the progression of ccRCC by regulating metabolic alterations. To fill this gap, here we integrated an extensive and systematic multi-omics profiling of kidney tissues from SETD2 deficiency-induced ccRCC mouse model and clinical ccRCC samples. Specifically, we performed transcriptomics, proteomics, lipid-metabolomics, metabolomics and metabolic mass spectrometry imaging to analyze ccRCC progression from PKD mouse. We thereby detected the anticipated strong alterations in metabolic pathways as well as previously undiscovered SETD2 biological impacts such as significantly upregulated sphingomyelin biosynthesis and profound dysregulation of amino acid metabolism, which was also confirmed in clinical ccRCC samples. This crucial role of sphingomyelin in SETD2 deficiency-induced ccRCC was confirmed with \textit{in vivo} demonstration that inhibiting sphingomyelin biosynthesis by myriocin which is the inhibitor of \textit{de novo} synthesis of sphingomyelin inhibited ccRCC development and body weight as well as improved survival in SETD2 deficiency-induced ccRCC.

\textbf{Results}
Schematic overview and multi-omics summary of the study.

To obtain a comprehensive molecular understanding of SETD2 deficiency-induced ccRCC, we generated Ksp<sup>Cre</sup><sup>MYC<sup>R26StopFL/+</sup> (KM) and Ksp<sup>Cre</sup><sup>MYC<sup>R26StopFL/+ Setd2<sup>flox/flox</sup> (KMS) mouse models which individually represented polycystic kidney disease (PKD) and clear cell renal cell carcinoma (ccRCC) (Fig. 1A and Extended Data Table 1). The phenotypes of PKD and ccRCC were verified by immunohistochemistry and oil red staining at P140 (postnatal day 140) and then kidney tissues were collected for multi-omics characterization (Fig. 1A and Extended Data Fig. 1A). Our previous work has revealed transcriptomic alterations associated with the Wnt/β-catenin signaling signatures and cellular metabolism in SETD2 deficiency-induced ccRCC<sup>19</sup>. To better understand SETD2-mediated systematic biological information, we further performed proteomics, lipid-metabolomics and metabolomics studies. Our liquid chromatography-mass spectrometry (LC-MS)-based proteomics profiles of laser capture microdissected (LCM) kidney tissues from nine PKD and five ccRCC mice (Extended Data Table 1). Proteomics analysis identified a total of 6581 proteins with an average of 4808 per sample. The data from transcriptomics and proteomics were of high quality as evaluated using standard measures (Fig. 1B-C, and Extended Data Fig. 1B-C). Further, we performed off-target metabolomics and off-target lipid metabolomics and identified 63 and 35 differential metabolites between PKD and ccRCC kidney tissues, respectively. The high quality of our metabolic data was revealed by the evaluation presented in Fig. 1D-1E and Extended Data Fig. 1D-E. Based on these comprehensive multi-omics datasets, we performed integrative pathway enrichment analysis to improve the system-level understanding of the cellular processes and pathways involved in SETD2 deficiency-induced ccRCC.
Figure 1. Schematic overview and multi-omics summary of the study.

A. Schematic representation of the multi-omics analyses of ccRCC, including Ksp<sup>Cre</sup>MYC<sup>R26StopFL/+</sup>(KM), Ksp<sup>Cre</sup>MYC<sup>R26StopFL/+</sup>Setd2<sup>flox/flox</sup> (KMS) mouse model preparation, identification of phenotypes of PKD and ccRCC by IHC and oil red staining, transcriptomics identification, protein identification, metabolites recognition, and function verification.

B. Heat map of the top 102 ranked transcriptomics genes between PKD (n = 5) and
ccRCC (n = 5) samples.

C. Heat map of the top 160 ranked glycolysis/gluconeogenesis, citrate cycle (TCA cycle), biosynthesis of amino acids, glycerolipid metabolism, glycerophospholipid metabolism, biosynthesis of unsaturated fatty acids, sphingolipid metabolism, cholesterol metabolism, protein digestion and absorption, carbohydrate digestion and absorption, pyruvate metabolism, fatty acid biosynthesis and fatty acid metabolism related proteins between PKD (n = 9) and ccRCC (n = 5) samples.

D. Heat map of the top 35 ranked metabolites between PKD (n = 5) and ccRCC (n = 3) samples.

E. Heat map of the top 63 ranked lipid-metabolites between PKD (n = 5) and ccRCC (n = 3) samples.

SETD2 deletion mutation induces dysregulation of multiple metabolic biological processes in ccRCC mice.

To investigate the biological function of SETD2 in ccRCC development, we examined the transcriptional profiles of freshly isolated renal tubular epithelial cells from PKD and ccRCC renal tissues\textsuperscript{19}. In total, expression levels of 1072 genes were identified as significantly different between PKD and ccRCC groups (FDR < 0.01, FC > 1.5) (Fig 2A), with 817 upregulated and 255 downregulated genes in the SETD2-deficient group (Fig 2B).

To further characterize host responses caused by SETD2 deficiency at the protein level, we performed LCM to isolate renal tubular epithelial cells from nine PKD and five ccRCC renal tissues and then conducted an LC-MS/MS proteomics study. For proteomic data analysis, Spearman’s correlation coefficient was calculated for all quality control (QC) runs using the mixture of all samples tested (Extended Data Fig. 1B). The average correlation coefficient of the QC samples was 0.95, demonstrating consistent stability of the mass spectrometry platform (Extended Data Fig. 1B). Principal component analysis (PCA) discriminated PKD from ccRCC mouse renal tubular epithelial cells based on the abundance of detected proteins (Extended Data Fig. 2A). In total, expression levels of 168 proteins were identified as significantly different between PKD and ccRCC groups (FDR < 0.01, FC > 1.5) (Fig. 2A), with 79 upregulated and 89 downregulated proteins in the SETD2 deficiency-induced ccRCC (Fig. 2B).

To identify the biological processes and pathways impacted by SETD2 deficiency, we performed enrichment analysis based on different genes and proteins. Gene Ontology (GO) enrichment analysis revealed significant upregulation of metabolic biological process in ccRCC renal tissues (Extended Data Fig. 2B). Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis indicated upregulation of metabolic pathways (Fig. 2C). Notably, metabolic reprogramming occurs in ccRCC, including tryptophan pathways, fatty acid oxidation and synthesis, glucose metabolism, TCA cycle, glutamine metabolism and arginine reprogramming\textsuperscript{12,32,40-43}. To validate and further explore this proposition, we performed KEGG enrichment analysis based on
different proteins between PKD and ccRCC renal tissues, which revealed significant upregulation of metabolic pathways again, including carbon metabolism, 2-oxocarboxylic acid metabolism, biosynthesis of amino acids, TCA cycle, propanoate metabolism, pyruvate metabolism, oxidative phosphorylation, glycerolipid metabolism and glutathione metabolism etcetera (Fig. 2D and Extended Data Fig. 2C). Proteomic-based analysis revealed multiple alterations in metabolic pathways compared with transcriptomics.

Furthermore, we integrated transcriptomics (RNA-sequencing) and proteomics datasets to discover a deep insight into the deregulation of biological processes in ccRCC caused by SETD2 deficiency. Gene set enrichment analysis (GSEA) pathway enrichment was adopted, we discovered multiple upregulated pathways, including processes of carbohydrate absorption, lipid transport and localization, fatty acid metabolic, glycerolipid metabolic and phosphatidylinositol metabolic, as well as downregulated metabolic pathways, such as the processes of TCA cycle, amino acid metabolism, lipase activity and epithelial cell development (Fig. 2E). Integrative analysis of proteomic and transcriptomic GO datasets adopting a GSEA approach revealed that SETD2 deletion mutation induces dysregulation of multiple metabolic biological processes in ccRCC.

Figure 2. SETD2 deletion mutation induces dysregulation of multiple metabolic biological processes in ccRCC mice.
A. The left volcano plot is part of the analysis performed on the transcriptomics data and show the dysregulation of mRNA between PKD and ccRCC kidney tissues. The right volcano plot is part of the analysis performed on the proteomics data and show the dysregulation of proteins between PKD and ccRCC kidney tissues. Significantly upregulated genes/proteins are colored in red and downregulated genes/proteins are colored in blue (FDR < 0.01, fold change (FC) > 1.5).

B. The proportion of differential genes and proteins between PKD and ccRCC kidney tissues.

C. Top enriched gene sets from Kyoto Encyclopedia of Genes and Genomes (KEGG) for differential transcriptomics genes between in PKD and ccRCC.

D. Significantly altered KEGG pathways between PKD and ccRCC. Enriched KEGG pathways based on differential proteins between in PKD and ccRCC kidney tissues (FDR < 0.01, fold change (FC) > 1.5) are shown.

E. Integrative GO analysis of proteomics and transcriptomics data. GSEA of upregulated and downregulated pathways between in PKD and ccRCC kidney tissues at mRNA (x axis) and protein (y axis) levels based on GO subset in AmiGO database. Normalized enrichment scores (NES) of GO terms are plotted.

SETD2 deficiency promotes alterations in multiple metabolic pathways in ccRCC mice.

Based on the above-mentioned findings that multiple metabolic processes were deregulated by SETD2, we conducted GSEA of KEGG metabolic pathways in PKD and ccRCC mouse renal tissues according to our proteomics data (Fig. 3A). We discovered that abundant upregulated pathways in SETD2-deficient renal tissues were frequently related to fatty acid, sphingolipid, protein, glycerolipid, ketone and lipoprotein metabolism, evidenced by upregulated proteins (FDR < 0.01), such as TECR, SLC16A10, SLC5A1, GPAM and MTND5 (Fig. 3B, Extended Data Fig. 2D-F and Extended Data Fig. 3A-C), whereas downregulated pathways included carbohydrate metabolism, glucose metabolism, citrate cycle, pyruvate metabolism, glutathione metabolism and amino acid metabolism, evidenced by downregulated proteins (FDR < 0.01), such as ALDH2, GOT2, IDH2 and GCLM (Fig. 3B, Extended Data Fig. 2D-F and Extended Data Fig. 3A-C).

To accurately investigate the system-level alterations in the deregulated metabolic pathways, we then re-analyzed our proteomics data. The up-regulation of fatty acid biosynthesis (HACD2, HACD3, TECR, THEM4 and ACOT7), carbohydrate absorption (SIS, OPP4, ACE2, SLC5A1, SLC2A2, SLC6A19, SLC16A10, ATP1A1 and ATP1A2), glycolysis (PGM2, PFKL, PFKM, PFKP, ENO3 and PKM) components that we observed at protein level revealed the potential deregulated bioenergy metabolism (Fig. 3C, Extended Data Fig. 4A-B and Extended Data Fig. 5). Additionally, the downregulation of citrate cycle (CS, CSL, ACO2, IDH1, IDH2, IDH3A, IDH3G, SUCLA2, SUCLG1, FH, MDH1, MDH2, ME1 and PCK1), amino acid metabolism (PSPH, SHMT2, GATM, GAMT, GPT2, GOT1, GOT2, BCAT1 and BCAT2)
components that we observed at protein level revealed the potential deregulated bioenergy metabolism again (Fig. 3C, Extended Data Fig. 4A-B and Extended Data Fig. 5). Both of the deregulated metabolic pathways are associated with the development of ccRCC\textsuperscript{43,44}. Notably, we detected significantly upregulated and downregulated proteins in the deregulated pathways in SETD2-deficient mouse renal tissues after normalization to individual protein levels, such as TECR, LPCAT3, SLC16A10, CEPT1, PFKP, PFKL, ATP1A1, KDSR, GLYCTK, PDHB, IDH3G, IDH2, BCAT2, GOT2, CS and FH. (Fig. 3D and and Extended Data Fig. 3D), suggesting the potential role of deregulated metabolic pathways in the progression of SETD2-deficiency caused ccRCC.
Figure 3. SETD2 deletion mutations promote alterations in multiple metabolic pathways in ccRCC mice.

A. Enriched KEGG pathways based on significantly differential proteins between in KMS and KM renal tissues. (FDR < 0.01 and fold change > 1.5).

B. Heat map and quantitative analysis of altered proteins (FDR < 0.01) in upregulated and downregulated pathways between KMS and KM renal tissues.

C. Reprogrammed metabolic pathways in KMS renal tissues. Protein changes of KMS renal tissues in comparison with KM renal tissues.

D. Relative abundance of representative proteins related with carbohydrate absorption, TCA cycle, amino acid metabolism and glycerolipid metabolism (n = 9 mice in KM group, n = 5 mice in KMS group). a.u., arbitrary unit. To calculate relative abundance, the peak area of each protein was divided by the sum peak areas of all detected proteins in the same sample and the resulting ratio value was used as the relative abundance of each protein. Data are shown as mean ± s.e.m. Two-tailed unpaired Student’s t-test followed by a Benjamini–Hochberg multiple comparison test were used for statistical analysis. Exact FDR values are indicated.

SETD2 deficiency promotes sphingomyelin accumulation in ccRCC mice.

