Reviewer’s report

Title: Systematically characterizing and prioritizing chemosensitivity related gene based on Gene Ontology and protein interaction network

Version: 1 Date: 18 April 2012

Reviewer: Yves Lussier

Reviewer’s report:

Major revisions or reject

Level of interest: An article of limited interest

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

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

Declaration of competing interests:

1. Is the question posed by the authors well defined?
   Yes the question is important: prediction of response to chemotherapy.

2. Are the methods appropriate and well described?

3. Are the data sound?
   yes

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
   This cannot be asserted as the authors failed to provide the table of curated CRGs.
   yes

5. Are the discussion and conclusions well balanced and adequately supported by the data?
   The prior work is well cited.

6. Are limitations of the work clearly stated?

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished? Yes for the chemo sensitivity genes, but “no” for the go analytics.

8. Do the title and abstract accurately convey what has been found?
   Yes, but the introduction to the abstract and the con
9. Is the writing acceptable? some sentences are a bit out of English or even incomprehensible (e.g. abstract “We documented 150 pairs [] from 492 LITTERATURES.”), while other sentences are a bit out of English. For example, the first sentence of the abstract comprises two ambiguities (capitalized): “It is IMPRESSIVE to predict chemotherapy RESPONSE BEFORE TREATMENT and to select alternative treatment regimens for chemotherapy-resistant patients.” The authors may prefer to use the terms “salient” or “judicious” or “valuable” instead of “impressive”, and perhaps the expression “predict response before initiating therapy” would better convey the meaning of the second ambiguous term.

Major revisions: Most importantly, the manuscript lacks methodological details to understand exactly how data is generated for the figures of the panels A,BC of Figure 1 as well as Figures 2,3, and 4. Further, the methods surrounding the GO ontology enrichment are particularly difficult to understand, provide no equations nor any references to pre-existing methods of GO prioritization.

1) The manuscript would benefit to be read by a native-English speaker or a professional English language writer/Editor. It is not publishable in the current form

2) The abstract should state the key results and p-value of the validation (predicting response to chemotherapy) in an independent dataset from the learning set or in a simulated dataset.

3) The conclusion is that the authors found correlation between drug activity and genes rather than response to therapy. Yet the first sentence of the abstract is misleading states that it is imperative to predict chemotherapeutic response, which is a different concept. The difference is significant in clinical care, many “active” drugs fail to create a clinically meaningful response. The authors must either restrict the statement to in vitro cellular chemosensitivity or add an additional sentence that brings the subject of the abstract back to in vitro cellular chemosensitivity before providing results. The third sentences of the abstract are superfluous, and instead the authors could mention their hypothesis. The first sentence of the results in the abstract does not provide a clear idea that the method employs

4) the majority of the letters and symbols used in Figure 1 are not documented in the figure’s legend, what does C1, C2 mean?

5) Figure 2 would benefit from a better definition in the legend of the vertical spaces corresponding to the numbers 80.6% and 91.9%

6) Figure 3 also need more definitions of the different thresholds, A,B,C,D, what does 21/62 mean (it is found in the main manuscript but not the figure legend – thus the figure legend is nearly useless to interpret the results)

7) The authors do not provide their annotated tables – only the ontologies about the tables. For reproducibility, the tables curated by the authors should be provided.