Reviewer’s report

Title: Complex structural rearrangements are present in high-grade dysplastic Barrett's oesophagus samples

Version: 0 Date: 17 May 2018

Reviewer: Matthew Stachler

Reviewer's report:

The authors present whole genome sequencing data and analysis from a group of Barrett's esophagus samples and EAC; including two patients with non-progressive disease (with two samples each), 7 non-dysplastic BE samples, and 5 BE samples with HGD that are then compared to 22 previously sequenced EAC samples. The paper is well written and presents some interesting data, especially concerning the structural variants in the non-dysplastic samples, but is limited due to the small sample size. Overall, there are some minor suggestions/questions that may improve the manuscript before publication.

1: It is this reviewer's opinion that the paper would read easier if the samples were relabeled in a simpler fashion (?something like NP1-1, NP1-2, NDBE-1, HGD-1, EAC-1,...) so that in reading the text the reader immediately knows what type of sample is being talked about. Additionally, if there is room, a simplified version of additional file 1 may be valuable as a non-supplementary table.

2: "The BE samples were comprised of samples from 2 non-progressor patients (OESO_P1_NP1 and 196 OESO_P7_NP2) with BE that had not progressed to cancer over a number of years." The authors nicely show the time between the sequenced samples but it would give added confidence in their non-progressor status if the authors also state the total follow up and follow up time from last biopsy sequenced for each patient.

3: "Five of the non-dysplastic samples, and one of the dysplastic samples, were from patients with EAC and biopsies were taken from a region adjacent to, but well separated from, the tumour at the time of surgery." Did the other two ND samples come from patients with concurrent HGD? If not, what separated these samples/patients from the non-progressors?

4: For the paired non-progressing samples, was each sample taken in the same location of BE? It is known that a patient's field of BE may not clonal and if the samples were not taken in the same location, this may very well be contributing. Additionally, one sample per time point really isn't adequate to judge how stable the samples were over time. All of these possible complicating factors need to be communicated better in the discussion.

5: For the 22 EAC samples, were these samples reanalyzed or is previously published work just being restated?
6: The finding of no signature 17 in the 2 non-progressing patients is very interesting and the authors do a good job of discussing the limitations of a very small sample set for this analysis. What was the range of sig 17 in the non-dysplastic samples from progressors (IE were any of these 0 as well)?

7: "In contrast, Stachler and co-workers, using exome sequencing, found a significant difference in the number of SNV/indels between dysplastic and non-dysplastic samples [13]. Such differences observed between studies may be due to the difficulties of accurately identifying the dysplastic stage of BE by histopathology, as considerable inter-observer variability in the diagnosis of dysplasia in BE has been reported [28]." Were there any differences in the mutation rate/burden between non-dysplastic and dysplastic samples if only the coding region was looked at? Given that mutation rates between coding and non-coding regions of the genome can vary and have more or less chances of being functionally relevant, I wonder if the type of sequencing could play a role? Also, were subclonal mutations able to be detected? One could speculate that subclonal structure could be more complex within HGD samples so some of the HGD mutations could be at lower allele fractions making them easier to miss?

8: The authors state a frozen section was used to confirm the diagnosis. Did they also determine the overall % of cells the epithelial cells that made up the diagnosis were? IE was there sufficient "tumor" percentage to accurately call mutations and copy number changes in all samples?

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.

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.

Not relevant to this manuscript
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