Dear Editors,

I appreciate the opportunity offered to submit the revised manuscript in response to the editor's and reviewers' suggestions. I would also like to express our sincere thanks to the editor and reviewers. Their comments have helped tremendously for improving the quality of the manuscript.

The point-by-point responses to the reviewers' comments are provided on the following page. The resulting new text is highlighted in yellow in the revised manuscript and also indicated in our responses.

I hope that we have addressed in the revised manuscript the concerns raised by the reviewers and the manuscript in its revised form merits consideration of publication in *World Journal of Gastroenterology.*

Sincerely,

Xue-Gong Fan and Ning Li

Department of Infectious diseases & Key Laboratory of Viral Hepatitis of Hunan province, Xiangya Hospital, Central South University, Changsha, Hunan 410008, P.R.China.

Tel: +86-731-84327392

Fax: +86-731-84327392

Email: xgfan@hotmail.com; nxli@hotmail.com.

Step 5: Peer-review report(s)

The authors must resolve all issues in the manuscript based on peer-review report(s) and make a point-to point response to the issues raised in the peer-review report(s) which listed below:

Reviewer #1:

Scientific Quality: Grade C (Good)

Language Quality: Grade B (Minor language polishing)

Conclusion: Major revision

Specific Comments to Authors: In view of a lack of suitable biomarkers for the diagnosis of hepatocellular carcinoma (HCC), authors develop a pattern recognition method based on the analysis of new metabolites identified from serum metabolites of a series od HCC patients. Proteome analysis was performed using ultra-performance liquid chromatography-spectrometry (UPLC-MS). Authors found that concentration of most metabolites were lower in patients diagnosed with HCC; although the levels of hydroxypurine were higher in patient cases. Authors have developed a model based on metabolic data that would be ideal for the discovery of new biomarkers that could be applied to HCC diagnosis. General comments:

The main results obtained by the authors is the definition of two models based on the metabolomes of three group of patients. From the scientific point of view, these findings might have certain relevance,

however the level the evidence achieved is not enough to consider that the results described are robust enough for being applied in a clinical context. Authors should do an effort in trying to describe better the results and highlight the relevant results omitting those descriptions that might lead to non-relevant information.

Reply: We are so sorry for the confusion. We have made correction according to the reviewer's instruction. In addition, we have asked a native English speaker to check the English expression. We hope that the language is now acceptable for the next review process.

Authors should clarify how the models have been constructed: using the whole proteome or with those metabolites that are differentially expressed between groups?

Reply: We use the 72 identified metabolites to construct the LDA model and other multivariate statistical analysis model.

In any case, the number of variables (metabolites) included in constructing the models is very high. From my view, the models proposed are quite complexes and my feeling is that the sample size is too small to generate an accurate diagnostic tool (20 samples for the training set and 10 for the validation) and the risk of overfitting is presumibly very high....Hence, It would be necessary to have an independent series of cases in order to validate these findings. Authors

should justify that the sample size is sufficient to validate the models that they have generated.

Reply: Thank you very much for your comments. Admittedly, the sample size of this study is a little small due to the difficulties in collecting blood samples from patients with informed consent. As suggested by the reviewer, an independent series of cases would be the best choice to validate these findings. However, the batch effects resulting from LC/MS analysis such as the drift of retention time and the differences in the response status of the instruments make it almost impossible to compare metabolomics data generated in different batches. Therefore, we utilized permutation analysis to validate these findings. Concisely, the classification labels of training samples were randomly shuffled, and the model construction and validation procedures were repeated 20 times. As shown in Table 1, the performance of models constructed from training samples with shuffled labels were significantly worse than that of the real model. Such results indicated that the model constructed in this study is meaningful. We will expand the sample size in future targeted studies of the two metabolites included in the current LDA model to verify the model.

Table 1 Permutation test of the LAD model (the training samples)

acc	Sensitivity	Specificity	mcc

1	0.526316	0.777778	0.3	0.088192
2	0.368421	0.444444	0.3	-0.25844
3	0.157895	0	0.3	-0.72457
4	0.263158	0	0.5	-0.56695
5	0.210526	0	0.4	-0.6445
6	0.736842	1	0.5	0.566947
7	0.631579	0.888889	0.4	0.327569
8	0.421053	0.222222	0.6	-0.19096
9	0.210526	0.222222	0.2	-0.57778
10	0.842105	0.888889	0.8	0.688889
11	0.842105	1	0.7	0.724569
12	0.842105	1	0.7	0.724569
13	0.315789	0.111111	0.5	-0.41773
14	0.473684	0.666667	0.3	-0.03581
15	0.736842	1	0.5	0.566947
16	0.210526	0	0.4	-0.6445
17	0.789474	0.777778	0.8	0.577778
18	0.263158	0	0.5	-0.56695
19	0.210526	0.222222	0.2	-0.57778
20	0.315789	0.111111	0.5	-0.41773
21	0.473684	0.222222	0.7	-0.08819
22	0.736842	0.666667	0.8	0.47194

