bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# Single-cell RNA sequencing-based characterization of resident lung mesenchymal stromal cells in bronchopulmonary dysplasia

I. Mižíková[1,2], F. Lesage[1,2], C. Cyr-Depauw[1,2], D. P. Cook [2,3], M. Hurskainen[1,2,4,5], S.M. Hänninen[6], A. Vadivel[1], P. Bardin[1,2], S. Zhong[1], O. Carpen[6], B. C. Vanderhyden[2,3,8], B. Thébaud[1,2,7]

[1] Sinclair Centre for Regenerative Medicine, Ottawa Hospital Research Institute, Ottawa, ON, Canada

[2] Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada

[3] Cancer Therapeutics Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

[4] Division of Pediatric Cardiology, New Children's Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, Finland

[5] Pediatric Research Center, New Children's Hospital, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[6] Precision Cancer Pathology, Department of Pathology and Research Program in Systems Oncology, University of Helsinki and HUS Diagnostic Center, Helsinki University Hospital, Helsinki, Finland

[7] Department of Pediatrics, Children's Hospital of Eastern Ontario (CHEO) and CHEO Research Institute, University of Ottawa, Ottawa, Ontario, Canada

[8] Department of Obstetrics and Gynecology, University of Ottawa/The Ottawa Hospital, Ottawa, Ontario, Canada

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**ABSTRACT (limit 250 words)**

Late lung development is a period of alveolar and microvascular formation, which is pivotal in ensuring sufficient and effective gas exchange. Defects in late lung development manifest in premature infants as a chronic lung disease named bronchopulmonary dysplasia (BPD). Numerous studies demonstrated the therapeutic properties of exogenous bone marrow and umbilical cord-derived mesenchymal stromal cells (MSCs) in experimental BPD. However, very little is known regarding the regenerative capacity of resident lung MSCs (L-MSCs) during normal development and in BPD. In this study we aimed to characterize the L-MSC population in homeostasis and upon injury. We used single-cell RNA sequencing (scRNA-seq) to profile *in situ Ly6a*$^{+}$ L-MSCs in the lungs of normal and $O_2$-exposed neonatal mice (a well-established model to mimic BPD) at three developmental timepoints (postnatal days 3, 7 and 14). Hyperoxia exposure increased the number, and altered the expression profile of L-MSCs, particularly by increasing the expression of multiple pro-inflammatory, pro-fibrotic, and anti-angiogenic genes. In order to identify potential changes induced in the L-MSCs transcriptome by storage and culture, we profiled 15,000 *Ly6a*$^{+}$ L-MSCs after *in vitro* culture. We observed great differences in expression profiles of *in situ* and cultured L-MSCs, particularly those derived from healthy lungs. Additionally, we have identified the location of L-MSCs in the developing lung and propose *Serpinf1* as a novel, culture-stable marker of L-MSCs. Finally, cell communication analysis suggests inflammatory signals from immune and endothelial cells as main drivers of hyperoxia-induced changes in L-MSCs transcriptome.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 1. INTRODUCTION

Late lung development represents an important period in lung maturation marked by an exponential increase in the gas exchange surface area by forming the most distal respiratory units, the alveoli. Within these units, respiration takes place across a thin (0.2 - 2μm) alveolo-capillary barrier. Formation of alveolar structures, a process known as alveolarization, is facilitated by spatially and temporarily coordinated interactions between diverse cell types and the pulmonary microenvironment [1]. Defects in late lung development in humans manifest as bronchopulmonary dysplasia (BPD), a multifactorial disease occurring as a consequence of premature birth, respiratory distress, and associated treatments in neonatal intensive care. BPD is the most common chronic disease in children and a leading cause of death in children under the age of 5 [1,2]. BPD is also associated with neurodevelopmental delay, increased incidence of asthma, re-hospitalizations and early-onset emphysema [3,4].

To date, multiple studies have demonstrated the lung protective effects of exogenous, bone marrow (BM)- or umbilical cord (UC)-derived, mesenchymal stromal cells (MSCs) in experimental BPD models [5–10]. The discovery of lung resident (L-)MSCs prompted questions regarding the apparent insufficient regenerative capacity of L-MSCs in lung injury [11]. Characterizing the L-MSC population in homeostasis and upon injury is pivotal in understanding the apparent contradiction between the therapeutic effects of exogenous MSCs, while the resident population fails to prevent neonatal lung injury from occurring. However, very little is currently known about the role of L-MSCs in postnatal lung development and in BPD. Lung stromal cells, including lipofibroblasts, myofibroblasts and matrix fibroblasts are a potent source of inter-cellular signaling and are known to play an important role in BPD pathogenesis [12]. However, how L-MSCs communicate with other cell populations and contribute to the development of BPD remains unknown.

While most authors report that L-MSCs can differentiate, to some extent, into chondroblasts, osteoblasts and adipocytes [13], form colonies *in vitro* [13,14], and express classical MSC markers THY1 (CD90), NT5E (CD73) and ENG (CD105) [13,15], no L-MSC-specific marker has yet been established. Due to the lack of standardization for L-MSC identification, as well as differences in expression profiles between species, no single marker has been broadly accepted. Lung mesenchymal progenitor cell markers have been proposed [13,15–

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

18], including LY6A, often referred to as SCA-1 (Stem cell antigen 1) [16–19]. LY6A was proposed as a defining progenitor marker for mesenchymal cell lineages in the lung [19] and $LY6A^+$ mesenchymal lung cells were shown to promote colony formation, proliferation and differentiation of epithelial progenitor cells [20].

In the study presented here we identify, for the first time, the transcriptome of *Ly6a*$^+$ L-MSCs in heathy and diseased developing mouse lungs. We hypothesized, that $O_2$-exposure (a well-established model to mimic BPD) significantly impacts the phenotype and function of L-MSCs, as well as cellular communication between L-MSCs and other cell populations in the developing lung. We identify perturbations to the phenotype and functional properties of L-MSCs in this model. Furthermore, we report extensive single-cell RNA sequencing (scRNA-seq) profiling of L-MSCs in the lungs of 36 healthy and $O_2$-exposed mice at three developmental timepoints (P3, P7, and P14). Finally, we investigate cultured *Ly6a*$^+$ L-MSCs and *Ly6a*$^-$ mouse lung stromal cells by scRNA-seq. We identify changes in L-MSCs transcription profile induced by storage and culture and present novel, culture-stable marker for this rare progenitor population.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 2. MATERIALS AND METHODS

### 2.1 Experimental animals

Pregnant C57BL/6N mice were purchased from Charles Rivers Laboratories, Saint Constant, QC, Canada at embryonic day (E)15. Mice were housed by the Animal Care and Veterinary Service of the University of Ottawa in accordance with institutional guidelines. All study protocols were approved by the animal ethics and research committee of the University of Ottawa (protocol OHRI-1696) and conducted according to guidelines from the Canadian Council on Animal Care (CCAC). Mouse pups born on the same day, were randomized at the day of birth [postnatal day (P)0] and divided into equal-sized litters of 6-8 pups/cage. Cages were then maintained either in room air (normoxia, 21% $O_2$), or in normobaric hyperoxia (85% $O_2$) until the day of harvest. The hyperoxic environment was maintained in sealed plexiglass chambers with continuous oxygen monitoring (BioSpherix, Redfield, NY). Mice were maintained in 12/12 hours light/dark cycle and received food ad libidum. In order to avoid confounding factors associated with oxygen toxicity, nursing dams were rotated between normoxic and hyperoxic group every 48 hours. Euthanasia was performed by an intraperitoneal (i.p.) injection of 10 µl/g Pentobarbital Sodium (CDMV, Saint-Hyacinthe, QC, Canada).

### 2.2 Lung isolation

Mouse pups designated for mean linear intercept (MLI) assessment or fluorescent in situ hybridization (FISH) were euthanized at P7 and P14, respectively. Following euthanasia, the chest was opened, mice were tracheotomized and lungs were installation-fixed for 5 minutes at 20cm $H_2O$ hydrostatic pressure. Lungs designated for histological assessment were fixed with 1.5% (w/v) paraformaldehyde (PFA) (Sigma-Aldrich, Oakville, ON, Canada) and 1.5% (w/v) glutaraldehyde (Sigma-Aldrich, Oakville, ON, Canada) in 150mM HEPES (Sigma-Aldrich, Oakville, ON, Canada). Lungs designated for FISH were fixed with 4% (w/v) PFA (Sigma-Aldrich, Oakville, ON, Canada). In both instances, lungs were kept in the fixation solution for 48 hours at 4°C and collected for embedding in paraffin. Paraffin-embedded tissue blocks designated for histological analyses were sectioned at 3 or 4µm as needed. Tissue dehydration, paraffin embedding and sectioning were performed by the University of Ottawa Louise Pelletier Histology Core Facility.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Mouse pups designated for lung cells isolation and fluorescence activated cell sorting (FACS) analyses were euthanized at P7. Mice also received an i.p. injection of 10 mU/g Heparin Sodium (LEO Pharma INc., Thornhill, ON, Canada). Following euthanasia, the chest was opened and the left atrium was perforated. Lungs were perfused through the right ventricle with 5 ml of 25 U/ml Heparin Sodium in DPBS supplemented with $Mg^{2+}$/$Ca^{2+}$ (ThermoFisher Scientific, Burlington, ON, Canada) until white. Lungs were removed from the thoracic cavity, dissected into individual lobes, and digested in enzyme mix at 37°C by gentleMACS™ Octo Dissociator (Miltenyi Biotech, Bergisch Gladbach, Germany). The detailed procedure, as well as enzyme mixture contents are provided in Supplementary Method S1. The suspension was then centrifuged and the resulting pellet was washed with 5 ml of 5% FBS (Sigma-Aldrich, Oakville, ON, Canada) in 1× DPBS (Lonza, Basel, Switzerland), filtered through 70 µm filter (Corning Life Sciences, Tewksbury, MA, USA) and centrifuged again. The resulting pellet was resuspended in 1ml of cold RBC lysis buffer (ThermoFisher Scientific, Burlington, ON, Canada) for 3 minutes at room temperature (RT). The cell suspension was then diluted with 5ml of 5% FBS solution, centrifuged and washed twice.

A detailed flowchart illustrating the allocation of each mice to respective experimental groups is depicted in Supplementary figure 1.

### 2.3 Mean linear intercept (MLI) measurement

Paraffin-embedded tissue blocks were sectioned at 4µm, stained with hematoxylin and eosin (H&E) stain, and scanned using the Axio Scan.Z1 (Zeiss, Oberkochen, Germany). The mean linear intercept (MLI) was estimated with Fiji/ImageJ software using a 64-point grid as described previously [21]. A total of 20 randomly selected 500µm×500µm fields of view were assessed in each lung.

### 2.4 Fluorescent activated cell sorting (FACS)

The number of cells in the single-cell suspension was estimated using the EVE NanoEnTek automatic cell counter and a total of $1\times10^6$ cells/sample were resuspended in 550 µl of FACS buffer (5% (v/v) FBS and 1mM EDTA in 1×DPBS). Cells were then incubated at RT in the dark

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with 2 µl/1×$10^6$ cells of CD16/32 antibody for 15 minutes. Following blocking, cells were centrifuged and pellets were resuspended in 1:100 mixture of panel of antibodies: FITC-CD31, AF647-CD45, Pe/Cy7- CD326, and BV421- LY-6A/E (Supplementary table 1). Cells were incubated with antibodies for 20 minutes in dark at RT, pelleted and washed 3x with FACS buffer. FACS was performed immediately using a MoFlo XDP (XDP, Beckman Coulter, Fullerton, CA, USA) and compensation and analysis was done using Summit v.5.4 at the Ottawa Hospital Research Institute (OHRI) StemCore facility.

### 2.5 Cell culture and storage

The detailed procedure is provided in Supplementary Method S2.

### 2.6 Colony formation assay

The detailed procedure is provided in Supplementary Method S3.

### 2.7 MSCs surface marker profiling

Cultured, passage 3 $CD31^-/CD45^-/EpCAM^-/LY6A^+$ L-MSCs were profiled for MSC surface markers by flow cytometry. Briefly, 3×$10^5$ cells/sample were resuspended in 200 µl of FACS buffer in 96-well plate and incubated at RT in the dark with 2 µl/1×$10^6$ cells of CD16/32 antibody for 15 minutes. Cells were then divided to 3 equal fractions, centrifuged and resuspended in one of the following 1:100 mixture of antibodies: i) BV421-CD31, Pe/Cy7-conjugated CD326, PE-CD73, and AF488- D105; ii) AF647- CD45, BV421-LY-6A/E, PE-conjugated CD34, and AF488-CD146; iii) PB-CD90.2 (Supplementary table 1). Cells were incubated with antibodies for 20 minutes in dark at RT, pelleted and washed 3x with FACS buffer. Flow cytometry was performed immediately using a MoFlo XDP (XDP, Beckman Coulter, Fullerton, CA, USA) and compensation and analysis was done using Summit v.5.4 at the OHRI core facility.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

### 2.8 Osteogenic differentiation

The detailed procedure is provided in Supplementary Method S4.

### 2.9 Adipogenic differentiation

The detailed procedure is provided in Supplementary Method S5.

### 2.10 Chondrogenic differentiation

The detailed procedure is provided in Supplementary Method S6.

### 2.11. Fluorescent in situ hybridization

The detailed procedure, as well as a list of used probes are provided in Supplementary Method S7.

### 2.12. Multiplexing samples for scRNA-seq

Multiplexing was performed according to the MULTI-seq protocol [22]. The detailed procedure is provided in Supplementary Method S8.

### 2.13. scRNA-seq library preparation and sequencing

Single-cell suspensions were processed using the 10x Genomics Single Cell 3' v3 RNA-seq kit by Ottawa Hospital Research Institute Stem Core Laboratories. Gene expression libraries were prepared according to the manufacturer's protocol. MULTI-seq barcode libraries were retrieved from the samples and libraries were prepared independently, as described previously[22]. Final libraries were sequenced on the NextSeq500 platform (Illumina) to reach an approximate depth of 20,000-25,000 reads/cell.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 2.14. scRNA-seq data analyses and quantification

### Processing and demultiplexing

Raw sequencing reads were processed using CellRanger v3.0.2 for lung homogenate sample and v3.1.0 for cultured cells, aligning reads to the mm10 build of the mouse genome. Except for explicitly setting --expect-cells=25000, default parameters were used for all samples. MULTI-seq barcode libraries were trimmed prior to demultiplexing to 28bp using Trimmomatic (v0.36). Demultiplexing was performed using the deMULTIplex R package (v1.0.2) as described previously[22,23]. Only cells positive for a single barcode were kept for further analysis and sample annotations were added to all cells in the data set.

### Quality control, integration, and clustering

All main processing steps were performed with Seurat v.4.0.0[24]. Quality control was performed independently on each library to find appropriate filtering thresholds. Expression matrices were loaded as Seurat objects into R. Only cells with > 200 genes detected and < 20% of UMIs mapped to mitochondrial genes were retained. Each unique sample was split based on MULTI-seq sample barcodes into a separate Seurat object. SCTransform[25] was used to normalize samples, select highly variable genes, and to regress out cell cycle and cell stress effects. To eliminate batch effects or biological variability effects on clustering, the data integration method implemented by Seurat for SCTransform-normalized data was performed, using the SelectIntegrationFeatures(), PrepSCTIntegration(), FindIntegrationAnchors(), and IntegrateData() functions. PCA was performed on the top 3000 variable genes and the data was clustered at a low resolution (dims=1:30, resolution=0.2 for lung homogenate data and 0.1 for cultured MSCs) with the Louvain algorithm implemented in the FindClusters() function in Seurat. Cell populations were identified with a simple Wilcoxon rank sum test with the FindAllMarkers() function in Seurat.

In the case of stromal cells from lung homogenates, a previously published, publicly available scRNA-seq dataset from newborn mice was re-analyzed[23]. A novel Ly6a$^{+}$ L-MSC population was identified based on the expression of Ly6a. New cell type labels for stromal populations were then added to the Seurat object containing all data.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

231

232 **Differential expression analysis (DSA), gene set enrichment analysis (GSEA) and functional**
233 **enrichment analysis**

234 To identify differentially expressed genes in response to hyperoxia or as a result of mouse age,
235 we used the R package muscat (v1.4.0). Pseudobulk expression profiles were generated for each
236 sample in each cluster and differential expression was tested between groups associated with the
237 experimental conditions. Genes with an adjusted p-value $< 0.05$ and a detection rate $\geq 10\%$ in at
238 least one of the conditions tested were considered significant. To further identify gene sets
239 associated with differentially expressed genes, we used the R package fgsea (v1.16.0). List of
240 gene sets comprised all GO terms, KEGG pathways, Reactome pathways, and the MSigDB
241 Hallmark gene sets acquired from the Molecular Signatures Database (v7.2)[26]. Gene sets with
242 an adjusted p-value $< 0.05$ were considered significantly enriched. Normalized enrichment score
243 (NES) was used to assess whether gene sets were associated with upregulated or downregulated
244 genes. Functional enrichment analysis (FEA) for selected ligands produced by Ly6a+ L-MSCs
245 were performed using the online Metascape tool[27]. Summary pathways relevant to lung were
246 considered.

247

248 **Cell communication inference**

249 To explore cell communication networks behind the developmental age, or hyperoxia-specific
250 effects, we utilized the R package nichenetr (v1.0.0), which uses information about expression of
251 cognate ligands, receptors, signaling pathways, and genomic targets to infer cell communication
252 patterns[28]. Differential gene expression analysis for P3 vs. P14, or hyperoxia vs. normoxia
253 groups were used in the NicheNet analysis. To prioritize results, analysis was limited to signaling
254 contributing to the effects in receiver cell types with >200 differentially expressed genes at P14
255 or in response to hyperoxia, but included all cell types as potential ligand senders. Background
256 expression of genes was specified with default approach used in NicheNet's pipeline, using all
257 genes with >10% detection in a given cluster. While using cells from both experimental
258 conditions, developmental age, or hyperoxia-induced ligands from cell types that increase in
259 proportion with age or in hyperoxia samples were prioritized. For each "receiver" cell

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

population, top 10 ligands predicted to drive developmental age, or hyperoxia-induced responses were selected based on the Pearson correlation coefficient between the ligand-target regulatory potential score of each ligand and the target indicator vector. Further, we assessed whether the expression of ligands and receptors was upregulated, or whether the populations expressing the ligands increased in proportion in P14 or hyperoxia samples, respectively. Finally, potential target genes were inferred. Summaries of ligand-receptor interactions are represented in circos plots.

