Author’s response to reviews

Title: Targeted sequencing of circulating cell-free DNA in stage II-III resectable oesophageal squamous cell carcinoma patients

Authors:

Pei Meng (p.meng@umcg.nl)
Jiacong Wei (j.wei@umcg.nl)
Yiqun Geng (yqgeng@126.com)
Shaobin Chen (Chensb535176@hotmail.com)
Miente Terpstra (m.m.terpstra.cluster@gmail.com)
Qiongyi Huang (huang_qiong_yi@126.com)
Qian Zhang (qianzi4706@163.com)
Zuoqing Su (forever.szq@163.com)
Wanchun Yu (904923208@qq.com)
Min Su (minsu@stu.edu.cn)
Klaas Kok (k.kok@umcg.nl)
Anke van den Berg (a.van.den.berg01@umcg.nl)
Jiang Gu (2523381625@qq.com)

Version: 5 Date: 26 Jun 2019

Author’s response to reviews:

Point-by-point response to comments
Dear Editor:

We appreciate allowing us to resubmit an adapted version of our manuscript (BCAN-D-18-00953R4) entitled “Targeted sequencing of circulating cell free DNA in stage II-III resectable oesophageal squamous cell carcinoma patients”. We have addressed the comments made by the editor and the reviewer in a point-by-point response below and have carried out revisions of our manuscript to clarify these issues.

We hope that the reviewer will be satisfied with the revised manuscript and will not raise new points. We have tried to address all points raised by the reviewer in this and two previous revisions. We sincerely think that we tried to address all remarks in a clear way. We hope that our revised manuscript will now be acceptable for publication.

Yours sincerely,

Editorial comments:

1. Please include a statement in the Authors' contributions section to the effect that all authors have read and approved the manuscript, and ensure that this is the case.

Response: We have adapted the Authors' contributions statement accordingly.

Comments of the reviewer:

The reviewer has raised multiple points that are mostly related to one essential topic, e.g. the background mismatch rate, and the inclusion of cfDNA of healthy individuals as a means to determine the error level of our sequencing approach. Although on hindsight this might have been a good approach, our pipeline already gives a good indication of the mismatch error rate for all our samples individually (Additional file 3), irrespective of the presence of mutations, including those samples for which no variants were detected in the tumour itself. Following up on the comments of the reviewer, we decided to check these numbers again for our sequence data. See also the table below. The average mismatch rates give the average percentage of bases that is misaligned on every position. For each position, this is the sum of three separate mismatches (average mismatch rate is 0.43% per position and 0.14% per alternative nucleotide).
<table>
<thead>
<tr>
<th>Sample</th>
<th>N =</th>
<th>Av mismatch rate</th>
<th>Stdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>White blood cells</td>
<td>18</td>
<td>0.42%</td>
<td>0.049</td>
</tr>
<tr>
<td>Tumour samples</td>
<td>18</td>
<td>0.44%</td>
<td>0.037</td>
</tr>
<tr>
<td>Plasma pre-surgery</td>
<td>15</td>
<td>0.44%</td>
<td>0.040</td>
</tr>
<tr>
<td>Plasma post-surgery</td>
<td>18</td>
<td>0.41%</td>
<td>0.039</td>
</tr>
<tr>
<td>All samples</td>
<td>69</td>
<td>0.43%</td>
<td>0.043</td>
</tr>
</tbody>
</table>

As is shown in this table, there is little variation in the mismatch error rate between samples of different origin, analysed in the same experimental set up. This tells us, that adding healthy controls would not have given us more information on the mismatch rates than we already can deduce from our current data set.

As the sequencing error is partly determined by the specifics of the sequence run, it will not help us to do an additional sequencing run with new samples at this point in time.

ABSTRACT

- Line 57, "...all somatic mutations disappeared or decreased after surgery.": How can the authors be sure they disappeared rather than they were not detectible? I would change this statement in view that most patients post-surgery have very low levels of cfDNA; they used very low plasma volumes and had high yields of cfDNA.

Response: Thanks for your suggestion. We agree that we should have formulated this differently. we have adapted the indicated text in the abstract and also at other points in the manuscript stating that “mutations had a MAF below the detection limit of our assay”.

METHODS

- Line 56, "DNA extraction, library preparation and sequencing" section: The technical aspect of the paper would have been strengthened by sequencing cfDNA from unrelated plasma controls to ascertain the error level of their whole sequencing approach.

Response: We have discussed this issue extensively at the beginning of our rebuttal. For each patient, we included tumour DNA as positive control and DNA from white blood cells (WBC) as negative control for each patient. Moreover, in analysing cfDNA we focussed on those mutations that detected in tumour tissue and not in normal, and were detected at a MAF above the background level of sequencing errors as determined by the average mismatch rate.
- Lines 6-9, "Plasma cfDNA was extracted with the QIAamp circulating nucleic acid kit…": Although the authors state how much blood was collected, there is no indication of how much plasma they used for cfDNA extraction. That should be included as 2-5 ml of blood will yield a maximum of 1-2 ml of plasma which might become relevant later on.

