Reviewer’s report

Title: DNA microarray integromics analysis platform

Version: 1
Date: 25 January 2015

Reviewer: Zhang Chen

Reviewer’s report:

The manuscript described a web-based platform for the integrative analysis of multiple microarray data. The platform is easy to use and provides many useful functions for microarray data analysis and visualization. Several features such as data sharing within the platform and importing data directly from EBI ArrayExpress make it a very user-friendly online tool for microarray-based data analysis.

Major Compulsory Revisions

1. The Integromics platform is not fully accessible for anonymous users. I can only test the platform as a “demo user” and I’m not able to repeat the analysis in the manuscript due to limited access. Can the authors upload their analysis results from the two case studies into the platform and make them accessible to the demo user? Or can the authors provide a user ID so that I can fully access this platform?

2. The analysis methods such as canonical analysis and partial least squares regression are not available on the Integromics platform (no such options under “Analysis Setup” tab), which is different from what Figure 4 shows.

3. Several figures and tables are not quite informative and I would suggest the authors consider removing them from the manuscript.
   (1) For example, the figures showing the analysis panel (Figure 5, 7, 9, 11) are not quite informative to me. The authors can simply state the parameter settings in the main text.
   (2) Table 1 is not quite necessary because only two out of nine cells have meaningful contents. The authors have already stated in the main text that only two pairs of comparisons gave significant results. I would suggest removing this table and simply stating the significant miRNAs for the two comparisons in the main text.
   (3) Figure 15 delivered very limited information. I would suggest the authors rank the genes by their loading values and only plot the top 10 or 20 genes with their gene names for each component so that we can know the genes with large loading values.
   (4) I’m not sure what useful information I can get from Figure 20. It gives a distribution of TargetScore probabilities. I think figures that can answer the three
biological questions are more relevant (see comment 5)

4. It is confusing to me that why the authors only showed a section of the table in Figure 17 but full table in Figure 19.

5. In case study 2, the author asked three interesting biological questions. As a reader/user/researcher, I'm more interested in how to use this platform answer those questions. Therefore, I would suggest the authors give a more directly elaboration and demonstration on how they can answer these questions with this Integromics platform, instead of just giving the descriptions about each output file without directly answering these three questions.

Minor Essential Revisions

1. Figure 6: can the authors provide a legend explaining the size and the color of the dots in the figure?

2. Figure 12: what is adjCV? Can the authors provide some explanation about it?

3. In case study 2, the authors stated that they used “a False Discovery Rate (FDR) threshold of 0”. Is this a typo?

Discretionary Revisions

1. Several other types of microarray data are also commonly used in the research: for example, ChIP-chip data, DNA methylation microarray and protein microarray data. Can the Integromics platform support such data type?

2. I would suggest the authors shorten the Implementation section and make the manuscript more concise.

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

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