Jumping on the Single-Cell RNA-Sequencing Bandwagon: Take Care Not to Put the Cart before the Horse

To the Editor:

Single-cell RNA sequencing (scRNA-seq) has the unprecedented advantage of resolving the signal from cell populations that otherwise would be hidden in the output from a bulk preparation. Therefore, this technique can provide insights into cellular heterogeneity in tissues to help identify previously unknown cell types (1), and define subpopulations within a known cell type (2) and their response to certain stimuli. Moreover, the identification of genes that drive the clustering of cells, based on the differential expression of those genes, can help identify specific markers for the cell population of interest.

Although the strengths of this technology should be acknowledged, it is equally important to be aware of its limitations. For example, cigarette smoke and IL-13 increase expression of the mucin gene MUC5AC in airway epithelial cells. This was shown many years ago and subsequently reproduced by many groups. Although it is reassuring that this latest (and expensive) technology has reproduced these findings, the pathways involved in MUC5AC expression may only seem identical because expression of SPDEF is induced by both stimuli (3). However, cigarette smoke activates EGFR to stimulate ERK1/2 phosphorylation (4), whereas IL-13 activates STAT6 to stimulate expression of MUC5AC (5). Because these upstream pathways are entirely affected by post-translational modifications, they would remain unknown if conclusions were drawn solely from scRNA-seq analyses.

Moreover, post-translational changes that occur on the protein level, including phosphorylation, ubiquitylation, sumoylation, and acetylation, to name a few, drive processes that ultimately form the phenotype of a cell but cannot be tested by RNA-seq approaches. Therefore, relying on RNA analyses that focus on a single component on the protein level can be misleading. Methods to capture post-translational modifications on the single-cell level are being explored and will, of course, be paradigm shifting.

As is the case with many new technologies, we are excited by the new breakthroughs and attempt to apply novel techniques with the hope of gaining insights into a previously unresolved question or problem, or just to “keep up with the latest science.” Also, individuals involved in peer reviews of papers or grants may suffer from the opinion that such techniques are “novel” or “innovative” and thus may be taken as indicators of research being at the forefront. It is important to remember that one should not use a new technology just to check off the “innovation box,” without carefully considering the appropriate techniques that would ultimately address the issues in a specific scenario.

Data from scRNA-seq have been used to generate hypotheses about the differentiation of stem cells and/or progenitor cells. It should be emphasized that without cell lineage tracing experiments, it is not appropriate to define which cell may be a precursor to a differentiated cell. Dying cells may have expression profiles similar

to those of differentiated cells and therefore be mislabeled as precursors. Promisingly, in a recent study, Xue and colleagues were able to resolve stem cells in caudal hematopoietic tissue from zebrafish in a spatiotemporal manner by combining data derived from temporal bulk RNA-seq and *in vivo* Gene Expression Omnibus sequencing and scRNA-seq (6). However, before the weaknesses noted above are addressed, we need to be cautious about how we view, approach, and apply this new technology.

Hence, scRNA-seq is only a tool that can help answer specific questions and is not a panacea that can solve all problems or prove/disprove any or all hypotheses. Many shortcomings of this technique, including the facts that expertise in analyzing large datasets is limited to certain groups, and that conclusions are drawn by extrapolation and without validation, need to be considered. Given the large cost of conducting scRNA-seq, it is imperative to pay attention to these details, as it is our scientific duty to use taxpayer dollars in a judicious manner. ■

Author disclosures are available with the text of this letter at www.atsjournals.org.

Venkataramana K. Sidhaye, M.D.
Johns Hopkins University
Baltimore, Maryland

Min Wu, M.B.B.S., Ph.D.
University of North Dakota
Grand Forks, North Dakota

Yohannes Tesfagzi, Ph.D.*
Brigham and Women's Hospital–Harvard Medical School
Boston, Massachusetts

ORCID ID: 0000-0002-3997-5839 (Y.T.).

*Corresponding author (e-mail: ytesfagzi@bwh.harvard.edu).

References

- Montoro DT, Haber AL, Biton M, Vinarsky V, Lin B, Birket SE, *et al*. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* 2018;560:319–324.
- Reyffman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, *et al*. Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. *Am J Respir Crit Care Med* 2019;199:1517–1536.
- Park KS, Korfhagen TR, Bruno MD, Kitzmiller JA, Wan H, Wert SE, *et al*. SPDEF regulates goblet cell hyperplasia in the airway epithelium. *J Clin Invest* 2007;117:978–988.
- Mebratu YA, Schwalm K, Smith KR, Schuyler M, Tesfagzi Y. Cigarette smoke suppresses Bik to cause epithelial cell hyperplasia and mucous cell metaplasia. *Am J Respir Crit Care Med* 2011;183:1531–1538.
- Thai P, Chen Y, Dolganov G, Wu R. Differential regulation of MUC5AC/Muc5ac and hCLCA-1/mGob-5 expression in airway epithelium. *Am J Respir Cell Mol Biol* 2005;33:523–530.
- Xue Y, Liu D, Cui G, Ding Y, Ai D, Gao S, *et al*. A 3D atlas of hematopoietic stem and progenitor cell expansion by multi-dimensional RNA-seq analysis. *Cell Rep* 2019;27:1567–1578, e5.

Copyright © 2020 by the American Thoracic Society

Author Contributions: Y.T. drafted the manuscript. V.K.S. and M.W. contributed substantially to finalizing the ideas.