Given that transcriptomics and proteomics integrative analysis clearly revealed the significant connection between metabolic reprogramming and SETD2 deficiency-caused ccRCC, we next conducted metabolomics analysis on mouse renal tubular epithelial cells from both KM and KMS groups. Mouse renal tubular epithelial cells were collected via laser capture microdissection (Extended Data Fig. 1A). Visualization of examined metabolites distinguished between KM and KMS by PCA analysis (Extended Data Fig. 6A-B). To discover pivotal metabolites and metabolic pathways deregulated by SETD2 deficiency, 35 differential metabolites were conducted to KEGG pathway enrichment analysis (Fig 4A and Extended Data Fig.6C). Furthermore, we next re-analyzed the differential metabolites to discover a deep insight into the significantly deregulated metabolites in ccRCC caused by SETD2 deficiency (Fig. 4B and Extended Data Fig. 6C). Of note, a significantly increased level of sphingomyelin was observed, suggesting deregulation of sphingomyelin biosynthesis in SETD2 deficiency-caused ccRCC renal tissues (Fig. 4B).

To confirm potential altered lipid metabolites in SETD2 deficiency-caused ccRCC, we next defined metabolites into seven lipid subclasses adopting the ClassyFire classification system \(^45\) (Fig. 4C). As expected, glycerolipids (85.74%) accounted for the largest proportion of the mouse renal tubules metabolome due to the high endogenous contents of ccRCC lipids ranging from triacylglycerol (TG, 82.54%), diacylglycerol (DG, 3.17%), cholesterol ester (CE, 3.17%), sphingomyelin (SM, 4.76%), glucosylceramide (GleCer, 1.59%), ceramide (Cer, 3.17%) and glycerophosphatidylcholine (GPC, 1.59%) (Fig. 4C). The significantly elevated levels of sphingomyelin in SETD2 deficiency-caused ccRCC mouse renal tissues (Fig. 4D)
suggest that SETD2 deficiency may promote the progression of ccRCC via deregulating sphingomyelin biosynthesis. Moreover, we performed Cox regression to identify associations between SM and SETD2-deficiency-caused ccRCC. The upregulation of TG (22:6 22:4 18:2), TG (49:4) and SM (d40:1) were associated with the development of SETD2 deficiency-caused ccRCC, while SM (d40:1) was the most significant (Fig 4E).

Given the above-mentioned findings that multiple upregulated proteins and metabolites were affected by SETD2, we systematically analyzed the correlation between differential proteins and metabolites (Fig. 4F). On the basis of the observation that the protein levels of TECR, KDSR and LPCAT3 were dramatically increased in KSM, which are key enzymes catalyzing the conversion of acetyl-CoA to sphingomyelin. Based on the metabolic data, we discovered that the palmitic acid which is an intermediate metabolite, sphingomyelin (d18:1/18:0), sphingomyelin (d40:1) and sphingomyelin (d41:1) were significantly increased in KMS group (Fig. 4F). Collectively, our results demonstrate sphingomyelin accumulation in SETD2 deficiency-caused ccRCC mice.
Figure 4. SETD2 deficiency induced sphingomyelin accumulation in ccRCC mice.

A. Altered KEGG metabolic pathways in KMS renal tissues compared with KM renal tissues enriched by significantly altered metabolites (FDR < 0.05).

B. Differential metabolites between KMS and KM renal tissues (FDR < 0.05). Top 9 of the most upregulated metabolites and top 3 of the most downregulated metabolites are shown.

C. Chemical composition of the mouse renal tubules metabolome using ClassyFire categories to classify the metabolite diversity of all annotated metabolites across assays.
D. Relative abundance of representative lipid metabolites between KM (n = 5) and KMS (n = 3). To calculate relative abundance, the peak area of each metabolite was divided by the sum peak areas of all detected metabolites in the same ample and the resulting ratio value was used as the relative abundance of each metabolite. Data are shown as mean ± s.e.m. Two-tailed unpaired Student’s t-test followed by a Benjamini-Hochberg multiple comparison test were used for statistical analysis. Exact FDR values are indicated.

E. Cox regression analysis of lipid metabolites values and significant arm-level lipid metabolites.

F. Schema of sphingomyelin biosynthesis with upregulated proteins in KMS renal tissues (FDR < 0.01 and fold change >1.5, n = 5 mice in KM group and n = 3 mice in KMS group).


SETD2 deletion mutation is associated with upregulated sphingomyelin biosynthesis in human ccRCC.

The Clinical Proteomic Tumor Analysis Consortium (CPTAC) has collected many landmark multiomics datasets. Next, we performed Cox regression to identify associations between differential proteins and SETD2 deficiency-caused ccRCC based on the CPTAC datasets. The upregulation of LPCAT3, KDSR, SLC16A1, CDIPT, DERL1 and PFKP and the downregulation of IDH3G, GOT2, HAAO, GLYCTK and FABP3 were associated with the development of SETD2 deficiency-caused ccRCC (Fig. 5A). The differential expression levels of proteins (PFKP, SLC16A1, KDSR, LPCAT3, CDIPT, HAAO, GOT2, FABP3, IDH3G and GLYCTK) was also shown in Fig. 5B and Extended Data 7. By surveying the CPTAC datasets, we discovered that citrate cycle and amino acid metabolism are positively associated with SETD2, while glycolysis, protein catabolic and lipids biosynthesis are negatively associated with SETD2 (Fig. 5C).

To fully validate the results based on the CPTAC datasets, we detected expression levels of metabolites in clinical ccRCC samples. The illustration of the workflow for mass spectrometry imaging (MS imaging) detection of metabolites in clinical ccRCC samples was shown in Fig. 5D and Extended Data Table 2. We next applied MS imaging method to detect metabolites in normal kidney tissue and ccRCC tissue samples from patients with ccRCC. Comparison of ccRCC tissues and normal kidney tissues showed significantly different sphingomyelin levels (Fig. 5D-F). Here, we demonstrated the representative MALDI-IMS images of significant sphingomyelin levels in ccRCC tissues derived from three patients including m/z 703.5741, m/z 731.605, m/z 815.697, m/z 675.5424, m/z 704.5775 and m/z 813.6827 which are SM (d18:0/16:1(9Z)), SM (d18:0/18:1(11Z)), SM (d18:0/24:1(15Z)), SM (d18:1/14:0), SM (d18:1/16:0) and SM (d18:1/24:1(15Z)) in Fig. 5D, m/z 791.5757 and m/z 795.6071 which are SM (40:7;3O) and SM (40:5;3O) in Fig. 5E and m/z 791.5757 and m/z
795.6071 which are SM (40:7;3O) and SM (40:6;3O) in Fig. 5F. We performed statistical analysis on the expression of sphingomyelin in the detected clinical tissues and found that the expression number of SM in the ccRCC tissues was significantly higher than that in the normal kidney tissues. Collectively, our results demonstrated that SETD2 deletion mutation is associate with upregulated sphingomyelin biosynthesis in human ccRCC.
Figure 5. SETD2 deletion mutation is associated with upregulated sphingomyelin biosynthesis in human ccRCC.

