Title: Identification of a multipotent lung progenitor for lung regeneration

Authors: Chava Rosen[1,2,3], Elias Shetzen[2], Irit Milman -Krentsis[2], Ran Orgad[1], Xiaohua Su[2], Raj Yadav[2], Michal Shemesh[4], Adi Biram[1], Ziv Shulman[1], Smadar Eventov-Friedman[5], Mukesh Maharjan [2], Yuan Qi[6], Jing Wang[6], and Yair Reisner[2*]

Affiliations:

[1] Department of Immunology, Weizmann Institute of Science, Rehovot, Israel

[2] Department of Stem Cell Transplantation and Cellular Therapy, University of Texas MD Anderson Cancer Center, Houston, Texas, USA.

[3] Department of Neonatology, Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Israel

[4] Cell Observatory of the MICC Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot, Israel

[5] Department of Neonatology, Hadassah Medical Center, Faculty of Medicine, Hebrew University of Jerusalem, Israel

[6] Department of Bioinformatics and Computational Biology, MD Anderson Cancer Center, Houston, Texas, USA

*Correspondence to YReisner@mdanderson.org

Summary

We recently showed that intravenous infusion of mouse or human, fetal or adult lung cells following conditioning of recipient mice leads to lung chimerism within alveolar and bronchiolar lineages, in distinct 'patches' containing both epithelial and endothelial cells. We show here, using *R26R-Confetti* mice as donors, that these multi-lineage patches are derived from a single lung progenitor. FACS of adult mouse lung cells revealed that the putative patch-forming progenitors co-express the endothelial marker CD31 (PECAM-1) and the epithelial marker CD326 (EPCAM). Transplantation of lung cells from transgenic Cre/lox mice expressing membrane GFP under the VEcad promoter (VEcad-Cre-mTmG), led to GFP+ patches comprising both GFP+ endothelial and epithelial cells in vivo, and in ex-vivo culture of CD326+CD31+ progenitors. Single cell mRNAseq of CD326+CD31+ lung cells revealed a subpopulation expressing canonical epithelial and endothelial genes. Such double positive GFP+NKX2.1+SOX17+ cells were also deteceted by immuno-histologial staining in lungs of VEcad-Cre-nTnG (expressing nulear GFP) mice in proximity to blood vessels. These findings provide new insights on lung progenitros and lung development and suggest a potential novel approach for lung regeneration.

Introduction

End stage respiratory diseases are among the leading causes of death worldwide, with more than 5.5 million deaths annually (World Health Organization data for 2020). Today, the only potential curative treatment for these conditions is by replacement of the damaged organ with a lung transplant. Due to a shortage of suitable organs, many patients die on the transplant waiting list, and therefore lung diseases are prime candidates for stem cell therapy.

Various cell populations have been shown to exhibit regenerative potential, including BM-derived cells (Weiss, Ortiz, 2013), lung-derived p63+ cells (Zuo et al., 2014), LNEP (lineage negative epithelial progenitors) (Vaughan et al., 2014) and mouse and human Sox9+ cells (Nichane et al., 2017) (Ma et al., 2018). Recently, we showed that fetal lung progenitors could potentially offer an attractive source for transplantation in mice, provided that the lung stem cell niche in the recipient is vacated of endogenous lung progenitors by adequate conditioning. Thus, in a procedure akin to bone marrow transplantation (BMT), a single cell suspension of mouse or human fetal lung cells harvested at the canalicular phase of gestation (20–22 weeks in human, and E15–E16 for mouse) and infused intravenously (I.V) following conditioning of recipient mice with naphthalene and 6Gy TBI, led to marked lung chimerism within alveolar and bronchiolar lineages. This chimerism was associated with significantly improved lung function (Rosen et al., 2015). More recently, this approach for lung chimerism induction was extended to transplantation of a single cell suspension of adult mouse lung donors (Milman Krentsis et al., 2018), (Hillel-Karniel et al., 2020) (Milman Krentsis et al., 2022) requiring about three fold higher cell doses to attain a similar level of chimerism (Milman Krentsis et al., 2018). Notably, a large proportion of the

donor-derived patches exhibit lung compartments comprised of multiple lineages, including epithelial and endothelial cells.

Considering that the lung patches are observed within confined borders similarly to spleen colonies which are known to be derived from a single multi-potent hemtopoietic progenitor, we hypothesized that these lung patches might represent a clonal derivation from a yet unknow multi-potent lung progenitor. To address this hypothesis we transplanted fetal or adult lung cells from *Rosa26-Confetti* mice bearing a multicolor Cre reporter system (Snippert et al., 2010). This four-color Cre recombination system provides a fetal or adult lung cell preparation in which each cell expresses just one randomly determined colour. Thus, the likelihood that each cell within a doublet in the transplanted cell population would be of the same colour is markedly reduced. We demonstrate that all donor-derived lung patches developing after transplantation are monochromatic, strongly supporting the clonal origin of donor-derived lung patches observed after transplantation, in striking resemblance to the spleen colony forming cells typically identified after BMT. These results provide definitive evidence of a single multi-potential lung progenitor capable of differentiating into diverse lung cell lineages.

In line with our observation that large proportion of the patches contain both endothelial and epithelial cells, and that each such patch is derived from a single progenitor, we searched for a putative multipotent lung progenitor capable of differentiating along these two distinct lineages. Further characterization of these progenitor cells by FACS showed that they exhibit a dual profile, expressing markers typical of both endothelial and epithelial fates. Furthermore, the presence of a unique cluster of double positive cells is supported by single cell RNAseq and by immune-histology of lungs of VEcad-Cre-nTnG transgenic mice.

Results

Long term chimerism in recipient lungs after transplantation of td-tomato labeled cells

In our previous studies we followed lung chimerism by immunohistology for up to 4 months after transplantation (Rosen et al., 2015), (Milman Krentsis et al., 2018). We have now extended the chimerism follow-up period to 6-8 months, showing by immunohistology the presence of donor-derived epithelial (club cells, AT1 and AT2 cells) and endothelial cells (Supplementary Fig. S1). Furthermore, as shown in Fig. 1, scRNAseq analysis of FACS purified CD45-Td-Tom+ (donor) or CD45-Td-Tom- (host) cells from lungs of chimeric mice (Fig. 1A,B), or CD45- cells from lungs of normal untransplanted mice, revealed comparable clusters of epithelial and endothelial cells with similar representation of each cluster (Fig.1C,D). Atribution of these clusters to different lung cell types was defined by expression of canonical gene markers of AT1 (AQP5, HOPX, PDPN), AT2 (SPC, LAMP3), ciliated (*FOXJ1*) and club (Scgb1a1) cells as well as endothelial (CDH5, SOX17, PECAM1) cells (Supplemntary Fig. S2-S5). Thus cluster 5 comprising predominantly AT1 cells represented 3.5, 5.1, 3.9 percent of the CD45- lung cells of normal control, sorted donor type cells and sorted host type cells, repectively. Cluster 10 comprising AT2 cells represented 0.6, 0.8 and 0.8 percent, respectively. Cluster 11comprising bronchial epitehelial cells cells represented 0.4, 0.5, and 0.4 percent respectively. Clusters 0, 2, 4 and 9 comprising different types of endothelial cells represented the majority of CD45- cells in all three samples.

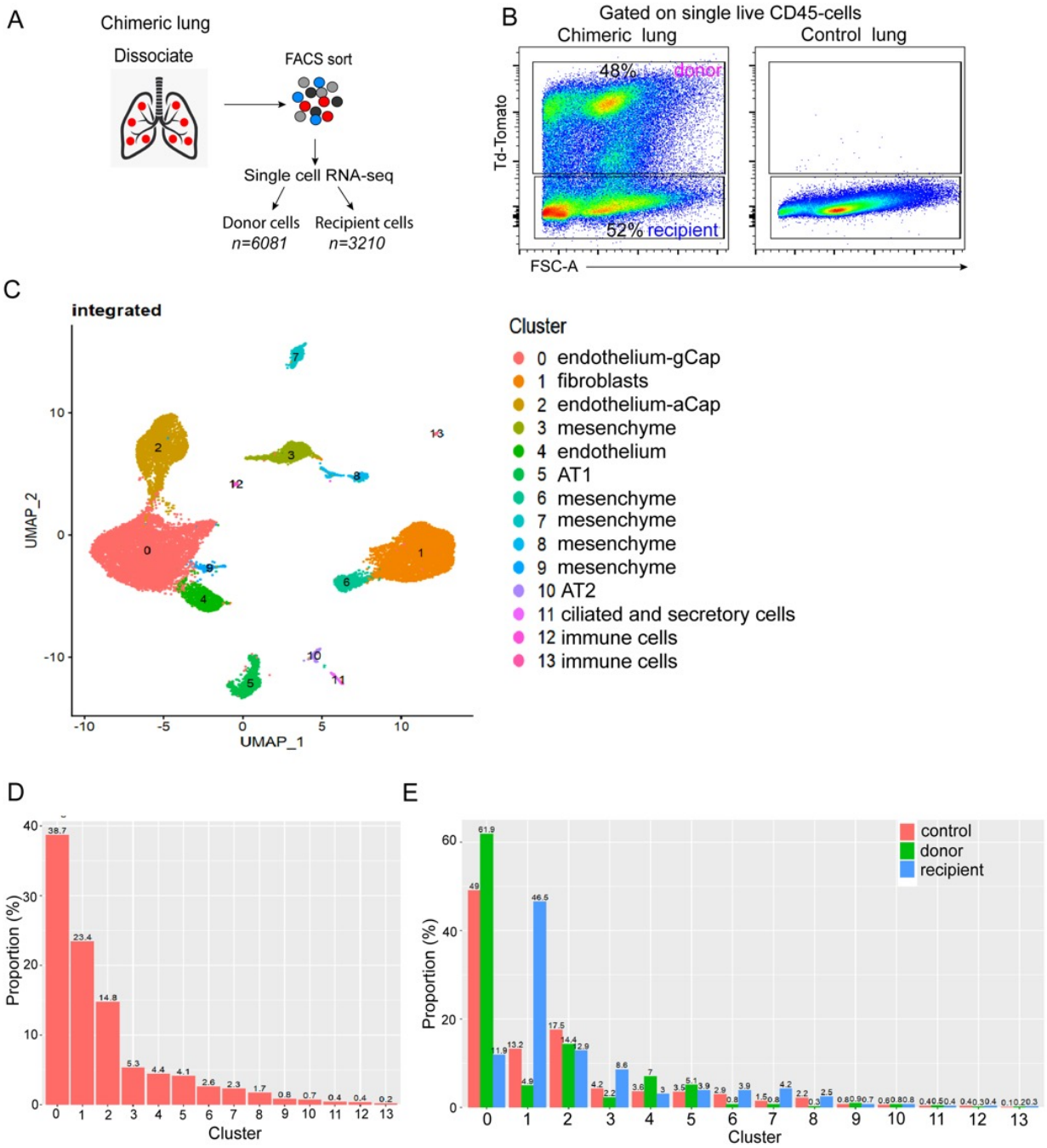

Fig. 1. sc-RNA seq analysis of chimerc lungs 6 months after transplantation of lung cells from Td -Tomato+ donors.

(A) Cell sorting scheme of host and donor lung cells from chimeric lungs. Three C57BL chimeric lungs transplanted with Td-Tomato donor cells were enzymatically dissociated and pooled for FACS purification.

(B) Donor and recipient derived cells were separated by Td -Tomato expression. Lung cells from C57BL non-transplanted mice was used as negative control.
(C) Transcriptome analysis of FACS purified lung cells described in (B). Clustering of transcriptomes, using UMAP: n=6081 donor, n=3210 recipient and n=6000 control lung cells were analyzed. The figure depicts annotations of integrated clusters of all three samples, based on gene signature. (D) Proportion of each cluster out of total cells in the merged dataset. (E) Proportion of each cluster within host, donor and control lung cells.

Transplantation of lung progenitors from fetal or adult Cag-Cre ER2 'Confetti' mice results in monochromatic patches. To unequivocally examine our hypothesis that the lung patches observed following transplantation are derived from a single progenitor, we carried out a set of experiments making use of the multicolor '*R26R-Confetti*' mouse reporter system (Snippert et al., 2010) as donors. In these mice, the recombination process is independent in each cell and stochastic. Daughter cells will all produce the same fluorescent protein, and if indeed all the cells within each patch are derived from a single cell, they should all be of the same color. Multiple studies used this system to demonstrate clonal behavior in organs such as the intestine (Snippert et al., 2010),(Schepers et al., 2012) lung (Desai, Brownfield & Krasnow, 2014), (Barkauskas et al., 2013), brain (Baggiolini et al., 2015), kidney, mammary gland, and others (Rios et al., 2014) (Scheele et al., 2017) (Jamieson et al., 2017) (Ganuza et al., 2017) (Lazzeri et al., 2018).
Thus, we began investigating lung patch formation following transplantation of lung cell suspension from E16 *R26R-Confetti* donors into RAG-2 recipients preconditioned with naphthalene and 6Gy TBI (see schematic overview in Supplementary Fig. S6A).

Tamoxifen (TMX) was administered to pregnant females at E12, the embryos were harvested at E16, and the fetal lungs, containing fluorescent cells were isolated under a fluorescent microscope. Approximately 5-6% of the harvested cells expressed one of the fluorescent tags. Two photon micrographs depict the typical population of monochromatic cells prior to transplantation, each expressing one of the four fluorescent tags (Supplementary Fig. S6B).

Transplantation of fetal lung cells from Tamoxifen-induced donors into naphthalene treated and irradiated recipient mice, resulted in discrete monochromatic lung patches, each expressing one of the four fluorescent proteins, strongly indicating that each patch is likely derived from a single lung progenitor (Supplementary Fig. S6C,D).

Next, based on our recent finding that patch forming progenitor cells are also present in the adult mouse lung, we applied the same approach to analyze the origin of lung patches after transplantation of adult lung cells. To this end, we used the protocol described above, but since the frequency of patch forming cells in the adult lung is about 3-4 fold lower compared to that found in E16 fetal lungs (Milman Krentsis et al., 2018), and only a fraction of the cells undergo the Cre recombination, we used a higher dose ($16x10^6$) of lung cells for transplantation (Fig. 2A). As shown in Fig. 1B and Supplementary Movie 1, following TMX treatment, monochromatic GFP or YFP or RFP or CFP fluorescent cells could be documented by two photon microscopy within the adult lung (Fig. 2B), and as shown by confocal microscopy, these tagged cells were randomly distributed throughout the entire lung (Fig. 2B). To evaluate the clonality of donor-derived patches at 8 weeks after transplantation, we used confocal, two photon, and light sheet microscopy.

Evaluation of chimeric lung slices by fluorescence microscopy demonstrated discrete monochromatic patches (Fig. 2C,D). Two photon snapshots of whole mount chimeric

lung, allowing detection of all 4 chomophores , documented monochromatic patches expressing either membranous CFP, nuclear GFP, or cytoplasmatic RFP or YFP (Fig. 2E). The full depth 2-photon microscopy scan of these monochromatic patches in the whole mount lung tissue is presented in Supplementary Movie 2. In addition, as shown in Fig. 2F,G and Supplementary Movie 3, light sheet microscopy of "cleared" chimeric lungs also revealed distinct monochromatic patches.

Altogether, we analyzed a total of 50 fields collected from 15 chimeric mice, including 12 transplanted with adult lung cells, and 3 with E16 fetal lung cells, and found the exclusive presence of monochromatic donor-derived patches. Testing our experimental distribution of single-color (n=50) vs. two-color (n=0) clones, against the theoretical distribution of 25% vs. 75% ($\chi 2_1 = 150$), by $\chi 2$ distribution test suggests that it is highly unlikely that any clone is derived from two cells ($p < 0.001$), demonstrating that the lung patches observed after transplantation of fetal or adult lung cells are derived from a single progenitor of *R26R-Confetti* donors (Supplementary Fig. S7, S8, S9).

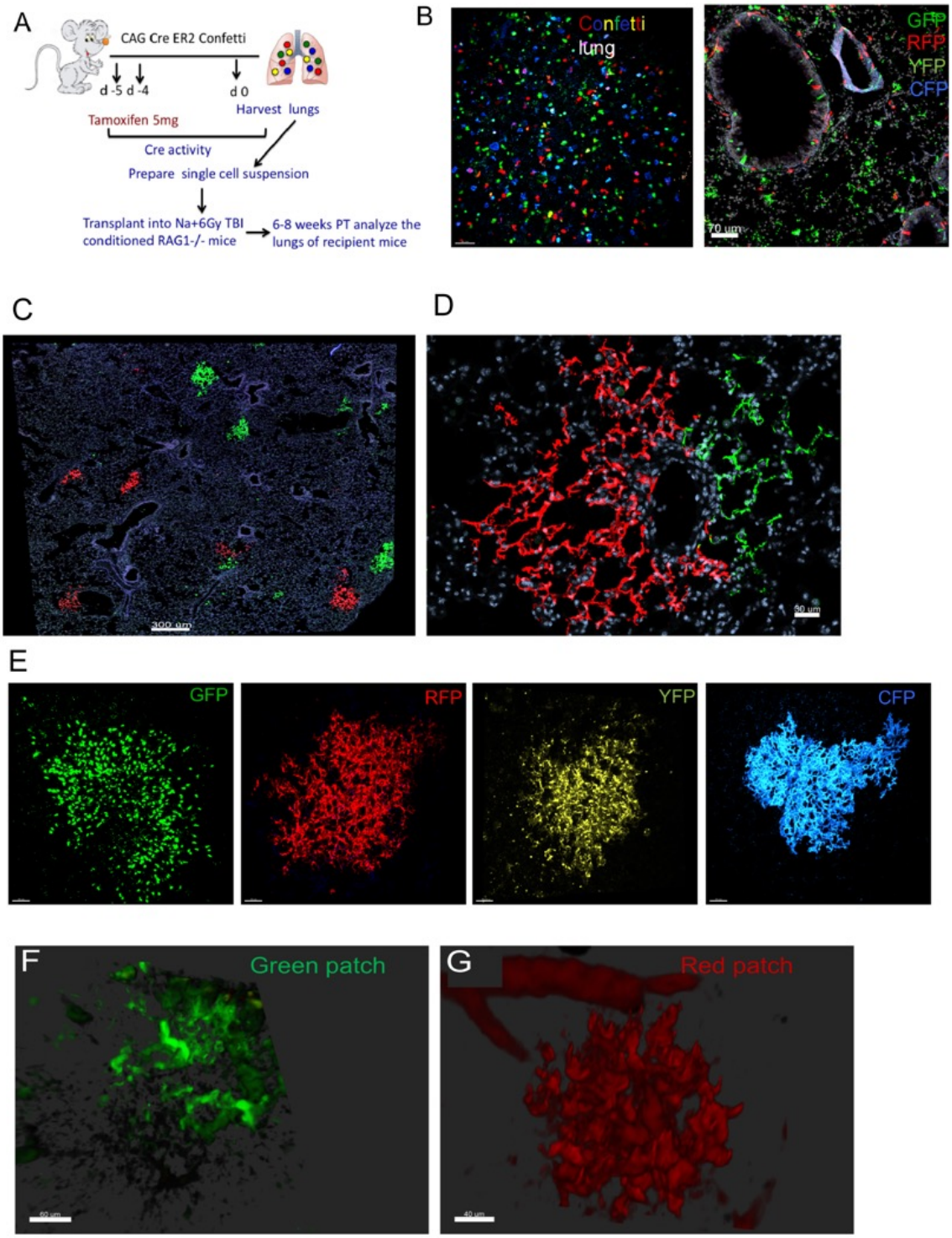

Fig. 2. Lung chimerism analysis at 8 weeks after transplantation of adult *R26R-Confetti* lung cells.

(A) Schematic presentation of the experimental procedure. (B, left image) Two photon microscopy of the adult lung prior to transplantation revealing single cells

each expressing one of the four tag colors after Cre recombination; n=3. (B, right image) Confocal image of adult *R26R-Confetti* lung slice demonstrating bronchial and alveolar parts of the lung tissue, and diffuse and random localization of the fluorescent cells within the donor lung. The results are representative of n=6 *R26R-Confetti* mice from 2 independent experiments. (C, D) Fluorescence microscopy of a section of chimeric lung transplanted with Confetti lung cells, demonstrating several single color patches (C, scale bar=200um), and a pair of neighboring single color patches (D, scale bar=50um). (E) Typical monochromatic patches, each exhibiting a distinct color in whole mount chimeric lung; n=4 chimeric lungs were evaluated by two-photon microscopy, scale bar=50um.

(F, G) Transparent chimeric lung produced by a clearing procedure was rendered and analyzed by light sheet microscopy, n=2 chimeric lungs were evaluated by LSM; in (F) scale bar= 60um and in (G), scale bar=40um.

Two distinct patch-forming lung cell progenitors

Considering that a large proportion of the observed patches contain both endothelial and epithelial cells and that each such a patch is derived from a single progenitor, we searched for a putative multipotent lung progenitor capable of differentiating along these two distinct lineages. As shown in Fig. 3A, FACS analysis of the CD45- lung cell population for expression of epithelial (CD326+) and endothelial (CD31+) markers clearly reveals the presence of a unique CD45-CD326+CD31+ double positive subpopulation exhibiting a wide range of CD326+ expression. This analysis revealed in 17 independent experiments, an average percentage of 4.62±2% CD45-CD326+CD31+ cells; 3.42±1.5% CD45-CD326+CD31- cells; 46.2±7.5% CD45-CD326- CD31+ cells; and 39.1±8.2% CD45-CD326-CD31- cells (Fig. 3B).

Next, we attempted to isolate by FACS the four subopulations as shown schematically in Fig. 3C, and define their capacity to form patches.

To trace donor-derived patches 8 weeks after transplantation of purified cell populations, we made use of labeled donor mice, including GFP (C57BL/6-Tg (CAG-EGFP)1Osb/J), mTmG (*Gt(ROSA)26Sor*$^{tm4(ACTB\text{-}tdTomato,\text{-}EGFP)Luo}$/J), and nTnG (*Gt(ROSA)26Sor*$^{tm1(CAG\text{-}tdTomato*,\text{-}EGFP*)Ees}$/J) mice (mTmG mice express membrane Td-tomato, and nTnG mice express nuclear Td-tomato) (Fig. 3D). The unique localization of the fluorescent tag either in the membrane or the nuclei of the sorted lung cells enabled tracking of membrane-bound, cytoplasmatic and nuclear epithelial and endothelial markers at the single cell level within the monochromatic patches of chimeric lungs, and to characterize the cellular composition of the patches after transplantation of the sorted lung cell subpopulations.

Notably, in a total of 16 experiments transplanting sorted cells, patch forming activity could be found only upon transplantation of CD45-CD326+CD31+ double positive (21 of 37 mice) or single positive CD45-CD326-CD31+ cells (28 out of 61 mice). Notably , no patch-forming activity was found in the single positive CD45-CD326+ CD31- epithelial cell fraction, or in the double negative CD45- CD326- CD31- population (Fig. 3E). It is possible that sorted cells might need additional facilitating cells in the donor cell preparation for the early steps of colonization in the recipient lungs. We therefore attempted to transplant 0.3-0.5 x10^6 sorted cells from lungs of mTmG or nTnG mice, together with (Fig. 3F) or without (Fig. 3G, H) 0.5x10^6 unsorted lung cells from GFP+ donors. These experiments clearly demonstrate that transplanted FACS sorted cells are capable of forming patches even in the absence of potentialy supporting cells from the co-transplanted unseparated GFP+ lung cell preparation (Fig. 3G, H).

Staining of the donor-derived patches for epithelial alveolar markers including AQP-5, HOPX and SPC, or for endothelial markers including CD31, ERG, and SOX17, demonstrated that the patches formed following transplantation of the single positive CD326-CD31+ endothelial cells were comprised predominantly of endothelial cells (Fig. 4A,B), while patches formed after transplantation of the double positive CD326+CD31+ subpopulation contained both endothelial and epithelial cells (Fig. 4C-G). This result is strongly supported by the use of nuclear staining to distinguish cellular borders between neighboring cells. To this end, co-localization of the donor-derived nT marker with nuclear ERG/SOX17 or with nuclear expression of HOPX, enabled us to identify donor-derived endothelial or AT1 epithelial cells, respectively (Fig. 4C-E; Supplementary Fig. S7). A similar distinction of AT2 cells could be attained by staining with anti-SPC antibody (Fig. 4F) or an smRNA FISH probe for SPC (Fig. 4G). Quantitative analysis of the cell numbers comprising the patches suggests that those found after transplantation of CD326+CD31+ lung cells were larger than those found after transplantation of CD326-CD31+ cells (Fig. 4H). This analysis confirmed the difference in cellular composition between the two types of patches, with co-localization of endothelial and epithelial cells mostly found in the patches formed following transplantation of the double positive CD326+CD31+ lung progenitors (Fig. 4I).

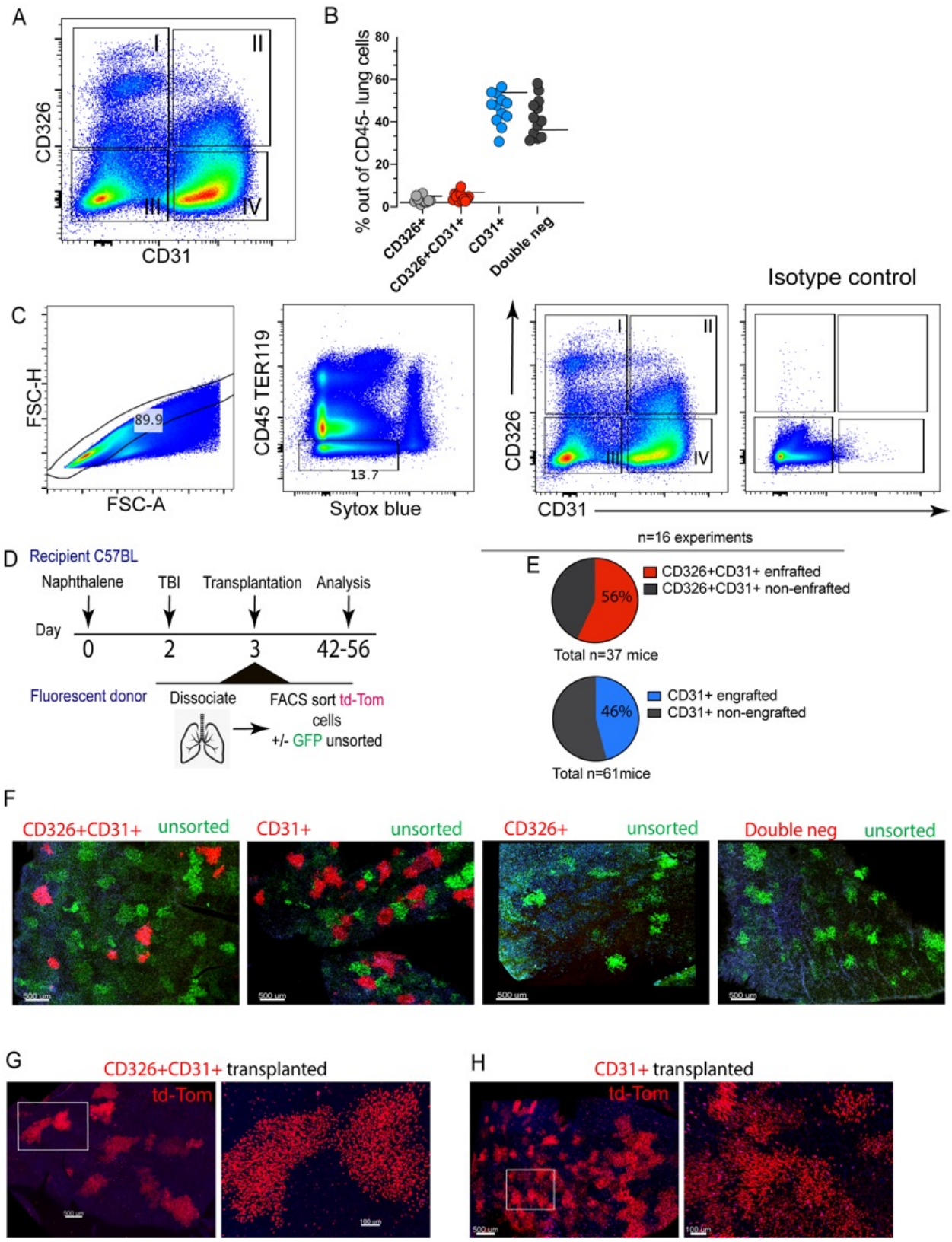

Fig. 3. Donor-derived lung patches after transplantation of different lung cell sub-populations. (A) Typical FACS analysis of CD45- lung cell subpopulations prior

to transplantation , including CD326+CD31-, CD326+CD31+, CD326-CD31+ and CD326-CD31- cells.

(B) Percentage of each subpopulation prot to transplantation, out of the CD45- non-hematopoietic lung cell population in 16 experiments.

(C) FACS strategy of lung cells: Lung cells were first gated on single cells based on forward scatter area and height, and thereafter on CD45- TER119- live cells (Sytox blue negative). The gated cells were then sorted into four fractions as follows: I) CD326+CD31-; II) CD326+CD31+; III) CD326-CD31-; IV) CD326-CD31+. (D) Schematic representation of the transplantation experiments using FACS purified cells. (E) Percentage of chimeric mice exhibiting donor-derived patches out of the total number of transplanted mice (n=16 experiments). (F) Donor-derived lung patches 6 weeks after transplantation of 0.3-0.5 x10^6 sorted double positive CD326+CD31+ cells, or single positive CD326-CD31+ endothelial cells from nTnG donors (red) mixed with 0.5 x10^6 unsorted cells from GFP+ donors (green). The whole mount confocal images shown for each group are representative of n=16 experiments, with at least n=10 mice in each group; scale bar=500um. Out of the total transplanted mice (n=98), 70% were transplanted with sorted td-tomato+ unsorted GFP cells, and 30% were transplanted only with the FACS td-tomato sorted cells. (G, H) Representative images of whole mount lungs of mice transplanted with 0.5X10^6 nTnG CD326+CD31+ or CD31+cells, in the absence of GFP+ co-transplanted unsorted lung cells. Donor-derived red patches express nuclear td-tomato. Images under low (scale bar=500 um) and high magnification (scale bar=100um) are representative of n=5 mice in each group.

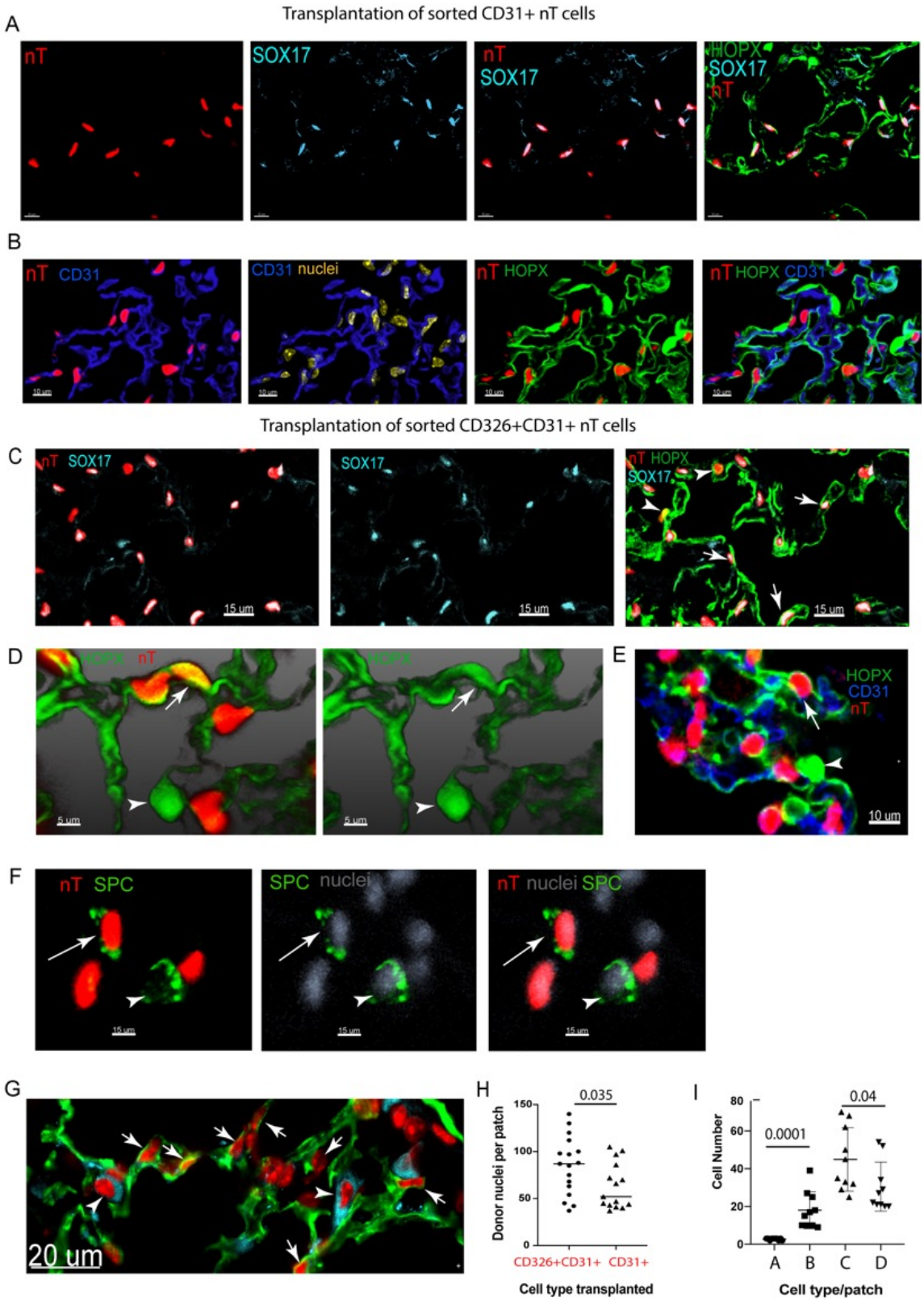

Fig. 4. Different cellular composition of donor-derived patches after transplantation of sorted CD326-CD31+ versus CD326+CD31+ lung cell subpopulations .

(A) Staining of typical lung patches derived from sorted nTnG CD326-CD31+ lung cells (red). Red cells positive for the endothelial nuclear marker SOX17(cyan) indicate donor-derived endothelial cells; scale bar=10um. (B) Staining for cell surface endothelial marker CD31(blue) and epithelial marker HOPX (green) in donor-derived patches formed after transplantation of CD326-CD31+ cells. Left and middle left - Staining for nTnG (red) and CD31 (blue) indicating donor-derived endothelial cells. Middle right -Staining for HOPX; right – Triple staining for nTnG+ (red), CD31 (blue) and HOPX (green). Donor-derived endothelial CD31+ cells reside in close proximity to recipient HOPX+ AT1 cells (indicated with arrow heads), while donor-derived AT1 cells were not detected; scale bar=10um.

(C-F) Typical staining of lung patches derived following transplantation of sorted nTnG CD326+CD31+ cells, showing donor-derived epithelial and endothelial cells. (C) Left: double staining for nTnG (red) and SOX17 (cyan); Middle: single staining for SOX17; Right: Triple staining for nTnG (red), SOX17 (cyan) and HOPX (green), showing donor-derived AT1 epithelial cells (arrowheads) and endothelial cells (arrows); scale bar =15um. (D) High magnification image of typical staining for HOPX showing donor (red, indicated with arrow) and host (indicated with arrow head) AT1 cells. (E) High magnification of triple staining of CD326+CD31+-derived patch, demonstrating the presence of donor-derived (red) AT1 and endothelial CD31+ cells, in close proximity. Arrow indicates donor-derived AT1 cells, and arrowhead indicates the host-derived AT1 cells; scale bar=10um in (D) and (E). Images are representative of n=3 mice. (F) Staining of a donor-derived patch with anti-SPC antibody (green), arrow indicates the donor derived nT-AT2 cell and arrow head shows the host AT2 cell. (G) Staining of a typical donor-derived patch for CD31 (green) and single molecule RNA FISH probe for SPC (cyan), demonstrating

endothelial and AT2 cells; scale bar=20um. (H, I) Graphical summary of quantitative differences between the composition of patches derived from transplantation of CD326+CD31+ and CD326- CD31+ cells. (H) Donor-derived nuclei/patch, showing that CD326+CD31+ progenitors produce larger patches compared to CD326- CD31+ cells; p=0.035, Student's t-test; n=20 patches were evaluated from n=3 mice in each group. (I) Absolute number of donor-derived epithelial and endothelial cells per patch, after transplantation from CD326+CD31+ or CD326-CD31+ cells: A and B - depict epithelial cells derived from transplanted CD326-CD31+ and CD326+CD31+ populations, respectively; p=0.0001, Student's t-test. C and D depict endothelial cells derived from transplanted CD326-CD31+ and CD326+CD31+ populations, respectively; p=0.04, Student's t-test; n=10 patches from 2 mice were evaluated for each group.

Fate tracing reveals differentiation of lung cells from VEcad-GFP donors towards epithelial fate

To further interrogate the potential multi-potency of the patch forming cells, we used VEcad-Cre mTmG and VEcad-Cre nTnG transgenic mice expressing GFP under the VE-cadherin (Cadherin 5) promoter. Transplantation of unsorted lung cells from such GFP+ donors led to green patches in which GFP+ endothelial and epithelial cells could be clearly identified (Fig. 5), further supporting the possibility that the patch forming lung cell is expressing a canonical endothelial marker such as VEcad, and its ability to generate epithelial AQP-5 + AT1 and Lamp-3+ AT2 alveolar cells. Furthermore, cells within the CD326+ CD31+ subpopulation purified by FACS from the lungs of adult mice, were also able to form epithelial colonies in-vitro with clear presence of vasculature within the organoids while CD326+ CD31- sorted cells

formed epithelial organoids without any sign of vascular formation. (Supplementary Fig. S10 A-C)

Culturing of FACS purified GFP+ and TD-Tom+ CD326+ CD31+ cell mixtures under 3D conditions in matrigel , resulted in generation of either green or red vascularized organoids, confirming the clonal properties of single CD326+ CD31+ cells (Supplementary Fig. S10D). Thus, these sorting experiments, using an ex-vivo assay, further support our hypothesis that the CD326+ CD31+ patch forming cell can uniquely give rise to both epithelail and endothelial cells .

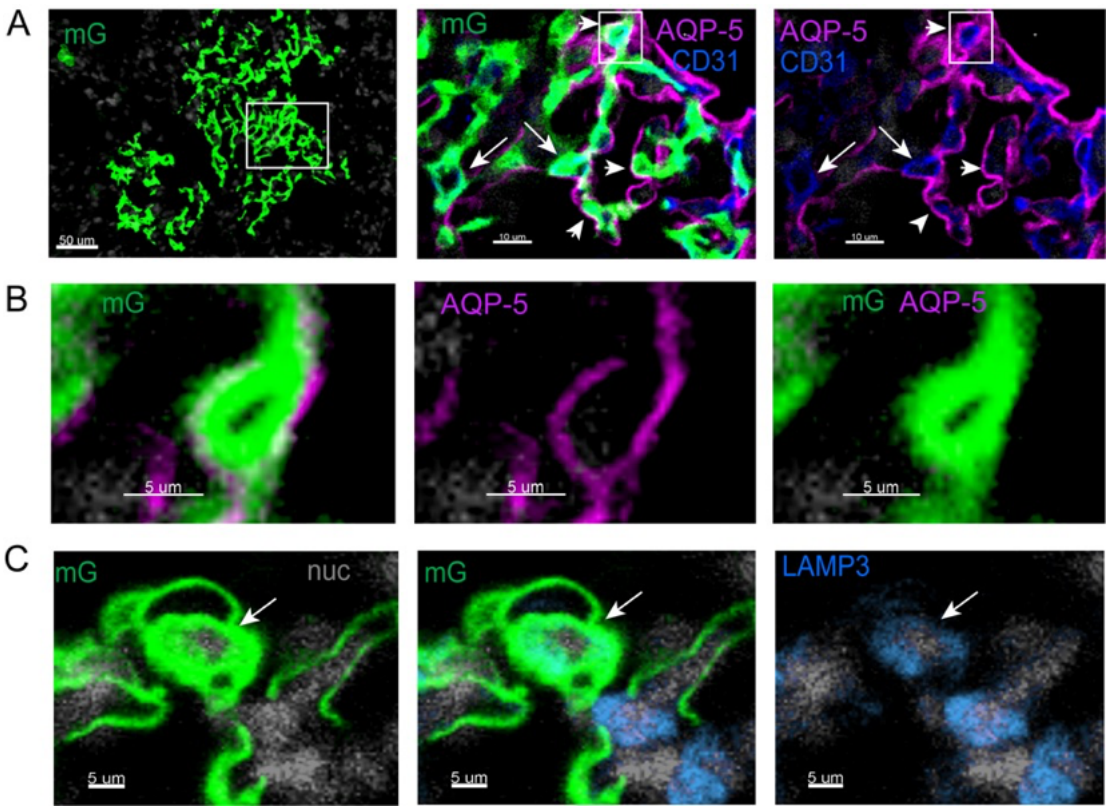

Fig. 5 Staining of GFP positive lung patches derived from transplanted VEcad Cre MTMG for epithelial and endothelial lineage differentiation markers.

(A) Left image shows GFP positive engrafted area of the lung;

scale bar=50 um. Middle and right images demonstrate close ups of the boxed area stained with anti-GFP (green), anti-AQP-5 (magenta) and anti CD31

(blue) antibodies; scale bar=10 um. (B) Close up of GFP+ AT1 donor-derived cells stained for AQP1(magenta); scale bar=5 um. (C) Donor-derived GFP+ AT2 cell stained for Lamp3 (cyan); scale bar=5 um. The images are representative of n=2 mice.

Single cell RNA seq analysis of sorted CD326+CD31+ lung cells reveal unique clusters expressing epithelial and endothelial canonical genes

While the sorting experiments clearly suggest that formation of the patches, which include high levels of donor-derived epithelial and endothelial cells, originate from the double positive CD326+CD31+ cell fraction, the FACS data shown in Fig. 3A and Fig. 3C indicated a wide spectrum of CD326 positivity ranging from dim to bright cells.

To further analyze the cell heterogeneity of the entire sorted CD45-CD31+CD326 lung cell population used in our transplantation assay, and to further interrogate potential expression of different canonical epithelial and endothelial genes within this broad CD326+CD31+ double positive population, we used single cell RNA-Seq analysis (Fig. 6 and Supplementary Fig. S11). As shown in Fig. 6, after sorting of the CD45-CD31+CD326+ progenitor cell population, this analysis revealed 12 clusters, out of which only clusters 3 and 9 represent a dual phenotype as demonstrated by feature plots for Epcam, Pecam-1, Cdh1, Ager, Cdh5, NKX2.1, and Podoplanin. In contrast, sorted single positive CD45-CD326-CD31+ cells do not exhibit substantial clusters expressing Cdh1, Ager, or NKx2.1 while all clusters express Cdh5. Taken together, clusters 3 and 9 which are spatially near each other represent about 9% of the CD45-CD326+CD31+ cells, and further confirm the presence of a lung cell subpopulation with dual expression of epithelial and endothelial genes within the sorted CD45-CD326+CD31+ lung cell subpopulation. Notably, although the CD45-CD326+CD31+ subpopulation was sorted based on broad range of staining staining for CD326, the majority of clusters, representing 91% of the cells, did not exhibit expression of the canonical epithelial genes, while they are highly positive for the endothelial ones.

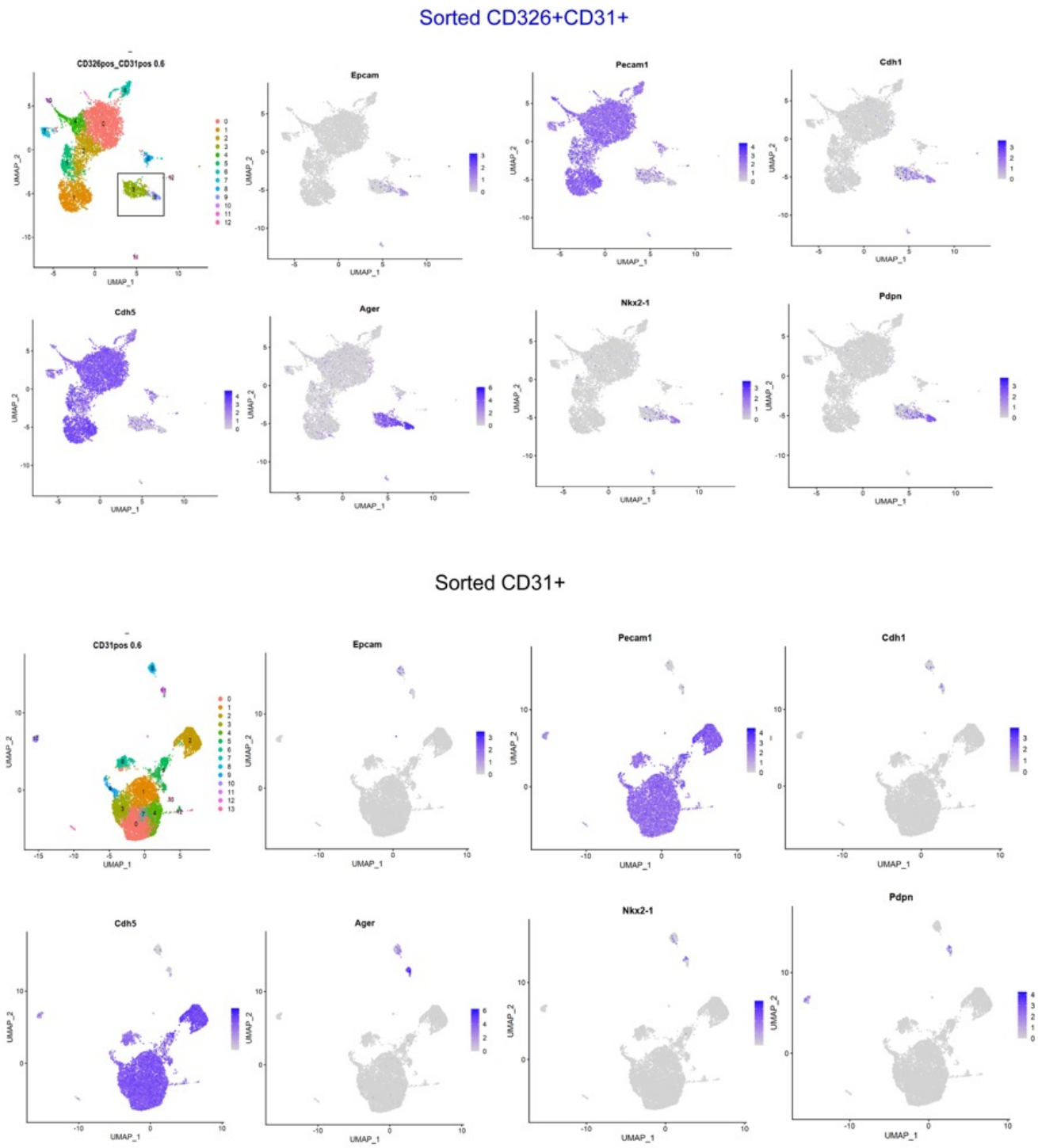

Fig. 6. Single cell RNA transcriptome analysis of FACs purified CD326+ CD31+ and CD326- CD31+ cells. Total of n=7168 CD326+ CD31+, and n=9431 CD326- CD31+ cells were analysed. Two cell populations were analyzed using 10x genomics plarform. Numbers of CD326+CD31+ and CD31+ cells were analyzed for QC

parameters and percentage of clusters defined within FACS-purified CD326+CD31+ cells (Supplementary Fig. S11). Clusters 3 and 9 represent about 9% of total cells analysed.

Identification by immunohistology of double positive lung cells in the mouse lung expressing canonical epithelial and endothelial markers.

As CD326 staining of frozen or fixed lung tissue is not effective, and based on our scRNAseq data, suggesting that canonical epithelial gene markers such as E-Cad and NKx2.1 can be used in combination with the endothelial marker VECad to distinguish the double positive cells from single positive epithelial or endothelial cells, we analyzed lungs of transgenic mice expressing GFP in the nucleus under the VECad promoter for the presence of putative double positive progenitors. Thus , as shown in Fig. 6A, VECad (GFP+) NKx2.1+SOX17+ cells could be clearly identified , predominantly in proximity to blood vessles. The expression of epithelial and endothelai markers could also ascertained by clear identification of VECad (GFP+) NKx2.1+CD31+ cells although nuclear staining for Sox17 is generally more reliable. Ror identification of these triple positive cells we pefromed high resolution confocal microscopy and traced single VECad (GFP+) NKx2.1+SOX17+ cells along the z-axis, and verified that indeed the VE-Cad (nGFP), NKx2.1 and SOX17 markers are expressed in the same cell (Fig. 7B and Supplementary Fig. S12).

Enumerating these triple positive cells in 10 fields revealed a frequency of about 0.04 % of these triple positive cells out of all counted lung cells. This low frequency is in line with the incidence of clusters 3 and 9 found by the scRNAseq analysis of FACS sorted cells. As shown in Fig. 2C, FACS purification included initial gating of

single cells (89.9%) out of which 13.9 % are included in the gate of CD45- TER119- cells; from this gate, the frequency of CD326+CD31+ cells is 4.62%. Thus, the estimated frequency of this subpopulation in the entire lung is around 0.57%, and therefore the frequency of cells of cluster 3 and 9 (9% of the sorted CD326+CD31+ population) is about 0.051%, similar to that found for the triple positive VECad (GFP+) NKx2.1+CD31+ cells in the entire lung by immunohistology. Thus both assays reveal the presence in the lung of a unique cell expressing both epiitheelial and endothelial canonincal genes, confirming that the identification by FACS anaylis of lung cells with dual expression of epithelial and endothelial genes is not merely an artifact. Further sorting and transplantation studies using lung cells of different transgenic mice, to dissect the broad range of CD326 positivity in the. CD326+CD31+ cell population and further characterize a more narrow patch forming subpopulation, are warranted.

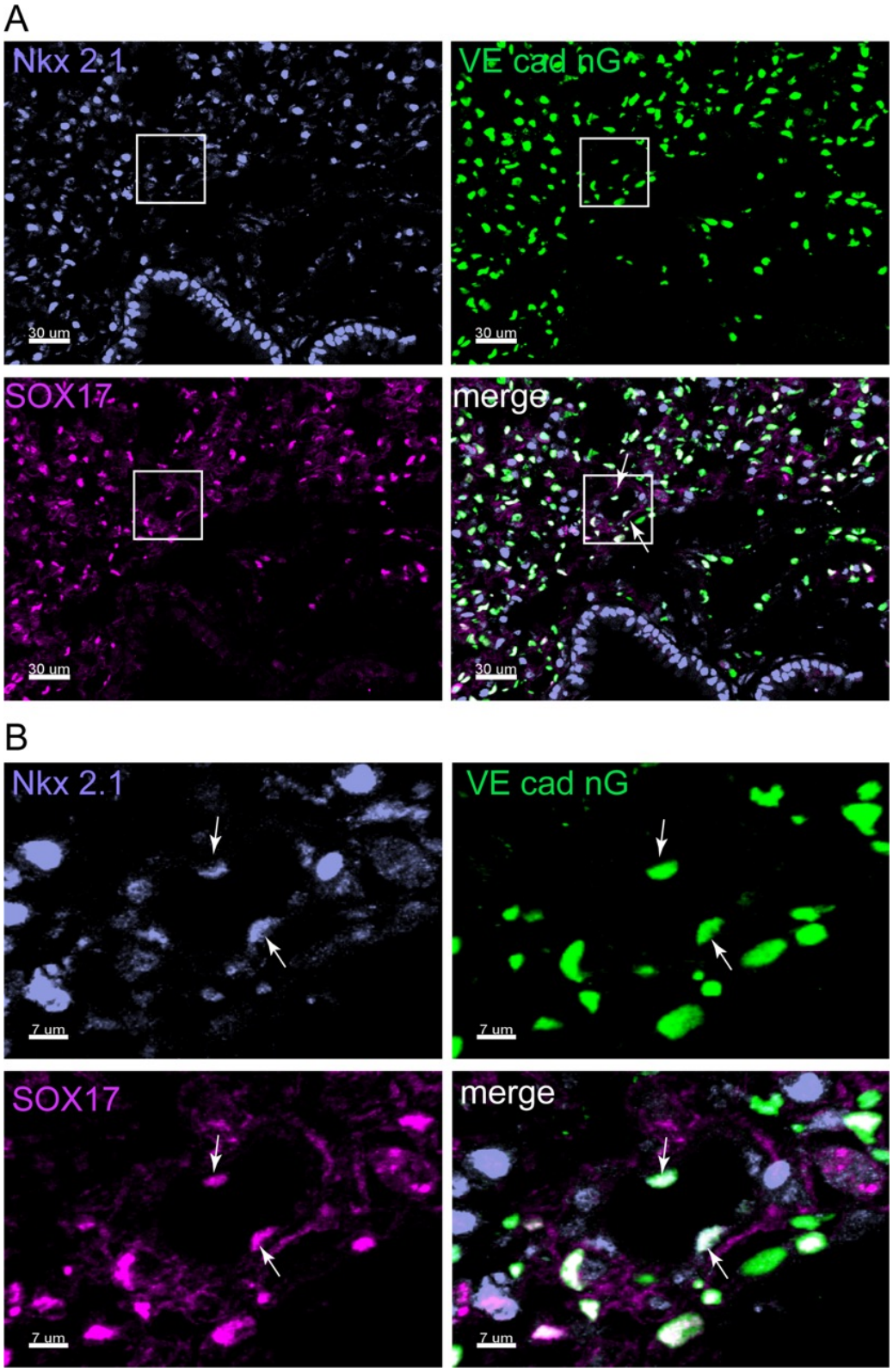

Fig. 7 (A) Typical Immunohistochemistry of a lung from transgenic VEcad nTnG mouse. (A) VEcad expression is depicted by GFP (green). Expression of Nkx2.1 and SOX17 is shown with anti- Nkx2.1 (violet), and anti- SOX17 (magenta)

antibody. Triple positive cells are indicated by arrows. Scale bar=30 um. In (B) is shown the close up of the boxed area in (A).

Discussion

We recently demonstrated that 6-8 weeks after transplantation of mouse or human fetal lung cells, harvested at the canalicular stage of gestation, into mice preconditioned with naphthalene and 6Gy TBI, numerous donor-derived patches comprising epithelial and endothelial cells could be detected throughout the lung (Rosen et al., 2015). These observations were recently confirmed by Shiraishi et al following injury with elastase and 6Gy TBI (Shiraishi et al., 2019). Similar findings were subsequently demonstrated following transplantation of adult lung cells, though the concentration of patch forming cells was 3-4 fold lower compared to that in the fetal lung (Milman Krentsis et al., 2018).

Considering that the lung patches are observed within confined borders similarly to spleen colonies which are known to be derived from a single multi-potent hemtopoietic progenitor, we hypothesized that these lung patches might represent a clonal derivation from a yet unknow multi-potent lung progenitor. Here, we show using *R26R-Confetti* donors, in which each fluorescent cell is labelled in one of four potential colors, by virtue of Cre-recombination, that at 8 weeks after transplantation all observed patches are monochromatic. This finding provides proof that each donor-derived lung patch found after transplantation of fetal or adult mouse lung cells is formed by colonization and differentiation of a single multi-potent lung progenitor. Furthermore, this finding enabled us to characterize the putative multi-potent lung progenitor by FACS purification. Thus, using our transplantation assay we have now found two distinct patch forming progenitors within the non-hematopoietic CD45-

lung compartment, namely, CD326+ CD31+ and CD326- CD31+ cells. However, the single positive CD31+ cells are predominantly associated with the formation of endothelial cells, while the double positive population is associated with numerous patches comprising both epithelial and endothelial cells.

Furthermore, we found that transplantation of lung cells from VEcad-Cre-mTmG and VEcad-Cre-nTnG transgenic donors leads to green patches comprising epithelial and endothelial cells expressing GFP, likely originating from a lung progenitor expressing a canonical endothelial marker such as VEcad, and capable of differentiating along both epithelial and endothelial lineages. It could be argued that prior to transplantation, GFP staining might leak in the VEcad-GFP donor mice into a small number of epithelial CD326+CD31- cells. However considering that our sorting experiments rule out patch forming activity by the single positive CD45-CD326+CD31- cells or the ability to generate CD326+ progeny from the single positive CD45-CD326-CD31+ cells, the results of these transplantation experiments are most consistent with the presence of a multi-potent lung progenitor within the double positive CD326+CD31+ population.

Furthermore, ex-vivo culture of FACS sorted CD326+CD31+ cells, in contrast to CD326+CD31- cells and CD326-CD31+, led to organoids containing epithelial cells, which also exhibited marked vasculature, further strengthening the hypothesis that CD326+CD31+ progenitor cells are capable of differentiating into both epithelial and endothelial cells.

A classic example of such cellular plasticity was demonstrated by Tata et al. (Tata, Purushothama Rao et al., 2013, Tata, P. R., Rajagopal, 2017), showing de-differentiation of fully mature secretory cells into basal stem cells with regenerative capacity. Another example of cellular plasticity, known as trans-differentiation or

trans-determination, occurs when stem cells from one region of the lung can convert into stem cells associated with the other lung region (Tata, Rajagopal, 2017). Moreover, evidence of multi-potential progenitors capable of developing into endothelial and epithelial linages, has also been described for human breast progenitors, which were found to express a significant level of 'Yamanaka' transcription factors (Pan et al., 2016, Roy et al., 2013).

While our sorting experiments strongly suggest that the multipotent patch forming lung progenitors are found within the CD45-CD326+CD31+ cell population, our FACS data also demonstrate marked heterogeneity of CD326 expression in this lung cell subpopulation. Notably, single cell RNAseq analysis of sorted CD45-CD326+CD31+ cells revealed two unique clusters, spatially near each other, exhibiting expression of canonical epithelial and endothelial genes. These cells represent about 9% of the entire FACS purified CD45-CD326+CD31+ cells, which in turn represents an average of 4.62±2 of the CD45- cells (in 17 experiments) in the mouse lung. Thus, the frequency of this subpopulation represents about 0.051% of the entire lung cell population. A similarly low level of lung cells expressing VEcad, Sox17 and NKx2.1 could also be detected by immunohistology of the entire mouse lung of transgenic VEcad-Cre-nTnG.

Taken together, all three assays used to characterize the cell fraction leading to multipotent patches upon transplantation, support the presence in this heterogenic lung cell fraction of a unique cell expressing canonical genes characteristic of both epithelial and endothelial lineages. It could be argued that the sorted CD45-CD326+CD31+ cell fraction might represent CD326+CD31- epithelial cells that somehow attach to CD326+CD31- cells and form cell doublets, and that upon transplantation of such doublets the formed donor derived patches include epithelail

cells originating from CD326+ CD31- progenitors and the endothelial cells originate from CD326-CD31+ progenitors. However this potential interpretation is negated by 2 major findings: 1) The patches derived from transplantation of cells from *R26R-Confetti* donors lead to monochromatic patches. 2) Transplantation of sorted Td-Tomato CD326+CD31- cells did not lead to any donor derived Td-Tomato patches in our assay even if combined with green unseparated lung cells (including CD326-CD31+), while this combined cell fraction leads to the formation of green patches.

Studies using transgenic mice to identify additional markers of the putative CD45-CD326+CD31+ lung progenitor are warranted. For example, it is possible to generate a genetic lineage-tracing system using dual recombinases (Cre-FLP) ((Liu et al., 2019), to specifically track and sort the putative patch forming lung progenitor population using markers identified by scRNAseq in clusters 3 and 9 (e.g. Cdh5 and NKx2.1)

In summary, based on the use of *R26R-Confetti* donors, we have established that donor-derived patches found after transplantation of lung cells into lung injured recipient mice, originate clonally from a single progenitor. In addition, FACS sorting data show that these progenitors are included within the CD45-CD326+ CD31+ lung cell sub-population, which is distinct from CD45-CD326-CD31-, CD45-CD326-CD31+ or CD45-CD326+CD31- cells, in its capacity to form patches comprising both epithelial and endothelial cells. This pluri-potent capacity of the patch forming cells is in striking contrast to previously described lung progenitors that are restricted to differentiation along epithelial lineages (Zuo et al., 2014), (Vaughan et al., 2014). Such pluripotency could be of particular value for lung injury repair, considering that

all the major lung diseases exhibit not only epithelial injuries, but are also associated with endothelial damage.

Furthermore, apart from its potential for translational studies aiming at the correction of different lung diseases, the identification of novel lung progenitors could also contribute to basic studies aimed at better understanding of fetal lung development, as well as understanding of steady state maintenance of different cellular lineages in the adult lung.

Acknowledgments: We thank Dr. Beata Toth and Dr. Shalev Itzkovitz for assistance with setting up the Confetti system and for helpful discussions, and Dr. Shelley Schwarzbaum for critical reading of the manuscript.

The imaging platform used for this work was supported by the De-Picciotto-Lesser Cell Observatory and the Elsie and Marvin Dekelboum Family Foundation.

Funding: This study was supported in part by a grant from the Flight Attendant Medical Research Institute (FAMRI), The Rambam Health Care Campus-Ernest and Bonnie Beutler Research Award for excellence in genomic medicine, by a research grant from Roberto and Renata Ruhman, and by Cancer Prevention and Research Institute of Texas, Grant/Award Number: CPRIT RR170008

Author contributions:

CR designed, performed and organized most of the experiments; analyzed and interpreted the data; and co-wrote the manuscript.

ES analyzed and interpreted the data and participated in discussions.

IMK, RO, XS, RY, MM assisted in performing experiments and analyzed the data.

MS, AB, ZS, SEF participated in microscopy data acquisition, analysis and discussions.

YQ, JW assisted with computational analysis of single cell RNA seq data

YR designed, coordinated and conducted the study, including analysis and interpretation of data, and co-wrote the manuscript

References

Baggiolini, A., Varum, S., Mateos, J.M., Bettosini, D., John, N., Bonalli, M., Ziegler, U., Dimou, L., Clevers, H. & Furrer, R. 2015, "Premigratory and migratory neural crest cells are multipotent in vivo", *Cell stem cell,* vol. 16, no. 3, pp. 314-322.

Barkauskas, C.E., Cronce, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W. & Hogan, B.L. 2013, "Type 2 alveolar cells are stem cells in adult lung", *The Journal of clinical investigation,* , pp. 0.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. 2018, "Integrating single-cell transcriptomic data across different conditions, technologies, and species", *Nature biotechnology,* vol. 36, no. 5, pp. 411-420.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L. & Steemers, F.J. 2019, "The single-cell transcriptional landscape of mammalian organogenesis", *Nature,* vol. 566, no. 7745, pp. 496-502.

Chen, M.J., Yokomizo, T., Zeigler, B.M., Dzierzak, E. & Speck, N.A. 2009, "Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter", *Nature,* vol. 457, no. 7231, pp. 887-891.

Chung, M. & Hogan, B.L. 2018, "Ager-CreER T2: a new genetic tool for studying lung alveolar development, homeostasis, and repair", *American journal of respiratory cell and molecular biology,* vol. 59, no. 6, pp. 706-712.

Desai, T.J., Brownfield, D.G. & Krasnow, M.A. 2014, "Alveolar progenitor and stem cells in lung development, renewal and cancer", *Nature,* vol. 507, no. 7491, pp. 190-194.

Ganuza, M., Hall, T., Finkelstein, D., Chabot, A., Kang, G. & McKinney-Freeman, S. 2017, "Lifelong haematopoiesis is established by hundreds of precursors throughout mammalian ontogeny", *Nature cell biology,* vol. 19, no. 10, pp. 1153.

Hänzelmann, S., Castelo, R. & Guinney, J. 2013, "GSVA: gene set variation analysis for microarray and RNA-seq data", *BMC bioinformatics,* vol. 14, no. 1, pp. 1-15.

Harfe, B.D., Scherz, P.J., Nissim, S., Tian, H., McMahon, A.P. & Tabin, C.J. 2004, "Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities", *Cell,* vol. 118, no. 4, pp. 517-528.

Hillel-Karniel, C., Rosen, C., Milman-Krentsis, I., Orgad, R., Bachar-Lustig, E., Shezen, E. & Reisner, Y. 2020, "Multi-lineage Lung Regeneration by Stem Cell Transplantation across Major Genetic Barriers", *Cell Reports,* vol. 30, no. 3, pp. 807-819. e4.

Jamieson, P.R., Dekkers, J.F., Rios, A.C., Fu, N.Y., Lindeman, G.J. & Visvader, J.E. 2017, "Derivation of a robust mouse mammary organoid system for studying tissue dynamics", *Development (Cambridge, England),* vol. 144, no. 6, pp. 1065-1071.

Lazzeri, E., Angelotti, M.L., Peired, A., Conte, C., Marschner, J.A., Maggi, L., Mazzinghi, B., Lombardi, D., Melica, M.E. & Nardi, S. 2018, "Endocycle-related tubular cell hypertrophy and progenitor proliferation recover renal function after acute kidney injury", *Nature communications,* vol. 9, no. 1, pp. 1344.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P. & Tamayo, P. 2015, "The molecular signatures database hallmark gene set collection", *Cell systems,* vol. 1, no. 6, pp. 417-425.

Liu, Q., Liu, K., Cui, G., Huang, X., Yao, S., Guo, W., Qin, Z., Li, Y., Yang, R. & Pu, W. 2019, "Lung regeneration by multipotent stem cells residing at the bronchioalveolar-duct junction", *Nature genetics,* vol. 51, no. 4, pp. 728-738.

Lyubimova, A., Itzkovitz, S., Junker, J.P., Fan, Z.P., Wu, X. & Van Oudenaarden, A. 2013, "Single-molecule mRNA detection and counting in mammalian tissue", *Nature protocols,* vol. 8, no. 9, pp. 1743.

Ma, Q., Ma, Y., Dai, X., Ren, T., Fu, Y., Liu, W., Han, Y., Wu, Y., Cheng, Y. & Zhang, T. 2018, "Regeneration of functional alveoli by adult human SOX9 airway basal cell transplantation", *Protein & cell,* vol. 9, no. 3, pp. 267-282.

McInnes, L., Healy, J. & Melville, J. 2018, "Umap: Uniform manifold approximation and projection for dimension reduction", *arXiv preprint arXiv:1802.03426,* .

Milman Krentsis, I., Rosen, C., Shezen, E., Aronovich, A., Nathanson, B., Bachar-Lustig, E., Berkman, N., Assayag, M., Shakhar, G. & Feferman, T. 2018, "Lung Injury Repair by Transplantation of Adult Lung Cells Following Preconditioning of Recipient Mice", *Stem cells translational medicine,* vol. 7, no. 1, pp. 68-77.

Muzumdar, M.D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. 2007, "A global double-fluorescent Cre reporter mouse", *genesis,* vol. 45, no. 9, pp. 593-605.

Nichane, M., Javed, A., Sivakamasundari, V., Ganesan, M., Ang, L.T., Kraus, P., Lufkin, T., Loh, K.M. & Lim, B. 2017, "Isolation and 3D expansion of multipotent Sox9 mouse lung progenitors", *Nature methods,* vol. 14, no. 12, pp. 1205.

Pan, D., Roy, S., Gascard, P., Zhao, J., Chen-Tanyolac, C. & Tlsty, T.D. 2016, "SOX2, OCT3/4 and NANOG expression and cellular plasticity in rare human somatic cells requires CD73", *Cellular signalling,* vol. 28, no. 12, pp. 1923-1932.

Prigge, J.R., Wiley, J.A., Talago, E.A., Young, E.M., Johns, L.L., Kundert, J.A., Sonsteng, K.M., Halford, W.P., Capecchi, M.R. & Schmidt, E.E. 2013, "Nuclear double-fluorescent reporter for in vivo and ex vivo analyses of biological transitions in mouse nuclei", *Mammalian genome,* vol. 24, no. 9, pp. 389-399.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A. & Trapnell, C. 2017, "Reversed graph embedding resolves complex single-cell trajectories", *Nature methods,* vol. 14, no. 10, pp. 979.

Rios, A.C., Fu, N.Y., Lindeman, G.J. & Visvader, J.E. 2014, "In situ identification of bipotent stem cells in the mammary gland", *Nature,* vol. 506, no. 7488, pp. 322.

Rosen, C., Shezen, E., Aronovich, A., Klionsky, Y.Z., Yaakov, Y., Assayag, M., Biton, I.E., Tal, O., Shakhar, G. & Ben-Hur, H. 2015, "Preconditioning allows engraftment of mouse and human embryonic lung cells, enabling lung repair in mice", *Nature medicine,* vol. 21, no. 8, pp. 869.

Roy, S., Gascard, P., Dumont, N., Zhao, J., Pan, D., Petrie, S., Margeta, M. & Tlsty, T.D. 2013, "Rare somatic cells from human breast tissue exhibit extensive lineage plasticity", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 110, no. 12, pp. 4598-4603.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F. & Regev, A. 2015, "Spatial reconstruction of single-cell gene expression data", *Nature biotechnology,* vol. 33, no. 5, pp. 495-502.

Scheele, C.L., Hannezo, E., Muraro, M.J., Zomer, A., Langedijk, N.S., Van Oudenaarden, A., Simons, B.D. & Van Rheenen, J. 2017, "Identity and dynamics of mammary stem cells during branching morphogenesis", *Nature,* vol. 542, no. 7641, pp. 313.

Schepers, A.G., Snippert, H.J., Stange, D.E., van den Born, M., van Es, J.H., van de Wetering, M. & Clevers, H. 2012, "Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas", *Science (New York, N.Y.),* vol. 337, no. 6095, pp. 730-735.

Shiraishi, K., Shichino, S., Tsukui, T., Hashimoto, S., Ueha, S. & Matsushima, K. 2019, "Engraftment and proliferation potential of embryonic lung tissue cells in irradiated mice with emphysema", *Scientific reports,* vol. 9, no. 1, pp. 3657.

Snippert, H.J., Van Der Flier, Laurens G, Sato, T., Van Es, J.H., Van Den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A.M., Van Rheenen, J. & Simons, B.D. 2010, "Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells", *Cell,* vol. 143, no. 1, pp. 134-144.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck III, W.M., Hao, Y., Stoeckius, M., Smibert, P. & Satija, R. 2019, "Comprehensive integration of single-cell data", *Cell,* vol. 177, no. 7, pp. 1888-1902. e21.

Takeda, N., Jain, R., LeBoeuf, M.R., Wang, Q., Lu, M.M. & Epstein, J.A. 2011, "Interconversion between intestinal stem cell populations in distinct niches", *Science (New York, N.Y.),* vol. 334, no. 6061, pp. 1420-1424.

Taniguchi, H., Lu, J. & Huang, Z.J. 2013, "The spatial and temporal origin of chandelier cells in mouse neocortex", *Science (New York, N.Y.),* vol. 339, no. 6115, pp. 70-74.

Tata, P.R. & Rajagopal, J. 2017, "Plasticity in the lung: making and breaking cell identity", *Development (Cambridge, England),* vol. 144, no. 5, pp. 755-766.

Tata, P.R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B.M., Vinarsky, V., Cho, J.L., Breton, S. & Sahay, A. 2013, "Dedifferentiation of committed epithelial cells into stem cells in vivo", *Nature,* vol. 503, no. 7475, pp. 218.

Traag, V.A., Waltman, L. & Van Eck, N.J. 2019, "From Louvain to Leiden: guaranteeing well-connected communities", *Scientific reports,* vol. 9, no. 1, pp. 1-12.

Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S. & Rinn, J.L. 2014, "The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells", *Nature biotechnology,* vol. 32, no. 4, pp. 381.

Vaughan, A.E., Brumwell, A.N., Xi, Y., Gotts, J.E., Brownfield, D.G., Treutlein, B., Tan, K., Tan, V., Liu, F.C. & Looney, M.R. 2014, "Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury", *Nature,* .

Weiss, D.J. & Ortiz, L.A. 2013, "Cell therapy trials for lung diseases: progress and cautions", *American journal of respiratory and critical care medicine,* vol. 188, no. 2, pp. 123-125.

Zuo, W., Zhang, T., Wu, D.Z., Guan, S.P., Liew, A., Yamamoto, Y., Wang, X., Lim, S.J., Vincent, M. & Lessard, M. 2014, "p63 Krt5 distal airway stem cells are essential for lung regeneration", *Nature,* .