Reviewer’s report

Title: Distinct signatures of lung cancer types: aberrant mucin O-glycosylation and compromised immune response

Version: 0 Date: 24 Dec 2018

Reviewer: Andrei Zinovyev

Reviewer’s report:

The manuscript by Marta Lucchetta et al, "Different molecular signatures in lung cancer types from integrative bioinformatic analyses of RNASeq data" describes a bioinformatics analysis of several lung cancer transcriptomic datasets with an objective of finding signatures for discriminating LUAD and LUSC lung cancer types. The authors found specific and common oncogenes and tumor suppressor genes in two cancer types, reported on existence of modules of co-expressed genes, made an observation on the function of mucins in LUAD and a deficit of immune response in LUSC.

Overall, I find the undertaken study convincing and performed at professional level, with careful choice of state-of-the-art methodology. The reported results might be of interest for other studies on lung cancer.

I think the manuscript deserves to be published in BMC Cancer, provided the authors will take into account suggestions listed below:

1) In the current form, it is relatively difficult to follow the analysis pipeline in the manuscript partly because of the absence of a clear graphical flow-chart of the applied methods. I strongly suggest to introduce such a Figure in the very beginning of the manuscript. The authors use some jargon words from their analyses (such as names of the datasets 'paired' and 'all', which definition is difficult to find in the manuscript when it is read sequentially).

2) I advise modifying the title of the manuscript to make it more specific with respect to undertaken analysis and findings.

It is not completely clear why the bioinformatics analysis is called "integrative" (only one type of data was analysed, even if in several datasets).

3) The initial objective of the study (distinction between two lung cancer subtypes) need to be better justified as a goal.

Current justification does not sound very clear to me. Why taking into account gene co-expression is so important?
4) I think the authors should make an effort to clearly distinguish novel findings and confirmation of already known facts on lung cancer biology. Also, it would be nice to make a statement on which of the found co-expressed modules are specific to lung cancer and which are common for many cancer types.

5) Since one of the conclusions of the authors is the important difference between LUAD and LUSC immune landscapes, I think it is obligatory to characterize them using one of the state-of-the-art deconvolution approaches and by estimating the presence of different types of immune and other stromal cells. For example, important question to answer is what amount of differential expression between two cancer types comes from the differences in the composition of their TMEs?

6) Some of the Figures are not optimally designed to deliver their messages. "Network hairball" plots in Figure 4 are not informative (and make the PDF file very difficult to manage). Left heatmap in Figure 7 when printed does not allow to make any conclusion (probably, due to a suboptimal choice of the colormap).

Some minor comments:

1) In the abstract, it would be better to explicitly specify that the authors study inter-tumoral heterogeneity.

2) Section 2.1. "Samples outliers" -> "Outlier samples"

3) Section 2.2. Not clear what is meant by "lack of distribution assumption"

4) Section 2.2. "bath" -> "batch"

5) Section 2.4. How is the threshold "m>0.56" justified?

6) Section 2.7. What it was decided to use "Interlogous Interaction Database" as a source of PPIs and not some other standard choices?

7) Section 3.4. The term "moonlight" genes has to be explained with a corresponding reference, otherwise its meaning is not clear.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.
Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

Yes

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?
4. Have you received reimbursements, fees, funding, or salary from an organisation that holds or has applied for patents relating to the content of the manuscript?
5. Do you have any other financial competing interests?
6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report
including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal