

EMBL-EBI workshop: Data and tools for transcriptomics and protein biology

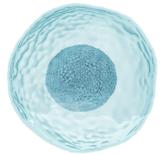
Single Cell Expression Atlas: a new tool for single cell transcriptomics

Laura Huerta, PhD

Senior Scientific Curator

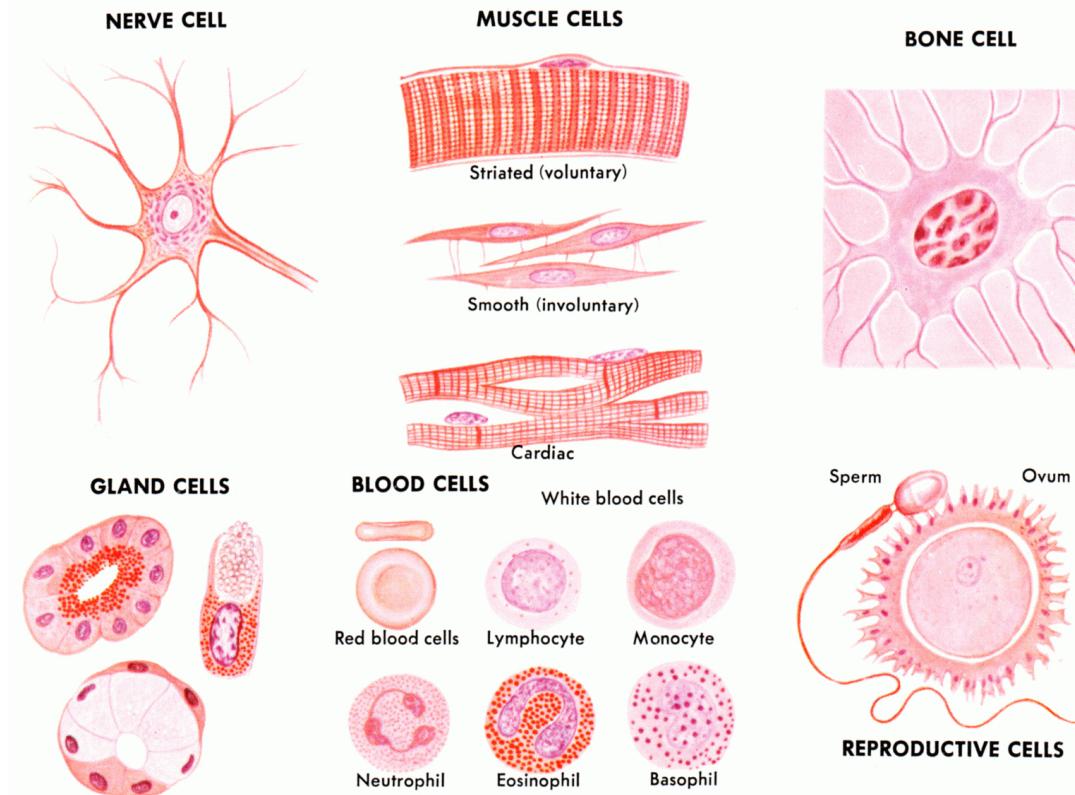
lauhuema@ebi.ac.uk

Malta, 31 July 2018

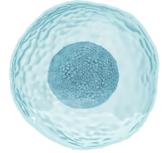


Why single cell gene expression?

The basic unit of life

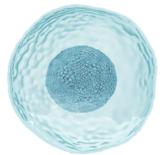


37 trillion

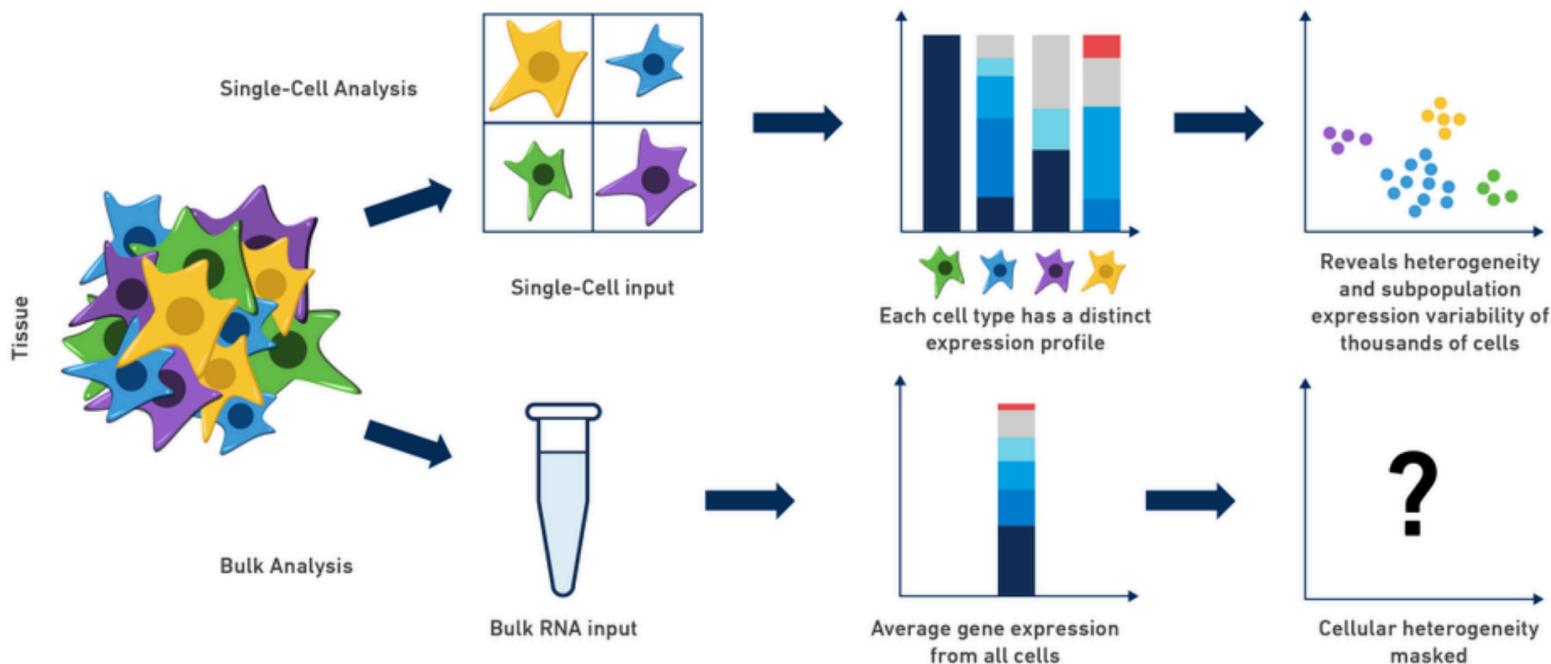


Why single cell gene expression?

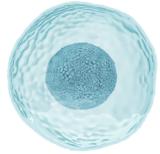




Why single cell gene expression?



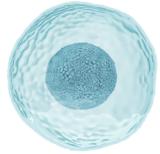
From 10x Genomics. Single cell RNA-seq reveals heterogeneity that is masked by bulk RNA-seq methods



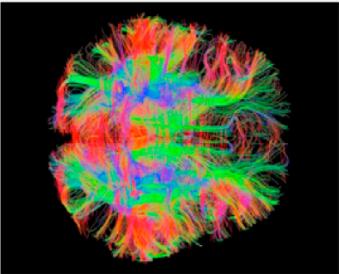
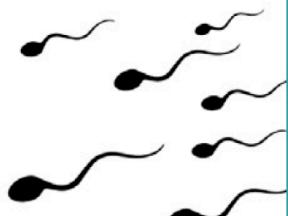
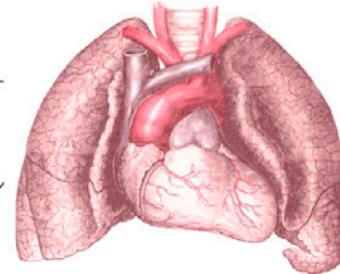
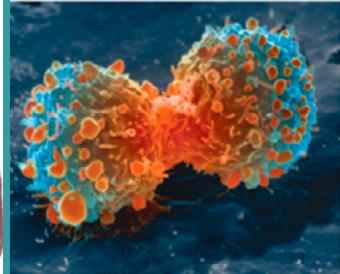
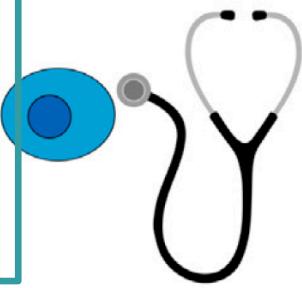
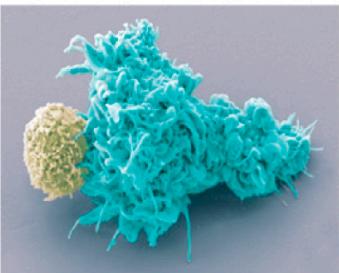
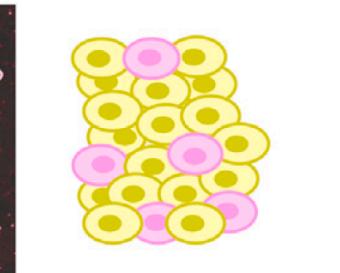
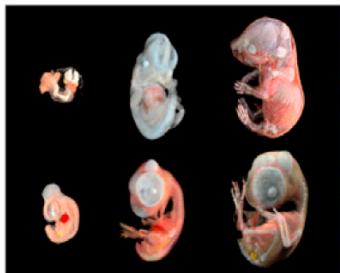
Why single cell gene expression?

Opens the door to several biological and clinical questions

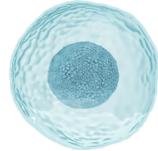
- ✓ Understanding heterogeneous samples:
 - ✓ Developmental biology, neuroscience and immunology: analyse cellular heterogeneity during immune or stem cell development
- ✓ Identification and analysis of rare cell types
 - ✓ Circulating tumor cells from liquid biopsy
- ✓ Understanding cellular transitions and switches in cell state
- ✓ Dissecting complex infections and revealing drug resistance genotypes



Why single cell gene expression?

Neurobiology	Germline Transmission	Organogenesis	Cancer biology	Clinical diagnostics
				
Immunology	Microbiology	Tissue Mosaicism	Embryology	Prenatal-genetic diagnosis
				

From Wang and Navin, 2015. doi:10.1016/j.molcel.2015.05.005



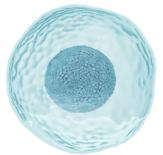
Single cell RNA-seq experiments

DNA-based sequencing methods



Challenge

- ✓ Amplification of minute amounts of mRNA from a cell:
 - ✓ Minimise RNA losses while its conversion into cDNA
 - ✓ Obtain enough DNA for sequencing after amplification (without introducing too much bias)



Direct RNA sequencing

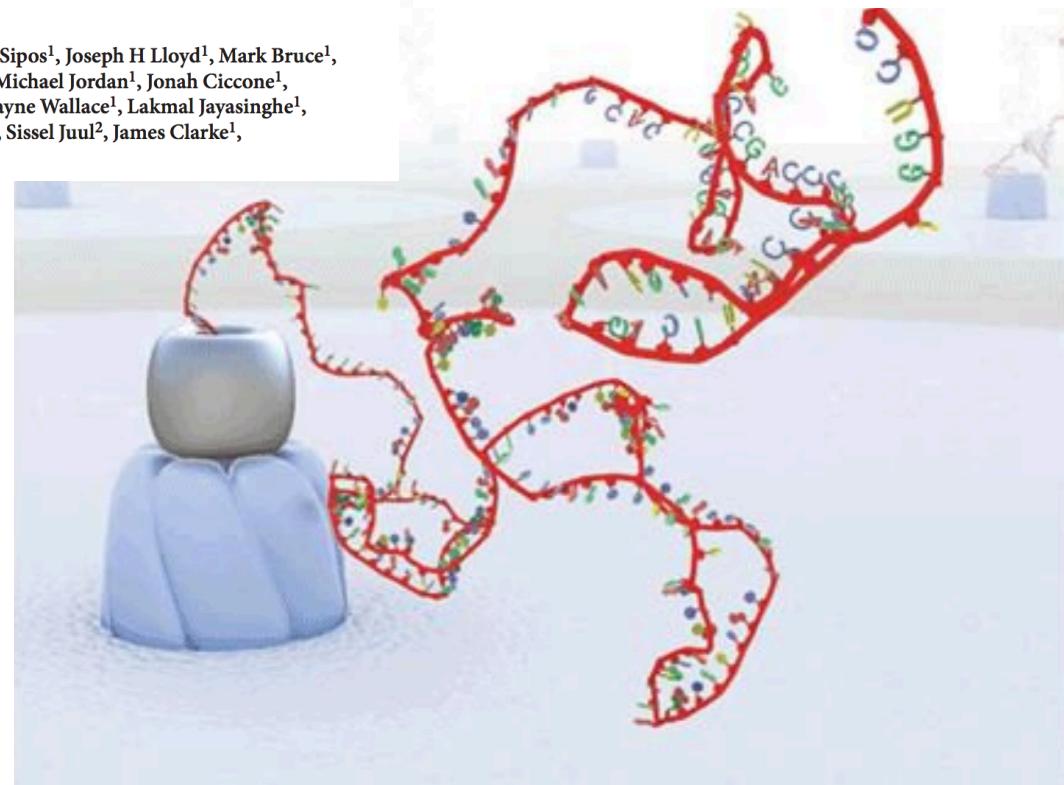
Highly parallel direct RNA sequencing on an array of nanopores

Daniel R Garalde¹, Elizabeth A Snell¹, Daniel Jachimowicz¹, Botond Sipos¹, Joseph H Lloyd¹, Mark Bruce¹, Nadia Pantic¹, Tigist Admassu¹, Phillip James¹, Anthony Warland¹, Michael Jordan¹, Jonah Ciccone¹, Sabrina Serra¹, Jemma Keenan¹, Samuel Martin¹, Luke McNeill¹, E Jayne Wallace¹, Lakmal Jayasinghe¹, Chris Wright¹, Javier Blasco¹, Stephen Young¹, Denise Brocklebank¹, Sissel Juul², James Clarke¹, Andrew J Heron¹ & Daniel J Turner¹ 

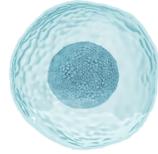
nature|methods

www.nature.com/naturemethods

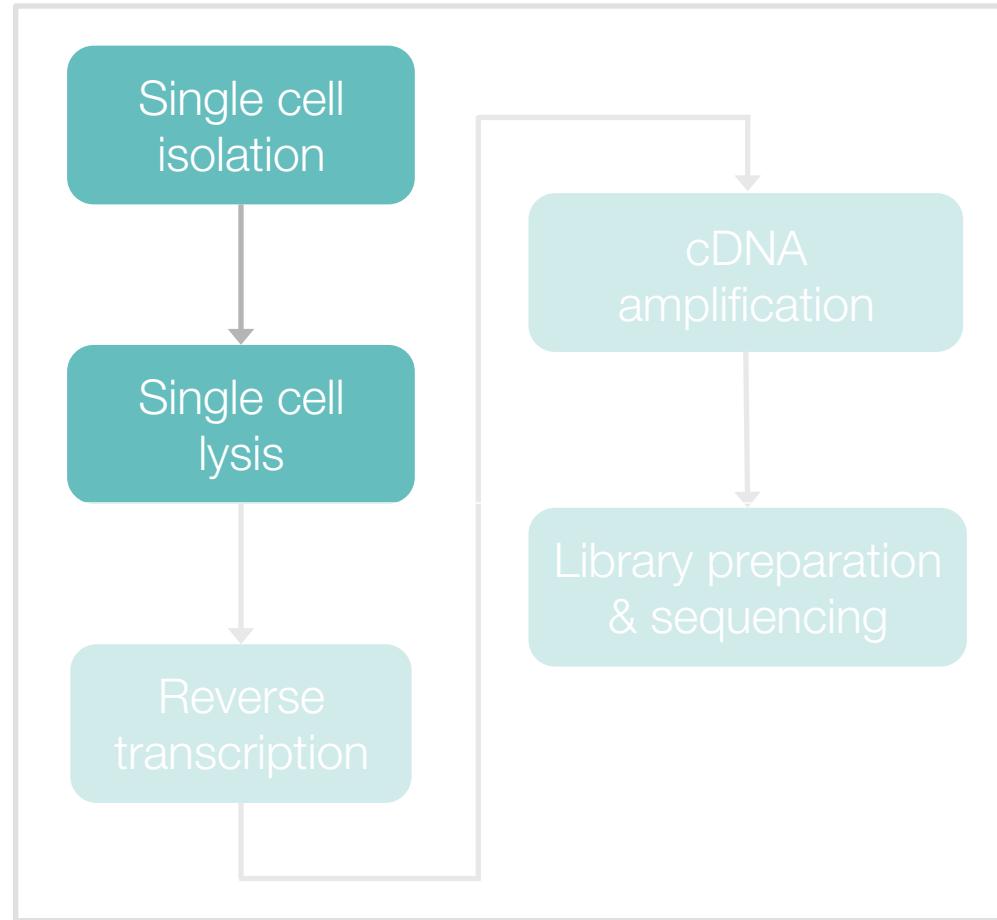
Techniques for life scientists and chemists

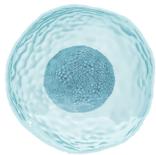


Garalde et al, 2018. doi:10.1038/nmeth.4577



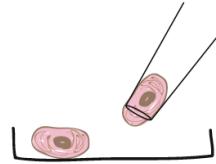
Single cell RNA-seq experiments



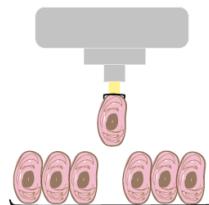


Single cell isolation and lysis

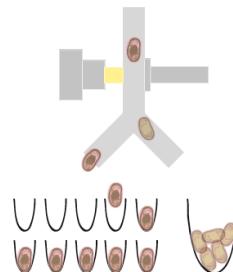
Micropipetting
Micromanipulation



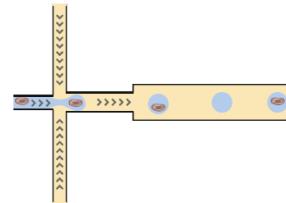
Laser Capture
Microdissection



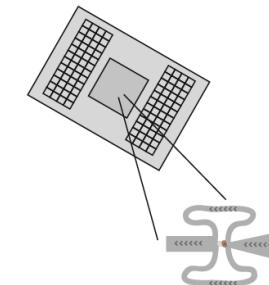
FACS



Microdroplets



Microfluidics
e.g. Fluidigm C1



low number of cells
any tissue

enables selection of cells
based on morphology or
fluorescent markers

visualisation of cells

time consuming

reaction in microliter
volumes

low number of cells
any tissue

enables selection of cells
based on morphology or
fluorescent markers

visualisation of cells

time consuming

reaction in microliter
volumes

hundreds of cells
dissociated cells

enables selection of cells
based on size or
fluorescent markers

fluorescence and light
scattering measurements

fast

reaction in microliter
volumes

large number of cells
dissociated cells

no selection of cells
(can presort with FACS)

fluorescence detection

fast

reaction in nanoliter
volumes

hundreds of cells
dissociated cells

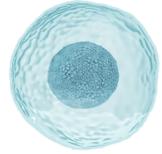
no selection of cells
(can presort with FACS)

visualisation of cells

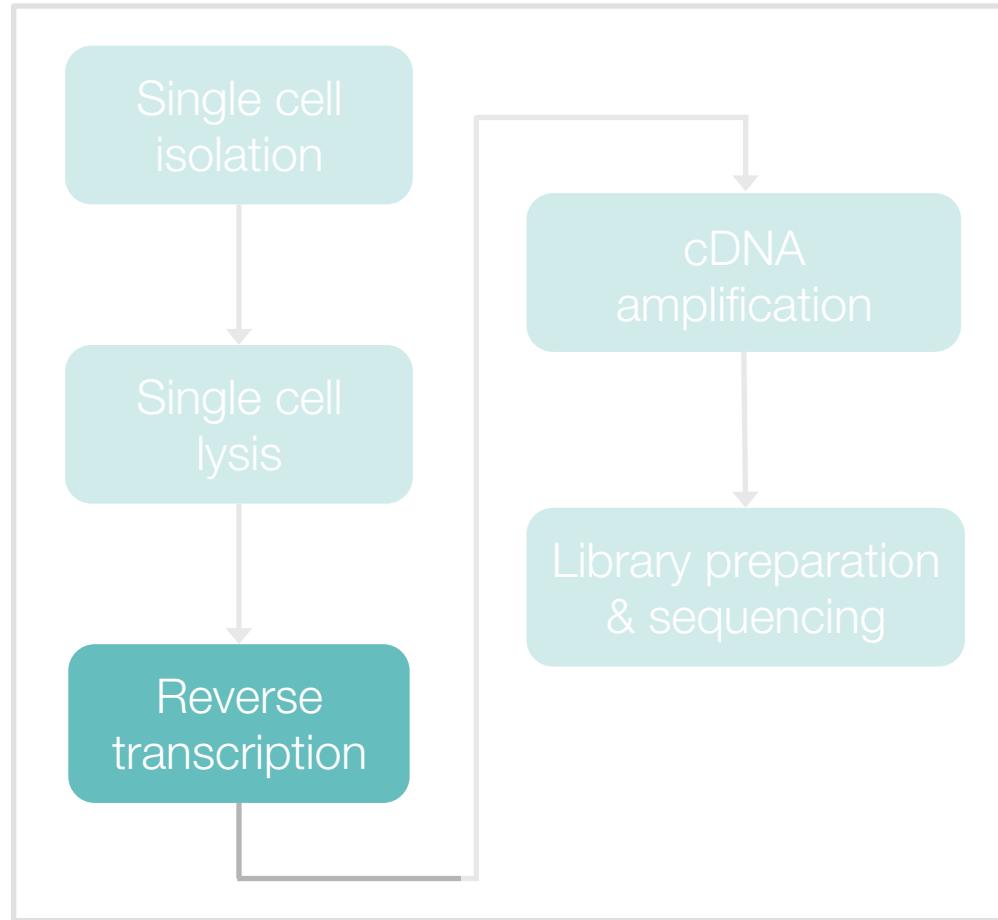
fast

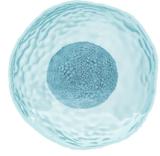
reaction in nanoliter
volumes

From Kolodziejczyk et al, 2015. doi:10.1016/j.molcel.2015.04.005

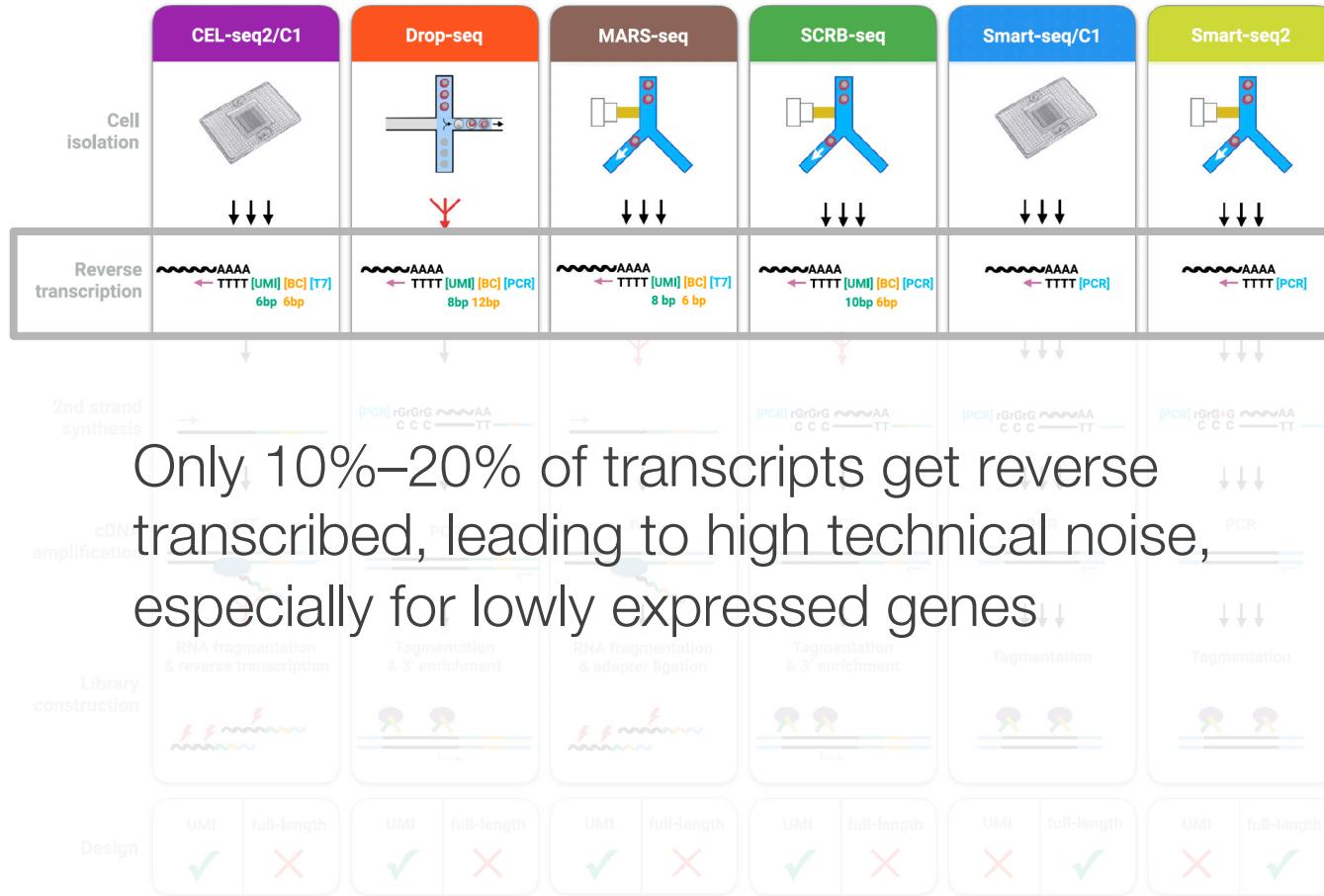


Single cell RNA-seq experiments

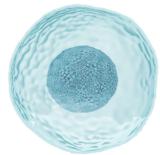




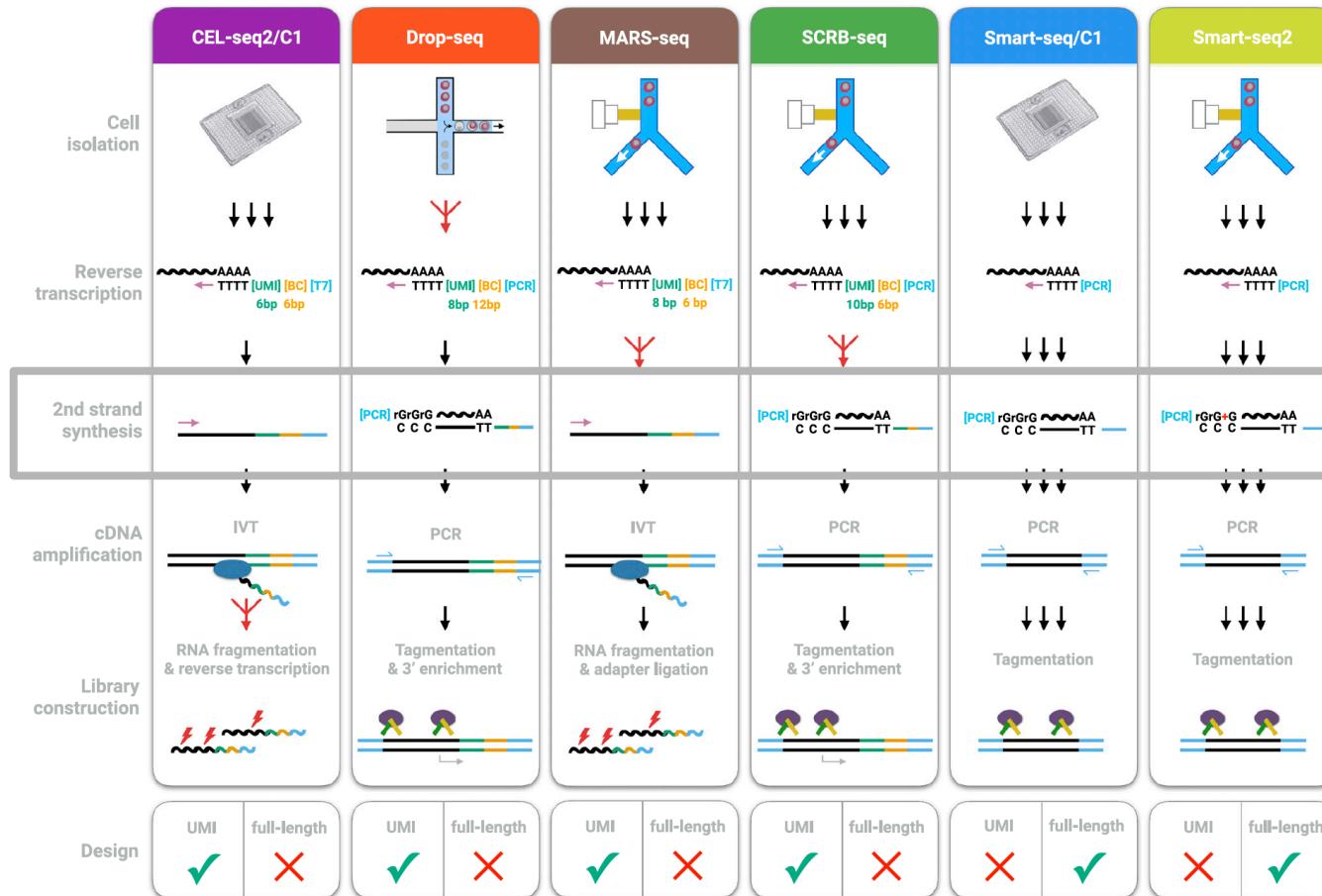
Single cell reverse transcription



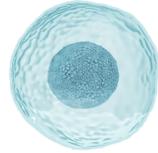
From Ziegenhain et al, 2017. doi:10.1016/j.molcel.2017.01.023



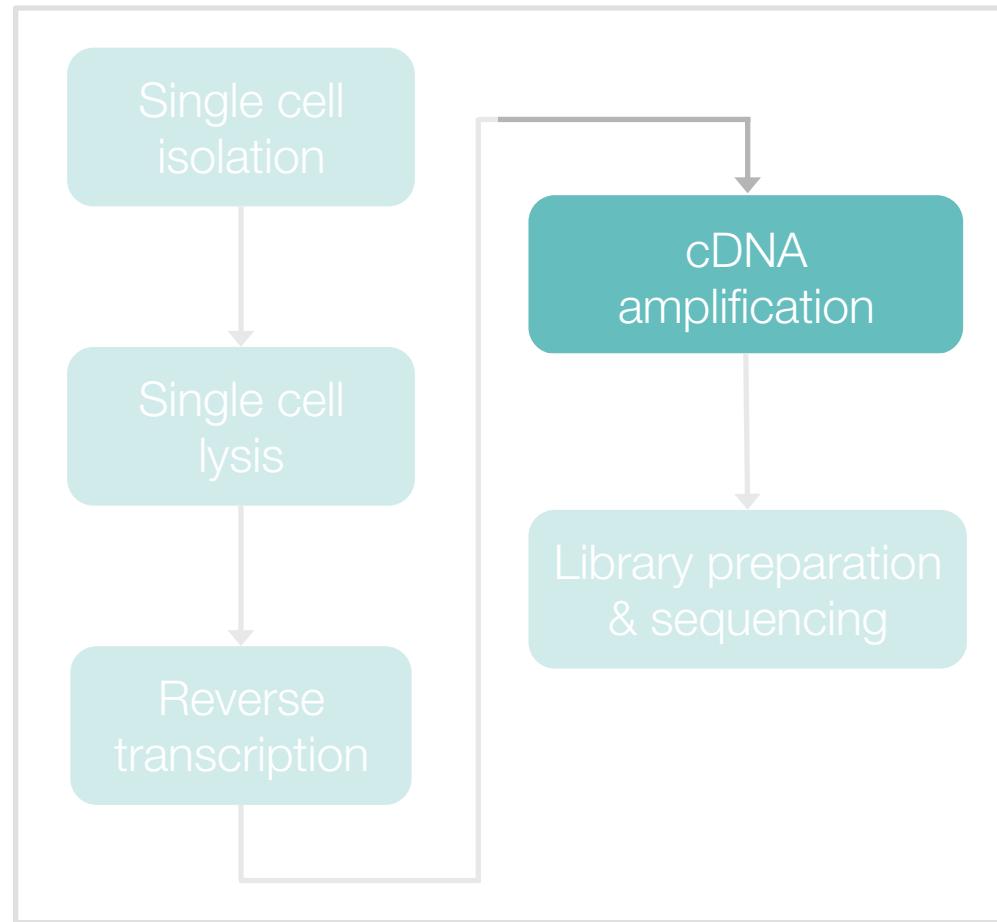
Single cell reverse transcription

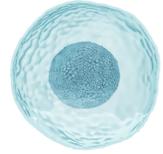


From Ziegenhain et al, 2017. doi:10.1016/j.molcel.2017.01.023

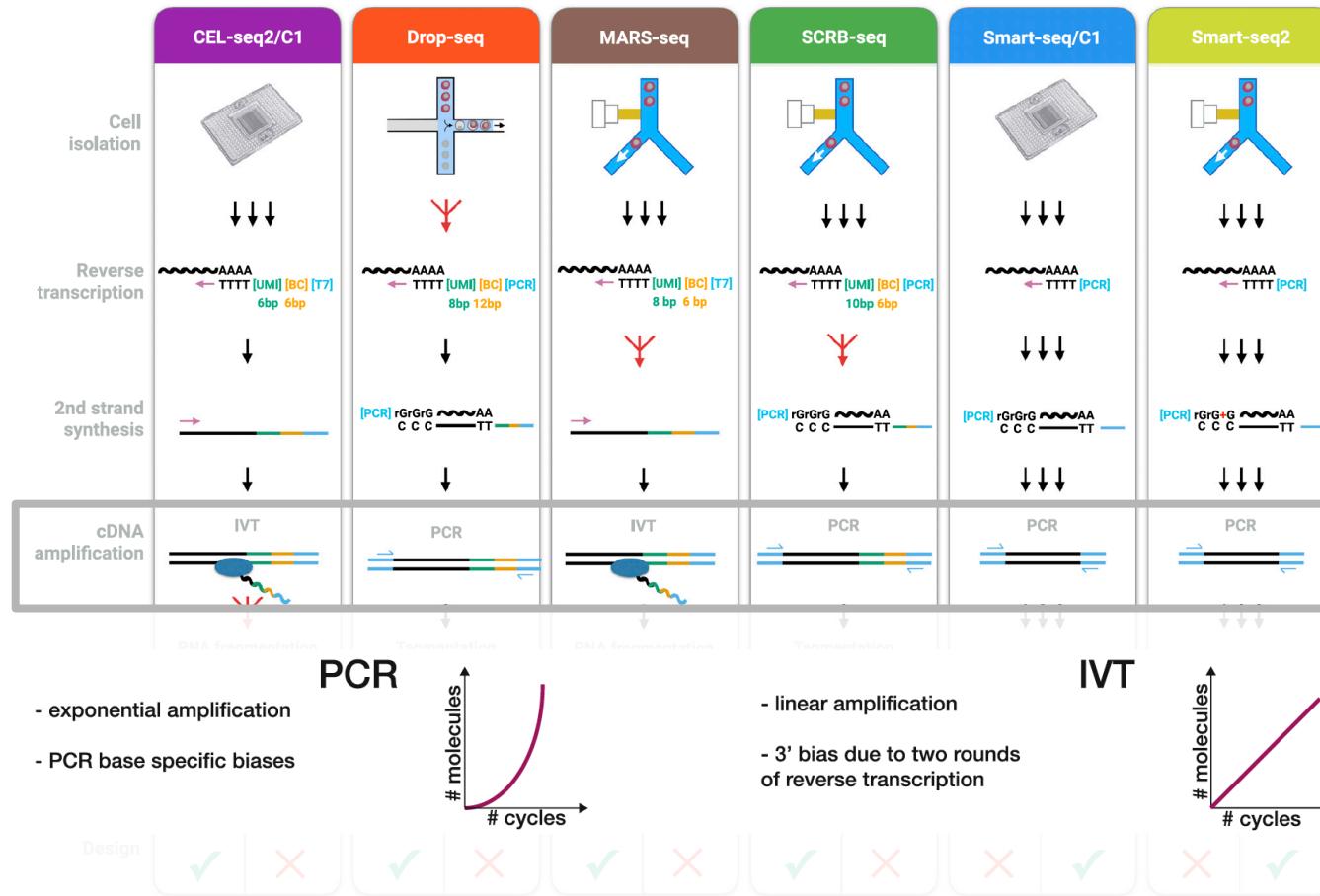


Single cell RNA-seq experiments

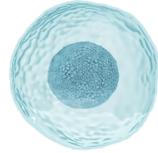




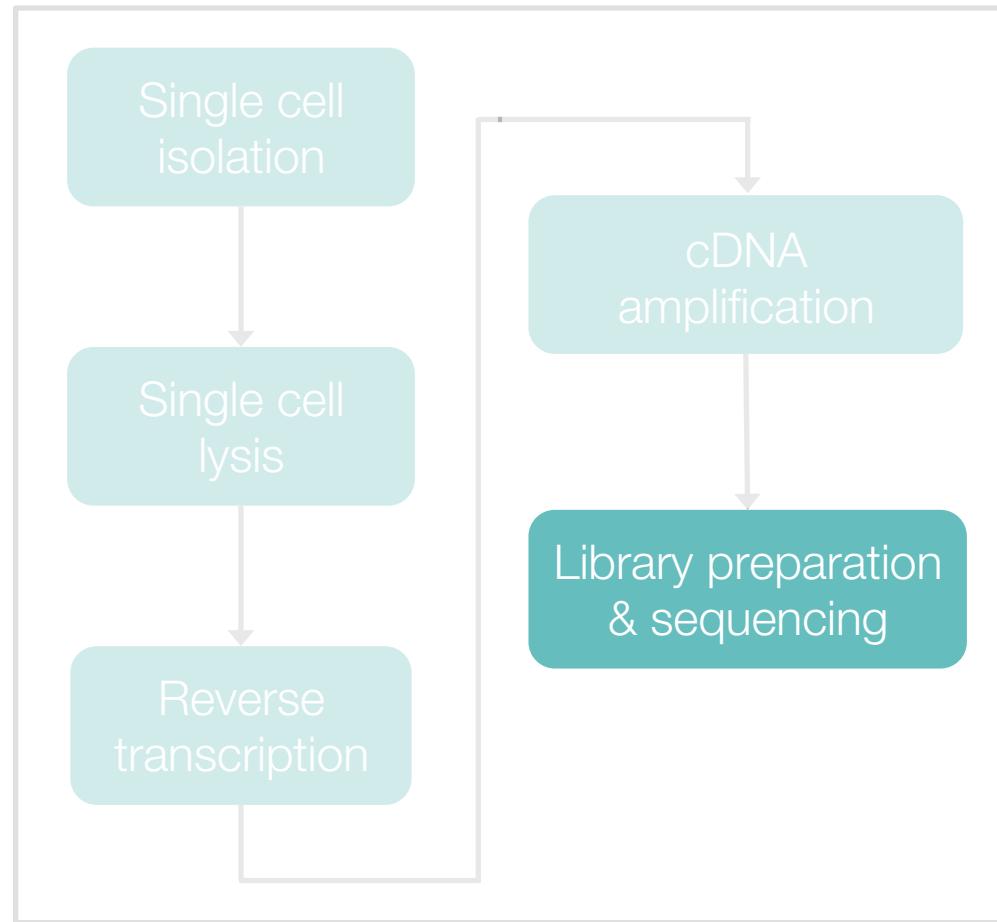
