Miscellaneous

Types of data / special analyses

- Multiplexed single-cell RNAseq
- Cite-seq (called also Total-seq)
- VDJ single-cell RNAseq
- VDJ 5' multiplexed single-cell RNAseq
- Spatial transcriptomics (not yet single-cell, but soon!)
- Multi-omics
- Single nuclei RNA-seq

Multiplexed single-cell RNA-seq

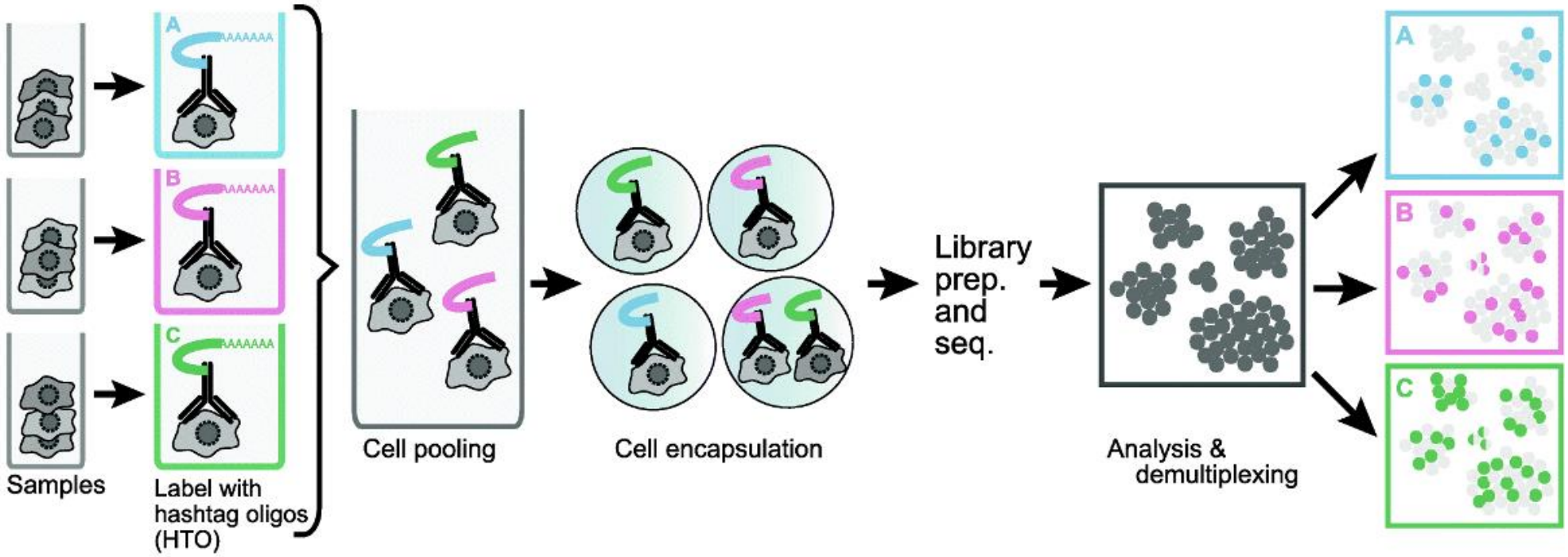
- Reduce the cost of the sequencing by using a labelling for each sample then pool different samples and then sequence.
- For the analysis, the only difference compared to standard single-cell RNA-seq data is in the cellranger part.
- Key advantages of Cell Multiplexing include:
- Increased sample throughput in a single experiment
- Increased number of cells assayed in a single experiment
- Increased number of possible replicates in a single experiment
- Detection of multiplets and their removal prior to analysis

support.10xgenomics.com/

Multiplexed single-cell RNA-seq-How it works

- Cell multiplexing oligos (sometimes called hashtag oligos) or CMO are added to the cells, one CMO per sample in a pool (this means the same CMO can be used in several different pools for different samples, and this is commonly used).
- The technique is similar to measuring cell surface proteins (that we will see with the total-seq and cite-seq methods).

Multiplexed single-cell RNA-seq-How it works



Stoeckius M et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol. Genome Biology; 2018;19:1–12.

cellranger multi --id Mysample345 --csv PATH_TO/Multi_file.csv

support.10xgenomics.com/

CSV file should contain

- The [gene-expression] section has two columns that specify parameters relevant to analysis of gene expression data, such as reference genome and cell-calling parameters, as well as other all-purpose parameters.
- The [libraries] section has three required columns that specify where the input FASTQ files may be found.
- The [samples] section has two required columns that specify sample information for Cell Multiplexing.

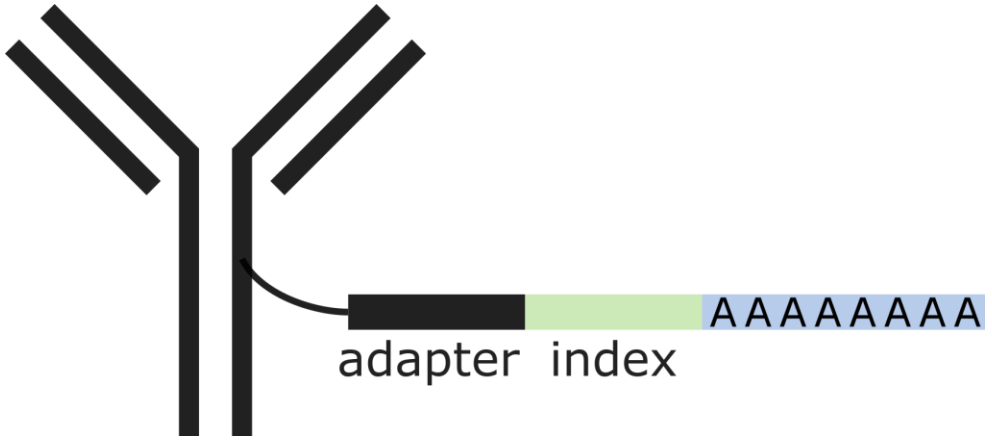
```
[gene-expression]
reference,/path/to/transcriptome

[libraries]
fastq_id,fastqs,feature_types
gex1,/path/to/fastqs,Gene Expression
mux1,/path/to/fastqs,Multiplexing Capture

[samples]
sample_id,cmo_ids
sample1,CMO301
sample2,CMO303
```

support.10xgenomics.com/

Total-seq / Cite-seq - REMINDER



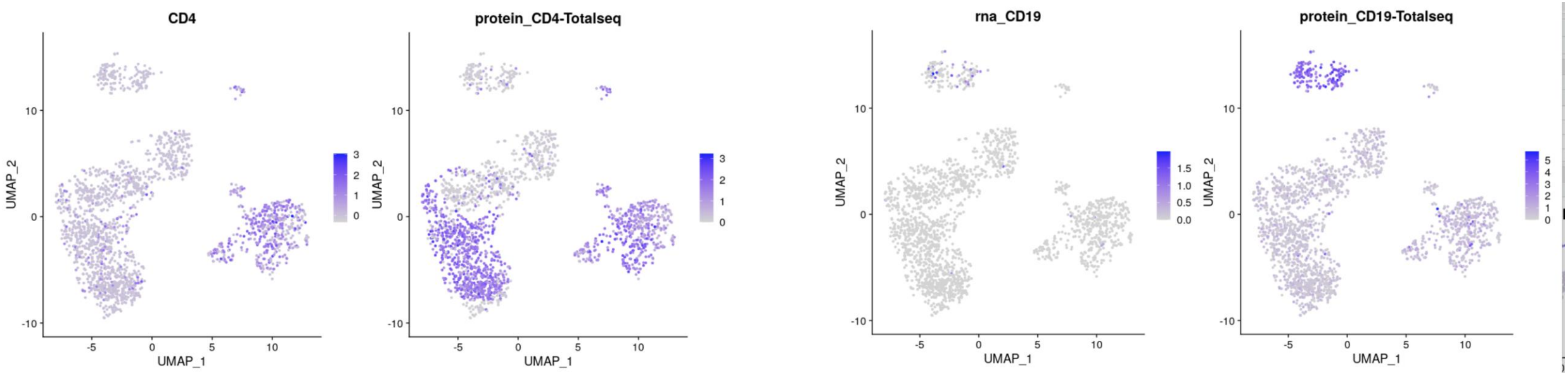
- Quantification of (cell-surface) proteins
- Together with transcriptome

Stoeckius M et al. Simultaneous epitope and transcriptome measurement in single cells. Nat Methods. 2017;14:865–8.

Total-seq / Cite-seq

- A little bit similar than multiplexing in terms of the file structure.
- This time each cell has a panel of surface proteins (or proteins) that are measured on top of the RNA-seq of each cell.
- This enables more precise annotation of cells, due to the high amount of dropouts in the RNA-seq part.

Courtesy Sarah Cattin



Total-seq cellranger

- **cellranger count --id=sample345 --libraries=library.csv --transcriptome=/path/refdata-gex-GRCh38-2020-A --feature-ref=feature_ref.csv**

library.csv file looks like this:

fastqs	sample	library_type
/opt/foo/	GEX_sample2	Gene Expression
/opt/foo/	Ab_sample2	Antibody Capture

… and feature_ref.csv like this

read	pattern
R2	5PNNNNNNNNNN(BC)

support.10xgenomics.com/

VDJ single-cell RNAseq

- **cellranger vdj --id=sample345 –fastq=PATH_TO_FASTQ --libraries=library.csv --transcriptome=/path/refdata-cellranger-vdj-GRCh38-alts-ensembl-7.0.0 --sample=sample-name**

VDJ 5' multiplexed single-cell RNAseq

- That is where the fun starts.
- Cellranger does not allow 5' multiplexed data. This is however exactly what you can do when you want to do VDJ analysis for lower cost.
- A "not too complicated" procedure is explained in this https://www.10xgenomics.com/resources/analysis-guides/demultiplexing-and-analyzing-5%E2%80%99-immune-profiling-libraries-pooled-with-hashtags article from the 10x genomics support this is for cellranger version 7 only.
- One demultiplexes using **cellranger multi** the single-cell RNAseq part (without touching the multiplexed VDJ fastqs). This generates mapped bam files for each sample.
- These bam files are returned to fastq files (bamtofastq) by making sure that only one fastq file is created.

The csv file will look like that

[gene-expression]
reference-path,/ref/refdata-gex-mm10-2020-A/
cmo-set,/data/CellRanger/feature_reference_W1.csv
[libraries]
fastq_id,fastqs,feature_types
W1_GEX,/data/fastq/nvid00011,Gene Expression
W1_FB,/data/fastq/nvid00019,Multiplexing Capture
[samples]
sample_id,cmo_ids
24_1,A0301
18_c,A0302
12_3,A0303
6_1,A0304

VDJ 5' multiplexed single-cell RNAseq

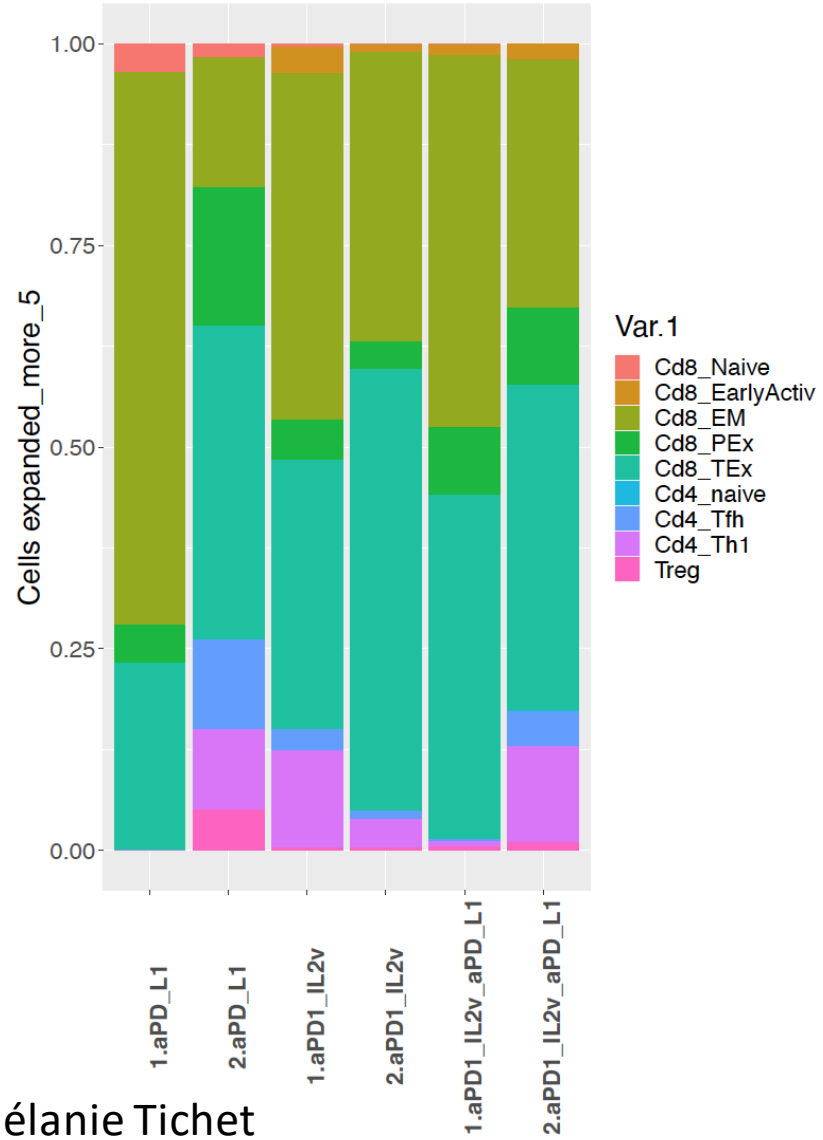
- Then the "per sample" single-cell RNAseq files are mapped back to the genome this time using the multi function with one part for the VDJ and the other for the RNAseq.
- Here again cellranger will complain as you do not have a similar amount of cells in both cases (indeed VDJ is still having all the reads from all the samples) so again here change cellranger's code and make sure this is OK.
- You then obtain VDJ and RNAseq results per sample as desired.

The csv file will look like that

[gene-expression]
reference-path,/ref/refdata-gex-mm10-2020-A/
force-cells,5000
chemistry, SC5P-R2
[vdj]
reference-path,/results/refdata-cellranger-vdj-GRCm38-alts-ensembl-5.0.0
[libraries]
fastq_id,fastqs,feature_types
bamtofastq,/results/FASTQ_per_sample/24_c/W4_GEX_0_1_HNC3HDRXY/,Gene Expression
W4_VDJ,/data/fastq/nvid00018/,VDJ

The things you can do with this

- Understand the composition, expansion and diversity of the clones in your samples.
- Get the differences in tcell composition of the expanded clones (are they naïve, memory or active t-cells?)

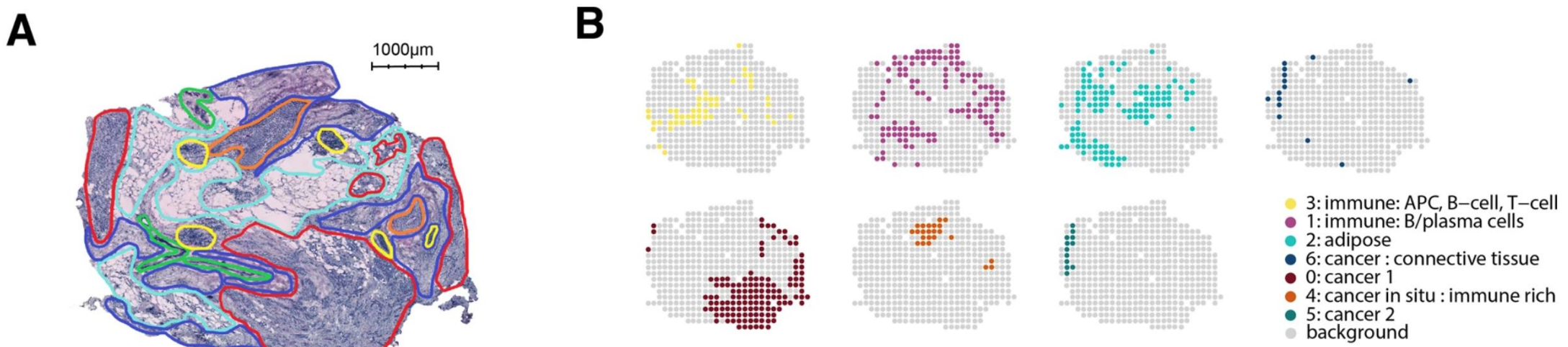


Courtesy Mélanie Tichet

Spatial transcriptomics

- On 10 micro-meter thick tissue slides you perform on spots of 55 micro-meter barcode and UMI libraries.
- This enables you to look at the resolution of 55 micro-meters the expression of cells in that spot.
- Instead of cellranger one uses for example spaceranger.
- The output files are very similar than the ones from cellranger. Except that you have another folder called spatial, with a txt file of spot and x-y coordinates of the section.
- Right now a lot of questions are still to be solved as it is not a single-cell resolution (how many cells are in one spots, how to normalise the data for example)
- Visium 10X announced that by the end of this year they will go down to 10 micro-meters of spots which then goes to a single-cell resolution.
- You can use Seurat for the analysis but also SpatialExperiment object or SingleCellExperiment object to deal with that data type.

Spatial transcriptomics example



Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions, Alma Andersson et al. Nature 2021

spaceranger count --slidefile yourslide.json

Multi-omics

- One can combine single-cell RNA-seq with single-cell ATAC-seq, having regulome and transcriptome information on similar cells. Single cell RNA-seq then helps the ATAC-seq single cell to be annotated by doing a label transfer.
- This label transfer functions using similar techniques then the one used for integration (merging k-tables).

Single Nuclei rnaseq-Reminder

- Alternative to scRNA-seq
- For tissues difficult to dissociate
- No ribosomes -> no translation of transcription factors during processing
- Lower representation of immune cells + surface proteins
- No difference in the cellranger part, in the QC filtering of cells with high mitochondrial content is unnecessary, otherwise it is treated in the same way as single-cell RNA-seq

Denisenko E, et al. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. Genome Biol.; 2020;21:1–25.