Reviewer's report

Title: Visualising associations between paired `omics' data sets

Version: 1 Date: 30 March 2012

Reviewer: William Bush

Reviewer's report:

This manuscript describes strategies and an R package for visualizing and interpreting paired -omics data, such as transcriptome and proteome datasets collected on the same subjects. This area is extremely important and underdeveloped, and the authors have developed a great tool.

Major Compulsory Revisions:

1. This manuscript could be greatly improved by focusing on a single example dataset (perhaps the Nutrimouse data) and walking through each of the visualizations (Correlation circle plots, Relevance Network, and Clustered Image Map). Currently, the demonstration of how the visualizations aid interpretation is masked somewhat by details of the individual datasets. The examples are compelling, but the text should describe in detail what scientific questions each of the analyses are trying to ask, and how the display of data in heatmap and network format aids the interpretation of the results from these analyses.

2. Does the method scale to larger sets of variables? Genomic and Transcriptomic data may contain many thousands of variables – please provide some sense of the scalability of the approach in the text.

3. The manuscript should be revised to more carefully tie the text to the figures -- for example, there appear to be missing panels from figure 1.

   Figure 2 would be easier to explain with a concrete example where readers have an intuitive sense about the correlation -- perhaps BMI and exercise or heart disease, etc.

   The text description of the correlations in figure 2 is a bit confusing because of the Yn nomenclature. I believe the explanation would be better by explicitly describing the relationships for Xb/Yb and Xc/Yc.

   The fatty acid clusters on figure 5 could be labeled on the figure itself.

   The authors describe the network topology for the Nutrimouse data, but this network is not presented as a figure?

   Please label each of the four subnetworks on figure 6 that you describe in the text.

Minor Essential Revisions:

1. In the Correlation circle plots section: “The use of such A graphical tool was then....”
2. In the Analysis Process section: "In the Nutrimouse data, IT cannot be assumed that..."

3. Please italicize the phrase "a priori" for easier reading.

Discretionary Comments:
For the relevance network, is it possible to somehow use the correlation structure in the layout of the graph? I realize graph layout engines are beyond the scope of the manuscript, but in my experience a major failing of graph representations is the tendency to create “hairballs” that are rather overwhelming... This could be alleviated somewhat by improving the layout of the graph itself.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

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