Single cell cDNA amplification

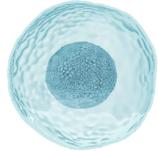


From Ziegenhain et al, 2017. doi:10.1016/j.molcel.2017.01.023

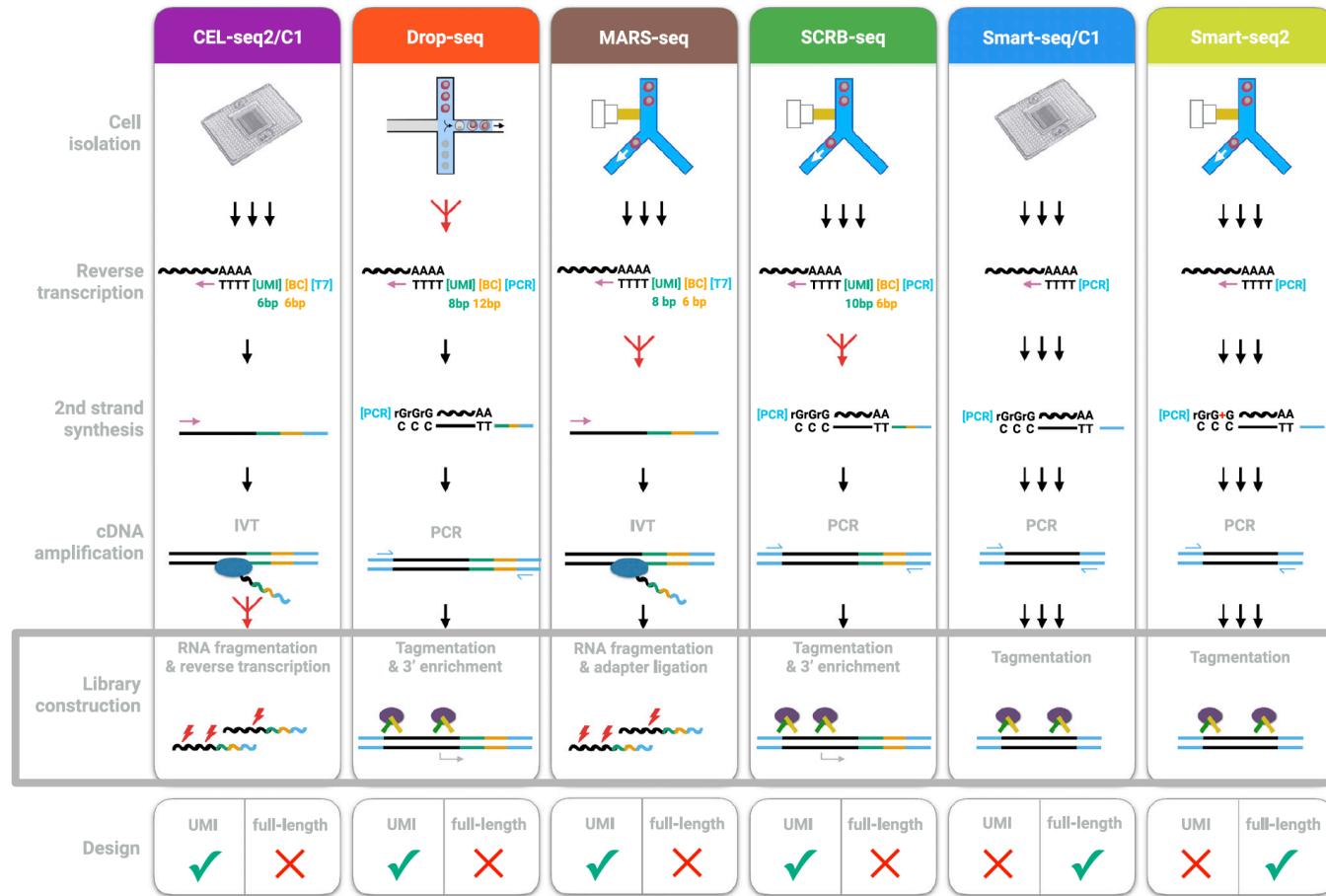


Single cell RNA-seq experiments

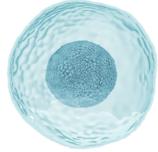




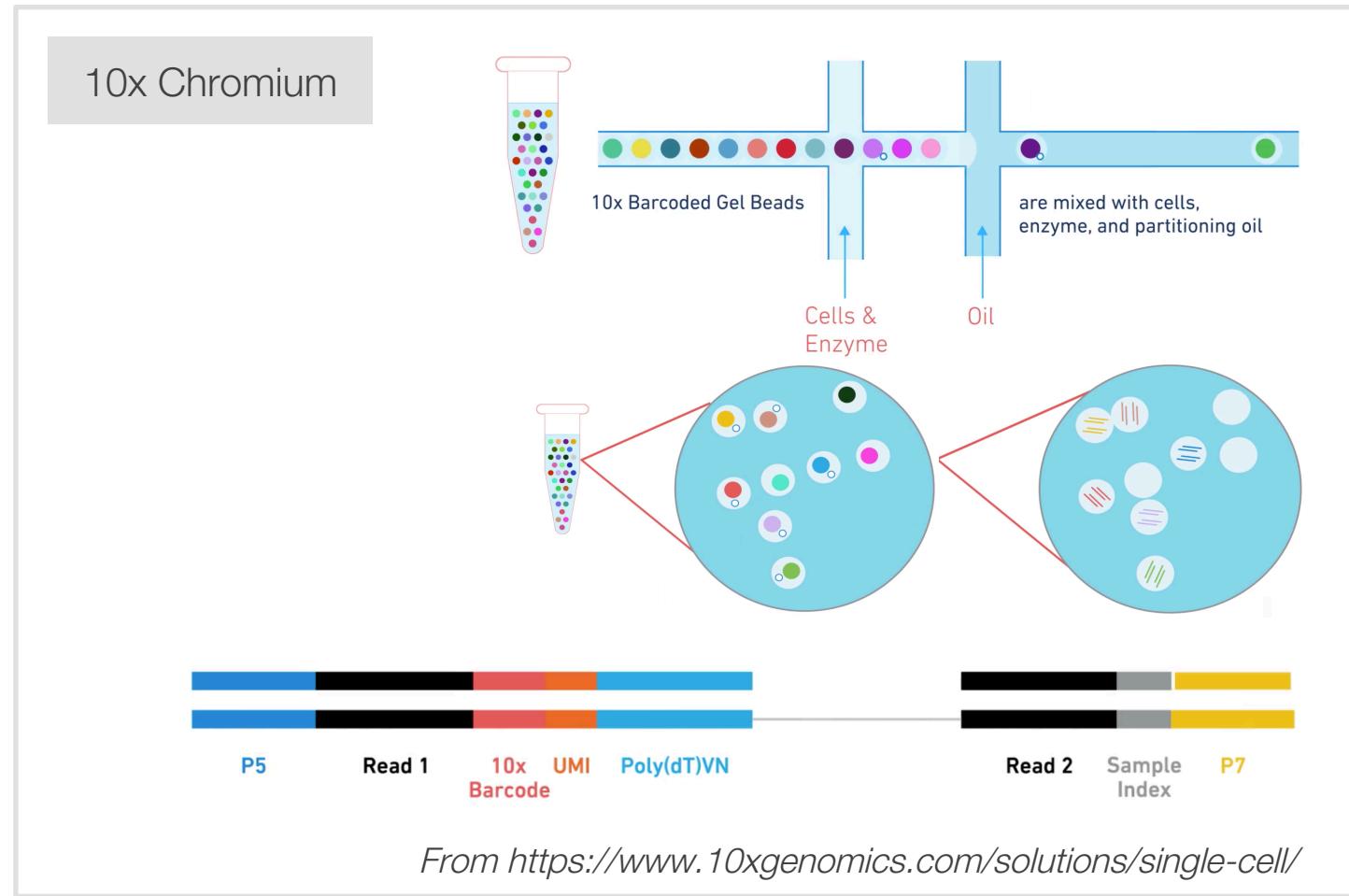
Single cell library preparation & sequencing

