Reviewer’s report

Title: Unbiased data mining identifies cell cycle transcripts that predict non-indolent Gleason score 7 prostate cancer

Version: 0 Date: 11 Jul 2018

Reviewer: Amrita Singh

Reviewer's report:

Johnston and colleagues provide interesting data in this study identifying at least 13 transcripts in Gleason score 7 patients that may help predict risk of developing advanced disease. The authors use different data visualization and analysis tools to interrogate multiple publicly available prostate cancer datasets. A test set of biomarkers was created based on set criteria and then validated. 11 transcripts were identified that are upregulated and may be associated with disease recurrence and advancement to castrate resistant prostate cancer. Overall, this study attempts to answer an important question in the field of prostate cancer which is to identify biomarkers that are predictive of outcome and help in clinical decision making for patients who fall in the intermediate risk group. The manuscript is well written and the data are well presented.

Comments-

1) Is the upregulation of 11 transcripts unique to Gleason score 7 patients or are these genes differentially regulated in low risk patients also? Compare transcript abundance of the 13 transcripts of interest between Gleason score for low risk and intermediate risk.

2) Perform multivariate regression analysis using PSA levels and transcript abundance as variables.

3) Is there a correlation between the effect of treatment drugs administered, age related, race related differences in different Gleason score 7 patients for therapy and expression of 11 transcripts that are up regulated?

4) Since the 11 cell cycle and mitosis genes are associated with HES6-dependent E2F1 transcription factor mediated CRPC,(in this study E2F2). Do the patients with Gleason score 7 tested in this study develop CRPC? This information will support robustness of these 11 transcripts to be used as biomarker for prediction of outcome.

Minor comments-

1) The references need to be cited in the recommended style of journal.

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 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