A. Cox regression analysis of glycolysis, citrate cycle, amino acid metabolism, protein catabolic and sphingomyelin biosynthesis values and significant arm-level metabolites. These data are based on clinical proteomic tumor analysis consortium (CPTAC).

B. Relative abundance of representative proteins related to glycolysis, citrate cycle, amino acid metabolism, protein catabolic and sphingomyelin biosynthesis. These data are based on clinical proteomic tumor analysis consortium (CPTAC).

C. The correlation of key protein expression abundances of glycolysis, citrate cycle, amino acid metabolism, protein catabolic and lipids biosynthesis and SETD2 mutant. These data are based on clinical proteomic tumor analysis consortium (CPTAC).

D. The illustration of the workflow for mass spectrometry imaging detection of clinical ccRCC samples.

E. Representative MALDI-IMS image of ccRCC tissue from the first patient. MALDI-IMS images are shown for m/z 703.5741, m/z 731.605, m/z 815.697, m/z 675.5424, m/z 704.5775 and m/z 813.6827 which are SM (d18:0/16:1(9Z)), SM (d18:0/18:1(11Z)), SM (d18:0/24:1(15Z)), SM (d18:1/14:0), SM (d18:1/16:0) and SM (d18:1/24:1(15Z)). The left side of the tissues is normal kidney tissue and the right side is ccRCC tissue. SM, sphingomyelin. Scale bar = 4mm.

F. Representative MALDI-IMS image of ccRCC tissue from the second patient. MALDI-IMS images are shown for m/z 791.5757 and m/z 795.6071 which are SM (40:7;3O) and SM (40:5;3O). Normal kidney tissue on the left and ccRCC tissue on the right. SM, sphingomyelin. Scale bar = 3mm.

G. Representative MALDI-IMS image of ccRCC tissue from the third patient. MALDI-IMS images are shown for m/z 791.5757 and m/z 795.6071 which are SM (40:7;3O) and SM (40:6;3O). Normal kidney tissue on the left and ccRCC tissue on the right. SM, sphingomyelin. Scale bar = 4mm.

Discussion:

Renal cell carcinoma is estimated to account for 2.2% of all cancer burdens globally, with more than 400,000 new diagnoses in 2020. ccRCC is the most common form of RCC, which is usually manifested as synchronous metastatic disease associated with poor prognosis. Although early detection of ccRCC can be successfully treated with surgery or ablation strategies, up to one-third of patients will still present or develop metastases. ccRCC is known to occur more frequently in patients with PKD than in the general population.
SETD2 deficiency-caused ccRCC has emerged as the majority of kidney cancer deaths\textsuperscript{20,23,26}; however, the mechanisms underlying SETD2 deficiency-caused ccRCC are largely unknown, highlighting an urgent need to expand our knowledge of its pathogenesis. This study systematically investigated SETD2 deficiency-caused ccRCC \textit{in vivo} based on comprehensive analysis of multi-omics. We conducted proteomics, lipidomics and metabolomics datasets, meanwhile adopted transcriptomics data sourced from our published work\textsuperscript{19} that was performed under exactly the same conditions as those of the current work. Our multi-omics analysis revealed various biological progression alters in mice with SETD2-deficient-caused ccRCC, including various metabolic reprogramming.

We and others have revealed SETD2 deficiency-caused metabolic reprogramming\textsuperscript{27,32,46,47,48} and our current study identified several metabolic pathway changes in SETD2 deficiency-caused ccRCC \textit{via} transcriptomics and proteomics, emphasizing the utility of multi-omics dataset. Tumor suppressors-deficiency are known to alter metabolic pathways to promote the proliferation of tumor cells and tumor metastasis\textsuperscript{15,49-52}. Furthermore, changes in cell metabolism caused by tumor suppressors-deficiency are mostly glucose transport, glycolysis, amino acid metabolism, tricarboxylic acid cycle, fatty acid biosynthesis, etc. which are related to nutrient formation, biomass assimilation and redox control\textsuperscript{53-57}.

Notably, our comprehensive analysis of proteomics and transcriptomics datasets demonstrated significant changes of metabolic pathways, including dramatically upregulated carbohydrate absorption, lipid transport and localization, fatty acid metabolism, glycerolipid metabolism and phosphatidylinositol metabolism and significantly downregulated TCA cycle and amino acid metabolism, indicating the changes in cell metabolism can support cancer initiation and tumor progression in ccRCC. A recent study demonstrated that SETD2 deficiency can lead to abnormal lipid metabolism in the liver and promote the initiation of HCC\textsuperscript{27}. Thus, our data suggest the possibility that SETD2 deficiency may cause deregulation of metabolic pathways in renal tissue and promote the progression of ccRCC. Deregulated fatty acid metabolism has been known to drive tumorigenesis and tumor progression\textsuperscript{58-62}. Of note, in our study, we identified dramatically upregulated components in fatty acid metabolism, including HACD2, HACD3, TECR, THEM4 and ACOT7 upon SETD2 deficiency.

Another Recent study has demonstrated that the deregulation of carbohydrate absorption, glycolysis, lipid biosynthesis, sphingolipid metabolism, TCA cycle and amino acid metabolism are needed to support tumor growth, progression and metastasis\textsuperscript{53,63-72}. Notably, in our study, we also identified dramatic deregulation of sphingolipid metabolism, carbohydrate absorption, glycolysis, lipid biosynthesis, TCA cycle and amino acid metabolism in SETD2 deficiency-caused ccRCC. Moreover, we also investigated the associations between differential proteins and SETD2 deficiency-caused ccRCC based on the CPTAC datasets and discovered significant deregulation.
of glycolysis, lipid biosynthesis, TCA cycle, amino acid metabolism and protein catabolism.

Furthermore, comprehensive pathways and metabolites analysis of metabolomics and lipidomics datasets revealed significant changes of lipid metabolism and chemical composition of lipid metabolites. Of note, in our study, we revealed dramatic sphingomyelin accumulation in SETD2 deficiency-caused ccRCC. Moreover, our MALDI-IMS image data of clinical ccRCC samples also confirmed that SETD2-deficient tumor tissues had a significant accumulation of sphingomyelin compared with normal tissues. These findings suggest that deregulation of sphingomyelin biosynthesis may be a critical incentive of ccRCC caused by SETD2 deficiency.

Collectively, we demonstrated that SETD2 deficiency promotes ccRCC formation by deregulating sphingomyelin biosynthesis. Several ongoing clinical trials are evaluating fatty acid and lipid-blocking therapies. Our data suggest that cancer patients with SETD2 deficiency, particularly polycystic kidney disease (PKD) patients that have potential to progress to clear cell renal carcinoma (ccRCC), should be considered for such therapy.

Acknowledgments
We thank Dr. Yongwei Wang and MS. Susu Xia at Bruker for providing mass spectrometry imaging detection and data analysis. This study was supported by funds from National Natural Science Foundation of China (81701353 to Yu Zhang). We thank Prof. Li Li and Dr. Hanyu Rao at Shanghai Jiao Tong University for providing mouse model and biology guide. We thank Prof. Xianting Ding at Shanghai Jiao Tong University for providing help.

Author contributions
Conceptualization was carried out by Y.Z. Experimental design was conducted by Y.Z. Software, formal analysis and visualization was conducted by Y.Z. Methodology and experimental investigation was conducted by Y.Z., C.L., T.Y., L.C., X.Y. H.A. and L.Z. C.L. assisted in mouse genotyping. T.Y. assisted in ultrasonic testing in mice. Y.Z. designed and performed proteomics, metabolomics, lipid metabolomics, mass spectrometry imaging, analyzed multi-omics data and interpreted results. Y.Z. was responsible for writing the draft manuscript. Y.Z. provided conceptualization, resources, methodology, supervising, validation, project administration, writing-review and editing. All authors reviewed and approved the manuscript.

Competing interests
None.

Additional information
None.
Methods:

1. Mouse strains
Setd2\textsuperscript{fl/fl} mice were purchased from Shanghai Biomodel Organism Co. which adopting conventional homologous recombination in embryonic stem (ES) cells. The Ksp\textsuperscript{Cre} mice (B6.Cg-Tg (Cdh16-cre) 9Igr/J) and MYC\textsuperscript{R26StopFL/p} (C57BL/6N-Gt (ROSA) 26Sor\textsuperscript{tm13(CAG-MYC-CD2)Rsky}) were generated by the Jackson Laboratory. Setd2\textsuperscript{fl/fl} mice were mated with Ksp\textsuperscript{Cre} mice to generate Setd2\textsuperscript{-KO} (Ksp\textsuperscript{Cre}; Setd2\textsuperscript{fl/fl}) mice in C57BL/6 back-ground. MYC\textsuperscript{R26StopFL/p} mice were mated with Ksp\textsuperscript{Cre} mice to generate MYC\textsuperscript{-OE} (Ksp\textsuperscript{Cre}; MYC\textsuperscript{R26StopFL/p}) mice in C57BL/6 back-ground. MYC\textsuperscript{-OE} mice were mated with Setd2\textsuperscript{-KO} mice to generate MYC\textsuperscript{-OE/ Setd2\textsuperscript{-KO}} (Ksp\textsuperscript{Cre}; MYC\textsuperscript{R26StopFL/p}; Setd2\textsuperscript{fl/fl}) mice housing under same condition. MYC\textsuperscript{-OE} (Ksp\textsuperscript{Cre}; MYC\textsuperscript{R26StopFL/p}) mice and MYC\textsuperscript{-OE/ Setd2\textsuperscript{-KO}} (Ksp\textsuperscript{Cre}; MYC\textsuperscript{R26StopFL/p}; Setd2\textsuperscript{fl/fl}) mice were defined as KM and KMS, respectively.

2. Patient samples
All patients included in the study underwent ccRCC surgery at the Ruijin Hospital Affiliated to Shanghai Jiaotong University. All human tissue samples were collected with informed consents and in compliance with the strategy of the Ethics Committees of the Ruijin Hospital, Shanghai Jiaotong University School of Medicine. This study was approved by the Ethics Committees and the clinical trial registry number is KY2018-153.

3. Tissue storage and slide preparation
Frozen mice kidneys and patient renal tissues were stored intact at -80 °C until LCM and MALDI MSI analysis. For mice kidneys, 12-µm and 18-µm thick sections were sectioned with a Cryostat (Thermo Fisher Scientific). 18-µm thick sections for LCM were thaw-mounted on MembraneSlide (1.0 PEN, ZEISS) and 12-µm thick sections for MALDI MSI analysis were thaw-mounted on ITO (indium tin oxide) slides (Bruker Daltonics). For patient renal tissues, 12-µm thick sections for MALDI MSI analysis were sectioned with a Cryostat (Thermo Fisher Scientific) and thaw-mounted on ITO slides (Bruker Daltonics).

4. Laser-capture microdissection (LCM)
Mice kidneys were prepared as previously described. For each tissues, 18-µm thick sections were sectioned with a Cryostat (Thermo Fisher Scientific) and thaw-mounted on MembraneSlide (1.0 PEN, ZEISS). ccRCC and PKD tissues depending on tissue phenotype were sectioned with a ZEISS PALM systems and ccRCC and PKD samples were thaw-mounted on ITO slides (Bruker Daltonics).
collected onto adhesive cap MiCroTube 500 (500 µL tubes). The collected samples were used for proteomics, metabolomics and lipidomics.

5. Protein preparation

The collected renal tissues were lysed by adding 50 µL lysis buffer containing RIPA (Beyotime Biotechnology) and cocktail protease inhibitor (50×, Beyotime Biotechnology). Samples were followed by sonication (SCIENTZ-IID) for 20 cycles of 5 seconds on and 5 seconds off and then boiled at 95°C for 2 minutes. Whole-proteome liquids were spun down at 10000g for 10 minutes at 4°C and collected the supernatant to new tubes. Protein concentration was determined by Pierce BCA Protein Assay Kit (Thermo Fisher Scientific) according to the manufacturer’s protocols. Then 50 µg proteins separated from each samples were transferred to new tubes and waiting for subsequent experiments. The separated proteins were mixed with cold acetone at a 1:4 volume ratio, then immediately precipitated at -20°C for overnight and spun down at 20000g for 10 minutes at 4°C. Then the subsequent was poured and the precipitant proteins were washed twice with cold acetone. Furthermore, proteins were reduced with dithiothreitol (DTT, Acros organics) followed by incubation for 1 hour at 55°C and alkylated with iodoacetamide (IAA, Acros organics) followed by incubation for 30 minutes in the dark at room temperature. Then proteins were transferred into ultrafiltration tubes (10kDa, Sartorius AG) and washed three times with 50 mM \( \text{NH}_4\text{HCO}_3 \) at 14000g for 15 minutes. Pellets were resuspended in 50 mM \( \text{NH}_4\text{HCO}_3 \) (Sigma-Aldrich) and digested 16 hours at 37°C with trypsin (Promega) and stopped by addition formic acid (FA, Thermo Fisher Scientific). Digested peptides were then desalted with macro spin column TARGA C18 (The Nest Group) and eluted with 60% acetonitrile (ACN, Thermo Fisher Scientific) solution containing 0.1% FA. Desalted peptides were determined by nanodrop (Thermo Fisher Scientific), and 5 µg cleaned peptides of each sample were dried in a Speed-Vac (Eppendorf).

