Author's response to reviews

Title: Research prioritization through prediction of future impact on biomedical science: a position paper on Inference-Analytics

Authors:

Madhavi K. Ganapathiraju (madhavi@pitt.edu)
Naoki Orii (norii@cs.cmu.edu)

Version: 3 **Date:** 10 July 2013

Author's response to reviews: see over

July 9, 2013

Dear Dr. Scott Edmunds,

Thank you again for the comments. We have completed the changes as suggested. Please see our responses below.

We are looking forward to seeing our paper in GigaScience journal!

Best Regards,
Madhavi Ganapathiraju

http://tonks.dbmi.pitt.edu/w

Editor's comments

1. GigaScience promotes more transparent and reproducible research, and we are not limited by space constraints in the text of the article or size of data we can host. In this experiment more methodological information could be included to make the protocol easier to understand, and including an additional figure as a ?flow diagram? going through the steps of the experiment could help to make the work much clearer.

We added two new figures to the manuscript to convey the concepts more clearly. (Figures 1 and 2). We added some text to state some aspects in Methods more clearly.

2. As we are able to host unlimited supporting information in our GigaDB repository (http://gigadb.org/), if there are useful datasets supporting or produced by this work we can integrate these into the paper, and these should be listed in an ?Availability of supporting data? section in the paper. We can include Binary biophysical PPIs you retrieved from HPRD and BioGRID for example.
If you have any code or scripts used in these protocols we can also help you copy them over to our GitHub repository (see:
https://github.com/gigascience/papers).

The interactions are downloaded from HPRD and BioGRID which have their own licenses; HPRD explicitly prohibits sharing of data with third party while it is unclear what BioGRID policy is. Hence, we are not able to share the data on the repository. We have however made available the results in their entirety as supplementary files with the manuscript. All the features are generated with existing python libraries (described in methods).

3. After the acknowledgments please add ?Competing Interests? and ?Authors? contributions? sections. For more see: http://www.gigasciencejournal.com/authors/instructions/research#formatting-competing

We added these sections.

Reviewer: Matthias Erwin Futschik

Reviewer's report: The authors addressed the concerns that I had in their response. It is not quite clear to me, why the additional analysis is not included in the main manuscripts or supplementary materials, but since their response will also be online, it is ok this way, too.

We feel that the additional analysis is more suitable in the reviewers comments and not included in the main manuscript for the following reason: the main point that we are driving in this paper is that there is a need for a new type of algorithms for *Inference Analytics* and when it comes to biology, one of the necessary inference analytics algorithms is that of *impact prediction*. Only to *demonstrate* the concept of impact prediction, we employed PPIs as a toy example. The datasets as-are-available today are not sufficient to develop a highly accurate model for impact prediction (as discussed in the manuscript). We are afraid that by discussing the impact prediction on a further reduced dataset (which yields poor performance) may shift the focus of the reader from *the need for impact prediction* to the results of PPI impact prediction.