Best practices guidelines for single cell RNA-seq analysis using the 10x Genomics Chromium system

NGS Platform and IBU, v.1.0, 9.5.2019

General info

We currently offer the following two protocols:

1) Single cell gene expression (GEX) profiling (3' tags)

   Full protocol:
   https://assets.ctfassets.net/an68im79xiti/51xGuiJhVKOeIIceW88gsQ/1db2c9b5c9283d183ff4599fb489a720/CG000183_ChromiumSingleCell3__v3_UG_Rev-A.pdf

   Workflow summary:
   Cells in tubes on ice→ NGS lab → (Ideally Cell Viability Assay. Currently not implemented) → GEM generation on 10x Chromium controler (6-7min) → (storage on ice possible up to 1 hour) → RT and cDNA amplification → (possible to freeze and store for max. 1 week) → Illumina DNA Nano library → Sequencer: 91bp (insert) + 26bp (10x barcode and UMI) + 8bp (Index I7) = 125 cycles. Customers will be charged for 100 cycles only.

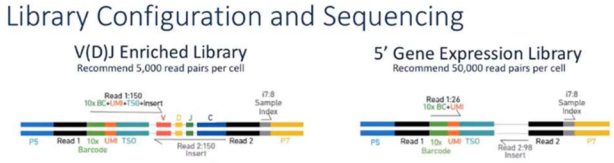
2) GEX (5' tags) profiling + VDJ (human and mouse only)

   Full protocol:
   https://assets.ctfassets.net/an68im79xiti/2IdqhhDPF6Swuocq8A8mws/a869b131a094180ec6ebd8cd5189c395/CG000086_ChromiumSingleCellV_D_J_ReagentKits_UG_Rev_J.pdf

   Workflow summary:
   Steps as above up to RT and cDNA amplification → Split into 2 parts

   Part 1) (5' GEX) Illumina DNA Nano library → Sequencing as above

   Part 2) Target enrichment (T or B cells) → VDJ amplicon library → Sequencer: 150bp (10x barcode + UMI + TSO + Insert) + 150bp (Insert)



Experimental design

Many considerations regarding experimental design are the same as for bulk RNA-seq, e.g. **replication, randomization, blocking**.

Single cell specific aspects:

1) What is important to consider when preparing samples and libraries?

The most important point for the single cell RNA seq experiments is to receive a perfect **single cell suspension without dead cells, debris and cell clumps**. The cells should be in **PBS** if possible, but there are alternatives.

Here is the link from 10x for sample prep:
https://support.10xgenomics.com/single-cell-gene-expression/sample-prep

The cell stock concentration should be between **700 and 1200 cells/ul** for optimal performance. The correct cell number is very important.

If possible, the customers should provide enough cells to have a **backup** sample because some samples (ca. 10%) fail at the first step (GEM generation) due to unknown problems. If there are enough cells left, we can repeat immediately. Therefore, if the problem was due to the chip it should work then, if there are cell clumps or other confounders it will not work again.

Batches: The Chromium system can process 8 samples at the same time (but it is no problem to do fewer than 8). When there are more than 8 samples, it becomes important to consider which libraries are prepared together. 1 batch takes ca. 2.5h of lab work so a maximum of 3 batches can be done per day.

Specific questions customers may have:

Can cells be stored?	No
Are there ways to work with fixed tissue?	Not established in Bern
Are there tissues/cell types that pose particular problems?	All that are not in suspension. Cell diameter must be <= 30 µm for 10x. Cells should ideally be round and not have “strange” shapes, e.g. neurons would be very difficult
How should the cells be brought to NGS platform	In PBS, on ice

Some background on dissolving tissue:
We have only worked with cell suspensions so far and have no experience with this.
There are commercial reagents for dissolving tissue but it is often unclear how this affects RNA. Also, it is unclear how quickly the transcriptome changes once a cell is no longer in its "normal" environment of neighbouring cells (Haque et al. 2017).

https://liorpachter.wordpress.com/2019/02/19/introduction-to-single-cell-rna-seq-technologies/: In fact, the formation of suspensions is a major bottleneck in the adoption of single-cell RNA-seq technology, as techniques vary by tissue and organism, and there is no general strategy.
Single nuclear RNA-seq (https://www.nature.com/articles/nmeth.4407) has proved to be important for experiments involving cells that are difficult to dissociate, e.g. human brain cells.

2) Requirements for cell quality and quantity?

- Proportion of live cells should be as high as possible
- For 10X, expected recovery is ca. 2/3 of input cells but this may vary between experiments/samples

3) Number of cells to be sequenced?

- Experiment specific. Depends e.g. on heterogeneity of input cells, minimum frequency of particular cell types of interest within the sample. Range: 500 -10K. Better not more than 5K because doublet rate increases.
- If we do not know anything about the sample, it may be best to perform a pilot study with high cell number and low sequencing depth
- Online tool to estimate how many cells we need to sample to see N cells of a particular cell type: https://satijalab.org/howmanycells

Source: Baran-Gale et al. 2018

4) Sequencing depth per cell?

- Depends on question and on number of cells that will be sequenced
- For GEX: ~50,000-100,000 reads/cell is now widely regarded as sufficient for most applications. ~1M reads per cell effectively means saturation
- For VDJ: min 5K per cell

Sources: Baran-Gale et al. 2018, Slides Tim, 10X recommendations

5) Costs

- Library prep:
  GEX: 2000 CHF per tube
  VDJ: 2000 + 500 CHF per tube = 2500 CHF
  VDJ only: 2000 CHF

- Sequencing:
  Required 100 cycle kit which is available only for NovaSeq SP, S1 and S2 flowcells. Total number of reads should be calculated with the goal of 100K reads/cell.
  For details see price list of NGS platform.

Some general notes

- Smart-seq2 could be of interest when full transcript is needed and/or when it is more important to have good coverage of the transcripts within each cell rather than having many cells. We cannot provide support with this.

Novel approaches that may be interesting

- Demuxlet - Haplotype-based deconvolution of cells: Pool multiple samples per barcode and demultiplex based on genotype. Requires genotype info for each sample. Kang et al. 2017 Nature Biotechnology 36: 89-94.
- Sci-RNA-seq / SPLiT-seq: Methods that rely on split pool combinatorial indexing (not sure how this works) and increase number of cells assayed to > 10K (see e.g. Chen et al. 2018)
- Review of other single-cell approaches: Packer & Trapnell 2018 Trends in Genetics: 34(9):653. For example: CITE-seq / REAP-seq where scRNA-seq is combined with analysis of cell surface proteins; chromatin accessibility; methylation.
- Overview of protocols available for Chromium: https://www.10xgenomics.com/solutions/
- Single nuclear RNA-seq (https://www.nature.com/articles/nmeth.4407)

References

Baran-Gale et al. 2018 Experimental design for single-cell RNA sequencing. Briefings in Functional Genomics 17(4): 233-239.

Chen, X., Teichmann, S.A., Meyer, K.B. (2018) From tissues to cell types and back: Single-cell gene expression analysis of tissue architecture. Annual Review of Biomedical Data Science 1: 29-51.

Haque et al. 2017 A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. Genome Medicine 9:75.

Pachter et al. 2019 Introduction to single-cell RNA-seq technologies.
https://liorpachter.wordpress.com/2019/02/19/introduction-to-single-cell-rna-seq-technologies/

Slides Tim: https://docs.google.com/presentation/d/1Ko57xMcJ2iuqnOcQ9MQnWZptq0LgeX3BgyPsNFE0AXQ/edit#slide=id.p

Other resources

http://10xqc.com/ - Quality control stats for 10X data from different labs