Reviewer's report

Title: Meta analysis of papillary thyroid carcinoma microarray data and independent validation: New insights from old data

Version: 1 Date: 29 December 2010

Reviewer: amit aggarwal

Reviewer's report:

The article “Meta analysis of papillary thyroid carcinoma microarray data and independent validation: New insights from old data” by Klemens et al is a short study showing how to discover and develop biomarkers of interest using large scale genomics studies. They focused on papillary thyroid carcinoma (PTC), which has several datasets available but none that appears to have sufficient number of patients across multiple histological subtypes. They used four datasets and performed integrated analysis using a published approach (Distance Weighted Discrimination (DWD)) to remove dataset specific biases, followed by a PAMR algorithm based feature selection based on malignant vs benign comparison. SERPINA1 was found to be the most discriminative gene and they conducted an independent validation of SERPINA1, which is likely the strongest part of the paper. The section is well written and the sensitivity/specificity plots appear quite convincing. As noted by the authors, the discrimination extends to other non-malignant types (not part of training set), suggesting a broader applicability of this gene and assay. The next challenge (as stated in the last part of the discussion) is to setup a similar assay in FNA materials. An independent validation study in FNAB’s would have made more definitive contribution to the area.

Major Compulsory Revisions

Minor Essential Revisions

Discretionary Revisions

1) The sensitivity/specificity of SERPINA1 assay in FNAB’s is likely a critical for its usability in clinical setting. See prior published report in use of sham-FNA data such as Durand et al, Evaluation of Gene Expression Profiles in Thyroid Nodule Biopsy Material to Diagnose Thyroid Cancer, Journal of Clinical Endocrinology & Metabolism, 93(4) 1195-1202, 2008.

SERPINA1, shows up in PTC vs non-malignant in figure 3b, and it may be possible to have sufficient discriminative ability in FNAB’s to distinguish PTC from non-malignant types.

2) I personally feel that the claim on “new insights” is bit weak as SERPINA1 and several other genes have previously been described by other researchers as a marker of PTC. The ‘bioinformatically’ driven functional analysis of top differentially expressed genes has not been developed or addressed in sufficient depth or detail.
3) The qPCR Ct data can be published as a supplemental data section.

4) The R-codes and data can be made available on website, for others to reproduce the results.

Conclusion: Accept after minor and discretionary revisions

Level of interest: An article of limited interest

My opinion is that this article may be of some interest to researchers as it addresses a very specific area of PTC detection using SERPINA1. As noted by Klemens et al, FNAB is the most common mode of diagnosis, and this assay (in its current form) may have limited applicability in routine pathology setting.

This article is likely of interest to researchers who are looking at the research methodology on using integrative analysis methods from large scale platforms (with small sample sizes) to discover and develop biomarkers.

My opinion is that this article may be of some interest to researchers as it addresses a very specific area of PTC detection using SERPINA1. As noted by Klemens et al, FNAB is the most common mode of diagnosis, and this assay (in its current form) may have limited applicability in routine pathology setting.

This article is likely of interest to researchers who are looking at the research methodology on using integrative analysis methods from large scale platforms (with small sample sizes) to discover and develop biomarkers.

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

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests