CrossMark

ORIGINAL ARTICLE

Comprehensive characterization and evaluation of hepatocellular carcinoma by LC–MS based serum metabolomics

Xin Lu[1,3] · Huan Nie[1] · Yiqun Li[1] · Chao Zhan[2] · Xia Liu[1] · Xiuyun Shi[1,3] ·
Ming Shi[1] · Yubao Zhang[2] · Yu Li[1,3]

Abstract The liver is an active organ in energy metabolism, and hepatocellular carcinomas (HCC) therefore unavoidably induce a series of metabolic alterations in the serum metabolome. In this study, 587 sera from HCC patients, hepatitis B virus (HBV) infected subjects and healthy control subjects were employed to characterize metabolic alterations using a liquid chromatography tandem mass spectrometry-based metabolomics approach. D-Galactose significantly increased whereas undecanoyl-L-carnitine and PE (P-18:0/0:0) decreased in HCC patients when compared with non-HCC controls (HBVs and NCs). These metabolites were finally identified as independent diagnostic factors for HCC discrimination. The combination of discriminative metabolites and alpha-fetoprotein could also significantly improve the diagnostic performance for HCC in terms of the sensitivity, specificity and area under the curve of the receiver operating characteristic curve; these three parameters displayed values of 0.96, 0.92 and 0.98, respectively. Furthermore, network and pathway analyses were conducted to explore the latent relationship among differential metabolites. In the network analysis, metabolites of diverse categories (including amino acids, dipeptides, phospholipids, fatty acids, steroids and acyl-carnitines) reflected the correlation between metabolite categories on material and energy conversion in HCC. Meanwhile, metabolites mapped in the pathways for primary bile acid biosynthesis and alanine, aspartate and glutamate metabolism revealed a centrality and significant variances. These results indicated that substantial biochemical perturbations occurred in these pathways in HCC patients. In summary, our study revealed the systematic landscape for metabolic alterations in the serum of HCC patients and provides useful information to clinicians and biologists.

✉ Yubao Zhang
zhyb88077@sina.com

✉ Yu Li
liyugene@hit.edu.cn

[1] School of Life Science and Technology, Harbin Institute of Technology, Room 310, Building 2E, Science Park of Harbin Institute of Technology, No. 2 Yikuang Street, Nangang District, Harbin 150001, China

[2] The Affiliated Tumor Hospital, Harbin Medical University, No. 150, Haping Road, Nangang District, Harbin 150001, China

[3] State Key Laboratory of Urban Water Resource and Environment, School of Life Science and Technology, Harbin Institute of Technology, Harbin, China

Abbreviations

UPLC	Ultra-performance liquid chromatography
MS	Mass spectrometry
TOF	Time-of-flight
NMR	Nuclear magnetic resonance spectroscopy
HCC	Hepatocellular carcinoma
HBV	Hepatitis B virus
AFP	α-Fetoprotein
ROC	Receiver operating characteristic
AUC	Area under the curve
PCA	Principle component analysis
HCA	Hierarchical cluster analysis
PLS-DA	Partial least square discriminant analysis

OPLS-DA	Orthogonal projection to latent structures discriminant analysis
RF	Random forest
VIM	Variable importance
QC	Quality control
RT	Retention time
m/z	Mass-to-charge ratio
ALT	Alanine aminotransferase
AST	Asparate aminotransferase
BPC	Basic peak chromatogram
TNM	Tumor node metastasis
GLS	Glutaminase
ROS	Reactive oxygen species
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
LPC	Lysophosphatidylcholine
LPE	Lysophosphatidylethanolamine
FA	Fatty acid
DP	Dipeptide
ST	Sterol lipid
AA	Amino acid
Se	Sensitivity
Sp	Specificity

1 Introduction

The liver is active in cellular metabolism and plays important roles in metabolic processes involving amino acids, carbohydrates, lipids and energy metabolism for maintaining vital activities (Anis and Irshad 2011). The liver is the largest exocrine gland; therefore, liver related diseases frequently trigger various alterations in bio-fluids, notably in the serum. Compared with other cancers, hepatocellular carcinoma (HCC) will therefore induce a series of more complicated metabolic changes, which could be partially reflected in the serum metabolome (Marquardt et al. 2012).

Emerging omics technologies have been rapidly developed during over the last few decades spanning from gene to metabolite levels. These technologies allow for the search of differential biomolecules for the diagnosis of and the explanation of diseases (Slany et al. 2013; Uto et al. 2010). Newly developed metabolomics could provide a tool for the monitoring of the progression of a disease and a language for interpreting the variation of upstream molecules (Nicholson et al. 1999). In recent years, several metabolomic studies have been performed to profile diseases, and these studies contributed to biomarker discovery and mechanism exploration. Montuschi et al. successfully discriminated between patients with unstable cystic fibrosis (CF), stable CF and healthy subjects by charactering the metabolic profiles of exhaled breath condensate using high-resolution nuclear magnetic resonance spectroscopy (NMR). (Montuschi et al. 2012; Motta et al. 2012). For metabolomics-based HCC profiling, Xu and colleagues established an integrated platform for both serum and tissue analysis and further investigated the feasibility of potential metabolite biomarkers by using multi-chromatography coupled high-resolution mass spectrometry, (Huang et al. 2013a; Zeng et al. 2014; Chen et al. 2013a). Diren et al. presented a combined metabolomics and transcriptomics study to profile a dynamic description of the aberrant energy metabolism in HCC (Beyoglu et al. 2013). However, because of the lack of a systematic evaluation of the clinical covariate factors and in-depth understanding of the associated mechanism, applications of these methods in clinical diagnosis and theoretical recognition are limited. Thus, a comprehensive characterization and evaluation of HCC will assist in the understanding of pathophysiologic processes and in the discovery of discriminative biomolecules for HCC diagnosis and individual-based treatment.

In this study, we obtained the metabolomic data based on the high-throughput UPLC/Q-TOF MS for 587 serum samples from HCC patients, HBV subjects and normal controls (NCs). Both an unsupervised method (principal component analysis, PCA) and a supervised method (partial least square discriminant analysis, PLS-DA) were employed to reveal the global metabolic changes in HCC, HBV and NC. Univariate and multivariate analyses were used to select the differential metabolites. To eliminate the potential effect of covariate factors, a multivariate logistic regression was performed with a covariate adjustment for gender, age and liver function status. The differential metabolites, which had independent predictive capacities for HCC discrimination alone or combined with AFP, could be determined as potentially discriminative metabolites. Therefore, three discriminative metabolites [D-galactose (Sac_1), undecanoyl-L-carnitine (C11) and PE(P-18:0/0:0)] were determined to be independent predictors for HCC diagnosis. To illustrate the relationship between the differential metabolites in HCC, correlation network and pathway analysis were both performed. Abnormal changes in the metabolites mapped in the acylcarnitine network and to the pathway for the primary bile acid biosynthesis displayed particular perturbations in material and energy metabolism in HCC patients.

2 Materials and methods

2.1 Chemicals and reagents

HPLC grade acetonitrile, methanol and formic acid were purchased from Sigma-Aldrich and Fluka (St. Louis, MO, USA); deionized water was produced by a Milli-Q ultrapure water system (Millipore, Billerica, USA). D-Galactose, succinic acid, taurine, DL-kynurenine, 3-indolelactic

acid and LysoPC(16:0) were also purchased from Sigma-Aldrich. All other standard compounds were supplied by AMRESCO (Rd. Cochran Solon, OH, USA).

2.2 Clinical samples

A total of 587 serum samples were collected from 267 HCC patients, 48 HBV subjects and 272 normal controls (NC). HCC patient samples were collected and pathologically confirmed in the Affiliated Tumor Hospital at the Harbin Medical University; other non-HCC samples originated from individuals receiving regular physical examinations at the Affiliated Hospital of Harbin Institute of Technology. All enrolled samples received a liver function test and were collected between December 2008 and July 2010. Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) simultaneously served as criteria to estimate the liver function status. A serum concentration of >200 ng/mL was considered as the cut-off for discriminating HCC patients (AFP-positive) (Bruix et al. 2005). All samples were divided into a training set and validation set according to the time sequence. The training set consisted of 47 HCC patients, 48 HBV patients and 48 normal controls, whereas the remaining 444 samples and additional 32 pairs of postoperative samples were assigned to the validation set. This study was approved by the ethical committees at the Affiliated Tumor Hospital in the Harbin Medical University and the Affiliated Hospital of Harbin Institute of Technology, respectively. The demographic and clinical characteristics of the subjects are presented in Table 1.

2.3 Sample preparation

After a 12-h fast, whole blood samples were collected with non-anticoagulant vacuum tubes; these samples were then centrifuged at $4000 \times g$ to separate the serum. The isolated sera were then stored at $-80\ °C$ until preparation. Prior to preparation, quality control samples (QCs) were prepared by mixing equal aliquots of supernatant from all of the samples in the identical corresponding dataset (i.e., training set or validation set).

The sample preparation followed a modified protocol as previously described (Bruce et al. 2009). In brief, a volume of 300 μl pre-cooled mixture of methanol/acetonitrile (1:1) was added to 100 μl serum for protein precipitation. Finally, the dried residue obtained after rotating-vacuuming was stored at $-80\ °C$ until analysis.

2.4 UPLC-Q/TOF MS analysis

In total, 100 μl of 50 % methanol was used for re-dissolving, and 5 μl of the supernatant was injected into a BEH C_{18} 2.1 mm × 100 mm, 1.7 μm column (Waters, Milford, USA) on a ultra-performance liquid chromatography (UPLC) system (Waters, Milford, USA). The mobile phase consisted of acetonitrile containing 0.1 % formic acid for canal A and deionized water containing 0.1 % formic acid for canal B. To avoid insufficient nebulization, the flow rate was set at 0.35 mL/min with a column temperature of 40 °C. A linear gradient for elution was set as follows: 1 % A for 0–0.5 min, 1–53 % A for 0.5–3.5 min, 53–70 % A for 3.5–7.5 min, 70–90 % A for 7.5–9 min, and then maintained at 90 % A for 9–13 min followed by alternating the gradient back to 1 % A from 13.1 to 15 min. To ensure stability during analysis, samples for quality control were analyzed at the beginning and the end of each running batch.

The MS acquisition and MS/MS identification were both performed in the positive and negative mode with a 6520 series accurate quadrupole time-of-flight mass spectrometer (Q-TOF MS) equipped with a dual electrospray ion source (Agilent, Santa Clara, CA, USA). The parameters for the acquisition were set as follows: the capillary voltage was set at 4 and 3.5 kV in the positive and negative mode, respectively; the gas temperature was 330 °C with a flow rate of 10 L/min; and the fragmentor and skimmer were maintained at 100 and 65 V, respectively. The mass range was from 70 to 1100 m/z with a scan rate of 1.5 spectra/s.

2.5 Data preprocessing

The data preprocessing, including the peak filtering, retention time (RT) correction, feature alignment and normalization, were performed in the R-project platform using the *xcms* package. The bandwidth was set at 15 s (bw = 15), and the peak width ranged from 5 to 30 s. The defaults were adopted for the other parameters.

2.6 Statistical analysis

Both a PCA and PLS-DA were conducted to reveal the global metabolic profiles of HCC patients, HBV subjects and NCs. To avoid overfitting, cross-validation was used to certify the stability and credibility of the PLS-DA models (Wiklund et al. 2007). Furthermore, we performed a univariate nonparametric Kruskal–Wallis rank sum test and designed a multivariate random forest (RF) model to select the potential differential metabolites according to the criteria of $p < 0.05$ and VIM > 1 (Huang et al. 2013b). An hierarchical cluster analysis (HCA) was applied to visualize the changes in the concentration levels of these metabolites in a heatmap. Multivariate logistic regression was used to evaluate whether the differential metabolites were the independent factors of the HCC diagnosis. The area under the receiver operating characteristics curve

Table 1 Demographic and clinical characteristics of the subjects in the training set and validation set

Charactertics	Training dataset			Validation samples		Post-operative ($n = 32$)
	HCC ($n = 47$)	HBV ($n = 48$)	NC ($n = 48$)	HCC ($n = 220$)	NC ($n = 224$)	
Age						
≤40	7	9	14	24	28	1
41–50	13	18	19	60	63	9
51–60	17	14	12	79	78	15
≥61	10	7	1	57	55	7
Gender						
Male	38	37	28	166	170	26
Female	10	11	20	54	54	6
Stage						
I	37	–	–	58	–	19
II	2	–	–	11	–	3
III	1	–	–	17	–	3
Unknown	7	–	–	134	–	7
Differentiation						
1	11	–	–	25	–	6
2	7	–	–	32	–	13
3	4	–	–	26	–	12
Unknown	25	–	–	137	–	1
AFP						
≤200	30	–	–	76	–	18
>200	9	–	–	122	–	6
Unknown	8	–	–	22	–	8

(AUC) was determined to evaluate the predictive power of the discriminative metabolites alone and together with AFP for HCC diagnosis. To further illustrate the latent relationship between the differential metabolites in HCC, a correlation network was designed and a pathway analysis was performed.

The PCA, PLS-DA and cross-validation were performed using SIMCA-p v.11.5 (Umetrics AB, Umea, Sweden). HCA and pathway analysis were conducted in MetaboAnalyst (http://www.metaboanalyst.ca/MetaboAnalyst/). The correlation network was constructed with Cytoscape v.3.1.0 (www.cytoscape.org) (Herrgard et al. 2008; Klamt et al. 2007). All other analyses and visualizations were performed using the R platform (Zhang et al. 2013).

2.7 Metabolite identification

The structural information of differential metabolites, including the RT, m/z and MS/MS spectrum, were well matched with those of authentic standards or confirmed spectrums in the Human Metabolome Database (HMDB, http://www.hmdb.ca/), Metabolite and Tandem MS Database (METLIN, http://metlin.scripps.edu/index.php) or Mass Bank Database (http://www.massbank.jp/).

3 Results

3.1 Metabolic profiling of HCC, HBV and NC

In the present study, all 587 serum samples attributed to the training set and validation set were analyzed using UPLC/Q-TOF MS. The workflow for the metabolomic data analysis is presented in Fig. 1. Because of a large proportion of HBV infections in the recruited HCC patients, we combined the HBV subjects and normal controls together as the non-HCC controls in the training set to eliminate the latent influence of HBV infection. The typical basic peak chromatograms (BPC) of the HCC, HBV and NC group in both the positive and negative mode are shown in Fig. S1.

The three-dimensional PCA score plot for subjects in the training set is shown in Fig. 2a, displaying no obvious outliers in these subjects. The PLS-DA score plot (Fig. 2b) indicated that HCC patients, HBVs and NCs could be separated from each other. Notably, HCC patients were far away from the two non-HCC controls (HBV and NC), whereas HBVs were located close to NCs. No obvious over-fitting was observed in the permutation test (Fig. 2c). Thus, this analysis suggested no obvious differences in the serum metabolic profiles between the HCC and two control

