MOLECULAR BIOLOGY

Accelerating a paradigm shift: The Common Fund Single Cell Analysis Program

Ananda L. Roy[1,2]*, Richard Conroy[1,2], Jessica Smith[1,2], Yong Yao[3], Andrea C. Beckel-Mitchener[3], James M. Anderson[2], Elizabeth L. Wilder[1,2]

It has become exceedingly important to understand the precise molecular profiles of the nearly 40 trillion cells in an adult human because of their role in determining health, disease, and therapeutic outcome. The National Institutes of Health (NIH) Common Fund–supported Single Cell Analysis Program (SCAP) was designed to address this challenge. In this review, we outline the original program goals and provide a perspective on the impact of the program as a catalyst for exploration of heterogeneity of human tissues at the cellular level. We believe that the technological advances in single-cell RNA sequencing and multiplexed imaging combined with computational methods made by this program will undoubtedly have an impact on broad and robust applications of single-cell analyses in both health and disease research.

INTRODUCTION

"The use of phage as an index to the state of an individual cell would appear to have general usefulness in a variety of problems in bacterial physiology where the proper interpretation of an overall observation depends upon the proverbial question of whether it represents a change in all or only some of the cells." This is the concluding statement from a landmark paper by Seymour Benzer in 1953 (*1*), which highlighted a ubiquitous question in the study of any population. Although the notion of cellular heterogeneity is not new, until recently, the appropriate methods to study this heterogeneity in mammalian cells were not well developed. While it has been known for quite some time that behavior and function of individual mammalian cells vary significantly from a cell population (*2*, *3*), a decade ago, it became clear that strategic development of next-generation tools and methods was necessary to understand and distinguish distinct cells in an apparent homogeneous population. This growing need for new technologies for studying heterogeneity in complex environments and awareness that heterogeneity has important functional implications led to the development of a single-cell analysis program by the National Institutes of Health (NIH) Common Fund (CF).

A major goal of the CF is to identify and address emerging scientific opportunities and outstanding challenges in biomedical research that no single NIH Institute or Center is poised to address on its own. In addition to supporting trans-NIH programs, the CF provides a "venture capital" space for high-risk, innovative research with potential for unusually high impact. These goal-driven programs are meant to catalyze a field in a span of 5 to 10 years. Planning for the Single Cell Analysis Program (SCAP; https://commonfund.nih.gov/singlecell) began in 2010, and it was formally launched in 2012 with the overall vision of accelerating the discovery, development, and translation of cross-cutting, innovative approaches for analyzing the heterogeneity of biologically relevant populations of human cells in situ. This program had the following goals: evaluate cellular heterogeneity through transcriptomic analysis, develop exceptionally innovative tools and technologies for single-cell analysis, accelerate the integration and translation of technologies to characterize biological processes at the single-cell level, and develop and translate single-cell methods for clinical diagnosis and therapy. An additional goal to measure and manipulate the heterogeneity of individual cells in situ was added later. This program ended in 2017; we provide a perspective here on the impact of the program as a catalyst for exploration of human tissues at the cellular level.

TECHNOLOGY DEVELOPMENT

Although cells are the basic unit of biological function, most methods for studying normal physiology and disease do not have resolution at the cellular level. Thus, rare cell types and their functions may be missed, dynamic events that lead to changing cellular states are difficult to detect, and variability between cells and the functional significance of the variability may be difficult to measure or overlooked altogether. While the importance of analyzing individual cells was becoming increasingly apparent, methods to enable analysis of individual cells from complex environments were not available. To overcome this challenge, the SCAP began in 2012 with an emphasis on technology development. At that time, RNA sequencing (RNA-seq) was starting to be used for ensemble measurement of gene expression in human tissues, while RNA and DNA fluorescence in situ hybridization (FISH) techniques were being pushed to subcellular resolution and higher sensitivity (although mostly for just a few genes) (*4*, *5*). Single-cell technologies—in particular, single-cell transcriptomics and live-cell imaging—have advanced significantly during the past 7 years, improving detection sensitivity, multiplexing, and assay throughput and allowing single-cell measurements in the context of the local microenvironment (*6*–*8*).

