bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1 **TITLE**

2 **SCRINSHOT, a spatial method for single-cell resolution mapping of cell states in tissue**
3 **sections**

4

5 **Authors**

6 Alexandros Sountoulidis[1,2 ‡], Andreas Liontos[1,2], Hong Phuong Nguyen[1,2], Alexandra B. Firsova[1,2],
7 Athanasios Fysikopoulos[3], Xiaoyan Qian[1,4], Werner Seeger[3], Erik Sundström[5], Mats Nilsson[1,4]
8 and Christos Samakovlis[1,2,3 ‡]

9 [1]Science for Life Laboratory, Solna, Sweden

10 [2] Department of Molecular Biosciences, Wenner-Gren Institute, Stockholm University, Stockholm,
11 Sweden

12 [3] Molecular Pneumology, Cardiopulmonary Institute, Justus Liebig University, Giessen, Germany

13 [4] Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

14 [5] Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm,
15 Sweden

16

17 [‡] Corresponding authors: Christos Samakovlis (christos.samakovlis@scilifelab.se), Alexandros
18 Sountoulidis (alexandros.sountoulidis@scilifelab.se)

19

20 Keywords: SCRINSHOT, padlock-probe, multiplex RNA FISH, spatial cell-type map, lung

21

22

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Changes in cell identities and positions underlie tissue development and disease progression. Although, single-cell mRNA sequencing (scRNA-Seq) methods rapidly generate extensive lists of cell-states, spatially resolved single-cell mapping presents a challenging task. We developed SCRINSHOT (Single Cell Resolution IN Situ Hybridization On Tissues), a sensitive, multiplex RNA mapping approach. Direct hybridization of padlock probes on mRNA is followed by circularization with SplintR ligase and rolling circle amplification (RCA) of the hybridized padlock probes. Sequential detection of RCA-products using fluorophore-labeled oligonucleotides profiles thousands of cells in tissue sections. We evaluated SCRINSHOT specificity and sensitivity on murine and human organs. SCRINSHOT quantification of marker gene expression shows high correlation with published scRNA-Seq data over a broad range of gene expression levels. We demonstrate the utility of SCRISHOT by mapping the locations of abundant and rare cell types along the murine airways. The amenability, multiplexity and quantitative qualities of SCRINSHOT facilitate single cell mRNA profiling of cell-state alterations in tissues under a variety of native and experimental conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

INTRODUCTION

Recent advances in single-cell RNA sequencing technologies (scRNA-Seq) enabled transcriptome analysis of individual cells and the identification of new cellular states in healthy and diseased conditions (1). These methods however fail to capture the spatial cellular organization in tissues due to cell dissociation. New spatial transcriptomic methods aim to circumvent the problem of lost cellular topology (2). They can be divided into two categories: First, targeted methods that directly detect specific mRNAs with single-cell resolution like ISS (3), MERFISH (4), osmFISH (5) and second global methods, which are based on barcode annotated positions and next generation sequencing to resolve RNA topology. The spatial resolution of global methods is still larger than the typical cellular dimensions (6, 7).

Targeted methodologies are based on nucleic acid probes (mainly DNA), complementary to the RNA species of interest, as in all *in situ* hybridization assays (8). Single-molecule fluorescence *in situ* hybridization (smFISH) is the most powerful among the spatial transcriptomic methods and has been used to supplement scRNA-Seq data with spatial information. It utilizes multiple fluorophore-labeled probes, which recognize the same RNA molecule along its length and visualize single RNA molecules as bright fluorescent dots (9, 10). Nonetheless, this method still retains some limitations such as low signal-to-noise ratio, reduced sensitivity on short transcripts, false positive signal due to unspecific-binding of the labeled probes and low capacity for multiplex detection of many RNA molecules (11-13). Multiplex detection with smFISH was initially addressed by the sequential fluorescence *in situ* hybridization (seqFISH) (14, 15) and the multiplexed error-robust FISH (MERFISH) (4). These approaches utilize sequential rounds of hybridization of FISH probes or barcode-based primary probes to detect multiple RNA species. The outstanding throughput of these methods makes them strong candidates for generation of spatial transcriptome maps in tissues. However, since the principle of these techniques is similar to smFISH, they require large number of gene-specific probes, confocal or super-resolution

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

71 microscopy to deconvolve the signals and complicated algorithms for both probe design and
72 analysis. Nevertheless, the low signal-to-noise ratio still remains a major technical challenge of
73 these methods, especially for tissue sections with strong auto-fluorescence from structural
74 extracellular matrix components like collagen and elastin (16). New strategies for signal
75 amplifications, such as branched-DNA amplification (RNAScope) (17) and hybridization chain
76 reaction (18, 19), have been recently combined with sophisticated probe design (AmpFISH) (20)
77 to increase sensitivity and specificity of smFISH.

78 Padlock probes have been successfully used to detect RNA species (21). They are linear DNA
79 molecules, with complementary arms to the target mRNA sequence and a common “backbone”.
80 Upon hybridization with the target sequence, they can be ligated, creating circular single-stranded
81 DNA molecules, which are used as templates for signal amplification using Φ29 polymerase-
82 mediated rolling circle amplification (RCA) (22). RCA products are large single stranded DNA
83 molecules containing hundreds of copies (23) of the complementary padlock-probe sequence.
84 They can be detected with fluorophore labelled oligos, which recognize either their RNA-specific
85 sequence or their backbone. Because each RCA-product contains hundreds of repeats of the
86 same detected sequence, the signal-to-noise ratio increases significantly, facilitating signal
87 detection by conventional epifluorescence microscopy. Also, multiplexity has been integrated into
88 the method by sequencing by ligation (24, 25). Since commercial ligases such as T4 DNA ligase,
89 T4 RNA ligase 2 and Ampligase show low activity on DNA/RNA hybrids, RNA has to be reverse-
90 transcribed to cDNA fragments before introducing padlock probes (3). However, cDNA synthesis
91 on fixed tissue sections is a challenging and expensive procedure (26), prompting new elegant
92 approaches trying to circumvent reverse-transcription by introducing Click-chemistry to ligate
93 DNA probes after hybridization on their RNA targets (ClampFISH (27)). More recently, “H-type”
94 DNA probes, which are hybridized to both RNA and padlock probes (PLISH (28)), or SNAIL-

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

design probes have been successfully used to facilitate intramolecular ligation of the padlock probes (STARmap (25)).

PBVC-1 DNA ligase (also known as SplintR ligase) shows strong ligase activity of DNA sequences in DNA/RNA hybrids. This enzyme is a DNA ligase encoded by *Paramecium bursaria Chlorella* virus 1, first discovered by Ho et al.,(29). A number of studies have successfully applied SplintR ligase for RNA species detection in cultured cells (26, 30-32). However, the fidelity of SplintR ligation in end-joining is questionable because it tolerates mismatches at the padlock probe junction and its usefulness is debated. In addition, SplintR ligase shows 1% of its ligation activity on a 1-nucleotide gap of a nicked duplex DNA substrate (26, 33).

We present SCRINSHOT an optimized protocol for multiplex RNA *in situ* detection on paraformaldehyde-fixed (PFA) tissue sections. We validated the sensitivity, specificity and multiplexity of SCRINSHOT and showed that it is quantitative over a broad range of gene expression levels. It is based on the *in situ* sequencing protocol (3) but bypasses the costly and inefficient reverse transcription on fixed tissue to gain higher detection efficiency, by utilizing SplintR ligase. To minimize false positive artifacts, we utilized 40nt long target-specific sequences in the padlock probes in combination with stringent hybridization conditions. SCRINSHOT performs on a variety of tissues including lung, kidney and heart and readily detects several epithelial, endothelial and mesenchymal cells. We tested the multiplexing of SCRINSHOT on mouse and human tissue sections, by simultaneous detection of characterized cell-type-selective markers. SCRINSHOT successfully identified distinct cell-types and helped to create spatial maps of large tissue areas at single cell resolution.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

RESULTS AND DISCUSSION

SCRINSHOT overview

SCRINSHOT evolved from our attempts to improve the detection sensitivity and reduce the cost of the *in situ* sequencing method, using PFA-fixed material. PFA fixation significantly improves the histology but makes RNA less accessible to enzymes and padlock probes (24, 26). We focused on stringent padlock probe hybridization to RNA targets and omitted the inefficient *in situ* cDNA synthesis step (Figure 1 and Supplementary Figure 1). Gene-specific padlock probes are first directly hybridized to 40nt long, unique sequences on the mRNA targets. Upon ligation of bound padlock probes, we amplify the padlock probe sequence using Φ29 polymerase dependent rolling cycle amplification. RCA products are subsequently detected by fluorophore-labelled oligos which recognize the gene specific part of the padlock, as previously described (3). Multiplexity is reduced compared to sequencing by ligation (24), which theoretically allows the detection of 256 transcript species in 4 hybridization runs, or to other barcode based approaches (4, 26). To increase the number of detected genes we used sequential hybridization cycles of fluorophore-labelled, uracil-containing oligos. After each detection cycle, fluorescent probes are removed by enzymatic fragmentation by uracil-N-glycosylase (UNG) and stringent washes. Sequential detection also enables the separate detection of low, medium and high abundant RNA species, increasing the dynamic range of detection (34). Our detection probes were each labelled by one of three commonly-used fluorophores and allowed up to 10 hybridization and imaging cycles, typically detecting 30 genes. After image acquisition, we utilized manual segmentation of nuclei and open access analysis tools for image stitching and signal quantification to construct a simple pipeline for quantitative mapping of expression counts for 20-30 genes to thousands of cells on tissue sections. The detailed protocol starting with tissue fixation and leading to mapping is presented in the “Additional File 1”.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

SCRINSHOT specificity depends on stringent hybridization of the padlock probe

The specificity of SCRINSHOT crucially relies on the specific targeting of the SplintR ligation activity to the correct sites of the interrogated mRNA. A recent study (33) reported that the fidelity of this ligase is poor as it tolerates mismatches at padlock probe junctions. In addition, SplintR shows 1% of its ligation activity on a 1-nucleotide gap of a nicked duplex DNA substrate and it is unable to join ends across a 2-nucleotide gap (35). These results question the specificity of SplintR ligase-based methods for *in situ* RNA detection. We reasoned that the choice of padlocks with high melting points (Tm around 70^{o}C) followed by stringent washes after DNA/RNA hybridization would circumvent SplintR promiscuity. We tested the dependence of SCRINSHOT signals on the hybridization and ligation steps of the padlock, by generating mutant padlocks, predicted to affect either the hybridization of the 3´padlock arm or the sequence of the ligation site. The 3´-scrambled arm of this padlock is expected to fail in hybridizing with the *Scgb1a1* mRNA, resulting in a linear, unligated padlock and therefore in a block in RCA. The single mismatch probe contains a single replaced nucleotide at the ligation site in 5´-end (C to G). This substitution was designed to address the effect of promiscuous ligation on the signal. In the same experiment, we included a slide, where we omitted the SplintR ligase from the reaction mixture to test whether padlock hybridization alone is sufficient to generate some RCA (SplintR^{neg}; Figure 2 D, D´). In all tested conditions we also used the *Actb* normal padlock probe. This transcript was detected at similar levels in all slides and served as a control for the reactions with the mutated padlocks. We first counted the dots of *Scgb1a1* and *Actb* signals in all airway cells of sequential lung tissue sections and plotted their ratios at the different conditions (Figure 2). The signal from single-mismatch padlock probe was reduced by 20%, compared to the normal *Scgb1a1* probe demonstrating the low SplintR fidelity for the sequence of the ligation site (Figure 2 B, B´, E). The *Scgb1a1* signal was lost when we used the 3´-scrambled padlock, indicating that padlock probe hybridization is necessary for circularization and subsequent RCA (Figure 2 C, C´, E). The

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

omission of SplintR from the ligation mixture resulted in undetectable signal for both *Scgb1a1* and *Actb* indicating the central role of ligation in signal amplification (Figure 2 D, D´, E). We noticed that the *Scgb1a1* signal showed significant crowding and a potential saturation leading to underestimation of the total number of RCA-products for this highly expressed gene (see below). This crowding was evident even when we added 5-fold less *Scgb1a1* padlock probe (0.01 µM) to the reactions in comparison to all other padlock probes. In an attempt to more accurately quantify the differences between the signals from normal and mutated padlocks we measured the overall fluorescence intensity (Raw Integrated Density) of the airway cell-ROIs. This showed that the single mismatch at the ligation site of *Scgb1a1* padlock probe causes 3-fold fluorescence signal reduction arguing that the *in situ* SplintR activity is substantially reduced, but not abolished by single nucleotide substitutions at the ligation site (Figure 2F). We conclude that the specificity of SCRINSHOT assay is largely provided by the hybridization stringency of the padlock probes, since SplintR is unable to ligate off-target padlock probes, but can ligate single-mismatch probes with low efficiency.

A *Scgb1a1* antisense oligonucleotide competes with padlock probe hybridization and signal detection

If the padlock hybridization is the critical step for signal generation then competition by an oligonucleotide, which recognizes the binding to the mRNA is expected to proportionally reduce the detected RCA signal. We used the *Scgb1a1* and *Actb* padlocks together with increasing concentrations of a competing, unlabeled oligonucleotide complementary to the mRNA sequence recognized by the *Scgb1a1* padlock. Inclusion of the competitor reduced the *Scgb1a1* signal, in a dose-dependent manner. Equal molar ratios of padlock probe and competitor caused signal reduction by 10-fold and the signal was eliminated when 5-fold excess of the competitor was used (Figure 3). This suggests that the SCRINSHOT signal is proportional to the target expression

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

levels because it can be proportionally competed with increasing concentrations of a synthetic oligo masking the hybridization site. It also highlights the importance of proper padlock design to achieve similarly high Tm values and hybridizations conditions for the different probes.

Application of SCRINSHOT in other organs

To evaluate SCRINSHOT applicability to other tissues, we performed the assay using PFA-fixed sections from adult mouse kidney and heart and human embryonic lung. On murine tissues, we used a common panel of padlocks targeting validated lung cell type markers. We targeted *Actb* as a generic marker, *Pecam1* as an endothelial cell marker, *Scgb1a1* as a club cell marker, *Sftpc and Napsa* as alveolar epithelial type II (AT2) cell markers and *Lyz2*, as a marker for AT2 cells, macrophages and neutrophils (36). As expected, *Actb* was uniformly expressed in kidney and heart while *Scgb1a1* and *Sftpc* were undetectable (Supplementary Figure 2 A, B). In both tissues, *Lyz2* was expressed by a few scattered cells, which presumably correspond to macrophages (37, 38) (Supplementary Figure 2). In a subset of kidney tubular structures, we detected *Napsa* (Supplementary Figure 2 A), which agrees with the previously described immunohistochemical detection of the marker in renal proximal tubule cells (39). In the vessel walls of the heart, we detected sparse signal for the endothelial cell marker *Pecam1* but the myocardial cells were negative for the marker (Supplementary Figure 2 B). In the human embryonic lung sections, we used probes targeting transcripts encoding 3 transcription factors, *SOX2*, *SOX9* and *ASCL1* (Supplementary Figure 3), which have previously been detected by antibody staining in subsets of epithelial cells (40, 41). In agreement with the published results, the SCRINSHOT signal for *SOX2* was confined mainly in the proximal part of the branching epithelium, whereas *SOX9* was selectively expressed in the distal tips. *ASCL1* expression overlapped with *SOX2* (Supplementary Figure 3). These experiments show that SCRINSHOT can be readily applied to map cell-type heterogeneity in a variety of tissues.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

SCRINSHOT generates quantitative gene expression profiles in single cells

We first tested the quantitative power of SCRINSHOT by correlating its detection performance with the fluorescence of a transgenic red fluorescent protein (RFP) in mouse lung tissue sections. In the *Sftpc-CreER;Rosa-Ai14* reporter mouse, the RFP expression is activated in AT2 cells, upon Tamoxifen induction of the Cre recombinase (42). Cre recombines out a transcriptional/translational STOP cassette (43) of the Rosa26 locus and allows RFP protein expression and fluorescence in AT2 cells (44). We injected pups with Tamoxifen on postnatal day 1 (P1) and analyzed the lungs on day P21. In the same experiment, we also used lungs from *Sftpc-CreER*^{neg}*-Rosa-Ai14*^{pos} mice sacrificed at P21 and lungs from *wild type* mice sacrificed at P60 as controls for the RFP induction and the potential effects of tissue autofluorescence in young and fully developed lungs.

We first analyzed 14167 *Sftpc-CreER*^{pos}*-Rosa-Ai14*^{pos} cells and found a reliable correlation (R^2=0.7233) between the endogenous RFP fluorescence and the SCRINSHOT-detected *RFP* mRNA molecules in each cell (Figure 4 B). By contrast, there was no correlation between *RFP* SCRINSHOT signal and the endogenous fluorescence (R^2=0.0671) in 3355 *Sftpc-CreER*^{neg}*-Rosa-Ai14*^{pos} cells. As expected, in the 1008 analyzed *wild type* cells, the correlation of RFP fluorescence and SCRINSHOT dots was very low. Upon closer inspection most of the signal in this lung was due to high auto-fluorescence from red blood cells, illustrating the specificity of SCRINSHOT (Figure 4 A, B).

Considering that the *Sftpc-CreER;Rosa-Ai14* reporter specifically labels the alveolar AT2 cells (42), we used SCRINSHOT to identify the *Sftpc*^{pos} AT2 cells and analyze their RFP expression in both RNA (SCRINSHOT) and protein fluorescence (Raw Integrated Density) level (Figure 5 A, B, D). In *Sftpc-CreER*^{pos} cells the mean number of SCRINSHOT detected *RFP* mRNA molecules was 6-fold higher compared to *Sftpc-CreER*^{neg} cells. In the same comparison the RFP

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

241 endogenous fluorescence was 5-fold higher. In *wild type* cells, the SCRINSHOT *RFP* signal was
242 hardly detectable. We observed a slight increase in RFP fluorescence, in *wild type* cells compared
243 to the ones from the *Sftpc-CreER*^{neg} mouse, presumably due to different levels of tissue
244 autofluorescence depending on developmental stage and fixation time of the tissues (4 hours for
245 P21 *Sftpc-CreER* and 8 hours for *wild type* P60).

246 Apart from the AT2 cells, the lung alveolus contains additional cells types, including endothelial,
247 inflammatory, epithelial AT1 cells and fibroblasts (45). In the *Sftpc-CreER*^{pos}*-Rosa-Ai14*^{pos} lung,
248 AT2 cells should be both *RFP*^{pos} and *Sftpc*^{pos} whereas the rest should be *RFP*^{neg}. To evaluate the
249 specificity and sensitivity of SCRINSHOT in the strict context of a transgenic, knock-in marker
250 expression, we analyzed the *Sftpc* signal in *RFP*^{pos} and *RFP*^{neg} alveolar cells and the
251 SCRINSHOT *RFP* signal in *Sftpc*^{pos} and *Sftpc*^{neg} alveolar cells from the *Sftpc-CreER*^{pos}*-Rosa-*
252 *Ai14*^{pos} mouse. Both genes are highly expressed, therefore we applied a threshold of 3 dots/cell
253 to score a cell as positive (see also Supplement for details on setting detection thresholds). Of
254 the 1429 analyzed *RFP*^{pos} cells only 1.1% did not express *Sftpc* and were thus scored as false
255 negatives. Among the 6848 *RFP*^{neg} cells, 7.4% were positive for *Sftpc* and were considered as
256 false positives. Conversely, we failed to detect *RFP* signal in 4% of the 1679 analyzed Sftpc^{pos}
257 cells and counted them as false negatives. Among the Sftpc^{neg} cells 7% scored as positive for
258 SCRINSHOT *RFP* signal and were designated as false positives (Figure 5 C, D). Thus, the levels
259 of false positive and false negative cell annotations by SCRINSHOT are on average 5%. This is
260 likely an overestimate because Tamoxifen-induced recombination is rarely 100% efficient and
261 because leaky RFP transcripts from the *Rosa* locus may escape nonsense-mediated RNA decay
262 (NMD) (46). In conclusion, the analysis of the inducible knock-in reporter shows that the
263 SCRINSHOT *RFP* signal highly correlates with RFP endogenous fluorescence intensity in *Sftpc-*
264 *CreER*^{pos} AT2 cells only. The low levels of false positive and false negative SCRINSHOT signals

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

for *RFP* and *Sftpc* argue for the efficiency of SCRINSHOT in the identification and quantification of gene expression in alveolar cell-types.

Multiplex performance and gene expression quantification using SCRINSHOT

To further explore the utility of SCRINSHOT in the spatial identification of cell types, we first tested if multiple rounds of hybridization and detection in lung tissue sections lead to loss of detection signal for the genes of interest, as seen in other transcriptomic approaches (5). We compared the detection signals of *Calca*, a known neuroendocrine gene marker during the first and the eighth cycles of hybridization and detection and found no significant loss of signal or decreased specificity (Supplementary Figure 4). In the same experiment, we also detected a pair of marker genes for airway secretory cells (*Scgb1a1* and *Cyp2f2*) and a pair of neuroendocrine cell markers (*Ascl1* and *Calca*) and mapped them relatively to each other in the bronchiolar epithelium (Figure 6B). In total we analyzed 15 genes encoding soluble secreted proteins (*Scgb1a1*, *Napsa*, *Lgi3*, *Calca*, *Sftpc* and *Lyz2*), cell surface proteins and receptors (*Cd74*, *Cldn18*, *Fgfr2* and *Ager*), a metabolic enzyme *Cyp2f2*, and signaling proteins and transcription factors (*Axin2, Spry2*, and *Etv5*, *Ascl1*) (Figure 6A, B). We first focused on *Sftpc-CreER*^{pos}*-Rosa-Ai14*^{pos} AT2 cells and on the analysis of 11 selective AT2 markers covering a broad spectrum of expression levels (Figure 6 A, C). To evaluate the utility of SCRINSHOT in complementing scRNA-Seq data with spatial information we compared the SCRISNSHOT analysis with available scRNA-Seq data from 156 AT2 cells (47). The mean values of SCRINSHOT dots in AT2 cells were proportional to the scRNA-Seq raw count values of AT2 cells. Spearman correlation analysis showed a strong correlation between the results of the two methodologies (ρ=0.9455) for highly-, moderately- and lowly-expressed genes (Figure 6C). As expected, there was no correlation of the SCRINSHOT data from 5615 *Sftpc*^{neg} alveolar cells when we compared them to the AT2 scRNA-Seq dataset (ρ=0.1636). The proportional mean gene expression levels in SCRISHOT and scRNA-Seq argue

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

that SCRINSHOT provides a suitable alternative for rapid, *in situ* evaluation of cell states detected by scRNA-Seq.

Spatial mapping of tracheal cell heterogeneity using SCRINSHOT

Recently, two studies addressed the heterogeneity of tracheal epithelium using sc-RNA Seq. These studies identified a new pulmonary cell type, which expresses *Cftr* and therefore considered to play a role in cystic fibrosis pathophysiology (48, 49). They also provided the detailed transcriptomic state of additional cell types, like basal, tuft and secretory cells, including club and two classes of goblet cells.

We tested the ability of SCRINSHOT to detect the above cell types and analyze tracheal epithelial cell heterogeneity with spatial, single-cell resolution. We used a panel of selective markers for club, goblet, basal, tuft and ionocytes, as identified in (48, 49). The 29 analyzed genes included (i) *Scgb1a1, Scgb3a1, Il13ra1, Reg3g, Lgr6* and *Bpifb1* as club-cell markers, (ii) *Foxq1, Gp2, Pax9, Spdef, Tff2, Lipf, Dcpp3* and *Dcpp1* as goblet-cell markers, (iii) *Trp5, Il25, Gng13, Six1, Alox5ap* and *Sox9* as tuft-cell markers, (iv) *Foxi1, Tfcp2l1, Cftr* and *Ascl3* as ionocyte markers, (v) *Ascl1* as a neuroendocrine cell marker (50) and (vi) *Krt5, Pdpn* and *Trp63* as basal cell markers. We also included *Muc5b* as a general secretory marker of the proximal airways (48, 49). For assignment of cell positions we utilized structural landmarks that separate the tracheal airway epithelium in three parts, the proximal, which extends until the end of the submucosal gland, the intermediate part, which spans eight cartilage rings deeper and the distal, which includes the remaining part of trachea epithelium, up to bronchial bifurcation (carina). We also assigned positions to proximal intra-lobar airway epithelial cells, which extend up to the L.L3 branching point (51) and to distal airway epithelial cells located at terminal bronchioles (Figure 7A).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

For cell-type annotation, we initially applied the following threshold criteria for the selected marker genes. We considered club cells, only the *Scgb1a1*^{pos} cells, which were negative for the neuroendocrine cell marker *Ascl1* and expressed up to one goblet, tuft, ionocyte and basal-cell markers. *Ascl1*^{pos} cells were considered as neuroendocrine cells. Similarly, we annotated the analyzed cells as goblet, basal, tuft cells and ionocytes, if they were positive for at least two characteristic genes of the respective type and only expressed up to one marker of the others. The previously described similarities in gene expression between goblet and club cells in (48), prompted us to consider a cell as goblet if it was positive for at least two of the identified goblet cell markers, regardless of *Scgb1a1* expression. We sampled 1068 cells in submucosal glands and 216 proximal trachea airway epithelial cells. In the intermediate part of the tracheal tube, we quantified 953 cells and 1164 in the distal. In the intra-lobar airway epithelium, we analyzed 551 cells in proximal and 484 in distal airways. SCRINSHOT detected all the previously described trachea cell-types. Club cells comprised 7% of total cells in proximal trachea and their proportion gradually increased to 77% towards the distal intra-lobar airways (Figure 7B). *Trp63*^{pos}, *Krt5*^{pos} and *Pdpn*^{pos} basal cells were primarily detected in the intermediate part of the trachea (21% of the measured cells) and became reduced towards the intra-lobar airways (Figure 7 B, C, D). Tuft cells expressing *Trmp5*, *Gng13 and Alox* were exclusively found in the tracheal epithelium (Figure 7 B, C, E). Ionocytes present a rare cell-type (<1% of trachea epithelium) implicated in the pathogenesis of cystic fibrosis. We detected sparse ionocytes expressing *Cftr* along with the transcription factors *Tfp2l1*, *Ascl3* and *Foxi1* (48, 49) in the tracheal airway epithelium (Figure 7 B, F) and even more rarely in the submucosal glands (Supplementary Figure 6). Their restricted positioning in the tracheal and submucosal gland epithelium highlights the importance of these lung regions in cystic fibrosis caused by *Cftr* mutations in experimental models and in patients (52, 53).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

The majority of the *Ascl1*^{pos} NE-cells were detected in the proximal-trachea, but as expected we also detected positive cells, scattered along the airways (Figure 7 B, C). Interestingly, 97% of the goblet cells were detected in the submucosal gland but not airway epithelium (Figure 7 B, F). In a hierarchical clustering of all annotated cells, the goblet cells were grouped together with the Muc5b^{pos} cells into 2 clusters (Figure 7C). In agreement with the identification of 65 goblet cells in the sc-RNA Seq data of Montoro et al. (48), *Gp2* is detected in the majority of goblet cells. We also detected *Tff2*^{pos} *Muc5b*^{pos} cells corresponding the described *goblet-1* sub-cluster and cells corresponding to the *Dcpp3*^{pos} *Lipf*^{pos} *goblet-2* sub-cluster. Interestingly, SCRINSHOT revealed some additional spatial heterogeneity of *goblet-2* cells in the submucosal glands. We also noticed Gp2^{pos}positive cells, expressing high levels of either *Dcpp3* or *Lipf* (Supplementary Figure 5 A) and second, a small subset of the Dccp3^{pos} cells also expressed *Dcpp1* (Supplementary Figure 5 B, arrowhead) (48). The spatial analysis of epithelial cell types in the trachea and lung airways demonstrate the utility of SCRINSHOT in the localization of rare cell types in a complex tissue.

Spatial mapping of airway and alveolar cells

We used the expression values of 15 genes, in 14167 cells, to generated a spatial map of macrophages, AT1 and AT2 cells in the alveolar compartment and club and neuroendocrine cells in the airways (Supplementary Figure 7). We annotated an airway cell as secretory if it was *Scgb1a1*^{pos} *Cyp2f2*^{pos} *Ascl1*^{neg} and neuroendocrine if it expressed *Ascl1*. In the alveolar compartment of the *Sftpc-CreER*^{pos}*-Rosa-Ai14*^{pos} lung, we annotated 1679 *Sftpc*^{pos} cells as AT2. More than 99% of them also contained more than 3 dots of *Lyz2*, another AT2 marker (54), which is also expressed by lung inflammatory cell-types (36, 55). 91.8% of the AT2 cells were also scored positive for *Cd74*. This gene is significantly enriched in AT2 cells but it is also expressed in hematopoietic lineages (56). Additionally, SCRINSHOT analysis distinguished a group of 588 *Lyz2*^{pos} *Cd74*^{pos} *Sftpc*^{neg}, which are probably macrophages. We also detected 312 cells positive

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

only for *Lyz2* and 450 expressing only *Cd74*. This suggests that the combinatorial expression of 3 genes defines four cell populations in the alveolar compartment, AT2 cells, *Lyz2*^{pos} *Cd74*^{pos} macrophages and single *Lyz2*^{pos} or single *Cd74*^{pos} immune cells. For identification of AT1 epithelial cells, we used the expression of *Ager* (57). We scored 1129 alveolar *Ager*^{pos} *Sftpc*^{neg} *Scgb1a1*^{neg} cells as AT1. This analysis argues that SCRINSHOT can readily distinguish known epithelial cell types in the lung airways and alveoli. Additional markers for other characterized cell-types, such as endothelial cells and fibroblasts, could be also included in future analyses to facilitate the creation of complete spatial maps of cells types in various organs and help to elucidate gene-expression and cell-type distribution patterns in healthy and diseased tissues.

CONCLUSION

SCRINSHOT simplifies multiplex, *in situ* detection of RNA in various tissues. The assay is optimized for high sensitivity and specificity. We present a robust analysis protocol with minimal requirements for extensive instrumentation or computational skills, which makes it user-friendly and accessible in many fields of biology. The protocol is suitable for validation of scRNA-seq results in small- or large-scale experiments or for the analysis of gene expression changes upon genetic or chemical manipulations. Importantly, SCRINSHOT is also a low-cost method compared to other *in situ* hybridization methods or antibody detection and the straight forward analysis protocol makes it suitable for routine laboratory use. Although SCRINSHOT multiplexity is reduced compared to "*in situ* sequencing" by ligation (24) or other barcode-based approaches (4, 58), the number of detection fluorophores or hybridization detection cycles can be increased to detect 50 genes per slide. Manual cell segmentation is the most time-consuming part of the analysis and we believe that automatic segmentation will provide a future solution. At present, our attempts with open access programs like Ilastik (59) or Cell Profiler (60) only scored a 50% success rate in segmentation of lung sections presumably because of the elongated and

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

overlapping cell shapes and the large cavities of the lung alveolar compartment. Confocal microscopy and higher magnifications may improve resolution but cost will also increase. The SCRINSHOT analysis of the mouse submucosal gland, trachea and proximal-lung airway epithelium for 29 genes and distal-lung for the expression of 15 genes demonstrates the capacity of the method to detect not only low abundance mRNAs but also rare cell-types, like ionocytes, tuft and neuroendocrine cells, facilitating the generation of cellular maps from complex tissues.

MATERIALS AND METHODS

Animals and Histology

All experiments with *wild type* (C57Bl/6J) mice were approved by the Northen Stockholm Animal Ethics Committee (Ethical Permit numbers N254/2014, N91/2016 and N92/2016).

For transgenic mouse experiments, *Sftpc-CreER*^{negative}(42);*Rosa26-Ai14*^{positive}(44) and *Sftpc-CreER*^{negative};*Rosa26-Ai14*^{positive} individuals were used according to German regulation for animal welfare at the Justus Liebig University of Giessen (Ethical Permit number GI 20/10, Nr. G 21/2017). The recombination and *Sftpc*^{pos} cell labelling of the *Sftpc-CreER*^{pos};*Rosa26-Ai14*^{pos} lung was done by one subcutaneous injection of Taxamoxifen, on postnatal day 1 (P1) and the tissues were collected on P21. An *Sftpc-CreER*^{neg};*Rosa26-Ai14*^{pos} littermate treated as above and used as negative control, in addition to a *wild type* P60 lung.

For lung tissue collection, the mice were anesthetized with a lethal dose of ketamine/xylazine. Lungs were perfused with ice cold PBS 1X pH7.4, through the heart right ventricle to remove red blood cells from the organ. A mixture (2:1 v/v) of (PFA 4% (Merck, 104005) in PBS 1X pH7.4):OCT (Leica Surgipath, FSC22) was injected into the lung, from the trachea using an insulin syringe with 20-24G plastic catheter B Braun 4251130-01), until the tip of the accessory lobe got inflated and the trachea was tied with surgical silk (Vömel, 14739). The lungs were removed and placed

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

411 in PFA 4% in PBS 1X pH7.4 for 4 hours for P21 lungs and 8 hours for adult, at 4°C in the dark,
412 with gentle rotation or shaking. The other analyzed organs were simply placed in PFA 4% in PBS
413 1X pH7.4 for 8 hours.

414 The tissues were transferred to a new tube with a mixture (2:1 v/v) of (30% sucrose in PBS 1X
415 pH7.4):OCT at 4°C for 12-16 hours, with gentle rotation or shaking. Thereafter, tissues were
416 embedded in OCT (Leica Surgipath, FSC22), using specific molds (Leica Surgipath, 3803025)
417 and frozen in a slurry of isopentane and dry ice. Tissue-OCT blocks were kept at -80°C until
418 sectioning. 10 µm thick sections were cut with a cryostat (Leica CM3050S) and placed on poly-
419 lysine slides (Thermo, J2800AMNZ), kept at room temperature for 3 hours with silica gel (Merck,
420 101969) and then stored at -80°C for further use.

421

422 **Embryonic human lungs**

423 Use of human fetal material from elective routine abortions was approved by the Swedish National
424 Board of Health and Welfare and the analysis using this material were approved by the Swedish
425 Ethical Review Authority (2018/769-31). After the clinical staff acquired informed written consent
426 by the patient, the tissue retrieved was transferred to the research prenatal material lung sample
427 was retrieved from a fetus at 8.5 w post-conception. The lung tissue was fixed for 8 hours at 4°C
428 and processed as above.

429

430 **Probe Design**

431 A detailed description of the Padlock probe design is provided in the Additional File 1. Briefly, the
432 PrimerQuest online tool (Integrated DNA Technologies: IDT) was used to select sequences of
433 Taqman probes (40-45 nucleotides) for the targeted mRNA. These sequences were then
434 interrogated against targeted-organism genome and transcriptome, with Blastn tool (NLM) to

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

guarantee their specificity. The Padlock Design Assistant.xlsm file was used to split the sequences in two and integrate them into the padlock backbone. Padlock probes were ordered from IDT as Ultramer DNA oligos and their sequences are provided in Additional File 2.

The 40-45 Taqman probe sequences were also used to prepare the fluorophore-labelled oligos. Using the IDT OligoAnalyzer tool the length of the sequences was adjusted to Tm 56^{o}C. To remove the fluorescent oligos after each detection cycle, we exchanged “T” nucleotides with “U” and treated with Uracil-DNA Glycosylase (Thermo, EN0362). Detection oligos were labelled at their 3’-end with fluorophores and manufactured by Eurofins Genomics. The sequences of fluorophore-labelled oligos are provided in Additional File 3.

Pretreatments of the slides

Slides were transferred from -80^{o}C to 45^{o}C to reduce moisture. A post fixation step with 4% PFA in PBS 1x pH7.4 was done, followed by washes with PBS-Tween20 0.05%. Permeabilization of tissues was done with 0.1M HCl for 3 minutes, followed by two washes with PBS-Tween20 0.05% and dehydration with a series of ethanol. SecureSeal^{TM} hybridization chambers (Grace Bio-Labs) were mounted on the slides and sections were preconditioned for 30 minutes at room temperature (R/T) with hybridization-reaction mixture of 1X Ampligase Buffer (Lucigen, A1905B), 0.05M KCl (Sigma-Aldrich, 60142), 20% deionized Formamide (Sigma-Aldrich, F9037), 0.2µg/ul BSA (New England Biolabs, B9000S), 1U/µl Ribolock (Thermo, EO0384) and 0.2µg/µl tRNA (Ambion, AM7119). To block unspecific binding of DNA we included 0.1µM Oligo-dT30 VN.

Padlock probe hybridization, ligation and RCA

Hybridization of the Padlock probes was done in the above solution, omitting Oligo-dT30 VN and adding 0.05 µM of each Padlock probe. We used three padlock probes for every targeted RNA-

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

459 species. For the highly abundant *Scgb1a1* mRNA, we used 0.01 µM of one padlock probe to
460 minimize molecular and optical saturation. Hybridization included a denaturation step at 55°C for
461 15 min and an annealing step at 45°C for 2 hours. Not hybridized padlock probes were removed
462 by washes with 10% Formamide in 2X SSC (Sigma-Aldrich, S6639). To minimize the effect of
463 the previously documented SplintR ligase nucleotide preferences (26, 33), padlock probe ligation
464 was performed overnight O/N at 25°C using the SplintR ligase (NEB, M0375) at a final
465 concentration of 0.5 Units/µl, T4 RNA ligase buffer (NEB, B0216) and 10µM ATP, according to
466 manufacturer recommendations (see New England Biolabs webpage).

467 Rolling Cycle Amplification (RCA) was done O/N at 30°C using 0.5 Units/µl Φ29 polymerase
468 (Lucigen, 30221-2). The reaction mixture contained also 1X Φ29 buffer, 5% Glycerol, 0.25 mM
469 dNTPs (Thermo, R0193), 0.2 µg/µl BSA and 0.1 µM RCA primers (RCA Primer1:
470 TAAATAGACGCAGTCAGT*A*A and RCA Primer2: CGCAAGATATACG*T*C). The "*" indicate
471 Thiophosphate-modified bounds to inhibit the 3-5 exonuclease activity of Φ29 polymerase (23).
472 A fixation step with 4% PFA for 15 minutes was done to ensure stabilization of the RCA-products
473 on the tissue. Sections were thoroughly washed with PBS-Tween20 0.05% before next step.

474

475 **Hybridization of detection-oligos**

476 The visualization of the RCA products was done with hybridization of the 3'-fluorophore labeled
477 detection oligos. The reaction mixture contained 2X SSC, 20% deionized Formamide, 0.2 µg/µl
478 BSA, 0.5 ng/µl DAPI (Biolegend, 422801), 0.4 µM FITC-labeled and 0.2 µM Cy3- and Cy5-labeled
479 detection probes for 1 hour at R/T. Washes were performed with 10% Formamide in 2X SSC,
480 followed by 6X SSC. Tissues sections were dehydrated with a series of ethanol, chambers were
481 removed and the slides were covered with SlowFade™ Gold Antifade Mountant (Thermo,
482 S36936) and a coverslip.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

After image acquisition, the coverslips were removed by placing the slides in 70% ethanol in 45°C. Then, sections were dehydrated using a series of ethanol to mount the hybridization chambers. After tissue rehydration with PBS-Tween20 0.05%, the detection-oligos were digested with Uracil-DNA Glycosylase for 1hour at 37°C. The reaction mixture contained 1X UNG buffer, 0.2 μg/μl BSA and 0.02 Units/μl Uracil-DNA Glycosylase (Thermo, EN0362). Destabilized oligos were stripped off by thorough washes with 65% deionized Formamide at 30°C. Multiple rounds of hybridization and imaging, as described above, were performed until all genes were imaged.

Image acquisition

Images were captured with a Zeiss Axio Observer Z.2 fluorescent microscope with a Colibri led light source, equipped with a Zeiss AxioCam 506 Mono digital camera and an automated stage, set to detect the same regions after every hybridization cycle.

Image analysis

The nuclear staining was used to align the images of the same areas between the hybridizations and multi-channel *.czi files, containing the images of all genes, were created using Zen2.5 (Carl Zeiss Microscopy GmbH). The images were analyzed as 16bit *.tiff files, without compression or scaling. Images were tiled using a custom script in Matlab (The MathWorks, Inc.). Manual nuclear segmentation was done with Fiji ROI Manager (61). The nuclear ROIs were expanded for 2 μm with a custom Cell-Profiler script and considered as cells. The signal-dots were counted in these cell-ROIs using Cell-Profiler 3.15 (60), Fiji (62, 63) and R-RStudio (58, 64-67) custom scripts (https://github.com/AlexSount/SCRINSHOT).

Thresholding

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A cell-ROI size criterion was applied to remove the outliers with very small or big surface. In particular, only cells included between two standard deviations of the mean size of the analyzed cells were further processed. In SCRINSHOT we considered a cell-ROI positive for an analyzed gene, we used a threshold strategy. First, we determined the maximum numbers of signal-dots per cell-ROI for all analyzed genes. A cell-ROI was considered as positive, if it contained more than 10% of the maximum number of signal-dots for the specific gene. The higher threshold was set to 3, which was applied for highly abundant genes with maxima over 31 signal dots.

Curation of the data

In general, the 2 µm nuclear expansion provides an underestimation of real signal dots and provides satisfactory results for airway cells (e.g. Figure 3A' merge-images with cell-ROI outlines) but the cellular segmentation of the alveolar region is more challenging mainly because of the irregular cell shapes and their overlap. This gave false positive cell-ROIs due to dots from adjacent true-positive cells being erroneously assigned to their neighbors. To reduce the noise, signal-dots of highly abundant RNA-species were used to visually inspect and remove the problematic cell-ROIs from further analysis.

Clustering

Annotated cells of submucosal gland, trachea and lung airway epithelium were clustered using hclust package in R (68). Log_2(dots/cell + 1) values were used to calculate Euclidean distances and clustering was done using ward.D2 method. Balloon plots were created by ggpubr package in R (69) and heatmaps with pheatmap package (70). Bootstrapping analysis was done by the clusterboot package (71). Clusters were considered as "stable" if bootstrap values were >0.5.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Analysis of *Sftpc-CreER*^{pos};*Rosa26-Ai14*^{pos} cells

For the identification of the RFP^{pos} alveolar cell-ROIs in the *Sftpc-CreER*^{pos};*Rosa26-Ai14*^{pos} lung, all analyzed alveolar cells from the RFP^{neg} *Sftpc-CreER*^{neg};*Rosa26-Ai14*^{pos} lung were used to determine the maximum endogenous fluorescence (in Raw Integrated Density values) and set it as a threshold. The RFP^{pos} and RFP^{neg} cell-ROIs were curated and the *Sftpc* signal-dots were measured in them.

For the correlation of RFP endogenous fluorescence with SCRINSHOT *RFP* signal, the endogenous fluorescence (in Raw Integrated Density values) of all segmented cell-ROIs from the analyzed tissue sections was correlated with the detected RFP signal-dots in the same cell-ROIs, using simple linear regression analysis in Graphpad Prism with default settings.

Comparison of SplintR and cDNA-based detection of RNA species

To compare the performance of padlock-probe hybridization (i) directly on RNA (SCRINSHOT) and (ii) after cDNA synthesis, we used sequential 10μm-thick sections from adult mouse lungs, fixed for 8 hours (see above). cDNA-based approaches in earlier publications have only used fresh frozen tissues and pepsin or proteinase K tissue treatments to increase RNA accessibility (3, 24). The padlock probes were designed to recognize exactly the same sequence of the analyzed genes. For the highly (*Scgb1a1, Sftpc*) and intermediate (*Actb*) expressed genes, we used only one padlock probe and for *Pecam1*, three probes because of its lower expression levels.

For cDNA synthesis, the RNA species of the tissue sections were transformed to cDNA by reverse transcription (RT) using random decamers. The RNA strands were degraded with RNaseH to let padlock probes hybridize to the corresponding cDNA sequences. Padlock probe ligation was

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

554 done with Ampligase DNA ligase. All other steps were done according to the provided
555 SCRINSHOT protocol.

556 For cDNA synthesis we used 20 U/µl of the SuperScript™ II Reverse Transcriptase (Thermo,
557 18064014), 1X SuperScript™ II RT buffer (Thermo), 0.5 mM dNTPs (Thermo, R0193), 10mers-
558 random primer (Thermo), 0.2 µg/µl BSA (NEB, B9000S), and 1 U/µl RiboLock RNase Inhibitor
559 (Thermo, EO0384), at 42^{o}C, O/N. The slides were post-fixed with 4% (w/v) PFA in PBS 1X pH7.4,
560 at RT for 30 min, following by 6 washes with PBS-Tween 0.05%.

561

562 **Evaluation of SCRINSHOT specificity using mutated padlock probes**

563 The experiment was done as described above, using 0.01 µM of *Scgb1a1* padlock probes and
564 0.05 µM of the *Actb*. The "mismatch" probe had the same sequence as the normal *Scgb1a1*
565 padlock probe with a C>G substitution at the 5'- ligation site and the "3'-scrambled" probe had the
566 same 5'-arm but the 3'-arm was scrambled. *Actb* padlock probe was used as internal control to
567 calculate the *Scgb1a1/Actb* ratios. The *Scgb1a1/Actb* ratios of cell-ROIs with zero *Actb* signal-
568 dots were considered as zeros. The detection of all *Scgb1a1* RCA-products was done using a
569 detection oligo, which recognizes the padlock-backbone of *Scgb1a1* but not of *Actb*. *Actb* RCA-
570 product was detected by a detection oligo, which recognizes its gene-specific sequence
571 (Additional File 3). The *Scgb1a1/Actb* fluorescence ratios were calculated using the Raw
572 Integrated Densities of the two genes in each cell-ROI.

573

574 **Evaluation of SCRINSHOT specificity using an antisense competitor**

575 To compete the binding of Scgb1a1 padlock probe on its target transcript, an antisense competitor
576 (A), which recognizes the Scgb1a1 mRNA between nucleotides 316-378 and masks the padlock
577 probe hybridization site, was used in the same hybridization mixture with the *Scgb1a1* padlock

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

probe (P) at different ratios: (i) P:A=1:0, (ii) P:A=1:1, (iii) P:A=1:5, keeping the *Scgb1a1* padlock probe concentration to 0.05μM.

Correlation with SCRINSHOT with published single cell RNA sequencing datasets

The GSE118891 dataset was used to retrieve gene expression values (raw counts) of all AT2, according to the cell annotation of provided metadata file (47). The genes of interest were selected and their Mean values were calculated and $\log_2$(dots+1) transformed. Similarly, the SCRINSHOT signal-dots per cell-ROI were $\log_2$(dots+1) transformed. Pearson correlation analysis was done using GraphPad Prism.

Statistical analysis

All statistical analyses were done with Graphpad Prism, using nonparametric tests, since the SCRINSHOT data do not follow canonical distributions. Multiple comparisons were done using ANOVA Kruskal-Wallis multiple comparison test, without multiple comparison correction ("*": P ≤ 0.05, "**": P ≤ 0.01, "***": P ≤ 0.001, "****": P ≤ 0.0001). For pairwise comparisons the statistical analysis was done using Mann-Whitney nonparametric t-test ("*": P ≤ 0.05, "**": P ≤ 0.01, "***": P ≤ 0.001, "****": P ≤ 0.0001). Spearman correlation was used to examine the correlation between SCRINSHOT and scRNA-Seq data.

Availability of data and materials

The datasets and analysis files of the current study have been deposited at Zenodo repository (DOI: 10.5281/zenodo.3634561). All scripts are available at https://github.com/AlexSount/SCRINSHOT.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Author contributions

CS and AS conceived the idea and designed the experiments with help by MN and WS. ABF and AF provided mouse tissues and ES the human fetal material. AS and HPN performed all hybridizations. AS, AL and XQ wrote the analysis scripts. AS, AL and ABF analyzed the datasets. All authors helped interpreting results. CS, AS and HPN wrote the manuscript and AS the detailed protocol (Additional file 1) with input from all other authors. All authors read and approved the final manuscript. CS provided all funding support.

Acknowledgments

We would like to thank the Associate Professor Qi Dai for the thorough reading and suggestions of the detailed SCRINSHOT protocol (Additional File 1).

Funding

This study was supported by grants from the Swedish Research Council (Vetenskapsrådet, 2016-05059) and the Swedish Cancer Society (Cancerfonden, 160499) to CS.

Competing interests

MN and XQ hold shares in CARTANA AB, a company that commercializes *in situ* sequencing technology.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure Legends

Figure 1. Schematic representation of SCRINSHOT. The major steps of the assay are (A) All padlock-probe hybridization for all the targeted RNA-species followed by ligation and RCA-amplification. (B) RCA-products are detected sequentially reading three per cycle, with FITC-, Cy3- and Cy5- labelled detection-oligos, which recognize the gene-specific part of the padlock probes. (C) Images from all detection cycles are aligned using DAPI nuclear staining and segmented to create the cell-ROIs and signal-dots are counted and registered to the cell-ROIs.

Figure 2. SCRINSHOT specificity relies on stringent hybridization of padlock probes to their target RNAs. Images of SCRINSHOT signal, using normal *Scgb1a1* padlock probe (A), a *Scgb1a1* padlock probe with a point mutation at its ligation site (B), a *Scgb1a1* padlock probe with 3'-scrambled arm (C) and normal padlock probe but omitting SplintR ligase (D). *Actb* normal padlock probe was used in all conditions as internal control. DAPI: blue, *Scgb1a1*: gray, *Actb*: red. "n" indicates the number of airway cells in the corresponding images. (A'-D') Magnified areas of the indicated positions (square brackets) of images in the left. Pink outlines show the 2 µm expanded airway nuclear ROIs, which are considered as cells. Scale-bar: 150 µm. (E) Violin plot of the Scgb1a1 and Actb signal-dots ratio in all airway cells. The ratio of cells with zero *Actb*-dots considered as zero. (F) Violin plot of the *Scgb1a1* and *Actb* fluorescence intensity ratio in all airway cells. SplintR^{pos} n=473, mismatch n=574, 3´-scrambled n=507 and SplintR^{neg} n=488.

Figure 3. Concentration-dependent SCRINSHOT detection efficiency. (A-C) Representative images of SCRINSHOT signals for *Scgb1a1* padlock probes in the absence or presence of an antisense competitor oligo targeting the binding site of *Scgb1a1* padlock probe. SCRINSHOT signal for *Scgb1a1* padlock probe in the absence of the antisense competitor oligo (A), when

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

mixed in equal concentration with the antisense competitor oligo (B) and when mixed with five times higher concentration of the antisense competitor oligo (C). (A'-C') Magnified areas of the indicated positions (square brackets) of images A-C. (D) Violin plot of the *Scgb1a1* and *Actb* signal-dots ratio in all airway cells of the three compared conditions. P: padlock probe, A: antisense competitor. Scale bars in (A-C) is 100 µm and (A'-C') is 10µm. All analyses were done using raw images and the same thresholds. For visualization purposes, brightness and contrast of *Scgb1a1* were set independently in the compared conditions to show the existence of signal in the presence antisense competitor and avoid signal saturation in its absence.

Figure 4. Efficient detection of *Rosa-Ai14* mRNA upon *Sftpc-CreER* driven recombination, using SCRINSHOT. (A) Representative images of alveolar regions from P21 *Sftpc-CreER*^{pos};*Rosa-Ai14*^{pos}, P21 *Sftpc-CreER*^{neg};*Rosa-Ai14*^{pos} and P60 wild-type (*Sftpc-CreER*^{neg};*Rosa-Ai14*^{neg}) lung sections. Asterisks indicate red blood cells with high auto-fluorescence, which give some false positive signals in wild-type lung, only. DAPI: blue, *RFP* mRNA detected with SCRINSHOT: magenta and endogenous fluorescence of RFP: green. Scale-bar: 10µm. (B) Regression analyses of RFP endogenous fluorescence (RAW integrated density) and *Ai14* SCRINSHOT dots, in the three analyzed lung sections, using Graphpad Prism. "n" indicates the number of total cells in the analyzed images, which used for the statistical analysis.

Figure 5. Correlation of *Sftpc* SCRINSHOT dots with RFP endogenous fluorescence and *RFP* mRNA signals. (A) Images of alveolar compartment from P21 *Sftpc-CreER*^{pos};*Rosa-Ai14*^{pos}, P21 *Sftpc-CreER*^{neg};*Rosa-Ai14*^{pos} and P60 wild-type (*Sftpc-CreER*^{neg};*Rosa-Ai14*^{neg}) lung sections, showing *Sftpc*^{pos} (gray) AT2 cells in correlation to *Ai14* SCRINSHOT dots (detected *RFP* mRNA) (green) and RFP endogenous fluorescence (red). Scale-bar: 10 µm. "n" designates the number of analyzed AT2 cells. (B) Violin plots of SCRINSHOT results for *RFP*, in addition to RFP

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

endogenous fluorescence (RAW integrated density) in AT2 cells of the analyzed lungs. (E) Bar plots of SCRINSHOT results for *Sftpc* in RFP^{pos} and RFP^{neg} alveolar cells, as indicated by endogenous fluorescence of *Sftpc-CreER*^{neg};*Rosa-Ai14*^{pos} alveolar cells (threshold 2304340 units of raw integrated density), in addition to RFP in *Sftpc*^{pos} and *Sftpc*^{neg} AT2 cells. Threshold is set to 3 dots and shown with red dotted lines.

Figure 6. SCRINSHOT multiplexity and correlation with single-cell mRNA sequencing. (A) Representative image of an alveolar region (Alv), showing the SCRINSHOT signal-dots of 11 AT2 selective genes. Merge image includes *Ager* (gray), *Sftpc* (green), *Cd74* (magenta) and DAPI (blue) for nuclear staining. The rest of the genes are indicated with italics and only single-channel images were used. (B) Image of an airway region (arw) with a neuroepithelial body (NEB), containing SCRINSHOT dots for the club-cell markers *Scgb1a1* (green) and *Cyp2f2* (yellow) and neuroendocrine (NE)-cell markers *Calca* (magenta) and *Ascl1* (gray). DAPI (blue) used for nuclear staining. (C) Correlation plot of the indicated genes, between the $\log_2$(raw counts+1) values of 156 AT2 cells from Liu et al., 2019 and SCRINSHOT signal dots of 1679 AT2 cells. Scale bar: 10μm.

Figure 7. Spatial mapping of tracheal cell types with SCRINSHOT. (A) Overview of analyzed trachea and lung tissue sections using nuclear staining. Analyzed areas are indicated by brackets, corresponding to proximal trachea, submucosal glands (GS), intermediate and distal trachea, followed by proximal and distal lung airways. Trachea image includes also thyroid glands. Scale bar: 500μm. (B) Balloon plot of the annotated cell-types at the analyzed positions. The balloon size indicates the percentage of the cell-type at the indicated position relatively to the total number of cells of the cell-type. The balloon intensity corresponds to the percentage of the specified cell-type, relative of the total number of cells at the indicated position. (C) Hierarchical clustering of

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the annotated cells. Heatmap shows the $\log_2$(dots/cell + 1) gene values in the analyzed cells. (D) Characteristic example of basal cells in tracheal epithelium. *Trp63* (green), *Ktr5* (red), *Pdpn* (gray) and DAPI (blue). Scale bar: 10µm. (E) Detection of a tuft cell in tracheal epithelium. *Sox9* (green), *Gng13* (red), *Alox5ap* (gray), *Trpm5* (yellow), *Six1* (magenta) and DAPI (blue). Scale bar: 10µm. (F) An ionocyte in trachea airway epithelium, detected with *Tfcp2l1* (green), *Foxi1* (red), *Cftr* (gray), *Ascl3* (yellow) and Dapi (blue). Scale bar: 10µm.

Supplementary Figure 1. Comparison of SplintR-based (SCRINSHOT) and the cDNA-based *in situ* hybridization assays for high, intermediate and low abundant genes in sequential PFA fixed lung sections. (A) Images of SplintR-based (SCRINSHOT) and cDNA-based *in situ* hybridization assays, in sequential lung sections. DAPI: blue, *Scgb1a1*: green, *Sftpc*: gray, *Actb*: red and *Pecam1*: cyan. Pink outlines show the 2 µm expanded airway nuclear ROIs, which are considered as cells. The square brackets indicate the magnified areas on the right. The “n” correspond to the number of counted cells in large images. Scale bar: 100µm. (B) Bar-plots of the analyzed gene signals, in the indicated tissue compartments, for SCRINSHOT and cDNA-based approaches. The differences between the two conditions are significant ($P < 0.0001$) for all analyzed genes. (C) Histograms of the analyzed genes. The Y-axes indicate the percentage of the cell-ROIs and the X-axes, the binned signal dots in each cell. In SplintR-condition, 350 cells localized in airways (arw) and 1624 in alveolar (alv) compartment. In cDNA-condition, there are 295 airway and 1706 alveolar cells. Analysis was done using raw images, with the same acquisition conditions and thresholds. Only for visualization purposes, signal intensity of *Scgb1a1* and *Sftpc* in cDNA-condition was set 5-times higher than SplintR.

Supplementary Figure 2. SCRINSHOT application on mouse kidney and heart. (A) Application of SCRINSHOT in adult mouse kidney sections shows signal for *Actb* (gray), *Napsa* (magenta)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

and *Lyz2* (cyan) but not *Scgb1a1* (green) and *Sftpc* (red). (B) Representative image from an adult mouse heart section, containing a vessel (v). *Pecam1* (magenta) is detected at the vessel walls, *Actb* (gray) is uniformly expressed and *Lyz2* (cyan) labels a few cells (presumably macrophages) but not *Scgb1a1* (green) and *Sftpc* (red) are not detected. DAPI: blue, scale bar: 50μm.

Supplementary Figure 3. SCRINSHOT application on fetal human lung. On the left, overview of a w8.5 whole left lung tissue section, showing SCRINSHOT signal for *SOX2* (green), *SOX9* (red) and *ASCL1* (gray). The square brackets correspond to the images on the right. DAPI (blue) was used as nuclear staining. Scale bar: 500μm. (A) Representative image of proximal epithelium, which is highly positive of *SOX2* and *ASCL1* but not *SOX9*. (B) Representative image of highly *SOX9* positive distal epithelium.

Supplementary Figure 4. Comparison of detected *Calca* RCA-products in first and eighth hybridizations. (A, B) Images on the left show the *Ascl1*^{pos} (gray) neuroendocrine cells of an airway neuroepithelial body, in relation to *Scgb1a1*^{pos} (red) club cells. DAPI: blue, scale bar: 10μm. (C) Note that neuroepithelial bodies are tightly packed cellular structures, as indicated by DAPI nuclear staining. The images on the right show the *Calca* RCA-products, detected in the first (D) and the eighth (E) detection cycles. (D'-F') Magnified areas of the indicated positions (brackets) of images D-F, respectively. (F) Overlay of *Calca* identified signal-dots in first (cyan) and eighth (yellow) detection cycles, using the same threshold in CellProfiler.

Supplementary Figure 5. SCRINSHOT mapping of submucosal glands reveals spatial heterogeneity in goblet cell population. Overview of the analyzed submucosal gland for the expression of 6 goblet cell markers shows their expression in submucosal gland but not airway

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

epithelium. *Muc5b* is detected along the airway epithelium of the proximal trachea, indicating that it is a general proximal epithelial cell marker. *Tff2*: green, *Muc5b*: red, *Gp2*: gray, *Dcpp3*: cyan, *Dcpp1*: magenta, *Lipf*: yellow, DAPI: blue and cell-ROIs: pink. Scale bar: 500μm. (A) Insert showing the previously described *Muc5b*^{pos} *Tff2*^{pos} goblet subtype (arrow) and the *Lipf*^{pos} *Dcpp3*^{pos} (asterisk). *Lipf*^{pos} *(hash)* and *Dcpp3*^{pos} (arrowhead) are detected in the same region, being positive for the general goblet cell marker *Gp2*. (B) Insert showing regionally restricted expression of *Dcpp1* in a subset of *Dcpp3*^{pos} cells. Insert scale bar: 10μm.

Supplementary Figure 6. Ionocyte in the submucosal gland. (Top) overview of the analyzed submucosal gland, showing SCRINSHOT signal for ionocyte markers, *Tfcp2l1* (green), *Cftr* (gray), *Foxi1* (red), *Ascl3* (yellow) and DAPI (blue). Scale bar: 500μm. (Bottom) Magnified area of the indicated square in overview image, showing a detected ionocyte in submucosal gland. Scale bar: 10μm.

Supplementary Figure 7. SCRINSHOT generates cell-type digital annotation maps of large tissue areas. (i) Overview image of SCRINSHOT fluorescence signal-dots for *Scgb1a1* (green), *Sftpc* (magenta) and *Ascl1* (gray) of a large area from P21 *Sftpc-CreER*^{pos};*Rosa-Ai14*^{pos} lung section, after Tamoxifen induction on P1 (RFP was not shown at that image). The image contains 14167 manual segmented nuclei, which were expanded for 2 μm and considered as cells. (ii) Spatial map of annotated cell types according to the indicated criteria. (A) Same airway area as in Figure 6B, showing club (*Scgb1a1*: green and *Cyp2f2*: yellow) and NE-cell (*Calca*: magenta and *Ascl1*: gray) markers, in an airway position with a neuro-epithelial body (NEB). (A') Cell-type digital annotation of the area corresponding to "A". The "*" indicate *Ascl1*^{pos} *Calca*^{neg} cells Club cells: green and NE-cells: red. (B) Same alveolar area as in Figure 6A showing *Ager*^{pos} (gray) *Sftpc*^{neg} (green) *Cd74*^{neg} (magenta) AT1 cells (arrow), *Ager*^{low} *Sftpc*^{pos} *Cd74*^{low} AT2 cells (asterisks)

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

771 and *Ager*^{neg} *Sftpc*^{neg} *Lyz2*^{pos} *Cd74*^{high} macrophages (mΦ, arrowhead). (B') Cell-type digital
772 annotation of the area corresponding to "B". AT1 cells: gray, AT2 cells: magenta and
773 macrophages: white. All not annotated cells are depicted with blue. Scale bar: 200μm.

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Literature

1. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative Analysis of Single-Cell RNA Sequencing Methods. Mol Cell. 2017;65(4):631-43 e4.

2. Strell C, Hilscher MM, Laxman N, Svedlund J, Wu C, Yokota C, et al. Placing RNA in context and space - methods for spatially resolved transcriptomics. FEBS J. 2019;286(8):1468-81.

3. Ke R, Mignardi M, Pacureanu A, Svedlund J, Botling J, Wahlby C, et al. In situ sequencing for RNA analysis in preserved tissue and cells. Nat Methods. 2013;10(9):857-60.

4. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. Science. 2015;348(6233):aaa6090.

5. Codeluppi S, Borm LE, Zeisel A, La Manno G, van Lunteren JA, Svensson CI, et al. Spatial organization of the somatosensory cortex revealed by osmFISH. Nat Methods. 2018;15(11):932-5.

6. Stahl PL, Salmen F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science. 2016;353(6294):78-82.

7. Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, et al. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. Science. 2019;363(6434):1463-7.

8. Gall JG, Pardue ML. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci U S A. 1969;63(2):378-83.

9. Femino AM, Fay FS, Fogarty K, Singer RH. Visualization of single RNA transcripts in situ. Science. 1998;280(5363):585-90.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

10. Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods. 2008;5(10):877-9.

11. Strell C, Hilscher MM, Laxman N, Svedlund J, Wu C, Yokota C, et al. Placing RNA in context and space – methods for spatially resolved transcriptomics. 2019;286(8):1468-81.

12. Nagendran M, Riordan DP, Harbury PB, Desai TJJE. Automated cell-type classification in intact tissues by single-cell molecular profiling. 2018;7:e30510.

13. Ramachandran Iyer EP, Punthambaker S, Liu S, Jindal K, Farrell M, Murn J, et al. Barcoded oligonucleotides ligated on RNA amplified for multiplex and parallel in-situ analyses. 2018:281121.

14. Eng CL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. Nature. 2019;568(7751):235-9.

15. Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L. Single-cell in situ RNA profiling by sequential hybridization. Nat Methods. 2014;11(4):360-1.

16. Shah S, Lubeck E, Schwarzkopf M, He TF, Greenbaum A, Sohn CH, et al. Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. Development. 2016;143(15):2862-7.

17. Wang F, Flanagan J, Su N, Wang L-C, Bui S, Nielson A, et al. RNAscope: A Novel in Situ RNA Analysis Platform for Formalin-Fixed, Paraffin-Embedded Tissues. The Journal of Molecular Diagnostics. 2012;14(1):22-9.

18. Choi HMT, Calvert CR, Husain N, Huss D, Barsi JC, Deverman BE, et al. Mapping a multiplexed zoo of mRNA expression. 2016;143(19):3632-7.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

836 19. Shah S, Lubeck E, Schwarzkopf M, He T-F, Greenbaum A, Sohn CH, et al. Single-
837 molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding
838 and clearing. 2016;143(15):2862-7.

839 20. Marras SAE, Bushkin Y, Tyagi S. High-fidelity amplified FISH for the detection and allelic
840 discrimination of single mRNA molecules. Proc Natl Acad Sci U S A. 2019;116(28):13921-6.

841 21. Larsson C, Grundberg I, Soderberg O, Nilsson M. In situ detection and genotyping of
842 individual mRNA molecules. Nat Methods. 2010;7(5):395-7.

843 22. Lizardi PM, Huang X, Zhu Z, Bray-Ward P, Thomas DC, Ward DC. Mutation detection and
844 single-molecule counting using isothermal rolling-circle amplification. Nat Genet. 1998;19(3):225-
845 32.

846 23. Dean FB, Nelson JR, Giesler TL, Lasken RS. Rapid amplification of plasmid and phage
847 DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. Genome Res.
848 2001;11(6):1095-9.

849 24. Svedlund J, Strell C, Qian X, Zilkens KJC, Tobin NP, Bergh J, et al. Generation of in situ
850 sequencing based OncoMaps to spatially resolve gene expression profiles of diagnostic and
851 prognostic markers in breast cancer. EBioMedicine. 2019.

852 25. Wang X, Allen WE, Wright MA, Sylwestrak EL, Samusik N, Vesuna S, et al. Three-
853 dimensional intact-tissue sequencing of single-cell transcriptional states. Science.
854 2018;361(6400).

855 26. Iyer EPR, Punthambaker S, Liu S, Jindal K, Farrell M, Murn J, et al. Barcoded
856 oligonucleotides ligated on RNA amplified for multiplex and parallel in-situ analyses. bioRxiv.
857 2018:281121.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

858 27. Rouhanifard SH, Mellis IA, Dunagin M, Bayatpour S, Jiang CL, Dardani I, et al. ClampFISH
859 detects individual nucleic acid molecules using click chemistry-based amplification. Nat
860 Biotechnol. 2018.

861 28. Nagendran M, Riordan DP, Harbury PB, Desai TJ. Automated cell-type classification in
862 intact tissues by single-cell molecular profiling. Elife. 2018;7.

863 29. Ho CK, Van Etten JL, Shuman S. Characterization of an ATP-dependent DNA ligase
864 encoded by Chlorella virus PBCV-1. J Virol. 1997;71(3):1931-7.

865 30. Jiang M, Liu L, Hong C, Chen D, Yao X, Chen X, et al. Single molecule chromogenic in
866 situ hybridization assay for RNA visualization in fixed cells and tissues. RNA. 2019;25(8):1038-
867 46.

868 31. Takahashi H, Ohkawachi M, Horio K, Kobori T, Aki T, Matsumura Y, et al. RNase H-
869 assisted RNA-primed rolling circle amplification for targeted RNA sequence detection. Sci Rep.
870 2018;8(1):7770.

871 32. Schneider N, Meier M. Efficient in situ detection of mRNAs using the Chlorella virus DNA
872 ligase for padlock probe ligation. RNA. 2017;23(2):250-6.

873 33. Krzywkowski T, Nilsson M. Fidelity of RNA templated end-joining by chlorella virus DNA
874 ligase and a novel iLock assay with improved direct RNA detection accuracy. Nucleic Acids Res.
875 2017;45(18):e161.

876 34. Shah S, Lubeck E, Zhou W, Cai L. In Situ Transcription Profiling of Single Cells Reveals
877 Spatial Organization of Cells in the Mouse Hippocampus. Neuron. 2016;92(2):342-57.

878 35. Ho CK, Van Etten JL, Shuman S. Characterization of an ATP-dependent DNA ligase
879 encoded by Chlorella virus PBCV-1. 1997;71(3):1931-7.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

880 36. Abram CL, Roberge GL, Hu Y, Lowell CA. Comparative analysis of the efficiency and
881 specificity of myeloid-Cre deleting strains using ROSA-EYFP reporter mice. J Immunol Methods.
882 2014;408:89-100.

883 37. Stadtfeld M, Ye M, Graf T. Identification of interventricular septum precursor cells in the
884 mouse embryo. Dev Biol. 2007;302(1):195-207.

885 38. Munro DAD, Hughes J. The Origins and Functions of Tissue-Resident Macrophages in
886 Kidney Development. Front Physiol. 2017;8:837.

887 39. Mori K, Shimizu H, Konno A, Iwanaga T. Immunohistochemical localization of napsin and
888 its potential role in protein catabolism in renal proximal tubules. Arch Histol Cytol. 2002;65(4):359-
889 68.

890 40. Danopoulos S, Alonso I, Thornton ME, Grubbs BH, Bellusci S, Warburton D, et al. Human
891 lung branching morphogenesis is orchestrated by the spatiotemporal distribution of ACTA2,
892 SOX2, and SOX9. Am J Physiol Lung Cell Mol Physiol. 2018;314(1):L144-L9.

893 41. Nikolic MZ, Caritg O, Jeng Q, Johnson JA, Sun D, Howell KJ, et al. Human embryonic
894 lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-
895 renewing organoids. Elife. 2017;6.

896 42. Chapman HA, Li X, Alexander JP, Brumwell A, Lorizio W, Tan K, et al. Integrin
897 alpha6beta4 identifies an adult distal lung epithelial population with regenerative potential in mice.
898 J Clin Invest. 2011;121(7):2855-62.

899 43. Lakso M, Sauer B, Mosinger B, Jr., Lee EJ, Manning RW, Yu SH, et al. Targeted oncogene
900 activation by site-specific recombination in transgenic mice. Proc Natl Acad Sci U S A.
901 1992;89(14):6232-6.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

902 44. Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, et al. A robust and
903 high-throughput Cre reporting and characterization system for the whole mouse brain. Nat
904 Neurosci. 2010;13(1):133-40.

905 45. Hogan BL, Barkauskas CE, Chapman HA, Epstein JA, Jain R, Hsia CC, et al. Repair and
906 regeneration of the respiratory system: complexity, plasticity, and mechanisms of lung stem cell
907 function. Cell Stem Cell. 2014;15(2):123-38.

908 46. Hug N, Longman D, Caceres JF. Mechanism and regulation of the nonsense-mediated
909 decay pathway. Nucleic Acids Res. 2016;44(4):1483-95.

910 47. Liu Q, Liu K, Cui G, Huang X, Yao S, Guo W, et al. Lung regeneration by multipotent stem
911 cells residing at the bronchioalveolar-duct junction. Nat Genet. 2019;51(4):728-38.

912 48. Montoro DT, Haber AL, Biton M, Vinarsky V, Lin B, Birket SE, et al. A revised airway
913 epithelial hierarchy includes CFTR-expressing ionocytes. Nature. 2018;560(7718):319-24.

914 49. Plasschaert LW, Zilionis R, Choo-Wing R, Savova V, Knehr J, Roma G, et al. A single-cell
915 atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature.
916 2018;560(7718):377-81.

917 50. McGovern S, Pan J, Oliver G, Cutz E, Yeger H. The role of hypoxia and neurogenic genes
918 (Mash-1 and Prox-1) in the developmental programming and maturation of pulmonary
919 neuroendocrine cells in fetal mouse lung. Lab Invest. 2010;90(2):180-95.

920 51. Metzger RJ, Klein OD, Martin GR, Krasnow MA. The branching programme of mouse lung
921 development. Nature. 2008;453(7196):745-50.

922 52. Birket SE, Davis JM, Fernandez CM, Tuggle KL, Oden AM, Chu KK, et al. Development
923 of an airway mucus defect in the cystic fibrosis rat. JCI Insight. 2018;3(1).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

924 53. Inglis SK, Wilson SM. Cystic fibrosis and airway submucosal glands. Pediatr Pulmonol.
925 2005;40(4):279-84.

926 54. Nabhan AN, Brownfield DG, Harbury PB, Krasnow MA, Desai TJ. Single-cell Wnt signaling
927 niches maintain stemness of alveolar type 2 cells. Science. 2018;359(6380):1118-23.

928 55. McCubbrey AL, Allison KC, Lee-Sherick AB, Jakubzick CV, Janssen WJ. Promoter
929 Specificity and Efficacy in Conditional and Inducible Transgenic Targeting of Lung Macrophages.
930 Front Immunol. 2017;8:1618.

931 56. Lee JH, Kim J, Gludish D, Roach RR, Saunders AH, Barrios J, et al. Surfactant protein-C
932 chromatin-bound green fluorescence protein reporter mice reveal heterogeneity of surfactant
933 protein C-expressing lung cells. Am J Respir Cell Mol Biol. 2013;48(3):288-98.

934 57. Chung MI, Hogan BLM. Ager-CreER(T2): A New Genetic Tool for Studying Lung Alveolar
935 Development, Homeostasis, and Repair. Am J Respir Cell Mol Biol. 2018;59(6):706-12.

936 58. Wickham H. Ggplot2 : elegant graphics for data analysis. New York: Springer; 2009. viii,
937 212 p. p.

938 59. Berg S, Kutra D, Kroeger T, Straehle CN, Kausler BX, Haubold C, et al. ilastik: interactive
939 machine learning for (bio)image analysis. Nat Methods. 2019.

940 60. McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, et al.
941 CellProfiler 3.0: Next-generation image processing for biology. PLoS Biol. 2018;16(7):e2005970.

942 61. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an
943 open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676-82.

944 62. Preibisch S, Saalfeld S, Tomancak P. Globally optimal stitching of tiled 3D microscopic
945 image acquisitions. Bioinformatics. 2009;25(11):1463-5.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

946 63. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image
947 analysis. Nat Methods. 2012;9(7):671-5.

948 64. Team RC. R: A language and environment for statistical computing. 2013.

949 65. Allaire J. RStudio: integrated development environment for R. Boston, MA. 2012;770.

950 66. Wickham H, Wickham MH. Package 'plyr'. Obtenido de https://cran rproject
951 org/web/packages/dplyr/dplyr pdf. 2016.

952 67. Peterson M, Malloy J, Buonaccorsi V, Marden J. Teaching RNAseq at undergraduate
953 institutions: a tutorial and R package from the Genome Consortium for Active Teaching.
954 CourseSource doi: https://doi org/1024918/cs. 2015.

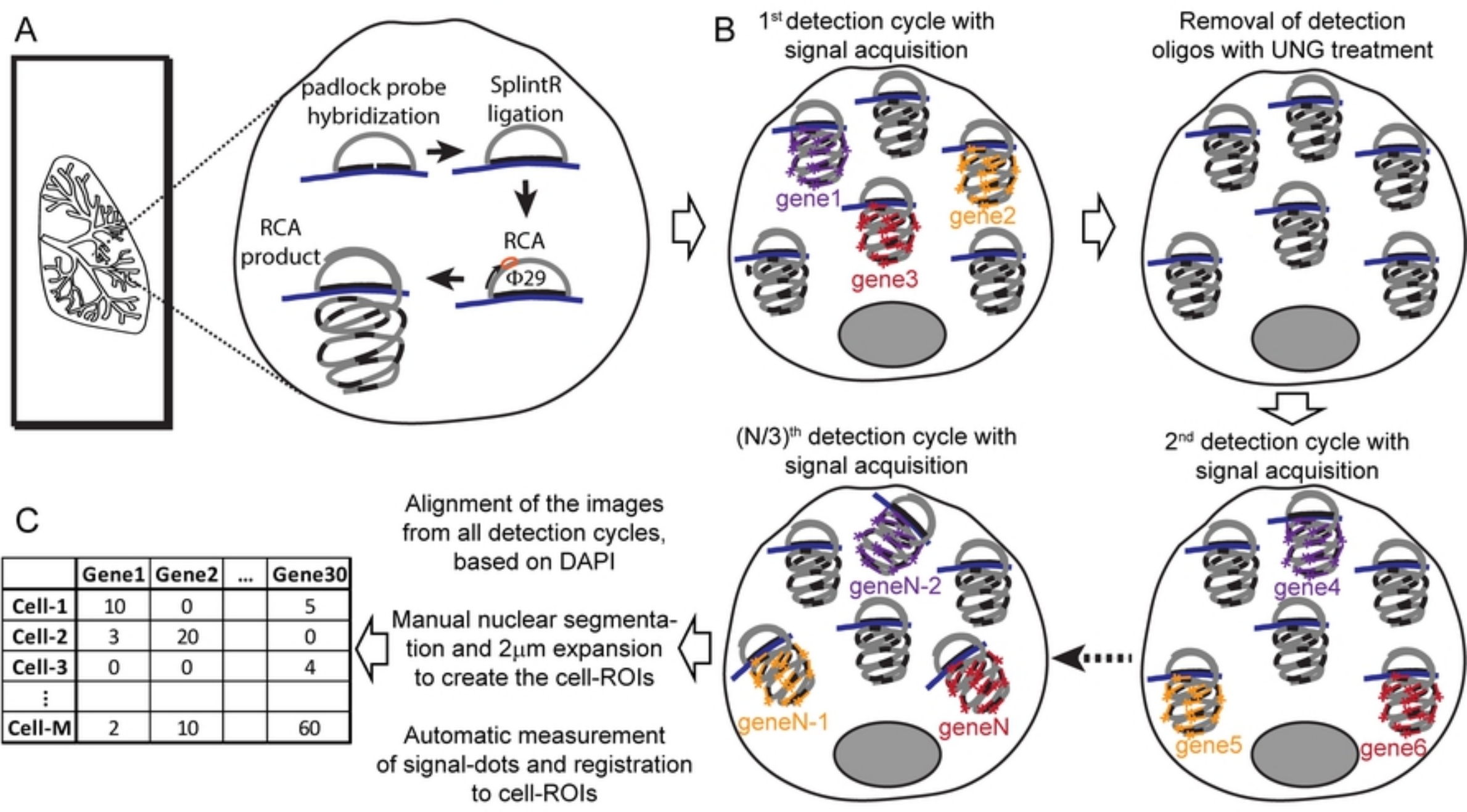
955 68. Mullner D. fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and
956 Python. J Stat Softw. 2013;53(9):1-18.