23	0.842105	1	0.7	0.724569
24	0.263158	0.333333	0.2	-0.47194
25	0.631579	0.888889	0.4	0.327569
26	0.157895	0	0.3	-0.72457
27	0.157895	0.111111	0.2	-0.68889
28	0.210526	0	0.4	-0.6445
29	0.210526	0	0.4	-0.6445
30	0.210526	0	0.4	-0.6445
31	0.789474	1	0.6	0.644503
32	0.526316	0.333333	0.7	0.035806
33	0.368421	0.444444	0.3	-0.25844
34	0.421053	0.222222	0.6	-0.19096
35	0.157895	0	0.3	-0.72457
36	0.842105	0.888889	0.8	0.688889
37	0.789474	1	0.6	0.644503
38	0.789474	0.777778	0.8	0.577778
39	0.631579	0.555556	0.7	0.258443
40	0.684211	0.888889	0.5	0.417734

Minor comments: -

Abstract: *Authors should include the number and type of analyzed patients as well as the type of samples.

Reply: We are so sorry for the confusion due to the too simplified statement. We have made correction according to the reviewer's comments as follows: Ultra-performance liquid chromatography-mass spectroscopy was used to characterize the serum metabolome of 30 hepatocellular carcinoma patients and 29 cirrhosis patients, and 31 health controls.

*Define AFP

Reply: We are so sorry for the inappropriate abbreviatio, we have used α-fetoprotein to instead of α-fetoprotein in the abstract.

*Check spelling errors *

Reply: Thank you very much for your suggestion. We have revised the whole manuscript carefully and tried to avoid any grammar or syntax error. In addition, we have asked a native English speaker to check the English. We hope that the language is now acceptable for the next review process.

The conclusion of the abstract is not in line with the aim of the study and it should be accordingly modified.

Reply: We are so sorry for the confusion. We have made corrections according to the reviewer's comments as follows: Hydroxypurine and proline might be novel biomarkers for hepatocellular carcinoma, and it could be diagnosed by the metabolomics model based on pattern recognition.

-Introduction: *Check references citations.

Reply: We are so sorry for the mistakes. We have revised the whole manuscript carefully and tried to avoid any reference errors.

Material and Methods:

*Patient and samples: was the study approved by an Ethical Committee? Did the patients sign an informed consent?

Reply: The Ethical statement and informed consent statement are listed at the end of the paper according to the journal’s requirements.

Institutional review board statement: The study was approved by the Ethics Committee of Xiangya Hospital, Central South University (Changsha, China). Informed consent statement: All patients gave

informed consent

Define the group of patients including the number of cases analyzed or belonging to each group;

Reply: We are so sorry for the confusion. We have made correction according to the reviewer's comments as follows: HCC patients (C group, N=30) all were with cirrhosis, and cirrhosis patients without HCC were Y group(N=29). The Child-Pugh Score of C group and Y group patients should be A or B. Healthy controls (N group, N=31) were chosen from the general population.

indicate the type of blood collection tube and the volume collected.

Reply: We are so sorry for the confusion. We have made correction according to the reviewer's comments as follows: Whole blood samples (3-5mL) were collected on an empty stomach in the morning in BD Vacutainer® blood specimen collection tubes (Weigao Group, Weihai, China). Whole blood samples were stored at 4℃ immediately after collection and were transported to the laboratory in <30 min

*Data processing and statistical analysis: Figure 1. Use the figure legit to detail the process of the data analysis and include the number of cases analyzed in each group.

Reply: We are so sorry for the confusion. We have made correction according to the reviewer's comments as follows: Figure 1. Road map of data analysis. Ordinary multivariate statistical analysis (PCA, PLS-DA, and OPLS-DA) were used to describe the meabolome of the three three group. Pattern recognition analysis based on sequential feature selection combined with LDA were used to diagnose HCC. Kruskal–Wallis test were used to identify difference metabolites.

-Results: *Figure 2: From my view this figure does not provide relevant information and I consider that it should be omitted in the main manuscript.

Reply: We have made correction according to the reviewer's comments, the Figure 2 has been omitted in the main manuscript and included in the Supplementary Material.

Reviewer #2:

Scientific Quality: Grade C (Good)

Language Quality: Grade B (Minor language polishing)

Conclusion: Major revision

Specific Comments to Authors: In this paper authors characterized the

serum metabolome of hepatocellular carcinoma to develop a new metabolomics diagnosis model and identify novel biomarkers useful for hepatocellular carcinoma screening. They based on pattern recognition method and ultra-performance liquid chromatography-mass spectroscopy to characterize the serum metabolome of patients with hepatocellular carcinoma and cirrhosis, followed by sequential feature selection combined with linear discriminant analysis to process the multivariate data. The paper is interesting, however there are some points that may be clarified before the publication.

