Reviewer's report

Title: Exome sequencing of primary breast cancers with paired metastatic lesions reveals metastasis-enriched mutations in the A-kinase anchoring protein family (AKAPs)

Version: 0 Date: 29 Nov 2016

Reviewer: Maurizio Callari

Reviewer's report:

The manuscript by Kjallquist et al. start from the interesting question of studying the differences in the mutational landscape of primary breast cancers and matched metastasis, identifying the AKAP gene family as more frequently mutated in the metastatic lesions. Although the topic is of interest, there are multiple major shortcomings that in my opinion need to be addressed:

- It is not clear whether the samples are frozen or FFPE. Details about sample characteristics, handling and method used for DNA extraction should be added in the method section. Importantly, if the samples (or some of them) are FFPE, this could give an alternative explanation for the high C>T/G>A rate observed (supp methods), in particular in metastatic lesions. Of note, all but not one of the mutations in AKAP genes in the first cohort are C>T/G>A. Moreover, the number of somatic mutations is surprisingly high for many samples (~50 mutations per exome are expected in breast cancer), while the number of overlapping mutations between primary and metastatic lesions is surprisingly low (e.g. fig 1c and 1d, patient Pat059). It looks like there are still a lot of false positives in the mutations called. The authors should demonstrate that their main observations are not caused/affected by these types of artefacts.

- I find inappropriate that the two cohorts are analysed using totally different bioinformatic approaches. Minor changes could be needed because of the differences between the two experiments, however the same approach should be used. Of the two, the second sounds more convincing and in keeping with current standards.

- The manuscript is poorly organised. Most of the background is about AKAP gene family, useful info that, however, are appropriate for the discussion section. Previous work studying matched primary and metastatic lesions should be acknowledged in the background, for example the enrichment of mutations in ESR1 in metastatic lesions (after HT) has been reported several time. In the method section, subsection titles do not match the subsection content (e.g. description of somatic mutation calling is under "Exome sequencing cohort I" and not under "Detection of mutations from whole-exome sequencing data".
- The enrichment analysis that led to the focus on AKAP genes is not described and the global results are not reported. Was it the strongest association found? Which ontology/set of genes was tested for enrichment?

- Several results are reported as "significant" but no statistical analysis was performed.

- I am not sure what is the point of the gene expression analysis showing that some members of the AKAP family are differentially expressed in different breast cancer subtypes.

- English need improvement, for some sentences it was hard to understand the meaning.

- Quite a few samples have a coverage >30x in 50% of the target region or less. Did the authors took it into consideration in some way? How this can affect the results (false negatives)?

- Figure 1a, 1e, 2a, 2b, 2c could be tables.

**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.

No

**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 am able to assess the statistics
Quality of written English
Please indicate the quality of language in the manuscript:

Not suitable for publication unless extensively edited

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