Author’s response to reviews

Title: Estimation of immune cell content in tumor using single-cell RNA-seq reference data

Authors:
Xiaoqing Yu (Xiaoqing.yu@moffitt.org)
Y. Chen (Ann.Chen@moffitt.org)
Jose Conejo-Garcia (Jose.Conejo-Garcia@moffitt.org)
Christine Chung (Christine.Chung@moffitt.org)
Xuefeng Wang (xuefeng.wang@moffitt.org)

Version: 1 Date: 01 Jul 2019

Author’s response to reviews:

June 27, 2019
RE: BCAN-D-19-01115

Title: Estimation of immune cell content in tumor using single-cell RNA-seq reference data

Dear Dr. Arora,

We would like to thank you and the reviewers for the very positive comments. We are very glad to know that our manuscript is potentially acceptable for publications. Our revision was delayed by the summer vacations but fortunately we were able to make it before the due day.

Below, please find our point-by-point response to the comments. We look forward to your favorable response and thanks again for your time and consideration!

Note: a clean copy of our revised manuscript is now uploaded upon request of Editorial Office.

Sincerely yours,

Xuefeng Wang, PhD
Reviewer reports:

Guy Adami (Reviewer 1): Estimation of immune cell content in tumor using single-cell RNA-seq reference data by Yu et al is an extremely well written paper that covers a method to use reference GEP matrices from SC RNAseq data of individual cell types, to deconvolute bulk tumor gene expression data. By combining t-SNE and ssGSEA the authors enhanced methods for differentiating T cell types in HNSCC samples.

Most of the analyses here are with simulated tumors not bulk tumor tissue. With scRNAseq and RNAseq of sorted cells there is a similar lag in time between enzymatic separation and isolation for single cells, cell sorting and placement in an RNA preservative. This at least in theory would make bulk tumor RNA levels for some genes different than that in sorted cells. Of course the gene lists used in the study to separate cell types, may already accommodate this problem.

Author’s Response: We thank the reviewer for the very positive comments!

Figure S 10 is very important to show the method works with bulk TCGA tumors. If there is some point for the survival curve, S11 it should be made as it is already well know that oral pharynx tumors with HPV are associated with longer survival.

Author’s Response: Thanks for the suggestion. As suggested, in Figure S11, we further added plots under HPV subgroups.

Page 6 Results Line 1 "However, results from scRNA-seq data suggests that the overall Treg expression signature may be underestimated in genomic projects that are biased towards tumors with higher purity, such as TCGA. "This point is likely to be true, but the alternative of studying samples with variable amounts of stroma is probably not ideal either unless similar methods to those described here can be used estimate tumor and stromal fractions in each sample. May want to address in the discussion. Histopath of tumors reveals lots of HNSCC with large scale T cell invasion in the tumor not in stroma.

Author’s Response: Thanks for the thoughtful comments. On a side note, we totally agree with the reviewer that there is no ideal alternative to purity-oriented sampling and most T cells are in the tumor not the stroma. However, the tumor-stroma margin region might be the missing spot. As suggested, we commented in the discussion section that “The single-cell HNSCC data is complementary to the TCGA bulk tumor data in that, while TCGA designs have been focused on tumor regions, scRNA-seq experiments can capture more immune cells in the surrounding
stroma or tumor margin, where a higher amount of lymphocytes such as Treg cells might reside in”.

Figure S1 - Should make it a regular figure as it adds clarity to understanding of the cluster diagram of Figure 1A when patient origin of sample is included.

Author’s Response: Agreed. We have only made Figure S1 a regular figure (Figure 1C).

As a non-mathematician I am not totally clear on how much true external validation was done, though the work depicted in S8, S9 and S10 is most clearly an example of external validation.

Figures 2D and 3B are hard to understand. May need better description.

Author’s Response: As suggested, more descriptions were added to Figure 2D and 3B. Please see the highlighted text in figure legends.

Hideo Baba, Ph.D., M.D. (Reviewer 2): The rapid development of single-cell RNA sequencing (scRNA-seq) provides unprecedented opportunities to directly profile the cell composition and understand tumor heterogeneity at a cellular level. The authors introduced a scheme for characterizing cell compositions from bulk tumor gene expression by integrating signatures learned from scRNA-seq data. They validated results by using in silico pooled bulk tumor samples, and also showed that single-cell-derived signatures provides the ability to separate T cell subtypes. This study may have considerable clinical implications. The data are very well presented. However, there are a few issues to be addressed before any possible acceptance of the manuscript.

Author’s Response: Thanks much for the positive comments.

1. The inclusion and selection criteria of this study should be described in more detail. Why did the authors select these 16 HNSCC patients?

Author’s Response: There were 18 patients in the original Puram et al. 2017 paper. We excluded two patients (MEEI9 and MEEI23) that have <=50 cells sequenced.
2. The authors need to demonstrate the information on clinicopathological factors in 16 HNSCC patients.

Author’s Response: We have added clinical characteristic of the 16 patients in Supplementary Table S1.

3. The authors should correct the following misspellings.

"scRNA-Req-derived" on line19, abstract.

"[CD8+ cells" on p10, line32.

Author’s Response: Corrected. Thanks.