A variety of single-cell technologies have been developed by the SCAP to analyze DNA sequence, DNA methylation, chromosome conformation, and chromatin state. This program supported the development of microdroplet and microchip-based single-cell RNA-seq protocols, helping to pave the way for marked expansion of their use across biology and medicine. The program also supported projects to improve the throughput of assays, increasing the scope and scale of experimental measures. There have also been many other independently developed methods that catalyzed the field of single-cell analysis. For instance, hundreds to thousands of unique RNA

[1]Office of Strategic Coordination, National Institutes of Health (NIH), 6001 Executive Boulevard, Bethesda, MD 20892, USA. [2]Division of Program Coordination, Planning, and Strategic Initiatives, Office of the Director, NIH, Bethesda, MD 20892, USA. [3]Division of Neuroscience and Basic Behavioral Science, National Institute of Mental Health, NIH, Bethesda, MD 20892, USA.
*Corresponding author. Email: ananda.roy@nih.gov

species can now be sequenced or imaged at single-copy sensitivity, and more than 40 unique protein species can be assayed simultaneously in a high-throughput manner using mass cytometry (*9*).

During the life cycle of the SCAP, technology development evolved around three broad themes: (i) droplet-based sequencing approaches [for example, inDrop for high-throughput single-cell transcriptomics (*10*)], (ii) enhanced spatial resolution via fluorescence-based techniques [FISSEQ (*11*)], and (iii) barcoding techniques to multiplex microscopic approaches [for example, seqFISH (*12*)]. Further advancements in sequencing and imaging techniques were achieved. These included isolated single nuclei in cells and postmortem tissues (*13*, *14*) as well as capture techniques such as transcriptome in vivo analysis (TIVA) tag for RNA capture (*15*) and the dual-view inverted selective plane illumination microscope (diSPIM) imaging system, which provides tracking of cell lineage in living multicellular organisms (*16*). The SCAP has also significantly moved the field of single-molecule FISH for RNA and protein detection forward (*17*), including the detection of single-nucleotide variants, multiplexing the detection of tens of RNAs and proteins, and RNA detection in live cells using CRISPR (clustered regularly interspaced short palindromic repeats) (*18*). Together, these techniques have contributed significantly to the tool kit for single-cell analysis and catalyzed the development of next-generation technologies, initiated commercialization of analytical products, and provided novel biological insights to better understand the role of cellular heterogeneity in the field of cancer, immunology, and neuroscience. Gratifyingly, SCAP-supported projects continue to yield novel methods for single-cell analysis even after the program officially ended—for instance, transcriptome-scale RNA SPOT has the potential to lower sequencing costs (*19*) as well as single-nucleus analysis of accessible chromatin in developing mouse brain (*20*) and analysis of nuclear lamins (*21*). Program-generated resources also include establishment of the SCAP-Transcriptome (SCRAP-T) Consortium project, which developed a public portal (www.scap-t.org) with phenotypic information and whole-transcriptome data from 56 human subjects. The SCAP-T Consortium established procedures for depositing metadata and RNA-seq data to the NIH National Center for Biotechnology Information (NCBI) database of Genotypes and Phenotype (dbGaP).

"FOLLOW THAT CELL" GRAND CHALLENGE

An outstanding scientific challenge for the single-cell analysis community was comprehensive dynamic measurement so that alterations in cell behavior and function of individual cells could be better assessed over time. To address this challenge, the SCAP instituted a grand challenge in 2014, called "Follow that Cell" (https://commonfund.nih.gov/singlecell/challenge), the largest multiphase prize organized by the NIH at that time. The primary purpose of the "Follow that Cell" Challenge was to stimulate the development of new tools and methods for analyzing dynamic states of individual cells that can serve as the basis for predicting alterations in cell behavior and function over time. The first phase lasted 4 months, and from the proposed concepts, a team of experts identified 16 compelling theoretical ideas that went forward to a reduction-to-practice second phase. Ten of these teams submitted experimental results 2 years later at the end of phase 2. A winner and a runner-up prize were awarded in June 2017. The winning project demonstrated a novel nanopipette technology that can be used to monitor the molecular properties of single cells over time and was convincingly demonstrated to repeatedly and nondestructively interrogate a population of single cells (*22*). The runner-up project developed a nondestructive way for a heterogeneous population of cells to "self-report" their internal transcriptional state in a quantitatively interpretable manner by exporting part of their cytoplasm in a barcoded capsid. Apart from these two winners,

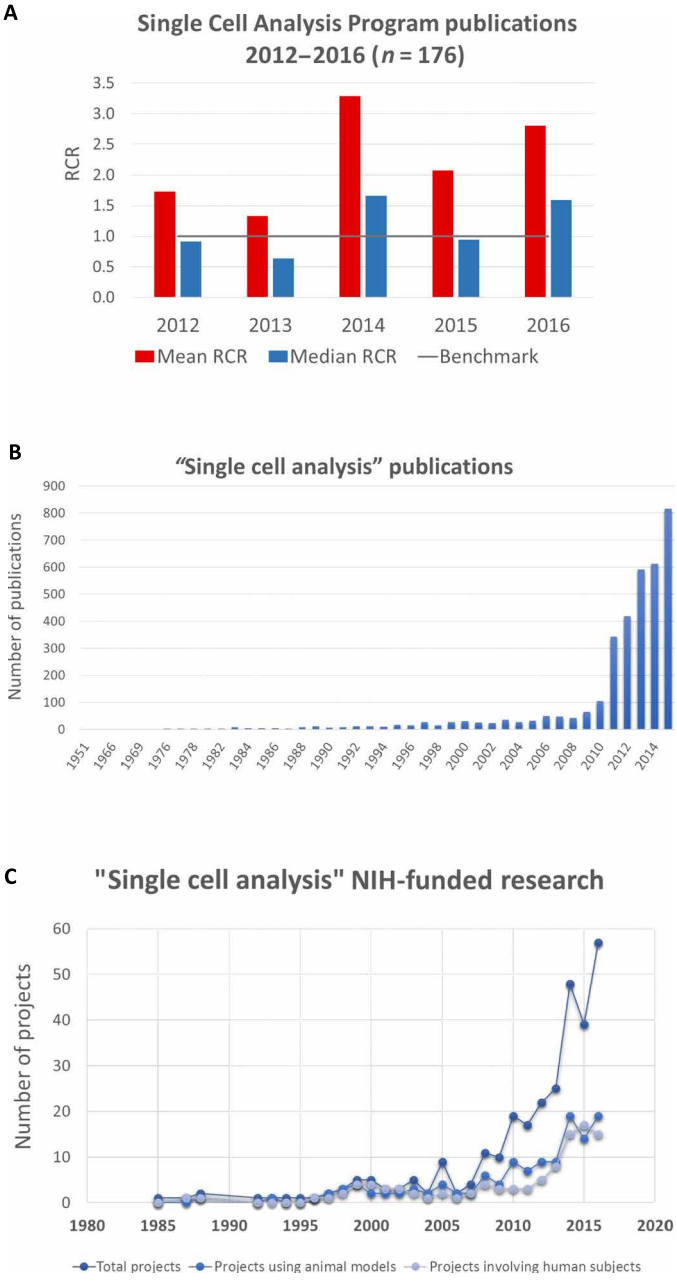

Fig. 1. Historical trends of single cell analysis publications and research projects. (**A**) Type 1 applications that were funded as part of the SCAP generated 218 publications. Of these publications, 176 are research articles that were published between 2012 and 2016. The mean and median RCR (Relative Citation Ratio) values are shown here. The RCR benchmark of 1.0 is based on the median RCR for all NIH-funded publications. For additional information about RCR, see (*17*). (**B**) A text search of PubMed for publications that use the term "single cell analysis" yields 3994 publications, with a 73% increase since 2012 (the year the SCAP was launched). (**C**) A text search of the Information for Management, Planning, Analysis, and Coordination II (IMPACII) database for awarded NIH applications that use the term "single cell analysis" resulted in 320 unique projects since 1985.

other notable solutions offered single-molecule barcode nanobiosensors for probing collective cell migration and photostable multiplexing nanoAssays for real-time molecular imaging of single live cells.

PROGRAM ACHIEVEMENTS AND LESSONS LEARNED

When the SCAP started, the field of single-cell analysis was at its nascent stages for studying heterogeneity of mammalian cells in complex environments. Although it is hard to quantify whether the field of single-cell analysis would have thrived without the support of the CF, it is clear that the SCAP helped catalyze and accelerate the field. For instance, between 2012 and 2016, the investigators in the SCAP published 216 papers. One hundred seventy-six of these papers are research articles that have received more than 4332 citations. The mean Relative Citation Ratio [RCR; (*23*)] for these papers is 2.50, indicating that publications from the SCAP were cited approximately 2.50 times more frequently than the average NIH-funded publication (Fig. 1A). It is worthwhile to note that, since the inception of the SCAP in 2012, there has also been a substantial increase in NIH-funded projects that use the term "single cell analysis" as well as a sharp increase in publications in this area (Fig. 1, B and C). An additional aspect of the program also emerged quickly after the program launched in 2012; in response to significant interest from the small business community, a companion Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) opportunities were created by the NIH SCAP Working Group, which resulted in 18 additional awards, and these awards were included in program-wide activities. The open annual investigator meetings also saw a significant rise in attendance throughout the program, with more than 350 researchers registering for the final meeting.

