Reviewer's report

Title: Breast Tumor Copy Number Aberration Phenotypes and Genomic Instability

Version: 1 Date: 21 March 2006

Reviewer: Charles Theillet

Reviewer's report:

General
This work presents the analysis of array-CGH data of 62 sporadic ductal carcinomas + 5 BRCA1 mutant tumors. Differences in genomic profiles could be distinguished when comparing ER+ and ER-, as well as P53wt and P53mt, however, the authors focused on their hierarchical clustering results. These showed the existence of 3 subgroups, distinguished by the type of CNC profiles; (1) low number of alterations (sometimes represented by 1q gain, 16p loss), (2) elevated mean fraction of CNC, which consisted of both gains and losses, but including few, if any, amplifications, (3) low levels of losses and gains associated to recurrent amplifications. Group 1 corresponded to low to medium grade and ER positive cancers, group 2 to a predominance of ER negative tumors, while group 3 was made of a mix of ER positive and negative cases. P53 mutations were evenly distributed among groups 2 and 3.

This paper presents an interesting analysis of array-CGH data on breast cancer. However, although data are valuable, the manuscript suffers from a number of conception problems, which make their interpretation difficult. Overall, this paper was felt as difficult to read and the data of complex interpretation, due to a rather muddled presentation. As a matter of fact, the information is delivered in fragmented portions, with some cores in the main text and some other, often essential to the general comprehension, in the Figure legends or in the supplementary data.

-------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

The authors propose in their conclusions that the amplification phenotype could be related to telomeric stress and to E2F driven gene expression profile. Although interesting, this suggestion can only be considered as a working hypothesis based on the data presented:

The telomere length can hardly be considered as conclusive.

The relation to E2F driven gene expression could not be granted, because these results have been extracted from a second dataset of 101 breast tumors, distinct from the one used to define the three clusters of genomic profiles. This dataset is presented as part of submitted work by one of the current coauthors (Chin) with no further information on how these tumors were analyzed and what the data looked like. It was felt, that either the data on these additional 101 tumors had to be included in the present paper or, if not possible, the whole argument on DNA damage and checkpoint and E2F driven expression taken out or presented as, what it looks to be, a working hypothesis.

The construction of the paper should be improved:
Most of the real data is presented in Figure 1 and supplementary data. It is felt that differences between ER+ and ER-, as well as those seen between p53wt and p53mt tumors should be included in the main core of the paper and the overlaps with the proposed subgroups based on the clustering analysis presented and commented. Furthermore, the difference the authors make between gains and amplification must be clearly presented in the text, as it is central to the definition of the...
clustering subgroups. It was, indeed, not obvious for this reader to class events shown in Figure 1C as “gains”. If these correspond to gains, then what do amplifications look like?

---

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

It was unclear to this reader how the data presented in Figure 2 contributed to the general argument in the paper.

Table 1 should be changed. The column presenting candidate genes in amplicons is incomplete and sometimes wrong. It should either be taken out or improved.

---

Discretionary Revisions (which the author can choose to ignore)

none

---

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes

**Declaration of competing interests:**

No competing interest