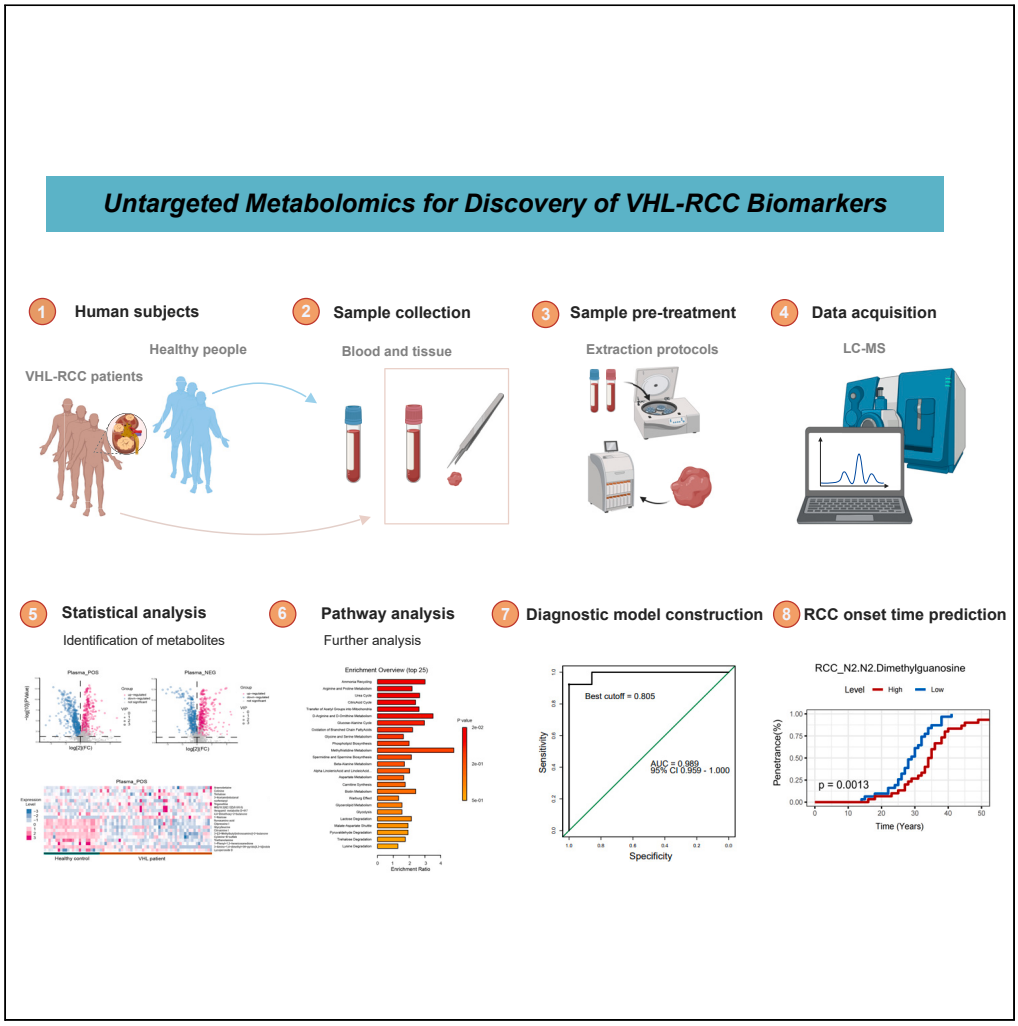
iScience

CellPress
OPEN ACCESS

Article

Metabolomic landscape of renal cell carcinoma in von Hippel-Lindau syndrome in a Chinese cohort



Zedan Zhang, Yi Wang, Wuping Yang, ..., Xiaohua Zhou, Yanqing Gong, Kan Gong

azhou@math.pku.edu.cn (X.Z.)
yqgong@bjmu.edu.cn (Y.G.)
gongkan_pku@126.com (K.G.)

Highlights

Untargeted metabolomics reveals key metabolic molecular profile in hereditary VHL-RCC

Top 10 DAMs establish a high-performance diagnostic model for hereditary VHL-RCC

Plasma level of N2,N2-dimethylguanosine predicts RCC onset timing in VHL patients

Zhang et al., iScience 27, 110357
July 19, 2024
https://doi.org/10.1016/j.isci.2024.110357

iScience

CellPress
OPEN ACCESS

Article

Metabolomic landscape of renal cell carcinoma in von Hippel-Lindau syndrome in a Chinese cohort

Zedan Zhang,[1,2,4] Yi Wang,[3,4] Wuping Yang,[1,2,4] Tao Liu,[1,2] Chuandong Wang,[1,2] Cong Huang,[1] Yawei Xu,[1,2] Xiaolin Chen,[1,2] Jingcheng Zhou,[1,2] Yizhou Wang,[1,2] Xiaohua Zhou,[3,*] Yanqing Gong,[1,2,*] and Kan Gong[1,2,5,*]

SUMMARY

Von Hippel-Lindau (VHL) syndrome is a rare autosomal dominant disorder, where renal cell carcinoma (RCC) serves as a significant cause of mortality. We collected peripheral blood from 61 VHL-RCC patients and 31 healthy individuals, along with 19 paired RCC tumor and adjacent non-malignant samples. Using liquid chromatography-mass spectrometry, we identified 238 plasma and 241 tissue differentially abundant metabolites (DAMs), highlighting key pathways such as arginine and proline metabolism. The top 10 of the 23 DAMs, common to both plasma and tissue, were instrumental in constructing a high-performance diagnostic model. These DAMs demonstrated significant correlations with *VHL* gene mutation types. Cox regression analysis revealed that plasma levels of N2,N2-dimethylguanosine were associated with the timing of RCC onset in VHL patients, acting as an independent predictive factor. This study enhances diagnostic accuracy for this rare condition and opens new avenues for exploring metabolic mechanisms of the disease and potential therapeutic directions.

INTRODUCTION

Von Hippel-Lindau (VHL) syndrome, a rare autosomal dominant disorder, affects approximately 1 in 36,000 to 53,000 births.[1,2] This condition arises from mutations in the VHL gene on chromosome 3, with a notable genetic penetrance exceeding 90% by age 70. The VHL gene encodes the VHL protein, a key component of the VBC E3 ubiquitin ligase complex, alongside elongation factors B and C. This complex plays a pivotal role in degrading hypoxia-inducible factors-α (HIF-α), which are downstream targets.[3] The fundamental mechanism underlying VHL syndrome involves the VHL protein's dysfunction, resulting in elevated levels of substrates like HIF-α. This elevation triggers the activation of various oncogenic factors, playing a central role in increasing the risk of various tumors, including central nervous system hemangioblastomas (CHB), renal cell carcinomas (RCC), retinal hemangioblastoma (RA), pheochromocytomas (PHEO), pancreatic cysts or tumors (PCT), and tumors in the genital system (GS) and endolymphatic sac.[1,4] With advancing age, the occurrence rate of VHL disease escalates, culminating in a penetrance rate of approximately 90% in patients above the age of 65.[2,5]

VHL disease, in clinical practice, is stratified into two types based on the presence or absence of PHEO, with this classification deeply anchored in genotype-phenotype correlations.[6] Type 1 VHL, often linked with truncating mutations, typically manifests with classic VHL symptoms like CHB and RCC, but notably lacks PHEO. On the other hand, Type 2 VHL, constituting about 7%–20% of VHL cases and characterized by missense mutations, includes PHEO in its spectrum. This type is further subdivided based on clinical presentation: Type 2A, without RCC; Type 2B, inclusive of RCC; and Type 2C, marked exclusively by the presence of PHEO.[7–9]

In VHL patients, RCC and CHB are the primary contributors to mortality.[10] The presence of these complications significantly elevates the risk of death in VHL disease, rendering the prognosis of this hereditary condition particularly complex and challenging.[5,11] In contrast to sporadic RCC, VHL-RCC often presents earlier, primarily between the ages of 30 and 50, and around 70% of VHL individuals are likely to develop RCC by their 60s[12,13] Characterized by its bilateral and multifocal traits, this variant of RCC in VHL pathology is also frequently accompanied by numerous renal cysts. These cysts are considered potential harbingers of clear cell RCC, a subtype closely and almost exclusively associated with VHL.[14] Moreover, the metastatic progression of RCC stands as a primary cause of mortality among individuals afflicted with VHL disease. VHL-RCC distinctly differs in its biological behavior from sporadic cases. Despite employing nephron-sparing surgery, the recommended first-line treatment, patients often face a high rate of recurrence post-operatively. Additionally, the necessity for

[1]Department of Urology, Peking University First Hospital, Beijing, China
[2]Hereditary Kidney Cancer Research Center, Peking University First Hospital, Beijing, China
[3]Beijing International Center for Mathematical Research and Department of Biostatistics, Peking University, Beijing, China
[4]These authors contributed equally
[5]Lead contact
*Correspondence: azhou@math.pku.edu.cn (X.Z.), yqgong@bjmu.edu.cn (Y.G.), gongkan_pku@126.com (K.G.)
https://doi.org/10.1016/j.isci.2024.110357

multiple surgeries can accelerate progression to end-stage disease, thereby significantly increasing both the physical and financial burden on patients.[15,16] Consequently, in individuals diagnosed with VHL disease, the goal of surgical intervention diverges from that in sporadic cases. The focus is more on preserving renal function and reducing the risk of metastasis, rather than solely aiming for complete tumor removal.

Preliminary retrospective studies have revealed that key factors such as age at onset, family history, type of mutation, and initial symptoms significantly influence patient survival.[17] Therefore, it is imperative to explore whether new biomarkers exist that can enable the early detection and prognostication of VHL-associated tumors, especially RCC. The identification of such markers would be pivotal in enhancing early diagnostic protocols and refining risk assessment strategies for RCC in this unique patient group.

RCC is considered a metabolic disease. When kidney cancer occurs, it disrupts the body's original metabolic habits causing changes in the body's related metabolic indicators. This characteristic provides a theoretical basis for the search for early tumor markers of RCC.[18–20] Metabolomics, a rapidly evolving field in biomedical research, focuses on the in-depth analysis of metabolites within a biological system. These metabolites, typically the end products of cellular activities, provide an immediate reflection of the physiological state of an organism. In oncology, metabolomics has revolutionized our understanding of cancer metabolism, revealing that these metabolic shifts are crucial in the onset, growth, and spread of tumors, rather than being mere byproducts of cancer.[21,22] Metabolomic research predominantly employs mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy. Key MS techniques include liquid chromatography-mass spectrometry (LC-MS) and gas chromatography-mass spectrometry (GC-MS), each with unique strengths and specific application scopes. For instance, Liu et al. has used LC-MS to demonstrate the potential of plasma metabolomics and lipidomics in diagnosing RCC, achieving high accuracy in distinguishing RCC from healthy and benign controls.[23] Maslov et al. conducted a metabolomic analysis using plasma samples from 51 healthy volunteers and 78 patients with various pathological types of RCC. Through this study, they developed a model with a higher diagnostic power and accuracy.[24] In addition to renal cancer, there is extensive research in the field of metabolomics across various cancer types.[25] The findings from these studies suggest the significant potential of metabolomics in identifying novel biomarkers for cancer diagnosis and screening.

In our study, to exclusively identify plasma DAMs primarily influenced by VHL-RCC and to eliminate the potential confounding effects of other VHL syndrome-related lesions such as pheochromocytomas,[26] pancreatic tumors,[27] and tumors in the genital system on plasma metabolite levels,[28] we collected peripheral blood from 61 VHL-RCC patients and 31 healthy individuals, along with 19 paired RCC tumor and adjacent non-malignant tissue (AN) samples from VHL patients. Using Liquid Chromatography-Mass Spectrometry, we analyzed low-molecular-weight metabolites, identifying key differences in plasma and tissue with high accuracy. This approach is integral to our objective of utilizing metabolomic sequencing to unveil distinct metabolic alterations in VHL-related tumors, particularly in VHL-related RCC. By doing so, we anticipate enhancing early detection, enabling prompt and effective interventions, and guiding the development of personalized, targeted treatment strategies for VHL-RCC patients.

RESULTS

Clinical and genetic characteristics

The study's methodology is depicted in Figure 1. A cohort containing 61 VHL patients with RCC and 31 healthy subjects was delineated by untargeted metabolomic profiling of plasma specimens in Table 1, Table 2, respectively. The cohort of VHL patients exhibited a balanced gender composition (49.2% males and 50.8% females). Notably, a significant majority of the patients (65.6%) reported a familial predisposition to the condition. Genetic profiling identified a spectrum of mutations; notably, PM (Point mutation)-Exon1 mutations were predominant, accounting for 41% of cases, succeeded by PM-Exon3 mutations in 29.5% and deletions (including partial complete deletions) in 23% of the patient population. The emergence of VHL-related symptoms predominantly transpired before the age of 30, encompassing 62.3% of the cases. The initial organ affected was most frequently RCC (42.6%), followed by CHB (26.2%). A meticulous evaluation of organ involvement revealed that patients with RCC also frequently presented with PCT in 70.5% of cases and CHB in 54.1%. In Table 2, the Chi-square analysis revealed no statistically significant differences in gender or birth year distributions between the group of heathy individuals and VHL patients. Table 3 details the clinical profile of 19 tissue-sampled VHL-RCC patients. The group displayed a male predominance (63.2%), with most diagnoses occurring before age 30 (57.9%). The majority had a family history of VHL (68.4%). Nearly all patients exhibited Type I VHL (89.5%), with RCC present in all cases and PCT in 63.2%. Additional clinical details are provided in Tables S1 and S2.

Quality control (QC) of the data

In plasma samples, the number of POS and NEG precursor molecules initially identified were 5,457 and 5,281, respectively, which reduced to 4,122 and 3,862 after the processing and filtering steps mentioned in the Methods section. Similarly, in tissue samples, the count for POS and NEG precursors decreased from 4,801 to 3,576 and from 5,187 to 4,076, respectively, following the same processing protocol. Theoretically, QC samples should be identical, but variations in substance extraction and analytical detection can lead to discrepancies among them. The smaller these differences, the greater the stability of the method and the higher the quality of the data. Figures S1A and S1B and Figures S1E and S1F exhibit remarkable clustering of the QC samples for plasma and tissue, respectively, which underscores the robustness and stability of the methodological approach. Figures S1C and S1D and Figures S1G and S1H illustrate that all QC samples are located within ± 2 STD (standard deviations), indicating reliable data quality in this experiment.

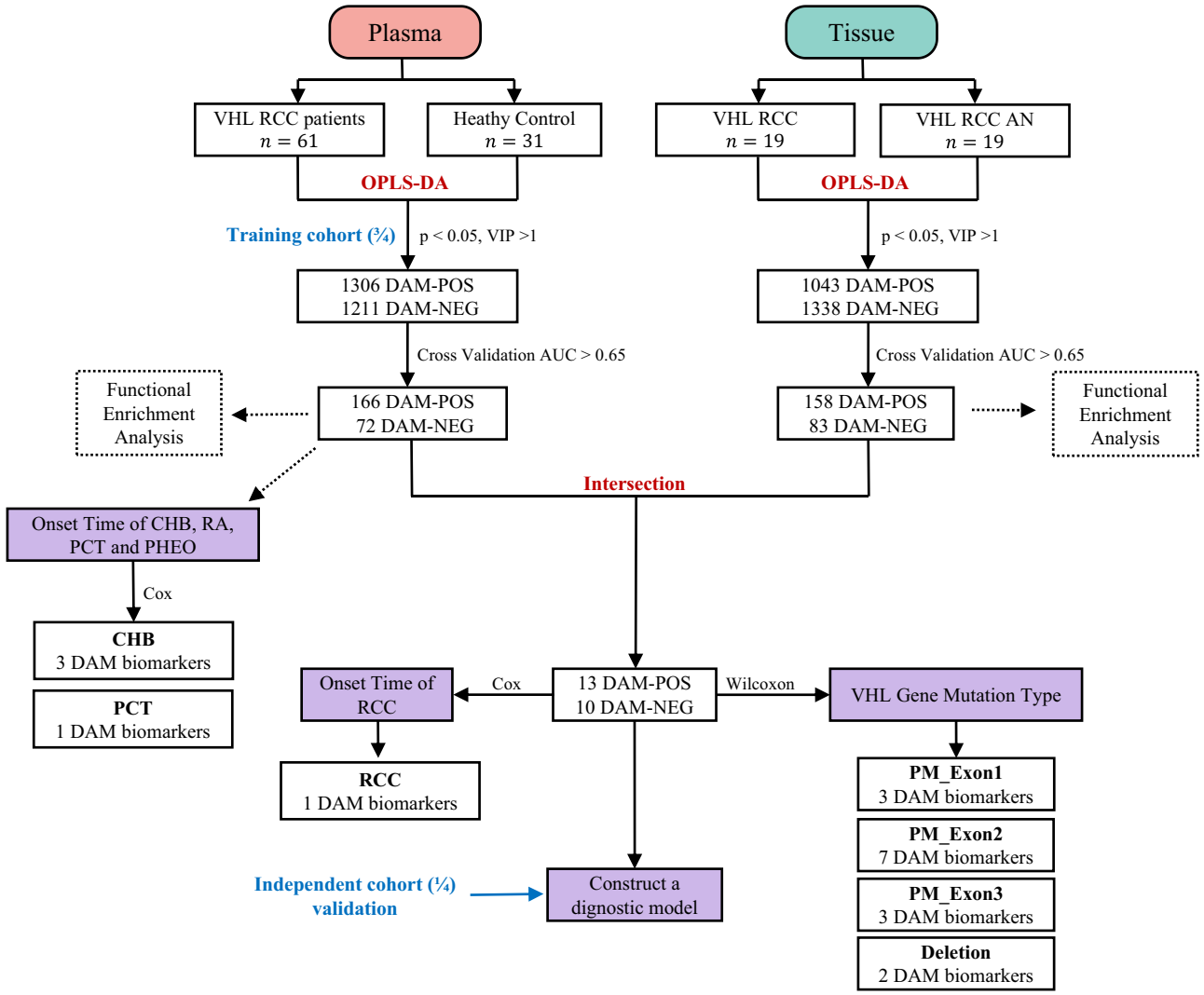


Figure 1. Flowchart of the study process

AN, adjacent non-malignant tissue; CHB, central nervous system hemangioblastomas; PHEO, pheochromocytomas; PCT, pancreatic cysts or tumors; RA, retinal hemangioblastoma; RCC, renal cell carcinoma; DAMs, differentially abundant metabolites; OPLA-DS, orthogonal projections to latent structures-discriminant analysis; POS, positive ionization mode; NEG, negative ionization mode; PM, point mutation.

DAM profiles in plasma and tissue samples

The metabolites detected from plasma and tissue samples were all used in OPLS-DA, which could effectively reduce model complexity and enhances interpretability to maximize the visualization of intergroup differences. The scatter point score plots from the OPLS-DA analysis for both plasma and tissue samples, in POS and NEG ionization modes respectively, are presented in Figures 2A–2D. These plots illustrate significant intergroup differences, while intragroup variations are relatively inconspicuous.

To assess the validity of our model and to protect against overfitting, a permutation test (n = 200 permutations) was conducted. The resulting permutation distribution for R^2 and Q^2 of both plasma and tissue samples in POS and NEG modes are depicted in histograms (Figures 2E–2H). The original model's R^2 value, indicating the explained variance, was significantly higher than that of the permuted models, with the actual value positioned at the far right end of the distribution. This demonstrates that the model explains the variance within the data to a degree that is unlikely to be due to chance. Similarly, the Q^2 value, denoting the predictive accuracy of the model, was notably superior to the permuted outcomes. The actual Q^2 was situated at the extreme of the permuted distribution, exceeding the 95th percentile of permuted Q^2 values. This stark contrast underscores the predictive reliability of the model and suggests that it has a high degree of validity in forecasting outcomes based on the metabolomic data.

Subsequent comparative analyses between the two groups utilized univariate analysis with the Mann-Whitney U test. Metabolites with P-value smaller than 0.05 from the test and those with a VIP value greater than 1 from the OPLS-DA model were intersected. This approach yielded a total of 1306 POS DAMs for the plasma group (Figure 3A) and 1211 NEG DAMs (Figure 3B). Similarly, for the tissue group, there were

Table 1. Clinical characteristics of VHL-RCC patients in untargeted plasma metabolomics analysis

Characteristics	Count	Ratio(%)
Overall	61	
Gender		
Male	30	49.2
Female	31	50.8
Birth year		
≤1980	21	34.4
>1980	40	65.6
Family history		
Yes	40	65.6
No	21	34.4
Gene type		
PM-Exon1	25	41
PM-Exon2	4	6.6
PM-Exon3	18	29.5
Deletion	14	23
Generation		
1	25	41
2	23	37.7
3	13	21.3
Type		
I	48	78.7
IIB	13	21.3
Onset age		
≤30	38	62.3
>30	23	37.7
First affected organ		
CHB	16	26.2
RA	4	7
RCC	26	42.6
PCT	7	11.5
PHEO	7	11.5
GS	1	1.7
Affected organ		
CHB	33	54.1
RA	14	23
RCC	61	100
PCT	43	70.5
PHEO	13	21.3
GS	6	9.8

VHL, von Hippel-Lindau disease; PM, point mutation; CHB, central nervous system hemangioblastoma; RA, retinal hemangioblastoma; RCC, renal cell carcinoma; PCT, pancreatic cyst or tumor; PHEO, pheochromocytoma; GS, genital system (epididymis or broad ligament).

1043 POS DAMs (Figure 3C) and 1338 NEG DAMs (Figure 3D) identified. For each comparison, heatmaps were constructed to display the top 10 upregulated and downregulated DAMs in the VHL patient and tumor group (Figures 3E–3H).

To further refine the selection of DAMs with higher specificity and sensitivity, those from both plasma and tissue sources with an AUC greater than 0.65 were filtered for subsequent analysis. After this selection process, the plasma group retained 166 DAMs in POS and 72 DAMs in NEG, while the tissue group retained 158 DAMs in POS and 83 DAMs in NEG.

Table 2. Clinical characteristics comparison of patients and healthy donors in untargeted plasma metabolomics analysis

	CON		VHL		
Characteristics	Patient	Ratio(%)	Patient	Ratio(%)	*p*-value
Gender					
Female	17	54.8	31	50.8	0.886
Male	14	45.2	30	49.2	
Birth year					
≤1980	5	16.1	21	34.4	0.11
>1980	26	83.9	40	65.6	

CON, healthy donors; VHL, von Hippel-Lindau-related renal cancer patients.

Pathway enrichment analysis of DAMs

To understand the characteristics of these DAMs, pathway enrichment analysis was conducted using the Small Molecule Pathway Database (SMPDB) of MetaboAnalyst.[29] The enrichment figure presents a comparative view of pathway enrichment analyses derived from two distinct sources: plasma (Figure 4A) and tissue (Figure 4B), each illustrating the top 25 metabolic pathways. Notably, both panels exhibit a subset of common pathways, such as arginine and proline metabolism and ammonia recycling, suggesting a shared metabolic emphasis. However, distinct pathways are also apparent. The plasma source uniquely highlights pathways like citric acid cycle and glycolysis, indicating specific metabolic activities in the circulatory milieu. Conversely, the tissue source underscores unique enrichments in histidine Metabolism and aspartate metabolism, reflecting the localized cellular processes. These differential enrichments emphasize the metabolic diversity between the systemic environment and the cellular context, which could have significant implications for understanding metabolic dynamics in health and disease.

DAMs specific to VHL-RCC

To identify DAMs specifically associated with VHL patient with RCC, we intersected the DAMs with an AUC greater than 0.65 from both plasma and tissue sources to obtain common DAMs. A Venn diagram illustrates the intersection numbers between different groups (Figure S2A; Table S1), showing that out of 238 DAMs from plasma and 241 DAMs from tissue, there are 23 DAMs in common. Lollipop charts present the AUC values of these 23 common DAMs originating from both POS (Figure 5A) and NEG (Figure 5B) ion modes across different tissues. For instance, in the POS mode, common DAMs include N2,N2-dimethylguanosine, Montecristin, 1-Kestose et al., whereas in the NEG mode, common DAMs include Indoxyl sulfate, Creatinine, N2-gamma-Glutamylglutamine et al.

Predictive power of DAMs for VHL-RCC

To accurately identify VHL patients with RCC within a extensive population, a diagnostic model was constructed using the top 10 DAMs selected from a common pool of 23, based on their importance scores specific to VHL-RCC calculated using the randomForest package (Figure 5C). This model demonstrated high efficiency in discriminating between healthy individuals and VHL patients with RCC in the training cohort (Figure S2B). Furthermore, the model's effectiveness was confirmed in an independent test cohort, exhibiting excellent validation metrics (AUC = 0.989, 95% CI = 0.959–1.000, Figure 5D).

Among these 10 DAMs, we identified N2,N2-dimethylguanosine as a potential predictor for the onset of RCC in VHL patients. Compared to the healthy population, its level is significantly downregulated in the plasma of patients with VHL-RCC ($p < 0.001$). In a comparison between RCC tumor and AN tissues in VHL patients, the level of N2,N2-dimethylguanosine was markedly lower in tumor tissues ($p < 0.001$)(Figure 6A). KM curve analysis revealed a significant association between the plasma levels of N2,N2-dimethylguanosine and the occurrence of RCC in patients (Figure 6B). Lower levels of N2,N2-dimethylguanosine were indicative of an earlier onset of RCC ($p = 0.0013$). Both univariate and multivariate Cox regression analyses confirmed that N2,N2-dimethylguanosine and patients birth year could independently predict the timing of RCC onset in VHL patients (Figure 6C). The internally validated model yielded a c-index of 0.81, indicating the strong predictive accuracy.

Correlation between DAMs and VHL mutation type

To investigate whether there is an association between *VHL* gene mutation sites and types and the levels of DAMs in the plasma of patients with VHL-associated RCC, Mann-Whitney tests were conducted to assess the significance of differences. Results from Figure 7A indicated that when point mutations occur in the first exon (PM-Exon1) of the *VHL* gene, compared to other mutation types, there was a significant reduction in the levels of Glycerophospholipids (PC(24:1(15Z)/18:4(6Z,9Z,12Z,15Z))) and Cysteine-S-sulfate ($p < 0.05$), and a significant elevation of Fatty Acyls ((10E,12Z)-(9S)-9-Hydroperoxyoctadeca-10,12-dienoic acid) ($p < 0.01$) in plasma. In the group with PM-Exon2 (Figure 7B; Figure S3), there was a significant elevation of Glycerophospholipids (PE(22:4(7Z,10Z,13Z,16Z)/14:0)), Glycerophospholipids (PE(16:0/18:2(9Z,12Z))), cysteine-S-sulfate, gamma-glutamylalanine, and Triazines (6-Chloro-N-(1-methylethyl)-1,3,5-triazine-2,4-diamine) ($p < 0.05$), while glycerophospholipids

Table 3. Clinical characteristics of VHL-RCC patients in untargeted tissue metabolomics analysis

Characteristics	Count	Ratio(%)
Overall	19	
Gender		
Male	12	63.2
Female	7	36.8
Birth year		
≤1980	11	57.9
>1980	8	42.1
Family history		
Yes	13	68.4
No	6	31.6
Gene type		
PM-Exon1	5	26.3
PM-Exon2	2	10.5
PM-Exon3	2	10.5
Deletion	10	52.6
Generation		
1	9	47.4
2	8	42.1
3	2	10.5
Type		
I	17	89.5
IIB	2	10.5
Onset age		
≤30	11	57.9
>30	8	42.1
Affected organ		
CHB	6	31.6
RA	6	31.6
RCC	19	100
PCT	12	63.2
PHEO	2	10.5
GS	4	21.1

VHL, von Hippel-Lindau disease; PM, point mutation; CHB, central nervous system hemangioblastoma; RA, retinal hemangioblastoma; RCC, renal cell carcinoma; PCT, pancreatic cyst or tumor; PHEO, pheochromocytoma; GS, genital system (epididymis or broad ligament).

(LysoPC(16:1(9Z)/0:0)) was significantly reduced ($p < 0.05$). In the group with PM-Exon3 (Figure 7C), both L-Malic acid and Fatty Acyls ((10E,12Z)-(9S)-9-hydroperoxyoctadeca-10,12-dienoic acid) were significantly downregulated ($p < 0.05$). In the group with deletions in the *VHL* gene, Glycerophospholipids (PC(24:1(15Z)/18:4(6Z,9Z,12Z,15Z))) were significantly upregulated ($p < 0.05$). These results indicate a possible correlation between various VHL gene mutation types and the levels of plasma DAMs. This, coupled with our prior discoveries that distinct VHL mutation types are associated with varying clinical presentations in patients,[30] points toward the opportunity for cohort expansion for additional validation. Consequently, these identified plasma DAMs could potentially be used for enhanced stratification of VHL patients in future studies.

Plasma DAM levels and lesion onset of non-renal organs

In line with the background information, VHL patients often present with concurrent pathologies in various organs. We aimed to investigate whether certain plasma-derived DAMs could serve as predictors for the onset of non-renal organ conditions such as CHB, PCT, PHEO, RA, and GS. Kaplan-Meier analyses identified a clear correlation between the plasma concentrations of hypoxanthine, dodecanoylcarnitine, and 4-Dihydroxy-2-hydroxymethyl-1-pyrrolidinepropanamide (4D2h1p) with the early onset of CHB (Figures 8A–8C). Elevated plasma hypoxanthine levels ($p = 0.0038$) and reduced levels of dodecanoylcarnitine ($p = 0.044$) and 4D2h1p ($p = 0.045$) were indicative of a sooner occurrence

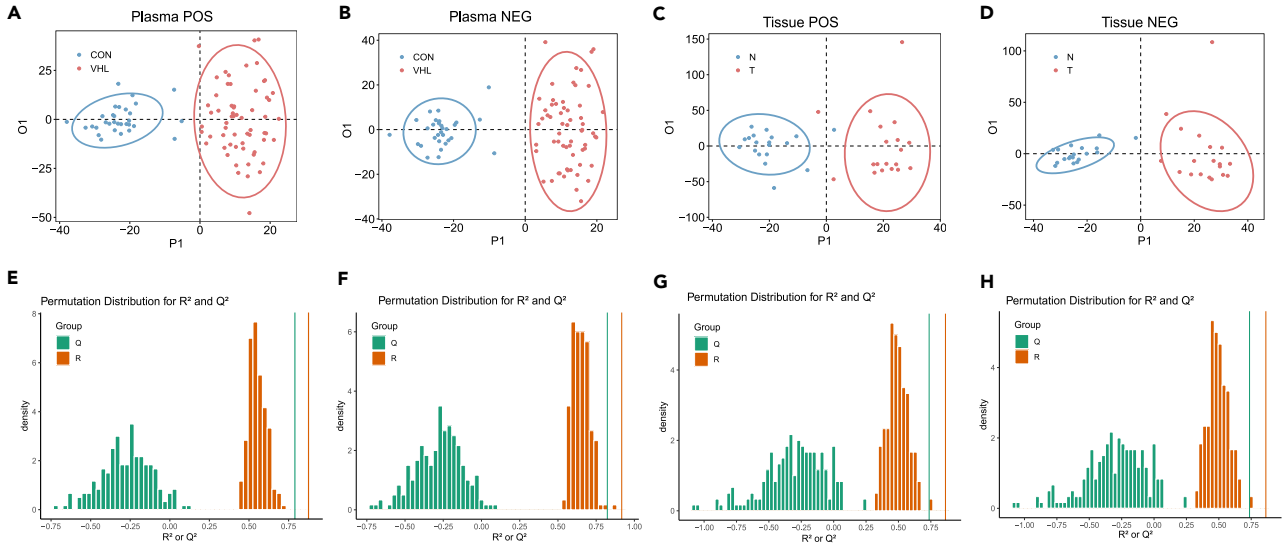


Figure 2. OPLS-DA model construction and permutation test analysis

Score plots of OPLS-DA for plasma samples under POS (A) and NEG (B), and tissue samples under POS (C) and NEG (D). The horizontal axis P1 represents the predictive principal component score of the first principal component, illustrating the differences between sample groups. The vertical axis O1 denotes the orthogonal principal component score, showing the variability within each sample group. Each scatter point represents an individual sample, with the color of the points indicating different experimental groupings. Histograms of permutation test for plasma samples under POS (E) and NEG (F), and tissue samples under POS (G) and NEG (H). They depict the permutation distribution for R^2 and Q^2 values, evaluating model robustness for the corresponding score plots. OPLS-DA, orthogonal projections to latent structures-discriminant analysis; POS, positive ionization mode; NEG, negative ionization mode.

of CHB. Moreover, a lower plasma concentration of Trehalose was associated with an earlier onset of PCT ($p < 0.029$)(Figure 8D). Owing to the small sample sizes of patients with PHEO, RA, and GS, no significant DAMs were discerned for these conditions.

DISCUSSION

Renal lesions in VHL disease exhibit a wide spectrum, extending from simple to complex cysts, and evolving into fully solid masses. This diversity in lesion types significantly elevates the risk of developing RCC, a hallmark of which includes bilateral and multifocal tumors, along with a high recurrence rate post-surgery. Such complexities often necessitate bilateral nephrectomy, leading to the subsequent need for dialysis in these patients.[31–34]

The complexity of RCC in VHL patients is evident not only in its pathological manifestations but also in the diversity of underlying molecular and cellular mechanisms. These varied mechanisms challenge conventional diagnostic and treatment approaches, prompting scientists to delve deeper, particularly into the molecular alterations in RCC. RCC recognized as a metabolic disease, is intricately linked to key genetic mutations such as *VHL*, *FH*, *MET*, and *BHD*.[35,36] These mutations play a crucial role in reprogramming metabolic pathways and networks during the progression of RCC.

The *VHL* gene generates two pVHL protein isoforms, pVHL30 and pVHL19, integral to the VBC E3 ubiquitin ligase complex.[37,38] This complex targets HIF-α for degradation, a process crucial for preventing uncontrolled gene activation under normoxic conditions. In patients with VHL syndrome, mutations in the *VHL* gene result in loss or reduction of VHL protein function, which impedes the normal degradation of HIF proteins, thus leading to abnormally high levels of HIF under hypoxic conditions. Particularly, the increase in HIF1α and HIF2α not only activates hypoxia-responsive genes associated with tumor growth, angiogenesis, and metastasis but also broadly affects multiple metabolic pathways including glycolysis, oxidative phosphorylation, fatty acid metabolism, glutamine breakdown, glutathione biosynthesis, amino acid metabolism, and the pentose phosphate pathway.[39,40] Within these pathways, HIF1α primarily enhances glycolysis to meet the energy demands in a low-oxygen environment, whereas HIF2α more significantly regulates iron metabolism and erythropoiesis, and promotes cell proliferation, growth, and migration in hypoxic conditions. More specifically, increased HIF activity reduces mitochondrial oxygen consumption and mitochondrial quality, and increases glycolysis, fatty acid synthesis, and glutaminolysis. The regulation of these metabolic processes is a key mechanism by which HIF activation enables cancer cells to adapt to the hypoxic microenvironment and promotes their proliferation and invasive capabilities.

Employing metabolomics to identify biomarkers of RCC illuminates these genetic and metabolic variations and provides valuable insights into tumor pathophysiology. This method greatly improves our ability to diagnose and prognosticate RCC and to discover non-invasive biomarkers. It enriches our understanding of the disease's core nature and helps identify potential therapeutic targets.[41,42] For instance, most metabolomic research on RCC to date has centered around kidney tissues and biofluids like plasma, serum, and urine.[23,43–45] Yet, these

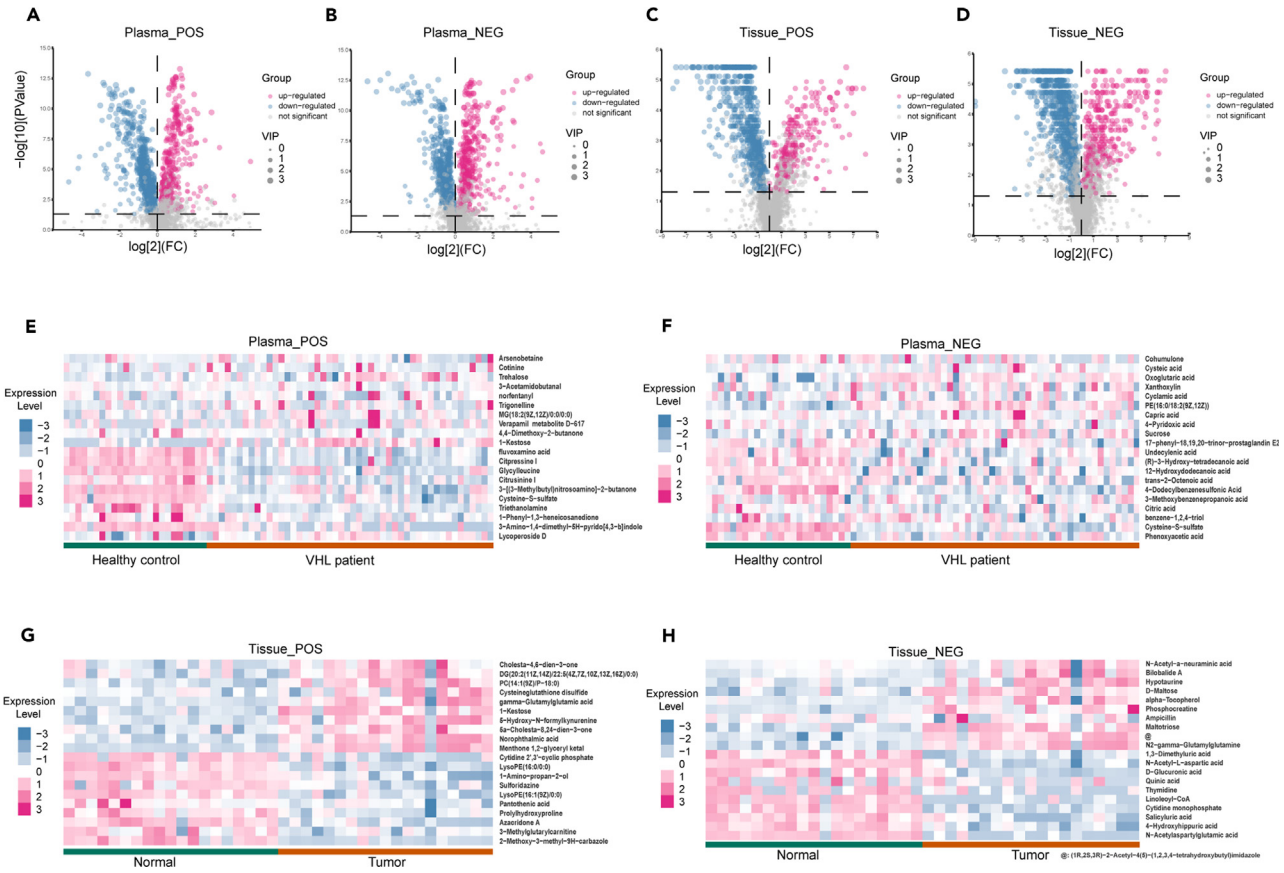


Figure 3. Differential metabolite analysis in plasma and tissue samples

Volcano plots of DAMs for plasma samples under POS (A) and NEG (B), and tissue samples under POS (C) and NEG (D). Metabolites with VIP score larger than 1 are highlighted, with upregulated metabolites in red and downregulated metabolites in blue. Heatmaps of DAMs for plasma samples under POS (E) and NEG (F), and tissue samples under POS (G) and NEG (H). DAMs, differentially abundant metabolites; POS, positive ionization mode; NEG, negative ionization mode; VIP, variable importance in the projection.

studies largely target RCC in a general sense and often overlook the distinct subset of hereditary kidney cancers, especially VHL-associated RCC.

Building upon this, we are exploring several key questions: Are there specific metabolic biomarkers that can distinguish cases of VHL patients with RCC within a broader population? In light of the critical role that the timing of organ-specific disease onset plays in the survival outcomes of VHL patients, especially concerning RCC, and our preliminary findings indicating a faster growth rate of RCC in VHL patients over the age of 35,[17,46] a question arises: Could plasma biomarkers be instrumental in predicting the onset of RCC in such patients? This could pave the way for early diagnosis and timely treatment, potentially enhancing both survival rates and quality of life.

In our study, we conducted LC-MS sequencing on plasma from VHL patients with RCC, comparing it against the peripheral blood of healthy subjects. We also analyzed RCC tumor tissues from VHL patients, using their matched AN tissues as a control. Our objective was to precisely identify metabolites uniquely linked to VHL patients with RCC, thereby providing deeper insights into the distinctive metabolic signature of the disease.

We identified key DAMs in plasma and tissue, linked primarily to arginine/proline metabolism, ammonic recycling, urea cycle, among others (Figure 4). These pathways are closely associated with the development and progression of RCC. Specifically, a general downregulation of genes involved in arginine and proline metabolism was noted in RCC, which is critical for tumor growth and survival.[47,48] Additionally, processes like ammonic recycling and the interconnected urea cycle, essential for nitrogen metabolism, play vital roles in RCC.[49] These mechanisms facilitate the repurposing of metabolic waste, like ammonia, for new amino acid synthesis, supporting the rapid proliferation of cancer cells in RCC.[50] Reflecting the robustness of our approach, the relative reliability of our sequencing results is further underscored by these findings.