In the years since the SCAP began, single-cell RNA analytic methods have come of age and are increasingly being incorporated into large studies. For instance, a collaboration with the Genotype-Tissue Expression (GTEx) program (*24*) resulted in SCAP projects using GTEx tissues to generate single-cell analysis data, while GTEx investigators used techniques and protocols developed by the SCAP and projects funded in the NIH BRAIN Initiative (www.braininitiative.nih.gov). The tools developed from the SCAP will also be useful for NIH programs like the BRAIN Cell Census Consortium (*25*), the National Cancer Institute's Tumor Cell Atlas, and a new CF program, the Human BioMolecular Atlas Program (HuBMAP; https://commonfund.nih.gov/HuBMAP). The broad objective of HuBMAP is to catalyze the development of an open, global framework for comprehensively mapping the human body at a cellular resolution, the basis for which was established by the technology development supported by the SCAP. However, additional investments in technology development will be required to enable analysis of proteins, protein modifications, metabolites, and intercellular signaling in a high-throughput and robust manner across multiple human tissues at cellular resolution. These technologies represent the next wave of innovation and will be a focus of the HuBMAP program. The field of single-cell analysis is poised to launch the next phase of research addressing cellular interactions, history, states, and profiling in human tissues, thereby providing an unprecedented view of cellular and molecular composition for more precise diagnosis and therapies (*8*).

