Author's response to reviews

Title: Serum diagnosis of diffuse large B-cell lymphomas and further identification of response to therapy using SELDI-TOF-MS and Tree Analysis Patterning

Authors:

Xing Zhang (xingzhangg@hotmail.com)
Bo Wang (wangb68377@sina.com)
Xiao-Shi Zhang (zxs617@163.com)
Zhi-Ming Li (wqjiang@yahoo.com)
Zhong-Zhen Guan (zzguan@21cn.com)
Wen-Qi Jiang (wqjiang@yahoo.com)

Version: 4 Date: 5 July 2007

Author's response to reviews: see over
Dear Dr. Edmunds,

Thank you for your kind decision and the reviewer’s comments. We have accordingly revised the manuscript. The revised version has been checked by a native English speaker from Elixigen Co., Ltd. Huntington Beach CA92648. Following is a point to point response to the reviewer’s comments.

**REVIEWER 1**

Comments and answers:

In Abstract and Introduction part, there are repeat mis-spelling term for SELDI-TOF-MS, the author use SELDI-TOP-MS, and this should be corrected as the SELDI-TOF-MS.

Answer: We have corrected the mistakes, and changed SELDI-TOP-MS to SELDI-TOF-MS in line 8 of page 2, line 13 of page 3, and line 15 of page 4. We are sorry for this mistake.

**REVIEWER 2**

Comments and answers:

Major Compulsory Revisions

Concerning methodology:
1. Reproducibility: the comparison of “10 selected M/Z peaks from another case study” is not enough explained in the 3.3 reproducibility and precision section. It seems that only single samples were analyzed, questioning about the quality of data obtained, which need at least duplicate experiments, randomly distributed in different chips in order to prevent methodology biases.

Answer: We did not describe it in detail in the original manuscript. All Samples including training set, test set and normal serum quality control (QC) sample were positioned randomly on the chips. To confirm the reproducibility of SELDI spectra in our study, we ran the pooled normal serum quality control (QC) sample in triplicate in the intra-assay analyses and on three different days in the inter-assay analyses. 10
selected M/Z peaks were randomly selected and compared to calculate the coefficient of variance. These mends were added in the results from line 17 to line 27 of page 10.

2. There is also question about the type of chips used in the study: in section 2.2 of material of methods, it is written at line 3 that the proteomics data set was generated with IMAC-3 chips, but at line 10 it appears that samples were loaded onto H4 chips and in tables 1 to 3 WCX2 chips are indicated. In section 3.1 of results, it is written that the WCX chips were the most discriminating for the construction of a decision tree. Which chips were really used? Does it means that Ciphergen Biomarker Wizard software analyzed data originating from different chips than Biomarker Pattern software? This point is confusing.

Answer: There were clerical errors, which WCX2 was written in IMAC-3 or H4 chips in original manuscript. We have corrected them in line 6, 13, and 19 of page 6 in the revised manuscript.

Concerning results:
1. No spectra showing the selected peaks are displayed, that would be important for the confidence into the results described.

Answer: We have added the representative spectra for the selected diagnostic peaks in the Supplementary figure 1 and 2 in the revised manuscript.

2. No protein identification of peaks selected was realized, weakening the study.

Answer: We believe it is important to identify the proteins or peptides of selected peaks. In the future study we will identify these peaks.

3. Three serum protein profile analysis were realized: controls against cancer, good prognosis against bad prognosis, early relapsing against late relapsing patients. 17 proteins of interest were detected, but only 2 (2954 and 4304 Da) were identified in 2 of the 3 analyses. This is questioning: why the peaks identified in bad response patients' sera are so weakly correlated with relapse markers? Why only one of the protein discriminating bad response or early relapse were retrieved in the analysis of cancer against healthy controls? This point is not even discussed in the text.

Answer: Peaks of interest in the diagnosis, prognosis and relapse were different. The reasons were due to the different purposes of the three different analyses, and the resources used by the diagnosis, prognosis and relapse analyses were different. Normal and tumor samples were compared in the diagnosis analysis. Only tumor samples were used for the other two analyses. The good prognosis samples and the
poor prognosis samples were compared in prognosis analyses, while the non-relapsed samples and the relapsed samples were compared in relapse analysis. Thus, the peaks identified in the three analyses may be different.

4. But the most important point is the following: why the protein peaks used for classification trees are different from the peaks selected in profile analyses? The mass value of tree nodes do not correspond to the mass values of the proteins listed in the text and in tables. For instance, in figure 1, the classification nodes peaks (cancer against healthy controls) are 5814, 3960, 14133, 5251, 4872 and 2503 Da. No one of these peaks is listed in the table 1 showing the significantly differently expressed peaks. This is the also true for the two other analyses. This seems to indicate that the discriminating peaks identified by Biomarker Wizard software are not the same than those identified by Biomarker Pattern software, questioning the relevance of the data obtained. Or that the profiles analyzed are not the same, questioning about the methodology employed in this study. Again no spectra is shown that may help to understand, and no identification was realized.