### 2.15. Statistical analysis

All statistical analyses were performed with GraphPad Prism 8.0. The presence of potential statistical outliers was determined by Grubbs' test. Data are presented as means ± SD. Differences in case of two-member groups were evaluated either by unpaired Student's *t*-test, or multiple unpaired Student's *t*-test with correction for multiple comparisons using the Holm-Šidák method. P values < 0.05 were considered as significant and depicted as following: P values < 0.05: *; P values < 0.01: **; P values < 0.001: ***; P values < 0.0001: ****.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## 3. RESULTS

### 3.1 The developing murine lung contains a population of L-MSC marked by the expression of *Ly6a*

In order to understand the expression patterns unique to LY6A$^+$ L-MSCs in the developing lung, we took advantage of a publicly available scRNA-seq dataset from newborn mice[23]. Within this dataset, we analyzed 7,994 stromal cells from normoxia or hyperoxia-exposed developing mouse pups on postnatal days (P)3, 7, and 14, clustered into 6 distinct populations (Fig. 1A). Based on the expression pattern of commonly used MSC markers (Supplementary fig. 2A) we selected *Ly6a* as most suitable marker to identify L-MSC in lung stroma. We then subsetted the *Ly6a*$^+$ cells, belonging almost exclusively to the *Col14a1*$^+$ fibroblasts, as a separate, seventh cluster (Fig. 1B). Differential gene expression analysis revealed that *Ly6a*$^+$ L-MSCs could be characterized by the expression of additional markers, including *Lum*, *Serpinf1*, or *Dcn*, with *Lum* being the single most unique identifier of the population (Fig. 1C, Supplementary table 2). It was previously shown to inhibit migration, invasion, and tube-formation in BM-MSCs[29], and was implicated in epithelial-mesenchymal transition and fibrocyte differentiation[30]. While *Ly6a*$^+$ L-MSCs expressed additional MSC markers *Mcam*, *Alcam* and *Eng*, their expression did not serve as a reliable indicator of *Ly6a*$^+$ L-MSCs (Fig. 1D).

### 3.2 The transcription profile and signaling activity of *Ly6a*$^+$ L-MSCs change significantly during postnatal lung development

We first aimed to understand how the L-MSC population changes in the postnatal developing lung. While the size of the population remained unchanged between P3 and P7, the second week of lung development in healthy mice was associated with an increase in the size of the *Ly6a*$^+$ stromal population (Figure 1E-F, Supplementary table 3). Similarly, differential state analysis (DSA) in normally developing lungs revealed that most changes in gene expression occurred in L-MSC between P7 and P14 (Fig. 1G, Supplementary table 4). Although the expression of genes such as *Apoe*, *Inmt*, *Klf9* and *Abca1* was drastically increased in L-MSCs, these genes were also considerably upregulated in *Ly6a*$^-$ stromal cells (Supplementary table 4). The largest L-MSC - specific expression changes were observed for *Mmp3*, *C1s1*, *Podn*, *Dlk1*, and *Agtr2* (Fig. 1H).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Gene set enrichment analysis (GSEA) identified extracellular matrix (ECM) formation, vascular development, and wound healing among the activated pathways (Fig. 1I, Supplementary table 5).

Next, to further understand how L-MSCs send and receive signals during postnatal development, we performed a cell communication analysis. We inferred developmental age-induced cellular communications between *Ly6a*$^{+}$ L-MSCs and other lung populations using the NicheNet tool [23,28] (Fig. 2A, Supplementary fig. 3-5, Supplementary table 6). During development L-MSCs received signals from several cell populations, including endothelial cells, interstitial macrophages (Int Mf), alveolar epithelial type 2 (AT2) cells, and stromal cells (Fig. 2A). *Col4a1*, *Fat1*, *Hmgb2*, *Vcam1* and *Hc* were identified as most potent ligands, targeting numerous downstream genes in the developing L-MSCs, including *Klf9*, *Top2a* and other strongly de-regulated genes (Fig. 2A-B, Figure 1G, Supplementary table 4). Furthermore, L-MSCs produced numerous ligands, targeting most lung cell populations, including itself (Fig. 2A), Among the most broadly acting ligands were *Agt*, *App*, and *Apoe* (Fig. 2C). Functional enrichment analysis (FEA) revealed, that the expression of the L-MSC-produced ligands was associated with pathways related to angiogenesis, cell migration, adhesion and chemotaxis, and ECM organization (Supplementary fig. 2B, Supplementary table 7).

### 3.3 The transcription profile and signaling activity of *Ly6a*$^{+}$ L-MSC change significantly during postnatal lung development in response to hyperoxia

Hyperoxia induced an increase in proportion of *Ly6a*$^{+}$ stromal cells as determined by scRNA-seq analysis at P14 (Fig. 1F, Supplementary table 3). This was consistent with increased proportion of LY6A$^{+}$ stromal cells in hyperoxia-exposed lungs at P7 as measured by flow cytometry (Supplementary fig. 2C-D). In order to identify hyperoxia-induced changes in gene expression specific to *Ly6a*$^{+}$ L-MSCs, we performed a DSA for both, *Ly6a*$^{+}$ L-MSC population and non-progenitor *Ly6a*$^{-}$ stromal cells (Supplementary table 8). Hyperoxia-induced expression changes most distinctive of *Ly6a*$^{+}$ L-MSCs are illustrated in Fig. 3A. Exposure to hyperoxia was associated with *Ly6a*$^{+}$ L-MSCs - specific increase in expression of multiple pro-inflammatory (*Cxcl1*, *Ccl2*), as well as pro-fibrotic and anti-angiogenic (*Timp1*, *Serpina3n*) genes (Fig. 3A, Supplementary table 8). GSEA of hyperoxia-induced changes in gene expression revealed an activation in inflammatory pathways, as well as decrease in pathways associated with arterial

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

development and morphogenesis (Fig. 3B, Supplementary table 9). When inspecting pathways altered by hyperoxia exclusively in *Ly6a*$^{+}$, but not *Ly6a*$^{-}$ stromal cells, activation of cytokine and chemokine signaling, cell cycle regulation, and senescence were most noticeable (Supplementary fig. 2E, Supplementary table 9).

To further understand the faith of *Ly6a*$^{+}$ L-MSCs in hyperoxia-induced injury, we performed a cell communication analysis using the NicheNet tool, inferring hyperoxia-induced cellular communications[23,28] (Fig. 3C, Supplementary fig. 6-7, Supplementary table 10). *Ly6a*$^{+}$ L-MSCs in hyperoxia-exposed lungs received signals from several cell populations, including immune cells, capillary and arterial endothelial cells, mesothelial cells and *Col13a1*$^{+}$ fibroblasts (Fig. 3C). Further, we inferred genes in *Ly6a*$^{+}$ L-MSCs most likely to be targeted by the received signals (Fig. 3D). Multiple ligands, such as *Apoe*, *Il1a*, *Ifng* and *Mmp9* were predicted to target the expression of pro-inflammatory, pro-fibrotic and anti-angiogenic genes discussed above, including *Timp1*, *Cxcl1* and *Icam1* (Fig. 3D). Expression of these target genes was elevated in *Ly6a*$^{+}$ L-MSCs by hyperoxia exposure (Fig. 3A). Finally, ligands produced by *Ly6a*$^{+}$ L-MSCs affected multiple cell populations, including alveolar macrophages, ciliated and AT2 cells, capillary and vein endothelium and other stromal populations. Among the most broadly acting ligands produced by *Ly6a*$^{+}$ L-MSCs were *Bmp4*, *Bmp5*, *Col4a1* and *Tnc* (Fig. 3A). Inferred target genes in receiving cells targeted by majority of these ligands included *Ccnd1*, *Cdkn1a*, *Icam1* and *Hmox1* (Fig. 3E). According to FEA, expression of the L-MSC-produced ligands were associated with pathways related to vessel morphogenesis, epithelial cell proliferation, cell chemotaxis, and immune homeostasis and response (Supplementary fig. 2F, Supplementary table 11).

### 3.4 Murine LY6A$^{+}$ L-MSCs localize to perivascular regions of the developing lung

Next, we aimed to localize the *Ly6a*$^{+}$ L-MSCs in the developing lung using FISH. L-MSCs were identified as *Ly6a*$^{+}$/*Col14a1*$^{+}$ cells. L-MSCs in both, normally and aberrantly-developing lungs localized to perivascular regions of large vessels with more double-positive cells observed in hyperoxia-exposed lungs (Fig. 4A).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Additionally, we aimed to validate some of the novel normoxic and hyperoxic L-MSC markers as suggested by scRNA-seq analysis (Fig. 1C, Fig. 3A). *Ly6a*$^{+}$ L-MSCs were co-stained for the hyperoxia-associated markers *Timp1* and *Serpina3n* (Fig. 4B and 4C, respectively). In both instances triple-positive cells were observed in the regions adjacent to large vessels (highlighted by white squares in low-magnification panels). These cells were not only more abundant in the lungs from BPD mice, but the expression levels of both, *Timp1* and *Seprina3n* were increased in the diseased lungs (see higher-magnification panels Fig. 4B-C).

### 3.5 Hyperoxia exposure does not impact clonal or differentiation potential of LY6A$^{+}$ L-MSCs

In order to verify their progenitor cell-like properties, we isolated and studied LY6A$^{+}$ L-MSCs from healthy and hyperoxia-exposed developing mouse pups. An arrest in lung development was induced by exposing newborn mouse pups to normobaric hyperoxia (85% $O_2$) (Fig. 5A). CD31$^{-}$/CD45$^{-}$/EpCAM$^{-}$/LY6A$^{+}$ L-MSCs were isolated from seven days-old healthy (21% $O_2$-exposed) or diseased (85% $O_2$-exposed) mouse pups (Fig. 5B) and examined for the hallmarks of the MSC phenotype *in vitro*. While lungs of hyperoxia-exposed pups consistently yielded higher numbers of LY6A$^{+}$ L-MSCs (Fig. 5B), no differences in the appearance (Fig. 5C), differentiation capacity (Fig. 5C), expression of surface markers (Fig. 5D), or clonal abilities (Fig. 5E) were observed between the cells isolated from healthy and diseased animals. LY6A$^{+}$ L-MSCs isolated from both healthy and hyperoxia-exposed mice had a fibroblast-like appearance and expressed classical markers of MSCs *in vitro* (Fig. 5C-D). In order to investigate their differentiation capacity, LY6A$^{+}$ L-MSCs were induced to differentiate along the osteogenic, chondrogenic, and adipogenic lineages. Both normoxia and hyperoxia-derived LY6A$^{+}$ L-MSCs produced osteogenic and chondrogenic matrix (Fig. 5C). However, only a single sample of normoxia-derived LY6A$^{+}$ L-MSCs produced a small number of adipocytes, and no lipogenic differentiation was observed in hyperoxia-derived LY6A$^{+}$ L-MSCs (data not shown). Postnatal hyperoxia exposure had no effect on colony-forming capacity of LY6A$^{+}$ L-MSCs as assessed by single-cell plating colony-forming assay. Both normoxia and hyperoxia-derived LY6A$^{+}$ L-MSCs produced colonies of various sizes. While larger colonies consisted of fibroblast-like spindle-shaped cells, smaller colonies were formed by cells with a large cytoplasm (Fig. 5E).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Inconsistent differentiation capacity and colony formation might suggest a heterogeneous nature of the $LY6A^+$ L-MSCs population.

### 3.6 Cell culture alters the gene expression profile of $LY6A^+$ L-MSCs

For therapeutic applications, MSCs are typically culture expanded, then frozen, over the short-, or long-term and thawed prior to administration. These various steps may alter the properties of the cell product. In order to understand changes in the L-MSCs expression profile induced by storage and culture, we performed a scRNA-seq analysis of cultured $LY6A^+$ and $LY6A^-$ lung stromal cells isolated from seven days-old healthy (21% $O_2$-exposed) or diseased (85% $O_2$-exposed) mouse pups (Fig. 6A, Supplementary fig. 2C-D). We sequenced over 15,000 cultured $CD31^-/CD45^-/EpCAM^-/LY6A^-$ and $CD31^-/CD45^-/EpCAM^-/LY6A^+$ cells and identified four distinct clusters (Fig. 6A-C, Supplementary tables 12-13). While normoxia and hyperoxia-derived $LY6A^-$ stromal cells contributed to all four clusters, very few $LY6A^+$ cells could be found in clusters 2 and 3 (Fig. 6B). The presence of distinct clusters within the L-MSCs population is consistent with the heterogeneous phenotype of cultured L-MSCs described above (Fig. 5E). In line with this finding, the highest levels of routine MSC markers, such as *Thy1*, *Eng*, *Alcam* or *Mcam*, were found in the largest cluster 0, while very little expression was seen in the two smallest clusters (Fig. 6D). While still expressing routine MSC markers to some level, cluster 1 was characterized by its distinct expression of *Cck*, previously found to attenuate *p53*-mediated apoptosis in lung cancer [31] (Figure 3A-C). Cluster 2 was distinguished by the expression of pro-adipogenic markers, such as *Igfbp2* and *Col4a1*, as well as markers of myofibroblasts (Des) and alveolar epithelium (*Krt8* and *Prnp2*) (Fig. 6C, Supplementary table 13). Cluster 3 was characterized by the expression of multiple osteogenic markers, including *Cryab*, *Postn* and *Ngfr*. Interestingly, the expression of both *Postn*, as well as another cluster 3 marker *Col18a1*, was previously reported in BPD patients and hyperoxia-exposed developing mice [32,33].

Next, we aimed to identify the best markers for cultured L-MSCs (Supplementary tables 14-18). We compared the gene expression profiles of $LY6A^+$ and $LY6A^-$ stromal cells (Supplementary tables 15-18) and identified differentially expressed genes between normoxia- and hyperoxia-derived subsets of these populations (Supplementary tables 17-18). In comparison

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

to LY6A$^{-}$ cells, LY6A$^{+}$ cells were characterized by high expression of *Actg2, Col1a2, Serpinf1, Prrx1* and *Lxn*, and by low expression of smooth muscle cell (SMC) marker *Tagln2*[34], alveolar progenitor marker *Tm4sf1*[35], and *Prdx6* (Fig. 6E, Supplementary tables 15-16). From these markers hyperoxia exposure further specifically increased the expression of *Actg2*, and decreased the expression of *Tagln2*, *Tm4sf1* and *Prdx6* in LY6A$^{+}$ cells. Hyperoxic LY6A$^{+}$ cells were additionally distinguished by expression of *Ptn, Adamts5, Rbp,* and *Col3a1* (Fig. 6E, Supplementary table 17). Expression of *Prrx1* and *Serpinf1* is known to favour an osteogenic phenotype, and *Serpinf1* is known to inhibit adipogenesis[36,37].

In order to identify L-MSCs expression patterns maintained after cell culture and storage, we next compared expression of the most promising markers of *Ly6a*$^{+}$ L-MSCs in both, *in situ* and *in vitro* datasets from cells isolated at P7 (Supplementary fig. 2G-H). This analysis revealed that a large portion of the expression profile characteristic for *Ly6a*$^{+}$ L-MSCs *in situ* (Supplementary fig. 2G-H) is lost when cells are frozen and cultured, including the expression of promising markers, such as *Lum*, *Ptn*, *Dcn*, or *Pi16* (Supplementary fig. 2H). Furthermore, while the expression pattern of some markers, such as *Serpina3n* or *C3* persisted in cultured cells, the portion of the cells expressing the gene was diminished (Fig. 7A-B). The most suitable *in situ* or *in vitro*-specific identifying markers of *Ly6a*$^{+}$ L-MSCs are depicted in Fig. 4A-B. Among the most stable markers of *Ly6a*$^{+}$ L-MSCs, resistant to changes induced by culture, were *Serpinf1* and *Postn* (Fig. 7A-B, Supplementary fig. 2G-H). In order to confirm the viability of *Serpinf1* as potential novel marker for L-MSCs we performed FISH in developing lungs at P14. Triple-positive cells could be found in lungs of both, normoxic and hyperoxic mice (Fig. 7C). No differences were apparent in *Serpinf1* expression intensity between the two groups.

Finally, new expression patterns arose particularly in hyperoxia-derived *Ly6a*$^{+}$ L-MSCs after cell culture. While a high *Ptn*, *Lum*, *Dcn, Col3a2* and *Col14a1* expression was initially characteristic of both, hyperoxia and normoxia-derived *Ly6a*$^{+}$ L-MSCs, in cultured L-MSCs this was true only for the hyperoxia-derived *Ly6a*$^{+}$ L-MSCs (Supplementary fig. 2G-H). This expression pattern denotes, that not only does the L-MSC transcriptome change in culture, but that the cells isolated from lungs of diseased mice tend to retain their expression profile and, potentially, progenitor-like nature longer.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## DISCUSSION

Our current knowledge regarding the identity and properties of tissue resident MSCs remains limited. Most studies analyse L-MSCs in culture after isolation with one, or several MSC markers. However, no explicit rules regarding which markers represent the L-MSC population the best exist to date. The progenitor-like characteristics of these cells have been established in culture [13,14], but it is not yet known why L-MSCs fail to prevent the lung injury or restore damage in the lung. While L-MSCs were previously found in bronchoalveolar lavage of BPD patients [50], it is not known whether this is due to increased apoptosis and subsequent shedding from the lung or is a sign of activation and proliferation of L-MSC and hence of increased numbers in BPD patients. Here, we provide an extensive scRNA-seq based analysis of L-MSCs in developing mouse lung, as well as in culture. We characterize the changes in trancriptomic profile induced in L-MSCs by developmental age, exposure to hyperoxia, and culture. Our study further provides an insight into communication between L-MSCs and other cell populations in the normally and abnormally developing lung. Finally, we propose novel markers for identification of L-MSCs in the developing lung.

The use of omics approaches to study tissue-specific MSCs *in vivo* has been previously proposed [38]. In the study presented here, we utilize scRNA-seq to study L-MSCs immediately after isolation (*in situ*) without confounding procedures, such as FACS, cell culture and storage, and hence preserve the in vivo activation status of the different lung populations as much as possible. We selected *Ly6a* to identify L-MSCs for 2 reasons: i) *Ly6a* is one of the most commonly used L-MSC markers and its expression has been shown in specific progenitor-like populations, ii) *Ly6a* was the only known MSC marker forming a visible subcluster within the lung mesenchyme of early postnatal mouse pups. We identified novel markers of L-MSCs, including *Lum*, *Serpinf1*, and *Dcn*. Next, we showed how the L-MSC's transcriptome changes during the course of normal lung development and in hyperoxia, and explored the communication between L-MSCs and other lung cell populations.

Hyperoxia-exposure, used as a model for BPD, was associated in *Ly6a*$^{+}$ L-MSCs with increased expression of multiple pro-inflammatory (*Cxcl1*, *Ccl2*), pro-fibrotic and anti-angiogenic (*Timp1*, *Serpina3n*) genes. Similarly, increased expression of both, *Timp1* and *Ccl2* was previously reported in hyperoxia-exposed rodents [39,40], and in plasma [41] or tracheal

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

aspirates (TA) [42] of BPD patients. *Timp1* expression was further increased in fibrotic foci in chronic BPD [43] and in the lungs of ventilated newborns [44]. GSEA further confirmed the activation of inflammatory and pro-fibrotic pathways, and a decrease in sprouting angiogenesis and vessel morphogenesis in the hyperoxia-exposed developing lungs. To further explore the role L-MSCs play in cell signaling during the development, we performed a cell communication inference analysis. L-MSCs in healthy developing lungs received ligands secreted mainly from endothelial, immune and other stromal cells. L-MSCs signalled back to the majority of lung cell populations with a selected set of ligands (Fig. 2). Upon hyperoxia exposure, L-MSCs received ligands primarily from immune and endothelial cells, including *Il1a*, *Mmp9*, *Ifng*, and *Fasl* (Fig. 3). Interestingly, multiple ligands received by L-MSCs were predicted to target the expression of pro-inflammatory, pro-fibrotic and anti-angiogenic genes increased in hyperoxia-exposed L-MSCs, such as *Timp1*, *Cxcl1* and *Icam1*. IFNγ and MMP9, which target the expression of both *Timp1* and *Cxcl1*, were previously implicated in development of alveolar hypoplasia [45] and an increased expression of IFNγ was reported in TA of BPD patients [27,46]. Development of BPD was also associated with increased TA and plasma protein levels of ICAM1 [47,48]. IL1A was also shown to induce an inflammatory phenotype in lung fibroblasts [49]. Additionally, *Fasl*$^{+}$ immune cells were shown to induce fibroblast cell death [50,51], and its overexpression was associated with alveolar apoptosis and disturbed alveolar and vascular development [52].

Next, we investigated how the L-MSCs' transcriptome changed due to culture and storage, both necessary steps for the preparation of a cell therapeutic product. ScRNA-seq analysis revealed, that following isolation, storage and culture, most L-MSCs retain the expression of MSC markers, including *Ly6a*$^{+}$. Cultured L-MSCs showed moderate ability to differentiate into chondrocytes and osteoblasts. However, we observed only one instance of successful differentiation along the adipogenic lineage, consistent with previous studies of L-MSCs in developing rats [13]. Inconsistent differentiation capacity could be attributed to heterogeneity within the L-MSC population as indicated by the variable size and morphology of L-MSC-derived colonies (Fig. 5). Importantly, such heterogeneity could indicate the existence of L-MSCs with varying progenitor-like capabilities, most likely impacting their therapeutic efficacy. Further, more detailed characterization of different L-MSCs subpopulations might be necessary in order to prepare a superior therapeutic product. ScRNA-seq revealed considerable changes in the transcriptome of L-MSCs in culture, implying that the cells studied and

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

maintained *in vitro* for the purposes of therapeutic interventions are appreciably altered compared to L-MSCs *in situ* (Fig. 6). Interestingly, we observed that the culture-induced transcription changes are less pronounced in L-MSCs derived from hyperoxia-exposed animals. This might suggest that hyperoxia primes L-MSCs to maintain certain characteristics, potentially in an attempt to trigger a repair mechanism. While the organism's own resident L-MSCs fail to prevent the hyperoxia-induced lung damage, a therapeutic use of injury-primed L-MSCs might be more beneficial than L-MSCs from healthy individuals. Interestingly, tissue origin and microenvironment were shown to significantly impact the behaviour and therapeutic efficacy of MSCs [53,54]. Moreover, conditioned media from BM-MSCs exposed *ex vivo* to hyperoxia exhibited superior therapeutic effects in the hyperoxia-induced rat BPD model when compared to media from BM-MSCs which were not pre-conditioned [55].

The localization of L-MSCs in the developing lungs has not yet been described. Here, we localized the L-MSC cells in the perivascular regions of both, heathy and diseased developing lungs by FISH. The *Ly6a*$^{+}$ L-MSCs in the hyperoxia-exposed lungs co-expressed *Timp1* and *Serpina3e*, confirming the results of scRNA-seq analysis. Finally, as *Ly6a* is not expressed in human tissues, we aimed to identify additional markers to label L-MSCs, both *in situ* and *in vitro*. *Lum*, identified as marker of L-MSCs *in situ*, is known to be produced by MSCs. Within the lung, it's expression was localized to peripheral lung and vessel walls [56]. While scRNA-seq revealed *Lum* as a promising L-MSCs marker *in situ*, it's expression in culture was preserved only in a small fraction of L-MSCs isolated from hyperoxia-exposed animals (Fig. 7A). In comparison, the expression of *Serpinf1* was well preserved *in vitro*, with the expression slightly increased in hyperoxic cells. Interestingly, *Serpinf1* expression was previously reported to be increased in hyperoxia-exposed newborn mice and *Serpinf1*$^{-/-}$ animals were protected from hyperoxia-induced lung injury [57]. *Serpinf1* is also known as an anti-angiogenic and anti-migratory marker associated with aging MSCs [57,58]. *In situ*, *Serpinf1* colocalized well with *Ly6a*$^{+}$/*Col14a1*$^{+}$ cells in both healthy and diseased lungs, suggesting *Serpinf1* as promising new marker for L-MSCs (Fig. 7).

To our knowledge, this is the first detailed report studying the characteristics and behaviour of L-MSC *in situ* and *in vitro*, during both health and disease. We unravelled the transcriptome and cellular communication of this lung resident cell population by scRNA-seq in

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

543 order to mechanistically understand its endogenous repair capabilities, as well as its potential use
544 as an exogenous cell therapeutic product. In addition, we have established several markers that
545 can be used to identify L-MSC *in vitro* and *in vivo*, both in healthy and diseased lungs.
546 Additional studies will be needed to further unravel the heterogeneity of this population, as well
547 as their therapeutic capabilities.

548

549 **ACKNOWLEDGMENTS**

550 The authors acknowledge Prof. Zev Gartner (University of California, San Francisco) for kindly
551 providing barcodes for multiplex labelling and Prof. Dr. Gloria Pryhuber from LungMAP
552 Human Tissue Core, Biorepository for Investigation of Neonatal Diseases of Lung-Normal
553 (BRINDL-NL) for kindly providing the human tissue material. This study was supported by the
554 Canadian Institutes of Health Research (CIHR), the German Research Foundation (Deutsche
555 Forschungsgemeinschaft), the Frederick Banting and Charles Best Doctoral Scholarship, the
556 Finnish Foundation for Pediatric Research, the Finnish Sigrid Juselius Foundation, the Canadian
557 Lung Association - Breathing as One, and the Molly Towel Perinatal Research Foundation.

558

559 **DICLOSURE OF POTENTIAL CONFICT OF INTERESTS**

560 The authors declare no conflicts of interest, financially or otherwise.

561

562

563

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# REFERENCES


1 Warburton D. Overview of Lung Development in the Newborn Human. Neonatology 2017;111:398–401.

2 Thébaud B, Goss KN, Laughon M, et al. Bronchopulmonary dysplasia. Nat Rev Dis Primers 2019;5:78.

3 Wong PM, Lees AN, Louw J, et al. Emphysema in young adult survivors of moderate-to-severe bronchopulmonary dysplasia. European Respiratory Journal 2008;32:321–328.

4 Fawke J, Lum S, Kirkby J, et al. Lung Function and Respiratory Symptoms at 11 Years in Children Born Extremely Preterm: The EPICure Study. Am J Respir Crit Care Med 2010;182:237–245.

5 O'Reilly M, Möbius MA, Vadivel A, et al. Late Rescue Therapy with Cord-Derived Mesenchymal Stromal Cells for Established Lung Injury in Experimental Bronchopulmonary Dysplasia. Stem Cells and Development 2020;29:364–371.

6 van Haaften T, Byrne R, Bonnet S, et al. Airway Delivery of Mesenchymal Stem Cells Prevents Arrested Alveolar Growth in Neonatal Lung Injury in Rats. Am J Respir Crit Care Med 2009;180:1131–1142.

7 Sutsko RP, Young KC, Ribeiro A, et al. Long-term reparative effects of mesenchymal stem cell therapy following neonatal hyperoxia-induced lung injury. Pediatr Res 2013;73:46–53.

8 Pierro M, Ionescu L, Montemurro T, et al. Short-term, long-term and paracrine effect of human umbilical cord-derived stem cells in lung injury prevention and repair in experimental bronchopulmonary dysplasia. Thorax 2013;68:475–484.

9 Augustine S, Cheng W, Avey MT, et al. Are all stem cells equal? Systematic review, evidence map, and meta-analyses of preclinical stem cell-based therapies for bronchopulmonary dysplasia. STEM CELLS Translational Medicine 2020;9:158–168.

10 Aslam M, Baveja R, Liang OD, et al. Bone Marrow Stromal Cells Attenuate Lung Injury in a Murine Model of Neonatal Chronic Lung Disease. Am J Respir Crit Care Med 2009;180:1122–1130.

11 Collins JJP, Thébaud B. Progenitor cells of the distal lung and their potential role in neonatal lung disease: Progenitor Cells of The Distal Lung. Birth Defects Research Part A: Clinical and Molecular Teratology 2014;100:217–226.

12 Ushakumary MG, Riccetti M, Perl AT. Resident interstitial lung fibroblasts and their role in alveolar stem cell niche development, homeostasis, injury, and regeneration. STEM CELLS Transl Med 2021:sctm.20-0526.

13 Collins JJP, Lithopoulos MA, dos Santos CC, et al. Impaired Angiogenic Supportive Capacity and Altered Gene Expression Profile of Resident CD146 $^{+}$ Mesenchymal Stromal Cells Isolated from Hyperoxia-Injured Neonatal Rat Lungs. Stem Cells and Development 2018;27:1109–1124.

14 Lama VN, Smith L, Badri L, et al. Evidence for tissue-resident mesenchymal stem cells in human adult lung from studies of transplanted allografts. J Clin Invest 2007;117:989–996.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

15 Rolandsson Enes S, Andersson Sjöland A, Skog I, et al. MSC from fetal and adult lungs possess lung-specific properties compared to bone marrow-derived MSC. Sci Rep 2016;6:29160.

16 Deng M, Li J, Gan Y, et al. Changes in the number of CD31$^{\−}$CD45$^{\−}$Sca-1$^{+}$ cells and Shh signaling pathway involvement in the lungs of mice with emphysema and relevant effects of acute adenovirus infection. COPD 2017;Volume 12:861–872.

17 Gong X, Sun Z, Cui D, et al. Isolation and characterization of lung resident mesenchymal stem cells capable of differentiating into alveolar epithelial type II cells: Isolation and characterization of LR-MSCs. Cell Biol Int 2014;38:405–411.

18 Shi C, Cao X, Chen X, et al. Intracellular surface-enhanced Raman scattering probes based on TAT peptide-conjugated Au nanostars for distinguishing the differentiation of lung resident mesenchymal stem cells. Biomaterials 2015;58:10–25.

19 McQualter JL, Brouard N, Williams B, et al. Endogenous Fibroblastic Progenitor Cells in the Adult Mouse Lung Are Highly Enriched in the Sca-1 Positive Cell Fraction. Stem Cells 2009;27:623–633.

20 McQualter JL, Yuen K, Williams B, et al. Evidence of an epithelial stem/progenitor cell hierarchy in the adult mouse lung. Proceedings of the National Academy of Sciences 2010;107:1414–1419.

21 Salaets T, Gie A, Jimenez J, et al. Local pulmonary drug delivery in the preterm rabbit: feasibility and efficacy of daily intratracheal injections. American Journal of Physiology-Lung Cellular and Molecular Physiology 2019;316:L589–L597.

22 McGinnis CS, Patterson DM, Winkler J, et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat Methods 2019;16:619–626.

23 Hurskainen M, Mižíková I, Cook DP, et al. Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage. Nat Commun 2021;12:1565.

24 Butler A, Hoffman P, Smibert P, et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 2018;36:411–420.

25 Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 2019;20:296.

26 Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems 2015;1:417–425.

27 Aghai ZH, Saslow JG, Mody K, et al. IFN-γ and IP-10 in tracheal aspirates from premature infants: Relationship with bronchopulmonary dysplasia. Pediatr Pulmonol 2013;48:8–13.

28 Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. Nat Methods 2020;17:159–162.

29 Malinowski M, Pietraszek K, Perreau C, et al. Effect of Lumican on the Migration of Human Mesenchymal Stem Cells and Endothelial Progenitor Cells: Involvement of Matrix Metalloproteinase-14. PLoS ONE 2012;7:e50709.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

635 30 Li L-F, Chu P-H, Hung C-Y, et al. Lumican Regulates Ventilation-Induced Epithelial-Mesenchymal
636 Transition Through Extracelluar Signal-Regulated Kinase Pathway. Chest 2013;143:1252–1260.

637 31 Han Y, Su C, Yu D, et al. Cholecystokinin attenuates radiation-induced lung cancer cell apoptosis
638 by modulating p53 gene transcription. Am J Transl Res 2017;9:638–646.

639 32 Bozyk PD, Bentley JK, Popova AP, et al. Neonatal Periostin Knockout Mice Are Protected from
640 Hyperoxia-Induced Alveolar Simplication. PLoS ONE 2012;7:e31336.

641 33 Mohamed WAW, Niyazy WH, Mahfouz AA. Angiopoietin-1 and Endostatin Levels in Cord Plasma
642 Predict the Development of Bronchopulmonary Dysplasia in Preterm Infants. Journal of Tropical
643 Pediatrics 2011;57:385–388.

644 34 Shudo Y, Cohen JE, Goldstone AB, et al. Isolation and trans-differentiation of mesenchymal
645 stromal cells into smooth muscle cells: Utility and applicability for cell-sheet engineering. Cytotherapy
646 2016;18:510–517.

647 35 Zacharias WJ, Frank DB, Zepp JA, et al. Regeneration of the lung alveolus by an evolutionarily
648 conserved epithelial progenitor. Nature 2018;555:251–255.

649 36 Yu S, Li P, Li B, et al. RelA promotes proliferation but inhibits osteogenic and chondrogenic
650 differentiation of mesenchymal stem cells. FEBS Lett 2020;594:1368–1378.

651 37 Gattu AK, Swenson ES, Iwakiri Y, et al. Determination of mesenchymal stem cell fate by pigment
652 epithelium-derived factor (PEDF) results in increased adiposity and reduced bone mineral content.
653 FASEB j 2013;27:4384–4394.

654 38 Sipp D, Robey PG, Turner L. Clear up this stem-cell mess. Nature 2018;561:455–457.

655 39 Hosford GE, Fang X, Olson DM. Hyperoxia Decreases Matrix Metalloproteinase-9 and Increases
656 Tissue Inhibitor of Matrix Metalloproteinase-1 Protein in the Newborn Rat Lung: Association with
657 Arrested Alveolarization. Pediatr Res 2004;56:26–34.

658 40 Kumar VHS, Lakshminrusimha S, Kishkurno S, et al. Neonatal hyperoxia increases airway
659 reactivity and inflammation in adult mice: Neonatal Hyperoxia and AHR in Mice. Pediatr Pulmonol
660 2016;51:1131–1141.

661 41 Schulz CG, Sawicki G, Lemke RP, et al. MMP-2 and MMP-9 and Their Tissue Inhibitors in the
662 Plasma of Preterm and Term Neonates. Pediatr Res 2004;55:794–801.

663 42 Fulton CT, Cui TX, Goldsmith AM, et al. Gene Expression Signatures Point to a Male Sex-Specific
664 Lung Mesenchymal Cell PDGF Receptor Signaling Defect in Infants Developing Bronchopulmonary
665 Dysplasia. Sci Rep 2018;8:17070.

666 43 Dik WA, de Krijger RR, Bonekamp L, et al. Localization and Potential Role of Matrix
667 Metalloproteinase-1 and Tissue Inhibitors of Metalloproteinase-1 and -2 in Different Phases of
668 Bronchopulmonary Dysplasia. Pediatr Res 2001;50:761–766.

669 44 De Paepe ME, Greco D, Mao Q. Angiogenesis-related gene expression profiling in ventilated
670 preterm human lungs. Experimental Lung Research 2010;36:399–410.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

671 45 Harijith A, Choo-Wing R, Cataltepe S, et al. A Role for Matrix Metalloproteinase 9 in IFNγ-
672 Mediated Injury in Developing Lungs: Relevance to Bronchopulmonary Dysplasia. Am J Respir Cell Mol
673 Biol 2011;44:621–630.

674 46 Hikino S, Ohga S, Kinjo T, et al. Tracheal aspirate gene expression in preterm newborns and
675 development of bronchopulmonary dysplasia: TAF cell gene expression predicting BPD. Pediatrics
676 International 2012;54:208–214.

677 47 Ballabh P, Kumari J, Krauss AN, et al. Soluble E-Selectin, Soluble L-Selectin and Soluble ICAM-1 in
678 Bronchopulmonary Dysplasia, and Changes With Dexamethasone. PEDIATRICS 2003;111:461–468.

679 48 Kojima T, Sasai M, Kobayashi Y. Increased soluble ICAM-1 in tracheal aspirates of infants with
680 bronchopulmonary dysplasia. The Lancet 1993;342:1023–1024.

681 49 Suwara MI, Green NJ, Borthwick LA, et al. IL-1α released from damaged epithelial cells is
682 sufficient and essential to trigger inflammatory responses in human lung fibroblasts. Mucosal Immunol
683 2014;7:684–693.

684 50 Nareznoi D, Konikov-Rozenman J, Petukhov D, et al. Matrix Metalloproteinases Retain Soluble
685 FasL-mediated Resistance to Cell Death in Fibrotic-Lung Myofibroblasts. Cells 2020;9:411.

686 51 Predescu SA, Zhang J, Bardita C, et al. Mouse Lung Fibroblast Resistance to Fas-Mediated
687 Apoptosis Is Dependent on the Baculoviral Inhibitor of Apoptosis Protein 4 and the Cellular FLICE-
688 Inhibitory Protein. Front Physiol 2017;8.

689 52 De Paepe ME, Gundavarapu S, Tantravahi U, et al. Fas-Ligand-Induced Apoptosis of Respiratory
690 Epithelial Cells Causes Disruption of Postcanalicular Alveolar Development. The American Journal of
691 Pathology 2008;173:42–56.

692 53 Abreu SC, Rolandsson Enes S, Dearborn J, et al. Lung inflammatory environments differentially
693 alter mesenchymal stromal cell behavior. American Journal of Physiology-Lung Cellular and Molecular
694 Physiology 2019;317:L823–L831.

695 54 Islam D, Huang Y, Fanelli V, et al. Identification and Modulation of Microenvironment Is Crucial
696 for Effective Mesenchymal Stromal Cell Therapy in Acute Lung Injury. Am J Respir Crit Care Med
697 2019;199:1214–1224.

698 55 Waszak P, Alphonse R, Vadivel A, et al. Preconditioning Enhances the Paracrine Effect of
699 Mesenchymal Stem Cells in Preventing Oxygen-Induced Neonatal Lung Injury in Rats. Stem Cells and
700 Development 2012;21:2789–2797.

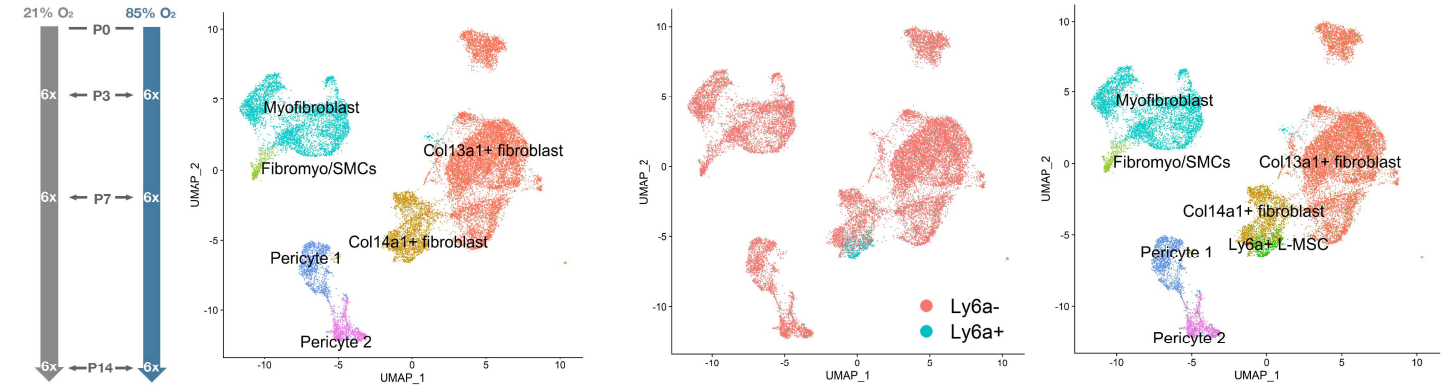
701 56 Dolhnikoff M, Morin J, Roughley PJ, et al. Expression of Lumican in Human Lungs. Am J Respir
702 Cell Mol Biol 1998;19:582–587.

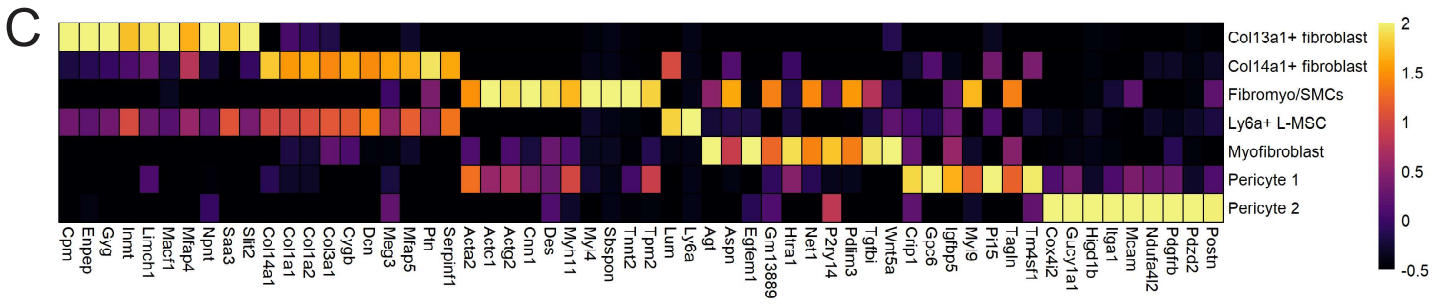
703 57 Chetty A, Bennett M, Dang L, et al. Pigment Epithelium–Derived Factor Mediates Impaired Lung
704 Vascular Development in Neonatal Hyperoxia. Am J Respir Cell Mol Biol 2015;52:295–303.

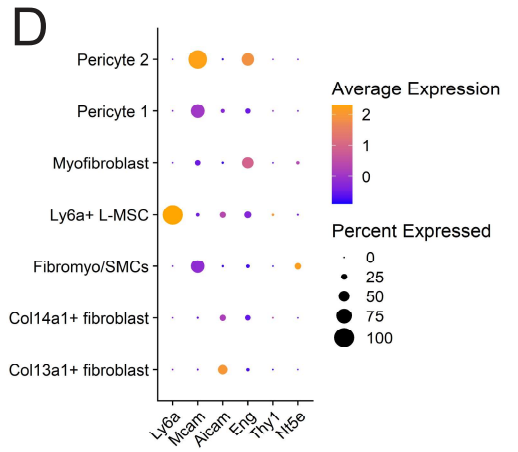
705 58 Liang H, Hou H, Yi W, et al. Increased expression of pigment epithelium-derived factor in aged
706 mesenchymal stem cells impairs their therapeutic efficacy for attenuating myocardial infarction injury‡.
707 European Heart Journal 2013;34:1681–1690.

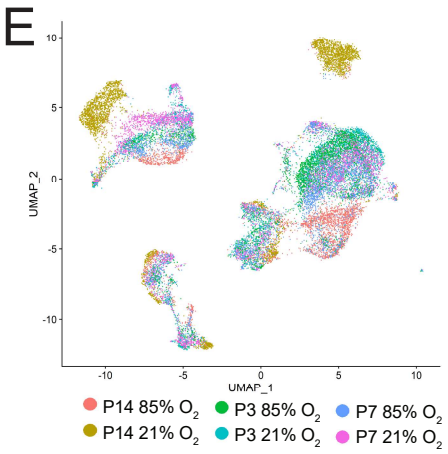
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

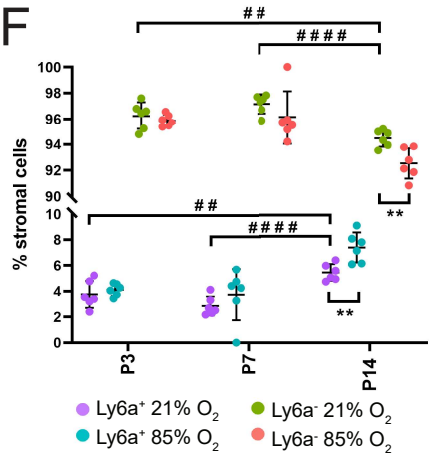
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

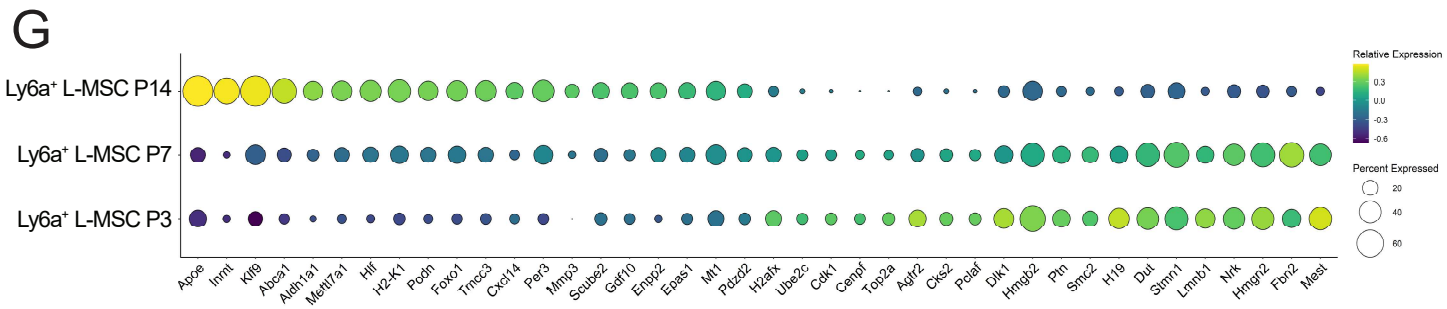


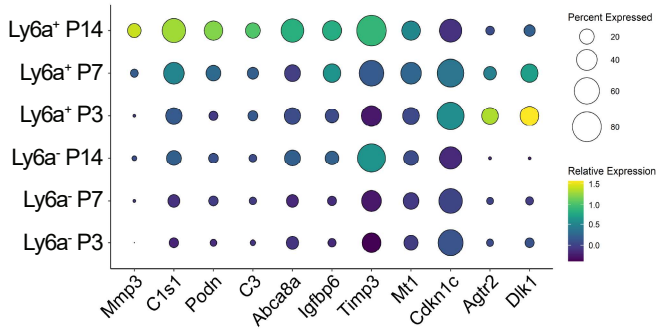



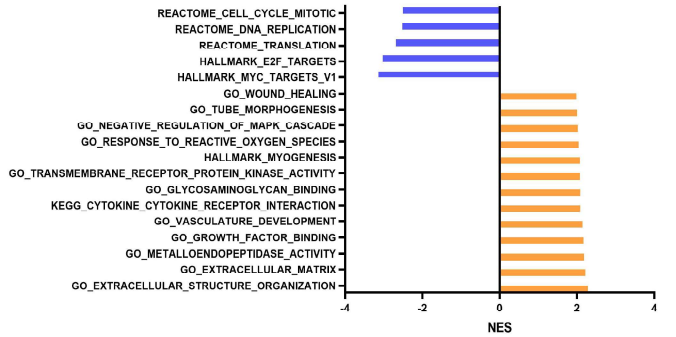
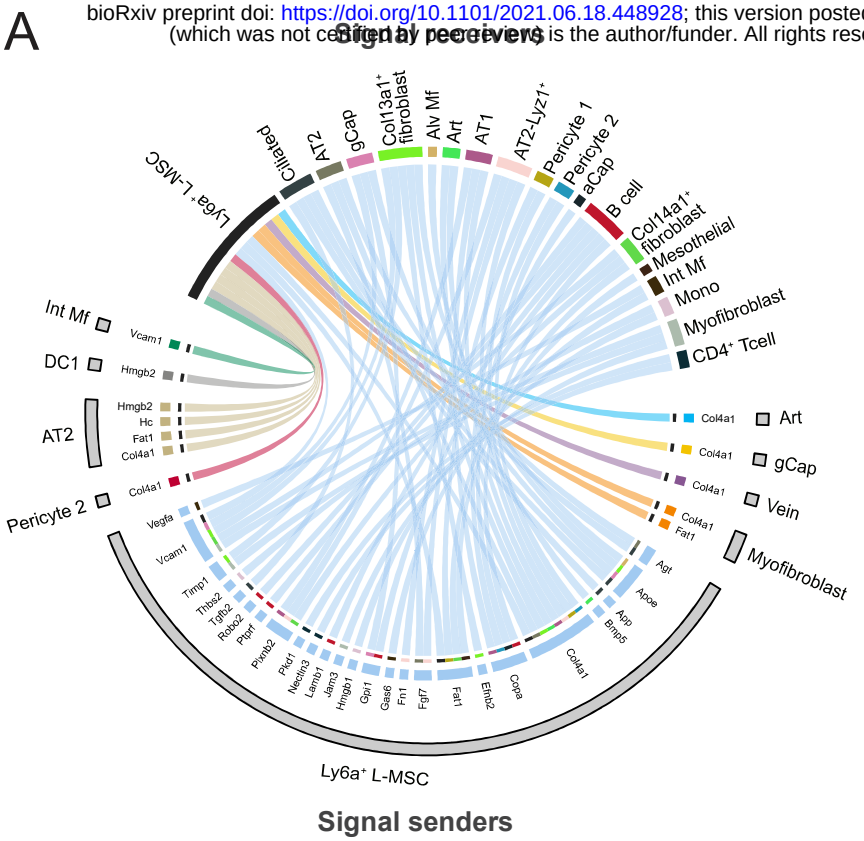
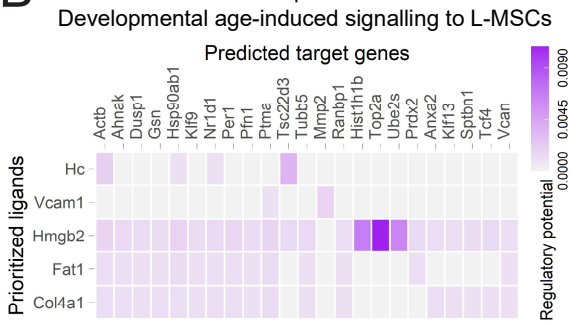




Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



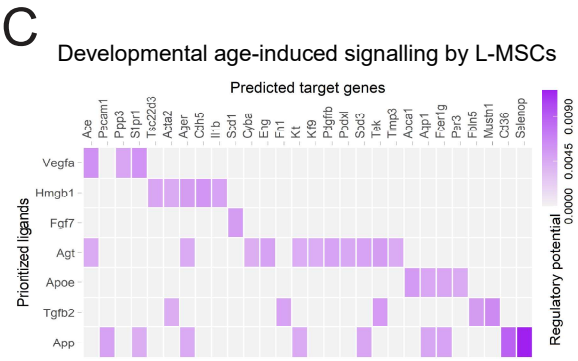
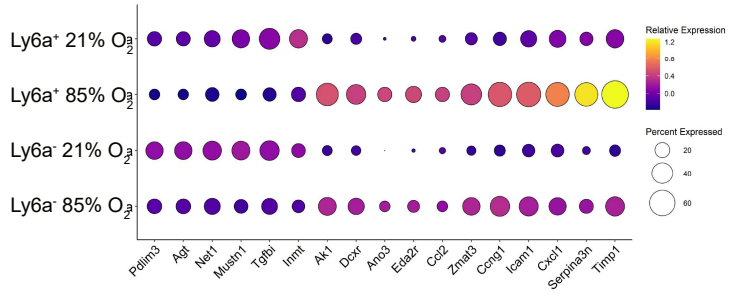
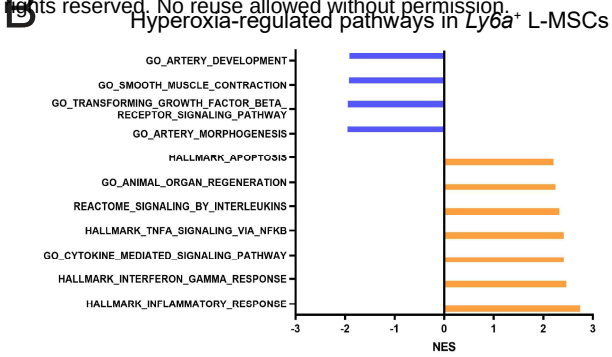
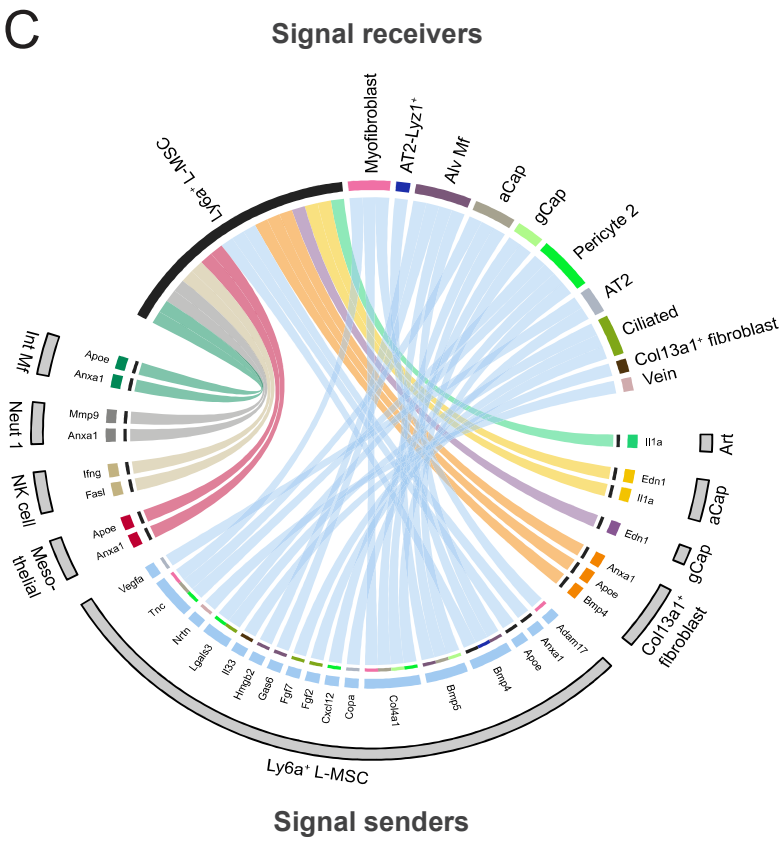



