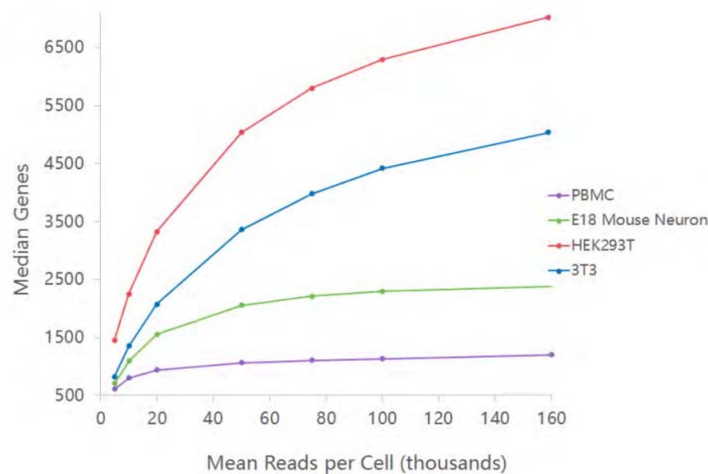
protocol v1.12

10x Genomics Chromium samples preparation guidelines for scRNA-seq

Important notes

Cell size

- Max cell diameter officially supported by 10x Genomics (10XG) is 30um, but the max theoretical limit is 65um (=size of the microcapillaries). These diameters are calculated for cells in suspension, not flattened on a petri/tissue surface.
- If your cells are at the upper limit of this size range, we can use "cheap" training beads to assess cells incorporation inside droplets with a microscope. In case your cells are too big, a special protocol exists for nuclei instead of cells. Nuclei isolation can also be useful when getting intact dissociated cells is not possible.
- Number of genes detected directly depends on the cell size (=amount of mRNA molecules per cell). Sequencing depth chosen may thus be modulated according to cell size. See plot from 10XG (for v2 reagents):



Targeted cell number

- Define the number of cells you would like to be ultimately sequenced (="targeted recovered cell number"), and indicate it to us when bringing the samples or before. Recovery rate is uncertain, in particular during first experiments with a given sample type, therefore target cells number is only indicative and number of actual recovered cells can differ significantly from it. See also comments in *Cells concentration* paragraph.
- The rate of doublets increases with targeted number of recovered cells (see table below), therefore we recommend not targeting more than 5'000 cells unless necessary.

Nb of Recovered Cells	Multiplet Rate (%)
500	~0.4%
1 000	~0.8%
2 000	~1.6%
3 000	~2.3%

4 000	~3.1%
5 000	~3.9%
6 000	~4.6%
7 000	~5.4%
8 000	~6.1%
9 000	~6.9%
10 000	~7.6%

- Each additional recovered cell comes with a sequencing cost, therefore calculate the number of cells needed thoughtfully. Consider as well that once the sample has been processed, it is not possible to sequence only a fraction of the recovered cells.
- The ideal number of targeted cells vary greatly with the biological question. For comparing transcriptional profiles of two already known populations of cells, 1'000 cells would be enough. At the other extreme, for identifying new rare subpopulations of cells (<1%), 5'000 or more cells may be needed. If unsure, we recommend 3'000 cells as a good starting point.

Number of reads/cell

The choice of how many reads you will want per cell can influence how many cells you would like to target, it is thus good to have a rough idea already before processing the cells. There are (at least) 3 factors that can influence the choice of number of reads/cell:

- the size of your cells. For very small cells (say 5-10um), mRNA levels are low and sequencing will saturate fast, thus 50-100'000 reads/cell capture much of the available information. In contrast very big cells (say >30um) contain a lot of mRNA and one can sequence up to 200'000 reads/cell and above and still capture new information.
- the biological question: if your question is simply to cluster cell in groups of known cell types, and/or to delineate their broad transcriptional profile or activated pathways, 50'000 reads/cells is generally enough (even for big cells). If you would like to the zoom in the different clusters and ask whether specific genes are expressed, then the more reads the better (up to complete sequencing saturation if needed, see above). If you also would like to delineate new cell type with subtle difference as compared to previously known ones, more reads are also better.
- the amount of money you are willing to spend, which depends on the number of samples and cells.

If you're not sure, 10XG recommends 50'000 reads/cells as a minimum starting point. We rather recommend a minimum 75'000 reads/cells.

Diverse important notes

- The standard protocol is intended for fresh unfixed/unfrozen cells, but a special protocol exists for methanol-fixed cells. Inquire if interested.
- If an mRNA from an ectopically-expressed transgene is highly overexpressed (GFP…) it may take a significant fraction of the reads, therefore take that into account when designing your experiment.
- Samples must not contain toxic/carcinogenic chemicals. Contact us in advance if it is not possible.

- Primary human cells must have been screened for absence of HIV, HBV and HCV infection (ideally also hCMV), and certificate must be provided together with the cells. More generally, cells must not be infected with an infectious agent. Contact us in advance if it is not possible.
- If VDJ reconstruction is needed, tell us in advance as the whole method/kit is different.
- If you have to prepare cells in different batches, make sure you avoid batch effects. Typically, avoid processing all controls in a batch and “treated” samples in a distinct batch.
- We can split the processing of cells in the instrument in several batches if this avoids some samples waiting on ice for very long times, and if this does not introduce batch effects.
- If you plan a big experiment, it may be worth first running just a single sample (pilot experiment) to assess quality of data and cell recovery rate.
- A general overview of 10XG experimental design can be obtained here: http://go.10xgenomics.com/getting-started-guide-facebook
- If the bioinformatics analysis is to be performed at the GECF, tell us in advance in case an ectopically expressed gene must be added to the reference set of genes. In particular, tags/markers that are inserted in 3'/C-ter of endogenous ORFs will require special thoughts, as the reads will likely lie within these added parts rather than the endogenous gene, and mapping on the endogenous reads will give no hit.
- 10XG recently released their “feature barcoding technology”, which allows to measure cell surface protein levels or perform CRISPR perturbation screens along to the usual scRNA-seq quantification. More information here: https://support.10xgenomics.com/single-cell-gene-expression/overview/doc/getting-started-single-cell-gene-expression-with-feature-barcoding-technology . If interested, contact us and/or look for BioLegend TotalSeq-B if doing 3'GE, or TotalSeq-C if doing 5'GE. CAUTION: this comes with a significantly more complicated library prep method and downstream data processing, therefore ask for a quote.
- 10XG regularly update their reagents and workflows/pipelines (currently v3.1). If you absolutely want us to use a specific version to be able to directly compare with one of your previously generated dataset, tell us ahead and we'll discuss what can be done. Yet we recommend to always use the latest version as improvements are in general significant.

Cells preparation

Cells must be submitted ready for subsequent steps in the procedure, according to the guidelines below and to the latest version of 10XG Cell Prep guide (document CG00053, https://support.10xgenomics.com/single-cell-gene-expression/sample-prep/doc/demonstrated-protocol-single-cell-protocols-cell-preparation-guide). This 10XG document explains how to wash and resuspend cells.

CAUTION: obtaining nicely dissociated and alive cells is likely the most important and often most tricky step of the whole 10XG experiment. Therefore, we strongly recommend that you test and if needed optimize the dissociation procedure before the real experiment. During that test process, include any putative step where the cells will have to wait on ice before being processed by us (waiting for other samples to be dissociated, journey towards GECF…).

Viability

- The key word for the 10XG Chromium is cell viability. It is very important to get good results downstream.
- MACS-sorting in general leads to higher viability than FACS-sorting, so prefer the former if possible.
- If you sort the cells by FACS before loading, include a viability stain to get rid of dead cells if possible.

- Miltenyi sells a kit to remove dead cells if needed ("Dead Cell Removal kit", 130-090-101), which works on mammalian cells and probably also on insects cells but to be tested to be sure.
- Before loading cells in the instrument, we'll check viability of the cells.
- The percentage of dead cells that can be considered acceptable depends on your experimental settings, and to which extent it can affect your downstream data analysis. Dead cells can generally be excluded bioinformatically as they contain a high percentage of mitochondrial mRNAs, but it may not be done easily if cells have just died or are in the process of dying. If your cells come from a healthy suspension cell line, anything more than 5-10% dead cells is probably a sign that something is bad, while if working from primary cells that underwent hours of dissection and sorting, 20% dead cells may be considered acceptable. We will move forward with any sample containing less than 20% dead cells, unless the user explicitly decides otherwise. When samples contain more dead cells, it is the user decision to either move forward anyway, or perform a dead cells removal procedure (see above), or cancel the experiment and try to improve cell viability for a subsequent experiment.

Dissociation

- Cells should be well dissociated from each others, and if relevant carefully gated on FACS to avoid doublets.
- If used, trypsin should be inactivated after use (serum, BSA…).
- High EDTA is not OK here as it will impede reverse transcription, so if you use EDTA to detach cells you will have to quench it and do an additional wash.
- To avoid clumps, cells can be passed through a cell strainer. This is usually not necessary if cells are FACS-sorted. Avoid doing that if it comes with a risk of skewing your data. Some examples:
  - BD Falcon tube with 35um strainer, #352235, not ideal for small volumes.
  - Flowmi pipette cell strainer of 40um or 70um (we have both at GECF).
  - Miltenyi Biotec 30 um PreSeparation Filter, cat#130041407….
- For dissociation of difficult tissues, consult publications from the relevant field, contact 10XG tech support, look at users-developed protocols (https://community.10xgenomics.com/t5/Customer-Developed-Protocols/ct-p/customer-protocols) and/or refer to 10XG webpage containing useful web links: https://kb.10xgenomics.com/hc/en-us/articles/218169563-How-do-I-dissociate-my-tissue-of-interest , such as the Worthington Tissue Dissociation Guide. Using a GentleMACS instrument may help, as well as accutase, collagenase, pronase or even better psychrophilic proteases which work at 4° (Potter lab, PMID: 28851704).
- As a general rule, remember that expression profiles can change rapidly if dissociation is lengthy, in particular if not performed at cold temperatures.

Cell concentration

- Your cells concentration should ideally be around 700-1'200 cells /ul (incl dead cells). FACS sorters often overestimate cells concentration, therefore take some safety margin and ask the sorting facility to give you a concentration a bit higher than that. 25ul are the minimum, but the more the better since we will also have to do a trypan blue… so if possible bring >50ul or even more. If you think you cannot reach these values, please contact us, as there is room for some flexibility. Nevertheless, concentrations range outside these values will make target cell number even less predictable.
- For concentrating or washing cells, 10Xg recommends spinning at 300xg for 5min at RT° for small cells (5-

10um), 250xg for 4min for medium cells (10-17um), and 150xg for 3 min at RT° for larger cells (17-25um). This will insure efficient pelleting while minimizing pelleting of debris or aggregation of cells. When doing it the first time, keep and check also supernatant to insure these settings are working for your cells. Also keep in mind that if you have or suspect to have a sub-population composed of small cells, you should use small cells settings to avoid losing it.

- Before loading cells in the instrument, we'll double-check the cell concentration.
- Bring some extra medium to perform cells dilutions if needed.
- When cells tend to aggregate, cell concentration should be kept on the lower side of the range (aim for 700 cells/ul rather than 1200 cells/ul).

Diverse

- Cells should be brought either in PBS 0.04% BSA, or alternatively can be kept in their preferred medium (FCS is not a problem, as long as it was filtered to remove putative large particles).
- Perform one wash, as described in the 10XG cell prep guide. This wash can be performed with your culture medium. In case cell amount is too low, this wash can exceptionally be skipped. More details in the cell prep guide.
- Place cells on ice once prepared.
- BioLegend and BD offer barcoded antibodies that allow to tag and multiplex different samples into one 10XG run (aka cells hashing). This is only available for human and mice for now. If interested, contact us and/or look for BioLegend TotalSeq-A if doing 3'GE, TotalSeq-C if doing 5'GE, or BD Single-Cell Multiplexing Kits. CAUTION: this comes with a significantly more complicated library prep method and downstream data processing so use that only when strictly necessary, and ask for a quote.

Versions log

- v1.01: initial release.
- v1.02: increased the minimal volume of cells to be submitted to 25ul. Clarified that toxic chemicals or infectious agents must not be present in the samples. Mentioned Miltenyi kit for dead cells removal.
- v1.03: added plot with gene number depending on cell size. Clarified that the dead cells removal kit works on all mammalian cells.
- v1.04: Clarified that above 20% dead cells, it is the user decision to move forward, perform a dead cell removal, or cancel the experiment.
- v1.05: Clarified that "recovery rate is uncertain, in particular during first experiments with a given sample type, therefore target cells number is only indicative and number of actual recovered cells can differ significantly from it. See also comments in *Cells concentration* paragraph". Also clarified that cells concentrations range outside recommended values will make target cell number even less predictable.
- v1.06: Clarified the rules for submitting primary human cells. Clarified that the cells must arrive at the GECF ready for subsequent steps in the procedure. Added a link to the experiment planner guide from 10XG.
- v1.07: Mentioned to inactivate trypsin after use; that expression profiles can change rapidly if dissociation is lengthy, in particular if not performed at cold temperatures; that if the bioinformatics analysis is to be performed at the GECF, we should be told in advance if an ectopically expressed gene must be added to the reference set of genes.

- v1.08: Indicated that we have both 40um and 70um flowmi at GECF. Added some more information regarding cells dissociation web resources. Clarified that this protocol is specifically written for scRNA-seq, a separate one will be written soon for scATAC-seq. Mentioned "feature barcoding" and "cell hashing" technologies.
- v1.09: Clarified that "obtaining nicely dissociated and alive cells is likely the most important and often most tricky step of the whole 10XG experiment. Therefore, we strongly recommend that you test and if needed optimize the dissociation procedure before the real experiment. During that test process, include any putative step where the cells will have to wait on ice before being processed by us (waiting for other samples to be dissociated, journey towards GECF…)." . Clarified that when cells tend to aggregate, cell concentration should be kept on the lower side of the range. Corrected a typo in the cells hashing description (TotalSeq_A instead of TotalSeq B). Added further description of protein quantification (feature barcoding).
- v1.10: Added a link to 10XG users-developed protocols. Mentioned that if you have to prepare cells in different batches, make sure you avoid batch effects. Typically, avoid processing all controls in a batch and "treated" samples in a distinct batch. Clarified that if you have or suspect to have a sub-population composed of small cells, you should use small cells centrifugation settings to avoid losing it.
- v1.11: Clarified the medium in which cells should be brought at GECF, and that the 10XG cell prep document is to be read for explanations regarding washes and cell resuspension.
- v1.12: Clarified that 10XG regularly update their reagents and workflows/pipelines (currently v3.1), and that if you absolutely want us to use a specific version to be able to directly compare with one of your previously generated dataset, tell us ahead and we'll discuss what can be done. Also clarified that, ectopic tags/markers that are inserted in 3'/C-ter of endogenous ORFs will require special thoughts, as the reads will likely lie within these added parts rather than the endogenous gene, and mapping on the endogenous reads will give no hit. Clarified how to determine ideal number of targeted cells. Added insights about how to define number of reads/cells.