Research on Tap:
Medicine in the Molecular Era: Single Cell Sequencing

October 1, 2018

DoM Single Cell Sequencing Core Facility

Yuriy Alekseyev, PhD

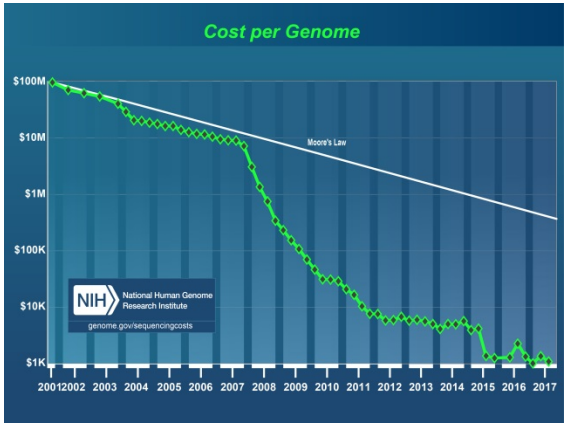
Director, Single Cell Sequencing Core
Director, Microarray and Sequencing Resource

Research Associate Professor, Department of Pathology and Laboratory Medicine

Recent Revolutionary Changes in Science and Medicine Driven by Technology

Revolution 1: Genomics

2003 Completion of the Human Genome Sequencing Project ($2.7B)
2007 Craig Venter's genome sequenced by Sanger sequencing ($10M)
2007 James Watson's genome sequenced by 454 technology ($2M)
2014 Illumina announces their new instrument HiSeqX capable of sequencing whole genome for less than $1000
2015 100,000 Genome project launched focusing on rare disease and cancer
2017 Illumina announces new instrument series NovaSeq capable of sequencing 48 genomes in less than 40h under $1000 per genome

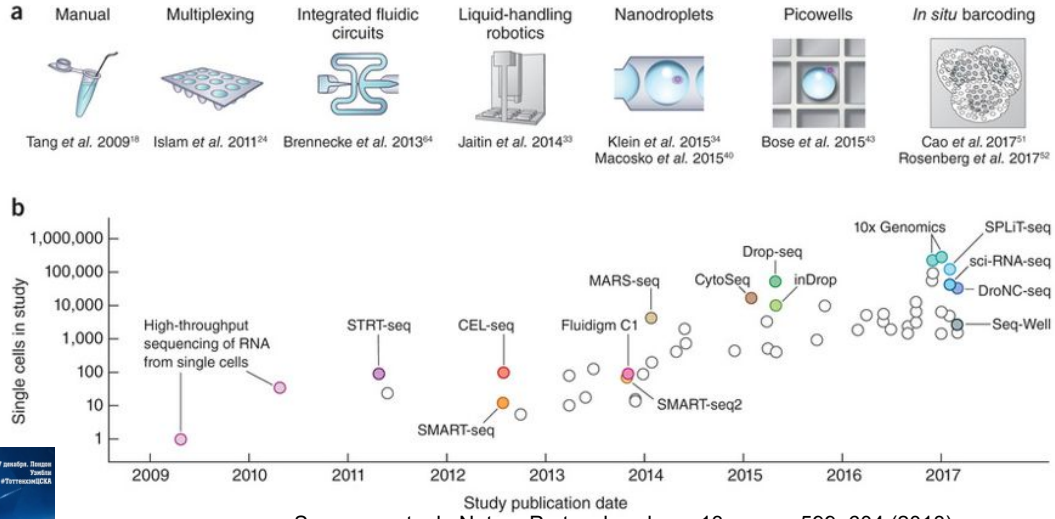


Wetterstrand KA. www.genome.gov/sequencingcostsdata

Revolution 2: Single Cell Genomics

Revolution in Molecular Biology: Total RNA minimal input for RNA-seq drops from ug amounts to 10-30 pg

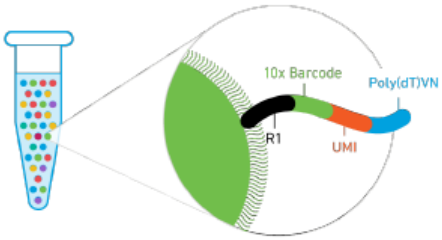
Development of methods for capturing Single Cells

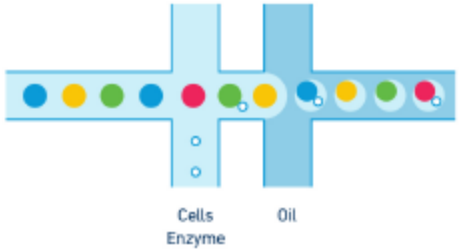


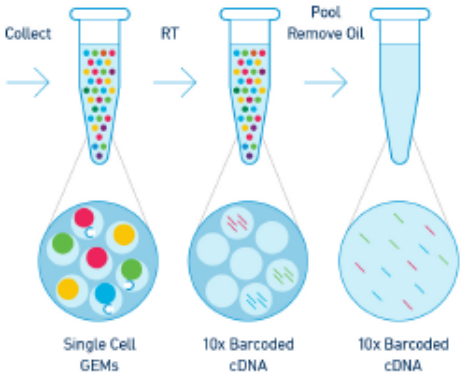
Svensson et. al., Nature Protocols volume 13, pages 599–604 (2018)

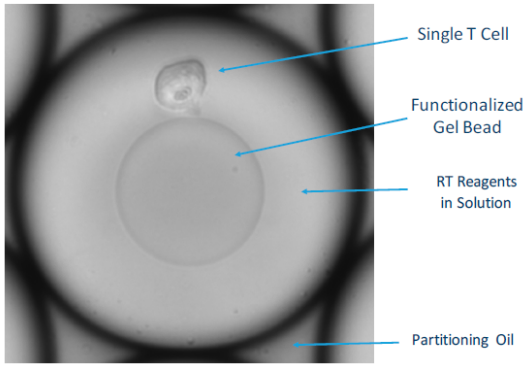
Why Single Cell Genomics ?

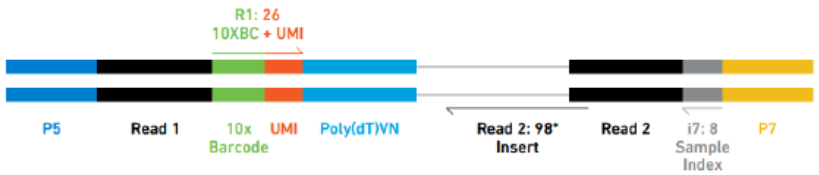
Current Leader in Single Cell Sequencing technology: 10x Genomics











DoM Single Cell Sequencing Core Services:

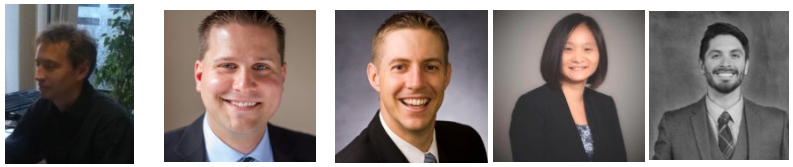
http://www.bumc.bu.edu/singlecell
http://www.bumc.bu.edu/microarray

Help with tissue dissociation

Single Cell Library Preparation (10x Genomics, Illumina/biorad ddSeq)
Coming soon: Manual plate-based single cell library preparation (Cel-seq2, NEB)
Sequencing (Microarray and Sequencing Resource)

Data Analysis

The Team:



Boston University Office of the Vice President and Associate Provost for Research

16+ Color Flow Cytometry with Multivariate Computational Analysis: Tools to Reveal Biomarkers and Mechanisms of Disease

Jennifer E. Snyder-Cappione, PhD

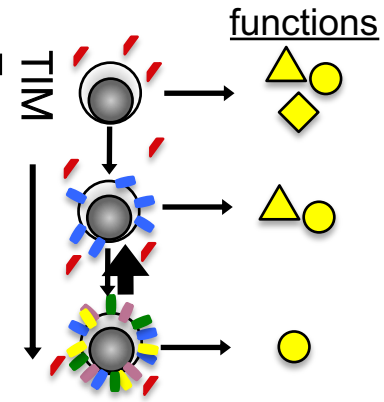
Assistant Professor, Department of Microbiology
Director, Flow Cytometry Core Facility
Boston University School of Medicine

BOSTON UNIVERSITY

Research Interest: Immune Cell Alterations in Chronic Conditions

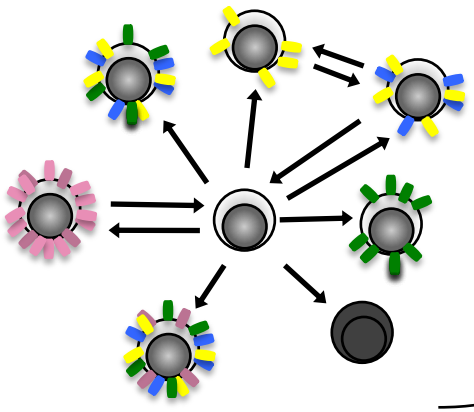
infections, autoimmunity, cancer

Immune Exhaustion
(T cells, NK cells)



Inhibitory Receptors (IRs)

Branched Differentiative Process



How do these expression patterns:

(1) vary with different immune cell subsets and chronic diseases?

(2) track with function?

Developed a 16-color flow panel to measure IR signatures on several immune cell subsets from one sample *Cytometry A.* 91(2):175-179, 2017

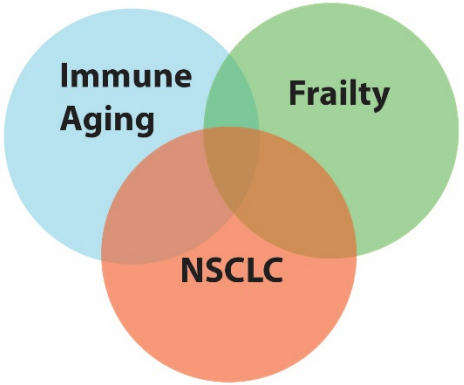
Ongoing Projects

(1) Mechanisms of age-induced inflammation +/- aviremic HIV infection

Rahm Gummuluru Ph.D., Nina Lin M.D., Manish Sagar M.D.

(2) Role of frailty, NSCLC on immune aging of older individuals

Rawad Elias, M.D.
UCONN

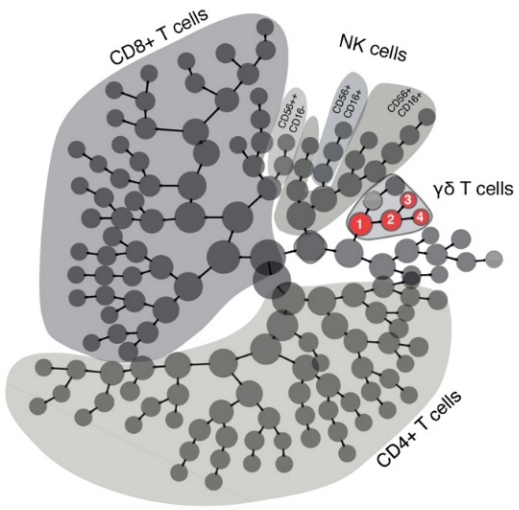


How can we:

(1) Analyze these enormous datasets?

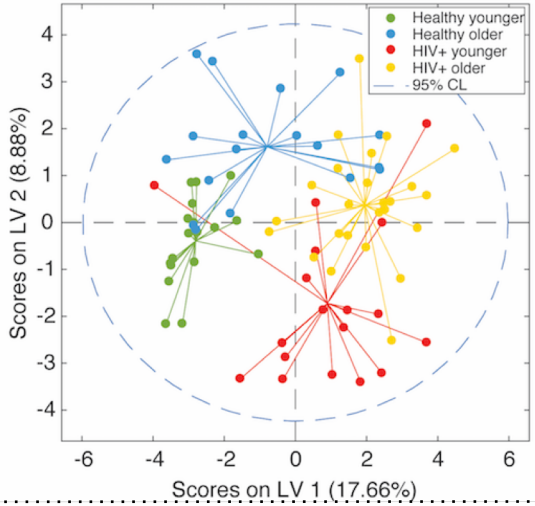
(1) Define the impact when conditions collide?

CITRUS reveals the cell signature that stratifies aviremic HIV+ subjects from controls



TIGIT on γδ T cells placed subjects into correct group with 89% efficiency

PLS-DA: Divergent 'Inflamm-aging' +/- HIV

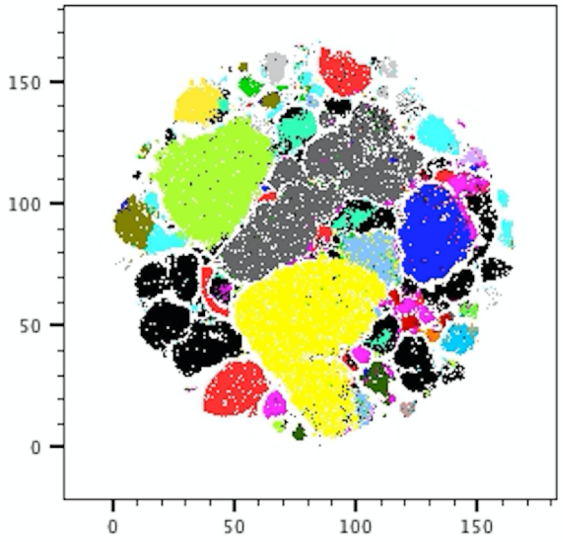


Lauffenberger lab
MIT

γδ T cell IR signatures, plasma cytokines used to train model

BUMC Flow Core: Advances in Single Cell Analysis

opt-SNE: Anna Belkina, M.D. Ph.D.



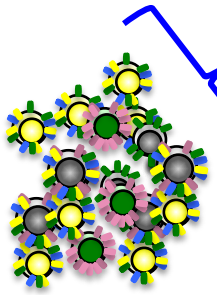
18 parameter dataset

>30 populations resolved

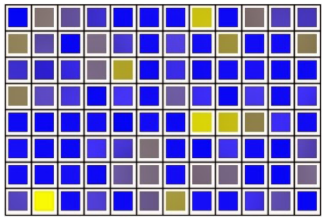
>20 million events included

Proteomic + Transcriptomic Sample Processing

digested tumors, blood, etc.



20+ parameter phenotyping



indexed single cell sorting

computational analysis
with Single Cell Core

scRNAseq data

Lung Cell Fate Trajectories Profiled by scRNA-Seq Time Series

Darrell N. Kotton, MD

David C. Seldin Professor of Medicine
Center for Regenerative Medicine

BU School of Medicine

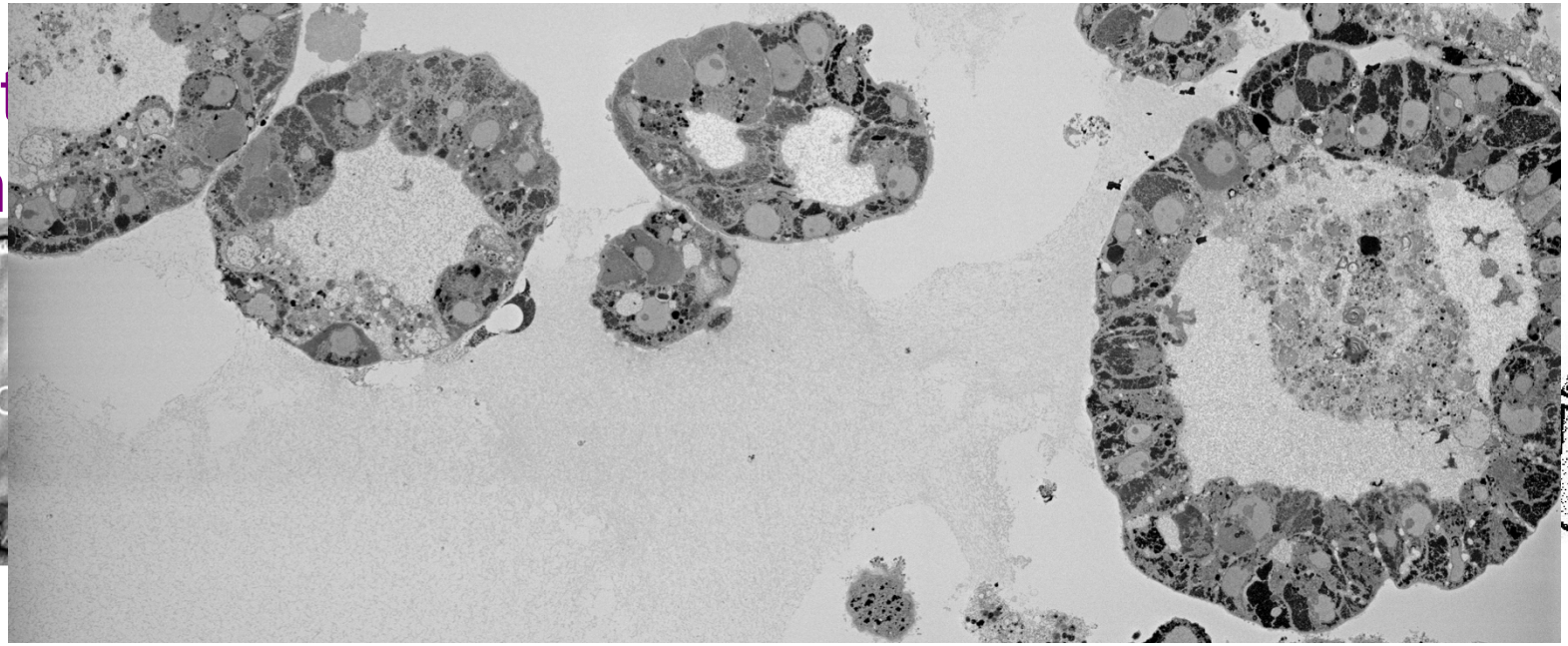
CReM
CENTER FOR REGENERATIVE MEDICINE

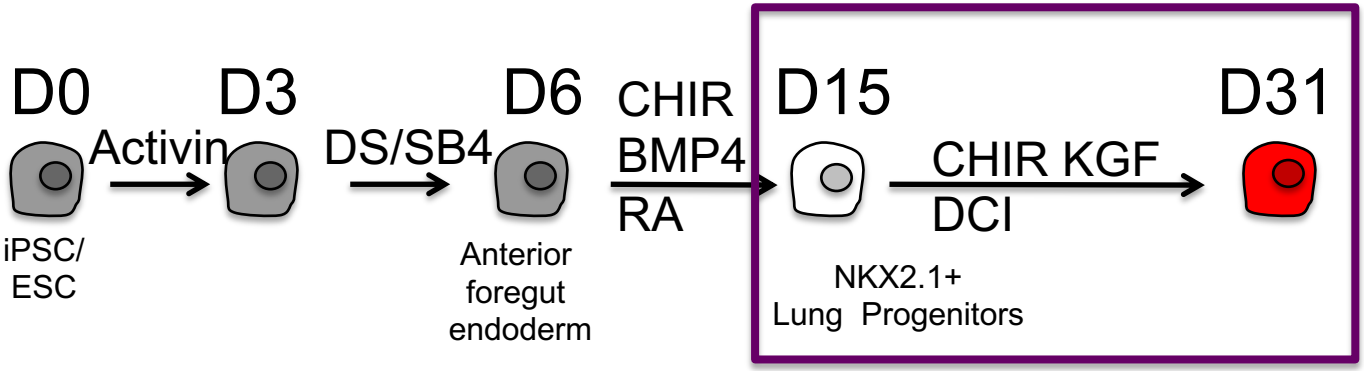
Boston Medical Center
EXCEPTIONAL CARE. WITHOUT EXCEPTION.

Boston University Office of the Vice President and Associate Provost for Research

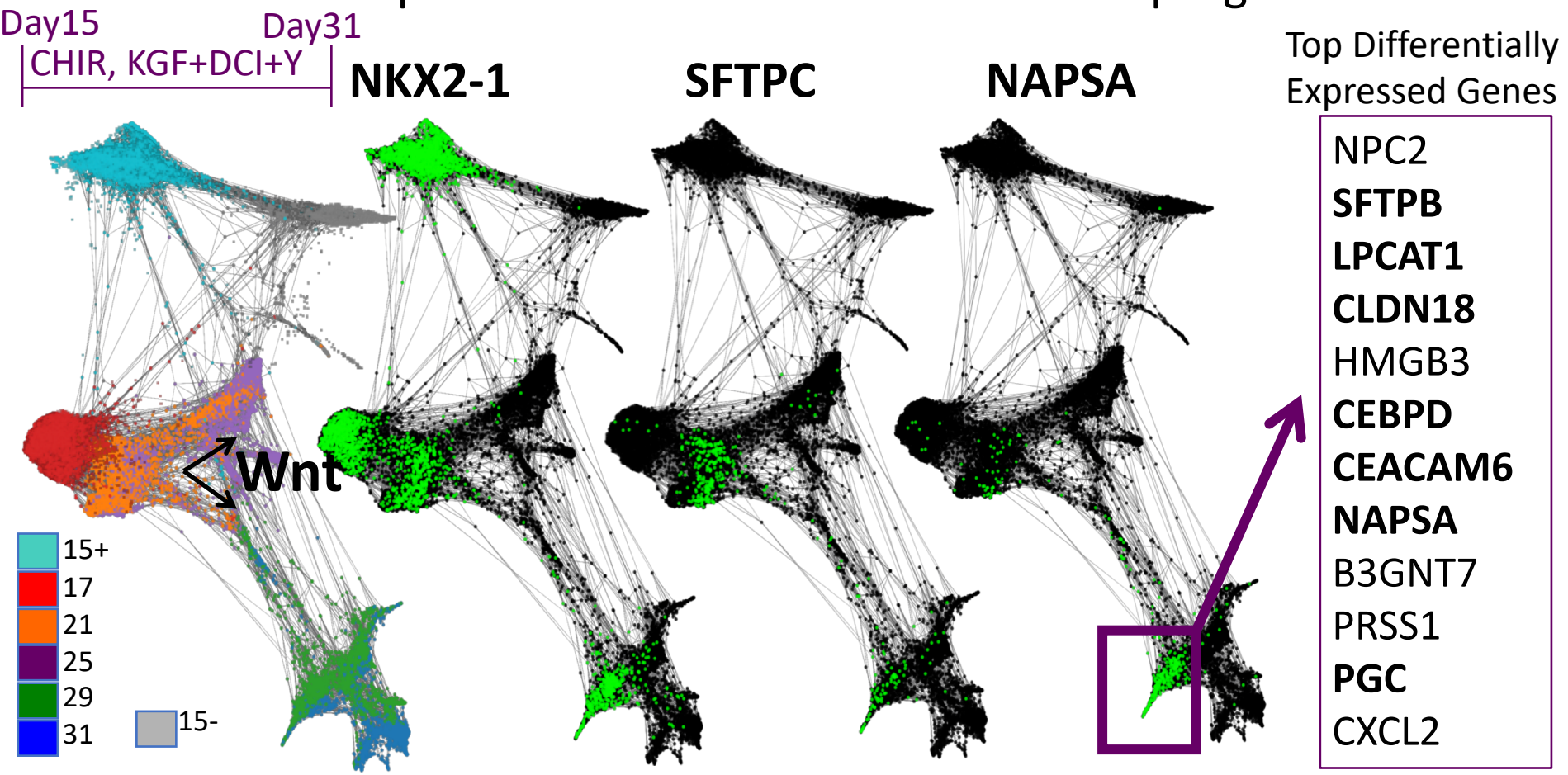
Boston University

Blast
Em




D0
D3
D6
CHIR
D15
D31
Activin
DS/SB4
BMP4
RA
CHIR KGF
DCI
iPSC/
ESC
Anterior
foregut
endoderm
NKX2.1+
Lung Progenitors

Time series data reveals differentiation and maturation of Alveolar epithelial cells from sorted NKX2-1+ progenitors



Thanks!

CReM

Kotton Lab

Killian Hurley

Nacho Caballero

Collaborators:
Ziv Bar-Joseph Lab

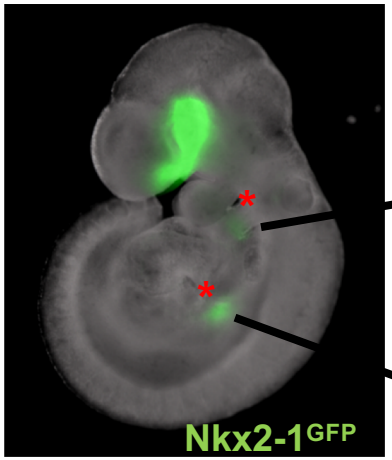
Epithelial Progenitor Heterogeneity in Lung Development

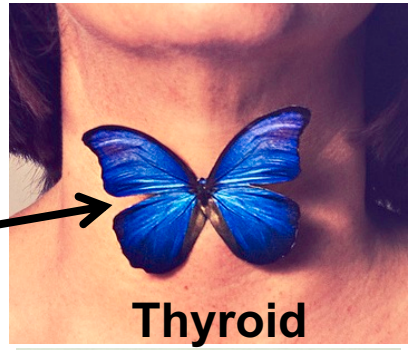
Laertis Ikonomou

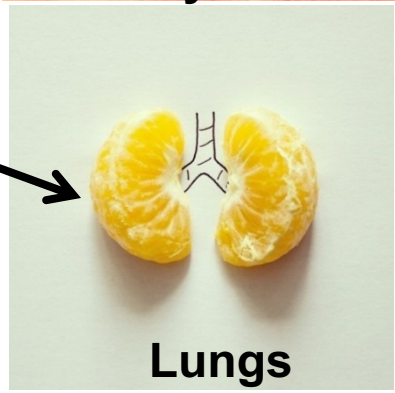
Assistant Professor
Center for Regenerative Medicine, Boston University and Boston Medical Center

BOSTON UNIVERSITY

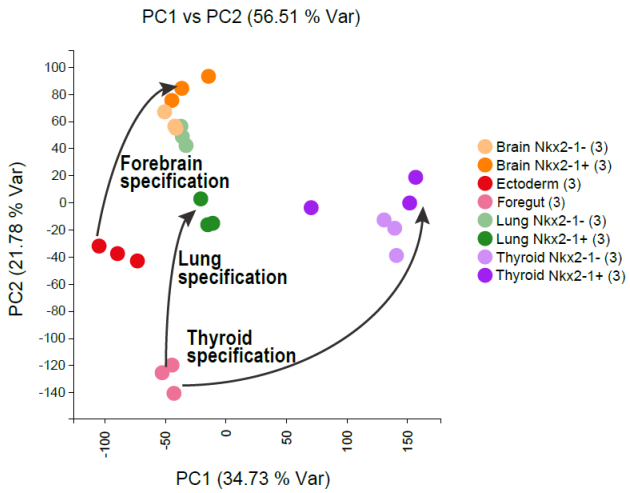
- **Nkx2-1⁺ lung and thyroid primordial progenitors arise as small clusters of cells* within the anterior foregut endoderm (AFE).**
- **Important intermediate cells in development and regenerative medicine (pluripotent stem cell directed differentiation): In vivo progenitors are "gold standard" for evaluating in vitro derived progenitor.**

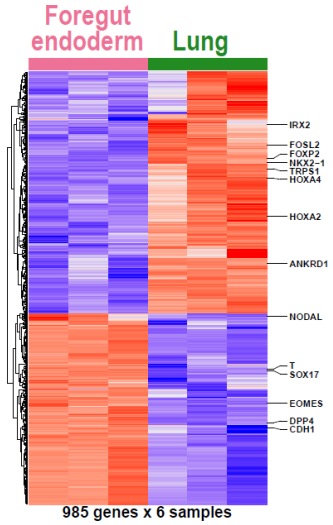


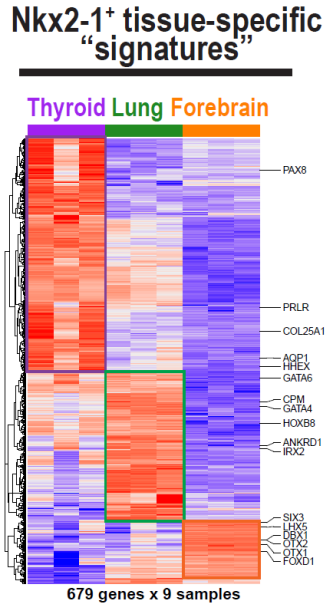




Population RNA-Seq



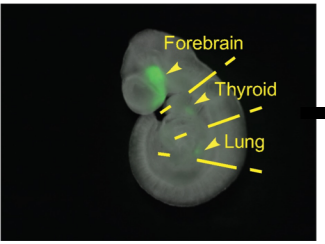


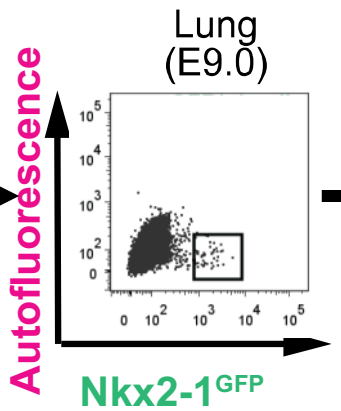


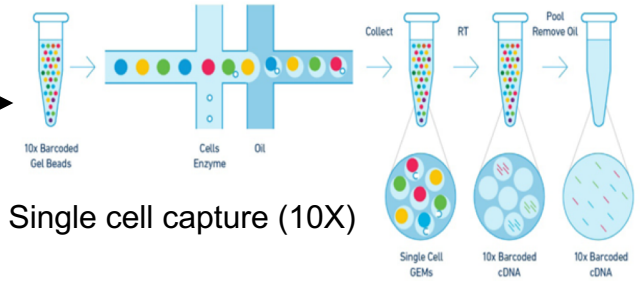
- **Progenitor transcriptional heterogeneity?**
- **Functional implications?**

BOSTON UNIVERSITY

Single-cell RNA-Seq (E9.0-E11.5)







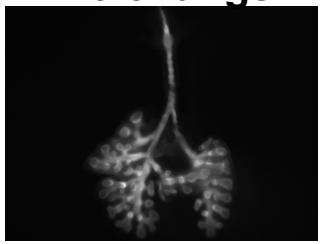
Single cell capture (10X)

Library prep & sequencing

Analysis

Proof-of-principle (E13.5 capture and seq)

E13.5 lungs

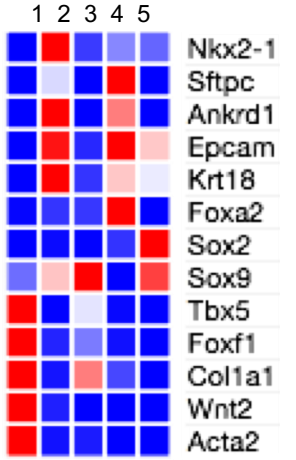


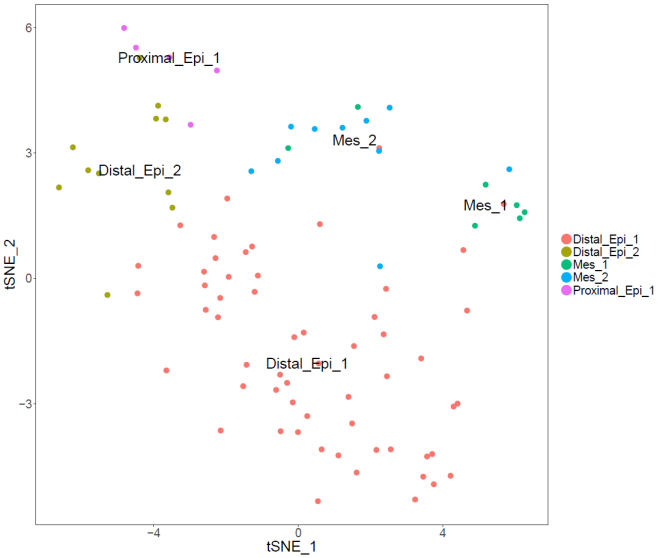
Digest, sort

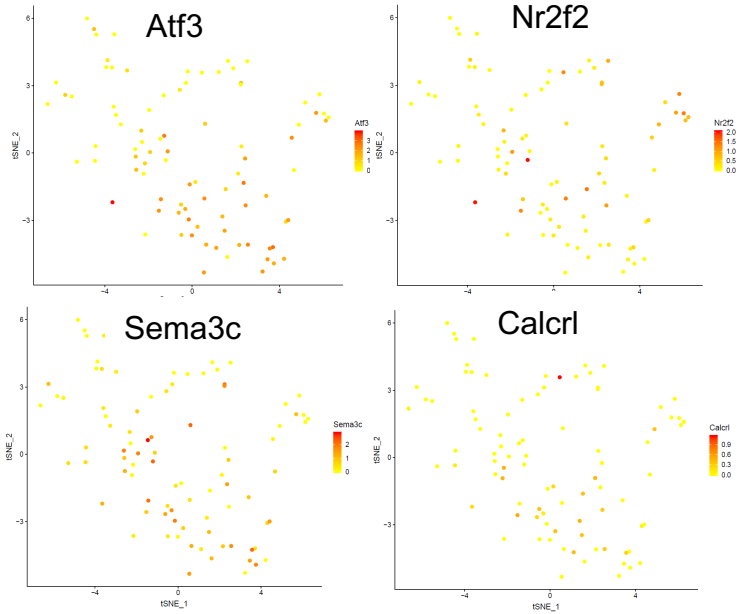
Mix Nkx2-1^{+} : Nkx2-1^{-} cells (90:10)

Capture

Library prep, Seq, Analysis







- **Special thanks to:**
  - Sara Lewandowski
  - Matt Lawton
  - Taylor Matte
  - Carlos Villacorta
  - Darrell N. Kotton
  - Paul Blainey/Ankur Kulshreshtha (Broad)

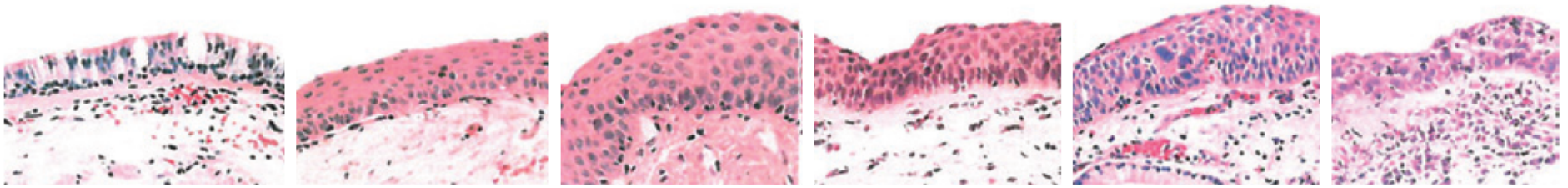
Characterizing Immune Heterogeneity Associated with Lung Cancer Premalignancy

Joshua Campbell

Assistant Professor
Dept. of Medicine, BU School of Medicine

BOSTON UNIVERSITY

Squamous Lung Carcinogenesis

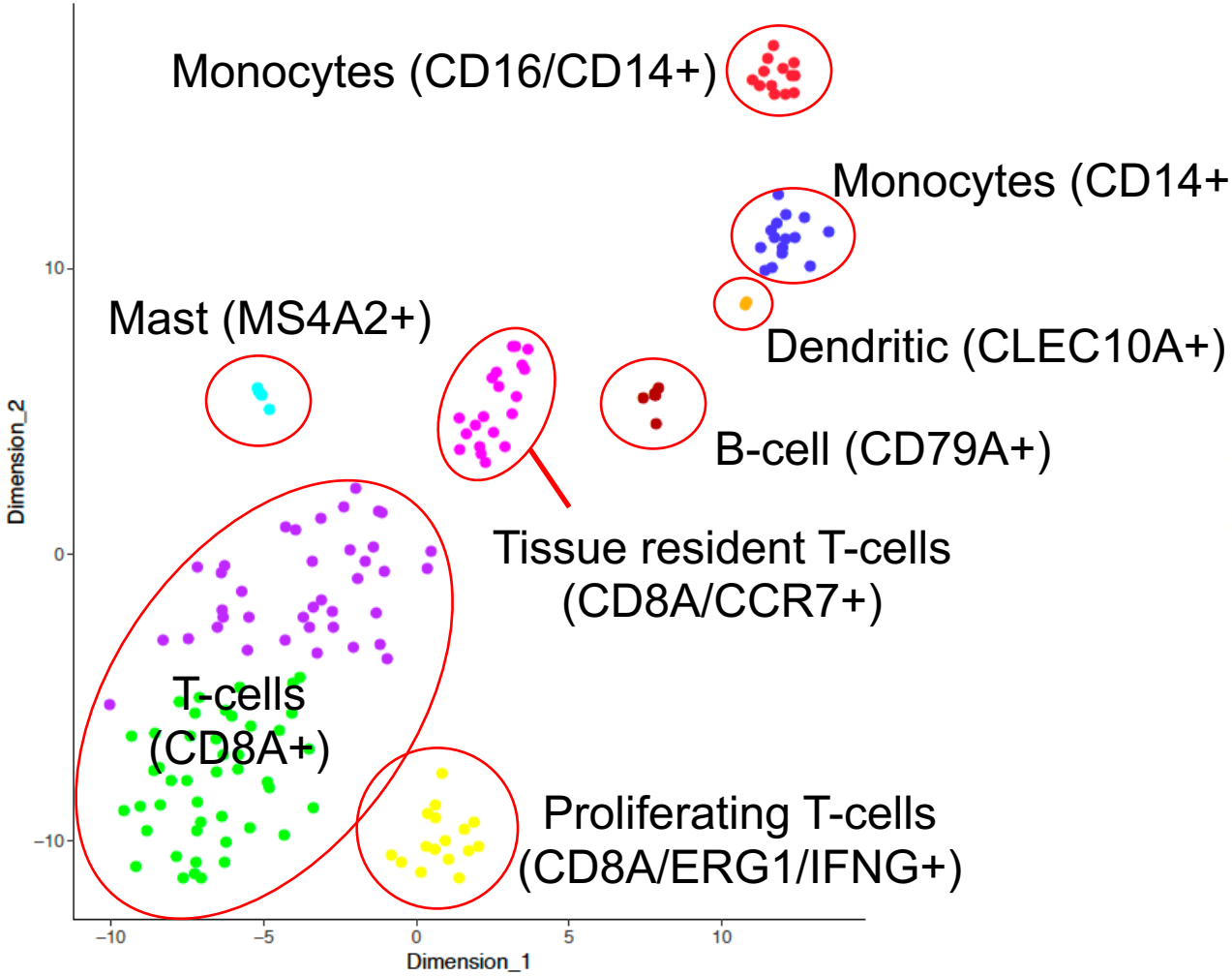


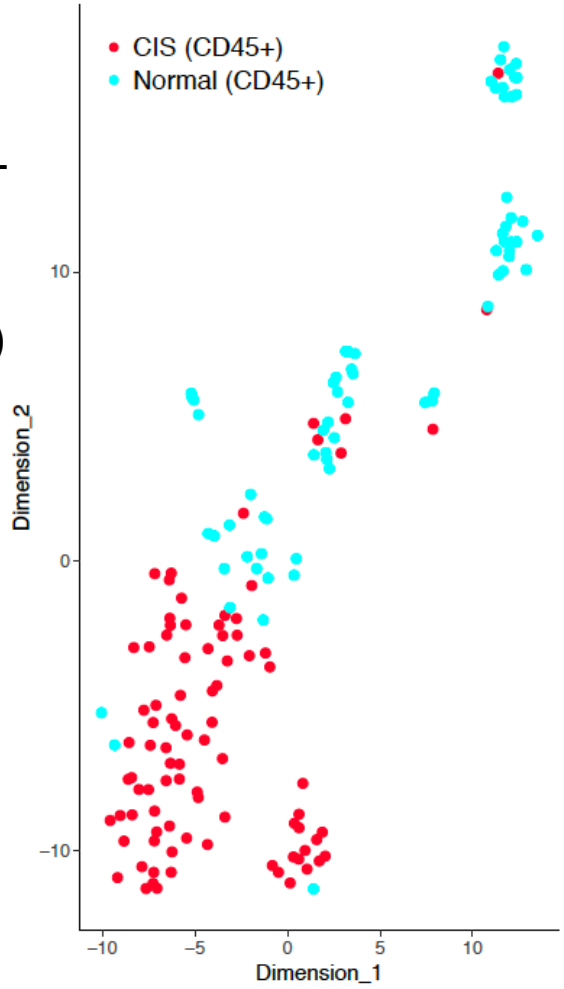
Normal | Squamous metaplasia | Mild dysplasia | Moderate dysplasia | Severe dysplasia | Carcinoma *in situ*

