bioinformatics.ca

Canadian Bioinformatics Workshops

www.bioinformatics.ca

bioinformaticsdotca.github.io

Supported by

McGill UNIVERSITY

Creative Commons

This page is available in the following languages:
Afrikaans български Català Dansk Deutsch Ελληνικά English English (CA) English (GB) English (US) Esperanto Castellano Castellano (AR) Español (CL) Castellano (CO) Español (Ecuador) Castellano (MX) Castellano (PE) Euskara Suomeksi français français (CA) Galego עברית hrvatski Magyar Italiano 日本語 한국어 Macedonian Melayu Nederlands Norsk Sesotho sa Leboa polski Português română slovenski jezik српски srpski (latinica) Sotho svenska 中文 華語 (台灣) isiZulu

creative commons

Attribution-Share Alike 2.5 Canada

You are free:

to Share — to copy, distribute and transmit the work

to Remix — to adapt the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Share Alike. If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar licence to this one.

- For any reuse or distribution, you must make clear to others the licence terms of this work.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- The author's moral rights are retained in this licence.

Disclaimer

Your fair dealing and other rights are in no way affected by the above.
This is a human-readable summary of the Legal Code (the full licence) available in the following languages:
English French

Learn how to distribute your work using this licence

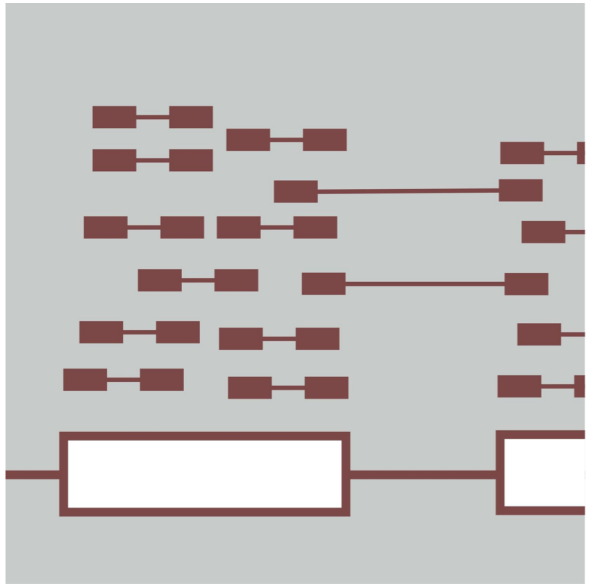
single-cell RNA-sequencing

Allegra Petti

Informatics for RNA-seq Analysis

June 17-19, 2020

bioinformatics.ca



Washington University School of Medicine in St. Louis

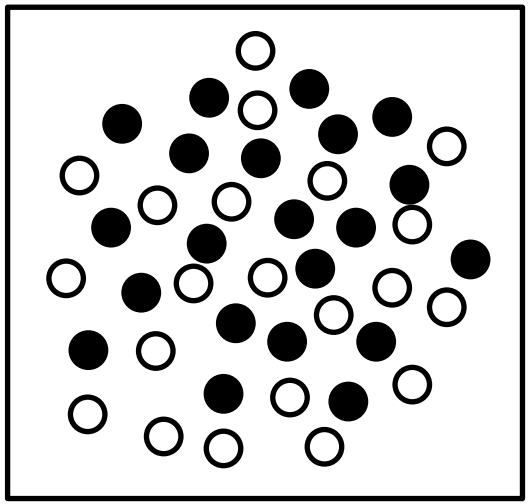
Part I: Introduction to scRNA-seq, cellranger, and the loupe browser

- Learning Objectives:
  - Understand the applications of scRNA-seq, and how it differs from bulk RNA-seq
  - Know the advantages of different scRNA-seq platforms
  - Understand the 10xGenomics technology
  - Learn the Cellranger commands for initial data processing of 10xGenomics data
  - Understand Cellranger output files
  - Learn to assess the success of your experiment
  - Use the loupe browser to perform initial data exploration

Single-cell RNA-seq captures expression heterogeneity

Identifies and counts unique transcripts in each cell

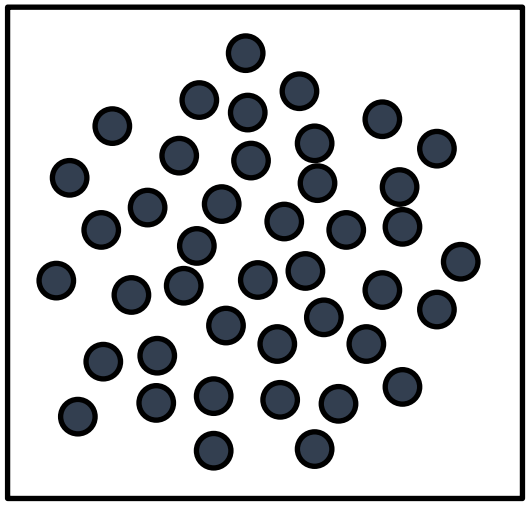
For one gene:

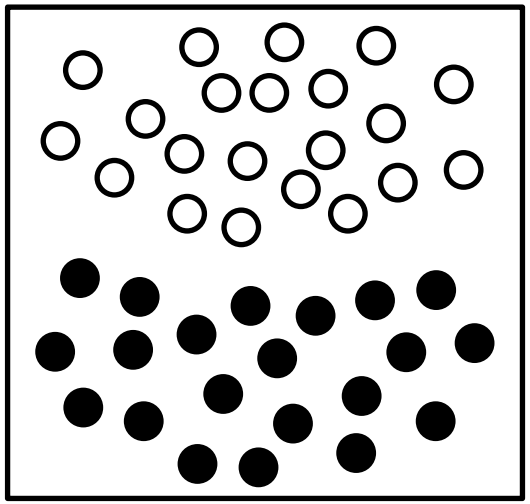


Gene X ON
Gene X OFF

Bulk

Single-cell





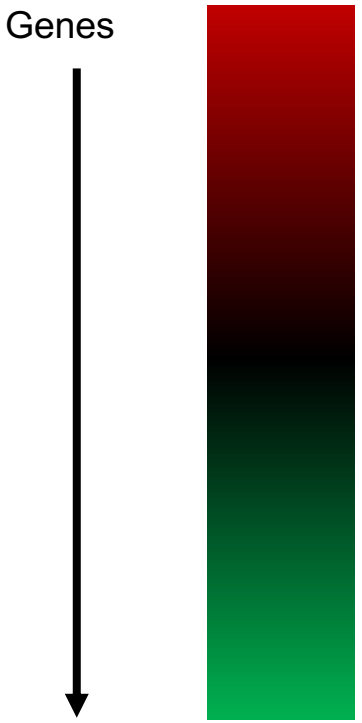
Bulk RNA-seq averages across the population

scRNA-seq reports per-cell expression and enables computational "sorting"

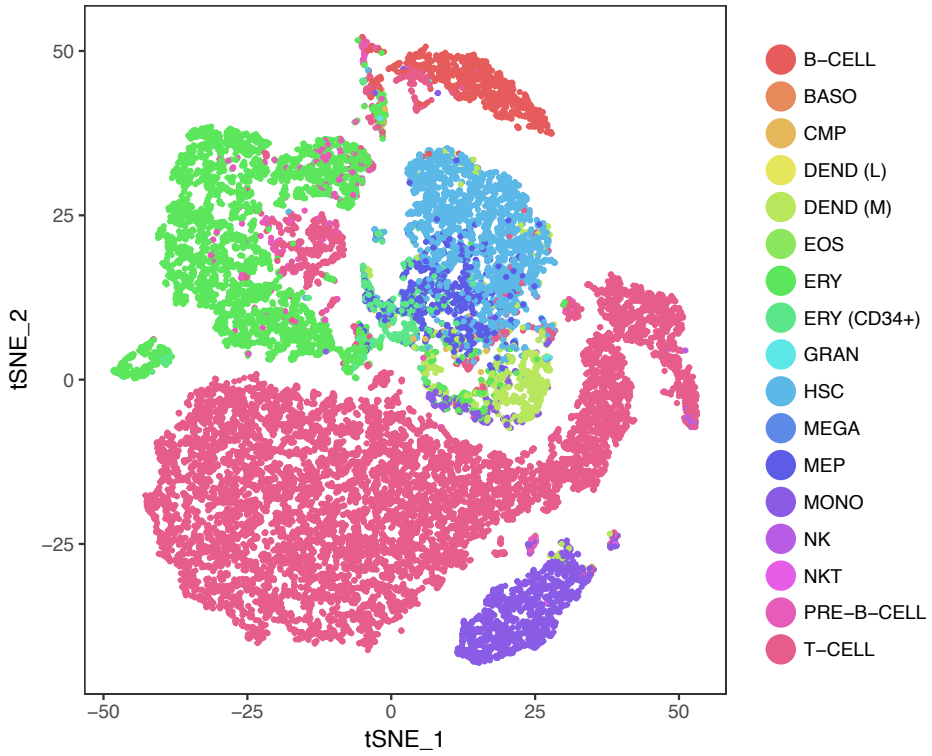
Single-cell RNA-seq captures expression (and genetic) heterogeneity

Identifies and counts unique transcripts in each cell

Bulk RNA-seq

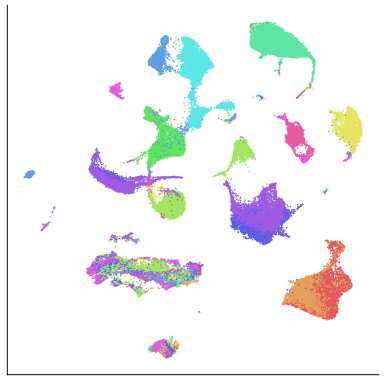


Single-cell RNA-seq
(tSNE/UMAP plot)

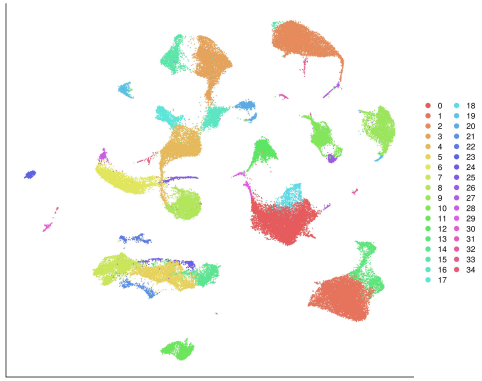


For many genes, multiple samples:

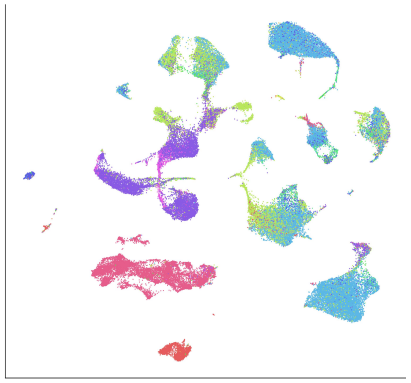
By Sample:



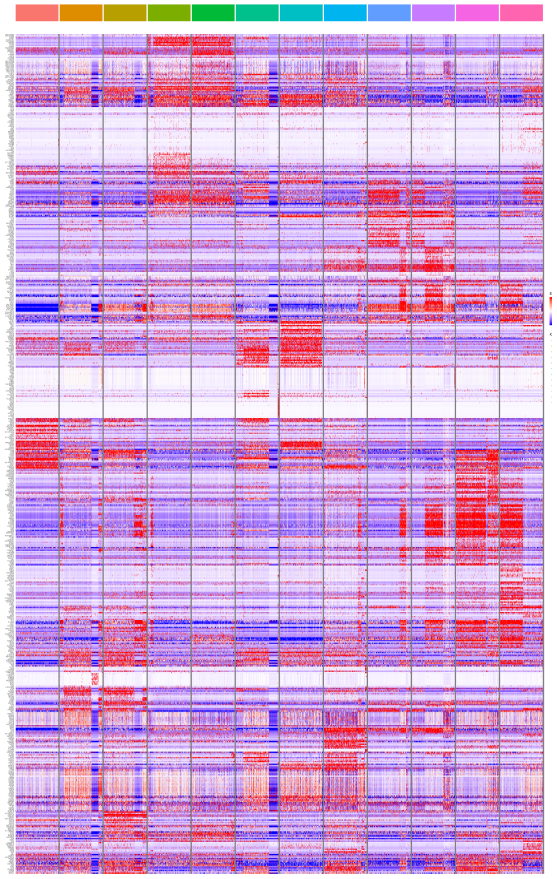
By Cluster:



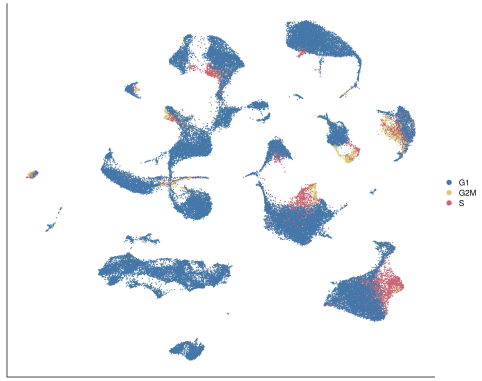
By Cell Type:



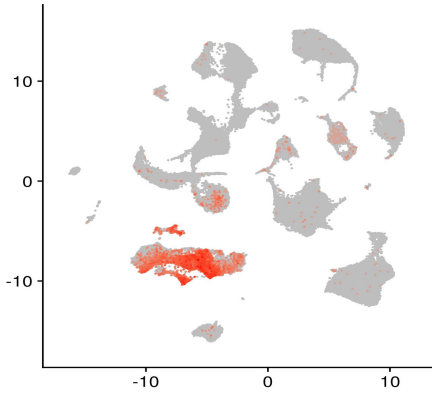
By Gene, Cell, and Sample:



By Cell Cycle Phase:



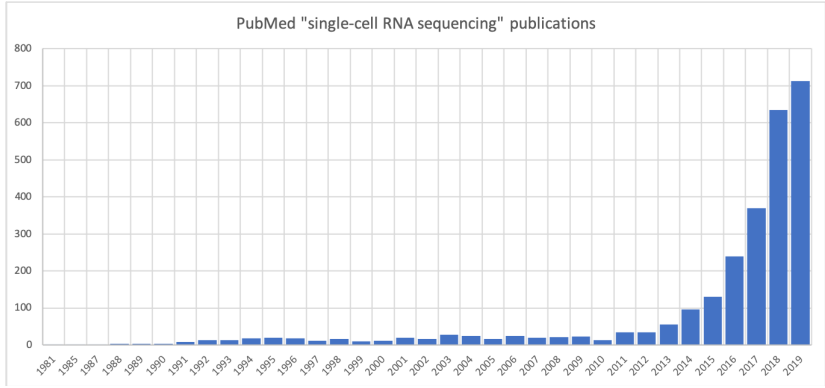
By Gene:



A new era in biology and medicine?

Community Goals

- Redefine cell “type”
- Redefine relationships among cell types
- Catalog all cell types in all diseased and normal tissues
- Discover/define new cell types



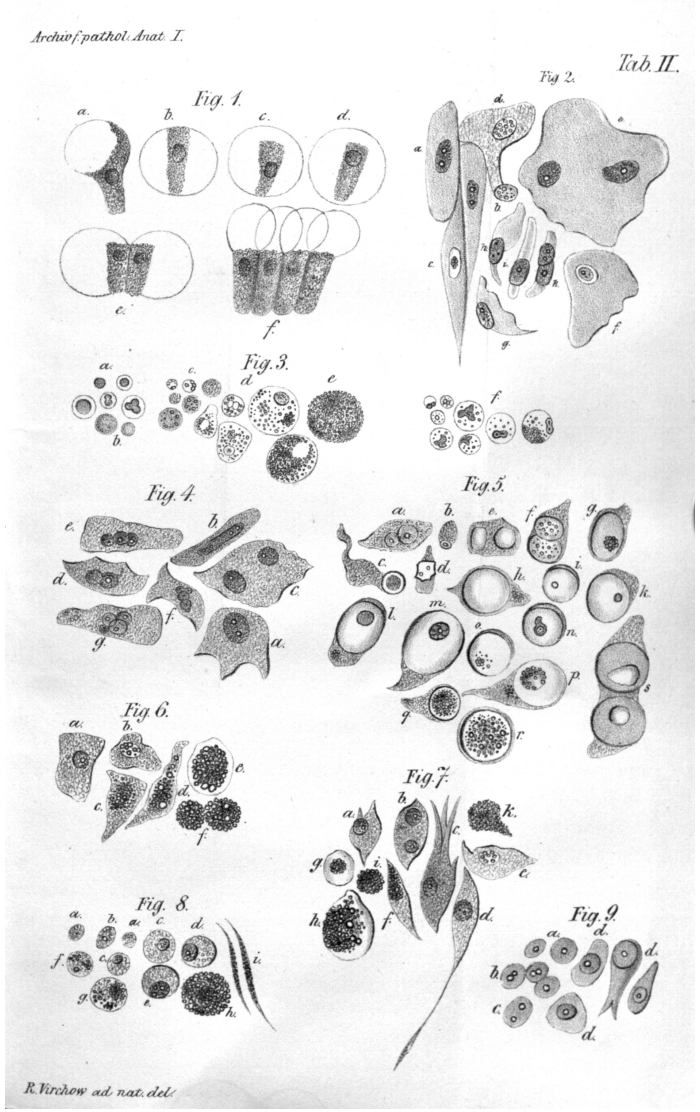
Personalized medicine

- Variation at the level of the individual - *between* individuals
- scRNA-seq: Variation at the level of the cell - *within* AND *between* individuals
  - High-resolution variation in diseased and normal cell types and states
  - Enables cross-patient correlations to be made at the level of individual cells

Why the optimism?

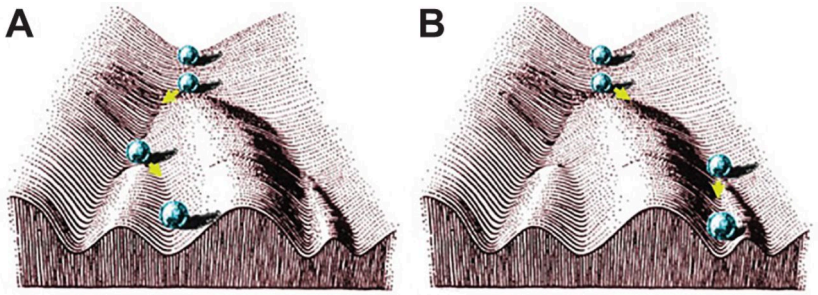
scRNA-seq in historical context of cell characterization:

1665	Hooke	Coined ”cell”
mid-1800s	assorted; dye industry	Histological stains
1855	Virchow	Cellular theory
1941	Coons	Immunohistochemistry
1994	Chalfie	Individual cells with GFP
1953 1968	Coulter Fulwyler	Flow cytometry: 17-18 features/cell
2009	U. Toronto, DVS	Mass cytometry (CyTOF): ~100 features/cell
2009-2015	Tang, Klein, Macosko	single-cell RNA-seq: 2-6K features/cell (~20K/sample)



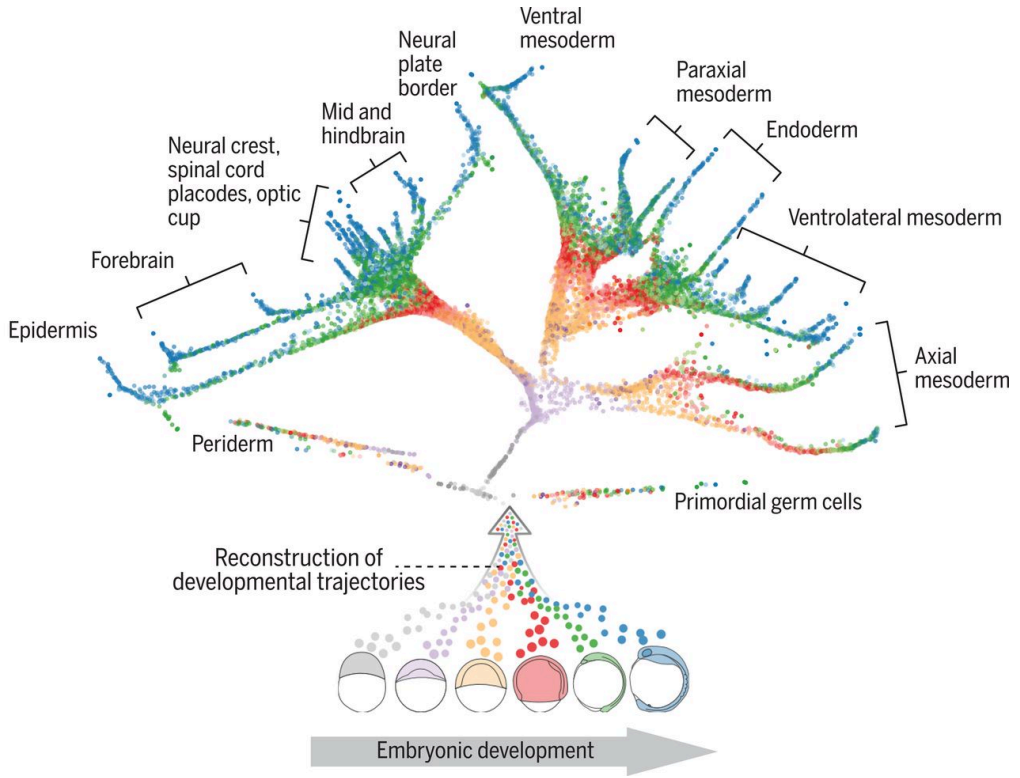
Revisiting concept and definition of “cell type”

- Cell type - stable, “hard-wired” (e.g. by transcription factors)
- Components:
  - Function/phenotype
  - Lineage
    - often continuous (not discrete)
  - State
    - Variable and continuous
    - reprogrammable, “soft-wired” (e.g. by environment)
    - Normal range of cell states vs. pathological range



Waddington Landscape

High dimensional scRNA-seq data permits detailed analysis and reconstruction of cell lineage and state:



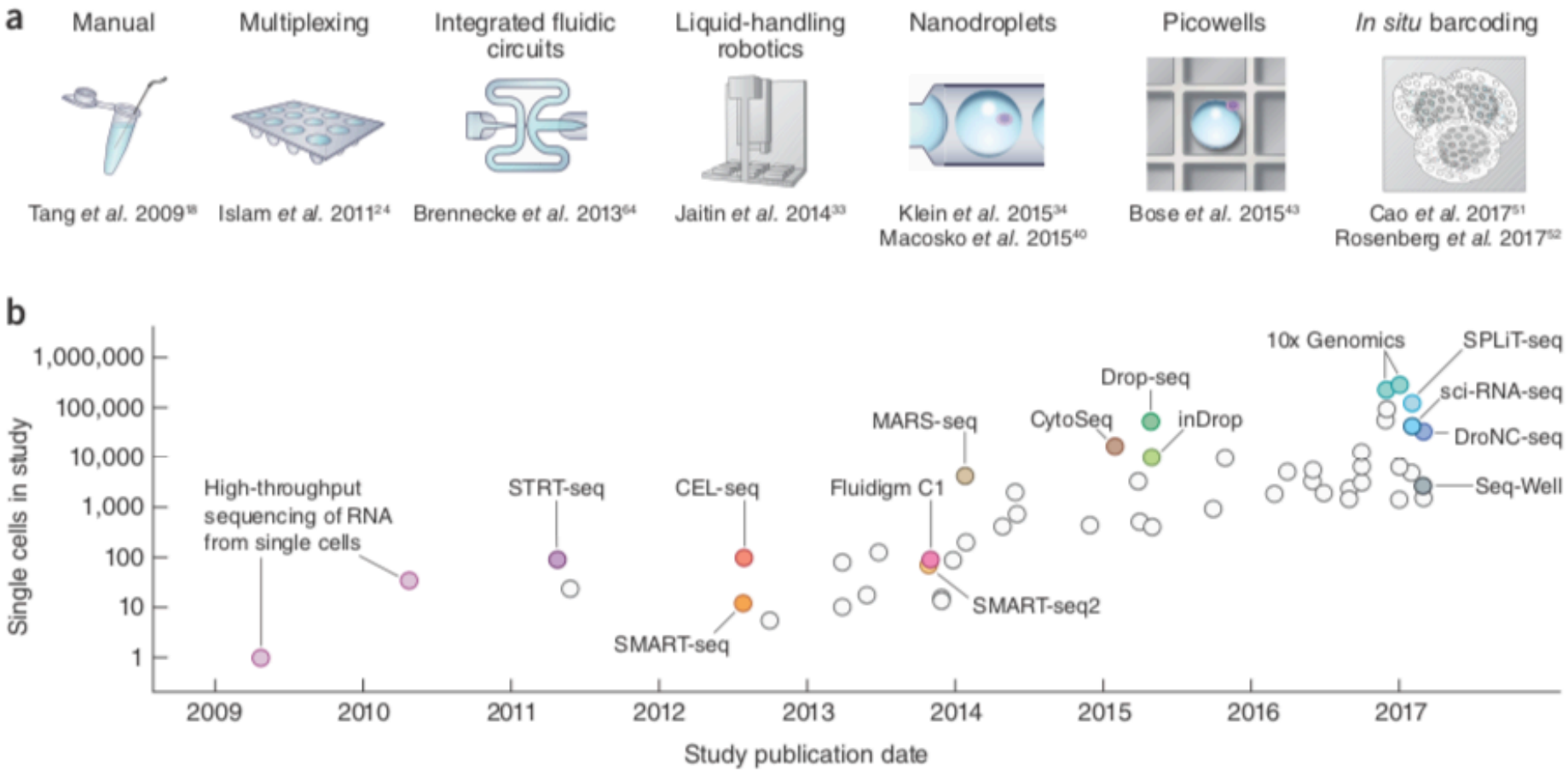
Developmental tree of early zebrafish embryogenesis

Farrell et al. Science 2018

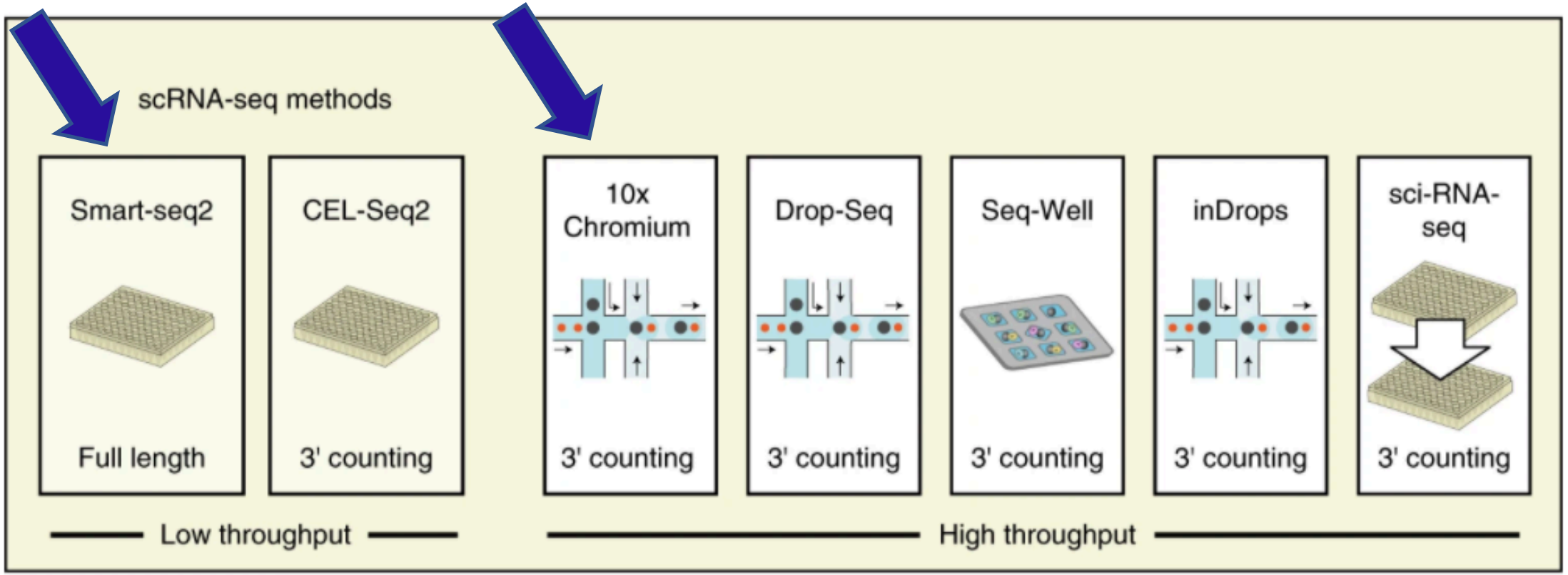
Camp et al, *Science* 2019; S.A. Morris, *Company of Biologists* 2019

Technology

Technology Development

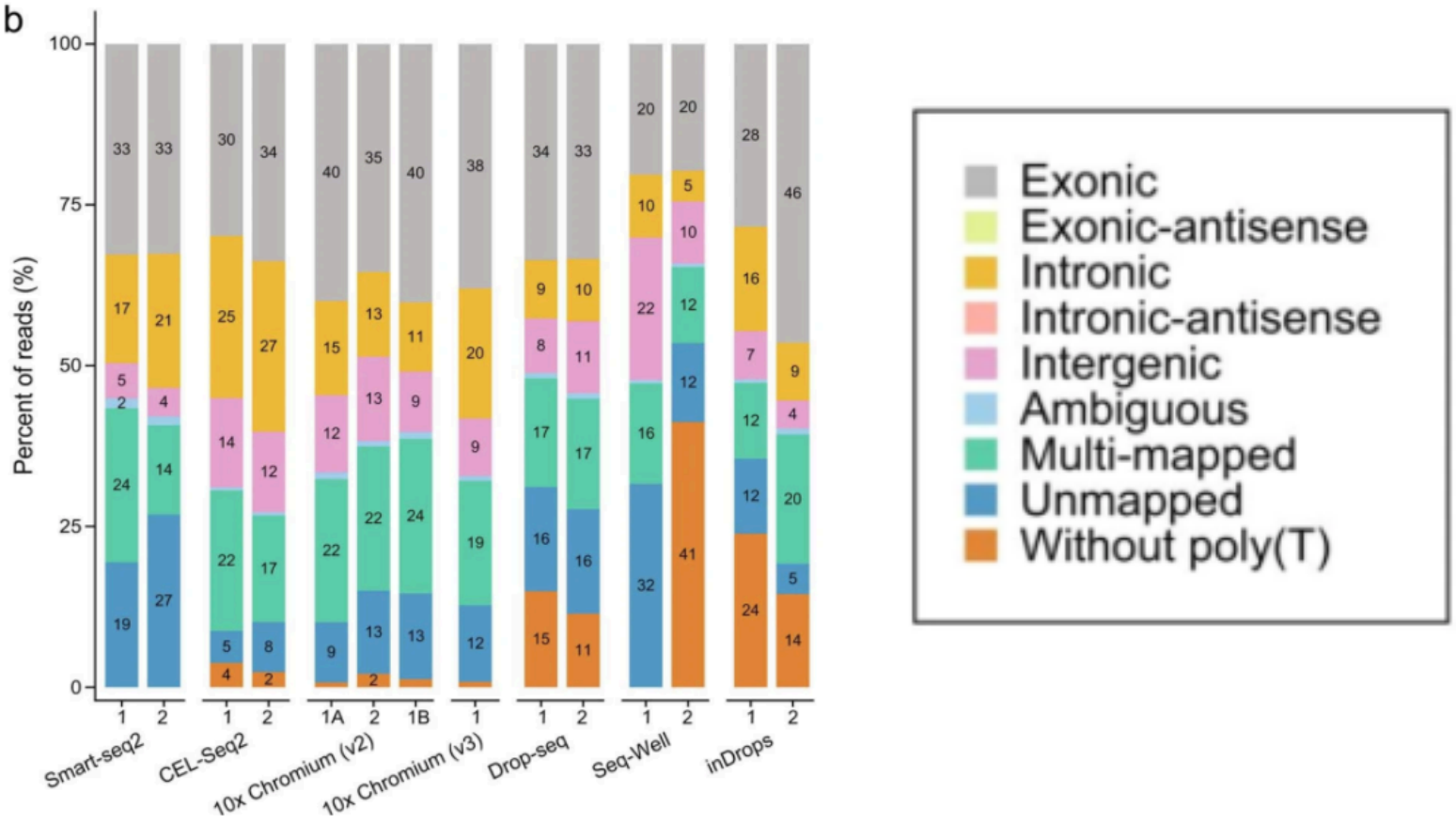


Technology comparison



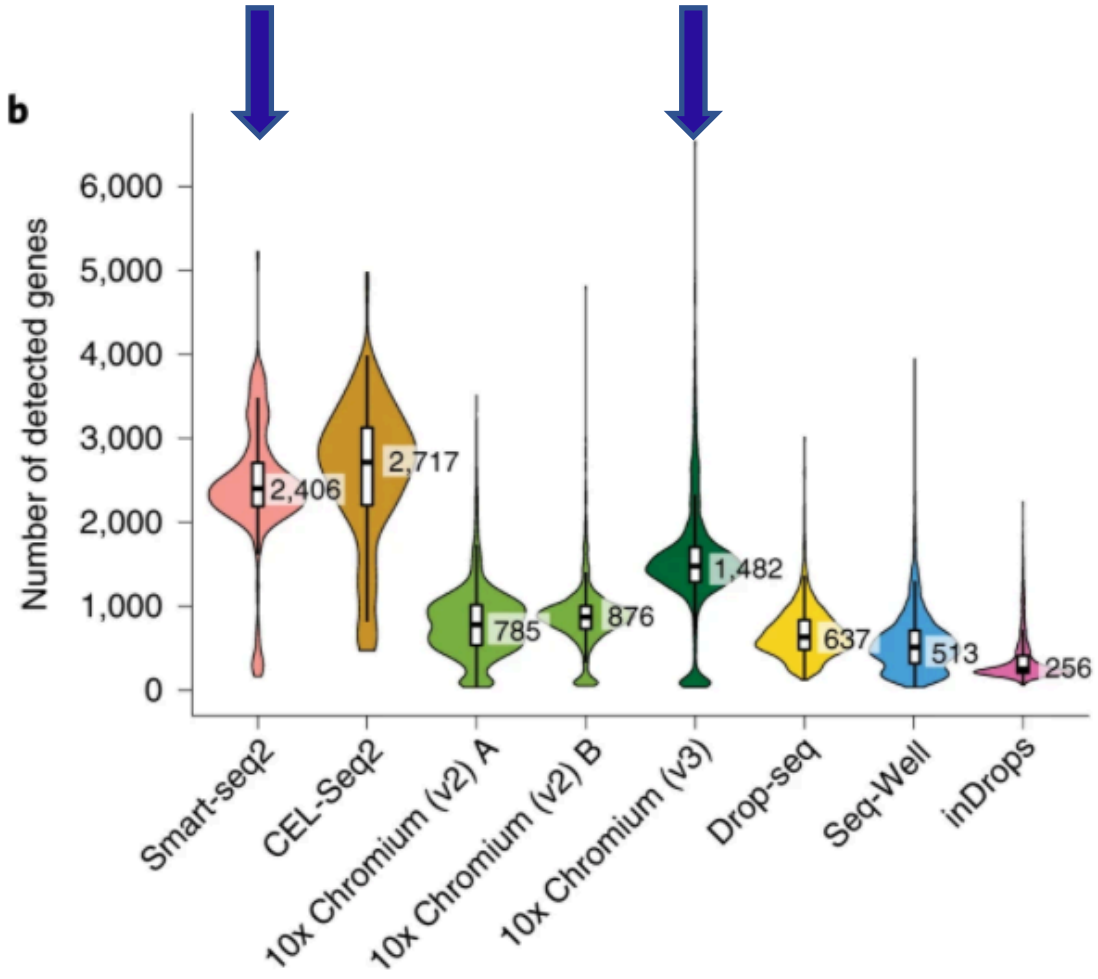
Ding J, et al. (2020) Nature Biotech. 38:737-746

What do these methods capture?



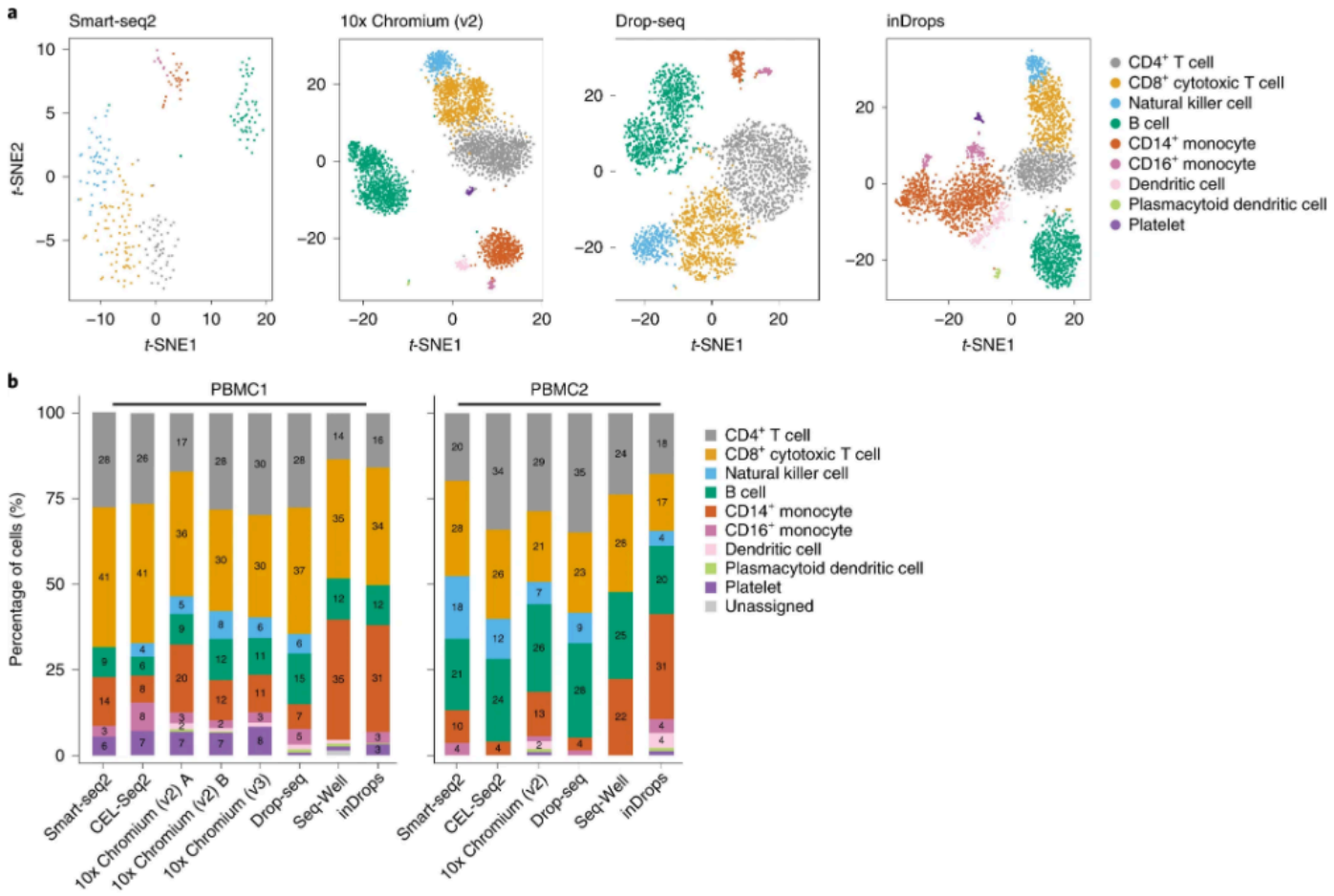
Ding J, et al. (2020) Nature Biotech. 38:737-746

Plate-based methods are more sensitive…



Ding J, et al. (2020) Nature Biotech. 38:737-746

…but are not necessarily better



Ding J, et al. (2020) Nature Biotech. 38:737-746

Popular commercial platforms: 10x Genomics vs Fluidigm SMART-Seq

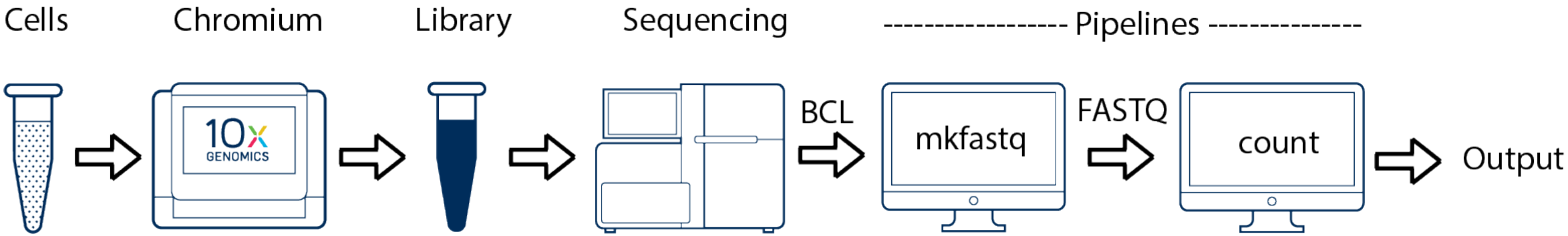
- **Drop-Seq: 10x Genomics**
  - 3’ V3.1 GEM Gene expression
  - 5’ V1 Gene expression (more sequence; less-biased coverage; TCR/BCR sequencing)

- Plate-based: Fluidigm
  - C1 SMART-Seq2: lower-throughput, more genes/cell, longer cDNAs, no UMIs
  - SMART-Seq3*: lower-throughput, more genes/cell, longer cDNAs (uses UMIs)

- All have limitations: must choose technology best suited to application
  - Lafzi et al, *Nature Protocols* 13:2742-2757

- Extensions/Variations
  - Single-nucleus RNA-sequencing for frozen or hard-to-dissociate tissues
  - CITE-seq (aka “feature barcoding”)
  - scATAC-seq

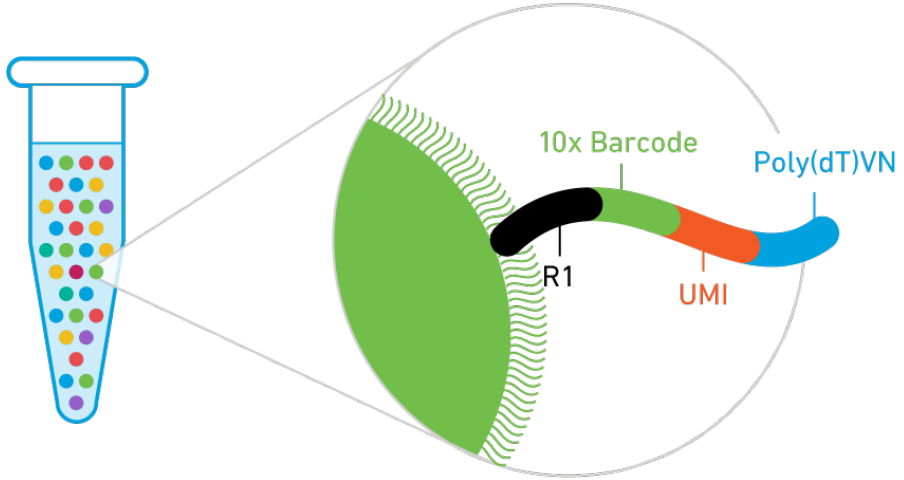
10xGenomics Technology, Pipeline, and Analysis

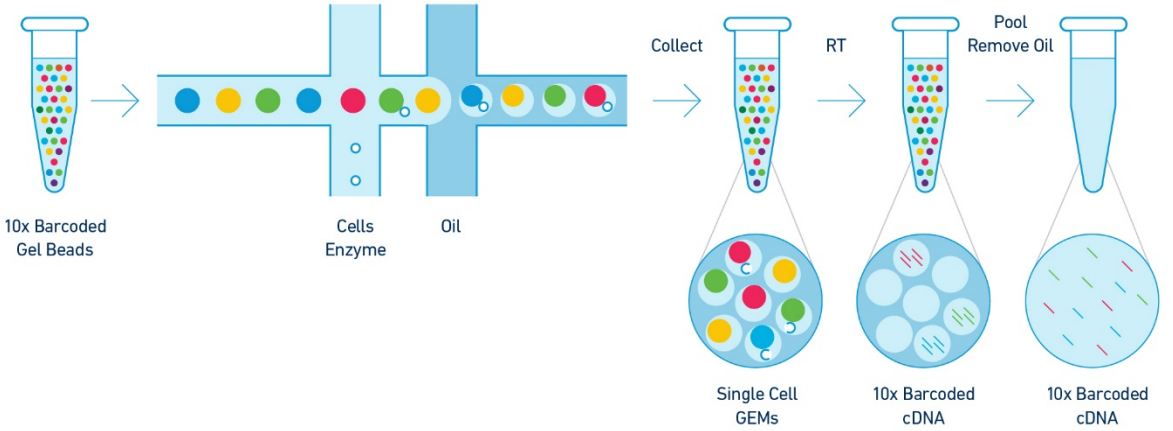


What happens in the Chromium instrument?

“Single Cell 3’ Solution” (10x Genomics)

Barcoded bead + cell = bar-coded cDNA library





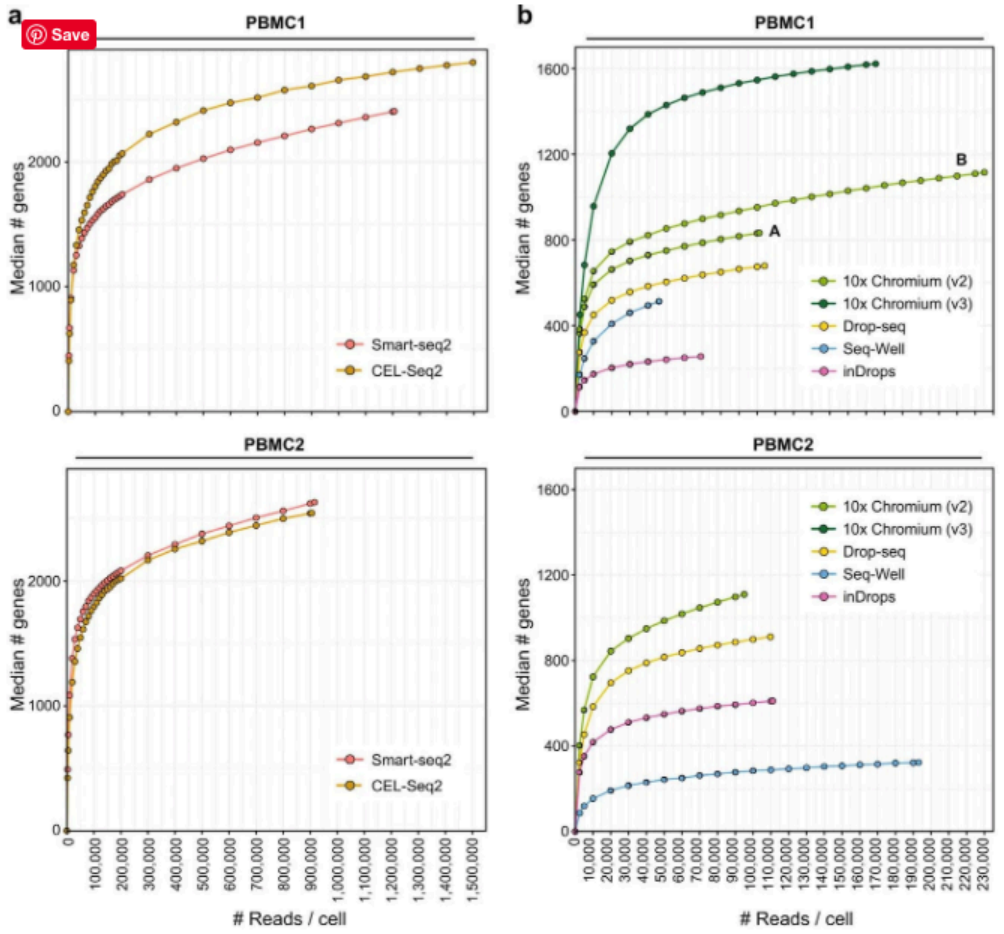
Pool
(sample index)
Amplify
Fragment*
Add primers

Sequence

How deeply do you need to sequence?

Rule of thumb:
Achieve 90% saturation

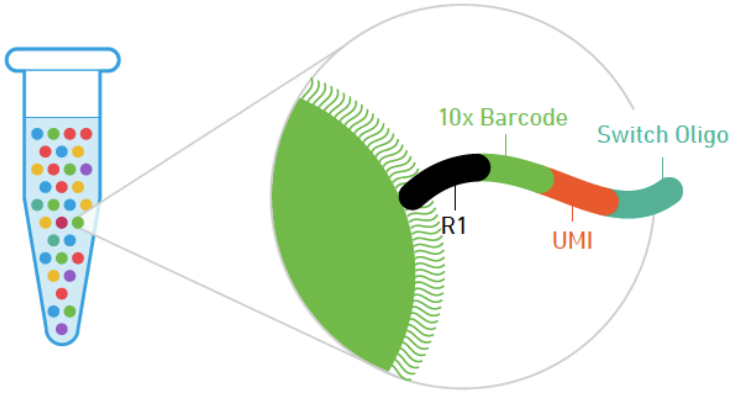
Official Recommendations (reads/cell):
3’ V3: 20K
3’ V2: 50K
5’ 20K
5’ with variant discovery: 200K
5’ V(D)J: 5K
Higher for cell lines



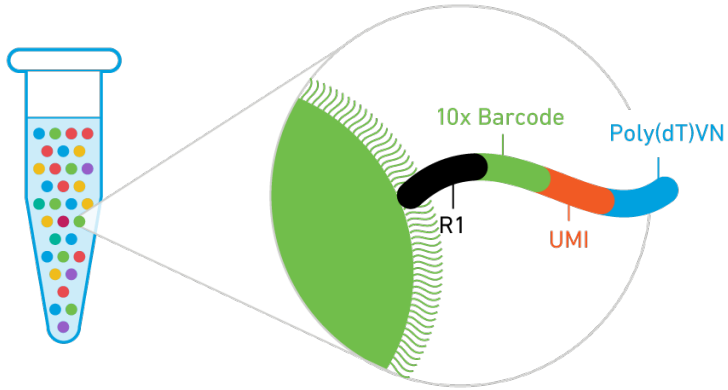
Ding J, et al. (2020) Nature Biotech. 38:737-746

5’ vs. 3’ Chemistry

5’ bead

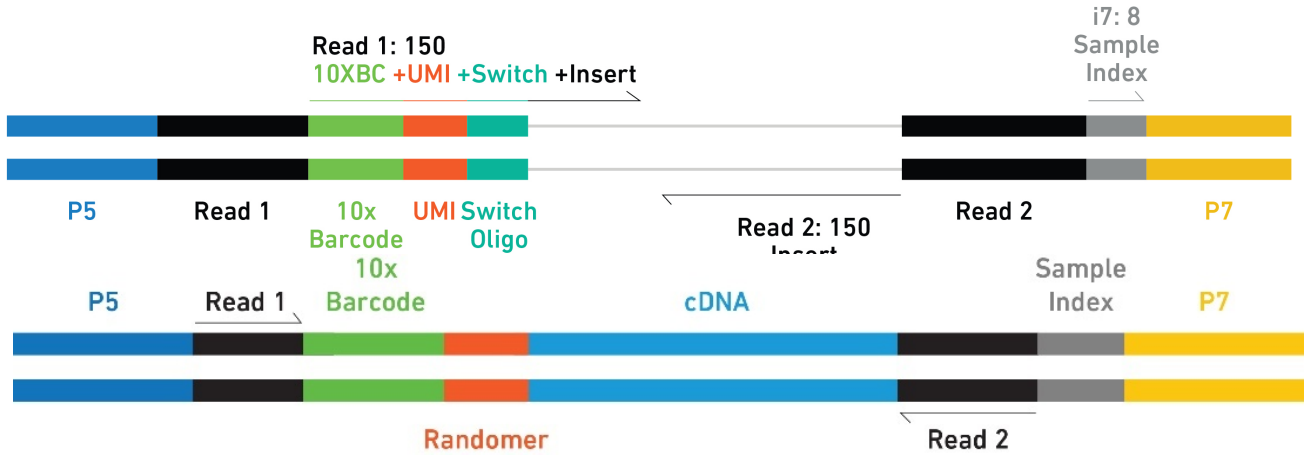


3’ bead



5’ library

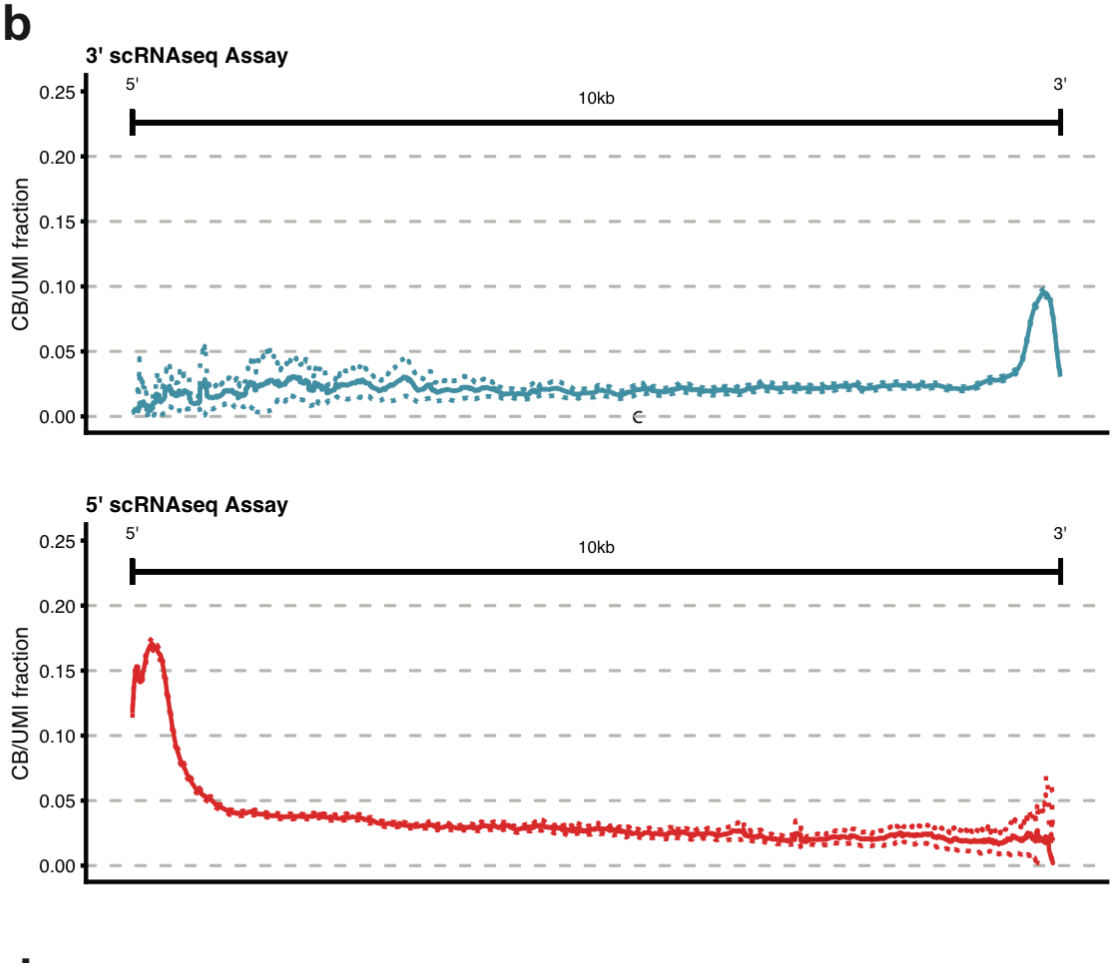
3’ library



And more! https://teichlab.github.io/scg_lib_structs/

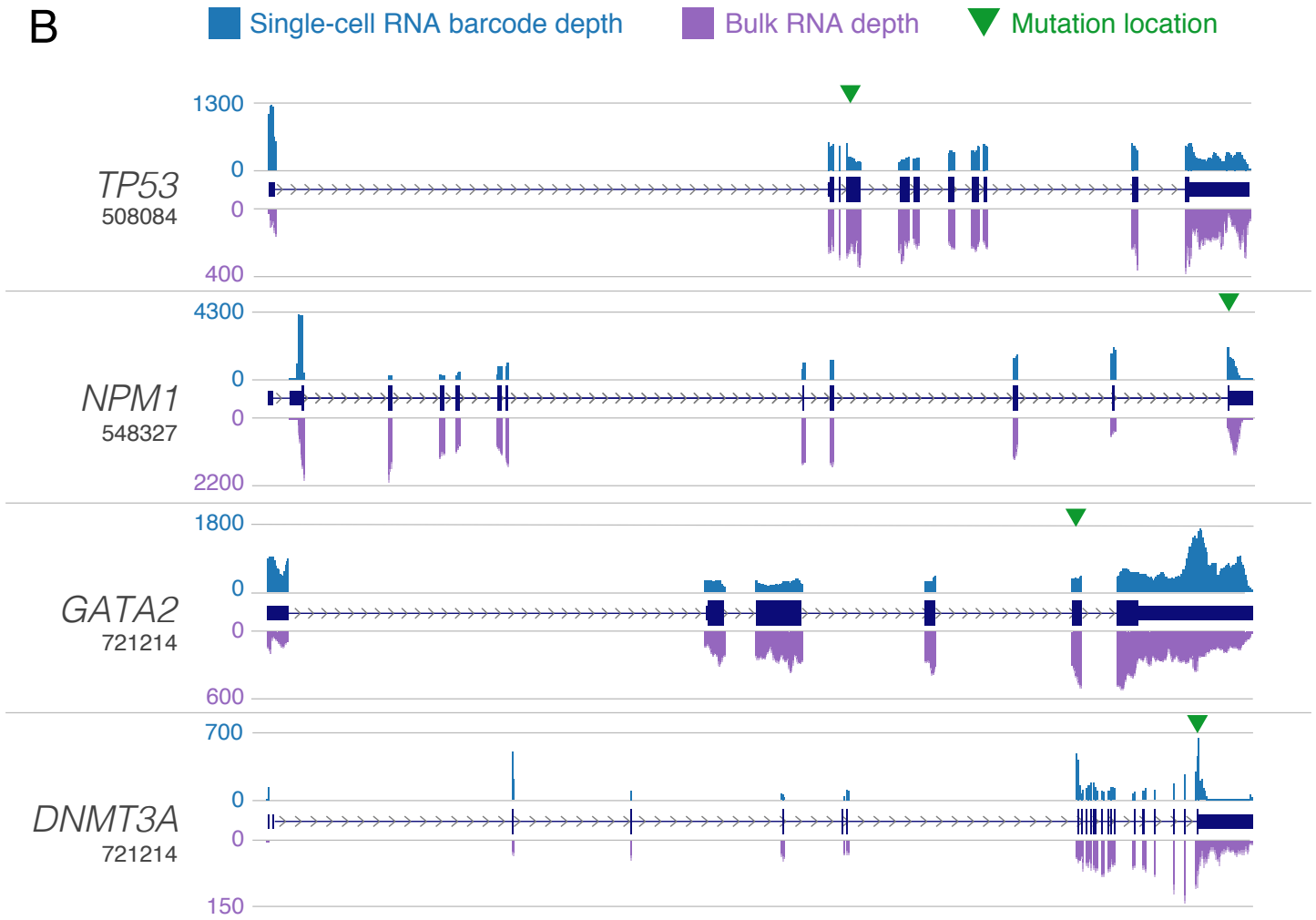
5’ vs. 3’ Transcript Coverage

Fraction of UMIs with coverage at indicated position



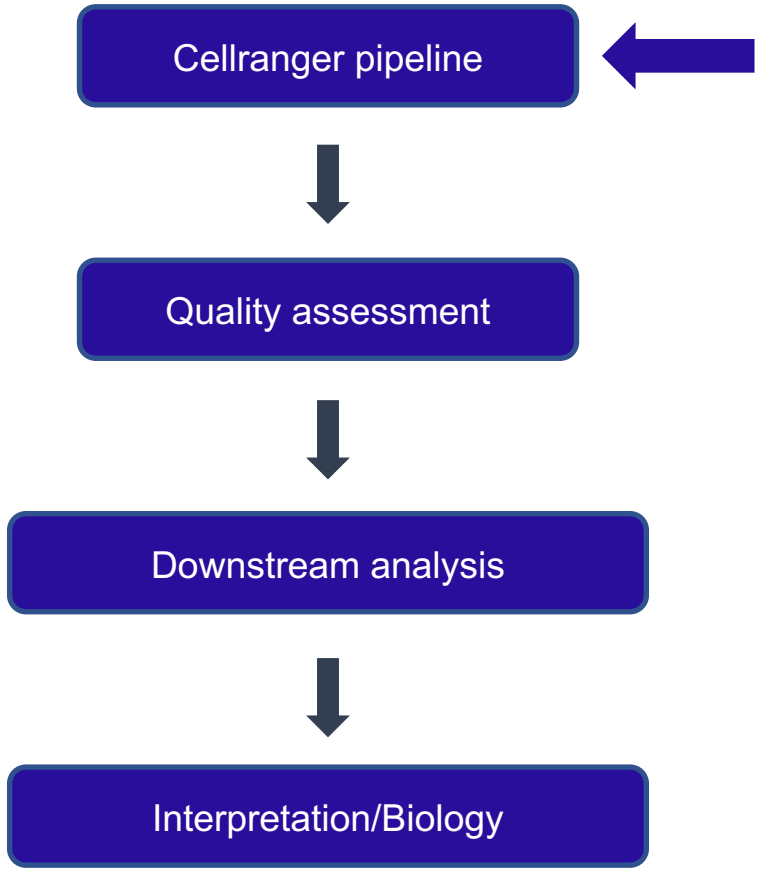
Petti A, et al. (2019) Nat. Comm. 10:3660

scRNA-seq recapitulates bulk transcript coverage



Petti A, et al. (2019) Nat. Comm. 10:3660

Post-sequencing workflow



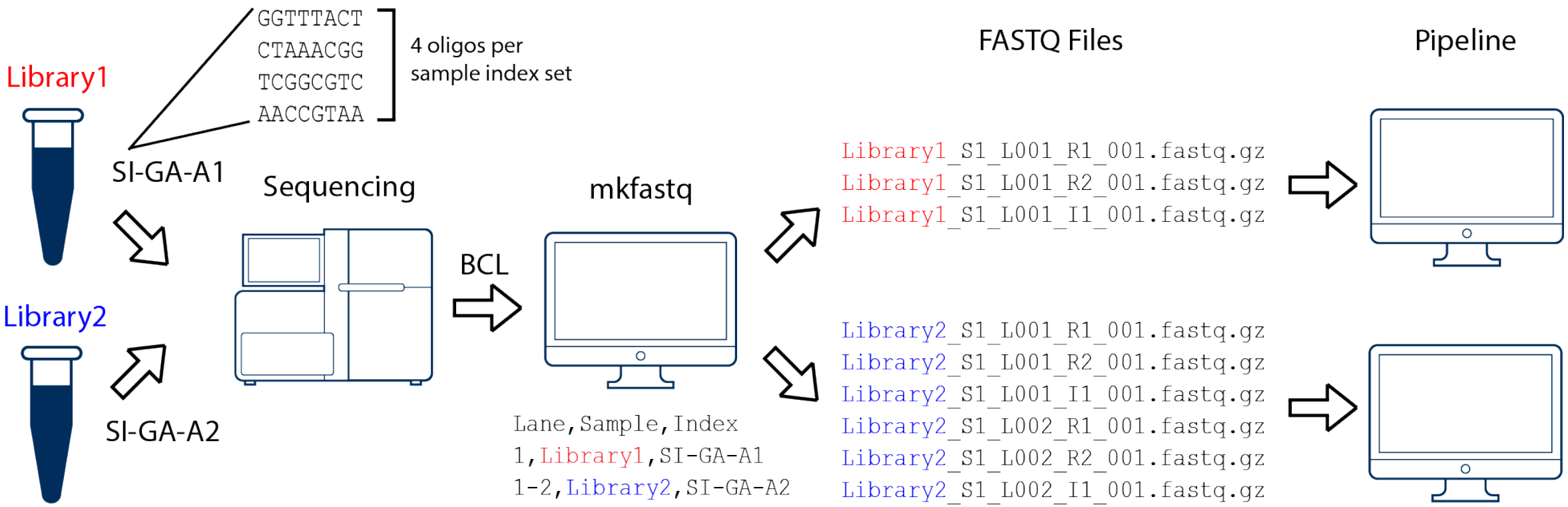
Alternative pipelines:

- kallisto bustools: https://www.kallistobus.tools/
- scumi: https://bitbucket.org/jerry00/scumi-dev/src/master/

Cellranger

- cellranger and all dependencies (e.g. reference transcriptomes) can be downloaded from the 10x Genomics website:
- https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest
- Extensive instructions are provided here: https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/using/tutorial_ov

Cellranger Step 1: Sample demultiplexing using ‘cellranger mkfastq’



https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/using/mkfastq

Running the ‘cellranger mkfastq’ command

Usually done by sequencing provider

```
cellranger mkfastq –id=SampleID –run=/path/to/machine/data/directory –
samplesheet=SampleSheet.csv –csv -qc
```

- cellranger mkfastq is a wrapper for Illumina’s bcl2fastq script, which converts basecall (bcl) files to fastq files
- -qc option is not available for NovaSeq sequencers
- If sequencing provider does this step, you should request the SampleSheet file
- Format of the SampleSheet.csv file (example only!):

```
Lane,Sample,Index
5,Sample1,SI-GA-E8
5,Sample2,SI-GA-E9
6,Sample1,SI-GA-E8
6,Sample2,SI-GA-E9
7,Sample1,SI-GA-E8
7,Sample2,SI-GA-E9
8,Sample1,SI-GA-E8
8,Sample2,SI-GA-E9
```

https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/using/mkfastq#fastq_output

Output of cellranger mkfastq

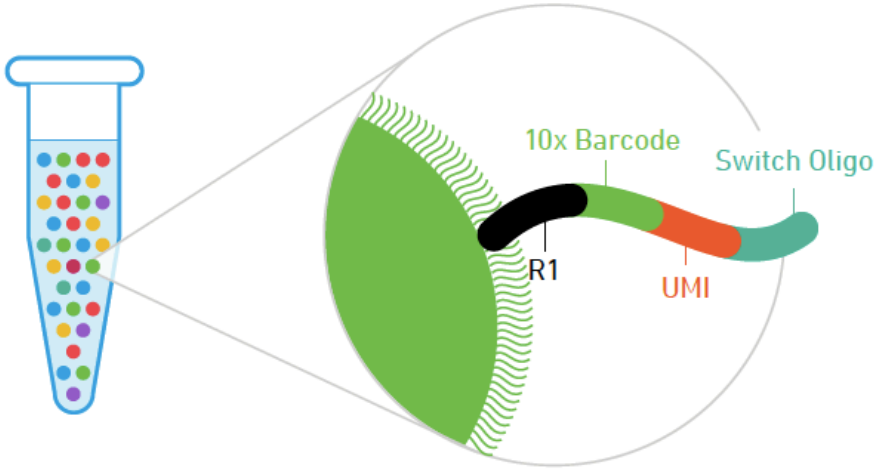
SampleID/FlowCellID/outs OR SampleID/outs/
- fastq_path
- input_samplesheet.csv
- interop_path
- **qc_summary.json** (not available for NovaSeq)

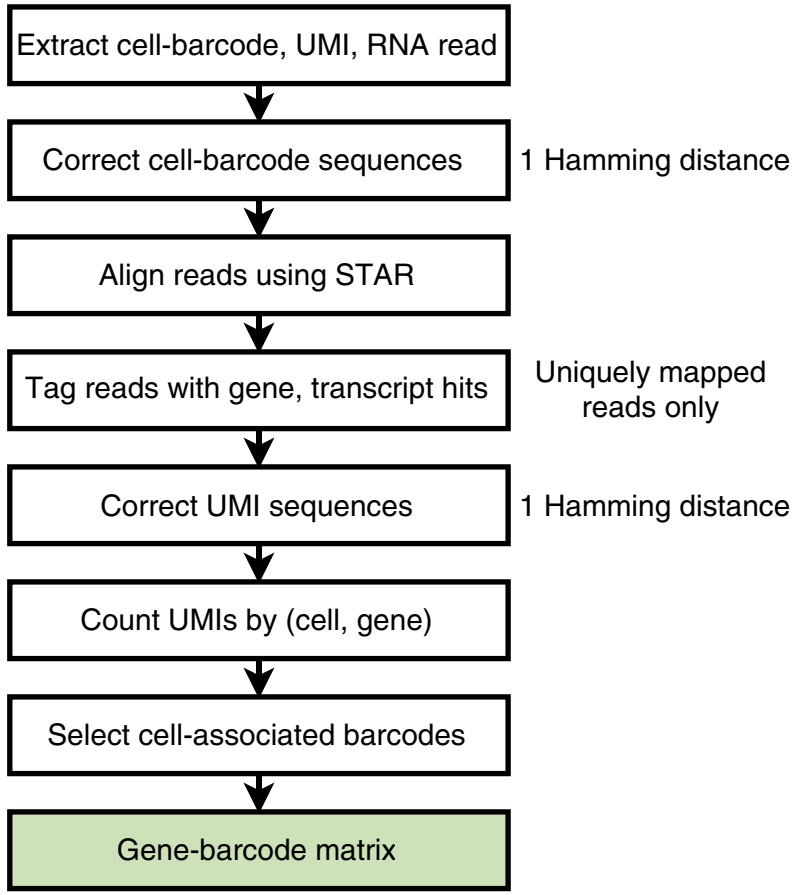

```
"sample_qc": {
    "M_FD-DNMT3A_HET_5_mo-DNMT3A_HET_5_mo_10x": {
        "1": {
            "barcode_exact_match_ratio": 0.9749976800017514,
            "barcode_q30_base_ratio": 0.9834758240855174,
            "bc_on_whitelist": 0.9839895504244381,
            "gem_count_estimate": 69920,
            "mean_barcode_qscore": 37.59103822723894,
            "number_reads": 32295283,
            "read1_q30_base_ratio": 0.9848257525600684,
            "read2_q30_base_ratio": 0.9071105899361747
        },
        "2": {
            "barcode_exact_match_ratio": 0.9734708523492642,
            "barcode_q30_base_ratio": 0.983418719089427,
            "bc_on_whitelist": 0.9839573344922891,
            "gem_count_estimate": 69928,
            "mean_barcode_qscore": 37.59170178631634,
            "number_reads": 32157686,
            "read1_q30_base_ratio": 0.9847541041737509,
            "read2_q30_base_ratio": 0.9074492953255405
        },
        "all": {
            "barcode_exact_match_ratio": 0.9742358959445298,
            "barcode_q30_base_ratio": 0.9834473325425862,
            "bc_on_whitelist": 0.9839734768463497,
            "gem_count_estimate": 72805,
            "mean_barcode_qscore": 37.59136929848026,
            "number_reads": 64452969,
            "read1_q30_base_ratio": 0.9847900048459545,
            "read2_q30_base_ratio": 0.9072795810893632
        }
    },
```


Base quality (q20 and q30 fraction by cycle) for barcode, UMI, read1, and read 2. Example:


```
"barcode_q30_fraction_by_cycle": [
    0.97418484019826268,
    0.97736041198282475,
    0.97889368916379427,
    0.98058003575024655,
    0.98025671019901617,
    0.97740489509564854,
    0.9847162345141085,
    0.98756393926824182,
    0.98745237160227162,
    0.98848508862667428,
    0.9885772992879841,
    0.9882130163901619,
    0.98774013989239307,
    0.98793212425760379,
    0.98716427813785512,
    0.98738571452544555
],
```

Cellranger Overview



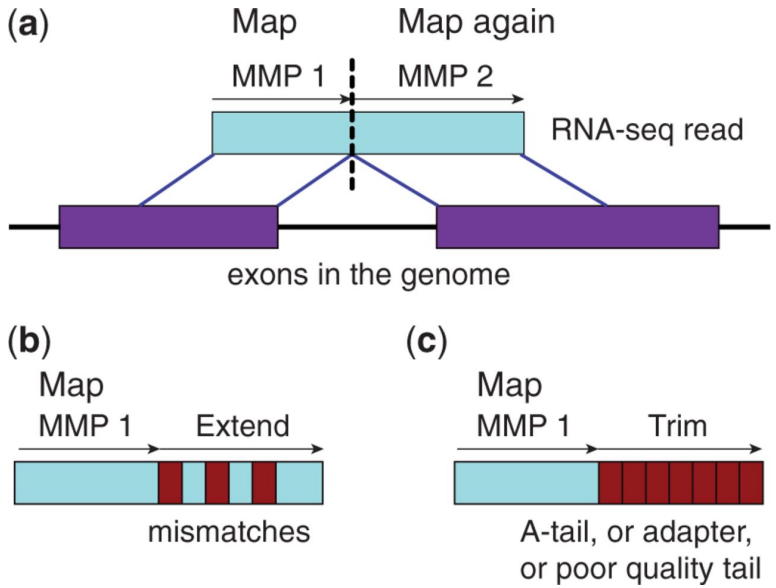


Zheng, et al. 2016. Nature Communications 8:14049

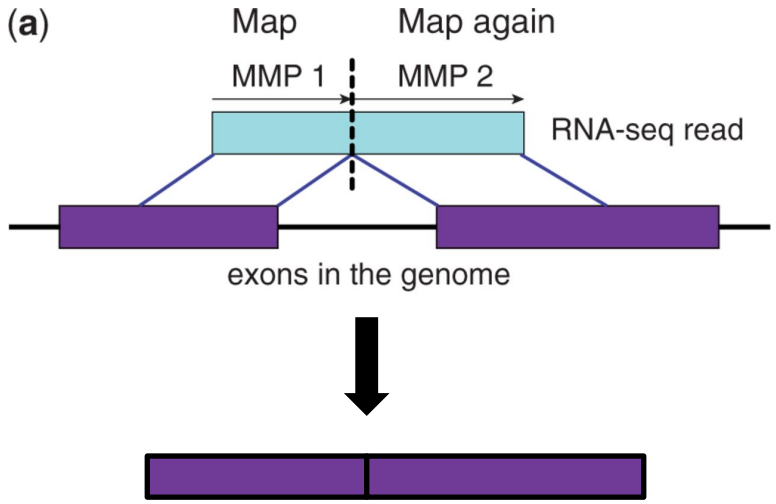
Cellranger: STAR alignment

STAR = Spliced Transcripts Alignment to a Reference
Aligns non-contiguous reads directly to the reference genome
Dobin A, et al. (2013) Bioinformatics 29(1):15-21.

Step 1: Seed search for Maximum Mappable Prefix (MMP)



Step 2: Cluster, stitch, and score the seeds from Step 1

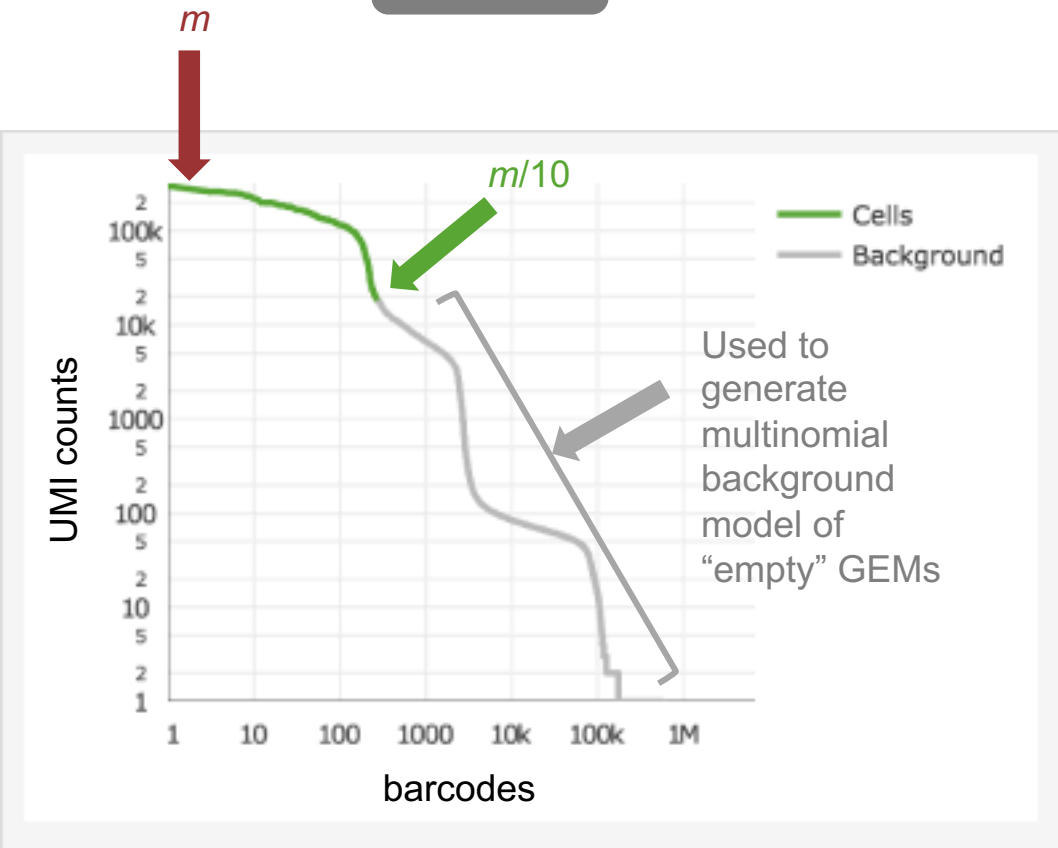


Processing aligned reads

- Uses transcript annotation GTF file to bin reads into exonic, intronic, and intergenic reads
- If at least 50% of the read overlaps an exon: exonic
- Otherwise, if it intersects an intron: intronic
- Otherwise, intergenic
- Exonic loci are prioritized in the event of multi-mapping
- If an exonic read corresponds to an annotated transcript, aligned to the same strand, and compatible with single-gene annotation, it is used for UMI counting

Cellranger: Selecting barcodes (cells)

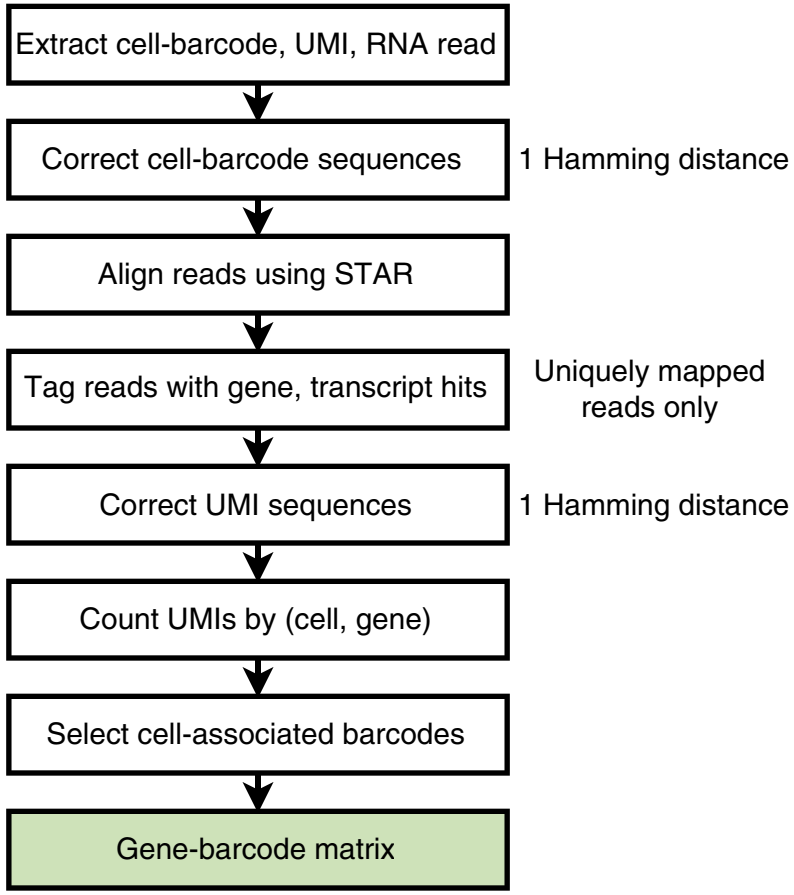
Step 1

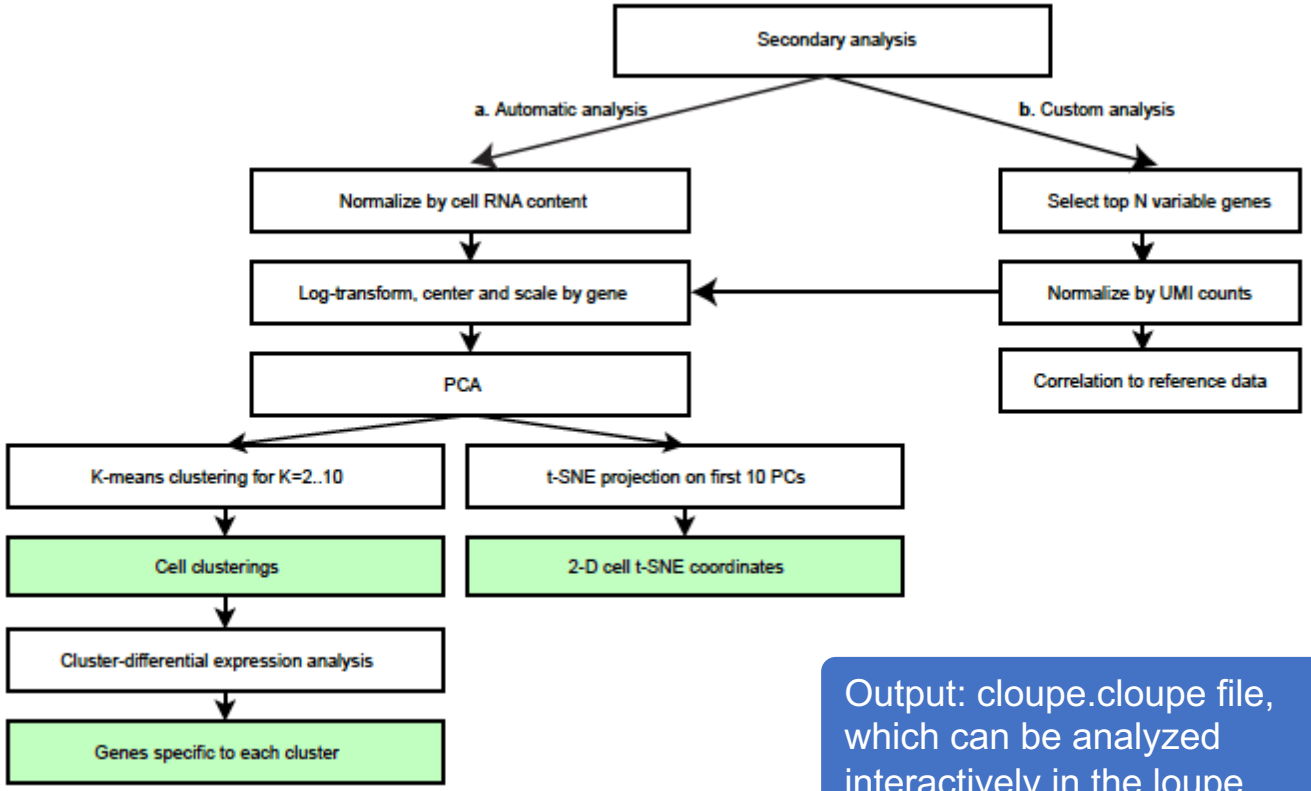


m = 99th %tile of expected cells

Analysis of Gene-barcode matrix (better done yourself)

Normalization, dimensionality reduction, data representation





Output: cloupe.cloupe file, which can be analyzed interactively in the loupe browser

Zheng, et al. 2016. Massively parallel digital transcriptional profiling of single cells. Nature Communications 8:14049

Running ‘cellranger count’ using the command line

https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome

Transcript alignment, counting, barcode selection, etc, to generate feature-barcode matrix:

```
cellranger count --id=$OutName --sample=$SampleName --fastqs=/path/to/fastqs –
indices=$SampleIndices --transcriptome=/path/to/refdata-cellranger-GRCh38-3.0.0 --
localmem=64 --localcores=12
```

Definitions:
$OutName = what you want the output directory to be called (using the sample name works well)
$SampleName = sample name provided to the sequencer; in fastq file name, e.g. SampleName_S1_L003_R1_001.fastq.gz
$SampleIndices = Set of four oligos, such as CAGTACTG,AGTAGTCT,GCAGTAGA,TTCCCGAC, *OR* a code like SI-GA-A2
Note that 10x Genomics provides oligo/code conversion files. 3’ files are here:
https://support.10xgenomics.com/single-cell-gene-expression/index/doc/specifications-sample-index-sets-for-single-cell-3

Cellranger output files

```
B115.mri.tgz  _invocation  outs                 _sitecheck  _vdrkill
_cmdline      _jobmode     _perf                _tags       _vdrkill._truncated_
_filelist     _log         _perf._truncated_    _timestamp  _versions
_finalstate   _mrosource   SC_RNA_COUNTER_CS    _uuid
```

analysis
- clustering - flat file clustering results
- diffexp – DEGs for each cluster
- pca – details about each principal component, projections, etc
- tsne – coordinates of each cell in t-SNE plot

cloupe.cloupe – input to loupe browser for interactive analysis

filtered_feature_bc_matrix
- barcodes.tsv.gz
- features.tsv.gz
- matrix.mtx.gz

filtered_feature_bc_matrix.h5

metrics_summary.csv – flat file QC information

molecule_info.h5

possorted_genome_bam.bam

possorted_genome_bam.bam.bai

raw_feature_bc_matrix – not filtered for cell-associated barcodes

raw_feature_bc_matrix.h5 – not filtered for cell-associated barcodes

web_summary.html – QC information and minimal interactive analysis

Did your experiment work?

Two key QC files:

- metrics_summary.csv
- web_summary.html

web_summary.html, metrics_summary.csv

The analysis detected some issues. Details »

Alert	Value	Detail
⚠ **Low Fraction Valid Barcodes**	63.7%	Ideal > 75%. This usually indicates a quality issue with the Illumina I7 read for Single Cell 3' v1 or the R1 read for Single Cell 3' v2. Application performance may be affected.
⚠ **Low Fraction Reads Confidently Mapped To Transcriptome**	26.1%	Ideal > 30%. This can indicate use of the wrong reference transcriptome, a reference transcriptome with overlapping genes, poor library quality, poor sequencing quality, or reads shorter than the recommended minimum. Application performance may be affected.
⚠ **High Fraction of Reads Mapped Antisense to Genes**	16.9%	Ideal < 10%. This can indicate use of an unsupported chemistry type (e.g. using Single Cell V(D)J for gene counting). Application performance may be affected.

Estimated Number of Cells

9,151

Mean Reads per Cell	Median Genes per Cell
147,263	2,206

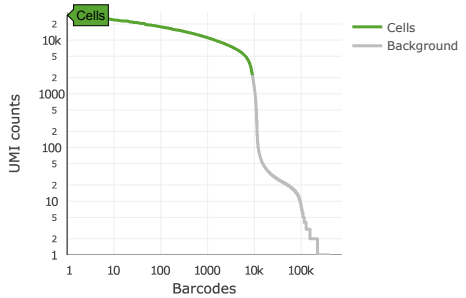
Sequencing

Number of Reads	1,347,604,104
Valid Barcodes	63.7%
Sequencing Saturation	89.0%
Q30 Bases in Barcode	94.9%
Q30 Bases in RNA Read	91.7%
Q30 Bases in RNA Read 2	87.5%
Q30 Bases in Sample Index	92.6%
Q30 Bases in UMI	94.7%

Mapping

Reads Mapped to Genome	92.8%
Reads Mapped Confidently to Genome	70.3%
Reads Mapped Confidently to Intergenic Regions	17.8%
Reads Mapped Confidently to Intronic Regions	10.1%
Reads Mapped Confidently to Exonic Regions	44.7%
Reads Mapped Confidently to Transcriptome	26.1%
Reads Mapped Antisense to Gene	16.9%

Cells



Estimated Number of Cells	9,151
Fraction Reads in Cells	91.9%
Mean Reads per Cell	147,263
Median Genes per Cell	2,206
Total Genes Detected	19,919
Median UMI Counts per Cell	6,407

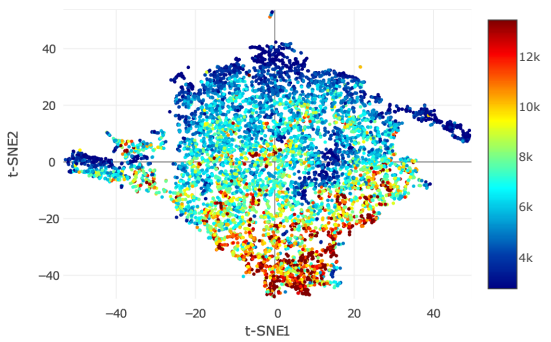
Sample

Name	115225_P_1_Nova
Description	
Transcriptome	GRCh38
Chemistry	Single Cell 5' PE
Cell Ranger Version	2.1.1

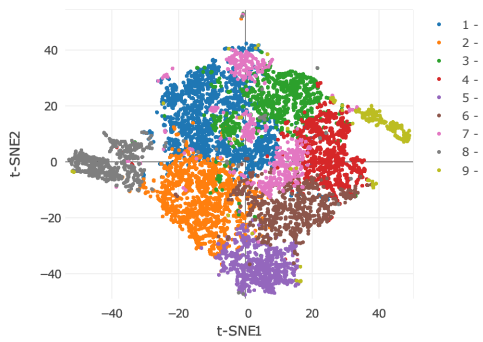
The analysis detected some issues. Details »

Clustering Type: Grap

t-SNE Projection of Cells Colored by UMI Counts



t-SNE projection of Cells Colored by Automated Clustering



Top Genes By Cluster (Log2 fold-change, p-value)

		Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		Cluster 7		Clus
Gene ID	**Gene name**	**L2FC**	**p-value**	**L2FC**	**p-value**	**L2FC**	**p-value**	**L2FC**	**p-value**	**L2FC**	**p-value**	**L2FC**	**p-value**	**L2FC**	**p-value**	**L2FC**
ENSG00000180573	HIST1H2AC	0.56	1e+00	-1.02	5e-01	0.48	1e+00	0.82	6e-01	-0.89	4e-01	-0.66	1e+00	0.29	1e+00	0.25
ENSG00000125652	ALKBH7	0.54	1e+00	-0.01	1e+00	0.19	1e+00	-0.07	1e+00	-0.80	6e-01	-0.32	1e+00	0.09	1e+00	0.33
ENSG00000257698	RP11-620J15.3	0.53	1e+00	-0.19	1e+00	0.01	1e+00	-0.00	1e+00	-0.23	1e+00	-0.36	1e+00	-0.01	1e+00	0.17
ENSG00000104894	CD37	0.48	1e+00	-0.27	1e+00	0.47	1e+00	0.26	1e+00	-0.74	6e-01	-0.43	1e+00	0.14	1e+00	0.22
ENSG00000150782	IL18	0.46	1e+00	-0.51	1e+00	0.86	9e-01	0.44	8e-01	-0.64	8e-01	-0.43	1e+00	0.02	1e+00	-0.49
ENSG00000267453	AC004791.2	0.46	1e+00	0.29	1e+00	-0.61	1e+00	-0.19	1e+00	-0.36	1e+00	-0.04	1e+00	0.14	1e+00	-0.37
ENSG00000196531	NACA	0.45	1e+00	-0.02	1e+00	0.18	1e+00	0.05	1e+00	-0.67	7e-01	-0.16	1e+00	0.16	1e+00	0.00
ENSG00000095932	SMIM24	0.44	1e+00	0.12	1e+00	0.15	1e+00	-0.06	1e+00	-0.13	1e+00	-0.14	1e+00	-0.00	1e+00	-1.74
ENSG00000204628	GNB2L1	0.44	1e+00	0.10	1e+00	0.12	1e+00	-0.08	1e+00	-0.51	9e-01	-0.21	1e+00	0.01	1e+00	-0.05
ENSG00000145708	CRHBP	0.43	1e+00	-0.26	1e+00	0.36	1e+00	0.52	8e-01	-0.55	9e-01	-0.04	1e+00	0.47	1e+00	-2.77
ENSG00000095917	TPSD1	0.42	1e+00	-0.53	1e+00	-1.03	1e+00	0.41	9e-01	-0.55	9e-01	0.06	1e+00	0.91	1e+00	0.41
ENSG00000105373	GLTSCR2	0.41	1e+00	0.06	1e+00	0.12	1e+00	-0.00	1e+00	-0.41	1e+00	-0.26	1e+00	-0.20	1e+00	-0.04
ENSG00000104408	EIF3E	0.40	1e+00	0.05	1e+00	0.23	1e+00	0.09	1e+00	-0.61	8e-01	-0.13	1e+00	0.02	1e+00	-0.15
ENSG00000263961	C1orf186	0.40	1e+00	-0.08	1e+00	0.40	1e+00	0.35	9e-01	-0.67	7e-01	-0.08	1e+00	0.25	1e+00	-1.46
ENSG00000170891	CYTL1	0.40	1e+00	-0.18	1e+00	0.86	9e-01	0.22	1e+00	-0.47	1e+00	-0.49	1e+00	0.13	1e+00	-1.12
ENSG00000269893	SNHG8	0.39	1e+00	0.18	1e+00	0.09	1e+00	-0.09	1e+00	-0.59	9e-01	-0.09	1e+00	0.17	1e+00	-0.33

Sequencing Saturation

Median Genes per Cell

Values from some real experiments

Metric	Human – Cryo. Bone Marrow	Human – Fresh Cell lines	Mouse – Cryo. Bone Marrow
Estimated Cells	7000	8500	5000
Target Reads/Cell	50K (expression), 200K (variants)		
Median Genes/Cell	2000	5600	2100
% Transcriptome mapping	>50%*	70	>70%
% Antisense Reads	~3%	~5	~3%
Fraction reads in cells	80-90%	80-90%	80-90%
Total Genes Detected	20,000	25,000	16,500
Median UMIs/Cell	5000-6000	25000	7000-8000

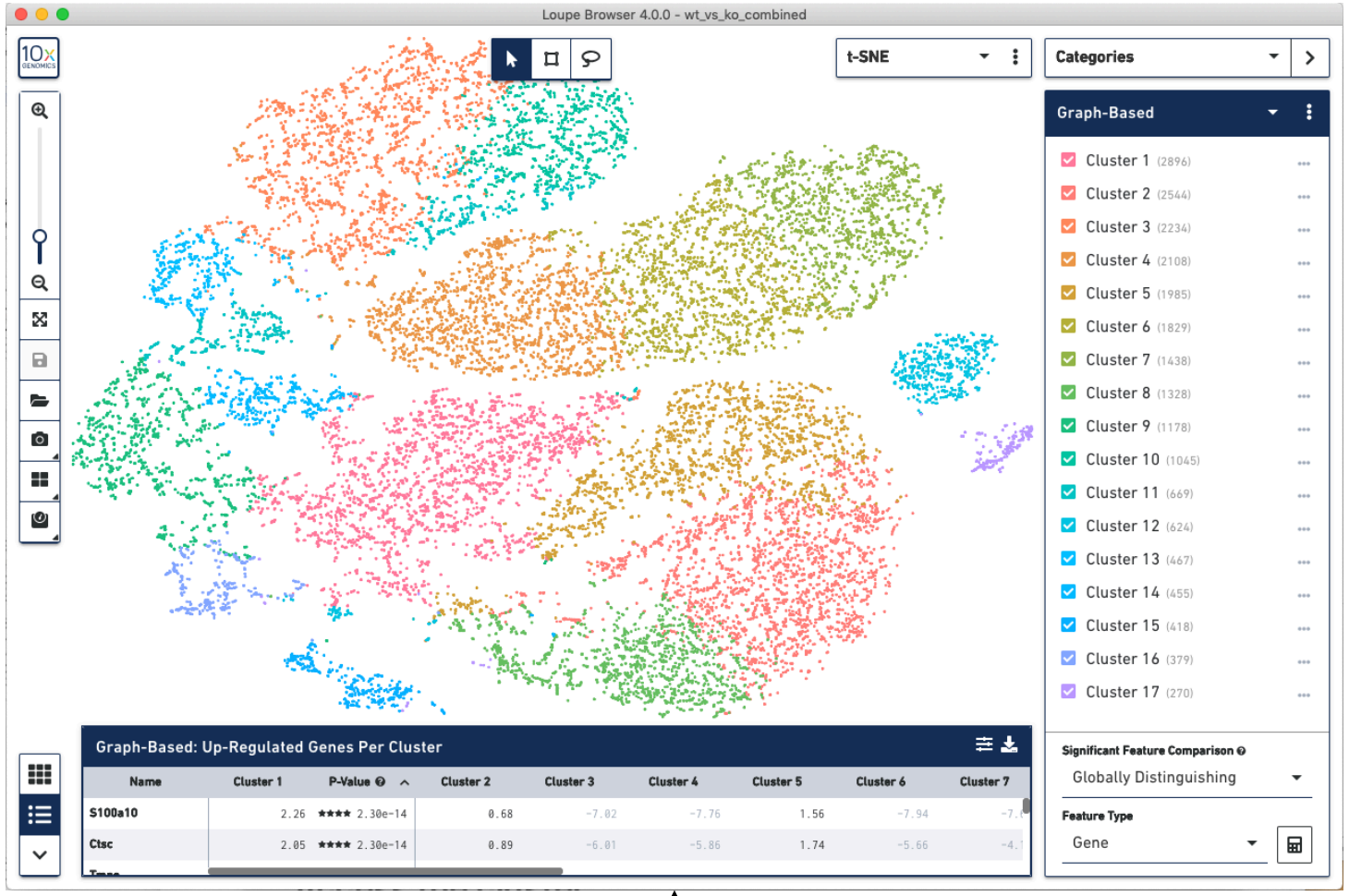
Possible reasons for low quality

Metric	Human
Estimated Cells	Low viability, lysed cells
Target Reads/Cell	Rarely problematic
% Transcriptome mapping	Wrong transcriptome, low sequence quality
% Antisense Reads	Wrong chemistry, low sequence quality
Fraction reads in cells	Lysed cells, extracellular RNA

Exploring the data using the loupe browser

Wild-type mouse (WT) compared to Knock-out (KO) mouse (2 replicates each)

File: wt_vs_ko_combined.cloupe



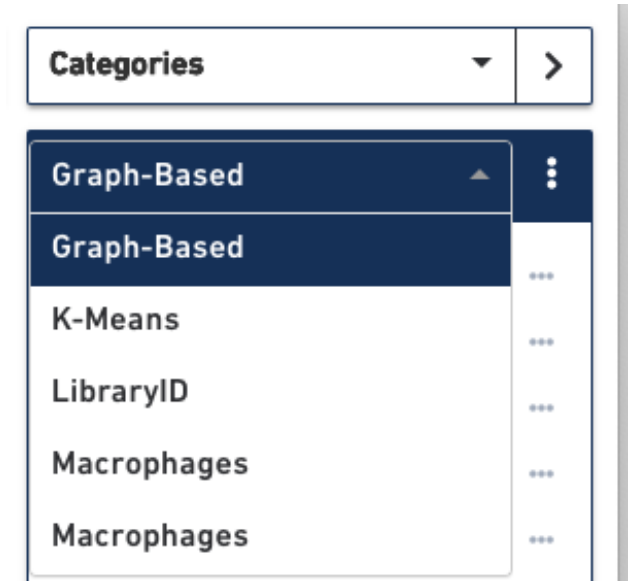
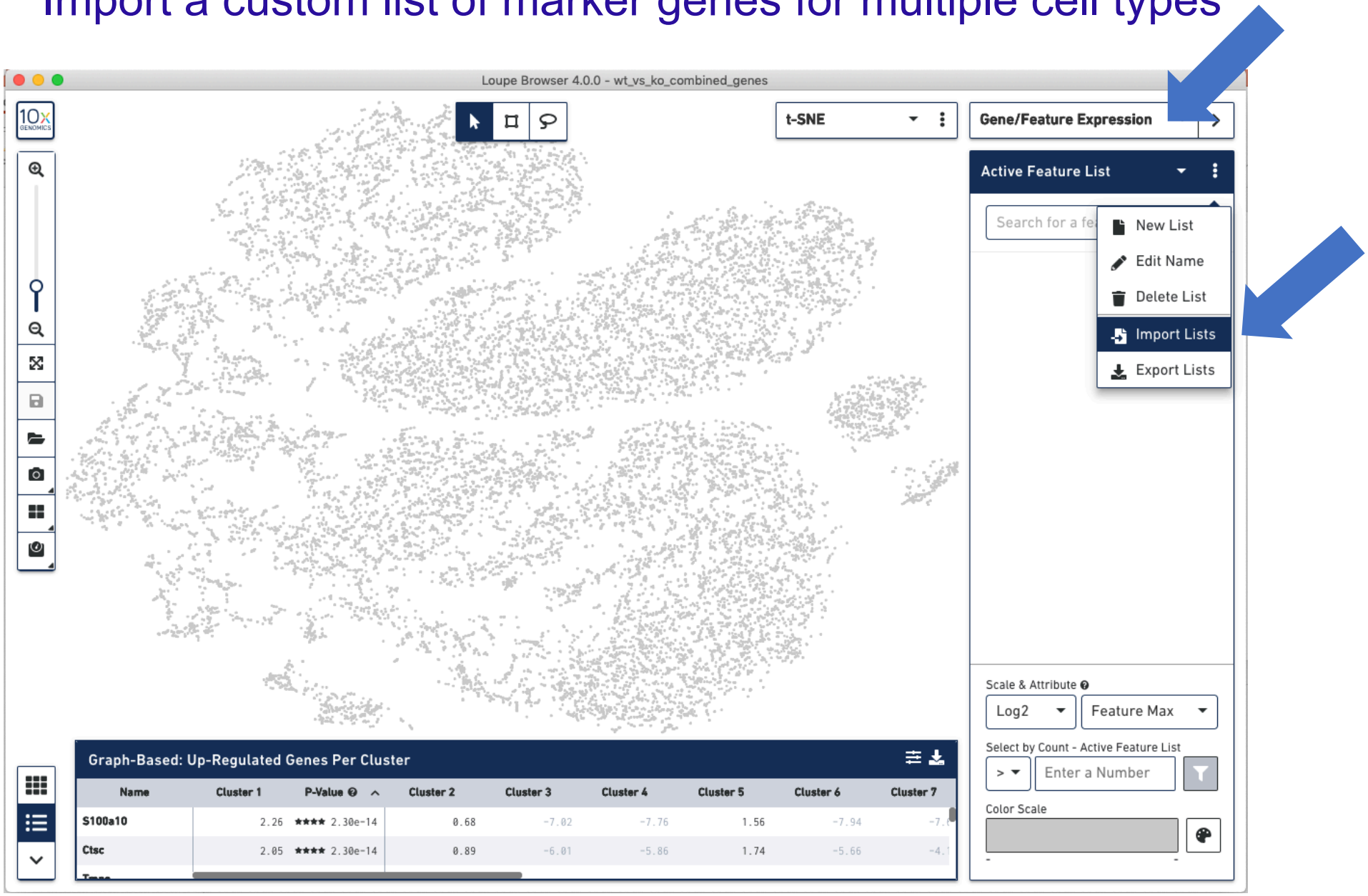


Table of differentially expressed genes in each cluster

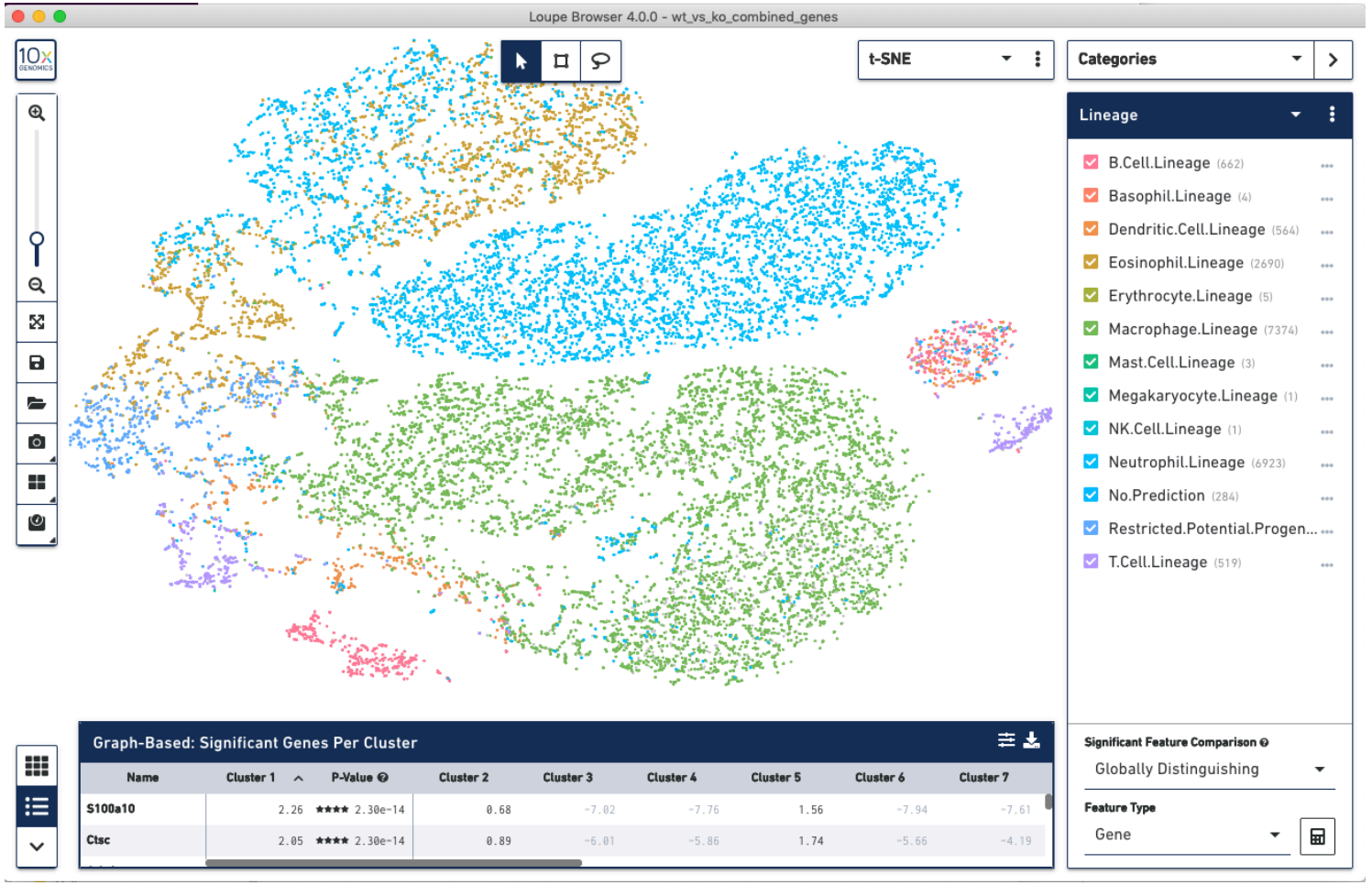

Loupe Browser 4.0.0 - wt_vs_ko_combined_genes
t-SNE
Categories
LibraryID
mouse31_Dnmt3a_KO (7122)
mouse31_Dnmt3a_KO_sort...
mouse32_Dnmt3a_WT (5378)
mouse32_Dnmt3a_WT_sort...
mouse31_Dnmt3a_KO
mouse31_Dnmt3a_KO_sorted
mouse32_Dnmt3a_WT
mouse32_Dnmt3a_WT_sorted
Graph-Based: Up-Regulated Genes Per Cluster
Name
Cluster 1
P-Value
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
S100a10
2.26
2.30e-14
0.68
-7.02
-7.76
1.56
-7.94
Ctsc
2.05
2.30e-14
0.89
-6.01
-5.86
1.74
-5.66
Significant Feature Comparison
Globally Distinguishing
Feature Type
Gene

Import a custom list of marker genes for multiple cell types



Examine some genes and gene sets…

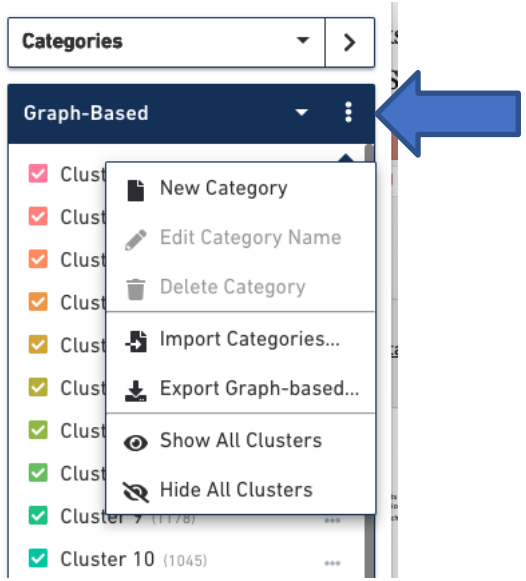
Import a custom list of inferred cell types*

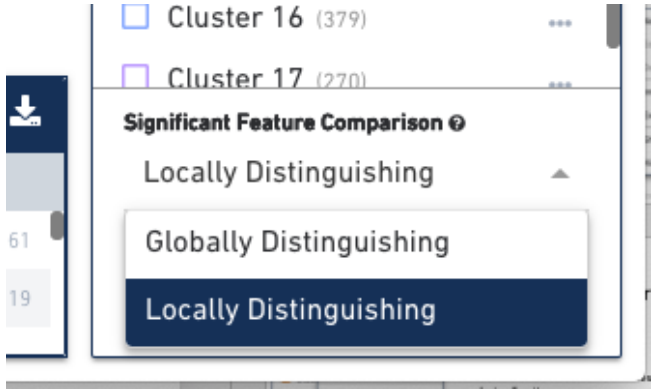


*These cell types were inferred using an in-house nearest-neighbor algorithm and the Haemopedia database. The SingleR package is similar.

Compare two clusters using a differential expression analysis

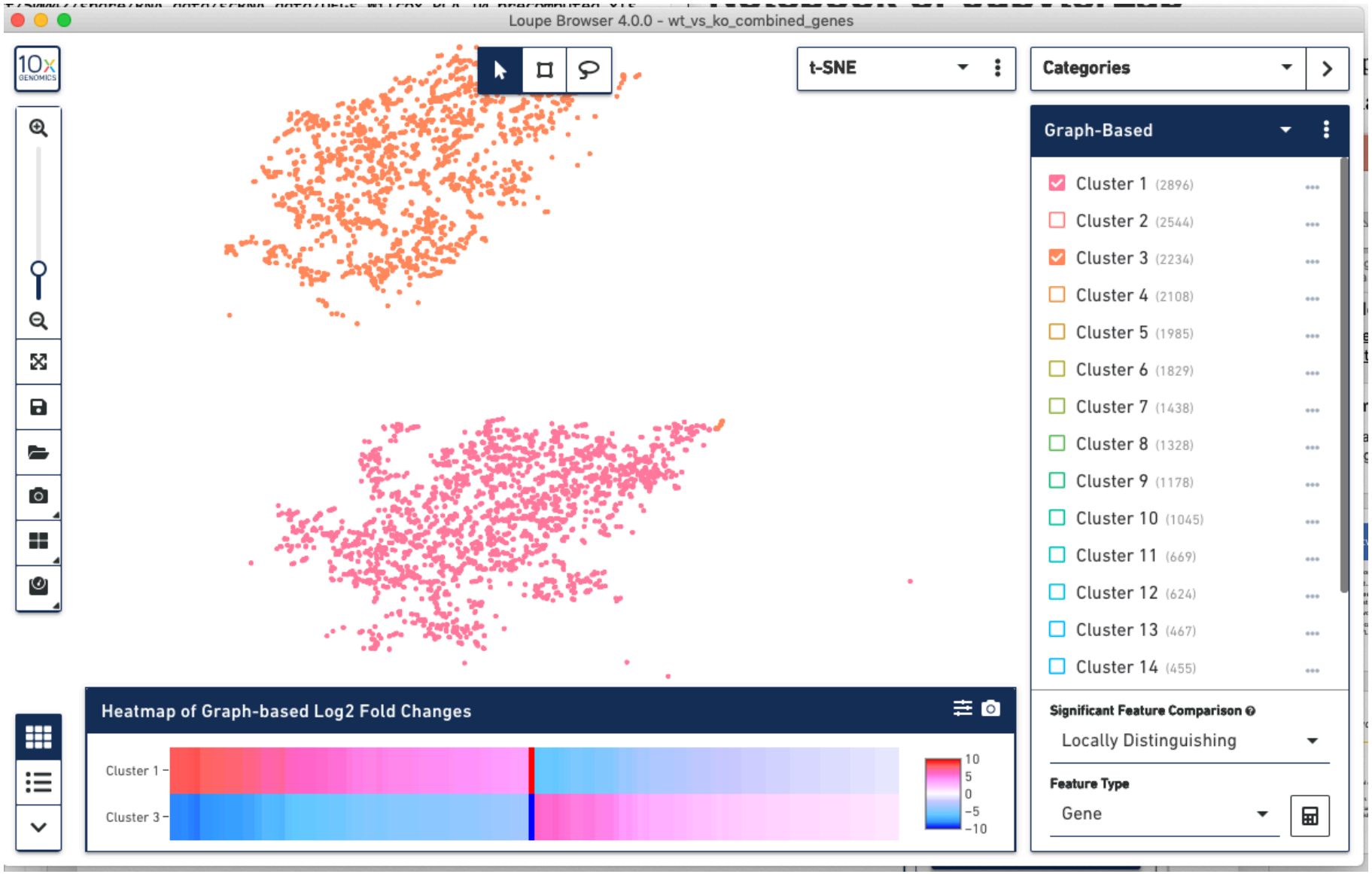
1. Select Categories…
2. Select Graph-Based…
3. Click the three vertical dots to access the “Hide All Clusters” option
4. Then click two clusters of your choice
5. In the Significant Feature Comparison panel, click “Locally distinguishing,” then press calculator icon





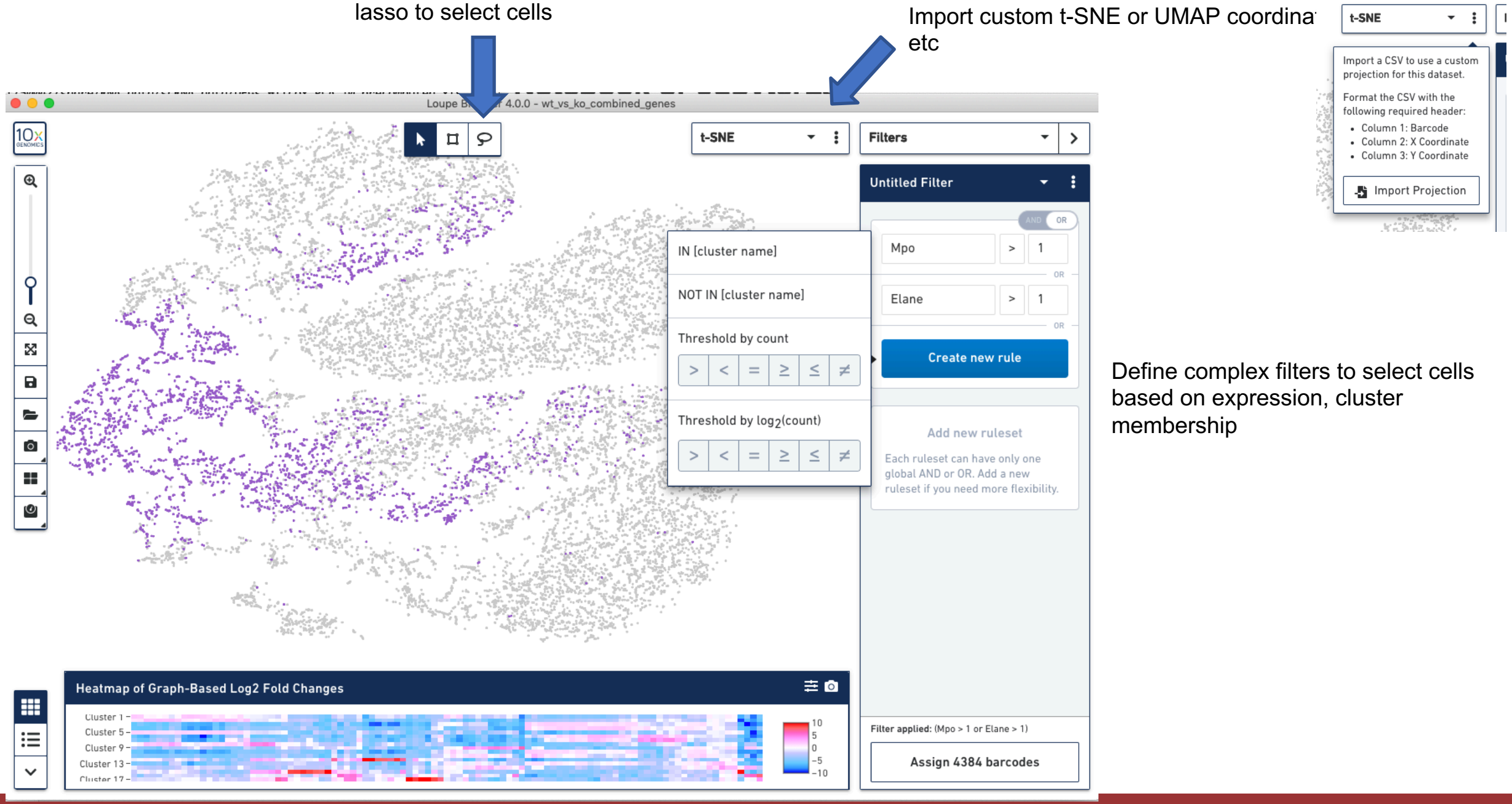
heatmap of results

table of results



lasso to select cells

Import custom t-SNE or UMAP coordina
etc



Define complex filters to select cells based on expression, cluster membership

Exercises

1. There are some significant differences in cell type composition between the WT and KO strains. What are they?
2. There are four clusters of macrophages in this data set. One cluster is missing from the KO mouse. Find this cluster, and generate a heatmap of genes that are differentially expressed across those four clusters. How is the “missing” cluster different from the others in terms of gene expression?

We are on a Coffee Break & Networking Session

Workshop Sponsors: