Metabolomics Study of Stepwise Hepatocarcinogenesis From the Model Rats to Patients: Potential Biomarkers Effective for Small Hepatocellular Carcinoma Diagnosis*

Yexiong Tan‡**, Peiyuan Yin§**, Liang Tang‡, Wenbin Xing¶, Qiang Huang§, Dan Cao‡, Xinjie Zhao§, Wenzhao Wang§, Xin Lu§, Zhiliang Xu§, Hongyang Wang‡¶, and Guowang Xu§

The aim of this study is to find the potential biomarkers from the rat hepatocellular carcinoma (HCC) disease model by using a non-target metabolomics method, and test their usefulness in early human HCC diagnosis. The serum metabolic profiling of the diethylnitrosamine-induced rat HCC model, which presents a stepwise histopathological progression that is similar to human HCC, was performed using liquid chromatography-mass spectrometry. Multivariate data analysis methods were utilized to identify the potential biomarkers. Three metabolites, taurocholic acid, lysophosphoethanolamine 16:0, and lysophosphatidylcholine 22:5, were defined as "marker metabolites," which can be used to distinguish the different stages of chemical hepatocarcinogenesis. These metabolites represented the abnormal metabolism during the progress of hepatocarcinogenesis, which could also be found in patients. To test their diagnostic potential 412 sera from 262 patients with HCC, 76 patients with cirrhosis and 74 patients with chronic hepatitis B were collected and studied, it was found that 3 marker metabolites were effective for the discrimination of small liver tumor (solitary nodules of less than 2 cm in diameter) patients, achieved a sensitivity of 80.5% and a specificity of 80.1%, which is better than those of α-fetoprotein (53 and 64%, respectively). Moreover, they were also effective for the discrimination of all HCCs and chronic liver disease patients, which could achieve a sensitivity of 87.5% and a specificity of 72.3%, better than those of α-fetoprotein (61.2 and 64%). These results indicate metabolomics method has the potential of finding biomarkers for the early diagnosis of HCC. Molecular & Cellular Proteomics 11: 10.1074/mcp.M111.010694, 1–12, 2012.

Hepatocellular carcinoma (HCC) is a type of malignancy with a high mortality rate worldwide, especially in the East Asian countries (1). In China, chronic hepatitis B virus (HBV) infection and subsequent liver cirrhosis are major precancerous lesions in the majority of HCC cases. An early diagnosis of small HCCs in these precancerous cases may greatly improve the outcome of HCC treatment (2). The current screening methods for the high risk population, such as ultrasound or the serum surveillance of tumor markers (mainly α-fetoprotein (AFP)), although effective, are far from ideal (3, 4). For example, a subset of patients with chronic hepatitis and/or cirrhosis exhibit modest elevations (10–500 ng/ml) of serum AFP, which may lead to an incorrect diagnosis (5). Hence, new biomarkers for monitoring hepatocarcinogenesis would be of great clinical importance.

Emerging platforms in the biomedical arena provide a new methodology to identify novel biomarkers. Within the framework of systems biology, metabolomics focuses on the quantitative measurement of holistic endogenous metabolites and is increasingly used in clinical fields that focus on the pathophysiological and diagnostic study of diseases (6). Metabolic fingerprinting and metabolite biomarkers have been studied for use in the discrimination or diagnosis of carcinoma (7, 8), diabetes mellitus (9) and inborn errors (10). Nontarget metabolomics approaches have also been used to search for new biomarkers and to explore the mechanism of carcinogenesis in hepatic diseases (11–20). According to these stud-

1 The abbreviations used are: HCC, hepatocellular carcinoma; CLD, chronic liver diseases; DEN, diethylnitrosamine; LPE, lysophosphoethanolamine; TCA, Taurocholic acid; PE, phosphoethanolamine; PC, phosphatidylcholine; LPC, lysophosphatidylcholine; HBV, hepatitis B virus; CT, computed tomography; MRI, magnetic resonance imaging; AFP, Alpha Fetoprotein; SD, Sprague-Dawley; PLS-DA, partial least squares discriminate analysis; TIC, total ion counts chromatogram; VIP, variable importance in the projection; HCA, hierarchical cluster analysis; ROC, receiver operating characteristic curve; AUC, area under the curve; ALT, Alanine Aminotransferase; AST, Aspartate Aminotransferase; PEMT 2, PE N-methyltransferase 2; PUFA, poly unsaturated fatty acid; FFA, Free fatty acid.

Want to cite this article? Please look on the last page for the proper citation format.
ies, several circulation metabolites have been reported to be predictors for HCC, such as hexanal, 1-octen-3-ol, octane (18), and a combination of multimetabolites (21). Soga et al. found that serum γ-glutamyl dipeptides can be used as biomarkers for discriminating different liver diseases (22). In our previous metabolomics studies on HCC based on the metabolic profiling of serum and urine (14, 19), we found that the levels of several metabolites, such as sphingosines in blood and carnitines in urine, were statistically significantly different between the patients and controls. These studies examined a range of metabolites that represented the metabolic deregulation of HCC patients and illustrated the ability of metabolomics to identify the potential biomarkers of HCC. However, no individual metabolites or their combination were evaluated for their value as biomarkers to monitor the hepatocarcinogenesis process and HCC early diagnosis.

The metabolome in the clinical samples fluctuates frequently depending on various genetic and environmental factors, and discovery of biomarkers using these clinical samples faces a great technical challenge because of the enormous complexity of the serum metabolome. In addition, it is not easy to obtain serial clinical samples of individual liver diseases that represent the stages of chronic hepatitis, cirrhosis, and HCC, respectively. In reverse, the different stages of liver disease development in the rat model of HCC can be well anticipated. It has been found that histological and genetic signatures of diethylnitrosamine (DEN)-induced hepatocarcinogenesis are similar to those of human HCC (23, 24). This rat model has also been used to investigate the progression of cirrhosis to HCC (25–27). Therefore, in the present study, a metabolomics approach was applied to the DEN-induced rat model to study the metabolic features of hepatocarcinogenesis using liquid chromatography-mass spectrometry (LC-MS). The “marker metabolites,” which were defined from the model rat, were applied to distinguish patients with chronic hepatitis, cirrhosis, or small HCC to evaluate their capacity for diagnosing small HCC.

**EXPERIMENTAL PROCEDURES**

**Animal Experiments**—The present study conformed to the Guide for the Care and Use of Laboratory Animals from the Second Military Medical University. The animal experiments were briefly described in our previous paper (28). A total of 80 male Sprague-Dawley (S.D.) rats were obtained from the Shanghai Experimental Animal Centre and were enrolled in the present study at the age of 42 days. The starting time for the animal experiment was defined as week 0. The model group included 52 rats, and 6 of these rats died of liver failure during the experiment. The control group consisted of 28 rats. The DEN administration (70 mg/kg body weight) was given to the model rats using intraperitoneal injections once a week after 14 days of habituation. In addition, saline injections of equivalent volumes were administered to the control animals. Ten injections were administered to each animal between week 2 and week 11.

To verify the histological progression of HCC, 4 rats from the model group and 2 rats from the control group were sacrificed under pentobarbital anesthesia every 2 weeks from week 2 until all of the surviving animals (n = 20) were sacrificed (20 weeks). The liver tissues were fixed in 10% buffered formalin and embedded in paraffin. The gross examinations and histopathological studies (hematoxylin–eosin staining) were used to monitor the progress of the carcinogenesis. The sera were sequentially collected after 8 h of overnight fasting. This collection was conducted from week 6 to week 20 once every 2 weeks. The sera were stored at −80 °C until analysis. The analyses were performed in a month after all samples were collected.

The detailed serum and tissue collection times are shown in Fig. 1.

**Collection of the Human Sera**—Human samples were obtained from National Liver Tissue Bank in the Second Military Medical University, Shanghai, China. All samples were collected and stored under the same standard operation protocol. Written informed consent was given by all participants. The present study was approved by the ethics committee of the Second Military Medical University, Shanghai, China. All of the serum specimens for the present study were collected with permission from 262 patients with HCC, 76 patients with cirrhosis, and 74 patients with hepatitis B. HCC was histopathologically diagnosed after the tumor excision. The age of HCC patients was 30 to 78 years including 39 female and 233 male cases. A total of 77 HCC patients were in early stages of HCC (Edmonson I, II) and 95 patients had a solitary nodule smaller than 2 cm. The diagnosis of chronic hepatitis B or cirrhosis was made using clinical, imaging, and laboratory evidence of hepatic decompensation or portal hypertension. The age and gender distributions of the participants in each group were matched, and the collecting time point and thereby the storage time of samples from different groups were similar. Detailed information about the participants is shown in Table I. The sera were stored at −80 °C until analysis. Sera were thawed at room temperature before analysis. Acetonitrile (400 μl) was added to the sera (100 μl) to quench the enzyme activity and precipitate the proteins. After 10 min of centrifugation at 10,000 × g, the supernatant (350 μl) from each sample was stored in an autosampler vial.

**Chemicals**—Acetonitrile and formic acid (HPLC grade) were purchased from Merck and TEDIA, respectively. lysophosphoethanolamine (LPE) 16:0, lysophosphatidylcholine (LPC) 16:0 and LPC 18:0 were purchased from Avanti polar lipids (Alabaster, AL), and taurocholic acid (TCA), glycocholic acid (GCA), cholic acid, tryptophan, docosahexaenoic acid, stearolic acid, oleic acid, clupanodonic acid and DEN were purchased from Sigma-Aldrich. The distilled water was filtered through a Milli-Q system (Millipore, Billerica, MA).

**Metabolic Profiling and the Target Metabolite Analysis**—The rat serum metabolic profiling analysis was performed using a LC-MS method on a 1200 Rapid Resolution Liquid Chromatography (RRLC) system that was coupled to a 6510 quadrupole time-of-flight (Q TOF) MS (Agilent, Santa Clara, CA). The used column was ZORBAX TM SB-AQ (10 cm × 2.1 mm 1.8 μm, Agilent, USA) running with the buffer consisting of water containing 0.1% formic acid (v/v) (A) and acetonitrile (B). The detailed LC and MS methods were described in our previous papers (14, 28) and the supplementary materials. The injection volume was 4 μl and the flow rate of the LC system was 0.3 ml/min. The total running time was 35 min including 11 min rinsing with acetonitrile and column equilibration.

The human serum metabolic profiling were acquired using a Thermo Fisher Accela LC system that was coupled to an LTQ Orbitrap XL system. The LC column and the elution gradient for this system were the same as those in the Agilent 1200 system. A high resolution MS of LTQ Orbitrap was used as the detector. The MS capillary temperature was held at 300 °C, and the spray voltage was 4.5 kV. The flow rate of the sheath gas was 35 arbitrary units, and the flow rate of the aux gas was 5 arbitrary units. The MS system was operated using the positive ion mode, and the mass range was set at 100–1000 m/z. The resolution of the Orbitrap was set at 60000, with a scan rate of 1 s/spectrum. At the defined operational conditions, the mass errors of the three “marker metabolites” were less than 2 ppm.
To obtain the contents of the three metabolites, all of the metabolites in a total ion chromatogram were extracted and aligned using the SIEVE software (V1.2, Thermo Fisher). The frame settings were 0.5 min (retention time) and 0.005 m/z (mass error). Next, the total peak area from one sample was assigned to a constant of 10000; the peak areas of all of the metabolites were normalized to this total area. After the confirmation of the three target metabolites, their normalized peak areas were exported to an Excel table (Microsoft, Redmond, WA) for the statistical analysis.

To ensure the stability and repeatability of the LC-MS systems, pooled quality control (QC) samples were used as in the literature (29–31). The QC samples were prepared from 10 µl of each sample and analyzed together with the other samples. A total of five runs of the QC samples were performed on the system before the sample sequence. The QC samples were also inserted and analyzed in every 10 samples.

Data Analysis—SIMCA-P 11.0 (Umetrics AB, Umeå, Sweden) was used for the chemometrics analysis. The multivariate pattern recognition technique of the partial least squares discriminate analysis (PLS-DA) with the pareto scaling was performed. The parameters of PLS-DA (R²Y, Q²Y) represent the goodness of the fit and the prediction ability of the models (32). Response permutation test was used to assess whether the model established exhibited overfitting because of the chance correlations.

SPSS 13.0 for Windows was used for the statistical analysis. The data were analyzed using the Wilcoxon Mann-Whitney Test, with p < 0.05 set as the level of statistical significance. A binary logistic regression was also performed using this software (33). After the regression, the values of the prediction probability were applied to the classification of the samples. Receiver operating characteristic curve (ROC) was made by using the SPSS software. The cutoff values were calculated based on the results of the ROC. The optimized cutoff values in this study were those corresponding with the highest accuracy (maximum sensitivity and specificity). Because different ROC curves were performed for animals, small HCC and HCC, 3 optimized cutoff values were defined. The Multi Experiment Viewer software (Version 4.5.1, http://www.tm4.org) was used for the hierarchical clustering analysis (HCA) and significance analysis for microarrays (SAM). The SAM method was performed according to Tusher et al. (34).

### Results

**Metabolic Profiling of the Rat HCC Model**—During DEN administration of 52 rats, six of these rats died of liver function failure or a hemorrhage of the tumor during the experiment. A total of 36 rats from the model group and 18 rats from the control group were sacrificed for histological observations (Fig. 1). According to the histological findings, all of the DEN treated rats that were alive at the end of the study (week 20) exhibited incidences of the liver tumor, as shown in the pathological analysis. The serial progression of hepatocarcinogenesis in this animal model was divided into three stages: the inflammation stage (week 4–8), the cirrhosis stage (weeks 10–14), and the HCC stage (week 16–20). The time points at “week 6,” “week 14,” and “week 20” were the characteristic histological changes of the inflammation, cirrhosis, and HCC stages, respectively (Fig. 1). In addition, “week 10” and “week 16” were time points that were between two stages with mixed features, depending on the individual animals (data not shown here). Finally, the sera from each stage of the 10 DEN treated rats and the 10 matched control rats were collected for a metabolic profile analysis (Fig. 1). Fig. 2A shows the rat serum total ion chromatograms (TIC) of the control group at the normal growth process and the diseased rats at the different pathological stages, including inflammation (week 6), cirrhosis (week 14), and carcinoma (week 20). These chromatograms show the metabolic alterations among the different groups.

After the peak detection and alignment of all TICs, a total of 1459 metabolite ions were enrolled in the final data set for the statistical analysis. To model and evaluate the systemic changes in the rat metabolome, a partial least squares discriminate analysis (PLS-DA) of the data for each typical stage of the lesions was performed. Five components were calculated, and the cumulative R²Y and Q² were 0.6 and 0.26, respectively. No overfitting of the data was observed based

### Table I

| Samples  | Hepatitis | Cirrhosis | HCC     |
|----------|-----------|-----------|---------|
| Sample No. | 74        | 76        | 262     |
| Age (yr)  | 47(22–71) | 55(35–79) | 52(30–78) |
| Sex (F/M) | 26/48     | 24/52     | 39/223  |
| AFP>20(µg/L) | 29        | 25        | 158     |
| ALT (U/L) | 175.6 ± 224.4 | 74.6 ± 128.1 | 63.7 ± 102.3 |
| AST (U/L) | 104.9 ± 127.0 | 81.3 ± 85.6 | 67.9 ± 100.4 |
| HBsAg (yes/no) | 74/0      | 76/0      | 223/39  |
| HCV (yes/no) | 0/74      | 0/76      | 1/261   |
| Edmonson stage | I, II  | III, IV   | 77      |
| Tumor diameter (cm) | <2       | 95       |
| 2–5       | 71       |
| >5        | 96       |
on the results of the permutation test (30) (the $R^2$-Y-intercept was 0.43, and the $Q^2$-intercept was $-0.34$). As shown in Fig. 2, the age-related metabolome changed along the direction of the second component, whereas the carcinogenesis-related metabolome changed mainly along the first component. The PLS-DA loadings indicated that metabolites such as palmityl-L-carnitine ($m/z$: 400.3423) were age-related. And the metabolites such as LPC 22:5 and LPE 16:0 were carcinogenesis-related.

As an ideal biomarker, it should be related to only the disease state, and have very few or no interference from the nondisease factors. To reduce the influence of the animal age, a Wilcoxon Mann-Whitney test was performed between the control and model groups with the same ages. It was found that 382, 427, and 445 variables significantly changed ($p < 0.05$) between the model and the control at the inflammation stage (week 6, 8), the cirrhosis stage (week 12, 14) and carcinoma (week 18, 20) stages, respectively. A total of 706 metabolites exhibited statistically significant differences ($p < 0.05$) in at least one stage of liver disease (Fig. 2C). These metabolites could better reflect the metabolic trends of the tumorigenesis and were also able to classify the three stages of the liver diseases (inflammation, cirrhosis, and carcinoma) using a PLS-DA model ($R^2$Y was 0.76, and $Q^2$ was 0.4, Fig. 2D). The PLS-DA model achieved a better classification between cirrhosis and carcinoma, which was overlapped when all the variables were used in Fig. 2B.

Selection and Identification of Important Differential Metabolites—Using the above PLS-DA model that was based on 706 ions with significant differences, the statistically important metabolites were studied. According to the variable importance in the projection (VIP), a total of 52 variables (ions) with a VIP $>2$ were selected, and the chemical structures of 44 ions (supplementary Table S1) were identified based on the metabolite identification strategy (35, 36). It was observed that these metabolites could also be found in the data set from the human metabolic profiling. Therefore, the results that were obtained from the rat models may be possibly extended to patients.

To further select the potential biomarkers from the 52 variables (ions) with a VIP $>2$, a heat map was constructed, providing the relative average contents of the selected ions in the model animals compared with the corresponding contents in the age-matched control animals (Fig. 3A). Several metabolites exhibited a characteristic trend of alterations that indicated the stage of the progression of hepatocarcinogenesis. For example, LPE 16:0 ($m/z$: 454.2928) and LPE 18:0 ($m/z$: 482.3245) exhibited higher contents from weeks 12 to 20. In addition, the level of the metabolite LPC 22:5 ($m/z$: 570.3547) increased at week 16 when tumors were observed in the models. Several other metabolites appeared at low levels in all of the model animals, such as docosahexaenoic acid ($m/z$: 329.2471), whereas other metabolites, such as TCA ($m/z$: 480.2776) and GCA ($m/z$: 412.2839), remained at higher levels in the model group.

To narrow down the scope of the biomarker pool, HCA was performed to understand the potential relationships among the metabolites. These metabolites were clustered according
to their Pearson correlation coefficients, which were shown on the plot at different colors (Fig. 3B). The closely related metabolites were clustered, six major clusters were observed (Fig. 3B, I–VI). Cluster I included two poly-unsaturated fatty acids (PUFAs) (clupanodonic acid and docosahexaenoic acid) and LPC 22:5. Cluster II was mainly fatty acids and carnitines. Cluster III was all LPCs with different carbon chains. Cluster IV consisted of LPE, and cluster V was mainly LPC. Cluster VI was mainly bile acids and related ions.

Based on the above results and our previous work (28), the representative characteristic metabolites were selected from each cluster including LPC 22:5, palmityl-L-carnitine, LPC 22:6, LPE 16:0, LPC O-16:0 and TCA, which correspond to clusters I to VI, respectively. The metabolites in the same cluster also had similar changing trends. The metabolic trajectories of these six characteristic metabolites during carcinogenesis are shown in Fig. 4. The level of LPC 22:5 appeared to increase significantly (Mann-Whitney test, p < 0.05) during the early stage of HCC (week 16) and greatly increased as the tumor progressed compared with the control animals (Fig. 4A). The level of palmityl-L-carnitine decreased with the age of the animals in both groups (Fig. 4B), however, it increased significantly (p < 0.05) in week 8 between two groups. The level of LPE 16:0 increased as the cirrhosis developed (weeks 8–14), and reached its peak during the advanced stage of HCC (weeks 20, Fig. 4C), and during this period, it increased significantly between the models and the controls (p < 0.05). The level of LPC 22:6 (Fig. 4D) decreased during the late stage of inflammation (weeks 8–10, p < 0.01), while increased in the late stage of HCC (weeks 20, p < 0.05). LPC O-16:0 (Fig. 4E) increased during the cirrhotic stage (weeks 12, p < 0.05) and in the advanced tumor stage (weeks 20, p < 0.001). The level of TCA significantly increased compared with that in the corresponding control (p < 0.01), especially at week 10 (Fig. 4F).

The “significance analysis for microarrays, (SAM)” method (34) was used to select the most significant metabolites.
Metabolomics Study of Early Biomarkers for HCC
results indicated that GCA, LPE 16:0, TCA, and LPE 18:0 were the most significant differential metabolites for the classification of the model and the control (supplemental Fig. S1A). And LPC 22:5 and LPC 20:4 are the most discriminant metabolites for distinguishing the carcinoma and CLD (supplemental Fig. S1B). According to the changing folds and correlations shown in Fig. 3A and 3B, LPE 16:0 and LPE 18:0 had a similar metabolic trajectory. Because LPE 16:0 had a smaller \( p \) value (t test), it was selected as a candidate marker from two LPEs. With the same reason for GCA and TCA, TCA was selected. For the selection of HCC related markers in supplemental Fig. S1B, LPC 20:4 was rejected because of its similar alteration between the model and the control, only in one or two stages had they a significant difference (similar metabolic trajectory to the metab-

**Fig. 3. Differential metabolites and their relationships.** A, Heat map of the 52 differential metabolites, with the degree of change in the models compared with the controls marked with colors. B, Hierarchical clustering of the differential metabolites. Pearson correlation coefficients of the 52 differential metabolites are marked on the plot. Six clusters were identified representing the different groups of metabolites. *Significantly changed metabolites during the progression of the liver cancer.

**Fig. 4. Alteration of the importantly differential metabolites in the animals.** The changing fold of each metabolite (mean ± S.E.) was the ratio of the average content in the model group to 10 control rats at week 6. A, LPC 22:5, B, Palmityl-L-carnitine, C, LPE 16:0, D, LPC 22:6, E, LPC O-16:0, F, TCA. Mann-Whitney test was applied to compare the contents of metabolites in the models and the controls from the same week. (*) \( p < 0.05 \), (**) \( p < 0.01 \), (***) \( p < 0.001 \).
olites in cluster III, such as LPC 22:6). The removal of LPC 20:4 can lower the risk of false positive. Finally, three metabolites including LPC 22:5, LPE 16:0, and TCA were selected as candidate markers for further validation.

**Discrimination of the Hepatocarcinogenesis Stages**—According to the relative contents of the metabolites (Fig. 4) and the SAM analysis (supplemental Fig. S1A, S1B), a discrimination flowchart is given in Fig. 5A.

Samples were randomly divided into a training data set (80% samples) to build the logistic model and a test data set (20% samples) for validation. LPE 16:0 and TCA were first applied to the classification of the control and diseased animals: 95.2% of the controls and 85.9% of the diseased animals were correctly discriminated in the training group, and 93.8% of the controls and 87.5% of the diseased animals in the test group were also correctly discriminated using the same model (cutoff value: 0.5, Fig. 5B). Next, the HCC animals were discriminated from other non-HCC rats using LPC 22:5, 81.2% of the HCC (weeks 18, 20), and 97.5% of the non-HCC (weeks 6–15) rats were correctly discriminated, and the accuracy was 80 and 75% in the test group, respectively (cutoff value: 0.5, Fig. 5C). Similarly, 93.8% of the inflammation (weeks 6, 8) and 93.8% of the cirrhosis (weeks 12, 14) animals were correctly identified using all three metabolites, and only 1 animal with hepatitis was wrongly classified (cutoff value: 0.5, Fig. 5D). Furthermore, Results from another DEN-induced HCC rat model cohort also provided similar ability of classification using the above logistic model (supplemental Fig. S2). The animals that were in the border stages (“week 10” and “week 16”) were also predicted using the above binary logistic regression methods with the data from the rat models. A. Flowchart of the discriminating liver diseases based on the three marker metabolites. B. Liver disease animal models (weeks 6–20) versus control animals (weeks 6–20). C. HCC (weeks 18, 20) versus CLD (weeks 6–14). The state in the border stage (week 16) was predicted using the above HCC and the CLD binary logistic regressions. D. Hepatitis (Hep, weeks 6–8) versus cirrhosis (Cir, weeks 12–14). The state in the border stage (10W, week 10) was predicted using the above hepatitis versus cirrhosis binary logistic regressions. T: training dataset, S: test data set.
regression, four of the 10 rats at week 10 were identified as having cirrhosis, and three of the 10 rats at “week 16” were predicted to be tumor-bearing.

**Preclinical Validation of the Three Marker Metabolites**—To test the usefulness of the three marker metabolites for human HCC diagnosis, especially discriminating small HCC from pre-cancer cirrhosis and chronic hepatitis, 412 serum samples from 262 patients with HCC, 76 patients with cirrhosis and 74 patients with chronic hepatitis B were analyzed using the Orbitrap MS, and the normalized contents of the three marker metabolites were calculated. For the 95 patients with small HCC (solitary nodules with a diameter of less than 2 cm), the three marker metabolites reached a sensitivity of 80.5% and a specificity of 80.1% (the cutoff value: 0.45, Fig. 6A), whereas the results of AFP (the cutoff value: 20 ng/ml) for these patients were 53 and 64%, respectively. Furthermore, the ROC curve analysis of the three marker metabolites yielded an AUC of 0.882, which was greater than that of AFP (0.648, Fig. 6B). However, the combination of metabolic markers and AFP achieved an AUC of 0.879.

When all HCC samples were enrolled, the sensitivity of the HCC diagnosis with three marker metabolites was 87.5%, and the specificity was 72.3% (the cutoff value: 0.65, Fig. 6C). When AFP was used as a biomarker for the same cohorts, 61.2% of the HCC group and 64% of the non-HCC group were correctly diagnosed (the cutoff value: 20 ng/ml). The ROC curve analysis revealed that the AUC of the three selected marker metabolites was 0.821, which was greater than that of AFP (0.678) (Fig. 6B).

The above results were calculated according to the optimized cut off values. Interestingly, there were also natural cut offs which were represented on Fig. 6. The natural cut offs are 0.35 for small HCC and 0.4 for all the patients. When these cutoffs were applied, a sensitivity of 92.7% and specificity of 59.7% could be achieved for the diagnosis of small HCC. And a sensitivity of 97.8% and specificity of 42.7% could be achieved for discrimination of all HCC patients.

When a cutoff value of 20 ng/ml AFP was used, 100 of 262 HCC patients were misdiagnosed. However, 91 of these 100 patients had a level of the metabolite markers higher than the
cutoff value (supplemental Fig. S3), implying the occurrence of HCC. On the other hand, for the 54 CLD patients with AFP above 20 ng/ml, only 16 cases showed positive for the “metabolite markers”. In addition, the use of the 3 marker metabolites plus AFP significantly increased the specificity and the diagnostic performance of the metabolic markers (AUC: 0.875, Fig. 6B).

DISCUSSION

Early or subclinical diagnosis of HCC would be helpful for the prevention and treatment. Because HCC patients with chronic liver diseases usually take a long time before HCC occurs, there should exist an opportunity to discover the early biomarkers of the occurrence and development of HCC. The nontarget metabolomics provides a global view of the organism and can be used to monitor the dynamic metabolic alterations that occur in different pathological processes (37).

The hypothesis of the present study was that a metabolomics study of hepatocarcinogenesis would provide early or even subclinical metabolic markers for HCC, the serum metabolic profiling analysis of DEN-induced rat models of HCC was the first step of this work. The tissue pathological results showed that the rat HCC model was successfully constructed. The serial progression of hepatocarcinogenesis could be found, including the inflammation stage, the cirrhosis stage, and the carcinoma stage.

The study was first focused on the systemic metabolic changes in the rat metabolome using a PLS-DA model. As shown in Fig. 1B, aging and carcinogenesis are major influence factors in model rats. Moreover, these two factors are parallel to the time axis. So the age-related metabolites were first removed using a statistical analysis (Mann-Whitney test) and the remaining 706 ions were also able to classify the stages of liver diseases (Fig. 2D), which provide possibility for the exploration of marker candidates. By using PLS-DA to 706 ions, 52 statistically important variables with VIP>2 were defined. The identified differential metabolites (supplemental Table S1) showed that many metabolic pathways including fatty acids, bile acids, amino acids, phospholipids etc. were influenced during carcinogenesis in the model rats. Among these metabolic alterations, the six clusters of metabolites (supplemental Table S1) represented the major metabolic disturbance.

Our results indicated that the deregulation of lipids metabolism is of great importance during the carcinogenesis procedure. Fatty acids and carnitines were included in the list of “most differential metabolites” which reflected the abnormal metabolism of lipids. Poly unsaturated fatty acids (PUFA) such as elapanodonic acid and docosahexaenoic acid, could play important roles in the modulation of cell proliferation through cellular lipids peroxidation (38). The deregulation of PUFA has been considered as the early events in the liver carcinogenesis (39). L-carnitine acts as a protector from liver carcinogenesis by decreasing free fatty acid in blood and lower the oxidative damage from lipids peroxidation (40, 41). Phospholipids take part in the metabolism of lipids. The balance of phospholipids constituting cellular membrane were kept by enzyme such as PE N-methyltransferase 2 (PEMT2), which catalyzes the conversion of PE to PC in the liver when dietary choline supply is inadequate. However, the expression of PEMT2 mRNA is reduced or absent in HCC and may cause the accumulation of PE or LPE, partly accounting for the high level of circulation PE or LPE (42). As evidence, the ratio of PE to PC in both blood and tissue elevated significantly, which indicated the low expression or activity in the model rats (data not shown). When clustered with the commonly used clinical liver injury indicator of aminotransferases (ALT, AST) and bilirubin, the reason for the increase of PE could partly be attributable to hepatocellular damage (supplemental Fig. S4).

Bile acids are synthesized from cholesterol in the liver, and cycle in gallbladder and intestine. The high levels of bile acids in the circulation would be useful references for hepatocellular damages. Liver injury in HCC and chronic liver diseases could also be reflected on the increase of blood bile acids (43). The deregulation of lipids metabolism, lipids peroxidation, and cellular damage were the major metabolic events in the model rats during carcinogenesis. The possibly relationship of differential metabolites and the hepatocarcinogenesis procedure was also given in supplemental Fig. S5 of the supplementary materials.

In the exploration of the disease mechanism, we hope to identify as many differential metabolites from a variety of pathways as possible. As a result, the strategy of a combination of several feature selection methods is required (28). However, for the diagnosis of diseases, redundant features are not necessary. In this study, although 52 ions were found to be of great value for the discrimination of hepatocarcinogenesis, additional methods were still necessary to narrow down the scope of the biomarker pool. With the HCA analysis 52 ions were clustered into six types. The results of SAM indicated that TCA and LPE 16:0 are the most significant metabolic features which are positively related to the model, and LPC 22:5 is the most significant metabolic feature of the carcinoma stage. These three metabolites, LPE 16:0, TCA, and LPC 22:5, referred to as the marker metabolites, were finally defined for their potential applications. Using the binary logistic regression method (33), the capacity of the above three marker metabolites to discriminate liver diseases in rats was evaluated. The results provided preliminary evidence for these metabolites to be used as markers for the classification of HCC and chronic liver diseases.

The model rats in the different disease stages were usually discriminated according to their age, which is not precise but is convenient to estimate the progression of HCC. However, the rats in the border stages exhibited variable pathological differences, depending on the individuals. The rats at week 10 suffered from inflammation or cirrhosis, and several rats at week 16 were tumor-bearing. Thus, the discrimination of the
rats at week 10 or week 16 was suggestive of the “early
diagnosis” of cirrhosis or carcinoma, respectively. As shown
in Figs. 5C and 5D, four rats at week 10 were classified as the
cirrhosis, and three rats at week 16 were classified as the
HCC. Although there is no direct pathological proof for
the resulting classifications, the above results are valuable for
guiding subsequent work aimed at discovering markers
for the warning or subclinical diagnosis of HCC.

To determine whether the three metabolic markers that
were identified using the rat models could be extended for
clinical HCC diagnosis, we performed a preliminary validation
using the sera from 412 patients with liver diseases, including
chronic hepatitis B, cirrhosis, and HCC. Because the diagno-
sis of small HCC is a challenge, more attention was taken to
it. Therefore, 95 cases of small HCC were first studied. Com-
pared with the traditional HCC biomarker AFP, the combina-
tional metabolic markers showed a better sensitivity. Simi-
larly, data from all 412 patients show three marker metabolites
have a better AUC value than the AFP (Fig. 6). It should be
pointed out that metabolites are regulated by diverse intrinsic
or extrinsic factors, the specificity of the metabolic markers is
still to be improved, the one of ways is the combinational use
of the AFP and the metabolic markers (Figs. 6B and 6D).

In summary, the results of the present metabolomics study
revealed the dynamic changes in hepatocarcinogenesis using the
DEN-induced rat model. However, to apply these results for
clinical use, the differences between rats and humans and
between chemical- and HBV-induced carcinogenesis must be
considered. Therefore, the validation using large scaled se-
rum samples from patients is very important. In the present
work, after acquiring the rat serum metabolic profile using
LC-MS, multivariate data analysis methods were utilized to
screen potential biomarkers. Of 52 differential metabolites,
three marker metabolites that provided the effective classifi-
cation of the disease stages of tumorigenesis were defined.
More importantly, these metabolites were shown to be effec-
tive in distinguishing patients with chronic hepatitis, cirrhosis
or HCC, especially those small HCCs. Further work is needed
for confirmation of these metabolites as early or subclinical
biomarkers. In addition, these metabolic features indicated
the deregulation of lipid metabolism during hepatocarcino-
genesis, which provides useful clues for future mechanism
exploration and identification of therapeutic targets of HCC.
The present study also highlights the ability of nontarget
metabolomics approaches to investigate the dynamic meta-
bolic alterations that occur during the complex biological
processes in carcinogenesis.

* This study has been supported by the State Key Science and
Technology Project for Infectious Diseases (2008ZX10002-017,
2008ZX10002-019) and China International Science and Technology
Cooperation Program (2009DAF141250) from State Ministry of Science
and Technology of China, and the key foundation (No. 20835006) and
the creative research group project (No.30921006, 21021004) from
National Natural Science Foundation of China.
high-resolution magic-angle spinning H-1 NMR spectroscopy in conjunction with multivariate data analysis. J. Proteome Res. 6, 2605–2614
21. Xue, R., Lin, Z., Deng, C., Dong, L., Liu, T., Wang, J., and Shen, X. (2008) A serum metabolomic investigation on hepatocellular carcinoma patients by chemical derivatization followed by gas chromatography/mass spectrometry. Rapid Commun. Mass Spectrom. 22, 3061–3068
22. Soga, T. S. M., Honma, M., Mori, M., Igarashi, K., Kashikura, K., Ikeda, S., Hirayama, A., Yamamoto, T., Yoshida, H., Otsuka, M., Tsuji, S., Yatomi, Y., Sakuragawa, T., Watanabe, H., Nihei, K., Saito, K., Kawata, S., Suzuki, H., Tomita, M., and Suematsu, M. (2011) Serum Metabolomics Reveals γ-Glutamyl Dipeptides as Biomarkers for Discrimination among Different Forms of Liver Disease. J. Hepatol. doi:10.1016/j.jhep.2011.01.031
23. Schiffer, E., Housset, C., Cachex, W., Wendum, D., Desbois-Mouthon, C., Rey, C., Clergue, F., Poupon, R., Barbu, V., and Rosmorduc, O. (2005) Gefitinib, an EGFR inhibitor, prevents hepatocellular carcinoma development in the rat liver with cirrhosis. Hepatology 41, 307–314
24. Qi, Y., Chen, X., Chan, C. Y., Li, D., Yu, F., Lin, M. C., Yew, D. T., Chen, Y., Aleksic, K., Lackner, C., Geigl, J. B., Schwarz, M., Auer, M., Ulz, P., Fischer, M., Trajanozki, Z., Otte, M., and Speicher, M. R. (2011) Evolution of genomic instability in diethylnitrosamine-induced rat hepatocellular carcinoma. Int. J. Cancer 122, 2862–2868
25. Aleksic, K., Lackner, C., Geigl, J. B., Schwarz, M., Auer, M., Ulz, P., Fischer, M., Trajanozki, Z., Otte, M., and Speicher, M. R. (2011) Evolution of genomic instability in diethylnitrosamine-induced hepatocarcinogenesis in mice. Hepatology 53, 895–904
26. Park, E. J., Lee, J. H., Yu, G. Y., He, G., Ali, S. R., Holzer, R. G., Osterreicher, C. H., Takahashi, H., and Karin, M. (2010) Dietary and genetic obesity promote liver inflammation and tumorigenesis by enhancing IL-6 and TNF expression. Cell 140, 197–208
27. Pitot, H. C., Dragan, Y. P., Teeguarden, J., Hsiia, S., and Campbell, H. (1996) Quantitation of multistage carcinogenesis in rat liver. Toxicol. Pathol. 24, 119–128
28. Lin, X. H., Wang, Q. C., Yin, P. Y., Tang, L., Tan, Y. X., Li, H., Yan, K., and Xu, G. (2011) A method for handling metabolomics data from liquid chromatography/mass spectrometry: combinational use of support vector machine recursive feature elimination, genetic algorithm and random forest for feature selection. Metabolomics DOI: 10.1007/s11306-011-0274-7
29. Blijses, S., Bobeldijk, I., Verheij, E. R., Ramaker, R., Kochhar, S., Macdonald, I. A., van Ommer, B., and Smilde, A. K. (2006) Large-scale human metabolomics studies: a strategy for data (pre-)processing and validation. Anal. Chem. 78, 567–574
30. Sangster, T., Major, H., Plumb, R., Wilson, A. J., and Wilson, I. D. (2006) A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabolic analysis. Analyst 131, 1075–1078
31. Gika, H. G., Theodoridis, G. A., Wingate, J. E., and Lundstedt, T. (2007) Chemometrics in metabolomics. J. Proteome Res. 6, 469–479
32. Trygg, J., Holmes, E., and Lundstedt, T. (2007) Chemometrics in metabolomics. J. Proteome Res. 6, 469–479
33. Xue, A., Scarlett, C. J., Chung, L., Butturini, G., Scarpa, A., Gandy, R., Wilson, S. R., Baxter, R. C., and Smith, R. C. (2010) Discovery of serum biomarkers for pancreatic adenocarcinoma using proteomic analysis. Br. J. Cancer 103, 391–400
34. Tisher, V. G., Tibshirani, R., and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. U.S.A. 98, 5116–5121
35. Chen, J., Zhao, X., Fritsche, J., Yin, P., Schmitt-Kopplin, P., Wang, W., Lu, X., Häring, H. U., Schleicher, E. D., Lehmann, R., and Xu, G. (2008) Practical approach for the identification and isomer elucidation of biomarkers detected in a metabolic study for the discovery of individuals at risk for diabetes by integrating the chromatographic and mass spectrometric information. Anal. Chem. 80, 1280–1289
36. Hu, C., van Dommelen, J., van der Heijden, R., Spijksma, G., Reimers, T. H., Wang, M., Sleen, E., Lu, X., Xu, G. W., van der Greef, J., and Hankemeier, T. (2006) RPLC-ion-trap-FTMS method for lipid profiling of plasma: method validation and application to p53 mutant mouse model. J. Proteome Res. 7, 4982–4991
37. Nicholson, J. K., and Lindon, J. C. (2008) Systems biology: Metabolomics. Nature 455, 1054–1056
38. Ji, C., Rouzer, C. A., Marnett, L. J., and Pietenpol, J. A. (1998) Induction of cell cycle arrest by the endogenous product of lipid peroxidation, malondialdehyde. Carcinogenesis 19, 1275–1283
39. Abel, S., Smuts, C. M., de Villiers, C., and Gelderblom, W. C. (2001) Changes in essential fatty acid patterns associated with normal liver regeneration and the progression of hepatocyte nodules in rat hepatocarcinogenesis. Carcinogenesis 22, 795–804
40. Furuno, T., Kanno, T., Arita, K., Asami, M., Utsumi, T., Doi, Y., Inoue, M., and Utsumi, K. (2001) Roles of long chain fatty acids and carnitine in mitochondrial permeability transition. Biochemical pharma- locy 62, 1037–1046
41. Al-Rejaie, S. S., Aleisa, A. M., Al-Yahya, A. A., Bakheet, S. A., Alsheikh, A., Fati, A. G., Al-Shabanah, O. A., and Sayed-Ahmed, M. M. (2009) Progression of diethylnitrosamine-induced hepatic carcinogenesis in carcinotine-depleted rats. World J. Gastroenterol. 15, 1373–1380
42. Tessitore, L., Marengo, B., Vance, D. E., Papotti, M., Mussa, A., Daidone, M. G., and Costa, A. (2003) Expression of phosphatidylethanolamine N-methyltransferase in human hepatocellular carcinomas. Oncology 65, 152–158
43. Pennington, C. R., Ross, P. E., and Bouchier, I. A. (1977) Serum bile acids in the diagnosis of hepatobiliary disease. Gut 18, 903–908

In order to cite this article properly, please include all of the following information: Tan, Y., Yin, P., Tang, L., King, W., Huang, Q., Cao, D., Zhao, X., Wang, W., Lu, X., Xu, Z., Wang, H., and Xu, G. (2012) Metabolomics Study of Stepwise Hepatocarcinogenesis From the Model Rats to Patients: Potential Biomarkers Effective for Small Hepatocellular Carcinoma Diagnosis. Mol. Cell. Proteomics 11(2):M111.010694. DOI: 10.1074/mcp.M111.010694.