Response: We now added the plasma volume in the text as proposed by the reviewer.

- Line 14, "Nanodrop": This is a bit crude for purity assessment given the limitations of spectrophotometry. For cfDNA this is not good and it would be interesting to have done so by BioAnalyzer instead, mostly to check the purity of the cfDNA and potential contamination by WBC released gDNA. This becomes relevant when you take together the low plasma volume they have obtained and the high yields of cfDNA that went into NGS (about 30…)

Response: We agree that measurement by nanodrop is no longer considered as the most optimal way to quantify the cfDNA yield. In our follow-up studies we have now switched to other methods for quantification of cfDNA yields. We have indicated the potential problem of lower ctDNA fraction in the cfDNA samples, especially after surgery in the discussion.

"Data analysis" section

- Lines 48-51, "This indicates the sequencing error rate at this specific site": Have you done several runs using health controls to ascertain this? I fear this might only reflect the error for this run and not systematic error associated with the approach.

Response: This has been addressed at the beginning of our rebuttal. The pipeline we used for variant calling in this study, has been validated extensively following quality guidelines for diagnostics laboratories and is in use in the diagnostic setting in the Department of Genetics (University Medical Centre Groningen). We observed that various projects that relied on hybridization-based target enrichment protocols and on the illumina platform for sequencing had comparable mismatch rates.
RESULTS

- Lines 40-46, "The median on-target coverage for the cfDNA samples was 613x (range 391x to 839x) pre-surgery and 752x (range 546x to 1932x) post-surgery": I am a bit concerned of the depth to which they sequence their cfDNA given the fact that the trend now is to sequence using error correction to a depth of 20000x minimum to be able to call mutations at AF of >0.3%. The becomes relevant when they claim mutations disappear, cannot it be they are not detecting them due to technical limitations? Authors should think about that alternative scenario.

Response: We considered the proposed alternative scenario with respect to differences in coverage and showed in figure 3 and additional files 3 and 4 that the post-surgery cfDNA had higher or similar coverage compared to the pre-surgery cfDNA. We did find the mutation in pre-surgery cfDNA, but did not find the mutation in post-surgery cfDNA. This indicates that the MAF is lower in post-surgery cfDNA. We agree that the disappearance is not correct in this context, and have adapted it in the manuscript.

DISCUSSION

- Lines 3-6, "Most of these mutations were not detected in cfDNA of blood samples obtained as early as 3-4 hours after surgery": It would be interesting for the authors to hypothesize why this is the case given that there is more cfDNA in post-surgery than pre-surgery (page 12, lines 48-51) and it might be expected a higher release of cfDNA following trauma (surgery) that could potentially mask cfDNA detection.

Response: Yes, we agree that the percentage of ctDNA in the cfDNA fraction might have been reduced due to surgery induced damage of cells, and as a consequence the release of cfDNA. We have now explained this in the manuscript (discussion section, second last paragraph, last 4 sentences).

- Lines 26-29, "To overcome this shortcoming we monitored sequencing error rates for all positions for which we identified mutations in cfDNA using our predefined criteria and showed that these sequencing error did not pass our criteria": Again, unless this has been tested extensively on healthy plasma this might just reflect the error for this specific run.

Response: On top of what we wrote above, sequence-specific error (SSE) are triggered by: (i) inverted repeats and (ii) GGC sequences for the Illumina sequencing platform [1]. The sequence error at different positions can be different and DNA quality may also influence sequencing error, so we think that checking the sequencing error at specific positions (additional file 5) and for each specific run on IGV is a good approach. In the clinic, molecular pathologist also check each reported mutation on IGV and check for errors by screening the other samples analysed in
the same run. As said, our pipeline has been validated and is in use in the diagnostic setting in the Department of Genetics (University Medical Centre Groningen). Sequencing errors to some extent is just normal statistics. Moreover, we have done extensive validations of NGS data by either Sanger sequencing or an independent NGS-based analysis in our previous studies. In both studies we observed a very high consistency with original NGS data and confirmed close to 100% of the mutations called with 4 or more reads in an independent validation [2, 3].

- Lines 51-53, "In our study, a stage IIA patient with a mean MAF of 2.4% in pre-surgery cfDNA developed a recurrence five months after surgery": Did this patient have mutations detected in cfDNA post-surgery? Authors should mention that.

Response: No, we did not find any mutations in post-surgery cfDNA in this patient. We now added the post-surgery cfDNA information in the manuscript as requested by the reviewer.