1- Pag 3, line 9: why hydroxypurine and purine? The authors may explain the role of these molecules in cirrhosis.

Reply: We used sequential feature selection combined with linear discriminant analysis (LDA) to construct a prediction model for identifying HCC from cirrhosis. Sequential feature selection utilized in this study means that features (metabolites) with best prediction performance was sequentially selected into the LDA model until the best performance with the least features was achieved. At last, we found the LDA model, which included hydroxypurine and proline, has the best performance with only two features (metabolites).

The decrease of hydroxypurine is associated with the increase of xanthine oxidase in cirrhosis.

The decrease of proline is associated with the collagen deposition in cirrhosis, because proline is raw materials of collagen.

2- Table 2 of Significantly altered metabolites should be better described in the Results section. A scatter plot with Metabolite differences should be shown to facilitate the reading and to better explain the differences.

Reply: A scatter plot with Metabolite differences has been listed in the Supplementary materials.

3- The appropriate control in this analysis is patients with cirrhosis. Authors found that glutamic acid, kynurenic acid, vanillic acid, and hydroxypurine (Figure 5B) were higher in patients with HCC than in patients with cirrhosis. Why they show only hydroxypurine and purine in Figure 5? Also differences of these other molecules should be shown.

Reply: A scatter plot with Metabolite differences has been listed in the Supplementary materials.

4- In the pattern recognition analysis for diagnosis of HCC the dataset was randomly split into a training set of 20 HCC samples and 20 cirrhosis samples and a validation set of 10 HCC samples and nine cirrhosis samples. The number of training set samples and the number of validation samples is too little and should be increased.

Reply: Thank you very much for your comments. Admittedly, the sample size of this study is a little small due to the difficulties in collecting blood samples from patients with informed consent. As suggested by the reviewer, an independent series of cases would be the best choice to validate these findings. However, the batch effects resulting from LC/MS analysis such as the drift of retention time and the differences in the response status of the instruments make it almost impossible to compare metabolomics data generated in different batches. Therefore, we utilized permutation analysis to validate these findings. Concisely, the classification labels of training samples were randomly shuffled, and the model construction and validation procedures were repeated 20 times. As shown in Table 1, the performance of models constructed from training samples with shuffled labels were significantly worse than that of the real model. Such results indicated that the model constructed in this study is meaningful. We will expand the sample size in future targeted studies of the two metabolites included in the current LDA model to verify the model.

Table 1. Permutation test of the LAD model (the training samples)

	acc	Sensitivity	Specificity	mcc	
1		0.526316	0.777778	0.3	0.088192
2		0.368421	0.444444	0.3	-0.25844

3	0.157895	0	0.3	-0.72457
4	0.263158	0	0.5	-0.56695
5	0.210526	0	0.4	-0.6445
6	0.736842	1	0.5	0.566947
7	0.631579	0.888889	0.4	0.327569
8	0.421053	0.222222	0.6	-0.19096
9	0.210526	0.222222	0.2	-0.57778
10	0.842105	0.888889	0.8	0.688889
11	0.842105	1	0.7	0.724569
12	0.842105	1	0.7	0.724569
13	0.315789	0.111111	0.5	-0.41773
14	0.473684	0.666667	0.3	-0.03581
15	0.736842	1	0.5	0.566947
16	0.210526	0	0.4	-0.6445
17	0.789474	0.777778	0.8	0.577778
18	0.263158	0	0.5	-0.56695
19	0.210526	0.222222	0.2	-0.57778
20	0.315789	0.111111	0.5	-0.41773
21	0.473684	0.222222	0.7	-0.08819
22	0.736842	0.666667	0.8	0.47194
23	0.842105	1	0.7	0.724569
24	0.263158	0.333333	0.2	-0.47194

25	0.631579	0.888889	0.4	0.327569
26	0.157895	0	0.3	-0.72457
27	0.157895	0.111111	0.2	-0.68889
28	0.210526	0	0.4	-0.6445
29	0.210526	0	0.4	-0.6445
30	0.210526	0	0.4	-0.6445
31	0.789474	1	0.6	0.644503
32	0.526316	0.333333	0.7	0.035806
33	0.368421	0.444444	0.3	-0.25844
34	0.421053	0.222222	0.6	-0.19096
35	0.157895	0	0.3	-0.72457
36	0.842105	0.888889	0.8	0.688889
37	0.789474	1	0.6	0.644503
38	0.789474	0.777778	0.8	0.577778
39	0.631579	0.555556	0.7	0.258443
40	0.684211	0.888889	0.5	0.417734

5- What happens if the dataset is again split in a new training set and validation set? Will be obtained the same results?

