Reviewer's report

Title: Molecular discrimination between follicular thyroid carcinoma and adenoma based on gene expression profiling in formalin-fixed paraffin-embedded samples

Version: 1 Date: 21 May 2013

Reviewer: Rehannah Borup

Reviewer's report:

Pfeifer et al. aim to derive a simple multi-gene mRNA-based classifier that would differentiate between FTC and FTA as well as preselect follicular tumors for increased risk of malignancy. The study comprises microarray and qRT-PCR analyses in addition to the use of online available microarray dataset.

Overall the study is relevant since Pfeifer et al. address the diagnostic dilemma of FTC by taking advantage of online available microarray data in combination with in house derived microarray data. The methods applied are in large well-established and the results presented in a comprehensible manner, although specific sections would benefit from revision.

Major Compulsory Revisions

1) One major issue is that the authors were unable to validate the defined 8-gene classifier in the FFPE samples which is the main purpose of the study, since only 5 of the 8 selected genes were amplified in the qPCR reaction. Furthermore it is unclear throughout the manuscript which classifier is the final, the 8-gene or the 5–gene classifier, since the two classifiers are referred to as the final interchangeably. This should be clarified and corrected.

2) Another issue is the small preoperative FNAB-based dataset used for validation purpose when Pfeifer et al. explicitly stresses that the approach for this study is different from other studies, in that the focus in this study is on the postoperative diagnostic dilemma (p.5, l.15-17). The paper could benefit from deleting the FNAB sections entirely to make it more “to the point”.

Minor Essential Revisions

Abstract

1) p.3, l.5-8: a rather disorganized sentence that have two aims, where in fact the study focus on a single aim; the differentiation of FTC and FTA, and not so much the preselecting of specimens with increased risk, which could easily be deleted. Moreover the sentence is missing the post-operative focus which should be included in the aim.

2) p.3, l.19-20: it is controversial if ITIH5 is significantly expressed in FFPE samples (according to Table 3).

3) p.3, l.23-25: The final conclusion does not support the results since the performance of the classifier is likely to be inferior to the performance of an experienced pathologist. The sentence should be rephrased.
Background

4) p.4, l.6: why has been?

5) p.5, l. 12-15: please verify if this paragraph is the opinion of Pfeifer et al. (in which case it does not belong under the “background” section) or in agreement with the paper by Li H, Robinson KA, Anton B, Saldanha IJ & Ladenson PW 2011 Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. Journal of Clinical Endocrinology and Metabolism 96 E1719–E1726, which could serve as a reference.

6) p.5, l. 24-26: should to be rewritten and backed up by references or deleted.

7) p.6, l. 2: delete additionally by qPCR analysis and replace with in frozen...

Methods

8) p.6, l. 4-14: before describing the array analysis, perhaps make a paragraph “Patients”, hereby separating the two entities. This may make the array paragraph more easily read.

9) p.6, l. 10-12: how many samples with concordant diagnosis of two pathologists?

10) p.6, l. 15: an independent set of samples.....would it be more correct to write the remaining samples were used as an independent set of.......?

11) p.7, l. 10-15: what is the purpose of including these samples? Consider leaving the FNAB section out of this study.


13) p.8, l. 2: 216 thyroid samples......please add of which XX originated from online available dataset...

14) p.8, l. 9-10: why not the most specific probe i.e. at_only probe sets versus _s or _f probe sets?

Results

15) p.8 l. 17: It is not clear whether the 5-gene classifier is the final classifier. This should be clarified.

16) P. 8 l. 25: insert: 5-gene classifier in “…validation of the 5-gene classifier was performed…

17) P. 9 l. 6. According to Figure 1, the 5-gene classifier and not the 8-gene classifier was validated on FNAB samples. Decide which classifier is deemed the final one and correct accordingly.

18) P. 10 l. 5-11: Several papers deal with the construction of classifiers with one or a small number of genes, for instance TSP (top-scoring pair(s) algorithms). In general these classifiers perform well if there are relative large differences in expression values between the genes in the two groups of samples that are classified, and some studies show that fold change effect the classification accuracy more than the number of significant genes included in the classifier. Furthermore, the stability of the genes selected across CV (cross validation loops) may be an important criterion for the evaluation of simple classifiers with...
small number of genes.
Since the 99 pre-selected genes have a log ratios above 1.5 with mean value above 5 in at least one of the groups in data set A, with corresponding log ratios levels in data set B, it is relevant to know the stability (percent occurrence) of the chosen genes within the CV loops as the number of genes is increased. The stability of the chosen genes for each CV loop may be relevant to include in the feature selection process, both for the formulation of classifiers with the pre-selection step and without the pre-selection as described in supplemental Information 1 page 3.

Discretionary Revisions

Abstract
1) p.3, l.12: FFPE has not been written out previously.
2) p.3, l. 17: other places in the manuscript the transcripts, named genes here, are written in italic. Please be consistent.

Background

Results
4) P.10 l. 6: “…, when over 20 genes were used”, change to: “…, when more than 20 genes were used within the preselected data set of 99 genes…”

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

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'