Reviewer's report

Title: Estrogen independent gene expression defines clinically relevant subgroups of estrogen receptor positive breast cancer

Version: 2
Date: 1 September 2014
Reviewer: Jan Budczies

Reviewer's report:

Overview

Hallett and Hassell introduce a new gene expression based stratification of hormone receptor positive breast cancer with clinical implications. Their approach is similar to the one of Perou and Sorlie that defines the ubiquitous molecular subtypes of breast cancer. In 2000, Perou and Sorlie employed unsupervised clustering with respect to an intrinsic gene set defined by strong inter-tumor variation compared to intra-tumor variation before and after doxorubicin. Here, Hallett and Hassell employ unsupervised clustering with respect to an intrinsic gene set defined by strong inter-tumor variation compared to intra-tumor variation before and after letrozole.

Major Compulsory Revisions

1. The intrinsic gene set that is used for clustering and definition of the new subgroups is identified in a cohort of post-menopausal woman patients treated with an aromatase inhibitor (letrozole). The role of the patient population and the role of the treatment should be discussed. How would the intrinsic gene set and the clustering change for a population of pre-menopausal women and/or for woman treated with an anti-estrogen (e.g. tamoxifen)?

2. A gene expression data set of 262 ER+ tumors was used as training set and a single gene expression data set of 298 ER+ tumors was used as validation set. A large number of breast cancer gene expression data set is publicly available, including large series generated by consortia such as METABRIC and TCGA. Analysis of additional validation sets would strongly increase validity and impact of the current study.

3. Figure 4B: The expression of RAD50 and BARD1 should be shown and tested for significance separately for each of the six subgroups.

4. The expression of RAD50 and BARD1 in the cell lines under investigation should be shown.

5. As only three cell lines are used as control group, it is unclear if the better response to etoposide is a property of membership to cluster #1 or if the result is biased by special properties of the cell lines used as controls. It would be more valid to investigate all 24 cell lines for response to etoposide.
6. How does the response to etoposide depend on the concentration of the drug? How does the concentration used in the cell culture experiments relate to clinical relevant concentrations?

7. A highly relevant clinical question in ER+ breast cancer is identification of low risk patients that can be treated with surgery and subsequent adjuvant endocrine therapy without need of cytotoxic hemotherapy. Can the new subtyping help to identify low risk patients? How does it compare to molecular test such as Oncotype DX and Endopredict?

8. How are the subgroups defined here related to the ten subgroups that were suggested by the METABRIC group (Curtis et al, Nature 2012)?

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests.