Introduction to Single-cell RNA-Seq

Wally the Welsh Corgi

Connecting & Computer Preliminaries

Make sure your workshop provided computer is connected to the "**Broad**" or "**Broad Internal**" wireless network.

Please do not connect your personal items

(laptop, phone, etc.) to these wireless networks; it will tax the wireless system and make the workshop less effective.

The password for computers is "password".

Center for Cell Circuits Computational Genomics Workshop 2017

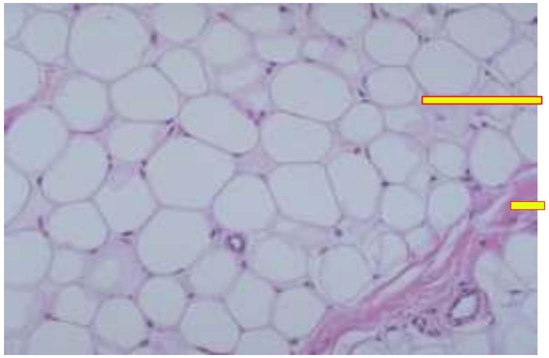
Introduction to single-cell RNA-Seq

Timothy Tickle
Brian Haas
Asma Bankapur

BROAD INSTITUTE

We Know Tissues are Heterogeneous

Adipose

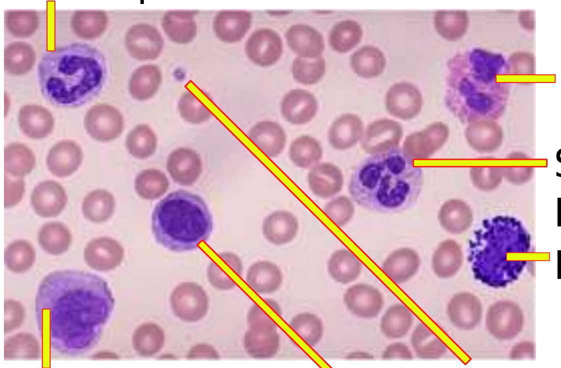


Fat

Connective tissue

Normal Peripheral Blood

Band Neutrophil



Eosinophil

Segmented Neutrophil

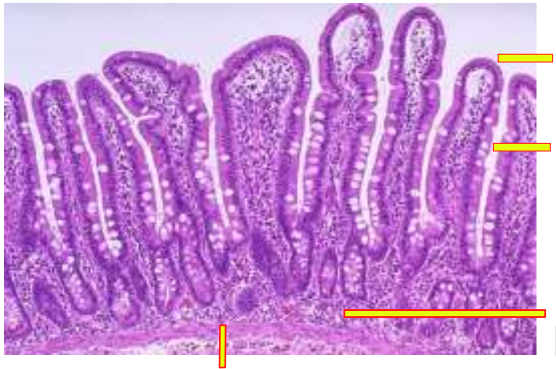
Basophil

Monocyte

Platelet

Lymphocyte

Small Intestine Mucosa



Epithelial cells

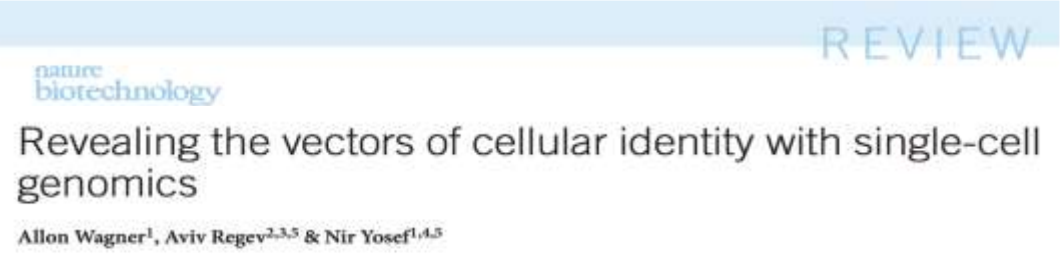
Goblet cells

Lamina propria

Muscularis mucosa (smooth muscle)

Created with figures from library.med.utah/WebPath/HISTHTML/HISTO.html

Cell Identity is More Than Histopathology



REVIEW

nature biotechnology

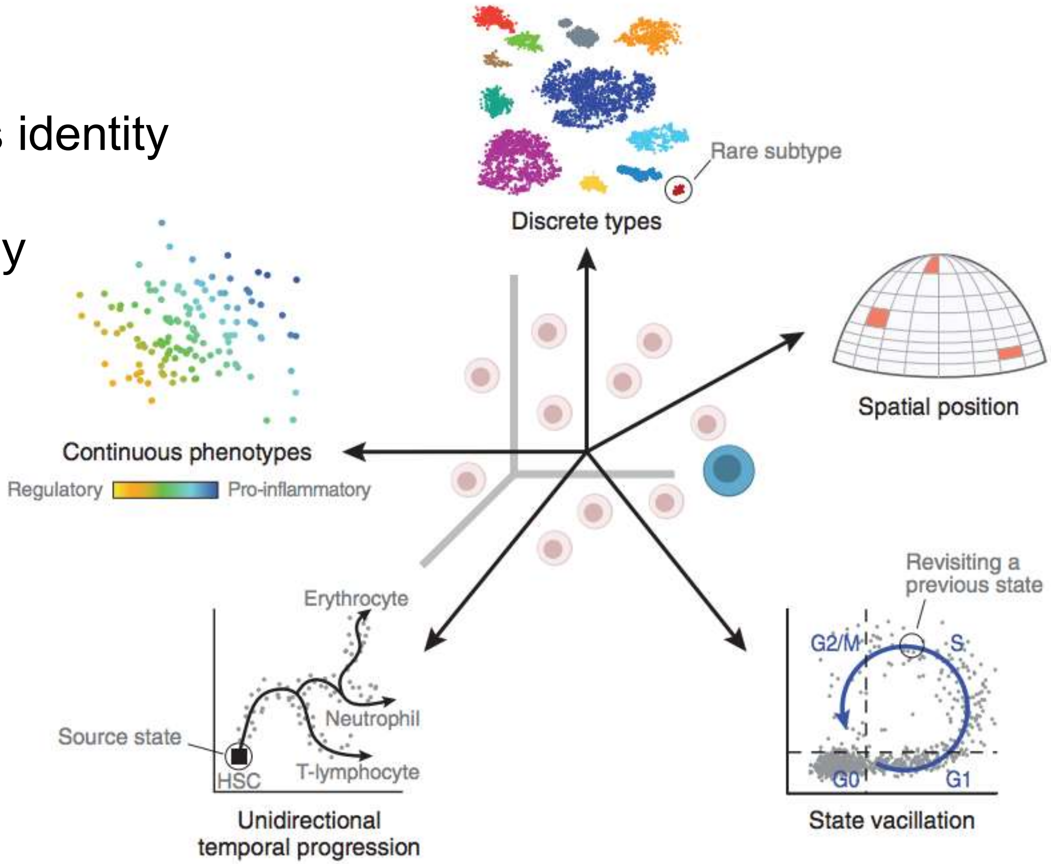
Revealing the vectors of cellular identity with single-cell genomics

Allon Wagner[1], Aviv Regev[2,3,5] & Nir Yosef[1,4,5]

A cell participates in multiple cell contexts.

Multiple factors shape a cell's identity

- Membership in a taxonomy of cell types
- Simultaneous time-dependent processes
- Response to the environment
- Spatial positioning



Before We Get Started

- Single-cell RNA-Seq (scRNA-Seq) analysis methodology is developing.
  - Give you a feel for the data.
  - Perform some analysis together.
- There is a vivid diversity of methodology.
  - These technique will grow as the field does.
  - Why were these specific tools chosen?
- This is a guided conversation through scRNA-Seq analysis.
  - Breadth and targeted depth.
  - There may be other opinions, if you have one, please speak up so we can all learn from it.

Before We Get Started

- Sections will be hands-on.
  - Much can be applied to other analysis.
  - Strengthen those R ninja skills!
  - If you need, cut and pasting is available.
    - cut_and_paste.txt

- There will be many cute corgi pictures.

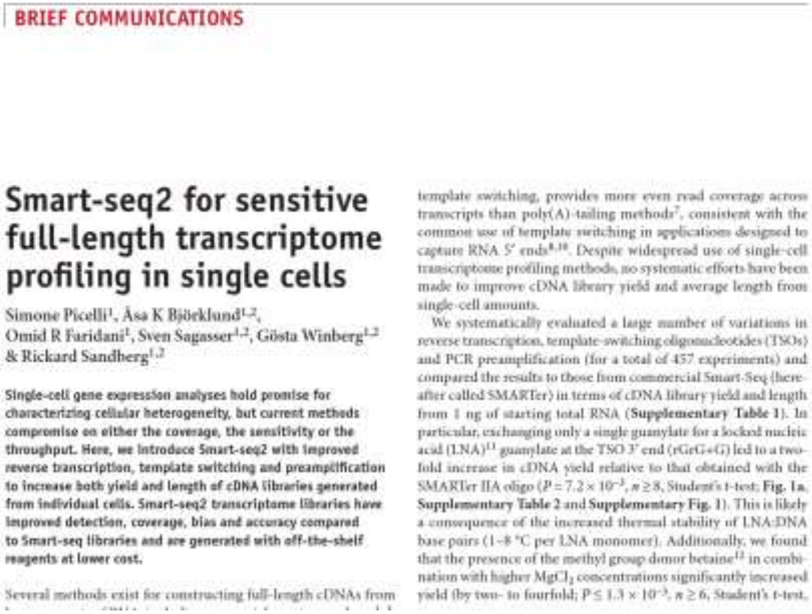
We Will Attempt to Cover

- Describe scRNA-Seq assays.
- Comparing assays.
- Sequence pipelines.
- How do measured counts behave?
- Concerns over study design.
- Initial data exploration.
- Gene and cell filtering.
- Plotting genes.
- Dimensional Reduction and plotting cells.
- Differential expression.
- Communicating your study.

Section: scRNA-Seq Assays

- There are many scRNA-Seq Assays, each differs:
  - Some commercialized
  - Full transcriptome vs 3’
  - Less or more automated
  - Different levels of throughput
  - Differences in cost

Smart-Seq2



BRIEF COMMUNICATIONS

Smart-seq2 for sensitive full-length transcriptome profiling in single cells

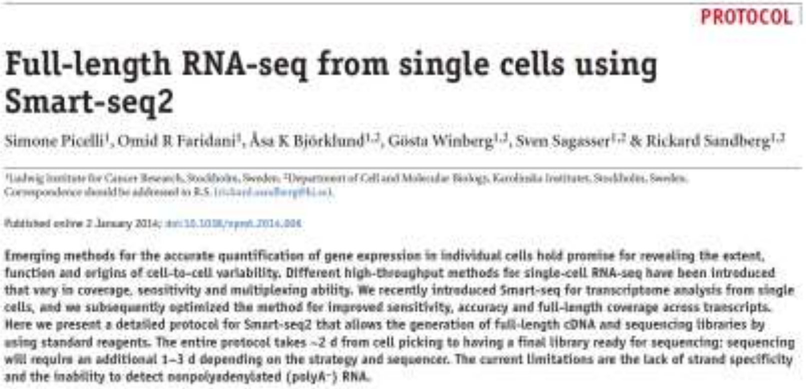
Simone Picelli[1], Åsa K Björklund[1,2], Omid R Faridani[1], Sven Sagasser[1,2], Gösta Winberg[1,2] & Rickard Sandberg[1,2]

Single-cell gene expression analyses hold promise for characterizing cellular heterogeneity, but current methods compromise on either the coverage, the sensitivity or the throughput. Here, we introduce Smart-seq2 with improved reverse transcription, template switching and preamplification to increase both yield and length of cDNA libraries generated from individual cells. Smart-seq2 transcriptome libraries have improved detection, coverage, bias and accuracy compared to Smart-seq libraries and are generated with off-the-shelf reagents at lower cost.

Several methods exist for constructing full-length cDNAs from

template switching, provides more even read coverage across transcripts than poly(A)-tailing methods[7], consistent with the common use of template switching in applications designed to capture RNA 5′ ends[8–10]. Despite widespread use of single-cell transcriptome profiling methods, no systematic efforts have been made to improve cDNA library yield and average length from single-cell amounts.

We systematically evaluated a large number of variations in reverse transcription, template-switching oligonucleotides (TSOs) and PCR preamplification (for a total of 457 experiments) and compared the results to those from commercial Smart-Seq (hereafter called SMARTer) in terms of cDNA library yield and length from 1 ng of starting total RNA (**Supplementary Table 1**). In particular, exchanging only a single guanylate for a locked nucleic acid (LNA)[11] guanylate at the TSO 3′ end (rGrG+G) led to a twofold increase in cDNA yield relative to that obtained with the SMARTer IIA oligo ($P = 7.2 \times 10^{-3}$, $n \geq 8$, Student's t-test; **Fig. 1a**, **Supplementary Table 2** and **Supplementary Fig. 1**). This is likely a consequence of the increased thermal stability of LNA:DNA base pairs (1–8 °C per LNA monomer). Additionally, we found that the presence of the methyl group donor betaine[12] in combination with higher $MgCl_2$ concentrations significantly increased yield (by two- to fourfold; $P \leq 1.3 \times 10^{-3}$, $n \geq 6$, Student's t-test,



PROTOCOL

Full-length RNA-seq from single cells using Smart-seq2

Simone Picelli[1], Omid R Faridani[1], Åsa K Björklund[1,2], Gösta Winberg[1,2], Sven Sagasser[1,2] & Rickard Sandberg[1,2]

[1]Ludwig Institute for Cancer Research, Stockholm, Sweden. [2]Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden. Correspondence should be addressed to R.S. (richard.sandberg@ki.se).

Published online 2 January 2014; doi:10.1038/nprot.2014.006

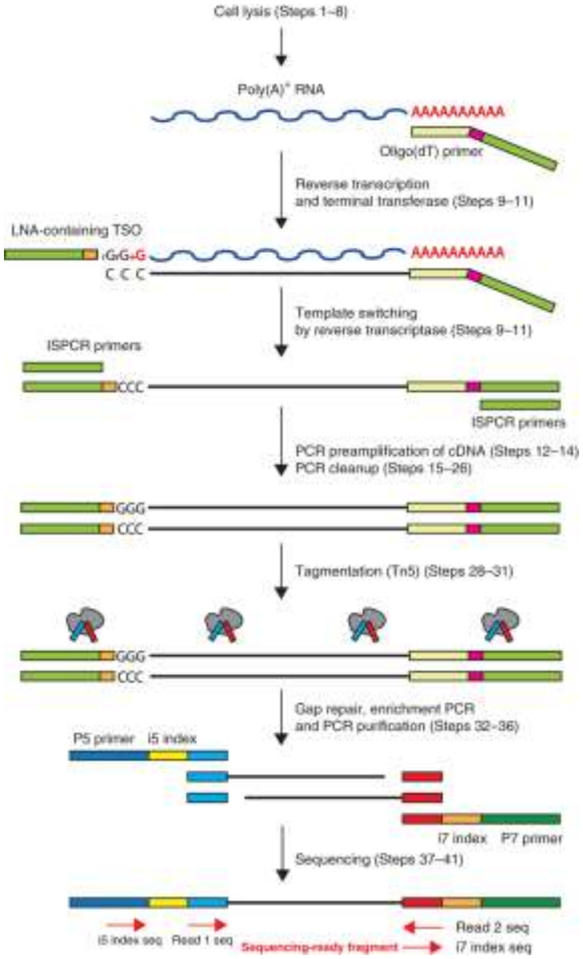
Emerging methods for the accurate quantification of gene expression in individual cells hold promise for revealing the extent, function and origins of cell-to-cell variability. Different high-throughput methods for single-cell RNA-seq have been introduced that vary in coverage, sensitivity and multiplexing ability. We recently introduced Smart-seq for transcriptome analysis from single cells, and we subsequently optimized the method for improved sensitivity, accuracy and full-length coverage across transcripts. Here we present a detailed protocol for Smart-seq2 that allows the generation of full-length cDNA and sequencing libraries by using standard reagents. The entire protocol takes ~2 d from cell picking to having a final library ready for sequencing; sequencing will require an additional 1–3 d depending on the strategy and sequencer. The current limitations are the lack of strand specificity and the inability to detect nonpolyadenylated (polyA⁻) RNA.

Smart-Seq2: Description

Full transcript scRNA-Seq

- Developed for single cell but can performed using total RNA.
- Selects for poly-A tail.
- Full transcript assay.
  - Uses template switching for 5' end capture.
- Standard illumina sequencing.
  - Off-the-shelf products.
- Hundreds of samples.
- Often do not see UMI used.

Smart-Seq2: Assay Overview

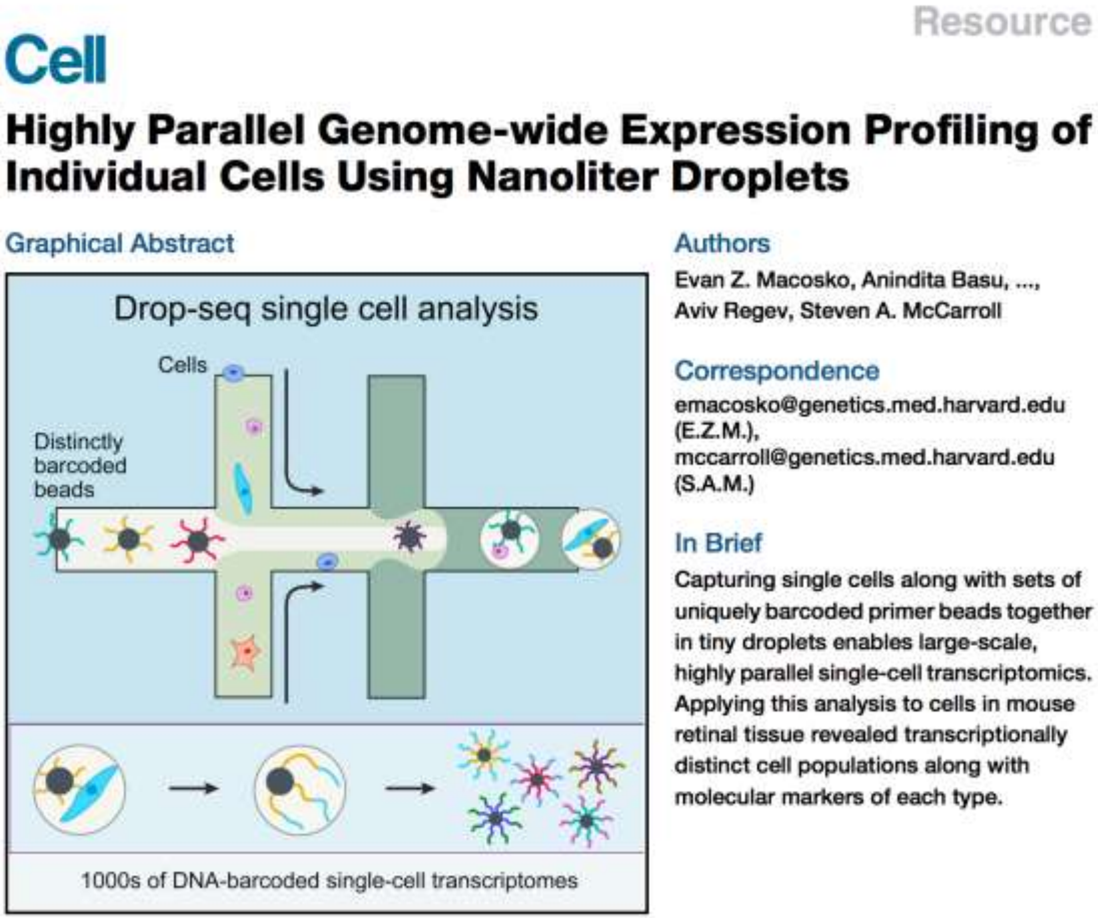


Cell lysis (Steps 1–8)
Poly(A)+ RNA
AAAAAAAAAA
Oligo(dT) primer
Reverse transcription and terminal transferase (Steps 9–11)
LNA-containing TSO
rGrG+G
CCC
Template switching by reverse transcriptase (Steps 9–11)
ISPCR primers
CCC
ISPCR primers
PCR preamplification of cDNA (Steps 12–14)
PCR cleanup (Steps 15–26)
GGG
CCC
Tagmentation (Tn5) (Steps 28–31)
GGG
CCC
Gap repair, enrichment PCR and PCR purification (Steps 32–36)
P5 primer
i5 index
i7 index
P7 primer
Sequencing (Steps 37–41)
i5 index seq
Read 1 seq
Sequencing-ready fragment
Read 2 seq
i7 index seq

- Poly-A capture with 30nt polyT and 25nt 5' anchor sequence.
- RT adding untemplated C
- Template switching
- Locked Nucleic Acid binds to untemplated C
- RT switches template
- Preamplification / cleanup
- DNA fragmentation and adapter ligation together.
- Gap Repair, enrich, purify.

Smart-Seq2: Equipment

Drop-seq



Resource

Cell

Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets

Graphical Abstract

Drop-seq single cell analysis
Cells
Distinctly barcoded beads
1000s of DNA-barcoded single-cell transcriptomes

Authors

Evan Z. Macosko, Anindita Basu, ..., Aviv Regev, Steven A. McCarroll

Correspondence

emacosko@genetics.med.harvard.edu (E.Z.M.), mccarroll@genetics.med.harvard.edu (S.A.M.)

In Brief

Capturing single cells along with sets of uniquely barcoded primer beads together in tiny droplets enables large-scale, highly parallel single-cell transcriptomics. Applying this analysis to cells in mouse retinal tissue revealed transcriptionally distinct cell populations along with molecular markers of each type.

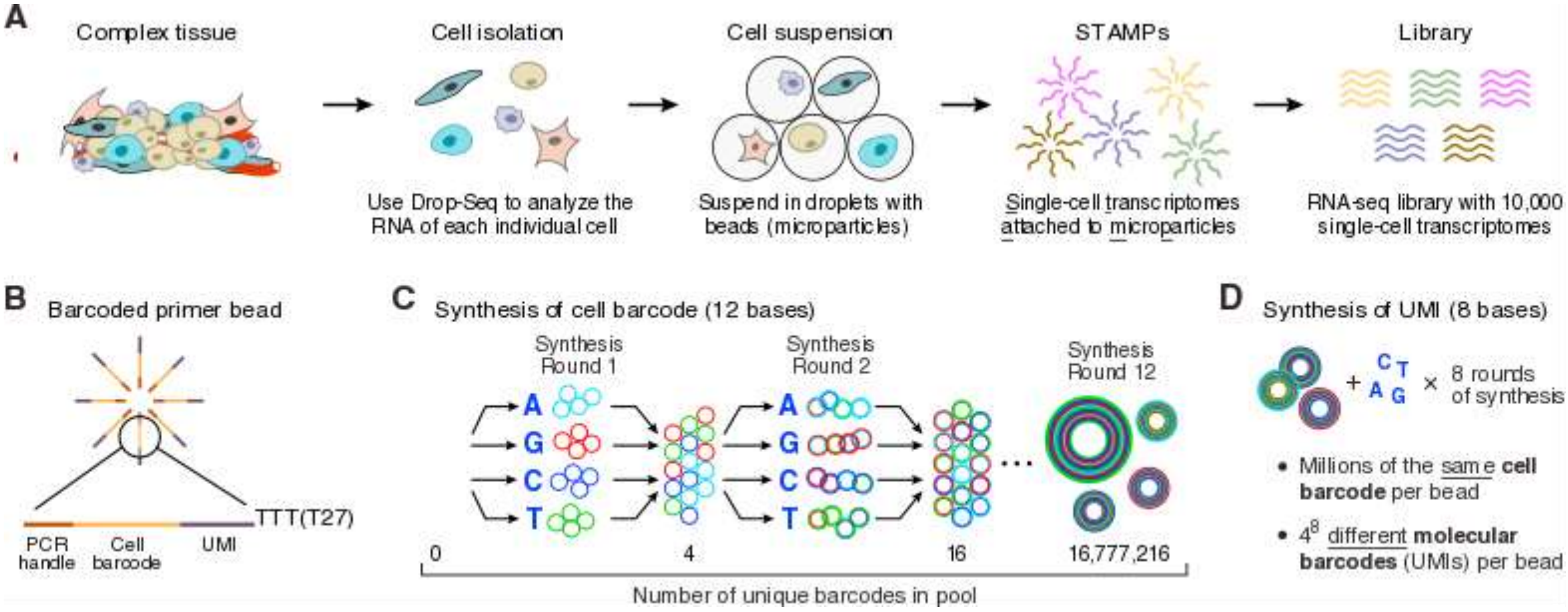
Drop-seq: Description

- Moved throughput from hundreds to thousands.
- Droplet-based processing using microfluidics
- Nanoliter scale aqueous drops in oil.
- 3' End
- Bead based (STAMPs).
- Single-cell transcriptomes attached to microparticles.
- Cell barcodes use split-pool synthesis.
- Uses UMI (Unique Molecular Identifier).
- RMT (Random Molecular Tag).
- Degenerate synthesis.

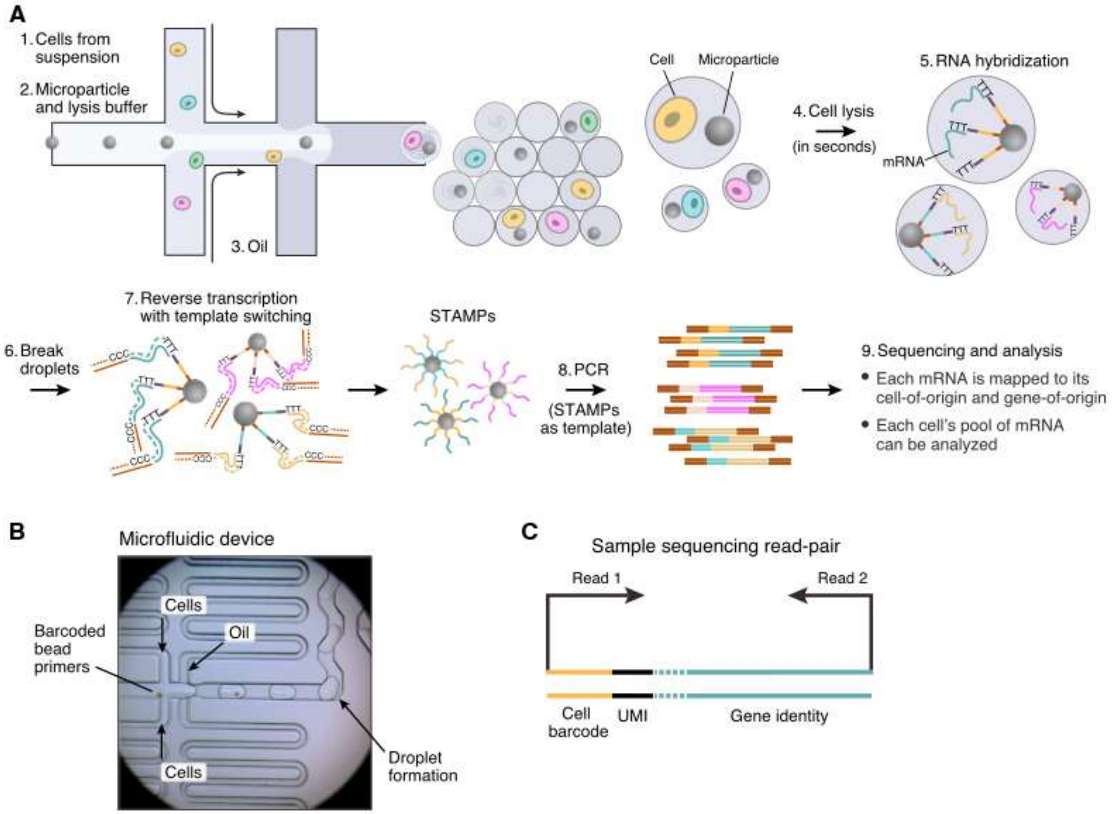
Drop-seq: Overview

- Click Here for Drop-seq Video Abstract

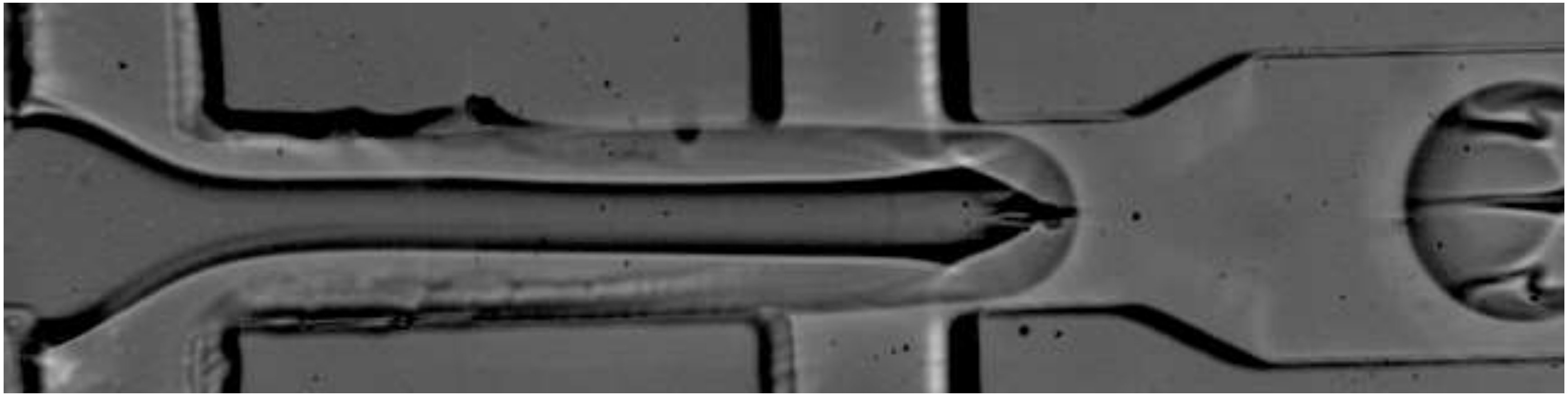
Drop-seq: Assay Overview


A
Complex tissue
Cell isolation
Cell suspension
STAMPs
Library
Use Drop-Seq to analyze the RNA of each individual cell
Suspend in droplets with beads (microparticles)
Single-cell transcriptomes attached to microparticles
RNA-seq library with 10,000 single-cell transcriptomes
B
Barcoded primer bead
TTT(T27)
PCR handle
Cell barcode
UMI
C
Synthesis of cell barcode (12 bases)
Synthesis Round 1
Synthesis Round 2
Synthesis Round 12
A
G
C
T
A
G
C
T
0
4
16
16,777,216
Number of unique barcodes in pool
D
Synthesis of UMI (8 bases)
+ C T A G × 8 rounds of synthesis
• Millions of the same cell barcode per bead
• 4^8 different molecular barcodes (UMIs) per bead

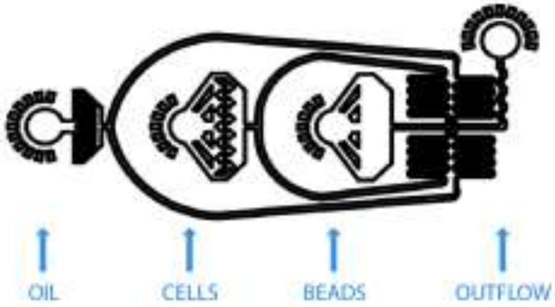
Drop-seq: Assay Overview



Drop-seq: Assay Overview



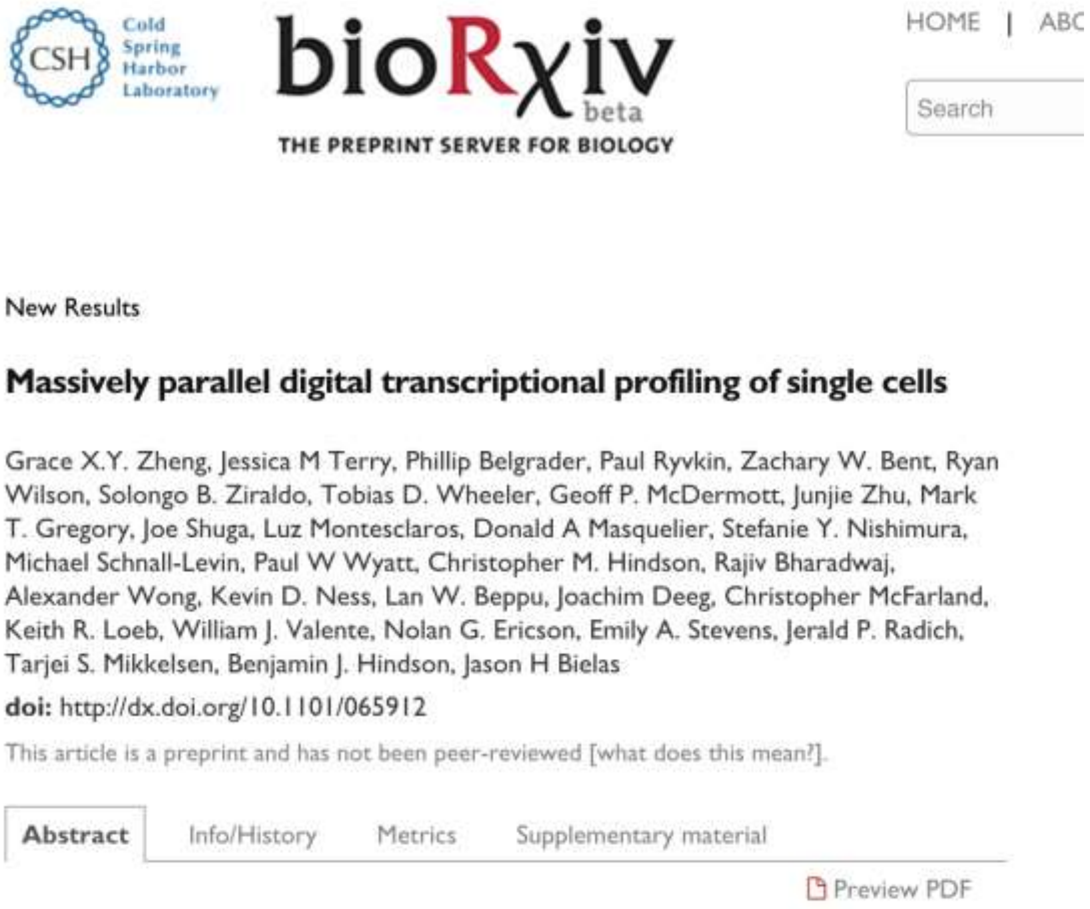
Drop-seq: Equipment



Drop-seq: Pointers

- Droplet-based assays can have leaky RNA.
- Before library generation wash off any medium (inhibits library generation).
- Adding PBS and BSA (0.05-0.01%) can protect the cell.
  - Too much produces a residue making harvesting the beads difficult.
- Filter all reagent with a 80 micron strainer before microfluidics.
- Some purchased devices add a hydrophobic coating.
  - Can deteriorate (2 months at best).
  - Recoating does work (in-house).

10X: Massively Parallel Sequencing


CSH
Cold Spring Harbor Laboratory
bioRxiv beta
THE PREPRINT SERVER FOR BIOLOGY
HOME | ABC
Search

New Results

Massively parallel digital transcriptional profiling of single cells

Grace X.Y. Zheng, Jessica M Terry, Phillip Belgrader, Paul Ryvkin, Zachary W. Bent, Ryan Wilson, Solongo B. Ziraldo, Tobias D. Wheeler, Geoff P. McDermott, Junjie Zhu, Mark T. Gregory, Joe Shuga, Luz Montesclaros, Donald A Masquelier, Stefanie Y. Nishimura, Michael Schnall-Levin, Paul W Wyatt, Christopher M. Hindson, Rajiv Bharadwaj, Alexander Wong, Kevin D. Ness, Lan W. Beppu, Joachim Deeg, Christopher McFarland, Keith R. Loeb, William J. Valente, Nolan G. Ericson, Emily A. Stevens, Jerald P. Radich, Tarjei S. Mikkelsen, Benjamin J. Hindson, Jason H Bielas

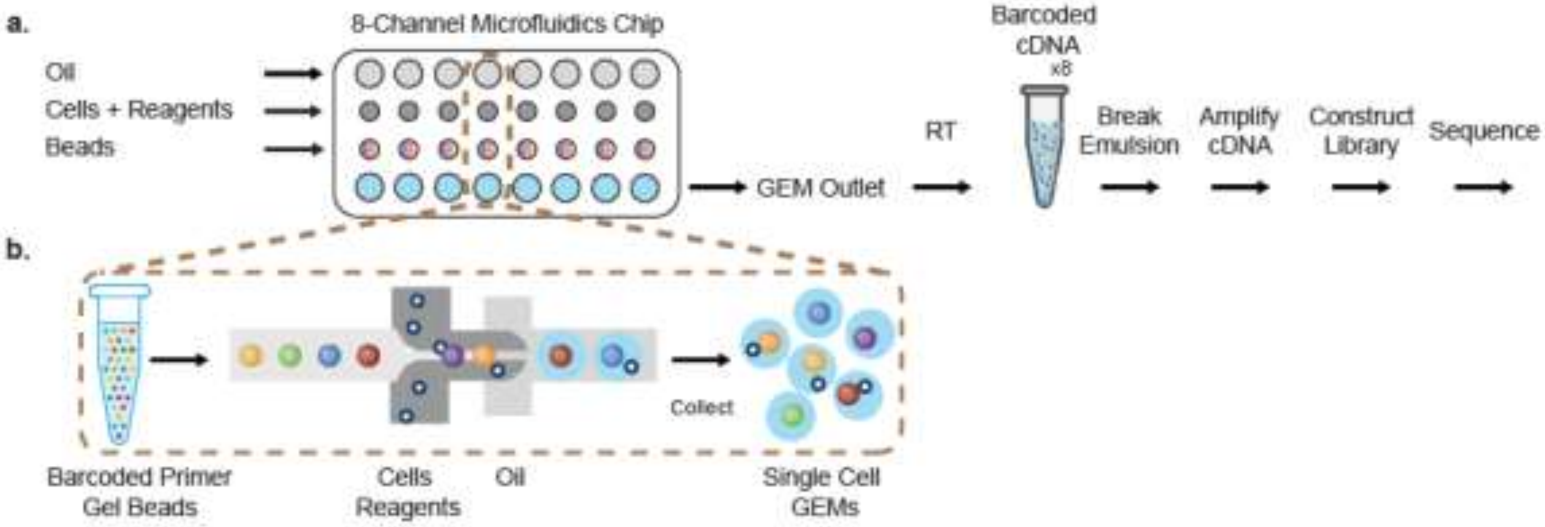
doi: http://dx.doi.org/10.1101/065912

This article is a preprint and has not been peer-reviewed [what does this mean?].

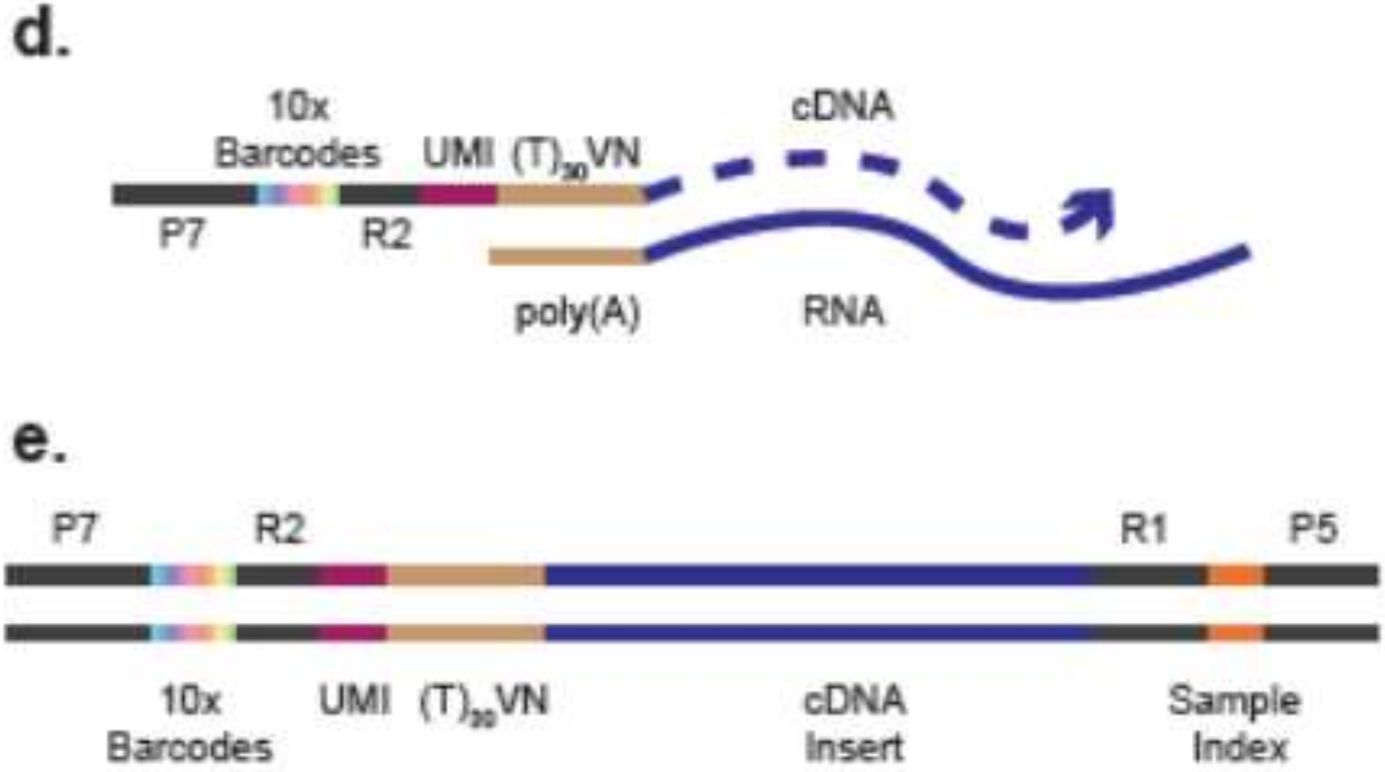
10X: Description

- Droplet-based, 3' mRNA.
  - GEM (Gel Bead in Emulsion)
- Standardized instrumentation and reagents.
- More high-throughput scaling to tens of thousands.
- Less processing time.
- Cell Ranger software is available for install.

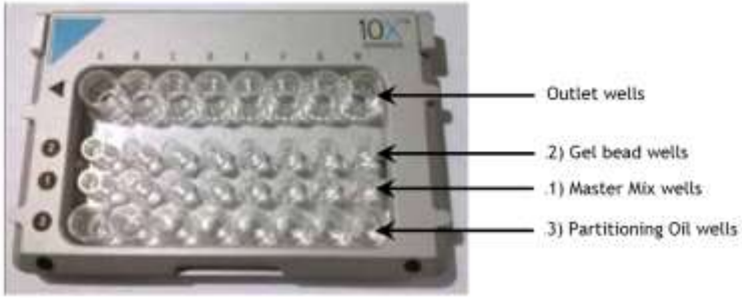
10X: Assay Overview



10X: Assay Overview



10X: Equipment

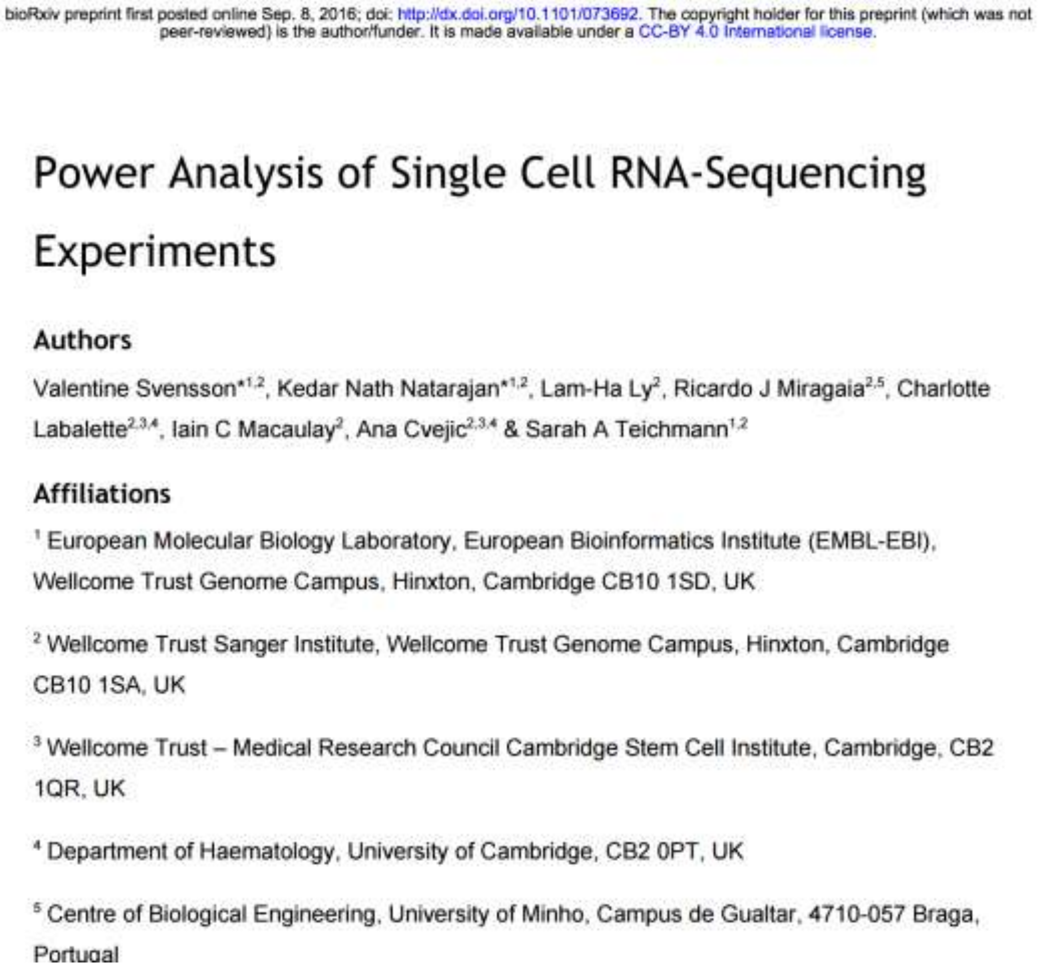


A Word on Sorting

- After disassociating cells cells can be performed.
- Know your cells, are they sticky, are they big?
  - Select an appropriate sized nozzle.
- Don't sort too quickly (1-2k cells per second or lower)
  - The slower the more time cells sit in lysis after sorting
  - 10 minutes max in lysis (some say 30 minutes)
- Calibrate speed of instrument with beads
  - Check alignment every 5-6 plates
- Afterwards spin down to make sure cells are in lysis buffer
  - Flash freeze
- Chloe Villani on sorting [click here]

Section: Comparing scRNA-Seq Assays

scRNA-Seq Assay Performance



Power Analysis of Single Cell RNA-Sequencing Experiments

Authors

Valentine Svensson*[1,2], Kedar Nath Natarajan*[1,2], Lam-Ha Ly[2], Ricardo J Miragaia[2,5], Charlotte Labalette[2,3,4], Iain C Macaulay[2], Ana Cvejic[2,3,4] & Sarah A Teichmann[1,2]

Affiliations

[1] European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK

[2] Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

[3] Wellcome Trust – Medical Research Council Cambridge Stem Cell Institute, Cambridge, CB2 1QR, UK

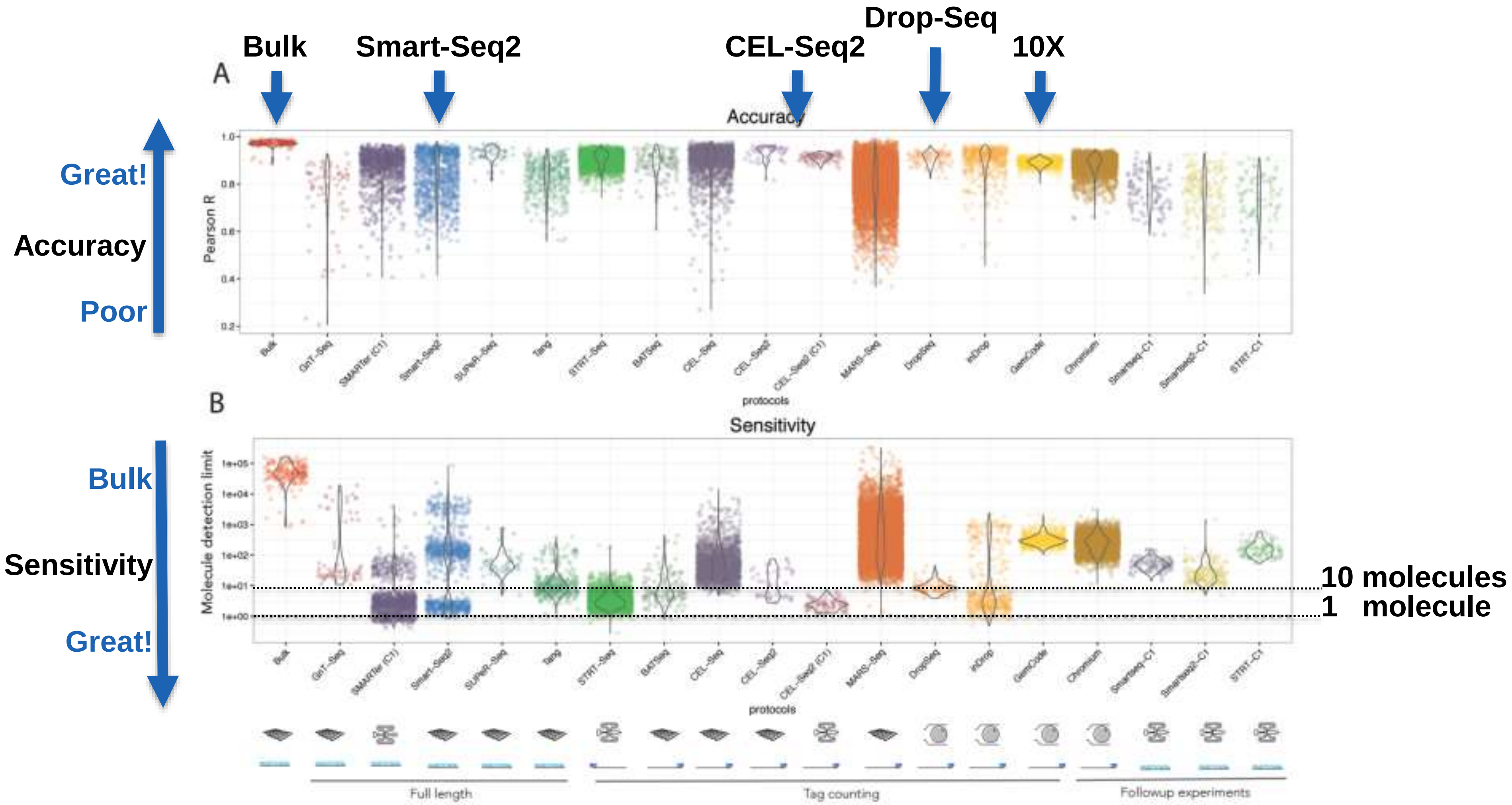
[4] Department of Haematology, University of Cambridge, CB2 0PT, UK

[5] Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal

ERCC-based Benchmarking

- Based on ERCC spike-ins.
  - Exogenous RNA-Spikins
  - No secondary structure
  - 25b polyA Tail
    - May be a conservative measurement given endogenous mRNA will have ~250b polyA.
- Accuracy
  - How well the abundance levels correlated with known spiked-in amounts.
- Sensitivity
  - Minimum number of input RNA molecules required to detect a spike-in.

Sensitivity and Specificity



Bulk

Smart-Seq2

CEL-Seq2

Drop-Seq

10X

Great!

Accuracy

Poor

Bulk

Sensitivity

Great!

10 molecules

1 molecule

Final Thoughts

- Different assays have different throughput.
  - SmartSeq2 < Drop-seq < 10X
- SmartSeq2 is full transcript.
- Plate-based methods get lysed in wells and so do not leak.
  - Droplet-based can have leaky RNA.
- In Drop-seq assays RT happens outside the droplets
  - Can use harsher lysis buffers.
  - 10X needs lysis buffers compatible with the RT enzyme.
- 10X is more standardized and comes with a pipeline.
  - Drop-seq is more customizable but more hands-on.
- Cost per library varies greatly.

Section: scRNA-Seq Pipelines

Sequences Differ So Pipelines Differ

- scRNA-Seq assays are different and produce different sequences
  - The sequence pipelines must be tailored to the sequence of interest.
  - Many pipelines are NOT compatible but many show similarities.

Start with FASTQ Sequences

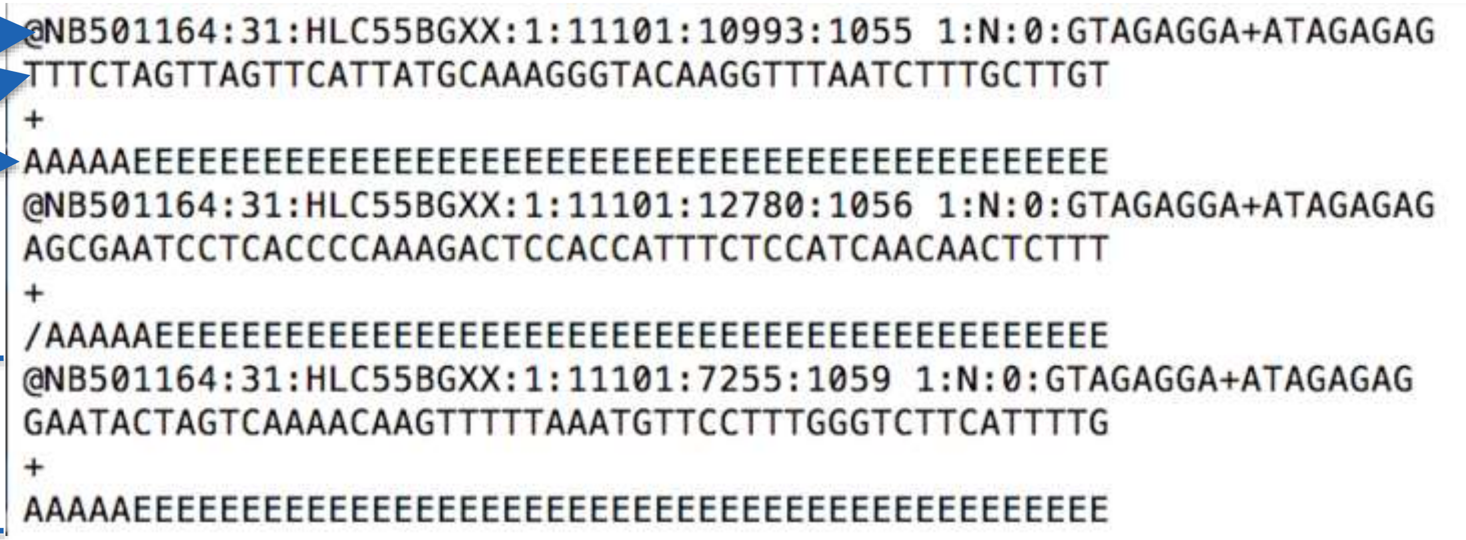
FASTQ File Format

Sequence Header

cDNA Sequence

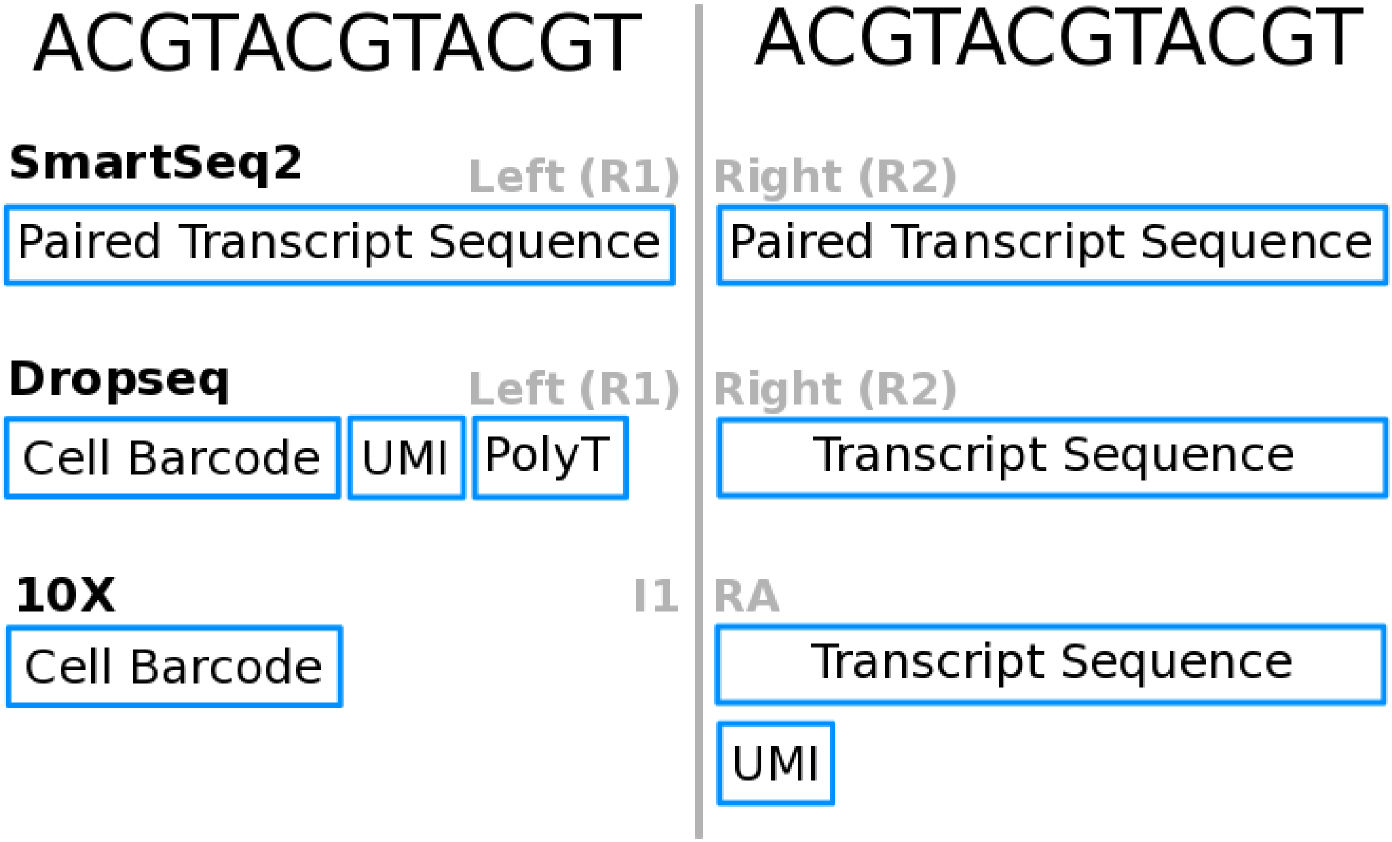
Base Quality

4 Lines are 1 sequence



```
@NB501164:31:HLC55BGXX:1:11101:10993:1055 1:N:0:GTAGAGGA+ATAGAGAG
TTTCTAGTTAGTTCATTATGCAAAGGGTACAAGGTTTAATCTTTGCTTGT
+
AAAAAEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
@NB501164:31:HLC55BGXX:1:11101:12780:1056 1:N:0:GTAGAGGA+ATAGAGAG
AGCGAATCCTCACCCCAAAGACTCCACCATTTCTCCATCAACAACTCTTT
+
/AAAAAEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
@NB501164:31:HLC55BGXX:1:11101:7255:1059 1:N:0:GTAGAGGA+ATAGAGAG
GAATACTAGTCAAAACAAGTTTTTAAATGTTCCTTTGGGTCTTCATTTTG
+
AAAAAEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE
```

Assays Differ in FASTQ Contents



ACGTACGTACGT

ACGTACGTACGT

SmartSeq2

Left (R1)

Paired Transcript Sequence

Right (R2)

Paired Transcript Sequence

Dropseq

Left (R1)

Cell Barcode

UMI

PolyT

Right (R2)

Transcript Sequence

10X

I1

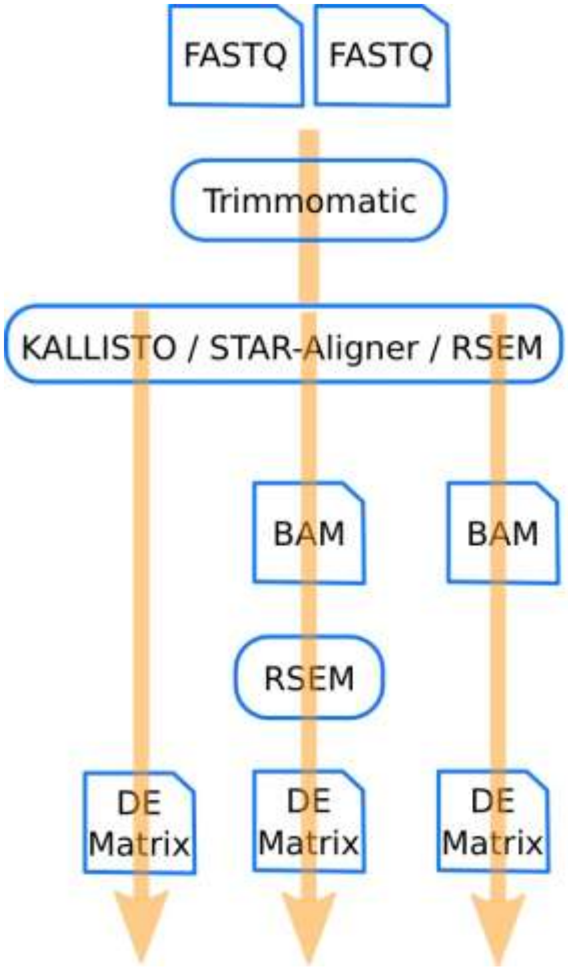
Cell Barcode

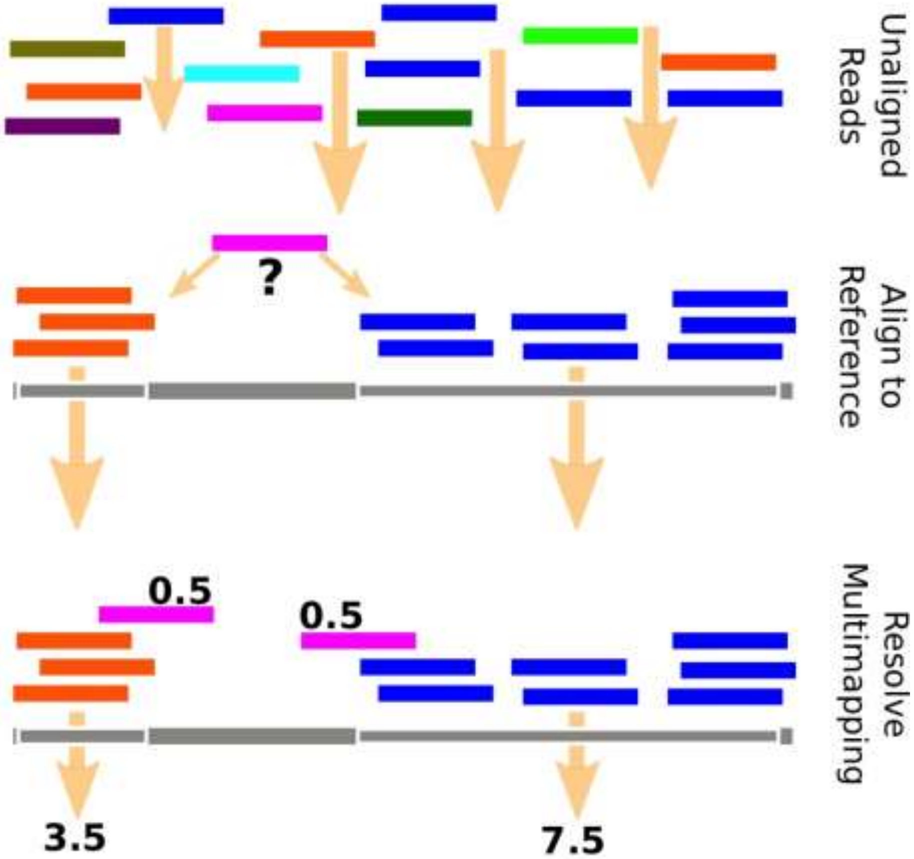
RA

Transcript Sequence

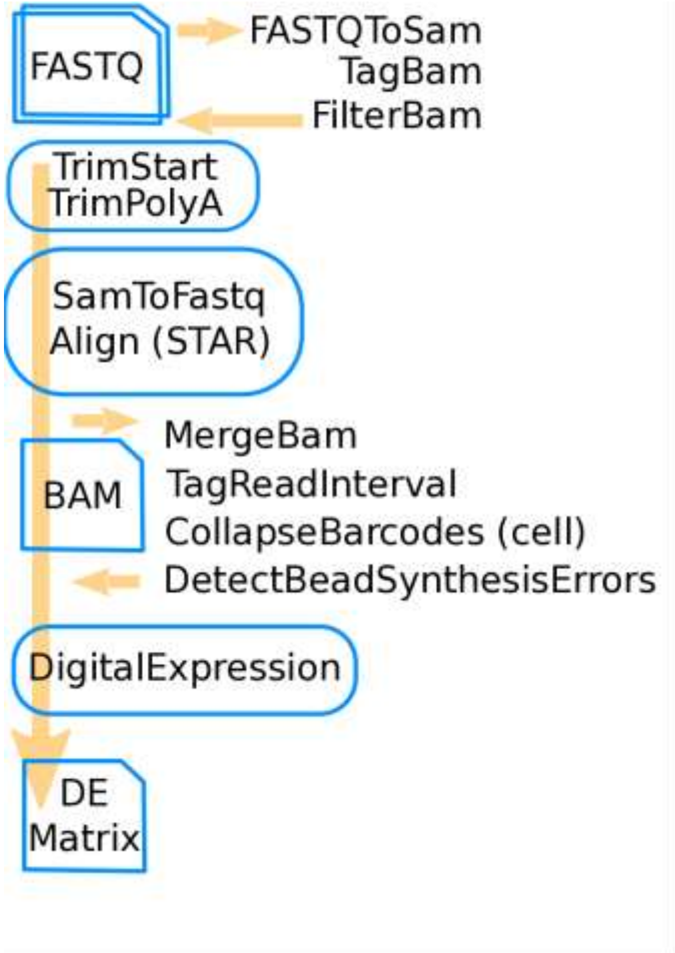
UMI

SmartSeq2: Pipeline Overview



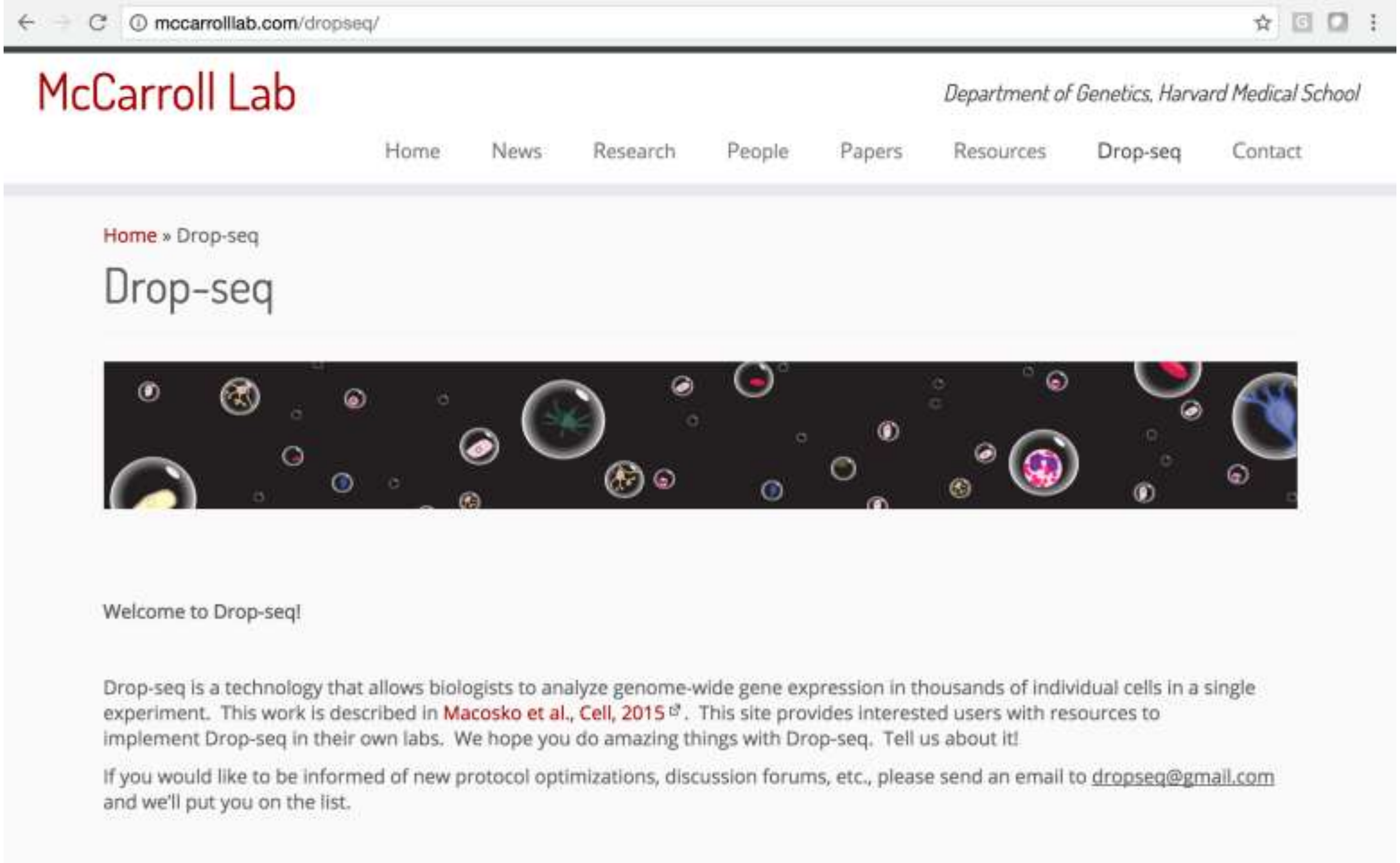


Drop-seq: Pipeline Overview



- Common functionality: trimming, alignment, generating count matrix.
- Adds book keeping for cell barcodes and UMIs, bead error detection, cell barcode collapsing, UMI collapsing.

Drop-seq: Further Help



McCarroll Lab

Department of Genetics, Harvard Medical School

Home » Drop-seq

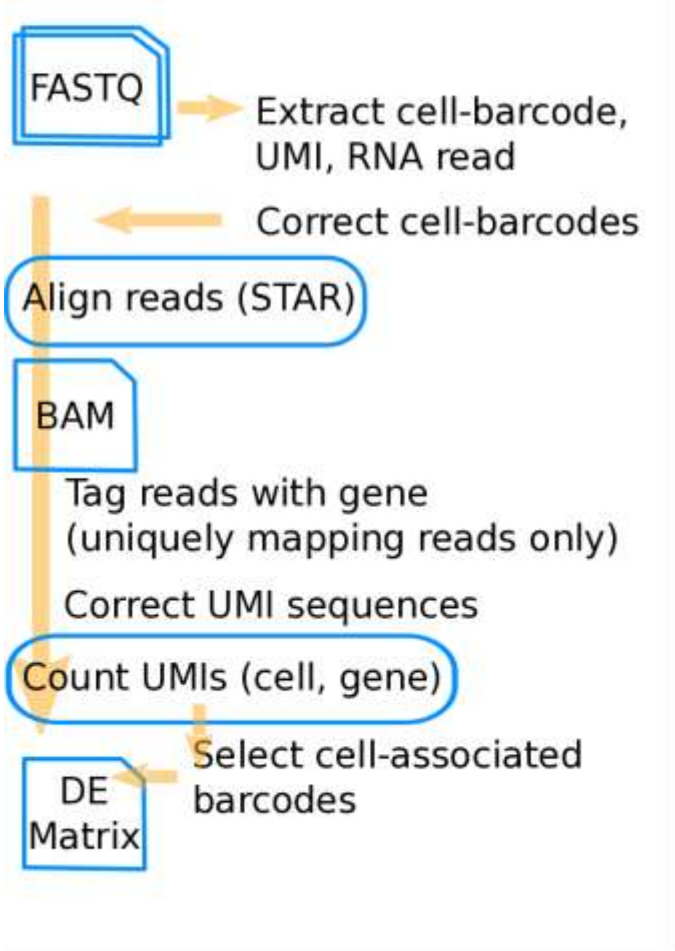
Drop-seq

Welcome to Drop-seq!

Drop-seq is a technology that allows biologists to analyze genome-wide gene expression in thousands of individual cells in a single experiment. This work is described in Macosko et al., Cell, 2015. This site provides interested users with resources to implement Drop-seq in their own labs. We hope you do amazing things with Drop-seq. Tell us about it!

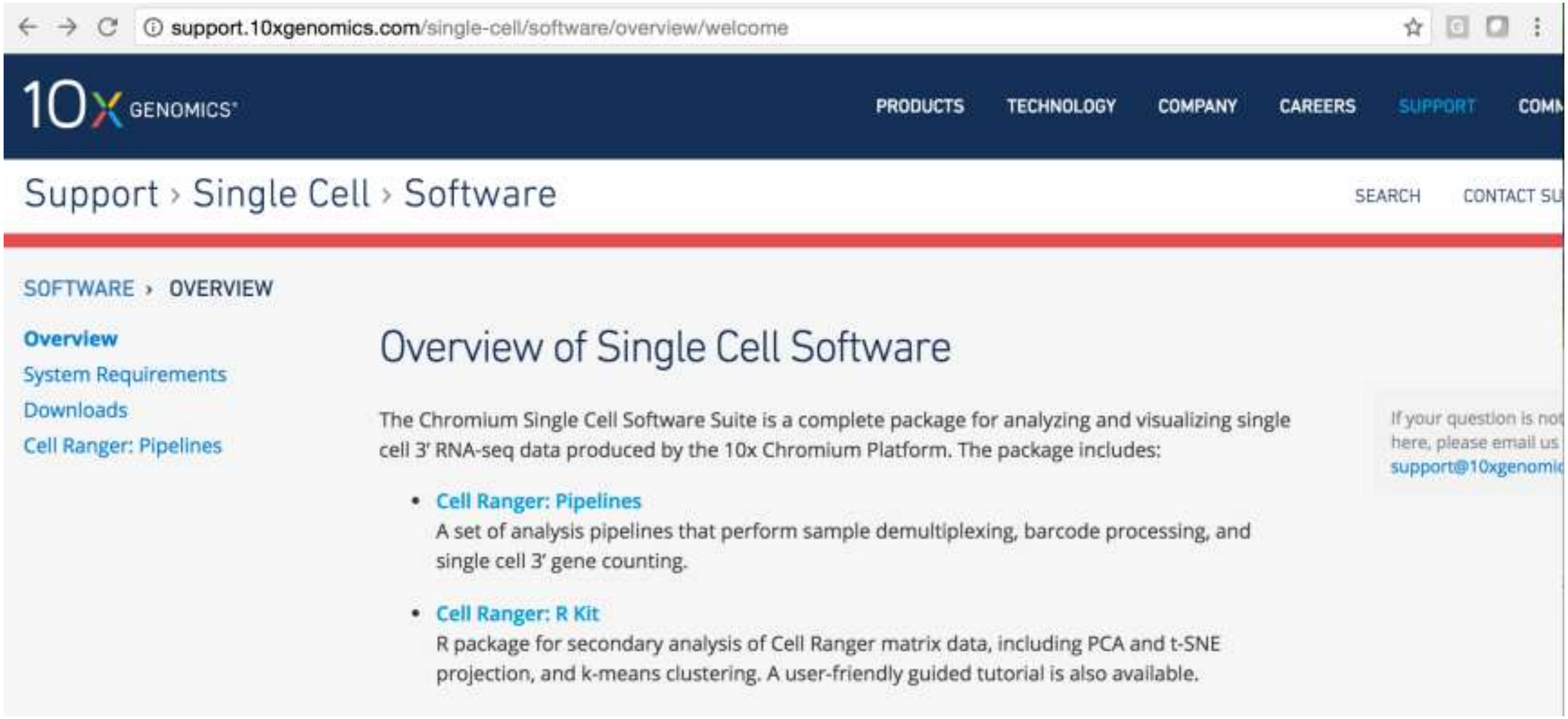
If you would like to be informed of new protocol optimizations, discussion forums, etc., please send an email to dropseq@gmail.com and we'll put you on the list.

10X: Pipeline Overview


FASTQ
Extract cell-barcode, UMI, RNA read
Correct cell-barcodes
Align reads (STAR)
BAM
Tag reads with gene (uniquely mapping reads only)
Correct UMI sequences
Count UMIs (cell, gene)
Select cell-associated barcodes
DE Matrix

- Steps conceptually similar to the Drop-seq pipeline.

10X: Further Help



Support › Single Cell › Software

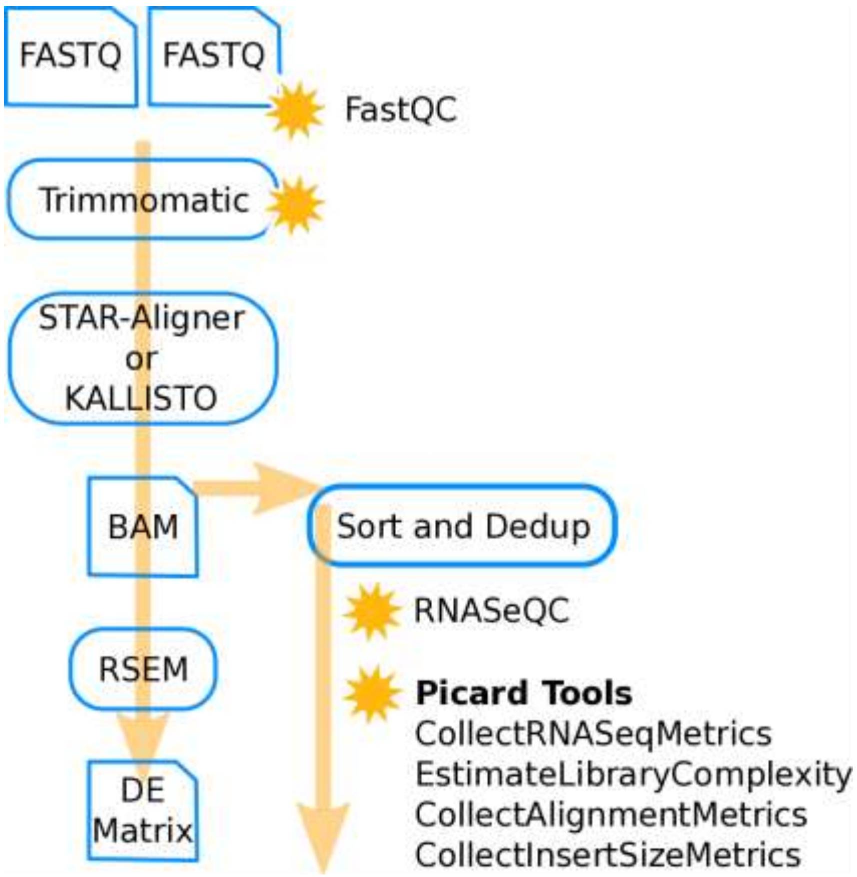
Overview of Single Cell Software

The Chromium Single Cell Software Suite is a complete package for analyzing and visualizing single cell 3′ RNA-seq data produced by the 10x Chromium Platform. The package includes:

- **Cell Ranger: Pipelines**
  A set of analysis pipelines that perform sample demultiplexing, barcode processing, and single cell 3′ gene counting.
- **Cell Ranger: R Kit**
  R package for secondary analysis of Cell Ranger matrix data, including PCA and t-SNE projection, and k-means clustering. A user-friendly guided tutorial is also available.

If your question is not
here, please email us
support@10xgenomic

Sequence Level Quality Control



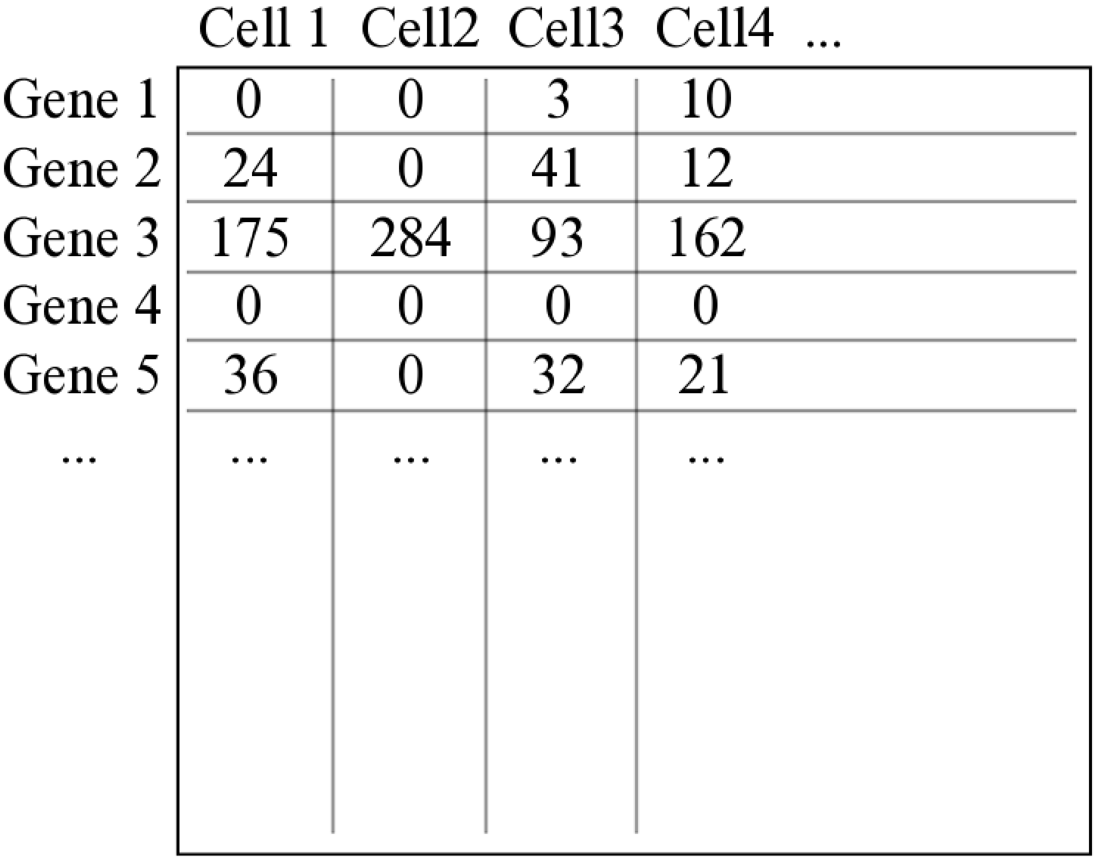
- Much of the QC that is performed is using traditional tools.

Pipeline Section Summary

- Single-cell RNA-Seq is a diverse ecosystem of assays.
  - Each assay has pros and cons.
- Sequences derived from these assays are complex and vary.
- Different pipelines are needed to address different sequence formats.
  - Common steps include:
    - Aligning
    - QC
    - Read counting.

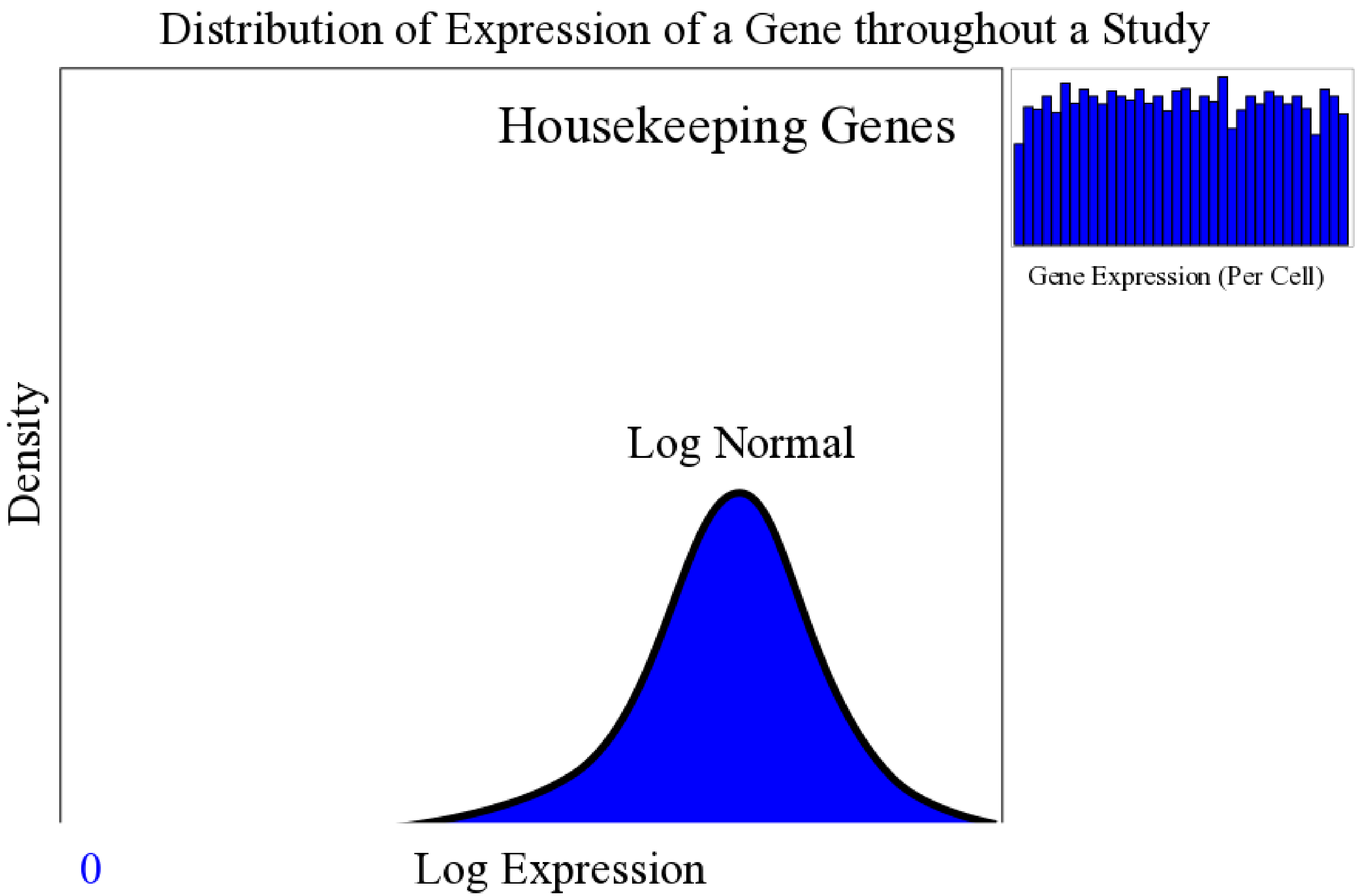
Section: scRNA-Seq Count Data

This is an Expression Matrix

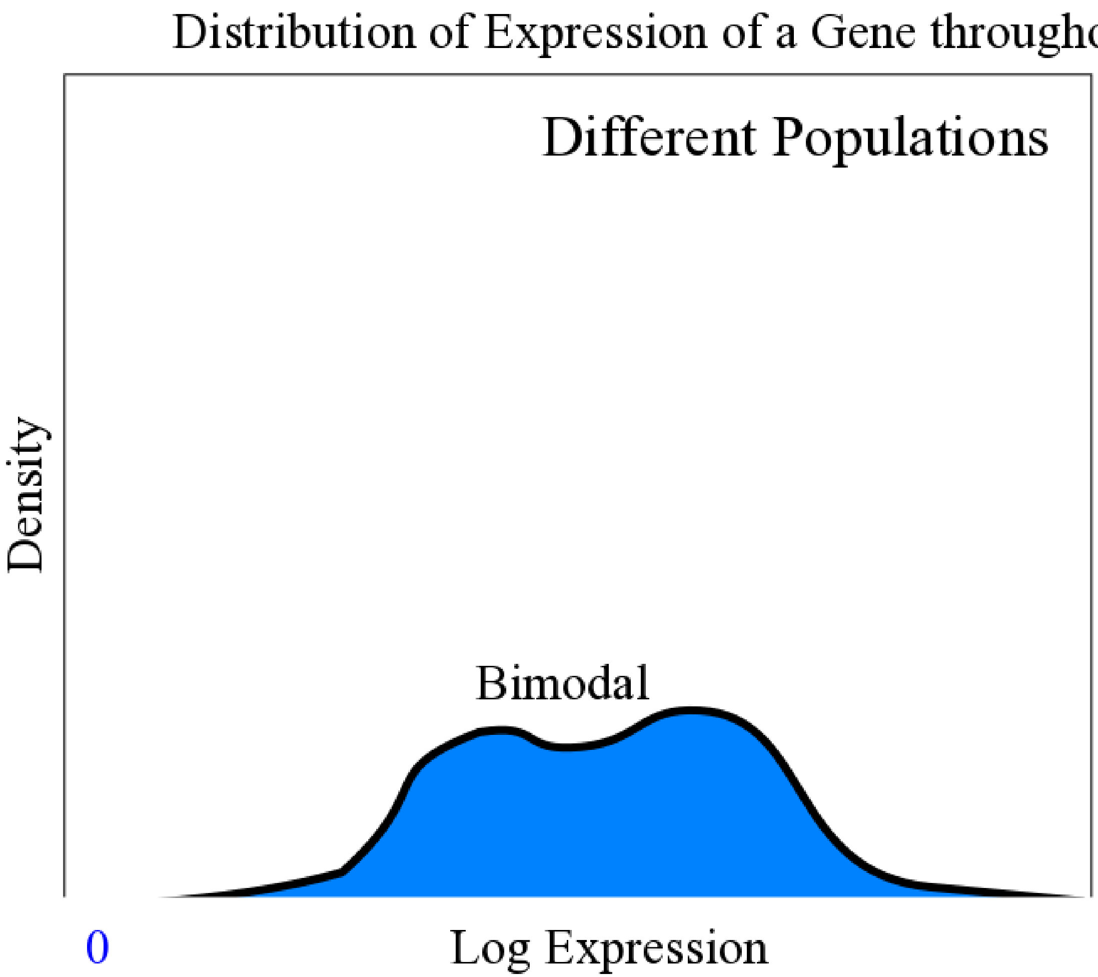


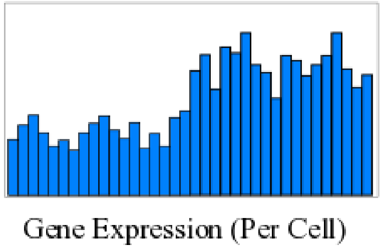
	Cell 1	Cell2	Cell3	Cell4	...
Gene 1	0	0	3	10	
Gene 2	24	0	41	12	
Gene 3	175	284	93	162	
Gene 4	0	0	0	0	
Gene 5	36	0	32	21	
...	...	...	...	...	

Genes Have Different Distributions

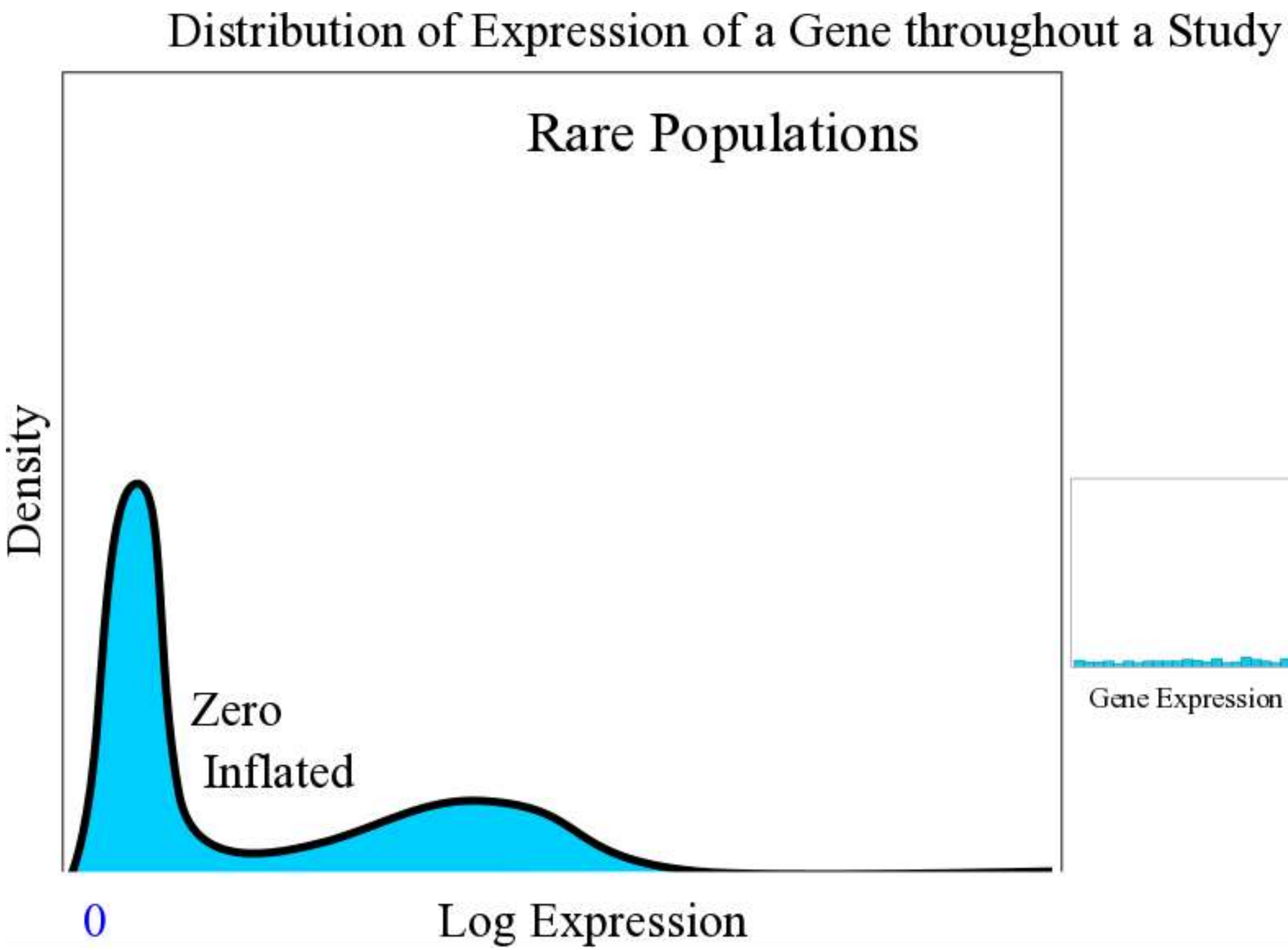


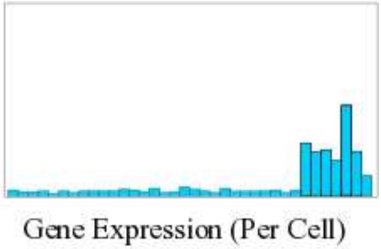
Genes Have Different Distributions



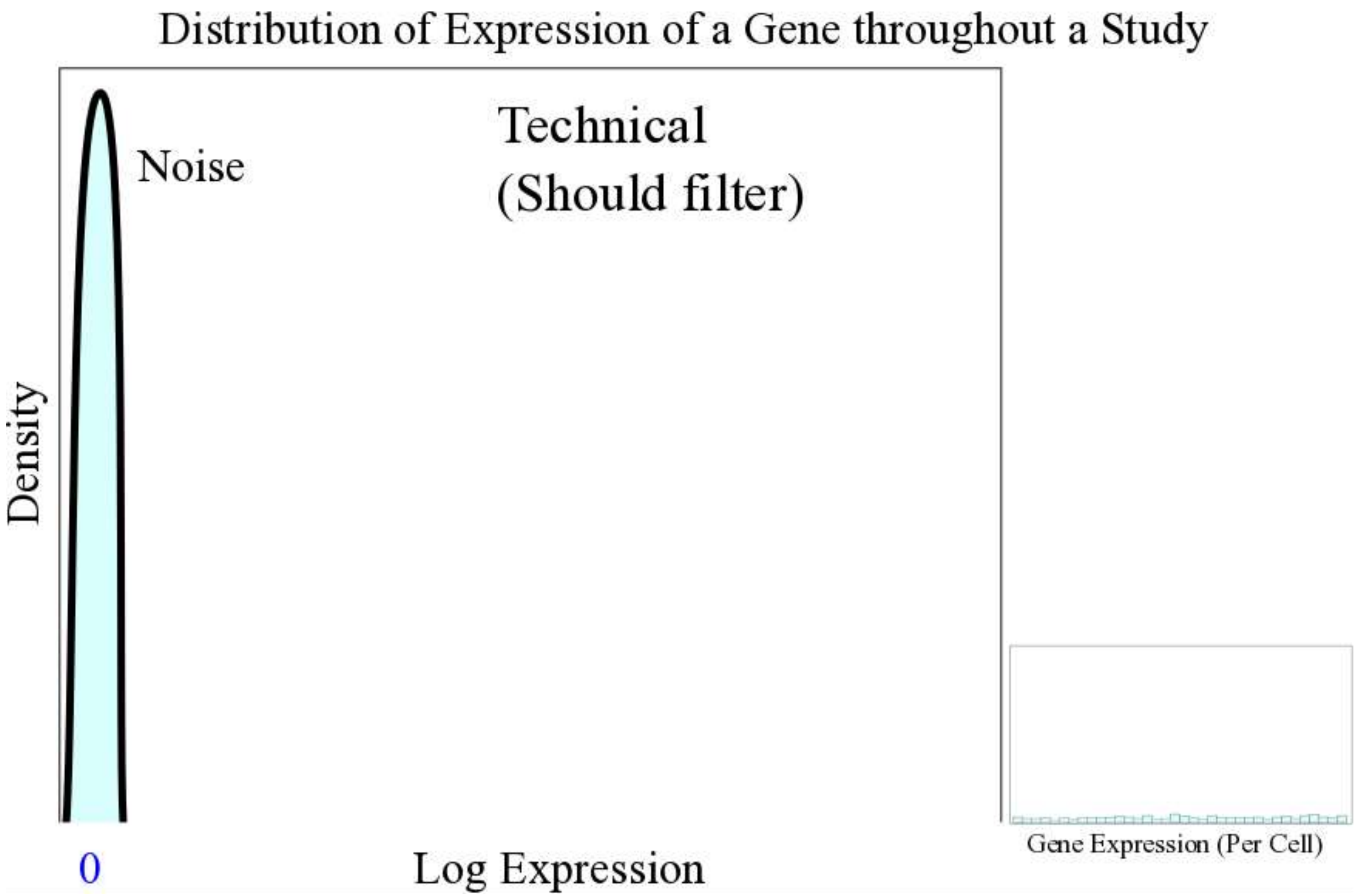


Genes Have Different Distributions

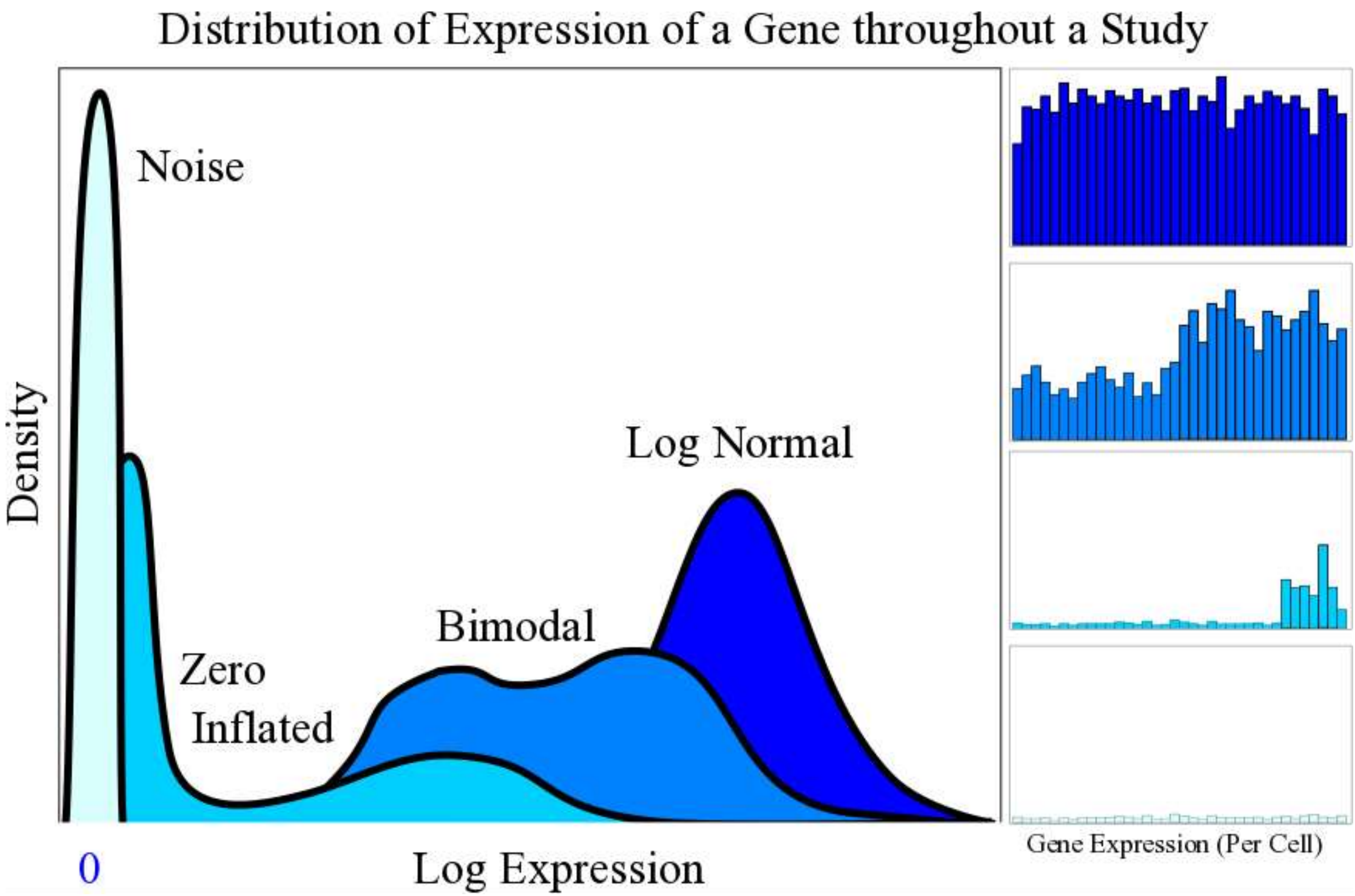




Genes Have Different Distributions

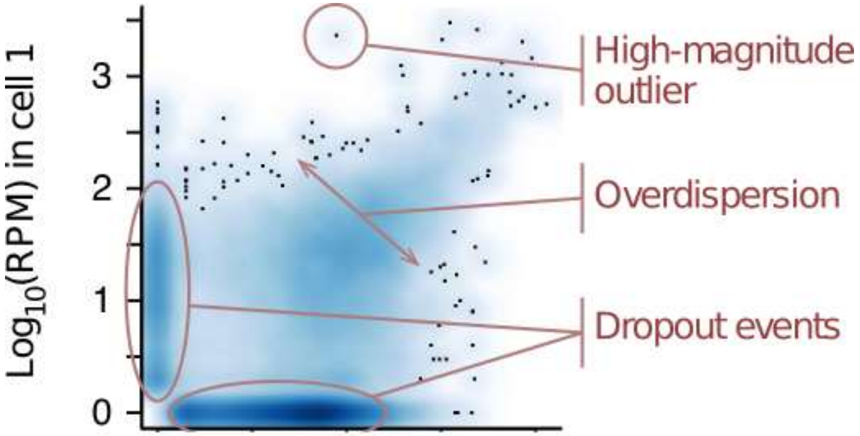


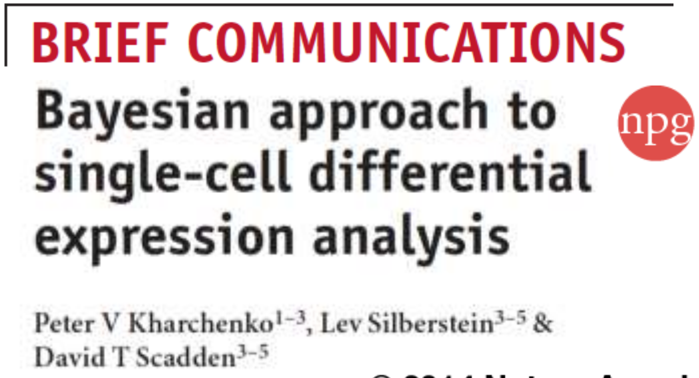
Genes Have Different Distributions



Underlying Biology

- Zero inflation.
  - Drop-out event during reverse-transcription.
  - Genes with more expression have less zeros.
  - Complexity varies.
- Transcription stochasticity.
  - Transcription bursting.
  - Coordinated transcription of multigene networks.
  - Over-dispersed counts.
- Higher Resolution.
  - More sources of signal



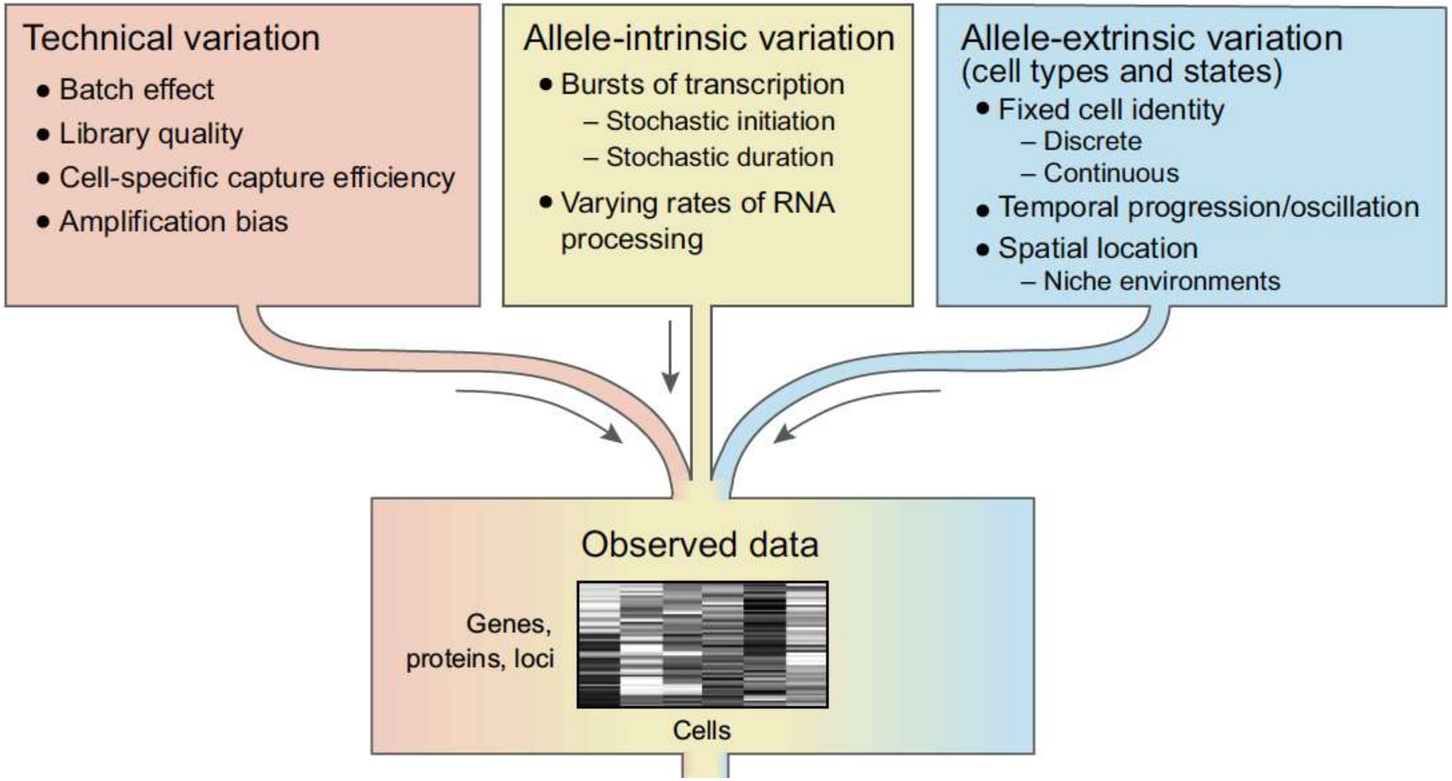


BRIEF COMMUNICATIONS

Bayesian approach to single-cell differential expression analysis

Peter V Kharchenko[1-3], Lev Silberstein[3-5] & David T Scadden[3-5]

Expression has Many Sources per Cell



Technical variation

- Batch effect
- Library quality
- Cell-specific capture efficiency
- Amplification bias

Allele-intrinsic variation

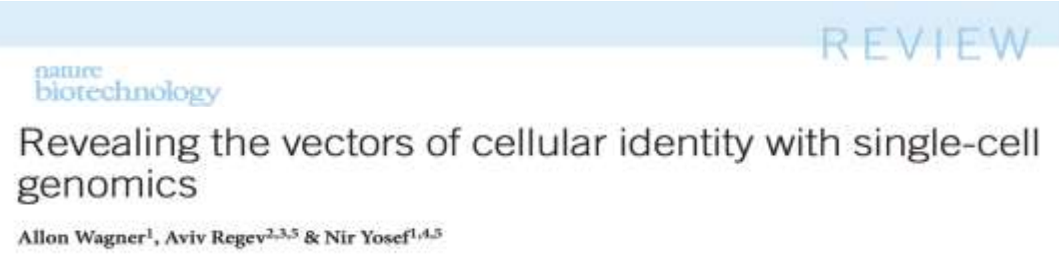
- Bursts of transcription
  - Stochastic initiation
  - Stochastic duration
- Varying rates of RNA processing

Allele-extrinsic variation (cell types and states)

- Fixed cell identity
  - Discrete
  - Continuous
- Temporal progression/oscillation
- Spatial location
  - Niche environments

Observed data

Genes, proteins, loci
Cells



REVIEW

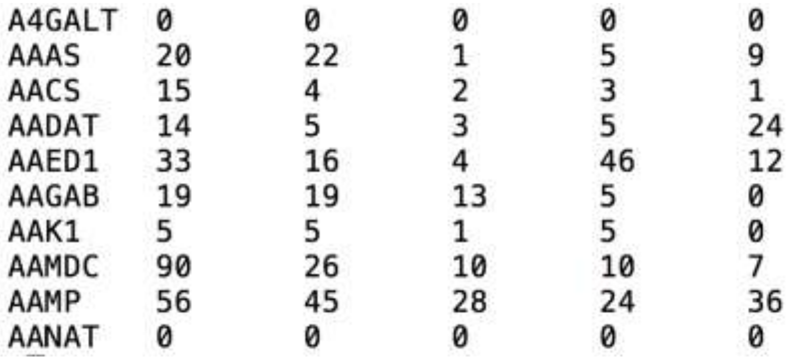
nature biotechnology

Revealing the vectors of cellular identity with single-cell genomics

Allon Wagner[1], Aviv Regev[2,3,5] & Nir Yosef[1,4,5]

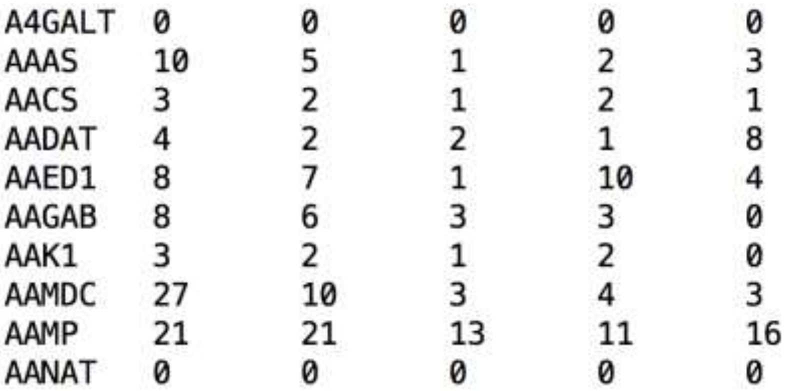
Data Analysis with UMIs

Read Counts



A4GALT	0	0	0	0	0
AAAS	20	22	1	5	9
AACS	15	4	2	3	1
AADAT	14	5	3	5	24
AAED1	33	16	4	46	12
AAGAB	19	19	13	5	0
AAK1	5	5	1	5	0
AAMDC	90	26	10	10	7
AAMP	56	45	28	24	36
AANAT	0	0	0	0	0

Counts by UMI



A4GALT	0	0	0	0	0
AAAS	10	5	1	2	3
AACS	3	2	1	2	1
AADAT	4	2	2	1	8
AAED1	8	7	1	10	4
AAGAB	8	6	3	3	0
AAK1	3	2	1	2	0
AAMDC	27	10	3	4	3
AAMP	21	21	13	11	16
AANAT	0	0	0	0	0

Collapsed but Not Linear

Summary of the Data

- We are still understanding scData and how to apply it.
  - Data can be NOT normal.
  - Data can be Zero-inflated.
  - Data can be very noisy.
  - Cells vary in library complexity.
  - Can represent many “basis vectors” or sources of expression simultaneously.
- Keeping these characteristics in analysis assumptions.
- Tend to filter more conservatively with UMIs.

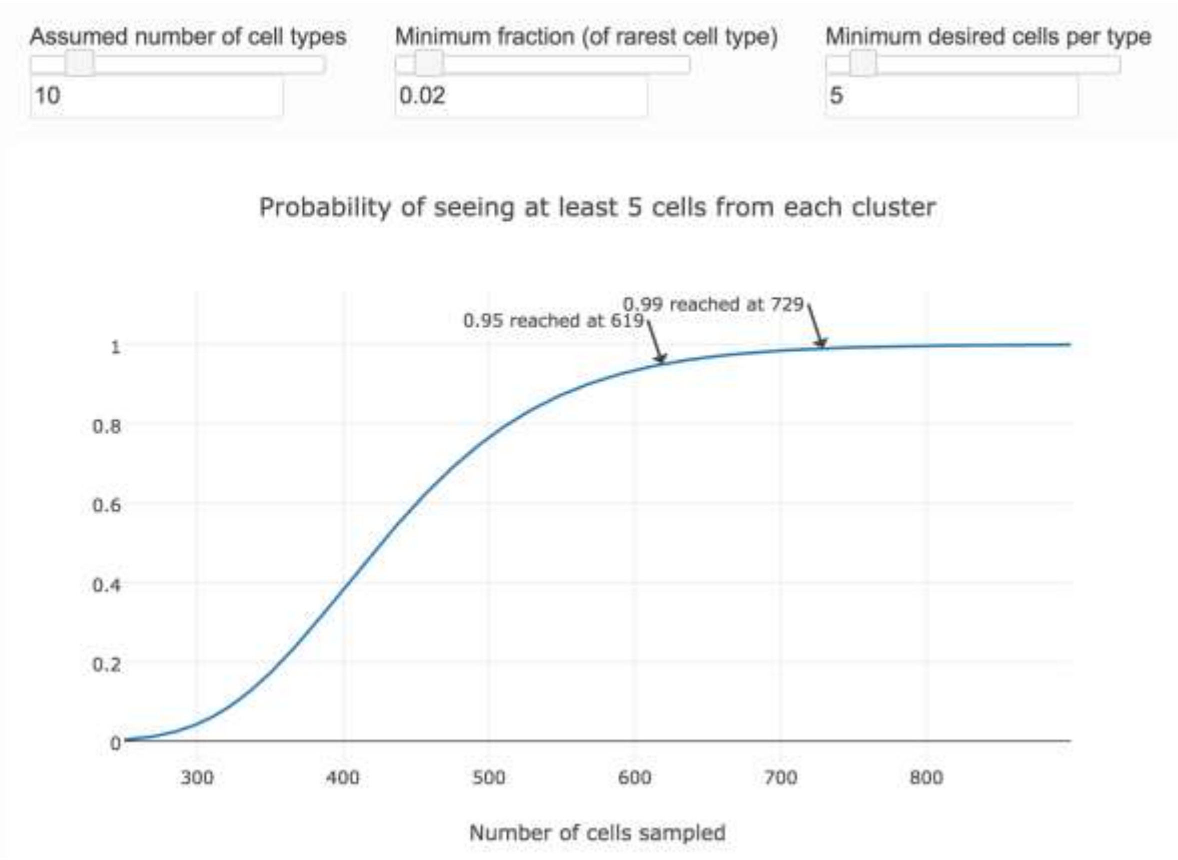
Section: Study Design and scRNA-Seq

scRNA-Seq Study Design

- How many cells?
  - Can change depending on the variability of the biology and the expectation of finding rare populations.
- How to design cell capture?
  - Single cell RNA-Seq is especially prone to technical batch affects (due to processing).
- Use of UMIs
- Use of ERCC spike-ins

How Many Cells?

- Satija lab online tool
  - satijalab.org/howmanycells



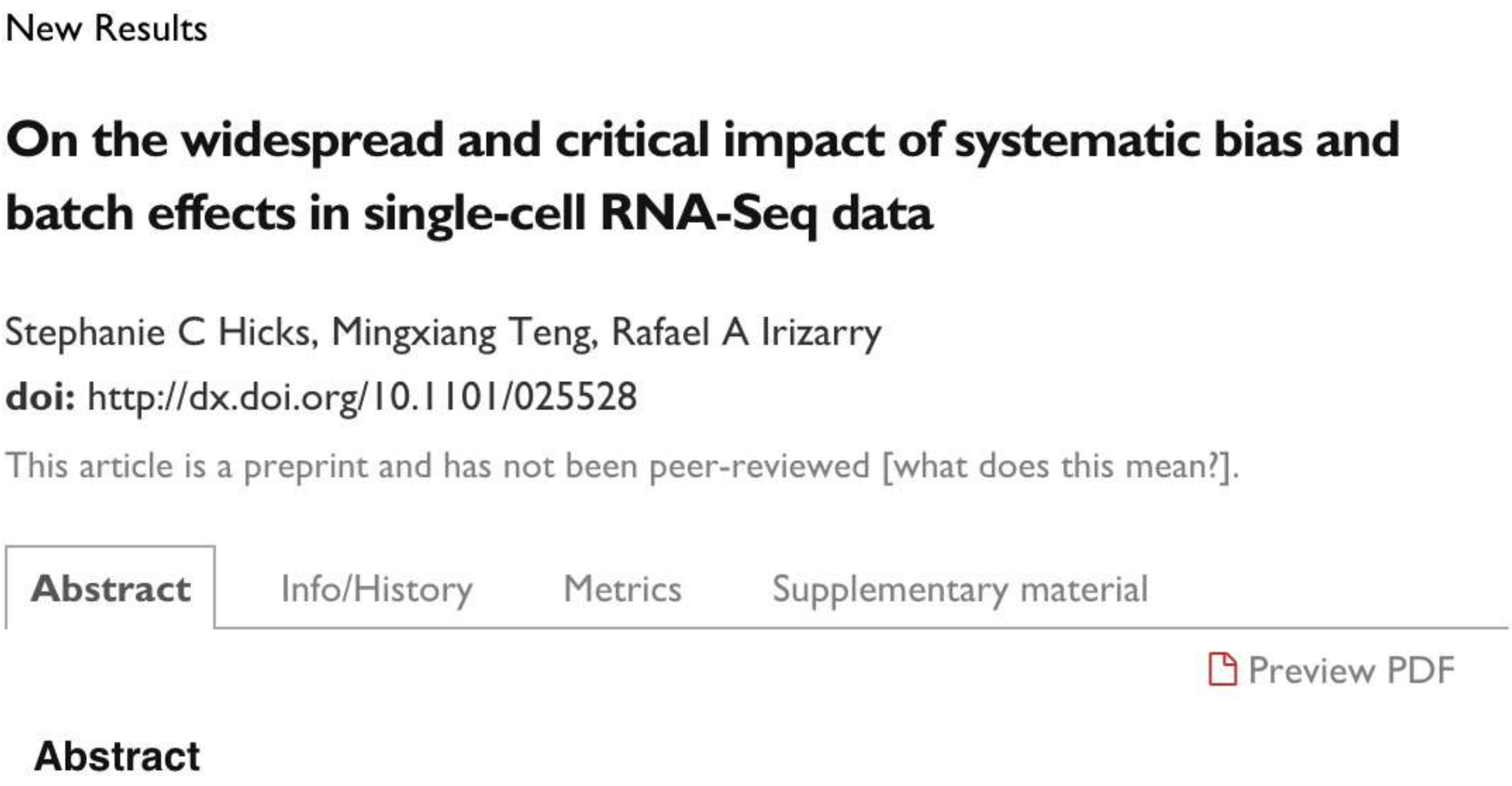
Assumed number of cell types: 10

Minimum fraction (of rarest cell type): 0.02

Minimum desired cells per type: 5

Probability of seeing at least 5 cells from each cluster

Single Cell RNA-Seq and Batch Affects



New Results

On the widespread and critical impact of systematic bias and batch effects in single-cell RNA-Seq data

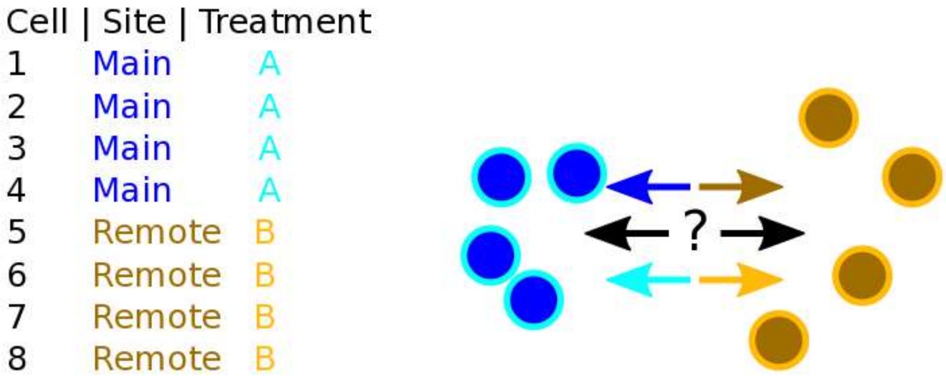
Stephanie C Hicks, Mingxiang Teng, Rafael A Irizarry

doi: http://dx.doi.org/10.1101/025528

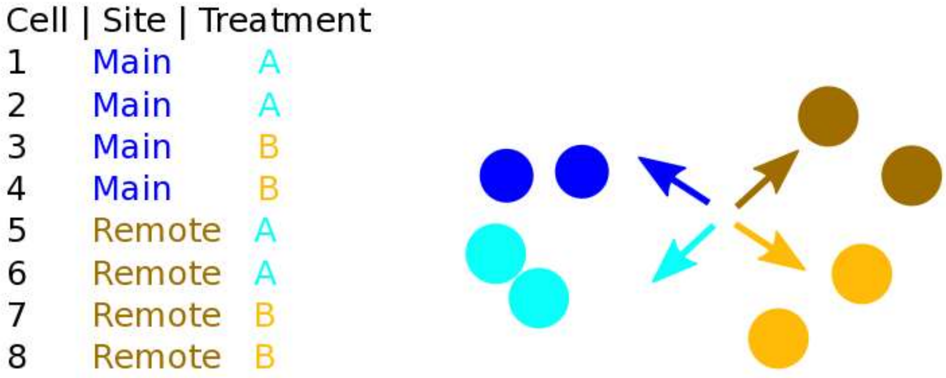
This article is a preprint and has not been peer-reviewed [what does this mean?].

Abstract

What is Study Confounding?

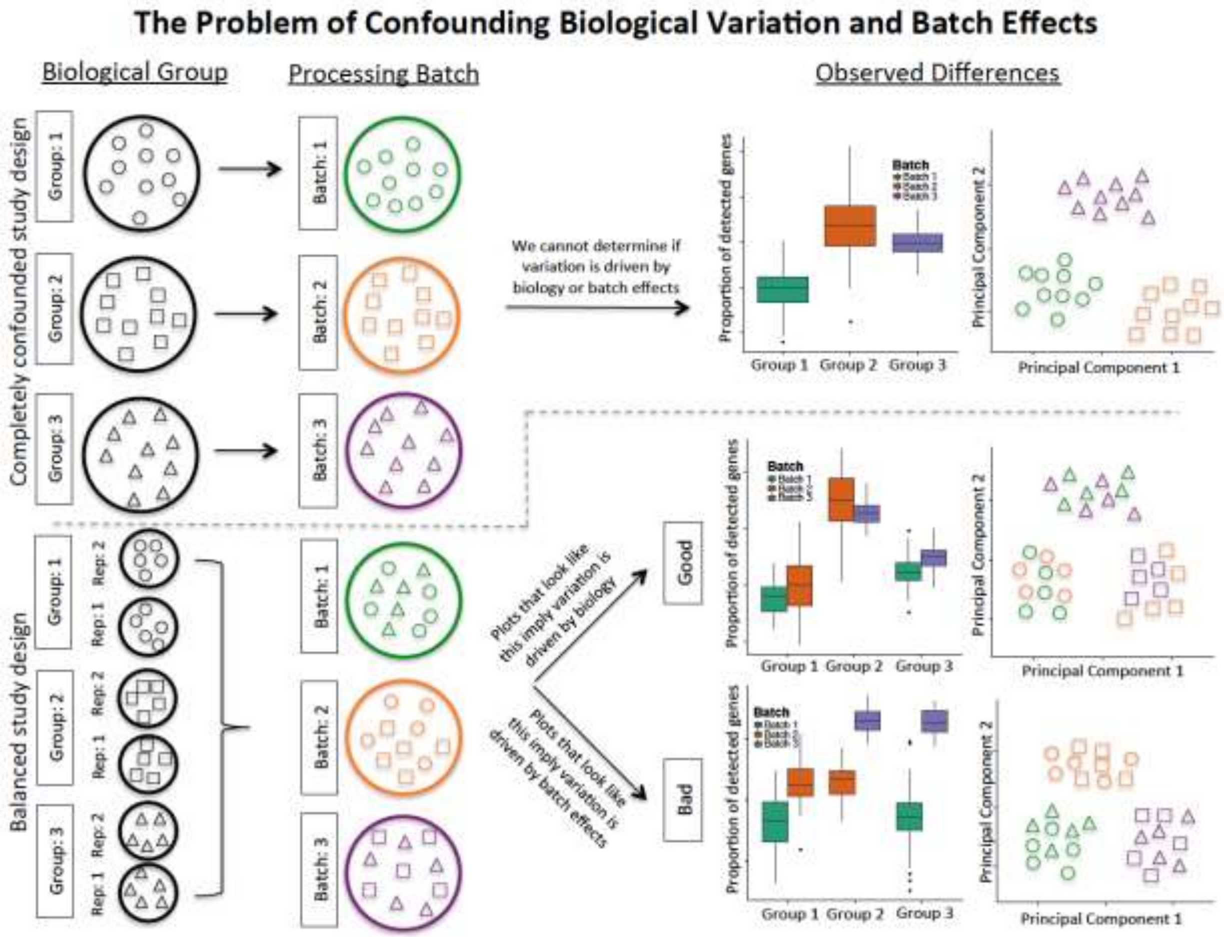


Cell	Site	Treatment
1	Main	A
2	Main	A
3	Main	A
4	Main	A
5	Remote	B
6	Remote	B
7	Remote	B
8	Remote	B



Cell	Site	Treatment
1	Main	A
2	Main	A
3	Main	B
4	Main	B
5	Remote	A
6	Remote	A
7	Remote	B
8	Remote	B

Confounding by Design

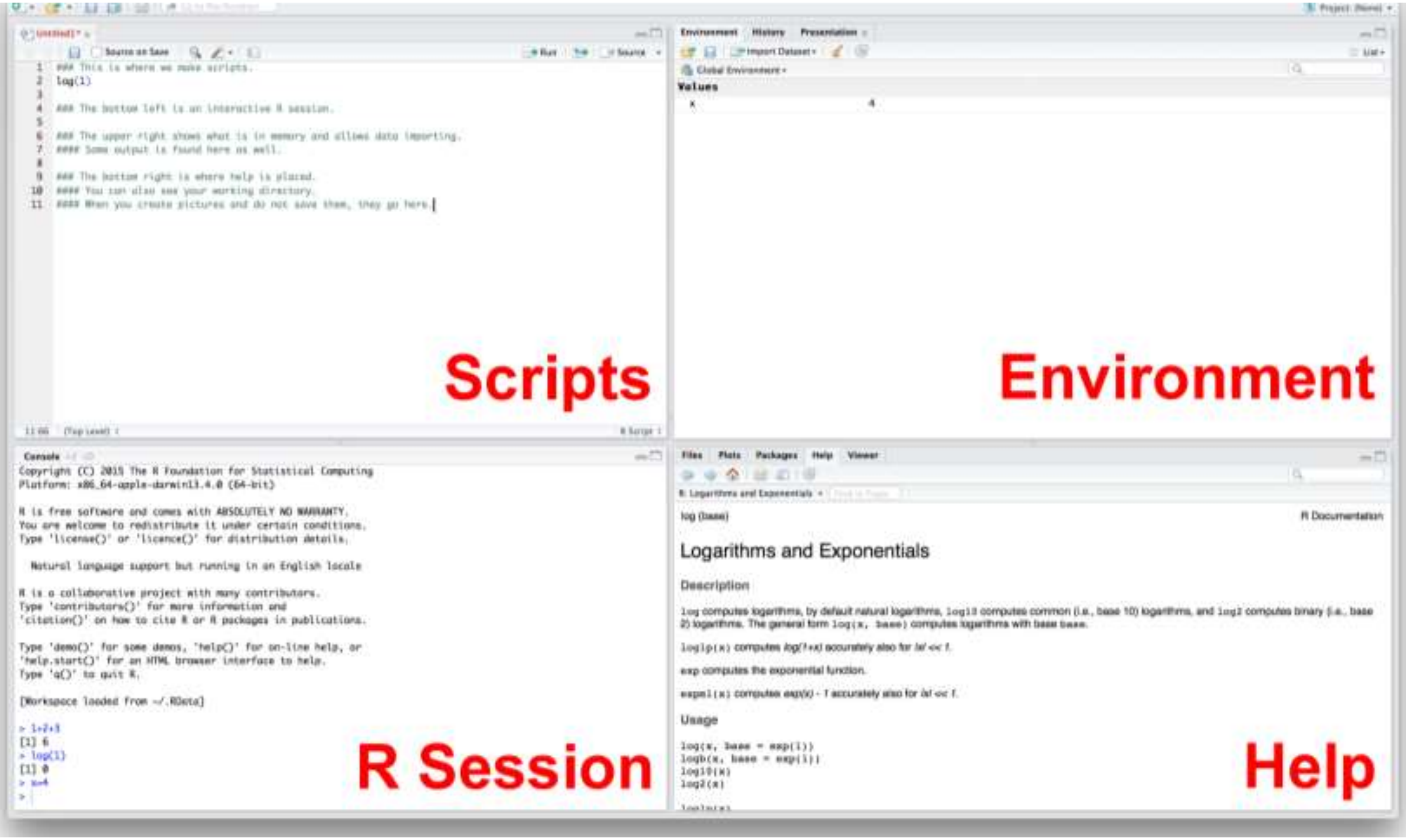

The Problem of Confounding Biological Variation and Batch Effects
Biological Group
Processing Batch
Observed Differences
Completely confounded study design
Group: 1
Group: 2
Group: 3
Batch: 1
Batch: 2
Batch: 3
We cannot determine if variation is driven by biology or batch effects
Proportion of detected genes
Batch
Batch 1
Batch 2
Batch 3
Group 1
Group 2
Group 3
Principal Component 2
Principal Component 1
Balanced study design
Rep: 1
Rep: 2
Plots that look like this imply variation is driven by biology
Plots that look like this imply variation is driven by batch effects
Good
Bad

Section: Initial Data Analysis

Motivation: Why Am I Using R?

- A lot of method development is happening in R.
- Free / open source / open science.
- Many supplemental computational biology packages.
- Data science is an art.
  - Data often requires one to create and manipulate analysis.
- This will allow you to experience key concepts in analysis.

RStudio (IDE)



Scripts

```
### This is where we make scripts.
log(1)

### The bottom left is an interactive R session.

### The upper right shows what is in memory and allows data importing.
#### Some output is found here as well.

### The bottom right is where help is placed.
#### You can also see your working directory.
#### When you create pictures and do not save them, they go here.
```

Environment

```
Values
  x                          4
```

R Session

```
Copyright (C) 2015 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

> 1+2+3
[1] 6
> log(1)
[1] 0
> x=4
>
```

Help

log (base) R Documentation

Logarithms and Exponentials

Description

`log` computes logarithms, by default natural logarithms, `log10` computes common (i.e., base 10) logarithms, and `log2` computes binary (i.e., base 2) logarithms. The general form `log(x, base)` computes logarithms with base `base`.

`log1p(x)` computes *log(1+x)* accurately also for *|x| << 1*.

`exp` computes the exponential function.

`expm1(x)` computes *exp(x) - 1* accurately also for *|x| << 1*.

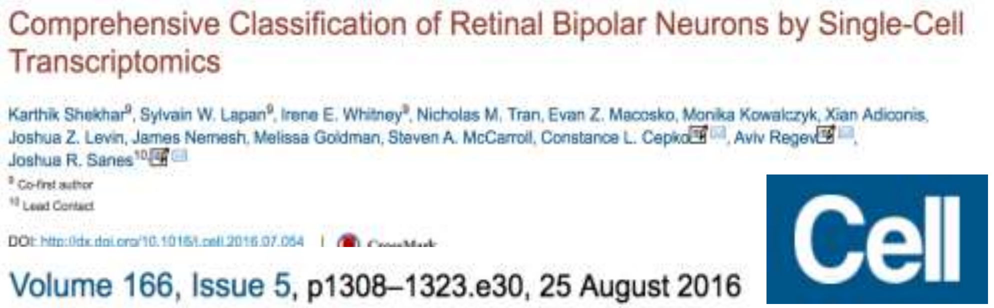
Usage

```
log(x, base = exp(1))
logb(x, base = exp(1))
log10(x)
log2(x)

log1p(x)
```

Initial Data Exploration

Today’s Data



Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics

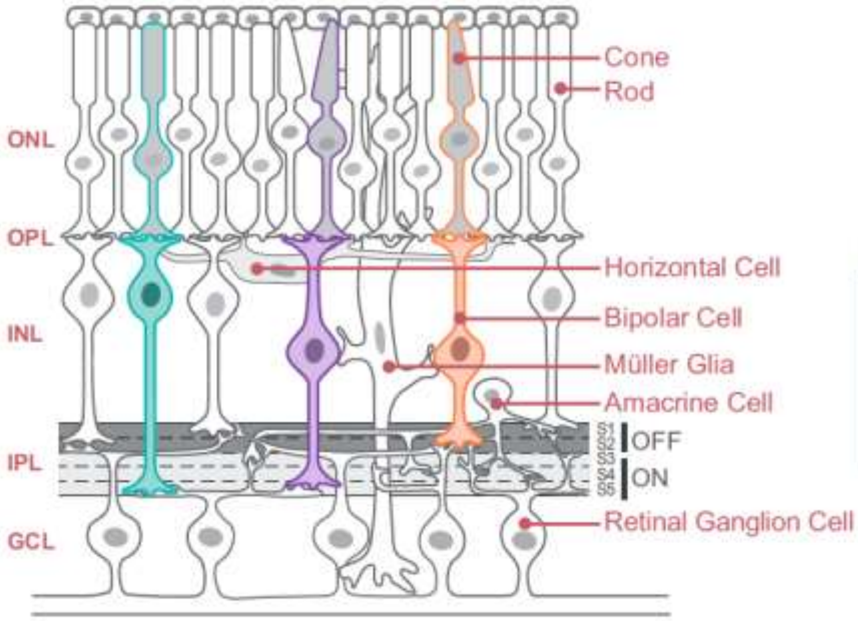
Karthik Shekhar[9], Sylvain W. Lapan[9], Irene E. Whitney[9], Nicholas M. Tran, Evan Z. Macosko, Monika Kowalczyk, Xian Adiconis, Joshua Z. Levin, James Nemesh, Melissa Goldman, Steven A. McCarroll, Constance L. Cepko, Aviv Regev, Joshua R. Sanes[10]

[9] Co-first author
[10] Lead Contact

DOI: http://dx.doi.org/10.1016/j.cell.2016.07.054 | CrossMark

Volume 166, Issue 5, p1308–1323.e30, 25 August 2016

Cell



- To generate a comprehensive, validated classification scheme for the bipolar cells of the mouse retina.
  - Cone or rod type, ON or OFF, 9-12 subtypes (morphological)
- ~44k cells from a transgenic mouse line marking BCs
  - After filtering 27k (we use 5k)

Logistics: Importing Code Libraries

- R Exercise

Representing Sparse Matrices

- R Exercise

What is a Sparse Matrix?

- Sparse Matrix
  - A matrix where most of the elements are 0.
- Dense Matrix
  - A matrix where most elements are not 0.
- Many ways to efficiently represent a sparse matrix in memory.
  - Here, the underlying data structure is a coordinate list.

2D Arrays vs Coordinate Lists

Can be optimal for dense matrices

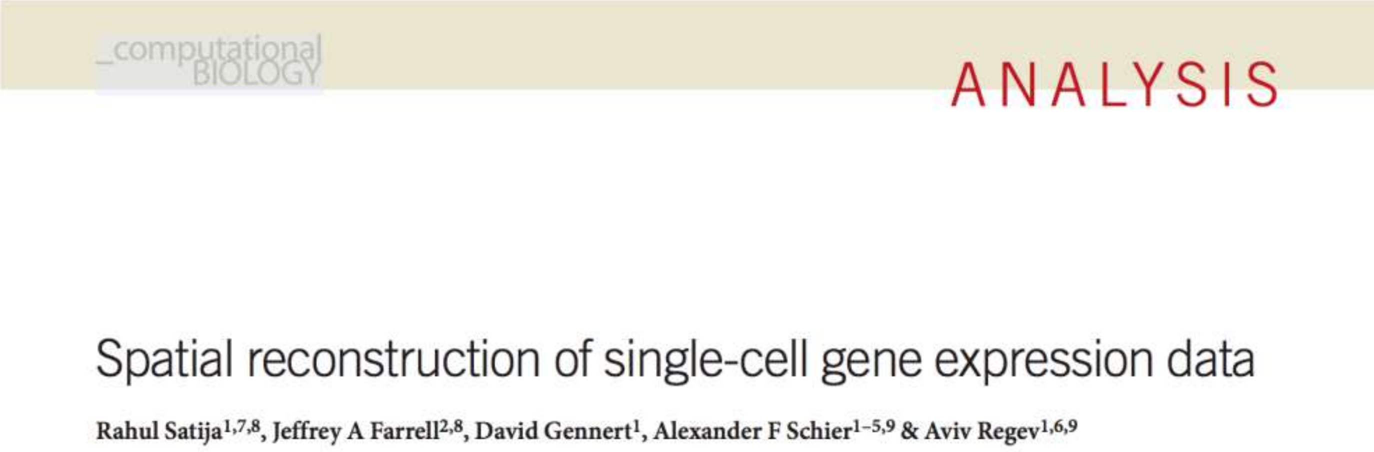
More optimal for sparse matrices

2D Arrays **vs** Coordinate List

1	2	3	4	5	6	7	8	
0	0	0	0	0	0	0	0	1
0	1	0	0	0	0	0	0	2
0	0	0	0	0	2	0	0	3
0	0	0	0	0	0	0	0	4
0	0	0	0	0	0	0	0	5
0	0	0	0	0	0	0	3	6

2	2	1
6	3	2
8	6	3

Seurat

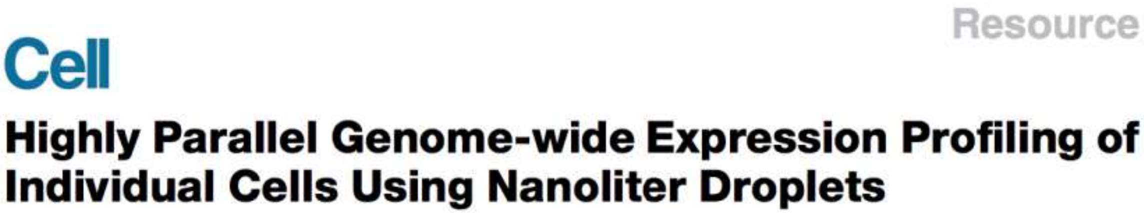


computational BIOLOGY

ANALYSIS

Spatial reconstruction of single-cell gene expression data

Rahul Satija[1,7,8], Jeffrey A Farrell[2,8], David Gennert[1], Alexander F Schier[1–5,9] & Aviv Regev[1,6,9]



Resource

Cell

Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets

https://github.com/satijalab/seurat

Create a Seurat Object

- R exercise

Expression: Bulk RNA-Seq Definition

In bulk RNA-Seq we learned counts are not expression.

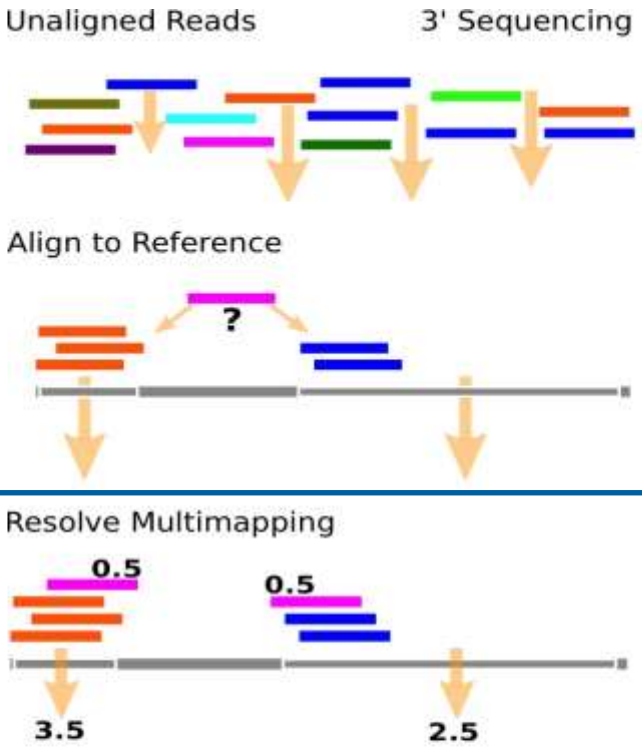
- Some counts belong to sequences which could go to many genes.
- Some transcripts are longer than other so they get sequenced more.
- Some samples are more deeply sequenced.
- The data is not normally distributed.

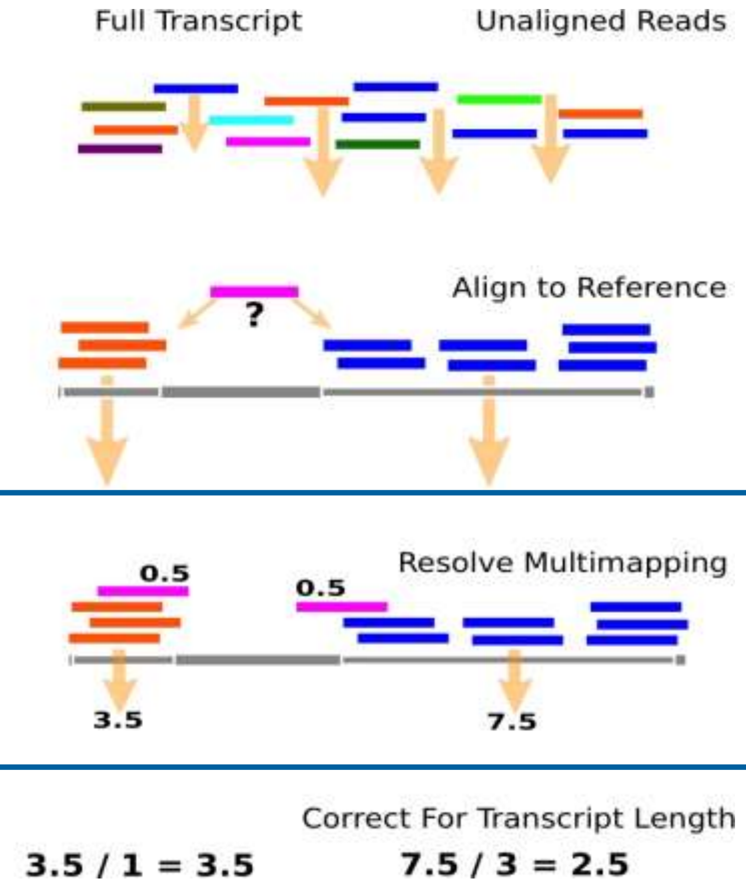
Depending on the scRNA-Seq assay these may be important.

Seurat has assumptions it makes with it’s defaults

- More appropriate for 3 prime assays.

Count Preparation is Different Depending on the Source





RSEM
KALLISTO

RSEM
KALLISTO

No transcript length correction

TPM

Seurat

Correct for Sequencing Depth
X / Column Total * 1E5 or 1E6

Correct for Sequencing Depth

TPM

Seurat

Log2() + 1

Log2() + 1

Seurat

Prepping Counts For Seurat

3 prime-

- Expected by Seurat.
- Counts collapsed with UMIs.
- Log2 transform (in Seurat).
- Account for sequencing depth (in Seurat).

Full Transcript Sequencing-

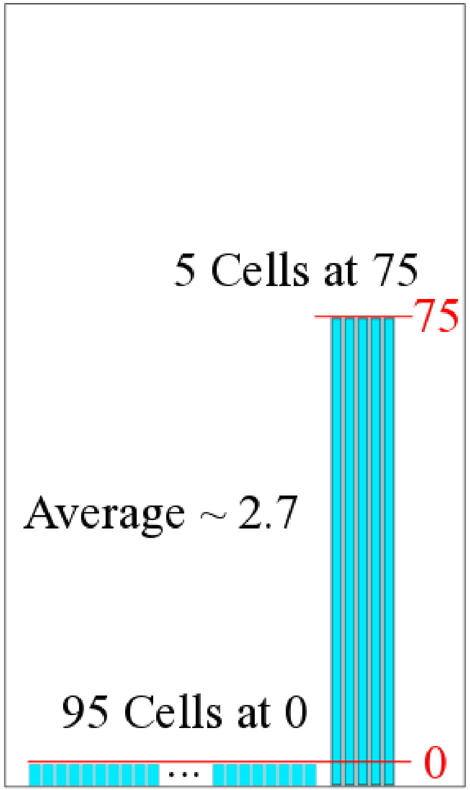
- Can be used in Seurat.
- TPM +1 transformed counts.
- Log2 transform (in Seurat).
- Sequencing depth is already accounted.

Sometimes Averages are Not Useful

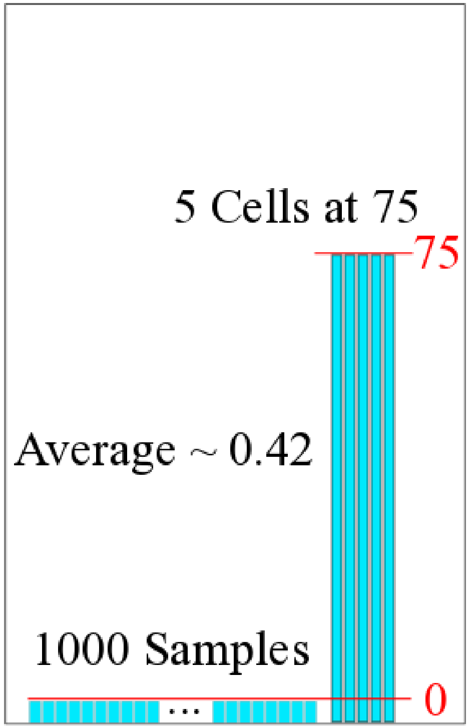
Say you were standing with one foot in the oven and one foot in an ice bucket. According to the percentage people, you should be perfectly comfortable. –Bobby Bragan

Filtering Genes: Averages are Less Useful

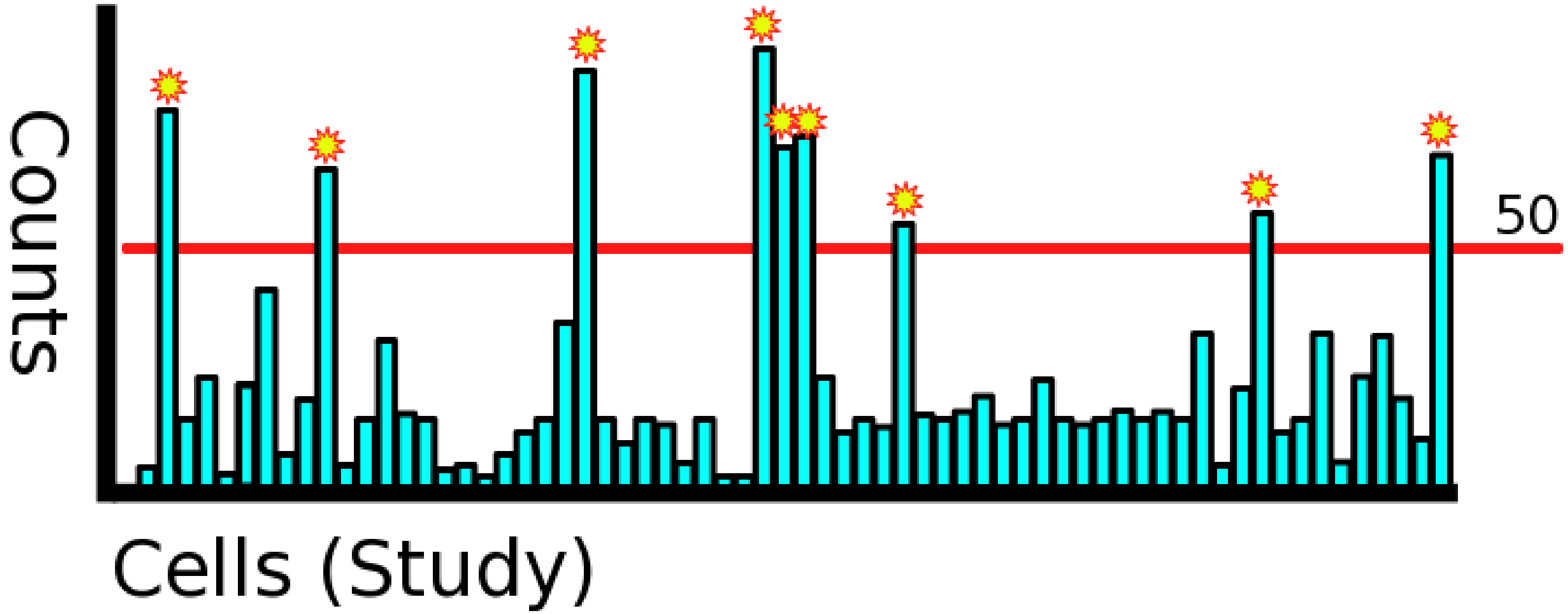
0s pull down average



Amount of 0s is arbitrary (study size, diversity)

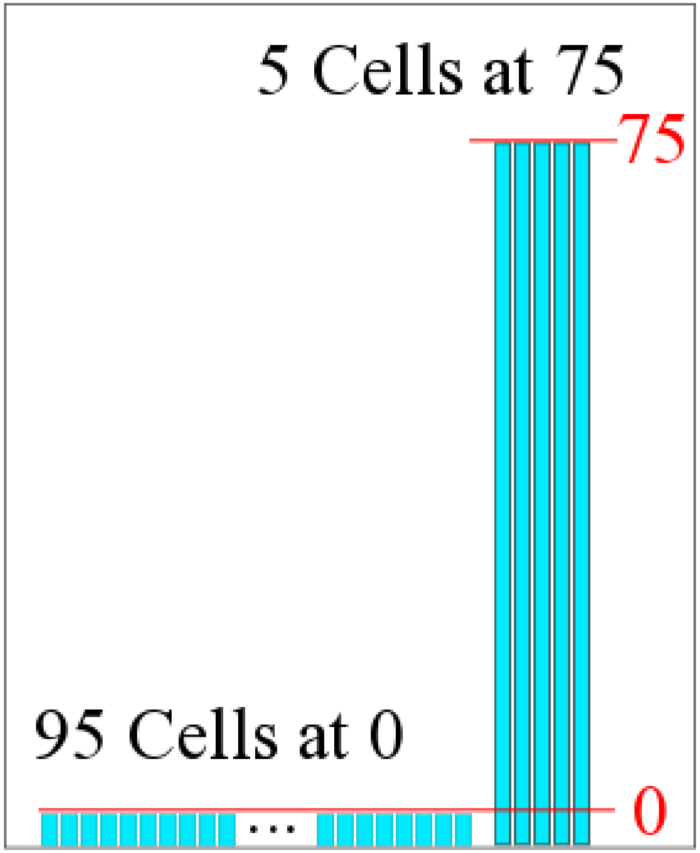


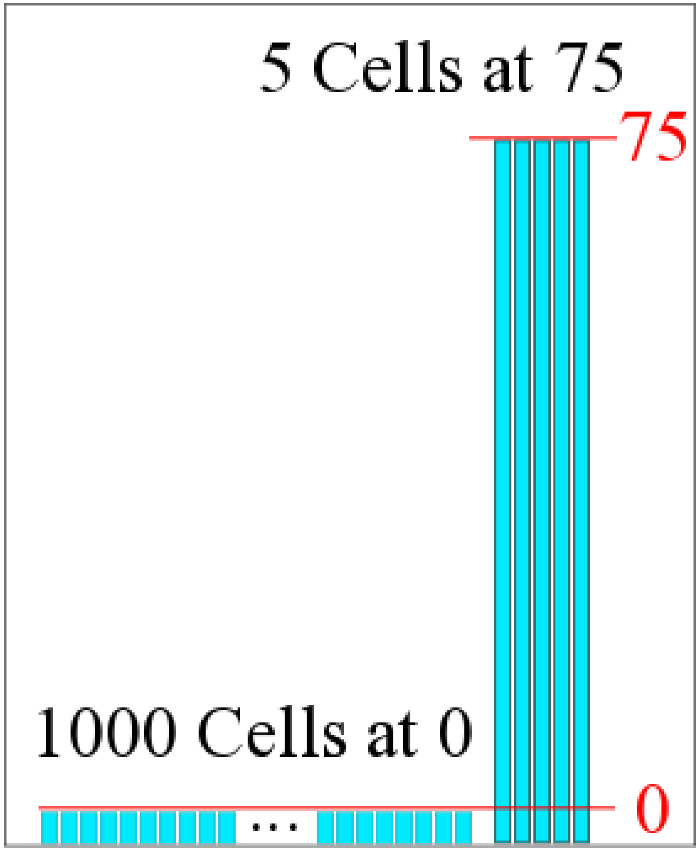
Filtering Genes: Using Prevalence



Filtering Genes: Using Prevalence

Filter: 5 cells must have 10 expression





Filtering Using Metadata

What is Metadata?

Other information that describes your measurements.

- Patient information.
  - Life style (smoking), Patient Biology (age), Comorbidity
- Study information.
  - Treatment, Cage, Sequencing Site, Sequencing Date
- Sequence QC on cells.
  - Useful in filtering.

Filtering Cells: Removing Outlier Cells

- Bulk RNA-Seq studies often do not remove outliers cells
  - scRNA-Seq often removes “failed libraries”.
- Outlier cells are not just measured by complexity
    - Percent Reads Mapping
    - Percent Mitochondrial Reads
    - Presence of marker genes
    - Intergenic/ exonic rate
    - 5' or 3' bias
    - other metadata …
- Useful Tools
  - Picard Tools and RNASeQC

Seurat: Filtering on Metadata

- R Exercise

Section: Plot Genes

Seurat: Viewing Specific Genes

- R Exercise

Section: Working with Batch Affects

Normalization and Batch Affect Correction

- The nature of scRNA-Seq assays can make them prone to confounding with batch affects.
  - Normalization and batch affect correction can help.
- Some are moving away from relying on a specific method.
  - Exploring the idea of combining or selecting from a collection of normalization or correction methods best for a specific study.
- Some believe UMI based analysis need not be normalized between samples given the absolute count of the molecules are being reported.
  - Be careful not to remove biological signal with good experimental design (avoiding confounding by design).

Seurat and Batch Affect Correction

- Using linear models one can regress covariates.
  - scale.data hold the residuals after regressing (z-scored)
- Dimensionality reduction and clustering.
- We use metadata we have.
  - One could imagine creating a metadata for cell cycle.

Seurat and Batch Affect Correction

- R exercise

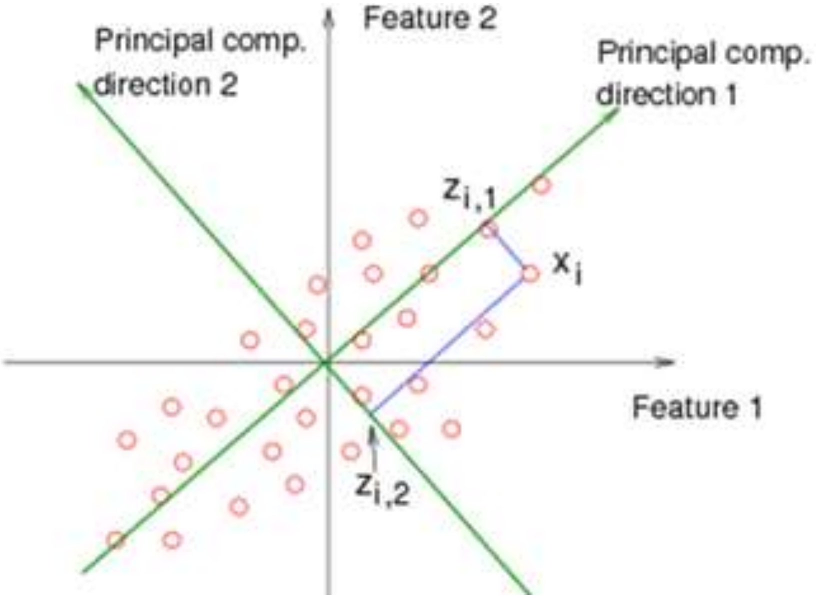
Section: Dimensionality Reduction and Plotting Samples

Dimensionality Reduction

- Start with many measurements (high dimensional).
  - Want to reduce to few features (lower-dimensional space).
- One way is to extract features based on capturing groups of variance.
- Another could be to preferentially select some of the current features.
  - We have already done this.
- We need this to plot the cells in 2D (or ordinate them)
- In scRNA-Seq PC1 may be complexity.

PCA: in Quick Theory

- Eigenvectors of covariance matrix.
- Find orthogonal groups of variance.
- Given from most to least variance.
  - Components of variation.
  - Linear combinations explaining the variance.



PCA: an Interactive Example

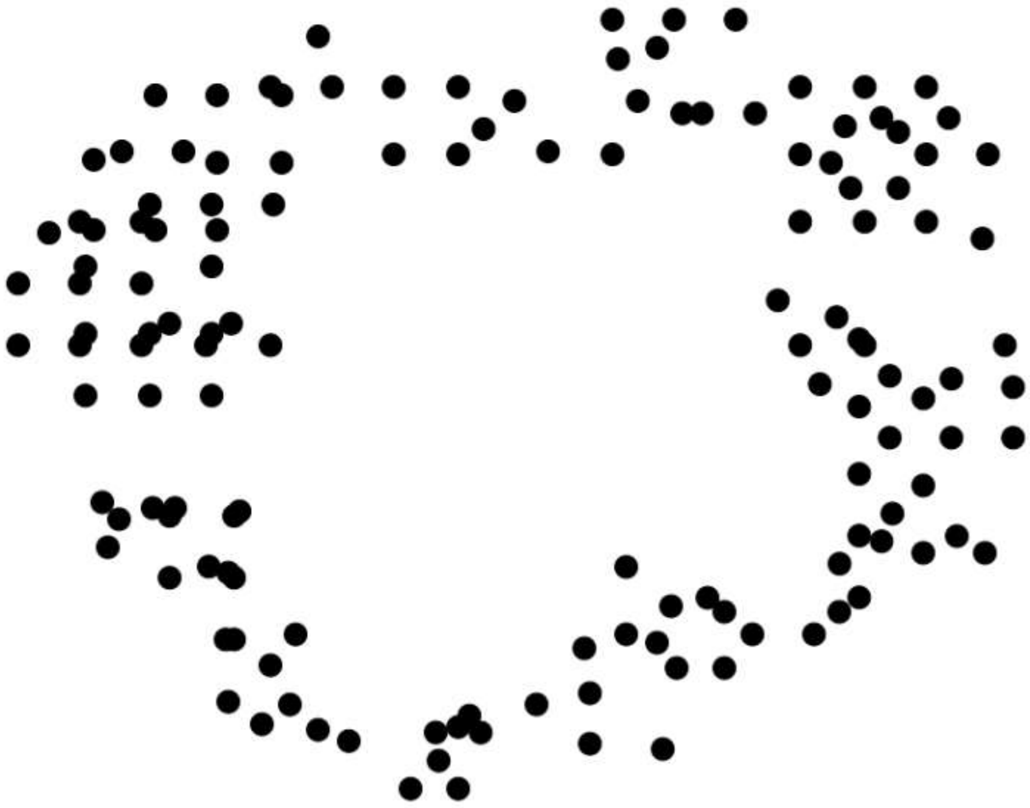
- PCA Explained Visually

PCA: in Practice

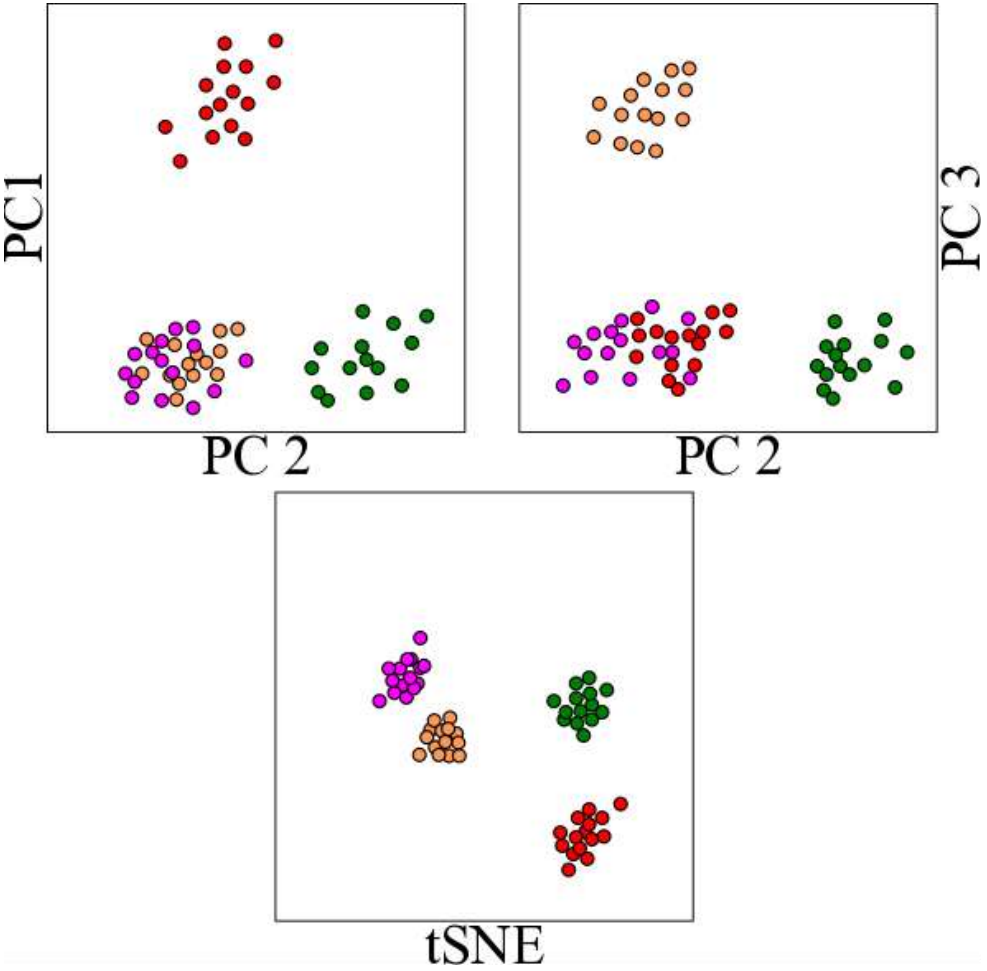
Things to be aware of-

- Data with different magnitudes will dominate.
  - Zero center and divided by SD.
    - (Standardized).
- Can be affected by outliers.
- Data is often first filtered to remove noise.

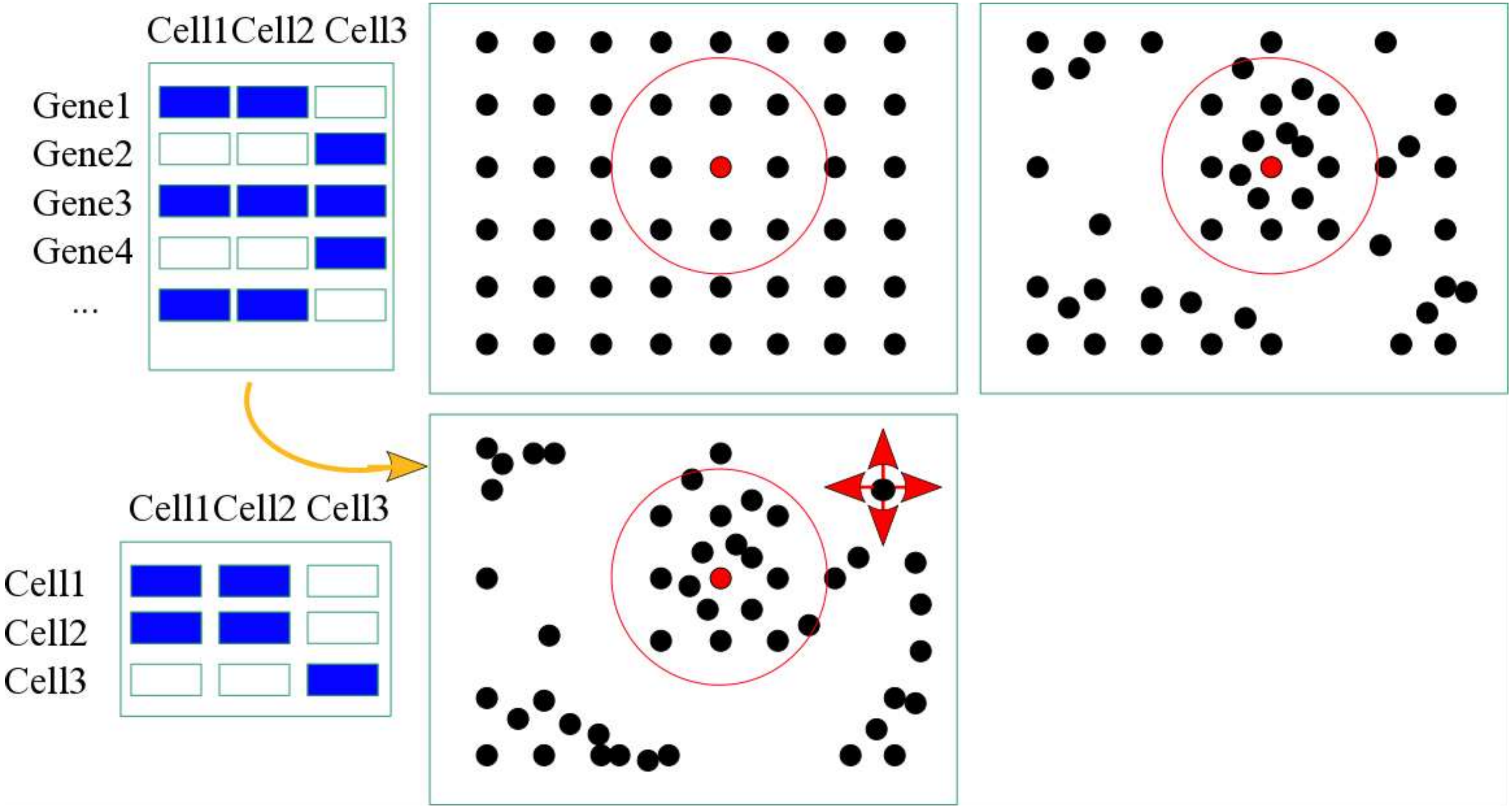
t-SNE: Nonlinear Dimensional Reduction



t-SNE: Collapsing the Visualization to 2D



t-SNE: How it works.



PCA and t-SNE Together

- Often t-SNE is performed on PCA components
  - Liberal number of components.
  - Removes mild signal (assumption of noise).
  - Faster, on less data but, hopefully the same signal.

Learn More About t-SNE

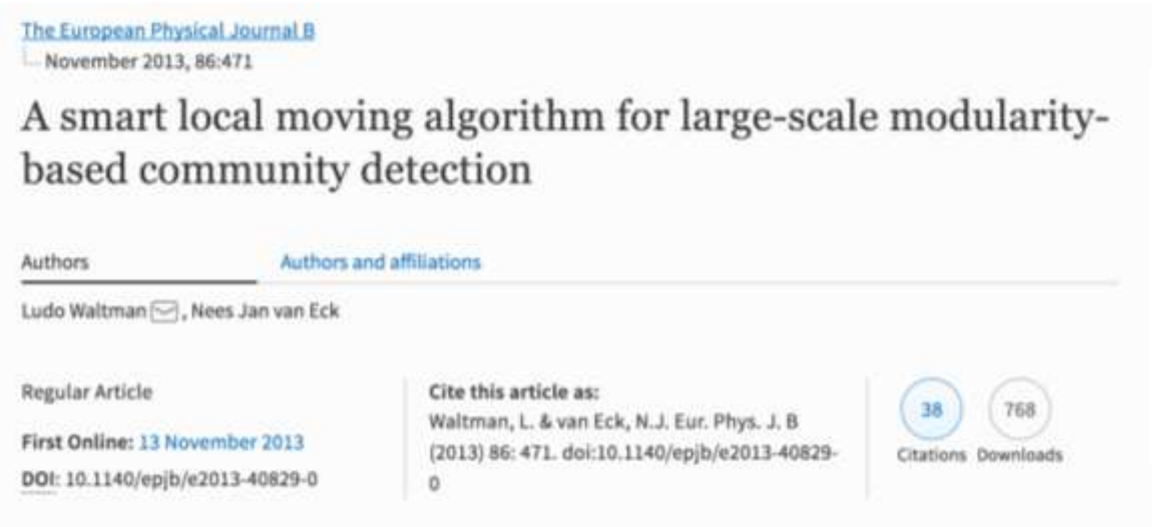
- Awesome Blog on t-SNE parameterization
  - http://distill.pub/2016/misread-tsne
- Publication
  - https://lvdmaaten.github.io/publications/papers/JMLR_2008.pdf
- Nice YouTube Video
  - https://www.youtube.com/watch?v=RJVL80Gg3lA
- Code
  - https://lvdmaaten.github.io/tsne/
- Interactive Tensor flow
  - http://projector.tensorflow.org/

Plotting Cells

Plotting Cells and Gene Expression

- R exercise.

Defining Clusters through Graphs



The European Physical Journal B
November 2013, 86:471

A smart local moving algorithm for large-scale modularity-based community detection

Authors   Authors and affiliations

Ludo Waltman, Nees Jan van Eck

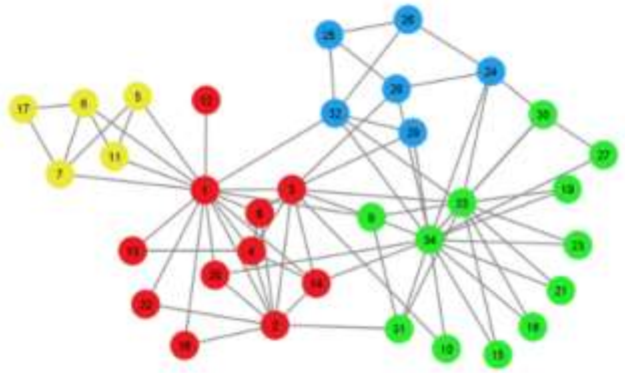
Regular Article

First Online: 13 November 2013

DOI: 10.1140/epjb/e2013-40829-0

Cite this article as:
Waltman, L. & van Eck, N.J. Eur. Phys. J. B (2013) 86: 471. doi:10.1140/epjb/e2013-40829-0

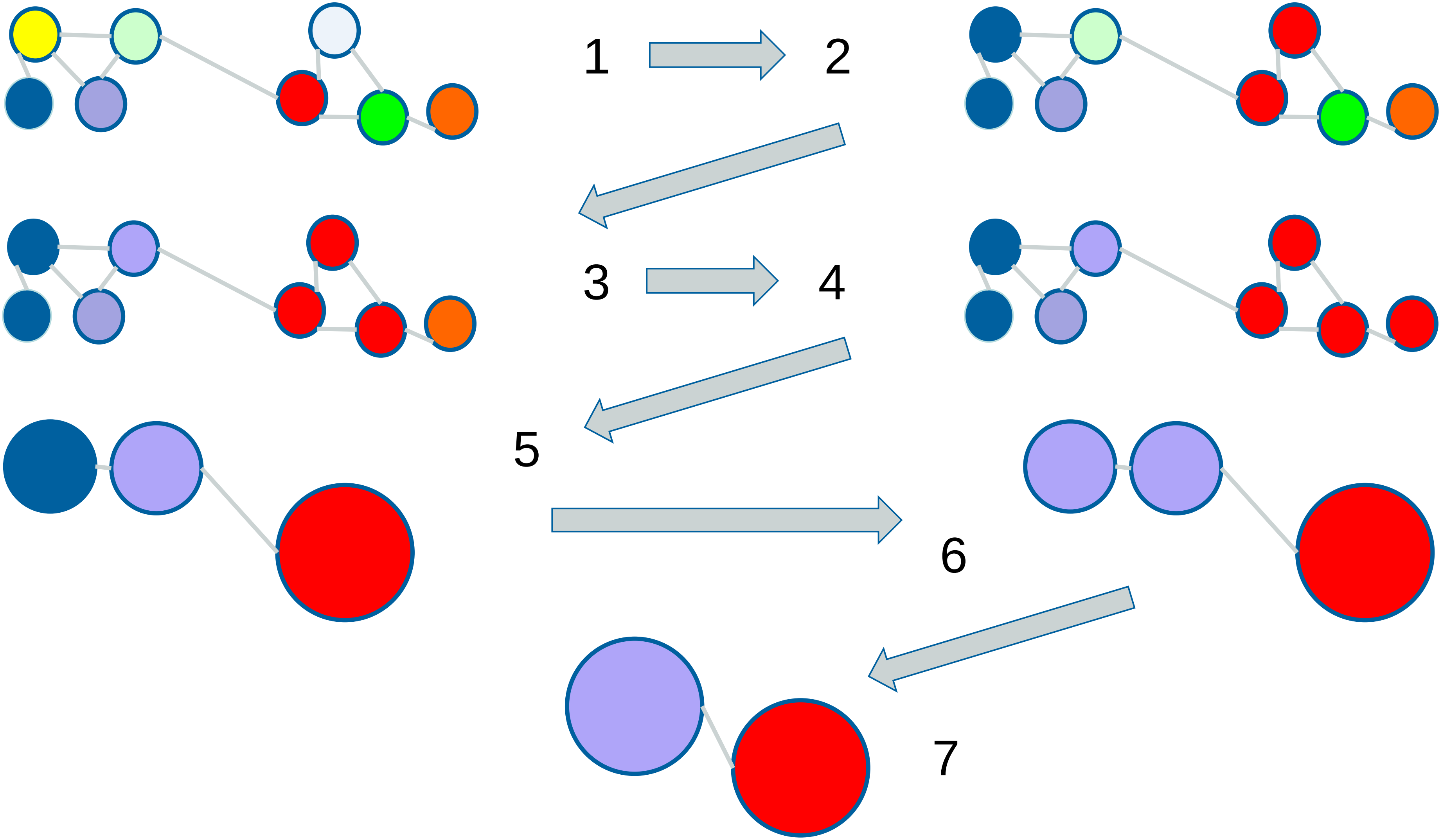
38 Citations   768 Downloads



- Smart Local Moving (SLM) algorithm for community (cluster) detection in large networks.
  - Can be applied to 10s of millions cells, 100s of millions of relationships.
  - Evolved from the Louvain algorithm

http://www.ludowaltman.nl/slm/

Local Moving Heuristic



Section Summary

- Dimensionality reduction help reduce data while hopefully keeping important signal.
  - t-SNE on PCA is often used in analysis
- Created several types of plot often seen in publications.
  - Plotting genes (through subgroups).
  - Ordinating cells in t-SNE space.
  - Heat maps of genes associated with PC components.
  - Plotting metadata on projects of data is an important QC tool.
- Cluster of cells are currently defined through graph, separate from the ordination (t-SNE / PCA).

Section: Differential Expression

Seurat: Differential Expression

- Default if one cluster again many tests.
  - Can specify an ident.2 test between clusters.
- Adding speed by exluding tests.
  - Min.pct - controls for sparsity
  - Min percentage in a group
  - Thresh.test - must have this difference in averages.

Seurat: Many Choices for DE

- bimod
  - Tests differences in mean and proportions.
- roc
  - Uses AUC like definition of separation.
- t
  - Student's T-test.
- tobit
  - Tobit regression on a smoothed data.

Seurat: DE and Plotting DE Genes

- R Exercise.

Dot plots

Size of circle

- Gene prevalence in cluster.

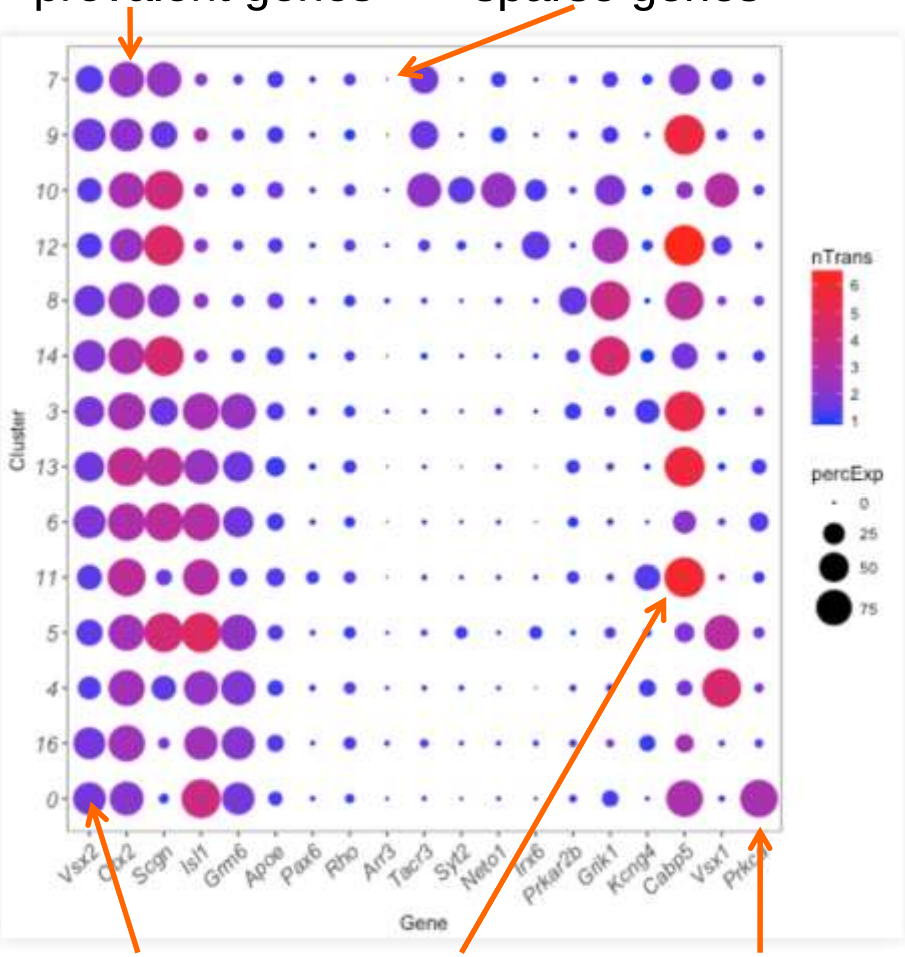
Color of circle

- More red, more expressed in cluster.

Scales well with many cells.

prevalent genes

sparse genes



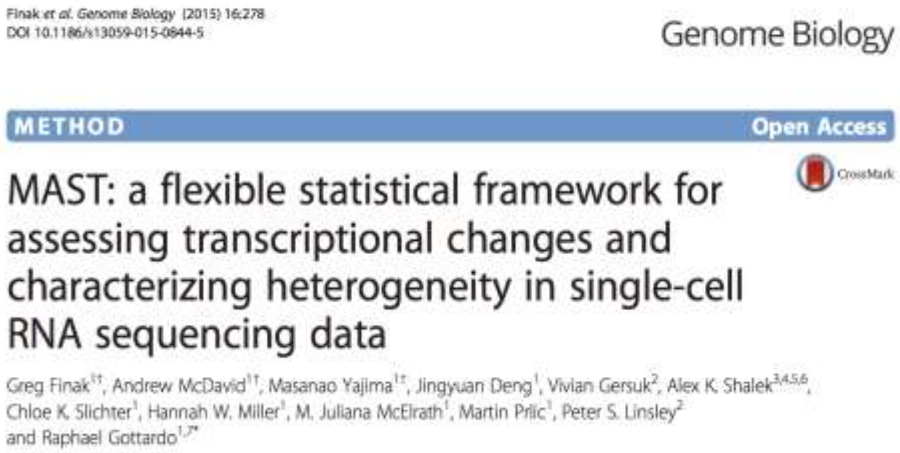
lowly expressed

highly expressed

very specific

Mast

- Uses hurdle model
  - Two part generalized linear model to address both rate of expression (prevalence) and expression.
  - GLM means covariates can be used to control for unwanted signal.
- CDR: Cellular detection rate
  - Cellular complexity
  - Values below a threshold are 0



Finak *et al. Genome Biology* (2015) 16:278
DOI 10.1186/s13059-015-0844-5

Genome Biology

METHOD Open Access

MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data

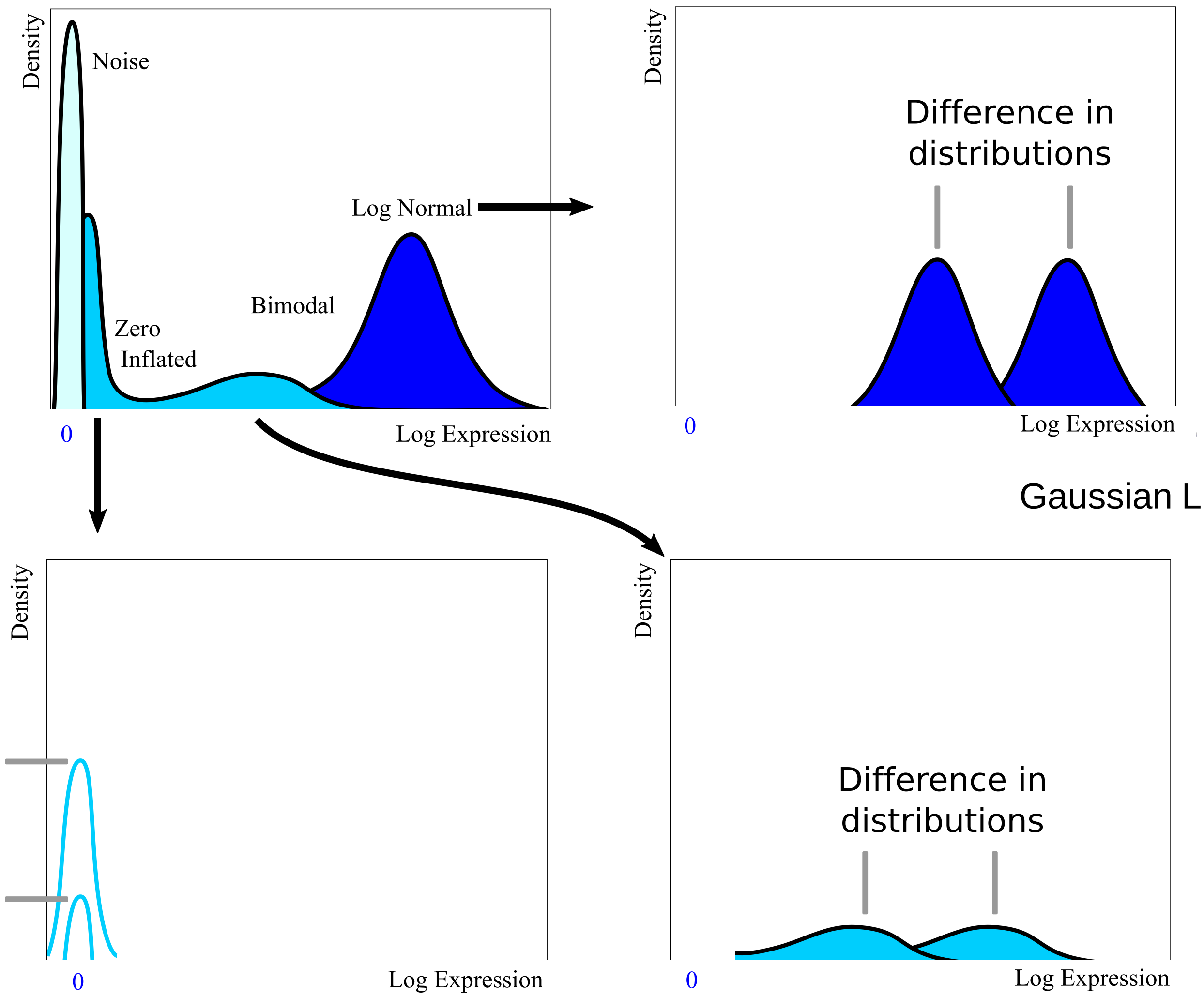
Greg Finak[1†], Andrew McDavid[1†], Masanao Yajima[1†], Jingyuan Deng[1], Vivian Gersuk[2], Alex K. Shalek[3,4,5,6], Chloe K. Slichter[1], Hannah W. Miller[1], M. Juliana McElrath[1], Martin Prlic[1], Peter S. Linsley[2] and Raphael Gottardo[1,7*]

- Additionally introduces a GSEA method.

https://github.com/RGLab/MAST

Mast: Hurdle Models

Distribution of Expression of a Gene throughout a Study



Logistic Regression

Gaussian Linear Model

Difference in distributions

Difference in number of zeroes

Difference in distributions

Mast: DE and Plotting DE Genes

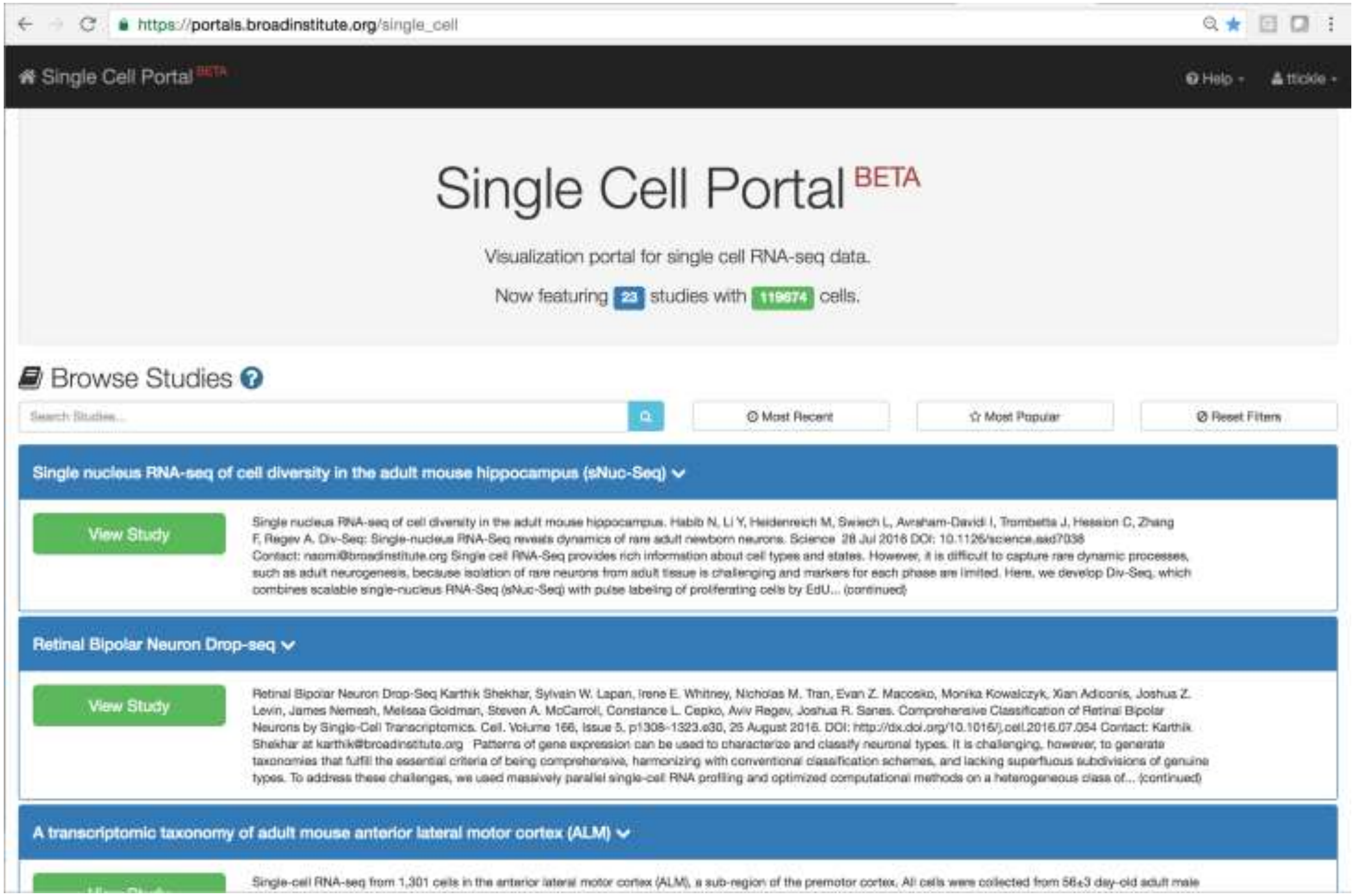
- R Exercise.

Section: Communicating Results to Collaborators

- Designing a study.
- Writing a grant.
- Performing experiments.
- Collecting data.
- Running sequencing pipelines.
- Performing some preliminary analysis.
- **Sharing ideas with private collaborators.**
- Refining analysis.
- Completing a paper.
- **Sharing analysis publicly.**

The Single Cell Portal

https://portals.broadinstitute.org/single_cell



Single Cell Portal BETA

Visualization portal for single cell RNA-seq data.

Now featuring 23 studies with 119674 cells.

Browse Studies

Single nucleus RNA-seq of cell diversity in the adult mouse hippocampus (sNuc-Seq)

Single nucleus RNA-seq of cell diversity in the adult mouse hippocampus. Habib N, Li Y, Heidenreich M, Swiech L, Avraham-Davidi I, Trombetta J, Hession C, Zhang F, Regev A. Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. Science 28 Jul 2016 DOI: 10.1126/science.aad7038 Contact: naomi@broadinstitute.org Single cell RNA-Seq provides rich information about cell types and states. However, it is difficult to capture rare dynamic processes, such as adult neurogenesis, because isolation of rare neurons from adult tissue is challenging and markers for each phase are limited. Here, we develop Div-Seq, which combines scalable single-nucleus RNA-Seq (sNuc-Seq) with pulse labeling of proliferating cells by EdU... (continued)

Retinal Bipolar Neuron Drop-seq

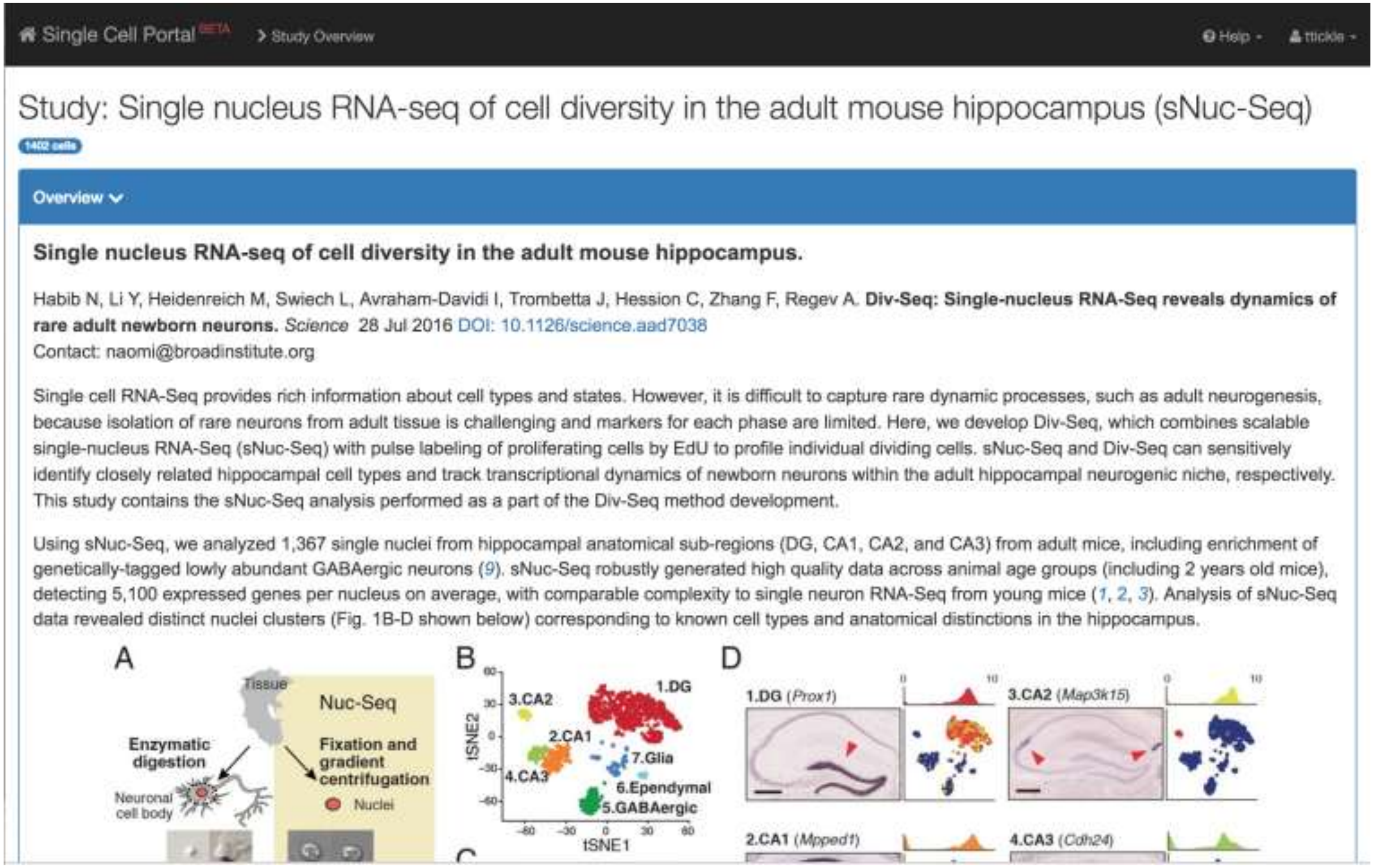
Retinal Bipolar Neuron Drop-Seq Karthik Shekhar, Sylvain W. Lapan, Irene E. Whitney, Nicholas M. Tran, Evan Z. Macosko, Monika Kowalczyk, Xian Adiconis, Joshua Z. Levin, James Nemesh, Melissa Goldman, Steven A. McCarroll, Constance L. Cepko, Aviv Regev, Joshua R. Sanes. Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. Cell. Volume 166, Issue 5, p1308-1323.e30, 25 August 2016. DOI: http://dx.doi.org/10.1016/j.cell.2016.07.054 Contact: Karthik Shekhar at karthik@broadinstitute.org Patterns of gene expression can be used to characterize and classify neuronal types. It is challenging, however, to generate taxonomies that fulfill the essential criteria of being comprehensive, harmonizing with conventional classification schemes, and lacking superfluous subdivisions of genuine types. To address these challenges, we used massively parallel single-cell RNA profiling and optimized computational methods on a heterogeneous class of... (continued)

A transcriptomic taxonomy of adult mouse anterior lateral motor cortex (ALM)

Single-cell RNA-seq from 1,301 cells in the anterior lateral motor cortex (ALM), a sub-region of the premotor cortex. All cells were collected from 56±3 day-old adult male

The Single Cell Portal

Study Descriptions Can Be Created



Study: Single nucleus RNA-seq of cell diversity in the adult mouse hippocampus (sNuc-Seq)

1402 cells

Overview

Single nucleus RNA-seq of cell diversity in the adult mouse hippocampus.

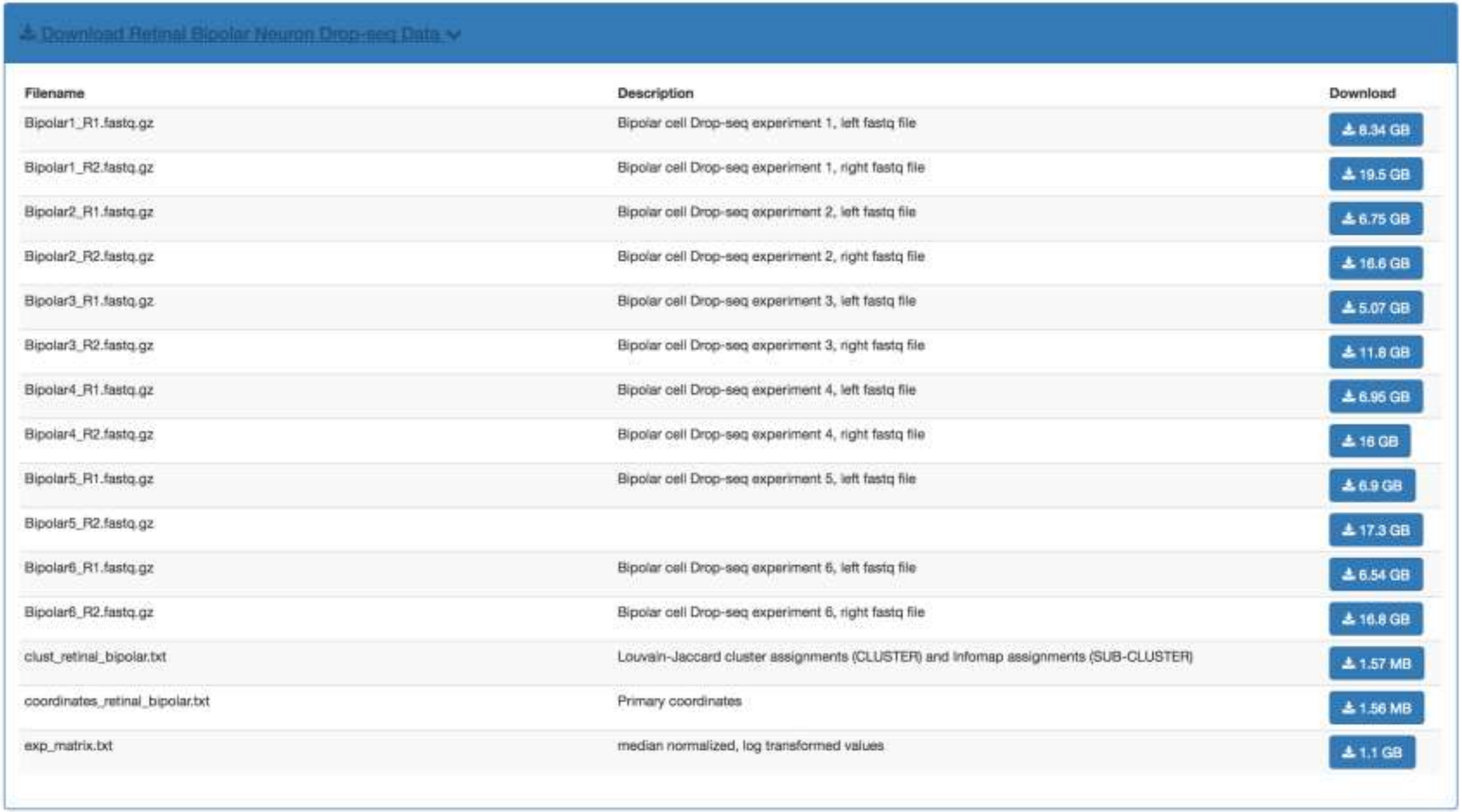
Habib N, Li Y, Heidenreich M, Swiech L, Avraham-Davidi I, Trombetta J, Hession C, Zhang F, Regev A. **Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons.** *Science* 28 Jul 2016 DOI: 10.1126/science.aad7038
Contact: naomi@broadinstitute.org

Single cell RNA-Seq provides rich information about cell types and states. However, it is difficult to capture rare dynamic processes, such as adult neurogenesis, because isolation of rare neurons from adult tissue is challenging and markers for each phase are limited. Here, we develop Div-Seq, which combines scalable single-nucleus RNA-Seq (sNuc-Seq) with pulse labeling of proliferating cells by EdU to profile individual dividing cells. sNuc-Seq and Div-Seq can sensitively identify closely related hippocampal cell types and track transcriptional dynamics of newborn neurons within the adult hippocampal neurogenic niche, respectively. This study contains the sNuc-Seq analysis performed as a part of the Div-Seq method development.

Using sNuc-Seq, we analyzed 1,367 single nuclei from hippocampal anatomical sub-regions (DG, CA1, CA2, and CA3) from adult mice, including enrichment of genetically-tagged lowly abundant GABAergic neurons (*9*). sNuc-Seq robustly generated high quality data across animal age groups (including 2 years old mice), detecting 5,100 expressed genes per nucleus on average, with comparable complexity to single neuron RNA-Seq from young mice (*1*, *2*, *3*). Analysis of sNuc-Seq data revealed distinct nuclei clusters (Fig. 1B-D shown below) corresponding to known cell types and anatomical distinctions in the hippocampus.

The Single Cell Portal

Data Can Be Shared

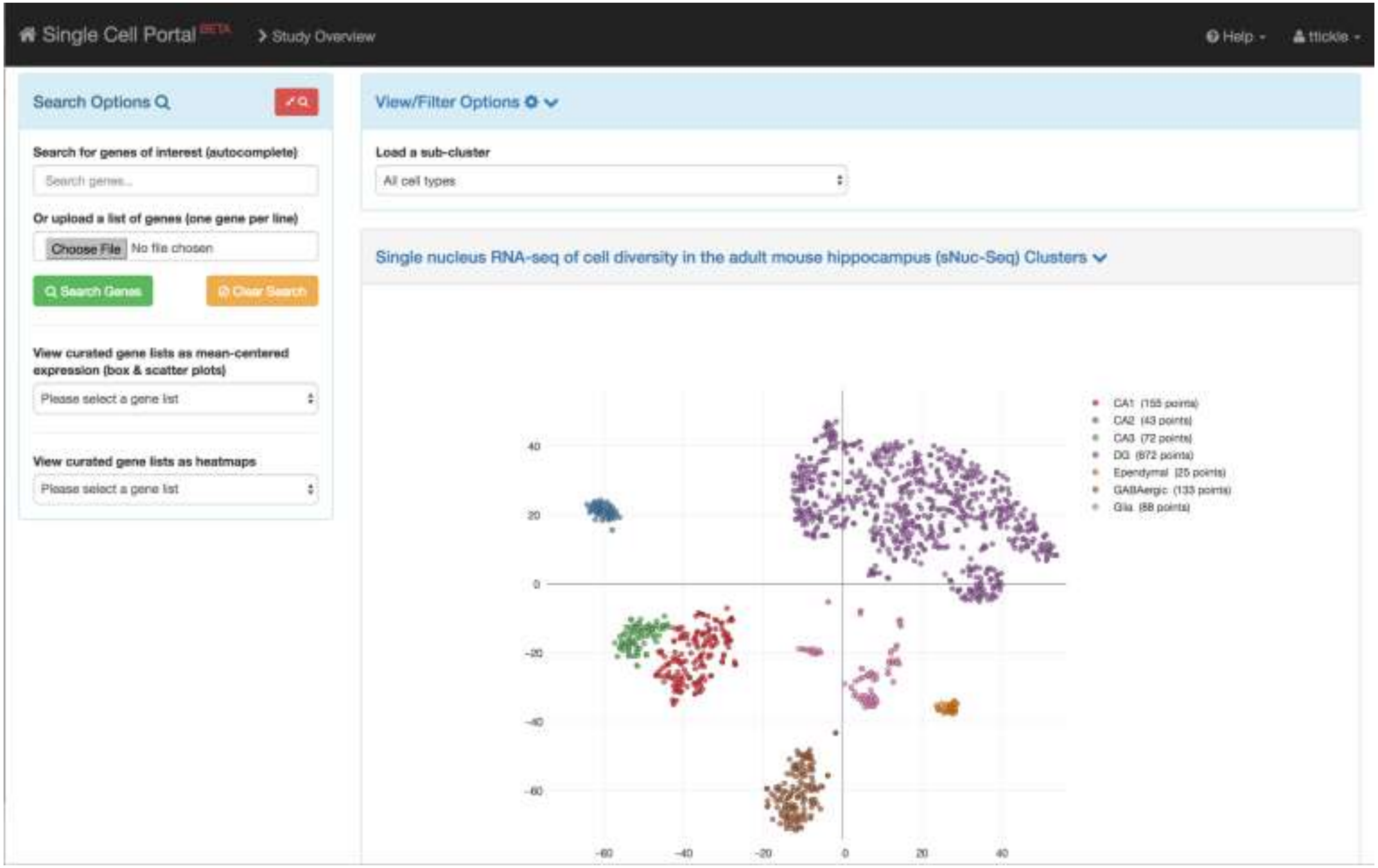


Download Retinal Bipolar Neuron Drop-seq Data

Filename	Description	Download
Bipolar1_R1.fastq.gz	Bipolar cell Drop-seq experiment 1, left fastq file	8.34 GB
Bipolar1_R2.fastq.gz	Bipolar cell Drop-seq experiment 1, right fastq file	19.5 GB
Bipolar2_R1.fastq.gz	Bipolar cell Drop-seq experiment 2, left fastq file	6.75 GB
Bipolar2_R2.fastq.gz	Bipolar cell Drop-seq experiment 2, right fastq file	16.6 GB
Bipolar3_R1.fastq.gz	Bipolar cell Drop-seq experiment 3, left fastq file	5.07 GB
Bipolar3_R2.fastq.gz	Bipolar cell Drop-seq experiment 3, right fastq file	11.8 GB
Bipolar4_R1.fastq.gz	Bipolar cell Drop-seq experiment 4, left fastq file	6.95 GB
Bipolar4_R2.fastq.gz	Bipolar cell Drop-seq experiment 4, right fastq file	16 GB
Bipolar5_R1.fastq.gz	Bipolar cell Drop-seq experiment 5, left fastq file	6.9 GB
Bipolar5_R2.fastq.gz		17.3 GB
Bipolar6_R1.fastq.gz	Bipolar cell Drop-seq experiment 6, left fastq file	6.54 GB
Bipolar6_R2.fastq.gz	Bipolar cell Drop-seq experiment 6, right fastq file	16.8 GB
clust_retinal_bipolar.txt	Louvain-Jaccard cluster assignments (CLUSTER) and Infomap assignments (SUB-CLUSTER)	1.57 MB
coordinates_retinal_bipolar.txt	Primary coordinates	1.56 MB
exp_matrix.txt	median normalized, log transformed values	1.1 GB

The Single Cell Portal

One Can Interact with Cell Clusters



Search for genes of interest (autocomplete)

Search genes...

Or upload a list of genes (one gene per line)

No file chosen

View curated gene lists as mean-centered expression (box & scatter plots)

Please select a gene list

View curated gene lists as heatmaps

Please select a gene list

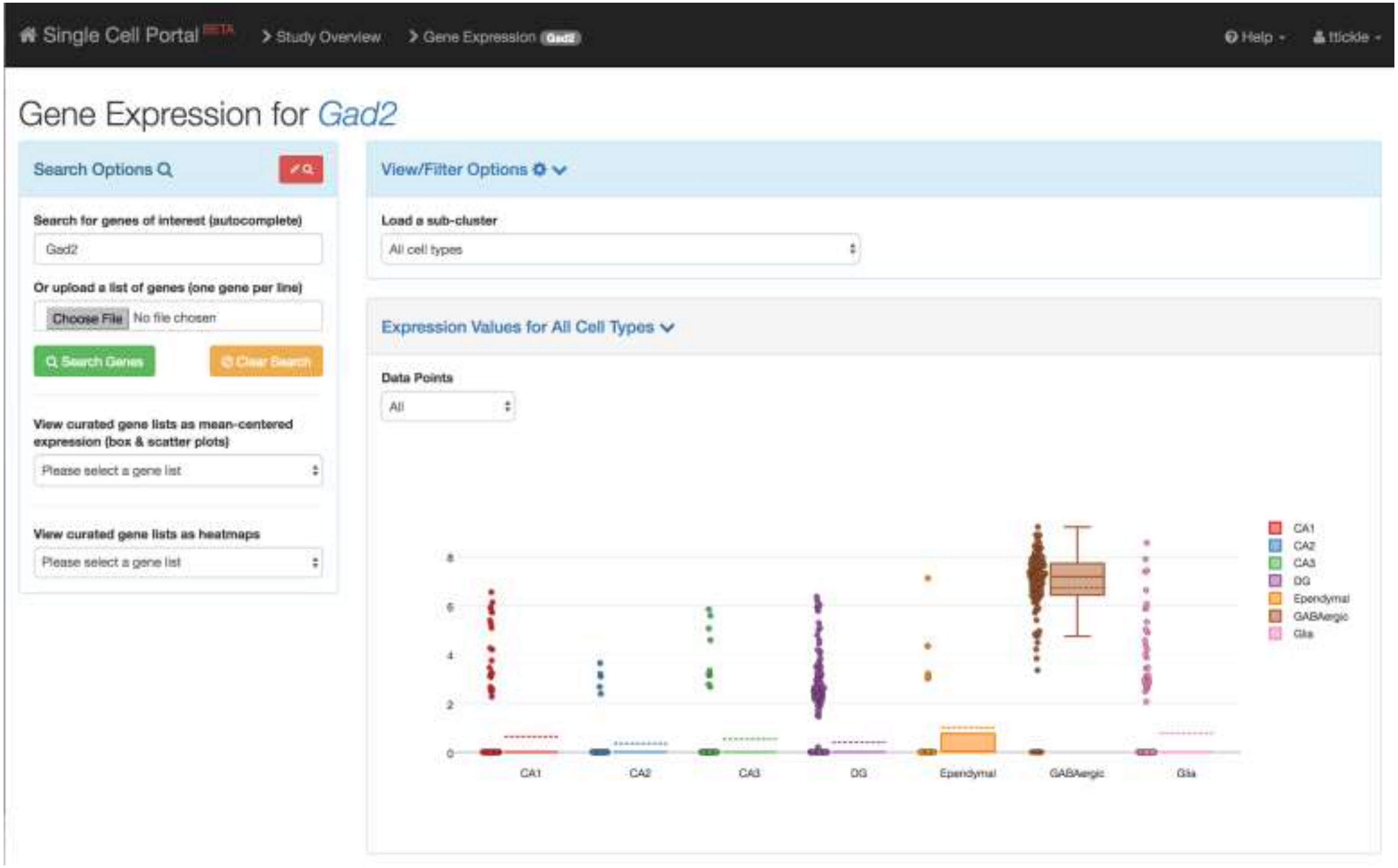
Load a sub-cluster

All cell types

Single nucleus RNA-seq of cell diversity in the adult mouse hippocampus (sNuc-Seq) Clusters

The Single Cell Portal

Gene Expression Can be Viewed Across Clusters



Gene Expression for *Gad2*

Search for genes of interest (autocomplete): Gad2

Or upload a list of genes (one gene per line): No file chosen

View curated gene lists as mean-centered expression (box & scatter plots): Please select a gene list

View curated gene lists as heatmaps: Please select a gene list

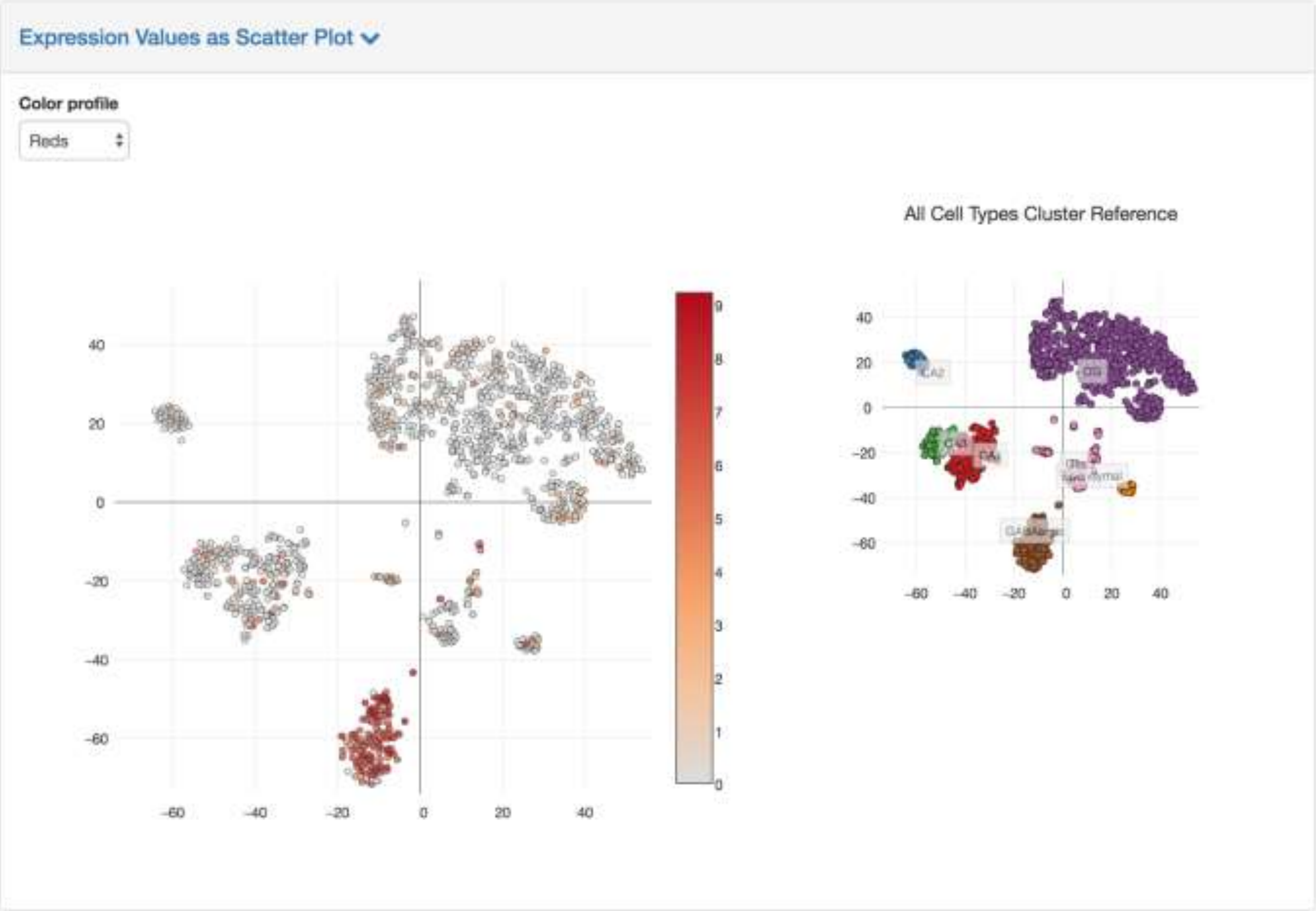
Load a sub-cluster: All cell types

Expression Values for All Cell Types

Data Points: All

The Single Cell Portal

Gene Expression Can be Viewed Across Clusters

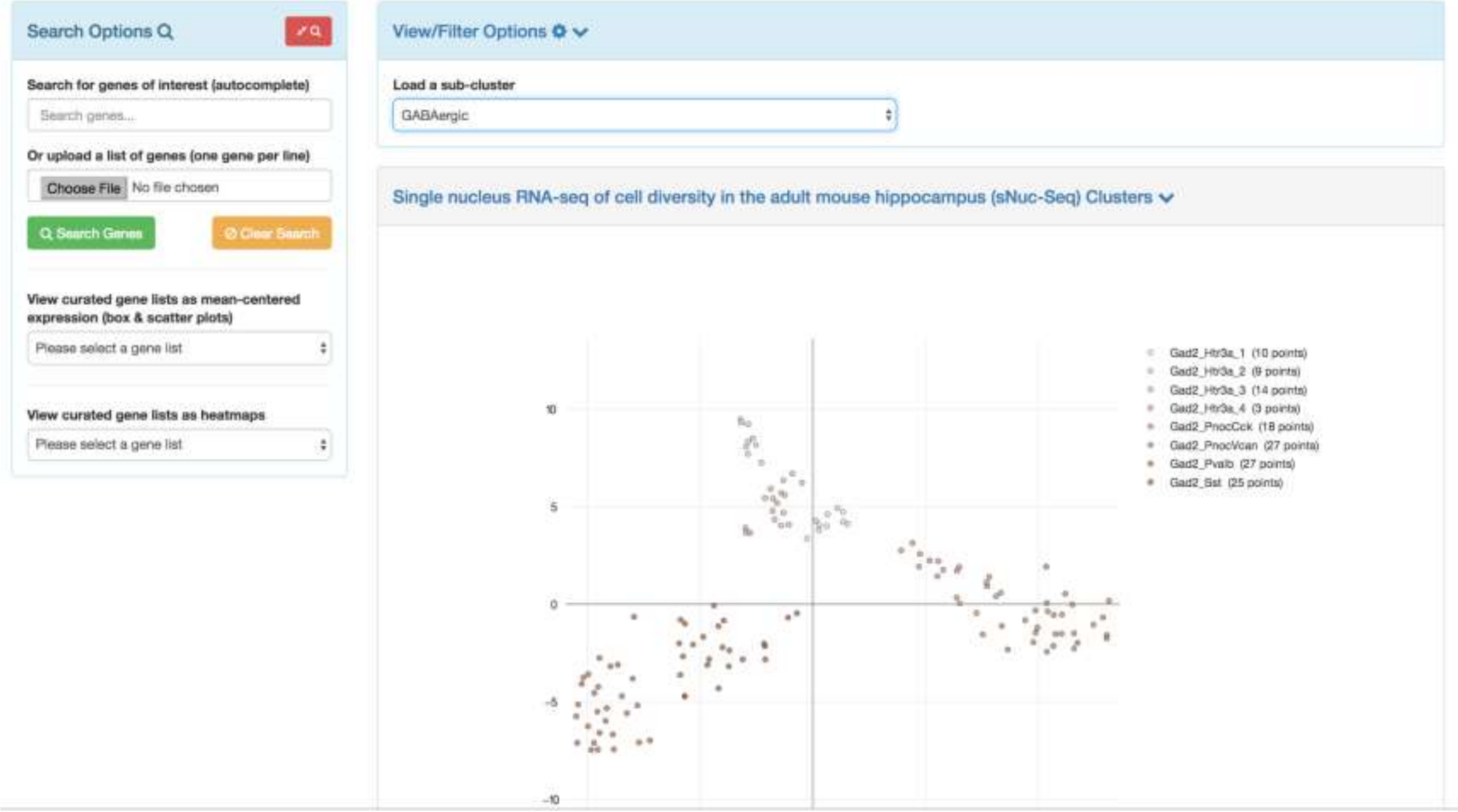


Expression Values as Scatter Plot

Color profile: Reds

The Single Cell Portal

Multiple Clustering Can be Used



Search Options

Search for genes of interest (autocomplete)

Search genes...

Or upload a list of genes (one gene per line)

Choose File No file chosen

Search Genes | Clear Search

View curated gene lists as mean-centered expression (box & scatter plots)

Please select a gene list

View curated gene lists as heatmaps

Please select a gene list

View/Filter Options

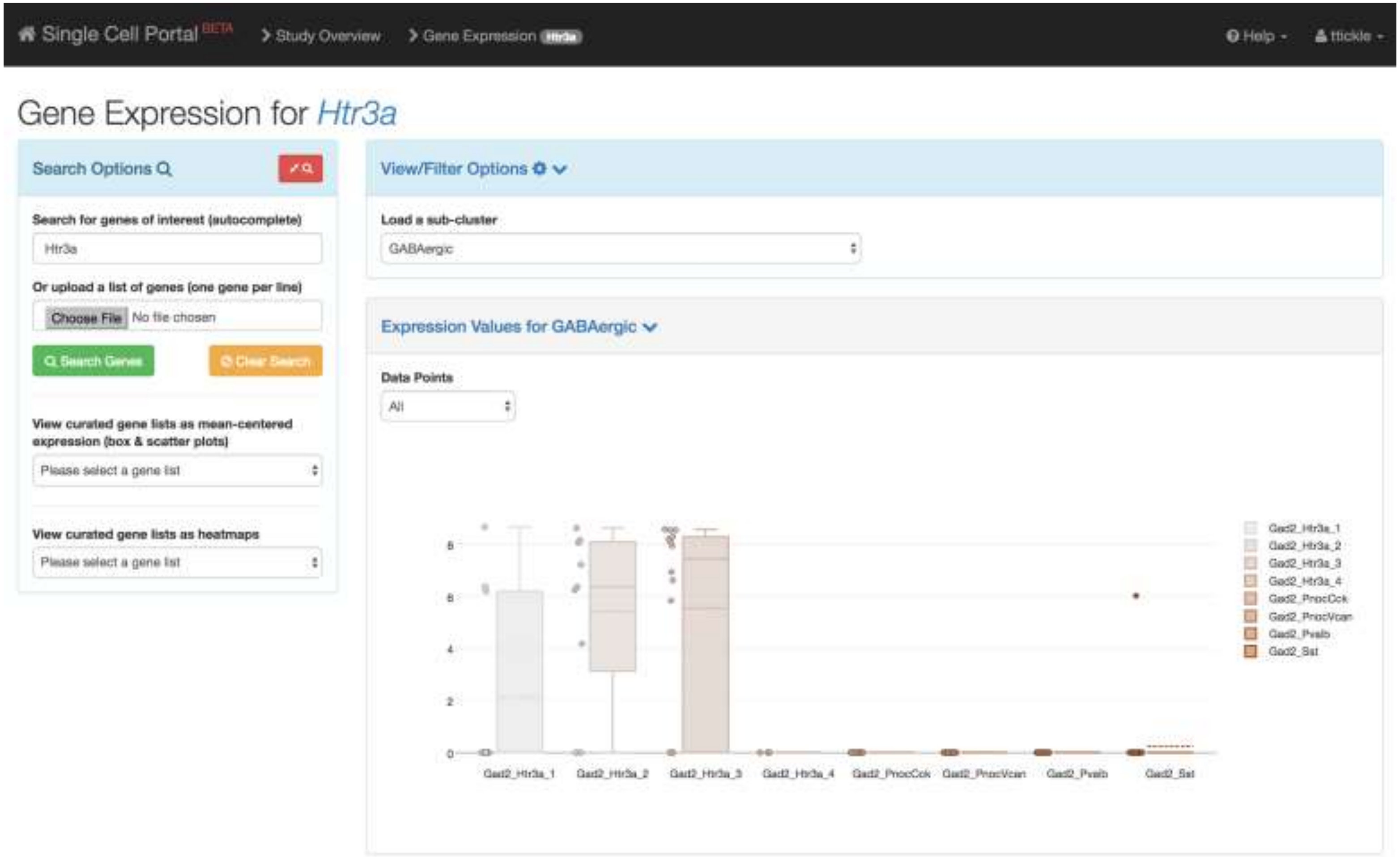
Load a sub-cluster

GABAergic

Single nucleus RNA-seq of cell diversity in the adult mouse hippocampus (sNuc-Seq) Clusters

The Single Cell Portal

Genes Can Be Viewed in Many Clusters



Gene Expression for *Htr3a*

Search for genes of interest (autocomplete): Htr3a

Or upload a list of genes (one gene per line): Choose File No file chosen

View curated gene lists as mean-centered expression (box & scatter plots): Please select a gene list

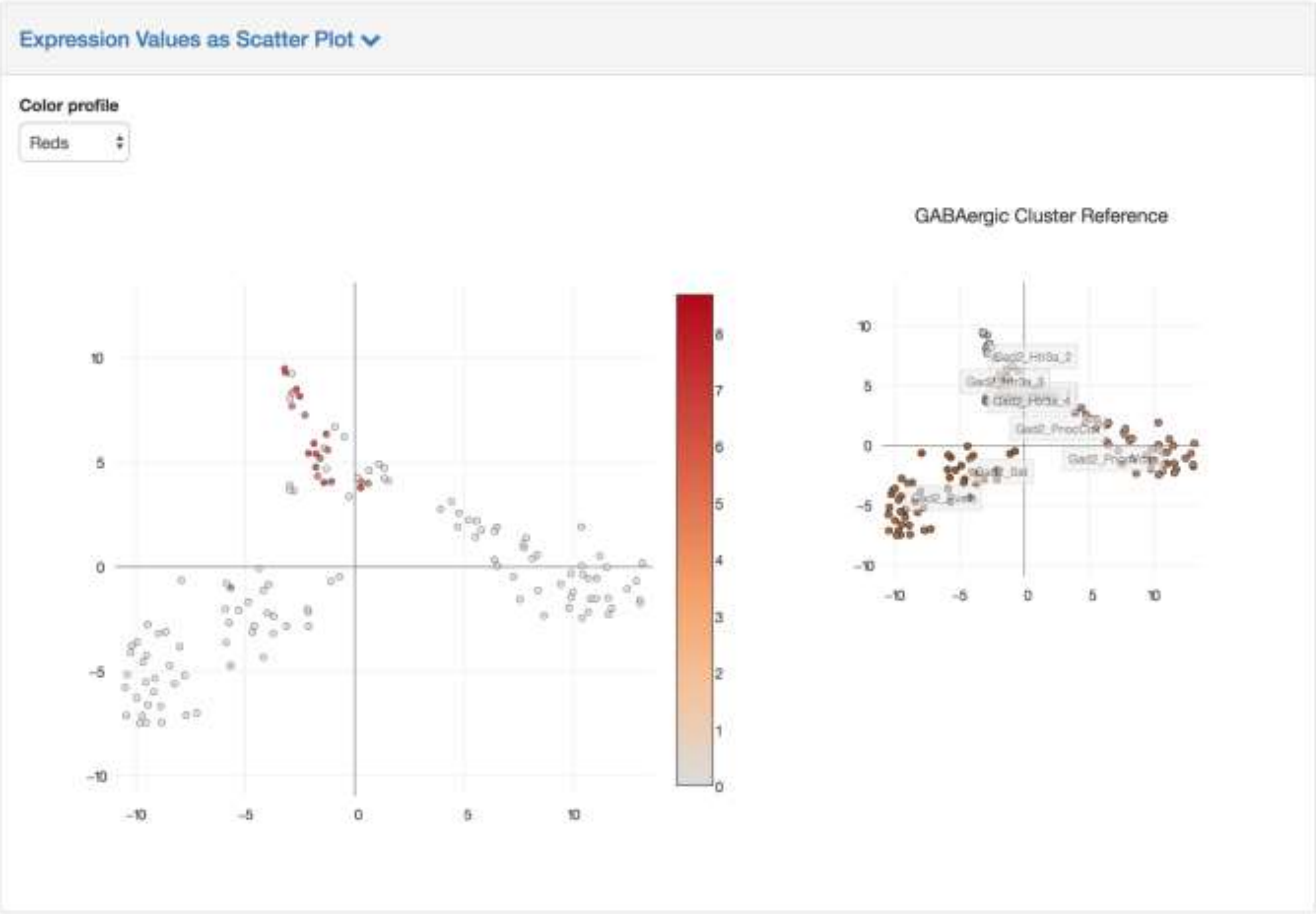
View curated gene lists as heatmaps: Please select a gene list

Load a sub-cluster: GABAergic

Data Points: All

The Single Cell Portal

Expression Can Be Shown in Many Clusterings



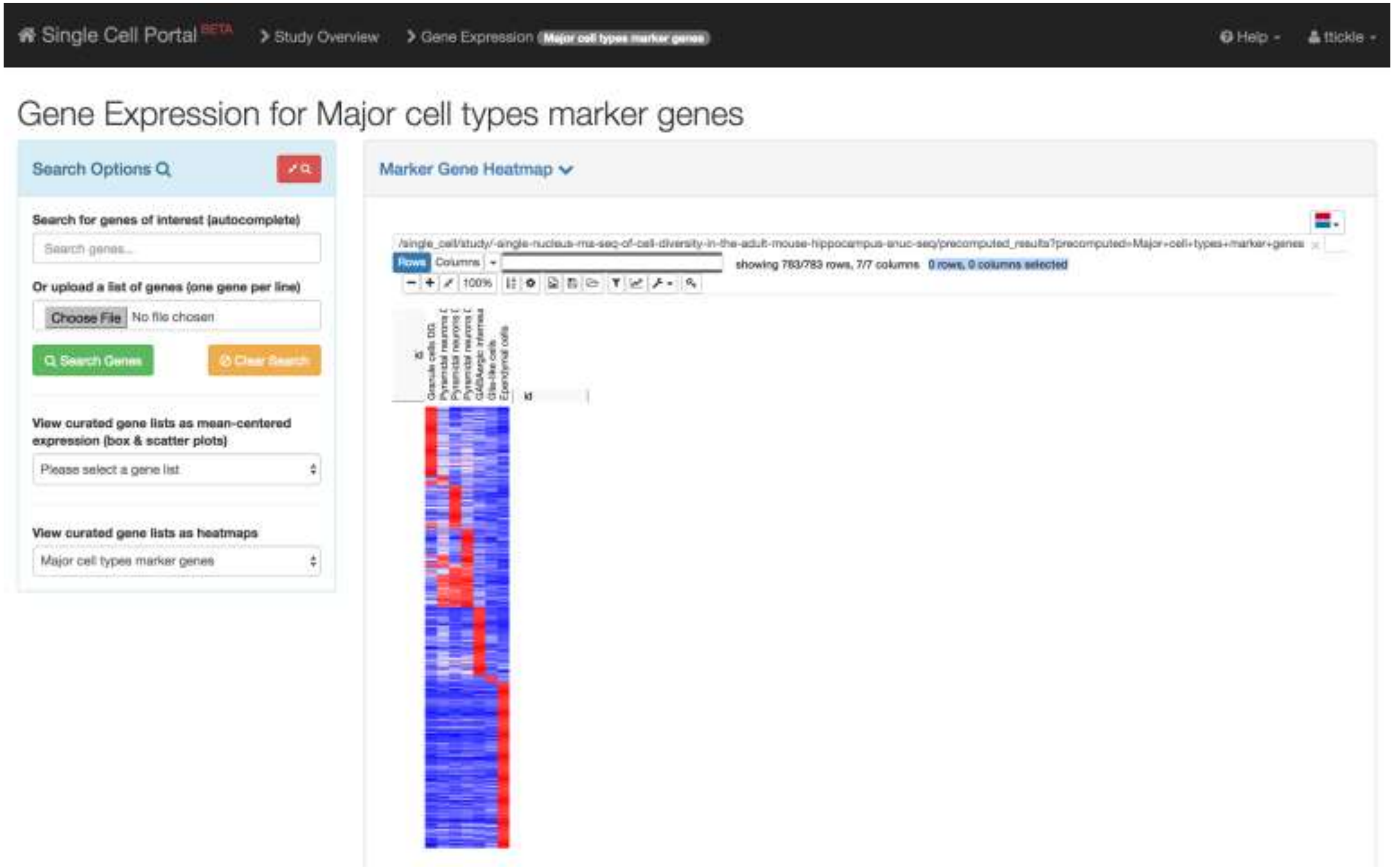
Expression Values as Scatter Plot

Color profile: Reds

GABAergic Cluster Reference

The Single Cell Portal

Expression in Clusters Can Also Be Shown as Heatmaps



Gene Expression for Major cell types marker genes

Search for genes of interest (autocomplete)

Search genes...

Or upload a list of genes (one gene per line)

No file chosen

View curated gene lists as mean-centered expression (box & scatter plots)

Please select a gene list

View curated gene lists as heatmaps

Major cell types marker genes

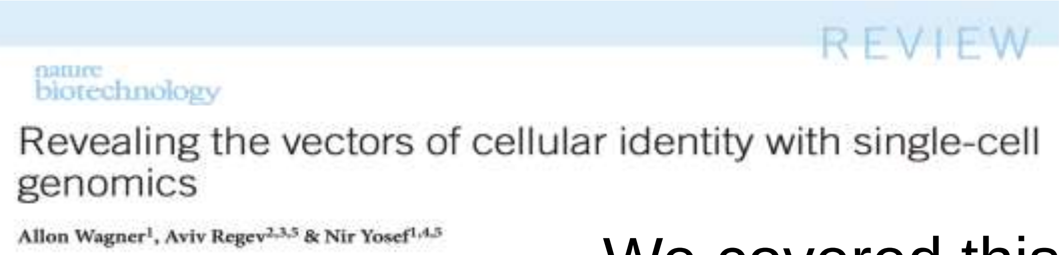
/single_cell/study/-single-nucleus-rna-seq-of-cell-diversity-in-the-adult-mouse-hippocampus-snuc-seq/precomputed_results?precomputed=Major+cell+types+marker+genes

showing 783/783 rows, 7/7 columns 0 rows, 0 columns selected

The Single Cell Portal

- Studies can be …
  - Private
  - Private but shared privately
  - Public but with data inaccessible
  - Public

Section: Wrapping Up
What Did We Miss (So Much)?



REVIEW

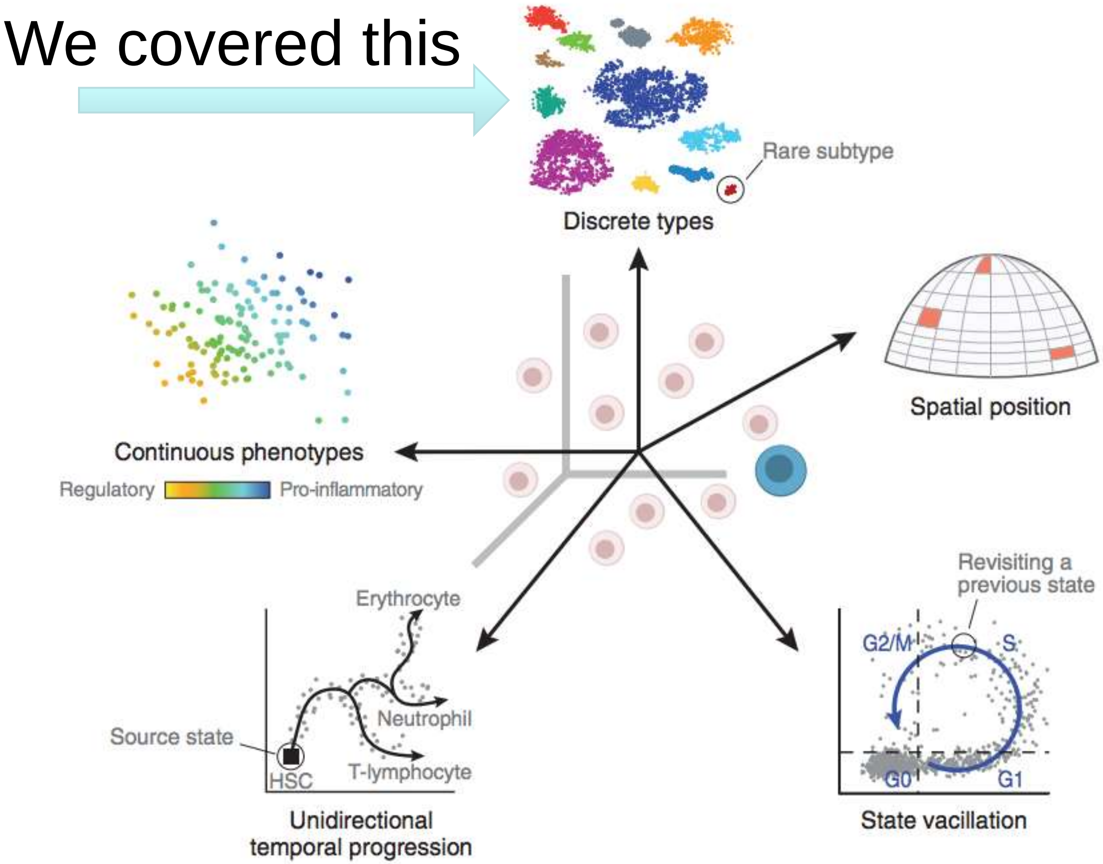
nature biotechnology

Revealing the vectors of cellular identity with single-cell genomics

Allon Wagner[1], Aviv Regev[2,3,5] & Nir Yosef[1,4,5]

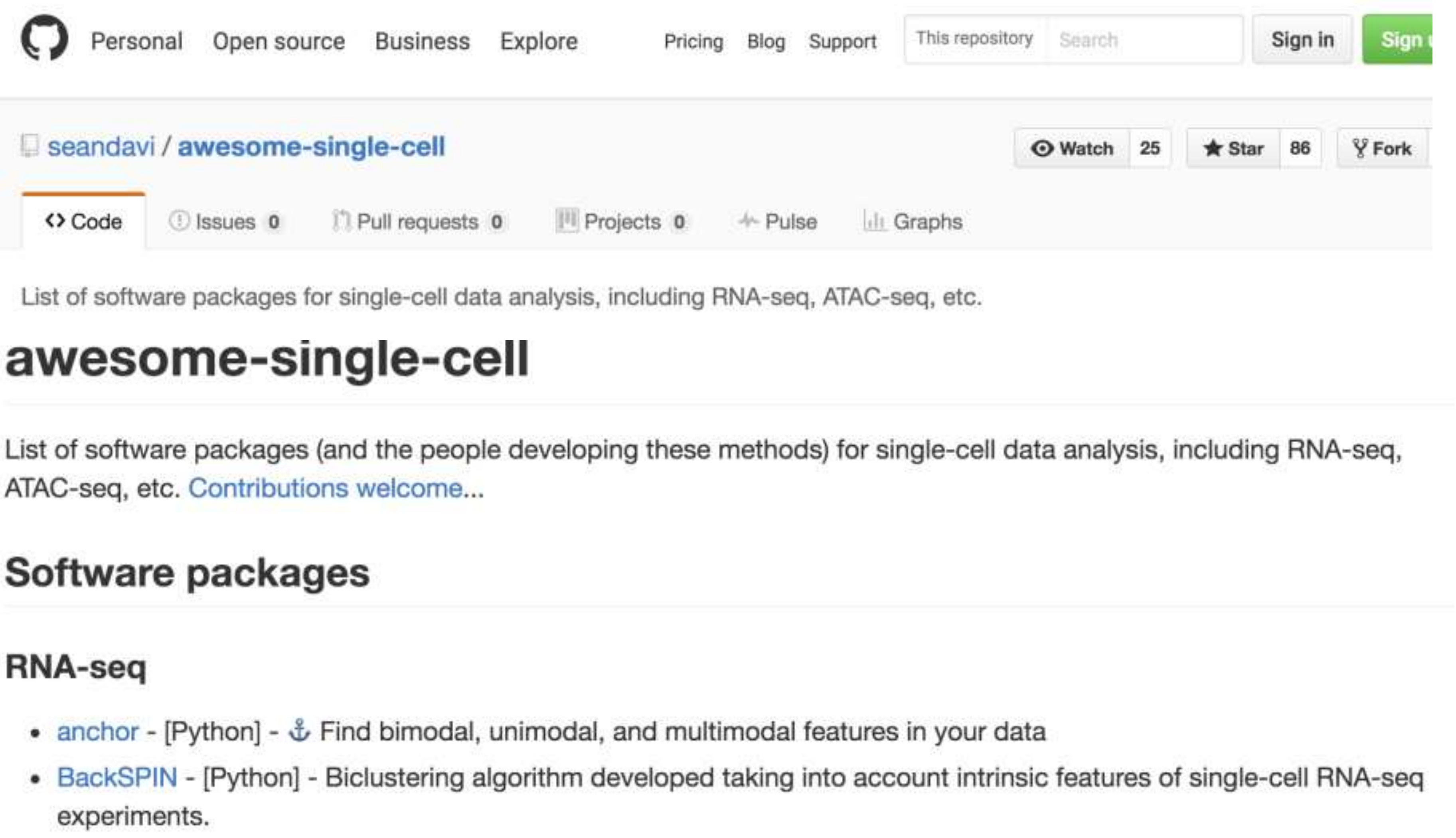
So much more to learn!

We covered this



Awesome List

https://github.com/seandavi/awesome-single-cell



List of software packages for single-cell data analysis, including RNA-seq, ATAC-seq, etc.

awesome-single-cell

List of software packages (and the people developing these methods) for single-cell data analysis, including RNA-seq, ATAC-seq, etc. Contributions welcome...

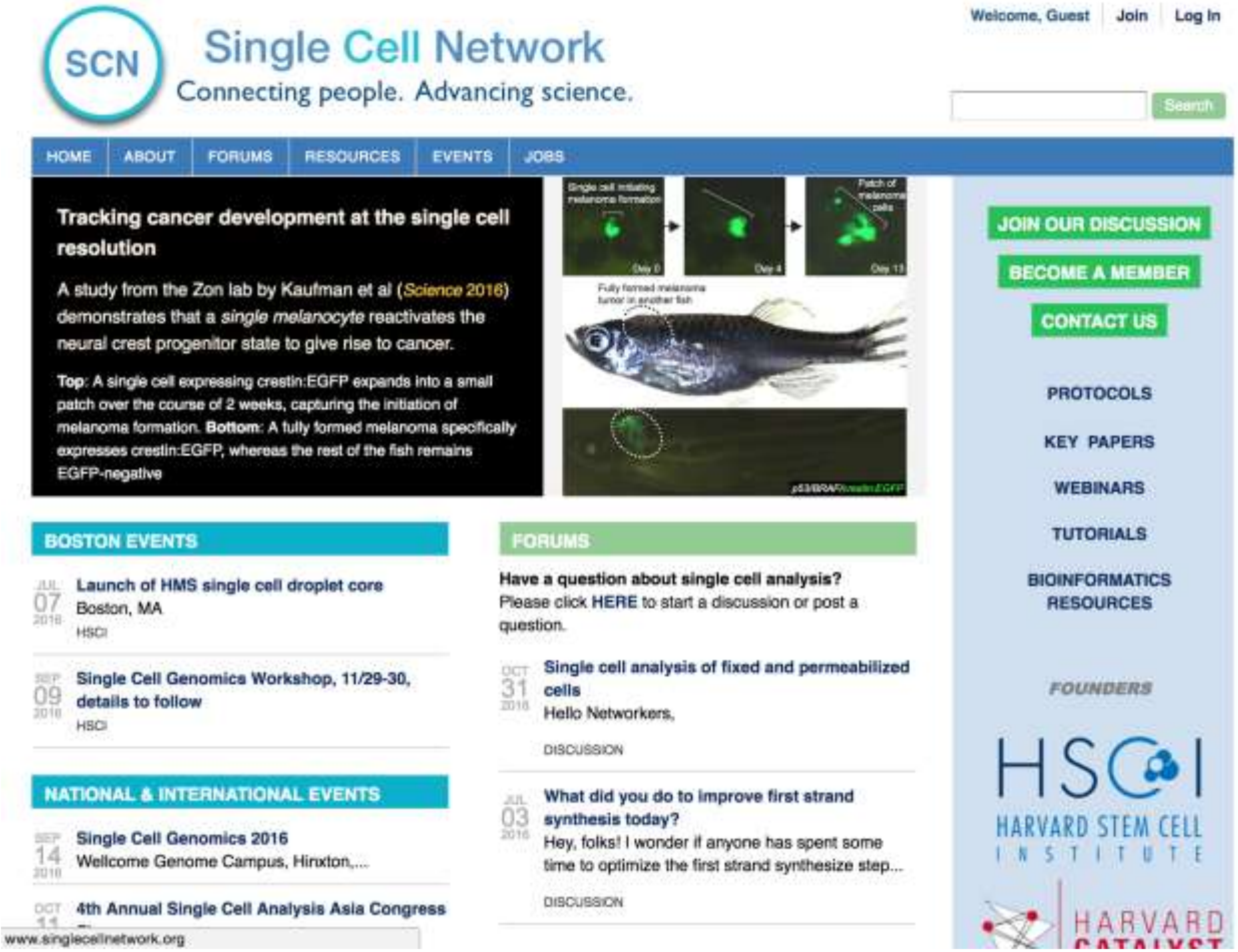
Software packages

RNA-seq

- anchor - [Python] - ⚓ Find bimodal, unimodal, and multimodal features in your data
- BackSPIN - [Python] - Biclustering algorithm developed taking into account intrinsic features of single-cell RNA-seq experiments.

Single Cell Network

www.singlecellnetwork.org



Welcome, Guest | Join | Log In

SCN

Single Cell Network

Connecting people. Advancing science.

Tracking cancer development at the single cell resolution

A study from the Zon lab by Kaufman et al (*Science* 2016) demonstrates that a *single melanocyte* reactivates the neural crest progenitor state to give rise to cancer.

Top: A single cell expressing crestin:EGFP expands into a small patch over the course of 2 weeks, capturing the initiation of melanoma formation. **Bottom**: A fully formed melanoma specifically expresses crestin:EGFP, whereas the rest of the fish remains EGFP-negative

Single cell initiating melanoma formation
Day 0
Day 4
Patch of melanoma cells
Day 13
Fully formed melanoma tumor in another fish
p53/BRAF/crestin:EGFP

BOSTON EVENTS

JUL 07 2016 **Launch of HMS single cell droplet core**
Boston, MA
HSCI

SEP 09 2016 **Single Cell Genomics Workshop, 11/29-30, details to follow**
HSCI

NATIONAL & INTERNATIONAL EVENTS

SEP 14 2016 **Single Cell Genomics 2016**
Wellcome Genome Campus, Hinxton,...

OCT 11 **4th Annual Single Cell Analysis Asia Congress**

FORUMS

Have a question about single cell analysis?
Please click **HERE** to start a discussion or post a question.

OCT 31 2016 **Single cell analysis of fixed and permeabilized cells**
Hello Networkers,
DISCUSSION

JUL 03 2016 **What did you do to improve first strand synthesis today?**
Hey, folks! I wonder if anyone has spent some time to optimize the first strand synthesize step...
DISCUSSION

FOUNDERS

www.singlecellnetwork.org

Thank You

Aviv Regev

Brian Haas

Adam Haber

Anindita Basu

Asma Bankapur

Chloe Villani

Karthik Shekhar

Kristine Schwenck

Matan Hofree

Michel Cole

Monika Kowalczyk

Nir Yosef

Sean Simmons

Regev Single Cell Working Group

Today's Attendees

Questions?