Author's response to reviews

Title: Identification of a robust gene signature that predicts breast cancer outcome in independent data sets

Authors:

James E. Korkola (korkolaj@mskcc.org)
Ekatarina Blaveri (eablaveri@gsk.com)
Sandy DeVries (devries@cc.ucsf.edu)
Dan H. Moore II (dmoore@cc.ucsf.edu)
E. SHELLEY Hwang (shelley hwang@ucsfmedctr.org)
Yunn-yi Chen (yunn-yi.chen@ucsf.edu)
Anne L.H. Estep (aestep@cc.ucsf.edu)
Karen L. Chew (kchew@cc.ucsf.edu)
Ron Jensen (rjensen@cc.ucsf.edu)
Frederic M. Waldman (waldman@cc.ucsf.edu)

Version: 4 Date: 20 October 2006

Author's response to reviews: see over
Dear Sir or Madam:

Please find enclosed a revised version of our manuscript entitled “Identification of a robust gene signature that predicts breast cancer outcome in independent datasets” by James E. Korkola et al., which we are resubmitting for consideration for publication in BMC Cancer (submission # 1345240735108392).

We apologize for the delays in revising the manuscript, and thank the editorial board for the extension for the revisions and their patience.

We wish to thank the reviewers for their thoughtful and insightful comments. We have made extensive changes to the manuscript based on the comments, as follows.

Reviewer 1 (Friend).

Major Issues

Reviewer Comment 1: The reviewer feels that the results of the 21 gene set classification are biased because we did not perform feature re-selection during each round of the leave-one-out classification in the training set.

Response: We agree that this was classification is biased, and have added a sentence to the results section to indicate that there is some bias in this result (page 11, last paragraph, bottom 2 lines). We have left the results with PAM and the correlation based cross-validation method as is, since both of these methods perform sampling at each stage of the cross-validation and are thus not biased. We have removed the classification rates for the SAM-derived gene set, since this set has the sampling bias (see page 11, 3rd paragraph). However, we have not changed the result of the 21 gene set. We have done this because re-sampling the genes at each stage of cross-validation would likely result in a slightly different set of 21 genes. We were interested in testing whether the set of genes (21 gene set) that were in common to all three methods of gene identification would have predictive ability in independent data sets. We have added a sentence (results, page 11, last paragraph, first line) explaining our rationale for doing this.

Reviewer Comment 2: The reviewer questions why the performance of the 21 gene set in the validation set is tested using a leave-one-out analysis rather than using the results of the predictive model that was built on the training set. Also, based on test Figure 4A appears to have too many samples.
Response: We have now excluded the training set. Our reason for making this change is because our validation set did not have the same endpoint as the training set (5 y DFS vs. 7 y DFS respectively). This was done because we did not have a sufficient number of specimens that had full 7 year followup. We felt this was a good approximation, since patients who have good 5 year outcomes should, in general, also have good 7 year outcomes. However, this was pointed out by both reviewers as an inconsistency. We have added a sentence to indicate that we did not have a sufficient number of specimens with long term follow-up to test the model directly (results, page 13, last paragraph, first line). We have also made the following changes in the Abstract, page 2: Minor changes, including deletion of the validation set from the materials and methods section; Deletion of validation set from results; addition of line indicating that the 21 gene set resulted in significant separation of patients based on survival in two independent data sets.

Figure 4A consists of the outcome prediction in the van’t Veer data set. This has been clarified in the text and in the figure legends.

Minor Issues

Reviewer minor issue 1: The reviewer questions on p.2 “21 genes in common to all three sets, had equal ability…” equal to what?

Response: We have changed this sentence (page 2, paragraph 1, line 6-7) to read, “21 genes in common to all three sets, also had the ability…”

Reviewer minor issue 2: Prognosis means prediction. Patients based on actual outcome should not be named as “Good Prognosis (GP) group” or “Poor Prognosis (PP) group”.

Response: We have changed this throughout the document to Good Outcome (GO) and Poor Outcome (PO) groups.

Reviewer minor issue 3: The training set had a 7 year selection cut, but the test set changed to 5 year. It seems to be inconsistent.

Response: As noted above, we have removed the validation set from the analysis.

Reviewer minor issue 4: Define log2(t/r) please.

Response: We have now defined log2(t/r) in the text (Materials and methods, page 7. Statistical analysis section, line 7) and listed it in the abbreviations (p.22).

Reviewer minor issue 4: p.11, bottom line: the “test set” seems to be “training set.”

Response: This was a mistake on our part and has been corrected (now on page 12, 2nd paragraph, 2nd line).
Reviewer 2 (Hu).

Reviewer Comment 1: In the “training” set, 7 years disease free survival was used as the cutoff for Good Prognosis and Poor Prognosis. In the “test” set, why was the cutoff changed to 5 years? This is not consistent.

Response: We have clarified this, as described above in major comment #2 and minor comment #3 from the first reviewer.

Reviewer Comment 2: page 6. Why were some samples purified using RNeasy? This inconsistent preparation of RNA samples for a data set of expression profile of tumors might affect abundance of some RNA species?

Response: We have added a line in the methods to explain why samples were purified with RNeasy and a second line to indicate that we were unable to see any changes in sample expression profiles based on RNeasy purification. See results, page 5 RNA Isolation section, line 5 and lines 7-8.

Reviewer Comment 3: page 10. What were the sample selection criteria? How does this criteria affect Kaplan-Meier analysis using ER status, stage, and grade?

Response: Samples were randomly selected, with the only criteria being the a tumor:normal cell ratio of greater than 70%. This is clarified in the methods section (p.5 paragraph 1, second sentence). Since they were randomly selected, there should be no effects on Kaplan-Meier analysis.

Reviewer Comment 4: page 11-line 5. What does multiple copies mean? If they are the same cDNA sequence of different spots on microarray to represent the same gene, collapse is a better statistic method instead of showing all copies in a gene set.

Response: Multiple copies refers to multiple clones representing the same gene on the array. We chose not to “collapse” them because the clones often are located in different parts of the gene. Thus, they represent different sequences of the same gene and as a result have measurements that are different. Furthermore, log(test/reference) ratios can differ dramatically for different clones for the same gene based on the presence of splice variants. Thus, we have left this result as is.

Reviewer Comment 5: page 18-line 5. Authors have proposed to use a common platform for expression profile of tumors. In reality, it is not an easy task to implement a common platform. However, a recent publication has largely addressed this issue using a statistic method “DWD” to merge datasets of breast tumors across different platforms (Hu Z. et al. BMC Genomics. 2006 Apr 27;7:96).

Response: We agree that in reality one common platform may not be easy to implement. We have added a line in the discussion to indicate that statistical and computational
methods exist to address this problem (page18 paragraph 1, last sentence), and have added the suggested reference.

Reviewer Comment 6: Finally, authors are encouraged to find out over-representation of GO categories of their 21 gene set in order to link outcome prediction with possible biological functions.

Response: We have added a section in the results to indicate the some of the GO categories of the 21 gene set (Results, page 13, Functional Annotation of Genes section).

We hope that these changes will be satisfactory for acceptance of the manuscript for publication. We look forward to hearing your response.

Sincerely,

F.M. Waldman, M.D., Ph.D.