Single cell analysis of lung endothelium identifies distinct sub-populations: implications for genetics of pulmonary arterial hypertension

Katharine Clapham, MD[1],Takaomi Adachi MD PhD[1], and Hyung J. Chun, MD[1].

[1]Yale Cardiovascular Research Center, Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut 06520, USA

Background: Endothelial cells (ECs) are thought to have a central role in the pathogenesis of pulmonary arterial hypertension (PAH). While disease-causing mutations in endothelial genes have been identified in individuals with PAH, the EC subpopulations perturbed by these alterations in signaling pathways have not been fully characterized. The advent of single cell RNA sequencing (scRNA-seq) technology has allowed for the recognition of the diversity of cellular subgroups.

Methods: *Cdh5(PAC)-CreERT^2:Rosa^{mTmG}* reporter mice were generated to GFP label ECs. Two weeks after tamoxifen induction of Cre recombinase, ECs were isolated from lungs by FACS sorting of GFP+ cells. scRNA-seq of ECs was performed with 10X Genomics GemCode technology on the Illumina HiSeq 4000 platform; the results were analyzed with Cell Ranger and R using Seurat. An estimated 3,159 cells were sequenced with 21,799 mean reads per cell.

Results: Eight distinct EC clusters were identified from analysis of the scRNA-seq data (Figure 1). Analyses of genes recently found to be mutated in human subjects with PAH (Graf, et.al., *Nature Communications* (2018)), as well as those associated with PAH in OMIM, grouped these genes into three broad categories: 1) genes expressed in nearly all endothelial subtypes (*Bmpr2, Acvrl1, Eng, Cav1, Sox17, Aqp1,* and*Tgfb1)* (Figure 2A), genes expressed in specific endothelial subpopulations (*Smad1, Kcnk3, Apln,* and *Birc5)* (Figure 2B), and 3) genes not appreciably expressed in lung ECs (*Atp13a3, Eif2ak4, Tbx4, Smad4,* and *Smad9*) (Figure 2C).

Conclusions: scRNA-seq allows for identification of EC populations with diverse gene expression profiles. Subcellular gene expression analysis demonstrates heterogeneity in the expression of PAH related genes amongst EC subpopulations. These findings will facilitate future validation and characterization of specific genes and their mutations in pulmonary vascular cell subtypes in PAH.

Figure 1: t-SNE plot of EC clusters.

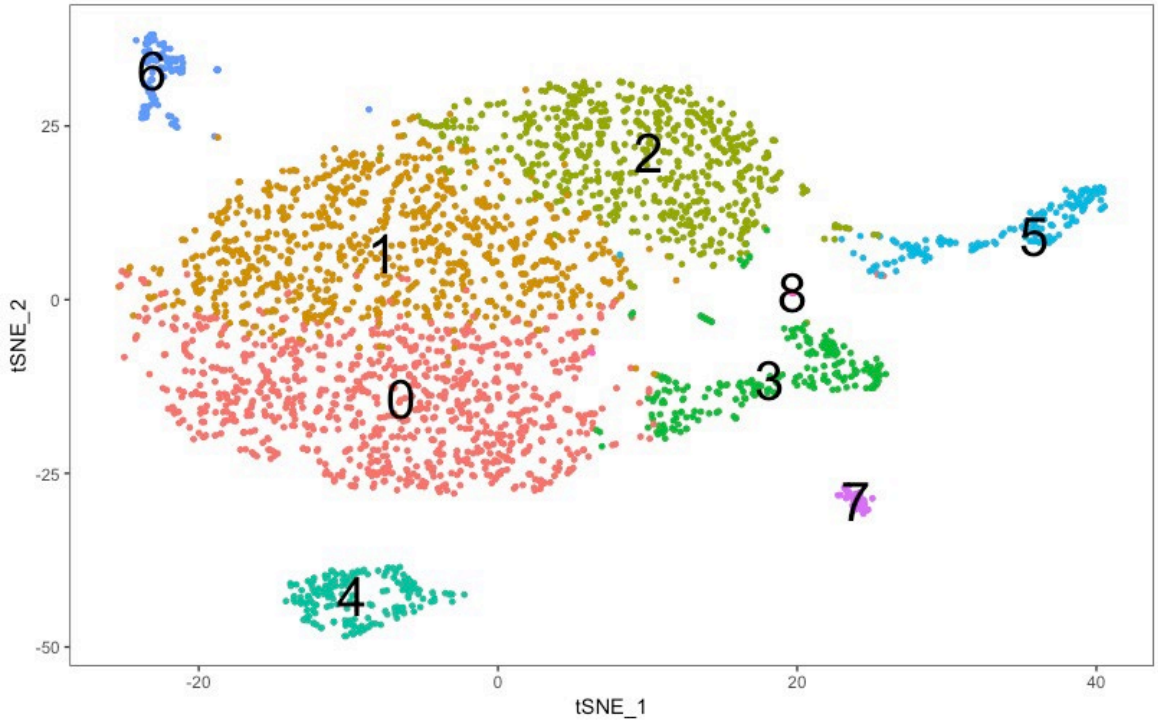
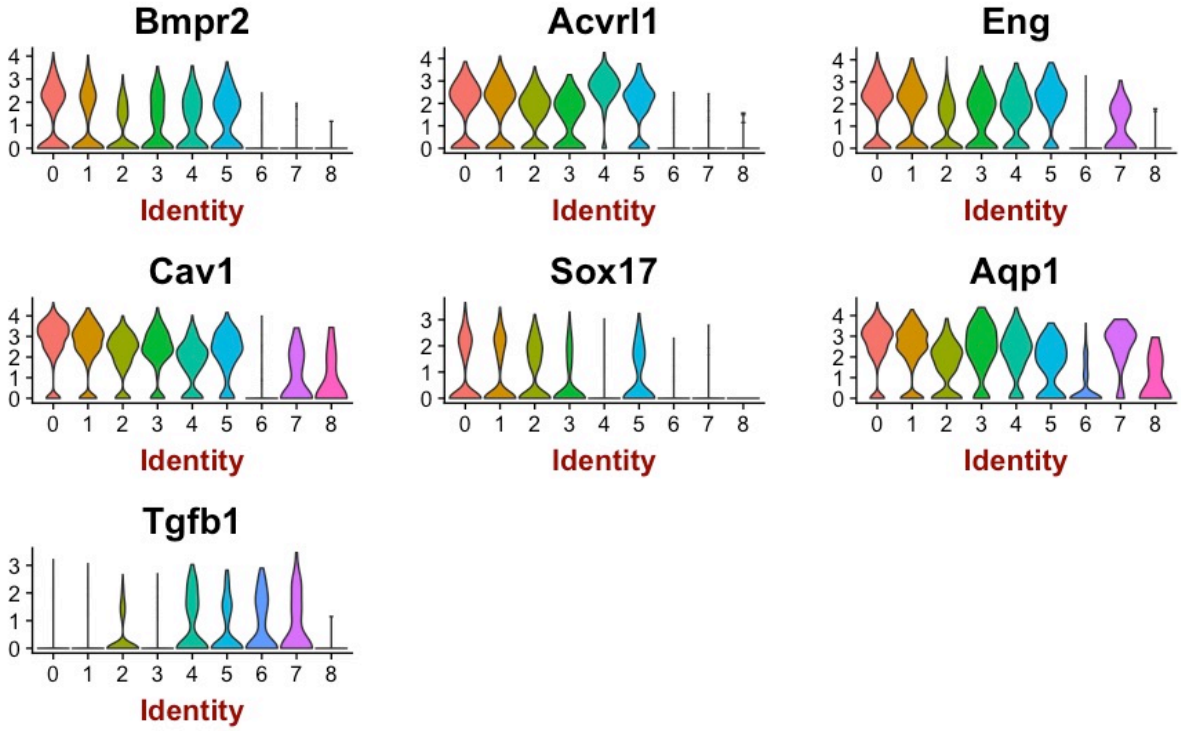
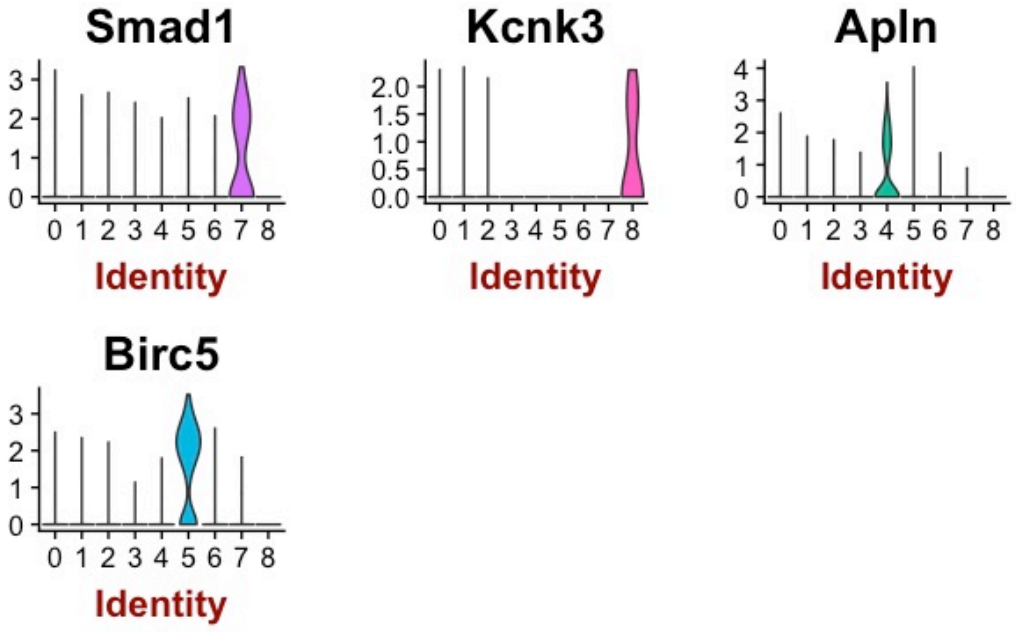


Figure 2: Violin plot of gene expression by EC cluster.
A.



B.



C.