6. Generation and processing of LC-MS/MS proteomic data

LC-MS was conducted with an EASY nanoLC 1200 (Thermo Fisher Scientific) coupled online to Q Exactive HF-X spectrometer (Thermo Fisher Scientific) with nano spray flex ion source. Dried peptides were resuspended with 0.1%FA and 500ng peptides were loaded on a HPLC-column (75-µm inner diameter, 2-cm length, 3-µm particle size, 100-Å pore size, Thermo Fisher Scientific) at a flow rate of 2 µL/min. Mobile phase A and B were 0.1%FA and 80% ACN (plus 0.1%FA) respectively at a flow rate of 300nL/min for 120 minutes. The gradient settings for the mobile phase were as follows: 8% mobile phase B duration 1 min, 28% mobile phase B duration 97 min, 36% mobile phase B duration 14 min and 100% mobile phase B duration 8 min. For MS analysis, mixed peptides derived from aliquots of each sample and analyzed on Q Exactive HF-X spectrometer in data-dependent acquisition (DDA) mode and peptides of each samples analyzed on Q Exactive HF-X spectrometer in data-independent acquisition (DIA) mode. Specifically, mixed peptides were conducted in DDA mode for ion library generation. The number of scan ranges for Q-Exactive HF-X was from 350 to 1200 m/z, which resolution was 60,000 and fixed first mass was...
230 m/z. At the same time, the automatic gain control (AGC) target was set as $3 \times 10^6$ and the maximum injection time (IT) was set as 20 ms.

**Raw DDA Data** were processed and analyzed by Spectronaut 15.0 (Biognosys AG, Switzerland) with default settings to identify an initial peptide identification which contained 77220 precursors, 35470 peptides, 6581 proteins and 6417 protein group. All protein database searches were conducted against uniprot-proteome-mus_musculus database (version 201907, 22290 entries), assuming that trypsin is the digestion enzyme. Carbamidomethyl of cysteine residues (C) was specified as the static modification. Oxidation of methionine (M) and Protein N-terminal acetylation (Acetyl) were considered as the variable modifications. The false discovery rate (FDR) cutoff on peptide-spectrum match (PSM), precursor and protein level were set at 1%. The mixed peptides were conducted in DDA mode for generating the proteome libraries.

**Raw DIA Data** were processed and analyzed by Spectronaut 15.0 (Biognosys AG, Switzerland) with default settings, and the retention time type was set to dynamic item response theory (iRT). Extracting data were determined adopting Spectronaut 15.0 depending on the extensive mass calibration. The ideal extraction window dynamically was determined using Spectronaut 15.0 based on iRT calibration and gradient stability. The FDR cutoff on precursor and protein level was set as 1%. Decoy generation was set to mutation, which was similar to scrambling, but only a random number of AA position swamps (min=2, max=length/2) will be applied. Normalization strategy was set as global normalization. The average of first three filtered peptides through the 1% FDR cut-off were used to calculate the quantities of major group. After one-way analysis of variance (ANOVA) test, differentially expressed proteins were selected based on p value <0.05 and absolute fold change >1.5.

7. **Metabolite preparation**

200 µL ddH$_2$O was added to the tube pre-added with tissues and homogenized with cooled N$_2$ gas flow from liquid nitrogen for three cycles, each cycle was centrifuged at 5500 r.p.m for 20s and repeat triplicate. Then, 800 µL of MeOH (methanol): ACN (v: v, 1:1) was added to 200 µL of the homogenate, vortexed for 30 s, centrifuged at 5500 r.p.m for 30s and sonicated for 10 min. Proteins will precipitate when incubated at 20 °C for 1h. Then the metabolite-containing sample was centrifuged at 13000 r.p.m. for 15 min at 4°C, and the supernatant containing metabolites was transferred into a new glass tube and evaporated to dryness in a speed vacuum concentrator. Dried metabolite pellets were then resuspended in 100 µL of ACN: H$_2$O (v: v,1:1) and sonicated for 10 min. The resuspended sample was spun at 13000 r.p.m. for 15 min at 4°C, and the metabolite-containing supernatant was transferred into a new glass tube and stored at -80°C until analysis.

8. **Generation and processing of LC-MS/MS metabolomics data**
The metabolite-containing supernatant was processed and analyzed by HPLC (High performance liquid chromatography)–MS/MS on the TripleTOF6600plus mass spectrometer (AB SCIEX, USA) combined with Agilent 1290 liquid chromatography (LC) system (Agilent, USA). For liquid chromatography separation, the ACQUITY UPLC BEH C18 column (100 mm × 2.1 mm i.d., 1.7µm; Waters) was employed and 2 µL of non-polar supernatant was injected and separated with a 12-minute gradient. The flow rate was set at 500 µL/min and the column temperature was 40°C. The chromatographic gradient program was shown in Table 1:

Table 1. The chromatographic gradient for metabolomics data analysis

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Flow rate (mL/min)</th>
<th>A (%)</th>
<th>B (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.5</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>7</td>
<td>0.5</td>
<td>35</td>
<td>65</td>
</tr>
<tr>
<td>8</td>
<td>0.5</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>0.5</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>9.1</td>
<td>0.5</td>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>12</td>
<td>0.5</td>
<td>5</td>
<td>95</td>
</tr>
</tbody>
</table>

Electrospray ionization mass spectrometry was obtained under positive ion mode and negative ion mode respectively. The full scan MS and MSMS information were simultaneously collected adopting information dependent acquisition (IDA). Mass spectrometry data were collected between m/z 60 and 1,200 Da. The ion spray voltage of positive ion mode was 5,000 V, and the negative ion mode ion spray voltage was 4000 V, while the temperature of heated-capillary was kept at 600°C. The curtain air flow, nebulizer and heater gas were set at 35, 60 and 60 arbitrary units, respectively. The collision energy was set at 30 V.

