July 25, 2014

Dear Drs. Montgomery and Gibson,

We are grateful for the opportunity to revise our manuscript entitled “Signature gene expression reveals novel clues on the molecular mechanisms of dimorphic transition in Penicillium marneffei”. We would like to thank reviewers for careful review of our manuscript and providing us with their comments and suggestion to improve the quality of the manuscript. The following responses have been prepared to address all of the reviewers’ comments in a point-by-point fashion.

We thank the reviewer for the constructive comments on the manuscript. We will detail in our response below how we plan to address the reviewer comments.

We thank all reviewers for their careful critique of our work. We are particularly appreciative of the suggestions for analyses and interpretations that they have put forward, and we are grateful to have been able to benefit from their expertise and insight. Through their comments, we feel that the paper is now a much stronger, more cohesive and clearer piece of work. Some suggestions have opened up new lines of exploration, and in this revised submission we have added several new components of work.

Briefly, these are:

* We have experimentally validated one of the variability markers for the blastocyst stage, HDDC2 using shRNA knockdowns in human iPSCs and a CRISPR Cas9-VP64 artificial transcriptional activator system in human ESCs.
* We have conducted a comprehensive comparison between using the SDC and the CV to measure expression variability for our analysis, including simulation studies to test the performance of these two statistics.
* We have assessed the functional impact of the stable genes identified by conducting a meta-analysis that used gene sets derived from studies on essentiality, ubiquitous expression, loss-of-variant genes, GWAS, haploinsufficiency, and human diseases.

We are grateful for the opportunity to improve our work so significantly through the comments of these reviewers.

**Associate Editor.**

*Comments:*

*1. While there is some value in providing integrated visual environments for scRNA-seq analysis, and MatLab reimplementations can be potentially useful for certain users, reviewers considered the features of the toolbox to be currently too limited to warrant publication. In particular, I concur with Reviewer 3 that the functionalities accessible through the GUI should be extended, to offer a interactive graphical environment for a wider range of use-cases. Moreover, the reimplementation of critical functions should be validated and shown to perform comparably (exactly or similarly in the case of stochastic methods) as previous implementation (within a supp mat). Also, Reviewer 2 mentioned comparable Matlab software, which should be mentioned and positioned against.*

While there is some value in providing integrated visual environments for scRNA-seq analysis, and MatLab reimplementations can be potentially useful for certain users, reviewers considered the features of the toolbox to be currently too limited to warrant publication. In particular, I concur with Reviewer 3 that the functionalities accessible through the GUI should be extended, to offer a interactive graphical environment for a wider range of use-cases. Moreover, the reimplementation of critical functions should be validated and shown to perform comparably (exactly or similarly in the case of stochastic methods) as previous implementation (within a supp mat). Also, Reviewer 2 mentioned comparable Matlab software, which should be mentioned and positioned against.

**Reviewer #1.**

*Comments:*

*1. Since this is a single-author paper, it should be "I" rather than "We" throughout the text.*

This has been changed.

*2. Please add a little more information to explain Figure 1.*

More information has been added to the legend text of Figure 1.

**Reviewer #2.**

*Major comments:*

*1. This note briefly describes a toolbox for analysis of scRNAseq data in MATLAB. A key features of this toolbox is its GUI, however the algorithms do not seem to be original, primarily wrappers for existing code in R. Even without original algorithms, the toolbox may be useful for some readers. However if it contains original algorithms, that should be mentioned as it would increase the paper's impact.*

More information has been added to the legend text of Figure 1.