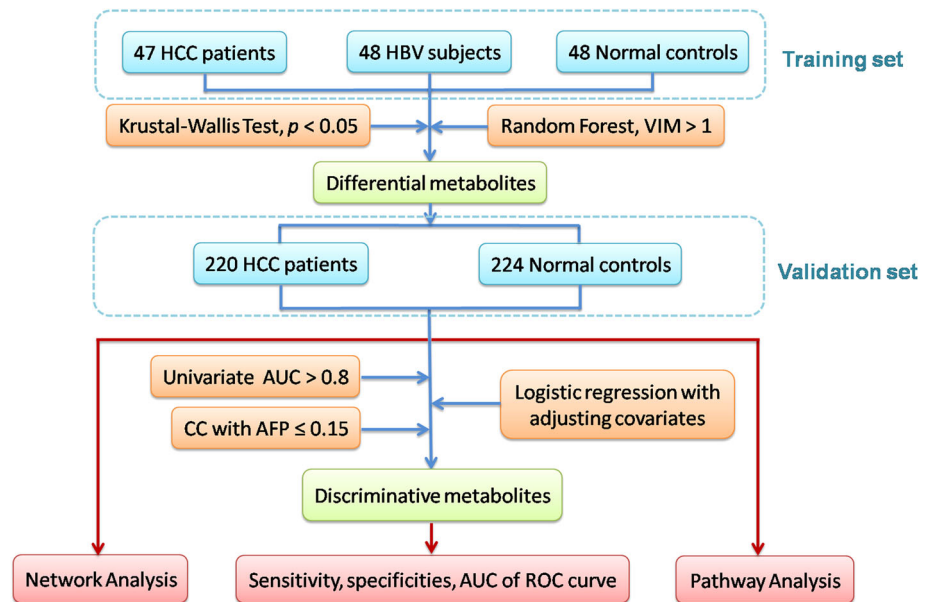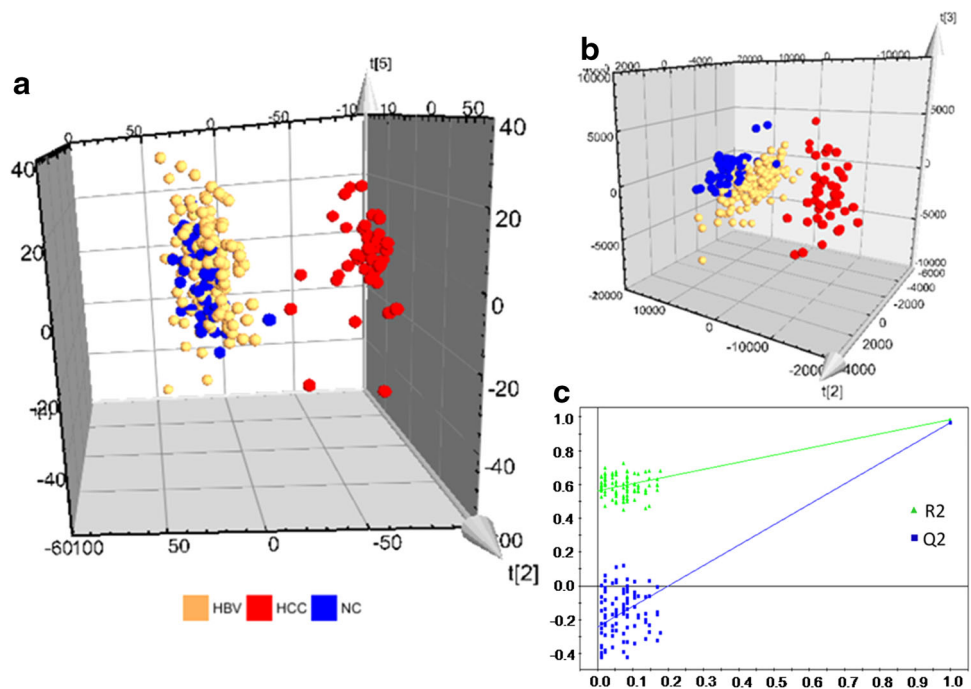**Fig. 1** The workflow of the metabolomic data analysis

Fig. 2 Three-dimensional score plots for the PCA (**a**) and PLS-DA model (**b**) to discriminate HCC ($n = 47$) from HBV ($n = 48$) and NC ($n = 48$) in the training set; and the cross-validation plot obtained from 100 permutation tests (**c**)

groups were noted, and the metabolic profiles of the two control groups were similar to each other.

3.2 Selection and identification of differential metabolites

To further explore the differential metabolites responsible for the differences in the metabolic profile between HCC and non-HCC controls, a univariate Kruskal–Wallis rank sum test and multivariate random forest were used to obtain the differential metabolites with the criteria of $p < 0.05$ and VIM > 1. In total, 35 differential metabolites in the positive ion mode (ESI+) and 54 in negative ion mode (ESI−) were selected as differential metabolites and are listed in Table S1.

The HCA-heatmap for all the differential metabolites and box plots of the 16 most significant metabolites (8 up-regulated in HCC and 8 down-regulated) are presented in Fig. 3. To be specific, D-galactose (Sac_1), taurocholate (ST_1), taurochenodeoxycholate (ST_2), glycocholate

(ST_3), taurine (OA_3), L-glutamate (AA_15), hexade-cenoyl-L-carnitine (C16:1), and (R)-3-hydroxy-hexade-canoic acid (FA_11) increased in HCC patients whereas undecanoyl-L-carnitine (C11), PE(P-18:0/0:0) (PE_2), β-aspartyl-phenylalanine (DP_5), phenylalanyl-phenylala-nine (DP_6), chenodeoxycholate (ST_8), LTB4 (FA_7), L-glutamine (AA_2), and PC(16:1(9Z)/2:0) (PC_7) de-creased. In the HCA-heatmap diagram, the HCC observa-tions were completely separated from the non-HCC control groups. Additionally, the HBV subjects were found to be generally scattered next to NCs. This organization was similar to the grouping patterns shown in the PLS-DA score plot. The Euclidean distances among the three groups also displayed substantial differences in the metabolic
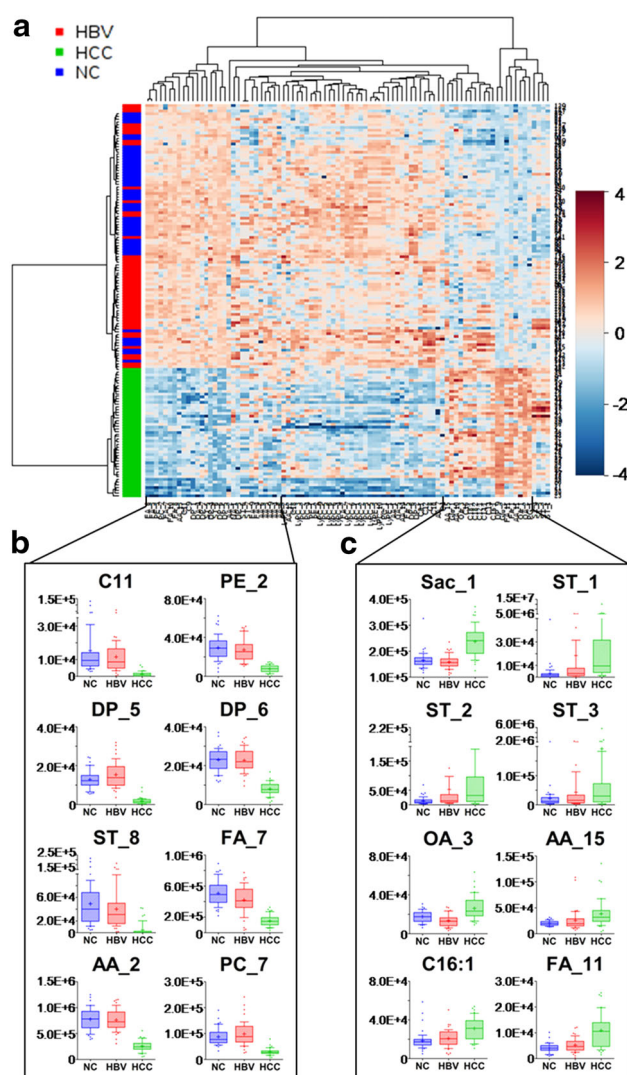

Fig. 3 HCA-heatmap plot and box plots indicating relative levels of differential metabolites in samples of the training set. **a** HCA-heatmap plot of 89 differential metabolites. **b** Box plots for the 8 most significant HCC-down-regulated metabolites. **c** Box plots for the 8 most significant HCC-up-regulated metabolites

profiles between the HCC and non-HCC controls and a similarity between HBVs and NCs. The details of the structural information of the identified differential metabolites are given in Fig. S2.