Data processing

ProteoWizard (version 3.0.6150) was employed to convert raw MS data files to mzXML format and convert raw MS2 data files to mgf format. Then R package “XCMS” (version 1.46.0) was adopted to process all of MS data files (mzXML format) for peak detection and calibration. The identification of metabolites was accomplished by MetDNA (http://metdna.zhulab.cn/) based on the MS1 peak table and MS2 data files (mgf format). In order to discover differential metabolites, the MS1 peak table was conducted to identify differential metabolites. Based on the normalized peak table, principal component analysis (PCA) and partial least squares discrimination analysis (PLSDA) were performed according to the total intensity to investigate the possible separation of metabolite spectrum between PKD (KM) and ccRCC (KMS), as well as fold changes and p-values (evaluated by Student’s t-test) were calculated simultaneously. The dataset contained 20405 features, and the differential features were selected according to VIP more than 1.

9. Lipid-metabolite preparation
PKD and ccRCC renal tissues were sectioned by LCM and homogenized in 200 µL ddH$_2$O for three cycles, each cycle was centrifuged at 5500 r.p.m for 20s and repeat triplicate. Then 200 µL of homogenized tissue suspension was transferred to a new glass tube. Furthermore, 380 µL of ddH$_2$O and 960 µL of MTBE (methyl tert-butyl ether) /MeOH (v:v, 5:1) were added to 200 µL of the homogenate, vortexed for 60 s, sonicated for 10 min and spun at 3000 r.p.m. for 15 min at 4 °C, and the 500µL of supernatant containing metabolites was transferred into a new glass tube. The extraction step was repeated triplicate, and the supernatant containing metabolites (1.5 mL) were combined and evaporated to dryness adopting a high-speed vacuum concentrator at 4°C. Dried metabolite pellets were then resuspended in 100 µL of DCM (Dichloromethane) /MeOH (v: v, 1:1).

10. Generation and processing of LC-MS/MS Lipidomics data
The lipid-metabolite-containing supernatant was processed and analyzed by HPLC–MS/MS on the TripleTOFTM6600 mass spectrometer (AB SCIEX, USA) combined with Agilent 1290 liquid chromatography system (Agilent, USA). For liquid chromatography separation, Kinetex C18 column ((particlessize,1.7µm;100mm(length) × 2.1mm(i.d.); PhenomenexKinetex) was employed and 2µL of non-polar supernatant was injected and separated with an 18-minute gradient. The flow rate was set at 300 µL /min and the column temperature was 55°C. The chromatographic gradient program was shown in Table 2:

Table 2. The chromatographic gradient for Lipidomics data analysis

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>Flow rate (mL/min)</th>
<th>A (%)</th>
<th>B (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.3</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>12</td>
<td>0.3</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>14</td>
<td>0.3</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>14.2</td>
<td>0.3</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>18</td>
<td>0.3</td>
<td>60</td>
<td>40</td>
</tr>
</tbody>
</table>

Electrospray ionization mass spectrometry was obtained under positive ion mode and negative ion mode respectively. The full scan MS and MSMS information were simultaneously collected adopting information dependent acquisition (IDA). Mass spectrometry data were collected between m/z 200 and 2,000 Da. The ion spray voltage of positive ion mode was 5,000 V, and the negative ion mode ion spray voltage was 4500 V, while the temperature of heated-capillary was kept at 600°C. The curtain air flow, nebulizer and heater gas were set at 30, 60 and 60 arbitrary units, respectively. The collision energy was set at 40 ± 25 V.

Data processing
ProteoWizard (version 3.0.6150) was employed to convert raw MS data files to mzXML format and convert raw MS2 data files to mgf format. Then R package
“XCMS” (version 1.46.0) was adopted to process all of MS data files (mzXML format) for peak detection and calibration. The identification of metabolites was accomplished by LipidIMMS Analyzer (http://imms.zhulab.cn/LipidIMMS/) based on the MS1 peak table and MS2 data files (mgf format). In order to discover differential lipid-metabolites, the MS1 peak table was conducted to SIMCAP (Version 14.1) to identify differential lipid-metabolites. Based on the normalized peak table, PCA and PLSDA were performed according to the total intensity to investigate the possible separation of metabolite spectrum between PKD (KM) and ccRCC (KMS), as well as fold changes and p-values (evaluated by student’s t-test) were calculated simultaneously. The dataset contained 7025 features, and the dramatically differentiated features were selected based on VIP greater than 1.

11. MALDI imaging sample preparation
The 12-µm thick frozen sections were previously thaw-mounted on ITO slides for MALDI MSI analysis. Then the samples were transferred into a vacuum desiccator for 20min to remove the water on the surface of tissues. For MALDI-MSI analysis, DHB (2,5-dihydroxybenzoic acid) matrix, prepared at 15mg/mL in 90% ACN plus 0.1%TFA was employed to coat on the surface of tissues via a HTX TM-Sprayer (HTX, USA) under the following optimized conditions: a LC pump, a 0.125mL/min plate rate, a 60 °C nozzle temperature, a 1200mm/min Z-arm velocity, a 14 no of passes, a CC moving pattern and a 3mm/s track spacing.

12. Generation and processing of MALDI MSI data
MALDI MSI acquisition was conducted adopting a TimsTOF Flex MALDI-2 mass spectrometer (Bruker). Imaging data were collected in full-scan mode to maximize sensitivity. Spectrometry were collected in positive mode under the following optimized instrument parameters: m/z300-1500 of mass range, 10000Hz laser frequency, 400 accumulated shots and 50-µm spatial resolution.

Then Bruker SCiLS Lab 2020a software was adopted for the analysis of the MALDI MSI data. The following statistical analysis was performed for further data mining. Then “Hypothesis Tests” was applied for determining whether a feature can be used to confidently discriminate between ccRCC and XXX. “Find discriminative m/z values (ROC analysis)” was used to find m/z values that can discriminates between ccRCC and X tissue sections. Then “segmentation” was conducted to similarities between spectra and groped into different clusters. Furthermore, “Component Analysis” was performed to distinguish the control group form the experiment group. As well as, Metaboscape 8.0.q software was used for identification of m/z values in the regions of the interest.

Reference:


20. Xie, Y., *et al.* SETD2 loss perturbs the kidney cancer epigenetic landscape to...


Extended Data
**Extended Data Fig. 1** | The scheme of multi-omics and quality assessments for transcriptomics, proteomics, lipidomics and metabolomics data. A. The workflow of samples preparation, multi-omics experiments and data analysis. B. Number of detected proteins in KM and KSM renal tissues. C-E, person correlations of proteomics (C), metabolomics (D) and lipidomics (E) data between samples from KM and KSM renal tissues.

**Extended Data Table 1.** The details information of laboratory mice

<table>
<thead>
<tr>
<th>Cases</th>
<th>Name</th>
<th>Genotype</th>
<th>Gender</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>KM_1</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Female</td>
<td>8W</td>
</tr>
<tr>
<td>2</td>
<td>KM_2</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Male</td>
<td>8W</td>
</tr>
<tr>
<td>3</td>
<td>KM_3</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Male</td>
<td>8W</td>
</tr>
<tr>
<td>4</td>
<td>KM_4</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Male</td>
<td>12W</td>
</tr>
<tr>
<td>5</td>
<td>KM_5</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Male</td>
<td>50W</td>
</tr>
<tr>
<td>6</td>
<td>KM_6</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Male</td>
<td>12W</td>
</tr>
<tr>
<td>7</td>
<td>KM_7</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Male</td>
<td>8W</td>
</tr>
<tr>
<td>8</td>
<td>KM_8</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Female</td>
<td>8W</td>
</tr>
<tr>
<td>9</td>
<td>KM_9</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;</td>
<td>Male</td>
<td>8W</td>
</tr>
<tr>
<td>10</td>
<td>KSM_1</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;Setd2&lt;sup&gt;flox/flox&lt;/sup&gt;</td>
<td>Female</td>
<td>40W</td>
</tr>
<tr>
<td>11</td>
<td>KSM_2</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;Setd2&lt;sup&gt;flox/flox&lt;/sup&gt;</td>
<td>Male</td>
<td>16W</td>
</tr>
<tr>
<td>12</td>
<td>KSM_3</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;Setd2&lt;sup&gt;flox/flox&lt;/sup&gt;</td>
<td>Male</td>
<td>16W</td>
</tr>
<tr>
<td>13</td>
<td>KSM_4</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;Setd2&lt;sup&gt;flox/flox&lt;/sup&gt;</td>
<td>Male</td>
<td>9W</td>
</tr>
<tr>
<td>14</td>
<td>KSM_5</td>
<td>Ksp&lt;sup&gt;Cre&lt;/sup&gt;MYC&lt;sup&gt;R26&lt;/sup&gt;Stop&lt;sup&gt;FL/+&lt;/sup&gt;Setd2&lt;sup&gt;flox/flox&lt;/sup&gt;</td>
<td>Male</td>
<td>16W</td>
</tr>
</tbody>
</table>
Extended data Fig.2 | Proteomics analysis of KM and KSM renal tissues.

A. Principal component analysis for the detected proteins in KM and KSM renal tissues.
B. The top ten most significantly altered GO terms including biological process, molecular function and cellular component, respectively, which enriched by significantly upregulated (red) and downregulated (blue) proteins (FDR <0.05 and FC>2).
C. The top 27 most significantly upregulated (red) KEGG pathways and top 22 most significantly downregulated (blue) KEGG pathways, respectively, which enriched by gene set enrichment analysis (FDR <0.05 and FC>2).
D-F. Heatmap of significantly altered proteins (FDR<0.05 and FC>2) related to oxidative phosphorylation (D),
cholesterol metabolism (E), ABC transporters (F) (n = 9 mice in KM group, n = 5 mice in KMS group).

Extended data Fig.3 | Metabolism-related proteomic analysis of KM and KMS renal tissues.

A-C. Heatmap of significantly altered proteins (FDR<0.05 and FC>2) related to pyruvate metabolism (A), glycerophospholipid metabolism (B) and biosynthesis of unsaturated fatty acids (C) (n = 9 mice in KM group, n = 5 mice in KMS group). D. Relative abundance of representative proteins related with pyruvate metabolism, glycerophospholipid metabolism, biosynthesis of unsaturated fatty acids, glycolysis, TCA, glucose and amino acid transport, amino acid metabolism (n = 9 mice in KM group, n = 5 mice in KMS group). a.u., arbitrary unit. To calculate relative abundance, the peak area of each protein was divided by the sum peak areas of all detected proteins in the same sample and the resulting ratio value was used as the relative abundance of each protein. Data are shown as mean ± s.e.m. Two-tailed unpaired Student’s t-tests followed by a Benjamini–Hochberg multiple comparison test were used for statistical analysis. Exact FDR values are indicated.
E xtended Data Fig. 4 | Protein-protein interaction in KEGG metabolism pathways and proteomic changes in SETD2-deficiency caused ccRCC renal tissues.

A. Heat map of proteomic changes of biosynthesis of unsaturated fatty acids, glycerophospholipid metabolism, fatty acid biosynthesis, protein digestion and absorption, fatty acid metabolism, sphingolipid metabolism, carbohydrate digestin and absorption, glycolysis/ gluconeogenesis, ABC transporters, TCA, pyruvate metabolism, Cholesterol metabolism and glycerolipid metabolism between KM and KMS renal tissues (n = 9 mice in KM group, n = 5 mice in KMS group). B. Correlations of altered metabolic related proteins between KM and KMS renal tissues. Metabolic related proteins were listed in the figure.
Extended Data Fig. 5 | Lipid metabolism-related differential proteins and KEGG pathways between KM and KSM renal tissues.

A. Hierarchical cluster analysis of alteration proteins abundance related to sphingolipid metabolism and glycerophospholipid metabolism between KM and KSM renal tissues (n = 9 mice in KM group, n = 5 mice in KMS group). B–C. Proteomic analysis of sphingolipid metabolism (B) and glycerophospholipid metabolism (C). Protein changes of metabolic enzymes in KM and KMS mouse renal were indicated. Proteins significantly upregulated in KMS mouse renal tissue were marked in red (FDR < 0.05).
Extended Data Fig. 6 | The PCA and person correlations of metabolic and lipidomic differences between KM and KSM renal tissues. The PCA analysis of differences metabolites based on lipidomic (A) and metabolic (B) between KM and KSM group. The person correlations of metabolomics (C) and lipidomics (D) differences metabolites data between samples from KM and KSM renal tissues (n = 5 mice in KM group, n = 3 mice in KMS group).
Extended Data Fig. 7 | Relative abundance of representative proteins related to glycolysis, citrate cycle, amino acid metabolism, protein catabolic and sphingomyelin biosynthesis. These data are based on clinical proteomic tumor analysis consortium (CPTAC).