Identification and isolation of EMI (Epithelial-mesenchymal intermediate) subpopulation from lung cancer stem cells

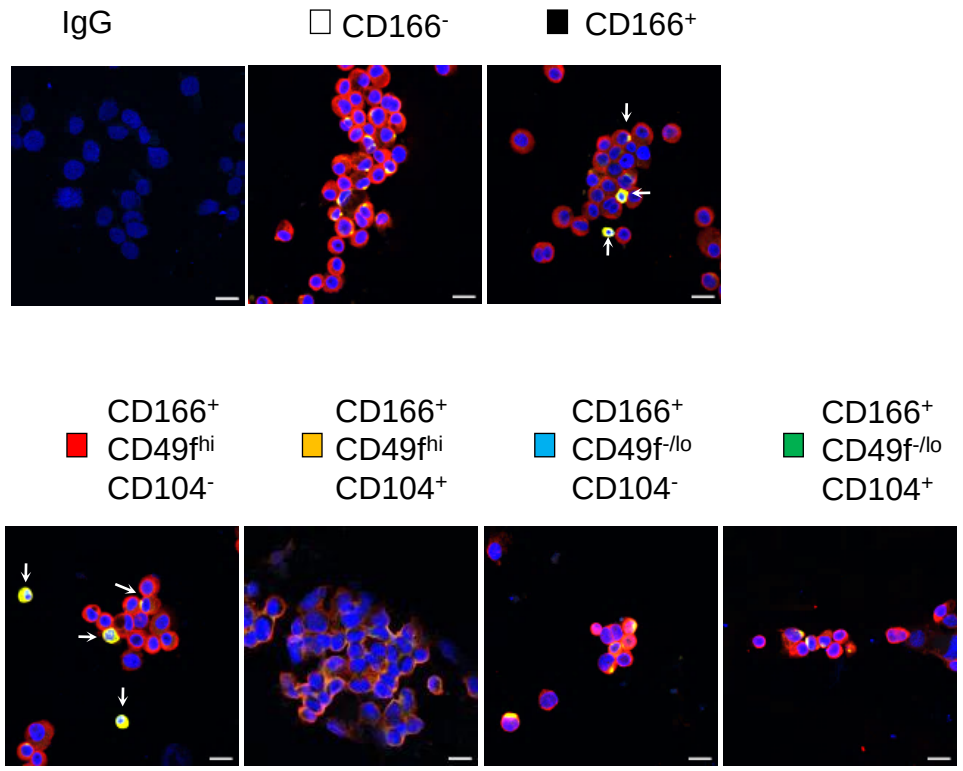
Identification of lung cancer stem cells (LCSCs)

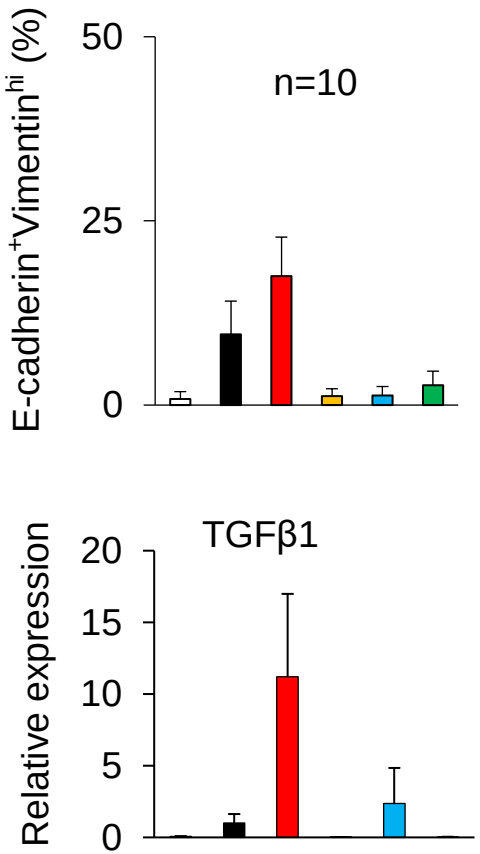
- $CD166^{+}CD49f^{hi}CD104^{-}Lin^{-}$ LCSC
  - Tumorigenic
  - Capable of self-renewal (tumor spheres in vitro, orthotopic lung tumors in immune-compromised mice)
- 1.4%

ALCAM (CD166+)
ITGA6 (CD46fhi)
ITGB4 (CD104-)
PDGFRA (CD140G-, Lin-)
PECAM1 (CD31-, Lin-)

Zhang et al, Cancer Research, 2015

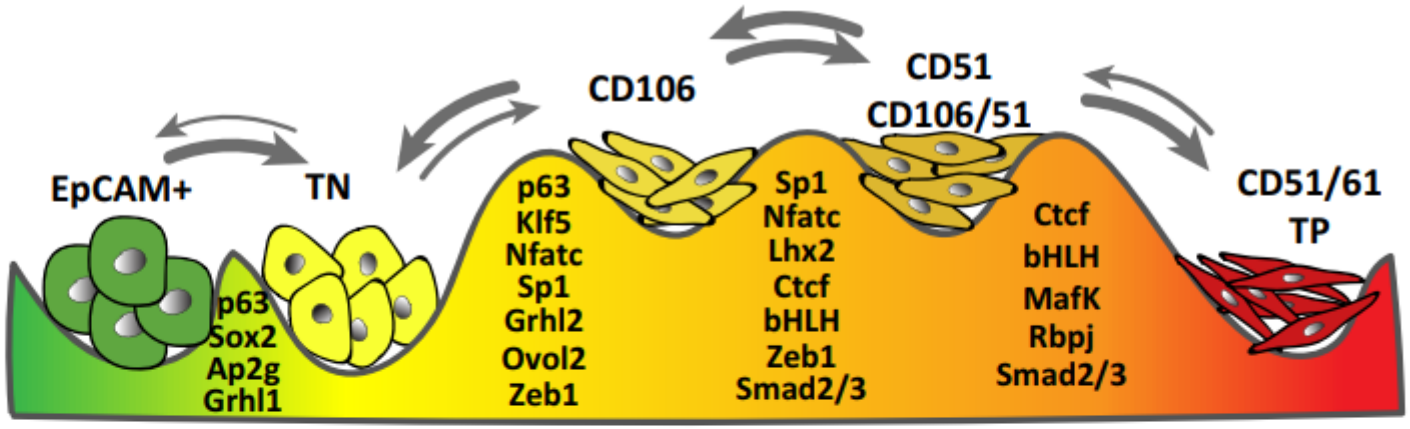
LCSCs Undergo a Special Phase----Epithelial-Mesenchymal Intermediate





Zhang et al, Unpublished

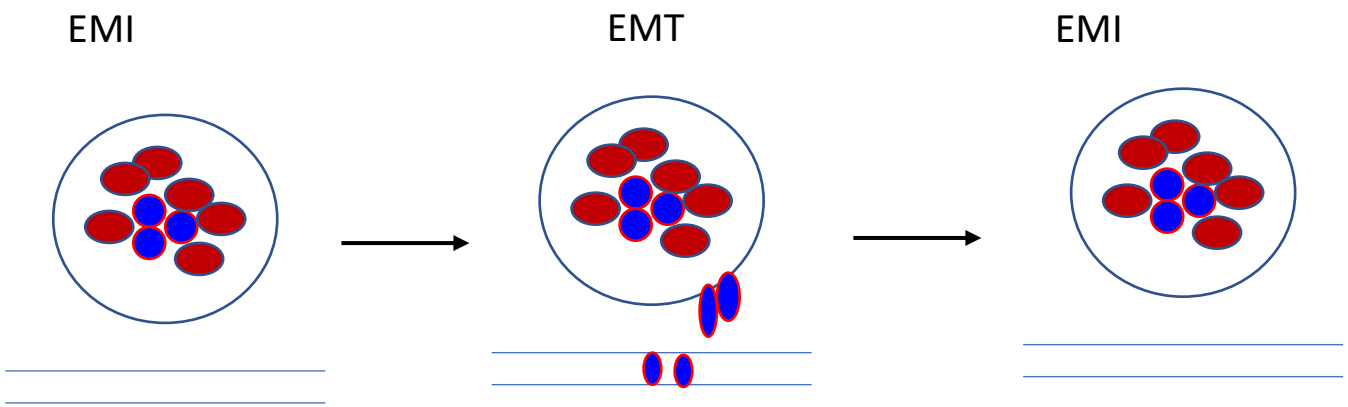
Different characteristics of cells at different stages of EMT



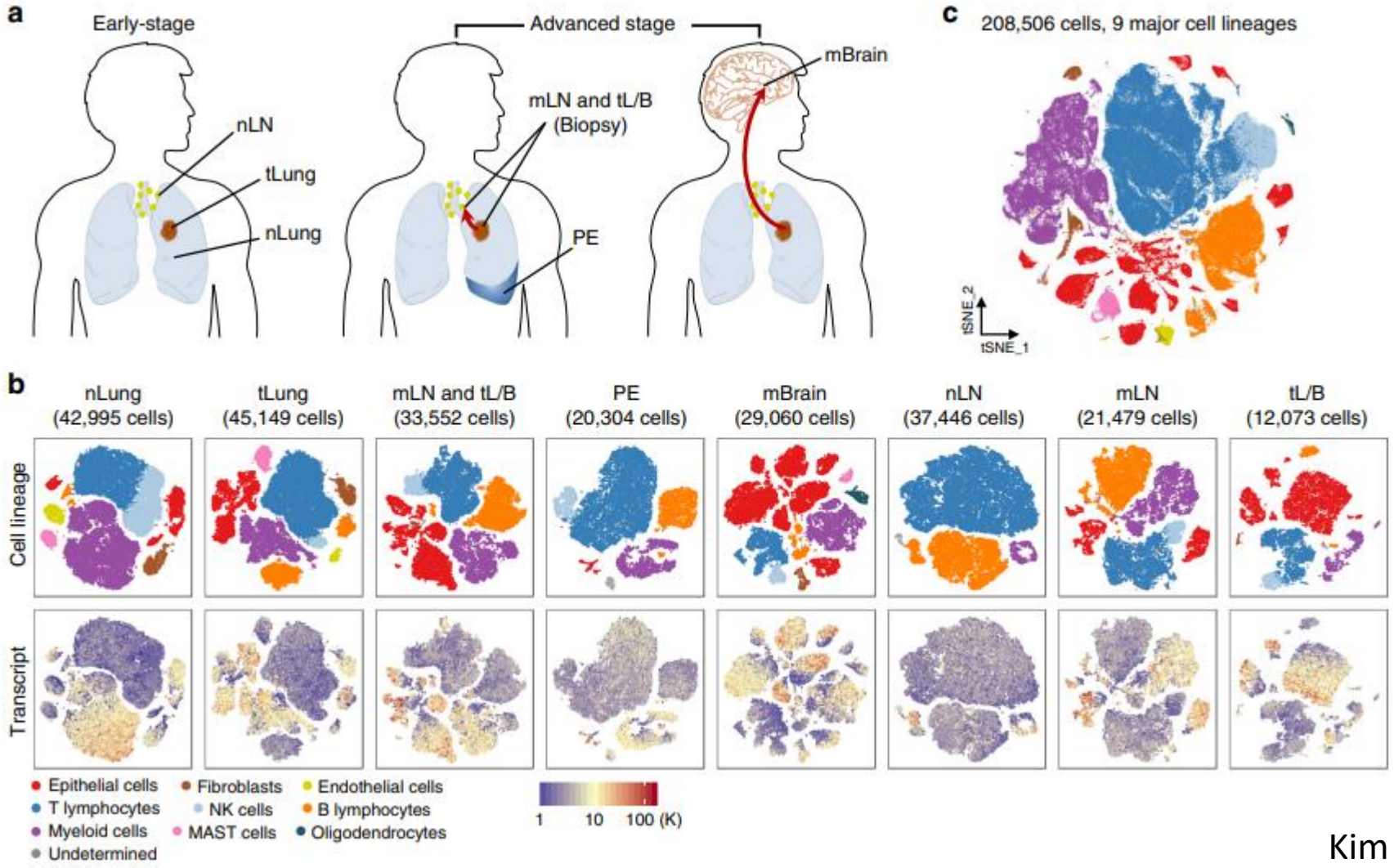
	Epithelial tumor cells	Early hybrid EMT state	Hybrid EMT state	Late hybrid EMT state	Mesenchymal tumor cells
Proliferation	+++++	++++	+++	++	+
Invasion	+	++	+++	++++	+++++
Plasticity	+	++	+++	++++	++
Stemness	+	+++	+++	+++	+++
Metastasis	+	++++	++++	++	+

Pastushenko et al., *Trends in Cell Biology*, 2019

A Hypothetic Model for EMI-EMT



The lung cancer scRNA-seq dataset



- 208,506 cells populating the normal tissues or early to metastatic stage cancer in 44 patients

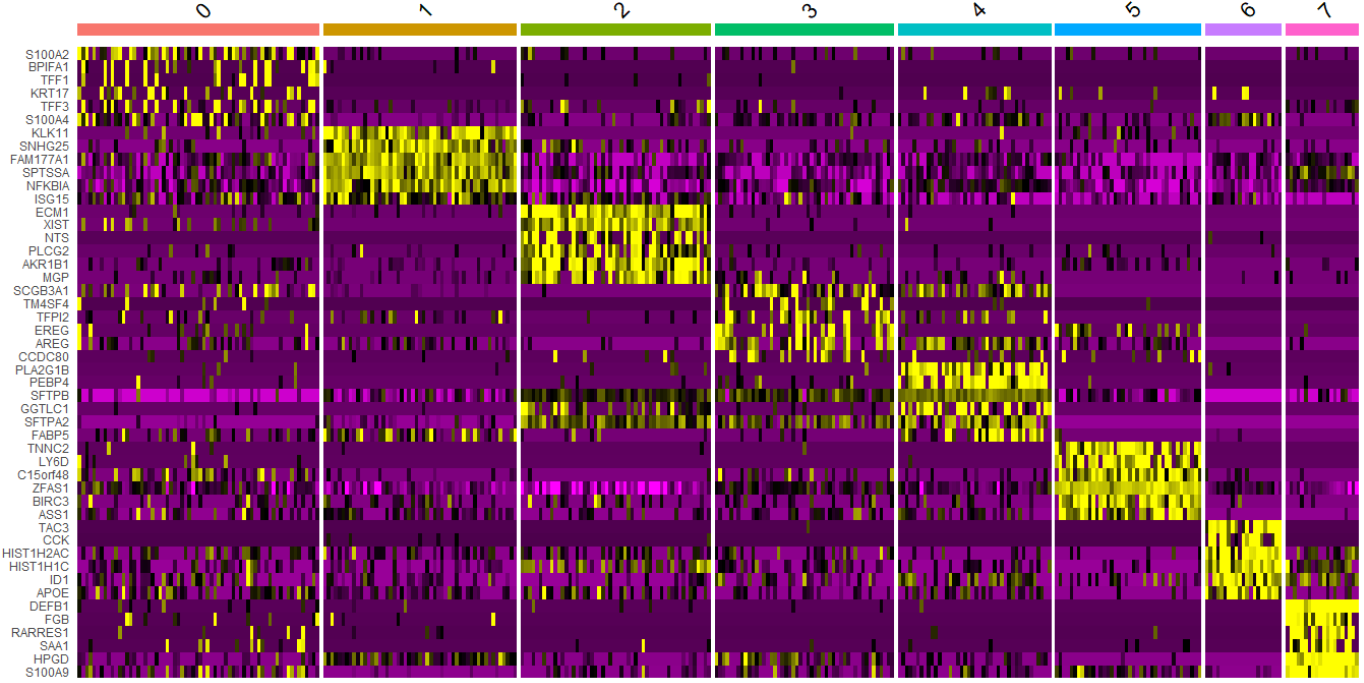
Kim et al, Nature Communications, 2020

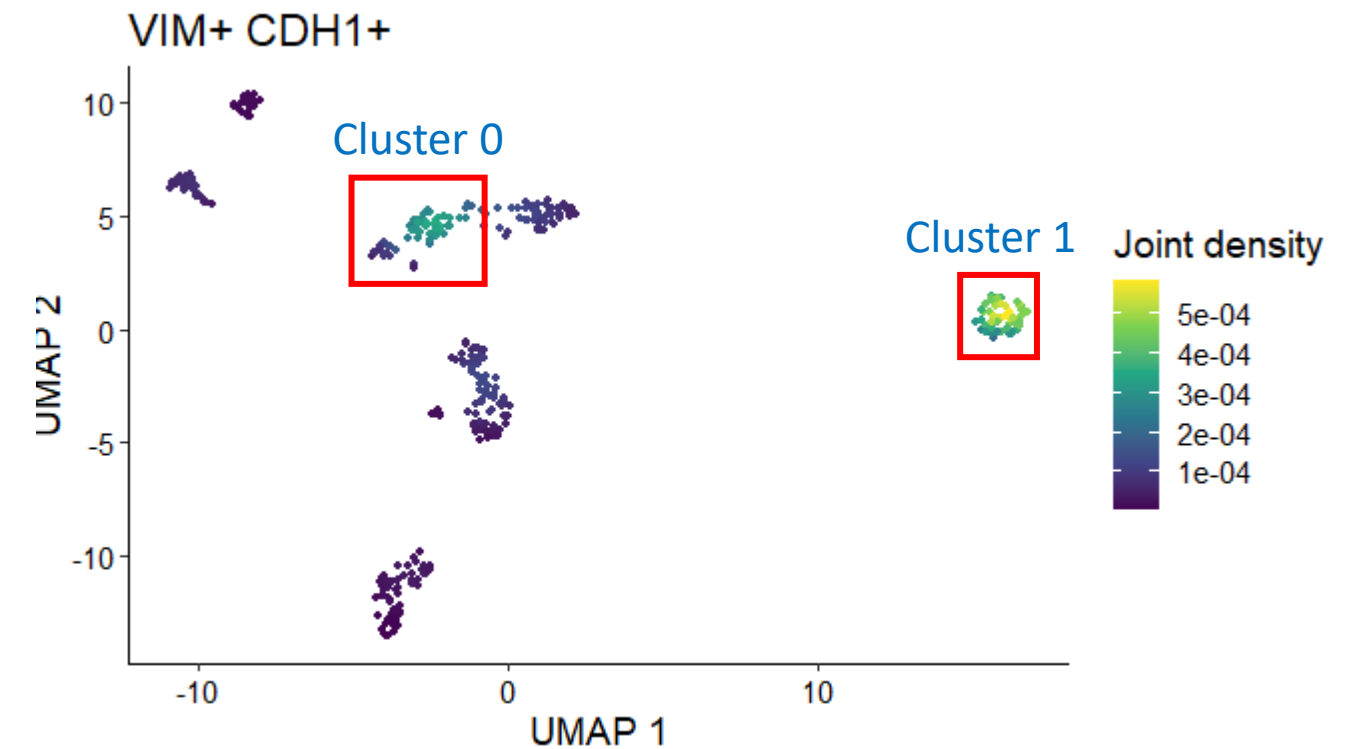
Select LCSCs using artificial threshold

- QC -> Normalization -> feature selection -> Scaling -> dimensional reduction (PCA, UMAP/t-SNE) -> clustering -> markers, DEG analysis…

- The total number of cancer epithelial cells: 36467

- The total number of cancer stem cells: 495
- The threshold:
  - ALCAM > 0.6
  - ITGA6 > 0.6
  - ITGB4 <= 0.01
  - PECAM1 <= 0.01
  - PDGFRA <= 0.01





Label EMI cells in LCSCs

- 2 clusters with both VIM and E-cad expressed
- 99 cells out of 495 cells with VIM > 1 & E-cad > 1

Validate the markers

- DEG selected using: abs(average log2FC) > 1.5 & adj p-value < 0.05
- Literature:
  - Cluster 0: upregulated markers associated with EMT
  - Cluster 1: upregulated markers not very related to EMT (still needs to be checked)
- Gene ontology
  - Cluster 0: downregulated: homeostatic process
  - Cluster 1: upregulated: mitochondria, cellular respiration. Downregulated: immune related

Next steps

- Further analysis of DEGs
  - GSEA using relevant gene sets
- Directly compare transcriptome of EMI cluster with non-EMI cluster
- Threshold is arbitrary, may not mimic FACS?
  - Adjust the threshold