Reviewer’s report

Title: Metrics to Estimate Differential Co-Expression Networks

Version: 0 Date: 09 Aug 2017

Reviewer: Teresa Maria Creanza

Reviewer’s report:

In this manuscript, the authors analyze 6 different measures of differential co-expression to detect differences in gene expression data. They consider 4 metrics known in literature and 2 novel metrics and propose a methodology to generate controlled datasets from real data for evaluate the performance of differential co-expression measures. The methods are also compared on TCGA breast tumor data. Moreover, they provide a package in R to implement differential co-expression analyses.

The authors should correct fundamental language errors. For example, in the definition of the metrics 4, they have to correct: "All metrics above selects" in "All metrics above select" and there are many other errors of this type in the text. In other cases, the uncorrect language makes difficult to understand the meaning of the sentences, for example:" The generated dataset will maintain the internal correlation structure than experimental cancer data but absent of differential expression."

The methods are clearly described except minor concerns that I list in the following:

1) In the description of the 3 categories of methods, the authors define the third category as "differential gene co-expression". This definition is to be modified to show the centrality of the gene in these approaches. For example, they can call this class "genes with differential co-expression patterns".

2) When the authors report the metrics in the reference [15] in the background section, they do not clarify which is the exponential beta. The beta parameter is described in the methods section but the authors have to clarify the terms in the formula when they introduce it in the background section. Moreover, they have to clarify that the sum is on the index j and sqrt is for the squared root.

However, I have major concerns about the procedure to generate controlled datasets. The authors build covariance matrices in order to generate data of genes linked in a network. They did not check for the positive definiteness of the covariance matrices. A criterium to ensure positive definiteness is the belonging of the covariance matrices to the class of diagonally dominant matrices.
Moreover, it is not clear how the addition of Gaussian noise of standard deviation \(s/3\) for a fixed number of genes and \(s/10\) for a different number of genes (where \(s\) is the standard deviation of the cancer data set) can preserve the correlation structure of the cancer data set.

Finally, I am not convinced of the need to generate two datasets that do not have differentially expressed genes but show only differences in gene correlations. In gene expression data sets, it happens that there are genes that change their correlations with many genes and we can indicate these as core genes. It could happen that the core gene is not differentially expressed but the genes that change their correlations with it result differentially expressed. It is important to consider differentially expressed and differentially co-expressed core genes at the same time. The differentially expressed genes represent deregulated functions in the cells. The differentially co-expressed core genes could represent key cancer driver gene as a result of mutations or other cancer modifications.

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of limited interest

**Quality of written English**
Please indicate the quality of language in the manuscript:

Not suitable for publication unless extensively edited

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

1. 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?

2. 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?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. 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?

5. Do you have any other financial competing interests?

6. 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.

I agree to the open peer review policy of the journal

Do you want to get recognition for reviewing this manuscript? Add a record of this review to Publons to track and showcase your reviewing expertise across the world’s journals. Signing up is quick, easy and free!

No