Only a subset of airway premalignant lesions progress to CIS/tumor

Many airway premalignant lesions will regress without intervention

Keith, R. L. & Miller, Y.E. *Nat. Rev. Clin. Oncol*. 2013
Breuer et al, *Clin Cancer Res*., 2005
van Boerdonk, *AJRCCM*, 2015

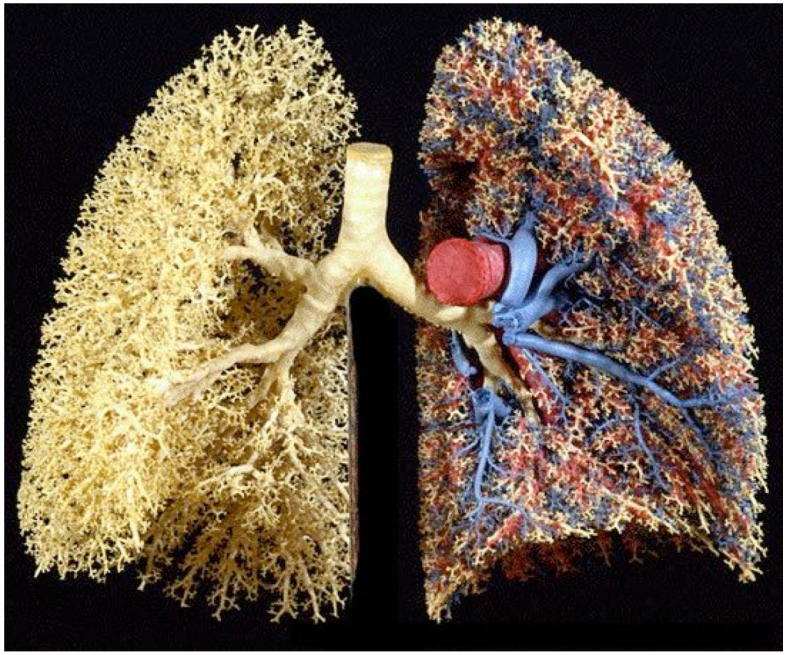

Monocytes (CD16/CD14+)
Monocytes (CD14+
Dendritic (CLEC10A+)
Mast (MS4A2+)
B-cell (CD79A+)
Tissue resident T-cells
(CD8A/CCR7+)
T-cells
(CD8A+)
Proliferating T-cells
(CD8A/ERG1/IFNG+)
Dimension_2
Dimension_1
10
0
−10
−10
−5
0
5
10


CIS (CD45+)
Normal (CD45+)
Dimension_2
Dimension_1
10
0
−10
−10
−5
0
5
10

Cellular diversity in the (regenerating) adult lung

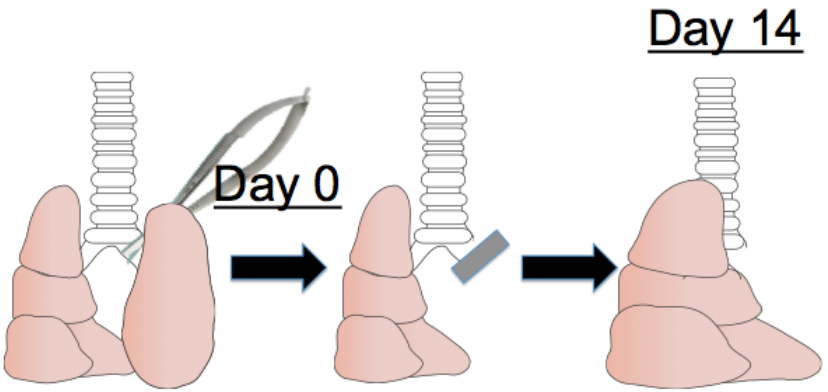
Jason Rock, PhD

Associate Professor
Department of Medicine/CReM, BUSM

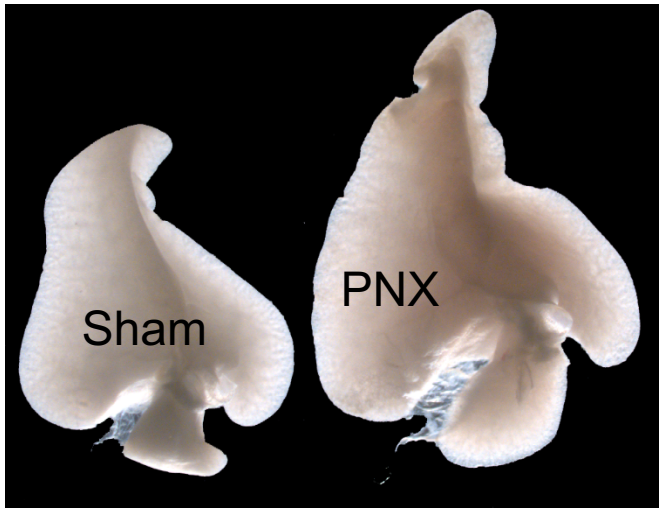


Adult lungs are complex –
~40 cell types
~400 million alveoli

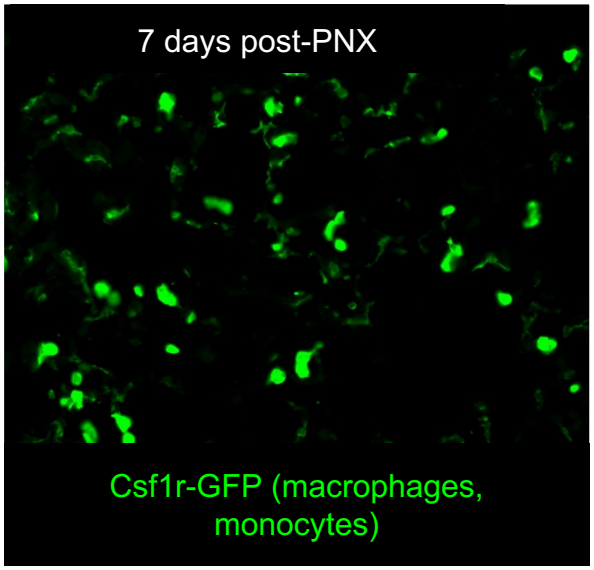
Adult lungs regenerate after mild to moderate injury (e.g., pneumonectomy)

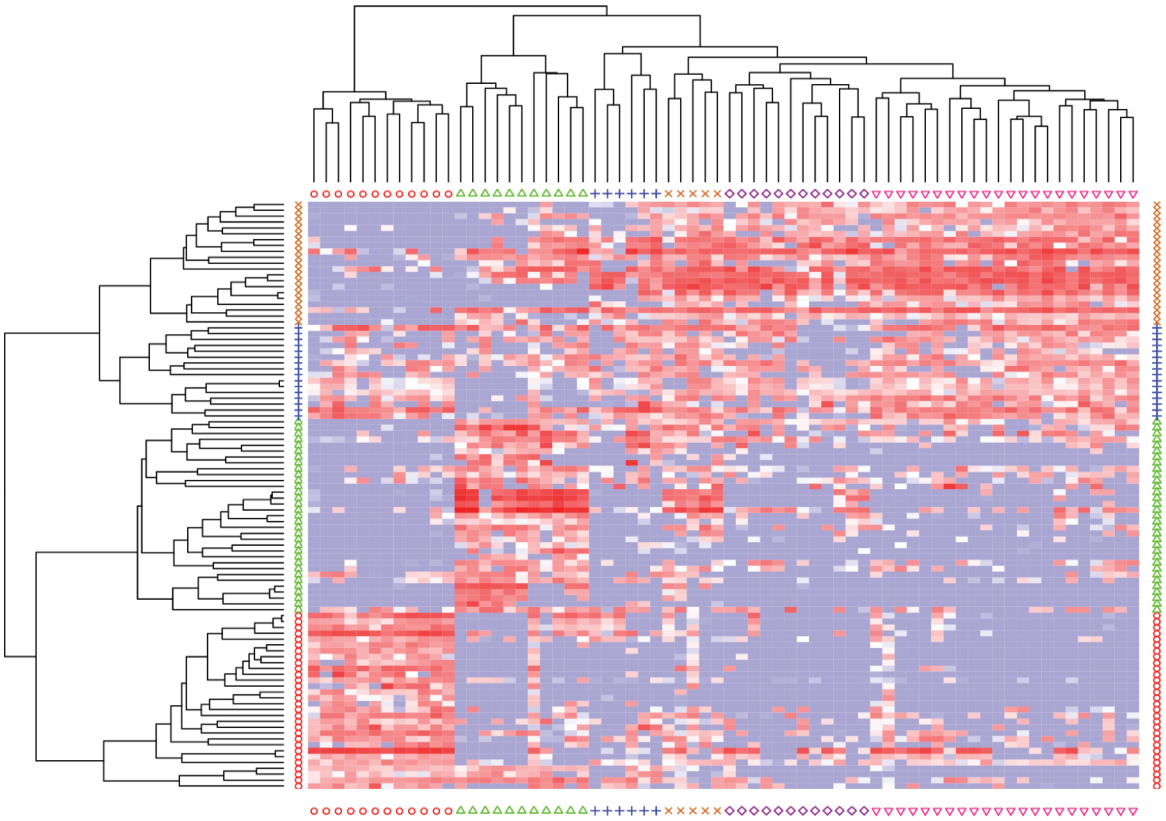


Increased lung volume, weight, and number of alveoli



Single cell sequencing identifies subsets of myeloid cells in the regenerating lung





7 day post-PNX Myeloid Subset Gene Groups

CCR2+	Eno1	MARCKS	Chil2
Ly6C	Sirpb1b	CCL 2, 4, 7	Chil3
Thbs1	Ndufa 4/8	CXCL12	Mrc1
Ifit 2, 3, 6	Mrpl18	CD74	IL18
CD52	Cox 3/6a1	H2-Ab1	CD164
	Cbx3	CD83	CD2
		Id2	CD9

Ongoing efforts to characterize epithelial cells, fibroblasts, endothelial cells and other leukocytes

Statistical Analyses of Single Cell Transcript with MAST

Masanao Yajima

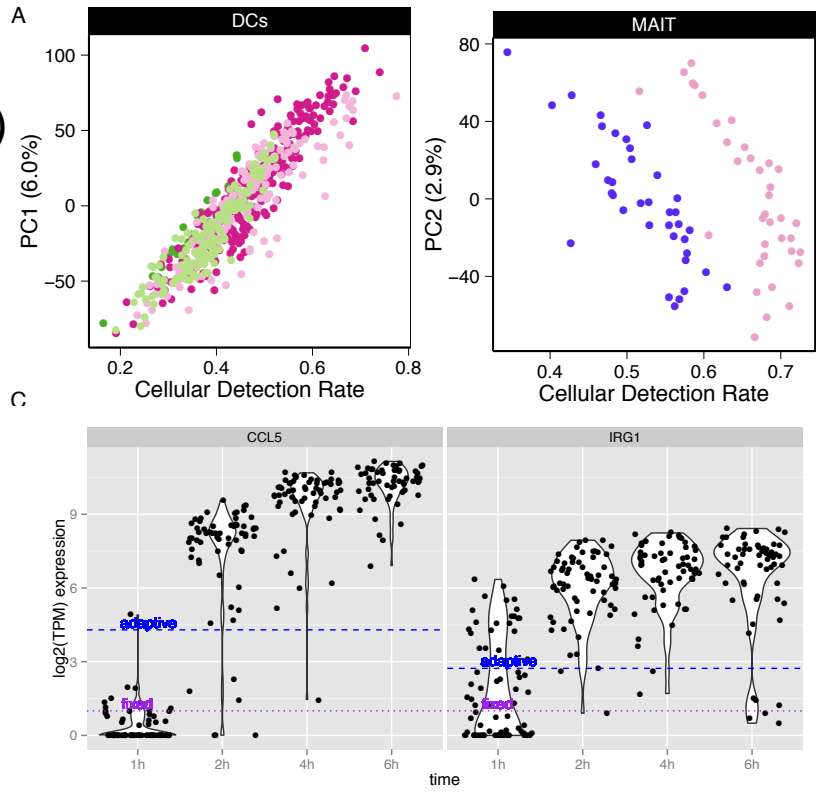
Associate Professor of Practice
Department of Mathematics and Statistics, CAS

BOSTON UNIVERSITY

Challenges of Single Cell Genomics : Biological, Technical and Statistical

- **Technical issues**
  - Unwanted cell-to-cell variability
  - Assay failure (e.g. due to cell capture, etc)
  - Batch effects
- **Bi-modality**
  - A gene can be off/on in a cell
  - Standard statistical models might not work
- **Large datasets and complex designs**
  - thousands of genes in thousands of cells with complex designs.
  - Computation matters.

