

RESEARCH HIGHLIGHT

Open Access



New skin for the old RNA-Seq ceremony: the age of single-cell multi-omics

Maayan Baron and Itai Yanai*

Abstract

New methods for simultaneously quantifying protein and gene expression at the single-cell level have the power to identify cell types and to classify cell populations.

While we are taught to never judge a book by its cover, covers may actually be revealing. In the case of a cell, the surface proteins on its “cover” are unique to particular cell types: for example, CD3 for T cells and CD19 for B cells. With such markers in hand, populations of cells can be classified into the cell types they contain, in particular using fluorescence-activated cell sorting (FACS) analysis with a panel of antibodies. Over the past 5 years, however, a newer technology for characterizing populations has emerged, known as single-cell RNA-seq. Similarly to FACS, cells can be clustered according to their transcriptomes and cell types and subpopulations become readily identifiable [1, 2]. For example, when we previously studied pancreatic tissue in mouse and human we identified 15 cell types and subpopulations of the ductal cells [3]. However, it was not known if the cell-surface markers and transcriptomes provided consistent information. What would the other measurable quantities reveal? Ultimately, it appears that many further insights could be gleaned from an analysis of cells with several high-throughput methods, all at once. We would want to integrate RNA-seq measurements with genome sequencing, protein profiles, post-transcriptional regulation, metabolomics and lipidomics, together with the cellular localization of each—all at single-cell resolution, of course. Such complete characterization of cells at the population level would be a true treasure trove for insights into cellular physiology and pathological states.

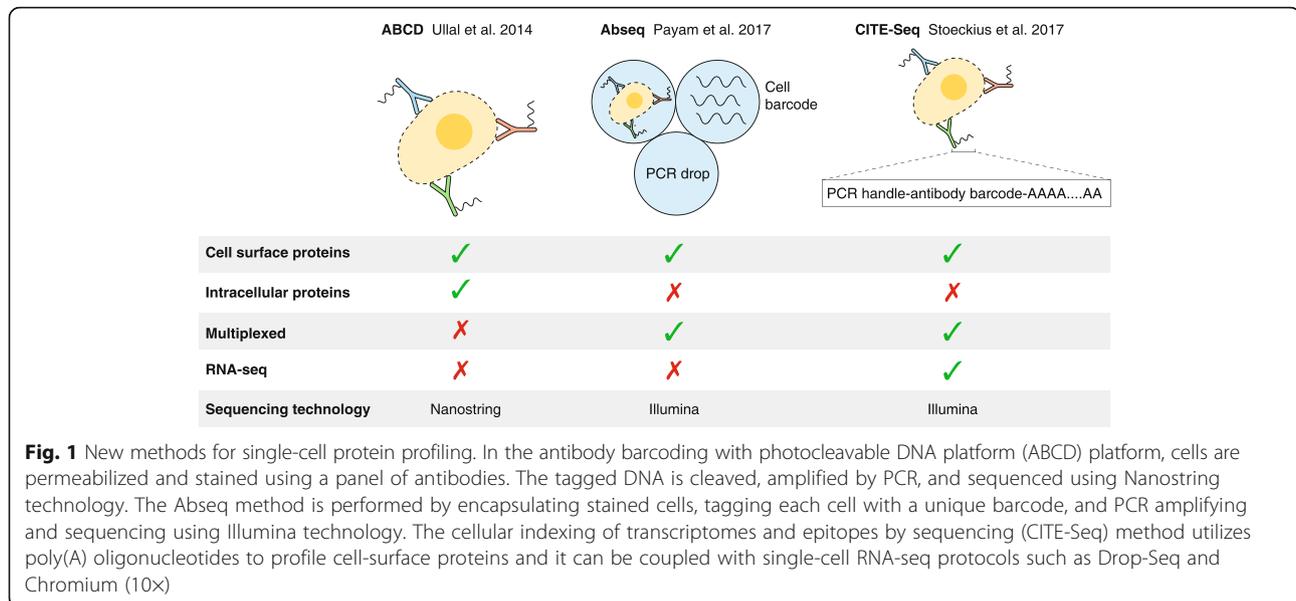
A recent work published in *Nature Methods* has made a significant step forward toward multi-omics [4] by producing both transcriptomes and cell-surface protein quantifications on populations of cells.

Cytometry by sequencing

The method—called CITE-Seq (cellular indexing of transcriptomes and epitopes by sequencing)—can be seen as a composite of two main concepts for how to derive cell-surface proteomics and transcriptomics from individual cells: DNA-conjugated antibodies and single-cell RNA-seq [4]. Detecting protein levels in individual cells is challenging as a result of low starting amounts and the lack of direct amplification methods common for nucleic acids. New techniques for protein profiling have been published in 2014 and earlier this year [5, 6]. The main insight for how to derive cell-surface proteomics is to tag proteins with antibodies conjugated to oligonucleotides (Fig. 1). By converting the detection of a protein to an oligonucleotide, the signal can then be amplified by exploiting Watson–Crick pairing of nucleic acids. This notion has been called “cytometry by sequencing” [4]. The identity of each protein is encoded in the oligonucleotides, which recapitulate a large number of distinguishable proteins: a sequence of length N corresponds to 4^N unique sequences, and therefore even a sequence of eight bases would theoretically suffice for capturing all cellular proteins.

Ullal et al. [5] first demonstrated this approach by developing the antibody barcoding with photocleavable DNA platform (ABCD; Fig. 1). The authors processed bulk samples of approximately 100 cells each, as well as samples containing individual cells, from a fine-needle aspiration and profiled over 90 proteins spanning cancer-relevant pathways. The cells are permeabilized before staining, thus enabling characterization of surface and intracellular proteins. In particular, in bypassing FACS and using only non-specialized instruments, ABCD is attractive for many clinical applications. The main constraint, however, is the lack of multiplexing of the different

* Correspondence: Itai.Yanai@nyumc.org
Institute for Computational Medicine, NYU School of Medicine, 430 East 29th St., New York, NY 10016, USA



samples/cells, which thereby limits the handling to only a few samples at a time. More recently, the Abseq method was introduced, which utilizes custom microfluidics devices [6] to achieve a multiplexed version of cytometry by sequencing (Fig. 1). First, cells are incubated with a variety of antibodies conjugated with oligonucleotides encoding the protein identity, followed by encapsulation in drops and pairing with additional oligonucleotides to barcode the cells. Altogether, this method requires three separate microfluidic chips, and constitutes an impressive technical feat for single-cell proteomics.

RNA-Seq and cell-surface proteomics in a drop

As in Abseq, cells in the CITE-Seq method are first incubated with cell-surface antibodies conjugated with oligonucleotides encoding the protein identity. CITE-Seq's second underlying concept is the application of single-cell RNA-seq. The particular novelty is the design of the oligonucleotides tagged to the antibodies which contain a poly(A) region compatible with existing single-cell RNA-seq methods, either well-based or droplet-based. To demonstrate its general applicability, Stoeckius et al. [4] successfully implemented CITE-Seq with two established high-throughput methods—Drop-Seq and 10x—to profile in parallel both the transcriptome and several cell-surface proteins of immune cells.

Using an antibody for CD8 as a proof-of-principle, Stoeckius et al. [4] demonstrated that the expression profile is comparable with the results detected by FACS. This is an important comparison since FACS has been the gold standard for the past two decades for profiling protein levels in millions of cells, for its speed, sensitivity, and capacity to profile tens of proteins at once (or even more if using multiplexing approaches) [7]. Extending to a

broader set of ten immune cell surface proteins, Stoeckius et al. [4] showed that identifying cell types using protein expression profiles generally corresponded to RNA expression. This marks a multi-omics first in quantifying both surface proteins and transcriptomes in a population of cells. Surprisingly, the correlations between mRNA and protein levels were low for individual cells ($0.02 < R < 0.53$), though higher when averaging across cell types ($0.58 < R < 0.95$). Moreover, the authors showed that using CITE-Seq enhanced the characterization of known subtypes of natural killer cells (CD56 bright and dim), which was not previously detected using single-cell RNA-seq methods. Thus, with the combination of both surface proteins and transcriptomics, new subpopulations may be revealed which would not be possible without their combination.

Future directions

Recently, the construction of a human cell atlas through a large-scale collaborative project has been proposed [8]. While single-cell RNA-seq is currently best positioned to provide the methodology for such an atlas, CITE-Seq now provides a possible additional layer of information. By classifying cells based upon a multi-omics approach, a refined atlas may be possible. How might CITE-Seq be scaled-up to provide a richer proteome beyond cell-surface proteins? A method for mild permeabilization seems a necessary addition to the protocol in order to also capture intracellular proteins. Practical issues regarding the number of antibodies used in parallel would also need to be addressed, in particular the issue of cross-reaction with unwanted epitopes. In addition, scaling-up to all proteins in a cell may introduce biases based upon differences in antibody affinities. Overall, these are exciting times in

which both the cover and the inside of the cell are used to characterize its biology.

Abbreviations

ABCD: Antibody barcoding with photocleavable DNA platform;
CITE-Seq: Cellular indexing of transcriptomes and epitopes by sequencing;
FACS: Fluorescence-activated cell sorting

Acknowledgements

None.

Funding

None.

Authors' contributions

MB and IY drafted and finalized the manuscript. Both authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Published online: 24 August 2017

References

1. Kowalczyk MS, Tirosh I, Heckl D, Rao TN, Dixit A, Haas BJ, et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. *Genome Res.* 2015;25:1860–72.
2. Zeisel A, Muñoz-Manchado AB, Codeluppi S, Lönnerberg P, La Manno G, Juréus A, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science.* 2015;347:1138–42. doi:10.1126/science.aaa1934.
3. Baron M, Veres A, Wolock SL, Faust AL, Gaujoux R, Vetere A, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. *Cell Syst.* 2016;3:346–60. doi:10.1016/j.cels.2016.08.011.
4. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods.* 2017. doi: 10.1038/nmeth.4380.
5. Ullal AV, Peterson V, Agasti SS, Tuang S, Juric D, Castro CM, Weissleder R. Cancer cell profiling by barcoding allows multiplexed protein analysis in fine needle aspirates. *Sci Transl Med.* 2014;6:219ra9. doi:10.1126/scitranslmed.3007361.
6. Shahi P, Kim SC, Haliburton JR, Gartner ZJ, Abate AR. Abseq: ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. *Sci Rep.* 2017;7:44447. doi:10.1038/srep44447.
7. Robinson JP, Roederer M. Flow cytometry strikes gold. *Science.* 2015;350:739–40.
8. Regev A, Teichmann S, Lander ES, Amit I, Benoist C, Birney E, et al. The human cell atlas. *bioRxiv.* 2017. doi:https://doi.org/10.1101/121202.