Reviewer's report

**Title:** Simultaneous copy number gains of NUPR1 and ERBB2 predicting poor prognosis in early-stage breast cancer

**Version:** 1  **Date:** 24 April 2012

**Reviewer:** Kylie Louise Gorringe

**Reviewer's report:**

1. Is the question posed by the authors well defined?
   Yes

2. Are the methods appropriate and well described?
   Yes.
   Minor revisions:
   - include description of how ER, PR and HER2 status defined
   - how was survival defined/determined – overall survival (death) or progression free survival?
   - define “UC” in Table 2

3. Are the data sound?
   Yes – minor concern that only >60% of tumour cells were required which may reduce sensitivity, however frequency of aberrations overall similar to previous studies.

   Major:
   Has mutiple testing correction been applied to comparisons between stage I and II cases? This may lead to no regions being significantly different. It should be stated if the regions do not pass multiple testing correction.

   Also, have regions been tested for copy number polymorphisms. In particular the high level amplifications on 1q and 16p are highly likely to be CNPs based on database of genomic variant information (DGV). This should be noted, and assessed for other high level amps/dels listed in supp data.

   Discretionary – in Additional data 3, would be helpful to have the cases with amps listed in case someone wanted to look at co-occurrences.

   Minor
   – in Table 2 how are “cancer-related genes” defined?
   - Figure 2 is not very clear. May be better as a heat map or removing CN-neutral cases
   - In Figure 3 the x-axis scale of the location on the right hand side is not even –
please adjust so that the graphs are scaled by Mb location. Also the x-axis label is “Proton” in A and “Portion” in B and C. Suggest this is changed to “Genomic position” for clarity.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?

- Would prefer raw data to be uploaded into GEO rather than normalized logratios on personal website.

5. Are the discussion and conclusions well balanced and adequately supported by the data?

   Major:
   - not convinced by additive effect of NUPR1 and ERBB2 on survival. Is it appropriate to combine these when multivariate analysis did not support an independent role for NUPR1?
   - study would be stronger if external data sets were used to validate findings. There are many of these now available, including for early-stage breast cancer:
     Thompson 2011 – early stage brca and recurrence
     Jonsson 2010
     Natrajan 2010
     Chin 2007
     Russnes 2010
     Chin 2006
     Curtis 2012

   The results should be compared to some/all of these studies as are appropriate. Given that several much larger studies have now been published, the authors could explore the population-specific aspect of their study to increase its relevance – are there any copy number changes that are more frequent in the Korean population than in those previously studied e.g. in UK, US, Scandinavia etc? e.g. Amplification of ERBB2 seems high (29%) compared to for example Chin 2007 who found 15.2% of cases to have amplification. This could be systematically explored e.g. by downloading datasets and undertaking an analysis in Nexus.

   - have the cases been similarly treated? Have any cases been treated with Herceptin for example?

6. Are limitations of the work clearly stated?

   Discretionary
   The authors mention missing data for ER, PR, HER2. HER2 could be inferred from copy number data. Could ER/PR be obtained for the missing FFPE samples by IHC? However, only 7 are missing this data in the validation set.
The authors should discuss the limitations in sample size that preclude a more sophisticated analysis taking into account well established breast cancer subtypes (lumA, lumB, basal etc) which have been shown to influence both survival and also the types of copy number changes observed (eg Thompson 2011, Jonsson 2010).

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
Limited acknowledgement of other breast cancer a-cgh studies

8. Do the title and abstract accurately convey what has been found?
I would dispute that the “picture of chromosomal alterations…not well studied” see list of studies above

9. Is the writing acceptable?
Yes, though should be read over carefully as a few minor errors present

**Level of interest:** An article of limited interest

**Quality of written English:** Needs some language corrections before being published

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

**Declaration of competing interests:**
I declare that I have no competing interests