3.3 Verification of discriminative metabolites

To obtain the discriminative metabolites with satisfactory predictive performances alone or combined with AFP (a biomarker commonly used in the clinical diagnosis of HCC), we followed the criteria for discriminative metabolite determination from differential metabolites: (1) employing a multivariate logistic regression adjusting for gender, age and liver function; (2) displaying a p value of the correlation coefficient with AFP smaller than 0.15 based on a Pearson correlation analysis (Huang et al. 2013a); (3) displaying a univariate AUC larger than 0.8 in the discrimination between HCC and NC (El Khouli et al. 2009; Wang et al. 2014). Three metabolites, i.e., D-galac-tose (Sac_1), undecanoyl-L-carnitine (C11) and PE(P-18:0/0:0) (PE_2), were determined to be potential discriminative metabolites for HCC prediction.

Moreover, we investigated the dynamic changes in the normal controls and HCC patients in different pathological stages (TNM staging system), and the call-back postop-erative tendencies for the three discriminative metabolites are shown in Fig. S3. The two panels of Fig. S3 revealed that D-galactose and PE(P-18:0/0:0) demonstrated more significant changes in stage I and II than that in stage III (Fig. S3a), whereas D-galactose and undecanoyl-L-carnitine showed clear call-back tendencies after surgery (Fig. S3b). These findings suggested that the discriminative metabo-lites might be useful for HCC early diagnosis and prognosis.

3.4 Diagnostic performance of the three metabolites in the external validation set

For the diagnostic performance of the discriminative metabolites, the predictive accuracy was assessed using the sensitivity (Se), specificity (Sp) and AUC in the external validation set. The Se, Sp and AUC were 0.62, 0.98 and 0.81 for AFP, whereas 0.94, 0.85 and 0.95 for the combi-nation of D-galactose, undecanoyl-L-carnitine and PE(P-18:0/0:0). The combination of the three metabolites to-gether with AFP in the prediction of HCC displayed a Se, Sp and AUC of 0.96, 0.92 and 0.98, respectively. For the AFP-negative patients, the combination of the three metabolites had an accuracy of 0.78.

These results indicated that the discriminative metabolites could provide a comparable diagnostic performance to AFP in the diagnosis of HCC and the prediction of AFP-negative

patients, which allows these metabolites to potentially contribute to HCC diagnosis in clinical practice.

3.5 Correlation network analysis

To exploit the latent relationships between the differential metabolites in HCC, a correlation network was built based on the criteria of a correlation coefficient ≥0.7, which ensured the robustness and reliability of the network construction. A total of 59 nodes and 129 edges were recruited in the network diagram in a circular layout (Fig. 4). The diagram showed that metabolites of diversified categories gathered in this diagram. Amino acids (AA), dipeptides (DP), glycerophospholipids (PC, LPC, PE, LPE), the majority of fatty acids (FA) and medium-chain acylcarnitines (Cn) showed down-regulated concentration levels (blue

nodes) in HCC patients whereas steroids (ST), a minority of fatty acids and long-chain acylcarnitines were up-regulated (red nodes). In accordance with the molecular composition and transforming relationship of metabolites of different classes, the entire network could be generally divided into four subnetworks: amino acids/dipeptides (Fig. 4a), glycerophospholipids/fatty acids (Fig. 4b), steroids (Fig. 4c) and acylcarnitines (Fig. 4d).

Although the enrollment of amino acids in the HCC related network analysis has been uncovered in a previous study (Huang et al. 2013a), we further extended this inter-category relationship with the recruitment of dipeptides in the present study. Notably, the five dipeptides enrolled were all structural with phenylalanine residuals; two of the dipeptides, i.e., L-beta-aspartyl-L-phenylalanine (DP_5) and phenylalanyl-phenylalanine (DP_6), both revealed a close

Fig. 4 Correlation network constructed with 59 differential metabolites (Pearson correlation analysis, $|r| \geq 0.7$). **a** Sub-network constructed with dipeptides (DP) and amino acids (AA). **b** Sub-network constructed with glycerophospholipids (LyPC/LyPE and PC/PE) and fatty acids (FA). **c** Sub-network constructed with steroids (ST). **d** Sub-network constructed with saturated acylcarnitines (Cn) and unsaturated acylcarnitines (Cn:1/n:2). **e** Line-plot indicating the concentration changes of the saturated medium- and long-chain acylcarnitines between HCC and NC. **f** Line-plot indicating the concentration changes of unsaturated medium- and long-chain acylcarnitines between HCC and NC. Nodes in *red* and *blue* represent the metabolites down-regulated and up-regulated in HCC, respectively. Nodes with different *color labels* represent metabolites from different categories, including DP, AA, FA, LyPC/LyPE, PC/PE, ST, Cn and Cn:1/n:2 *$p < 0.05$; **$p < 0.0001$ (Color figure online)
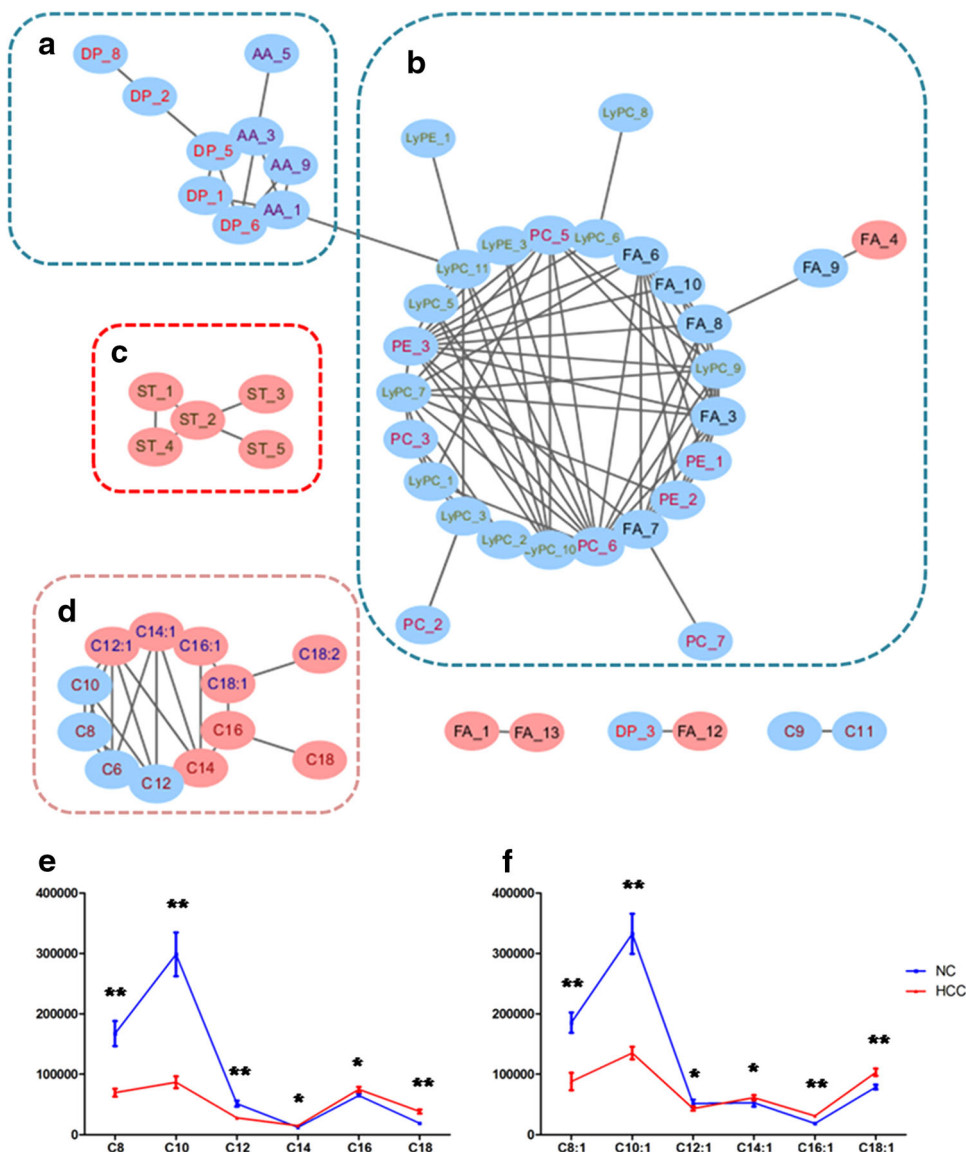
correlation with phenylalanine (AA_3) as shown in Fig. 4a, implying the potential key role of phenylalanine in the transformation between dipeptides and amino acids in HCC.

Glycerophospholipids, including LPC, LPE, PC and PE, were found in the largest subnetwork (Fig. 4b) with no further clustering between subclasses. The fatty acids were observed adhering to glycerophospholipids, and a clear intra-category gathering pattern was demonstrated between these categories as shown in the diagram. Such cross-linkages between glycerophospholipids and fatty acids also might reflect the disordered aliphatic transition for fulfilling the aberrant energy requirement during tumor proliferation (Morita et al. 2013; Huang et al. 2013a).

Steroids and acylcarnitines were both found to gather independently in their own subnetworks, suggesting their specificities to metabolic alterations in HCC. Notably, 12 acylcarnitines were found scattered in sequential arrangement according to the acyl-chain length and unsaturated degree, in which short and medium-chain acylcarnitines (C6-C12) decreased in HCC patients, whereas long-chain acylcarnitines (C14-C18) increased with the elongation of the acyl-chain in both the saturated (Cn) and unsaturated acylcarnitines (Cn:1), as presented in Fig. 4e, f.

In brief, the intra-category gathering landscapes could be clearly observed in metabolites of different classes in the network diagram, suggesting the underlying transformation of substances and energy in HCC.

3.6 Pathway analysis

To further investigate the biochemical perturbation correlated with HCC, a pathway analysis was also performed based on all of the differential metabolites identified in the multivariate analysis which revealed significant differences between HCC patients and controls. A total of 42 metabolic pathways were mapped in this study, and two pathways, i.e., primary bile acid biosynthesis and alanine, aspartate and glutamate metabolism, were demonstrated as notable pathways because of the topological centralities between the metabolites mapped in them.

In particular, 7 differential metabolites were mapped in the primary bile acid biosynthesis pathway (7/47). Figure 5 revealed a significant accumulation of 4 conjugated bile acids (ST_1–4) and a reduction in free bile acid (ST_8) in HCC patients. The elevated fold-changes (medians) in HCC patients were 6.07, 3.19, 2.55, 1.28 and 0.02 for ST_1 (Taurocholate), ST_2 (Taurochenodeoxycholate), ST_3 (Glycocholate), ST_4 (Glycochenodeoxycholate) and ST_8 (Chenodeoxycholate), respectively.

Therefore, ST_1 revealed a more significant up-regulation than ST_3, whereas ST_2 showed a more significant up-regulation than ST4. ST_1 and ST_2 were both conjugated with taurine, whereas ST_3 and ST_4 were both conjugated with glycine. Similarly, ST_1 also revealed a more significant up-regulation than ST_2, whereas ST_3 showed a more significant up-regulation than ST_4. Additionally, ST_1 and ST_3 both belonged to cholates, whereas ST_2 and ST_4 were chenodeoxycholates. The two taurine conjugated bile acids (ST_1 and ST_2) both showed more significant up-regulations than their glycine conjugated counterparts (ST_3 and ST_4). The two cholates (ST_1 and ST_3) both showed more significant up-regulations than their chenodeoxycholic counterparts (ST_2 and ST_4).

For the alanine, aspartate and glutamate metabolism pathway (http://www.kegg.jp/), HCC patients displayed lower levels of L-glutamine and L-aspartic acid and higher levels of L-glutamate and succinate. The probable transformation from L-glutamine to L-glutamate catalyzed by a phosphate-dependent glutaminase (GLS) suggests the accelerated material conversation of nitrogen in HCC.

4 Discussion

HCC has a high mortality and morbidity rate, remaining the fifth most common and the third leading cause of cancer-related death worldwide (Beyoglu et al. 2013; Chen and Haas-Kogan 2010). Therefore, HCC remains a challenge in mechanism explanation and clinical diagnosis. Over the past 5 years, metabolomic studies aimed at attaining biomarkers for HCC diagnosis have widely investigated the serum, urine and even tissue samples (Chen et al. 2011, 2013b; Ressom et al. 2012; Wu et al. 2009; Huang et al. 2013a; Beyoglu et al. 2013). In majority of these studies, multivariate analysis methods (i.e., PCA, PLS-DA and OPLS-DA) were commonly conducted to directly select potential biomarkers that showed the best diagnostic performance. Our study employed this conventional route, benefiting from a large cohort population (nearly 600 subjects) with comprehensive demographic and clinical characteristics. The strategy we adopted for selecting discriminative metabolites was believed to provide a reasonable process for enhancing the clinical reliability of discriminative predictors because the influence from covariate factors (gender, age and abnormal liver function) was considered from a clinical point of view and from epidemiological studies (Fairweather and Rose 2004; Mostertz et al. 2010; Chen et al. 2011; Fitian et al. 2014).

For the discriminative metabolites, an aldohexose, d-galactose was found up-regulated in HCC patients. Under the phosphorylation catalyzed by a galactokinase, D-galactose could be transformed into D-galactose-1-phosphate, which plays a key role in an exchange with glucose-1-phosphate for energy production in glycolysis (Liu et al.
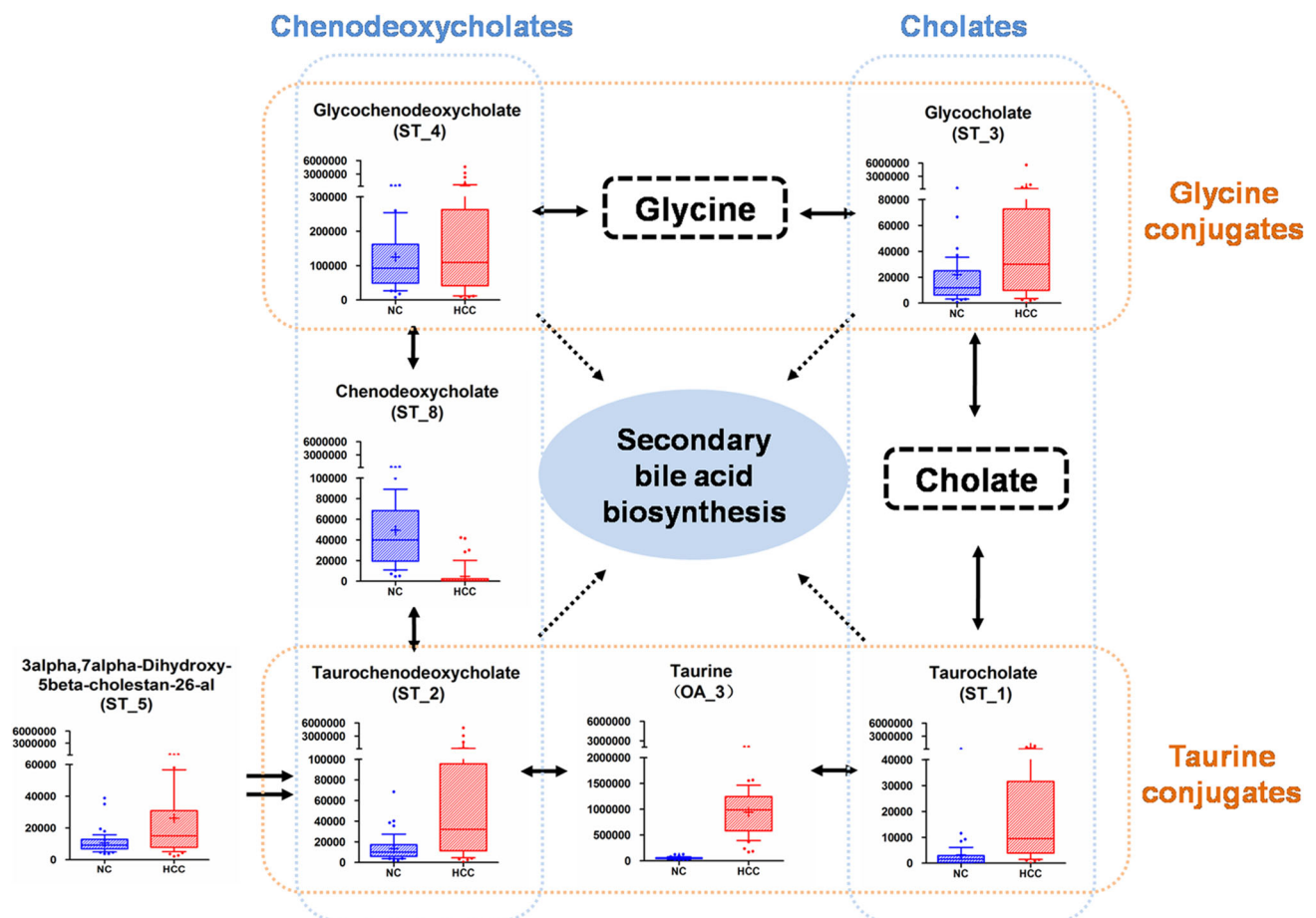
Fig. 5 Metabolic changes between HCC and NC of differential metabolites in the primary bile acid pathway. Bile acid conjugates in *orange boxes* indicate glycine and taurine conjugates, whereas those in *blue boxes* indicate the chenodeoxycholates and cholates (Color figure online)

