Single-cell lineage tracing with a focus on cancer metastasis using macsGESTALT

Pinello Lab Journal Club

Friday | June 25th, 2021

Overview of today's presentation

Paper we are covering today:

Cancer Cell

Article

Single-cell lineage tracing of metastatic cancer reveals selection of hybrid EMT states

First we will cover some history, background, motivation, and context:

RESEARCH ARTICLE

CELL LINEAGE TRACING

Whole-organism lineage tracing by combinatorial and cumulative genome editing

Aaron McKenna,[1]* Gregory M. Findlay,[1]* James A. Gagnon,[2]* Marshall S. Horw
Alexander F. Schier,[2,4,5,6]† Jay Shendure[1,7]†

Developmental Cell

CellPress

Review

Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development

Science

RESEARCH ARTICLES

Cite as: Weinreb *et al.*, *Science* 10.1126/science.aaw3381 (2020).

Lineage tracing on transcriptional landscapes links state to fate during differentiation

riguez-Fraticelli[2,3]*, Fernando D. Camargo[2,3]†, Allon M. Klein[1]†‡

SINGLE-CELL OMICS

Building a lineage from single cells: genetic techniques for cell lineage tracking

Mollie B. Woodworth[1–3], Kelly M. Girskis[1–3] and Christopher A. Walsh[1–3]

Check for updates

Lineage tracing meets single-cell omics: opportunities and challenges

Daniel E. Wagner[1,2] and Allon M. Klein[1]

Article

Lineage Tracing in Humans Enabled by Mitochondrial Mutations and Single-Cell Genomics

Leif S. Ludwig,[1,2,15,*] Caleb A. Lareau,[1,2,3,4,15] Jacob C. Ulirsch,[1,2,4,15] Elena Christ
Karin Pelka,[1,6,7] Will Ge,[1] Yaara Oren,[1,8] Alison Brack,[1] Travis Law,[1] Christopher
Genevieve M. Boland,[6,10] Nir Hacohen,[1,6,7] Orit Rozenblatt-Rosen,[1] Martin J. Ary
Aviv Regev,[1,13,*] and Vijay G. Sankaran[1,2,14,16,*]

Cell

ARTICLE

https://doi.org/10.1038/s41586-018-0744-4

Single-cell mapping of lineage and identity in direct reprogramming

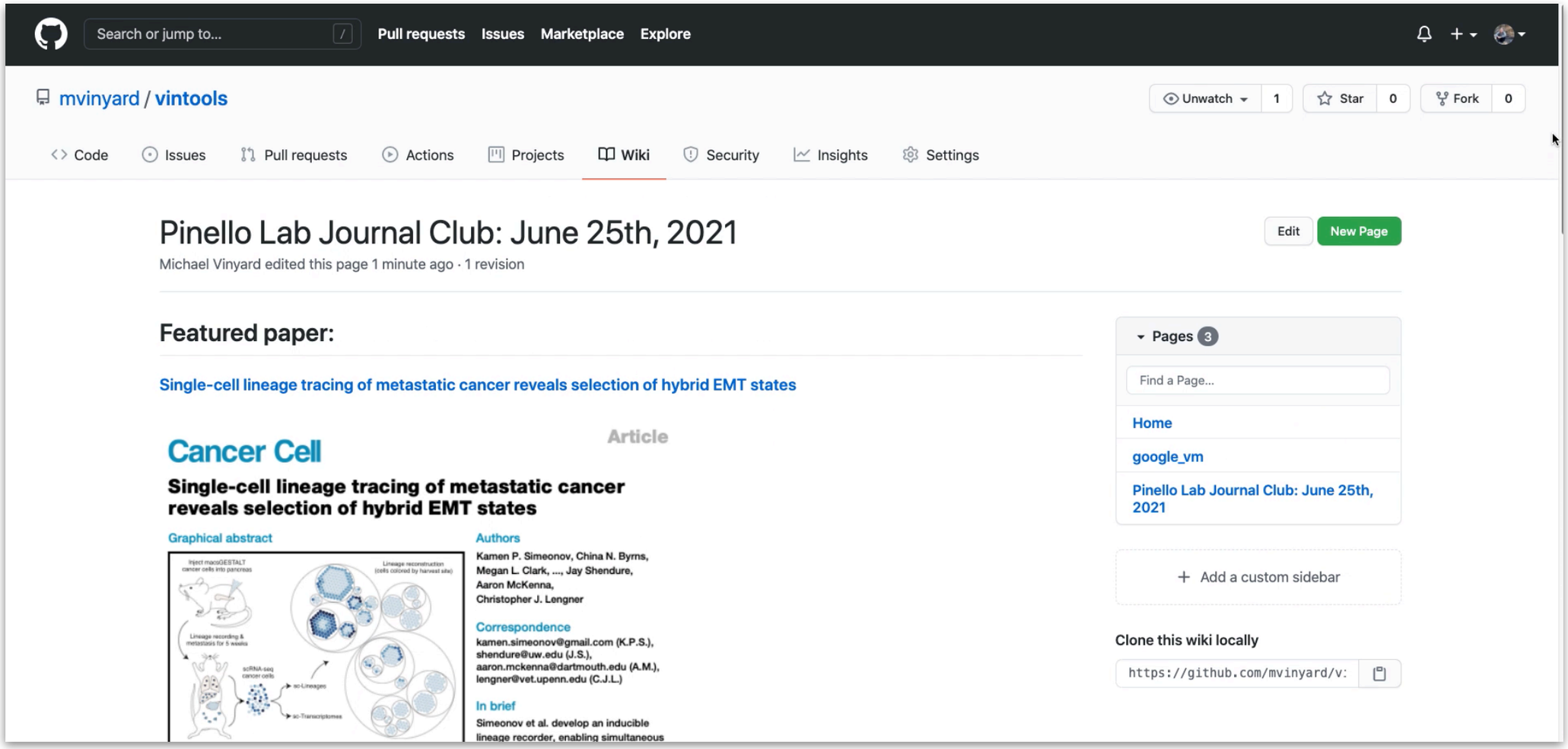
Brent A. Biddy[1,2,3], Wenjun Kong[1,2,3], Kenji Kamimoto[1,2,3], Chuner Guo[1,2,3], Sarah E. Waye[1,2,3], Tao Sun[1,2,3,4] & Samantha A. Morris[1,2,3]*

age reconstruction from clonal correlations

inreb[a] and Allon M. Klein[a,1]

Overview of today's presentation

Access all sides with links to papers and notes:



GitHub.com/mvinyard/vintools

Overview of today's presentation

1. Brief history lesson and overview of tools / datasets / technologies available

2. Overview of GESTALT, the precursor technology to macsGESTALT

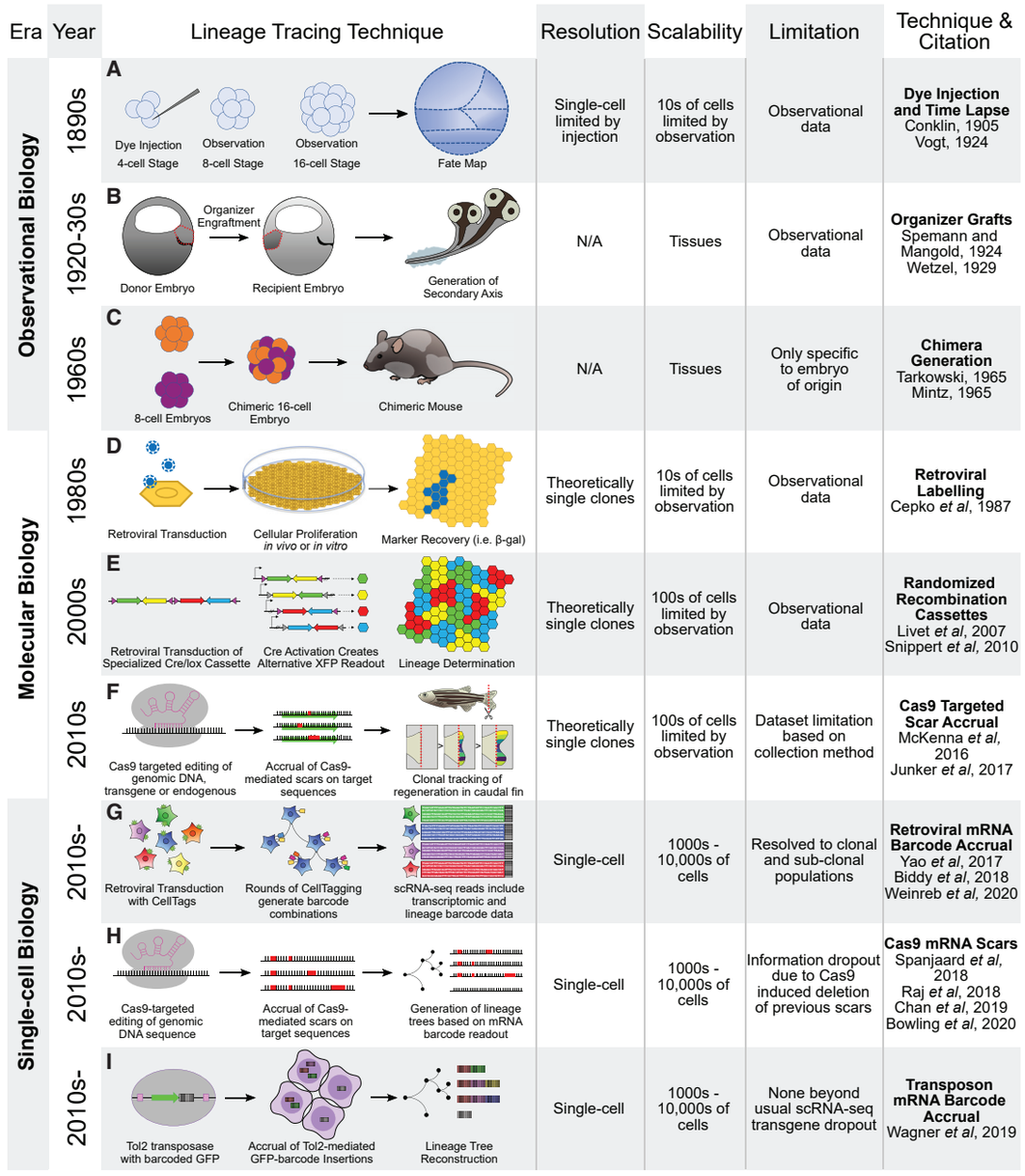
3. In-depth coverage of the macsGESTALT paper

Brief history of fate mapping and lineage tracing

Fate Map

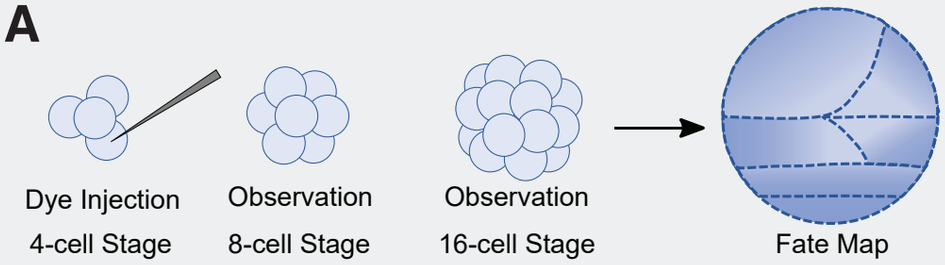
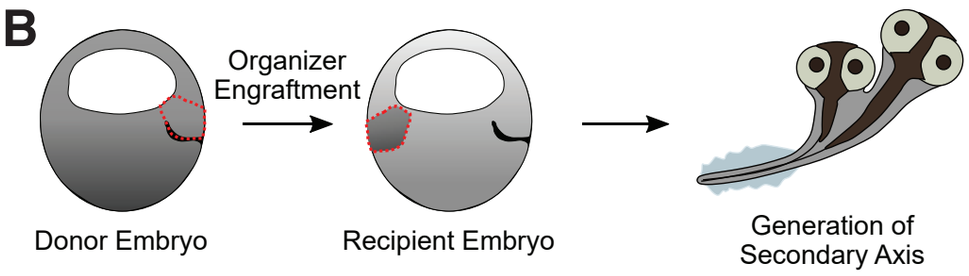
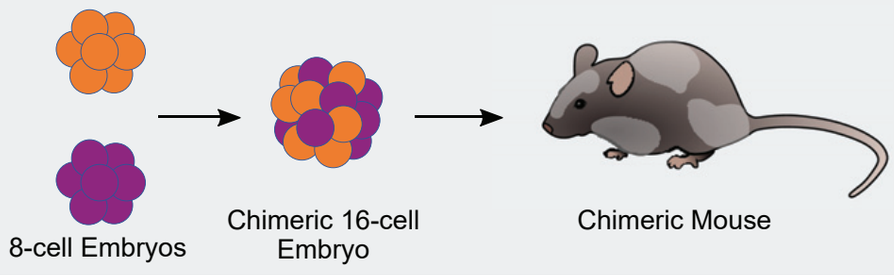
Fate maps = *schematics* of developmental potential

- Fate mapping and lineage tracing are related but **distinct**
- Lineage tracing identifies progeny from a given ancestor cell



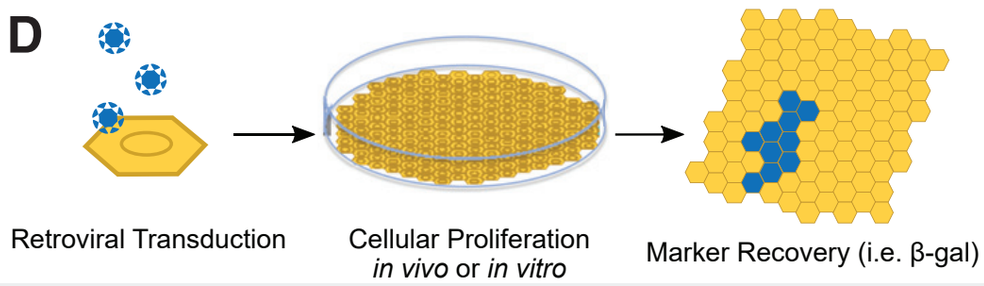
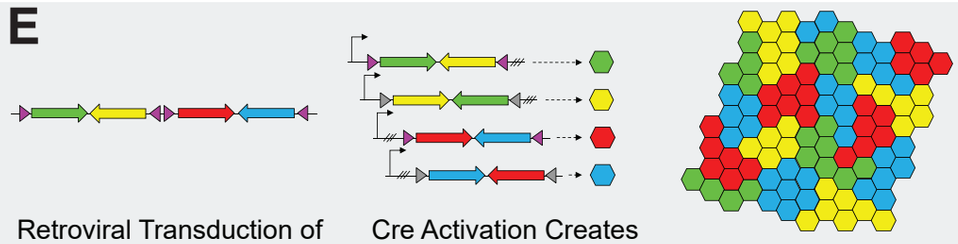
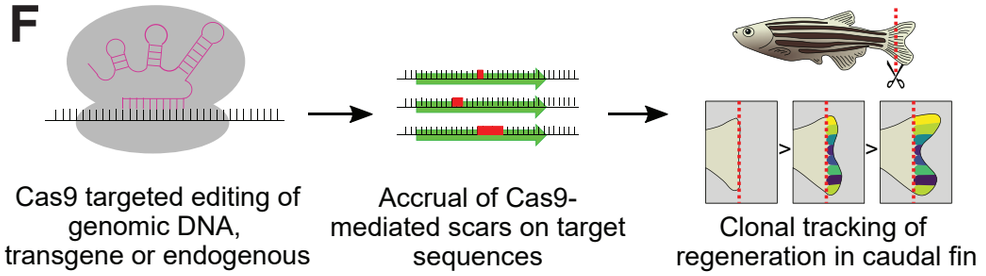
Figures adapted from Figures 1, 2 of VanHorn et al., Dev Cell (2020)

Fate mapping and lineage tracing

Era	Year	Lineage Tracing Technique	Resolution	Scalability	Limitation	Technique & Citation
Observational Biology	1890s	**A**  	Single-cell limited by injection	10s of cells limited by observation	Observational data	**Dye Injection and Time Lapse** Conklin, 1905 Vogt, 1924
	1920-30s	**B**  	N/A	Tissues	Observational data	**Organizer Grafts** Spemann and Mangold, 1924 Wetzel, 1929
	1960s	**C**  	N/A	Tissues	Only specific to embryo of origin	**Chimera Generation** Tarkowski, 1965 Mintz, 1965

From Figure 2 of VanHorn et al., Dev Cell (2020)

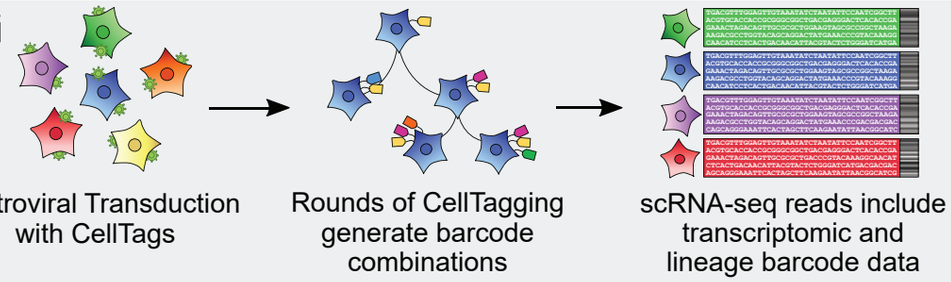
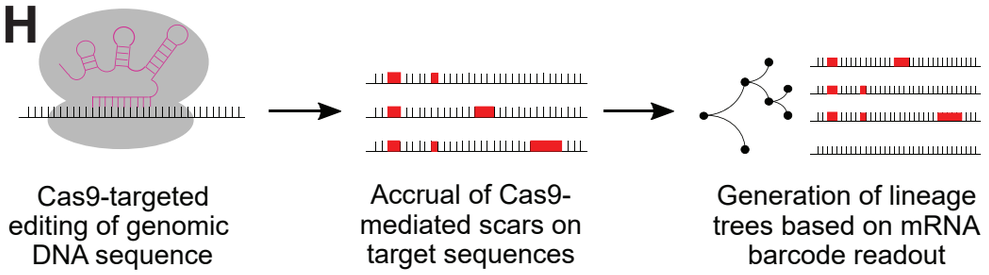
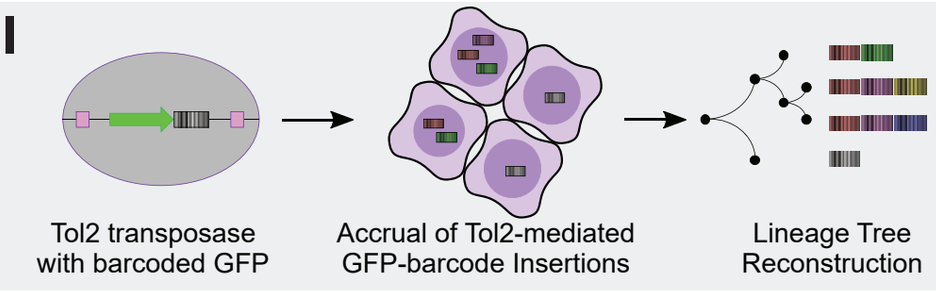
Fate mapping and lineage tracing

Era	Year	Lineage Tracing Technique	Resolution	Scalability	Limitation	Technique & Citation
Molecular Biology	1980s		Theoretically single clones	10s of cells limited by observation	Observational data	**Retroviral Labelling** Cepko *et al*, 1987
	2000s		Theoretically single clones	100s of cells limited by observation	Observational data	**Randomized Recombination Cassettes** Livet *et al*, 2007 Snippert *et al,* 2010
	2010s		Theoretically single clones	100s of cells limited by observation	Dataset limitation based on collection method	**Cas9 Targeted Scar Accrual** McKenna *et al,* 2016 Junker *et al*, 2017

Still pre-“single-cell revolution” takeover

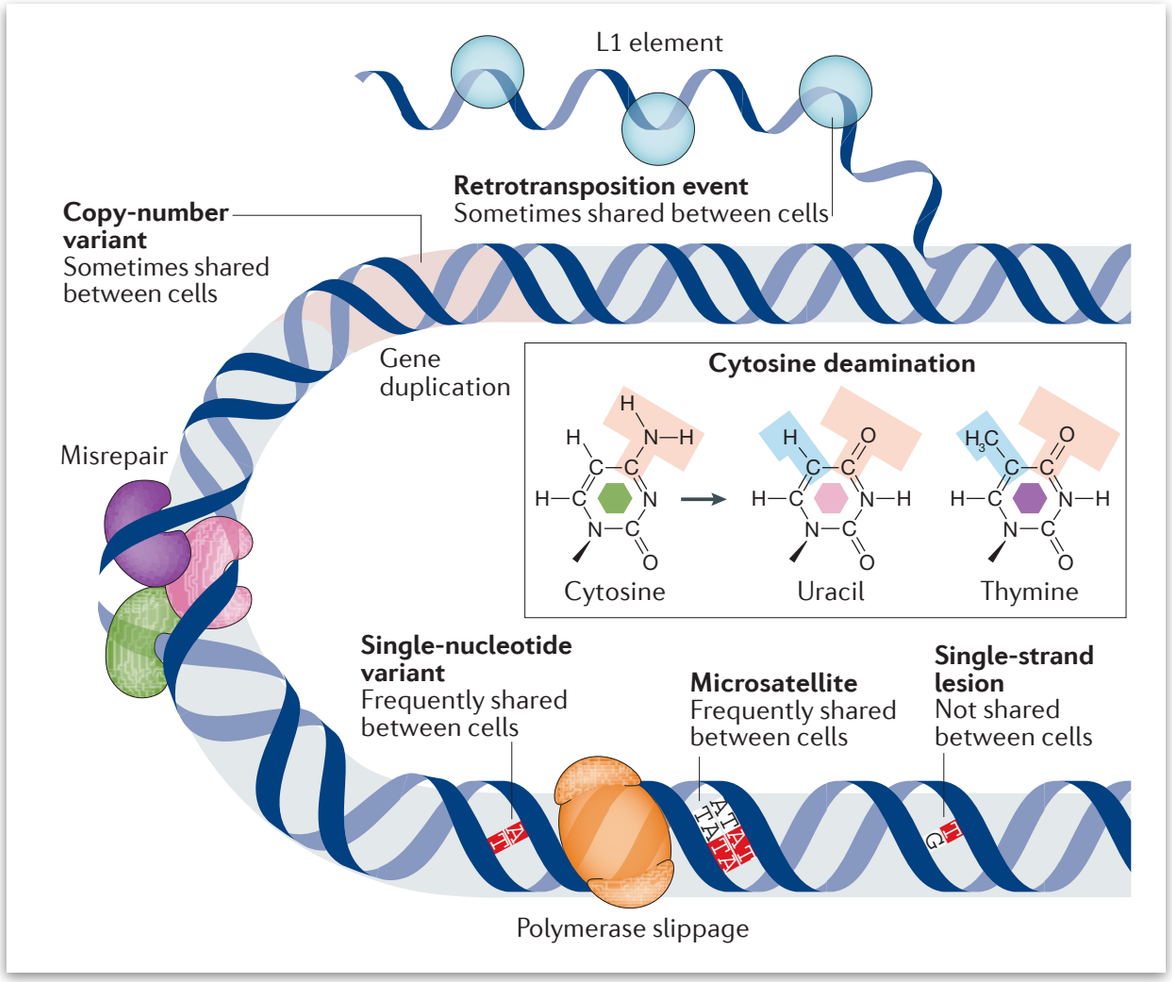
From Figure 2 of VanHorn et al., Dev Cell (2020)

Fate mapping and lineage tracing

Era	Year	Lineage Tracing Technique	Resolution	Scalability	Limitation	Technique & Citation
Single-cell Biology	2010s-	G 	Single-cell	1000s - 10,000s of cells	Resolved to clonal and sub-clonal populations	**Retroviral mRNA Barcode Accrual** Yao *et al*, 2017 Biddy *et al*, 2018 Weinreb *et al,* 2020
	2010s-	H 	Single-cell	1000s - 10,000s of cells	Information dropout due to Cas9 induced deletion of previous scars	**Cas9 mRNA Scars** Spanjaard *et al,* 2018 Raj *et al*, 2018 Chan *et al*, 2019 Bowling *et al*, 2020
	2010s-	I 	Single-cell	1000s - 10,000s of cells	None beyond usual scRNA-seq transgene dropout	**Transposon mRNA Barcode Accrual** Wagner *et al*, 2019

From Figure 2 of VanHorn et al., Dev Cell (2020)

Using somatic mutations



From Figure 4 of Woodworth et al., Nat Rev Gen (2017)

Advantage: already “in the data” or “free"

Two main **limitations** of using somatic variation in lineage tracing:

1. **WGS** required ($$$, unscalable)
2. Inherent read **sparsity** of scRNA-seq

Using mtDNA mutations

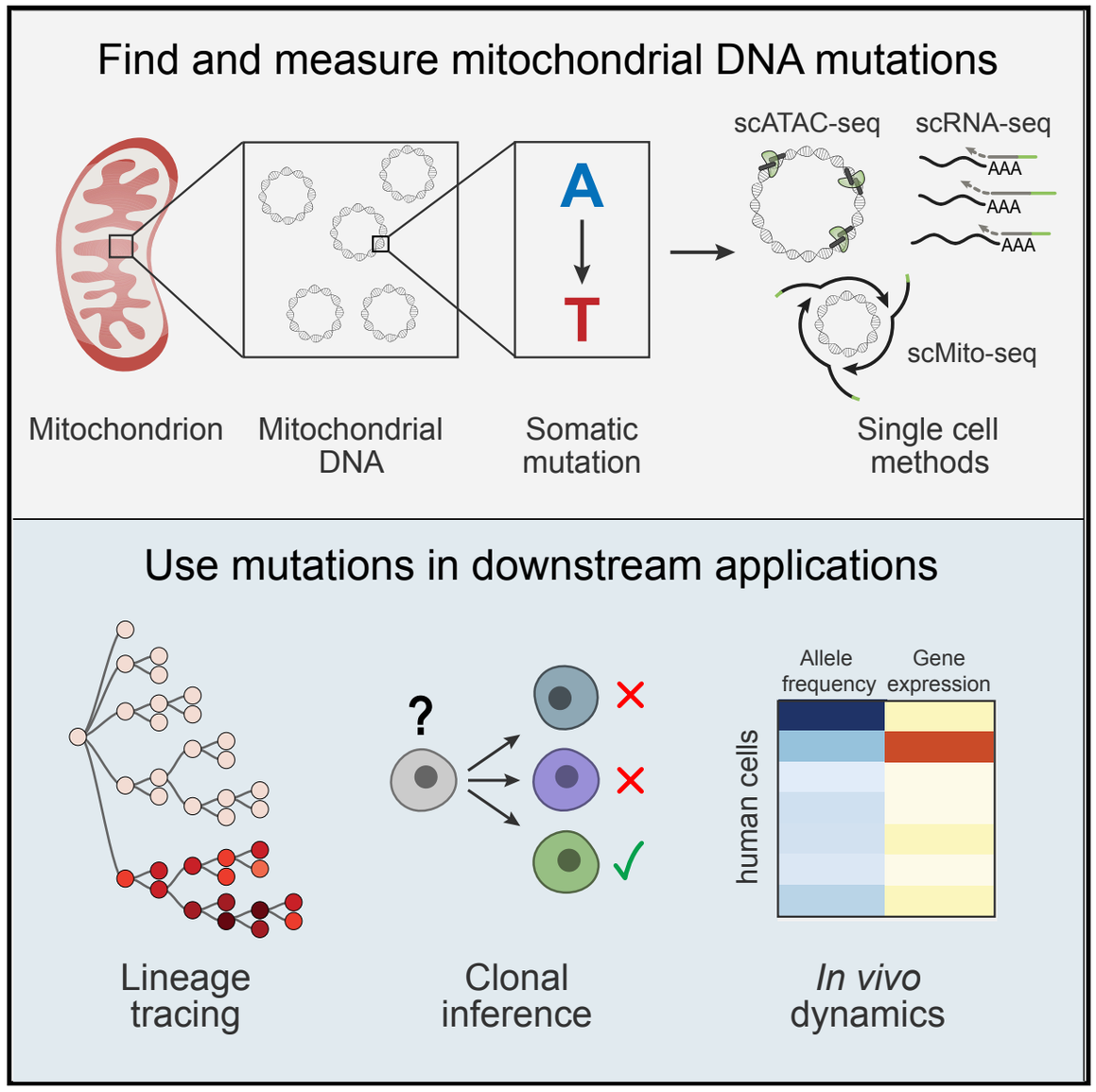

Article

Cell

Lineage Tracing in Humans Enabled by Mitochondrial Mutations and Single-Cell Genomics

Leif S. Ludwig,[1,2,15,*] Caleb A. Lareau,[1,2,3,4,15] Jacob C. Ulirsch,[1,2,4,15] Elena Christian,[1] Christoph Muus,[1,5] Lauren H. Li,[1,2] Karin Pelka,[1,6,7] Will Ge,[1] Yaara Oren,[1,8] Alison Brack,[1] Travis Law,[1] Christopher Rodman,[1] Jonathan H. Chen,[1,9] Genevieve M. Boland,[6,10] Nir Hacohen,[1,6,7] Orit Rozenblatt-Rosen,[1] Martin J. Aryee,[1,3,11] Jason D. Buenrostro,[1,12] Aviv Regev,[1,13,*] and Vijay G. Sankaran[1,2,14,16,*]

Graphical Abstract



From Graphical Abstract of Ludwig et al., Cell (2021)

Advantages:

1. Again, “free”
2. mtDNA mutation rates > gDNA mutation rates
3. Readily paired GEX / CA with clonal lineage information
4. No WGS or fancy barcoding required

Limitations:

1. Current approaches offer limited coverage of mito genome
2. Potential for horizontal gene transfer in mito genome (unclear to what extent)

GESTALT (**g**enome **e**diting of **s**ynthetic **t**arget **a**rrays for **l**ineage **t**racing)

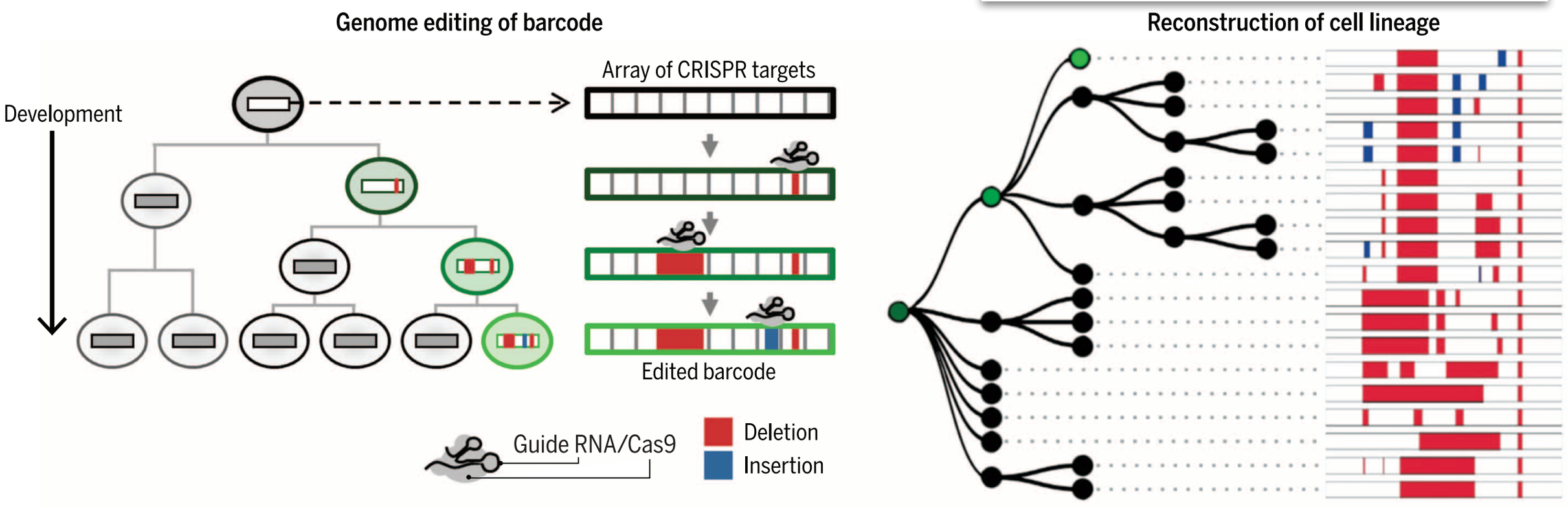
- This paper sets a precedent for large-scale lineage tracing - how? Tracing must…
  1. Impart unique marks over division
  2. Marks must accumulate over time
  3. Easy single-cell readout

RESEARCH ARTICLE

CELL LINEAGE TRACING

Whole-organism lineage tracing by combinatorial and cumulative genome editing

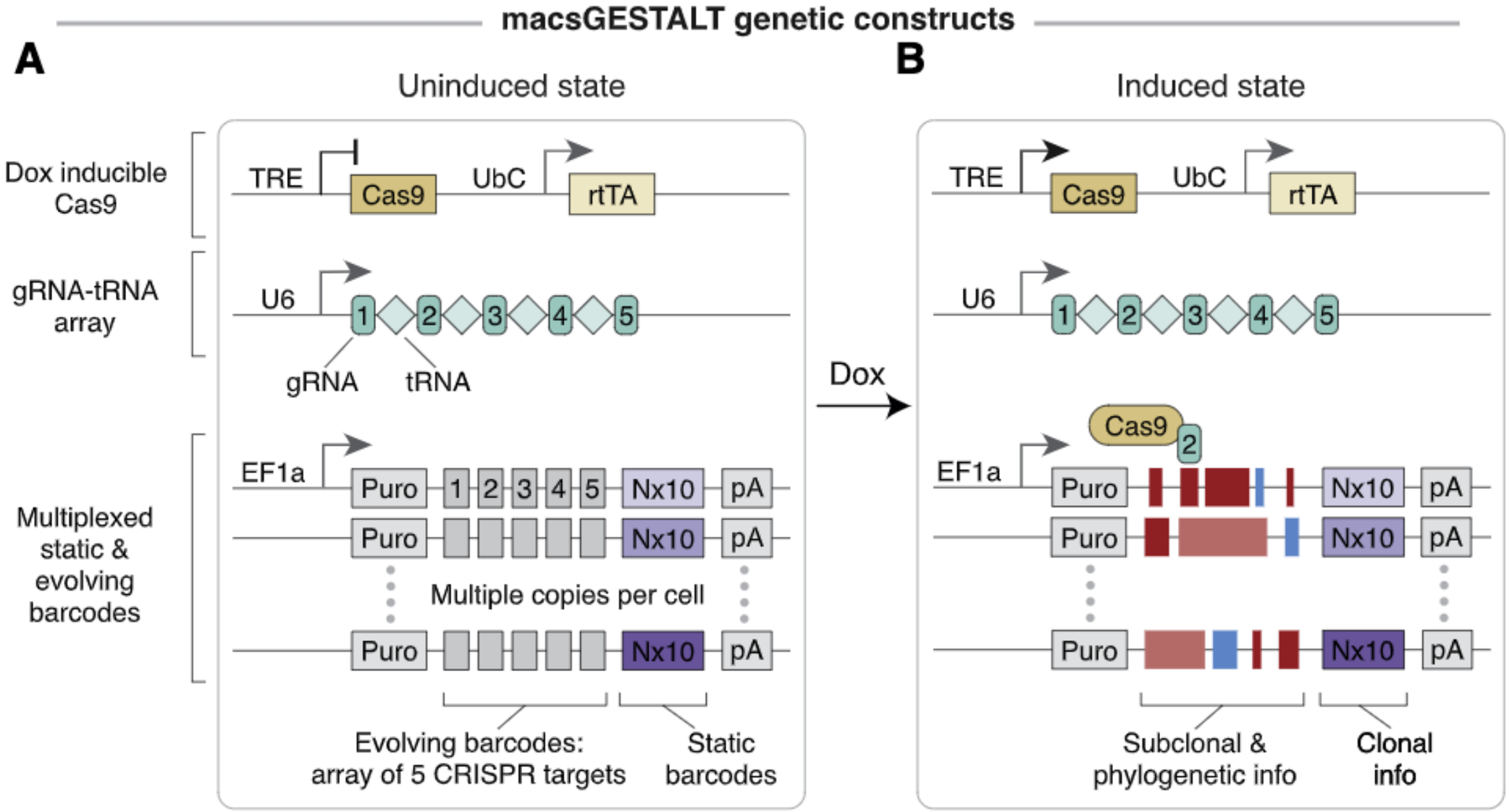
Aaron McKenna,[1]* Gregory M. Findlay,[1]* James A. Gagnon,[2]* Marshall S. Horwitz,[1,3] Alexander F. Schier,[2,4,5,6]† Jay Shendure[1,7]†



- Applied to cell culture and zebrafish
- Found that most cells in adult organs derive from relatively few embryonic ancestor cells

From Summary Figure of McKenna et al., Science (2016)

macsGESTALT for high-res lineage tracing

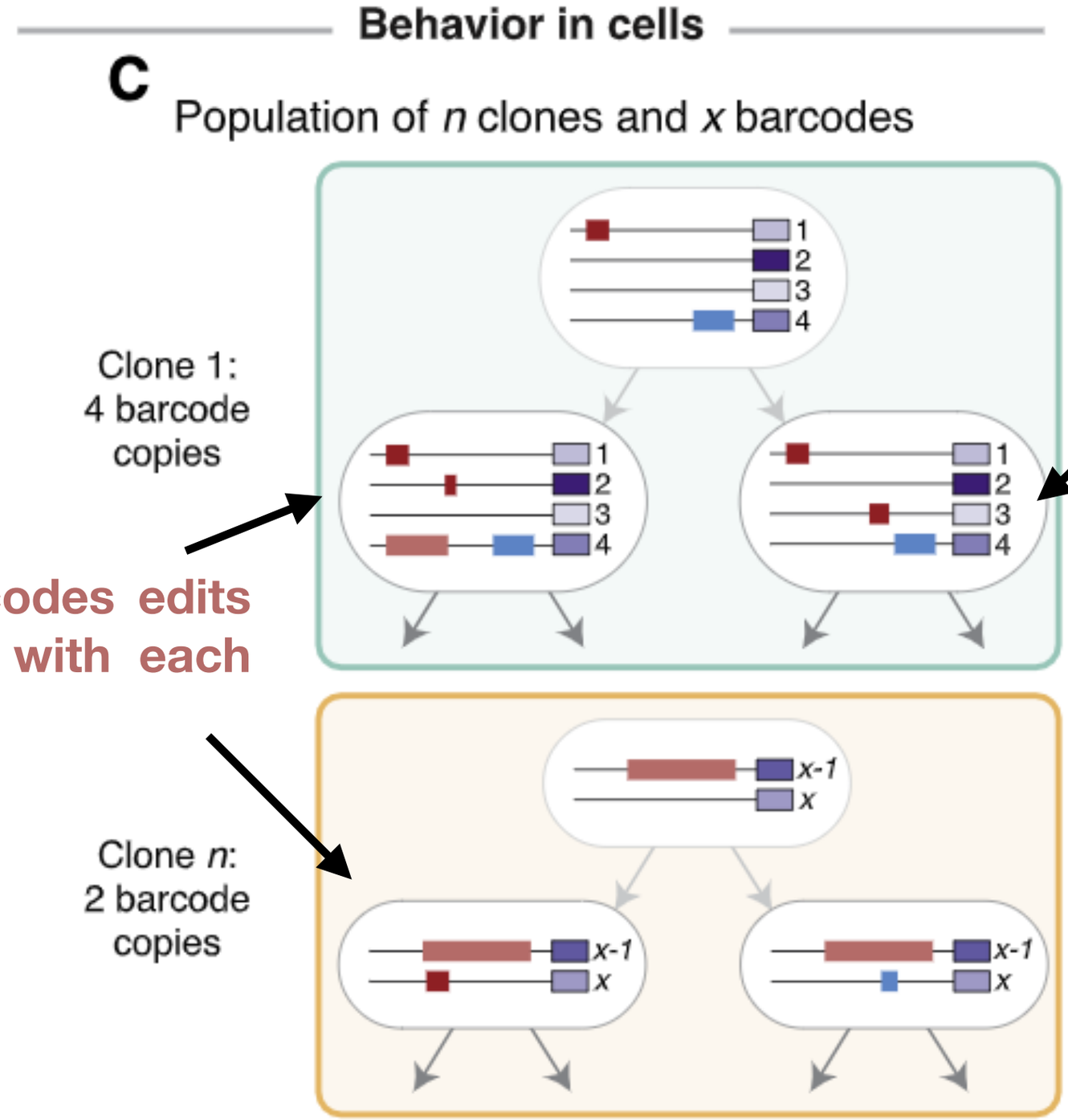


Static barcodes track clonal information

Evolving barcodes (via indel mutagenesis) track sub-clonal phylogenetic information

From Figure 1 of Simeonov et al., Cancer Cell (2021)

macsGESTALT for high-res lineage tracing



From Figure 1 of Simeonov et al., Cancer Cell (2021)

macsGESTALT for high-res lineage tracing

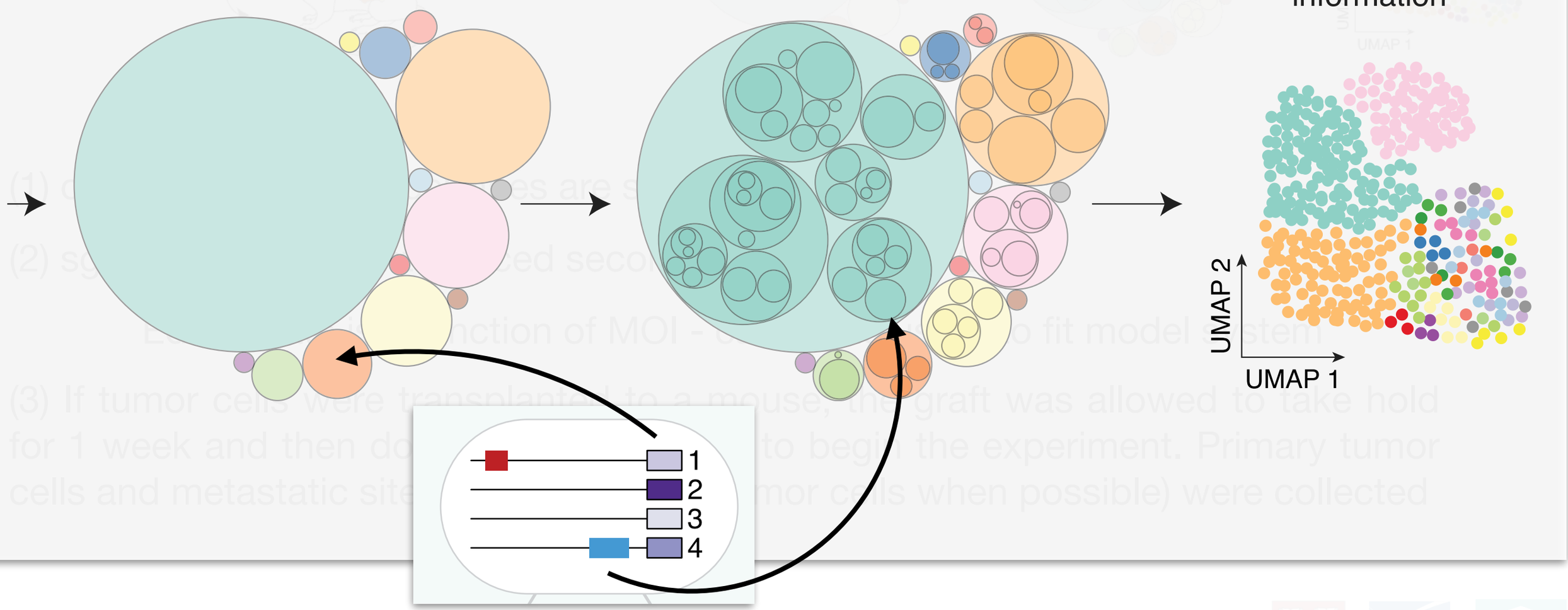
Analysis workflow

E

Clone discovery via static barcode presence

Subclonal reconstruction via evolving barcodes

Integration with transcriptional & functional information



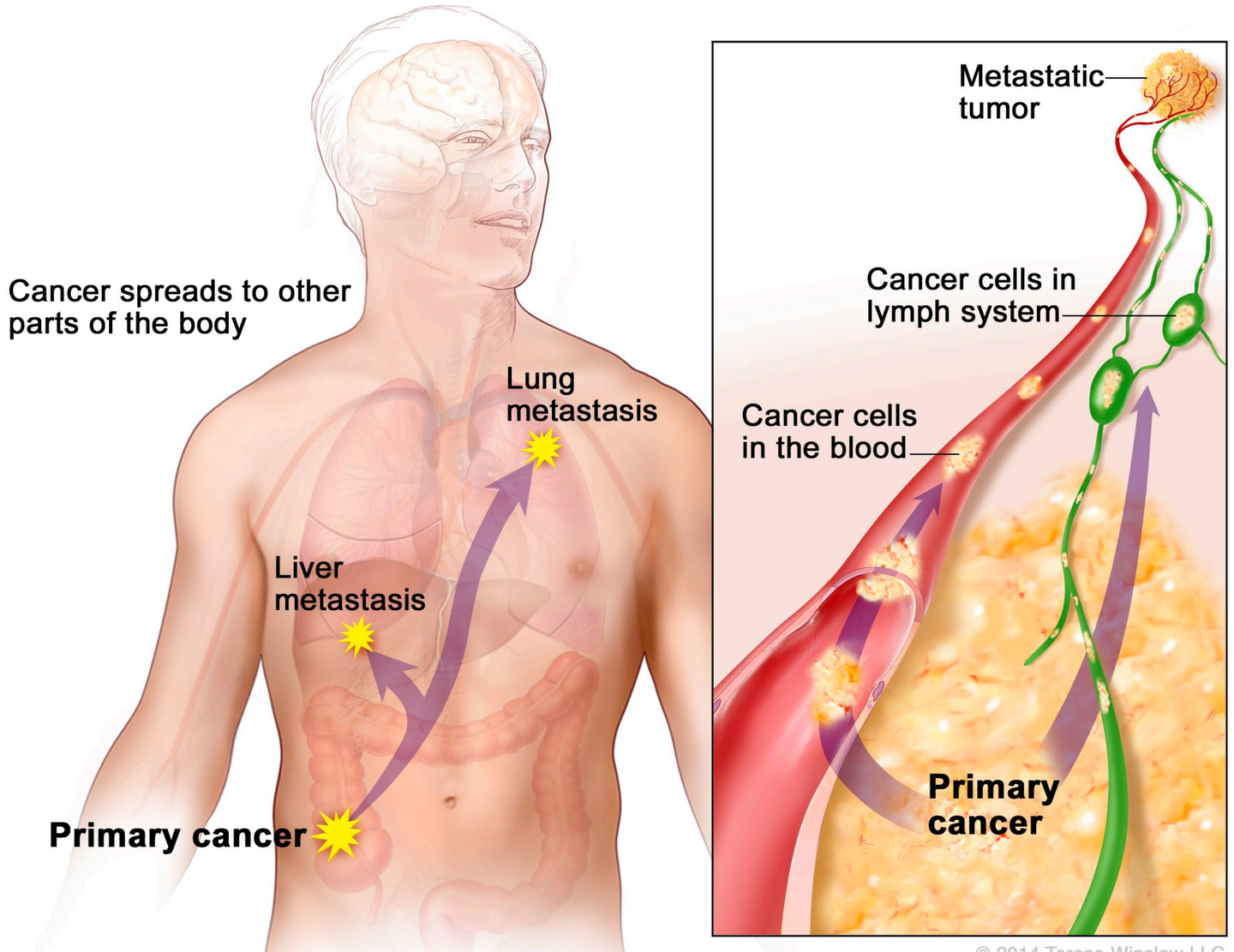
From Figure 1 of Simeonov et al., Cancer Cell (2021)

What is metastasis and EMT?

- Metastasis = most cancer deaths

- EMT ~ metastasis

- EMT thought to play a role across many cancer types



***Source**: National Cancer Institute*

Most metastases arise from rare clones

Experiment: combine macsGESTALT with scRNA-seq

From Figure 2 of Simeonov et al., Cancer Cell (2021)

PDAC is very deadly (5-year survival rate of 9%)

KPCY mouse tumor model cell line transplanted into non-tumor-bearing mice

Good model for two reasons:

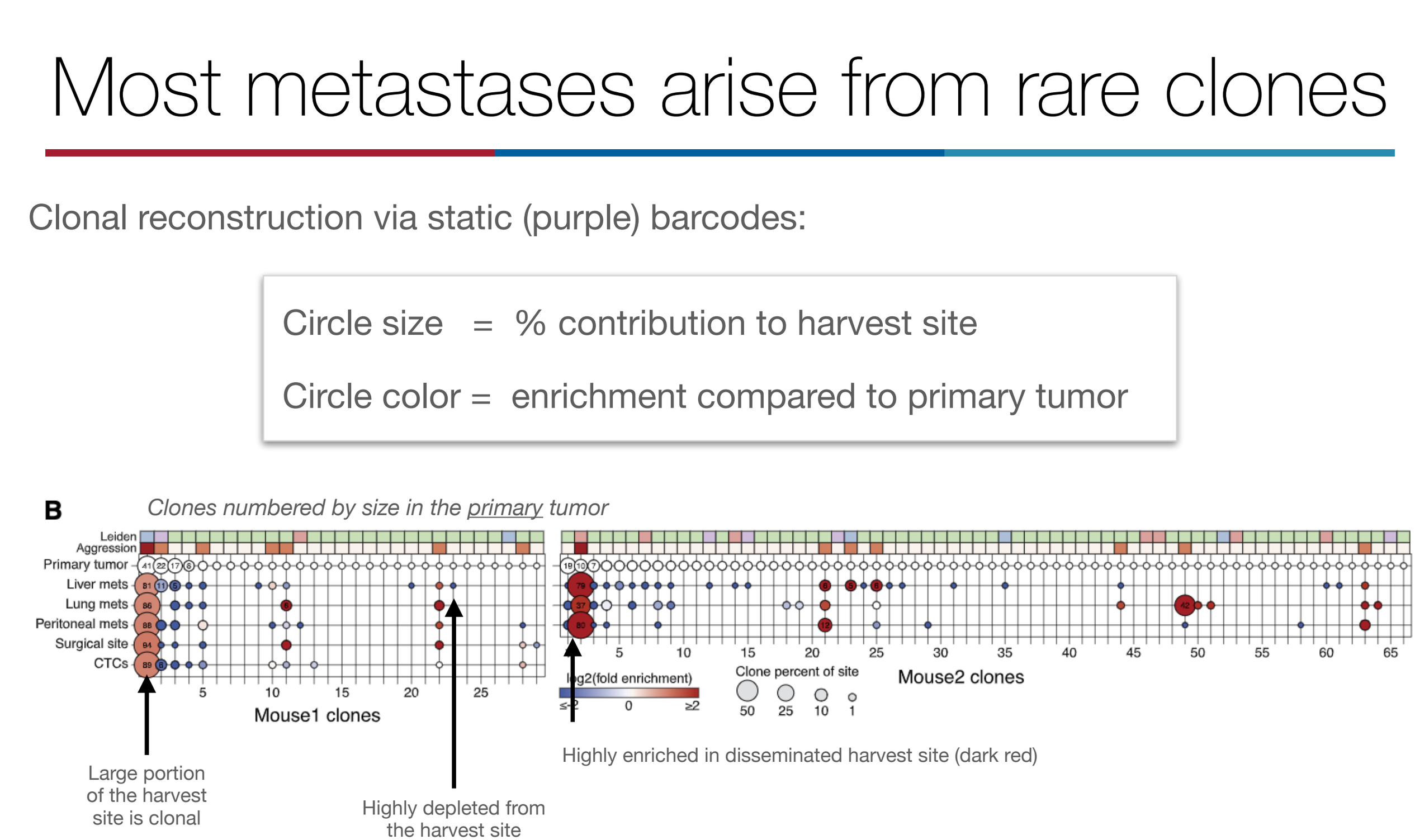
(1) This model exhibits consistent metastasis kinetics
(2) Good model of human disease (Kras GoF and p53 LoF are the most common drivers of human PDAC)
(3) minimal *in vitro* cell line culture time
(4) pancreatic focal lesion disseminates to the same sites as in human PDAC (incl. liver and lung)

Most metastases arise from rare clones

Clonal reconstruction via static (purple) barcodes:

Circle size = % contribution to harvest site

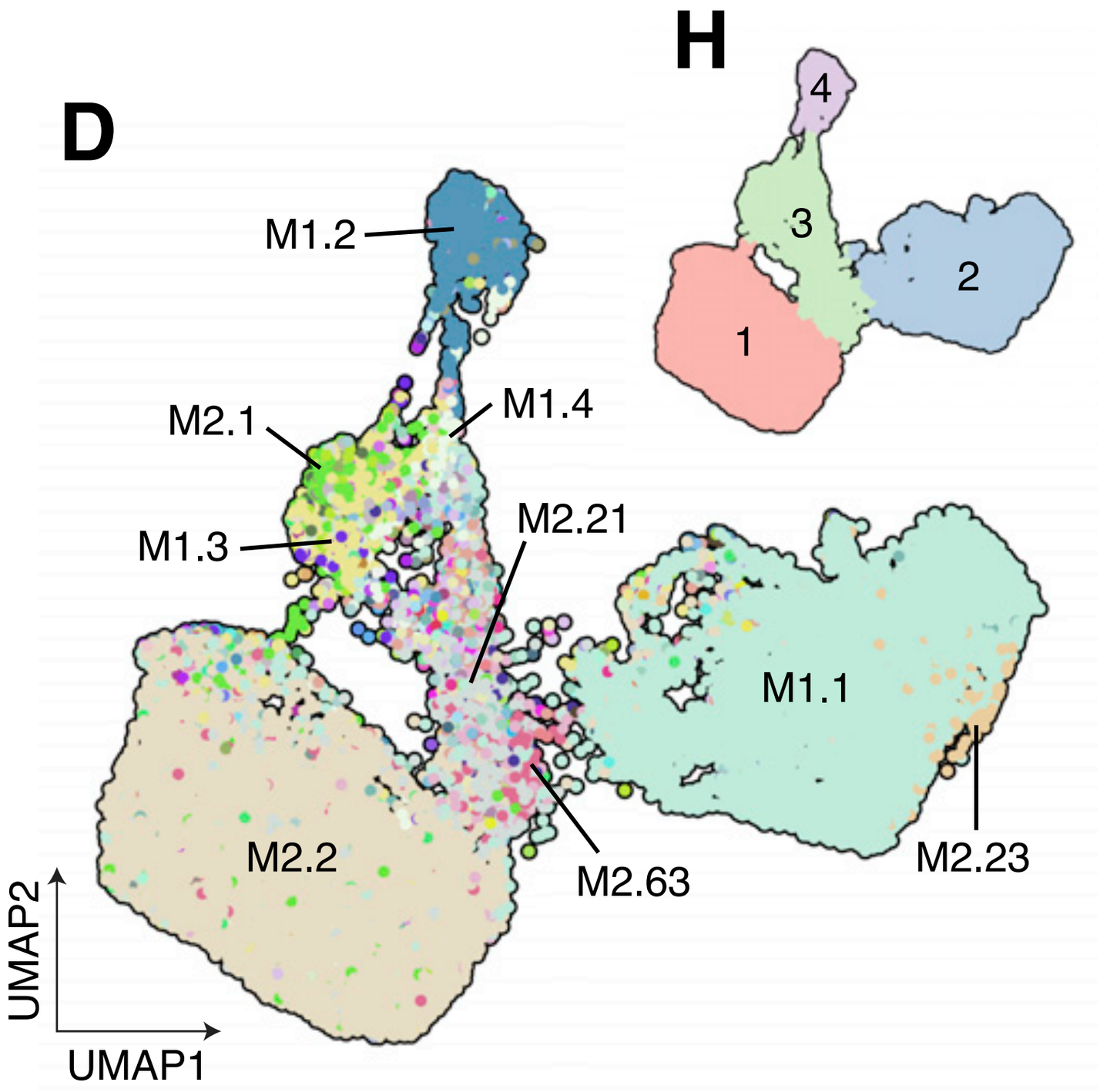
Circle color = enrichment compared to primary tumor



"51% of clones (48/95) failed to metastasize at all, suggesting that mutations in Kras and p53 alone do not ensure metastatic success."

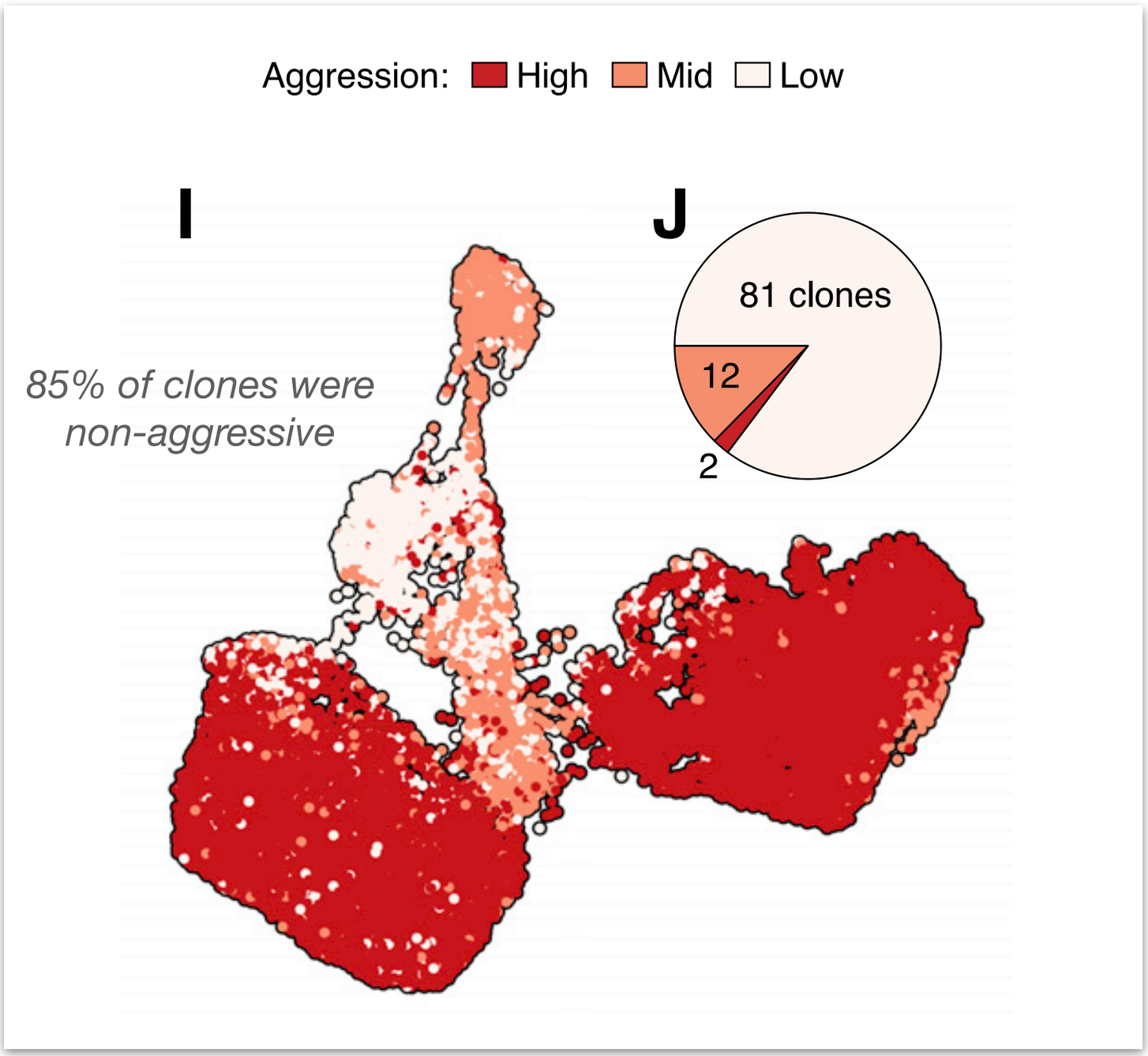
From Figure 2 of Simeonov et al., Cancer Cell (2021)

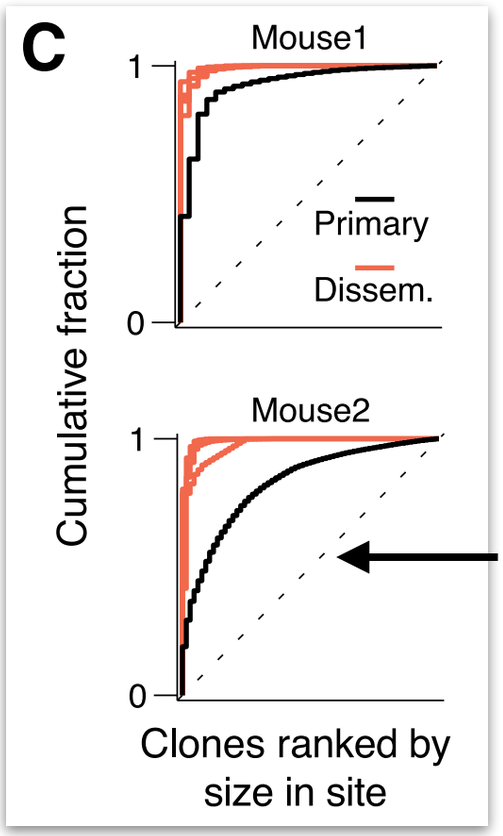
Most metastases arise from rare clones



From Figures 2 of Simeonov et al., Cancer Cell (2021)

Most metastases arise from rare clones

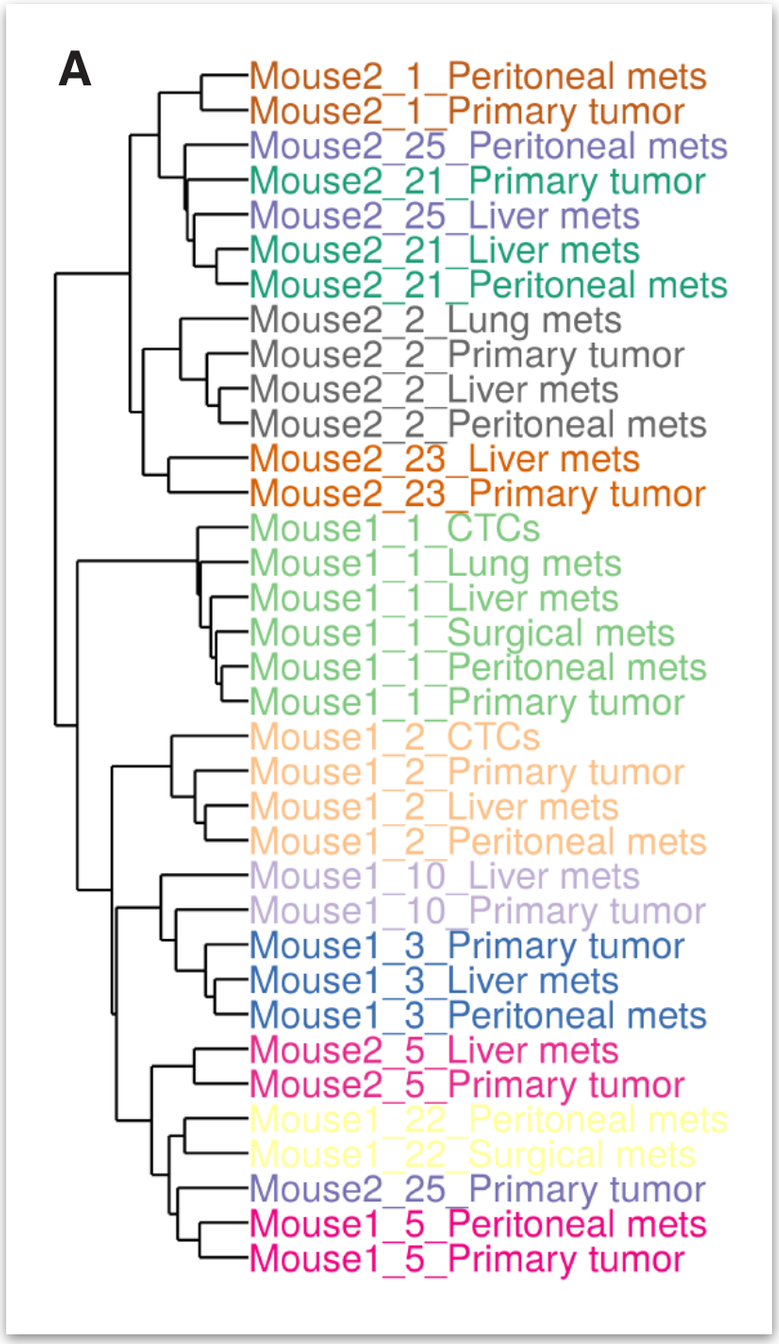




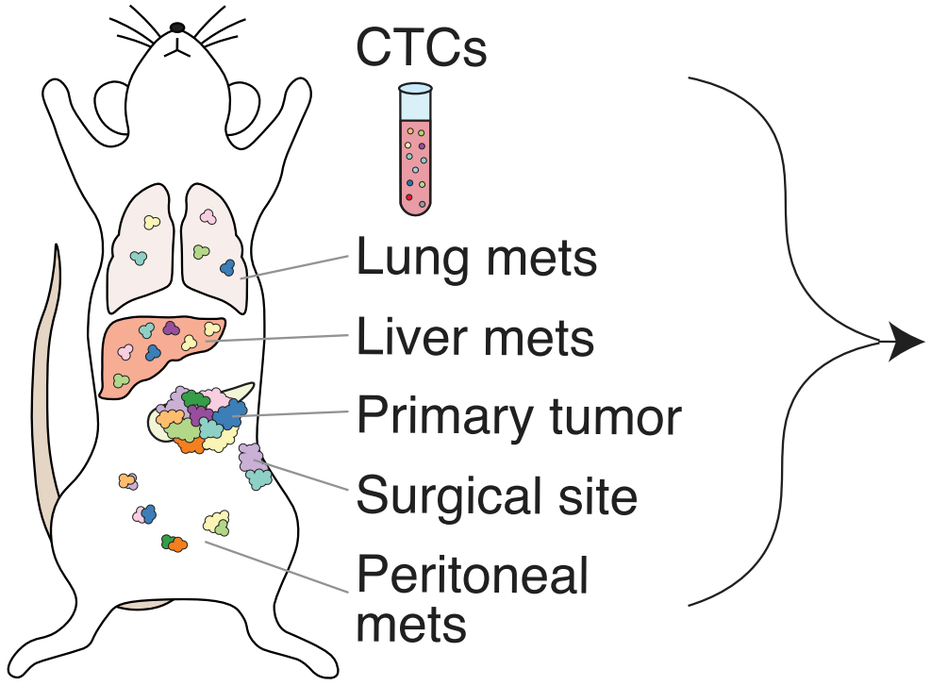
*Just **two clones** dominate most of the cell population!*

From Figures 2 of Simeonov et al., Cancer Cell (2021)

Cells retain transcriptional identity after metastasis



- Clone-site pseudobulk analysis - samples colored by clone
- Samples hierarchically clustered (based on scRNA-seq) - cluster preferentially by clone rather than harvest site

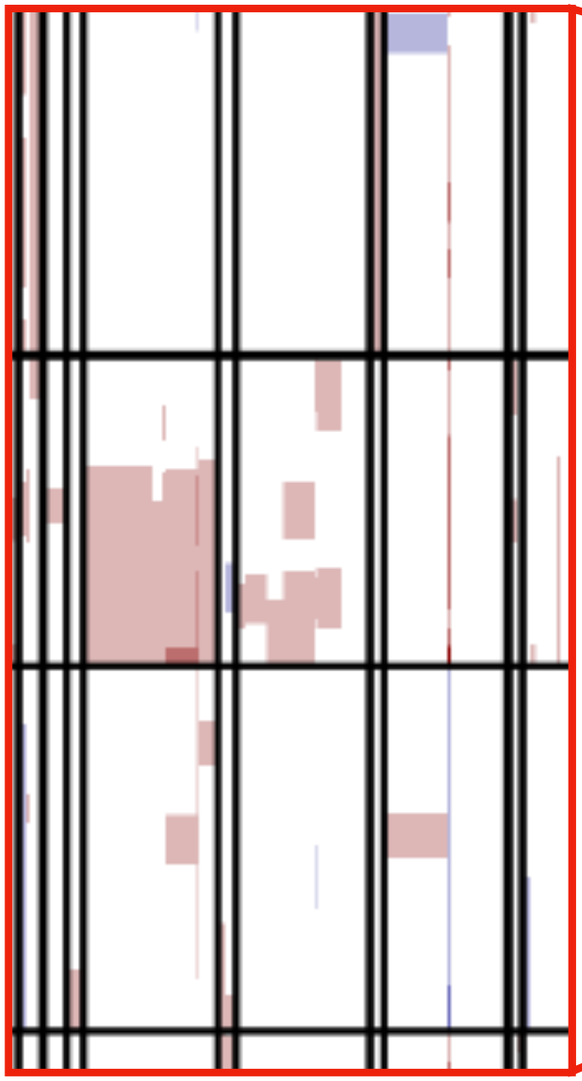


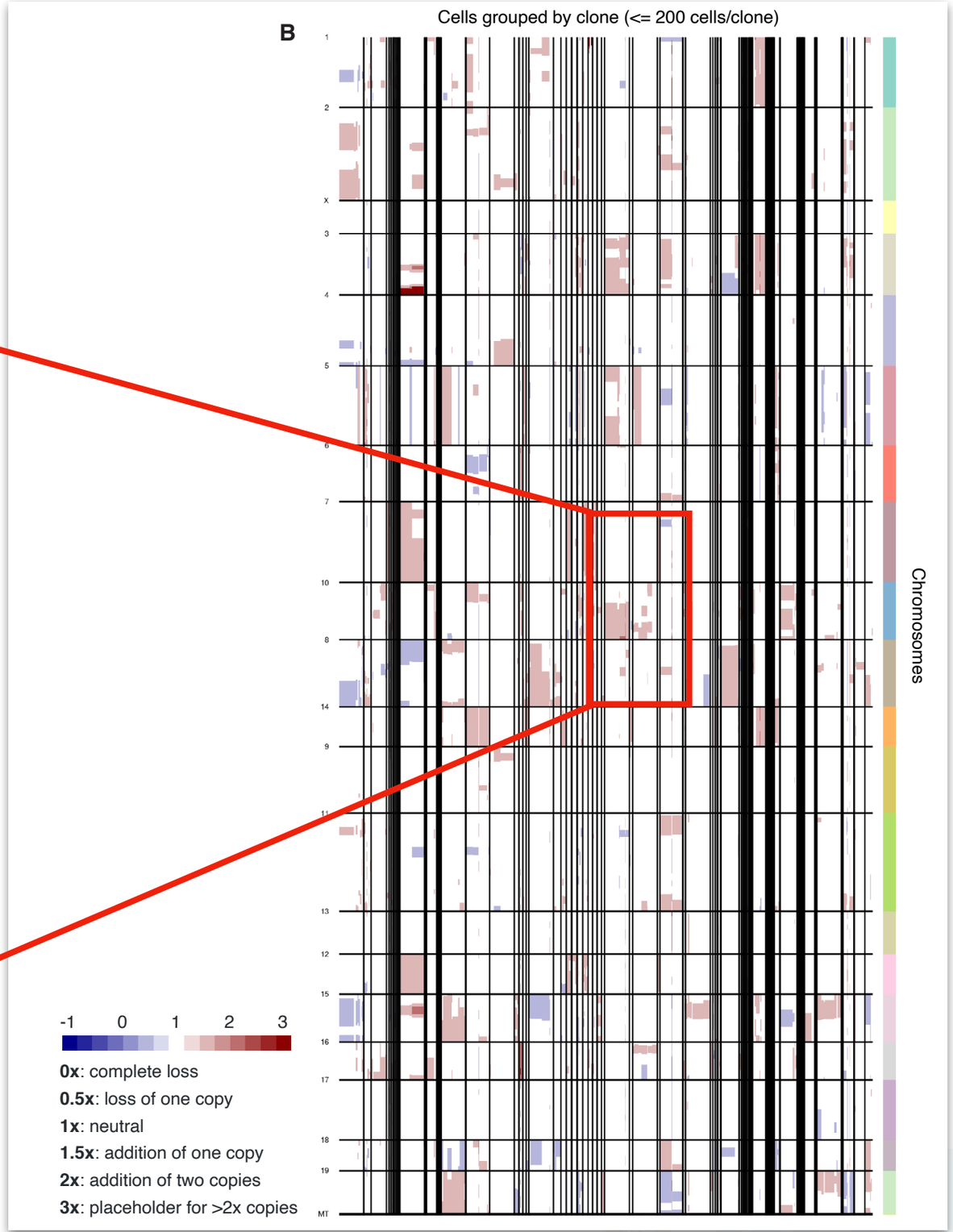
1. FACS sort GFP+ PDAC cells
2. sc-Seq transcriptome & barcodes

From Figures 2, S3 of Simeonov et al., Cancer Cell (2021)

Cells retain transcriptional identity after metastasis

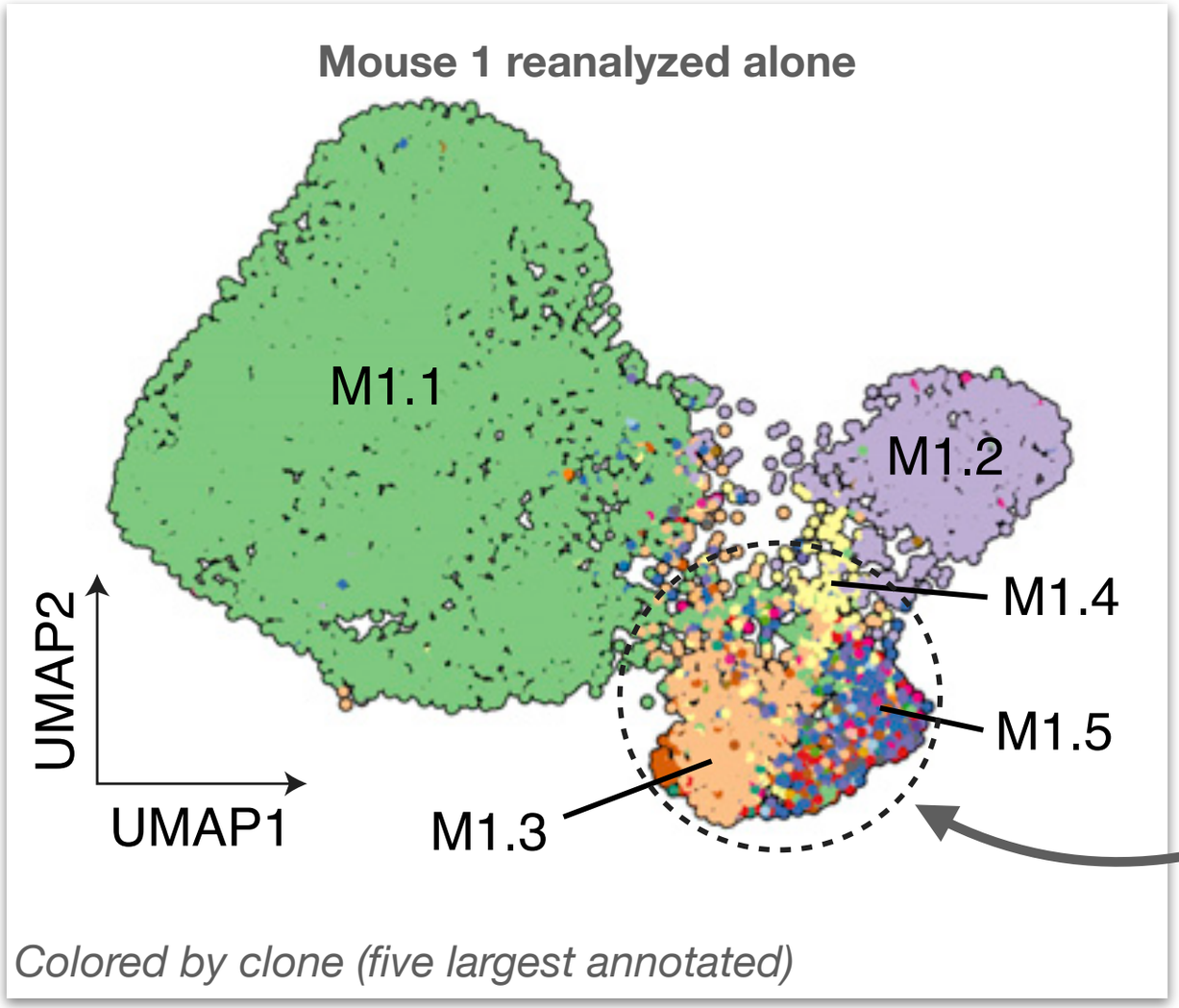
- Performed *InferCNV* analysis

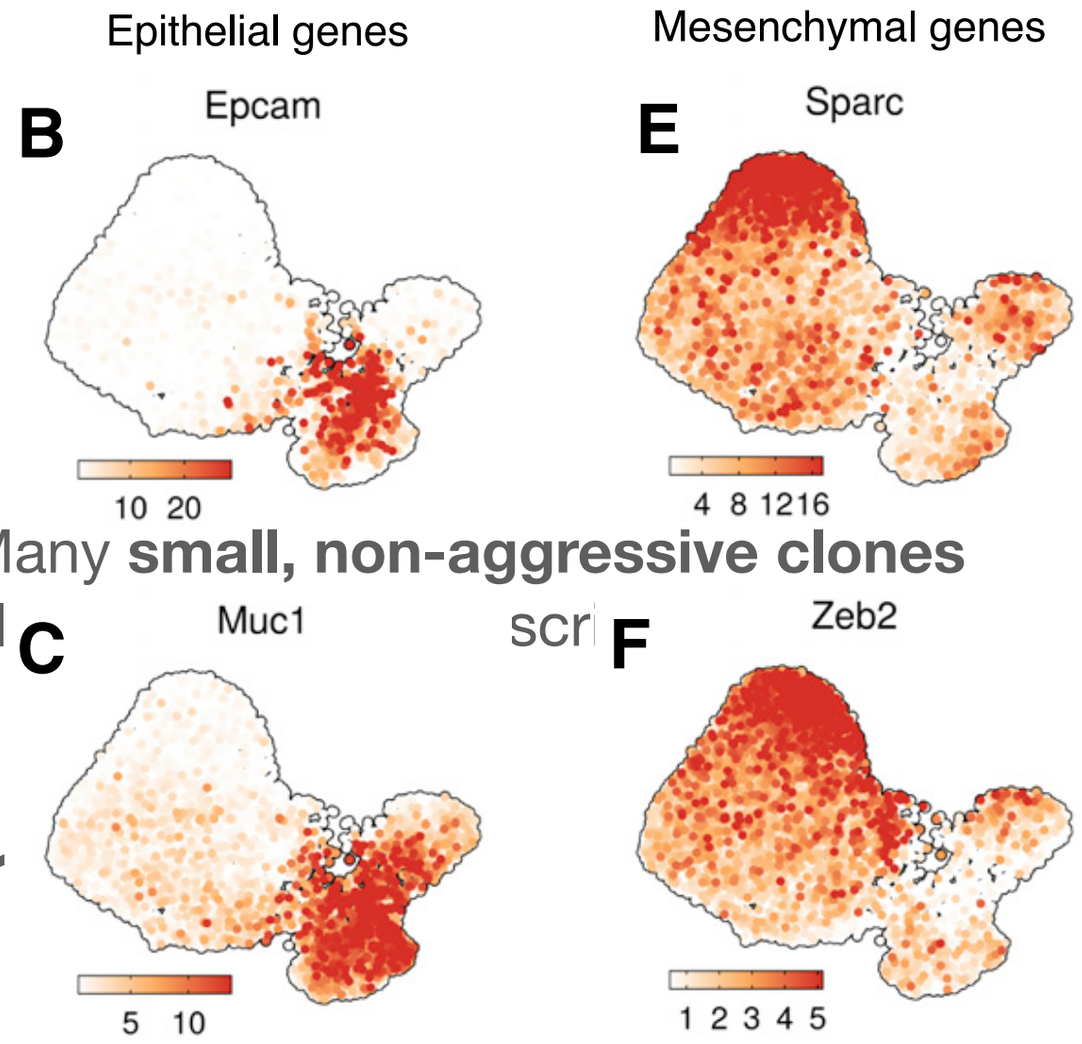




From Figures S3 of Simeonov et al., Cancer Cell (2021)

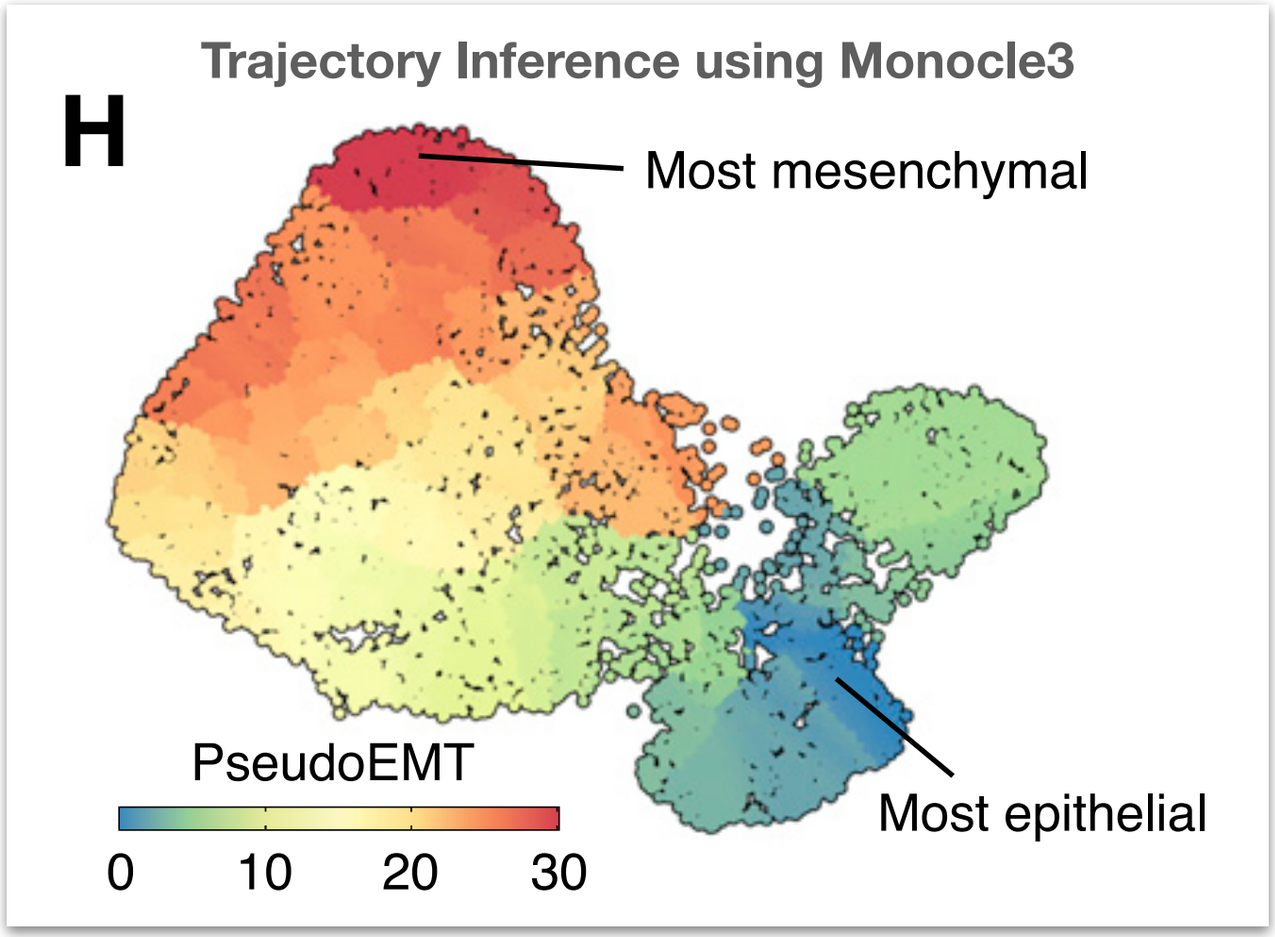
Transcriptional EMT continuum *in vivo*

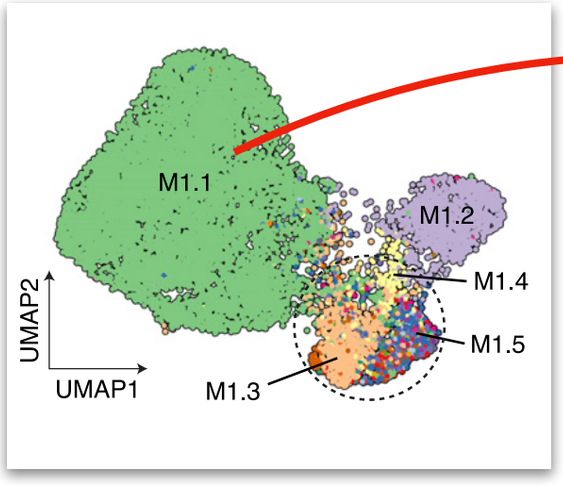


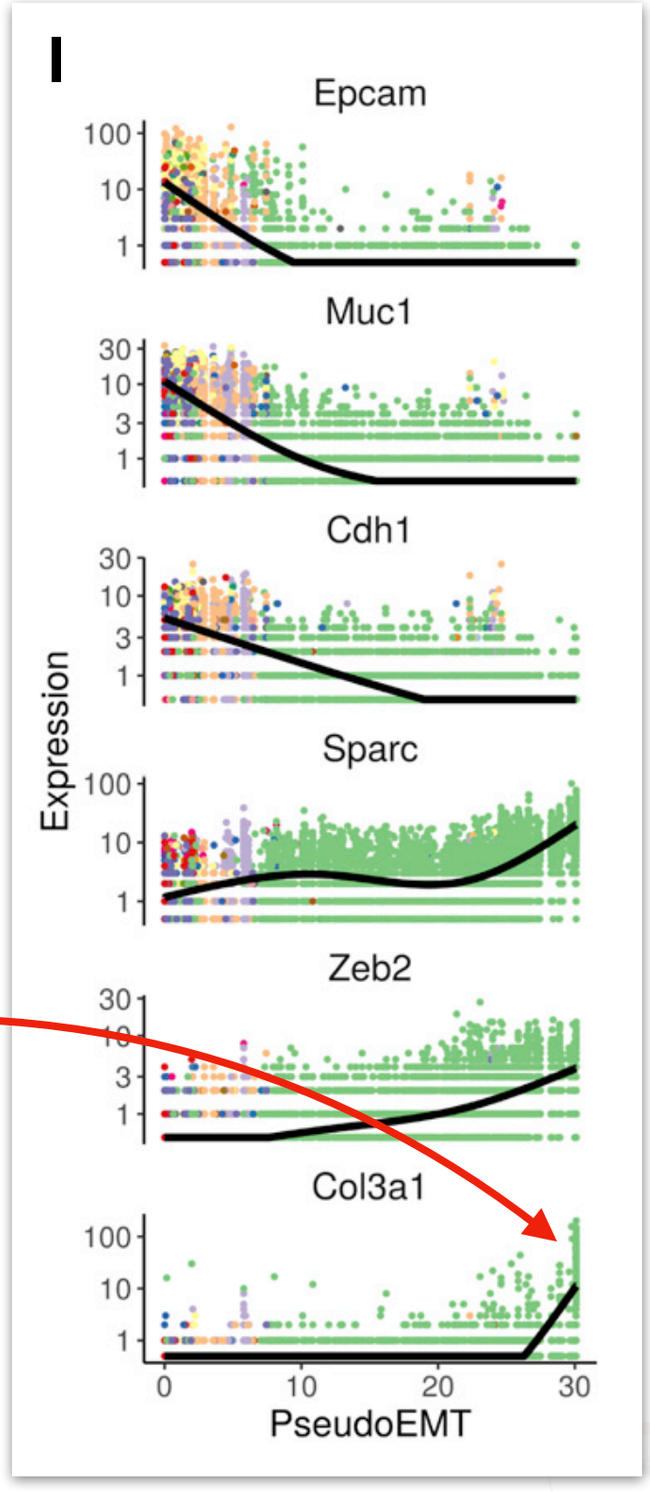


From Figure 3 of Simeonov et al., Cancer Cell (2021)

Transcriptional EMT continuum *in vivo*



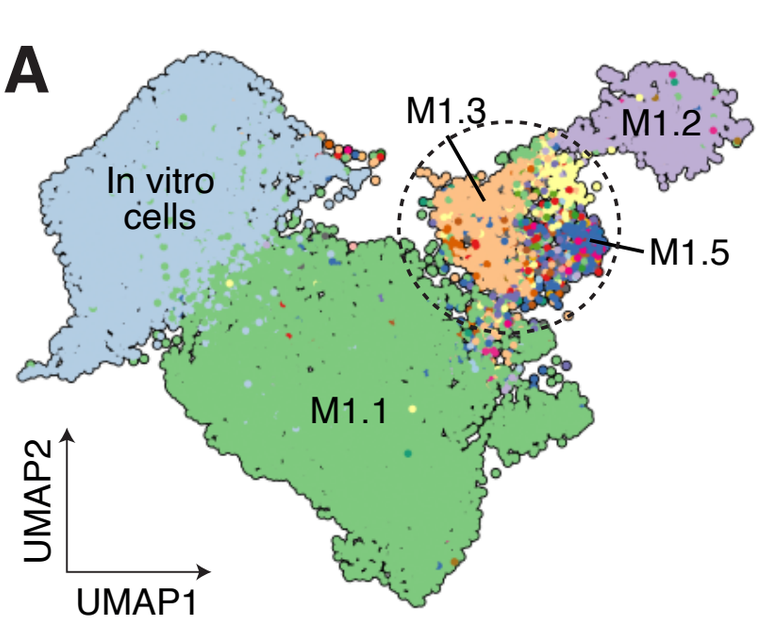


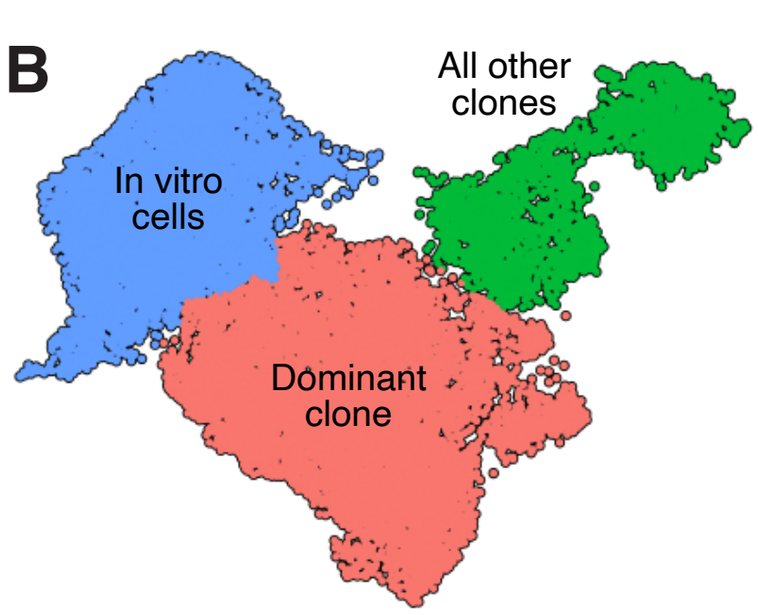


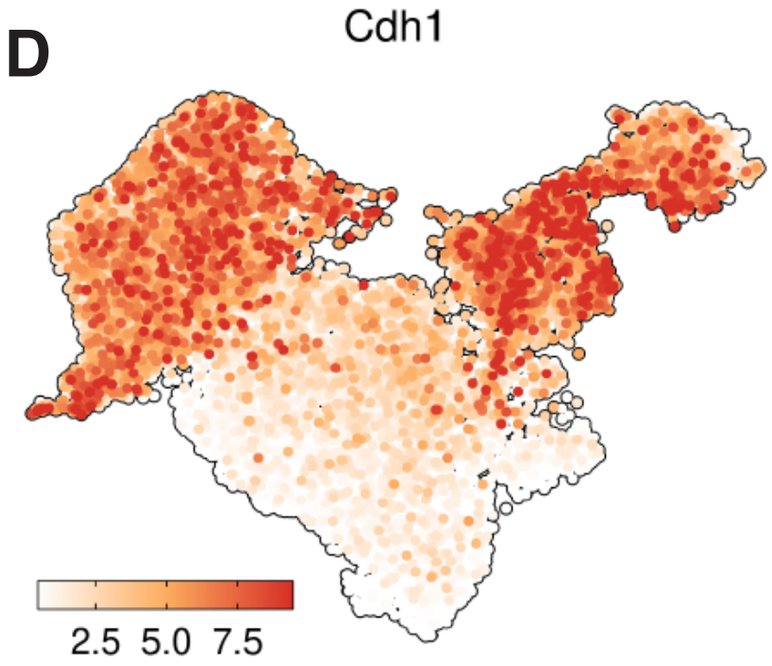
From Figure 3 of Simeonov et al., Cancer Cell (2021)

Transcriptional EMT continuum *in vivo*

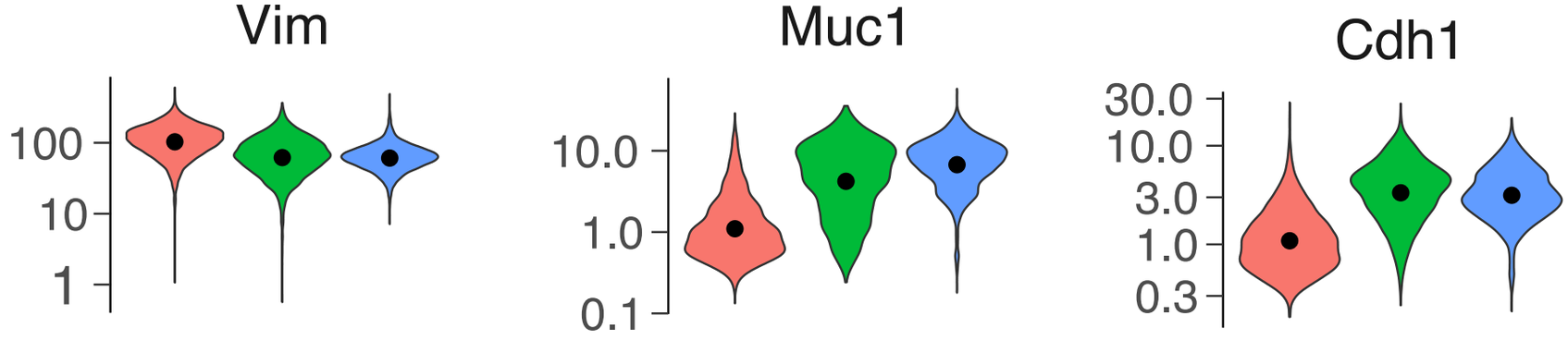
- 27/29 clones are epithelial - is this the default transcriptional state?
- scRNA-seq of *in vitro* cultured cells -> 40 distinct clones





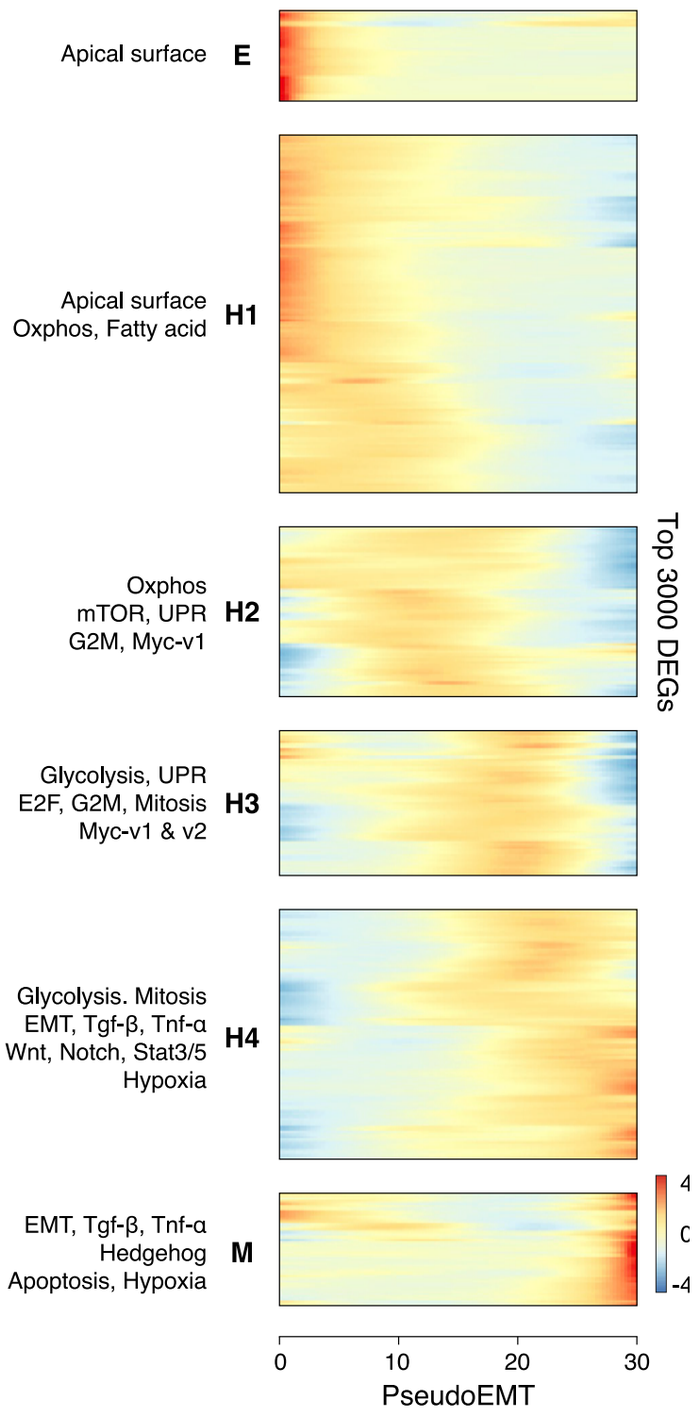


- Compared to PDAC / EMT, *in vitro* cultured cells **strikingly epithelial**

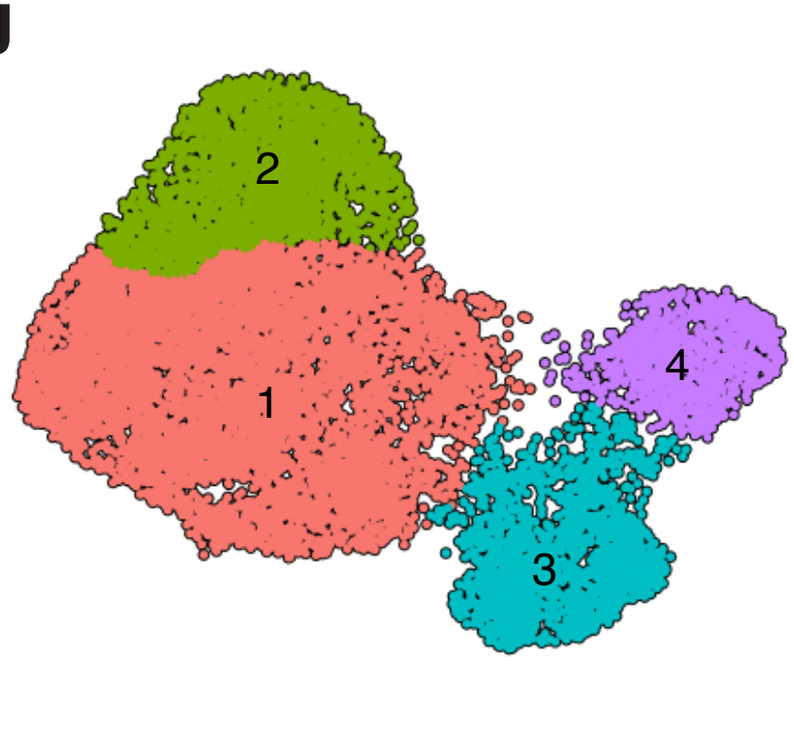


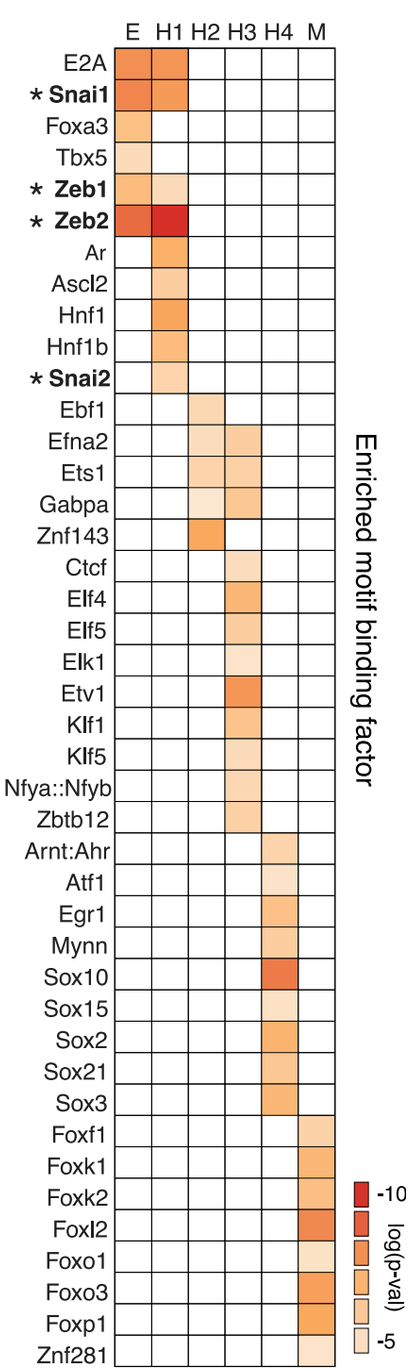
From Figure S5 of Simeonov et al., Cancer Cell (2021)

Transcriptional EMT continuum *in vivo*



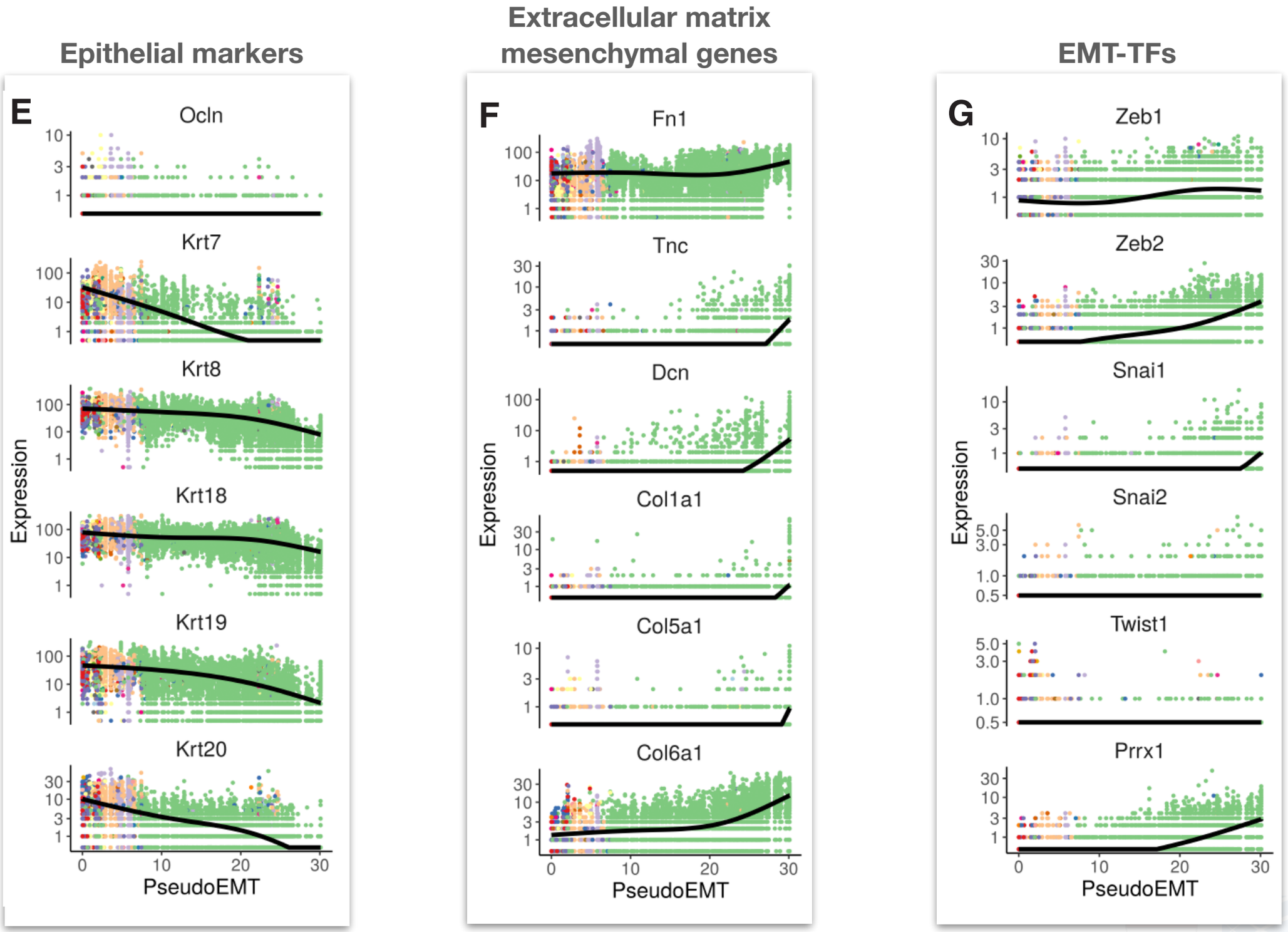
- Remember: not an EMT model but a PDAC model
- Cancer proliferation ~ EMT
- Possible ties between common cancer targets like Myc / mTOR / Wnt / Notch / etc and EMT





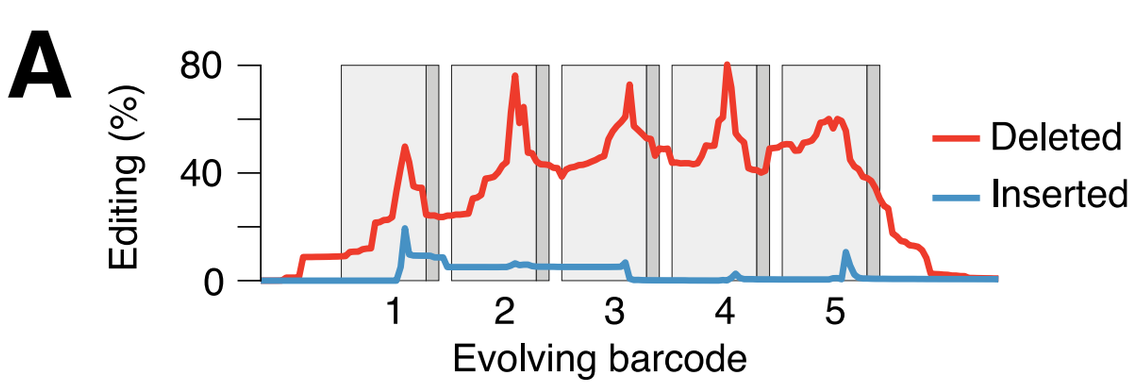
From Figure 3, S5 of Simeonov et al., Cancer Cell (2021)

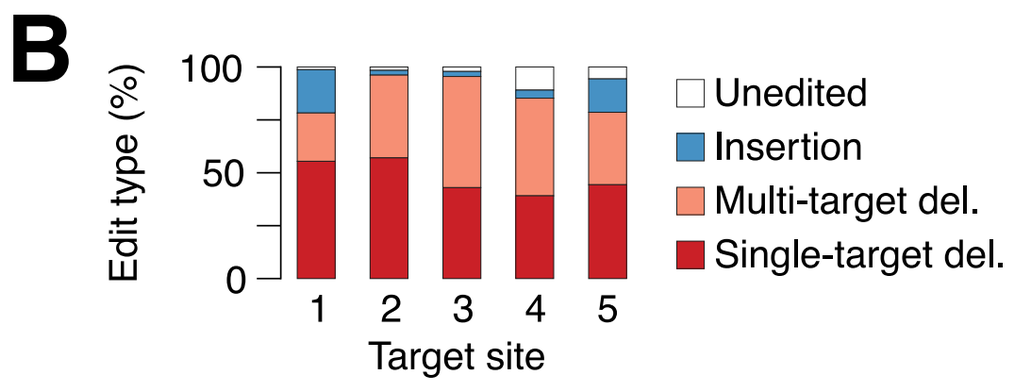
Transcriptional EMT continuum *in vivo*



From Figure S4 of Simeonov et al., Cancer Cell (2021)

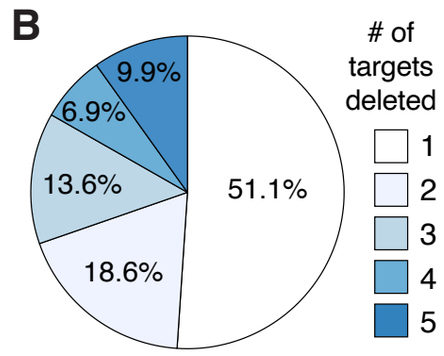
High-res subclonal lineage reconstruction

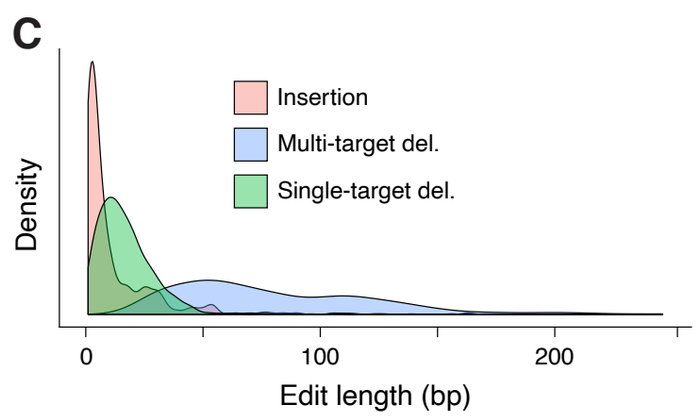




A

	Mouse 1	Mouse 2	Both mice
Total barcodes	50,269	26,705	76,974
Total barcodes with edit	50,132	25,968	76,100
% of barcodes with edit	99.7%	97.2%	98.9%
Total targets	251,345	133,525	384,870
Total targets with edit	246,232	122,013	368,245
% of targets with edit	98.0%	91.4%	95.7%
Distinct edits	2,104	1,383	3,487 (633)
Distinct evolving barcodes	3,886	1,696	5,582 (61)
Distinct barcode-of-barcodes (subclones)	3,997	2,062	6,059 (4)



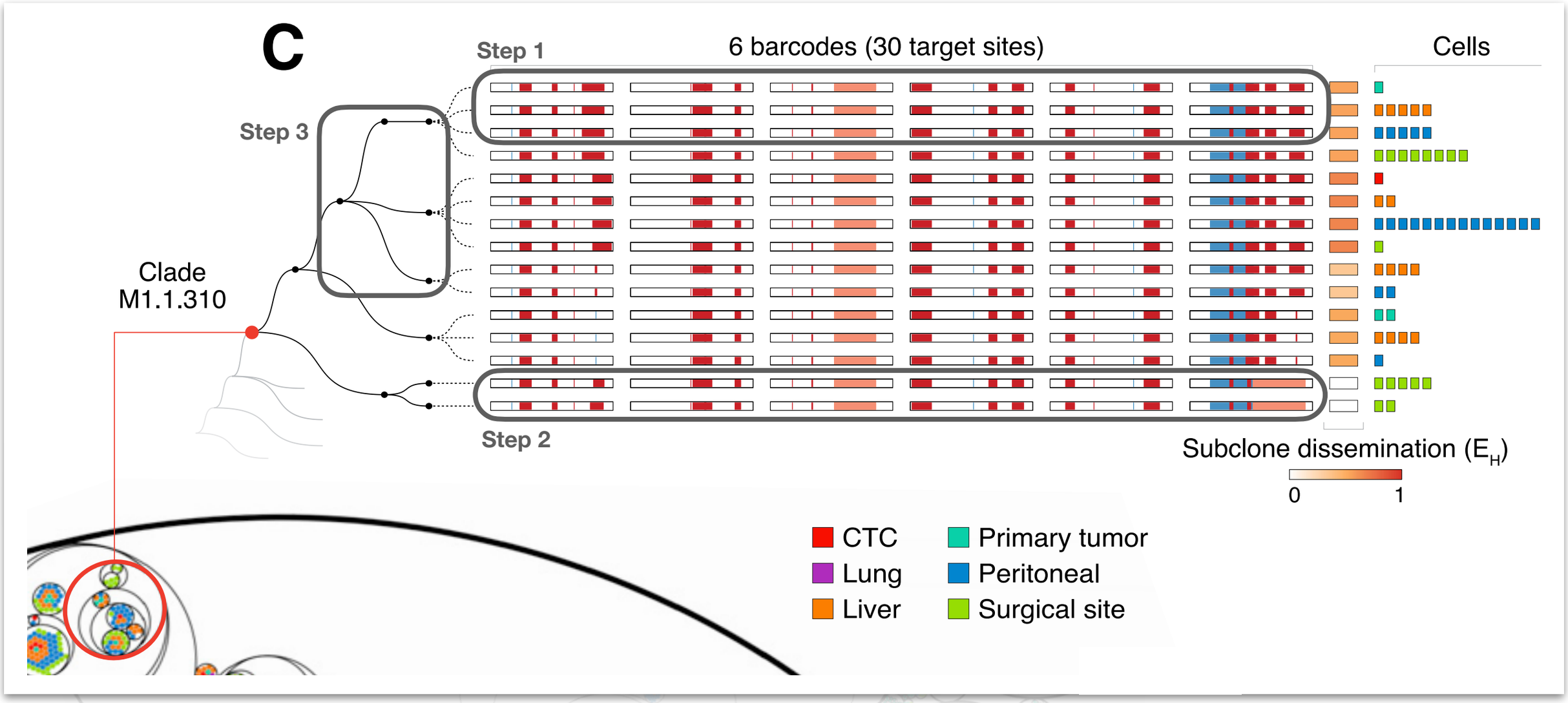


Recovered on average:

- **Mouse 1**: 18.5 target sites (3.7 barcodes) / cell
- **Mouse 2**: 8.5 target sites (1.7 barcodes) / cell

From Figures 4, S6 of Simeonov et al., Cancer Cell (2021)

High-res subclonal lineage reconstruction



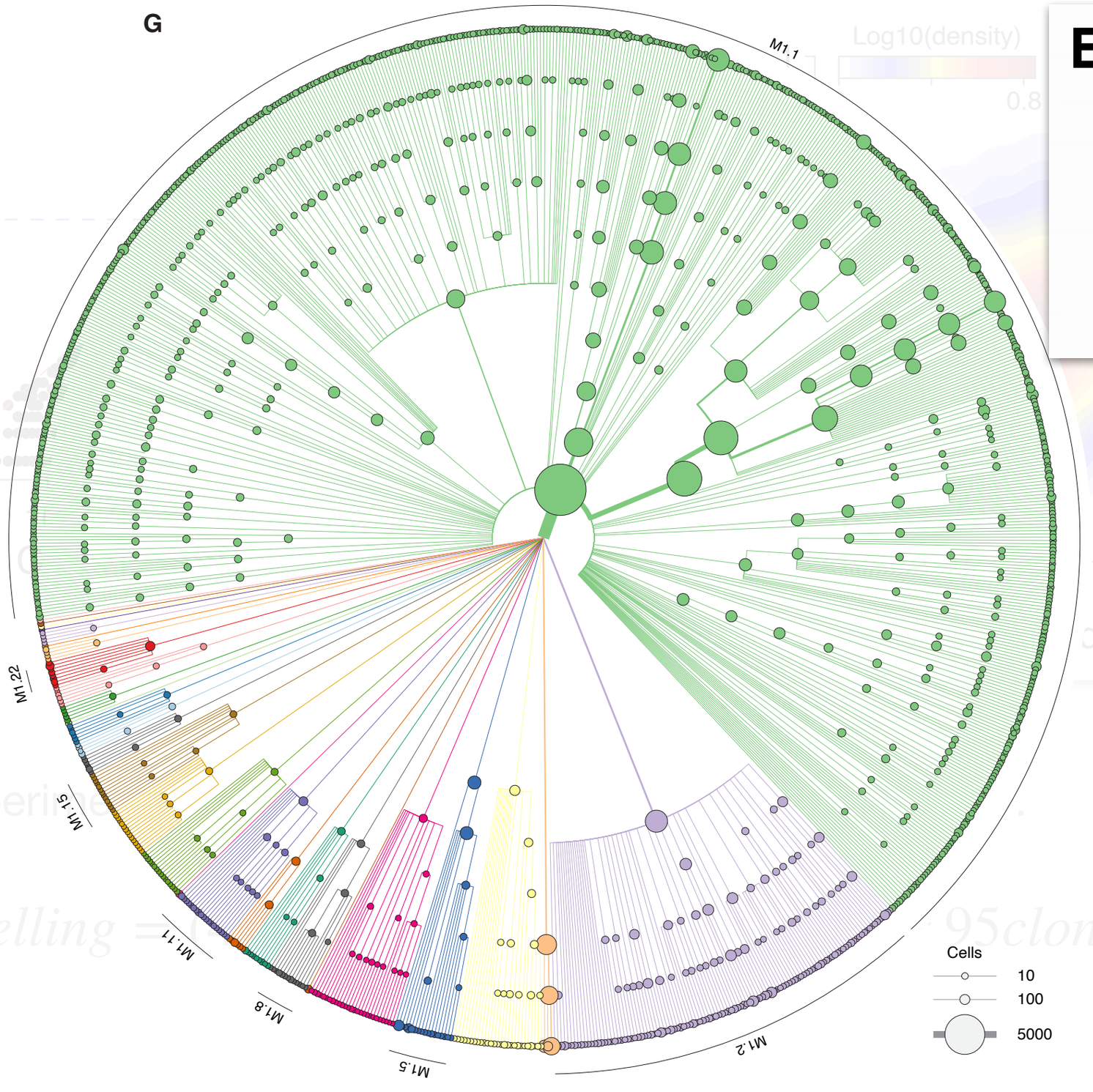
Step 1: assemble “barcode of barcodes”

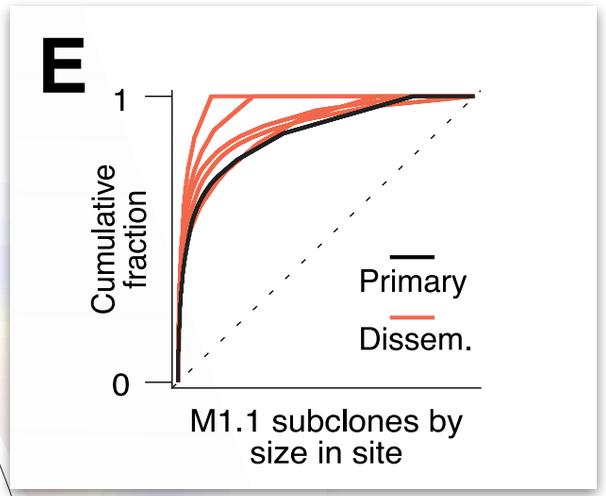
Step 2: group cells into subclones

Step 3: reconstruct phylogenetic relationships between subclones

From Figure 4 of Simeonov et al., Cancer Cell (2021)

High-res subclonal lineage reconstruction

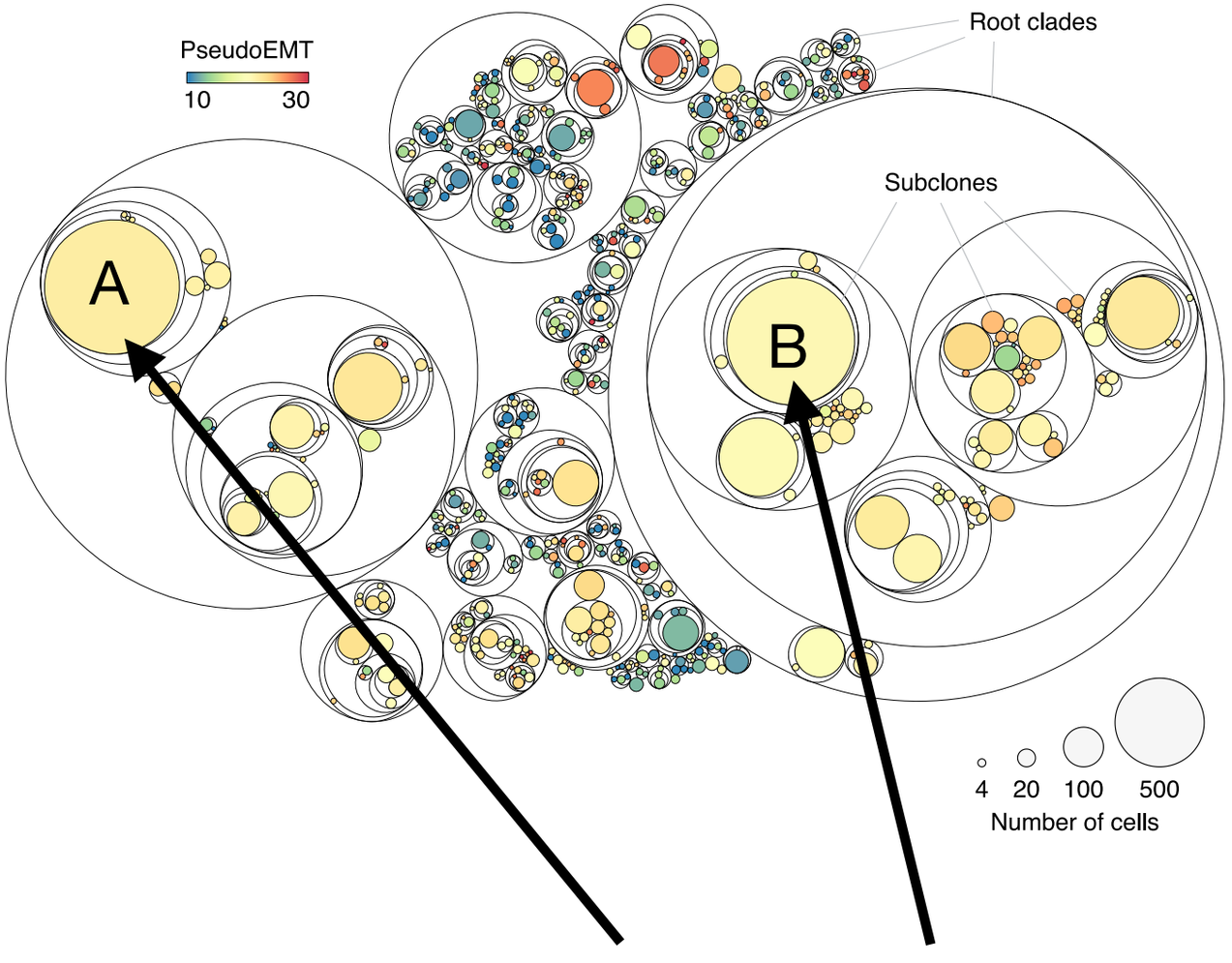


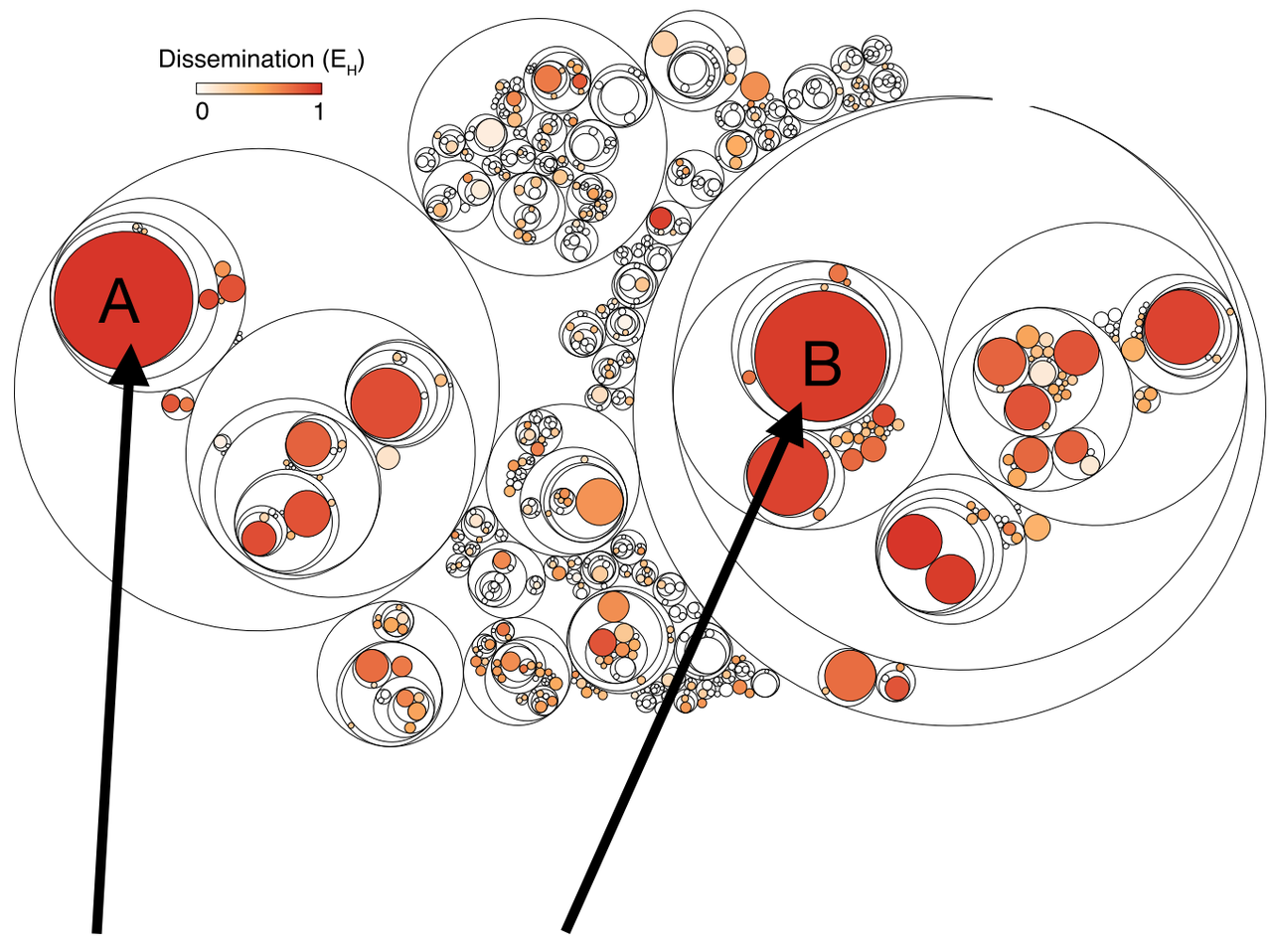


From Figures 4, S6 of Simeonov et al., Cancer Cell (2021)

Peak metastatic aggression corresponds to late-hybrid EMT states

How does the range of **intraclonal EMT states** relate **to subclonal behavior?**



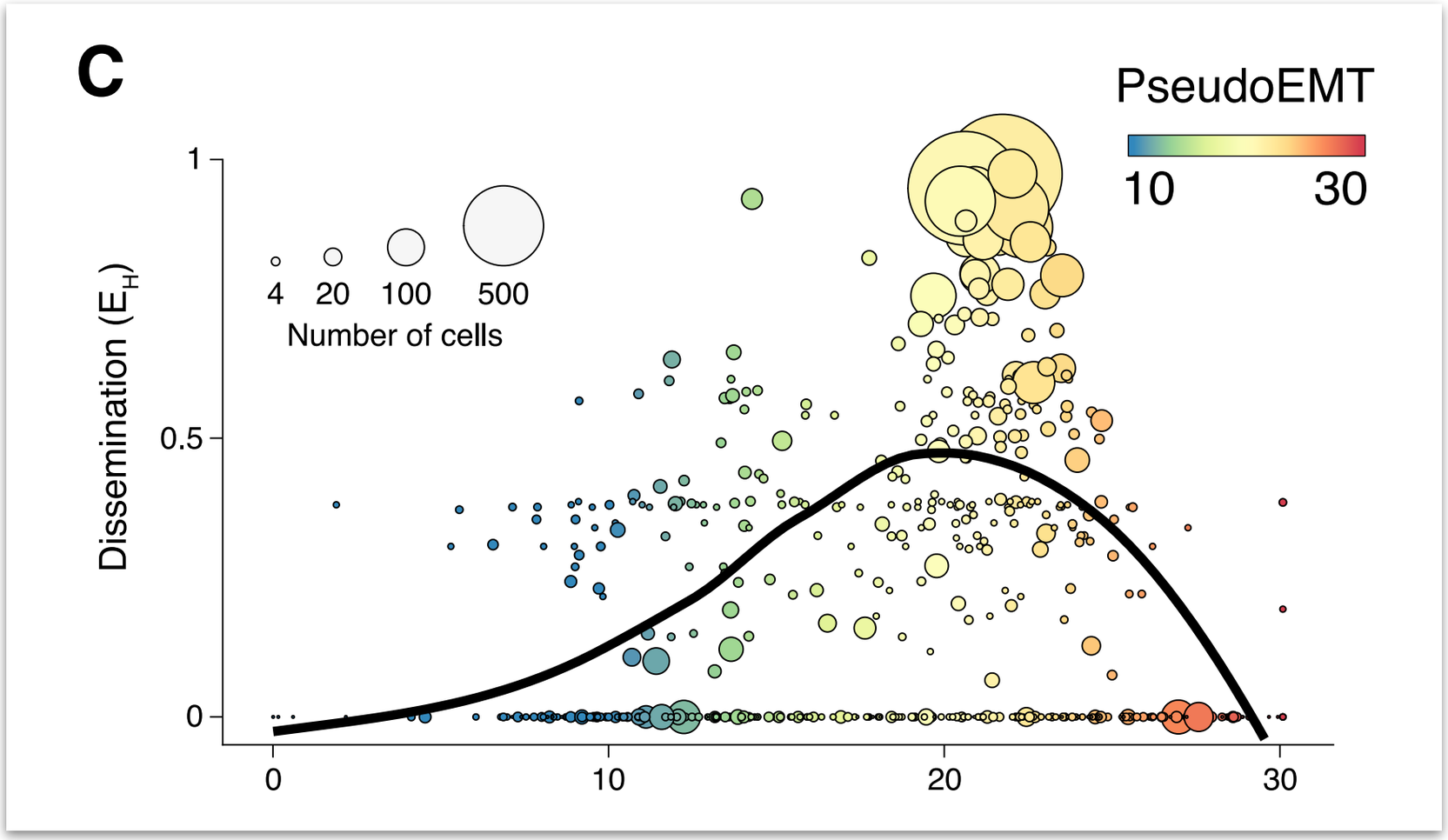


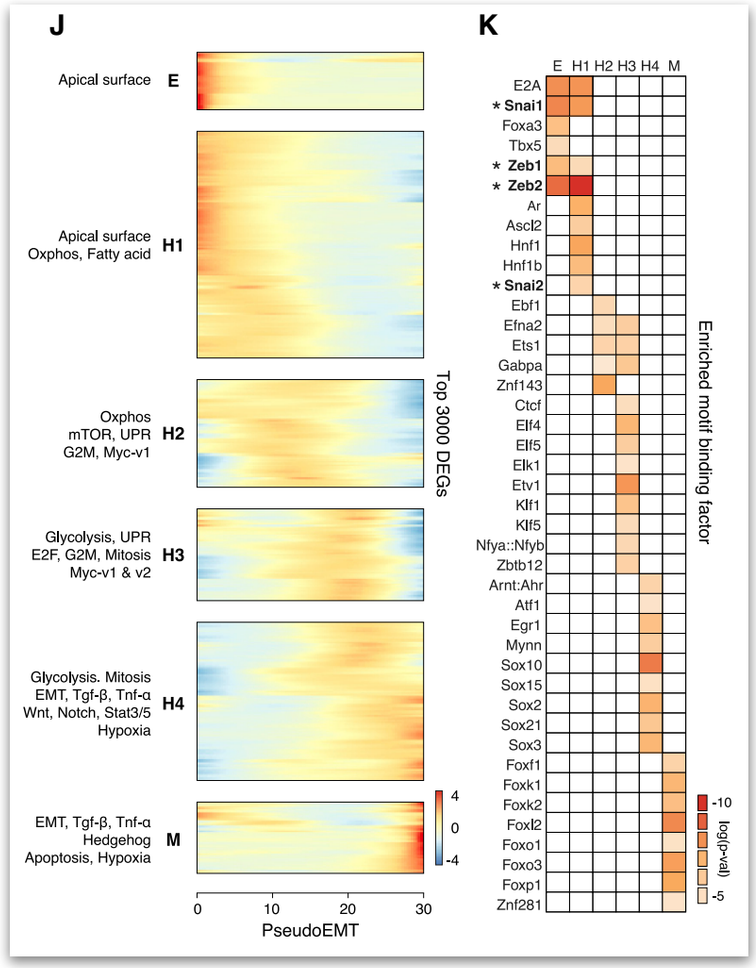
Direct evidence that **hybrid EMT** states are **more metastatic than EMT extremes**

From Figures 3, 5 of Simeonov et al., Cancer Cell (2021)

Peak metastatic aggression corresponds to late-hybrid EMT states

Remember from Figure 3….

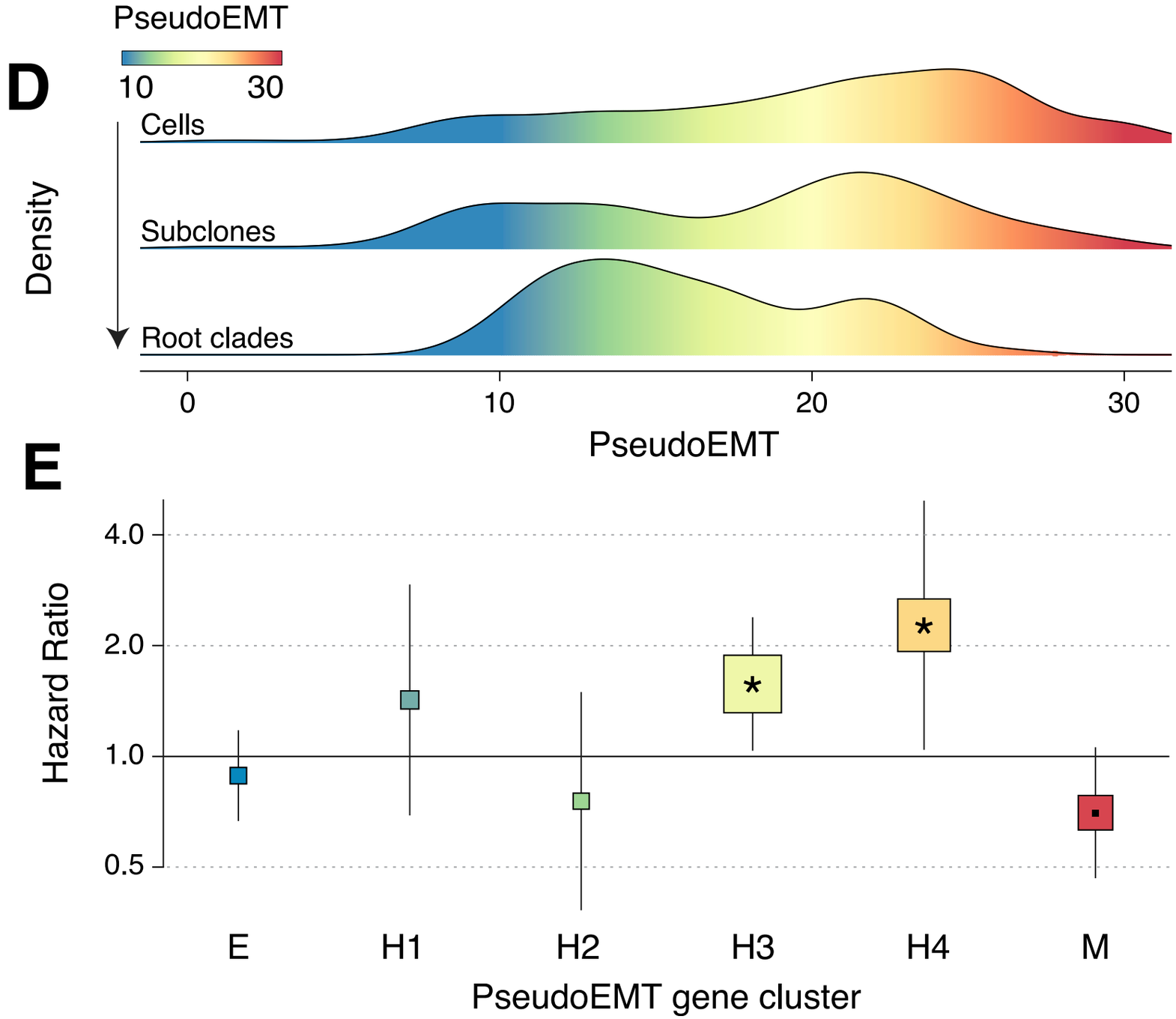




***Important theme**: this paper provides a lot of evidence that EMT, while its own process is very inherent to cancer metastasis - remember we're studying a cancer model not an EMT model*

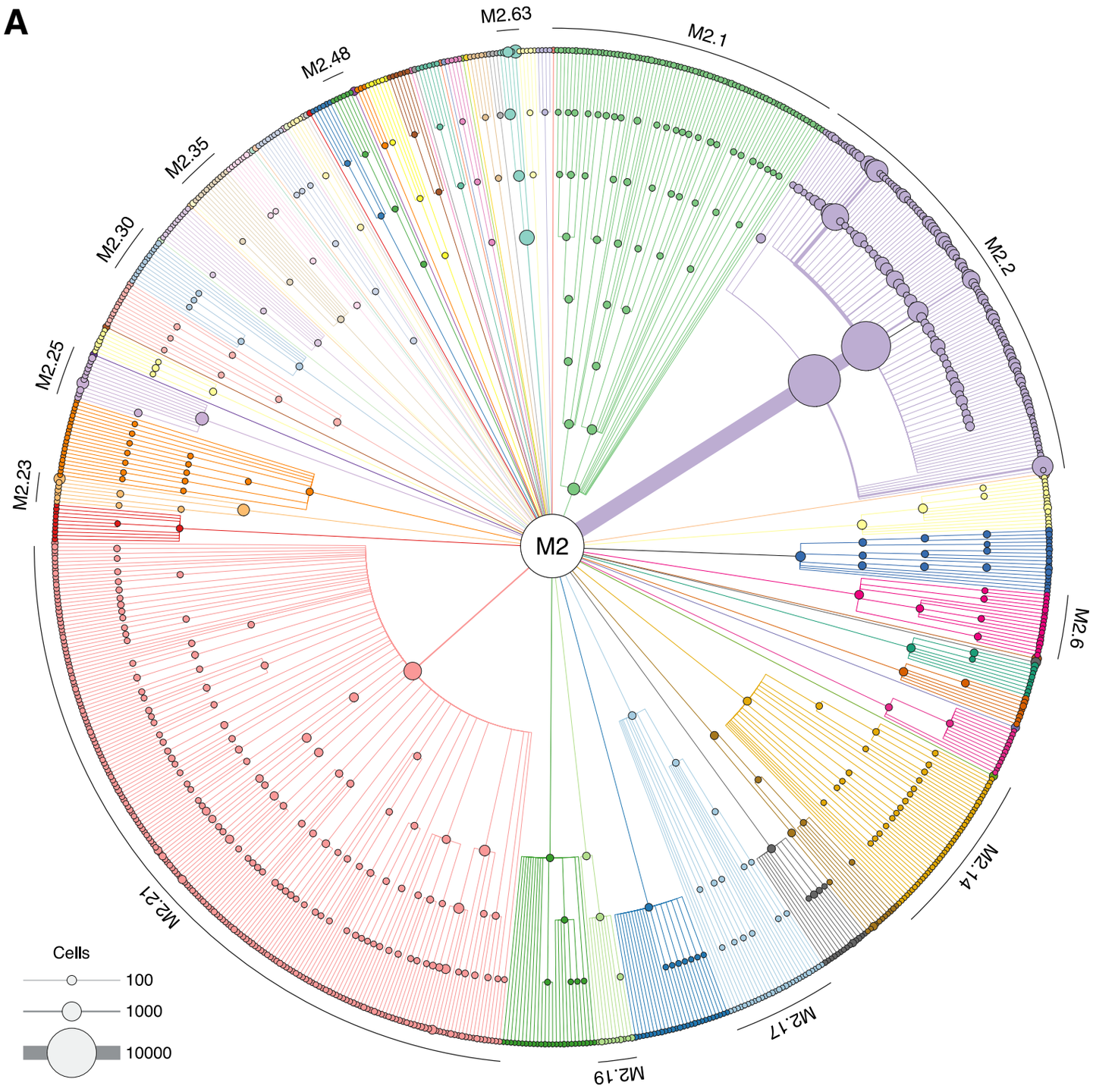
From Figure 3, 5 of Simeonov et al., Cancer Cell (2021)

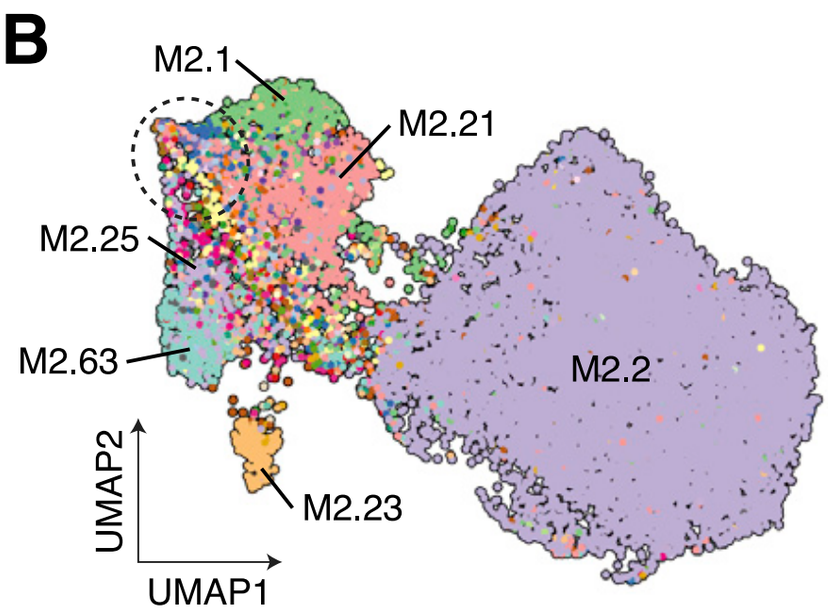
Peak metastatic aggression corresponds to late-hybrid EMT states

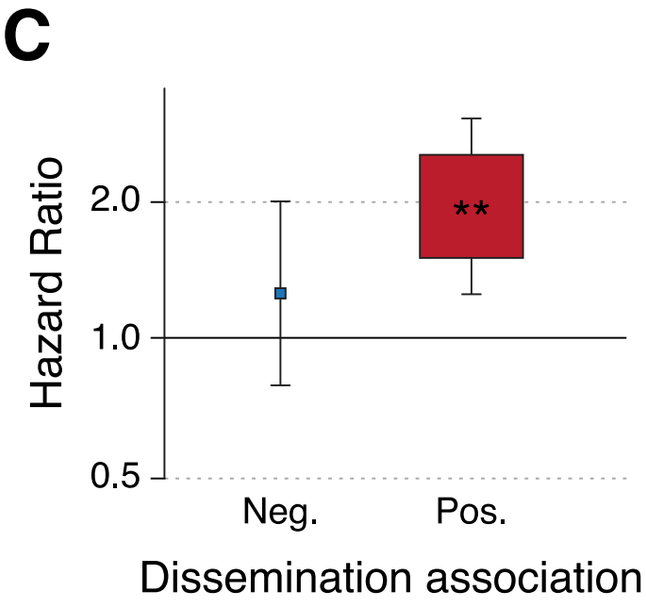


From Figure 5 of Simeonov et al., Cancer Cell (2021)

Complementary process to canonical EMT

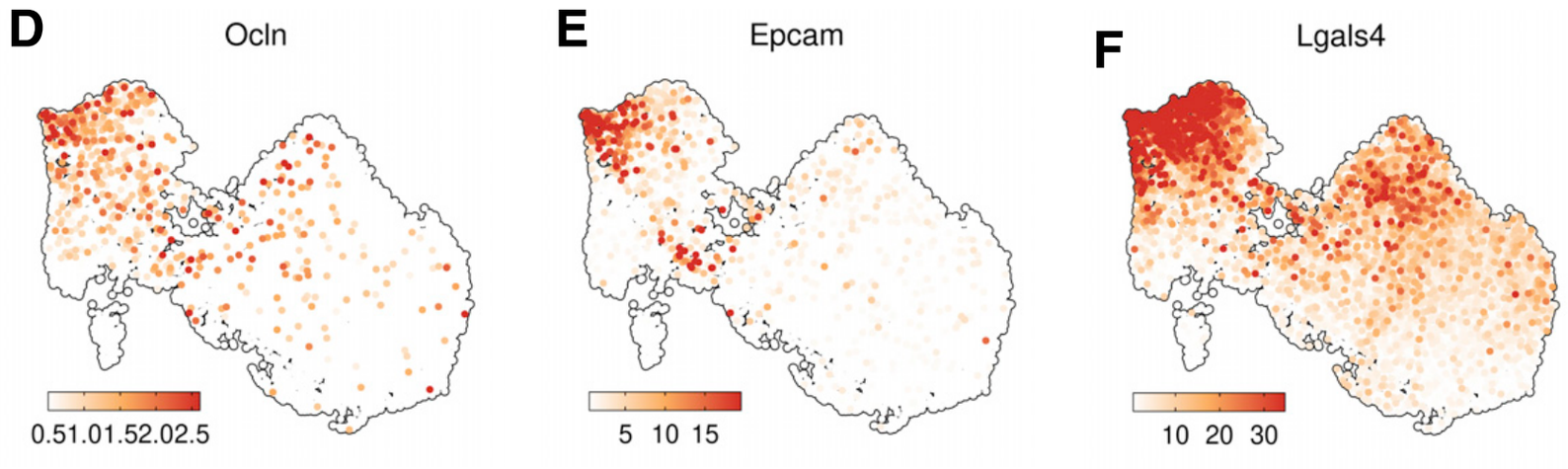




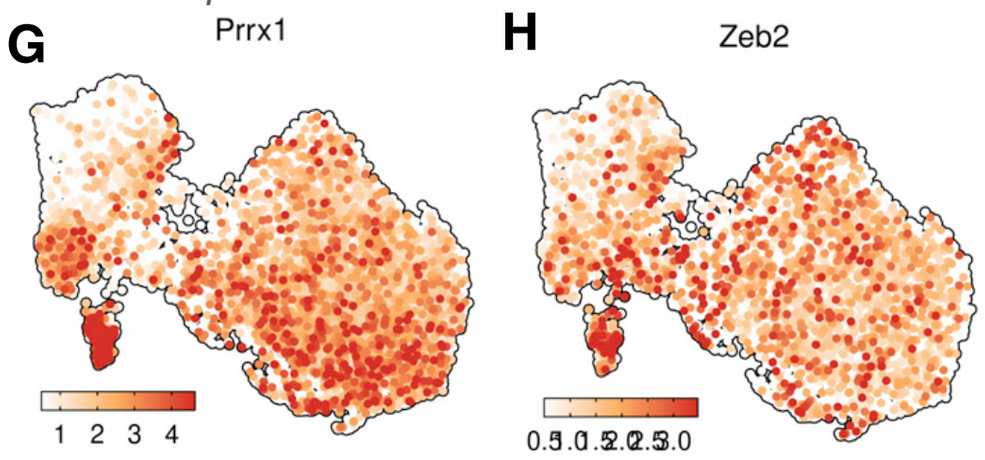


From Figure 6 of Simeonov et al., Cancer Cell (2021)

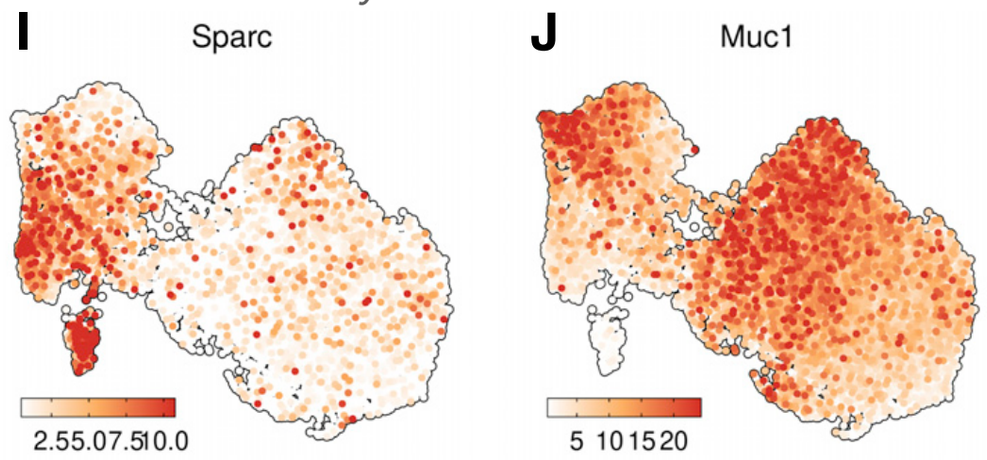
Complementary process to canonical EMT



GEX as expected

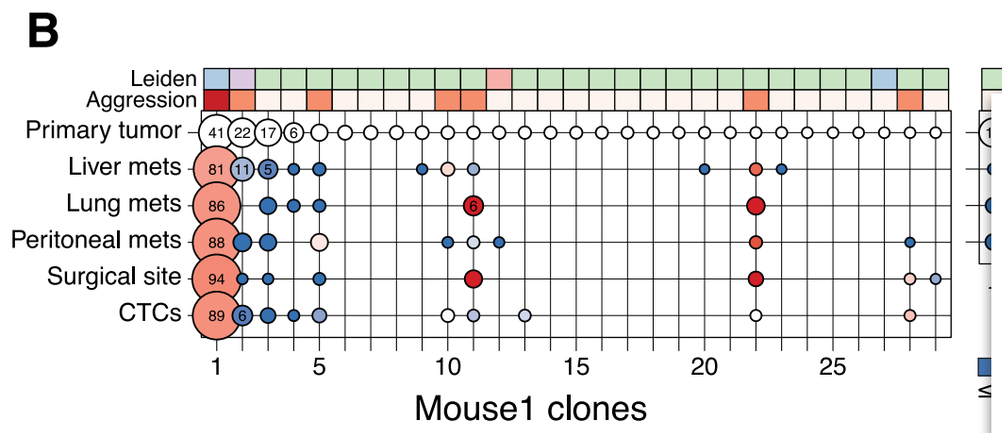


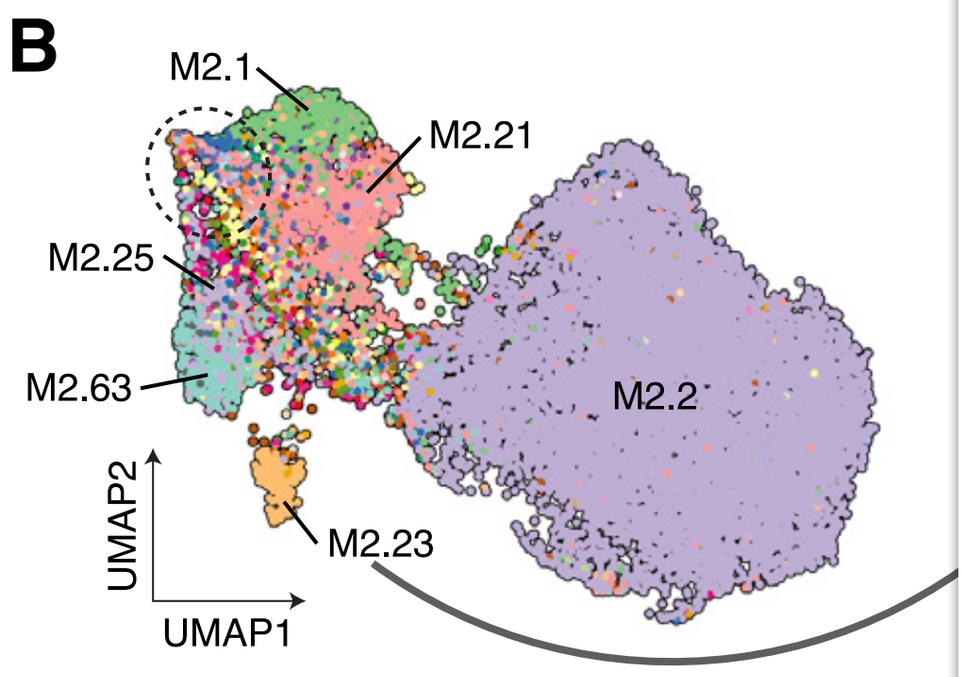
GEX contradictory

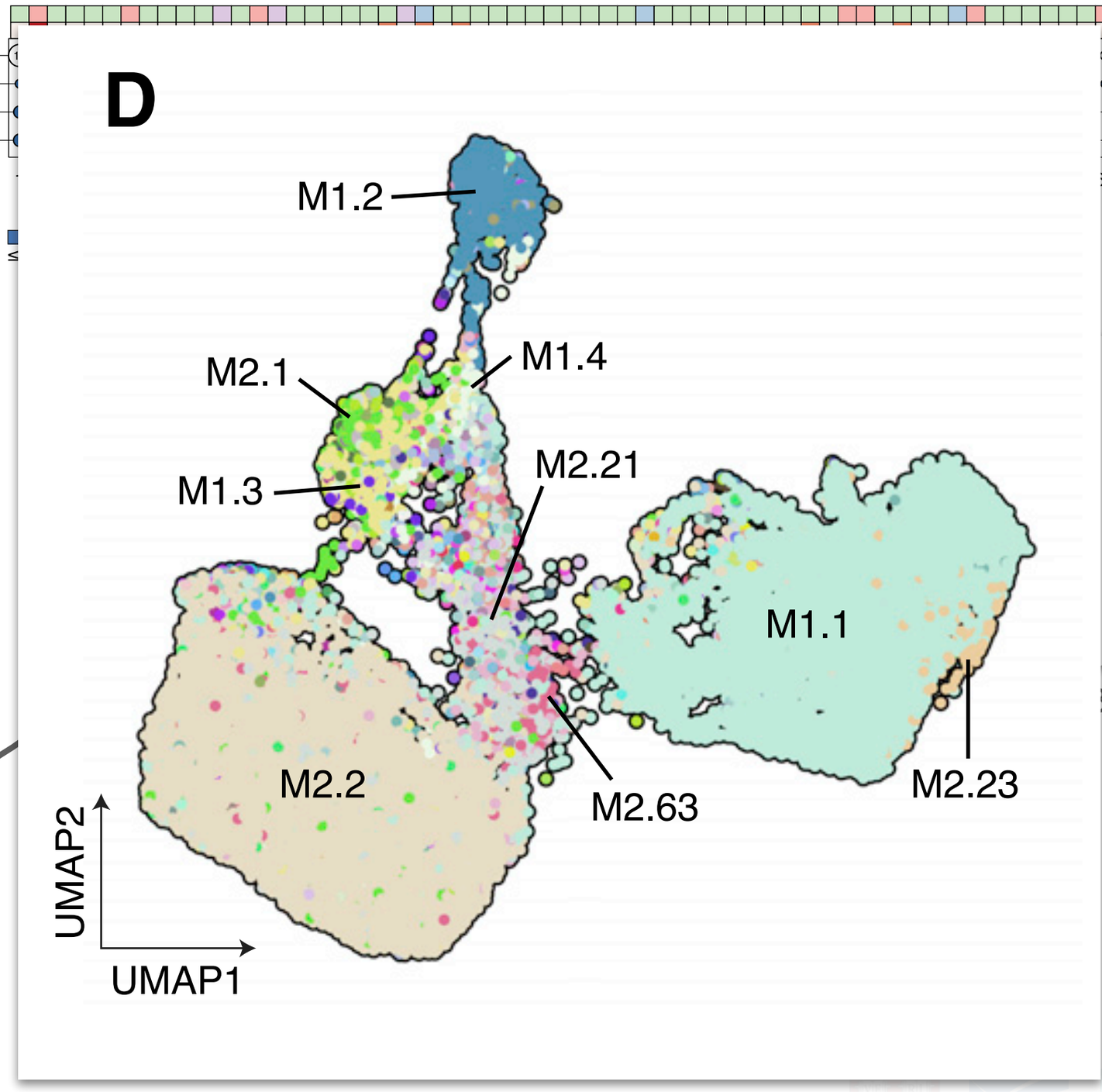


From Figure 6 of Simeonov et al., Cancer Cell (2021)

Complementary process to canonical EMT

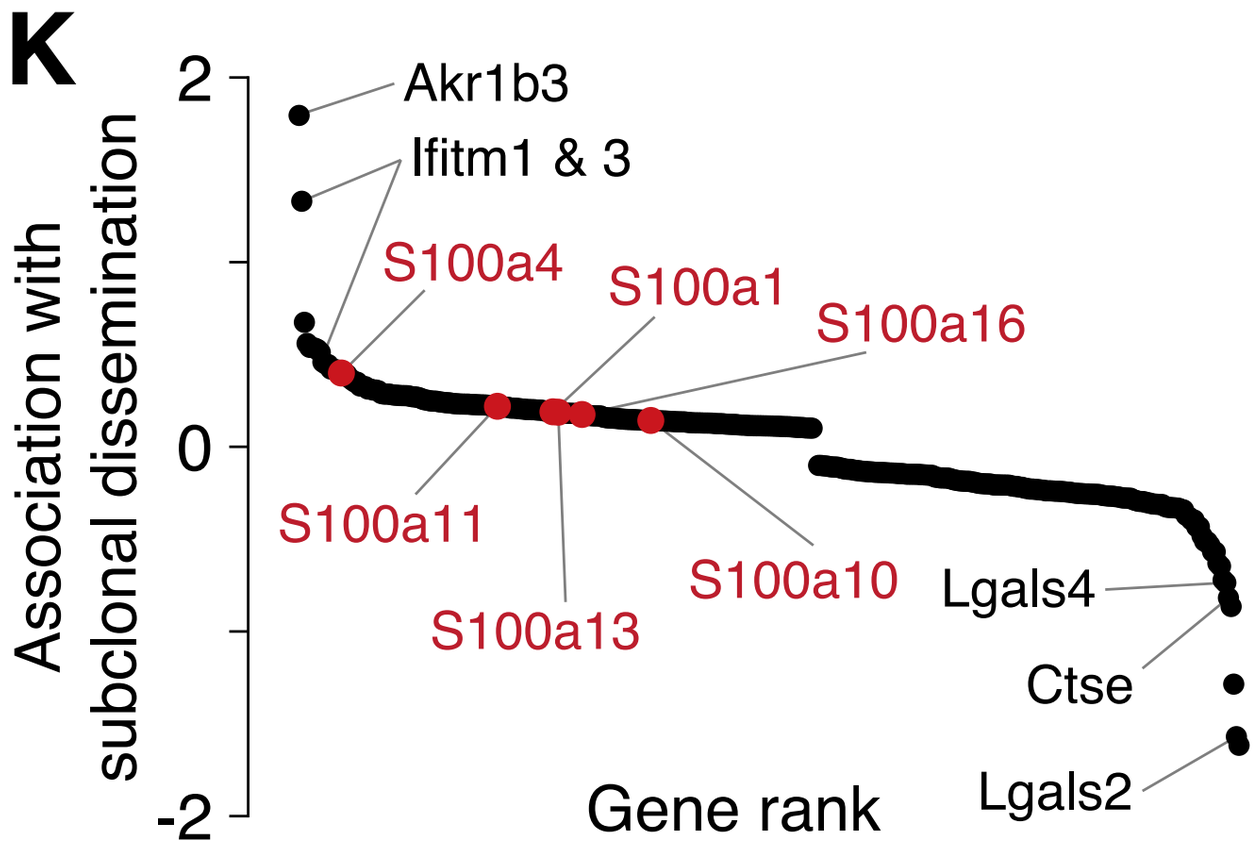


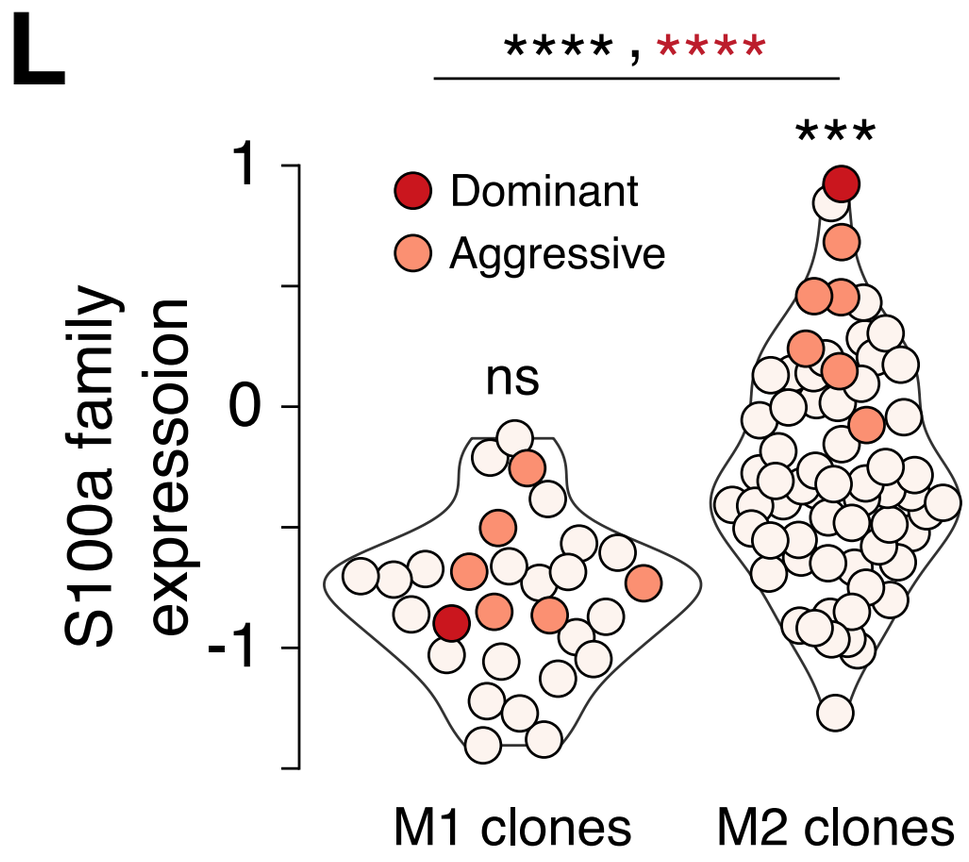




From Figures 2,6 of Simeonov et al., Cancer Cell (2021)

Complementary process to canonical EMT

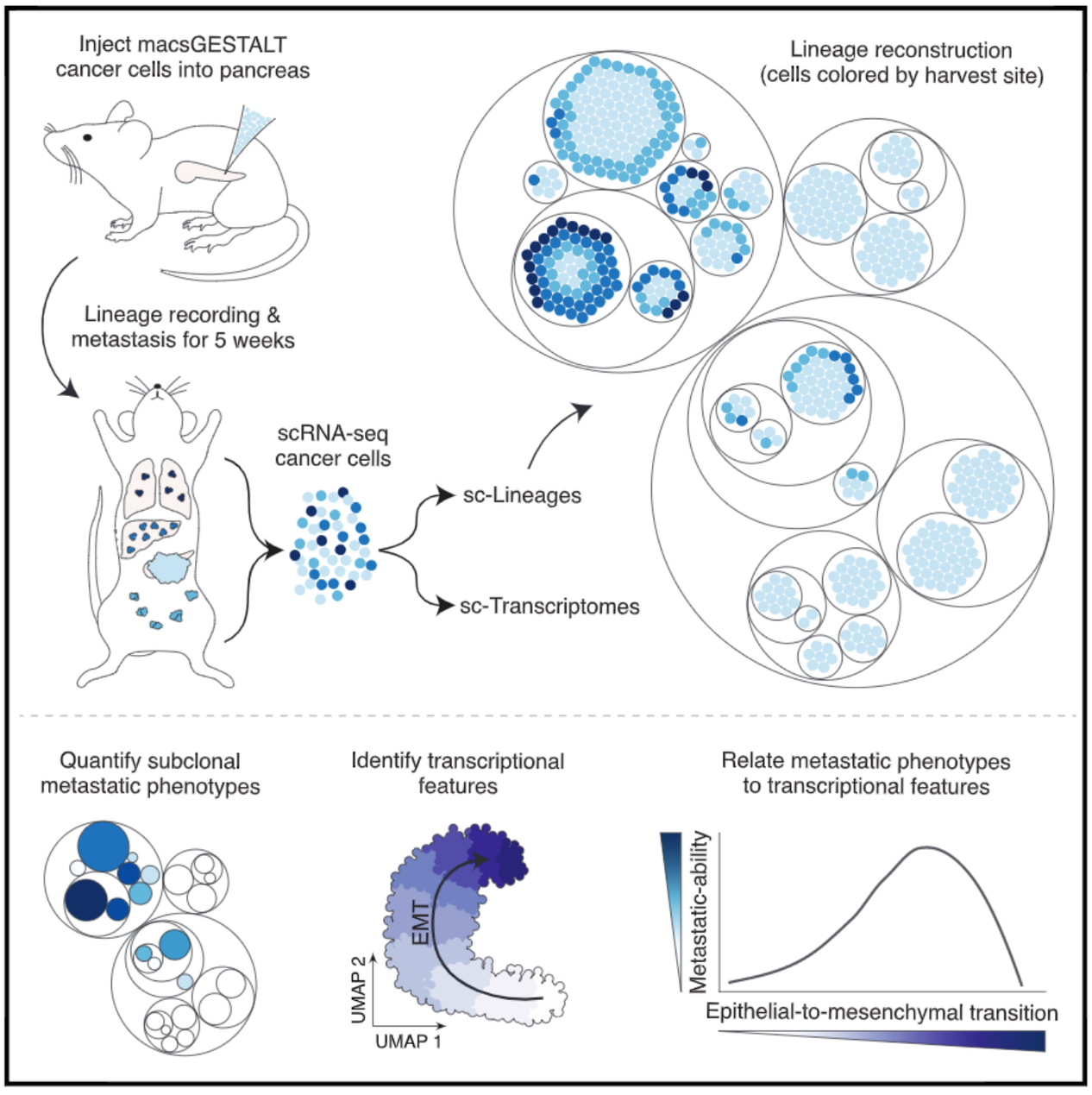




From Figures 6 of Simeonov et al., Cancer Cell (2021)

Major contributions of macsGESTALT

Graphical abstract



- macsGESTALT is an inducible lineage-recorder
- In vivo model of pancreatic cancer metastasis
- Finding recurrent drivers across cancers remains elusive
- Metastatically competent model wherein most clones do NOT metastasize
- In EMT, metastatic aggression rises and peaks during a late intermediate hybrid stage
- ID gene sets within EMT hybrid stages that are predictive of human survival outcome (in 2 cancers, not in 3)
- *S100* genes were found across metastatic subpopulations

Thank you for listening!

Access all sides with links
to papers and notes:

Scan me!

GitHub.com/mvinyard/vintools