Reviewer's report

Title: Class Prediction Between Tumor and Normal Tissues Based on The Pair-wise Gene Expression Ratio

Version: 2 Date: 2 April 2004

Reviewer: Inge Jonassen

Reviewer's report:

General

------------------------------------------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

The presentation is not very clear. A number of aspects need to be more detailed described. At the same time the manuscript is too verbose and containing too many figures and tables. A major revision is needed and it would need to be reviewed again in light of more detailed description of critical aspects of the methodology applied.

Some critical aspects:

- it is not clear whether feature selection is performed only using the training data or on the full data set before the leave one out cross validation is applied using the features (genes or gene ratios) selected by analysing the full data set. See Ambroise and McLachlan PNAS 2002 for a discussion.

- it is not clear what method is used for class prediction

- it is claimed that the gene ratio approach alleviates the need for normalisation - the argument for this is not convincing. The method is validated on one data set only. This aspect should be addressed in a more careful manner.

- the manuscript fails to refer to and discuss relationship to other papers considering pairs of genes or combinations of genes in the supervised analysis of gene expression data. For example Bø and Jonassen (Genome Biology 2002) proposed a method to consider pairs of genes in combination effectively able to discover pairs of genes whose ratios is informative about (e.g.) clinical outcome (that approach is not limited to ratios)

- in the analysis of gene ratios, not all possible gene ratios are considered since this is claimed to be too computationally demanding. I can hardly see that this is the case as for example Bø and Jonassen (see above) consider all gene pairs explicitely.

- The language and presentation should be improved

- The reference list is in my view too long and includes many very remotely relevant papers.
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

The number of tables, figures, supplementary material is excessive and not justified by the limited contribution that the paper makes.

Discretionary Revisions (which the author can choose to ignore)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Not suitable for publication unless extensively edited

Statistical review: Yes

Declaration of competing interests:

none