2000; Lai and Klapa 2004). Undercanoyl-L-carnitine (C11) belongs to an acylcarnitine with an acyl-chain of odd carbons; this molecule could be viewed as an indicator of phytanic acid metabolism in peroxisomal β-oxidation (van Vlies et al. 2005). PE(P-18:0/0:0) is subject to glycerophosphoethanolamines (PEs) that have been found correlated with increased fatty acid β-oxidation during HCC development (Beyoglu and Idle 2013). The diagnostic performance Se, Sp and AUC of AFP was greatly enhanced to 0.96, 0.92 and 0.98 after the combination with the three metabolites (from 0.62, 0.98 and 0.81 for AFP alone). For the AFP-negative patients, an accuracy of 0.78 was obtained with the combination of these molecules. The core function of the discriminative metabolites is to provide a complement to the analysis for when AFP performed unsatisfactorily.

Network analyses have been widely applied in metabolomic studies to capture the relationships in global changes. Notably, all five dipeptides enrolled in our network carried a phenylalanine residue and were correlated with a reduction in HCC patients. Although dipeptides

were generally considered as incomplete breakdown products of protein digestion or protein catabolism, the specific metabolic mechanism of dipeptides in HCC patients remains rarely reported. Previous studies have demonstrated that phenylalanine dipeptides and derivatives revealed better activities for antioxidation and inhibition on the replication of HBV DNA than positive control reagents (Qiu et al. 2011; Xu et al. 2009). Soga et al. found significantly altered serum levels of γ-glutamyl dipeptides in different liver diseases involving HCC (Soga et al. 2011). The results suggest that the alterations of dipeptides might be induced by the depletion of GSH to neutralize the excessively generated reactive oxygen species (ROS). Therefore, in oxidative stress environments in HCC patients (Beyoglu and Idle 2013), the reduction of phenylalanine dipeptides might be caused by the interaction between phenylalanine dipeptides and overly generated ROS.

As an essential intermediate for transporting long chain fatty acids into mitochondrial matrix for β-oxidation, carnitine plays a key role in maintaining energy supplements

for the activities of the human body. Thus, carnitine could also be viewed as an indicator of acetyl-CoA metabolism. In the present network, medium- and long-chain acylcarnitines were found to scatter correlatively in an acyl-chain length related arrangement for both saturated (Cn) and unsaturated acylcarnitines (Cn:1). Medium-chain acylcarnitines displayed decreased levels in HCC patients and long-chain acylcarnitines increased with the elongation of the acyl-chain. Notably, several reports showed that long-chain acylcarnitines were linked to a poor clinical status whereas short-chain and medium-chain acylcarnitines were associated with positive effects (Redman et al. 2011; Lum et al. 2011). Simultaneous behaviors of medium and long-chain acylcarnitines have been found in select pathological statuses, including long-term hemodialysis and insulin sensitivity in healthy but over-weight adults (Reuter et al. 2005; Redman et al. 2011). These results were also consistent with the stepped gradient from the down-regulated medium-chain acylcarnitines to the up-regulated long-chain acylcarnitines. Particularly, changes in the medium and long-chain acylcarnitines were observed in knockout mice and were found to be related with mitochondrial dysfunction and incomplete β-oxidation (Koves et al. 2008; Noland et al. 2009). The accumulations of long-chain acylcarnitines will likely inhibit the activity of citrate lysase and increase the activity of citrate synthase, further resulting in an elevation of cytosolic citrate. This elevation might additionally contribute to the energy supplement for the cell proliferation of HCC.

A metabolic pathway analysis can be used to characterize the disturbed metabolic states in a category-integrated perspective and further elucidate the associated fluctuation of cellular functions. In our study, the majority of perturbed metabolic pathways are the primary bile acid biosynthesis and the lesser studied alanine, aspartate and glutamate metabolism pathway. The metabolites mapped in the primary bile acid biosynthesis pathway showed a well-formed centrality in the pathway topology analysis, and the accumulation of conjugated bile acids and a reduction of free bile acid were both observed in HCC patients.

Primary bile acid, which refers to the bile acid directly transformed from cholesterol in hepatocytes, could be divided into cholic acid, chenodeoxycholic acid and their conjugated forms of glycine and taurine. These forms of primary bile acid are generally believed to play important roles in both the absorption of lipids and the regulation of energy homeostasis by maintaining the activity of the thyroid hormone (Nagana Gowda et al. 2009). Previous reports also revealed the potential activities of primary bile acids as signaling molecules with systematic endocrine functions (Brendel et al. 2002; Claudel et al. 2005; Cui et al. 2003; De Fabiani et al. 2003). Therefore, these

molecules could be viewed as a barometer of the hepatic status. Abnormal concentration levels of bile acids and its conjugates have been observed in several studies of hepatic diseases, such as HBV infection, liver injury, cirrhosis and HCC (Yang et al. 2006; Yin et al. 2009; Ressom et al. 2012). One potential mechanism responsible for the accumulation of conjugated bile acids in HCC patients was the abnormally promoted activity of the rate-limiting enzyme CYP7A1, which could modulate the transformation from cholesterol to 7α-hydroxycholesterol. This enzyme therefore could play a key role in the catalysis of primary bile acids biosynthesis (Garcia-Canaveras et al. 2011). We also hypothesized that the increase in conjugated bile acids might be related with their stabilities in HCC patients because of their dissociation properties. Because of the higher pKa values of the conjugates (0.28, 0.18, −0.04 and −0.29 for ST_1, ST_2, ST_3 and ST_4, respectively, http://www.hmdb.ca/) compared with the free bile acid (−0.54 for ST_8, http://www.hmdb.ca/), the acidic environment caused by activated glycolysis in HCC (Harjes et al. 2012) tended to enhance the stabilities of the conjugated bile acids rather than those of thee free acids, further altering the metabolic balance between the conjugates and free bile acids. This imbalance would finally induce the excessive generation of bile acid conjugates and reduction of free bile acids.

Additionally, two taurine conjugated bile acids revealed a more significant up-regulation than their glycine conjugated counterparts, potentially because of the higher pKa values of taurine conjugates than the glycine counterparts. By contrast, the two cholates displayed more significant up-regulation than their chenodeoxycholic counterparts, potentially caused by the differences in the water solubility between the cholates (0.077 g/L for ST_1 and 0.025 g/L for ST_3) and chendeoxycholates (0.0075 g/L for ST_2 and 0.0079 g/L for ST_4, http://www.hmdb.ca/) in the acidic environment (Harjes et al. 2012). Therefore, we presented an explanation to the alterations in the bile acids from the perspective of the relationship between the chemical properties of the metabolites and internal microenvironment in HCC patients. This analysis provides a potential route to alleviate the toxicity of over-secreted bile acids.

Glutamine was widely recognized as a substrate that provides nitrogen for the biosynthesis of various biomolecules. The reaction from glutamine to glutamate was found to be activated because of the highly expressed glutaminase (GLS) in tumor cells (Gao et al. 2009). With the generation of glutamate in tumor cells resulting from the enhanced GLS, a large amount of α-nitrogen was obtained via the transformation from glutamine to glutamate. This process resulted in a substantial storage of α-nitrogen in the glutamate pool. Furthermore, under the activities of aminotransferases, the over-generated nitrogen tended to be

distributed into other pools of nonessential amino acids for protein synthesis (DeBerardinis and Cheng 2010). Therefore, the vast amount of glutamine and the accelerated GLS in tumor cells will unavoidably contribute to fulfill the large gaps in the building blocks for cell replication.

Notably, due to the non-targeted metabolic profiling of this study, potential metabolites in the pathways in which the three discriminative metabolites reside may not be detected. Therefore, a subsequent study of targeted HCC metabolomics upon metabolites in these pathways might be necessary in the future.

5 Concluding remarks

The discriminative metabolite selection strategy in our study was reliable because the assessment was conducted based on a multivariate analysis that adjusted for clinical covariate factors and on a correlation with AFP. Moreover, the correlation network and pathway analysis provided additional information about the relationships between differential metabolites and the possible biological mechanism for metabolites in HCC development. These results will assist biologists and clinicians in exploring mechanisms and inferring clinical diagnoses.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Anis, M., & Irshad, A. (2011). Imaging of hepatocellular carcinoma: Practical guide to differential diagnosis. *Clinics in Liver Disease, 15*(2), 335–352.

Beyoglu, D., & Idle, J. R. (2013). The metabolomic window into hepatobiliary disease. *Journal of Hepatology, 59*(4), 842–858.

Beyoglu, D., Imbeaud, S., Maurhofer, O., Bioulac-Sage, P., Zucman-Rossi, J., Dufour, J. F., et al. (2013). Tissue metabolomics of hepatocellular carcinoma: Tumor energy metabolism and the role of transcriptomic classification. *Hepatology, 58*(1), 229–238.

Brendel, C., Schoonjans, K., Botrugno, O. A., Treuter, E., & Auwerx, J. (2002). The small heterodimer partner interacts with the liver X receptor alpha and represses its transcriptional activity. *Molecular Endocrinology, 16*(9), 2065–2076.

Bruce, S. J., Tavazzi, I., Parisod, V., Rezzi, S., Kochhar, S., & Guy, P. A. (2009). Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. *Analytical Chemistry, 81*(9), 3285–3296.

Bruix, J., Sherman, M., Practice Guidelines Committee, & American Association for the Study of Liver Diseases. (2005). Management of hepatocellular carcinoma. *Hepatology, 42*(5), 1208–1236.

Chen, C. P., & Haas-Kogan, D. (2010). Neoplasms of the hepatobiliary system: Clinical presentation, molecular pathways and diagnostics. *Expert Review of Molecular Diagnostics, 10*(7), 883–895.

Chen, S., Hoene, M., Li, J., Li, Y., Zhao, X., Haring, H. U., et al. (2013a). Simultaneous extraction of metabolome and lipidome with methyl tert-butyl ether from a single small tissue sample for ultra-high performance liquid chromatography/mass spectrometry. *Journal of Chromatography A, 1298*, 9–16.

Chen, S., Kong, H., Lu, X., Li, Y., Yin, P., Zeng, Z., et al. (2013b). Pseudotargeted metabolomics method and its application in serum biomarker discovery for hepatocellular carcinoma based on ultra high-performance liquid chromatography/triple quadrupole mass spectrometry. *Analytical Chemistry, 85*(17), 8326–8333.

Chen, T., Xie, G., Wang, X., Fan, J., Qiu, Y., Zheng, X., et al. (2011a). Serum and urine metabolite profiling reveals potential biomarkers of human hepatocellular carcinoma. *Molecular & Cellular Proteomics, 10*(7), M110.004945.

Chen, F., Xue, J., Zhou, L., Wu, S., & Chen, Z. (2011b). Identification of serum biomarkers of hepatocarcinoma through liquid chromatography/mass spectrometry-based metabonomic method. *Analytical and Bioanalytical Chemistry, 401*(6), 1899–1904.

Claudel, T., Staels, B., & Kuipers, F. (2005). The Farnesoid X receptor: A molecular link between bile acid and lipid and glucose metabolism. *Arteriosclerosis, Thrombosis, and Vascular Biology, 25*(10), 2020–2030.

Cui, J., Huang, L., Zhao, A., Lew, J. L., Yu, J., Sahoo, S., et al. (2003). Guggulsterone is a farnesoid X receptor antagonist in coactivator association assays but acts to enhance transcription of bile salt export pump. *Journal of Biological Chemistry, 278*(12), 10214–10220.

De Fabiani, E., Mitro, N., Gilardi, F., Caruso, D., Galli, G., & Crestani, M. (2003). Coordinated control of cholesterol catabolism to bile acids and of gluconeogenesis via a novel mechanism of transcription regulation linked to the fasted-to-fed cycle. *Journal of Biological Chemistry, 278*(40), 39124–39132.

DeBerardinis, R. J., & Cheng, T. (2010). Q's next: The diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene, 29*(3), 313–324.

El Khouli, R. H., Macura, K. J., Barker, P. B., Habba, M. R., Jacobs, M. A., & Bluemke, D. A. (2009). Relationship of temporal resolution to diagnostic performance for dynamic contrast enhanced MRI of the breast. *Journal of Magnetic Resonance Imaging, 30*(5), 999–1004.

Fairweather, D., & Rose, N. R. (2004). Women and autoimmune diseases. *Emerging Infectious Diseases, 10*(11), 2005–2011.

Fitian, A. I., Nelson, D. R., Liu, C., Xu, Y., Ararat, M., & Cabrera, R. (2014). Integrated metabolomic profiling of hepatocellular carcinoma in hepatitis C cirrhosis through GC/MS and UPLC/MS-MS. *Liver Int, 34*(9), 1428–1444.

Gao, P., Tchernyshyov, I., Chang, T. C., Lee, Y. S., Kita, K., Ochi, T., et al. (2009). c-Myc suppression of miR-23a/b enhances

mitochondrial glutaminase expression and glutamine metabolism. *Nature, 458*(7239), 762–765.

Garcia-Canaveras, J. C., Donato, M. T., Castell, J. V., & Lahoz, A. (2011). A comprehensive untargeted metabonomic analysis of human steatotic liver tissue by RP and HILIC chromatography coupled to mass spectrometry reveals important metabolic alterations. *Journal of Proteome Research, 10*(10), 4825–4834.

Harjes, U., Bensaad, K., & Harris, A. L. (2012). Endothelial cell metabolism and implications for cancer therapy. *British Journal of Cancer, 107*(8), 1207–1212.

Herrgard, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., et al. (2008). A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nature Biotechnology, 26*(10), 1155–1160.

Huang, Q., Tan, Y., Yin, P., Ye, G., Gao, P., Lu, X., et al. (2013a). Metabolic characterization of hepatocellular carcinoma using nontargeted tissue metabolomics. *Cancer Research, 73*(16), 4992–5002.