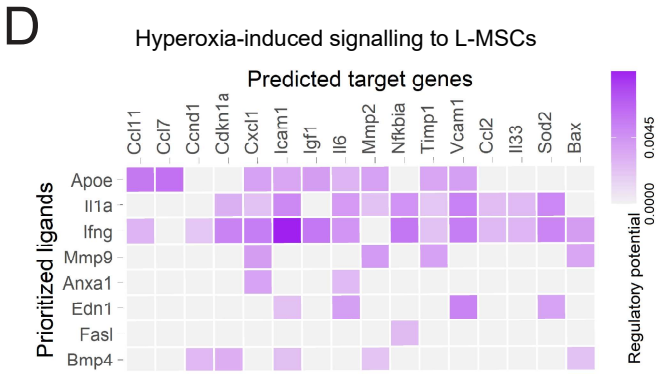

Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



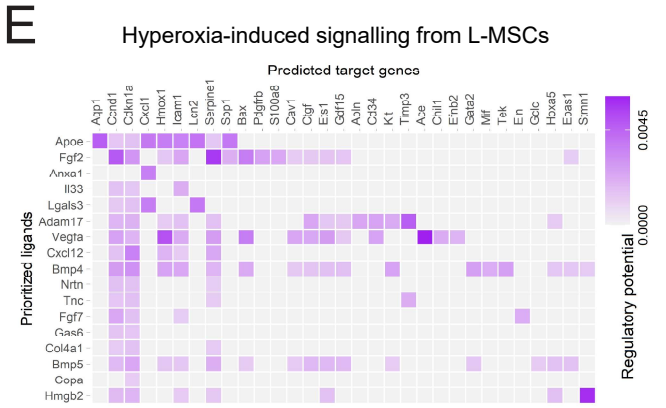




Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

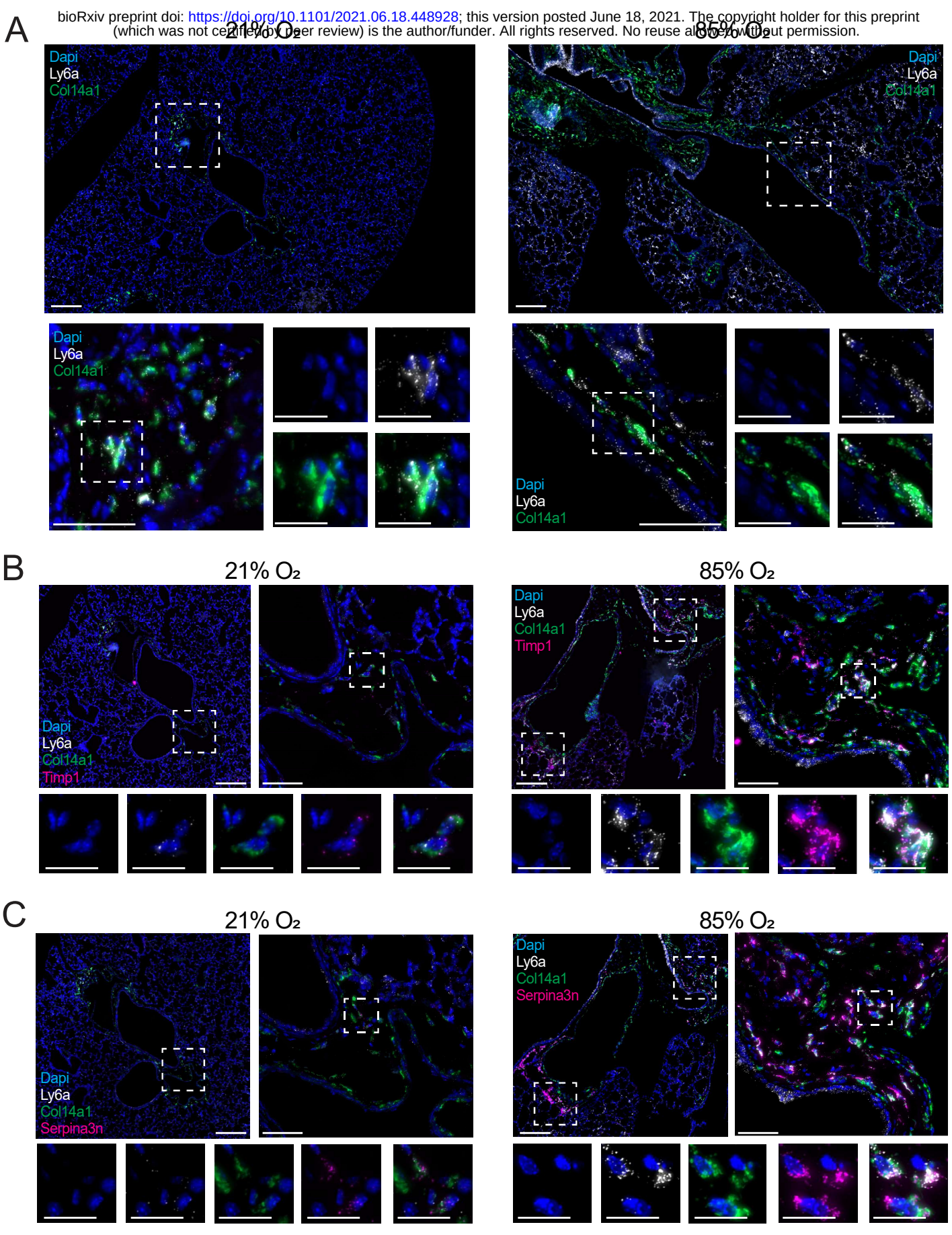
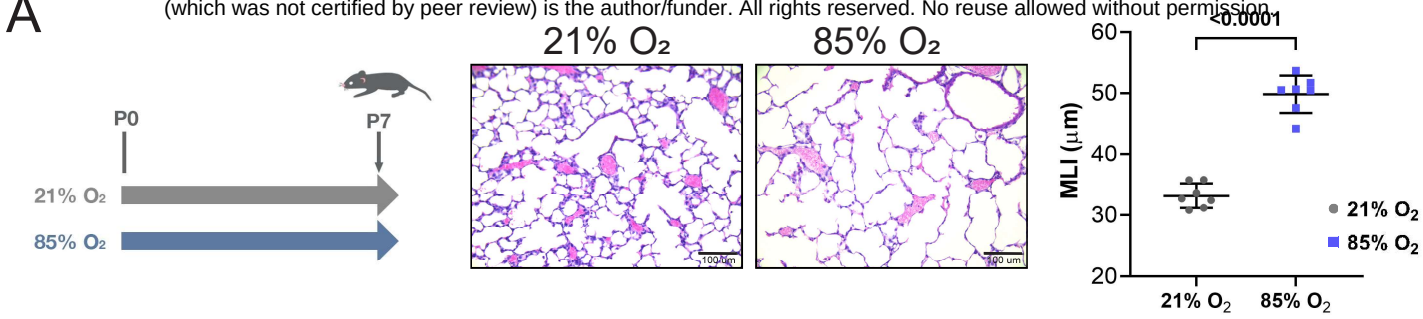
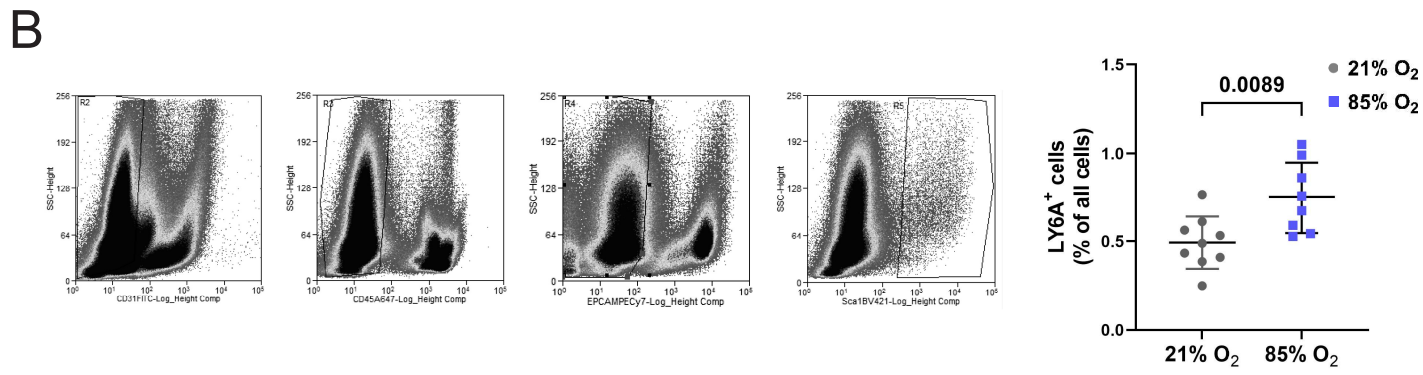
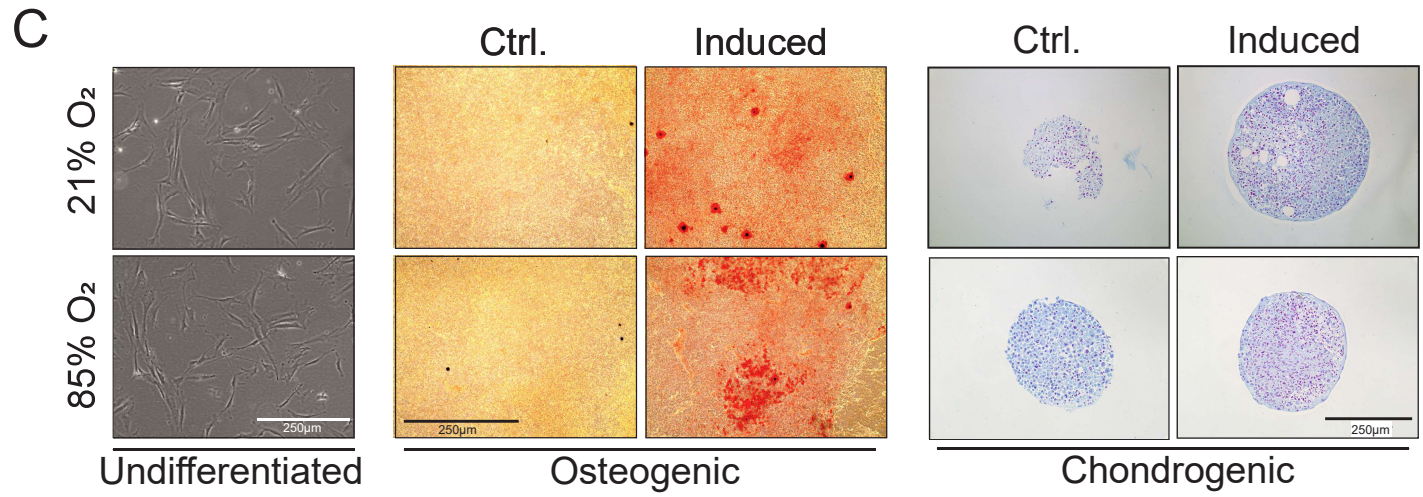



Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



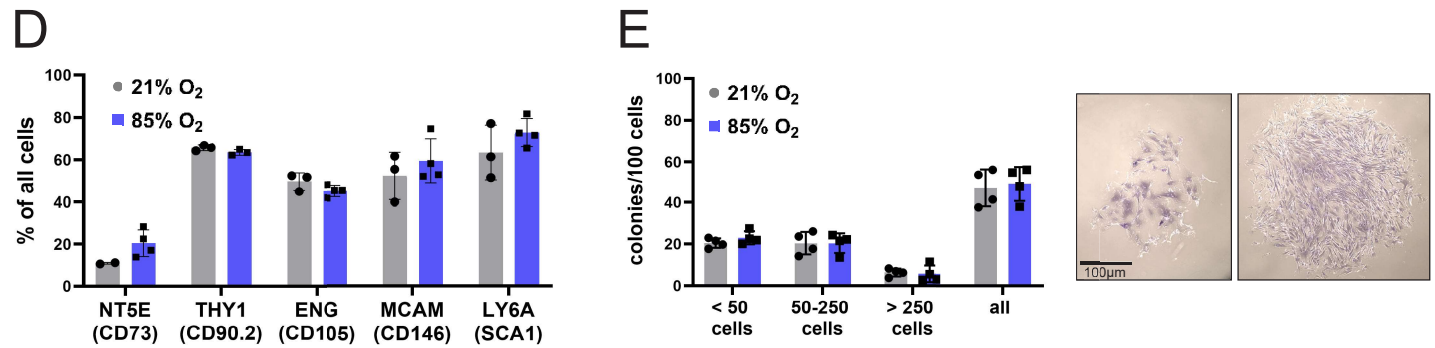
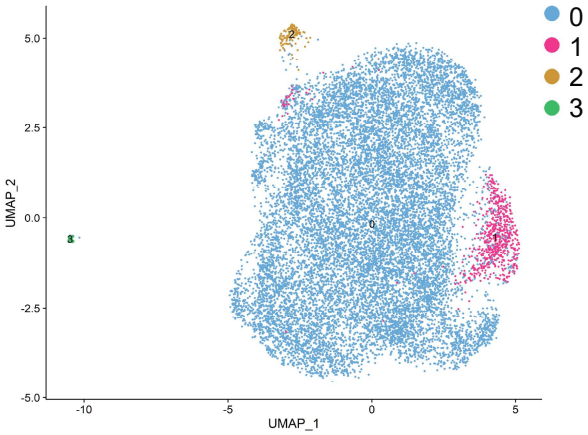




Figure 5

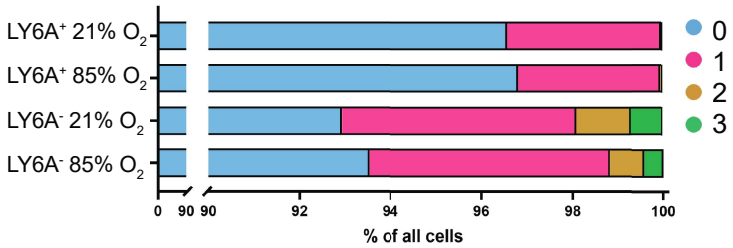
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A

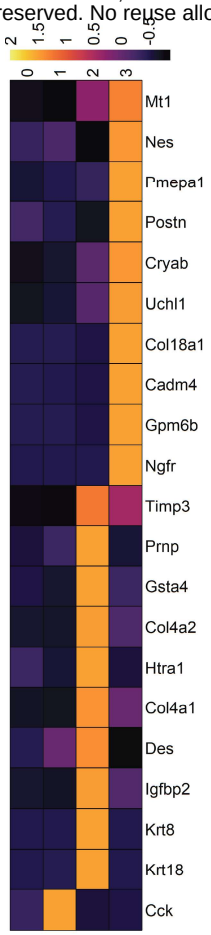
| Cell population | 21% $O_2$ | 85% $O_2$ | |
|---|---|---|---|
| LR-MSC ($LY6A^+$) | 2579 | 2565 | 5144 |
| Lung stroma ($LY6A^-$) | 6084 | 3793 | 9877 |



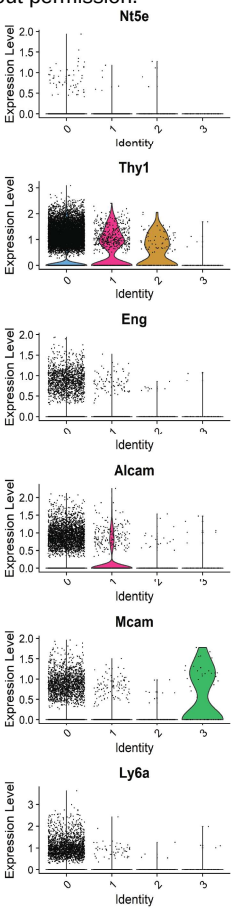

B




C




D




E

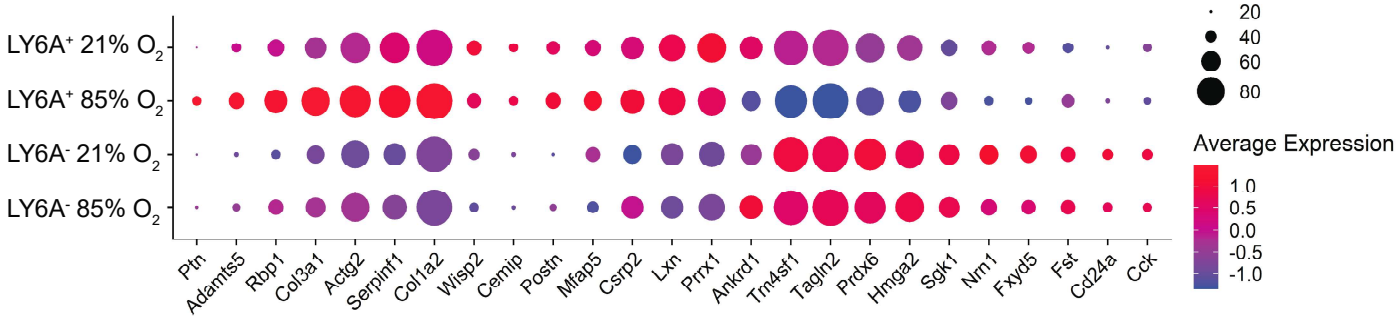
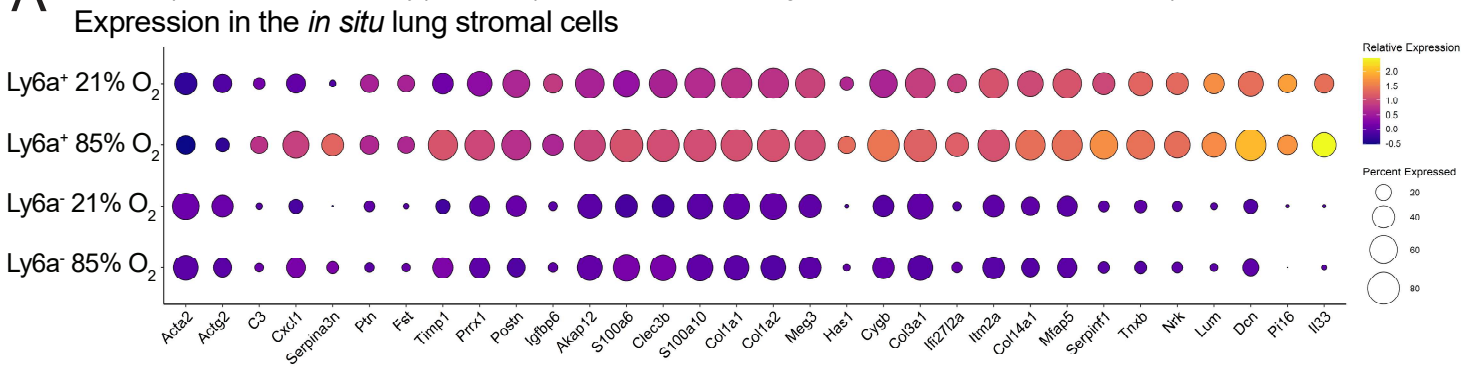


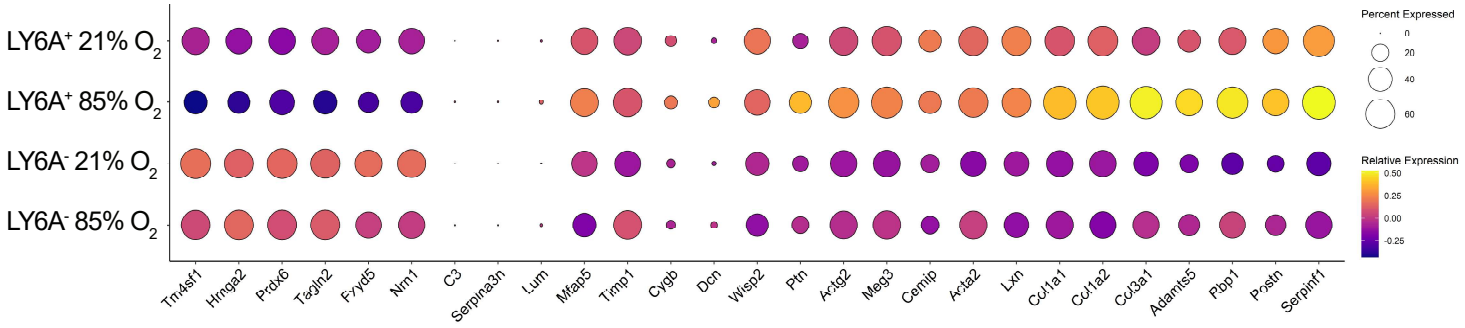

Figure 6

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A




B




C

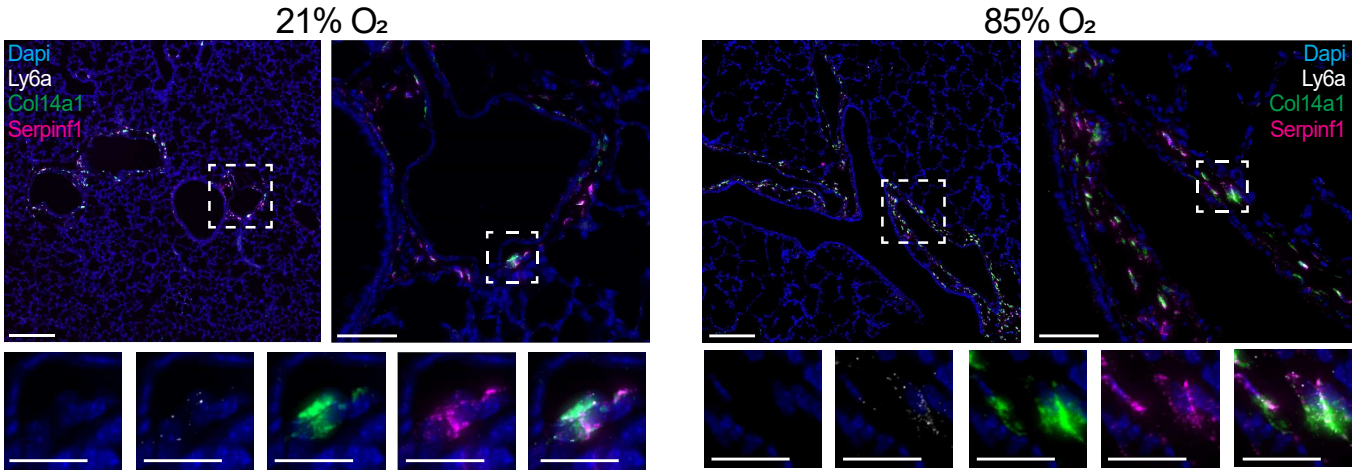
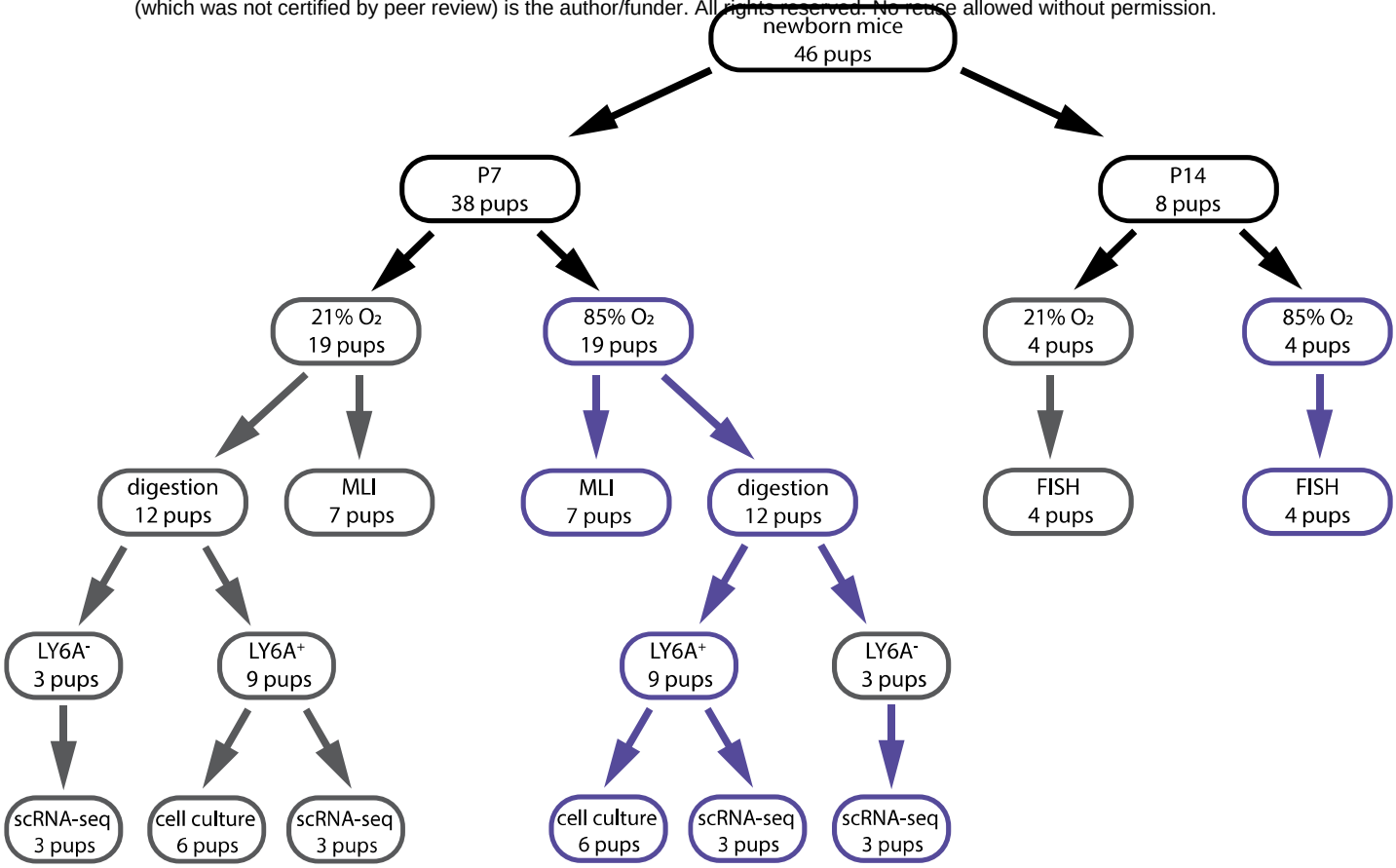


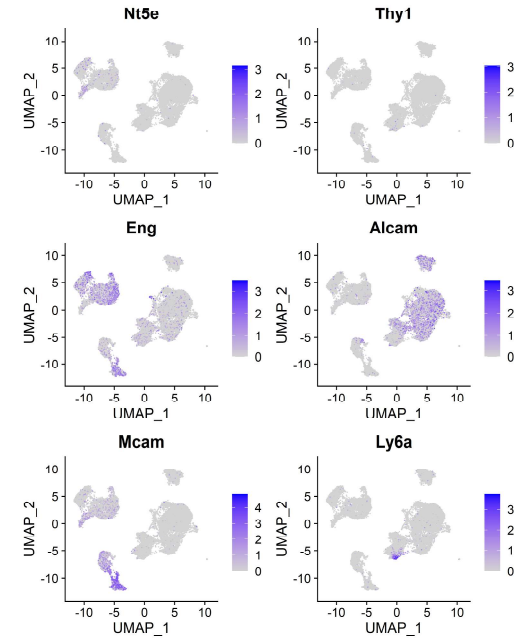

Figure 7

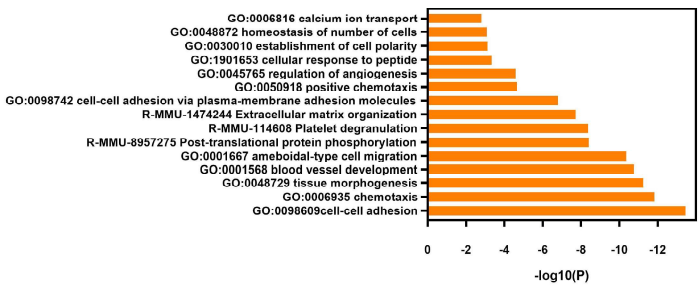
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

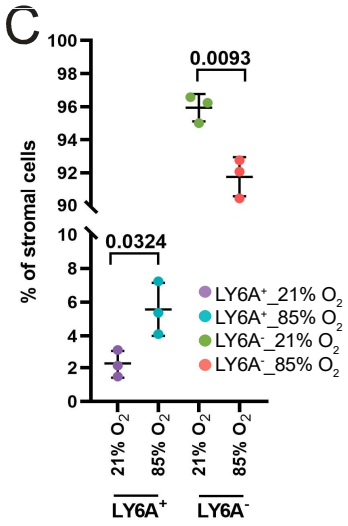


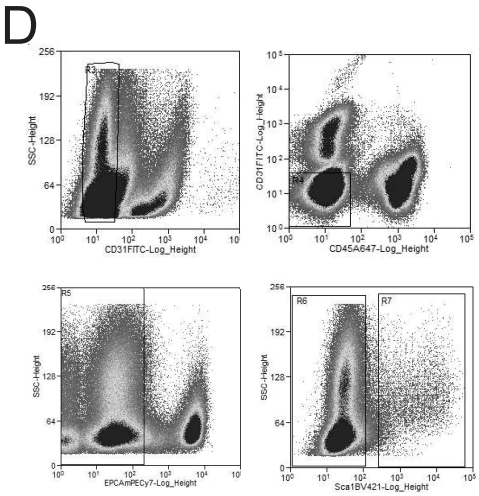

Supplementary Figure 1

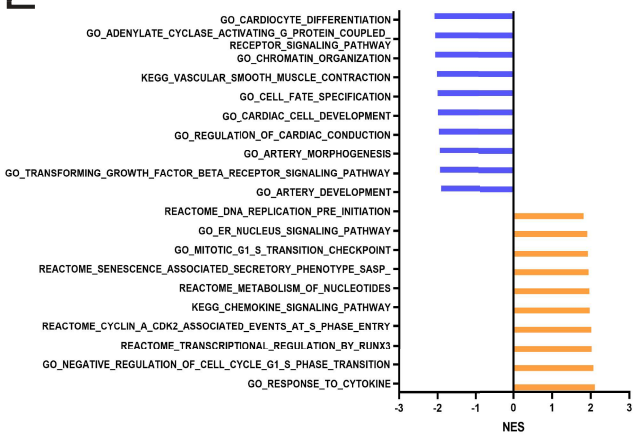
bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