Our analysis identified 23 common DAMs derived from both plasma and tissue sources (Figures 6A and 6B; Table S1). Some of these metabolites or their associated metabolic enzymes, such as gamma-glutamylalanine, decanoylcarnitine, LysoPC(16:0), indoxyl sulfate, L-asparagine, and L-malic acid, have been previously reported in the plasma or tissue of patients with sporadic RCC in various metabolomic

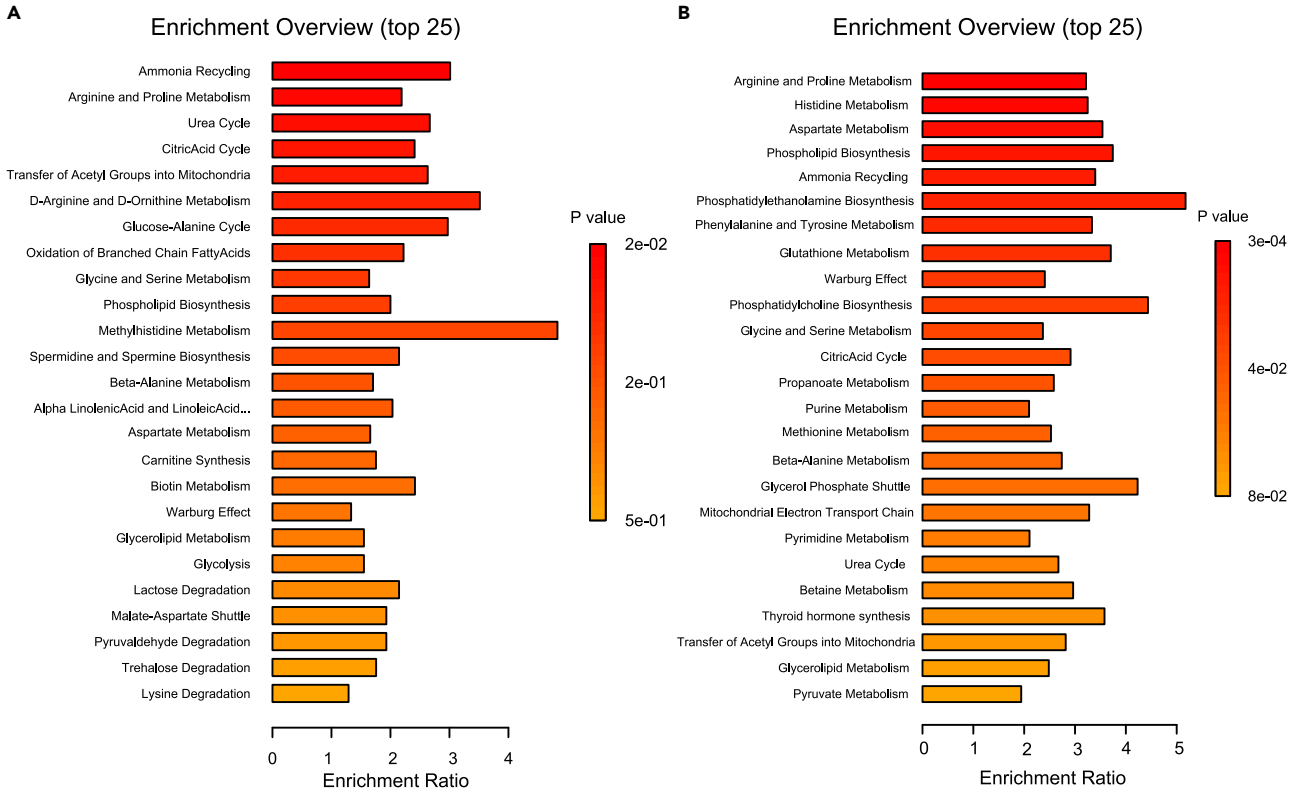


Figure 4. Pathway enrichment analysis of DAMs

(A) The top 25 enriched pathways of DAMs from plasma samples.

(B) The top 25 enriched pathways of DAMs from tissue samples. DAMs, differentially abundant metabolites.

studies or molecular mechanism research.[51–55] The concurrence of certain DAMs in our findings indirectly substantiates the reliability of our research. Additionally, the identification of DAMs not previously reported in sporadic RCC studies prompts us to consider whether these non-consistent DAMs could be indicative of the development of extra-kidney lesions. To maximize the focus on DAMs associated with VHL-RCC, our study design included a combined differential analysis of both plasma and tissue samples. While we cannot definitively state that the identified DAMs are solely due to significant differences caused by RCC, screening for common DAMs in both patient plasma and renal cancer tissues, especially in cases involving multi-organ lesions, offers a robust approach to predict the onset of RCC. Another indication that the non-consistent DAMs, remaining from the initial 23 common DAMs, are unrelated to the progression of lesions in other organs is derived from our analysis. We examined the correlation between all plasma-derived DAMs (not just the 23 common ones) and the timing of extrarenal lesion occurrence in these patients. Significant results are displayed in Figure 8, while non-significant findings are omitted. The results did not show any clear association between these non-consistent DAMs and the progression of lesions in other organs. However, these observations might also be attributed to variations in detection methods, tissue samples, and other heterogeneities, necessitating further validation in subsequent studies.

In our study, we identified 23 DAMs significantly associated with RCC, involved primarily in fatty acid, amino acid, and carbohydrate metabolism (Table S1). Specifically, Oleamide and (10E,12Z)-(9S)-9-hydroperoxyoctadeca-10,12-dienoic acid, belonging to fatty acyls, as well as PC(16:0/16:0) and PC(24:1(15Z)/18:4(6Z,9Z,12Z,15Z)) from glycerophospholipids, exhibited significant increases in both renal cancer tissues and plasma of VHL patients. PC(16:0/16:0), a key phosphatidylcholine in cell membranes,[56] demonstrated strong discriminative power with AUC values above 0.9 in our renal cancer study. It was notably elevated in Warthin tumor lymphoid stroma, metastases from colon and breast cancers, underscoring its potential importance in cancer pathophysiology.[57–59] The elevated levels of certain substances observed in both plasma and tissue may result from *VHL* mutations, which increase HIF activity in tumor cells. This heightened activity stimulates lipogenesis, involving essential fatty acids like PC(16:0/16:0) that are critical for membrane synthesis, signaling molecule production, and lipid storage, thus supporting cellular growth and proliferation.[40] As tumor cells ramp up lipid production, some of these lipids are released into the bloodstream, potentially explaining the concurrent rise in metabolic processes in both compartments.[60]

Additionally, certain glycerophospholipids such as subclasses of lysophosphatidylcholine (LysoPC) including LysoPC(16:1(9Z)/0:0) and LysoPC(16:0) have shown significant decreases in levels within both plasma and tissues. LysoPC is primarily found as a component of oxidatively modified low-density lipoprotein across various cell types. In normal cells, LPC may play a role in routine cellular processes, including cell signaling via binding to G protein-coupled receptors and Toll-like receptors. This interaction could contribute to foundational regulatory

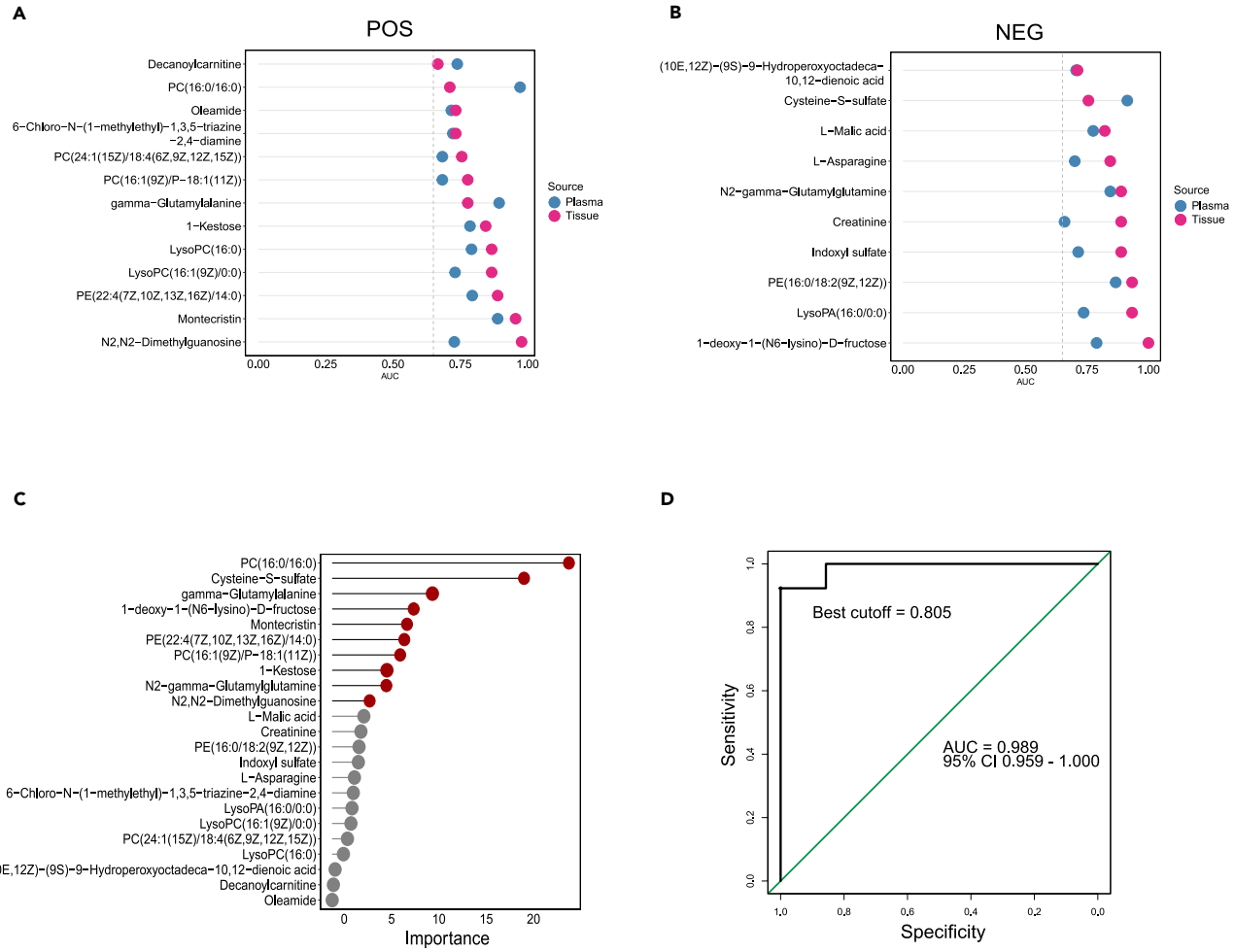


Figure 5. Construction of the diagnostic model for the patients with VHL-RCC disease

Lollipop plots display DAMs with AUC greater than 0.65 from plasma and tissue samples under POS (A) and NEG (B).

(C) The common 23 DAMs ranked by their contributions to classification accuracy using plasma cohort.

(D) ROC curve and AUC of the diagnostic model (generated using top 10 ranked DAMs) calculated by using the independent test set. VHL, von Hippel-Lindau; RCC, renal cell carcinoma; DAMs, differentially abundant metabolites; ROC, receiver operating characteristic; AUC, area under the curve; POS, positive ionization mode; NEG, negative ionization mode.

functions in immune surveillance and cellular communication.[61] Beyond our renal cancer research, reductions in Lysophosphatidylcholine (LysoPC) levels have been noted across several types of cancer, including lung, ovarian, laryngeal, and pancreatic cancers.[62–65] Research indicates that LysoPC may play a role in suppressing tumor initiation and progression.[62] The enzyme LysoPC acyltransferase (LPCAT), which catalyzes the transformation of LysoPC to phosphatidylcholine (PC), is reported to exhibit elevated expression in various malignancies.[66,67] Our findings (Table S1) reveal significant elevations in PC metabolites such as PC(16:0/16:0). This observation might explain the general decline in LysoPC levels in both plasma and tissue samples from our study. It also provides an alternative perspective on the potential pathways leading to the observed increase in PC products in both plasma and tissue levels in our study. However, considering the naturally high levels of LysoPC in normal cells, another plausible explanation for this decline could be a decrease in the proportion of normal cells within the tumor tissue.

We also noted significant declines in amino acids such as Cysteine-S-sulfate and L-Asparagine in both plasma and tissue samples. Cysteine-S-sulfate, a potent NMDA receptor agonist and a derivative of cysteine, is associated with neurological impairments due to its accumulation in rare sulfur amino acid metabolism disorders.[68,69] Research in oncology also reveals that both glioma and breast cancer patients have lower serum levels of cysteine and Cysteine-S-sulfate, with a more pronounced decrease observed in malignant cases compared to benign tumors.[70,71] This pattern is likely due to the dysregulation of amino acid metabolism by HIF under low oxygen conditions, which influence the expression of Cysteine dioxygenase 1.[72] This shift in enzyme activity directs cysteine metabolism toward cysteine sulfinic acid, thus inhibiting the production of Cysteine-S-sulfate. L-Asparagine, a non-essential amino acid, is synthesized from aspartate and glutamine under the catalysis of asparagine synthetase (ASNS). It is essential for protein biosynthesis and plays pivotal roles in numerous physiological

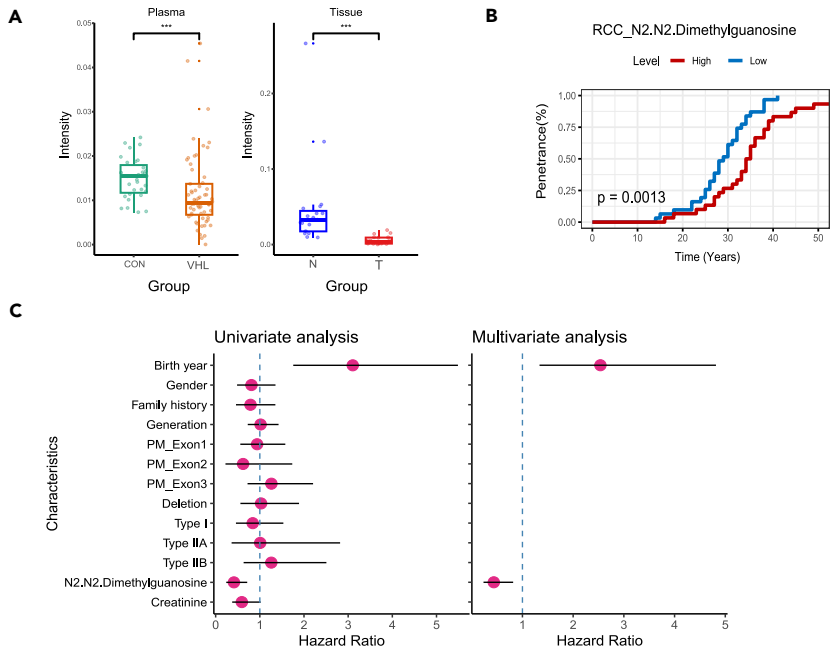


Figure 6. N2,N2-dimethylguanosine is a potential predictor for the onset of RCC in VHL patients

(A) The differential levels of N2,N2-dimethylguanosine in plasma and tissue samples.

(B) The comparison of age-related tumor risk in high and low levels of N2,N2-dimethylguanosine in plasma cohort.

(C) Univariate and multivariate Cox regression analysis of N2,N2-dimethylguanosine in plasma cohort.

processes.[73] Although studies indicate elevated ASNS expression in renal cancer cells theoretically suggesting increased tissue levels of L-Asparagine, our findings paradoxically show a decline.[54] This decrease could be due to tumor cells consuming large amounts of amino acids, including L-Asparagine, to support rapid proliferation. The heightened demand likely exceeds the production capacity of ASNS, leading to a net reduction in L-Asparagine levels.

Subsequently, we utilized the top 10 DAMs, based on their importance scores out of the 23 identified, to construct a predictive model for diagnosing VHL-RCC. The model demonstrated excellent performance, with an AUC of 1 for the training set and an AUC of 0.989 for the independent test set, indicating its strong capability to distinguish between healthy individuals and VHL patients with RCC. However, it's important to note that this diagnostic model does not reliably differentiate between VHL-RCC and sporadic RCC. To address this, our future work will involve expanding the sample size and conducting metabolomic sequencing under the same standards for these two types of cancer.

Among these 23 DAMs, our attention was particularly drawn to the clinical significance of N2,N2-dimethylguanosine. N2,N2-dimethylguanosine, a noteworthy modified nucleoside present in transfer RNA (tRNA) and ribosomal RNA (rRNA), serves as an essential element of the metabolome and has gained prominence as a crucial biomarker in diverse health contexts. Studies have linked elevated levels of N2,N2-dimethylguanosine to a nearly 2-fold increased risk of type 2 diabetes,[74] indicated its role in distinguishing acute respiratory distress syndrome in pneumonia patients,[75] and shown its utility in estimating RNA turnover rates, particularly in preterm infants.[76] Furthermore, in pulmonary arterial hypertension, its alteration correlates with disease severity and patient outcomes.[77] These findings underscore the importance of N2,N2-dimethylguanosine in metabolomics research, highlighting its potential as a diagnostic and prognostic tool across multiple medical conditions. However, its exploration in the context of cancer remains unreported. In our study, we discovered that N2,N2-dimethylguanosine, which was notably reduced in both plasma and tissue of VHL patients, not only served as a biomarker to distinguish VHL-associated RCC patients from healthy controls (AUC = 0.728, Table S1) but also independently predicted the earlier onset of VHL-associated RCC (Figure 7C), which suggests that lower levels of N2,N2-dimethylguanosine might be associated with an earlier development of VHL-associated RCC. This highlights its potential as a crucial early diagnostic and therapeutic target in managing this form of cancer.

In our team's previous research, we discovered a correlation between different mutation locations in the *VHL* gene and the affected organs in VHL patients.[78] Additionally, various mutation types in this gene were found to be associated with different prognoses among patients.[79] Interestingly, in this study, we unexpectedly observed a connection between different mutation locations in the *VHL* gene and the levels of certain DAMs in plasma (Figures 8, S3). Notably, distinct mutation locations resulted in variations in the same metabolite levels. For instance, the level of Cysteine-S-sulfate was reduced in the PM-Exon1 group but elevated in the PM-Exon2 group. These findings prompt the question of whether the differential functionality of the VHL protein caused by mutations at different sites or mutation types may impact the manifestation of organ involvement in patients through their influence on specific metabolic pathways. These preliminary discoveries provide further impetus for exploring the mechanisms behind genotype-phenotype correlations in previous research.

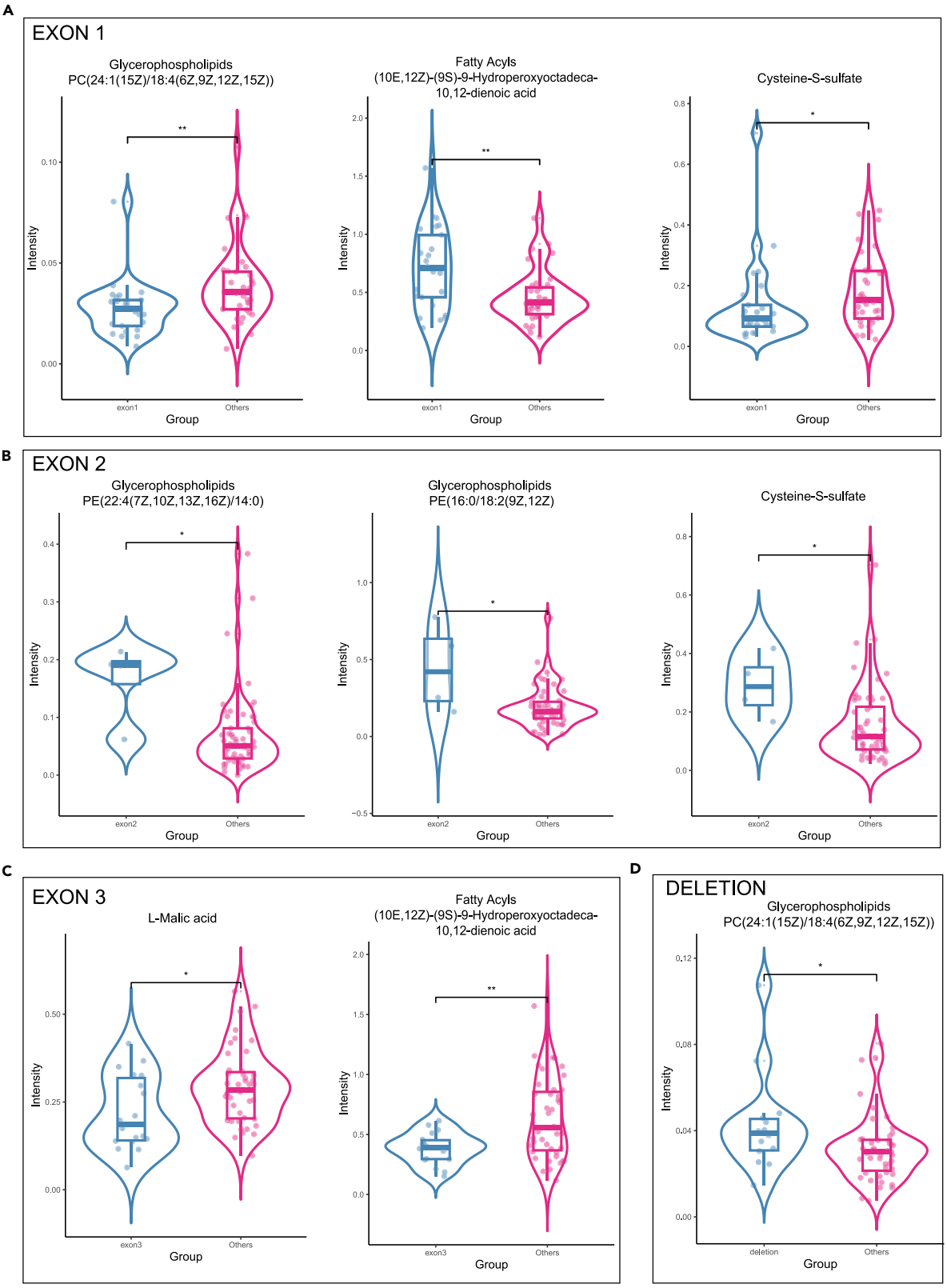

A
EXON 1
Glycerophospholipids
PC(24:1(15Z)/18:4(6Z,9Z,12Z,15Z))
Fatty Acyls
(10E,12Z)-(9S)-9-Hydroperoxyoctadeca-
10,12-dienoic acid
Cysteine-S-sulfate
Intensity
Group
exon1
Others
B
EXON 2
Glycerophospholipids
PE(22:4(7Z,10Z,13Z,16Z)/14:0)
Glycerophospholipids
PE(16:0/18:2(9Z,12Z)
Cysteine-S-sulfate
exon2
C
EXON 3
L-Malic acid
Fatty Acyls
(10E,12Z)-(9S)-9-Hydroperoxyoctadeca-
10,12-dienoic acid
exon3
D
DELETION
Glycerophospholipids
PC(24:1(15Z)/18:4(6Z,9Z,12Z,15Z))
deletion

Figure 7. The relationship between VHL gene mutation sites or types and the levels of DAMs in the plasma of VHL patients

(A) The mutation site on exon 1.

(B) The mutation site on exon 2.

(C) The mutation site on exon 3.

(D) The mutation type is deletion (including partial and complete deletions). Mann-Whitney tests were conducted to assess the significance of differences. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Overall, VHL-RCC is often underdiagnosed or misdiagnosed due to its rarity. Our study, analyzing the largest cohort of VHL patients to date using untargeted metabolomics sequencing of blood plasma and tissues, has developed a highly accurate diagnostic prediction model. This model identifies specific metabolic biomarkers for VHL-RCC, demonstrating exceptional performance in an independent test set and significantly improving diagnosis rates. It also links metabolite variations to specific *VHL* gene mutations, with N2,N2-dimethylguanosine pinpointed as a crucial predictor for RCC onset. Monitoring this metabolite could substantially enhance patient care and outcomes, offering a profound insight into the metabolic intricacies of VHL-RCC. This research not only illuminates the path for better management of this rare cancer but also sets a new standard in the metabolomic study of VHL-RCC.

Limitations of the study

The study's main limitation lies in its small sample size. Upcoming research will involve a larger cohort, including more VHL patients and those with sporadic RCC, to better delineate their metabolic profiles. This will enhance diagnostic and prognostic capabilities for VHL RCC. Another limitation is that the study has not yet conducted experiments to delve into the relationship between metabolites and RCC progression. Future work could use primary renal cancer cells from VHL patients for a more precise evaluation of the potential link between these metabolites and RCC development.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY

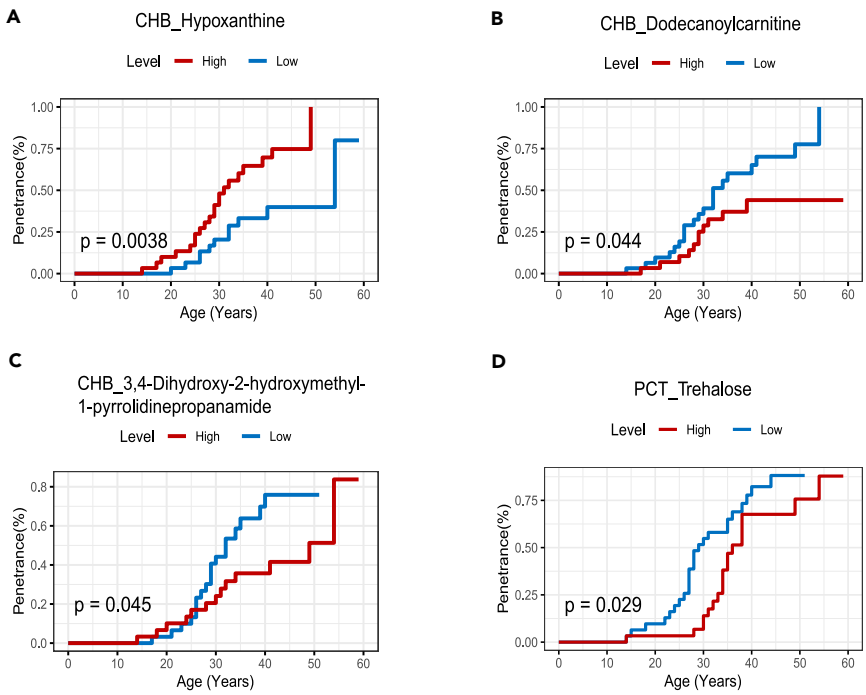


Figure 8. Comparison of age-related tumor risk by high and low levels of DAM in patients with VHL disease

(A) Hypoxanthine levels in CHB.

(B) Dodecanoylcarnitine levels in CHB.

(C) 3,4-Dihydroxy-2-hydroxymethyl-1-pyrrolidinepropanamide levels in CHB.

(D) Trehalose levels in PCT. CHB, central nervous system hemangioblastomas; PCT, pancreatic cysts or tumors; VHL, von Hippel-Lindau.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110357.