*Also note that there is other MATLAB code for transcriptomic analysis already in existence, forexample at https://urldefense.proofpoint.com/v2/url?u=https-3A\_\_github.com\_cortex-2Dlab\_Transcriptomics&d=DwIFaQ&c=u6LDEWzohnDQ01ySGnxMzg&r=dzRP0h5ZWyh3FOHMTgCOAg&m=vJlip4uybg8wkXmMgaF6rMAGsLpcsgPpTr0MR5nI4ck&s=q7Q\_W8hcjxk1NN9aSPDhPIDGIjjKxcT55N3uaw9WPp8&e= ,*

SCell (Diaz, et al., 2016) and SCUBA (Marco, et al., 2014), We thank the reviewer for the very constructive suggestion, and for drawing our attention to the Guo et al. paper. This proved to be a terrific resource and we were able to verify our key findings with this data. We have included a supplemental text (Text S8) which details the results we obtained by applying the same approach used on the human embryo data set. As the reviewer notes, the Guo data set includes a smaller number of transcripts and we noticed that ourcomparison of specific genes that were stable in both species

**Reviewer #3.**

*General comments:*

*1. Single-cell RNA sequencing (scRNA-seq) offers gene expression measurements at single-cell resolution and makes it possible to study molecular mechanisms at the single-cell level. A large number of methods for the analysis of scRNA-seq data have been developed (mainly in R or Python). Here, the authors developed a Matlab toolbox for scRNA-seq data analysis. This toolbox includes many categories of functions that can be used for scRNA-seq data analysis, such as gene filtering, data normalization and visualization. The authors implemented a GUI application for facilitating users to perform scRNA-seq data analysis. However, only a small subset of functions included in the tool box can be accessed through GUI. Therefore, many of the functions (such as SC3 for clustering) still need to be run from command line, thus not providing any additional convenience compared to directly using the native tool (say the SC3 R package). In summary, the contribution of this toolbox to the field is very limited. My major concerns are detailed in the following.*

*Major comments:*

*1. As described in Fig. 1, only six types of functions (filter, normalization, batch correction, imputation, feature selection and visualization) can be accessed through GUI, with many of the functions included in the toolbox such as cell clustering, trajectory analysis and network construction are not implemented in the GUI. This makes the value of this work very limited. Taking SC3 as an example, I believe that it would be difficult to persuade users to use the re-implemented Matlab version of SC3 instead of the native SC3 R package.*

More information has been added to the legend text of Figure 1.

*2. I would suggest to add more functions to aid data analysis. For example, the authors may consider adding functions to compare the data before and after the removal of batch effect, which may help users to investigate whether batch effects are removed or not.*

More information has been added to the legend text of Figure 1.

*3. The authors are suggested to move the description of the algorithm of each functional category In the Results section to the Method section.*

More information has been added to the legend text of Figure 1.

*4. I also suggest to add more demo examples/scripts to illustrate how to use the toolbox.*

More information has been added to the legend text of Figure 1.

*5. As some algorithms such as the “sc\_sc3.m” is re-implemented in Matlab. The authors need to prove that the re-implemented version can reproduce the results obtained by the native one.*

More information has been added to the legend text of Figure 1.

*6. In addition to the commonly used analysis tools, the most recent algorithms for each category of functions may be added. For example, the authors re-implemented two feature selection methods but both of them were proposed before 2016.*

More information has been added to the legend text of Figure 1.

*7. Figure 1 is of poor quality. The screen shots of the GUI are obscure. In addition, it is necessary to add figure captions to describe each sub-figure..*

We added a new Supplementary Table to include the specific information of RMD values in each of these four groups, including more detailed statistics for these distributions.

Again, we appreciate all your valuable comments. We worked hard to be responsive to them. We have uploaded two versions of the updated manuscript to the system: a version of our manuscript with all edits incorporated, and a version of our manuscript with all edits visible and all comments visible (ms\_visual\_edit.pdf). We recommend that reviewers open the second document and read through this version to identify the place where we addressed the corresponding comments. Thank you for taking the time and energy to help us improve the paper.

Sincerely,

James Cai, Ph.D.

Associate Professor

Department of Veterinary Integrative Biosciences

Texas A&M University

4458 TAMU

College Station, TX 77843-4458

Tel (979) 458-5482 | Fax (979) 847-8981

jcai@tamu.edu

www.genomezoo.net