



Applications of Single-Cell Profiling Methods to Enhance Mechanistic Understanding of Toxicological Responses

Society of Toxicology Symposium

David Gallegos, U.S. EPA, Chair

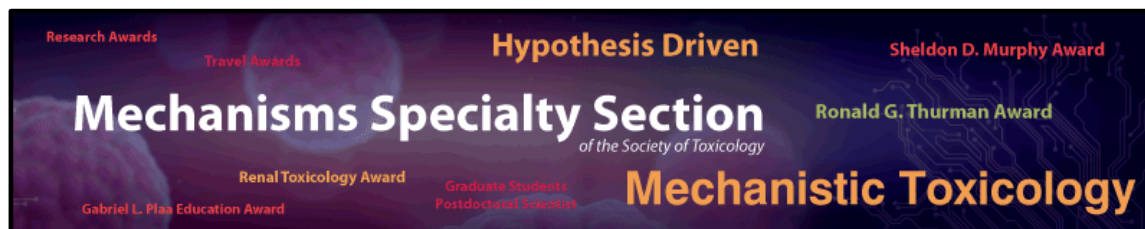
Kelly Bakulski, University of Michigan, Co-Chair

March 30, 2022

SOT | Society of
Toxicology
Creating a Safer and Healthier World by Advancing
the Science and Increasing the Impact of Toxicology

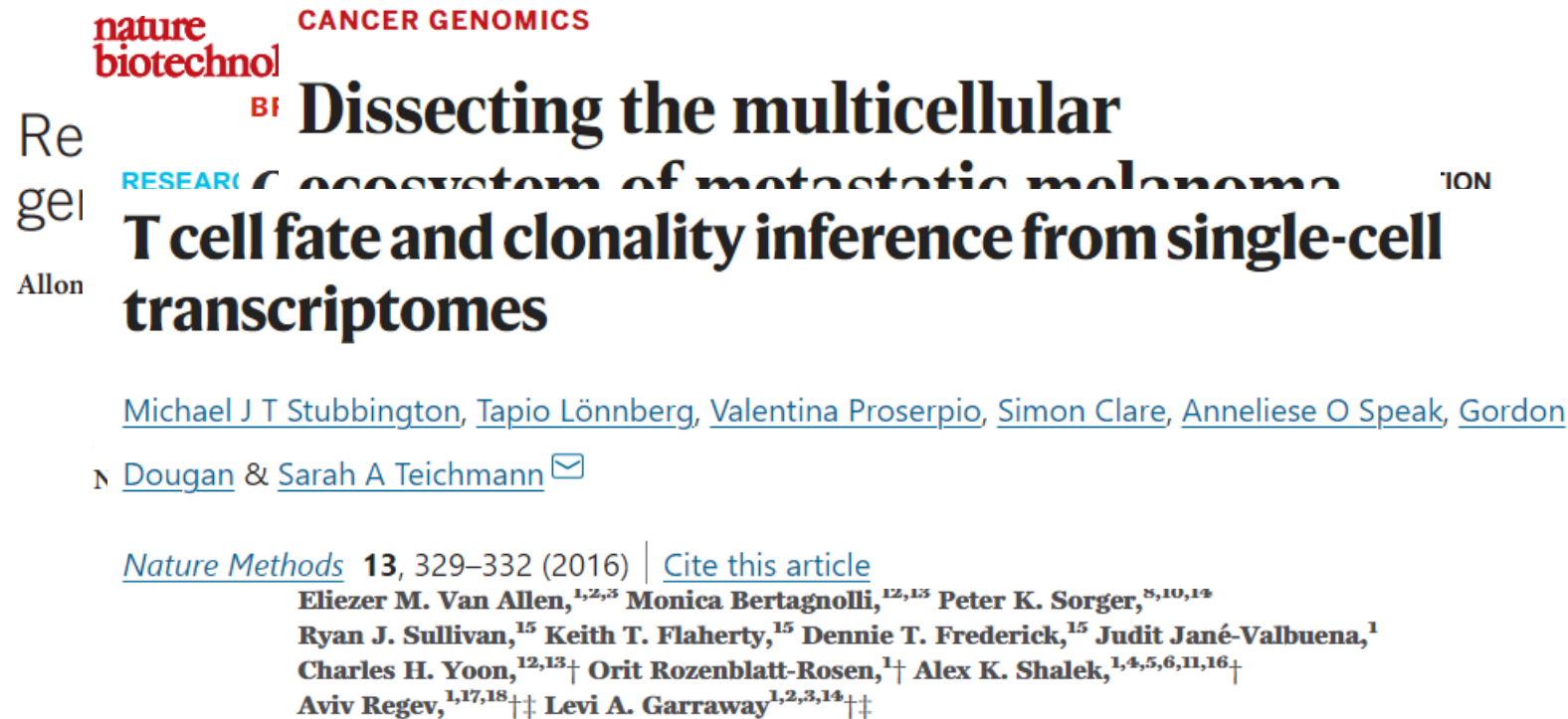
Session Introduction: Advances and Technical Considerations in Single-Cell Profiling for Use in Mechanistic Toxicology

Thanks to our sponsoring specialty sessions

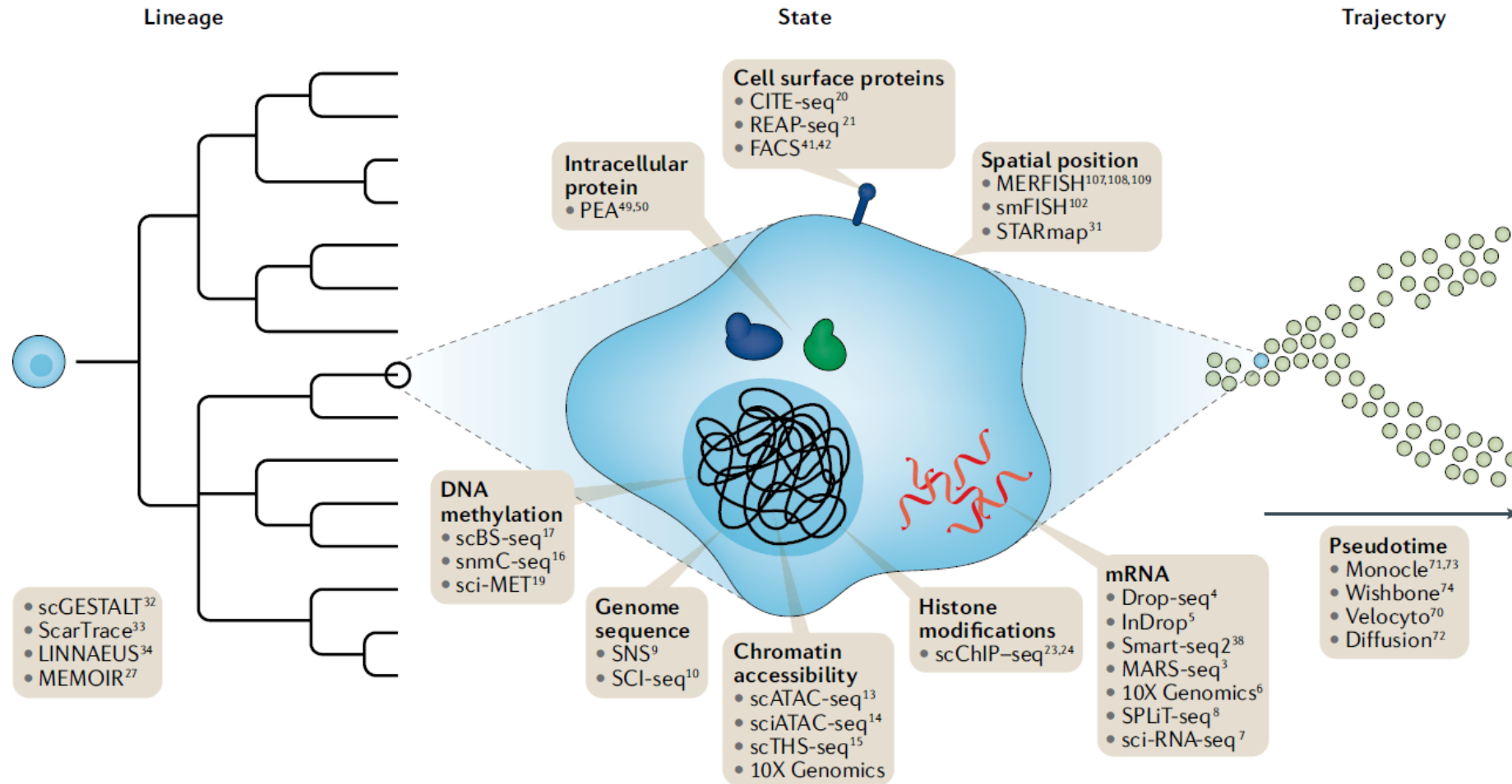


Disclaimer: This presentation does not represent U.S. EPA policy or endorsement

SINGLE-CELL SEQUENCING CONCEPTUAL UTILITY



RAPID EXPANSION IN SINGLE CELL PROFILING



Stuart and Satija, *Nature Reviews Genetics*, 2019

PRINCIPLES AND STEPS OF SINGLE-CELL SEQ WORKFLOW

I. Tissue Procurement



Source:

- Primary human
- Model organism
- Cell culture

Key considerations:

- Biological variation
- Sampling/handling variation
- Duration of sourcing

Study design:

- Biological replicates
- Technical replicates
- Cell number calculation
- Workflow optimization

II. Tissue Dissociation



Method:

- Mechanical mincing
- Enzymatic digestion
- Automated blending
- Microfluidics devices

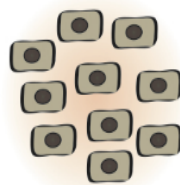
Key considerations:

- Experimental consistency
- Shortest duration
- Highest cell/nucleus quality
- Representation of all cell types

Quality control:

- FACS analysis
- qPCR for marker genes
- Imaging of cell integrity
- RNA quality (RIN)

III. Cell Enrichment (optional)



Method:

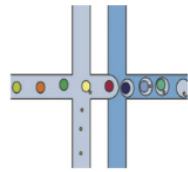
- Differential centrifugation, sedimentation, filtration
- Antibody labeling for positive/negative selection
- Flow cytometry or bead-based enrichment
- Dead cell removal

Key considerations:

- Additional handling
- Longer duration
- Loss of RNA quality
- Transcriptome changes

PRINCIPLES AND STEPS OF SINGLE-CELL RNA-SEQ WORKFLOW

IV. Single Cell RNAseq Platform



Method:

- Droplet-based
- Tube-based after FACS
- Microwell-based
- Microfluidics-enabled

Key considerations:

- Cell throughput and handling time
- Gene coverage and cell type detection
- Whole transcript versus 3'end counting
- Imaging capability for doublet detection

V. Library Sequencing



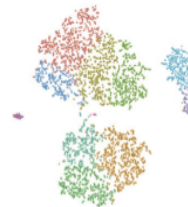
Method:

- Illumina NGS
- Compatible with cDNA library

Sequencing depth considerations:

- 3'end counting: low depth ~50K RPC
- Whole transcript: high depth ~1M RPC
- Alternative splicing: ~20-30M RPC
- Iterative optimization for biological system

VI. Computational Analysis



Key considerations:

- Separation of *batch* and *condition*
- Technical vs. biological variation

Sample Batch correction approaches:

- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST

TISSUE/SAMPLE PROCUREMENT

Difficulties in Single Cell Genomics:

Sensitive Preparations
Preparation/Isolation Artifacts

High Technical Variability and Noise
Substantial Dropout/ Zero Counts

I. Tissue Procurement



Source:

- Primary human
- Model organism
- Cell culture

Key considerations:

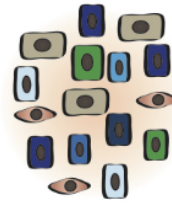
- Biological variation
- Sampling/handling variation
- Duration of sourcing

Study design:

- Biological replicates
- Technical replicates
- Cell number calculation
- Workflow optimization

SINGLE CELLS VS. SINGLE NUCLEI

II. Tissue Dissociation



Method:

- Mechanical mincing
- Enzymatic digestion
- Automated blending
- Microfluidics devices

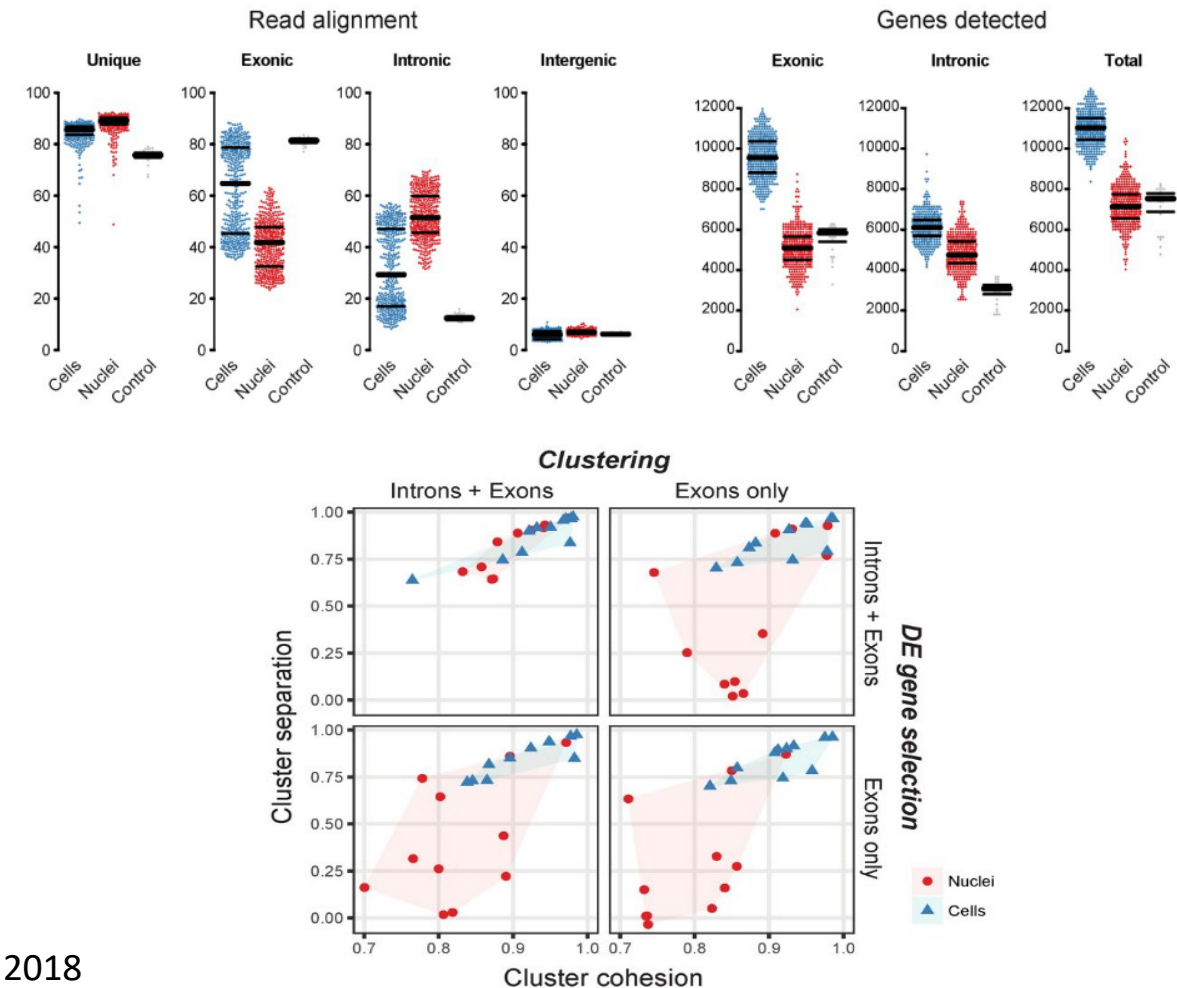
Key considerations:

- Experimental consistency
- Shortest duration
- Highest cell/nucleus quality
- Representation of all cell types

Quality control:

- FACS analysis
- qPCR for marker genes
- Imaging of cell integrity
- RNA quality (RIN)

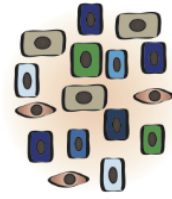
- Cell Type/State and Harvest Limitations
- Frozen/Archived/Post-Mortem Tissue
- Total Detectable Transcripts
- Intronic Sequence Utility
- Nuclear Sequence Proportion
- Transcript Types and Differential Enrichment (ncRNA)



Bakken et al, *PLOS One*, 2018

CELL AND NUCLEUS DISSOCIATION

II. Tissue Dissociation



Method:

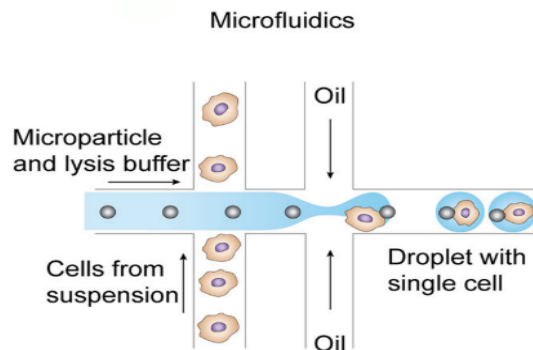
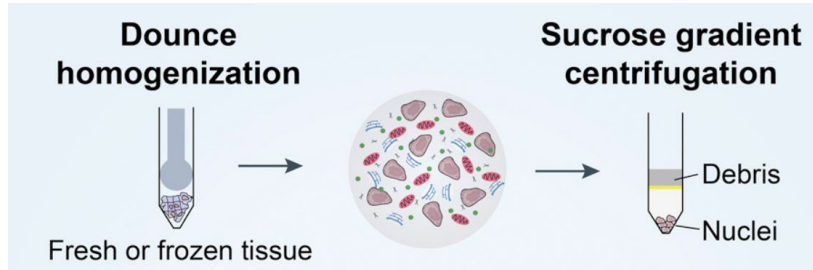
- Mechanical mincing
- Enzymatic digestion
- Automated blending
- Microfluidics devices

Key considerations:

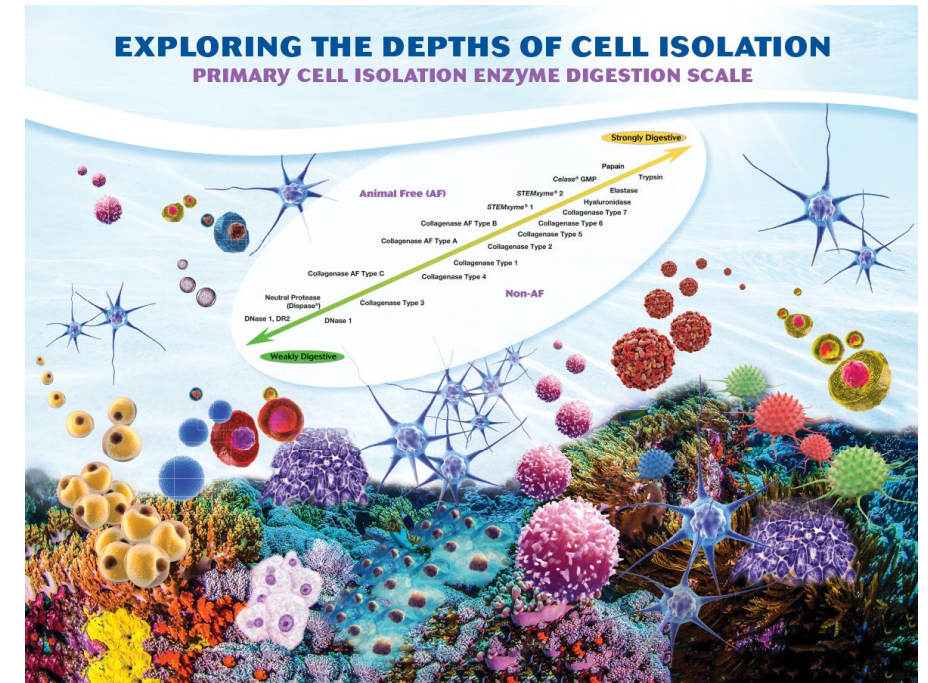
- Experimental consistency
- Shortest duration
- Highest cell/nucleus quality
- Representation of all cell types

Quality control:

- FACS analysis
- qPCR for marker genes
- Imaging of cell integrity
- RNA quality (RIN)



- 1) *Tissue Dissection*
- 2) *Mechanical Mincing*
- 3) *Enzymatic/proteolytic (ECM) digestion*
- 4) *Mechanical Agitation*
- 5) *Optional Enrichment*

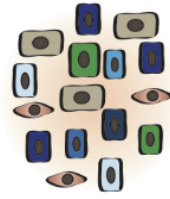


Hu et al, *Molecular Cell*, 2017

Hwang et al, *Experimental & Molecular Medicine*, 2018

CELL AND NUCLEUS DISSOCIATION

II. Tissue Dissociation



Method:

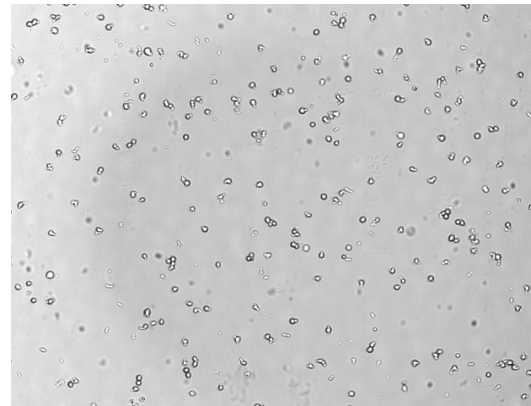
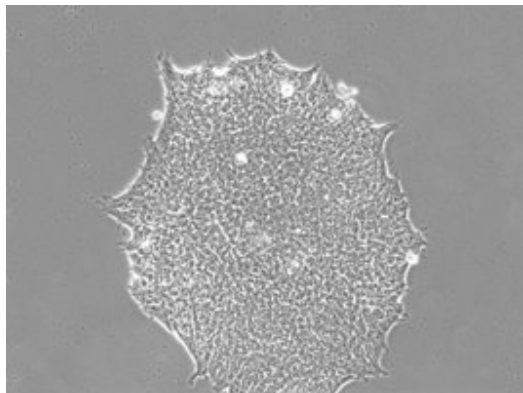
- Mechanical mincing
- Enzymatic digestion
- Automated blending
- Microfluidics devices

Key considerations:

- Experimental consistency
- Shortest duration
- Highest cell/nucleus quality
- Representation of all cell types

Quality control:

- FACS analysis
- qPCR for marker genes
- Imaging of cell integrity
- RNA quality (RIN)



Results:

Count

FL1: 1121
FL2: 43

Viability
96.4%

Concentration

3.24×10^6 cells/mL
 1.22×10^5 cells/mL

Mean Diameter

18.1 micron
11.9 micron

Suspension Buffer Options

Basic: 1XPBS + 0.04% BSA

Enhanced for Transcript Stasis: 1XPBS + 0.04% BSA
+100 mM D-AP5, 5 mg ml⁻¹ of actinomycin D, 20 mM Triptolide, 10 mg ml⁻¹ of anisomycin

Enhanced for Cell Suspension and Separation: 1X PBS + 0.6% BSA + 15% Optiprep Density Gradient Medium

Other Possible Reagents and Changes: 1 nM tetrodotoxin citrate, Kynurenic Acid, Itraconazole

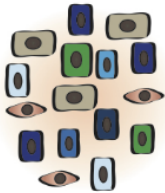
Key Considerations:

Viscosity

Pausing

Downstream Platforms
Transcriptional/Translational
Minimizing Processing Artifacts
Cell Clumping

CELL AND NUCLEUS DISSOCIATION



II. Tissue Dissociation

Method:

- Mechanical mincing
- Enzymatic digestion
- Automated blending
- Microfluidics devices

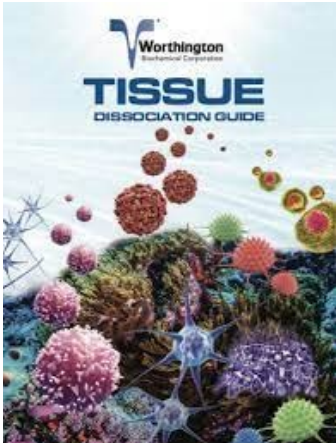
Key considerations:

- Experimental consistency
- Shortest duration
- Highest cell/nucleus quality
- Representation of all cell types

Quality control:

- FACS analysis
- qPCR for marker genes
- Imaging of cell integrity
- RNA quality (RIN)

Tissue Tables (references, grouped by tissue type and species)



Adipose/Fat	Adrenal	Bone	Brain
Cartilage	Colon	Endothelial	Epithelial
Eye	Heart	Intestine	Kidney
Liver	Lung	Lymph nodes	Mammary
Miscellaneous	Muscle	Neural	Pancreas
Parotid	Pituitary	Prostate	Reproductive
Scales	Skin	Spleen	Stem
Thymus	Thyroid/Parathyroid	Tonsil	Tumor

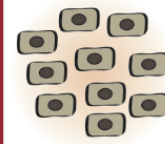
<https://www.worthington-biochem.com/tissuedissociation/>

Brain						Brain
Species	Species Detail	Cell(s)	Enzyme(s)	Medium	Reference	
Human	Human	Microglia	Collagenase Type 1: 300 u/ml Trypsin: 0.125%	DMEM	Mizee, M., Miedema, S., van der Poel, M., Adelia , S., van Strien, M., Melief, J., Smolders, J., Hendrickx, D., Heutinck, K., Hamann, J. and Huitinga, I.: Isolation of Primary Microglia from the Human Post-Mortem Brain: Effects of Ante- and Post- Mortem Variables., <i>Acta Neuropathol</i> 5, 16, 2017 (11604)	
	Human, adult	Neuronal	Papain: 20 u/ml	Neurobasal	Spaethling, J., Na, Y., Lee, J., Ulyanova, A., Baltuch, G., Bell, T., Brem, S., Chen, H., Dueck, H., Fisher, S., Garcia, M., Khaladkar, M., Kung, D., Lucas, T., O'Rourke, D. and Stefanik, D.: Primary Cell Culture of Live Neurosurgically Resected Aged Adult Human Brain Cells and Single Cell Transcriptomics., <i>Cell Rep</i> 18, 791-803, 2017 (11673)	
	Human, fetal and mature	Astrocytes and neurons	Papain: 7.5-20 u/ml	RPMI	Zhang, Y., Sloan, S., Clarke, L., Caneda, C., Plaza, C., Blumenthal, P., Vogel, H., Steinberg, G., Edwards, M., Li, G., Duncan, J., Cheshier, S., Shuer, L., Chang, E., Grant, G., Gephart, M. and Barres, B.: Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse., <i>Neuron</i> 89, 37-53, 2016 (11490)	

CELL ENRICHMENT AND ISOLATION

- Limiting Dilution
- Micromanipulation
- Differential Centrifugation
 - Debris Removal and Cell-Size Separation
- Flow Cytometry/Fluorescent Activated Sorting (FACS)
 - Debris Removal
 - Cell-Type Targeting
 - Doublet/Multiplet Filtering
- Immunoprecipitation (IP)
 - Cell-Type Targeting
- Concern: Time
- Goals:
 - Increase Purity/Viability
 - Decrease Debris
 - Target Specific Cell/Nuclei Populations

III. Cell Enrichment (optional)

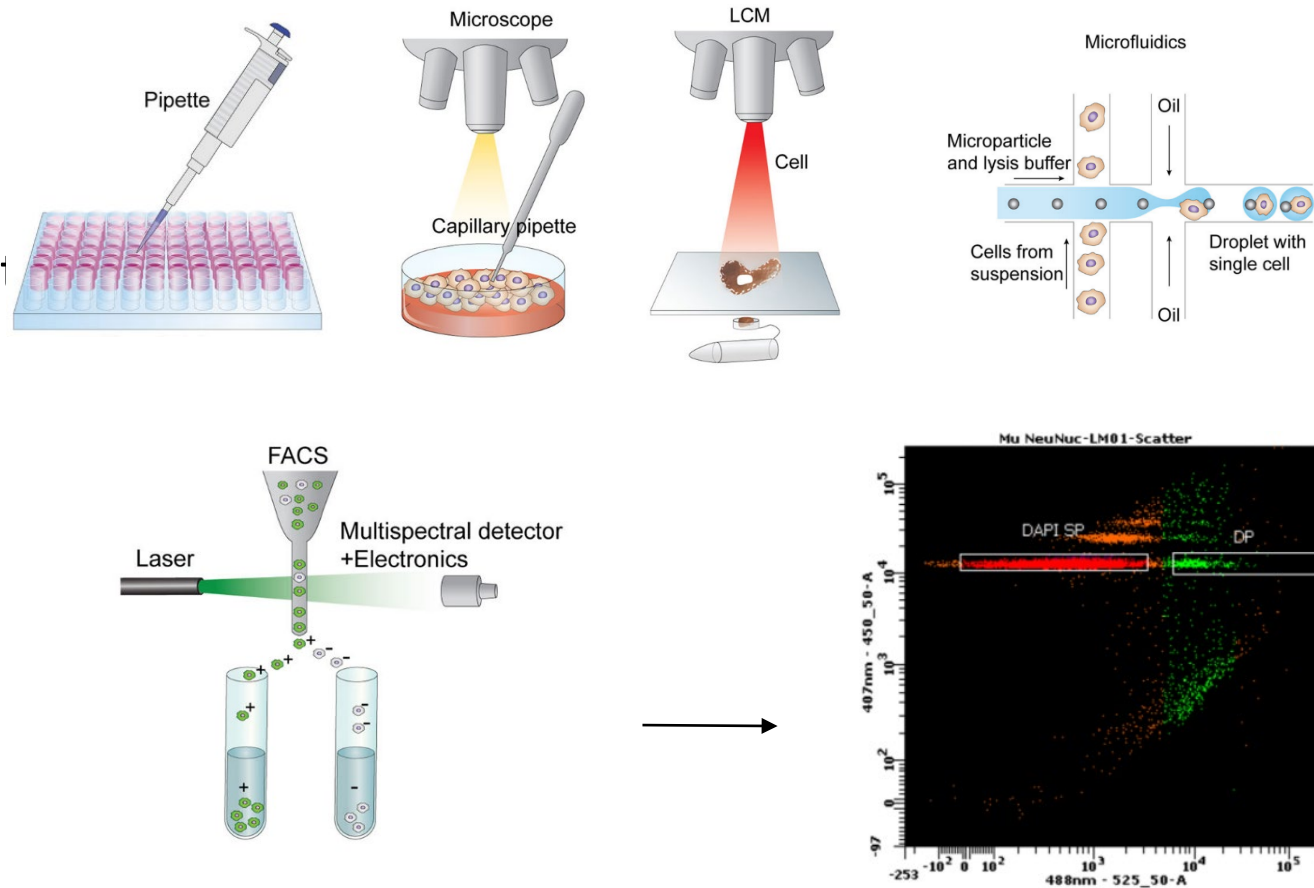


Method:

- Differential centrifugation, sedimentation, filtration
- Antibody labeling for positive/negative selection
- Flow cytometry or bead-based enrichment
- Dead cell removal

Key considerations:

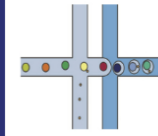
- Additional handling
- Longer duration
- Loss of RNA quality
- Transcriptome changes



Hwang et al, *Experimental & Molecular Medicine*, 2018

SINGLE-CELL RNA-SEQ PLATFORMS

IV. Single Cell RNAseq Platform

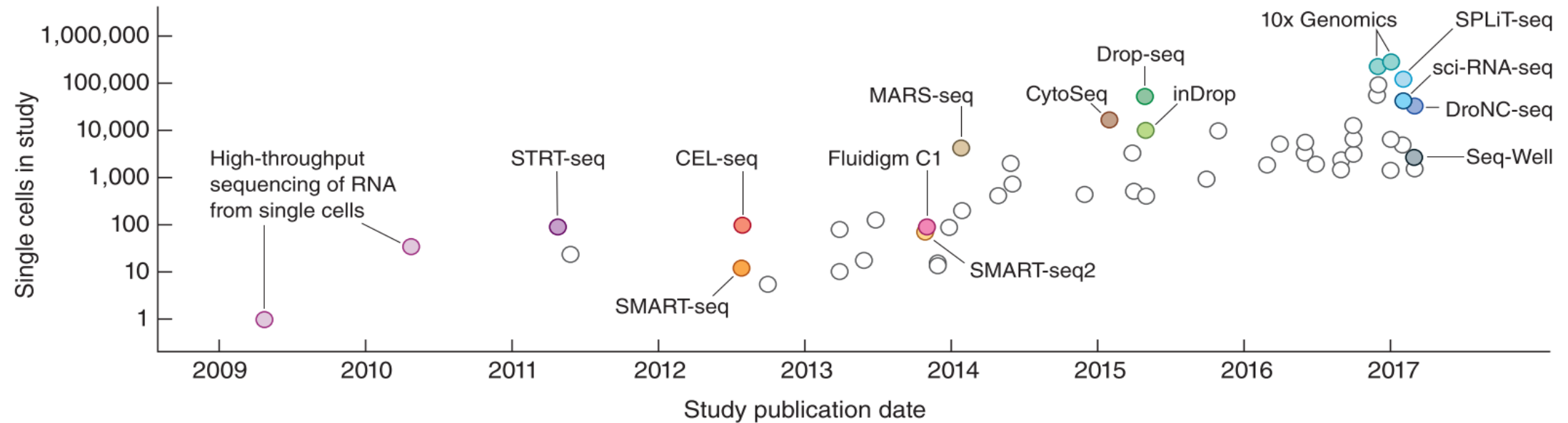
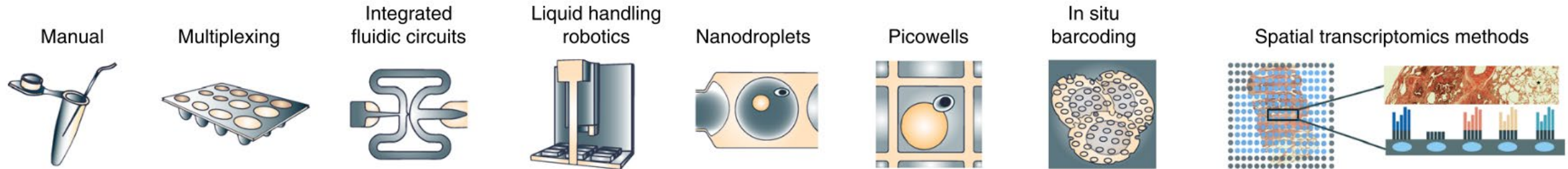


Method:

- Droplet-based
- Tube-based after FACS
- Microwell-based
- Microfluidics-enabled

Key considerations:

- Cell throughput and handling time
- Gene coverage and cell type detection
- Whole transcript versus 3' end counting
- Imaging capability for doublet detection



Aldridge and Teichmann, *Nature Communications*, 2020

SC-RNASEQ LIBRARY PRINCIPLES

V. Library Sequencing

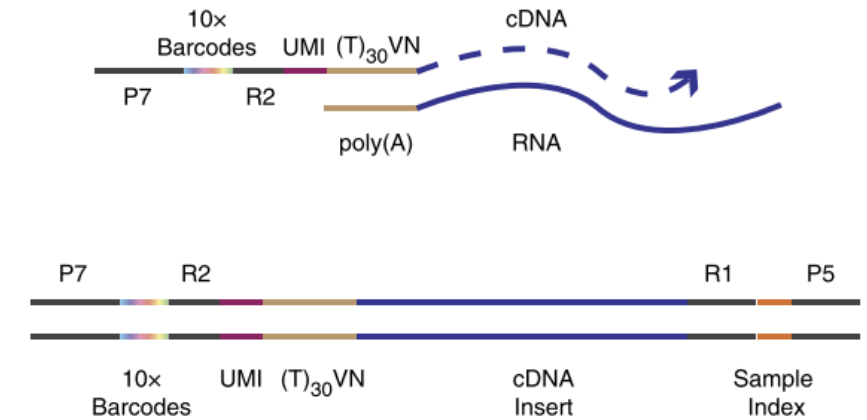
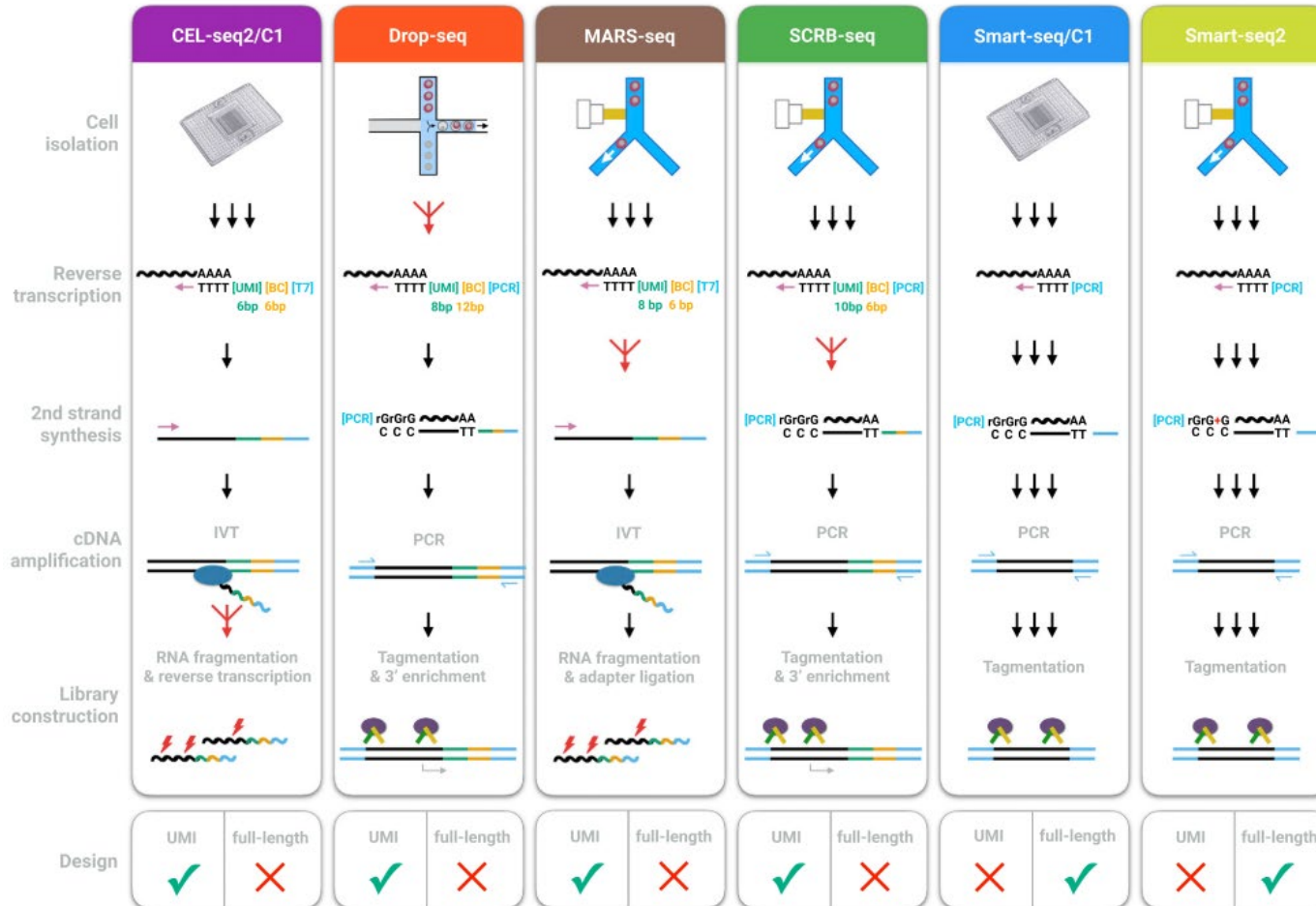


Method:

- Illumina NGS
- Compatible with cDNA library

Sequencing depth considerations:

- 3' end counting: low depth ~50K RPC
- Whole transcript: high depth ~1M RPC
- Alternative splicing: ~20-30M RPC
- Iterative optimization for biological system



Ziegenhain et al, *Molecular Cell*, 2017
 Zheng et al, *Nature Communications*, 2017

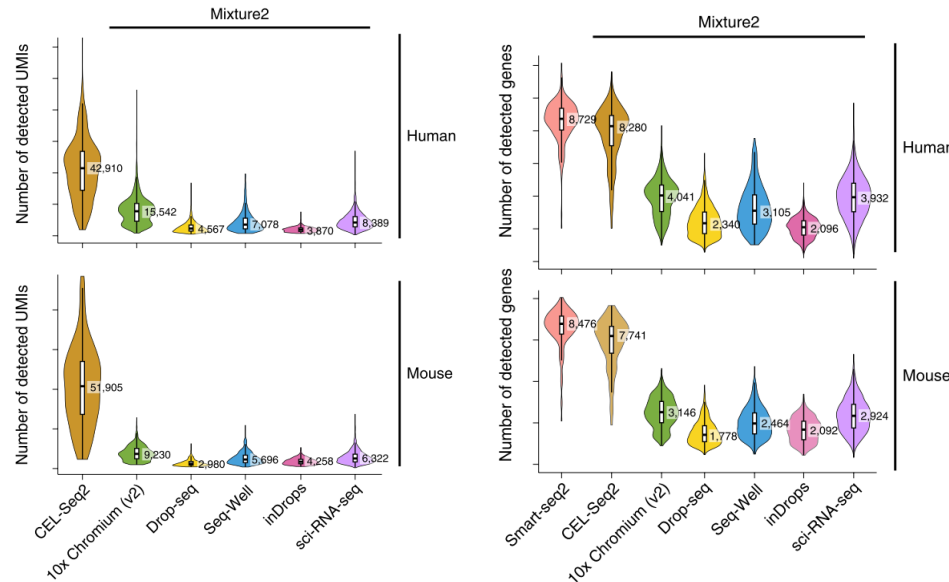
SINGLE-CELL RNA-SEQ PLATFORMS

3' (and 5') End

Modern Examples: CEL-seq2, MARS-seq, Drop-seq, InDrop, 10X Chromium, SPLiT-seq, Quartz-Seq2, sci-RNA-seq

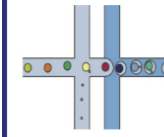
Features: Higher-throughput, higher scalability, lower cost per cell, require lower sequencing depth

Important Controls: Unique Molecular Identifiers (UMIs)



Reagent Cost &
Cost/Cell
Speed
Scalability
Sensitivity

IV. Single Cell RNAseq Platform



Method:

- Droplet-based
- Tube-based after FACS
- Microwell-based
- Microfluidics-enabled

Key considerations:

- Cell throughput and handling time
- Gene coverage and cell type detection
- Whole transcript versus 3'end counting
- Imaging capability for doublet detection

Full-Length Transcript

Modern Examples: Smart-seq2, SUPeR-seq, MATQ-seq

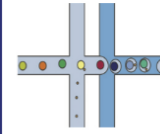
Features: Detect more expressed genes, better coverage of low-abundance genes, can incorporate nascent transcripts, allows for more downstream analysis

Important Controls: Spike-In RNA

Ding et al, *Nature Biotechnology*, 2020

SINGLE-CELL RNA-SEQ PLATFORMS

IV. Single Cell RNAseq Platform



Method:

- Droplet-based
- Tube-based after FACS
- Microwell-based
- Microfluidics-enabled

Key considerations:

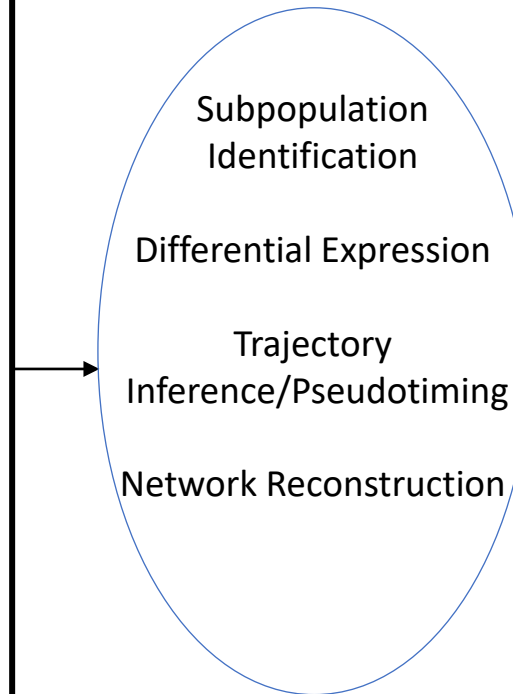
- Cell throughput and handling time
- Gene coverage and cell type detection
- Whole transcript versus 3' end counting
- Imaging capability for doublet detection

3' (and 5') End

Modern Examples: CEL-seq2, MARS-seq, Drop-seq, InDrop, 10X Chromium, SPLiT-seq, Quartz-Seq2, sci-RNA-seq

Features: Higher-throughput, higher scalability, lower cost per cell, require lower sequencing depth

Important Controls: Unique Molecular Identifiers (UMIs)



Full-Length Transcript

Modern Examples: Smart-seq2, SUPeR-seq, MATQ-seq

Features: Detect more expressed genes, better coverage of low-abundance genes, can incorporate nascent transcripts, allows for more downstream analysis

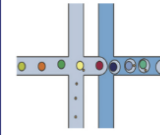
Important Controls: Spike-In RNA



DROPLET-BASED PLATFORMS

- Reaction Specificity
- Poisson Inclusion
- Ambient RNA correction

IV. Single Cell RNAseq Platform

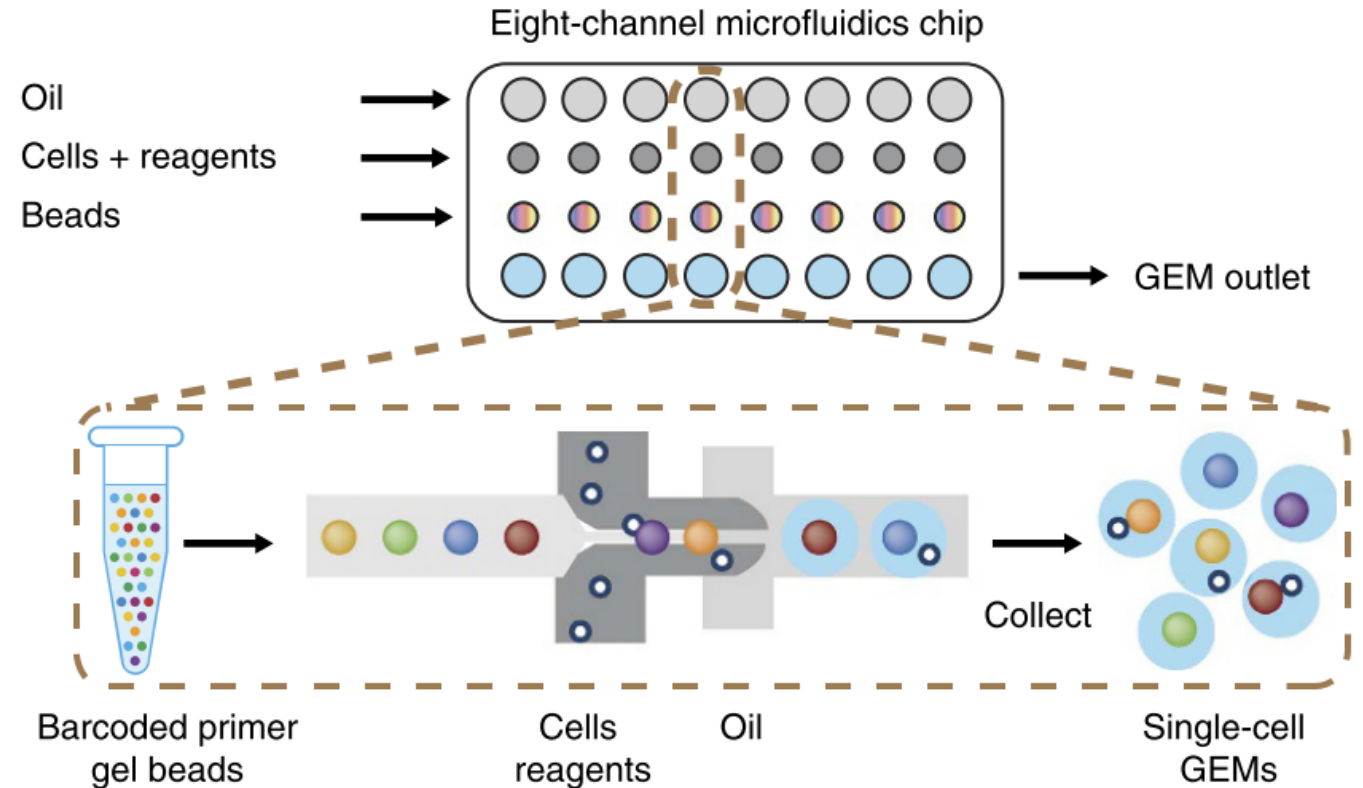


Method:

- Droplet-based
- Tube-based after FACS
- Microwell-based
- Microfluidics-enabled

Key considerations:

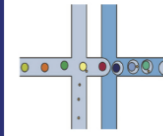
- Cell throughput and handling time
- Gene coverage and cell type detection
- Whole transcript versus 3' end counting
- Imaging capability for doublet detection



Zheng et al, *Nature Communications*, 2017

COMBINATORIAL INDEXING PLATFORMS

IV. Single Cell RNAseq Platform



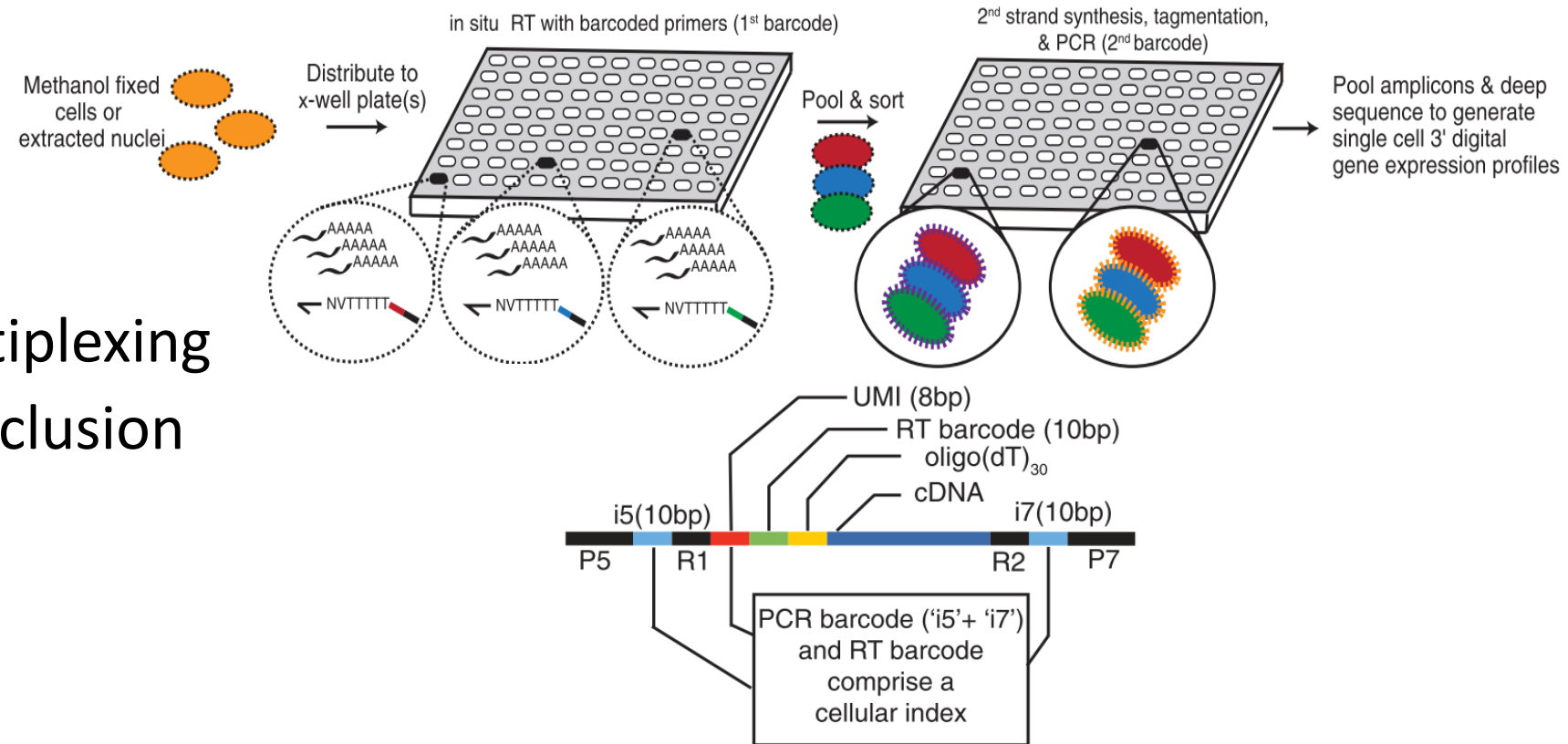
Method:

- Droplet-based
- Tube-based after FACS
- Microwell-based
- Microfluidics-enabled

Key considerations:

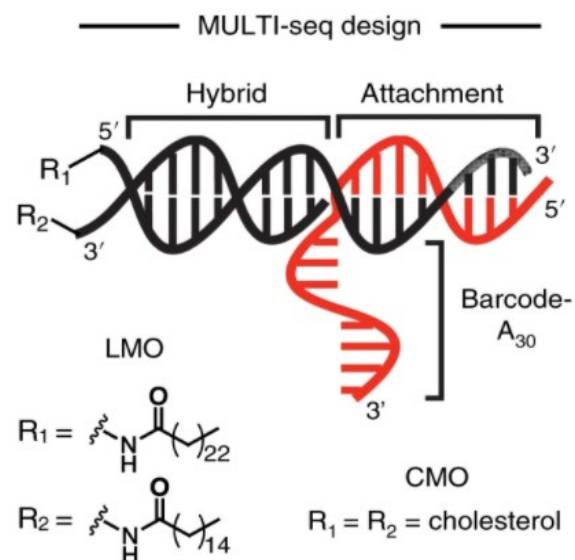
- Cell throughput and handling time
- Gene coverage and cell type detection
- Whole transcript versus 3' end counting
- Imaging capability for doublet detection

- Highly Scalable
- Simplified Internal Multiplexing
- Maximizes Condition Inclusion
- Batch Combination

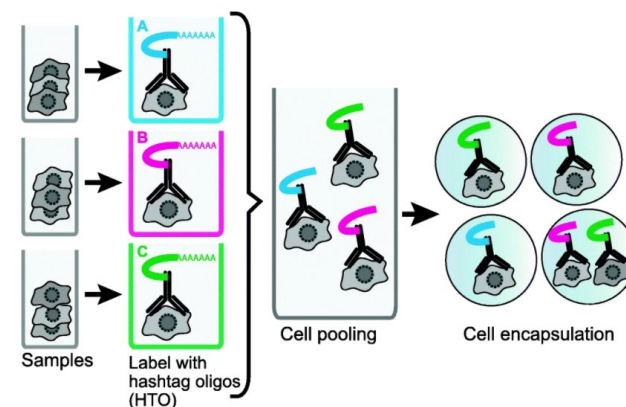


Cao et al, *Science*, 2017

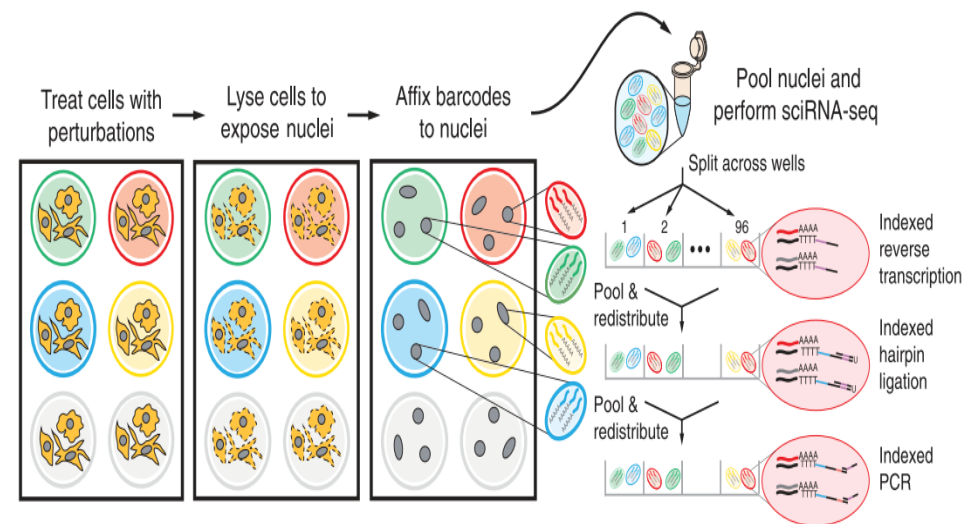
PLATFORM CONSIDERATIONS: STRATEGIC MULTIPLEXING ADVANCES



Multi-Seq
McGinnis et al, *Nature Methods*, 2019



CITE-Seq/Cell Hashing
Stoeckius et al, *Genome Biology*, 2019



Sci-Plex
Srivatsan et al, *Science*, 2020

SEQUENCING DEPTH CONSIDERATIONS

General Benchmarks:

Full-length Transcript: $>1 \times 10^6$ reads per cell
Min: 5×10^4 reads per cell

3' End Sequencing: $>2 \times 10^4$ reads per cell
Min: 1×10^4 reads per cell

Alternative Splicing: $>15 \times 10^6$ reads per cell

V. Library Sequencing

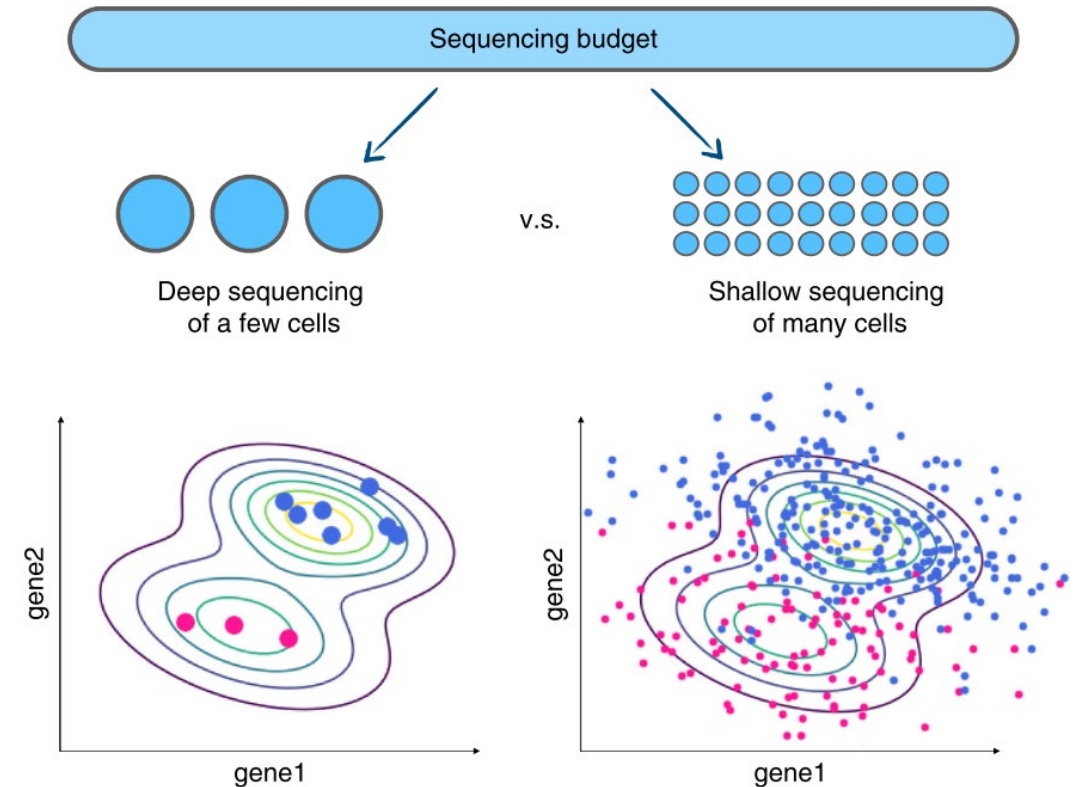


Method:

- Illumina NGS
- Compatible with cDNA library

Sequencing depth considerations:

- 3' end counting: low depth $\sim 50K$ RPC
- Whole transcript: high depth $\sim 1M$ RPC
- Alternative splicing: $\sim 20-30M$ RPC
- Iterative optimization for biological system



Zhang et al, *Nature Communications*, 2020

VI. Computational Analysis



Key considerations:

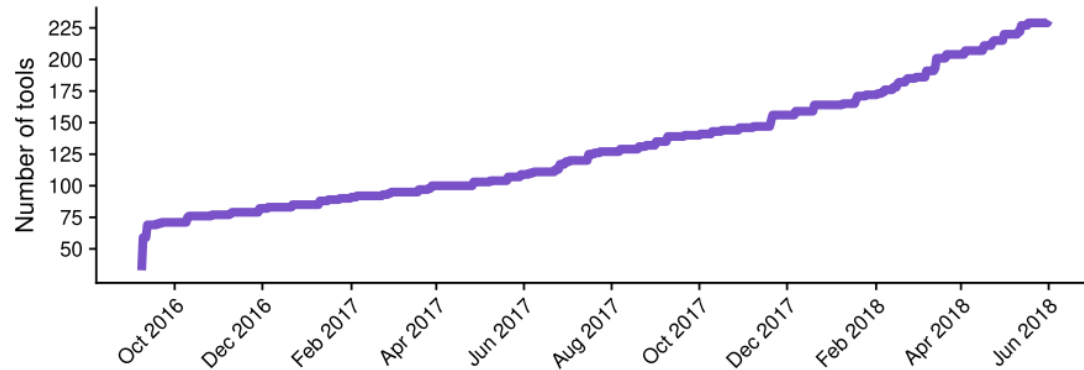
- Separation of *batch* and *condition*
- Technical vs. biological variation

Sample Batch correction approaches:

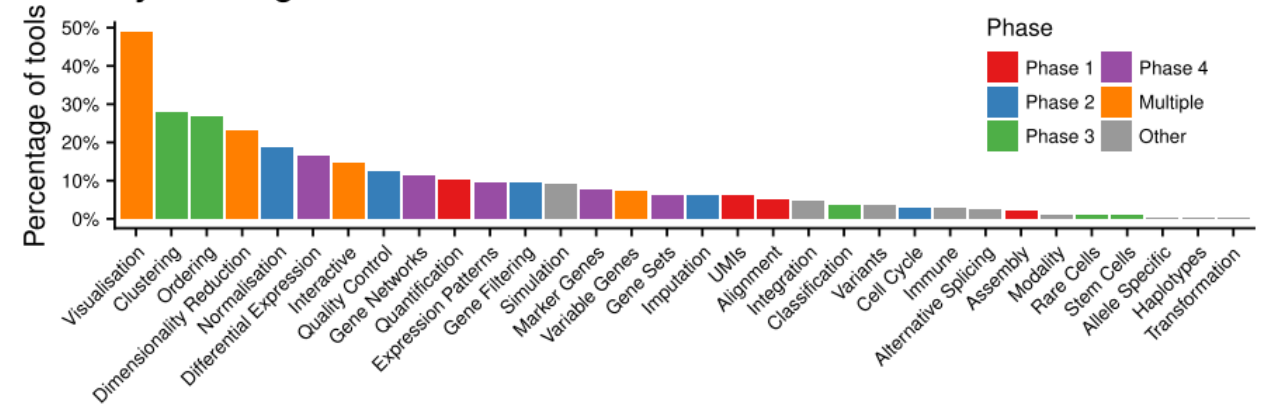
- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST

SC-RNASEQ DATA ANALYSIS: TOOLS

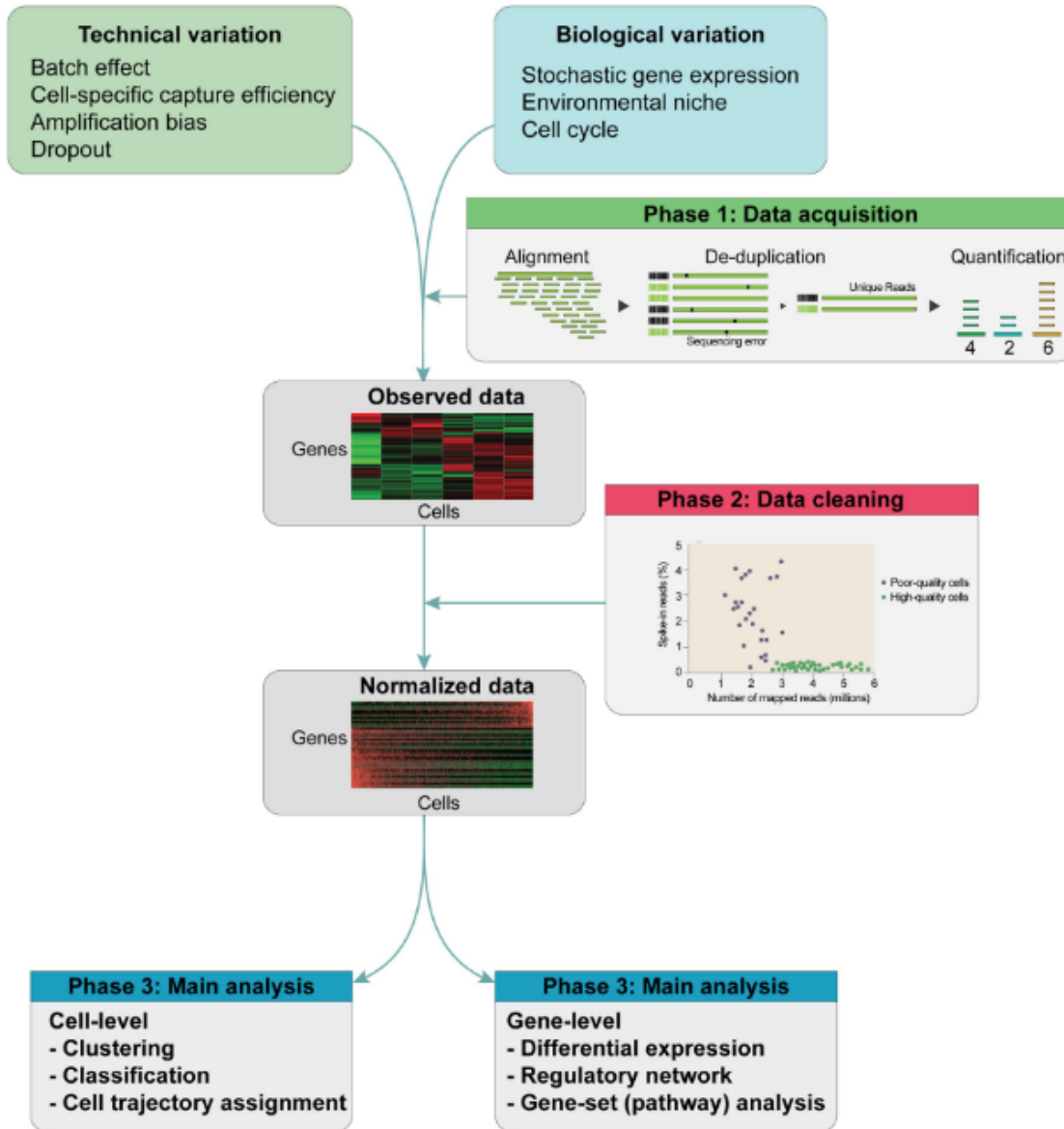
Increase in tools over time



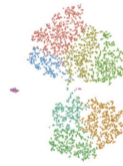
Analysis categories



Zappia et al, *PLOS Computational Biology*, 2018



VI. Computational Analysis



Key considerations:

- Separation of *batch* and *condition*
- Technical vs. biological variation

Sample Batch correction approaches:

- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST

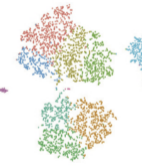
SC-RNASEQ DATA ANALYSIS: PRACTICE

PRE- PROCESSING: QUALITY CONTROL AND COVARIATES

- Raw data quality and Demultiplexing
- Counts per Barcode (Cell)
- Genes per Barcode (Cell)
- Fraction of Mitochondrial Reads
- Doublet Detection
- Multi-Variate Assessment
 - Permissive vs Conservative
 - Revisiting and Sample Differences

Luecken & Theis, *Molecular Systems Biology*, 2019

VI. Computational Analysis

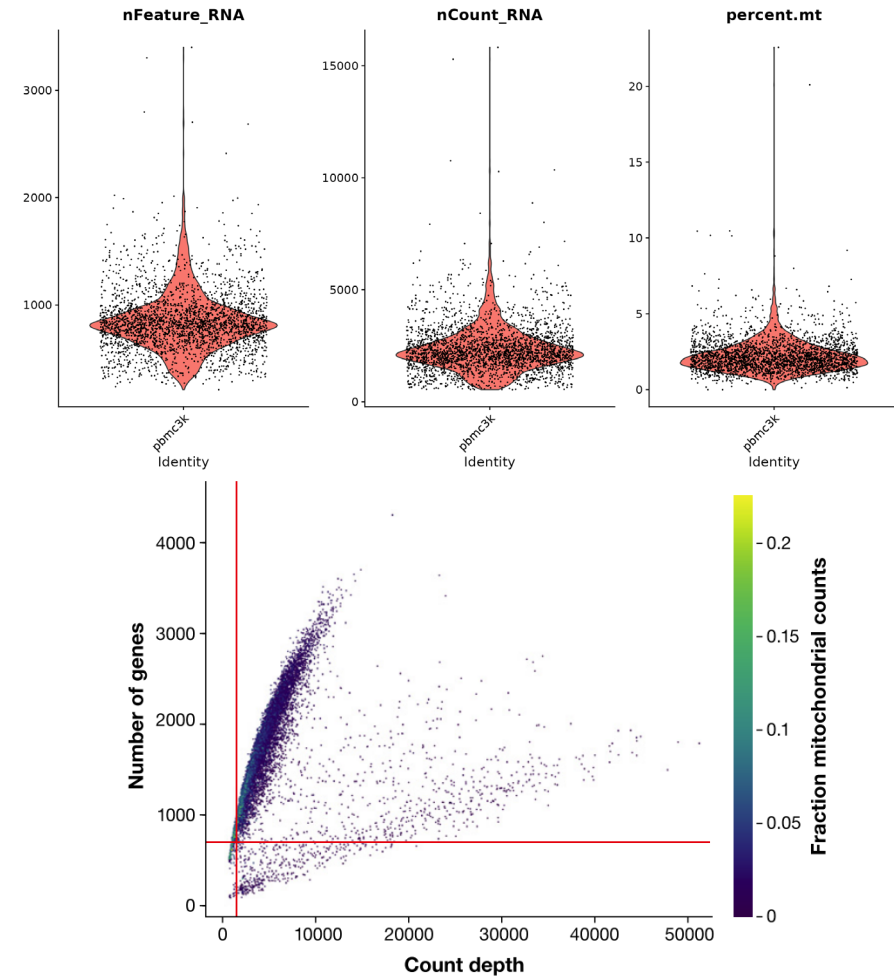


Key considerations:

- Separation of *batch* and *condition*
- Technical vs. biological variation

Sample Batch correction approaches:

- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST



PRE-PROCESSING: NORMALIZATION, SCALING AND FEATURE DETECTION

- Critical to obtaining relative gene expression between cells
- Account for variability in count depth and gene dropout
- Within Sample Normalization
 - CPM/C Normalization and *scran*
 - TPM – full length vs non
 - Non-Linear Normalization
 - Log transformation
- Scaling Mean and Variance
- Effect Corrections
 - Biological – e.g. Cell Cycle
 - Technical – e.g. Batch, Imputation

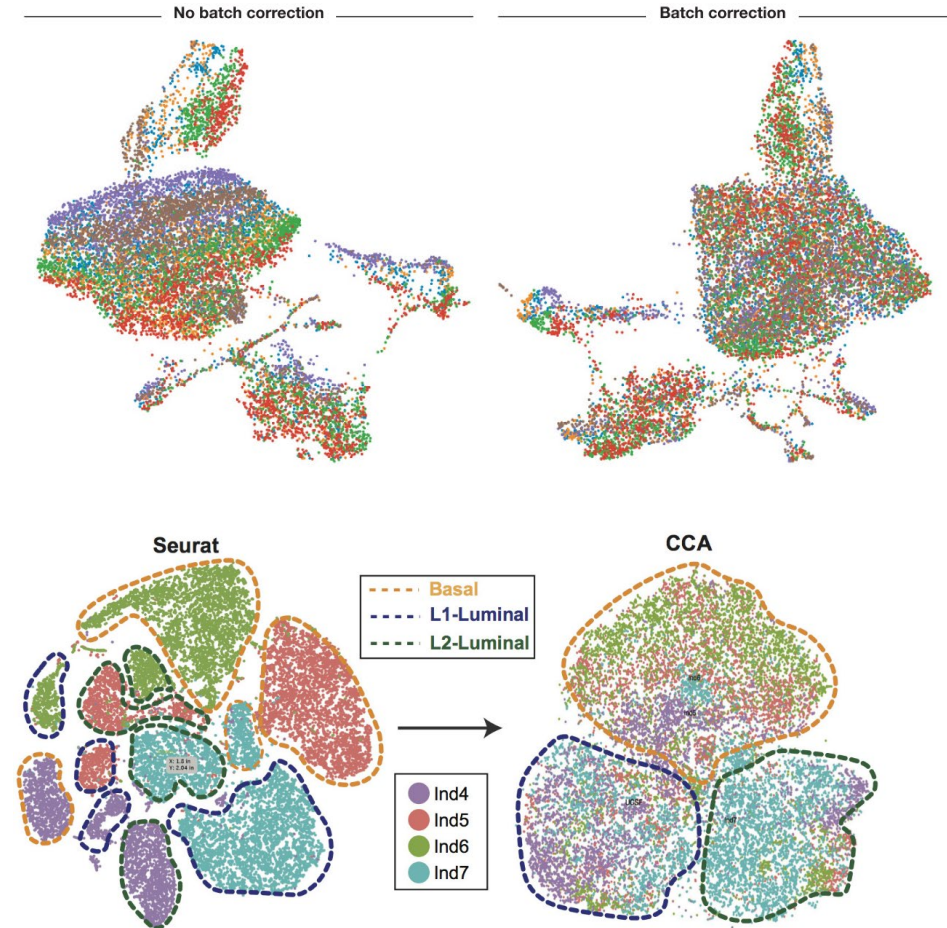
VI. Computational Analysis

Key considerations:

- Separation of *batch* and *condition*
- Technical vs. biological variation

Sample Batch correction approaches:

- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST

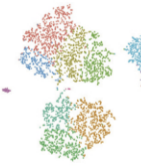


Luecken & Theis, *Molecular Systems Biology*, 2019
Nguyen et al, *Frontiers in Cell and Developmental Biology*, 2018

DIMENSIONALITY REDUCTION AND VISUALIZATION

- Unsupervised Feature Detection
 - Highest Variable Genes (Seurat FVG)
 - Spike-In Based (BASiCS)
 - Dropout Based (M3Drop)
- Dimensionality Reduction and Visualization
 - Principal Component Analysis (PCA)
 - Latent Semantic and Jaccard
 - T-Distributed Stochastic Neighbor Embedding (t-SNE)
 - Uniform Approximation and Projection (UMAP)
 - Visualization vs Summarization

VI. Computational Analysis

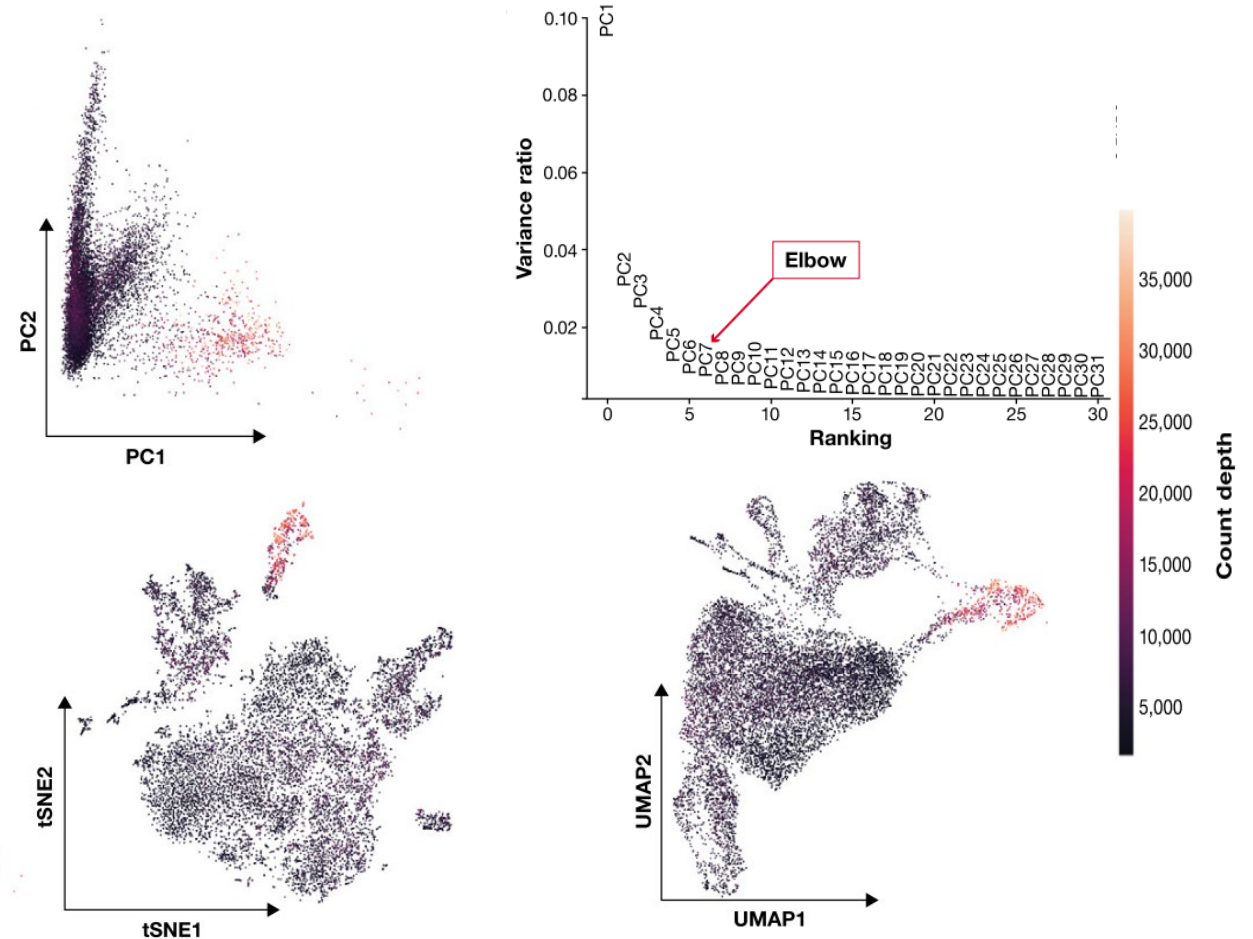


Key considerations:

- Separation of *batch* and *condition*
- Technical vs. biological variation

Sample Batch correction approaches:

- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST

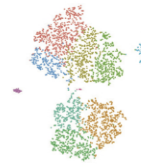


Luecken & Theis, *Molecular Systems Biology*, 2019

CELL CLUSTERING

- Clustering and Community Detection
 - K-means
 - Hierarchical
 - K-Nearest Neighbor
 - Sub-clustering
- “Marker” Genes and “Cell Types”
 - Secondary methods
- Atlases and Databases
 - Human Cell Atlas

VI. Computational Analysis

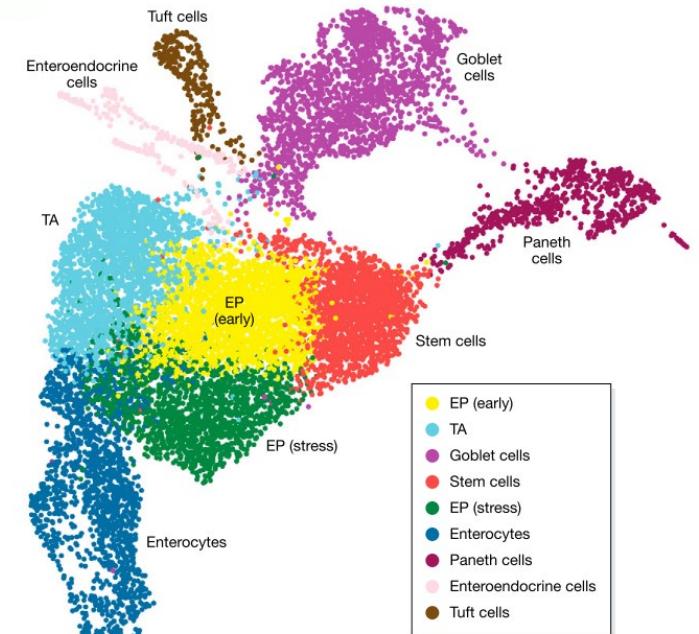
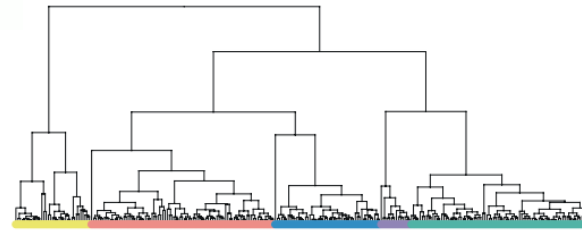
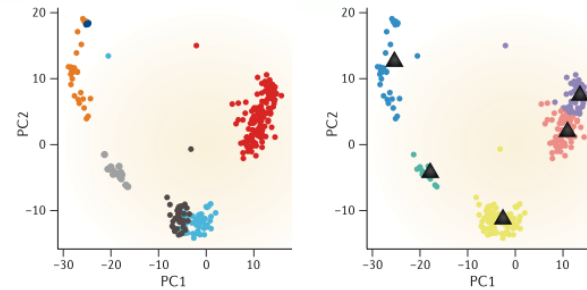


Key considerations:

- Separation of *batch* and *condition*
- Technical vs. biological variation

Sample Batch correction approaches:

- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST



Kiselev, Andrews, & Hemberg *Nature Reviews Genetics*, 2019

DIFFERENTIAL EXPRESSION AND TRAJECTORY INFERENCE

- Differential Expression

- Bulk Tools vs SC Tools
- MAST

- Trajectory Inference/Pseudotiming

- Differentiation and Progressive Cell Cha
- Linear Minimum variability
- Topological Assumptions
- Multiple Confirmation and Further Perturbation

Luecken & Theis, *Molecular Systems Biology*, 2019
Saelens, *Nature Biotechnology*, 2019

VI. Computational Analysis

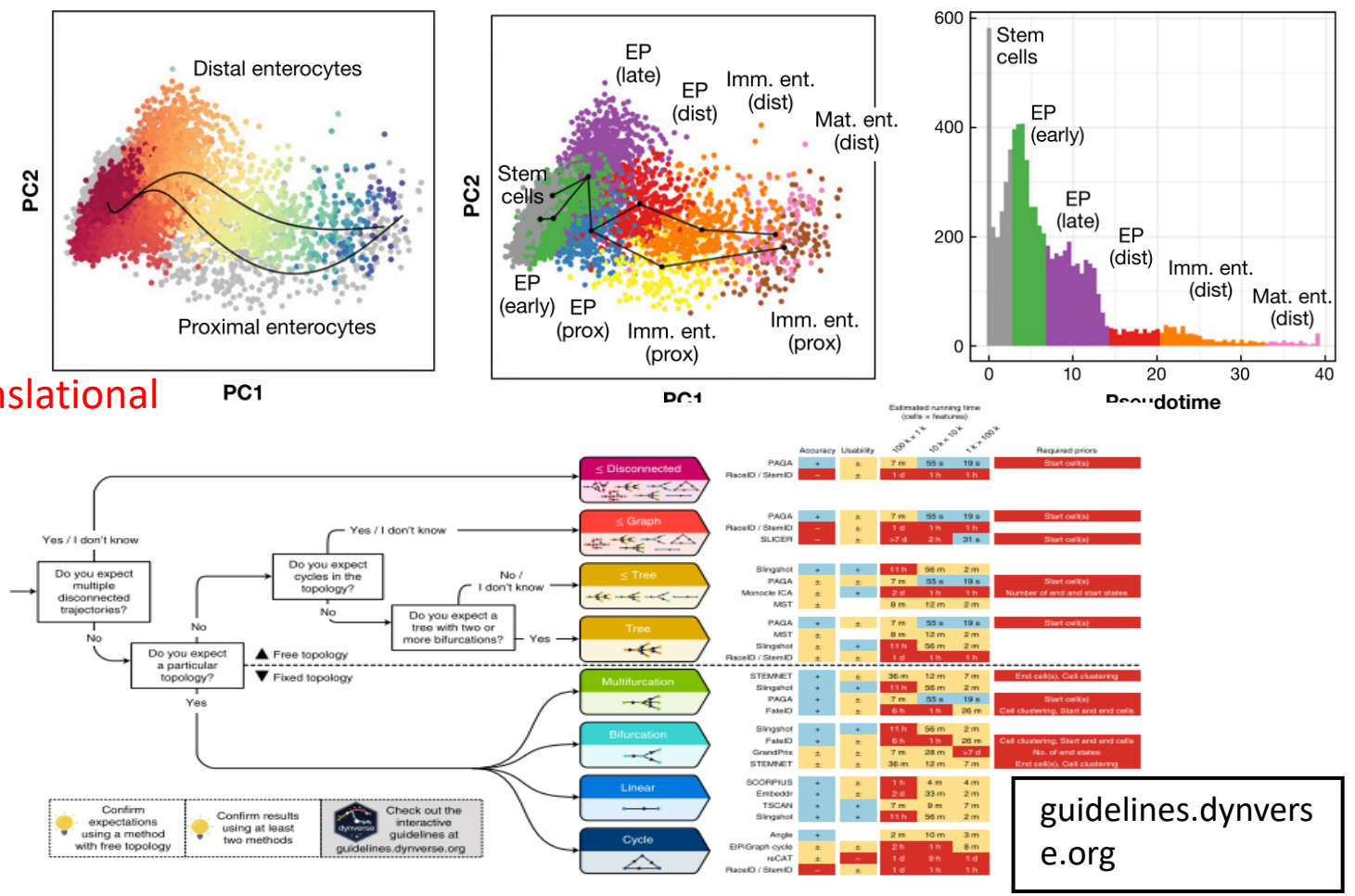


Key considerations:

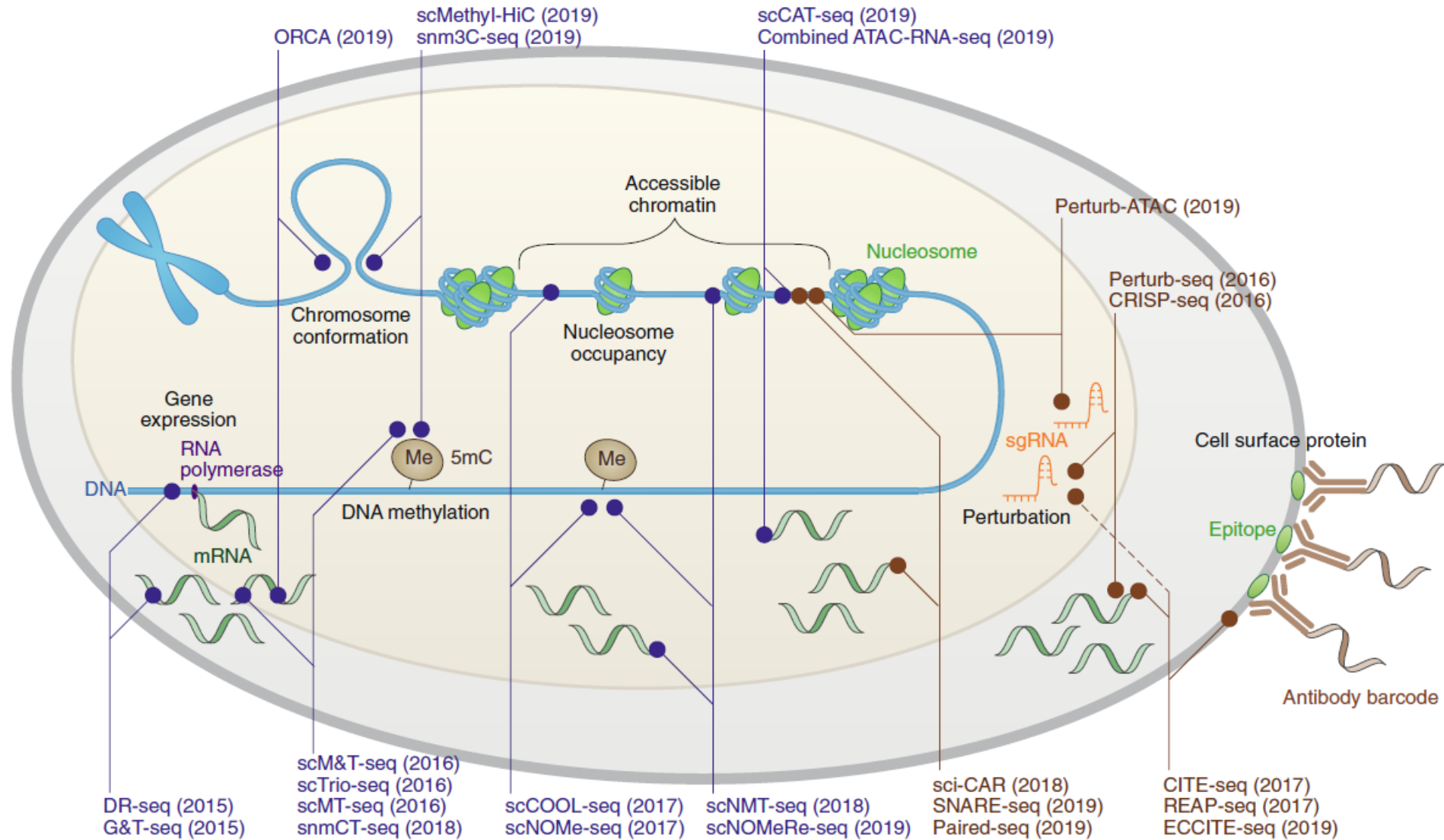
- Separation of *batch* and *condition*
- Technical vs. biological variation

Sample Batch correction approaches:

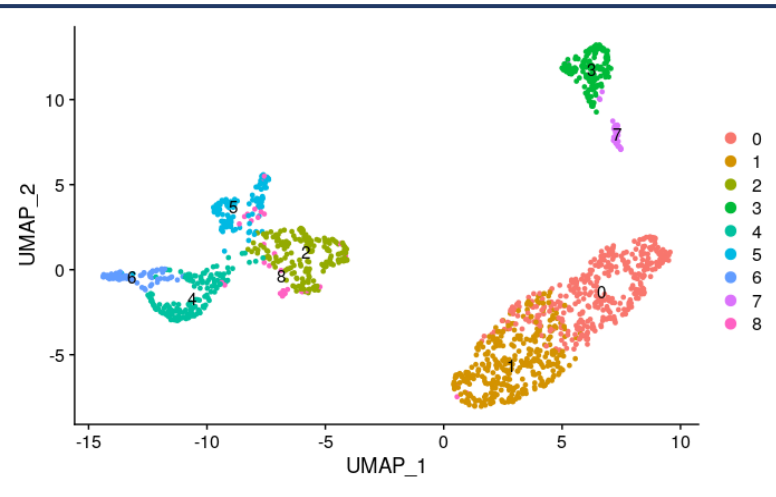
- Cell Hashing
- Demuxlet
- Canonical correlation analysis (CCA)
- MAST



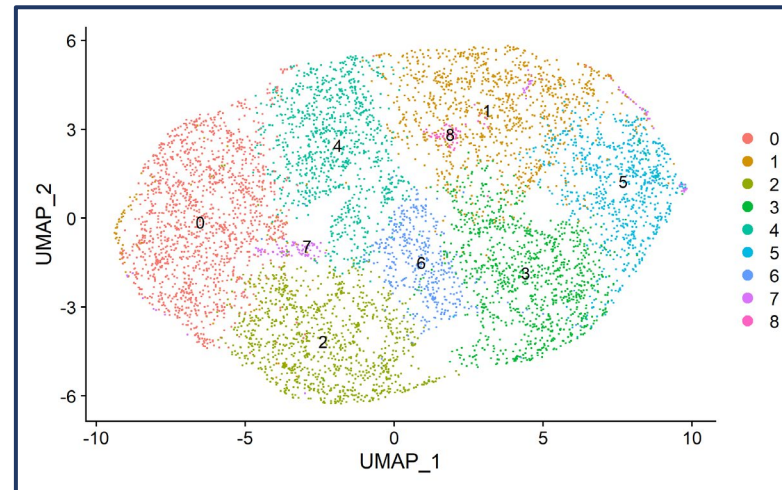
MULTI-MODAL DATA ADVANCEMENTS



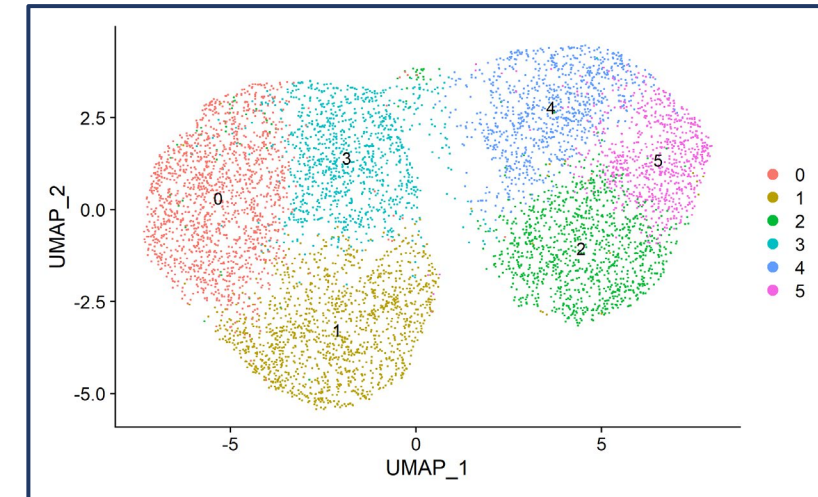
SINGLE-CELL TOXICOLOGICAL PROFILING AT EPA



Neurons
**Developmental
Neurotoxicity**
Dr. Tim Shafer
Dr. Brian Chorley



HepaRG Cell Lines
**Hepatic Toxicity Gene
Networks**
Dr. Imran Shah
Dr. Brian Chorley



*Human Embryonic Stem
Cells*
**Predictive Toxicology
Using Stem Cell Models**
Dr. Sid Hunter

Applications of Single-Cell Profiling Methods to Enhance Mechanistic Understanding of Toxicological Responses

Dr. Kelly Bakulski,
University of Michigan

Single-Cell Analysis of the Gene Expression Effects of Developmental Lead (Pb) Exposure on the Mouse Hippocampus

Dr. Britton Goodale,
Dartmouth College

Single-cell RNA-seq Analysis Reveals That Prenatal Arsenic Exposure Results in Long-term, Adverse Effects on Immune Gene Expression and in Response to Influenza A Infection

Dr. Rance Nault,
Michigan State University

Applications of Single-Cell Transcriptomics in Dose-Response Assessments of the Effects of Dioxin

Dr. Joseph Wu,
Stanford University

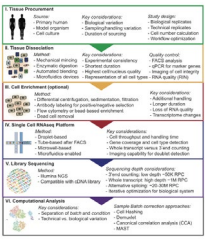
Single-Cell RNA Sequencing of Human Embryonic Stem Cell Differentiation Delineates Adverse Effects of Toxicants on Embryonic Development

ACKNOWLEDGMENTS

- US EPA – Chorley Lab
 - Dr. Brian Chorley
 - Gleta Carswell
 - Gail Nelson
 - Nyssa Tucker
- US EPA – Hunter Lab
 - Dr. Sid Hunter
 - Susan Jeffay
 - Sarah Park
 - Jessica Conley
- Duke University – West Lab
 - Dr. Anne West
 - Melyssa Minto
 - Luke Bartelt



QUESTIONS AND RESOURCES

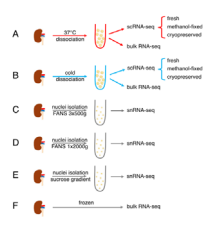


Comprehensive Multi-Stage
Nguyen et al, *Frontiers in Cell and Developmental Biology*, 2018

Chen et al, *Frontiers in Genetics*, 2019

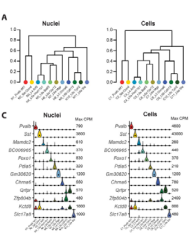
TABLE 1 | Summary of widely used scRNA-seq technologies.

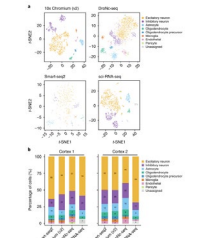
Method	Throughput	Library complexity	Library size	References
Tang method	Low	No	No	Tang et al., 2009
QuantSeq	Low	No	No	Bergmann et al., 2015
10x Genomics	High	No	No	McKenney et al., 2016
Smart-seq	Low	No	No	Brar et al., 2015
Smart-seq2	Low	No	No	Brar et al., 2015
MIST-seq	Low	No	No	Sheng et al., 2017
SPRINT-seq	Low	No	No	Sheng et al., 2017
CEL-seq	Low	No	No	Sheng et al., 2017
CEL-seq2	Low	No	No	Sheng et al., 2017
Drop-seq	Low	No	No	Sheng et al., 2017
Drop-seq2	Low	No	No	Sheng et al., 2017
Drop-seq3	Low	No	No	Sheng et al., 2017
Drop-seq4	Low	No	No	Sheng et al., 2017
Drop-seq5	Low	No	No	Sheng et al., 2017
Drop-seq6	Low	No	No	Sheng et al., 2017
Drop-seq7	Low	No	No	Sheng et al., 2017
Drop-seq8	Low	No	No	Sheng et al., 2017
Drop-seq9	Low	No	No	Sheng et al., 2017
Drop-seq10	Low	No	No	Sheng et al., 2017
Drop-seq11	Low	No	No	Sheng et al., 2017
Drop-seq12	Low	No	No	Sheng et al., 2017
Drop-seq13	Low	No	No	Sheng et al., 2017
Drop-seq14	Low	No	No	Sheng et al., 2017
Drop-seq15	Low	No	No	Sheng et al., 2017
Drop-seq16	Low	No	No	Sheng et al., 2017
Drop-seq17	Low	No	No	Sheng et al., 2017
Drop-seq18	Low	No	No	Sheng et al., 2017
Drop-seq19	Low	No	No	Sheng et al., 2017
Drop-seq20	Low	No	No	Sheng et al., 2017



Cell/Nucleus Preparation
Bakken et al, *PLOS One*, 2020

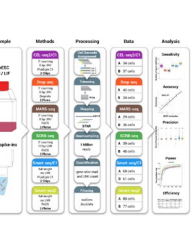
Denisenko, *Genome Biology*, 2019

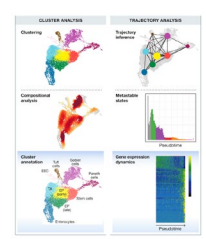




sc-RNAseq Platforms
Ding et al, *Nature Biotechnology*, 2020


Ziegenhain et al, *Molecular Cell*, 2017

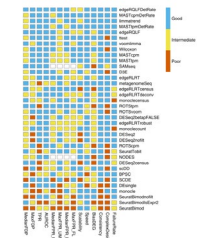




Data Analysis & Clustering
Luecken & Theis, *Molecular Systems Biology*, 2019

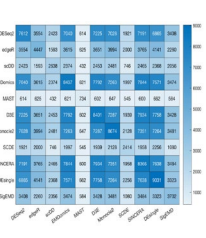
Kiselev, Andrews, & Hemberg *Nature Reviews Genetics*, 2019

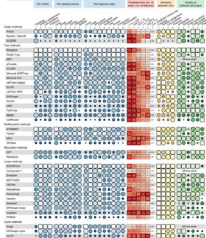




Differential Expression
Soneson & Robinson, *Nature Methods*, 2018

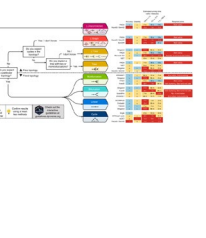
Wang et al, *BMC Bioinformatics*, 2019





Trajectory Inference
Saelens, *Nature Biotechnology*, 2019

Deaconinck et al, *Current Opinions in Systems Biology*, 2021



COMMON PLATFORMS AND ANALYSIS TOOLS

Platforms

Methods	Transcript coverage	UMI possibility	Strand specific	References
Tang method	Nearly full-length	No	No	Tang et al., 2009
Quartz-Seq	Full-length	No	No	Sasagawa et al., 2013
SUPeR-seq	Full-length	No	No	Fan X. et al., 2015
Smart-seq	Full-length	No	No	Ramskold et al., 2012
Smart-seq2	Full-length	No	No	Picelli et al., 2013
MATQ-seq	Full-length	Yes	Yes	Sheng et al., 2017
STRT-seq and STRT/C1	5'-only	Yes	Yes	Islam et al., 2011, 2012
CEL-seq	3'-only	Yes	Yes	Hashimshony et al., 2012
CEL-seq2	3'-only	Yes	Yes	Hashimshony et al., 2016
MARS-seq	3'-only	Yes	Yes	Jaitin et al., 2014
CytoSeq	3'-only	Yes	Yes	Fan H.C. et al., 2015
Drop-seq	3'-only	Yes	Yes	Macosko et al., 2015
InDrop	3'-only	Yes	Yes	Klein et al., 2015
Chromium	3'-only	Yes	Yes	Zheng et al., 2017
SPLIT-seq	3'-only	Yes	Yes	Rosenberg et al., 2018
sci-RNA-seq	3'-only	Yes	Yes	Cao et al., 2017
Seq-Well	3'-only	Yes	Yes	Gierahn et al., 2017
DroNC-seq	3'-only	Yes	Yes	Habib et al., 2017
Quartz-Seq2	3'-only	Yes	Yes	Sasagawa et al., 2018

Alignment Tools

Tools	Category	URL	References
TopHat2	Read mapping	https://ccb.jhu.edu/software/tophat/index.shtml	Kim et al., 2013
STAR	Read mapping	https://github.com/alexdobin/STAR	Dobin and Gingeras, 2015
HISAT2	Read mapping	https://ccb.jhu.edu/software/hisat2/index.shtml	Kim et al., 2015
Cufflinks	Expression quantification	https://github.com/cole-trapnell-lab/cufflinks	Trapnell et al., 2010
RSEM	Expression quantification	https://github.com/deweylab/RSEM	Li and Dewey, 2011
StringTie	Expression quantification	https://github.com/gpertea/stringtie	Pertea et al., 2015

Chen et al, *Frontiers in Genetics*, 2019

COMMON PLATFORMS AND ANALYSIS TOOLS

Cluster Identification

Methods	URL	References
SC3	http://bioconductor.org/packages/SC3	Kiselev et al., 2017
ZIFA	https://github.com/epierson9/ZIFA	Pierson and Yau, 2015
Destiny	https://github.com/theislab/destiny	Angerer et al., 2016
SNN-Cliq	http://bioinfo.uncc.edu/SNNCliq/	Xu and Su, 2015
RaceID	https://github.com/dgrun/RaceID	Grun et al., 2015
SCUBA	https://github.com/gcyuan/SCUBA	Marco et al., 2014
BackSPIN	https://github.com/linnarsson-lab/BackSPIN	Zeisel et al., 2015
PAGODA	http://hms-dbmi.github.io/scde/	Fan et al., 2016
CIDR	https://github.com/VCCRI/CIDR	Lin et al., 2017
pcaReduce	https://github.com/JustinaZ/pcaReduce	Zurauskiene and Yau, 2016
Seurat	https://github.com/satijalab/seurat	Satija et al., 2015
TSCAN	https://github.com/zji90/TSCAN	Ji and Ji, 2016

Differential Expression

Methods	Category	URL	References
ROTS	Single cell	https://bioconductor.org/packages/release/bioc/html/ROTS.html	Seyednasrollah et al., 2016
MAST	Single cell	https://github.com/RGLab/MAST	Finak et al., 2015
BCseq	Single cell	https://bioconductor.org/packages/devel/bioc/html/bcSeq.html	Chen and Zheng, 2018
SCDE	Single cell	http://hms-dbmi.github.io/scde/	Kharchenko et al., 2014
DEsingle	Single cell	https://bioconductor.org/packages/DEsingle	Miao et al., 2018
Census	Single cell	http://cole-trapnell-lab.github.io/monocle-release/	Qiu et al., 2017
D3E	Single cell	https://github.com/hemberg-lab/D3E	Delmans and Hemberg, 2016
BPSC	Single cell	https://github.com/nghiavtr/BPSC	Vu et al., 2016
DESeq2	Bulk	https://bioconductor.org/packages/release/bioc/html/DESeq2.html	Love et al., 2014
edgeR	Bulk	https://bioconductor.org/packages/release/bioc/html/edgeR.html	Robinson et al., 2010
Limma	Bulk	https://bioconductor.org/packages/release/bioc/html/limma.html	Ritchie et al., 2015
Ballgown	Bulk	http://www.bioconductor.org/packages/release/bioc/html/ballgown.html	Frazee et al., 2015

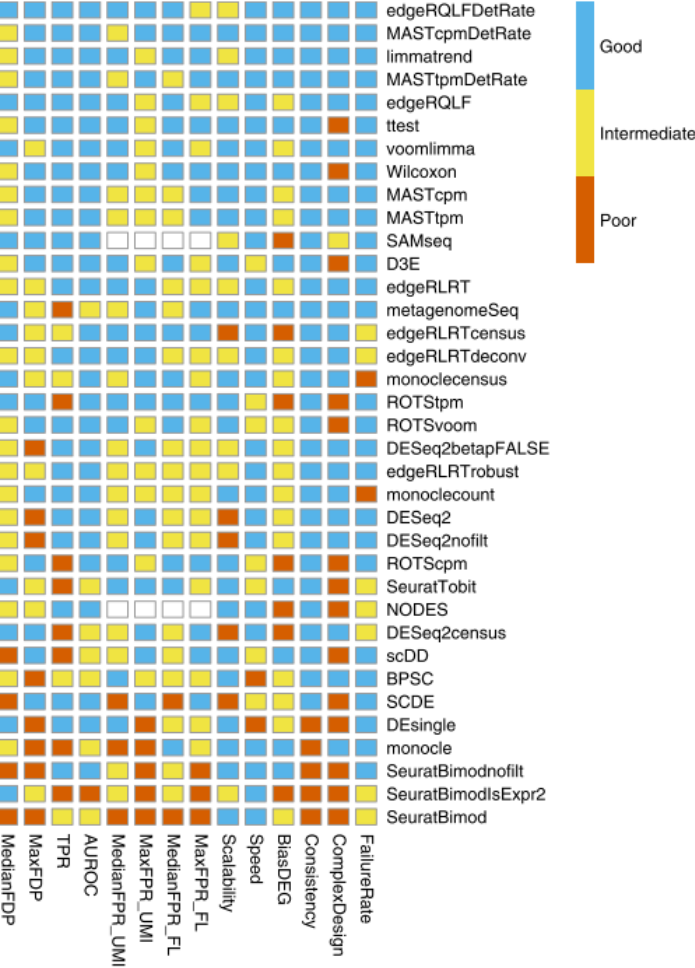
Trajectory Inference

Tools	Dimensionality reduction	URL	References
Monocle	ICA	http://cole-trapnell-lab.github.io/monocle-release/	Trapnell et al., 2014
Waterfall	PCA	https://www.cell.com/cms/10.1016/j.stem.2015.07.013/attachment/3e966901-034f-418a-a439-996c50292a11/mmc9.zip	Shin et al., 2015
Wishbone	Diffusion maps	https://github.com/ManuSetty/wishbone	Setty et al., 2016
GrandPrix	Gaussian Process Latent Variable Model	https://github.com/ManchesterBioinference/GrandPrix	Ahmed et al., 2019
SCUBA	t-SNE	https://github.com/gcyuan/SCUBA	Marco et al., 2014
DPT	Diffusion maps	https://media.nature.com/original/nature-assets/nmeth/journal/v13/n10/extref/nmeth.3971-S3.zip	Haghverdi et al., 2016
TSCAN	PCA	https://github.com/zji90/TSCAN	Ji and Ji, 2016
Monocle2	RGE	http://cole-trapnell-lab.github.io/monocle-release/	Qiu et al., 2017
Slingshot	Any	https://github.com/kstreet13/slingshot	Street et al., 2018
CellRouter	Any	https://github.com/edroaldo/cellrouter	Lummertz da Rocha et al., 2018

EXTRAS

Differential Expression Tools

Soneson & Robinson, *Nature Methods*, 2018



Clustering Tools

Kiselev, Andrews, & Hemberg *Nature Reviews Genetics*, 2019

Name	Year	Method type	Strengths	Limitations
scanpy ⁴	2018	PCA + graph-based	Very scalable	May not be accurate for small data sets
Seurat (latest) ³	2016			
PhenoGraph ³²	2015			
SC3 ³²	2017	PCA + k-means	High accuracy through consensus, provides estimation of k	High complexity, not scalable
SIMLR ²⁴	2017	Data-driven dimensionality reduction + k-means	Concurrent training of the distance metric improves sensitivity in noisy data sets	Adjusting the distance metric to make cells fit the clusters may artificially inflate quality measures
CIDR ²⁵	2017	PCA + hierarchical	Implicitly imputes dropouts when calculating distances	
GiniClust ²⁵	2016	DBSCAN	Sensitive to rare cell types	Not effective for the detection of large clusters
pcaReduce ²⁷	2016	PCA + k-means + hierarchical	Provides hierarchy of solutions	Very stochastic, does not provide a stable result
Tasic et al. ²⁸	2016	PCA + hierarchical	Cross validation used to perform fuzzy clustering	High complexity, no software package available
TSCAN ⁴¹	2016	PCA + Gaussian mixture model	Combines clustering and pseudotime analysis	Assumes clusters follow multivariate normal distribution
mpath ⁴⁵	2016	Hierarchical	Combines clustering and pseudotime analysis	Uses empirically defined thresholds and a priori knowledge
BackSPIN ²⁶	2015	Bicustering (hierarchical)	Multiple rounds of feature selection improve clustering resolution	Tends to over-partition the data
RaceID ²³ , RaceID ²¹⁵ , RaceID ³	2015	k-Means	Detects rare cell types, provides estimation of k	Performs poorly when there are no rare cell types
SINCERA ⁵	2015	Hierarchical	Method is intuitively easy to understand	Simple hierarchical clustering is used, may not be appropriate for very noisy data
SNN-Cliq ⁸⁰	2015	Graph-based	Provides estimation of k	High complexity, not scalable