Reply: We split the dataset randomly in to new training sets and validation sets for 40 times, the hydroxypurine and proline are the best performance metabolites combination in these new LDA model. The efficiency of these model is listed below.

	acc_test	Sensitivity test	Specificity _test	mcc_test	
1		0.789474	1	0.6	0.644503
2		0.789474	0.888889	0.7	0.595543
3		0.736842	0.888889	0.6	0.506048
4		0.789474	1	0.6	0.644503
5		0.842105	0.888889	0.8	0.688889
6		0.789474	0.888889	0.7	0.595543
7		0.842105	0.777778	0.9	0.685437
8		0.789474	0.888889	0.7	0.595543
9		0.684211	1	0.4	0.489898
10		0.842105	0.888889	0.8	0.688889
11		0.842105	0.888889	0.8	0.688889

6- I have the impression that in the LDA model the differences are due only to Hydroxypurine and Proline. What happens if Hydroxypurine and Proline are excluded?

Reply: If Hydroxypurine and Proline are excluded is the LDA model, the efficiency of LDA model decreased significantly. The most efficiency model are based on 6 metabolites (L-Tyrosine, Creatine, L-Glutamic acid,

3-indoleacrylic acid, L-Leucine, p-Cresyl sulfate”). The Model performance is illustrated as follows:

For training set : LOOCV accuracy: 0.925; sensitivity: 0.9; specificity: 0.95; ppv: 0.9473684; npv: 0.9047619; mcc: 0.8510645; Area under the curve: 0.925.

For test set: LOOCV accuracy: 0.5789474; sensitivity: 0.5555556; specificity:0.6; ppv: 0.5555556; npv:0.6; mcc:0.1555556; Area under the curve: 0.5778.

Pag 7, line 10: what scientist understand better about pathogenesis of HCC except of metabolic way? It is not clear.

Reply: In the famous review- Hallmarks of Cancer: The Next Generation (Cell. 2011;144(5):646-74.), written by Douglas Hanahan and Robert A. Weinberg. The ten hallmarks of cancer are Self-Sufficiency in Growth Signals, Insensitivity to Antigrowth Signals, Evading Apoptosis, Limitless Replicative Potential, Sustained Angiogenesis, Tissue Invasion and Metastasis, Avoiding Immune Destruction, Tumor Promotion Inflammation, Deregulating Cellular Energetics, Genome Instability and Mutation. Among them, the change of metabolic way (Deregulating

Cellular Energetics) could provide material and energy support for other features.

In the study, only the LDA test gave significant results about different metabolites between HCC and cirrhosis. It is enough to define the results of the paper?

Reply: At first, we intended to establish a PLS-DA model or OPLS-DA model with the aim of distinguishing patients with HCC from patients with cirrhosis. However, since the metabolomes of HCC and cirrhosis are not very different, the efficiency of the models was not robust enough to discriminate the two groups using ordinary PLS-DA or OPLS-DA models. Therefore, we used pattern recognition, an advance data processing method, to achieve our aim. This is not just a simple LDA model. we have performed the following works to optimize the LDA model and confirm the model efficiency. Firstly, sequential feature selection combined with LDA to search most suitable biomarkers. In other words, metabolites with best prediction performance were sequentially selected into the LDA model until the best performance was achieved. Secondly, we have utilized an external validation set and performed the permutation test to validate the model efficiency.

7- In figure of Histopathological examination (HE staining), authors should include the negative control.

Reply：We have added the liver biopsy HE figures of from a health control to the histopathological examination figure, and the figure has been move to Supplementary materials according to the comments of Reviewer #1.

10-In some sentences the English language need to be revised.

Reply: Thank you very much for your suggestion. We have revised the whole manuscript carefully and tried to avoid any grammar or syntax error. In addition, we have asked a native English speaker to check the English. We hope that the language is now acceptable for the next review process.

Step 6: Editorial Office's comments

5 Issues raised: (1) The authors did not provide the approved grant application form(s). Please upload the approved grant application form(s) or funding agency copy of any approval document(s);

Reply: We have uploaded the approval document(s) of funding agency.

(2) The authors did not provide original pictures. Please provide the original figure documents. Please prepare and arrange the figures using PowerPoint to ensure that all graphs or arrows or text portions can be reprocessed by the editor.

Reply: We have uploaded the original figure documents according to your comments.

(3) PMID and DOI numbers are missing in the reference list. Please provide the PubMed numbers and DOI citation numbers to the reference list and list all authors of the references. Please revise throughout;

Reply: We have added the PMID and DOI numbers to each reference.

(4) The “Article Highlights” section is missing.

Reply: We have added the “Article Highlights” to the manuscripts.