Reviewer's report

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

Version: 4 Date: 11 February 2005

Reviewer: Brian Haab

Reviewer's report:

This paper examines the question of the choice of reference for gene array studies to determine abnormal expression in tumor tissue. This issue is important because of evidence that "normal" tissue taken from the same organ as a cancer, in the uninvolved areas, may have altered expression relative to tissue in which no cancer exists. This question is very difficult to answer since we can't routinely obtain tissue from the organs of healthy individuals – the closest we can get is from organ donors, which is the strategy employed here. The paper shows that the expression profiles of the normal tissue from donors is different from both the tumors and the matched normals, with greater differences in comparison to the tumor. The authors also examine the ontologies of the differential genes to look for clues to the sources of the differences.

The main potential confounding factor in this study, which is directly acknowledged and addressed in the paper, is that the normal tissue samples were collected under different conditions than the other samples. Thus, the differences observed in these samples may be due to sample handling, not biology. Also, the normal samples were from a different age range. The question of the influence of age range was addressed by comparing only samples that had matched ages. The question of whether the similarities between the tumor tissue and the matched normals were due to the fact that they were from the same patients was nicely addressed by splitit the samples into non-overlapping patients, and observing that range of correlations was similar to that when the whole sets were used.

So the fundamental question remained the influence of sample preparation. The paper makes at least two arguments to support the concept that the observed differences are real and not artifactual. One argument was that the tumors are more different from the donors than the matched normal, and another is that the gene ontologies comprising those differences more resemble cancer-related alterations when comparing tumor to donor. Both arguments are consistent with their hypothesis but of course not sufficient to prove it. But it looks like the authors have thoroughly considered this question and all aspects of the data analysis, and the supporting data referenced in the discussion very much bolster their arguments. Therefore I think the conclusions are valid and properly qualified, and they raise important considerations for this field.

Just a few points (discretionary revisions):

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?

Regarding sample collection, it was mentioned that all of the tumor tissues and matched normals were taken from the transition zone of the prostate. Were the donor samples also taken from the transition zone?

The authors could also include recommendations for reducing the experimental bias in the sample
collection protocols, since the only real way to remove suspicion of artifactual bias is to control for it experimentally. If you could change anything, what would it be? Is there any way to improve the sample collection process?

What next?: Accept after discretionary revisions

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

I declare that I have no competing interests’