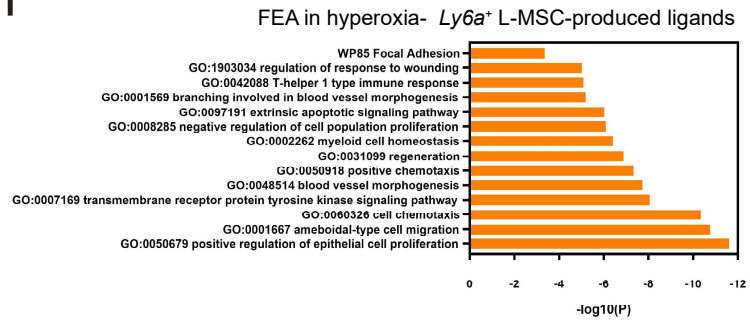


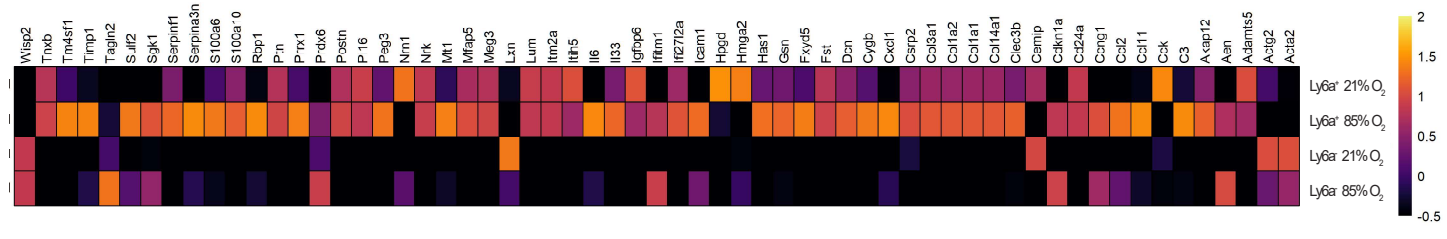



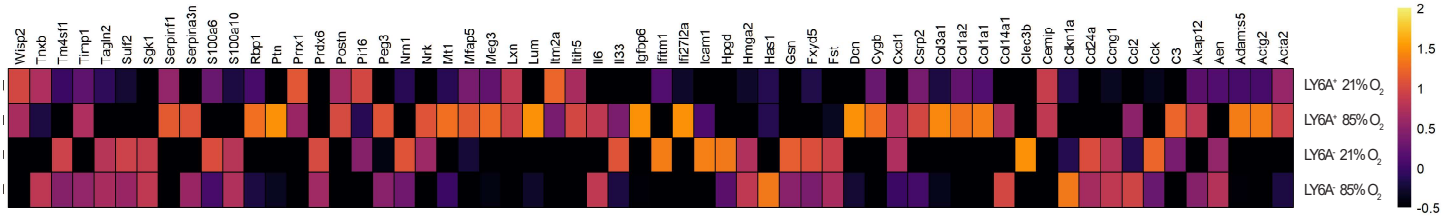



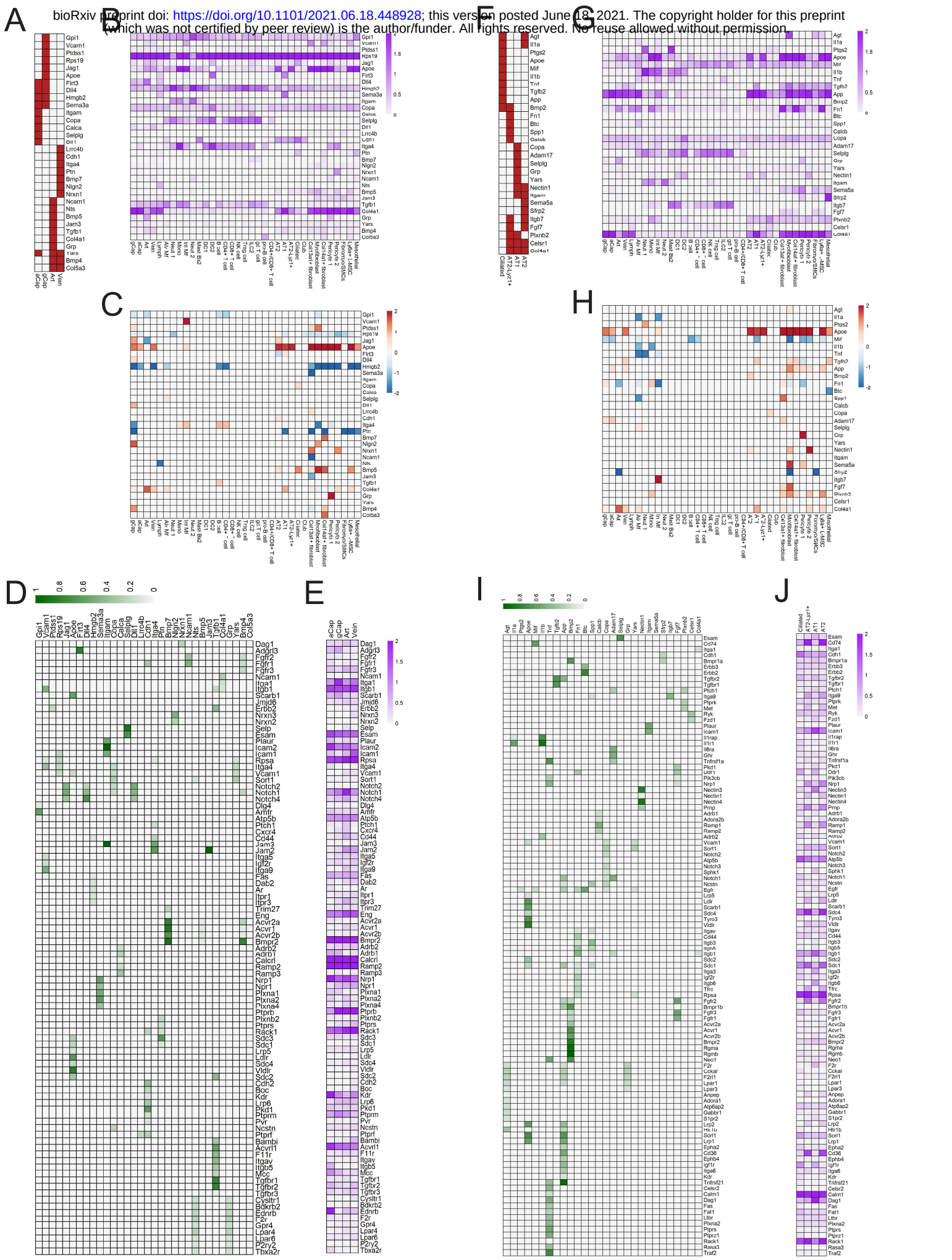



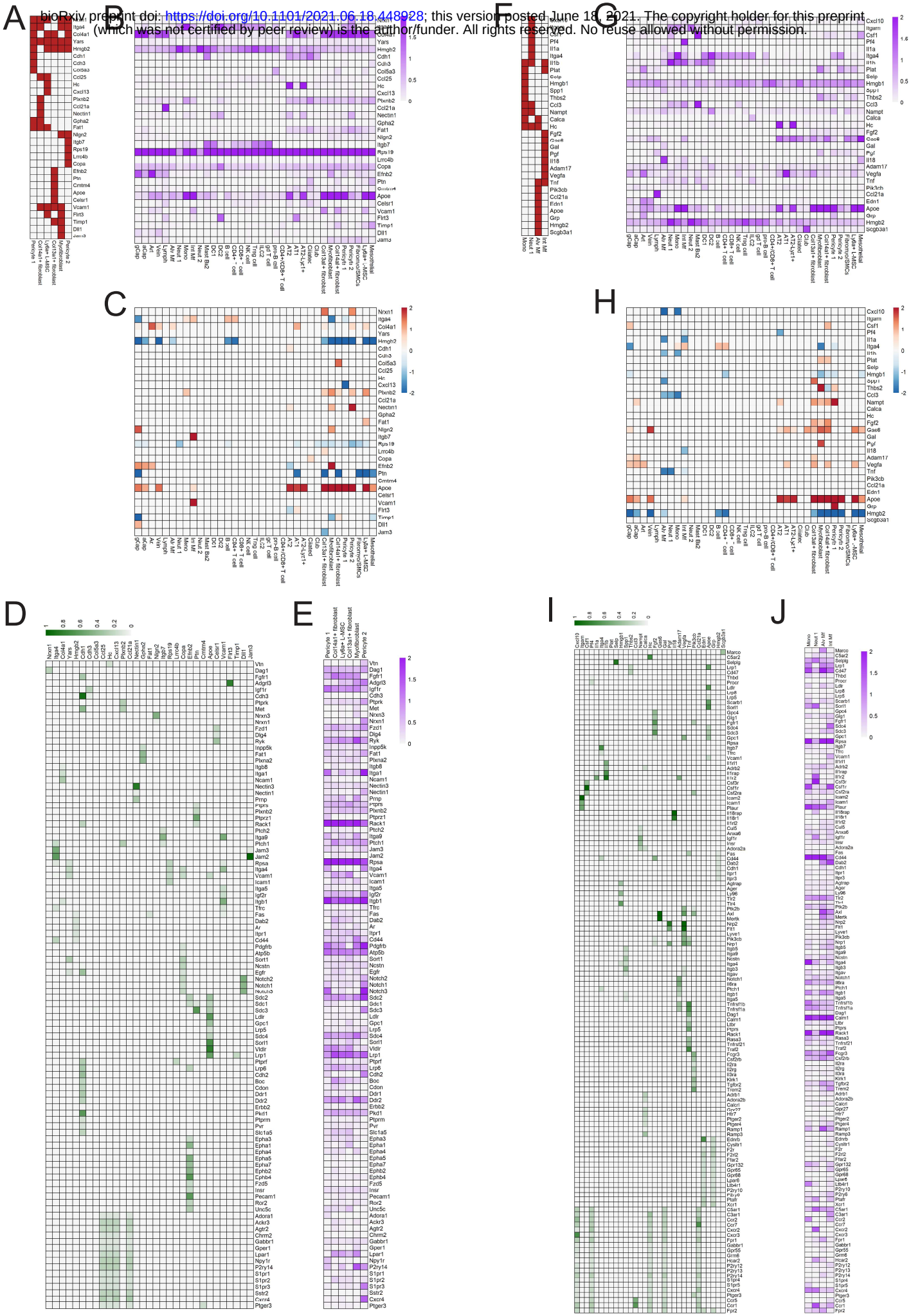

Supplementary Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



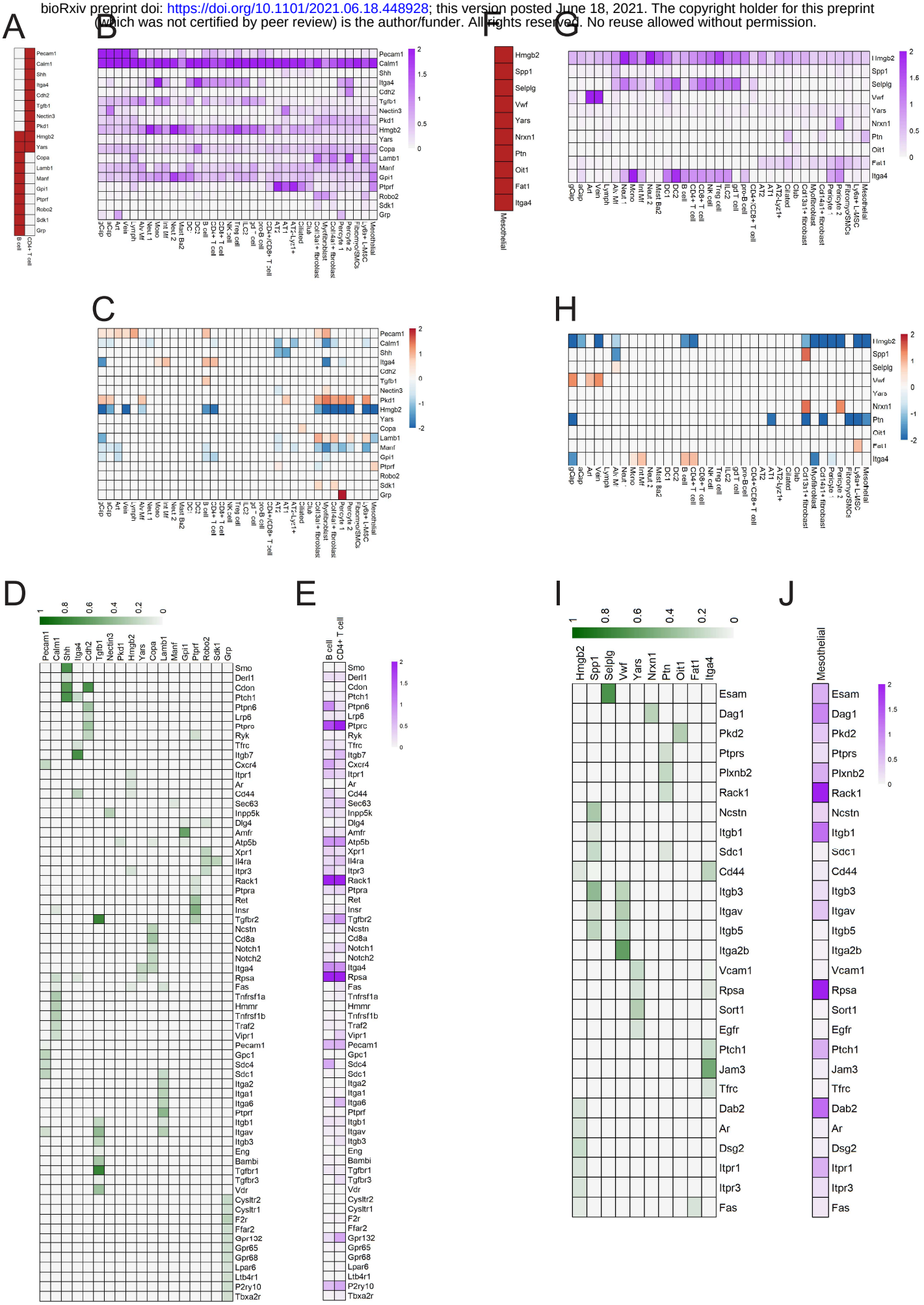

Supplementary Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



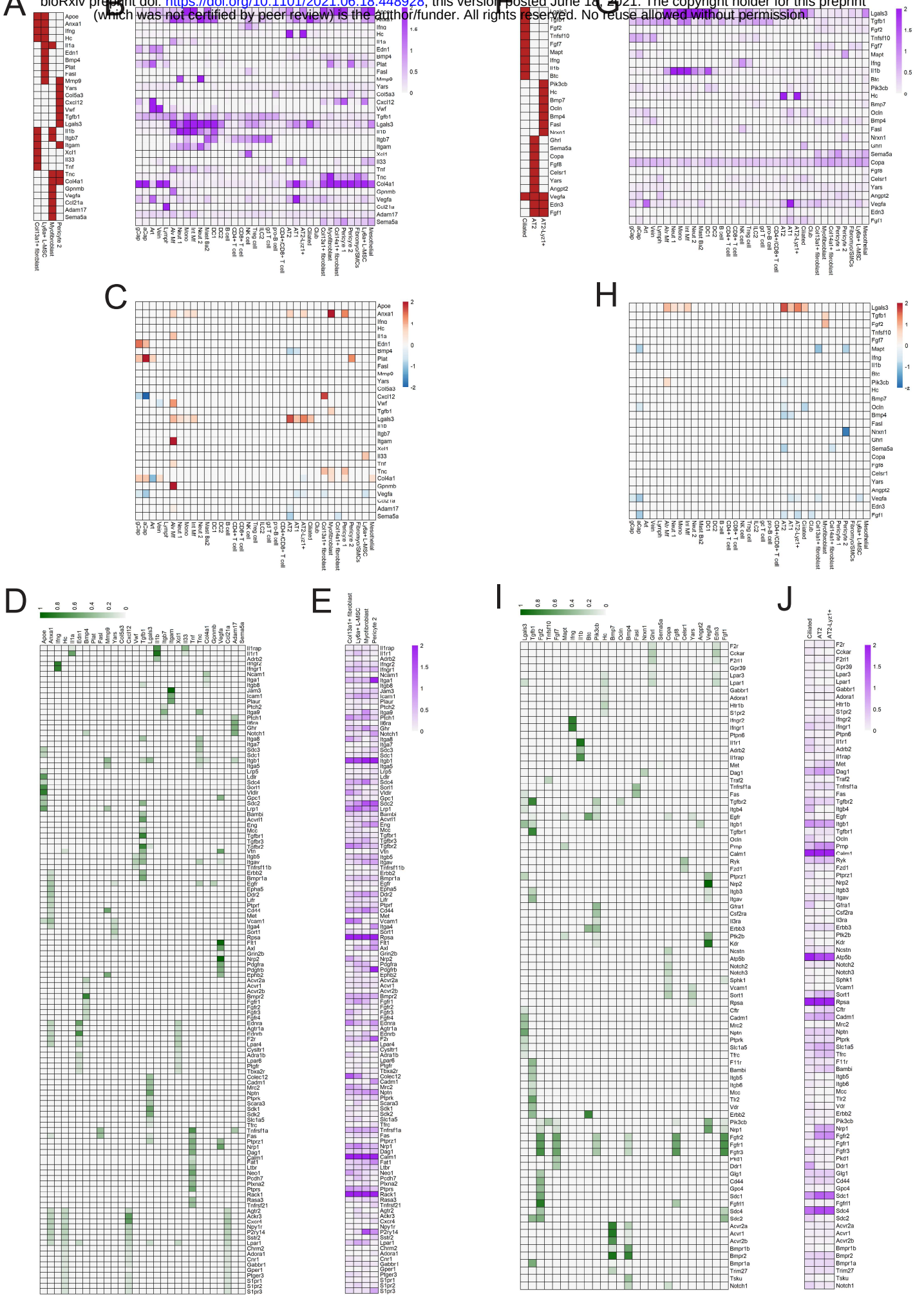

Supplementary Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



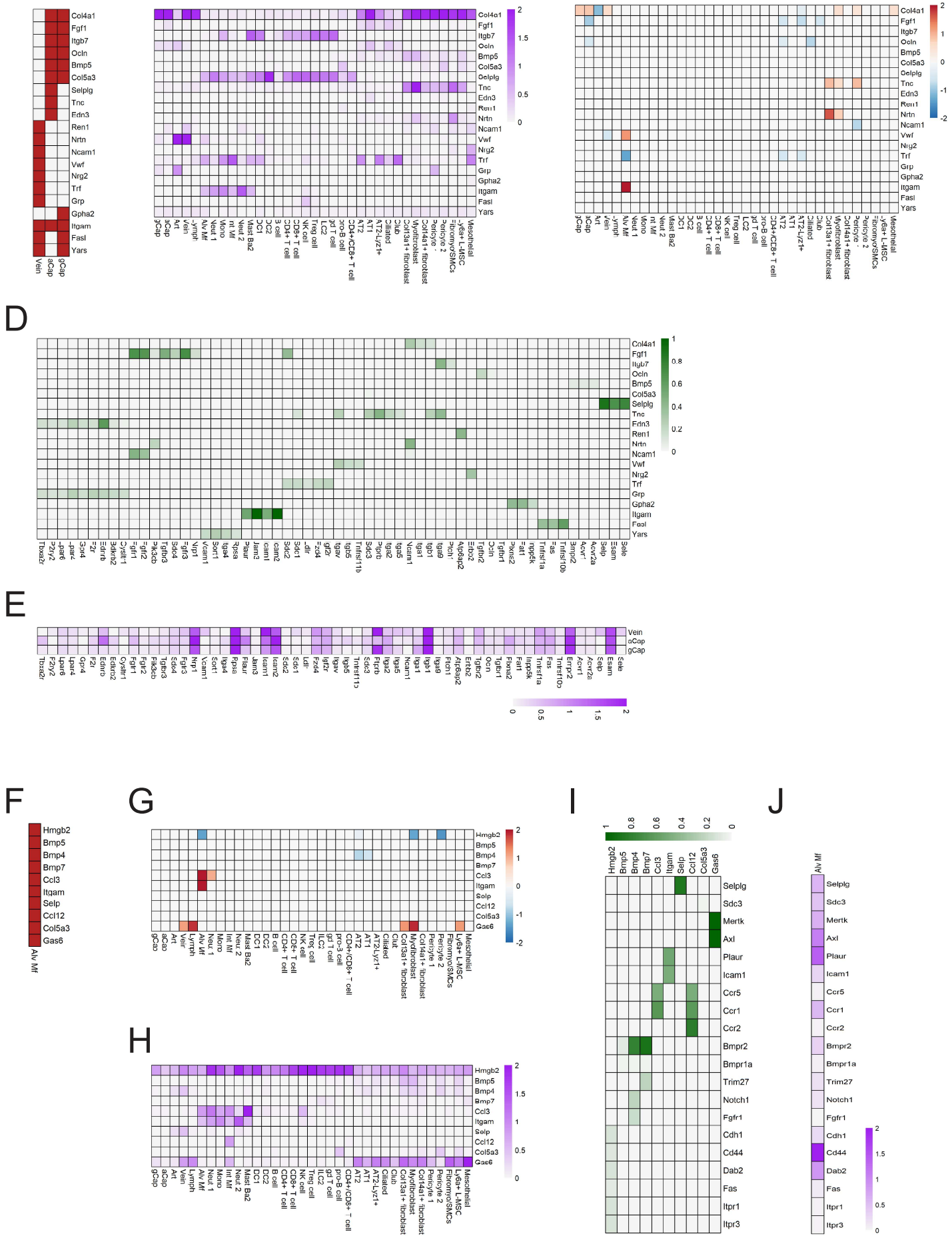

# Supplementary Figure 5

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




Supplementary Figure 6

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.




# Supplementary Figure 7

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Figure 1. Gene expression profile of *Ly6a*$^+$ L-MSCs during late lung development. (A)** Six clusters of stromal cells were previously identified in developing lungs. In the dataset re-analyzed here mice were exposed to room air (21% $O_2$) or hyperoxia (85%$O_2$) from P1 onwards and lungs were harvested at P3, P7 and P14. **(B)** UMAP plots showing the expression of *Ly6a* mRNA (left panel) within the lung stroma and new cluster identities, including the *Ly6a*$^+$ L-MSCs. **(C)** Heatmap of top ten most differentially expressed genes across stromal clusters depicted in panel (B). **(D)** Dotplot depicting expression of routine MSC markers in lung stromal populations. **(E)** UMAP plots depicting the distribution of lung stromal cells based on the developmental age and oxygen exposure. **(F)** Relative contribution of *Ly6a*$^+$ and *Ly6a*$^-$ cells in developing lung stroma at P3, P7 and P14. n = 6 animals/group. Data are presented as means ± SD. Statistical analyses were performed with GraphPad Prism 8.0 and the presence of potential statistical outliers was determined by Grubbs' test. Significance was evaluated by multiple unpaired Student's *t*- test with Holm-Šidák correction for *Ly6a*$^+$ and *Ly6a*$^-$ cells separately. *P* values < 0.05 were considered significant and are depicted. **(G)** Dotplot depicting the expression of most differentially expressed genes in *Ly6a*$^+$ L-MSCs during normal lung development. **(H)** Dotplot depicting the expression of genes that are differentially expressed specifically in *Ly6a*$^+$ L-MSCs and not in other lung stromal clusters during normal lung development. **(I)** Selected developmental age-associated signalling pathways in the *Ly6a*$^+$ L-MSC cluster identified by gene set enrichment analysis (GSEA). All terms are significantly enriched (adjusted p-value < 0.05). Normalized enrichment scores (NES) values were computed by gene set enrichment analysis on fold change-ranked genes. Expression values in Heatmap represent Z-score-transformed log(TP10k+1) values. Expression levels in Dotplots and UMAP plots are presented as log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k.

Data depicted in 1A were adapted from Hurskainen M, Mižíková I, Cook DP, et al. *Single cell transcriptomic analysis of murine lung development on hyperoxia-induced damage.* Nat Commun 2021;12:1565.

**Figure 2. Age-associated gene expression and signalling in the developing *Ly6a*$^+$ L-MSCs. (A)** Circos plot showing inferred cell communications between *Ly6a*$^+$ L-MSCs and other populations in the developing mouse lung. Cell communications associated with increasing developmental age are depicted. Cell types in the top right correspond to receiver populations with the largest

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expression changes in response to increasing age. These cell types are connected to the sender cell types expressing ligands predicted to promote this response. Ligands expressed by the same cell population are coloured the same. **(B)** Heatmap depicting predicted target genes for ligands most likely to be received by normally developing *Ly6a*$^+$ L-MSC population as indicated in **(A)**. The intensity of expression is indicated as specified by the colour legend. **(C)** Heatmap depicting predicted target genes for ligands sent by *Ly6a*$^+$ L-MSC population in normally developing lungs as indicated in **(A)**. The intensity of expression is indicated as specified by the colour legend.

**Figure 3. Hyperoxia-induced gene expression and signalling in the developing *Ly6a*$^+$ L-MSCs. (A)** Dotplot depicting the expression of markers specifically altered by hyperoxia exposure in *Ly6a*$^+$ and *Ly6a*$^-$ cells in the developing mouse lung. **(B)** Selected hyperoxia-regulated signalling pathways in the *Ly6a*$^+$ L-MSC cluster identified by gene set enrichment analysis (GSEA). All terms are significantly enriched (adjusted p-value < 0.05). Normalized enrichment scores (NES) values were computed by gene set enrichment analysis on fold change-ranked genes. **(C)** Circos plot showing inferred cell communications between *Ly6a*$^+$ L-MSCs and other populations in the developing mouse lung. Cell communications induced by exposure to hyperoxia are depicted. Cell types in the top right correspond to receiver populations with the largest expression changes in response to hyperoxia. These cell types are connected to the sender cell types expressing ligands predicted to promote this response. Ligands expressed by the same cell population are coloured the same. **(D)** Heatmap depicting predicted target genes for ligands most likely to be received by *Ly6a*$^+$ L-MSC population in hyperoxia as indicated in **(C)**. The intensity of expression is indicated as specified by the colour legend. **(E)** Heatmap depicting predicted target genes for ligands sent by *Ly6a*$^+$ L-MSC population in hyperoxia as indicated in **(C)**. The intensity of expression is indicated as specified by the colour legend.

**Figure 4. Identification of *Ly6a*$^+$ L-MSCs in the developing lung. (A)** Fluorescent RNA *in situ* hybridization showing localization of L-MSCs identified by the co-expression of *Ly6a* (white) and *Col14a1* (green) in lungs of room air (21% $O_2$) or hyperoxia (85%$O_2$)-exposed developing mice. Scale bar = 200µm for low-magnification (5×, top panels) windows, 50µm for higher-magnification (40×, bottom left panels) windows, and 20µm for high-magnification (63×, bottom right panels) windows. Four 14-days old animals/group were analysed. Expression levels in

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Dotplot are presented as log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k. **(B)** Fluorescent RNA *in situ* hybridization showing co-expression of *Ly6a* (white), *Col14a1* (green), and *Timp1* (pink) in lungs of room air (21% $O_2$) or hyperoxia (85%$O_2$)-exposed developing mice. Scale bar = 200µm for low-magnification (5×, top left) windows, 50µm for higher-magnification (40×, top right) windows, and 20µm for high-magnification (63×, bottom panels) windows. Four 14-days old animals/group were analysed. **(C)** Fluorescent RNA *in situ* hybridization showing co-expression of *Ly6a* (white), *Col14a1* (green), and *Serpina3n* (pink) in lungs of room air (21% $O_2$) or hyperoxia (85%$O_2$)-exposed developing mice. Scale bar = 200µm for low-magnification (5×, top left) windows, 50µm for higher-magnification (40×, top right) windows, and 20µm for high-magnification (63×, bottom panels) windows. Four 14-days old animals/group were analysed.

**Figure 5. Characterization of LY6A$^+$ L-MSCs in normal and impaired mouse lung development. (A)** Mouse pups were exposed to room air (21% $O_2$, grey) or hyperoxia (85%$O_2$, blue) from P1 onwards. Mice were harvested on postnatal day (P)7. Representative histological sections from lungs developing in 21% $O_2$ or 85% $O_2$. Lung morphometry was quantified by the mean linear intercept (MLI) measurement. n = 7 animals/group. Scale bar = 100µm. **(B)** LY6A$^+$ L-MSCs were identified by flow cytometry as CD45-AF647$^-$/CD31-FITC$^-$/CD326(EPCAM)-PeCy7$^-$/LY6A(SCA1)-BV421$^+$ cells and their proportion in lung homogenates was quantified. n = 8-9 animals/group. **(C)** Representative images of undifferentiated LY6A$^+$ L-MSCs and LY6A$^+$ L-MSCs differentiated towards osteogenic and chondrogenic lineages and stained with Alizarin Red S or Alcian Blue, respectively. Scale bar = 250µm. Experiments were performed in quadruplets. **(D)** Expression of routine MSCs surface markers in cultured LY6A$^+$ L-MSCs isolated from room air (21% $O_2$, grey bars) or hyperoxia-exposed (85%$O_2$, purple bars) developing pups as determined by flow cytometry. n = 3-4 animals/group. **(E)** Quantification and representative images of colony formation of cultured LY6A$^+$ L-MSCs isolated from room air (21% $O_2$, grey bars) or hyperoxia-exposed (85%$O_2$, purple bars). n = 4 animals/group. Scale bar =100µm. All data are presented as means ± SD. Statistical analyses were performed with GraphPad Prism 8.0. The presence of potential statistical outliers was determined by Grubbs' test. Significance was evaluated by unpaired Student's *t*-test for analysis in panels (A) and (B), and by multiple unpaired Student's *t*-

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

test with Holm-Šidák correction in panels (D) and (E). *P* values < 0.05 were considered significant and are depicted.

**Figure 6. Gene expression profile of cultured normoxia and hyperoxia-derived LY6A$^+$ L-MSCs. (A)** LY6A$^+$ and LY6A$^-$ stromal cells isolated from lungs of room air (21% $O_2$) or hyperoxia (85%$O_2$)-exposed developing mice were frozen, cultured and sequenced at passage 3. n = 3 animals/group. scRNA-seq identified four clusters of cultured LY6A$^+$ and LY6A$^-$ stromal cells. **(B)** Relative distribution of room air (21% $O_2$) or hyperoxia (85%$O_2$)-derived LY6A$^+$ and LY6A$^-$ cells to the four different clusters. n = 3 animals/group. **(C)** Heatmap of top ten most differentially expressed genes across clusters depicted in panel (A). **(D)** Violin plots depicting expression of routine MSC markers in cultured stromal populations. **(E)** Dotplot depicting expression of oxygen-specific markers in LY6A$^+$ and LY6A$^-$ cultured lung stromal cells. Expression values in Heatmap and violin plots represent Z-score-transformed log(TP10k+1) values. Expression levels in Dotplot and UMAP plot are presented as log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k.

**Figure 7. Identification of novel markers for *in situ* and cultured *Ly6a*$^+$ L-MSCs.**
**(A)** Identifying markers were first established in the *in situ*, or cultured *Ly6a*$^+$ and *Ly6a*$^-$ lung stromal cells based on Supplementary figures 1G-H. Dotplot depicts the expression levels of those markers, most suitable for identification of *Ly6a*$^+$ and *Ly6a*$^-$ lung stromal cells *in situ* in normoxic or hyperoxic animals at P7. (**B**) Identifying markers were first established in the *in situ*, or cultured *Ly6a*$^+$ and *Ly6a*$^-$ lung stromal cells based on Supplementary figures 1G-H. Dotplot depicts the expression levels of those markers, most suitable for identification of normoxia-derived and hyperoxia-derived LY6A$^+$ and LY6A$^-$ lung stromal cells in culture. **(C)** Fluorescent RNA *in situ* hybridization showing co-expression of *Ly6a* (white), *Col14a1* (green), and *Serpinf1* (pink) in lungs of room air (21% $O_2$) or hyperoxia (85%$O_2$)-exposed developing mice. Scale bar = 200µm for low-magnification (5×, top left) windows, 50µm for higher-magnification (40×, top right) windows, and 20µm for high-magnification (63×, bottom panels) windows. Four 14-days old animals/group were analysed.

**Supplementary figure 1. Flowchart depicting the allocation of mice to experimental groups.**

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Flowchart illustrating the group identity of mice sacrificed for the purpuses of the present study. Purple color depicts the groups allocated to hyperoxia (85% $O_2$). Mice sacrificied as a part of previously published scRNA-seq dataset from developing newborn mice are not included.

**Supplementary figure 2. Identification of *Ly6a*$^+$ L-MSCs in the developing lung.**

**(A)** UMAP plots depicting the expression of commonly used MSC markers in lung stromal cells isolated from lungs of room air (21% $O_2$) or hyperoxia (85%$O_2$)-exposed developing mice. **(B)** Metascape functional enrichment analysis for ligands indicated in Fig. 2C**.** Developmental age-associated summary pathways relevant to lung are depicted. **(C)** Quantification of FACS results in lung homogenates. n = 3 animals/group. Data are presented as means ± SD. Significance was evaluated by ordinary one-way ANOVA with Tukey multiple comparisons correction test. *P* values < 0.05 were considered significant and are depicted. LY6A$^+$ **(D)** L-MSCs were isolated from developing mice at P7 and identified by flow cytometry as CD45$^-$/CD31$^-$/CD326(EPCAM)$^-$ /LY6A(SCA1)$^+$ cells. n = 3 animals/group. **(E)** Selected hyperoxia-regulated signalling pathways specific only to the *Ly6a*$^+$ L-MSC cluster identified by gene set enrichment analysis (GSEA). All terms are significantly enriched (adjusted p-value < 0.05). Normalized enrichment scores (NES) values were computed by gene set enrichment analysis on fold change-ranked genes. **(F)** Metascape functional enrichment analysis for ligands indicated in Fig. 3E**.** Hyperoxia-regulated summary pathways relevant to lung are depicted. **(G)** Identifying markers were first established in the P7 *in situ*, or cultured *Ly6a*$^+$ and *Ly6a*$^-$ lung stromal cells based on Supplementary tables 4, 15, 16, 17 and 18, as well as Fig. 3A and 6E. The expression levels of these identifying markers are depicted here in the *in situ* lung stromal cells from the room air or hyperoxia-exposed mice at P7. The intensity of expression is indicated as specified by the colour legend. **(H)** Identifying markers were first established in the P7 *in situ*, or cultured *Ly6a*$^+$ and *Ly6a*$^-$ lung stromal cells based on Supplementary tables 4, 15, 16, 17 and 18, as well as Fig. 3A and 6E. The expression levels of these identifying markers are depicted here in the cultured lung stromal cells from the room air or hyperoxia-exposed mice. The intensity of expression is indicated as specified by the colour legend. Expression values in Heatmap represent Z-score-transformed log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**Supplementary figure 3. Developmental age-associated ligand and receptor activity affecting lung endothelial and epithelial populations.** Panels A-E relate to endothelial populations, panels F-J relate to epithelial populations. **(A, F)** Heatmap depicting top 10 ligands predicted to affect the listed lung cell populations (coloured red). **(B, G)** Heatmap depicting average log(TP10k+1) expression values of ligands for each cell population in the P14 samples (depicted in violet). **(C, H)** Heatmap depicting the log(fold change) expression of ligands in the P14 samples (depicted in red/blue). **(D, I)** Heatmap depicting putative receptors for each ligand according to the prior interaction potential in NicheNet's model (depicted in green). **(E, J)** Heatmap depicting average log(TP10k+1) expression values of receptors for each cell population (depicted in violet). Expression values in violin plots represent Z-score-transformed log(TP10k+1) values. Expression levels in UMAP plots and Dotplots are presented as log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k.

**Supplementary figure 4. Developmental age-associated ligand and receptor activity affecting lung stromal and myeloid populations.** Panels A-E relate to stromal populations, panels F-J relate to myeloid populations. **(A, F)** Heatmap depicting top 10 ligands predicted to affect the listed lung cell populations (coloured red). **(B, G)** Heatmap depicting average log(TP10k+1) expression values of ligands for each cell population in the P14 samples (depicted in violet). **(C, H)** Heatmap depicting the log(fold change) expression of ligands in the P14 samples (depicted in red/blue). **(D, I)** Heatmap depicting putative receptors for each ligand according to the prior interaction potential in NicheNet's model (depicted in green). **(E, J)** Heatmap depicting average log(TP10k+1) expression values of receptors for each cell population (depicted in violet). Expression values in violin plots represent Z-score-transformed log(TP10k+1) values. Expression levels in UMAP plots and Dotplots are presented as log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k.

**Supplementary figure 5. Developmental age-associated ligand and receptor activity affecting lung lymphoid and mesothelial populations.** Panels A-E relate to lymphoid populations, panels F-J relate to mesothelial populations. **(A, F)** Heatmap depicting top 10 ligands predicted to affect the listed lung cell populations (coloured red). **(B, G)** Heatmap depicting average log(TP10k+1)

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expression values of ligands for each cell population in the P14 samples (depicted in violet). **(C, H)** Heatmap depicting the log(fold change) expression of ligands in the P14 samples (depicted in red/blue). **(D, I)** Heatmap depicting putative receptors for each ligand according to the prior interaction potential in NicheNet's model (depicted in green). **(E, J)** Heatmap depicting average log(TP10k+1) expression values of receptors for each cell population (depicted in violet). Expression values in violin plots represent Z-score-transformed log(TP10k+1) values. Expression levels in UMAP plots and Dotplots are presented as log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k.

**Supplementary figure 6. Hyperoxia-induced ligand and receptor activity affecting lung stromal and epithelial populations.** Panels A-E relate to stromal populations, panels F-J relate to epithelial populations. **(A, F)** Heatmap depicting top 10 ligands predicted to affect the listed lung cell populations (coloured red). **(B, G)** Heatmap depicting average log(TP10k+1) expression values of ligands for each cell population in the hyperoxia samples (depicted in violet). **(C, H)** Heatmap depicting the log(fold change) expression of ligands in hyperoxia samples (depicted in red/blue). **(D, I)** Heatmap depicting putative receptors for each ligand according to the prior interaction potential in NicheNet's model (depicted in green). **(E, J)** Heatmap depicting average log(TP10k+1) expression values of receptors for each cell population (depicted in violet). Expression values in violin plots represent Z-score-transformed log(TP10k+1) values. Expression levels in UMAP plots and Dotplots are presented as log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k.

**Supplementary figure 7. Hyperoxia-induced ligand and receptor activity affecting lung endothelial and myeloid populations.** Panels A-E relate to endothelial populations, panels F-J relate to myeloid populations. **(A, F)** Heatmap depicting top 10 ligands predicted to affect the listed lung cell populations (coloured red). **(B, G)** Heatmap depicting average log(TP10k+1) expression values of ligands for each cell population in the hyperoxia samples (depicted in violet). **(C, H)** Heatmap depicting the log(fold change) expression of ligands in hyperoxia samples (depicted in red/blue). **(D, I)** Heatmap depicting putative receptors for each ligand according to the prior interaction potential in NicheNet's model (depicted in green). **(E, J)** Heatmap depicting average log(TP10k+1) expression values of receptors for each cell population (depicted in violet).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.18.448928; this version posted June 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Expression values in violin plots represent Z-score-transformed log(TP10k+1) values. Expression levels in UMAP plots and Dotplots are presented as log(TP10k+1) values. Log(TP10k+1) corresponds to log-transformed UMIs per 10k.