Answer: However, the peaks in biomarkers and tree pattern were not the same, which could be explained by that the biomarkers had been identified as a single peak which showed the highest discrepancy between two groups, while the patterns stressed synergistic effects of peaks in-group. Moreover, peaks in the pattern still showed significant difference in two groups but the degree of difference was less than those in biomarkers. We have added the representative spectra of the selected diagnostic peaks for the classification tree in the Supplementary figure 2 in the revised manuscript.

We believe it is important to examine the identifications of the selected peaks. In the future study we hope to identify the most important peaks.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
Concerning discussion: The results are not discussed, but rather the utility of mass spectrometry bioinformatic profile analysis is described (see above).

Answer: We had added the discussion about the reproducibility of SELDI technology on page 13 to page 15. We also added the discussion about the reasons why the peaks in predicting diagnosis, prognosis and relapse may be different in line 25 to line 29 of page 12.

2 studies should have been cited:

Answer: We have added these two studies in the reference on page 4.

REVIEWER 3

Comments and answers:
1. SELDI-TOF-MS analysis has been employed in several cases by now for the diagnosis/prognosis of various diseases in serum and urine. In all cases, high accuracy rates in disease detection were received. Nevertheless, the published studies also revealed that the technique faces various reproducibility problems and is prone to artifacts. There is lack of discussion on these issues in the manuscript. It is recommended that the authors discuss the reproducibility aspects of the SELDI profiling assay, its employment as a “biomarker discovery” versus “pattern analysis” tool, how reproducibility problems are being addressed in recent studies and also how the authors addressed some of these issues in the presented study.

Answer: We have added this part about reproducibility in discussion as suggested on page 13 to page 15.

2. The authors report the statistically “top-scored “peaks and also the peaks utilized for the generation of the decision tree. There is very limited overlap between the two ie, only 2 peaks found to differ significantly were employed by the algorithm for the generation of the decision tree. The authors should try to explain/discuss on this issue.

Answer: However, the peaks in biomarkers and tree pattern were not the same, which could be explained by that the biomarkers had been identified as a single peak which showed the highest discrepancy between two groups, while the patterns stressed synergistic effects of peaks in-group. Moreover, peaks in the pattern still showed significant difference in two groups but the degree of difference was less than those in biomarkers.

3. Were the samples positioned randomly on the chips? Randomization is very important when performing this type of analysis, as shown by multiple studies. This information is missing from “materials and methods".
Answer: Samples were positioned randomly on the chips. We have added this in the materials and methods from line 14 to 16 on page 6.

4. How many peaks (or peak clusters following the second pass peak selection) in total were detected?

Answer: To search for the serum proteins/polypeptides significantly different between DLBCL and normal patients, two hundred and ten peaks identified in the training set, three hundred and twenty one peaks were identified in the training set of discriminating good prognosis from poor prognosis patients, and three hundred and eleven peaks were identified in the training set of discriminating non-relapse and relapse patients. Some of these peaks were then used to construct the decision tree classification algorithm.

5. It is not very clear how the reproducibility study was conducted. The authors report that they “compared 10 selected M/Z peaks from an unaffected case study”. Do they mean that they run one control sample multiple (how many?) times? Also were all samples analyzed in duplicate?

Answer: We did not describe it in detail in the original manuscript. All Samples including training set, test set and normal serum quality control (QC) sample were positioned randomly on the chips. To confirm the reproducibility of SELDI spectra in our study, we ran the pooled normal serum quality control (QC) sample in triplicate in the intra-assay analyses and on three different days in the inter-assay analyses. 10 selected M/Z peaks were randomly selected and compared to calculate the coefficient of variance. These mends were added in the results from line 17 to line 27 of page 10.

Minor Essential Revisions

1. In “methods” the authors mention that they applied the sample on H4 chips and they describe the respective technique whereas in the “results” they report that they employed WCX. This discrepancy should be clarified.

Answer: There were clerical errors, which WCX2 was written in IMAC-3 or H4 chips in manuscript. We have corrected them in line 6, 13, and 19 of page 6 in the revised manuscript.

2. In the abstract the authors should mention that the accuracy rates they provide refer to the classification of the blinded test sets.

Answer: The test group is exactly the blinded sets in our study. The accuracy rates of the test set had been mentioned in the original manuscript.
3. A proteomic data set reported by Adam et al was used to classify the samples in this study.” The meaning of this sentence is not clear. Do the authors want to say that they employed the experimental protocol described in the study by Adam et al?

Answer: Yes, we followed the experimental protocol as reported by Adam et al.

We hope you will find the revised manuscript acceptable for publication in BMC Cancer.

Best regards.

Sincerely yours,

Wenqi Jiang, M.D.
Professor
Department of Medical Oncology, Cancer Center, Sun Yet-sen University,
651 DongFeng East Road, Guangzhou, P. R. China. 510060
Tel: 86(20)87343352
Fax: 86(20)87343392
E-mail: wqjiang@yahoo.com and wqjiang@mail.sysu.edu.cn