ACKNOWLEDGMENTS

We thank all the volunteers who participated in this study. We are also grateful to Shanghai Biotree Biomedical Biotechnology Co., Ltd., for their assistance with metabolomic sequencing. This work was supported by the National Natural Science Foundation of China (No. 82141103), Sino-Russian Mathematics Center, National High Level Hospital Clinical Research Funding (High Quality Clinical Research Project of Peking University First Hospital, 2022CR75), National Natural Science Foundation of China (Nos. 82172617, 82172665, 81872081, 82103153), Beijing Natural Science Foundation of Beijing (No. 7232176; QY23068), and Capital's Funds for Health Improvement and Research (2022-2-4074).

AUTHOR CONTRIBUTIONS

K.G. and Y.G. conceived the project. Z.Z., W.Y., and Y.W. analyzed the metabolomic sequencing data. W.Y., C.W., T.L., C.H., and X.C. collected clinical samples and detailed information of all subjects. Z.Z. drafted the manuscript. X.Z. provided guidance on data analysis. Y.X., J.Z., and Y.W. refined the manuscript. K.G. and Y.G. provided clinical expertise. K.G., Y.G., and X.Z. supervised the work and established the workflow of the analysis. All authors have read and agreed to the final version of the manuscript.

DECLARATION OF INTERESTS

We declare there are no any competing financial interests in relation to this work.

Received: February 21, 2024
Revised: May 10, 2024
Accepted: June 20, 2024
Published: June 22, 2024

REFERENCES

1. Maher, E.R., Neumann, H.P., and Richard, S. (2011). von Hippel-Lindau disease: a clinical and scientific review. Eur. J. Hum. Genet. *19*, 617–623. https://doi.org/10.1038/ejhg.2010.175.
2. Kaelin, W.G. (2007). Von hippel-lindau disease. Annu. Rev. Pathol. *2*, 145–173. https://doi.org/10.1146/annurev.pathol.2.010506.092049.
3. Gossage, L., Eisen, T., and Maher, E.R. (2015). VHL, the story of a tumour suppressor gene. Nat. Rev. Cancer *15*, 55–64. https://doi.org/10.1038/nrc3844.
4. Molino, D., Sepe, J., Anastasio, P., and De Santo, N.G. (2006). The history of von Hippel-Lindau disease. J. Nephrol. *19*, S119–S123.
5. Peng, S., Shepard, M.J., Wang, J., Li, T., Ning, X., Cai, L., Zhuang, Z., and Gong, K. (2017). Genotype-phenotype correlations in Chinese von Hippel-Lindau disease patients. Oncotarget *8*, 38456–38465. https://doi.org/10.18632/oncotarget.16594.
6. McNeill, A., Rattenberry, E., Barber, R., Killick, P., MacDonald, F., and Maher, E.R. (2009). Genotype–phenotype correlations in VHL exon deletions. Am. J. Med. Genet. *149A*, 2147–2151. https://doi.org/10.1002/ajmg.a.33023.
7. Barontini, M., and Dahia, P.L.M. (2010). VHL disease. Best Pract. Res. Clin. Endocrinol. Metab. *24*, 401–413. https://doi.org/10.1016/j.beem.2010.01.002.
8. Nordstrom-O'Brien, M., van der Luijt, R.B., van Rooijen, E., van den Ouweland, A.M., Majoor-Krakauer, D.F., Lolkema, M.P., van Brussel, A., Voest, E.E., and Giles, R.H. (2010). Genetic analysis of von Hippel-Lindau disease. Hum. Mutat. *31*, 521–537. https://doi.org/10.1002/humu.21219.
9. Cascón, A., Escobar, B., Montero-Conde, C., Rodríguez-Antona, C., Ruiz-Llorente, S., Osorio, A., Mercadillo, F., Letón, R., Campos, J.M., García-Sagredo, J.M., Benítez, J., et al. (2007). Loss of the actin regulator HSPC300 results in clear cell renal cell carcinoma protection in Von Hippel-Lindau patients. Hum. Mutat. *28*, 613–621. https://doi.org/10.1002/humu.20496.
10. Bausch, B., Jilg, C., Gläsker, S., Vortmeyer, A., Lützen, N., Anton, A., Eng, C., and Neumann, H.P.H. (2013). Renal cancer in von Hippel-Lindau disease and related syndromes. Nat. Rev. Nephrol. *9*, 529–538. https://doi.org/10.1038/nrneph.2013.144.
11. Ploussard, G., Droupy, S., Ferlicot, S., Ples, R., Rocher, L., Richard, S., and Benoit, G. (2007). Local recurrence after nephron-sparing surgery in von Hippel-Lindau disease. Urology *70*, 435–439. https://doi.org/10.1016/j.urology.2007.04.040.
12. Carlo, M.I., Hakimi, A.A., Stewart, G.D., Bratslavsky, G., Brugarolas, J., Chen, Y.-B., Linehan, W.M., Maher, E.R., Merino, M.J., Offit, K., et al. (2019). Familial kidney cancer: implications of new syndromes and molecular insights. Eur. Urol. *76*, 754–764. https://doi.org/10.1016/j.eururo.2019.06.015.

13. Crespigio, J., Berbel, L.C.L., Dias, M.A., Berbel, R.F., Pereira, S.S., Pignatelli, D., and Mazzuco, T.L. (2018). Von Hippel–Lindau disease: a single gene, several hereditary tumors. J. Endocrinol. Invest. *41*, 21–31. https://doi.org/10.1007/s40618-017-0683-1.
14. Duffey, B.G., Choyke, P.L., Glenn, G., Grubb, R.L., Venzon, D., Linehan, W.M., and Walther, M.M. (2004). The relationship between renal tumor size and metastases in patients with von Hippel-Lindau disease. J. Urol. *172*, 63–65. https://doi.org/10.1097/01.ju.0000132127.79974.3f.
15. Jilg, C.A., Neumann, H.P.H., Gläsker, S., Schäfer, O., Leiber, C., Ardelt, P.U., Schwardt, M., and Schultze-Seemann, W. (2012). Nephron sparing surgery in von Hippel-Lindau associated renal cell carcinoma; clinicopathological long-term follow-up. Fam. Cancer *11*, 387–394. https://doi.org/10.1007/s10689-012-9525-7.
16. Kim, E., and Zschiedrich, S. (2018). Renal cell carcinoma in von hippel–lindau disease—from tumor genetics to novel therapeutic strategies. Front. Pediatr. *6*, 16. https://doi.org/10.3389/fped.2018.00016.
17. Wang, J.Y., Peng, S.H., Li, T., Ning, X.H., Liu, S.J., Hong, B.A., Liu, J.Y., Wu, P.J., Zhou, B.W., Zhou, J.C., et al. (2018). Risk factors for survival in patients with von Hippel-Lindau disease. J. Med. Genet. *55*, 322–328. https://doi.org/10.1136/jmedgenet-2017-104995.
18. di Meo, N.A., Lasorsa, F., Rutigliano, M., Loizzo, D., Ferro, M., Stella, A., Bizzoca, C., Vincenti, L., Pandolfo, S.D., Autorino, R., et al. (2022). Renal cell carcinoma as a metabolic disease: an update on main pathways, potential biomarkers, and therapeutic targets. Int. J. Mol. Sci. *23*, 14360. https://doi.org/10.3390/ijms232214360.
19. Yuk, H.D., Hwang, E.C., Park, J.Y., Jeong, C.W., Song, C., Seo, S.I., Byun, S.-S., Kwak, C., Hong, S.-H., Kang, M., et al. (2020). The number of metabolic features as a significant prognostic factor in patients with metastatic renal cell carcinoma. Sci. Rep. *10*, 6967. https://doi.org/10.1038/s41598-020-63816-9.
20. Gupta, A., Nath, K., Bansal, N., and Kumar, M. (2020). Role of metabolomics-derived biomarkers to identify renal cell carcinoma: a comprehensive perspective of the past ten years and advancements. Expert Rev. Mol. Diagn. *20*, 5–18. https://doi.org/10.1080/14737159.2020.1704259.
21. Johnson, C.H., Ivanisevic, J., and Siuzdak, G. (2016). Metabolomics: beyond biomarkers and towards mechanisms. Nat. Rev. Mol. Cell Biol. *17*, 451–459. https://doi.org/10.1038/nrm.2016.25.
22. Rinschen, M.M., Ivanisevic, J., Giera, M., and Siuzdak, G. (2019). Identification of bioactive metabolites using activity metabolomics. Nat. Rev. Mol. Cell Biol. *20*, 353–367. https://doi.org/10.1038/s41580-019-0108-4.
23. Liu, X., Zhang, M., Liu, X., Sun, H., Guo, Z., Tang, X., Wang, Z., Li, J., He, L., Zhang, W., et al. (2020). Investigation of plasma metabolic and lipidomic characteristics of a Chinese cohort and a pilot study of renal cell carcinoma biomarker. Front. Oncol. *10*, 1507. https://doi.org/10.3389/fonc.2020.01507.
24. Maslov, D.L., Trifonova, O.P., Lichtenberg, S., Balashova, E.E., Mamedli, Z.Z., Alferov, A.A., Stilidi, I.S., Lokhov, P.G., Kushlinskii, N.E., and Archakov, A.I. (2022). Blood plasma metabolome profiling at different stages of renal cell carcinoma. Cancers *15*, 140. https://doi.org/10.3390/cancers15010140.
25. Schmidt, D.R., Patel, R., Kirsch, D.G., Lewis, C.A., Vander Heiden, M.G., and Locasale, J.W. (2021). Metabolomics in cancer research and emerging applications in clinical oncology. CA A Cancer J. Clin. *71*, 333–358. https://doi.org/10.3322/caac.21670.
26. Eisenhofer, G. (2012). Screening for pheochromocytomas and paragangliomas. Curr. Hypertens. Rep. *14*, 130–137. https://doi.org/10.1007/s11906-012-0246-y.
27. Fahrmann, J.F., Bantis, L.E., Capello, M., Scelo, G., Dennison, J.B., Patel, N., Murage, E., Vykoukal, J., Kundnani, D.L., Foretova, L., et al. (2019). A plasma-derived protein-metabolite multiplexed panel for early-stage pancreatic cancer. J. Natl. Cancer Inst. *111*, 372–379. https://doi.org/10.1093/jnci/djy126.
28. Zeleznik, O.A., Eliassen, A.H., Kraft, P., Poole, E.M., Rosner, B.A., Jeanfavre, S., Deik, A.A., Bullock, K., Hitchcock, D.S., Avila-Pacheco, J., et al. (2020). A prospective analysis of circulating plasma metabolites associated with ovarian cancer risk. Cancer Res. *80*, 1357–1367. https://doi.org/10.1158/0008-5472.CAN-19-2567.
29. Pang, Z., Chong, J., Zhou, G., de Lima Morais, D.A., Chang, L., Barrette, M., Gauthier, C., Jacques, P.É., Li, S., and Xia, J. (2021). MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res. *49*, W388–w396. https://doi.org/10.1093/nar/gkab382.
30. Liu, S.J., Wang, J.Y., Peng, S.H., Li, T., Ning, X.H., Hong, B.A., Liu, J.Y., Wu, P.J., Zhou, B.W., Zhou, J.C., et al. (2018). Genotype and phenotype correlation in von Hippel-Lindau disease based on alteration of the HIF-alpha binding site in VHL protein. Genet. Med. *20*, 1266–1273. https://doi.org/10.1038/gim.2017.261.
31. Iacovelli, R., Arduini, D., Ciccarese, C., Pierconti, F., Strusi, A., Piro, G., Carbone, C., Foschi, N., Daniele, G., and Tortora, G. (2022). Targeting hypoxia-inducible factor pathways in sporadic and Von Hippel-Lindau syndrome-related kidney cancers. Crit. Rev. Oncol. Hematol. *176*, 103750. https://doi.org/10.1016/j.critrevonc.2022.103750.
32. Poulsen, M.L.M., Budtz-Jørgensen, E., and Bisgaard, M.L. (2010). Surveillance in von Hippel-Lindau disease (vHL). Clin. Genet. *77*, 49–59. https://doi.org/10.1111/j.1399-0004.2009.01281.x.
33. Maher, E.R., Yates, J.R., Harries, R., Benjamin, C., Harris, R., Moore, A.T., and Ferguson-Smith, M.A. (1990). Clinical features and natural history of von Hippel-Lindau disease. Q. J. Med. *77*, 1151–1163. https://doi.org/10.1093/qjmed/77.2.1151.
34. Zhang, J., Pan, J.H., Dong, B.J., Xue, W., Liu, D.M., and Huang, Y.R. (2012). Active surveillance of renal masses in von Hippel-Lindau disease: growth rates and clinical outcome over a median follow-up period of 56 months. Fam. Cancer *11*, 209–214. https://doi.org/10.1007/s10689-011-9503-5.
35. Lucarelli, G., Loizzo, D., Franzin, R., Battaglia, S., Ferro, M., Cantiello, F., Castellano, G., Bettocchi, C., Ditonno, P., and Battaglia, M. (2019). Metabolomic insights into pathophysiological mechanisms and biomarker discovery in clear cell renal cell carcinoma. Expert Rev. Mol. Diagn. *19*, 397–407. https://doi.org/10.1080/14737159.2019.1607729.
36. Chappell, J.C., Payne, L.B., and Rathmell, W.K. (2019). Hypoxia, angiogenesis, and metabolism in the hereditary kidney cancers. J. Clin. Invest. *129*, 442–451. https://doi.org/10.1172/jci120855.
37. Kang, B., Camps, J., Fan, B., Jiang, H., Ibrahim, M.M., Hu, X., Qin, S., Kirchhoff, D., Chiang, D.Y., Wang, S., et al. (2022). Parallel single-cell and bulk transcriptome analyses reveal key features of the gastric tumor microenvironment. Genome Biol. *23*, 265. https://doi.org/10.1186/s13059-022-02828-2.
38. Kaelin, W.G., Jr. (2004). The von Hippel-Lindau tumor suppressor gene and kidney cancer. Clin. Cancer Res. *10*, 6290S-5S. https://doi.org/10.1158/1078-0432.CCR-sup-040025.
39. Schödel, J., Grampp, S., Maher, E.R., Moch, H., Ratcliffe, P.J., Russo, P., and Mole, D.R. (2016). Hypoxia, Hypoxia-inducible Transcription Factors, and Renal Cancer. Eur. Urol. *69*, 646–657. https://doi.org/10.1016/j.eururo.2015.08.007.
40. Taylor, C.T., and Scholz, C.C. (2022). The effect of HIF on metabolism and immunity. Nat. Rev. Nephrol. *18*, 573–587. https://doi.org/10.1038/s41581-022-00587-8.
41. Silva, C., Perestrelo, R., Silva, P., Tomás, H., and Câmara, J.S. (2019). Breast Cancer Metabolomics: From Analytical Platforms to Multivariate Data Analysis. A Review. Metabolites *9*, 102. https://doi.org/10.3390/metabo9050102.
42. Plewa, S., Horała, A., Dereziński, P., Nowak-Markwitz, E., Matysiak, J., and Kokot, Z.J. (2019). Wide spectrum targeted metabolomics identifies potential ovarian cancer biomarkers. Life Sci. *222*, 235–244. https://doi.org/10.1016/j.lfs.2019.03.004.
43. Jing, L., Guigonis, J.-M., Borchiellini, D., Durand, M., Pourcher, T., and Ambrosetti, D. (2019). LC-MS based metabolomic profiling for renal cell carcinoma histologic subtypes. Sci. Rep. *9*, 15635. https://doi.org/10.1038/s41598-019-52059-y.
44. Falegan, O.S., Ball, M.W., Shaykhutdinov, R.A., Pieroraio, P.M., Farshidfar, F., Vogel, H.J., Allaf, M.E., and Hyndman, M.E. (2017). Urine and serum metabolomics analyses may distinguish between stages of renal cell carcinoma. Metabolites *7*, 6. https://doi.org/10.3390/metabo7010006.
45. Sato, T., Kawasaki, Y., Maekawa, M., Takasaki, S., Shimada, S., Morozumi, K., Sato, M., Kawamorita, N., Yamashita, S., Mitsuzuka, K., et al. (2020). Accurate quantification of urinary metabolites for predictive models manifest clinicopathology of renal cell carcinoma. Cancer Sci. *111*, 2570–2578. https://doi.org/10.1111/cas.14440.
46. Peng, X., Chen, J., Wang, J., Peng, S., Liu, S., Ma, K., Zhou, J., Hong, B., Zhou, B., Zhang, J., et al. (2019). Natural history of renal tumours in von Hippel-Lindau disease: a large retrospective study of Chinese patients. J. Med. Genet. *56*, 380–387. https://doi.org/10.1136/jmedgenet-2018-105567.
47. Perroud, B., Lee, J., Valkova, N., Dhirapong, A., Lin, P.Y., Fiehn, O., Kültz, D., and Weiss, R.H. (2006). Pathway analysis of kidney cancer using proteomics and metabolic profiling. Mol. Cancer *5*, 64. https://doi.org/10.1186/1476-4598-5-64.
48. Pandey, N., Lanke, V., and Vinod, P.K. (2020). Network-based metabolic characterization of renal cell carcinoma. Sci. Rep. *10*, 5955. https://doi.org/10.1038/s41598-020-62853-8.
49. Kaushik, A.K., Tarangelo, A., Boroughs, L.K., Ragavan, M., Zhang, Y., Wu, C.Y., Li, X., Ahumada, K., Chiang, J.C., Tcheuyap, V.T., et al. (2022). In vivo characterization of glutamine metabolism identifies therapeutic

targets in clear cell renal cell carcinoma. Sci. Adv. *8*, eabp8293. https://doi.org/10.1126/sciadv.abp8293.

50. Spinelli, J.B., Yoon, H., Ringel, A.E., Jeanfavre, S., Clish, C.B., and Haigis, M.C. (2017). Metabolic recycling of ammonia via glutamate dehydrogenase supports breast cancer biomass. Science *358*, 941–946. https://doi.org/10.1126/science.aam9305.
51. Priolo, C., Khabibullin, D., Reznik, E., Filippakis, H., Ogórek, B., Kavanagh, T.R., Nijmeh, J., Herbert, Z.T., Asara, J.M., Kwiatkowski, D.J., et al. (2018). Impairment of gamma-glutamyl transferase 1 activity in the metabolic pathogenesis of chromophobe renal cell carcinoma. Proc. Natl. Acad. Sci. USA *115*, E6274–e6282. https://doi.org/10.1073/pnas.1710849115.
52. Nizioł, J., Bonifay, V., Ossoliński, K., Ossoliński, T., Ossolińska, A., Sunner, J., Beech, I., Arendowski, A., and Ruman, T. (2018). Metabolomic study of human tissue and urine in clear cell renal carcinoma by LC-HRMS and PLS-DA. Anal. Bioanal. Chem. *410*, 3859–3869. https://doi.org/10.1007/s00216-018-1059-x.
53. Du, Y., Wang, Q., Zhang, X., Wang, X., Qin, C., Sheng, Z., Yin, H., Jiang, C., Li, J., and Xu, T. (2017). Lysophosphatidylcholine acyltransferase 1 upregulation and concomitant phospholipid alterations in clear cell renal cell carcinoma. J. Exp. Clin. Cancer Res. *36*, 66. https://doi.org/10.1186/s13046-017-0525-1.
54. Gan, X., Liu, R., Cheng, H., Mao, W., Feng, N., and Chen, M. (2022). ASNS can predict the poor prognosis of clear cell renal cell carcinoma. Front. Oncol. *12*, 882888. https://doi.org/10.3389/fonc.2022.882888.
55. Caboni, P. (2018). Preliminary Metabolomic Study of Urine Samples in Patients Affected by Renal Clear Cell Cancer by GC-MS. Int. J. Clin. Urol. *2*, 1–5.
56. Van Meer, G., Voelker, D.R., and Feigenson, G.W. (2008). Membrane lipids: where they are and how they behave. Nat. Rev. Mol. Cell Biol. *9*, 112–124. https://doi.org/10.1038/nrm2330.
57. He, Q., Takizawa, Y., Hayasaka, T., Masaki, N., Kusama, Y., Su, J., Mineta, H., and Setou, M. (2014). Increased phosphatidylcholine (16: 0/16: 0) in the folliculus lymphaticus of Warthin tumor. Anal. Bioanal. Chem. *406*, 5815–5825. https://doi.org/10.1007/s00216-014-7890-9.
58. Shimma, S., Sugiura, Y., Hayasaka, T., Hoshikawa, Y., Noda, T., and Setou, M. (2007). MALDI-based imaging mass spectrometry revealed abnormal distribution of phospholipids in colon cancer liver metastasis. J. Chromatogr. B *855*, 98–103. https://doi.org/10.1016/j.jchromb.2007.02.037.
59. Chughtai, K., Jiang, L., Greenwood, T.R., Glunde, K., and Heeren, R.M.A. (2013). Mass spectrometry images acylcarnitines, phosphatidylcholines, and sphingomyelin in MDA-MB-231 breast tumor models [S]. J. Lipid Res. *54*, 333–344. https://doi.org/10.1194/jlr.M027961.
60. Jeong, D.-W., Lee, S., and Chun, Y.-S. (2021). How cancer cells remodel lipid metabolism: strategies targeting transcription factors. Lipids Health Dis. *20*, 163. https://doi.org/10.1186/s12944-021-01593-8.
61. Liu, P., Zhu, W., Chen, C., Yan, B., Zhu, L., Chen, X., and Peng, C. (2020). The mechanisms of lysophosphatidylcholine in the development of diseases. Life Sci. *247*, 117443. https://doi.org/10.1016/j.lfs.2020.117443.
62. Dong, J., Cai, X., Zhao, L., Xue, X., Zou, L., Zhang, X., and Liang, X. (2010). Lysophosphatidylcholine profiling of plasma: discrimination of isomers and discovery of lung cancer biomarkers. Metabolomics *6*, 478–488.
63. Kim, S.C., Kim, M.K., Kim, Y.H., Ahn, S.A., Kim, K.H., Kim, K., Kim, W.K., Lee, J.H., Cho, J.Y., and Yoo, B.C. (2014). Differential levels of L-homocysteic acid and lysophosphatidylcholine (16: 0) in sera of patients with ovarian cancer. Oncol. Lett. *8*, 566–574. https://doi.org/10.3892/ol.2014.2214.
64. Zhang, X., Hou, H., Chen, H., Liu, Y., Wang, A., and Hu, Q. (2018). Serum metabolomics of laryngeal cancer based on liquid chromatography coupled with quadrupole time-of-flight mass spectrometry. Biomed. Chromatogr. *32*, e4181. https://doi.org/10.1002/bmc.4181.
65. Xu, H., Zhang, L., Kang, H., Liu, J., Zhang, J., Zhao, J., and Liu, S. (2021). Metabolomics identifies biomarker signatures to differentiate pancreatic cancer from type 2 diabetes mellitus in early diagnosis. Internet J. Endocrinol. *2021*, 9990768. https://doi.org/10.1155/2021/9990768.
66. Nakanishi, H., Shindou, H., Hishikawa, D., Harayama, T., Ogasawara, R., Suwabe, A., Taguchi, R., and Shimizu, T. (2006). Cloning and characterization of mouse lung-type acyl-CoA: lysophosphatidylcholine acyltransferase 1 (LPCAT1): expression in alveolar type II cells and possible involvement in surfactant production. J. Biol. Chem. *281*, 20140–20147. https://doi.org/10.1074/jbc.M600225200.
67. Lin, C. (2016). Recent Advances of Relationship Between miRNA-related SNPs and Lung Cancer Susceptibility. Cancer Research on Prevention and Treatment *43*, 1090–1094.
68. Nguyen, M., Le Mignon, M., Schnellbächer, A., Wehsling, M., Braun, J., Baumgaertner, J., Grabner, M., and Zimmer, A. (2023). Mechanistic insights into the biological activity of S-Sulfocysteine in CHO cells using a multi-omics approach. Front. Bioeng. Biotechnol. *11*, 1230422. https://doi.org/10.3389/fbioe.2023.1230422.
69. Zecchini, M., Lucas, R., and Le Gresley, A. (2019). New insights into the cystine-sulfite reaction. Molecules *24*, 2377. https://doi.org/10.3390/molecules24132377.
70. Huang, J., Weinstein, S.J., Kitahara, C.M., Karoly, E.D., Sampson, J.N., and Albanes, D. (2017). A prospective study of serum metabolites and glioma risk. Oncotarget *8*, 70366–70377. https://doi.org/10.18632/oncotarget.19705.
71. Liu, J., Zhou, Y., Liu, H., Ma, M., Wang, F., Liu, C., Yuan, Q., Wang, H., Hou, X., and Yin, P. (2022). Metabolic reprogramming enables the auxiliary diagnosis of breast cancer by automated breast volume scanner. Front. Oncol. *12*, 939606. https://doi.org/10.3389/fonc.2022.939606.
72. Warnhoff, K., Bhattacharya, S., Snoozy, J., Breen, P.C., and Ruvkun, G. (2024). Hypoxia-inducible factor induces cysteine dioxygenase and promotes cysteine homeostasis in Caenorhabditis elegans. Elife *12*, RP89173. https://doi.org/10.7554/eLife.89173.
73. Chiu, M., Taurino, G., Bianchi, M.G., Kilberg, M.S., and Bussolati, O. (2019). Asparagine synthetase in cancer: beyond acute lymphoblastic leukemia. Front. Oncol. *9*, 1480. https://doi.org/10.3389/fonc.2019.01480.
74. Ottosson, F., Smith, E., Gallo, W., Fernandez, C., and Melander, O. (2019). Purine metabolites and carnitine biosynthesis intermediates are biomarkers for incident type 2 diabetes. J. Clin. Endocrinol. Metab. *104*, 4921–4930. https://doi.org/10.1210/jc.2019-00822.
75. Xiong, F., Jiang, K., Chen, J., Yan, Y., Zhou, Y., Chen, Z., Zheng, H., Li, Y., and Gao, H. (2023). Metabolomics Study Revealing Purines as Potential Diagnostic Biomarkers of Acute Respiratory Distress Syndrome in Patients with Community– Acquired Pneumonia. J. Proteome Res. *22*, 2558–2569. https://doi.org/10.1021/acs.jproteome.2c00788.
76. Sander, G., Hülsemann, J., Topp, H., Heller-Schöch, G., and Schöch, G. (1986). Protein and RNA turnover in preterm infants and adults: a comparison based on urinary excretion of 3-methylhistidine and of modified one-way RNA catabolites. Ann. Nutr. Metab. *30*, 137–142. https://doi.org/10.1159/000177186.
77. Rhodes, C.J., Ghataorhe, P., Wharton, J., Rue-Albrecht, K.C., Hadinnapola, C., Watson, G., Bleda, M., Haimel, M., Coghlan, G., Corris, P.A., et al. (2017). Plasma metabolomics implicates modified transfer RNAs and altered bioenergetics in the outcomes of pulmonary arterial hypertension. Circulation *135*, 460–475. https://doi.org/10.1161/CIRCULATIONAHA.116.024602.
78. Hong, B., Ma, K., Zhou, J., Zhang, J., Wang, J., Liu, S., Zhang, Z., Cai, L., Zhang, N., and Gong, K. (2019). Frequent Mutations of VHL Gene and the Clinical Phenotypes in the Largest Chinese Cohort With Von Hippel-Lindau Disease. Front. Genet. *10*, 867. https://doi.org/10.3389/fgene.2019.00867.
79. Xie, H., Ma, K., Zhang, J., Hong, B., Zhou, J., Li, L., Zhang, K., Gong, K., and Cai, L. (2020). Novel genetic characterisation and phenotype correlation in von Hippel-Lindau (VHL) disease based on the Elongin C binding site: a large retrospective study. J. Med. Genet. *57*, 744–751. https://doi.org/10.1136/jmedgenet-2019-106336.
80. Dunn, W.B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J.D., Halsall, A., Haselden, J.N., et al. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat. Protoc. *6*, 1060–1083. https://doi.org/10.1038/nprot.2011.335.
81. Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A., Dong, L., Zhang, Z., Yu, C., Sun, Y., et al. (2021). The Genome Sequence Archive Family: Toward Explosive Data Growth and Diverse Data Types. Dev. Reprod. Biol. *19*, 578–583. https://doi.org/10.1016/j.gpb.2021.08.001.
82. CNCB-NGDC Members and Partners (2024). Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2024. Nucleic Acids Res. *52*, D18–d32. https://doi.org/10.1093/nar/gkad1078.
83. Wang, X., Zhang, N., Ning, X., Li, T., Wu, P., Peng, S., Fan, Y., Bu, D., and Gong, K. (2014). Higher prevalence of novel mutations in VHL gene in Chinese von Hippel-Lindau disease

patients. Urology *83*, 675.e1–675.e6755. https://doi.org/10.1016/j.urology.2013.09.069.

84. Wang, J., Zhang, T., Shen, X., Liu, J., Zhao, D., Sun, Y., Wang, L., Liu, Y., Gong, X., Liu, Y., et al. (2016). Serum metabolomics for early diagnosis of esophageal squamous cell carcinoma by UHPLC-QTOF/MS. Metabolomics *12*, 1–10.

85. Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal. Chem. *78*, 779–787. https://doi.org/10.1021/ac051437y.

86. Trygg, J., and Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). J. Chemom. *16*, 119–128.

87. RColorBrewer, S., and Liaw, M.A. (2018). Package 'randomforest' (University of California).

88. 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 Bioinf. *12*, 77–78. https://doi.org/10.1186/1471-2105-12-77.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
SALSA MLPA Probemix	MRC Holland	P016-C2 VHL kit
Deposited data		
Metabolomic raw data	This paper	https://ngdc.cncb.ac.cn/omix: accession no.OMIX005906
Code for metabolomic analysis	This paper	https://github.com/singleces/Metabolomics_VHL
Sanger sequencing		https://ngdc.cncb.ac.cn/omix: accession no.OMIX006689
Software and algorithms		
Xcalibur	Thermo Fisher Scientific	RRID:SCR_014593
ProteoWizard	Smith et al.[80]	http://metlin.scripps.edu/download/
SIMCA	Sartorius Stedim Data Analytics AB, Umea, Sweden	http://umetrics.com/products/simca
R package	CRAN	N/A
Other		
UHPLC system	Thermo Fisher Scientific	Vanquish
Q Exactive HFX mass spectrometer	Thermo Fisher Scientific	Orbitrap MS

RESOURCE AVAILABILITY

Lead contact

Further information and requests for the data should be directed to and will be fulfilled by the lead contact, Kan Gong (gongkan_pku@126.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Metabolomics data reported in this paper have been deposited in the OMIX, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences[81,82] (https://ngdc.cncb.ac.cn/omix: accession no.OMIX005906) and is publicly accessible. Sanger sequencing data has been deposited at OMIX database under accession number OMIX006689. As per local regulations, the data is placed under controlled access and access can be requested via following the appropriate procedure.
- Original code used for metabolomic analysis is available at: https://github.com/singleces/Metabolomics_VHL.
- Any additional information required to reanalyze the data reported in this study will be provided upon request to the lead contact.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human and tissue subjects

Patients diagnosed with VHL syndrome at Peking University First Hospital, the only international VHL Alliance clinical care center in China, were enrolled in this study. All patients were from diverse regions of China and were of Han Chinese ethnicity. The inclusion criteria for enrollment were as follows: (1) Genetic confirmation of VHL gene point mutations or fragment deletions via Sanger sequencing or next-generation sequencing; (2) Clinical diagnosis of VHL syndrome, with at least one family member having undergone genetic testing; (3) Radiological evidence (ultrasound, CT, MRI) suggesting renal occupancy and suspicion of RCC; (4) Without history of undergoing radiotherapy, radical surgical procedures, or any palliative surgical interventions prior to participation in the study.

Peripheral blood samples were collected from patients diagnosed with VHL-RCC and from healthy individuals at the same institution for comparison. RCC tumor and AN samples were obtained from VHL patients undergoing surgical treatment for kidney cancer. The collection of these samples was integral to the study, enabling subsequent comparative analyses between plasma of VHL patients and healthy subjects, as well as between tumor and AN samples. Patient demographics, including gender, age, family history, generation of onset, age at disease onset, and pathological type, were acquired through the clinical registration system or via direct consultation with the patients and their relatives. The age of onset was defined as the age at which the patient first exhibited symptoms related to VHL manifestation (As detailed in Table S2).

This project adhered to the principles and spirit of the Declaration of Helsinki and was approved by the Medical Ethics Committee of Peking University First Hospital, Beijing, China. Informed consent was obtained from all participants after they were fully informed about the research process and potential outcomes.

METHOD DETAILS

Genetic testing

Genomic DNA was extracted from the peripheral blood leukocytes of patients suspected to VHL disease using a DNA extraction kit (Tiangen, China). The VHL germline mutation status was assessed by PCR amplification with previously described VHL primers, and the PCR products were analyzed through Sanger sequencing.[83] This analysis primarily identified missense mutations, small insertions or deletions, frameshift mutations, and splice site mutations. If Sanger sequencing revealed no mutations, further analysis for large fragment deletions (LDs) was performed using either next-generation sequencing (NGS) or multiplex ligation-dependent probe amplification (MLPA) with the SALSA MLPA Probemix P016-C2 VHL kit (MRC Holland, Amsterdam, The Netherlands) methods. All results were carefully reviewed by two different individuals to ensure the accuracy of the mutation identification. The mutation spectrum of the patients with VHL disease is presented in Table S2.

Sample preprocessing and preservation

Whole blood samples from all study subjects were collected in the morning on an empty stomach to minimize the impact of food and diurnal variations on the plasma levels of low molecular weight metabolites. These samples were then placed in tubes containing ethylenediaminetetraacetic acid (EDTA) as an anticoagulant. Subsequently, the samples underwent centrifugation at 3500 rpm for 10 min to separate the plasma. After centrifugation, all samples were systematically cataloged and stored at a temperature of −80°C for future experimental analysis.

RCC tissue and corresponding AN samples were immediately collected post partial or radical nephrectomy. The samples were rapidly cleansed with PBS to remove any surface blood clots, flash-frozen in liquid nitrogen, and subsequently transferred to a −80°C freezer for long-term preservation, pending further sequencing analysis.

Metabolites extraction and LC-MS/MS analysis

For the extraction of metabolites, both plasma and tissue samples underwent a similar chemical process. Initially, 50 μL of each sample was transferred into an EP tube. For tissue samples, an additional step of homogenization or pulverization was performed prior to further processing to ensure effective cell lysis and metabolite release. Following this, 200 μL of a 1:1 acetonitrile-methanol mixture containing an isotopic standard was added to each sample, whether plasma or tissue. The mixture was then vortexed for 30 s and sonicated in an ice-water bath for 10 min. Afterward, the samples were incubated at −40°C for 1 h to facilitate protein precipitation, followed by centrifugation at 4 °C at 12,000 rpm for 15 min. The resulting supernatant from both plasma and tissue samples was then collected for subsequent analysis. Finally, a quality control (QC) sample was prepared by combining equal volumes of supernatants from all the processed samples.[80]

LC-MS/MS analyses were conducted on an UHPLC system (Vanquish, Thermo Fisher Scientific), interfaced with a Q Exactive HFX Orbitrap mass spectrometer (Thermo). This platform employs an electrospray ionization source, which operates in two ionization modes: positive ion mode (POS) and negative ion mode (NEG). When applied in metabolomics, the concurrent use of both ionization methods enhances the coverage of metabolites, resulting in superior detection performance. In the subsequent data analysis, the datasets from the two ionization modes are analyzed separately to ensure the comprehensiveness and accuracy of the metabolic profiling.

Chromatographic separation was achieved using a UPLC BEH Amide column measuring 2.1 mm × 100 mm with 1.7 μm particles.[84] The elution was performed with a mobile phase comprising 25 mmol/L ammonium acetate and 25 ammonia hydroxide in water (pH adjusted to 9.75) as solvent A, and acetonitrile as solvent B. Samples were maintained at 4°C in the auto-sampler, and the injection volume for analysis was set at 3 μL. The QE HFX mass spectrometer, controlled by Xcalibur software (Thermo), was selected for its proficiency in obtaining MS/MS spectra via information-dependent acquisition (IDA) mode. In this configuration, the software perpetually analyzes the full scan MS spectrum. Parameters for the ESI source were established as follows: sheath gas flow at 30 Arb, auxiliary gas flow at 25 Arb, capillary temperature at 350°C, a full MS resolution of 60000, MS/MS resolution at 7500, and varying collision energies of 10, 30, and 60 in NCE mode. The instrument was operated with a spray voltage of 3.6 kV in positive mode and −3.2 kV in negative mode.

Data processing and model development

Raw data conversion into mzXML format was executed using ProteoWizard software.[85] Subsequent processing stages, including peak identification, extraction, alignment, and integration, were carried out using a custom-developed program based on the R language and utilizing the XCMS platform. For the step of metabolite annotation, an internally developed MS2 database (BiotreeDB) was utilized, with a set annotation threshold of 0.3.

Peak filtering was performed to eliminate noise and to filter out deviations based on the relative standard deviation or the coefficient of variation. Criteria for peak retention included having no more than 50% missing values in a single group or across all groups in the peak area data. Missing values in the raw data were addressed by imputation, replacing them with half the minimum value detected. Data normalization was conducted using internal standards (IS) for uniformity. Subsequently, Orthogonal Projections to Latent Structures-Discriminant Analysis (OPLS-DA) is utilized for a more efficient analysis.[86]

Data were logarithmically transformed and UV-scaled using SIMCA software (Version 16.0.2, Sartorius Stedim Data Analytics AB, Umea, Sweden). Initially, an OPLS-DA model was constructed focusing on the first principal component. The model's quality was assessed using 7-fold cross-validation. The effectiveness of the model was evaluated based on R^2Y (the model's interpretability of the categorical variable Y) and Q^2 (predictive ability of the model) obtained from cross-validation. Finally, a permutation test was conducted by randomly altering the order of the categorical variable Y multiple times to generate various random Q^2 values, providing further validation of the model's effectiveness. After model construction, a permutation test is performed to validate the model's statistical significance and predictive power. This test involves randomly shuffling class labels and rebuilding the model multiple times to assess if the original model performs significantly better than these random iterations, thus confirming its robustness and reliability.

Differentially abundunt metabolites (DAMs) identification

DAMs were identified based on criteria of VIP (Variable Importance in the Projection) $\geq$ 1 and P-value $\leq$0.05, utilizing the VIP from the first principal component projection and P-values obtained from Mann-Whitney tests. To identify metabolites with high sensitivity and specificity, the methodology involved a 5-fold cross-validation approach for logistic regression modeling, implemented to analyze metabolomic data. For each fold, the model was trained on a subset of the data and validated on a separate test set, calculating the Area Under the Curve (AUC) for each metabolite. The average AUC values across all folds were computed for each metabolite. This process evaluates the efficacy of each metabolite in distinguishing between different conditions such as VHL patient and control groups. Metabolites with an AUC greater than 0.65 were selected as DAMs for subsequent analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

To utilize plasma metabolites for accurately distinguishing healthy individuals from patients with VHL patients with RCC, a diagnostic model was developed through a series of analyses. Subjects in the plasma group were randomly divided in a 4:1 ratio into a training cohort and a test cohort. The training cohort comprised 48 VHL patients and 24 healthy volunteers, while the independent test set included 13 VHL patients and 7 healthy volunteers (Table S3). Using the randomForest package in R,[87] the importance scores for 23 differentially abundant metabolites (DAMs) specific to RCC, shared by both plasma and tissue, were calculated in the training cohort. The top 10 DAMs ranked by the important scores were selected to establish the diagnostic model based on the training cohort. The model's efficiency was evaluated using the area under the receiver operating characteristic (ROC) curve, calculated via the pROC package in R.[88] To mitigate overfitting, the predictive performance of the model was further assessed using the randomForest package on the independent test set.

To identify DAMs that were associated to age-related tumor risk of VHL patients, the relationship between the levels of DAMs and the onset age of various VHL manifestations was investigated. In this study, the plasma cohort of VHL patients was randomly divided into a training set and an validation set in an 4:1 ratio, allowing for both univariate and multivariate Cox regression analyses to examine the associations of interest. The predictive accuracy of a risk prediction model was assessed using the concordance index (c-index). To further explore DAMs specifically associated with the onset of RCC, the intersection of DAMs from both plasma and tissue was analyzed again using Cox regression.

Additionally, to explore the relationship between the levels of DAMs and *VHL* gene mutation locations and types, Mann-Whitney tests were employed to compare the abundunce of each metabolite in cases with specific types of mutations against all other cases. Kaplan-Meier curves, dividing patients into two groups based on the median plasma DAMs abundunce, were plotted using ggsurvplot to analyze the relationship between the abundunce of DAMs and initial disease site onset, supplemented by log rank analysis. Pathway enrichment analysis of the DAMs was performed using MetaboAnalyst 5.0 based on the Small Molecule Pathway Database (SMPDB).[29] The comparison of clinical parameters between the control group and VHL patients in plasma group, as well as between the training and validation sets, was calculated using the Chi-square test. All analyses were conducted using R language (version 4.2.2). A P-value <0.05 was considered statistically significant.