CDR vs PCA



MAST: A unified computational framework
***Genome Biology* (2015)**

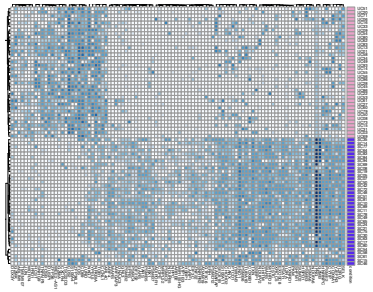
BOSTON UNIVERSITY

MAST: A unified computational framework *Genome Biology* (2015)

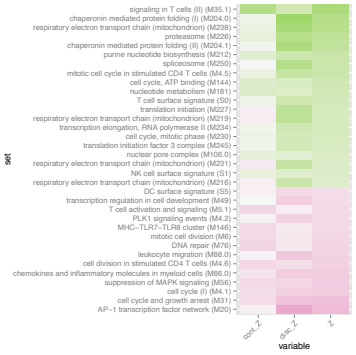
Generalized linear Model-based Analysis for Single-cell Transcriptomics

- Support for multiplexed-qPCR, NanoString, and scRNA-seq
- Thresholding and filtering methodology
- Semi-continuous model for estimation and inference
- Gene set enrichment analysis
- Implemented in Single Cell Toolkit (SCKT) (Jenkins et.al)

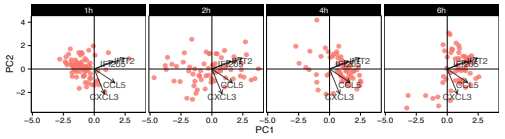
Differential Expression

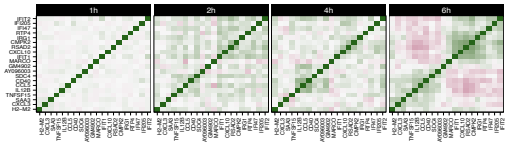


Gene set enrichment



Residual Analysis





MSSP – A Statistical Hub

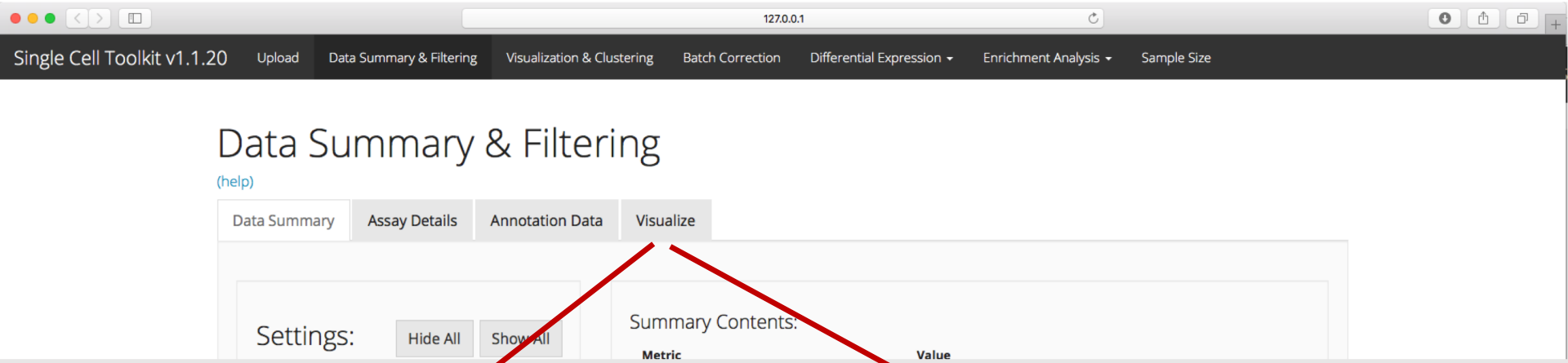
- MS in Statistical Practice (42 students)
  - Practice Centric Data Science Program
  - Statistical Consulting (**FREE**)
    1. One and done
    2. Limited Duration (10~20h student time)
    3. Collaboration
  - Supervised by a PhD student supervised by faculties
- Always looking for ways to help you with your research.

http://sites.bu.edu/mssp-consulting/

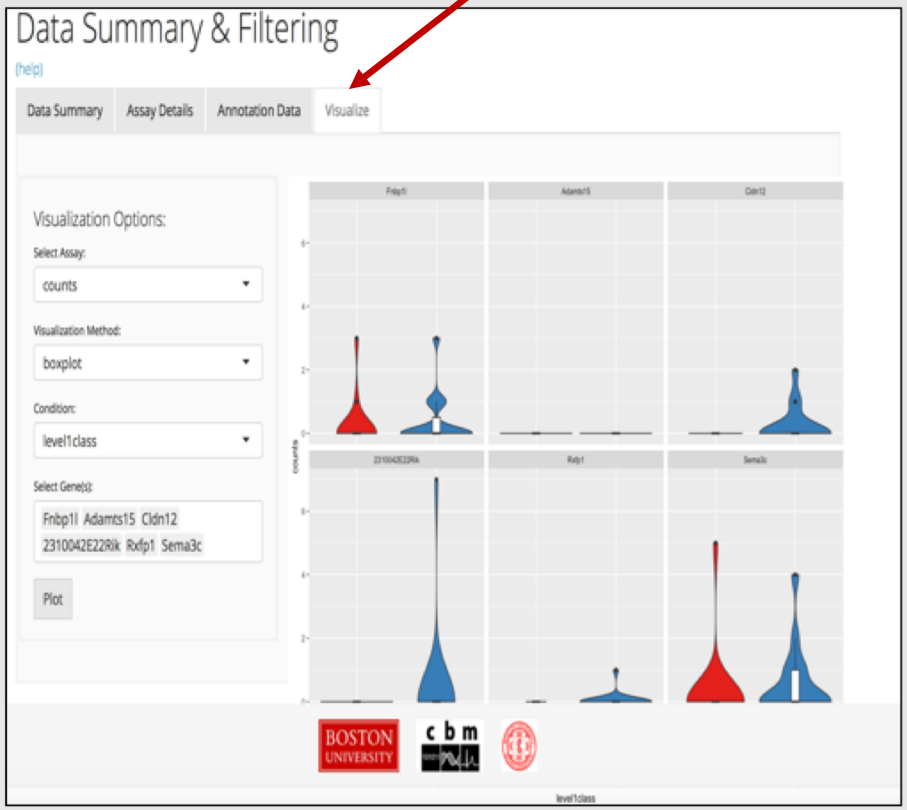
Interactive single cell RNA-Seq analysis with the Single Cell Toolkit (SCTK)

W. Evan Johnson

Associate Professor,
Division of Computational Biomedicine
Department of Medicine



- ***Visualize each gene of interest with a specific condition.***
- ***Methods such as boxplot, scatterplot, boxplot and heatmap***



Single Cell Toolkit v0.3.9 | Upload | Data Summary and Filtering | DR & Clustering | Batch Correction | Differential Expression | Pathway Activity Analysis | Sample Size

Sequencing Depth | Number of cells | Snapshot

Minimum readcount to detect gene

10

Minimum number of cells with nonzero expression to detect gene

3

Number of bootstrap iterations.

10

Maximum log10(number of simulated reads)

3 5 12

3 4 5 6 7 8 9 10 11 12

how many values to simulate

5 10 100

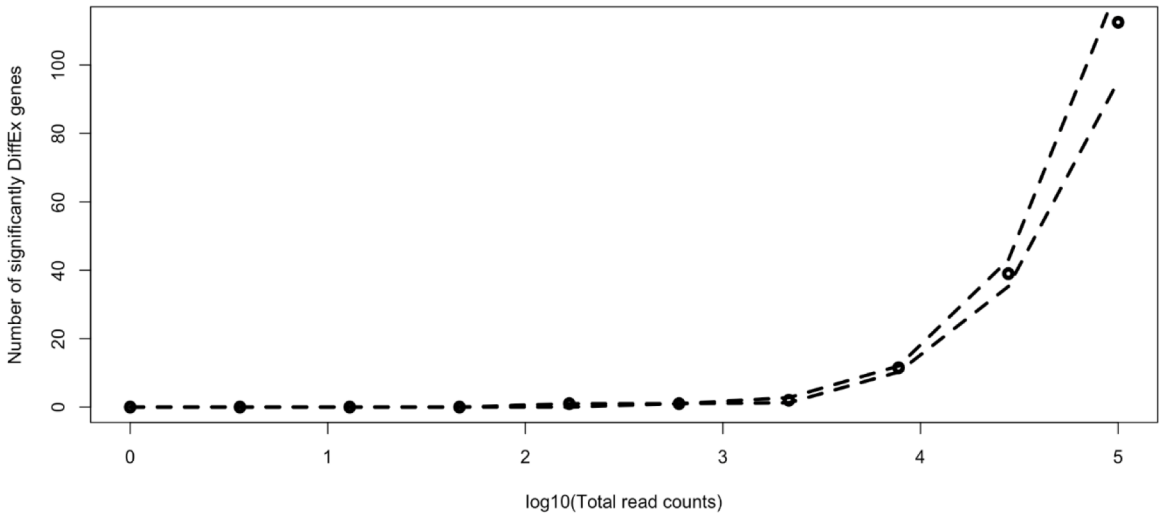
5 15 25 35 45 55 65 75 85 95 100

Condition for diffex

condition

Run subsampler

Genes Detected | Minimum Detectable Effect Size | Number of Diffex Genes



BOSTON UNIVERSITY | c b m

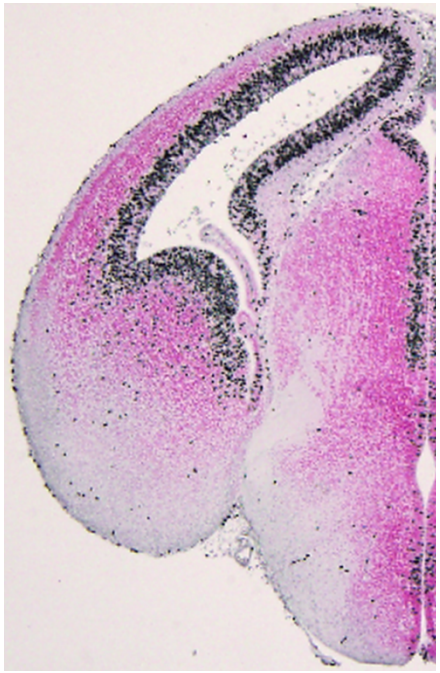
127.0.0.1:3236/#tab-8388-6

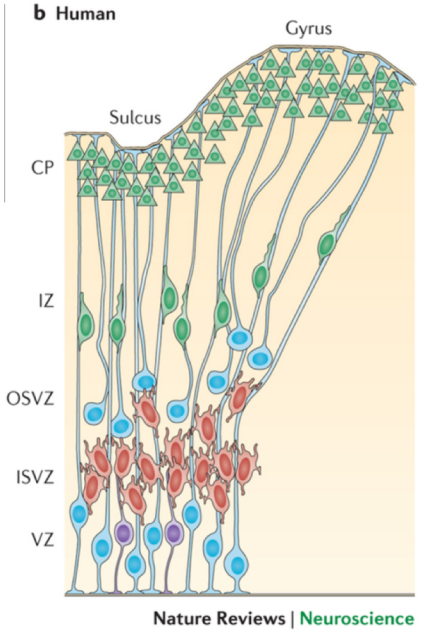
Identification of Precursor Heterogeneity in the Neocortex Using Single Cell Transcriptomics

