Author's response to reviews

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

Authors:
Seung-Hyun Jung (hyun@catholic.ac.kr)
Ah-Won Lee (klee@catholic.ac.kr)
Seon-Hee Yim (lyra@catholic.ac.kr)
Hae-Jin Hu (haejin.hu@gmail.com)
Chungyoul Choi (chungychoe@gmail.com)
Yeun-Jun Chung (yejun@catholic.ac.kr)

Version: 2 Date: 3 June 2012

Author's response to reviews: see over
Dear Editor,

Regarding: MS 1968760341634343

Please consider the uploaded manuscript which has been revised for publication in *BMC Cancer*.

We have addressed all the points raised by the reviewers. All the changes are marked in blue font in the revised manuscript. Our revised manuscript was checked and corrected by a native English editor (E-WorldEditing Inc., USA). We also revised the supplementary materials.

Points by point replies to the reviewer’s comments are included in the ‘response to the reviewer’s comment’ below.

All authors listed in this manuscript have read and approved the content of the revised manuscript, new authorship and agreed to its submission.

Thank you very much for your helpful comments.

Yours Faithfully,

Yeun-Jun Chung M.D., Ph.D.
Integrated Research Center for Genome Polymorphism
The Catholic University of Korea, 505 Banpo-dong, Socho-gu
Seoul 137-701, Korea
Email: yejun@catholic.ac.kr
Phone : +82-2-2258-7343
FAX : +82-2-537-0572
Response to reviewer’s comments

Reviewer 1

Minor:
- include description of how ER, PR and HER2 status defined

We added the brief description in the Methods section of the revised manuscript (page 5).

In the original manuscript, the hormone receptor status for ER, PR and HER2 was obtained through a medical record and there were some missing information (7.2~31.3% of the study subjects). In this revision, we performed immunohistochemical staining for ER, PR and HER2 for all the cases without hormone receptor data and completed our data. Based on the complete hormone receptor information, we re-performed the down-stream analysis and could categorize our cases into the molecular subtypes.

- how was survival defined/determined – overall survival (death) or progression free survival?

All the survival rates in this study are overall survival. We added this description in the Methods section (page 8) and legends of Fig 4 and 5.

- define “UC” in Table 2

UC meant unclassified RARs, neither earlier nor later event. Definition of UC is described in the footnote of Table 2.

Major:
Has multiple 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.

We performed a FDR correction for the association analysis between CNA regions and clinical features listed in supplementary table 2 and also for the association analysis between RARs and clinical features. We added the Method and the Results in the revised manuscript (page 8-10) and presented the FDR P values to the supplementary table 2 and 4.

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.

We discussed this point in the Discussion section (page 15-16) and changed the Table 3 and supplementary table 3.

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

We added the information in the revised supplementary table 3

Minor
– in Table 2 how are “cancer-related genes” defined?
All the genes searched by the key word ‘tumor or cancer’ in the ‘NCBI Genes’ were defined as cancer-related genes. Among them, the genes located in the alteration regions are listed as cancer-related genes located in the specific region.

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.

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

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?

As the reviewer commented, NUPR1 was not validated in the replication. Although we did not mention in our original manuscript, we had explored the association rules between RAR markers and the death events using the CPAR (Classification based on Predictive Association Rules) algorithm. As a result, eight combination rules had been found to be associated with death events (Laplace accuracy score >0.75). All the eight rules had contained ‘RAR-G12 and RAR-G13 positives’. This result had driven us to explore the synergistic effect of these two genes. Regarding this, we added the association rule mining method in the Methods section (page 7-8) and Results in the ‘RARs associated with prognosis in EBC’ section (page 11) in the Result and the data was provided as supplementary table 6. We also discussed this point in the Discussion section (page 15-16).

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
Natrajran 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.

: According to the reviewer’s comment, we compared our data with recently reported breast cancer copy number analysis data listed above. We described this point in the discussion section (page 13, 15).

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

: All the EBCs examined in this study were collected during 1998 to 2002. At that time, there was no Herceptin treatment in Korea. All the patients who underwent breast conservation surgery were given adjuvant radiotherapy: Patients with pT>1 cm or axillary lymph node metastasis or hormone receptor negatives were given adjuvant chemotherapy with CMF (Cyclophosphamide, Methotrexate, 5FU) or FAC (5FU, Doxorubicin, cyclophosphamide) regimen. If necessary, hormone therapy was done with Tamoxifen or Toremifen.

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.

: As described above, we performed immunohistochemical staining for ER, PR and HER2 for all the cases with missing receptor information in this revision. We described the ER, PR, HER2 staining and molecular subtyping (Luminal A, Luminal B, HER2 and TNBC) in the Methods section (page 5). Table 1 and Table 4 were also changed according to the new hormone receptor information and subtypes.

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).

: Using the complete hormone receptor information, we grouped our cases into the molecular subtypes (Luminal A, Luminal B, HER2 and TNBC). We described the overall CNA frequency profiles for stage I and II groups, and among the molecular subtypes in the Result (page 9, supplementary figure 1). The distribution of the RARs by subtype was provided as supplementary table 5 and described in the Results section (page 10-11). We also examined the prognostic implications of the RARs in the four molecular subtypes but did not see any significant association. We discussed this limitation in the Discussion section (page 15-16).

I would dispute that the “picture of chromosomal alterations…not well studied” see list of studies above

: We changed the expression and cited relevant references (page 3).
Reviewer 2

1. Authors described that gDNA from a healthy female was used as normal reference for all array-CGH experiments. But, more information is needed.

: We gave a more detailed account about the universal reference DNA (page 5). Because we used a single individual as universal reference, we cannot rule out the possibility that some of the high-level CNAs identified in this study can be CNVs. We discussed this limitation in the Discussion section (page 15~16).

2. The authors identified that simultaneous gains of NUPR1 and ERBB2 can be a significant predictor of poor prognosis in EBCs. However, the NUPR1 gain was not replicated in the larger replication set. This result need to be discussed. What might be the reason of failure of replication? Is there any possible explanation for cross activity between NUPR1 and ERBB2 gene function?

: As the reviewer commented, NUPR1 was not validated in the replication. Although we did not mention in our original manuscript, we had explored the association rules between RAR markers and the death events using the CPAR (Classification based on Predictive Association Rules) algorithm. As a result, eight combination rules had been found to be associated with death events (Laplace accuracy score >0.75). All the eight rules had contained ‘RAR-G12 and RAR-G13 positives’. This result had driven us to explore the synergistic effect of these two genes. Regarding this, we added the association rule mining method in the Methods section (page 7-8) and results in the ‘RARs associated with prognosis in EBC’ section (page 11) in the Result and the data was provided as supplementary table 6. We also discussed this point in the Discussion section (page 15-16).

3. In discussion, authors stated that “Of the RARs identified in this study, 15 RARs were commonly detected in both stages I and II, which suggests that these copy number alterations are acquired at an earlier stage including precancerous stage.”. However, precancerous stage samples were not tested in this experiment. This description should be limited in stage I and II EBC.

: We agree to the reviewer’s point. We changed the description (page 13).