Molecules and Cells

Editorial

Single-Cell Analysis: Technology, Data Analysis, and Applications

Daehee Hwang*

School of Biological Sciences, Seoul National University, Seoul 08826, Korea
*Correspondence: daehee@snu.ac.kr
https://doi.org/10.14348/molcells.2023.0020
www.molcells.org

The volume of publications on single-cell analysis has recently increased, with thousands of articles being published every year. In contrast to bulk analysis, which averages the signals from different cell types, single-cell analysis has enabled us to unveil the cellular heterogeneity present among different cell types and even within isogeneic cell populations. This progress has been driven by the development of single-cell isolation, barcoding, and sequencing techniques, but also of computational methods allowing the analysis of tremendous amounts of single-cell level data that have unique features, such as zero inflation (dropouts), over-dispersion and different count distributions from bulk RNA sequencing (RNA-seq) datasets. Single-cell analysis has led to the discovery of new cellular phenomena and enabled the investigation of disease microenvironments, thereby identifying new cellular and molecular targets for the diagnosis and treatment of refractory diseases.

Several review papers summarizing single-cell technologies, their applications, and/or data analysis methods have been published. However, new technologies and data analysis methods as well as innovative applications of these technologies continue to emerge. The present special issue on single-cell analysis includes five review papers that introduce the latest advances in technology, application, and data analysis.

Two of these reviews summarize new technologies for single-cell multiomic and epigenomic analyses. Ik Soo Kim first presents single-cell molecular barcoding methods for decoding multimodal information (e.g., DNA, RNA, and protein), which defines precise cellular states. Molecular barcoding technologies allow to extract multimodal information at the single-cell level by inserting barcodes into the molecules of isolated cells and separating each cell from the others. The review also summarizes recently developed single-cell multiomic approaches to acquire the genome, epigenome, and protein profiles simultaneously with the transcriptome. In particular, it focuses on methods used to anchor or tag molecules from a cell and improve throughputs with sample multiplexing. Furthermore, it discusses the future developments of these technologies. Kim and Lee review single-cell epigenomic technologies. The genome is almost identical in all cells of one's body. However, the functions and morphology of each cell are different, which are determined by the genes and proteins expressed in the cell. Over the last decade, technologies have been developed to extract epigenetic information, such as DNA methylation, histone modifications, chromatin accessibility, and chromatin conformation, at the single-cell level. Single-cell epigenetic analysis improves the understanding of how epigenetic factors contribute to gene regulation in individual cells and how they differ from cell to cell.

Two manuscripts in the present special issue describe recent data analysis methods developed for key informatics tasks, namely (1) cell-type deconvolution to unfold bulk data by single-cell information and (2) integration of multiple single-cell datasets. Im and Kim provide a comprehensive overview of cell-type deconvolution methods and their applications for characterizing tumor microenvironments. Since single-cell technologies are still not cost effective, scientists

Received January 17, 2023; accepted January 24, 2023; published online February 27, 2023

eISSN: 0219-1032

have developed many cell-type deconvolution methods to delineate cellular compositions from bulk transcriptome data. Im and Kim's review categorizes 20 cell-type deconvolution techniques according to three primary criteria (characteristics of the methodology, use of prior knowledge of cell types, and outcome of the methods) and highlights the advantage of the recent methods using probabilistic models. It also presents two scenarios of the common application of the cell-type deconvolution methods to study tumor microenvironments. Ryu et al. review the methods developed for integrating multiple single-cell RNA-seq datasets. The occurrence of a batch effect while analyzing different datasets is inevitable due to the differences in cell isolation and handling protocols, library preparation technologies, and sequencing platforms. Several methods have been developed to remove these batch effects for the effective integration of multiple single-cell RNA-seq datasets. These methods are based on diverse concepts and approaches, which are proven useful to examine whether cellular features (cell subpopulations and marker genes) identified from a certain dataset are consistently present or whether their condition-dependent variations are consistently observed in different datasets generated under similar or distinct conditions. Ryu et al. summarize the concepts and approaches of the integration methods and discuss their advantages and limitations. Thus, these two reviews provide a guideline for selecting the appropriate method for cell-type deconvolution and integrating single-cell RNA-seq datasets.

Finally, Jung and Lee review single-cell genomics methods for investigating the pathogenesis of inflammatory diseases, including coronavirus disease 2019 (COVID-19) infection. Unbiased profiling of single cells under inflammation conditions provides an understanding of immune networks activated in heterogeneous cell populations. Although single-cell profiling technologies, such as flow cytometry, have been used to evaluate protein expression levels of a few dozen antigens, recent technical advances enable unbiased transcriptomic and epigenetic analyses of each cell. In this context, the human cell atlas, the most comprehensive reference map of the molecular state of cells, has been established using methods for high-throughput single-cell profiling. Immunologists have actively used single-cell genomic technologies to rapidly and comprehensively investigate the pathogenesis of COVID-19 during the recent pandemic. Jung and Lee introduce recent progress in understanding inflammatory diseases by applying single-cell genomic technologies and propose new directions for future single-cell genomic studies.

In summary, the present special issue is a collection of two reviews on emerging technologies, two reviews on data analysis methods, and one review on the application of single-cell analysis to inflammatory diseases. It provides a broad spectrum of (1) key problems and solutions in single-cell technology and data analysis, (2) a guideline for selecting the appropriate technologies and data analysis methods, and (3) new directions for single-cell genomic studies. Therefore, this special issue offers unique opportunities for readers to become familiar with single-cell analysis.

CONFLICT OF INTEREST

The author has no potential conflicts of interest to disclose.

ORCID

Daehee Hwang https://orcid.org/0000-0002-7553-0044