BMC Cancer

RESEARCH

Open Access

Comprehensive molecular characteristics of hepatocellular carcinoma based on multi-omics analysis

Ying-Ying Wang[1†], Wan-Xia Yang[1†], Jiang-Ying Cai[1], Fang-Fang Wang[1] and Chong-Ge You[1*]

Abstract

Background The high heterogeneity of hepatocellular carcinoma (HCC) poses challenges for precision treatment strategies. This study aims to use multi-omics methodologies to better understand its pathogenesis and discover biomarkers.

Methods Quantitative proteomics was used to investigate hepatocellular carcinoma tissues (HCT) and their corresponding adjacent non-tumor tissues (DNT), obtained from six HCC patients. Untargeted metabolomics was applied to analyze the metabolic profiles of HCT and DNT of ten HCC patients. Statistical analyses, such as the Student's t-test, were performed to identify differentially expressed proteins (DEPs) and metabolites (DEMs) between the two groups. The functions and metabolic pathways involving DEPs and DEMs were annotated and enriched using the gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) databases. Bioinformatics methods were then utilized to analyze consistency between proteomics and metabolomics results, leading to identification of potential biomarkers along with key altered pathways associated with HCC.

Results This study identified 1556 DEPs between HCT and DNT samples. These DEPs were primarily enriched in crucial biological pathways such as amino acid degradation, fatty acid metabolism, and DNA replication. Subsequently, the analysis of metabolomics identified 500 DEMs that mainly participated in glycerophospholipid metabolism, the phospholipase D signaling pathway, and choline metabolism related to cancer. Integrated analysis of proteomics and metabolomics data unveiled significant dysfunctions in bile secretion, multiple amino acid and fatty acid metabolic pathways among HCC patients. Further investigation revealed that five proteins (PTP4A3, B4GALT5, GAB1, ME2, and PKM) along with seven metabolites (PI(6 keto-PGF1alpha/16:0), 13, 16, 19-docosatrienoic acid, PA(18:2(9Z, 12Z)/20:1(11Z)), Citric Acid, PG(20:3(6, 8, 11)-OH(5)/18:2(9Z, 12Z)), Spermidine, and N2-Acetylornithine) exhibited excellent diagnostic efficiency for HCC and could serve as its potential biomarkers.

†Ying-Ying Wang and Wan-Xia Yang contributed equally to this study and share first authorship.

*Correspondence:
Chong-Ge You
youchg@lzu.edu.cn

Full list of author information is available at the end of the article

Conclusion Our integrated proteome and metabolome analysis revealed 10 key HCC-related pathways and proposed 12 potential biomarkers, which may enhance our understanding of HCC pathophysiology and be helpful in facilitating early diagnosis and treatment strategies.

Keywords Hepatocellular carcinoma, Proteomics, Metabolomics, Multi-omics combined analysis, Biomarkers

Introduction

Hepatocellular carcinoma (HCC) is the predominant histological subtype of primary liver cancer, ranking sixth in global malignant tumor incidence and third as a leading cause of cancer-related mortality [1]. Annually, there are approximately 740,000 new cases of HCC and 640,000 deaths worldwide [2]. Despite significant advancements in the diagnosis, treatment, and prevention of HCC in recent years, its intricate pathogenesis and diverse clinical manifestations remain focal points and challenges for scientific research. The majority of HCC patients are diagnosed at an advanced stage where radical treatment options are not feasible. Studies have revealed that the 5-year survival rate of HCC patients undergoing surgical resection in the intermediate and advanced stages ranges from only 50–68%, with a recurrence rate as high as 50-70%. Conversely, HCC patients diagnosed and treated at an early stage exhibit a more than 70% 5-year survival rate, accompanied by a mere 10% recurrence rate [3, 4]. Consequently, employing advanced research methodologies for an in-depth exploration of HCC pathogenesis and the identification of novel diagnostic markers holds immense significance in enhancing early detection rates, overall survival rates, and quality of life among HCC patients.

In recent years, the rapid development of post-genomic technology has made omics research methods an indispensable tool for studying disease pathogenesis due to their multidimensional, systematic, and diverse advantages. Proteomics enables comprehensive analysis of protein expression, structure, and function in organisms to reveal dynamic changes and regulatory networks [5]. Metabolomics provides a deep understanding of cellular metabolism changes and their relationship with disease occurrence and development by studying all metabolic substances and pathways in an organism [6]. Using proteomics methods to analyze liver tissue and plasma of HCC patients, researchers identified alterations in spliceosome pathway activation, cholesterol homeostasis disorder, and up-regulation of coagulation and complement related pathways in HCC patients. Additionally, it is suggested that aldo-keto reductase family 1 member B10 (AKR1B10), cathepsin A (CTSA), sterol o-acyltransferase1 (SOAT1), pyrroline-5-carboxylate reductase 2 (PYCR2), and alcohol dehydrogenase 1 A (class I), alpha polypeptide (ADH1A) have potential as biomarkers for the diagnosis and prognosis of HCC [7–9]. Based on metabolomics methods, researchers investigated the plasma and liver tissues of HCC patients and observed abnormalities in metabolic pathways including tricarboxylic acid cycle (TCA cycle), glycolysis, purine metabolism, fatty acid metabolism, and retinol metabolism. Free fatty acids such as betaine and propionylcarnitine, retinol derivatives like retinaldehyde, uric acid levels along with bile acids and elaidic acid were found to play significant roles in these metabolic pathways. These metabolites hold promise as potential diagnostic markers for HCC [10–16]. The application of liver tissue proteomics and metabolomics is particularly crucial in the study of HCC. Utilizing paired omics analysis of different tissues in patients not only effectively mitigates individual differences such as age, gender, and region but also significantly enhances the accuracy of screening organ-specific biomarkers. This approach enables a more accurate reflection of the physiological and pathological state of the liver itself, independent from other organs or systemic influences [6].

With the advancement of science and technology, multi-omics integrated analysis has emerged as a more precise, comprehensive, and efficient approach for disease research. By integrating transcriptomics, proteomics, and metabolomics analyses, researchers have conducted thorough investigations into the pathogenesis of malignant tumors such as liver cancer, colorectal cancer, and breast cancer. Consequently, they have identified potential diagnostic markers, novel therapeutic targets, and prognostic indicators. This innovative methodology offers enhanced accuracy and effectiveness in the prevention, diagnosis, treatment, and prognosis of these diseases [17–19]. Despite the progress made, there remains a dearth of comprehensive exploration into the pathogenesis and biomarkers of HCC through the integration of liver tissue proteomics and metabolomics. Moreover, the overall research status concerning protein and metabolite regulatory networks in HCC patients remains unclear, with highly specific diagnostic markers yet to be identified. Therefore, this study aims to comprehensively apply liver tissue proteomics and metabolomics technologies to compare and analyze paracancerous liver tissues with HCC tissues' protein and metabolite expression profiles. The objective is to identify differentially expressed proteins and metabolites closely associated with HCC while investigating their correlation and role in its pathogenesis. Ultimately, this study aims to provide novel insights for early diagnosis and personalized treatment of HCC.

Materials and methods

Participants and sample collection

Ten patients with HCC who were admitted to the Second Hospital of Lanzhou University between March 2023 and April 2024 were included in this study. Non-targeted metabolomics analysis was conducted on selected samples of hepatocellular carcinoma tissues (HCT) and their corresponding adjacent non-tumor tissues (DNT). All patients had a confirmed pathological diagnosis of HCC and had not undergone any prior treatment, including chemotherapy, radiotherapy, immunotherapy, or other therapies. The metabolic profile changes in ten HCC patients were analyzed using the non-targeted metabolomics approach. Additionally, the data-independent acquisition (DIA) quantitative proteomics method was employed to further investigate the protein and metabolomic profiles of six HCC patients by analyzing their HCT and DNT samples jointly. The HCT samples were collected from the central region of surgically resected liver tissue containing solid tumors, while DNT samples were obtained from an area located 5 cm outside the boundary of the solid tumor. Immediately after surgical resection, all tissue specimens were flash-frozen in liquid nitrogen to preserve their biological activity and stored at -80 °C until analysis.

DIA quantitative proteomic analysis

Extraction of protein

Under frozen conditions, excise tissue weighing 100 ± 0.5 mg and transfer it to a grinding tube. Add an appropriate volume of protein lysis buffer (8 M urea + 1% SDS) supplemented with protease inhibitors. Agitate the mixture three times using a cryogenic grinder for 40 s each time. Place the mixture on ice for 30 min, vortexing for 5–10 s every 5 min to facilitate lysis. Subsequently, centrifuge at 4 °C, 16,000 g for 30 min and collect the supernatant. Quantify the protein content in the supernatant using Bradford assay and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

Peptide sample preparation and DIA mass spectrometry detection

Firstly, a protein sample of 100 μg was taken and mixed with lysis buffer and triethylammonium bicarbonate buffer (TEAB) to achieve a final concentration of 100 mM. Tris (2-carboxyethyl) phosphine (TCEP) was then added to reach a final concentration of 10mM, and the reaction was carried out at 37 °C for 60 min. Subsequently, Iodoacetamide (IAM) was added to attain a final concentration of 40mM, and the mixture was incubated in darkness at room temperature for 40 min. Afterwards, the mixture was subjected to centrifugation after adding six volumes of pre-cooled acetone and allowing it to stand at -20 °C for four hours in order to separate the precipitate. The precipitate was dissolved in 100μL of TEAB with a concentration of 100mM and digested overnight at 37 °C using trypsin at a mass ratio of 1:50. The resulting peptide mixture was then dried under vacuum to remove solvents, followed by resuspension in trifluoroacetic acid (TFA) solution with a concentration of 0.1%. Desalting using an HLB column was performed to eliminate impurities from the desalted peptides which were subsequently dried again under vacuum to remove any remaining solvents. Finally, quantification of peptides took place using NANODROP ONE UV-Visspectrophotometer (Thermo Fisher Scientific, San Jose, USA).

Equal amounts of peptides were initially dissolved in a loading buffer for mass spectrometry and subsequently injected into a vanquish Neo chromatography system (Thermo Fisher Scientific, San Jose, USA). High-performance liquid chromatography (HPLC) separation was conducted using a uPAC High Throughput column (75 μm × 5.5 cm, Thermo Fisher Scientific, San Jose, USA). A gradient elution method was employed, utilizing mobile phase A containing 2% acetonitrile and 0.1% formic acid, along with mobile phase B consisting of 80% acetonitrile and 0.1% formic acid. During an 8-minute chromatographic separation, the ratio between mobile phases A and B varied according to a predefined gradient to optimize sample separation. Subsequent to the separation process, the peptides entered an astral mass spectrometer (Thermo Fisher Scientific, San Jose, USA) for ionization and measurement of their mass-to-charge ratios, thereby generating mass spectrometry data. Data acquisition and initial processing were managed by Thermo xcalibur software (Thermo Fisher Scientific, San Jose, USA), employing a SWATH-based DIA approach throughout the entire procedure.

Differentially expressed proteins analysis

The proteome discoverer software (Thermo Fisher Scientific, San Jose, USA) was utilized for hierarchical database construction and data analysis. Subsequently, the database search results were imported into Spectronaut™ software (Biognosys AG, Schlieren, Switzerland) to establish the spectral library. The Spectronaut™ software was employed to extract DIA raw data peaks based on ion pair information from the spectrum library. Default settings of the software and iRT correction were applied during analysis. Up to 6 specific peptides per protein and up to 6 child ions per peptide were selected. Protein quantification results underwent preprocessing by eliminating duplicate values and filling missing values. Differential expression proteins (DAPs) were identified using univariate analysis and other methods with a screening threshold of the absolute value of the logarithm base 2 of the fold change ($|Log_2FC| >= 1$ and adjusted p-value < 0.05.

Metabolomics analysis

Metabolite extraction

The tissue sample weighing 100 mg was placed in a 2mL centrifuge tube, followed by the addition of a ground bead with a diameter of 6 mm. Subsequently, 400μL of an extraction solution consisting of methanol and water in a volume ratio of 4:1, containing an internal standard L-2-chlorophenylalanine at a concentration of 0.02 mg/mL, was added to the centrifuge tube for metabolite extraction. Next, the sample solution was homogenized using a frozen tissue grinder for 6 min (-10 °C, 50 Hz) and then subjected to low temperature ultrasound extraction for 30 min (5 °C, 40 kHz). After sonication, the centrifuge tube was cooled at -20 °C for 30 min and subsequently centrifuged for another 15 min (4 °C, 13000 g). Finally, the supernatant was transferred into an injection vial equipped with an inner cannula for analysis.

LC-MS/MS analysis

The LC-MS analysis was performed using an ultra-high performance liquid chromatography system coupled with a Q Exactive hybrid mass spectrometer (UHPLC-Q Exactive). Chromatographic conditions: 2μL of the sample was injected onto an HSST3 column for separation prior to mass spectrometric detection. Mobile phase A consisted of 95% water and 5% acetonitrile (containing 0.1% formic acid), while mobile phase B was a mixture of 47.5% acetonitrile, 47.5% isopropanol, and 5% water (also containing 0.1% formic acid). The separation gradient started with mobile phase B linearly increasing from 0 to 5% in 0–0.1 min, then from 5 to 25% in 0.1–2 min. From 2 to 9 min, B linearly rose to 100%, staying constant at 100% till 13 min. In 13.0-13.1 min, B linearly decreased to 0%. The flow rate was set at 0.40 mL/min, and the column temperature was maintained at 40 °C during the experiment. The mass spectrometric data was collected using a Thermo UHPLC-Q Exactive mass spectrometer equipped with an electrospray ionization (ESI) source operating in either positive or negative ion mode. The optimal conditions were set as follows: heater temperature, 400 °C; capillary temperature, 320 °C; sheath gas flow rate, 40arb; auxiliary gas flow rate, 10arb; ion-spray voltage floating (ISVF), -2800 V in negative mode and 3500 V in positive mode, respectively. Data acquisition was performed in Data-Dependent Acquisition (DDA) mode with the following settings: normalized collision energy, 20-40-60 V rolling for MS2; isolation window, 1.2 m/z; maximum injection time, 100ms; AGC target, 3e6; dynamic exclusion, 10s. The full MS resolution was set to 70,000, and the MS2 resolution was set to 17,500. The detection was carried out over a mass range of 70–1050 m/z for both MS1 and MS2 scans.

Differentially expressed metabolites analysis

The raw LC-MS data were imported into Progenesis QI (Waters Corporation, Milford, USA) for extensive preprocessing, including baseline filtering, peak identification, integration, retention time correction, and peak alignment. This comprehensive workflow resulted in a structured data matrix, consisting of retention times, mass-to-charge ratios (m/z), and peak intensities. Subsequently, the MS spectra within this data matrix were matched against both public database and Majorbio's proprietary database to identify and annotate the metabolites presented in the samples. After searching the database, the matrix data was preprocessed, and the 80% rule was used to remove the missing values [20]. In order to reduce the error caused by sample preparation and instrument instability, the sum normalization method was used to normalize the response intensity of the spectral peak of the sample, and the normalized data matrix was obtained. At the same time, the variables with relative standard deviation (RSD) > 30% of quality control samples were deleted, and log10 logarithmic processing was performed to obtain the data matrix for subsequent analysis [20, 21]. Principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA), and univariate analysis were utilized to screen for differentially expressed metabolites (DEMs) between groups. DEMs were identified based on the following criteria: FC value ≥ 1, variable importance in projection (VIP) score from the PLS-DA model ≥ 1, and a statistical significance level with adjusted p-value < 0.05 [22, 23].

Statistical analysis

Statistical analyses were performed using R (version 4.4.1). Normality was assessed using the Shapiro-Wilk test from the "stats" package, and group comparisons were conducted using Student's t-test. For multivariate statistical analysis, PCA and PLS-DA were performed using the "ropls" packages (version 1.6.2). Data were scaled (unit variance scaling and mean centering) for PCA, and PLS-DA components were determined by 10-fold cross-validation. Spearman correlation analysis was conducted using the "stats" package. Data visualization was performed with "ggplot2" package (version 3.4.0). The gene ontology (GO) and human metabolome database (HMDB) databases were used for functional annotation analysis of all DEPs and DEMs. Kyoto encyclopedia of genes and genomes (KEGG) database metabolic pathways involved in the differences between proteins and metabolites annotation analysis. GO enrichment analysis was conducted with Goatools (version 1.4.4), using the Fisher exact test with Benjamini-Hochberg false discovery rate (FDR) correction (adjusted p-value < 0.05 considered significant). KEGG pathway enrichment analysis was performed using the "scipy"

package (version 1.7.3) in Python (version 3.8), with the Fisher exact test and FDR correction applied. Relevant networks were constructed And visualized using Cytoscape software (version 3.7.2). All analyses were performed on the Majorbio cloud platform (cloud.majorbio.com). All statistical tests were two-sided, and adjusted p-value < 0.05 was considered statistically significant.

Result

Proteomics

Screening of deps

In this study, DIA quantitative proteomic analysis was performed on HCT and DNT tissues from six HCC patients, and a total of 126,081 peptides were identified and 8435 proteins were quantified. Based on the screening criteria ($|Log_2FC| >=1$ and adjusted p-value < 0.05), a total of 1556 DEPs were identified, of which 1148 proteins showed an up-regulated trend and 408 proteins showed a down-regulated trend in HCC patients (Fig. 1A). The results of cluster analysis showed that there were significant differences in DEPs distribution between HCT and DNT, and the intra-group reproducibility was relatively stable (Fig. 1B).

Functional analysis of deps

In order to deeply understand and analyze the specific functions and roles of DEPs in biological processes, we performed detailed subcellular localization prediction for 1556 DEPs, and performed comprehensive protein function annotation and enrichment analysis based on GO and KEGG databases. The results of subcellular localization showed that the DEPs were mainly concentrated in the cytoplasm (41%), nucleus (20%) and mitochondria (15%) (Fig. 2A). The GO annotation results showed that the top 20 GO terms were mainly focused on biological process (BP) such as cellular process, metabolic process and biological regulation, followed by molecular function (MF) such as protein binding, catalytic activity and molecular function regulation activity. While the cellular component (CC) contains cellular anatomical entities and protein complexes (Fig. 2C). KEGG annotation results showed that 50% of the top 20 KEGG metabolic pathways in abundance belonged to the metabolic category, and among these metabolic pathways, the number of DEPs involved in amino acid metabolism, carbohydrate metabolism, and lipid metabolism was particularly significant (Fig. 2D). According to the adjusted p-value, GO and KEGG enrichment analysis were performed on the expression of DEPs. GO enrichment analysis showed that the DEPs were mainly enriched in cytochrome P450 epoxidase pathway, primary alcohol metabolism process and glutamine family amino acid metabolism process. In terms of cellular components, it mainly involves DNA prereplication initiation complex, CMG complex, mitochondrial large ribosomal subunit and organelle large ribosomal subunit. In terms of molecular functions, the main activities were arachidonic acid monooxygenase activity, arachidonic acid cyclooxygenase activity,

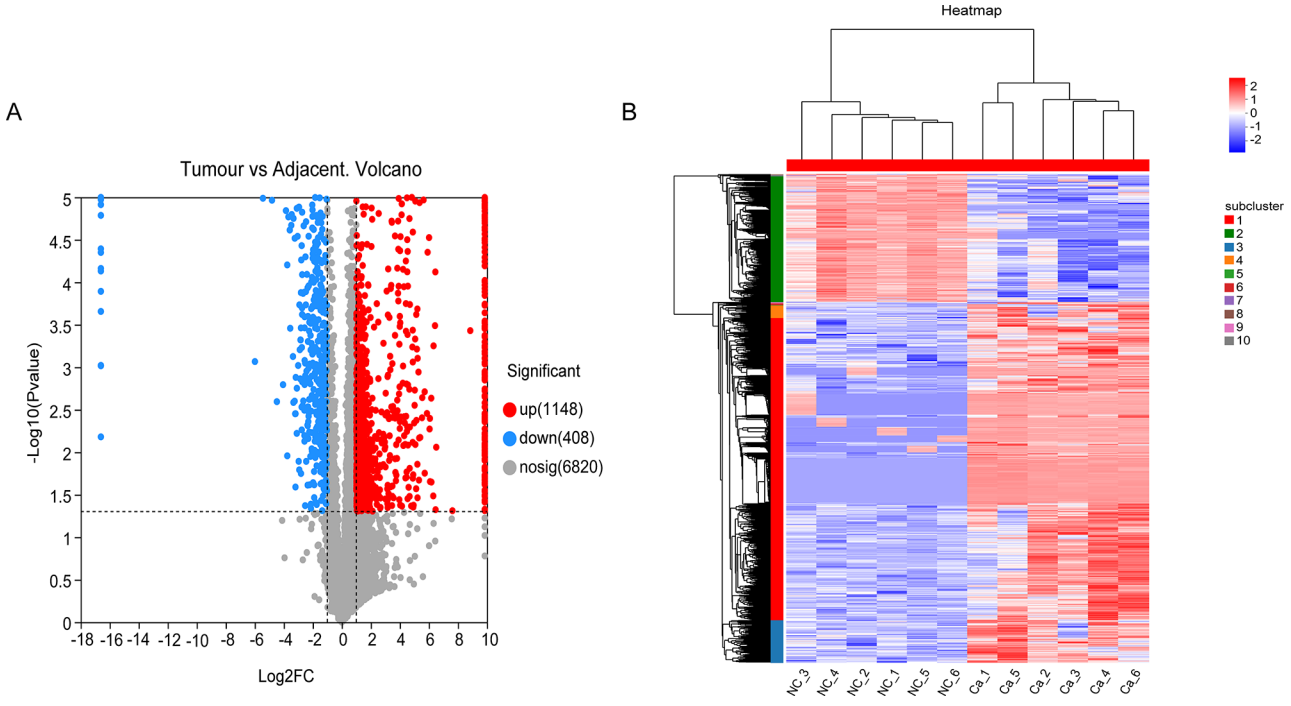

Fig. 1 Characterization of differentially expressed proteins. (**A**) Volcano plot illustrating the differential expression of proteins. (**B**) Hierarchical clustering heatmap depicting the differential protein profiles. The abbreviation NC denotes paracancerous tissue, while Ca refers to hepatocellular carcinoma tissue

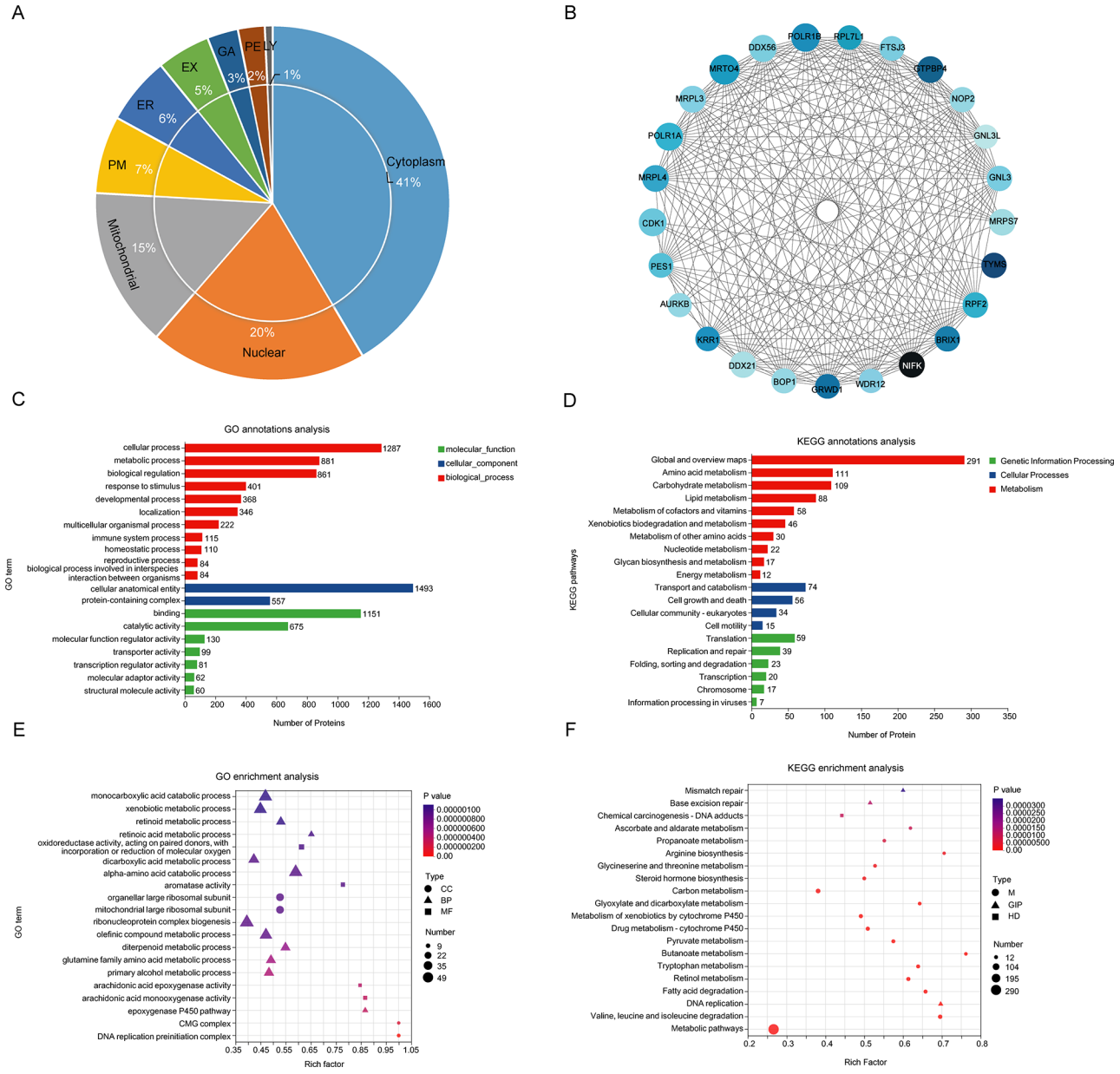

Fig. 2 Functional analysis of differentially expressed proteins. (**A**) Subcellular localization mapping of differentially expressed proteins. Abbreviations: PM, plasma membrane; ER, endoplasmic reticulum; EX, extracellular; GA, golgi apparatus; PE, Periplasm; LY, lysosomal. (**B**) PPI network analysis of differentially expressed proteins. The intensity of the color of the circle in the figure corresponds to the number of nodes connected by the protein; a darker shade indicates a higher number of nodes connected by the protein. (**C**) GO term analysis for the differentially expressed proteins in HCC. (**D**) KEGG annotation for the differentially expressed proteins in HCC. (**E**) GO enrichment analysis for the identification of enriched terms among the differentially expressed proteins. The abscissa represents the enrichment rate, which signifies the ratio of enriched proteins in a specific GO term to the total number of proteins annotated to that term. A higher ratio indicates a greater level of enrichment. The ordinate denotes the GO term itself. Each bubble in the figure corresponds to a distinct GO Term, and its size is proportional to the number of concentrated proteins enriched by that particular term. The varying colors of the bubbles represent adjusted *p*-values, with smaller values indicating redder hues. Abbreviations: CC, cellular component; BP, biological process; MF, molecular function. (**F**) KEGG pathway enrichment analysis to identify significantly enriched pathways among the differentially expressed proteins. Abbreviations: M, metabolism; GIP, genetic information processing; HD, human diseases

aromatase activity and oxidoreductase activity (Fig. 2E). KEGG enrichment analysis showed that the DEPs were mainly enriched in valine, leucine and isoleucine degradation, fatty acid degradation, DNA replication and retinol metabolism pathways (Fig. 2F).

In order to determine the interaction between DEPs, we used the STRING (https://string-db.org/) database to predict the functional correlation between two groups of DEPs from the genomic correlation of their coding genes. The top 25 proteins with the highest connectivity

degree were selected to draw the protein-protein interaction (PPI) network diagram (Fig. 2B). All the 25 proteins showed a trend of up-regulation in HCC, and the 10 proteins with the lowest adjusted *p*-value and high connectivity were POLR1B, MRTO4, GTPBP4, NIFK, TYMS, POLR1A, MRPL4, CDK1, BRIX1, and GRWD1 (Table 1).

Metabolomics

Screening of DEMs

In this study, PCA and PLS-DA analysis were used for pattern recognition of HCT and DNT in 10 HCC patients, respectively. PCA and PLS-DA analysis showed that there was a good separation trend between HCT and DNT (Fig. 3A and B). In order to judge the model quality, 200 permutation tests were performed on the PLS-DA model, and the results showed that the model was good and did not overfit (Fig. 3C). According to the screening conditions (FC ≥ 1, VIP of the PLS-DA model ≥ 1, and adjusted *p*-value < 0.05), a total of 500 DEMs were screened in this study, of which 195 showed an up-regulated trend and 305 showed a down-regulated trend in HCC patients (Fig. 3D). The HMDB compound classification results showed that the DEMs were classified into 11 categories. Among them, lipids and lipoid molecules (189), organic acids and their derivatives (95), organic heterocyclic compounds (66), and organic oxygen-containing compounds (35) were in the majority (Fig. 4).

Among the top 20 DEMs with the lowest adjusted *p*-value, 10 were up-regulated in HCC (PI (6 keto-PGF1alpha/16:0), 13,16, 19-docosatrienoic acid and PA (18:2(9Z,12Z)/20:1(11Z)), etc.), and 10 were down-regulated (N-Amidino-L-aspartate, Lupex and 2,3,4, 5-tetrahydro-2-pyridinecarboxylic acid, among others) (Table 2). At the same time, correlation heat map showed that most of the above 20 DEMs had obvious positive correlation (Fig. 5A). PLS-DA was used as the supervised model, and 7-fold cross validation was used to test the prediction of different changes in pairing. The VIP score analysis of the first principal component identified the top 20 metabolites differentially associated with HCC occurrence. Among these metabolites, 4 were found to be up-regulated in HCC: D-glucurono-6,3-lactone, Pectic acid, Tryptophyl-Valine, and Senkyunolide N. Additionally, 16 of them were down-regulated, including galactosyl lactate, 2,3,4,5-tetrahydro-2-pyridinecarboxylic acid, and N, N-dimethyl-N'-(8-methyl-5 H-pyrimido[5,4-b] indol-4-yl) ethane-1,2-diamine (Fig. 5B).

Table 1 The PPI analysis identified the top 25 proteins with the highest degree of connectivity

Symbol	Protein name	Trend	Adjusted *p*-value	Degree	Log_2 FC
POLR1B	DNA-directed RNA polymerase I subunit RPA2	up	< 0.001	120	1.31
MRTO4	mRNA turnover protein 4 homolog	up	< 0.001	128	1.17
GTPBP4	GTP-binding protein 4	up	< 0.001	111	1.53
NIFK	MKI67 FHA domain-interacting nucleolar phosphoprotein	up	< 0.001	107	1.35
TYMS	Thymidylate synthase	up	< 0.001	104	5.21
POLR1A	DNA-directed RNA polymerase I subunit RPA1	up	< 0.001	117	1.04
MRPL4	39 S ribosomal protein L4	up	< 0.001	111	1.10
CDK1	Cyclin-dependent kinase 1	up	< 0.001	117	2.06
BRIX1	Ribosome biogenesis protein BRX1 homolog	up	< 0.001	104	1.31
GRWD1	Glutamate-rich WD repeat-containing protein 1	up	< 0.001	101	1.21
MRPL3	39 S ribosomal protein L3	up	< 0.001	112	1.01
RPF2	Ribosome production factor 2 homolog	up	< 0.001	105	1.67
DDX56	Probable ATP-dependent RNA helicase DDX56	up	0.01	113	1.18
KRR1	KRR1 small subunit processome component homolog	up	< 0.001	100	1.23
RPL7L1	60 S ribosomal protein L7-like 1	up	< 0.001	100	1.07
MRPS7	28 S ribosomal protein S7	up	0.02	110	1.12
FTSJ3	pre-rRNA 2'-O-ribose RNA methyltransferase FTSJ3	up	< 0.001	101	1.55
GNL3	Guanine nucleotide-binding protein-like 3	up	0.01	104	1.37
DDX21	Nucleolar RNA helicase 2	up	0.02	109	1.71
PES1	Pescadillo homolog	up	< 0.001	101	1.32
NOP2	Probable 28 S rRNA (cytosine (4447)-C (5))-methyltransferase	up	0.01	104	1.46
BOP1	Ribosome biogenesis protein BOP1	up	0.01	103	1.40
WDR12	Ribosome biogenesis protein WDR12	up	< 0.001	99	1.21
AURKB	Aurora kinase B	up	0.02	98	9.86
GNL3L	Guanine nucleotide-binding protein-like 3-like protein	up	0.02	98	1.88