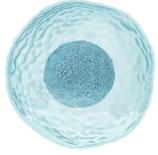


From Ziegenhain et al, 2017. doi:10.1016/j.molcel.2017.01.023

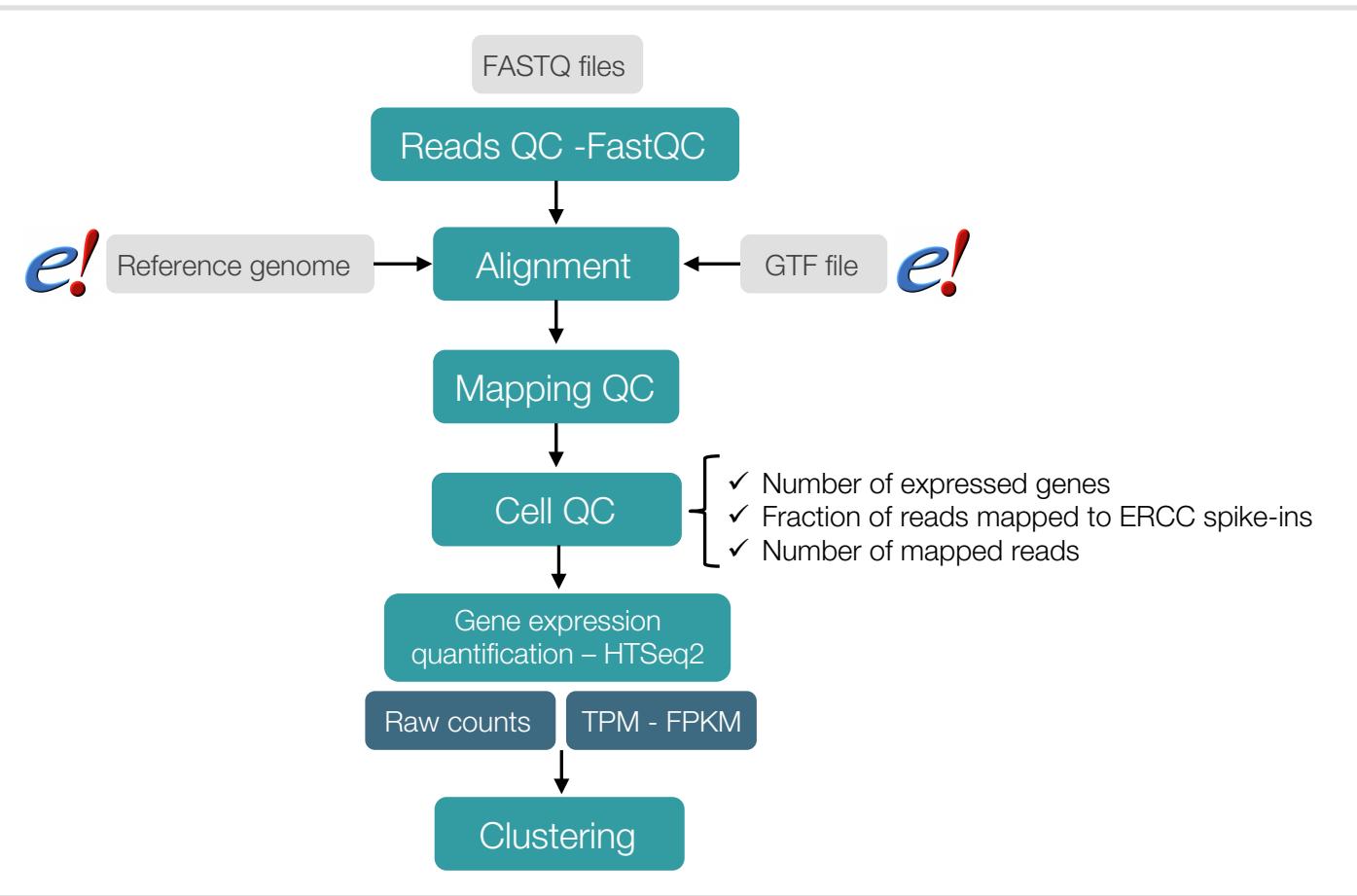


Single cell RNA-seq experiments



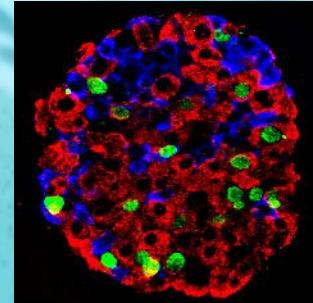
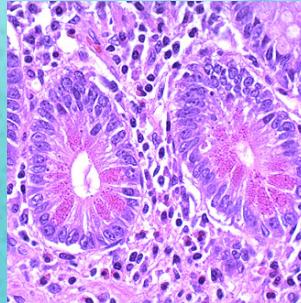
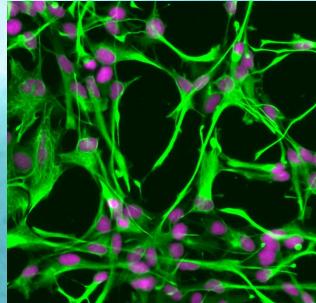


Single cell RNA-seq analysis



Single Cell Expression Atlas

- ✓ Launch in May 218
- ✓ Human, mouse, zebrafish and common fruit fly data
- ✓ Analysis results for ~48,000 cells
- ✓ Including ~200 cell types



New resource - Single Cell Expression Atlas

The screenshot shows the homepage of the Single Cell Expression Atlas. At the top, there's a navigation bar with links for Home, Browse experiments, Help, and Feedback. Below the navigation is a large search bar with the placeholder "Gene ID, gene name or gene feature". A user has typed "defa23" into the search field. To the right of the search bar is a dropdown menu labeled "Species" set to "Any". Below the search bar, there's a section for "MGD SYMBOL (MOUSE)" which shows "Defa23". Underneath this, a list of five experimental results is displayed, each with a "Cells" count, a "Results" button, and a timestamp. The results are:

- 2,244 cells | Results | 11-05-2018 | Single-cell transcriptome analysis of precursors of human CD4+ cytotoxic T lymphocytes *Homo sapiens*
- 387 cells | Results | 26-04-2018 | Single-cell RNA-seq analysis of 389 cells of the small intestinal epithelium in response to *Salmonella enterica* and to the parasitic helminth *Heligmosomoides polygyrus* *Mus musculus*
- 1,521 cells | Results | 26-04-2018 | Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium *Mus musculus*
- 3,588 cells | Results | 26-04-2018 | Single-Cell RNAseq analysis of diffuse neoplastic infiltrating cells at the migrating front of human glioblastoma *Homo sapiens*
- 648 cells | Results | 26-04-2018 | Single cell RNA-sequencing of human tonsil Innate lymphoid cells (ILCs) *Homo sapiens*

Discover and interpret gene expression analysis results at single cell level

wwwdev.ebi.ac.uk/gxa/sc/

New resource - Single Cell Expression Atlas

Search results

Gene ID or gene symbol

defa23

Search

Marker genes

Experiments with marker genes

Species

Mus musculus

Inferred cell type

Select...

Organism part

Select...

defa23 is expressed in:

26-05-2018 | *Mus musculus*

E-GEOID-99235: Single cell RNA-seq of mouse lung vascular transcriptomes

About the experiment: By analyzing about 1500 vascular single cell transcriptomes from adult mouse lung, we provide molecular definitions of the principal vascular (endothelial and mural cells), and vascular-associated cell classes. We used different transgenic mouse reporter lines with fluorescent expression in different populations of vascular cells, Cldn5(BAC)eGFP mice that strongly labels lung endothelial cells, mice positive for Pdgfrb(BAC)eGFP and NG2dsRed. We chose to focus our efforts to adult mice 10 to 19 w... [\(Read more\)](#)

25-05-2018 | *Mus musculus*

E-GEOID-103334: Single cell RNA-seq of microglia cells isolated from the hippocampus of a mouse model of severe neurodegeneration with AD-like phenotypes

About the experiment: In this study, we used single-cell RNA-sequencing to gain unprecedented insight into the phenotypic heterogeneity and the transcriptional dynamics of microglia cells during the progression of neurodegeneration. Briefly, by using a severe neurodegeneration mouse model with Alzheimer's-like pathology and phenotypes (CK-p25 model), we surveyed microglia activation by RNA sequencing longitudinally at fine temporal- and single-cell resolution. In summary, our work identified previously unobserved het... [\(Read more\)](#)

25-05-2018 | *Mus musculus*

E-ENAD-13: Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium

- [View marker gene in clusters 15 for k = 16](#)
- [View marker gene in clusters 6 for k = 6](#)
- [View marker gene in clusters 7 for k = 7](#)
- [View marker gene in clusters 2 for k = 8](#)

Defa23 gene expression across mouse cells from different single cell experiments

New resource - Single Cell Expression Atlas

Search results

Gene ID or gene symbol

defa23

Search

Marker genes

Experiments with marker genes

Species

Mus musculus

Inferred cell type

Select... ▾

Organism part

Select... ▾

Single cell mouse experiments in which *defa23* is a marker gene

defa23 is expressed in:

25-05-2018 | *Mus musculus*

E-ENAD-13: Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium

- [View marker gene in clusters 15 for k = 16](#)
- [View marker gene in clusters 6 for k = 6](#)
- [View marker gene in clusters 7 for k = 7](#)
- [View marker gene in clusters 2 for k = 8](#)
- [View marker gene in clusters 9 for k = 9](#)
- [View marker gene in clusters 10 for k = 10](#)
- [View marker gene in clusters 7 for k = 11](#)
- [View marker gene in clusters 7 for k = 12](#)
- [View marker gene in clusters 13 for k = 13](#)
- [View marker gene in clusters 14 for k = 14](#)
- [View marker gene in clusters 15 for k = 15](#)

About the experiment: By analyzing about 1500 vascular single cell transcriptomes from adult mouse lung, we provide molecular definitions of the principal vascular (endothelial and mural cells), and vascular-associated cell classes. We used different transgenic mouse reporter lines with fluorescent expression in different populations of vascular cells, Cldn5(BAC)eGFP mice that strongly labels lung endothelial cells, mice positive for Pdgfrb(BAC)eGFP and NG2dsRed. We chose to focus our efforts to adult mice 10 to 19 w... [\(Read more\)](#)

