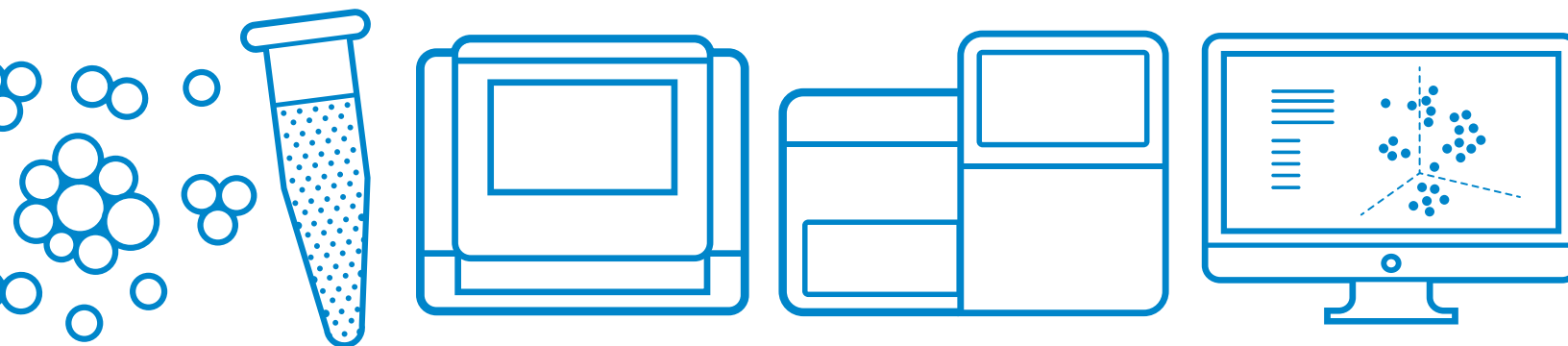


# Getting Started with Single Cell Gene Expression

## Experiment Planning Guide



Sample Type  
& Preparation

Library Preparation

Sequencing

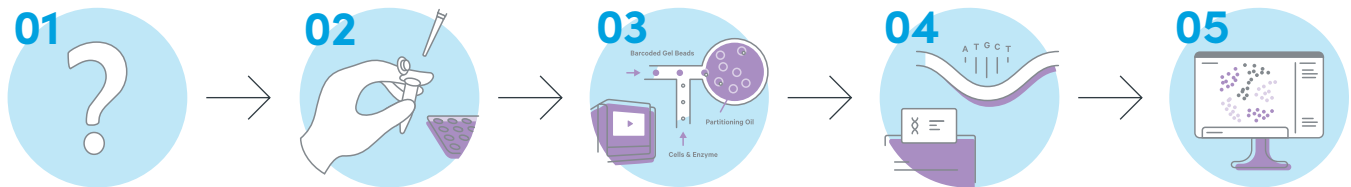
Data Analysis

# Preparing for single cell gene expression experiments

**Go beyond traditional gene expression analysis to more deeply characterize cell populations, including individual cell types and states in heterogeneous tissues.**

Differences in gene expression across organisms, tissues, and disease states have historically been quantified using approaches such as in situ hybridization (ISH), microarrays, or total RNA sequencing (RNA-seq). However, these methods can be low throughput, in the case of ISH, or provide only an average readout across cell populations. Single cell transcriptomic technologies, including Chromium Single Cell Gene Expression, allow the direct measurement of gene expression at the single cell level to quantify cell population heterogeneity and characterize cell types, cell states, and dynamic cellular transitions cell by cell.

This guide will help you get started with your single cell gene expression experiments and should serve as a roadmap to help you design your experiments, optimize experimental parameters, and identify appropriate computational/analytical tools to analyze your single cell gene expression data.



Prior to starting your single cell experiments, we recommend you walk through a five-step process that will help guide your experimental design and determine how to best answer your research questions.

**01 What scientific questions do I want to answer?**

**02 What are the best practices for preparing and processing my sample?**

**03 How many cells and replicates do my experiments require?**

**04 What depth of sequencing do I need?**

**05 How do I analyze and visualize my data?**

---

## 01 What scientific questions do I want to answer?

Single cell gene expression analysis can provide answers to many different types of research questions, including but not limited to:

**How heterogeneous is my cell population?** See publications: [3](#), [4](#), [7](#), [10](#), [12](#)

**What novel cell types and states reside in my tissue?** See publications: [1](#), [2](#), [8](#)

**Are there low abundance transcripts driving cell states or differentiation?** See publications: [7](#), [14](#)

**How do cells respond to therapy or contribute to stratification of disease state?** See publications: [5](#), [11](#), [13](#)

---

## 02 What are the best practices for preparing and processing my sample?



### Sample type & preparation

It is critical that you obtain a clean single cell suspension free of cell debris, with minimal cell aggregates and high viability (>70%). It is also important to know the size range of the cells studied. The cell size is usually correlated with the number of transcripts expressed in the cell. A wide range of cell sizes (up to 30  $\mu\text{m}$ ) are compatible with Chromium Single Cell Next GEM Chips. In general, cell preparation protocols will vary depending on the tissue's origin and the cell types being studied. Each tissue type is unique, and thus, it is critical to optimize sample preparation before starting any single cell experiment. We recommend beginning by reviewing our [Cell Preparation Guide](#). Find more answers to common questions about sample prep on our Sample Prep FAQs page. We also have sample prep-focused webinars available for viewing in our [Videos Library](#).



### Sample procurement and storage

If samples can be processed immediately, using fresh tissue and cell isolation is preferred. If there must be a delay between collection and processing, [freezing may be appropriate](#). In either case, fast processing or preservation is critical to maintain sample biology. If the sample needs to be shipped to another site, the best method is to cryopreserve the cells and [ship on dry ice](#). Refer to this [Demonstrated Protocol](#) for guidance on tissue dissociation.



### Cell enrichment

When characterizing rare cell populations or samples with low viability, enriching for live cells of interest prior to generating single cell partitions can help ensure adequate numbers of your cells of interest. The Chromium system is compatible with [FACS](#) and bead- or column-based enrichment methods. See our [Demonstrated Protocol](#) for additional guidance on cell enrichment.



### Cells versus nuclei

Whether to use cells or nuclei as input depends on a few factors, including sample type and sample handling logistics. If fresh tissue can be obtained, cell isolation is preferred. For frozen tissue or archival samples, nuclei must be isolated directly following the guidance in our [Demonstrated Protocol](#). Alternatively, tissue can be dissociated and cryopreserved immediately after it is received, using this [Demonstrated Protocol](#), enabling cells to be stored long term.

### Sample preparation best practices

- Optimize sample preparation first
- Maintain single cell viability of greater than 70%
- Ensure cell or nuclei suspensions are clean and free of debris
- Count cells after all cleanup, sorting, and enrichment steps are done

## 03 How many cells and replicates do my experiments require?

### Number of cells

Deciding on the number of cells required depends on the expected cellular heterogeneity in the sample, the number of cells available, the minimum frequency expected of desired subpopulations, and the minimum number of cells of each cell type desired for data analysis (see [online tool](#)). If the sample diversity is not known, a high number of cells at low sequencing depth may be the most flexible option to obtain a representative proportion of the cell population and meaningful biological information. Often, greater cell number, rather than sequencing depth, improves cell classification ability (Ding et al. 2020). For highly heterogeneous samples, thousands of cells may be required to resolve each subpopulation fully.

### Number of replicates

Determining the number of replicates depends on the research project, the type of sample, and the number of cells required in the study. The matter of biological replicates is still an [open question](#) in the field. In some studies, one sample alone may be sufficient—with each cell representing a biological replicate and different samples from different individuals accounting for the variability of a particular biological process. In other studies, to mitigate biological variability in small cell populations across time, it can be beneficial to computationally aggregate cells from different samples to cover all aspects of the cell population being studied. Other cases may require that multiple replicates be derived from a single sample to increase the study's total number of cells.

### Batch effects

Batch effects can be introduced at any stage of the workflow and are primarily due to logistical constraints that result from different preparation times, operators, and handling protocols. The 10x Genomics Chromium system demonstrates minimal technical variability across a variety of [technical replicates](#). A number of computational tools, including [Seurat](#), [scrn](#), and [scrone](#), can correct batch effects. Cell Ranger can also perform batch correction for libraries generated with multiple versions of our Single Cell 3' Gene Expression chemistry, as described [here](#).

## Choosing the right single cell gene expression product for your experiment

### Single Cell Gene Expression LT Kit

- Target 100–1,000 cells per sample
- Obtain preliminary data for pilot studies or grant applications
- Optimize experimental design or sample preparation
- Perform studies with low cell throughput needs

### Single Cell Gene Expression Kit

- Target 500–10,000 cells per sample
- Characterize complex cell populations and tissues
- Discover rare cell types
- Analyze precious samples with limited cell number
- Perform single cell CRISPR screens
- Address most research questions

---

## 04 What depth of sequencing do I need?

The sequencing depth per experiment for gene expression libraries is dependent on total mRNA content in individual cells, and the diversity of mRNA species. In general, at the same transcript diversity, cells expressing a low amount of mRNA will require less sequencing depth than cells expressing a large amount of mRNA. When sequencing cost or capacity is limiting, there is often a [trade-off](#) between [breadth versus depth](#), i.e. sequencing a higher number of cells at low depth versus sequencing a lower number of cells with more reads.

10x Genomics single cell libraries are compatible with short-read sequencers and are available in a dual indexing configuration. Our single cell gene expression workflows use unique molecular identifiers (UMIs) to barcode each transcript molecule before amplification takes place, resulting in a digital gene expression profile that accounts for PCR amplification bias.

---

## 05 How do I analyze and visualize my data?

### Process and analyze sequencing data

Single Cell Gene Expression from 10x Genomics comes with free, easy-to-use software for data analysis and visualization. Cell Ranger is our suite of analysis pipelines that will align reads, filter, count barcodes and UMIs, generate Feature Barcode matrices, and perform clustering and gene expression analysis to turn your raw sequencing data into results. Cell Ranger can aggregate outputs from multiple experiments, normalize to the same sequencing depth, and re-analyze the combined data. Cell Ranger will also provide actionable quality control metrics for your data, letting you know how clean your sample was and how robust your results are. The primary output of Cell Ranger is a count matrix, consisting of columns for every cell barcode and rows for all measured features, such as genes for transcriptome analysis. For each feature, the number of unique molecules detected is provided as a digital readout. Data processed by Cell Ranger can then be exported to third-party analysis tools or visualized directly in Loupe Browser, our visualization software. Cell Ranger can be downloaded and run as a command-line tool in a Linux environment or in the 10x Genomics Cloud using an intuitive point-and-click interface.

### 10x Genomics Cloud Analysis

- Process data for free\* using Cell Ranger analysis pipelines with [Cloud Analysis](#)
- Start and manage your analyses through a simple web interface
- Get your results quickly with a fast and scalable cloud platform
- Available in the United States only

**See if Cloud Analysis is [right for you](#)**





\*See 10x Genomics Cloud [Terms of Use](#) for restrictions and details

### Interactively explore single cell results

Loupe Browser is a desktop application designed for quick, interactive single cell data visualization. Built to accelerate data exploration, Loupe Browser allows you to determine cell types, discover rare cell populations, explore biological substructure, and identify new marker genes ([see online tutorial](#)). A number of community-developed analysis tools can be used to extend your data analysis. Data can be exported from Cell Ranger and imported into commonly used tools such as [Seurat](#), [Scanpy](#), [Monocle](#), or [Bioconductor](#). For an overview of single cell analysis best practices, see Luecken and Theis (2019).

## Use case examples

| Publication   |  Experiment Snapshot   |  Aim   |  Impact  |
|---|---|---|---|
| <p><b>Altered human oligodendrocyte heterogeneity in multiple sclerosis.</b></p> <p>Jäkel S, et al. <i>Nature</i> 566: 543–547, 2019.</p>   | <p><b>Research area:</b> Neurodegenerative disease</p> <p><b>Organism:</b> Human</p> <p><b>Sample type:</b> Nuclei from brain white matter</p>  | <p>Examine the role of oligodendrocyte heterogeneity in severity of multiple sclerosis (MS)</p>   | <p>Identified subclusters of oligodendroglia in control human white matter that were underrepresented or enriched in MS tissue, suggesting altered oligodendroglial heterogeneity may impact disease severity.</p>  |
| <p><b>A cell atlas of human thymic development defines T cell repertoire formation.</b></p> <p>Park JE, et al. <i>Science</i> 367: eaay3224, 2020.</p>  | <p><b>Research area:</b> Immunobiology</p> <p><b>Organism:</b> Human and mouse</p> <p><b>Sample type:</b> Cells from thymic spanning fetal to adult development, and bone marrow and spleen</p> | <p>Identify all cell types and states in the developing and adult human thymus</p>  | <p>Characterized over 50 cell states in the human thymus, changes in cell type abundance, and gene expression profiles across development. Novel populations were mapped in situ using unique marker gene signatures identified with single cell RNA-seq and TCR sequencing plus pseudotime analysis revealed a bias in V(D)J usage that correlated with developmental stage.</p> |
| <p><b>Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer.</b></p> <p>Stewart CA, et al. <i>Nat Cancer</i> 1: 423–436, 2020.</p> | <p><b>Research area:</b> Targeted therapies</p> <p><b>Organism:</b> Human</p> <p><b>Sample type:</b> Cells from xenograft tumors and circulating tumor cells isolated from blood</p>            | <p>Study molecular and cellular responses to chemotherapy</p>   | <p>Observed an increase in intratumoral heterogeneity following treatment resistance, ruling out the effects of biological replicates and the xenograft model with the resolution provided by single cell sequencing. Findings were recapitulated using limiting numbers of circulating tumor cells collected directly from patient blood.</p>                                    |
| <p><b>Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma.</b></p> <p>Wang R, et al. <i>Nat Med</i> 27: 141–151, 2021.</p>                        | <p><b>Research area:</b> Cancer biology</p> <p><b>Organism:</b> Human</p> <p><b>Sample type:</b> Malignant ascites from peritoneal cavity of gastric carcinoma patients</p>                     | <p>Determine the origins and consequences of intratumoral heterogeneity</p>   | <p>Discovered that diversity in tumor cell lineage and cell state composition are key contributors to intratumoral heterogeneity by comparing single cell data to the Human Cell Landscape. Enabled conclusions about correlation of intratumoral heterogeneity and tumor cell origins with patient survival.</p>   |
| <p><b>BACH2 enforces the transcriptional and epigenetic programs of stem-like CD8<sup>+</sup> T cells.</b></p> <p>Yao C, et al. <i>Nat Immunol</i> 22: 370–380, 2021.</p>                                     | <p><b>Research area:</b> Immunotherapies</p> <p><b>Organism:</b> Mouse</p> <p><b>Sample type:</b> Spleen</p>  | <p>Characterize the molecular mechanisms responsible for CD8<sup>+</sup> T-cell differentiation and cell state during chronic infection</p> | <p>Compared control, BACH2 overexpressing, and BACH2 deficient CD8<sup>+</sup> T cells in mice after infection with lymphocytic choriomeningitis virus (LCMV) and identified single cell gene expression signatures that suggest BACH2 establishes a stem-like transcriptional program in a cell-intrinsic manner and prevents cell differentiation.</p>                          |

|  <b>Sample prep</b>   |  <b>Library prep</b>  |  <b>Sequencing</b>  |  <b>Analysis tools</b>  |
|--|--|--|--|
| <p>Nuclei isolated from white matter areas of fresh-frozen post mortem human brain</p> <p>20 samples total, consisting of 5 control individuals with 1 sample each, and 4 MS patients with up to 5 regions sampled for each patient</p>  | <p>Chromium Single Cell 3' kit v2</p> <p>8,000 nuclei targeted per sample</p> <p>20 libraries</p>  | <p>17,799 total nuclei</p> <p>1,096 mean genes per nucleus</p>   | <p>Seurat2: nuclei clustering and canonical correlation analysis between control and MS samples</p>  |
| <p>Cells were sorted based on marker gene expression, with even numbers of CD45<sup>+</sup> and CD45<sup>-</sup> cells sampled to enrich for stromal population</p> <p>34 samples total, consisting of 15 embryonic and fetal thymi, 9 postnatal thymi, and 10 matched bone marrow or spleen samples</p>                 | <p>Chromium Single Cell 3' kit v2</p> <p>Chromium Single Cell 5' kit</p> <p>Chromium Single Cell V(D)J Enrichment kit, Human T cell</p> <p>8,000 cells targeted per sample</p> <p>13 3' Gene Expression libraries</p> <p>17 5' Gene Expression libraries</p> <p>17 TCR libraries</p> | <p>255,901 total cells</p> <p>Illumina HiSeq 4000</p> <p>Up to 8 libraries were multiplexed and sequenced together, distributed over 4 lanes</p>   | <p>Cell Ranger: alignment and quantification of single cell data</p> <p>Scanpy: downstream analysis, including normalization, PCA, and clustering</p>  |
| <p>Human cells isolated from circulating tumor cell (CTC)-derived xenografts and blood</p> <p>19 samples total, consisting of two replicates, each of 8 different CTC-derived xenografts (CDX) grown in different mice, and CTC from patient with SCLC at 3 timepoints. Replicate data were pooled after sequencing.</p> | <p>Chromium Single Cell 3' kit v2</p> <p>Loaded 12,200 cells per sample</p> <p>19 libraries</p>  | <p>36,353 total cells, consisting of ~2,000-9,000 cells for each CDX model and 712 CTC cells</p> <p>HiSeq Illumina 4000</p> <p>Final data were downsampled to 2,000 cells per CDX model for comparison</p> | <p>Seurat: accounting for cell-cycle effects, PCA, and identification of clusters</p>  |
| <p>Peritoneal carcinomatosis (PC) cells from patients with gastric carcinoma</p> <p>15 samples total, from 10 short-term survivors and 10 long-term survivors. 5 patients were excluded due to low tumor cell recovery (&lt;50)</p>  | <p>Chromium Single Cell 3' kit v2</p> <p>15 libraries</p>  | <p>45,048 total cells</p> <p>Illumina NovaSeq6000</p> <p>All libraries sequenced together on one S2 flow cell</p>  | <p>Seurat: PCA, cell clustering, and cell type annotation</p> <p>Monocle 3: clustering analysis and construction of single cell trajectories</p> <p>inferCNV: inference of copy number variants for arm-level events and construction of a phylogenetic tree</p> |
| <p>Genetically modified P14 CD8<sup>+</sup> T cells expressing TCR that recognize LCMV were transferred to mice; mice were infected with LCMV; P14 cells were sorted from spleen 7 days after infection</p> <p>4 samples total, each with over 4,000 cells pooled from more than 3 mice each</p>                         | <p>Chromium Single Cell 3' kit v2</p> <p>4 libraries</p>   | <p>21,520 total cells</p> <p>Illumina HiSeq 3000</p>   | <p>Cell Ranger, Seurat, and Monocle: analysis of single cell data</p>  |

## References

1. Al-Dalahmah O, et al. Single-nucleus RNA-seq identifies Huntington disease astrocyte states. *Acta Neuropathol Commun* 8:19, 2020.
2. Gillen AE, et al. Single-cell RNA sequencing of childhood ependymoma reveals neoplastic cell subpopulations that impact molecular classification and etiology. *Cell Rep* 32: 108023, 2020.
3. Huang W, et al. Origins and proliferative states of human oligodendrocyte precursor cells. *Cell* 182: 594–608.e11, 2020.
4. Jäkel S, et al. Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* 566: 543–547, 2019.
5. Kim D, et al. Targeted therapy guided by single-cell transcriptomic analysis in drug-induced hypersensitivity syndrome: a case report. *Nat Med* 26: 236–243, 2020.
6. Luecken MD & Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol* 15: e8746, 2019.
7. Nehar-Belaid D, et al. Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat Immunol* 21: 1094–1106, 2020.
8. Park JE, et al. A cell atlas of human thymic development defines T cell repertoire formation. *Science* 367: eaay3224, 2020.
9. Richards LM, et al. Gradient of Developmental and Injury Response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity. *Nat Cancer* 2: 157–173, 2021.
10. Schlesinger Y, et al. Single-cell transcriptomes of pancreatic preinvasive lesions and cancer reveal acinar metaplastic cells' heterogeneity. *Nat Commun* 11: 4516, 2020.
11. Stewart CA, et al. Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. *Nat Cancer* 1: 423–436, 2020.
12. Wang R, et al. Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. *Nat Med* 27: 141–151, 2021.
13. Wang S, et al. Anti-human TREM2 induces microglia proliferation and reduces pathology in an Alzheimer's disease model. *J Exp Med* 217: e20200785, 2020.
14. Yao C, et al. BACH2 enforces the transcriptional and epigenetic programs of stem-like CD8+ T cells. *Nat Immunol* 22: 370–380, 2021.

Contact us  
[10xgenomics.com](https://10xgenomics.com) | [info@10xgenomics.com](mailto:info@10xgenomics.com)

© 2021 10x Genomics, Inc. FOR RESEARCH USE ONLY. NOT FOR USE IN DIAGNOSTIC PROCEDURES.  
LIT000027 - Rev E - Brochure - Getting Started with Single Cell Gene Expression

