Reviewer's report

Title: A comparison and evaluation of five biclustering algorithms by quantifying goodness of biclusters for gene expression data

Version: 1 Date: 7 March 2012

Reviewer: Lijing Xu

Reviewer's report:

This manuscript described the evaluation of five established biclustering algorithms using Gene Ontology annotation and protein-protein interaction. The authors proposed a Weighted Enrichment Score. They concluded that different algorithms performed differently on different datasets. They also pointed out that some algorithms, such as BIMAX and SAMBA, are less data-dependent, while others, such as ISA, are gene-dependent.

Major comments:
The whole Gene Ontology Weighted Enrichment Score section (on page 6, 7, 8) is poorly presented and has some fundamental problems. It is true that Li's method (using the best P-value of GO term as the P-value of one bicluster which includes multiple GO terms) is not appropriate, since the P-value of a bicluster obtained in this way is size-dependent. But this problem could be easily solved by normalizing with the number of genes which have GO terms within one bicluster. Instead, the authors proposed to assign more weights to GO terms that include more genes. The reason for such a weight assignment is not specified in the paper, and it seems hard to justify. The P-value of one GO term obtained by a hyper-geometric test already reflects the size effect. Why should one GO term including more genes in one bicluster receive more weight? To justify the weight assignment, the author should make a comparison with the simple normalization solution, i.e., calculating the enrichment score of one bicluster as the sum of the enrichment scores of all GO terms divided by the total number of genes that have GO terms in this bicluster.

Minor comments:
1. Page 4 and page 5 have redundant full names of the five algorithms.
2. The interpretation of P-value on page 7 is very confusing.
3. The results part needs to be rewritten.
4. The use of the past tense through the paper would be preferred.

Level of interest: An article of limited interest

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.