

## Reviewer Report

**Title: Correcting for experiment-specific variability in expression compendia can remove underlying signals**

**Version: Original Submission**    **Date: 5/26/2020**

**Reviewer name: Hao Li**

### Reviewer Comments to Author:

AJ Lee et al. developed a neural network-based approach to generate gene expression compendia and then used the simulated compendia to test the influence of batch effects in rediscovering underlying signals. The authors reached a very interesting conclusion that when combining a large number of experiments, correcting for batches would harm the performance. This topic is of great interest to the readers of GigaScience. I only have a few minor comments:

1. The figures S1 and S2 were switched.
2. The authors should also cite PEER (PMID: 22343431), which is currently one of the most widely used package to correct for experimental noises.
3. It would be interesting if the authors could explore the influence of experiment numbers in the noise effects.
4. Much of the description regarding the analyses and discussions were included in Methods part, the authors might consider to move some of the text to Results and Discussion to facilitate the understanding of readers.

### Methods

Are the methods appropriate to the aims of the study, are they well described, and are necessary controls included? Choose an item.

### Conclusions

Are the conclusions adequately supported by the data shown? Choose an item.

### Reporting Standards

Does the manuscript adhere to the journal's guidelines on [minimum standards of reporting?](#) Choose an item.

Choose an item.

### Statistics

Are you able to assess all statistics in the manuscript, including the appropriateness of statistical tests used? Choose an item.

### **Quality of Written English**

Please indicate the quality of language in the manuscript: Choose an item.

### **Declaration of Competing Interests**

Please complete a declaration of competing interests, considering the following questions:

- Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
- Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
- Do you hold or are you currently applying for any patents relating to the content of the manuscript?
- Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?
- Do you have any other financial competing interests?
- Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (<http://creativecommons.org/licenses/by/4.0/>). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

Choose an item.

To further support our reviewers, we have joined with Publons, where you can gain additional credit to further highlight your hard work (see: <https://publons.com/journal/530/gigascience>). On publication of this paper, your review will be automatically added to Publons, you can then choose whether or not to claim your Publons credit. I understand this statement.

Yes Choose an item.