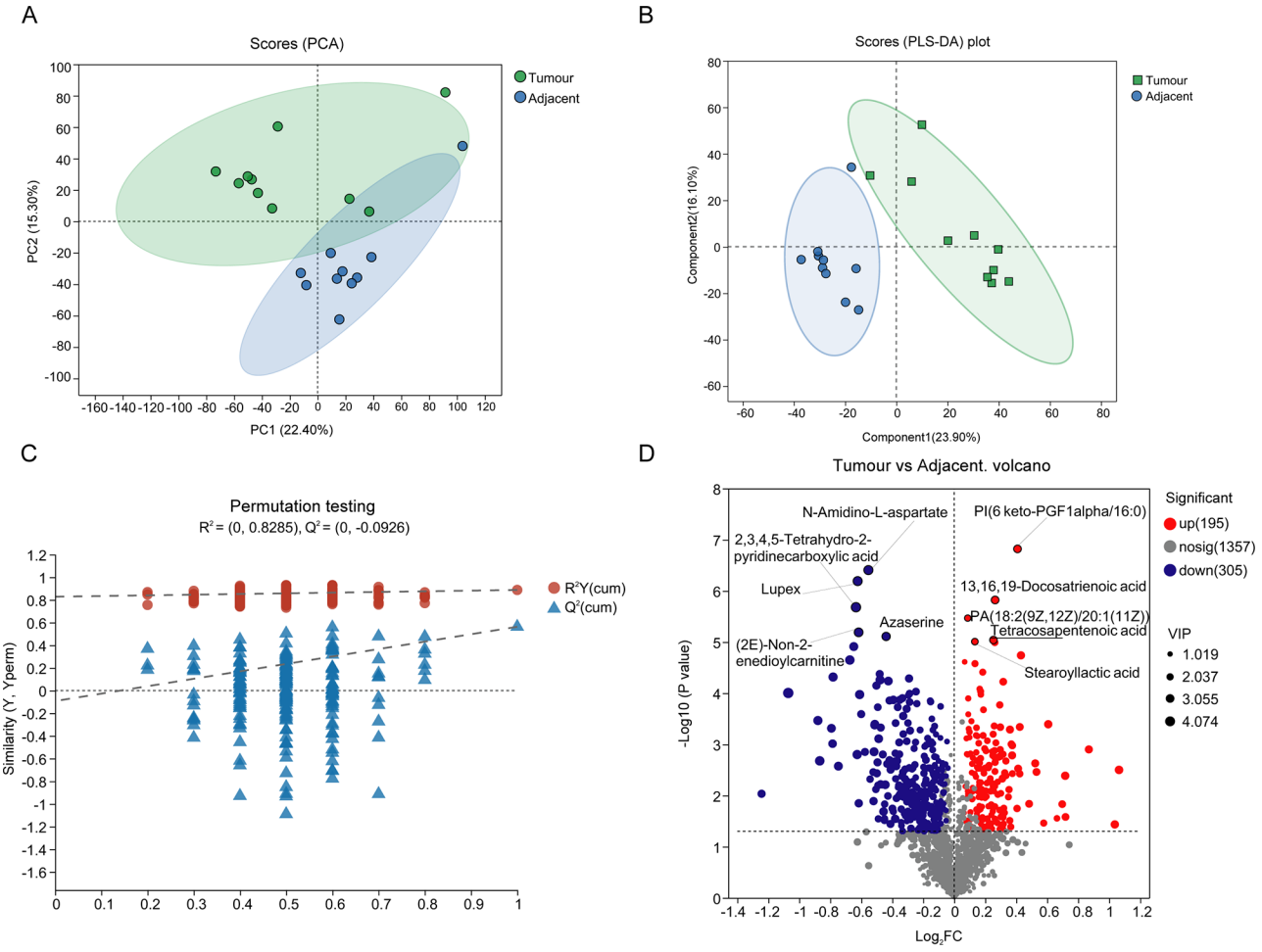

Fig. 3 Characterization of metabolomics between HCT and DNT. (**A**) Plot of PCA scores between the two groups. After conducting dimension analysis, the reduced samples were plotted as relative coordinate points on the principal components p1 and p2. The distance between each point reflects the level of aggregation and dispersion among samples, with shorter distances indicating higher similarity and longer distances suggesting greater dissimilarity. (**B**) Plot of PLS-DA scores between the two groups. The PLS-DA score plot is commonly used to visually demonstrate the model's classification effectiveness, where a larger separation between the two sample groups indicates a more pronounced classification effect. (**C**) PLS-DA permutation test plot. The abscissa represents the permutation retention of the test, while the ordinate represents R^2 (red dot) and Q^2 (blue triangle) values. The dashed lines correspond to regression lines for R^2 and Q^2 respectively. Notably, the regression line for R^2 is positioned above that of Q^2. Moreover, the intercept between the regression line for Q^2 and y-axis is -0.0926, indicating a well-fitted model with high predictability suitable for subsequent data analysis. (**D**) Volcano plot of differentially expressed metabolites. The horizontal axis represents the $\log_2$FC in expression between the two groups, while the vertical axis represents the -log10(p-value) of statistical test for metabolite expression difference. Higher values indicate greater significance in expression differences. Each dot on the plot corresponds to a specific metabolite, with dot size indicating VIP value. Red dots represent up-regulated metabolites, blue dots represent down-regulated metabolites, and gray dots represent non-differential metabolites. The position of each dot reflects its level of significance in terms of expression difference

Discussion

The objective of this study is to investigate the potential pathogenesis and biomarkers of HCC by conducting a combined omic analysis on HCT and DNT tissues from HCC patients. By performing functional and association analyses on DEPs and DEMs, along with their biological significance, this study identified 5 proteins such as PTP4A3 and 7 metabolites such as PI (6 keto-PGF1alpha/16:0) that may assist in the diagnosis of HCC. Furthermore, metabolic pathways such as tryptophan metabolism, butanoate metabolism, pyruvate metabolism, and abnormal bile secretion were found to be closely associated with symptoms like immune escape, abnormal energy metabolism, and lipid peroxidation in HCC patients.

The proteins PTP4A3, B4GALT5, GAB1, ME2, and PKM were found to exhibit high expression levels in HCC and demonstrated good diagnostic potential for HCC (AUC > 0.90). PTP4A3 is a phosphatase that plays a crucial role in cellular signaling by catalyzing the dephosphorylation of tyrosine or serine/threonine residues in protein molecules [24]. Its involvement in

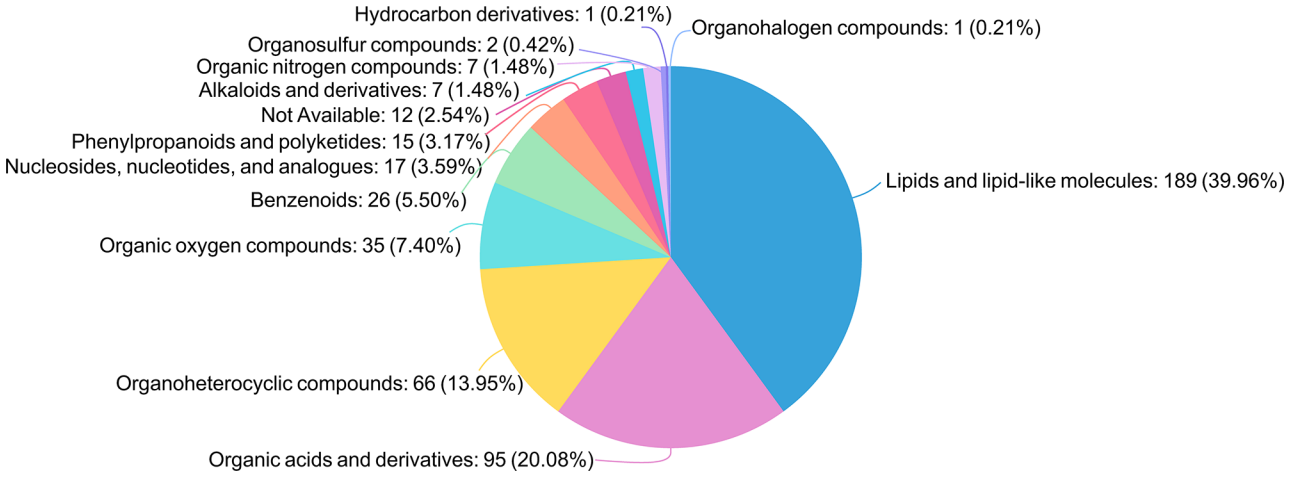

Fig. 4 The classification of compounds in HMDB

Table 2 The top 20 metabolites with significant differences between HCT and DNT

Metabolite	Trend	Mode	VIP	FC	Log_2FC	Adjusted *p*-value
PI (6 keto-PGF1alpha/16:0)	up	ESI-	2.4039	1.3273	0.4085	1.514E-07
N-Amidino-L-aspartate	down	ESI+	2.9828	0.6812	-0.5538	3.926E-07
Lupex	down	ESI-	2.9952	0.6492	-0.6233	6.446E-07
13,16,19-Docosatrienoic acid	up	ESI-	2.3091	1.2007	0.2639	0.000001507
2,3,4,5-Tetrahydro-2-pyridinecarboxylic acid	down	ESI-	3.1523	0.6443	-0.6342	0.000002089
PA (18:2(9Z,12Z)/20:1(11Z))	up	ESI-	1.4519	1.0623	0.0872	0.000003408
(2E)-Non-2-enedioylcarnitine	down	ESI-	2.7138	0.6521	-0.6168	0.000006465
Azaserine	down	ESI+	2.5226	0.7374	-0.4395	0.000007775
Tetracosapentenoic acid	up	ESI-	2.2002	1.1919	0.2533	0.000009136
Stearoyllactic acid	up	ESI-	1.5042	1.0968	0.1333	0.00000983
5,8,11-Eicosatrienoic acid	up	ESI-	1.9709	1.199	0.2618	0.00001021
Abacavir	down	ESI-	2.8114	0.6381	-0.6481	0.00001227
Cytidine-5′-diphosphocholine	up	ESI-	2.6111	1.3479	0.4307	0.00001821
(6R)-Folinic acid	down	ESI-	2.9997	0.6269	-0.6737	0.00002248
Citric Acid	up	ESI-	1.1145	1.0479	0.0675	0.00002444
PG (18:2(9Z,12Z)/18:1(9Z))	up	ESI-	1.7352	1.0973	0.134	0.00002644
PG (20:3(6,8,11)-OH(5)/18:2(9Z,12Z))	up	ESI-	1.8584	1.136	0.184	0.00003883
6-Lactoyltetrahydropterin	down	ESI-	2.2779	0.7165	-0.481	0.00004214
Pirbuterol	down	ESI-	2.877	0.5816	-0.7819	0.00004833
L-Hexanoylcarnitine	down	ESI+	2.2113	0.8171	-0.2914	0.00005157

Abbreviations: VIP, variable projection importance; FC, fold chang

the development, metastasis, and recurrence of various cancers has been confirmed, particularly in solid tumors such as colorectal cancer, breast cancer, and HCC where its overexpression often indicates a poor prognosis [25, 26]. In our study, PTP4A3 showed a significant difference between HCT and DNT, with an adjusted p-value < 0.01 Correlation analysis further revealed a strong positive correlation between PTP4A3 and key metabolites 13-16-19-docosatrienoic acid and citric acid. These findings collectively emphasize the important role of PTP4A3 in HCC development and progression. The combined analysis of proteomics and metabolomics in this study revealed that B4GALT5 exhibited the highest FC (FC: 926.6) among the five potential protein biomarkers identified in this study. It was also the only one found to be significantly associated with PG (18:2 (z, z) / 18:1 (z) 9), PG (David,8,11(6)-OH no more(5)/(z, z)), PI (6 keto-PGF1alpha/16:0), and PA (18:2(9Z,12Z)/20:1(11Z)). These findings were consistent with Luo et al.'s research [27], which demonstrated that B4GALT5 was highly expressed in HCC patients and enriched in lipid metabolism-related pathways. Additionally, KEGG enrichment analysis indicated that B4GALT5 primarily participated in sphingolipid metabolism and mucin type O-glycan biosynthesis pathways. This suggests a potential role for B4GALT5 in promoting HCC cell proliferation,

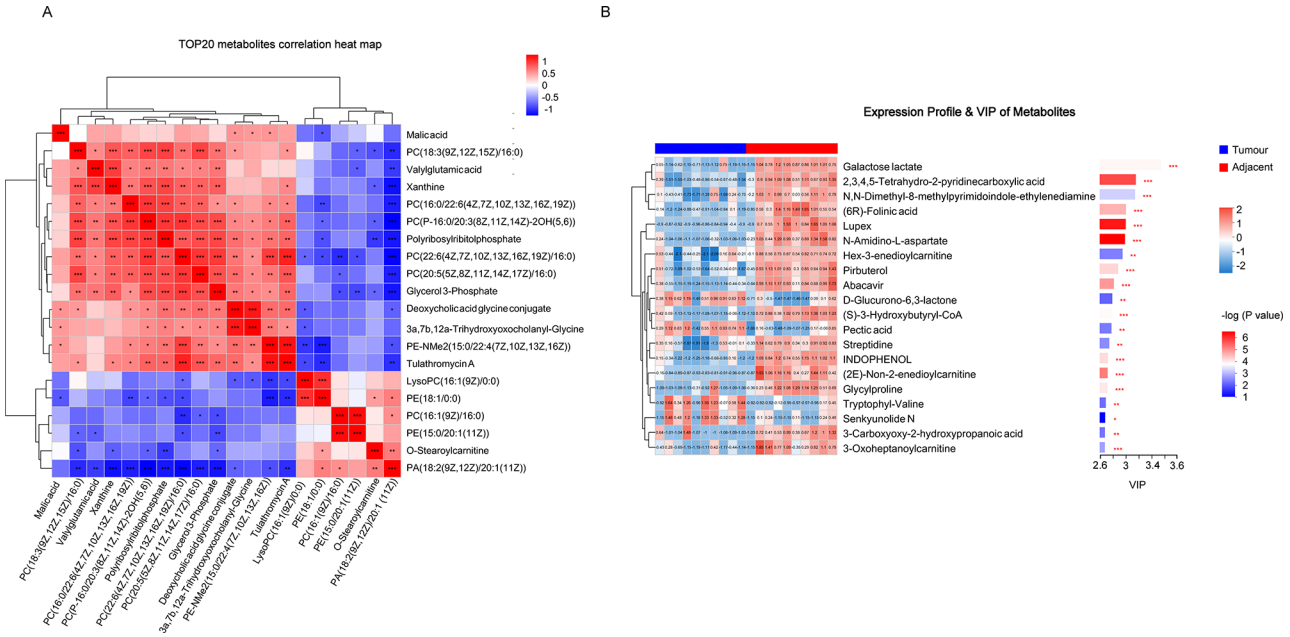

Fig. 5 The metabolite features with the top 20 differences between HCT and DNT. (**A**) A correlation heatmap was generated for the top 20 differentially expressed metabolites. (**B**) VIP value analysis was performed on the top 20 differential metabolites. On the left side, a dendrogram of metabolite clustering is presented, where closer branches indicate similar expression patterns among all metabolites in the samples. Each column represents a sample. Each row represents a metabolite, and color indicates its relative expression level in this group of samples. The gradient color block provides a visual representation of the correspondence between color gradient and numerical size. On the right side, a VIP bar chart displays the contribution value of each metabolite to the difference between two groups (threshold set at 1). A longer bar indicates a greater difference in abundance between HCT and DNT groups. The color of each bar reflects significance levels (adjusted *p*-value) for differences between groups: darker colors represent smaller adjusted *p*-values (-log10(adjusted *p*-value)). Additionally, asterisks (*) denote $P < 0.05$, ** denotes $P < 0.01$, and *** denotes $P < 0.001$

metastasis, invasion, and ultimately contributing to poor prognosis among HCC patients [28]. As a versatile adaptor protein, GAB1 can respond to various stimuli such as extracellular growth factors, cytokines, and immune cell receptors. It is activated by multiple kinases, thereby regulating tumor proliferation, cell cycle progression, evasion of apoptosis, invasion and migration, as well as angiogenesis through the PI3K/AKT and Ras/MAPK signaling pathways [29]. In this study, GAB1 was found to be involved in several signaling pathways including the ras signaling pathway, phospholipase D signaling pathway, and ErbB signaling pathway. The expression level of GAB1 showed a positive correlation with PI(6-keto-PGF1alpha/16:0) and Cytidine-5'-diphosphocholine levels, indicating its potential importance in HCC. Additionally, previous studies have demonstrated that over expression of GAB1 was associated with poor prognosis in patients with anterior skull base meningioma, HCC and breast cancer, suggesting its potential utility as a biomarker for HCC diagnosis and prognosis [30–32]. Our study revealed that ME2 primarily participates in the conversion of malic acid to pyruvate and plays a crucial role in carbon metabolism. Moreover, an interaction was observed between ME2 and citric acid as well as malic acid, which are two vital components of the tricarboxylic acid (TCA) cycle. Numerous studies have also substantiated that ME2 serves as a significant source of NADPH, thereby regulating cellular reduction-oxidation (REDOX) balance and indirectly supporting fatty acids and nucleotide synthesis. Over expression of ME2 is closely associated with metabolic reprogramming in cancer, including glycolysis, amino acid metabolism, and lipid synthesis. Simultaneously, it exerts a pivotal function in the initiation, invasion, and metastasis of various malignant tumors [33]. Furthermore, specific investigations on ME2 in HCC have demonstrated its substantial involvement in glycolysis and lipid synthesis among HCC patients [34, 35]. Pyruvate kinase (PK) is a pivotal enzyme in the glycolysis pathway, catalyzing the final step by transferring a phosphate group from phosphoenolpyruvic acid (PEP) to ADP, resulting in the production of pyruvate and ATP [36]. The PKM gene encodes two isoforms, PKM1 and PKM2, with PKM2 being overexpressed in HCC and other cancers. It plays a crucial role in promoting tumor cell proliferation and metastasis, directly influencing the Warburg effect in tumor metabolism [37]. This study specifically focused on liver tissues from HCC patients and revealed high expression levels of PKM (without specifying subtypes). Furthermore, it demonstrated that PKM was involved in several key metabolic pathways including pyruvate metabolism, carbon metabolism, biosynthesis of amino acids, and

glycolysis/gluconeogenesis. Metabolically linked to citric acid and malic acid - important components of the TCA cycle - through its product pyruvate. Considering PKM2 is predominantly found in tissues with high glucose metabolism such as liver tissue, and previous studies have suggested its potential use as a biomarker for malignant tumors like liver cancer [38]. Our findings further support the significance of members within the PKM family as potential biomarkers for HCC.

Recent studies have demonstrated the dynamic nature of metabolic reprogramming during carcinogenesis and cancer development. Cancer cells necessitate extensive cellular metabolic alterations to survive and proliferate under diverse environmental conditions. Consequently, metabolomics has emerged as a prominent research avenue in tumor cell biology for investigating metabolic reprogramming in tumors [39]. As a prototypical metabolism-related tumor, HCC exhibits distinct characteristics compared to normal liver tissues concerning glucose metabolism, lipid metabolism, and protein metabolism. Under normal physiological conditions, cells efficiently utilize glucose as an energy source through orderly metabolic processes such as glycolysis, the TCA cycle, and oxidative phosphorylation to generate ATP supporting various cellular activities. However, HCC displays significant changes in glucose metabolism. HCC cells exhibit enhanced glucose uptake capacity leading to heightened glycolytic activity and increased lactate production even in oxygen-rich environments. This results in an elevated ratio of lactate to pyruvate—an abnormal phenomenon known as the "Warburg effect" [37]. Furthermore, the abundant metabolic intermediates generated during glycolysis not only fulfill immediate energy demands but also provide essential biosynthetic precursors facilitating rapid proliferation and division of HCC cells. The comprehensive analysis combining proteomics and metabolomics conducted in this study further elucidated the intricate nature of energy metabolism in HCC. Among all detected metabolic pathways, pyruvate metabolism pathway holds significant importance for energy production in HCC with its enrichment ranking among the top 10 pathways identified. It is noteworthy that citric acid serves as a pivotal intersection point connecting multiple energy metabolism pathways [40, 41].

In this study, PI (6 keto-PGF1alpha/16:0), 13,16, 19-docosatrienoic acid, PA (18:2(9Z,12Z)/20:1(11Z)) and PG (20:3(6,8,11)-OH(5)/18:2(9Z,12Z)) exhibited significant up-regulation in HCC. They demonstrated a strong diagnostic ability for HCC (AUC > 0.90) and were found to be correlated with several key node proteins such as B4GALT5. Although no studies have been conducted on these four molecules specifically, they are likely to play a pivotal role in the rearrangement of lipid metabolism in HCC as representatives of different subclasses within the lipid metabolism network. Furthermore, multiple metabolomics studies on HCC have indicated an up-regulation of phosphatidylinositol, phosphatidic acid, glycerophospholipid, polyunsaturated fatty acids and their derivatives [12, 42, 43, 44]. These abnormal alterations in lipid metabolism may be closely associated with liver cancer cell proliferation and tumor angiogenesis. Reprogramming of amino acid metabolism is a crucial characteristic of tumor metabolic abnormalities. Research has revealed that amino acids not only serve as nitrogen donors to provide raw materials for the synthesis of biological macromolecules, such as nucleotides required for tumor cell proliferation, invasion, and immune evasion but also act as vital metabolic substances for immune cells to activate and exert an anti-tumor role within the tumor microenvironment. Abnormal alterations in amino acid metabolism are closely associated with the occurrence and progression of tumors and tumor immunity. Furthermore, certain key proteins or enzymes involved in these metabolic pathways can be utilized as biomarkers for tumor diagnosis and treatment [45, 46]. Among the top 10 metabolic pathways identified in KEGG enrichment analysis conducted in this study, 4 were associated with amino acid metabolism. These included tryptophan metabolism, alanine, aspartate and glutamate metabolism, arginine biosynthesis, and arginine and proline metabolism. Previous studies have also demonstrated the crucial role of these four metabolic pathways in HCC, specifically in cell proliferation, energy supply, and immune evasion. Notably, impaired tryptophan metabolism can hinder NAD + synthesis in the liver and promote liver tumorigenesis through DNA damage [47]. The metabolism of alanine, aspartate, and glutamate is not only associated with changes in early HCC patient characteristics but also linked to cell proliferation and intracellular signal transduction in HCC [45, 48]. Dirk et al. [49] demonstrated that arginine biosynthesis was enhanced in HCC patients, accompanied by RNA binding motif protein 39 (RBM39) regulation on metabolic-related genes, promoting the reprogramming of metabolic pathways in HCC. In this study, N2-acetylornithine was considered as a potential biomarker for HCC due to its critical role in arginine biosynthesis. The metabolism of arginine and proline is not only closely associated with ATP production, protein synthesis, and nucleotide synthesis in tumor cells, but also plays a role in maintaining intracellular REDOX homeostasis [50, 51]. The degradation of arginine during the urea cycle produces several metabolites including urea, ornithine, and polyamines. Polyamines are a class of small organic polycations that are crucial for DNA replication, protein translation, and cell proliferation [52]. Among them, spermidine-a key member of the polyamine family—exerts essential biochemical characteristics in controlling cell growth, regulating the

cell cycle, providing antioxidant protection mechanisms, facilitating cell signal transduction and gene expression regulation [53]. Therefore, spermidine is considered as a potential biomarker for HCC in this study.

This study conducted a comprehensive protein-metabolite analysis to explore the molecular characteristics of HCC. Our findings indicated the presence of abnormal changes in 10 pathways, including tryptophan metabolism, and suggested the potential of 12 biomarkers as candidates for HCC diagnosis or prognosis. These results suggested novel avenues for understanding the pathogenesis of HCC and could potentially contribute to the discovery of HCC-specific biomarkers. However, it is important to acknowledge certain limitations in our study. Firstly, this study was conducted at a single center with a relatively small sample size. The limited sample size may somewhat impact the generalizability of our findings. Secondly, we did not integrate the omics results with the clinical characteristics of HCC patients in this study. Further research should aim to incorporate clinicopathological features and patient prognosis data for a more comprehensive association analysis and evaluation of the clinical value of these biomarkers. Thirdly, the results of this study have not been validated in other cohorts, and extensive prospective cohort studies are required to further validate our findings.

Conclusions

The present study presented a comprehensive analysis of the intricate alterations in HCC at the proteome and metabolome levels through an integrated multi-omics investigation. We have identified 10 aberrant metabolic pathways and discovered 12 potential diagnostic markers, thereby providing novel insights into the pathogenesis of HCC and offering promising candidates for HCC biomarkers.

Functional analysis of DEMs

In order to determine the specific functions and roles of DEMs in biological processes, this study performed pathway annotation and enrichment analysis of 500 DEMs based on KEGG database. KEGG annotation results showed that 41% of the annotated pathways belonged to the metabolic category, in which lipid metabolism, amino acid metabolism, cofactor and vitamin metabolism, and carbohydrate metabolism had the largest number of metabolites (Fig. 6A). KEGG enrichment analysis showed that the DEMs were mainly enriched in glycerophospholipid metabolism, choline metabolism in cancer, phospholipase D signaling pathway, and aldosterone synthesis and secretion pathways (Fig. 5B). Among the top 20 pathways exhibiting the lowest adjusted *p*-value, the overall expression of DEMs involved in 8 specific pathways, including glycerophospholipid metabolism, choline metabolism in cancer, phospholipase D signaling pathway, and arginine and proline metabolism, tended to be up-regulated. Conversely, the overall expression of DEMs involved in another 12 pathways, such as aldosterone synthesis and secretion, arachidonic acid metabolism, arginine biosynthesis, and oxidative phosphorylation, exhibited a tendency for down-regulation (Fig. 6C).

Integrated analysis of proteomics and metabolomics

Spearman correlation analysis

In this study, both proteomics and metabolomics analyses were conducted on HCT and DNT samples from six HCC patients, and the resulting datasets were merged for a comprehensive two-omics analysis. The two-way orthogonal partial least squares (O2PLS) method was

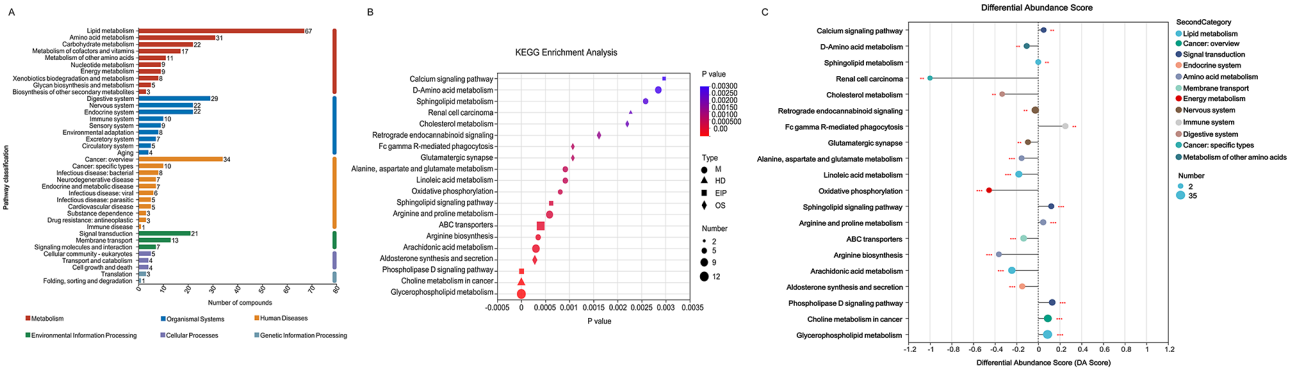

Fig. 6 Functional analysis of differentially expressed metabolites. (**A**) KEGG annotation analysis was performed to identify differentially expressed metabolites. (**B**) KEGG enrichment analysis was carried out to determine the significance of these differentially expressed metabolites in metabolic pathways. Abbreviations: M, metabolism; HD, human diseases; EIP, environmental information processing; OS, organismal systems. (**C**) Differential abundance score map of KEGG pathways. The horizontal axis represents the DA Score and the vertical axis indicates the name of each metabolic pathway in KEGG database. The DA Score reflects overall changes in all annotated metabolites within a given pathway, with a score of 1 indicating up-regulation and −1 indicating down-regulation for all annotated differential metabolites within that pathway. The length of line segments corresponds to absolute values of DA Scores, while dot sizes indicate numbers of differential metabolites annotated in each pathway

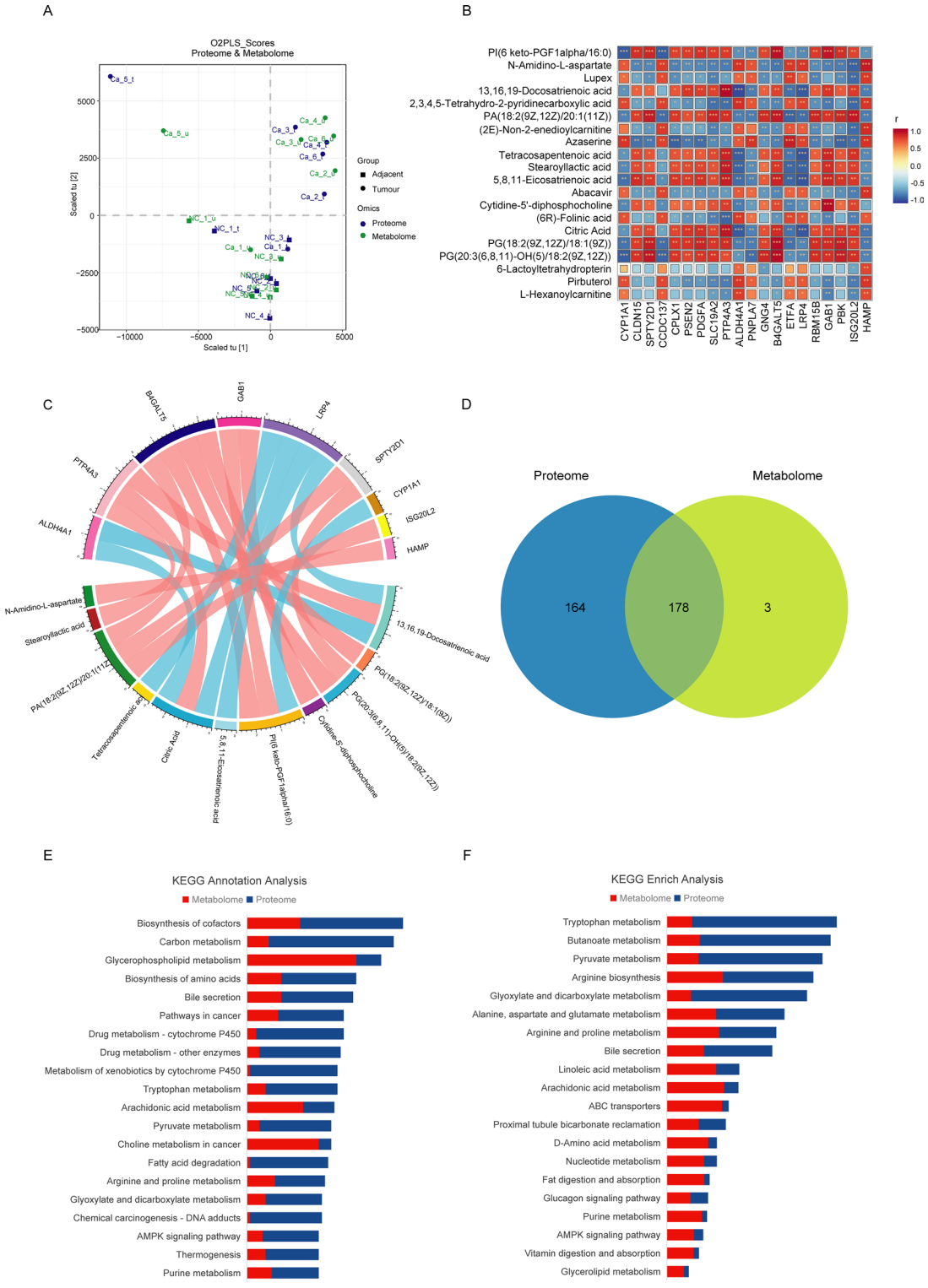

Fig. 7 Consistency analysis of differentially expressed proteins and metabolites. (**A**) Combined sample score plot of the O2PLS model for the protein-metabolism-set. The squares and dots represented samples from different groups. The colors of the dots indicated different omics samples, with blue representing protein samples (denoted by 't' after the sample name) and green representing metabolome samples (denoted by 'u' after the sample name). In this figure, the horizontal and vertical coordinates represented the combined score of the metabolome and proteome. (**B**) Correlation analysis on the top 20 differentially expressed proteins and metabolites between HCT and DNT. r represents the correlation value. (**C**) Chord diagram of the top 20 protein-metabolite pairs ranked by their correlation scores. Red represents positive correlation and blue represents negative correlation. (**D**) Venn diagram of differentially expressed proteins and metabolites involved in the KEGG pathway. (**E**) KEGG annotation analysis of differentially expressed proteins and metabolites. (**F**) KEGG enrichment analysis of differentially expressed proteins and metabolites

Table 3 Differentially expressed metabolites involved in the top 10 enrichment pathways

ID	Metabolite name	Adjusted *p*-value	Log_2 FC
C05660	5-Methoxyindoleacetate	0.008	0.354
C10164	Picolinic Acid	0.027	0.316
C01144	(S)-3-Hydroxybutyryl-CoA	< 0.001	-0.610
C05834	3-Methyldioxyindole	0.027	-0.256
C00108	2-Aminobenzoic Acid	0.021	-0.232
C05635	5-Hydroxyindoleacetic acid	0.021	-0.130
C00025	N-Choloylglycine	0.015	-0.272
C00025	L-Glutamic Acid	0.019	-0.340
C02630	2-Hydroxyglutarate	< 0.001	0.144
C00122	Fumaric acid	0.004	-0.112
C01384	Maleic Acid	0.011	-0.201
ID	Metabolite name	Adjusted *p*-value	Log_2 FC
C00546	APC	0.009	0.189
C00149	Malic acid	0.003	-0.077
C05993	Acetyl adenylate	0.021	-0.197
C00158	Citric Acid	< 0.001	0.068
C01182	D-ribulose-1,5-bisphosphate	0.001	-0.249
C00064	L-Glutamine	0.005	-0.378
C00077	L-Ornithine	0.009	-0.073
C00437	N2-Acetylornithine	0.002	0.334
C05466	Chenodeoxycholylglycine	0.043	-0.223
C02538	Estrone	< 0.001	-0.378
C00639	Dinoprost	0.016	-0.182
C01844	Pravastatin	0.003	-0.317
C01921	Glycocholic Acid	0.011	-0.228
C00315	Spermidine	< 0.001	0.090
C07151	1,1-Dimethylbiguanide	0.016	0.263
C03033	Phenethylamine glucuronide	0.005	-0.267
C01042	N-acetylaspartate	0.002	0.116
C04281	Pyrroline hydroxycarboxylic acid	0.013	-0.240
C00019	S-Adenosylmethionine	0.003	0.413
C01165	L-Glutamic gamma-semialdehyde	0.028	0.133
C04137	D-Octopine	0.002	0.248
C03415	N2-Succinyl-L-ornithine	0.027	0.182
C14767	9(S)-HODE	0.018	-0.073
C14827	9(S)-HpODE	0.048	-0.116
C04717	13-L-Hydroperoxylinoleic acid	0.007	-0.126
C04056	Octadeca-9,11-dienoic acid	0.003	-0.126
C00157	PC (22:6(4Z,7Z,10Z,13Z,16Z,19Z)/16:0)	< 0.001	-0.088

employed to assess the intrinsic correlation between the two datasets. The combined score plot shows great similarity between the two data matrices (Fig. 7A). To gain a more intuitive understanding of the co-regulatory relationship between DEPs and DEMs, spearman correlation coefficient calculations were performed between each DEP and DEM, resulting in a correlation coefficient matrix. Heat maps were then generated to visualize the correlation strengths, focusing on the top 20 DEPs from the proteomics dataset and top 20 DEMs from the metabolomics dataset, respectively, based on their smallest adjusted *p*-values. To gain a more intuitive understanding of the co-regulatory relationship between DEPs and DEMs, a correlation coefficient matrix was derived by computing the spearman correlation coefficients between each pair of DEPs and DEMs. This matrix was then visualized using heatmaps, which displayed the correlation strengths for the top 20 DEPs and top 20 DEMs with the smallest adjusted *p*-values from their respective omics datasets (Fig. 7B). Furthermore, within this matrix, the top 20 protein-metabolite pairs with the highest absolute correlation coefficients were selected, and their correlations were plotted using a chord diagram to illustrate the strong interactions (Fig. 7C). The results showed that B4GALT5 was positively correlated with four metabolites, including PG (18:2(9Z,12Z)/18:1(9Z)), PG

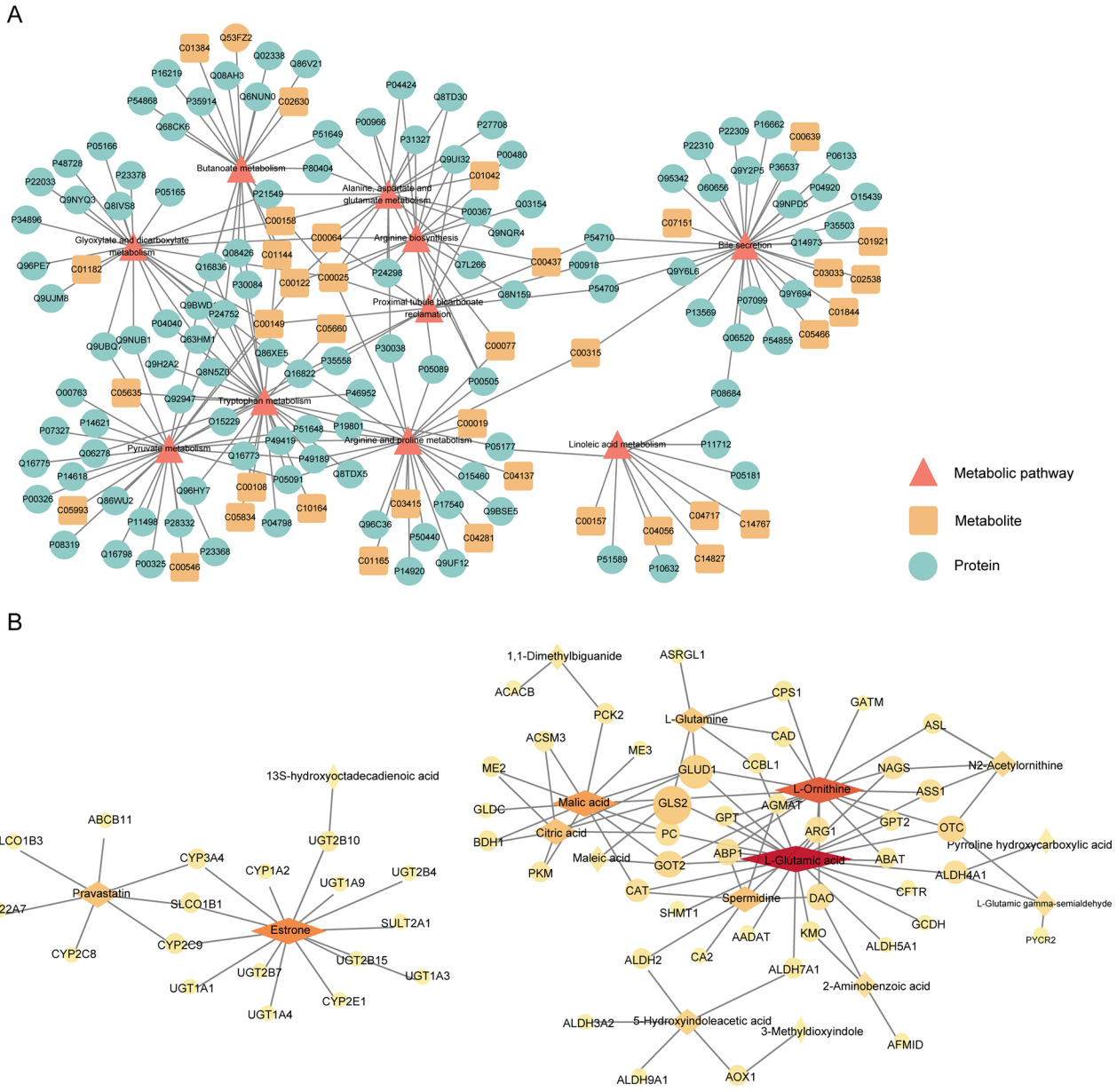

Fig. 8 Interaction analysis of the differentially expressed proteins and metabolites involved in the top 10 enrichment pathways. (**A**) Distribution of the differentially expressed proteins and metabolites involved in the top 10 enrichment pathways. (**B**) The interaction between the differentially expressed proteins and metabolites involved in the top 10 enrichment pathways. The diamonds represent metabolites, while the circles represent proteins. The size and color of the figures indicate the number of interacting molecules with them, with larger and darker figures representing more interacting molecules

(20:3(6,8,11)-OH(5)/18:2(9Z,12Z)), PI (6 keto-PGF1alpha/16:0), and PA (18:2(9Z,12Z)/20:1(11Z)), while LRP4 was negatively correlated with 5,8,11-Eicosatrienoic acid, Tetracosapentenoic acid, Citric Acid, and 13,16,19-Docosatrienoic acid.

Combined analysis of biological functions

Based on the KEGG database, this study compared the pathways involved by DEPs and DEMs and found that they jointly participate in 178 pathways (Fig. 7D). The KEGG annotation results indicated that the most significant pathways, considering the number of proteins and metabolites involved, include Bile secretion, Biosynthesis of amino acids, Glycerophospholipid metabolism, and Tryptophan metabolism (Fig. 7E). The KEGG enrichment analysis revealed that the top 10 pathways exhibiting significant enrichment including tryptophan metabolism, butanoate metabolism, pyruvate metabolism, and bile secretion. Notably, both tryptophan metabolism and bile secretion pathways demonstrated a high number of DEPs and DEMs as well as substantial degrees of enrichment (Fig. 7F, Additional file 1 and

Table 4 The diagnostic value of potential biomarkers

Name	FC	Adjusted *p*-value	AUC
Protein			
PTP4A3	22.53	1.657E-07	1.000
B4GALT5	926.6	5.159E-07	1.000
GAB1	49.76	7.712E-07	1.000
ME2	2.037	0.007719	0.944
PKM	2.344	0.04095	0.917
Metabolite			
PI (6 keto-PGF1alpha/16:0)	1.3273	1.514E-07	0.995
13,16,19-Docosatrienoic acid	1.2007	0.000001507	0.990
PA (18:2(9Z,12Z)/20:1(11Z))	1.0623	0.000003408	0.980
Citric Acid	1.0479	0.00002444	0.990
PG (20:3(6,8,11)-OH(5)/18:2(9Z,12Z))	1.136	0.00003883	0.960
Spermidine	1.0647	0.0001266	0.970
N2-Acetylornithine	1.2602	0.002112	0.900

Table 3). Subsequently, this study investigated the interaction between DEPs and DEMs in the top 10 enriched pathways. The results revealed that L-Glutamic acid, L-Ornithine, Malic acid, and Citric acid exhibited interactions with GLUD1 and other proteins. Notably, Citric acid demonstrated an up-regulation trend in HCC, while the other three metabolites displayed a down-regulation trend in HCC. Among these differential metabolite-interacting proteins, GLS2 showed the highest level of interaction and was found to be down-regulated in HCC (Fig. 8A and B).

Screening of potential biomarkers

To further identify potential biomarkers highly associated with the occurrence of HCC, the following screening criteria were formulated in this study: (1) The DEPs (DEMs) between HCT and DNT exhibited statistically significant differences (adjusted p-value < 0.05) and demonstrated a trend of high expression in HCT. (2) The absolute correlation coefficient between the DEPs and DEMs was greater than 0.85. (3) These biomarkers participated in the top 10 enriched pathways, and their interaction degree was greater than or equal to 2. (4) The area under curve (AUC) of these molecules for distinguishing HCC was greater than or equal to 0.90. Based on the above screening conditions and literature review results, a total of twelve potential biomarkers for HCC were identified in this study, including five proteins (PTP4A3, B4GALT5, GAB1, ME2, and PKM) and seven metabolites (PI(6 keto-PGF1alpha/16:0), 13,16,19-docosatrienoic acid, PA(18:2(9Z,12Z)/20:1(11Z)), Citric Acid, PG(20:3(6, 8, 11)-OH(5)/18: 2(9Z, 12Z)), Spermidine, and N2-Acetylornithine) (Table 4 and Additional file 2).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-13952-0.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

We would like to express our gratitude to all participants involved in this study, as well as Gansu Province and The Second Hospital of Lanzhou University for their generous financial support.

Author contributions

All authors contributed to the article and approved the submitted version. YY W and WX Y served as co-first authors, conceiving and designing the study. JY C and FF W collected the samples and analyzed the data. Chong-Ge You supervised the entire study and revised the manuscript.

Funding

This work was supported by the Cuiying Scientific and Technological Innovation Program of The Second Hospital and Clinical Medical School of Lanzhou University (grant number CY2022-QN-A18 and CY2021-BJ-A16), and Lanzhou science and Technology Development Guiding program project (grant number 2023-ZD-85).

Data availability

The original data sets is available on request from the corresponding author at youchg@lzu.edu.cn.

Declarations

Ethics approval and consent to participate

This study is strictly in accordance with the Declaration of Helsinki and was approved by the authors' Institutional Review Board (Medical Ethics Committee of The Second Hospital & Clinical Medical School, Lanzhou University, China; No. 2024 A-075). Informed consent was obtained from all participants for this study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Laboratory Medicine Center, The Second Hospital & Clinical Medical School, Lanzhou University, Lanzhou, Gansu Province 730030, China

Received: 17 October 2024 / Accepted: 17 March 2025
Published online: 30 March 2025

References

1. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. Nat Reviews Disease Primers. 2021;7(1):6. https://doi.org/10.1038/s41572-020-00240-3.
2. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Cancer J Clin. 2024;74(3):229–63. https://doi.org/10.3322/caac.21834.
3. Singal AG, Kanwal F, Llovet JM. Global trends in hepatocellular carcinoma epidemiology: implications for screening, prevention and therapy. Nat Reviews Clin Oncol. 2023;20(12):864–84. https://doi.org/10.1038/s41571-023-00825-3.
4. Sugawara Y, Hibi T. Surgical treatment of hepatocellular carcinoma. Biosci Trends. 2021;15(3):138–41. https://doi.org/10.5582/bst.2021.01094.
5. Moldogazieva NT, Mokhosoev IM, Zavadskiy SP, Terentiev AA. Proteomic profiling and artificial intelligence for hepatocellular carcinoma translational medicine. Biomedicines. 2021;9(2). https://doi.org/10.3390/biomedicines9020159.

