Reviewer's report

Title: Gene expression meta-analysis identifies metastatic pathways and transcription factors in breast cancer

Version: 1 Date: 29 June 2008

Reviewer: Ole Christian Lingjaerde

Reviewer's report:

The authors propose to combine results from several GSEA or pathway analyses by considering the average rank of a gene set. This allows, for example, the identification of gene sets that are consistently ranked high on different data sets while not ranked high enough to be significant on any individual data set. My main criticism to the paper is that the properties of the method used to calculate the significance of a gene set in the meta-analysis are poorly described. This makes it hard to assess the importance of the results obtained. Also, the fact that the patients included in the study are quite diverse (in terms of treatment, lymph node status and measured response) deserves considerably more attention in the paper and requires convincing arguments in favor of interpreting the results as being related specifically to metastasis.

Major Compulsory Revisions
- What is the null hypothesis of their test of significance in the meta analysis? This hypothesis should be clearly stated and its relevance in the current context should be explained.
- What is the interpretation of the mean rank value, and what does it actually mean that a gene set is significant in the meta-analysis? In the GSEA analysis gene sets were ranked according to the normalized enrichment score, and the input to the meta-analysis are these ranked lists. However, the ranks thus obtained say nothing about the actual level of significance of gene lists according to GSEA. Indeed, on one particular data set one may find none or just a few significant gene sets, while on another one may find many significant gene sets. Thus, a gene set may e.g. be assigned a low rank value on one data set even though it is not significantly up-regulated, and at the same time be assigned a high rank value on another data set even though it is significantly
- The authors should explain why their proposed method of repeatedly drawing random ranking values and calculating the mean indeed simulates the distribution under the stated null hypothesis.

- The inclusion in their study of patients with quite diverse characteristics requires more thorough discussion and justification.

Minor Essential Revisions
There are many language errors (for example in the conclusion). These must be corrected prior to publication.

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

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

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**
I declare that I have no competing interests