REFERENCES AND NOTES

1. S. Benzer, Induced synthesis of enzymes in bacteria analyzed at the cellular level. *Biochim. Biophys. Acta* **11**, 383–395 (1953).
2. A. Raj, A. van Oudenaarden, Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008).
3. C. Trapnell, Defining cell types and states with single-cell genomics. *Genome Res.* **25**, 1491–1498 (2015).
4. A. M. Femino, F. S. Fay, K. Fogarty, R. H. Singer, Visualization of single RNA transcripts in situ. *Science* **280**, 585–590 (1998).
5. A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, S. Tyagi, Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **5**, 877–879 (2008).
6. T. Bartfai, P. T. Buckley, J. Eberwine, Drug targets: Single-cell transcriptomics hastens unbiased discovery. *Trends Pharmacol. Sci.* **33**, 9–16 (2012).
7. O. Symmons, A. Raj, What's luck got to do with it: Single cells, multiple fates, and biological nondeterminism. *Mol. Cell* **62**, 788–802 (2016).
8. A. Tanay, A. Regev, Scaling single-cell genomics from phenomenology to mechanism. *Nature* **541**, 331–338 (2017).
9. M. H. Spitzer, G. P. Nolan, Mass cytometry: Single cells, many features. *Cell* **165**, 780–791 (2016).
10. A. M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D. A. Weitz, M. W. Kirschner, Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
11. J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, T. C. Ferrante, R. Terry, B. M. Turczyk, J. L. Yang, H. S. Lee, J. Aach, K. Zhang, G. M. Church, Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat. Protoc.* **10**, 442–458 (2015).
12. A. F. Coskun, L. Cai, Dense transcript profiling in single cells by image correlation decoding. *Nat. Methods* **13**, 657–660 (2016).
13. B. B. Lake, R. Ai, G. E. Kaeser, N. S. Salathia, Y. C. Yung, R. Liu, A. Wildberg, D. Gao, H.-L. Fung, S. Chen, R. Vijayaraghavan, J. Wong, A. Chen, X. Sheng, F. Kaper, R. Shen, M. Ronaghi, J.-B. Fan, W. Wang, J. Chun, K. Zhang, Neuronal subtypes and diversity revealed by single nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
14. Y. Wang, J. Waters, M. L. Leung, A. Unruh, W. Roh, X. Shi, K. Chen, P. Scheet, S. Vattathil, H. Liang, A. Multani, H. Zhang, R. Zhao, F. Michor, F. Meric-Bernstam, N. E. Navin, Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
15. D. Lovatt, B. K. Ruble, J. Lee, H. Dueck, T. K. Kim, S. Fisher, C. Francis, J. M. Spaethling, J. A. Wolf, M. S. Grady, A. V. Ulyanova, S. B. Yeldell, J. C. Griepenburg, P. T. Buckley, J. Kim, J.-Y. Sul, I. J. Dmochowski, J. Eberwine, Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. *Nat. Methods* **11**, 190–196 (2014).
16. A. Kumar, Y. Wu, R. Christensen, P. Chandris, W. Gandler, E. McCreedy, A. Bokinsky, D. A. Colón-Ramos, Z. Bao, M. McAuliffe, G. Rondeau, H. Shroff, Dual-view plane illumination microscopy for rapid and spatially isotropic imaging. *Nat. Protoc.* **9**, 2555–2573 (2014).
17. P. Ginart, J. M. Kalish, C. L. Jiang, A. C. Yu, M. S. Bartolomei, A. Raj, Visualizing allele-specific expression in single cells reveals epigenetic mosaicism in an *H19* loss-of-imprinting mutant. *Genes Dev.* **30**, 567–578 (2016).
18. B. Chen, J. Hu, R. Almeida, H. Liu, S. Balakrishnan, C. Covill-Cooke, W. A. Lim, B. Huang, Expanding the CRISPR imaging toolset with *Staphylococcus aureus* Cas9 for simultaneous imaging of multiple genomic loci. *Nucleic Acids Res.* **44**, e75 (2016).
19. C.-H. L. Eng, S. Shah, J. Thomassie, L. Cai, Profiling the transcriptome with RNA SPOTs. *Nat. Methods* **14**, 1153–1155 (2017).
20. S. Preissl, R. Fang, H. Huang, Y. Zhao, R. Raviram, D. U. Gorkin, Y. Zhang, B. C. Sos, V. Afzal, D. E. Dickel, S. Kuan, A. Visel, L. A. Pennacchio, K. Zhang, B. Ren, Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat. Neurosci.* **21**, 432–439 (2018).
21. J.-K. Kim, A. Louhghalam, G. Lee, B. W. Schafer, D. Wirtz, D.-H. Kim, Nuclear lamin A/C harnesses the perinuclear apical actin cables to protect nuclear morphology. *Nat. Commun.* **14**, 2123 (2017).
22. E. N. Tóth, A. Lohith, M. Mondal, J. Guo, A. Fukamizu, N. Pourmand, Single-cell nanobiopsy reveals compartmentalization of mRNAs within neuronal cells. *J. Biol. Chem.* **293**, 4940–4951 (2018).
23. B. I. Hutchins, X. Yuan, J. M. Anderson, G. M. Santangelo, Relative Citation Ratio (RCR): A new metric that uses citation rates to measure influence at the article level. *PLOS Biol.* **14**, e1002541 (2016).
24. The GTEx Consortium, The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
25. J. R. Ecker, D. H. Geschwind, A. R. Kriegstein, J. Ngai, P. Osten, D. Polioudakis, A. Regev, N. Sestan, I. R. Wickersham, H. Zeng, The BRAIN Initiative Cell Census Consortium: Lessons learned toward generating a comprehensive brain cell atlas. *Neuron* **96**, 542–557 (2017).

Acknowledgments
Funding: No funding was received for this manuscript. **Author contributions:** A.L.R. and E.L.W. wrote the article with input from R.C., Y.Y., A.C.B.-M., and J.M.A. J.S. did the analysis that led to the figure. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the materials cited herein. Additional data related to this paper may be requested from the authors.

Submitted 12 April 2018
Accepted 20 June 2018
Published 1 August 2018
10.1126/sciadv.aat8573

Citation: A. L. Roy, R. Conroy, J. Smith, Y. Yao, A. C. Beckel-Mitchener, J. M. Anderson, E. L. Wilder, Accelerating a paradigm shift: The Common Fund Single Cell Analysis Program. *Sci. Adv.* **4**, eaat8573 (2018).

Accelerating a paradigm shift: The Common Fund Single Cell Analysis Program

Ananda L. Roy, Richard Conroy, Jessica Smith, Yong Yao, Andrea C. Beckel-Mitchener, James M. Anderson and Elizabeth L. Wilder

Sci Adv **4** (8), eaat8573.
DOI: 10.1126/sciadv.aat8573