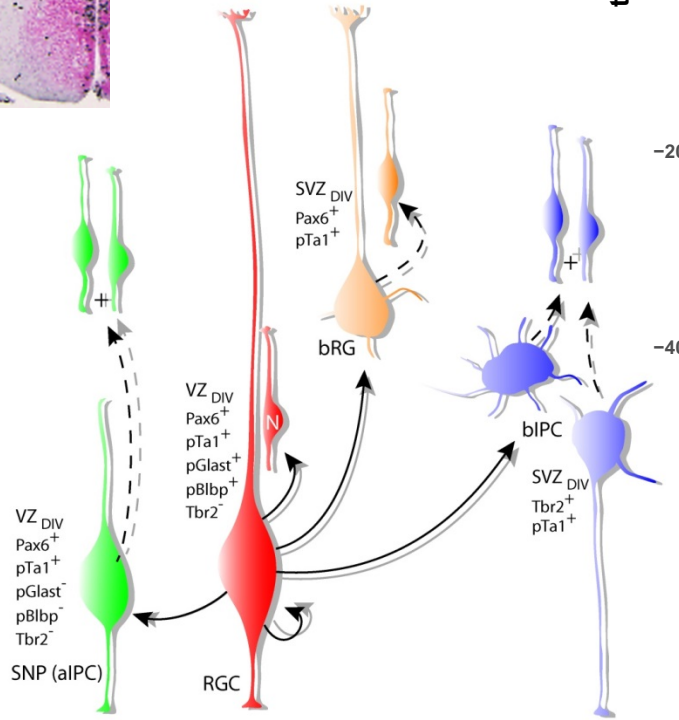
Tarik Haydar, PhD

Professor
Department of Anatomy and Neurobiology, BU-Med

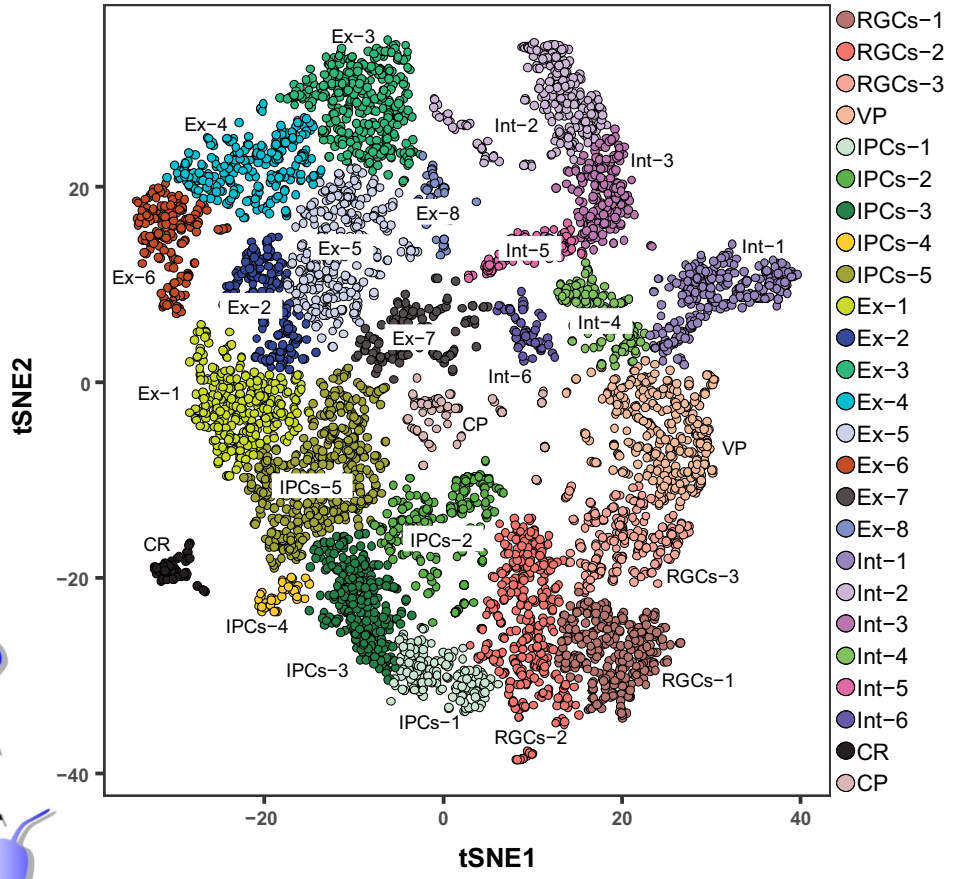
BOSTON UNIVERSITY

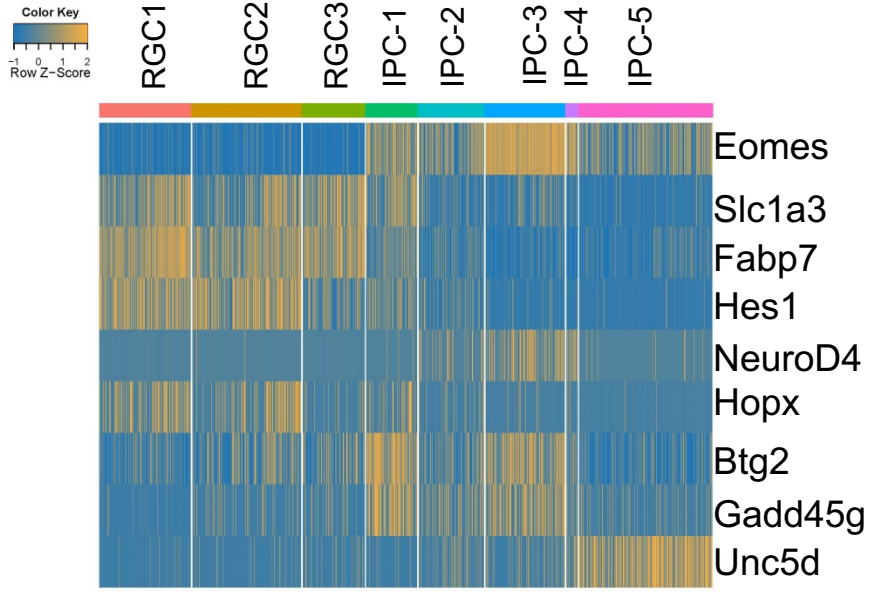


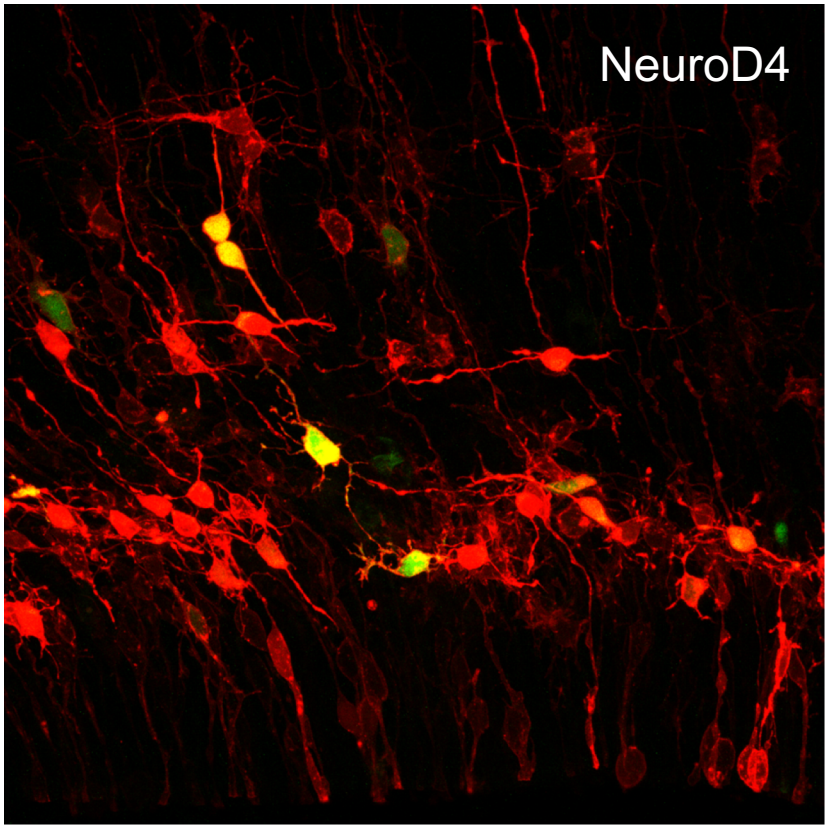




Mouse E15.5 fetal neocortex

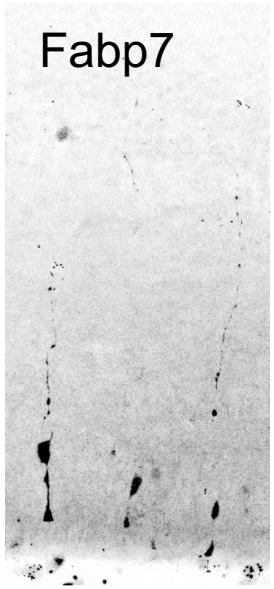


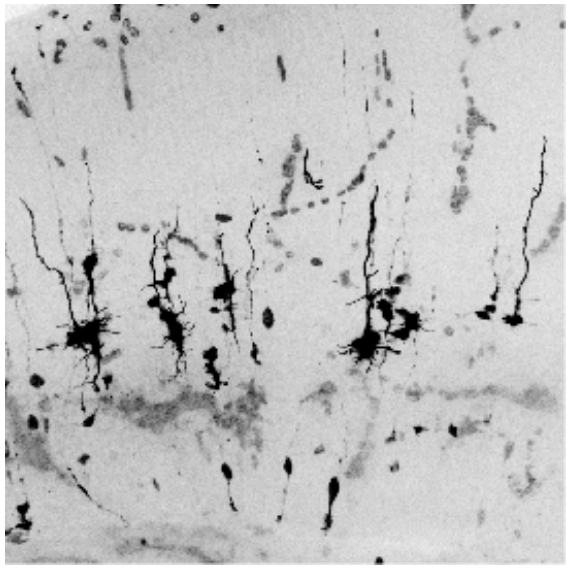


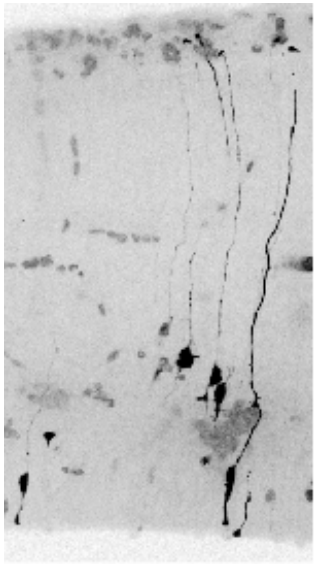


15hrs

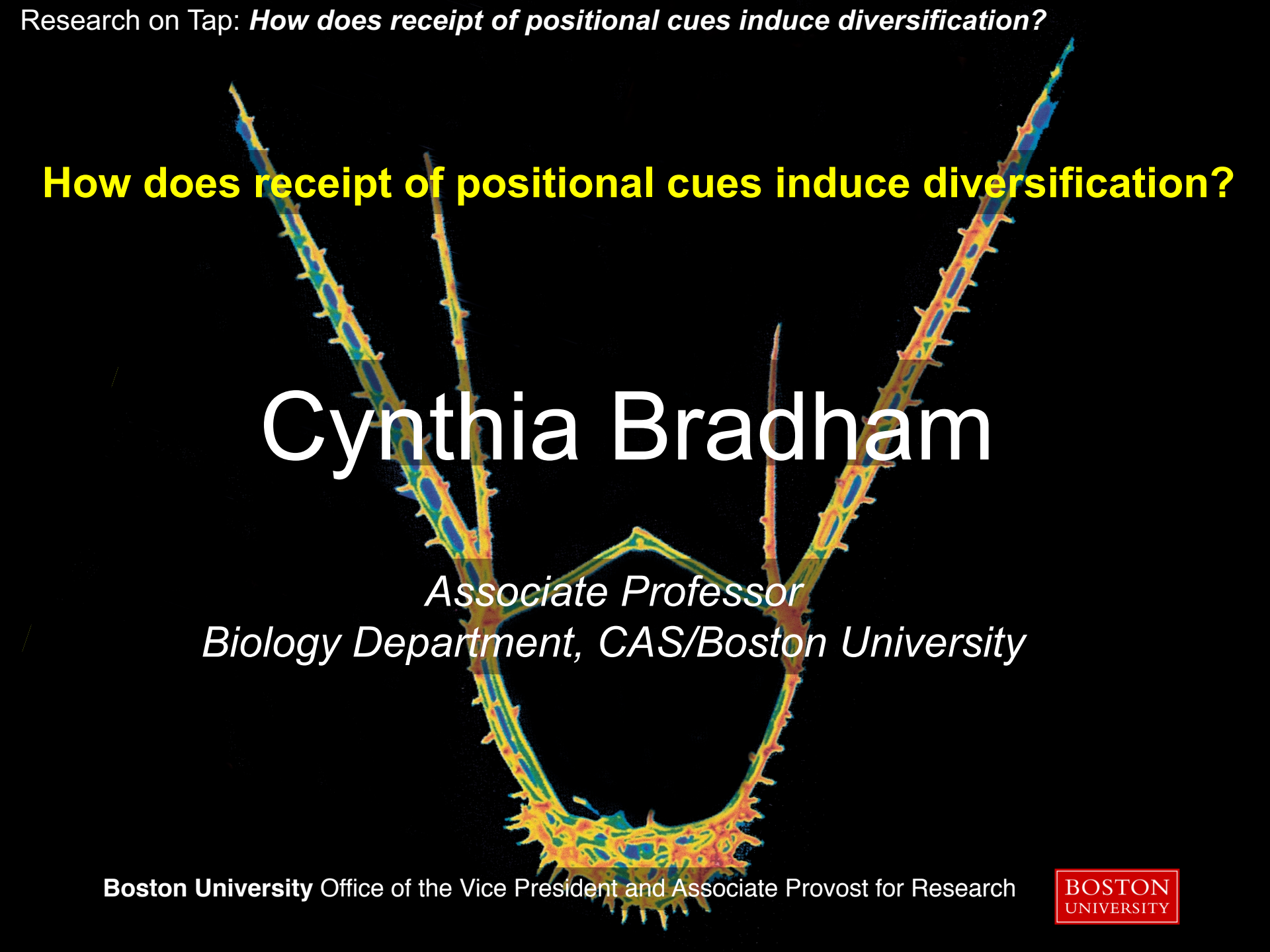
24hrs







pTbr2-Flpe
+
pFabp7-FNF-Cre
+
pCAG-LNL-GFP



How does receipt of positional cues induce diversification?

Cynthia Bradham

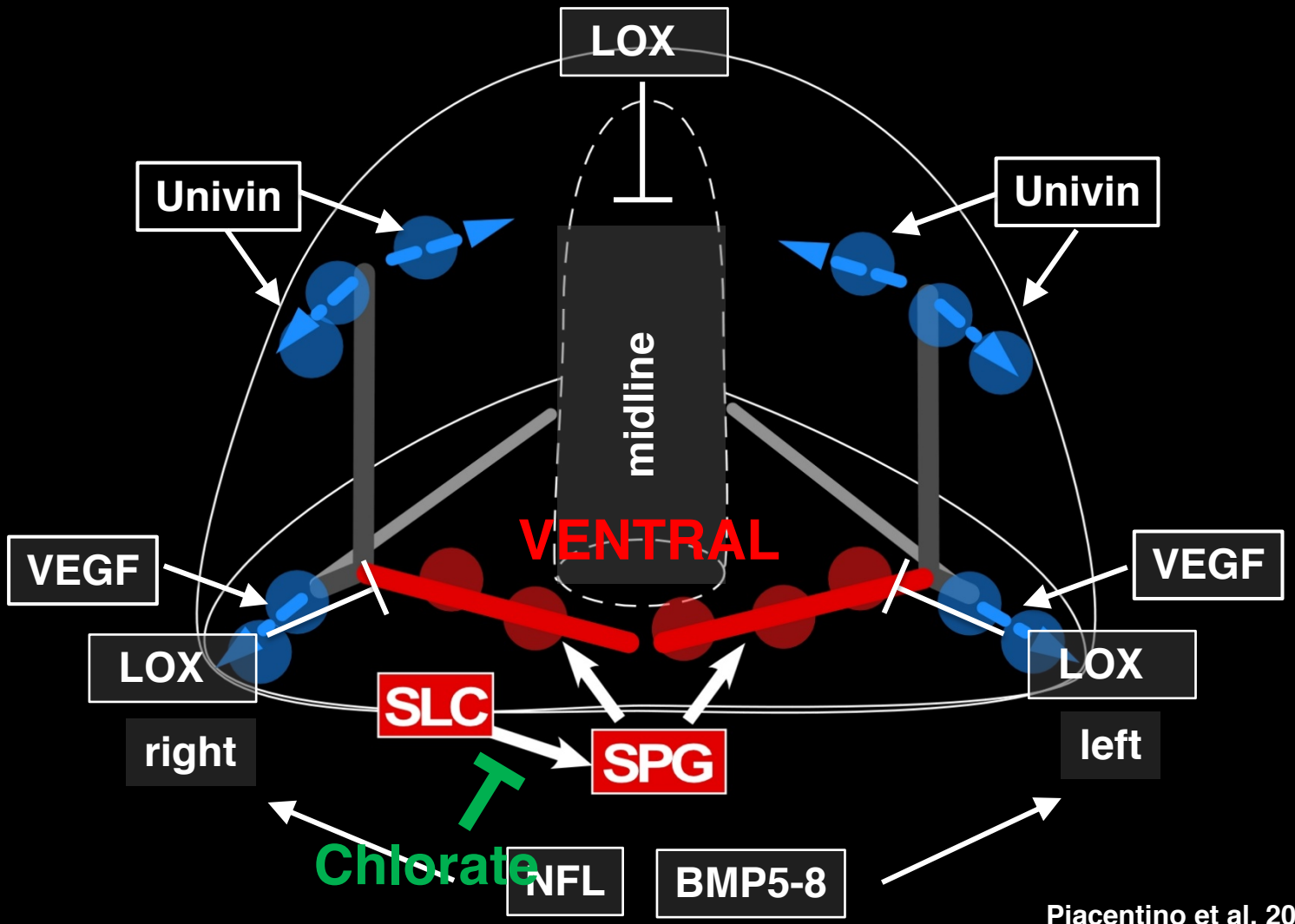
Associate Professor
Biology Department, CAS/Boston University

Boston University Office of the Vice President and Associate Provost for Research

BOSTON UNIVERSITY

We discovered multiple ectodermal cues via an RNA-seq-based screen

The Skeleton is secreted by PMCs

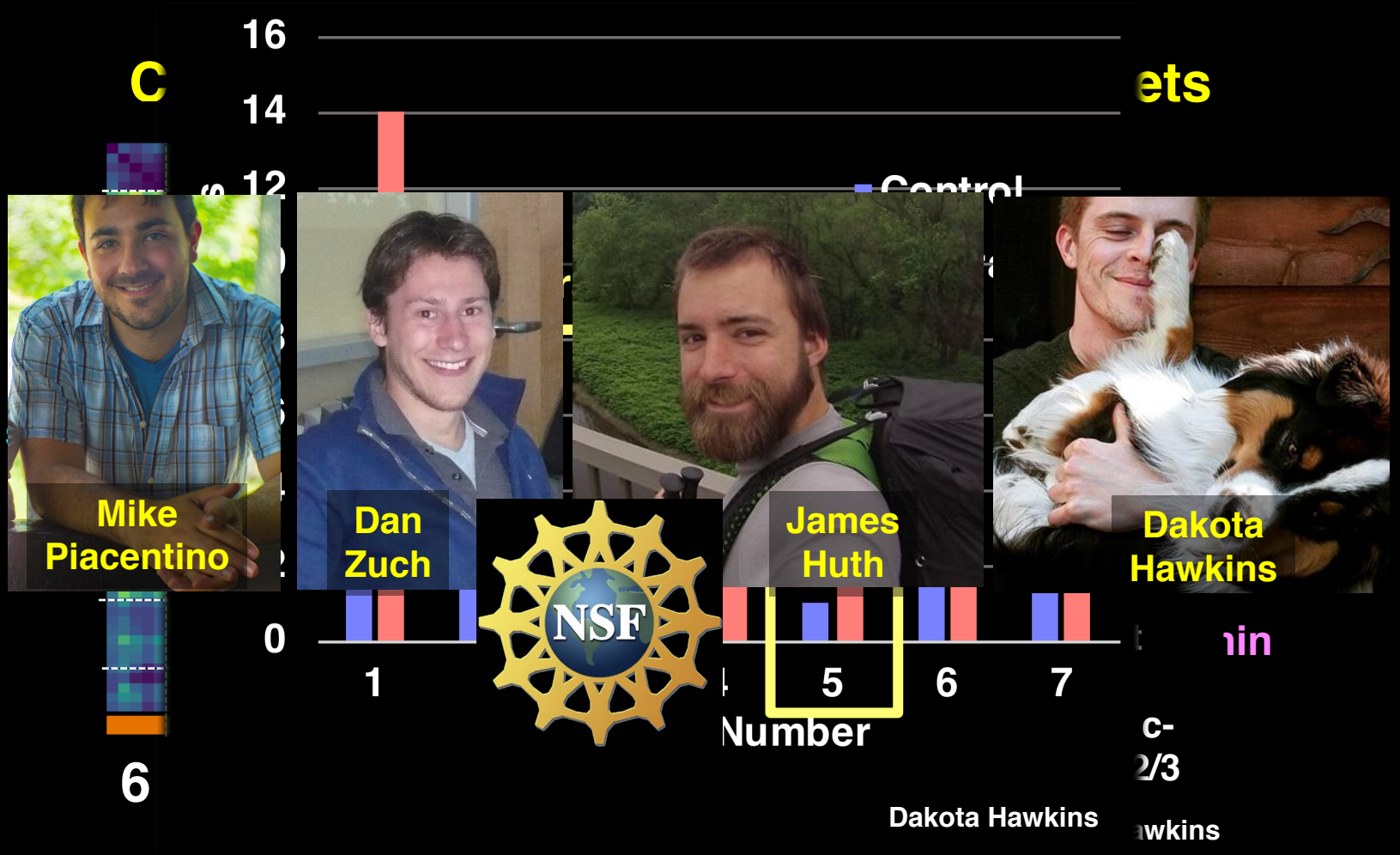


Piacentino et al. 2016 Dev. Biol.

Piacentino et al. 2016 Development

Piacentino, Ramachandran, and Bradham 2015 Development

Cue reception is required for induction of PMC subsets



What is the network for cue-mediated PMC diversification?

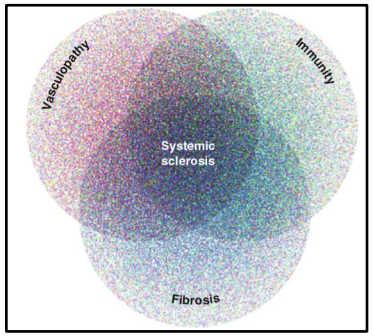
Molecular phenotyping of endothelial cells and fibroblasts in the fibrotic tissues

Maria Trojanowska, PhD

Professor
Arthritis Center/Medicine

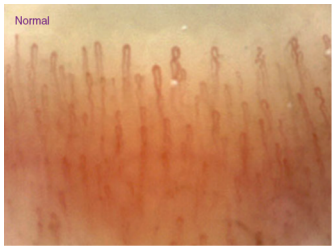
BOSTON UNIVERSITY

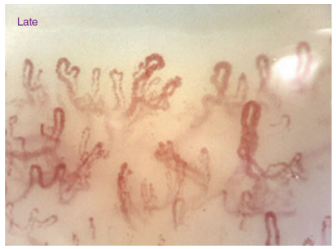
Rheumatology: M. Trojanowska, R. Simms, L. Stawski
Computational Medicine: Evan Johnson



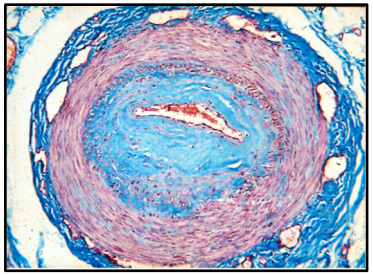
The scleroderma triad of intermediate pathophenotypes

Manifestations of vascular disease in scleroderma

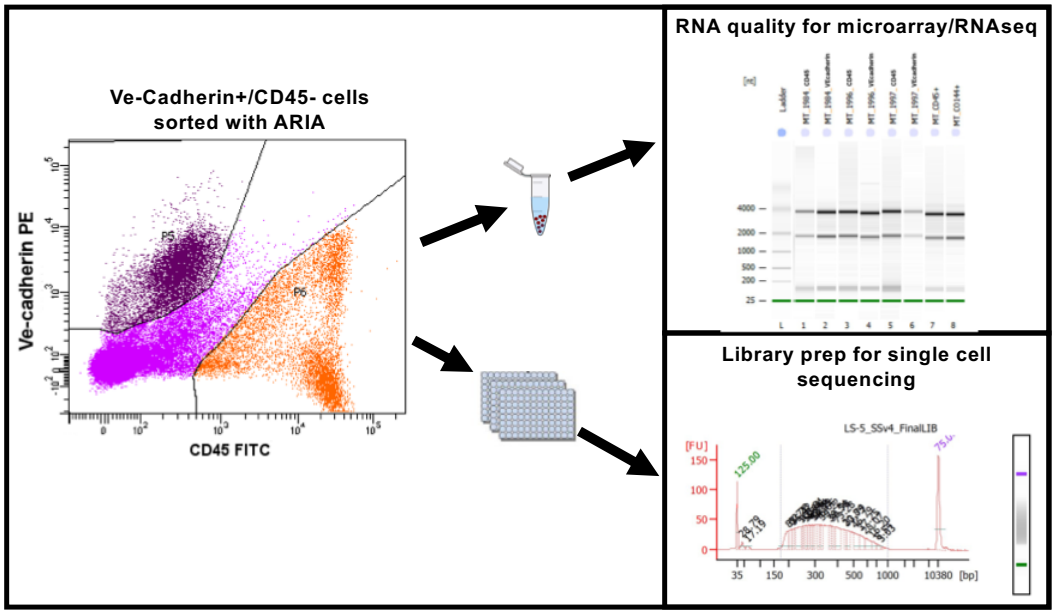




Nailfold capillaries



Occluded digital artery



HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION

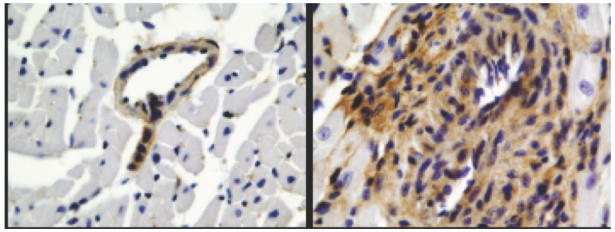
		HC1	HC2	SSc1	SSc2
SFRP4	secreted frizzled-related protein 4	1.88619046	2.49608294	4.86950151	4.37886584
CXCL12	C-X-C motif chemokine ligand 12	5.15950334	4.97193875	5.76993783	5.99452213
VCAM1	vascular cell adhesion molecule 1	2.85353298	2.55931188	3.56332813	3.89626639
ADAM12	ADAM metallopeptidase domain 12	0.34222201	0.33721108	0.93444935	0.88185846
SPP1	secreted phosphoprotein 1	1.64887725	2.8974006	6.25373394	4.75726297
CADM1	cell adhesion molecule 1	0.8044578	0.93123805	1.78734588	1.42739699
IGFBP3	insulin like growth factor binding protein 3	3.39512996	3.85548056	5.37660439	4.58603905
TNC	tenascin C	2.25770427	3.03301895	5.7374457	4.26543395
POSTN	periostin	7.4116745	4.27434881	9.3419971	9.83264105
CDH11	cadherin 11	4.21114805	3.58845274	4.81902883	4.63696323
IGFBP2	insulin like growth factor binding protein 2	1.4172434	1.17192003	1.83500454	2.46775834
SCG2	secretogranin II	0.66848945	0.41519596	0.98612653	1.40574313
FAS	Fas cell surface death receptor	4.36620056	3.822436	4.68753486	4.98209428
NTM	neurotrimin 2.892535	1.79562491	1.58247125	2.29585545	2.04108564
FGF2	fibroblast growth factor 2 (basic)	4.64249067	4.44607799	6.27059248	5.06218279
ACTA2	actin, alpha 2, smooth muscle, aorta	4.25165869	3.28369884	5.14545828	4.4785878
TIMP1	TIMP metallopeptidase inhibitor 1	8.70369254	8.62581803	9.30833448	9.00414864
COL6A3	collagen type VI alpha 3 chain	4.70340404	4.15270593	4.95181376	6.14089082
CTHRC1	collagen triple helix repeat containing 1	4.83892116	4.23735161	5.04052281	5.16975902
TPM2	tropomyosin 2 (beta)	5.70441195	5.44340121	7.41761541	6.04154074
NID2	nidogen 2 (osteonidogen)	2.65198552	2.55478584	2.83485827	3.34573104
COL5A2	collagen, type V, alpha 2	1.03889487	1.14042395	1.26406247	1.6592996
MYLK	myosin light chain kinase	2.29004875	1.76381163	3.0589322	2.31304163
COL6A2	collagen, type VI, alpha 2	6.42827388	5.85159098	6.50367918	6.64549482

BOSTON UNIVERSITY

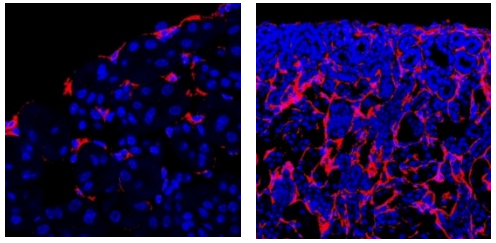
Fibrosis ARC: X. Varelas, W. Lu, M. Layne, M. Trojanowska, A. Bujor, J. Browning, K. Ravid, M. Kukuruzinska, P. Trackman, S. Monti

Purpose: To identify common and tissue specific factors in organ fibrosis

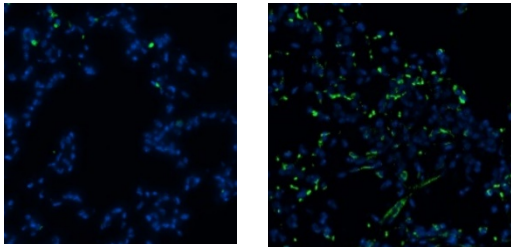
PDGFRβ+ cells are increased in various fibrotic tissues



heart



kidney



lungs

Fibrotic organs
Skin
Kidney
Lung
Heart
Adipose Tissue
Bone marrow

Isolate PDGRβ+ cells

Determine transcriptional profile of PDGFRβ+ profibrotic cells by single cell RNAseq

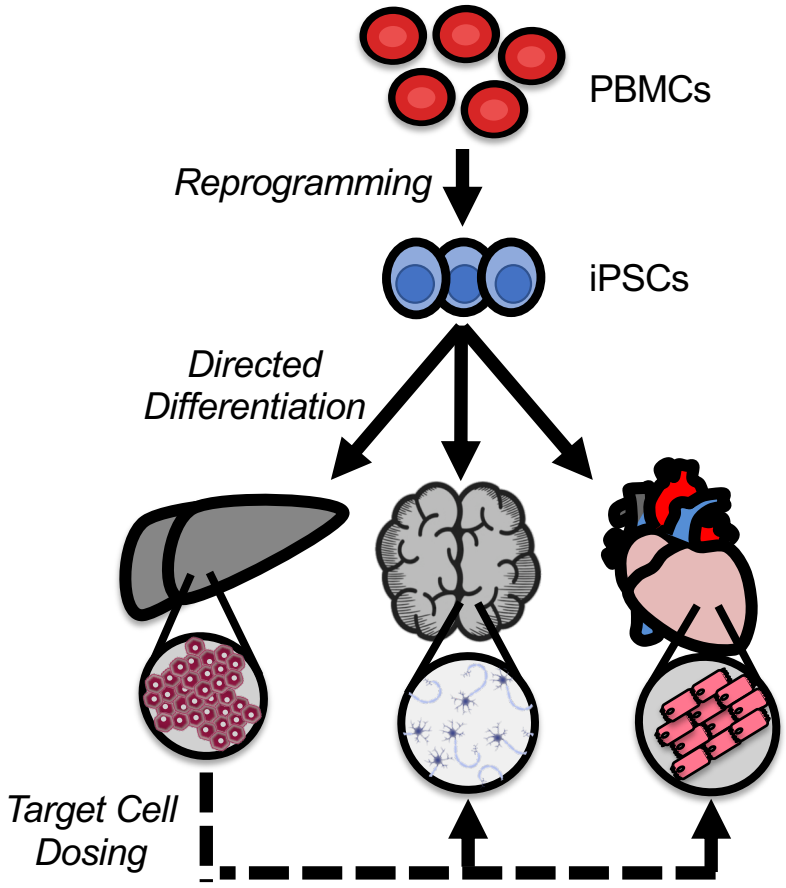
BOSTON UNIVERSITY

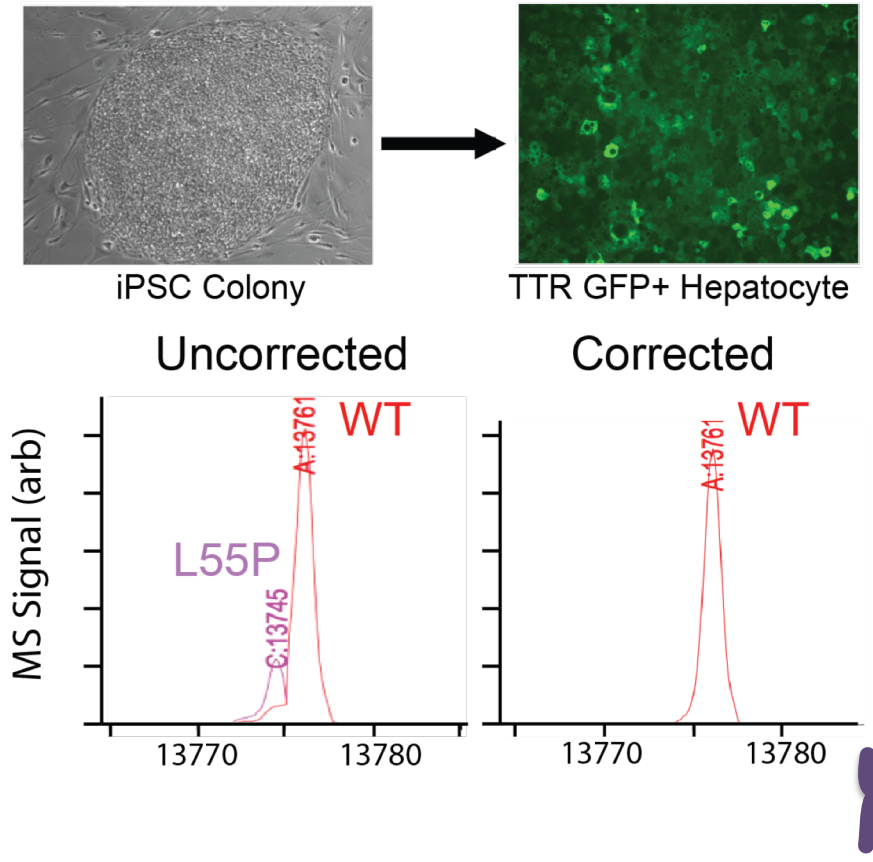
Single Cell RNA Sequencing (scRNAseq) Reveals A Novel Hepatic Disease Signature In ATTR Amyloidosis

George J. Murphy, PhD

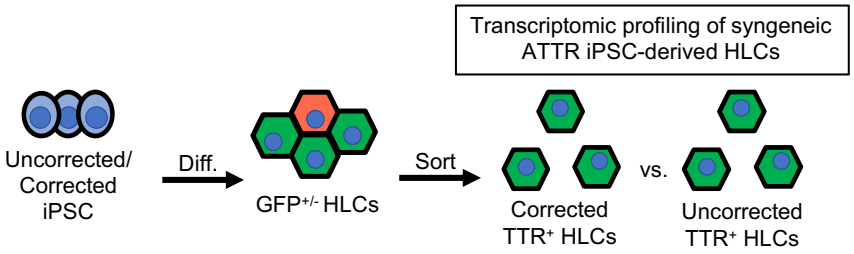
Associate Professor
Department of Medicine, Division of Hematology-Oncology
Co-Director BU and BMC Center for Regenerative Medicine (CReM)

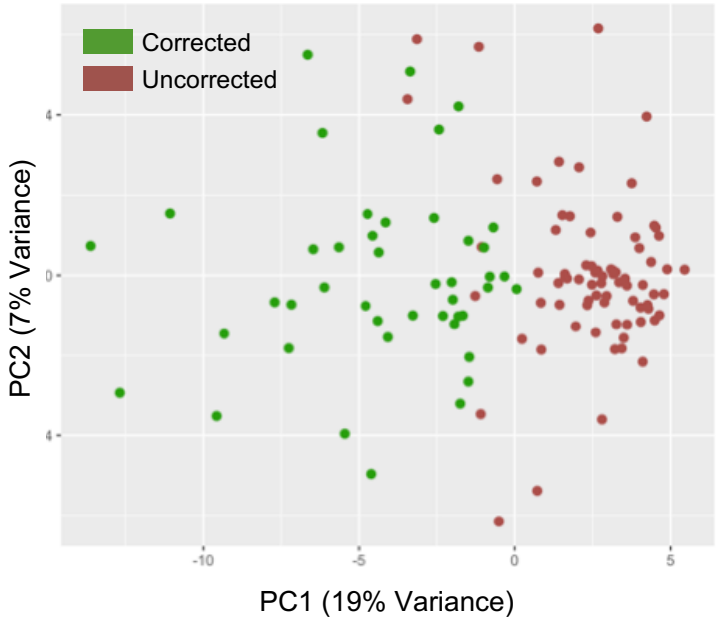
Single cell RNA sequencing (scRNAseq) of corrected vs. uncorrected syngeneic iPSC-derived HLCs reveals a novel hepatic disease signature.

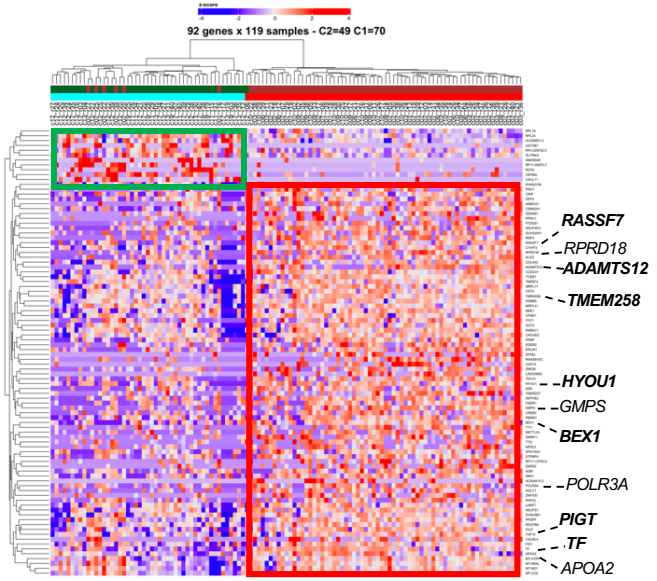




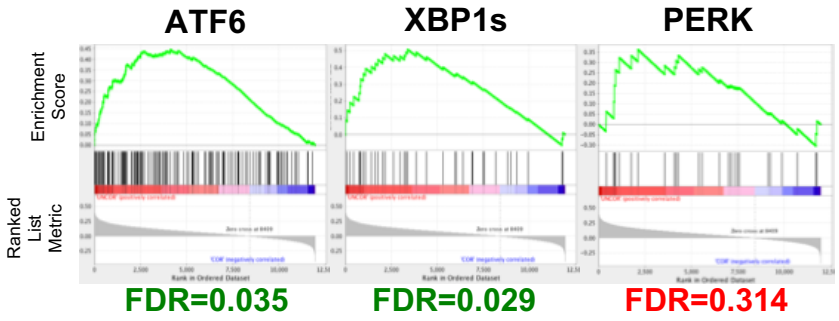
Single cell RNA sequencing (scRNAseq) of corrected vs. uncorrected syngeneic iPSC-derived HLCs reveals a novel hepatic disease signature.







Corrected/Uncorrected



FDR=0.035 FDR=0.029 FDR=0.314

BOSTON UNIVERSITY