Reviewer's report

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

Version: Date: 28 April 2007

Reviewer: Laurent Mauvieux

Reviewer's report:

General

The paper is generally well written and the subject is of importance: detection of diffuse large cell B lymphoma (the most frequent type of lymphoma) and identification of good and bad responders to chemotherapy using SELDI-TOF technology in a series of 132 patients and 75 healthy controls sera. This is the first study using SELDI-TOF in diffuse large cell lymphoma, even if this technology was already used in different types of lymphoma (not cited in this manuscript). But many questions prevent in my opinion its acceptance in BMC cancer, that are explained below.

-------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Concerning methodology:

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.

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.

Concerning results:

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

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

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.

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.
In conclusion, in my opinion, there is too much question about methodology and relevance of the results described in this study for publication in BMC cancer.

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).

Discretionary Revisions (which the author can choose to ignore)

2 studies should have been cited:


What next?: Reject because scientifically unsound

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests