Reviewer’s report

Title: Immunogenomic Pathways Associated with Cytotoxic Lymphocyte Infiltration and Survival in Colorectal Cancer.

Version: 0 Date: 23 Sep 2019

Reviewer: Deborah Neklason

Reviewer's report:

Immunogenomic Pathways Associated with Cytotoxic Lymphocyte Infiltration and Survival in Colorectal Cancer by Shen, Guan, Hummel et al.

Manuscript describing analysis of TCGA gene expression data from colorectal cancers to identify potential therapeutic targets that affect survival. Specifically targeting analysis of cytotoxic lymphocyte infiltration to better understand immune checkpoint inhibitor vulnerabilities.

Major concerns

1. Paper needs to be revised to not have jargon and make more broadly understandable to the reader. Problem areas include:
   a. Non-standard abbreviations should be spelled out throughout. It is very difficult to read with so many. Like CRC and TCGA (standard); MCP, RSC, DEGs and many more not standard.
   b. What is Level 3 mRNA data

2. It is not clear if there is a circular argument and/or the input and output are not independent. Could you discuss if the genes that make up differential expression are the same or similar genes that are used to define CL using the MCP Counter method? In particular, "Immunoregulatory interactions between lymphoid and non-lymphoid cell". Specifically, what genes defined the CL abundance and how do these compare to the genes differentially expressed in your analysis.

3. Related to #2, would be helpful to provide more detail about CL classification using MCP-Counter method. What goes into the classification, how different are tissues as far as CL presence? How powerful is this classification for the data set you are using - do tissues fall into clear groups. What is the median CL score and distribution for your data set?

4. Is the differential expression in right sided cancers (hi vs lo CL) due to the MSI tumors. Is the differential expression still present if the MSI tumors are excluded?

5. Could you expand upon survival analysis? Could you graphically represent gene expression vs survival? What do you consider significantly different survival? Possibly report statistics supporting the genes reported in Table 2.

6. How many genes are in the Immunoregulatory pathway (Table 2); what is the denominator?

7. Table 2, Pathways column is not helpful. If you had a companion model of CL interacting with non-lymphoid cell, might make some sense.

8. Shorter discussion - more to the point.

9. The CD48 discussion is confusing. Since it is differentially expressed in both metastatic groups - but opposite relationship to survival.
Other Suggestions
1. Need to better define early, localized, metastatic. There is overlap in staging. Are you using specific TNM and not stage? What is it?
2. Looks like you used RSEM normalized RNA-seq data, but talk about download of data off of Illumina HiSeq and Illumina-GA (raw data). Not sure how and when this downloaded other data is used. If not used, no need to mention.
3. Make sure there is sufficient detail in methods that someone could reproduce your analysis.

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

No

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.

No

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 recommend additional statistical review

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 organization 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