New resource - Single Cell Expression Atlas

Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium

Single-cell RNA-Seq mRNA baseline

Organism: *Mus musculus*

Publication:

• Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K et al. (2017) *A single-cell survey of the small intestinal epithelium.*



Very high *defa23* expression
in Paneth cells

New resource - Single Cell Expression Atlas

Results Experiment Design Supplementary Information Downloads

Metadata is manually curated
and annotated to ontologies

Cell	Sample Characteristics										Experimental Variables
	age	cell type	facs marker	genotype	individual	inferred cell type	organism	organism part	strain	single cell identifier	
SRR5102341	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	late enterocyte progen...	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 1	
SRR5102342	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	enterocyte	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 2	
SRR5102343	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 3	
SRR5102344	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 4	
SRR5102345	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 5	
SRR5102346	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 6	
SRR5102347	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 7	
SRR5102348	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 8	
SRR5102349	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	goblet cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 9	
SRR5102350	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 8	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 10	
SRR5102351	7 to 10 week	epithelial cell	LGR5-High	Lgr5-GFP	Mouse 7	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 11	
SRR5102352	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	late enterocyte progen...	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 12	
SRR5102353	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	goblet cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 13	
SRR5102354	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	stem cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 14	
SRR5102355	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 15	
SRR5102356	7 to 10 week	epithelial cell	Epcam	wild type	Mouse 4	transit amplifying cell	Mus musculus	small intestine	C57BL/6J	Atlas WT single cell 16	

New resource - Single Cell Expression Atlas

Details on the methods followed to re-analyse each experiment

Results	Experiment Design	Supplementary Information	Downloads
---------	-------------------	---------------------------	-----------

Analysis Methods

Analysis	Software	Version	Citation
QC	FASTQC	0.11.7	http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Filtering and trimming	FASTX	0.0.13	http://hannonlab.cshl.edu/fastx_toolkit/
Pipeline	IRAP	1.0.0b	Nuno A. Fonseca, Robert Petryszak, John Marioni, Alvis Brazma. iRAP - an integrated RNA-seq Analysis Pipeline. bioRxiv 2014. doi: http://dx.doi.org/10.1101/005991
Clustering/Gene markers	SC3	1.4.2	Kiselev, Vladimir Yu, Kristina Kirschner, Michael T. Schaub, Tallulah Andrews, Andrew Yiu, Tamir Chandra, Kedar N. Natarajan et al. SC3: consensus clustering of single-cell RNA-seq data. Nature methods 14, no. 5 (2017): 483.
Reads alignment	bowtie2	2.3.2	Langmead, Ben, and Steven L. Salzberg. Fast gapped-read alignment with Bowtie 2. Nature methods 9, no. 4 (2012): 357-359.
QC	fastq_utils	0.14.7	https://github.com/nunofonseca/fastq_utils
Gene/transcript quantification	kallisto	0.44.0	Nicolas L Bray, Harold Pimentel, Pál Melsted, Lior Pachter. Near-optimal probabilistic RNA-seq quantification. Nature Biotechnology (2016).

New resource - Single Cell Expression Atlas

Download curated metadata
and analysis results

Single-cell RNA-seq analysis of 1,522 cells of the small intestinal epithelium

Single-cell RNA-Seq mRNA baseline

Organism: *Mus musculus*

Publication:

- Haber AL, Biton M, Rogel N, Herbst RH, Shekhar K et al. (2017) [*A single-cell survey of the small intestinal epithelium.*](#)

Results

Experiment Design

Supplementary Information

Downloads

Metadata files

 SDRF file (.txt)

 IDF file (.txt)

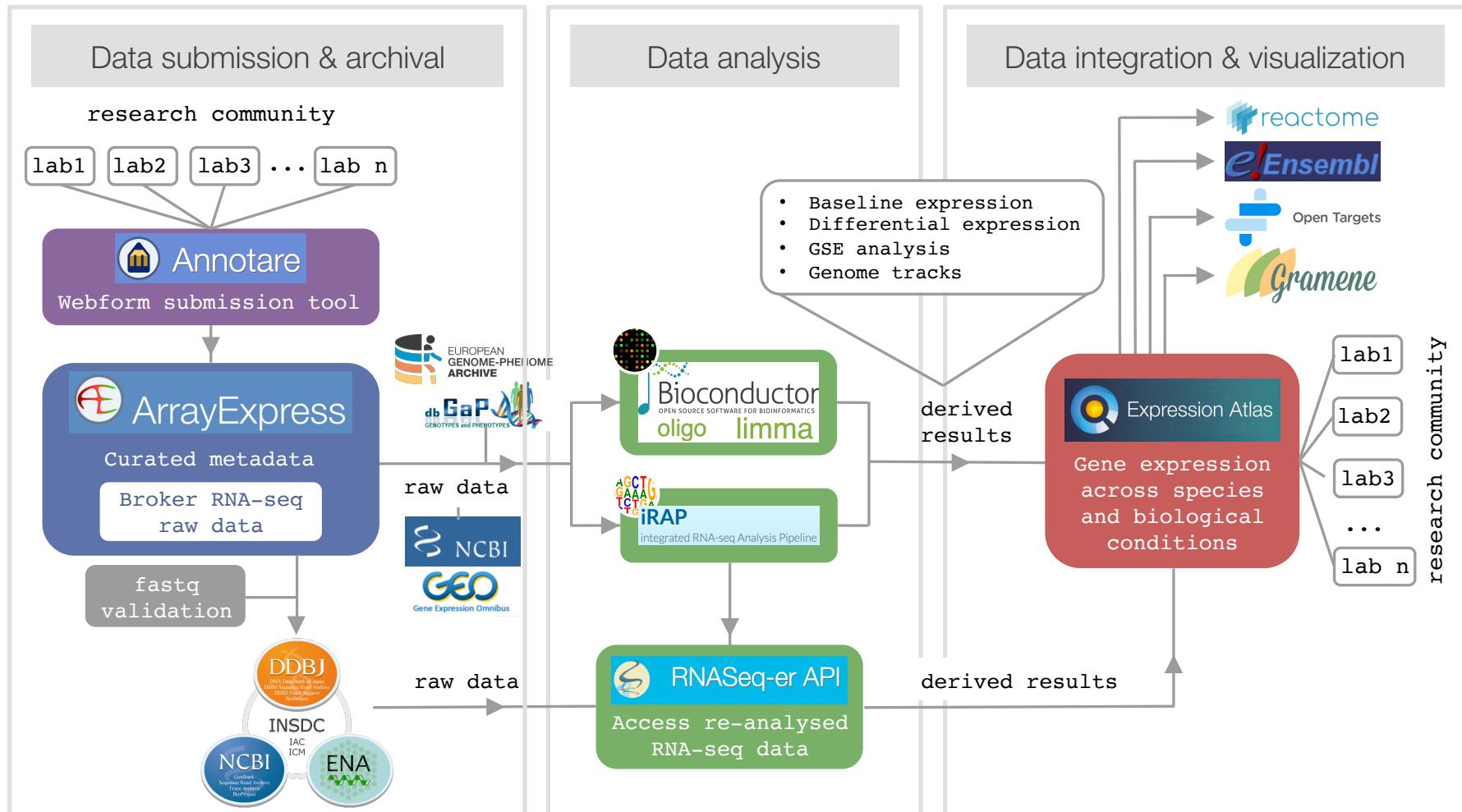
 Experiment design file (.tsv)

Result files

 Clustering file (.tsv)

 Filtered quantification files (MatrixMarket)

Functional genomics resources at EMBL-EBI





Expression Atlas: Who we are?



Irene
Papatheodorou
Team Leader

Data curation



Laura Huerta

Data analysis



Pablo Moreno

User Interface



Alfonso Fuentes



Monica Jianu



Nancy George



Anja Fullgrabe



Jon Manning



Suhaib Mohammed



Haider Iqbal



Acknowledgements

- ✓ Gramene collaborators
- ✓ EBI teams:
 - SPOT
 - Reactome
 - ENA
 - Ensembl
 - Ensembl Genomes
 - PRIDE
- ✓ Chan Zuckerberg Initiative
- ✓ Human Cell Atlas
- ✓ Wellcome Trust



EMBL-EBI workshop: Data and tools for transcriptomics and protein biology

Single Cell Expression Atlas: a new tool for single cell transcriptomics

Laura Huerta, PhD

Senior Scientific Curator

lauhuema@ebi.ac.uk

Malta, 31 July 2018