Author's response to reviews

Title: Differences in gene expression in prostate cancer, normal appearing prostate tissue adjacent to cancer and prostate tissue from cancer free organ donors

Authors:

Uma R Chandran (chandran@pitt.edu)
Rajiv Dhir (dhirr@msx.upmc.edu)
Changqing Ma (chmst40@pitt.edu)
George Michalopoulos (michalopoulosgk@msx.upmc.edu)
Michael Becich (becichmj@msx.upmc.edu)
John Gilbertson (gilbertsonjr@msx.upmc.edu)

Version: 5 Date: 1 March 2005

Author's response to reviews: see over
The changes made to the manuscript are detailed below. The reviewer’s comments are italicized and next to it our response and changes are described.

Revisions suggested by Dr. Brian Haab:

1. In the first paragraph of the discussion, comparison to other gene expression studies is mentioned, that they are qualitatively similar. A few genes have appeared in many of the studies. Were those major genes found here also, such as Hepsin and PIM1, and others?

   We have added a sentence at the end of paragraph 1 of Discussion describing some of the genes found in common with other studies and added the reference 30.

2. The donor prostates were taken from the same zone as the tumor prostates. Added a line in the methods section.

3. We have added a sentence at the end of discussion, paragraph 2 stating that all steps of microarray analysis from tissue acquisition to sample hybridization need to be documented.

Revisions suggested by Dr. Remi Houlgatte

1. The normal and donor populations have a very different age distribution. 7% of tumour cases are under 50, whereas 60 % are under 40 for the normal donors. This point is discussed in a short paragraph of page 7, but no data are shown. This is an important issue, and some data should be provided to the reader. This could be obtained either by
   - Classification methods: Does the Donor above 40 closer to adjacent samples than those below?
   - ANOVA: What is the impact of age in variance analysis?
   - Prediction methods: If the genes differentially expressed between donor below 40 and adjacent tissue, are used, in which class, donor above 40 are predicted?

   Or another appropriate method ...

   We had previously performed SAM analysis of over 40 and under 40 differences and found no significant differences between these groups. Upon the reviewer’s suggestions we also performed classification analysis using PAM (Tibshirani et al, PNAS 99: 6567). For the classification, we trained the classifier using younger than 40 donors and adjacent normals and then tested the older donors. We also performed the reverse analysis by training on older than 40 donors and adjacent normals and testing the younger donors. In either case the error rates were similar for the test donor set, regardless of their age. The test donors classified with the training donors rather than with the adjacent normals. We have not included these figures, since we do not believe it adds to the study. However, if the reviewer or editors wish to see the data, we can certainly make it available.
2. During preparation of the manuscript, we had discussions with pathologists in the group about the effects of warm ischemia and other processing conditions on gene expression differences between tumors, adjacent normals and donors. Based on literature about expression changes related to warm ischemia, we expected that up regulation of genes such as fos/jun, may be observed in donors since they are likely subjected to ischemic and other conditions that affect early response genes. However, unexpectedly, we see up regulation of these genes in tumors and not donors. Since careful records of processing conditions are not usually maintained in most institutions, these observations are anecdotal. Therefore, we did not expand on this subject and but mention that further studies need to be carried out to address these issues. In fact, documentation and quality control assessment of all steps of microarray studies from tissue acquisition to hybridization are currently being carried in our group as part of standard operating procedure development for a cancer biomarker laboratory.

3. These controversial points (Major points 1 & 2) make necessary that the author make their data publicly available. This should be done according to MIAME guidelines.

This study was funded by the NCI’s “Molecular ReClassification of Prostate Cancer” grant; we are therefore required to submit the data to the Director’s Challenge database in a MIAME compliant form. The microarray data management tools for NCI, including the Director’s Challenge consortium (http://caarray.nci.nih.gov/) were release on 1/31/2005 and we are the first users of the system. Unfortunately, at this time, the user interface for the caArray system is unable to upload the large number of files produced by our study. We are working directly with the director and developers of caArray to deposit our data into the system. If after a reasonable length of time (a month or so), the software issues are not resolved, we will deposit the data into another MIAME compliant database such as GEO from NCBI (http://www.ncbi.nlm.nih.gov/geo/) or make the raw data files available on our group’s website. We are committed to open access to our microarray data.

4. Page 3 Method Section: The distribution of age is missing for adjacent normal samples

The ages have been added in the Methods section.

5. Page 4 bottom “Therefore, in the interests of clarity, we will focus on the MAS 5.0 results in the remainder this paper.” This sentence is not correct: figure 1 shows dCHIP plots.

Figures representing dCHIP (Figures 1D and 1E and 1F) have been removed.

6. Page 9 Paragraph 2: “Our donors did not have prostate cancer or PIN identified …” PIN ? Prostatic intraepithelial neoplasia ?

PIN has been changed to prostatic intraepithelial neoplasia.