BMC Gastroenterology

RESEARCH Open Access

Integration of bile proteomics and metabolomics analyses reveals novel insights into different types of gallstones in a high-altitude area

Xiaofeng Jing[1†], Ying Ma[2†], Defu Li[2], Tiecheng Zhang[1], Haiqi Xiang[1], Fan Xu[1*] and Yonghong Xia[2*]

Abstract

Background To explore the pathogenesis of different subtypes of gallstones in high-altitude populations from a molecular perspective.

Methods We collected bile samples from 20 cholesterol gallstone disease (CGD) patients and 20 pigment gallstone disease (PGD) patients. Proteomics analysis was performed by LC/MS DIA, while metabolomics analysis was performed by UPLC- Q-TOF/MS.

Results We identified 154 up-regulated and 196 down-regulated differentially expressed proteins, which were significantly enriched in neurodegenerative diseases, energy metabolism, amino acid metabolism etc. In metabolomics analysis, 20 up-regulated and 63 down-regulated differentially expressed metabolites were identified, and they were significantly enriched in vitamin B6 metabolism. Three pathways of integrated proteomics and metabolomics were significantly enriched: porphyrin and chlorophyll metabolism, riboflavin metabolism and aminoacyl-tRNA biosynthesis. Remarkably, 7 differentially expressed proteins and metabolites showed excellent predictive performance and were selected as potential biomarkers.

Conclusion The findings of our metabolomics and proteomics analyses help to elucidate the underlying mechanisms of gallstone formation in high-altitude populations.

Keywords Bile, Cholesterol gallstone, Pigment gallstone, High altitude, Metabolomics, Proteomics

†Xiaofeng Jing and Ying Ma these authors contributed equally to this work.

*Correspondence:
Fan Xu
xufan@cmc.edu.cn
Yonghong Xia
XyhQhsjtyy@126.com
[1]Department of Evidence-Based Medicine and Social Medicine, School of Public Health, Chengdu Medical College, Xindu avenue 783, Chengdu, Sichuan 610500, China
[2]Department of Hepatobiliary Surgery, Qinghai Provincial Traffic Hospital, Traffic lane 7, Xining 810001, China

Introduction

Gallstones, also known as cholelithiasis, are a common digestive tract disorder. While the prevalence of cholelithiasis in the United States is remarkably high at 20%, Qinghai Province in western China exhibits a prevalence rate similar to that of developed countries, approximately 15–20% [1, 2]. Due to the high-cholesterol diet prevalent in high-altitude regions, the incidence of gallstones is significantly higher than in flatland areas [3]. A major characteristic of gallstones is that not all patients experience symptoms. Sometimes, the stones may migrate near the opening of the bile duct and obstruct bile flow, potentially leading to gallbladder tension and typical biliary colic [4, 5]. Regarding the physical examination, most patients with chronic cholecystitis or gallstones may have no positive signs, while a few patients may exhibit tenderness or pain upon palpation in the right upper abdomen [6].

Currently, laparoscopic cholecystectomy is the standard treatment method [7]. However, laparoscopic cholecystectomy-associated complications such as postoperative bile duct injury, incisional hernia, and abdominal pain cannot be overlooked within clinical practice [8, 9]. Furthermore, drug treatment often leads to high recurrence rates by merely addressing symptoms rather than tackling the root cause, consequently causing physical and psychological distress for patients [10–12]. Ultrasonography is the most common imaging modality for diagnosing gallstones, but imaging methods can only identify patients who have already been affected [4, 13]. Therefore, identifying potential gallstone disease patients from a molecular perspective and intervening early can help reduce the incidence rate in the region.

Bile is a unique and vital aqueous secretion of the liver that is formed by the hepatocyte and modified downstream by absorptive and secretory properties of the bile duct epithelium [14]. When gallstones are formed, the bile reflects the changes in the pathogenesis of the body. Excessive liver cholesterol secretion will cause cholesterol supersaturation in gallbladder bile, which may lead to acute and chronic cholecystitis and common bile duct stones [15, 16]. Previous studies reported that osteopontin (OPN) in bile may participate in the formation of cholesterol gallstones by regulating the metabolism of cholesterol and bile acids, altering bile composition [17]. Although molecular diagnostic techniques can predict the risk of gallstones in patients, their clinical application remains unclear.

With the rapid development of chromatographic technology, liquid chromatography-mass spectrometry (LC-MS) data-independent acquisition (DIA) have become the mainstream methods for the separation and detection of proteomics [18, 19], which have the characteristics of high sensitivity, good specificity and a low detection limit. Our research intends to combine proteomics and metabolomics of bile to find possible biomarkers and construct a prediction model to provide a theoretical basis for early risk identification and prediction of cholelithiasis.

Method

Ethics statement

This study was conducted in accordance with the Declaration of Helsinki. All patients provided written informed consent. This study was approved by the Ethics Committee of the Qinghai Provincial Traffic Hospital, Qinghai, China ([2022] no. 727).

Subjects

Twenty patients with cholesterol-type cholelithiasis in Qinghai Provincial Traffic Hospital from June to November 2023 were screened as the case group, and 20 patients with bile pigment-type gallstones were included as the control group; their bile specimens and basic information were collected. The inclusion and exclusion criteria for cholelithiasis patients were as follows. Inclusion criteria: cholelithiasis diagnosed by abdominal imaging examination, abdominal pain of varying degrees at admission with or without fever and jaundice etc., informed about the study and signed the consent form, good cognitive function and ability to express themselves. Exclusion criteria: combined hepatic and renal insufficiency, the presence of biliary or pancreatic tumours, patients with psychiatric diseases, people with communication disorders and poor cognitive expression and people with cardiopulmonary dysfunction. The basic characteristics of the two groups were comparable. Two original bile samples were combined into one test sample for LC-MS DIA proteomics analysis, considering the influence of gender and age.

Overall design

This study included 20 cholesterol gallstone disease (CGD) patients and 20 pigment gallstone disease (PGD) patients aged 24 to 71 years. Their bile samples were collected and stored, and a quantitative analysis was then performed by LC-MS and UPLC-Q-TOF/MS. The pathogenesis of cholelithiasis was subsequently studied through multi-omics with a visual analysis. The details are shown in Fig. 1.

Bile sample collection

Bile samples from total 40 patients in two groups were collected after surgical section. During the surgical progress, a portion from the total 10–20 mL of bile was aspirated into a 10 mL test tube under sterile conditions and stored in a refrigerator at −80 ℃.

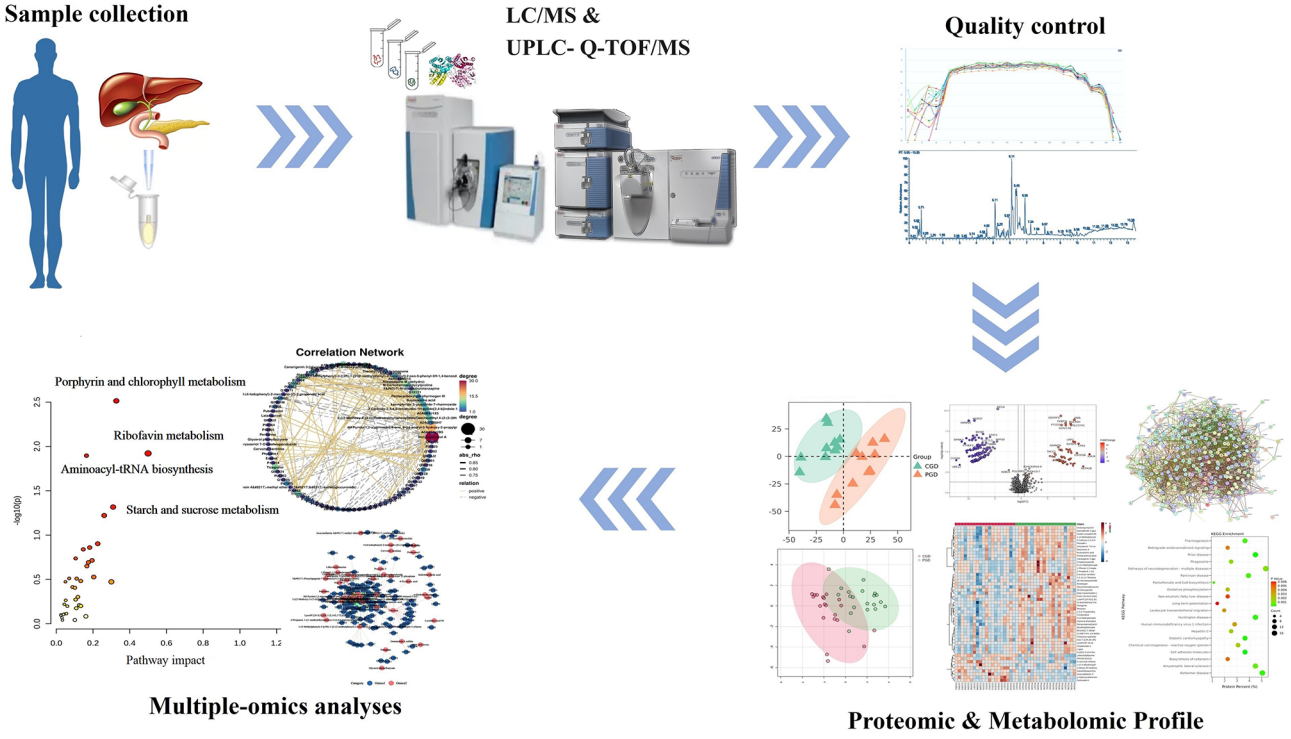


Fig. 1 General workflow. LC-MS, liquid chromatography–mass spectrometry. UPLC- Q-TOF/MS, ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry

Proteomic analysis by LC/MS DIA

Proteomic analysis was performed on 20 pairs of combined bile samples. The LC-MS data-independent acquisition (DIA) method was used for the untargeted proteome. The proteomics analysis was conducted by Shanghai Shisuan Biotechnology Co., Ltd. The experimental protocol on sample pre-processing, DIA data collection and database retrieval is provided in the Supplementary material.

Proteomic data analysis

Partial least squares discriminant analysis (PLS-DA) was applied for the discrimination among CGD and PGD patients of differentially expressed proteins (DEP). Screening of DEPs was to select samples for comparison according to the experimental design (pairwise comparison analysis was performed by default), and the average relative quantitative value of each protein in each group of samples in the comparison sample pair was calculated. The ratio of the mean value of each group of samples in the comparison sample pair was the fold change (FC). At the same time, the t-test was used to calculate the p-value to judge the significance of the difference. DEP was defined as FC>1.5 & FC<0.67, $P<0.05$, false discovery rate (FDR)correction was applied to control the expected proportion of false positives in all significant results. DEP pathway enrichment analysis and geneset enrichment analysis were performed using the KEGG pathway database and GO resource [20–22]. Protein-protein interaction (PPI) network analysis of DEP was constructed by STRING (https://string-db.org) [23].

Metabolomic analysis by UPLC-Q-TOF/MS

Metabolomic analysis of the 40 samples was conducted using the UPLC-Q-TOF/MS (Ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry) analysis platform. Initially, the samples underwent preprocessing to remove proteins and extract metabolites. Subsequently, mass spectrometry detection was performed separately in positive- and negative-ion modes. The resulting data underwent metabolite identification to generate a data matrix containing retention time, peak area, mass-to-charge ratio and identification information. The metabolomics analysis was conducted by Shanghai Shisuan Biotechnology Co., Ltd. Details are shown in the Supplementary files.

Metabolomic data analysis

The data matrix was uploaded into the MetaboAnalyst for further analysis (www.metaboanalyst.ca). They were subjected to partial least square discriminant analysis (PLS-DA) based on metabolite effects. We calculated the VIP (variable importance in the projection) value to determine its contribution to the classification. The metabolites with VIP>1 and $P<0.05$ and FC>1.5 or FC<0.67 were considered to be differentially expressed

metabolites (DEMs), false discovery rate (FDR)correction was also applied to control the expected proportion of false positives in all significant results. The metabolic pathways enrichment of DEMs was performed by searching the KEGG database [20].

Correlation analysis of proteomics and metabolomics

The correlations between DEPs and DEMs were identified by Pearson correlation analysis. The correlation network was constructed based on a correlation coefficient > 0.7 & $P < 0.05$. KEGG pathway enrichment analysis was conducted on DEPs and DEMs performed using MetaboAnalyst (www.metaboanalyst.ca) [24]. Combined biomarkers from proteomics and metabolomics were identified using a random forest model, and their predictive performance was assessed through 10-fold cross-validation and receiver operating characteristic curve analysis [25, 26].

Statistical analysis

All the statistical analyses were performed using StataMP 17.0. Continuous data with a normal distribution were shown as mean ± sd, and comparisons between groups were analysed using a *t*-test. Non-normally distributed continuous data were shown as M (P_{25}, P_{75}), and comparisons between groups were analysed using the Wilcoxon rank-sum test. Categorical data are shown as N (%), and comparisons between groups were analysed using a chi-square test. A P value < 0.05 was considered statistically significant.

Results

We collected information on the clinical characteristics of two types of gallstone disease patients. The result showed that only GgT levels were significantly different between the two groups, while other demographic and clinical information did not exhibit significant differences (Table 1).

Table 1 Clinical characteristic of CGD and PGD patients [(mean ± sd) / M(P_{25},P_{75})]. Neu: neutrophile granulocyte; Tcho: total cholesterol; Ldlc: low density lipoprotein cholesterol; Hdlc: high density lipoprotein cholesterol; wbc: white blood cell; Neut: neutrophilic granulocyte percentage; tg: triglyceride; Tbil: total bilirubin; Dbil: direct bilirubin; Indbil: indirect bilirubin; Ggt: gamma-glutamyl transpeptidase; alp: alkaline phosphatase; Alt: glutamic-pyruvic transaminase; ast: aspartate aminotransferase; tba: total bile acid

	CGD	PGD	t/z/x^2	*P*
Sex			0.100	0.752
male	10	9		
female	10	11		
Ethnic			4.206	0.240
Han	13	12		
Zhang	5	3		
Hui	1	5		
Mongol	1	0		
age	44.5 ± 10.4	48.45 ± 11.44	-1.143	0.260
bmi	24.47 ± 4.2	24.05 ± 2.76	0.380	0.705
Gra	65.95 ± 9.82	66.68 ± 10.87	-0.223	0.824
Tcho	3.99 ± 1.15	3.6 ± 1.69	0.846	0.402
Ldl-c	2.29(1.94,2.51)	2.6(2.29,2.94)	-1.677	0.094
Hdl-c	1.24(1.19,1.57)	1.62(1.31,3.63)	-1.746	0.081
Wbcc	6.82(5.22,7.6)	5.43(3.91,6.84)	1.758	0.079
Neut	4.2(3.09,5.19)	3.48(2.32,5.29)	1.447	0.148
Tg	1.88(1.1,2.39)	1.41(1.11,1.99)	0.757	0.449
Tbil	17.45(14.15,24.1)	18(14.05,20.8)	0.068	0.946
Dbil	5.7(4.65,8)	5.5(4.9,7.05)	0.027	0.978
Indbil	11.6(9.3,16.1)	12.5(8.7,14.45)	0.881	0.888
Ggt	45(25,133.5)	24.5(17.5,40.5)	2.369	0.017*
Alp	99(83.5,117.5)	94.5(69,102)	1.272	0.203
Alt	27.5(19,41.5)	25(17.5,40.5)	0.528	0.597
Ast	21(19,28)	21(17,24)	0.705	0.481
Tba	4.4(2.65,8.3)	3.85(2.65,7.3)	0.501	0.616

Neu: neutrophile granulocyte; Tcho: total cholesterol; Ldlc: low density lipoprotein cholesterol; Hdlc: high density lipoprotein cholesterol; Wbc: white blood cell; Neut: neutrophilic granulocyte percentage; Tg: triglyceride; Tbil: total bilirubin; Dbil: direct bilirubin; Indbil: Indirect bilirubin; Ggt: gamma-glutamyl transpeptidase; Alp: alkaline phosphatase; Alt: glutamic-pyruvic transaminase; Ast: aspartate aminotransferase; Tba: total bile acid

Proteomics

The PLS-DA score plot revealed a clear separation between CGD patients and PGD patients (Fig. 2a); a combination of the first principal component (PC1) and second principal component (PC2) could explain 86.75% of the variance. We further employed the 5-fold cross-validation method, the results indicated that the accuracy across the five folds is all above 0.9. The Q2 values range from a minimum of 0.559 to a maximum of 0.815 across different folds (all >0.5), suggesting a strong predictive capability of the model (Fig. 2b). A volcano plot was also obtained by plotting $\log_2$FC as the abscissa and the negative $\log_{10}$-based *P-value* ($-\log_{10}p$) as the ordinate after the proteins in the CGD and PGD groups were further analysed using *t*-test. Under the condition of FC>1.5 & FC<0.67, $P<0.05$, a total of 154 up-regulated and 196 down-regulated DEP were discovered in CGD patients compared to PGD patients (Fig. 2c). Hierarchical clustering analysis of the Top 100 DEPs, between the CGD and PGD groups, showed that the expression patterns in the CGD group differed from those in the PGD group, and protein expression in each group was clustered (Fig. 2d).

In total, 350 dysregulated DEPs were identified between the two groups, including 154 up-regulated and 196 down-regulated proteins. For biological processes, these proteins were mainly involved in cellular processes, metabolic processes, response to stimulus, dephosphorylation and anabolism of purines (Fig. 3a). For molecular functions, the proteins were primarily involved in binding and catalytic activity (Fig. 3b). For the cellular component, the proteins were primarily involved in the cellular anatomical entity, organelle and membrane (Fig. 3c). The results of KEGG pathway analyses revealed that these DEPs were significantly enriched in pathways of neurodegenerative diseases, such as prion disease and Parkinson's disease, energy metabolism, such as thermogenesis and amino acid metabolism, such as pantothenate and CoA biosynthesis (Fig. 3d). PPI network analyses of the DEPs were obtained from the STRING 11.0 database (Fig. 4). PPI network analysis systematically analyses the interactions among many proteins in biological systems.

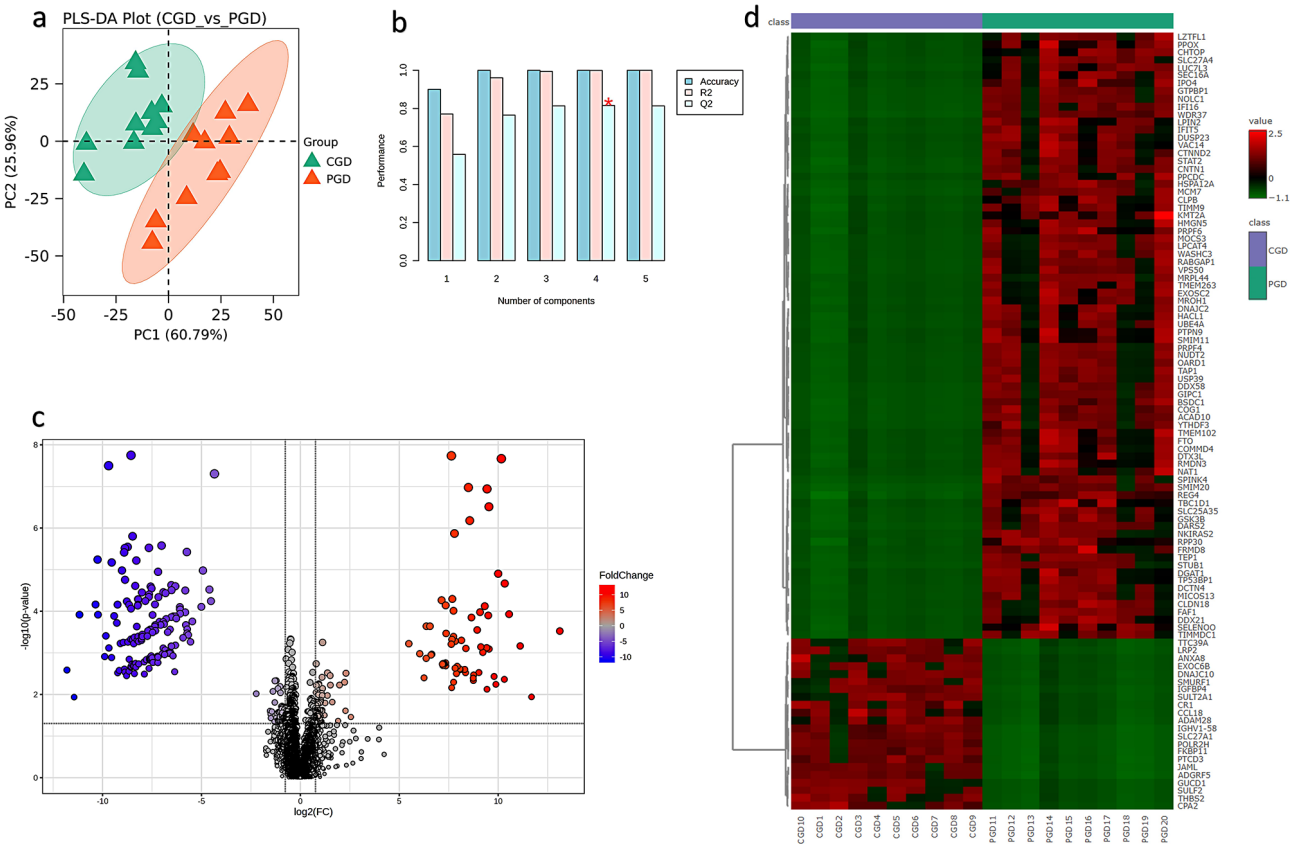


Fig. 2 Differential analysis of protein expression levels between CGD and PGD **a**) Partial least squares-discriminant analysis (PLS-DA) of proteomic data in CGD and PGD. **b**) cross-validation for PLS-DA. Q2 is an estimate of the predictive ability of the model, and is calculated via cross-validation **c**) Volcano plot of all proteins. The abscessus represents log2 (Fold Change), the ordinate represents -log10 (p-value), and each point represents a protein. Red dots represent proteins with a significant fold change (FC) > 1.5 ($p < 0.05$); blue dots represent proteins with a significant FC < 0.67 ($p < 0.05$); black dots, no obvious changes in the protein level. **d)** HCA showing the top 100 differentiated expressed proteins (DEPs). The red and green colors in the heatmap denote higher gene expression and lower gene expression, respectively

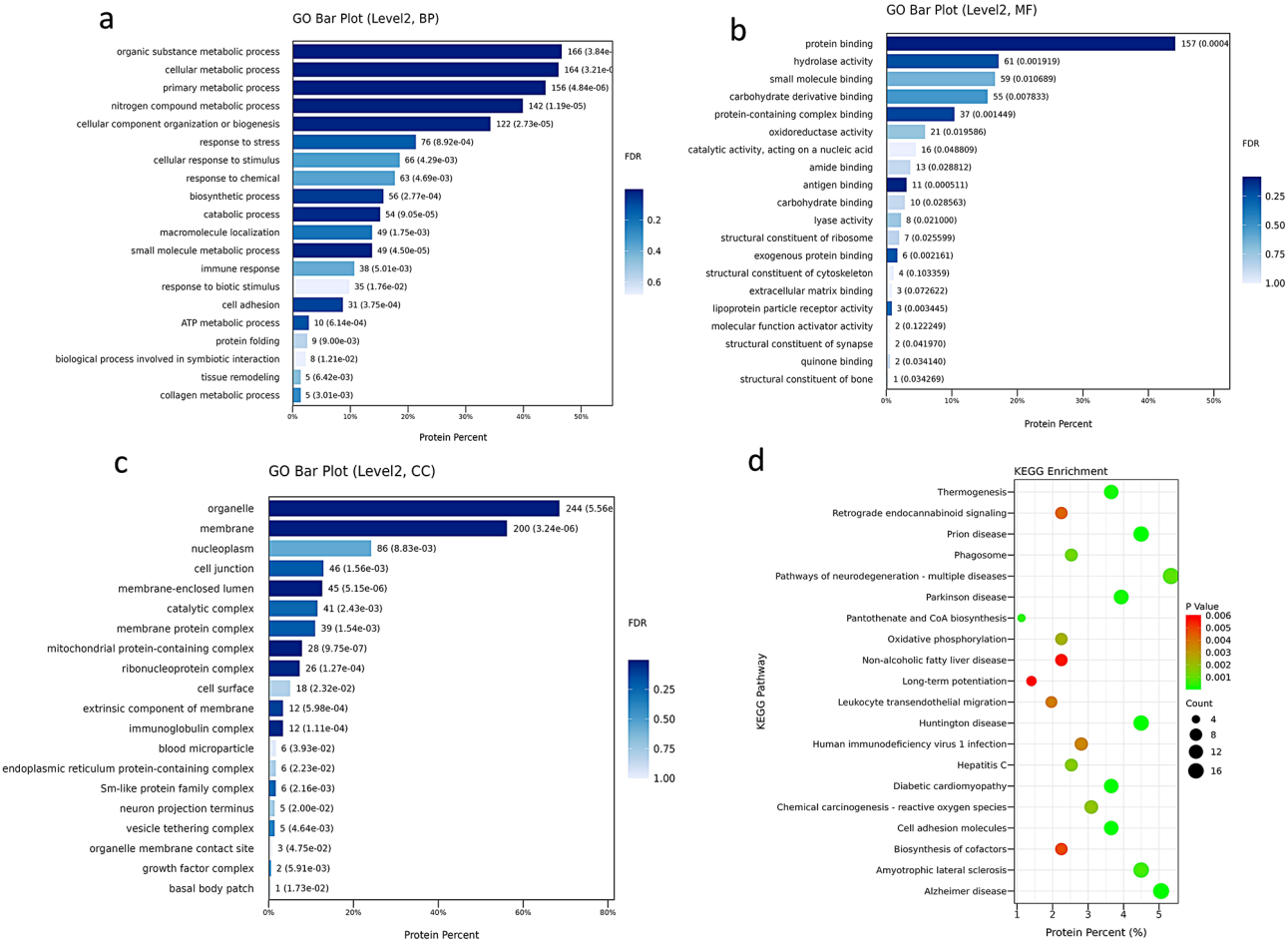


Fig. 3 Gene Ontology annotation and KEGG enrichment analysis of differentiated expressed proteins (DEPs). **a**) Biological processes (BP) functional annotation analysis of top 20 DEPs. **b**) Molecular functions (MF) functional annotation analysis of top 20 DEPs. **c**) Cellular component (CC) functional annotation analysis of top 20 DEPs. **d**) Kyoto Encyclopedia of Genes and Genomes (KEGG) functional annotation analysis of top 20 DEPs

In sum, enrichment analyses of DEPs were performed to investigate the potential correlations between them, and the results showed either strong or weak correlations between these proteins.

Metabolomics

Twenty bile samples per group were detected using UPLC-Q-TOF/MS, and the Human Metabolome Database and METLIN Metabolite Database were used to identify DEMs [27, 28]. A total of 1,504 metabolites were identified, including 895 in positive-ion mode and 610 in negative-ion mode. The PLS-DA model showed the separation between the two groups of bile samples (Fig. 5a). We further employed a permutation test to evaluate the accuracy of the PLS-DA model (Fig. 5b). The R2Y in the model was close to 1, indicating that the established model conforms well to the true situation of the sample data. The Q2 in the model was also close to 1, suggesting that if new samples were added to the model, a similar distribution can be obtained. Overall, the original model can effectively explain the differences between the two sets of samples. Metabolites with significant differences between the two groups were screened by fold change, and the *P*-value of the *t*-test (fold change > 2 or fold change < 0.67, and $p < 0.05$). The volcano plot showed that a total of 20 up-regulated metabolites and 63 down-regulated metabolites were significantly different (Fig. 5d). The cluster heatmap revealed clustering patterns of both total and selected metabolites in positive and negative ion modes (Fig. 5c). KEGG pathway analysis showed significant enrichment in one metabolic pathway, vitamin B6 metabolism (Fig. 5e).

Proteomic and metabolomic association analysis

We next performed integrated analysis of differential metabolites and proteins between the two groups of gallstone patients, to further evaluate the association between them. Three KEGG pathways were found to be involved in both metabolomics and proteomics, including porphyrin and chlorophyll metabolism, riboflavin

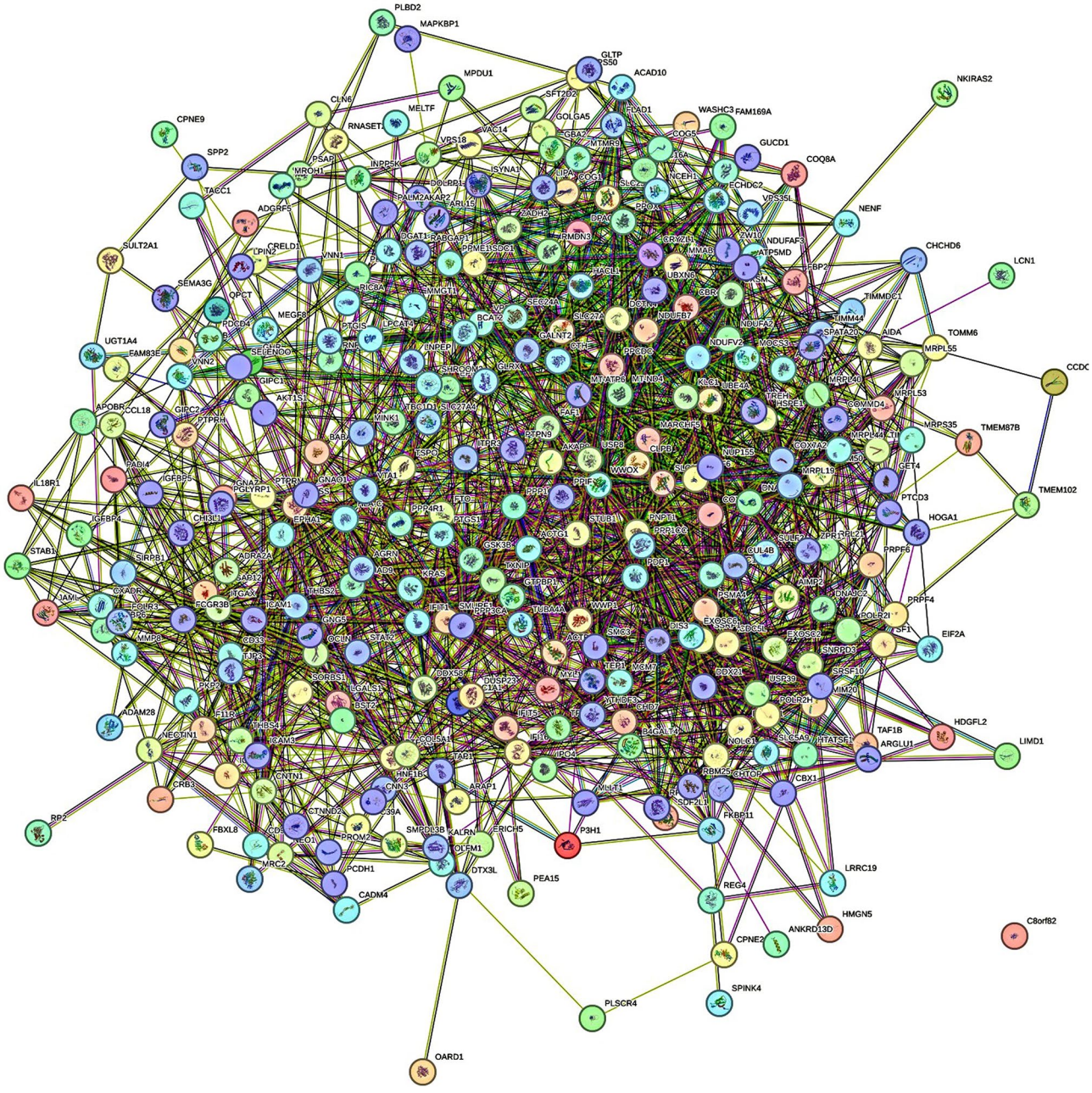


Fig. 4 Protein-protein interaction (PPI) network analysis. Every nodes represent a protein, red color represents the queried protein, and the other colors represent other proteins that have an interaction relationship with the queried protein

metabolism, and aminoacyl-tRNA biosynthesis (Fig. 6c). We performed a Pearson correlation heatmap to objectively reflect the expression patterns of DEPs and DEMs, which identified relatively strong or weak protein-metabolites correlations, and similar expression patterns involved in a similar biological process were grouped into the same cluster (Fig. 6a). Besides, we constructed a correlation network graph to demonstrate the main interactions of DEPs and DEMs ($r>0.7$, $p<0.05$). The correlation network displayed complex correlations between differential proteins and metabolites (Fig. 6b).

Screening of biomarkers by machine learning

We further employed the random forest model and ROC analysis to identify potential biomarkers and validate the associations between proteomics and metabolomics. By screening and selecting based on the mean decrease accuracy index, we identified seven DEPs and DEMs as potential biomarkers which exhibited good predictive

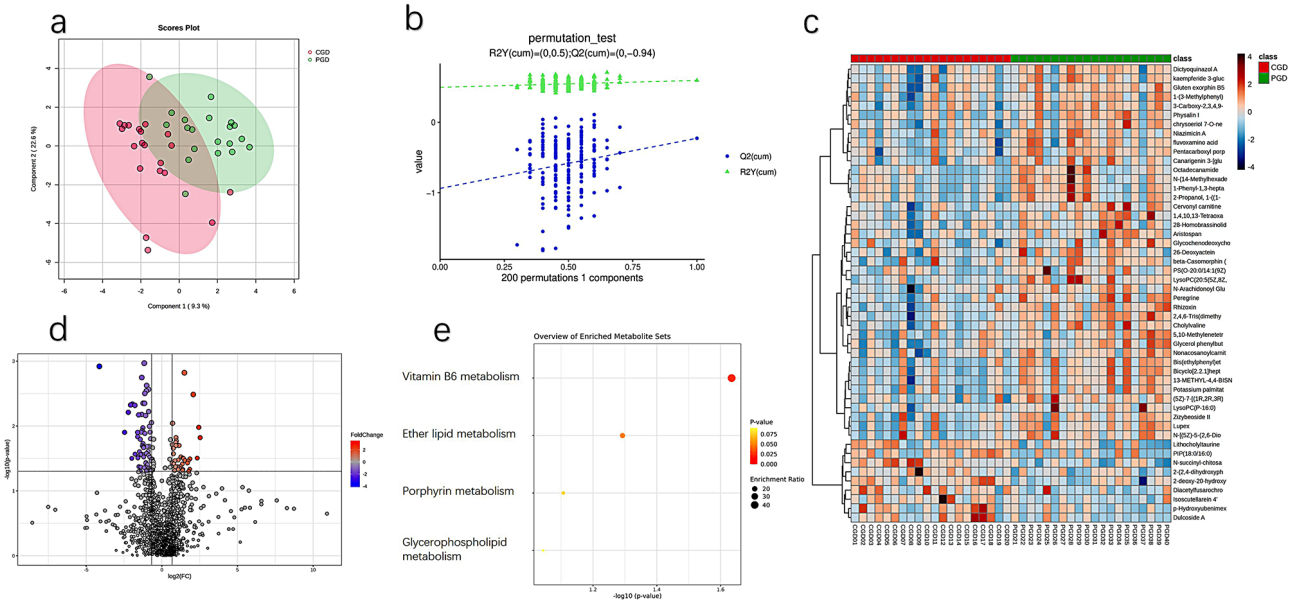
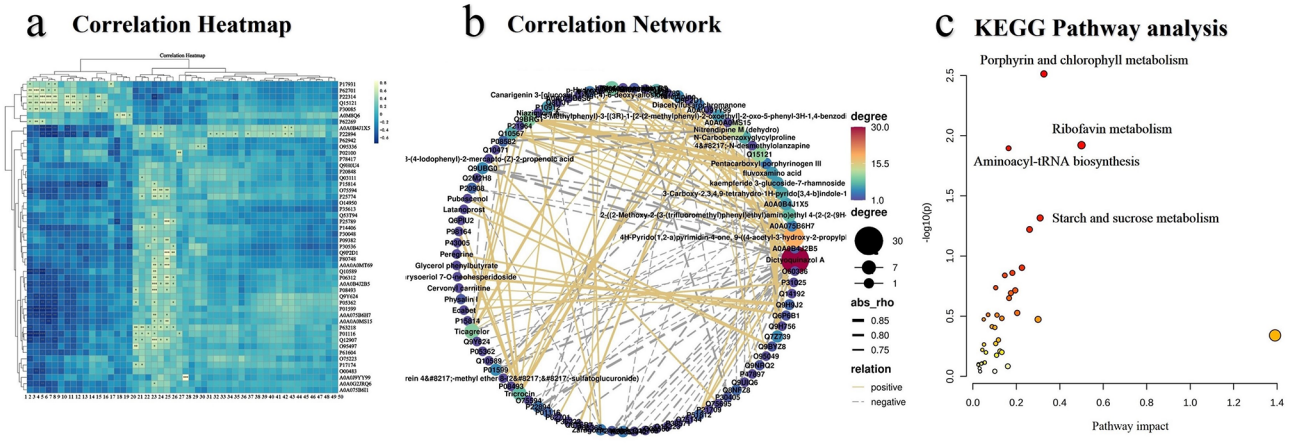


Fig. 5 Differential analysis of metabolite expression levels between CGD patients and PGD patients. **a**) Partial least-squares discriminant analysis model of the metabolomes from the two groups. **b**) Permutation test for PLS-DA model. R2Y and Q2 are used as a measure of whether the model is overfitted **c**) Heatmap of metabolites with fold changes of > 1.5 and < 0.67 between the two groups. Each group represents a single column. Red indicates greater abundance of the metabolite. **d**) Volcano plot of dysregulated urine metabolism in patients with gallstones. **e**) Metabolic pathways with the highest enrichment of differentially expressed metabolites



1.Gluten exorphin B5, 2. 1-(3-Methylphenyl)-3-[(3R)-1-[2-(2-methylphenyl)-2-oxoethyl]-2-oxo-5-phenyl-3H-1,4-benzodiazepin-3-yl]urea, 3. 3-Carboxy-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole-1-propanoic acid, 4. 2-((2-Methoxy-2-(3-(trifluoromethyl)phenyl)ethyl)amino)ethyl 4-(2-(2-(9H-fluoren-9-yl)acetamido)ethyl)benzoate, 5.Zaragozic acid A, 6.chrysoeriol 7-O-neohesperidoside, 7.hemsleyanoside, 8.Niazimicin A, 9.LysoPC(20:5(5Z,8Z,11Z,14Z,17Z)), 10.3’-Prenylapigenin 7-[rhamnosyl-(1->6)-glucoside], 11.Octadecanamide, 12.Chlorin E6, 13.2-Propanol, 1-((1-methylethyl)amino)-3-(2,4,6-tripropylphenoxy)-, 14.Pepstatin, 15.Peregrine, 16.28-Homobrassinolide, 17.Cervonyl carnitine, 18.Sulfolithocholic acid, 19.Lithocholyltaurine, 20.Leu Glu Pro His Ile, 21.Diacetylfusarochromanone, 22.Precorrin 5, 23.Tricrocin, 24.p-Hydroxyubenimex, 25.N-succinyl-chitosan, 26.Isoscutellarein 4’-methyl ether 8-(2’,’-sulfatoglucuronide), 27.Artemether, 28.Dulcoside A, 29.(3S,5R,6R,6’S)-6,7-Didehydro-5,6-dihydro-3,5,6’-trihydroxy-13,14,20-trinor-3’-oxo-beta,epsilon-caroten-19’,11’-olide 3-acetate, 30.N-Arachidonoyl Glutamic acid, 31.PE(14:0/14:0), 32.Benzylpenicilloyl Polylysine, 33.(5Z)-7-[(1R,2R,3R)-3-Hydroxy-2-[(1E,3S)-3-hydroxy-5-phenylpent-1-en-1-yl]-5-oxocyclopentyl]hept-5-enoylcarnitine, 34.Brincidofovir, 35.26-Deoxyactein, 36.Rhizoxin, 37.Sialorphin, 38.Lupex., 39.Zizybeoside II, 40.Bicyclo[2.2.1]heptan-2-one, 1,7,7-trimethyl-3-(phenylmethylene)-, 41.LysoPC(P-16:0), 42.LysoPC(20:2(11Z,14Z)), 43.LysoPC(20:0/0:0), 44.Biliverdin, 45.5,10-Methylenetetrahydromethanopterin, 46.1,4,10,13-Tetraoxa-7,16-diazacyclooctadecane, 47.Cholylvaline, 48.2,4,6-Tris(dimethylaminomethyl)phenol, 49.Glycochenodeoxycholic acid 3-glucuronide, 50.Sarcodon scabrosus Depsipeptide

Fig. 6 Proteomics and metabolomics correlation analyses. **a**) Correlation heatmap analyses of the top 50 DEPs and DEMs with the correlation over 0.6. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. **b**) Correlation network of DEPs and DEMs. Degree is reflected in the proximity between a node and other nodes in the network, solid yellow lines represent a positive correlation, and hrey dashed line represent a negative correlation. The line thickness indicates the correlation strength. **c**) Pathway analysis of DEPs and DEMs

performance (Fig. 7a). The five DEPs were Q9HCE7, Q12888, Q2M2I3, Q9NPL8, and Q96CD2, each with an AUC of 1.00, specificity of 1.00, and sensitivity of 1.00 (Fig. 7b; Table 2). The two DEMs were N-Arachidonoyl Glutamic acid (AUC=0.89, specificity=0.90, sensitivity=0.90) and Pentacarboxyl porphyrinogen III (AUC=0.88, specificity=0.80, sensitivity=1) (Fig. 7b; Table 2). As expected, the combination of these biomarkers implied a high possibility of the presence of gallstones (AUC=0.99) (Fig. 7c).

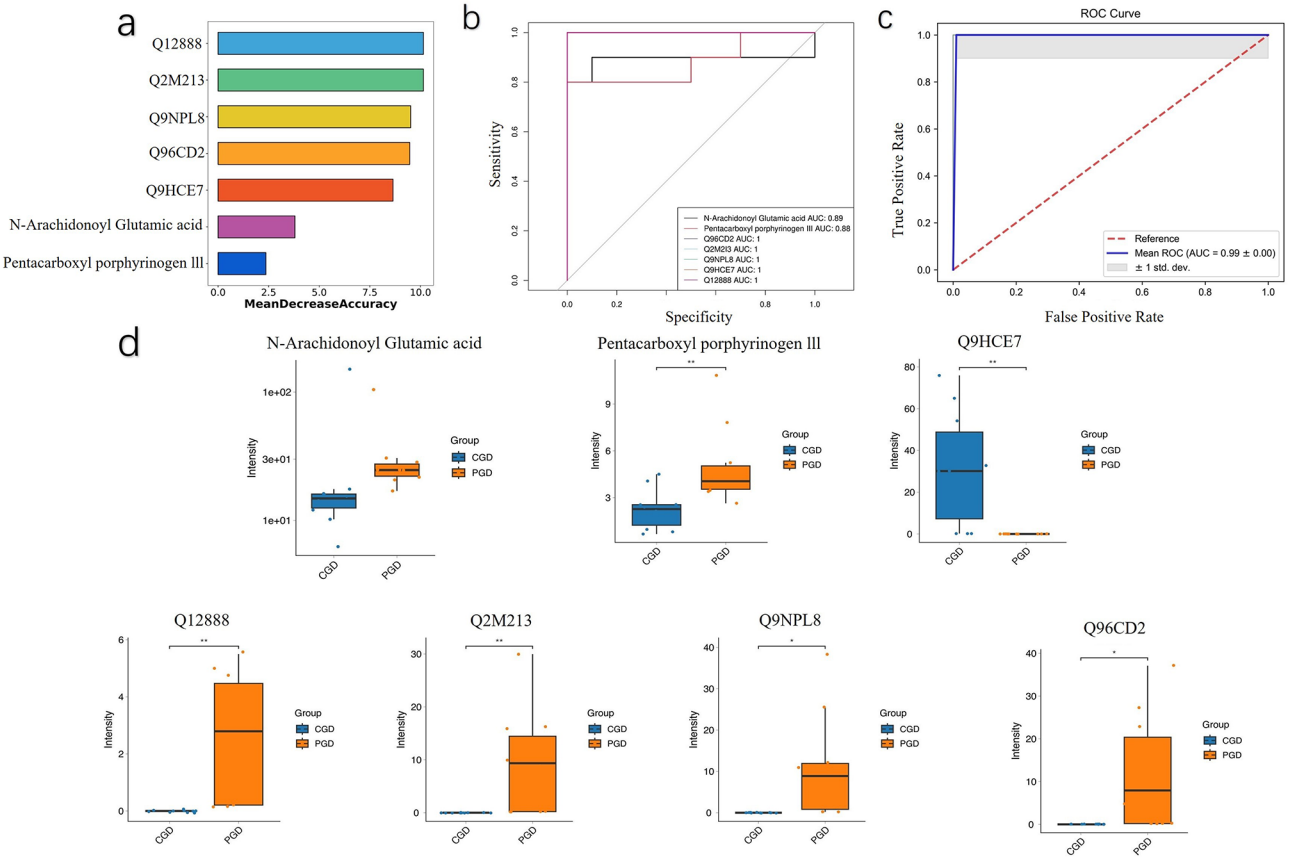


Fig. 7 Identification of potential biomarkers by machine learning. **a**) Among the 7 candidate biomarkers in the random forest analysis based on the mean decrease in the Accuracy index. **b**) ROC curve of the combination of five proteins and two metabolites. **c**) ROC curve of the combination of 7 biomarkers. **d**) Relative concentrations of 7 biomarkers selected by random forest analysis. ROC, receiver operating characteristic; AUC, area under the curve; std. dev., standard deviation; CGD, cholesterol gallstone disease; PGD, pigment gallstone disease

Table 2 Predictive performances of the 7 biomarkers. DEPs, different expressed proteins; DEMs, different expressed metabolites; AUC, area under the curve

DEPs/DEMs	AUC	Specificity	Sensitivity	cutoff	95%ci
N-Arachidonoyl Glutamic acid	0.89	0.90	0.90	2.687	0.694-1.000
Pentacarboxyl porphyrinogen III	0.88	0.80	1.00	2.638	0.711-1.000
Q96CD2	1.00	1.00	1.00	0.206	1.000–1.000
Q2M2I3	1.00	1.00	1.00	0.206	1.000–1.000
Q9NPL8	1.00	1.00	1.00	0.206	1.000–1.000
Q9HCE7	1.00	1.00	1.00	0.000	1.000–1.000
Q12888	1.00	1.00	1.00	0.206	1.000–1.000

Discussion

We identified a total of 3712 human proteins, of which 154 were up-regulated and 196 were down-regulated. KEGG pathway enrichment analysis of important DEPs indicated the presence of several distinct metabolic pathways. Mitochondrial thermogenesis is of great significance as a target for the treatment of metabolic disorders [29]. Apoptotic brown adipocytes enhance the expression of thermogenic programs in healthy adipocytes [30]. A previous study reported that hepatic glucocorticoid receptor deficiency affects dietary fat absorption and brown adipose tissue activation and predisposes to cholesterol gallstone formation [31], which may reflect the role of thermogenesis in cholesterol gallstone formation. Acetyl-CoA is crucial for cell metabolism and protein modifications, impacting cell activity and epigenetic inheritance [32]. Acyl-CoA formation initiates cellular fatty acid metabolism, and recent studies implicated ACOTs (Acyl-CoA thioesterases) in the pathogenesis of metabolic diseases [33, 34]. This may suggest a metabolic disorder in patients with gallstones, potentially impacting the pathways of pantothenic acid and CoA biosynthesis.

The significantly upregulated proteins mainly involve the immune response, while the significantly downregulated proteins mainly involve energy metabolism and protein synthesis and degradation. IGFBPs are currently biomarkers for many diseases, including diabetes and cancer [35–37]. The bile proteins of patients with gallstones were significantly enriched in hepatitis C. IGFBP4 is involved in genes involved in HCV cell growth and malignancy while also being involved in adipogenesis [38, 39]. The significant upregulation of IGFBP4 in gallstone patients may indicate an increase in cholesterol concentration. We noticed that Cox8a was significantly down-regulated in gallstone bile. High caveolar cholesterol levels inhibit tyrosine-induced phosphorylation of caveolin proteins, which may reflect impaired oxidative phosphorylation in gallstone patients [40].

Metabolomics is an effective method to assess the potential role of metabolites in diseases. Compared to PGD patients, the CGD group exhibited significant alterations in 83 bile metabolites. Although KEGG enrichment analysis showed few pathways enriched for DEMs, we still found one significant pathway. Vitamin B6 is an essential element for maintaining human life functions [41]. Previous studies have confirmed its ability to reduce total serum cholesterol levels in mice, and vitamin B6 levels were negatively correlated with total cholesterol and non-high-density lipoprotein levels in patients with hyperlipidaemia [42, 43]. The enriched pathway of vitamin B6 metabolism involves significant upregulation of 4-pyridoxic acid, a major breakdown metabolite of vitamin B6 metabolism [44]. The upregulation of its levels may indicate dysregulation of vitamin B6 metabolism in patients with gallstones [45].

Correlation analysis has been shown to aid in the identification of putative key regulatory elements in network modulation [46, 47]. Under the condition of a correlation coefficient over 0.6, there were 583 correlations among 285 proteins and 83 metabolites. Further, KEGG Pathway enrichment analysis was performed on the joint DEPs and DEMs, revealing three enriched pathways, including porphyrin and chlorophyll metabolism, riboflavin metabolism and aminoacyl-tRNA biosynthesis. Porphyrins play a crucial role in the production, storage, and utilization of oxygen in life processes, while chlorophyll, as an indispensable pigment in photosynthesis, always absorbs light energy and subsequently transfers it to reaction centres [48, 49]. Previous studies reported disturbances in porphyrin and chlorophyll metabolism in mice with liver damage [50]. Bilirubin is the main component of gallstones, and extreme bilirubin levels are a causal risk factor for symptomatic gallstone disease [51, 52]. Bilirubin was also an important molecule in the porphyrin and chlorophyll metabolic pathways [53]. In our study, bilirubin exhibited significant down-regulation. Changes in bilirubin may contribute to the significant enrichment of bile in porphyrin and chlorophyll metabolism in gallstone patients. Riboflavin is an essential component of the coenzymes flavin adenine dinucleotide and flavin mononucleotide, which serve as electron carriers in fatty acid metabolism [54–56]. Riboflavin deficiency affects lipid metabolism and reduces the low-density lipoprotein cholesterol level [55, 57]. Tang et al. reported that riboflavin deficiency increases cholesterol synthesis and leads to fatty liver [58]. Therefore, the dysregulation in riboflavin metabolism in patients with gallstones may subsequently affect the formation of cholesterol stones. Aminoacyl-tRNA biosynthesis was significantly changed when the CGD group was compared with the PGD group. Aminoacyl-tRNA synthetases are enzymes crucial for protein synthesis [59]. While the role of amino acids in gallstone formation remains unclear, recent studies have demonstrated that alterations in the bile amino acid composition influence the formation of gallstones [60]. Significant enrichment of this pathway may be associated with gallstone formation. Furthermore, we employed a machine learning model to identify molecular features related to proteins and metabolites associated with gallstones. We discovered 5 proteins and 2 metabolites as potential biomarkers. Remarkably, these 7 biomarkers demonstrated an excellent AUC of 0.99, highlighting their potential for predicting gallstone formation.

Limitations

First, the small sample size may have promoted the identification of specific proteins and metabolites. Larger samples are needed to validate our findings. Second, due to the invasiveness of bile specimen extraction, we were unable to collect relevant samples from healthy people.

Conclusion

Our study conducted a combined analysis of proteomics and metabolomics to provide new insights into microbial alterations in two subtypes of gallstones. Comprehensive multi-omics data suggest that characteristic pathways and dysregulated molecules identified in CGD and PGD proteomic or metabolomic analyses may help elucidate the underlying mechanisms of gallstone formation in high-altitude populations. Besides, 7 differentially expressed proteins and metabolites showed excellent predictive performance and were selected as potential biomarkers.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12876-024-03422-5.

Supplementary Material 1

Acknowledgements
This study was support by Qinghai Provincial Department of Science and Technology, Qinghai province traffic hospital and Chengdu medical college.

Author contributions
Author contributions: JXF, MY and XF wrote the manuscript. MY, LDF and XYH collected, extracted and cleaned the data. JXF, ZTC and XHQ cleaned the raw data. JXF and XF analysed and interpreted the data. XF and XYH proofread the manuscript.

Funding
This study were supported by the Qinghai Province 2023 focus research and development and transformation plan (2023-SF-117) and Key discipline project in the School of Public Health, Chengdu Medical College (No.21).

Data availability
The data that support the findings of this study are available from https://osf.io/wtsgq/. But restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the corresponding authors Yonghong Xia upon reasonable request and with permission of The open science framework.

Declarations

Ethical approval and consent to participate
This study was conducted in accordance with the Declaration of Helsinki. All patients provided written informed consent. This study was approved by the Ethics Committee of the Qinghai Traffic Hospital, Qinghai, China ([2022] no. 727).

Consent for publication
Not applicable.

Conflict of interest
The authors declare that they have no competing interests as defined by BMC, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Image copyright
Not applicable.

Received: 29 June 2024 / Accepted: 16 September 2024
Published online: 30 September 2024

References

1. Mounir I, Shashank S, Gareth M-S, Maged R, Amit B, Matthew R. W, Gallstones: watch and wait, or intervene? Cleve Clin J Med. 2018;85(4).
2. China Association Of Integrative Medicine Emergency Medicine Committee, Traditional ECOCJOI, Western Medicine, Li Z, Wang D, Li Y. [Expert consensus on diagnosis and treatment of septic shock with integrated traditional Chinese and western medicine]. Zhonghua Wei Zhong Bing Ji jiu Yi Xue. 2019;31(11):1317–23.
3. Giordani MT, Giaretta R, Scolarin C, Stefani MP, Pellizzari C, Tamarozzi F, et al. Ultrasound and infections on the Tibetan Plateau(). J Ultrasound. 2012;15(2):83–92.
4. Carsten G, Simon S, Frank L. The treatment of Gallstone Disease. Dtsch Arztebl Int. 2020;117.
5. Katrin H, Kristin Huld H, Pall Helgi M. [Gallstones - review]. Laeknabladid. 2020;106.
6. Cynthia AS. Compounded bioidentical hormone therapy: new recommendations from the 2020 National Academies of Sciences, Engineering, and Medicine. Menopause. 2021;28(5).
7. Hirajima S, Koh T, Sakai T, Imamura T, Kato S, Nishimura Y, et al. Utility of laparoscopic subtotal cholecystectomy with or without cystic Duct ligation for severe cholecystitis. Am Surg. 2017;83(11):1209–13.
8. Bogdan C, Lech C. [Laparoscopic cholecystectomy-advantages and disadvantages]. Wiad Lek. 2003;56(0).
9. Aun A, Summaya S, Rabel K, Sunil Sadruddin S, Farah Naz F. Difficulties in laparoscopic cholecystectomy: Conversion versus surgeon's failure. J Ayub Med Coll Abbottabad. 2017;28(4).
10. Farzad A, Natalie S, Subhashini A. Gallstone disease: Cholecystitis, Mirizzi Syndrome, Bouveret Syndrome, Gallstone Ileus. Surg Clin North Am. 2019;99(2).
11. Janice YJL, Margaret GK, Stephen P. Diagnosis and treatment of gallstone disease. Practitioner. 2015;259(1783).
12. Irfan A, Jemma H, Karen I, Rodolfo H, Katie G, Rebecca B, et al. Effectiveness of conservative management versus laparoscopic cholecystectomy in the prevention of recurrent symptoms and complications in adults with uncomplicated symptomatic gallstone disease (C-GALL trial): pragmatic, multicentre randomised controlled trial. BMJ. 2023;383:0.
13. C F P-V, M G-M, C R, E N-M, C A B-L LF, dlC-T et al. Gallstone ileus: an overview of the literature. Rev Gastroenterol Mex. 2017;82(3).
14. James LB. Bile formation and secretion. Compr Physiol. 2013;3(3).
15. Hao S, Jonathan W, James Y, Yu J, Shaolong H, Wei H et al. Factors Influencing Gallstone Formation: Rev Literature Biomolecules. 2022;12(4).
16. Agostino DC, David Q-HW, Piero P. An update on the pathogenesis of cholesterol gallstone disease. Curr Opin Gastroenterol. 2017;34(2).
17. Jing L, Ming L, Wei-Qing S, Zong-You C, Wen-Wei Z, Lu L et al. Osteopontin Deficiency alters biliary homeostasis and protects against gallstone formation. Sci Rep. 2016;6(0).
18. Christoph S, Linda S. After another decade: LC-MS/MS became routine in clinical diagnostics. Clin Biochem. 2020;82(0).
19. He L, Yansong X, Xin Z, Taiyang J, Ziru W, Yuansong S et al. DIA-Based proteomic analysis of plasma protein profiles in patients with severe Acute Pancreatitis. Molecules. 2022;27(12).
20. M K, S G. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 1999;28.
21. Minoru K, Yoko S. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci. 2019;29(1).
22. C A B MA, J A B DB, H B JMC et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1).
23. Damian S, Rebecca K, Mikaela K, Katerina N, Farrokh M, Radja H et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2022;51.
24. Jie Y, Fanxia M, Fangping H, Fei C, Wangxiao B, Yamei Y et al. Metabolic abnormalities in patients with chronic disorders of consciousness. Aging Dis. 2021;12(2).
25. Bo S, Xiao Y, Yaoting S, Xiaojie B, Juping D, Chao Z et al. Proteomic and metabolomic characterization of COVID-19 patient sera. Cell. 2020;182.
26. Karimollah H-T. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Casp J Intern Med. 2013;4.
27. David SW, AnChi G, Eponine O, Fei W, Afia A, Harrison P et al. HMDB 5.0: the human metabolome database for 2022. Nucleic Acids Res. 2022;50(0).
28. Carlos G, Rafael J, Xavier M-B, Amelia D-A, Benedikt P, Gerrit W. H, METLIN: A Technology platform for identifying knowns and unknowns. Anal Chem.90(5).
29. Ambre MB, Yuriy K. Mitochondrial H(+) leak and thermogenesis. Annu Rev Physiol. 2021;84(0).
30. Birte N, Saskia H-B, Laura P, Martin F, Michelle YJ, Anne H et al. Apoptotic brown adipocytes enhance energy expenditure via extracellular inosine. Nature. 2022;609(7926).
31. Adam JR, Mauricio BD, Anja R, Johanna K, Tessa W, Anja K-H et al. Molecular control of systemic bile acid homeostasis by the liver glucocorticoid receptor. Cell Metab. 2011;14(1).
32. Weijing H, Qingguo L, Xinxiang L. Acetyl-CoA regulates lipid metabolism and histone acetylation modification in cancer. Biochim Biophys Acta Rev Cancer. 2022;1878(1).
33. Jessica ME, Caitlyn EB, Michael JW. Metabolic and tissue-specific regulation of acyl-CoA metabolism. PLoS ONE. 2015;10(3).
34. Veronika T, Stefan EHA, David EC. Deactivating fatty acids: Acyl-CoA thioesterase-mediated control of lipid metabolism. Trends Endocrinol Metab. 2017;28(7).
35. Hana Th A, Mohamed A-F, Eman A, Sriraman D, Zahra A, Ashraf E et al. Increased levels of circulating IGFBP4 and ANGPTL8 with a prospective role in Diabetic Nephropathy. Int J Mol Sci. 2023;24(18).
36. Chunjing L, Yu C, Li Z, Jierong L, Huayan W, Fengsheng L et al. LncRNA IGFBP4-1 promotes tumor development by activating Janus kinase-signal transducer and activator of transcription pathway in bladder urothelial carcinoma. Int J Biol Sci. 2020;16(13).

37. Liye T, Yali W, Zefeng S, Jingwei C, Junhao Z, Shunjie X et al. Activation of IGFBP4 via unconventional mechanism of miRNA attenuates metastasis of intrahepatic cholangiocarcinoma. Hepatol Int. 2023;18(1).
38. Aunji P, Shwetha S, Shuchismita D, Amit G, Rakesh A, Saumitra D. Exosome-associated microRNA-375 induces cell proliferation by regulating IGFBP4 upon hepatitis C virus infection. Mol Microbiol. 2022;118(5).
39. Yunhua LM, Michael S, Samantha D, Khushdeep B, Cigdem K, Sayuko K et al. Functional characterization of a novel p.Ser76Thr variant in IGFBP4 that associates with body mass index in American indians. Eur J Hum Genet. 2022;30(10).
40. G M M JB. M C C. Roles of cholesterol and bile salts in the pathogenesis of gallbladder hypomotility and inflammation: cholecystitis is not caused by cystic duct obstruction. Neurogastroenterol Motil. 2013;25(4).
41. Pramod R. Preventing vitamin B6-Related neurotoxicity. Am J Ther. 2023;29(6).
42. Sibel T, Emre S, Melahat D. Vitamin B6 supplementation improves oxidative stress and enhances serum paraoxonase/arylesterase activities in streptozotocin-induced diabetic rats. ScientificWorldJournal. 2014;2014:0.
43. Madleen Nabil A-Q, Wajdi Khalaf Jamil AM, Rasha MH. Association of Vitamins D, B6, and B12 deficiencies with hyperlipidemia among Jordanian adults. Rep Biochem Mol Biol. 2024;12(3).
44. Stephen PC, Robert DR, Dennis J, Wayne M, Yao ES, Karen W. L E, Elevated plasma 4-pyridoxic acid in renal insufficiency. Am J Clin Nutr. 2002;75(1).
45. Rima O, Juergen G, Wilfred AN. 4-Pyridoxic Acid/Pyridoxine ratio in patients with type 2 diabetes is related to Global Cardiovascular Risk scores. Diagnostics (Basel). 2019;9(1).
46. Jun R, Xianghui W, Fangfang T, Qinsi H, Aihua X, Xinyi C et al. Molecular characterization of Advanced Colorectal Cancer using serum proteomics and Metabolomics. Front Mol Biosci. 2021;8(0).
47. Zhi Z, Qingfeng W, Xianghui W, Xiaoming Z, Lijuan L, Jiquan Z, et al. Correlation analysis between Trace Elements and Colorectal Cancer metabolism by Integrated Serum Proteome and Metabolome. Front Immunol. 2022;13:0.
48. Mathias OS, Natalia NS, Karl JH. Classic highlights in porphyrin and porphyrinoid total synthesis and biosynthesis. Chem Soc Rev. 2021;50(7).
49. Xu L, Wei Z, Di N, Xiaomin L. Effects of abiotic stress on chlorophyll metabolism. Plant Sci. 2024;342(0).
50. Shizhao X, Fanli K, Zhengwu S, Yalin X, Fei Q, Jianzhi S. Hepatoprotective effect and metabonomics studies of radix gentianae in rats with acute liver injury. Pharm Biol. 2021;59(1).
51. Stephan B, Clemens S, Henry V, Marcus S, Juan FM, Silvia CS et al. Loci from a genome-wide analysis of bilirubin levels are associated with gallstone risk and composition. Gastroenterology. 2010;139(6).
52. Stefan S, Ruth F-S, Børge GN, Anne T-H. Extreme bilirubin levels as a causal risk factor for symptomatic gallstone disease. JAMA Intern Med. 2013;173(13).
53. Junxiang W, Yinyin C, Yang L, Fenhua Z, Meifen Y, Jin X et al. Metabolomics analysis of the serum from children with urolithiasis using UPLC-MS. Clin Transl Sci. 2021;14(4).
54. Helene M, Kristina P, Mary W. Causes and clinical sequelae of Riboflavin Deficiency. Annu Rev Nutr. 2023;43(0).
55. Xiangyu B, Weina G, Yawen W, Zhanxin Y, Qingao X, Changjiang G et al. Riboflavin deficiency affects lipid metabolism partly by reducing apolipoprotein B100 synthesis in rats. J Nutr Biochem. 2019;70(0).
56. Jiajia Y, Xuewei P, Chen Y, Yuxuan D, Tolbert O, Taowei Y et al. Microbial production of riboflavin: biotechnological advances and perspectives. Metab Eng. 2021;68(0).
57. Larry AC, Danielle NS, Xiaofei W, Nares T, Wilfrid C, Harold B 3rd. W. Transcriptional profiling of liver in riboflavin-deficient chicken embryos explains impaired lipid utilization, energy depletion, massive hemorrhaging, and delayed feathering. BMC Genomics. 2018;19(1).
58. Jing T, Maria AH, Jian H, Ming X, Wenbiao S, Yong J et al. Severe riboflavin deficiency induces alterations in the hepatic proteome of starter Pekin ducks. Br J Nutr. 2017;118(9).
59. Miguel Angel RG, Michael I. Aminoacyl-tRNA synthetases. RNA. 2020;26(8).
60. Jolanta B, Joanna B, Artur P, Krystyna S. Biliary amino acids and telocytes in Gallstone Disease. Metabolites. 2023;13(6).

Publisher's note