Huang, J. H., Yan, J., Wu, Q. H., Duarte Ferro, M., Yi, L. Z., Lu, H. M., et al. (2013b). Selective of informative metabolites using random forests based on model population analysis. *Talanta, 117*, 549–555.

Klamt, S., Saez-Rodriguez, J., & Gilles, E. D. (2007). Structural and functional analysis of cellular networks with Cell NetAnalyzer. *BMC Systems Biology, 1*, 2.

Koves, T. R., Ussher, J. R., Noland, R. C., Slentz, D., Mosedale, M., Ilkayeva, O., et al. (2008). Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance. *Cell Metabolism, 7*(1), 45–56.

Lai, K., & Klapa, M. I. (2004). Alternative pathways of galactose assimilation: Could inverse metabolic engineering provide an alternative to galactosemic patients? *Metabolic Engineering, 6*(3), 239–244.

Liu, G., Hale, G. E., & Hughes, C. L. (2000). Galactose metabolism and ovarian toxicity. *Reproductive Toxicology, 14*(5), 377–384.

Lum, H., Sloane, R., Huffman, K. M., Kraus, V. B., Thompson, D. K., Kraus, W. E., et al. (2011). Plasma acylcarnitines are associated with physical performance in elderly men. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 66*(5), 548–553.

Marquardt, J. U., Galle, P. R., & Teufel, A. (2012). Molecular diagnosis and therapy of hepatocellular carcinoma (HCC): An emerging field for advanced technologies. *Journal of Hepatology, 56*(1), 267–275.

Montuschi, P., Paris, D., Melck, D., Lucidi, V., Ciabattoni, G., Raia, V., et al. (2012). NMR spectroscopy metabolomic profiling of exhaled breath condensate in patients with stable and unstable cystic fibrosis. *Thorax, 67*(3), 222–228.

Morita, Y., Sakaguchi, T., Ikegami, K., Goto-Inoue, N., Hayasaka, T., Hang, V. T., et al. (2013). Lysophosphatidylcholine acyltransferase 1 altered phospholipid composition and regulated hepatoma progression. *Journal of Hepatology, 59*(2), 292–299.

Mostertz, W., Stevenson, M., Acharya, C., Chan, I., Walters, K., Lamlertthon, W., et al. (2010). Age- and sex-specific genomic profiles in non-small cell lung cancer. *JAMA, 303*(6), 535–543.

Motta, A., Paris, D., Melck, D., de Laurentiis, G., Maniscalco, M., Sofia, M., et al. (2012). Nuclear magnetic resonance-based metabolomics of exhaled breath condensate: Methodological aspects. *European Respiratory Journal, 39*(2), 498–500.

Nagana Gowda, G. A., Shanaiah, N., Cooper, A., Maluccio, M., & Raftery, D. (2009). Bile acids conjugation in human bile is not random: New insights from (1)H-NMR spectroscopy at 800 MHz. *Lipids, 44*(6), 527–535.

Nicholson, J. K., Lindon, J. C., & Holmes, E. (1999). 'Metabonomics': Understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica, 29*(11), 1181–1189.

Noland, R. C., Koves, T. R., Seiler, S. E., Lum, H., Lust, R. M., Ilkayeva, O., et al. (2009). Carnitine insufficiency caused by aging and overnutrition compromises mitochondrial performance and metabolic control. *Journal of Biological Chemistry, 284*(34), 22840–22852.

Qiu, J., Xu, B., Huang, Z., Pan, W., Cao, P., Liu, C., et al. (2011). Synthesis and biological evaluation of Matijing-Su derivatives as potent anti-HBV agents. *Bioorganic & Medicinal Chemistry, 19*(18), 5352–5360.

Redman, L. M., Huffman, K. M., Landerman, L. R., Pieper, C. F., Bain, J. R., Muehlbauer, M. J., et al. (2011). Effect of caloric restriction with and without exercise on metabolic intermediates in nonobese men and women. *Journal of Clinical Endocrinology and Metabolism, 96*(2), E312–E321.

Ressom, H. W., Xiao, J. F., Tuli, L., Varghese, R. S., Zhou, B., Tsai, T. H., et al. (2012). Utilization of metabolomics to identify serum biomarkers for hepatocellular carcinoma in patients with liver cirrhosis. *Analytica Chimica Acta, 743*, 90–100.

Reuter, S. E., Evans, A. M., Faull, R. J., Chace, D. H., & Fornasini, G. (2005). Impact of haemodialysis on individual endogenous plasma acylcarnitine concentrations in end-stage renal disease. *Annals of Clinical Biochemistry, 42*(Pt 5), 387–393.

Slany, A., Haudek-Prinz, V., Zwickl, H., Stattner, S., Grasl-Kraupp, B., & Gerner, C. (2013). Myofibroblasts are important contributors to human hepatocellular carcinoma: Evidence for tumor promotion by proteome profiling. *Electrophoresis, 34*(24), 3315–3325.

Soga, T., Sugimoto, M., Honma, M., Mori, M., Igarashi, K., Kashikura, K., et al. (2011). Serum metabolomics reveals gamma-glutamyl dipeptides as biomarkers for discrimination among different forms of liver disease. *Journal of Hepatology, 55*(4), 896–905.

Uto, H., Kanmura, S., Takami, Y., & Tsubouchi, H. (2010). Clinical proteomics for liver disease: a promising approach for discovery of novel biomarkers. *Proteome Science, 8*, 70.

van Vlies, N., Tian, L., Overmars, H., Bootsma, A. H., Kulik, W., Wanders, R. J., et al. (2005). Characterization of carnitine and fatty acid metabolism in the long-chain acyl-CoA dehydrogenase-deficient mouse. *Biochemical Journal, 387*(Pt 1), 185–193.

Wang, Q., Gao, P., Wang, X., & Duan, Y. (2014). The early diagnosis and monitoring of squamous cell carcinoma via saliva metabolomics. *Scientific Reports, 4*, 6802.

Wiklund, S., Nilsson, D., Eriksson, L., Sjostrom, M., Wold, S., & Faber, K. (2007). A randomization test for PLS component selection. *Journal of Chemometrics, 21*(10–11), 427–439.

Wu, H., Xue, R., Dong, L., Liu, T., Deng, C., Zeng, H., et al. (2009). Metabolomic profiling of human urine in hepatocellular carcinoma patients using gas chromatography/mass spectrometry. *Analytica Chimica Acta, 648*(1), 98–104.

Xu, B., Huang, Z., Liu, C., Cai, Z., Pan, W., Cao, P., et al. (2009). Synthesis and anti-hepatitis B virus activities of Matijing-Su derivatives. *Bioorganic & Medicinal Chemistry, 17*(8), 3118–3125.

Yang, J., Zhao, X., Liu, X., Wang, C., Gao, P., Wang, J., et al. (2006). High performance liquid chromatography-mass spectrometry for metabonomics: Potential biomarkers for acute deterioration of liver function in chronic hepatitis B. *Journal of Proteome Research, 5*(3), 554–561.

Yin, P., Wan, D., Zhao, C., Chen, J., Zhao, X., Wang, W., et al. (2009). A metabonomic study of hepatitis B-induced liver cirrhosis and hepatocellular carcinoma by using RP-LC and HILIC coupled with mass spectrometry. *Molecular BioSystems, 5*(8), 868–876.

Zeng, Z., Liu, X., Dai, W., Yin, P., Zhou, L., Huang, Q., et al. (2014). Ion fusion of high-resolution LC-MS-based metabolomics data to discover more reliable biomarkers. *Analytical Chemistry, 86*(8), 3793–3800.

Zhang, T., Wu, X., Ke, C., Yin, M., Li, Z., Fan, L., et al. (2013). Identification of potential biomarkers for ovarian cancer by urinary metabolomic profiling. *Journal of Proteome Research, 12*(1), 505–512.