6. Schmidt DR, Patel R, Kirsch DG, Lewis CA, Vander Heiden MG, Locasale JW. Metabolomics in cancer research and emerging applications in clinical oncology. Cancer J Clin. 2021;71(4):333–58. https://doi.org/10.3322/caac.21670.
7. Du Z, Liu X, Wei X, Luo H, Li P, Shi M, et al. Quantitative proteomics identifies a plasma multi-protein model for detection of hepatocellular carcinoma. Sci Rep. 2020;10(1). https://doi.org/10.1038/s41598-020-72510-9.
8. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, et al. Integrated proteogenomic characterization of HBV-Related hepatocellular carcinoma. Cell. 2019;179(2):561–77. https://doi.org/10.1016/j.cell.2019.08.052.
9. Jiang Y, Sun A, Zhao Y, Ying W, Sun H, Yang X, et al. Proteomics identifies new therapeutic targets of early-stage hepatocellular carcinoma. Nature. 2019;567(7747):257–61. https://doi.org/10.1038/s41586-019-0987-8.
10. Huang Q, Tan Y, Yin P, Ye G, Gao P, Lu X, et al. Metabolic characterization of hepatocellular carcinoma using nontargeted tissue metabolomics. Cancer Res. 2013;73(16):4992–5002. https://doi.org/10.1158/0008-5472.Can-13-0308.
11. Han J, Han MI, Xing H, Li Z, Dy Y, Wu H, et al. Tissue and serum metabolomic phenotyping for diagnosis and prognosis of hepatocellular carcinoma. Int J Cancer. 2019;146(6):1741–53. https://doi.org/10.1002/ijc.32599.
12. Ferrarini A, Di Poto C, He S, Tu C, Varghese RS, Kara Balla A, et al. Metabolomic analysis of liver tissues for characterization of hepatocellular carcinoma. J Proteome Res. 2019;18(8):3067–76. https://doi.org/10.1021/acs.jproteome.9b00185.
13. Fang C, Wang H, Lin Z, Liu X, Dong L, Jiang T, et al. Metabolic reprogramming and risk stratification of hepatocellular carcinoma studied by using gas Chromatography–Mass Spectrometry-Based metabolomics. Cancers. 2022;14(1). https://doi.org/10.3390/cancers14010231.
14. Rashid MM, Varghese RS, Ding Y, Ressom HW. Biomarker discovery for hepatocellular carcinoma in patients with liver cirrhosis using untargeted metabolomics and lipidomics studies. Metabolites. 2023;13(10). https://doi.org/10.3390/metabo13101047.
15. Tambay V, Raymond V-A, Goossens C, Rousseau L, Turcotte S, Bilodeau M. Metabolomics-Guided identification of a distinctive hepatocellular carcinoma signature. Cancers. 2023;15(12). https://doi.org/10.3390/cancers15123232.
16. Tang L, Zeng J, Geng P, Fang C, Wang Y, Sun M, et al. Global metabolic profiling identifies a pivotal role of proline and hydroxyproline metabolism in supporting hypoxic response in hepatocellular carcinoma. Clin Cancer Res. 2018;24(2):474–85. https://doi.org/10.1158/1078-0432.Ccr-17-1707.
17. Beyoğlu D, Imbeaud S, Maurhofer O, Bioulac-Sage P, Zucman-Rossi J, Dufour J-F, et al. Tissue metabolomics of hepatocellular carcinoma: tumor energy metabolism and the role of transcriptomic classification. Hepatology. 2013;58(1):229–38. https://doi.org/10.1002/hep.26350.
18. Menyhárt O, Győrffy B. Multi-omics approaches in cancer research with applications in tumor subtyping, prognosis, and diagnosis. Comput Struct Biotechnol J. 2021;19:949–60. https://doi.org/10.1016/j.csbj.2021.01.009.
19. Ng CKY, Dazert E, Boldanova T, Coto-Llerena M, Nuciforo S, Ercan C, et al. Integrative proteogenomic characterization of hepatocellular carcinoma across etiologies and stages. Nat Commun. 2022;13(1). https://doi.org/10.1038/s41467-022-29960-8.
20. Malinowska JM, Palosaari T, Sund J, Carpi D, Bouhifd M, Weber RJM, et al. Integrating in vitro metabolomics with a 96-well high-throughput screening platform. Metabolomics. 2022;18(1). https://doi.org/10.1007/s11306-021-01867-3.
21. Märtens A, Holle J, Mollenhauer B, Wegner A, Kirwan J, Hiller K. Instrumental drift in untargeted metabolomics: optimizing data quality with intrastudy QC samples. Metabolites. 2023;13(5). https://doi.org/10.3390/metabo13050665.
22. Chen Z, Yang C, Guo Z, Song S, Gao Y, Wang D, et al. A novel PDX modeling strategy and its application in metabolomics study for malignant pleural mesothelioma. BMC Cancer. 2021;21(1). https://doi.org/10.1186/s12885-021-08980-5.
23. Wang R, Kang H, Zhang X, Nie Q, Wang H, Wang C, et al. Urinary metabolomics for discovering metabolic biomarkers of bladder cancer by UPLC-MS. BMC Cancer. 2022;22(1). https://doi.org/10.1186/s12885-022-09318-5.
24. Wei M, Korotkov KV, Blackburn JS. Targeting phosphatases of regenerating liver (PRLs) in cancer. Pharmacol Ther. 2018;190:128–38. https://doi.org/10.1016/j.pharmthera.2018.05.014.
25. Lin C, Yu M, Wu X, Wang H, Wei M, Zhang L. Targeting moonlighting enzymes in cancer. Molecules. 2024;29(7). https://doi.org/10.3390/molecules29071573.
26. Zhou Q, Zhou Q, Liu Q, He Z, Yan Y, Lin J, et al. PRL-3 facilitates hepatocellular carcinoma progression by co-amplifying with and activating FAK. Theranostics. 2020;10(22):10345–59. https://doi.org/10.7150/thno.42069.
27. Han Y, Li Z, Wu Q, Liu H, Sun Z, Wu Y, et al. B4GALT5 high expression associated with poor prognosis of hepatocellular carcinoma. BMC Cancer. 2022;22(1):392. https://doi.org/10.1186/s12885-022-09442-2.
28. Wang D, Zhu ZZ, Jiang H, Zhu J, Cong WM, Wen BJ, et al. Multiple genes identified as targets for 20q13.12-13.33 gain contributing to unfavorable clinical outcomes in patients with hepatocellular carcinoma. Hep Intl. 2015;9(3):438–46. https://doi.org/10.1007/s12072-015-9642-0.
29. Pérez-Baena MJ, Cordero-Pérez FJ, Pérez-Losada J, Holgado-Madruga M. The role of GAB1 in cancer. Cancers (Basel). 2023;15(16). https://doi.org/10.3390/cancers15164179.
30. Boetto J, Bielle F, Tran S, Marijon P, Peyre M, Rigau V, et al. GAB1 as a marker of recurrence in anterior skull base meningioma. Neurosurgery. 2023;92(2):391–7. https://doi.org/10.1227/neu.0000000000002209.
31. Furuta K, Yoshida Y, Ogura S, Kurahashi T, Kizu T, Maeda S, et al. Gab1 adaptor protein acts as a gatekeeper to balance hepatocyte death and proliferation during acetaminophen-induced liver injury in mice. Hepatology. 2016;63(4):1340–55. https://doi.org/10.1002/hep.28410.
32. Wang X, Peng J, Yang Z, Zhou PJ, An N, Wei L, et al. Elevated expression of Gab1 promotes breast cancer metastasis by dissociating the PAR complex. J Experimental Clin cancer Research: CR. 2019;38(1):27. https://doi.org/10.1186/s13046-019-1025-2.
33. Sarfraz I, Rasul A, Hussain G, Hussain SM, Ahmad M, Nageen B, et al. Malic enzyme 2 as a potential therapeutic drug target for cancer. IUBMB Life. 2018;70(11):1076–83. https://doi.org/10.1002/iub.1930.
34. Zhang S, Cheng ZM, Yu JL, Lu K, Xu SJ, Lu Y, et al. Malic enzyme 2 promotes the progression of hepatocellular carcinoma via increasing triglyceride production. Cancer Med. 2021;10(19):6795–806. https://doi.org/10.1002/cam4.4209.
35. Zhou W, Zhang S, Cai Z, Gao F, Deng W, Wen Y, et al. A glycolysis-related gene pairs signature predicts prognosis in patients with hepatocellular carcinoma. PeerJ. 2020;8:e9944. https://doi.org/10.7717/peerj.9944.
36. Zhu S, Guo Y, Zhang X, Liu H, Yin M, Chen X, et al. Pyruvate kinase M2 (PKM2) in cancer and cancer therapeutics. Cancer Lett. 2021;503:240–8. https://doi.org/10.1016/j.canlet.2020.11.018.
37. Fukushi A, Kim HD, Chang YC, Kim CH. Revisited metabolic control and reprogramming cancers by means of the Warburg effect in tumor cells. Int J Mol Sci. 2022;23(17). https://doi.org/10.3390/ijms231710037.
38. Zhang S, Liao Z, Li S, Luo Y. Non-metabolic enzyme function of PKM2 in hepatocellular carcinoma: A review. Med (Baltim). 2023;102(42):e35571. https://doi.org/10.1097/MD.0000000000035571.
39. Yang F, Hilakivi-Clarke L, Shaha A, Wang Y, Wang X, Deng Y, et al. Metabolic reprogramming and its clinical implication for liver cancer. Hepatology. 2023;78(5):1602–24. https://doi.org/10.1097/HEP.0000000000000005.
40. Hu B, Lin JZ, Yang XB, Sang XT. Aberrant lipid metabolism in hepatocellular carcinoma cells as well as immune microenvironment: A review. Cell Prolif. 2020;53(3):e12772. https://doi.org/10.1111/cpr.12772.
41. Jin HR, Wang J, Wang ZJ, Xi MJ, Xia BH, Deng K, et al. Lipid metabolic reprogramming in tumor microenvironment: from mechanisms to therapeutics. J Hematol Oncol. 2023;16(1):103. https://doi.org/10.1186/s13045-023-01498-2.
42. Buchard B, Teilhet C, Abeywickrama Samarakoon N, Massoulier S, Joubert-Zakeyh J, Blouin C, et al. Two metabolomics phenotypes of human hepatocellular carcinoma in Non-Alcoholic fatty liver disease according to fibrosis severity. Metabolites. 2021;11(1). https://doi.org/10.3390/metabo11010054.
43. Ismail IT, Fiehn O, Elfert A, Helal M, Salama I, El-Said H. Sugar alcohols have a key role in pathogenesis of chronic liver disease and hepatocellular carcinoma in whole blood and liver tissues. Cancers (Basel). 2020;12(2). https://doi.org/10.3390/cancers12020484.
44. Tambay V, Raymond VA, Goossens C, Rousseau L, Turcotte S, Bilodeau M. Metabolomics-Guided identification of a distinctive hepatocellular carcinoma signature. Cancers (Basel). 2023;15(12). https://doi.org/10.3390/cancers15123232.
45. Du D, Liu C, Qin M, Zhang X, Xi T, Yuan S, et al. Metabolic dysregulation and emerging therapeutical targets for hepatocellular carcinoma. Acta Pharm Sin B. 2022;12(2):558–80. https://doi.org/10.1016/j.apsb.2021.09.019.
46. Foglia B, Beltrà M, Sutti S, Cannito S. Metabolic reprogramming of HCC: A new microenvironment for immune responses. Int J Mol Sci. 2023;24(8). https://doi.org/10.3390/ijms24087463.
47. Tummala KS, Gomes AL, Yilmaz M, Graña O, Bakiri L, Ruppen I, et al. Inhibition of de Novo NAD(+) synthesis by oncogenic URI causes liver tumorigenesis through DNA damage. Cancer Cell. 2014;26(6):826–39. https://doi.org/10.1016/j.ccell.2014.10.002.

48. Zhang F, Wu J, Zhang L, Zhang J, Yang R. Alterations in serum metabolic profiles of early-stage hepatocellular carcinoma patients after radiofrequency ablation therapy. J Pharm Biomed Anal. 2024;243:116073. https://doi.org/10.1016/j.jpba.2024.116073.
49. Mossmann D, Müller C, Park S, Ryback B, Colombi M, Ritter N, et al. Arginine reprograms metabolism in liver cancer via RBM39. Cell. 2023;186(23):5068–83. https://doi.org/10.1016/j.cell.2023.09.011.
50. Chen CL, Hsu SC, Ann DK, Yen Y, Kung HJ. Arginine signaling and cancer metabolism. Cancers (Basel). 2021;13(14). https://doi.org/10.3390/cancers13143541.
51. Geng P, Qin W, Xu G. Proline metabolism in cancer. Amino Acids. 2021;53(12):1769–77. https://doi.org/10.1007/s00726-021-03060-1.
52. Murthy D, Attri KS, Suresh V, Rajacharya GH, Valenzuela CA, Thakur R, et al. The MUC1-HIF-1α signaling axis regulates pancreatic cancer pathogenesis through polyamine metabolism remodeling. Proc Natl Acad Sci U S A. 2024;121(14):e2315509121. https://doi.org/10.1073/pnas.2315509121.
53. Fan J, Feng Z, Chen N. Spermidine as a target for cancer therapy. Pharmacol Res. 2020. https://doi.org/10.1016/j.phrs.2020.104943.

Publisher's note