

New frontiers in single-cell genomics

Nicholas E. Navin,^{1,2,3} Orit Rozenblatt-Rosen,⁴ and Nancy R. Zhang⁵

¹Department of Genetics, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA; ²Graduate School of Biomedical Sciences, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA; ³Department of Bioinformatics, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA; ⁴Genentech, Inc., South San Francisco, California 94080, USA; ⁵Department of Statistics and Data Science, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

The burgeoning field of single-cell genomics has undergone enormous progress since its inception more than a decade ago. The rapid growth of the field has been fueled by the development of innovative technologies, novel computational analysis methods, and a growing number of diverse applications across many fields of biology and biomedicine. Advances in single-cell profiling technologies now allow us to chart cells at unprecedented resolution. Today, through the development of high-throughput methods including droplets, nanowell, and combinatorial indexing methods, we can profile thousands of cells in parallel. Single-cell and single-nucleus RNA-seq are now used for data generation by many laboratories. Genomic methods for measuring chromatin states and for multimodal profiling are also maturing and there is a growing number of spatial technologies for profiling cells in the tissue context. Much progress has been made in the analysis and annotation of single-cell data and its integration with spatial data, and many algorithms, analysis pipelines, and tools are now accessible as open-source software. These experimental and computational methods are now an invaluable part of the molecular biologist's toolkit.

In this special issue of *Genome Research*, we have assembled a diverse collection of research and review articles in the field of single-cell genomics. While much progress has been made on analysis methods for scRNA-seq data, the common tasks of quality control, batch correction, and cell type classification still represent major challenges that are always improving. Thus, the Methods section of this issue begins with four articles (Heiser et al. 2021; Lakkis et al. 2021; Aevermann et al. 2021; Kimmel and Kelley 2021) that address these challenges with new computational algorithms. With the accumulation of single-cell reference data across many tissues, we are now well-positioned to study how they are reprogrammed in disease. The integration of single-cell references with population-level bulk sequencing data is a component in many study designs. The next two articles, by Erdmann-Pham et al. (2021) and Wang et al. (2021a), tackle this challenge by developing novel statistical models for deconvolution of bulk RNA-seq data.

The next five articles report new technologies for single-cell epigenetics (Ohnuki et al. 2021 and Ku et al. 2021), spatial transcriptomics (Miller et al. 2021), single-cell post-transcriptional regulation (Gao et al. 2021), and single-cell metabolomics (Alghamdi et al. 2021). Through novel experimental protocols, bioinformatic tools, and/or mathematical modeling, these Methods expand and enrich the capabilities of single-cell research.

Corresponding authors: NNavin@mdanderson.org, rozenblattrosen.orit@gene.com, nzh@wharton.upenn.edu

Article and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.276129.121>. Freely available online through the *Genome Research* Open Access option.

Applications of single-cell methods are rapidly accumulating in every branch of biomedicine, which we highlight with two articles of very different flavors. The study by Wang et al. (2021b) applies single-cell RNA and chromatin accessibility sequencing to reconstruct the lineage-specific regulatory landscape of hair cells. In a study by Xu et al. (2021), a multimodal methodology for single-cell/nucleus genomics-informed drug discovery was developed and applied to the identification of potential treatments for Alzheimer's disease.

We have also assembled four Resource articles that develop single-cell reference atlases that will serve as valuable resources for our community. Nieto et al. (2021) presents a comprehensive pan-cancer immune cell atlas that integrates single-cell and spatial data sets, Baker et al. (2021) presents an atlas of the *Drosophila* brain exposed to cocaine, Slaidina et al. (2021) presents a single-cell atlas of *Drosophila* ovaries and, finally, Durham et al. (2021) presents an atlas of the tissue-specific chromatin landscape in *Caenorhabditis elegans*.

In addition to these Methods and Resources, we have invited several experts in the field to write Perspective review pieces. Dr. Andrew Adey has contributed a technology review on the history and future use of transposase to develop single-cell genomics approaches (Adey 2021), while Dr. Guo-Cheng Yuan and colleagues have contributed a review on the cutting-edge computational approaches that are being used to study spatial transcriptomic data (Dries et al. 2021). Dr. Itai Yanai and colleagues have written a piece on understanding the diversity of cancer cell states in tumors, and how their emerging properties lead to cancer phenotypes (Barkley et al. 2021). Dr. Mingyao Li and colleagues have written a comprehensive review on the application of single-cell genomic methods to the study of common diseases and population-level variation (Auerbach et al. 2021). These review articles provide both an overview of the studies published in the areas of technology development, cancer research, human disease, and computational methods, as well as a timely outlook on where the field is heading in these areas.

We hope you enjoy reading this special issue, and would like to thank the authors, reviewers, and the Editorial Board members for their hard work and contributions. We look forward to seeing what the next decade of the field will hold, as spatial and multi-omic technology development mature and more biological and translational applications are explored. While we could loosely speculate on the state of the field in another 10 years, we prefer to leave you with some advice from Dr. Einstein: "I never think of the future—it comes soon enough."

© 2021 Navin et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Competing interest statement

N.E.N., O.R.-R., and N.R.Z. served as Guest Editors for this issue of *Genome Research*, and had access to all papers prior to publication. O.R.-R. is an employee of Genentech, Inc.

References

- Adey AC. 2021. Tagmentation-based single-cell genomics. *Genome Res* (this issue) **31**: 1693–1705. doi:10.1101/gr.275223.121
- Aevermann B, Zhang Y, Novotny M, Keshk M, Bakken T, Miller J, Hodge R, Lelieveldt B, Lein E, Scheuermann RH. 2021. A machine learning method for the discovery of minimum marker gene combinations for cell type identification from single-cell RNA sequencing. *Genome Res* (this issue) **31**: 1767–1780. doi:10.1101/gr.275569.121
- Alghamdi N, Chang W, Dang P, Lu X, Wan C, Gampala S, Huang Z, Wang J, Ma Q, Zang Y, et al. 2021. A graph neural network model to estimate cell-wise metabolic flux using single-cell RNA-seq data. *Genome Res* (this issue) **31**: 1867–1884. doi:10.1101/gr.271205.120
- Auerbach BJ, Hu J, Reilly MP, Li M. 2021. Applications of single-cell genomics and computational strategies to study common disease and population-level variation. *Genome Res* (this issue) **31**: 1728–1741. doi:10.1101/gr.275430.121
- Baker BM, Mokashi SS, Shankar V, Hatfield JS, Hannah RC, Mackay TFC, Anholt RRH. 2021. The *Drosophila* brain on cocaine at single-cell resolution. *Genome Res* (this issue) **31**: 1927–1937. doi:10.1101/gr.268037.120
- Barkley D, Rao A, Pour M, Franca GS, Yanai I. 2021. Cancer cell states and emergent properties of the dynamic tumor system. *Genome Res* (this issue) **31**: 1719–1727. doi:10.1101/gr.275308.121
- Dries R, Chen J, del Rossi N, Khan MM, Sistig A, Yuan G-C. 2021. Advances in spatial transcriptomic data analysis. *Genome Res* (this issue) **31**: 1706–1718. doi:10.1101/gr.275224.121
- Durham TJ, Daza RM, Gevirtzman L, Cusanovich DA, Bolonduro O, Noble WS, Shendure J, Waterston RH. 2021. Comprehensive characterization of tissue-specific chromatin accessibility in L2 *Caenorhabditis elegans* nematodes. *Genome Res* (this issue) **31**: 1952–1969. doi:10.1101/gr.271791.120
- Erdmann-Pham DD, Fischer J, Hong J, Song YS. 2021. Likelihood-based deconvolution of bulk gene expression data using single-cell references. *Genome Res* (this issue) **31**: 1794–1806. doi:10.1101/gr.272344.120
- Gao Y, Li L, Amos CI, Li W. 2021. Analysis of alternative polyadenylation from single-cell RNA-seq using scDaPars reveals cell subpopulations invisible to gene expression. *Genome Res* (this issue) **31**: 1856–1866. doi:10.1101/gr.271346.120
- Heiser CN, Wang VM, Chen B, Hughey JJ, Lau KS. 2021. Automated quality control and cell identification of droplet-based single-cell data using dropkick. *Genome Res* (this issue) **31**: 1742–1752. doi:10.1101/gr.271908.120
- Kimmel JC, Kelley DR. 2021. Semisupervised adversarial neural networks for single-cell classification. *Genome Res* (this issue) **31**: 1781–1793. doi:10.1101/gr.268581.120
- Ku WL, Pan L, Cao Y, Gao W, Zhao K. 2021. Profiling single-cell histone modifications using indexing chromatin immunocleavage sequencing. *Genome Res* (this issue) **31**: 1831–1842. doi:10.1101/gr.260893.120
- Lakkis J, Wang D, Zhang Y, Hu G, Wang K, Pan H, Ungar L, Reilly MP, Li X, Li M. 2021. A joint deep learning model enables simultaneous batch effect correction, denoising, and clustering in single-cell transcriptomics. *Genome Res* (this issue) **31**: 1753–1766. doi:10.1101/gr.271874.120
- Miller BF, Bambah-Mukku D, Dulac C, Zhuang X, Fan J. 2021. Characterizing spatial gene expression heterogeneity in spatially resolved single-cell transcriptomic data with nonuniform cellular densities. *Genome Res* (this issue) **31**: 1843–1855. doi:10.1101/gr.271288.120
- Nieto P, Elosua-Bayes M, Trincado JL, Marchese D, Massoni-Badosa R, Salvany M, Henriques A, Nieto J, Aguilar-Fernández S, Mereu E, et al. 2021. A single-cell tumor immune atlas for precision oncology. *Genome Res* (this issue) **31**: 1913–1926. doi:10.1101/gr.273300.120
- Ohnuki H, Venzon DJ, Lobanov A, Tosato G. 2021. Iterative epigenomic analyses in the same single cell. *Genome Res* (this issue) **31**: 1819–1830. doi:10.1101/gr.269068.120
- Slaidina M, Gupta S, Banisch TU, Lehmann R. 2021. A single-cell atlas reveals unanticipated cell type complexity in *Drosophila* ovaries. *Genome Res* (this issue) **31**: 1938–1951. doi:10.1101/gr.274340.120
- Wang J, Roeder K, Devlin B. 2021a. Bayesian estimation of cell type-specific gene expression with prior derived from single-cell data. *Genome Res* (this issue) **31**: 1807–1818. doi:10.1101/gr.268722.120
- Wang S, Lee MP, Jones S, Liu J, Waldhaus J. 2021b. Mapping the regulatory landscape of auditory hair cells from single-cell multi-omics data. *Genome Res* (this issue) **31**: 1885–1899. doi:10.1101/gr.271080.120
- Xu J, Zhang P, Huang Y, Zhou Y, Hou Y, Bekris LM, Lathia J, Chiang C-W, Li L, Pieper AA, et al. 2021. Multimodal single-cell/nucleus RNA sequencing data analysis uncovers molecular networks between disease-associated microglia and astrocytes with implications for drug repurposing in Alzheimer's disease. *Genome Res* (this issue) **31**: 1900–1912. doi:10.1101/gr.272484.120