957 69. Kassambara A. ggpubr:"ggplot2" based publication ready plots. R package version 01.
958 2017;6.

959 70. Kolde R. Pheatmap: pretty heatmaps. 2015. R package version. 2017;1(10).

960 71. Hennig C. fpc: Flexible procedures for clustering. R package version 2.1-5. 2013.

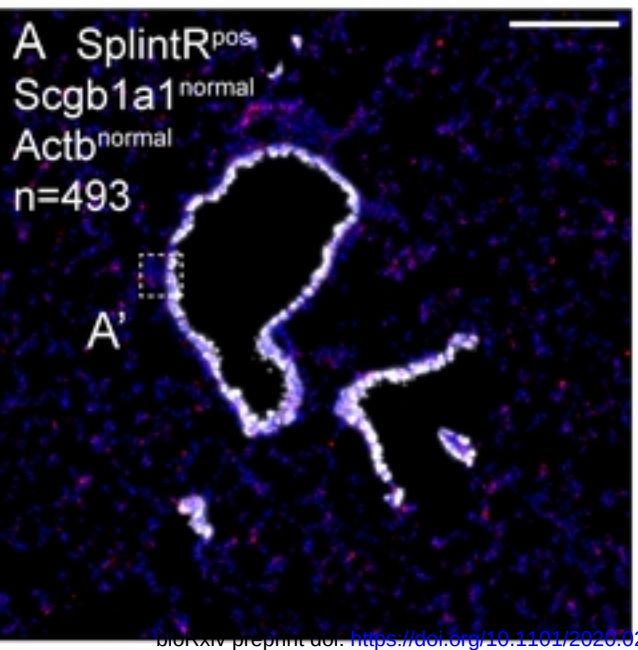
961

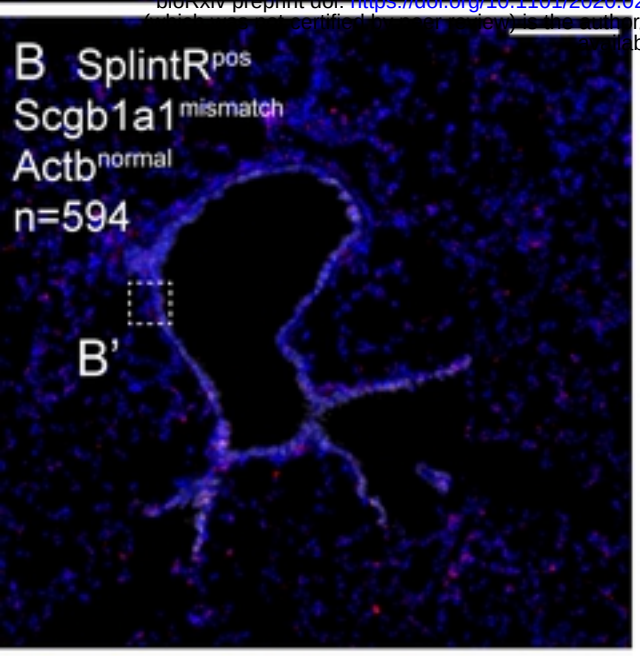


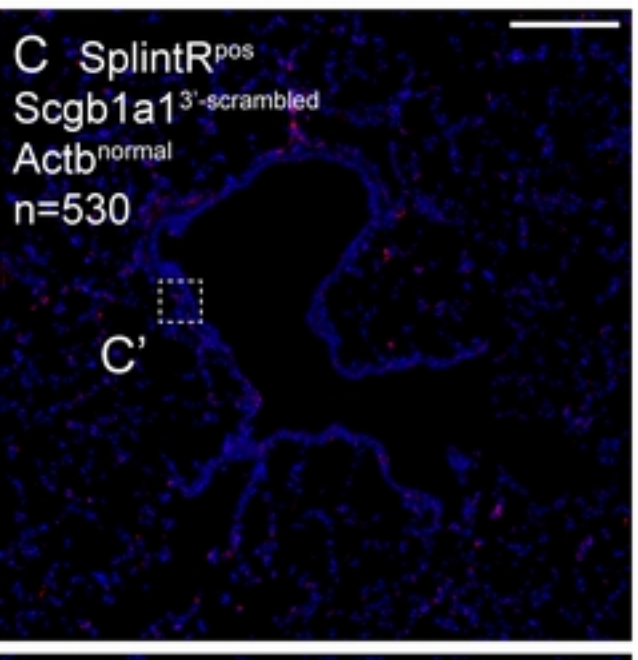
	Gene1	Gene2	...	Gene30
Cell-1	10	0		5
Cell-2	3	20		0
Cell-3	0	0		4
⋮				
Cell-M	2	10		60

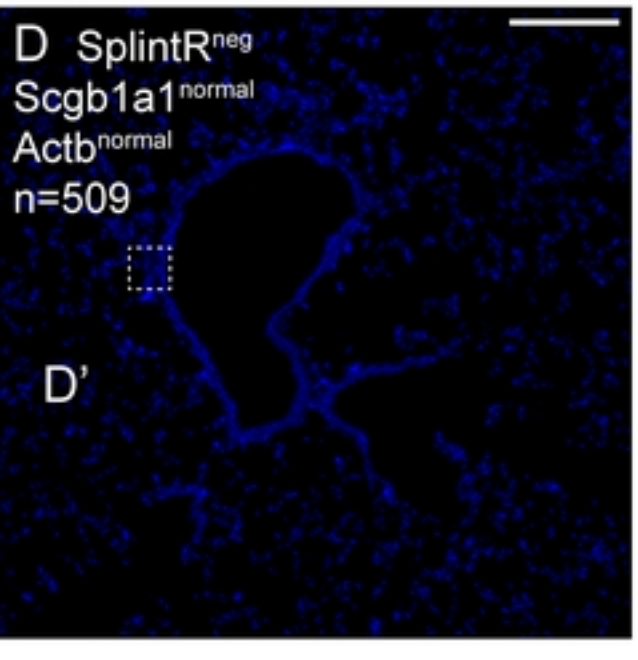
Figure_1

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted Febru... 0...
...certified by peer review) is the author/funder, who has granted bioRxiv a lice... di...
...ilable under aCC-BY 4.0 International license



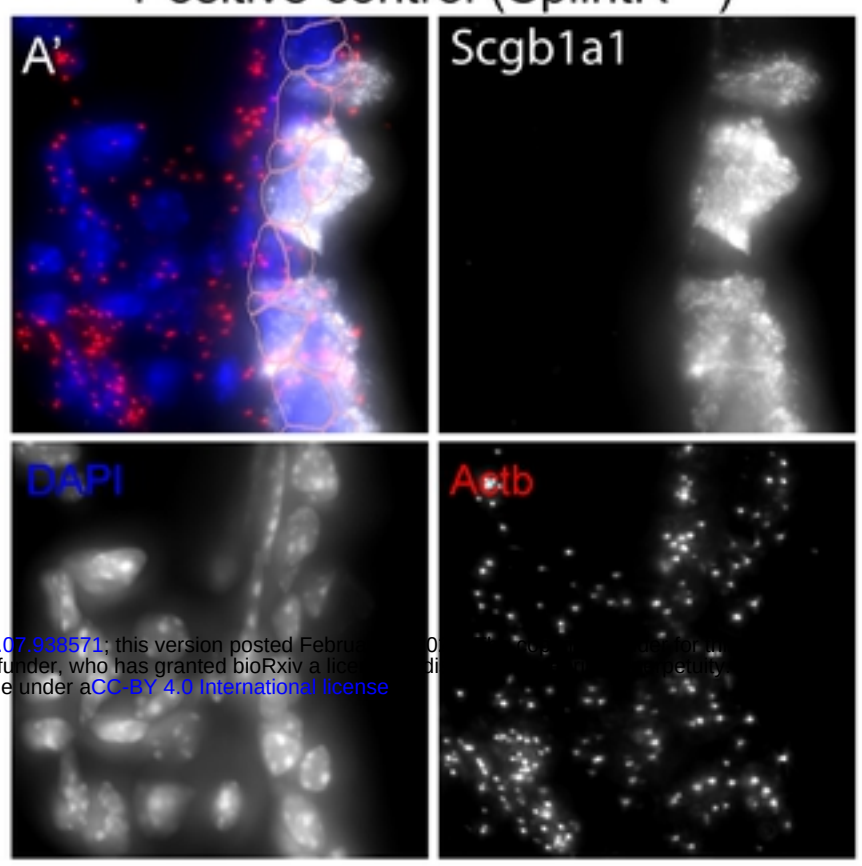


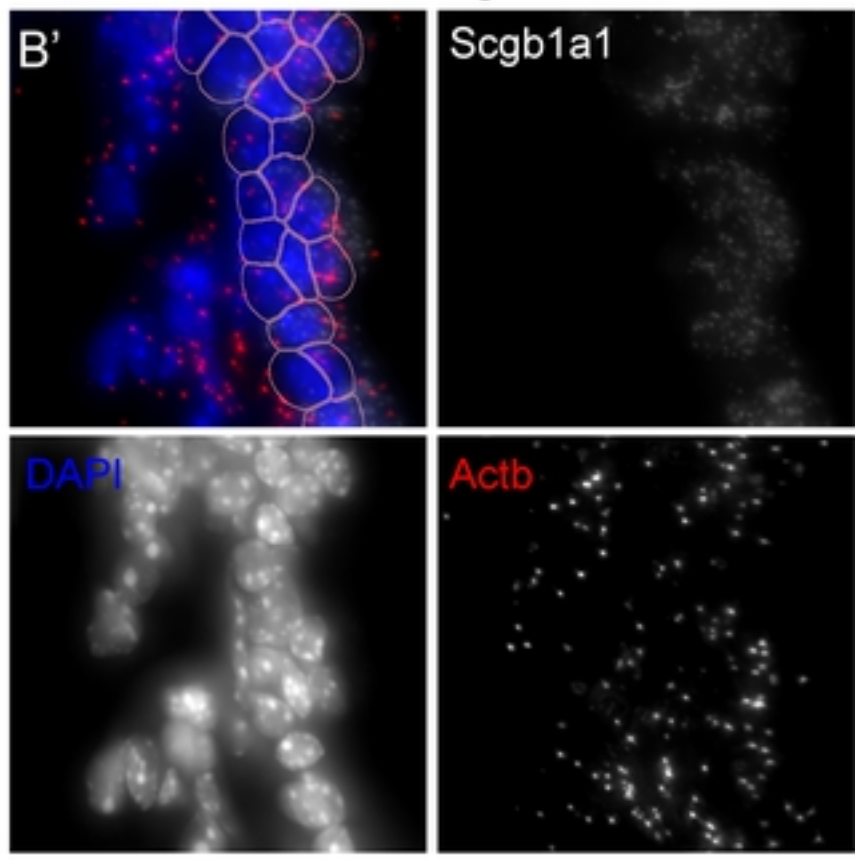




Positive control (SplintR^pos)

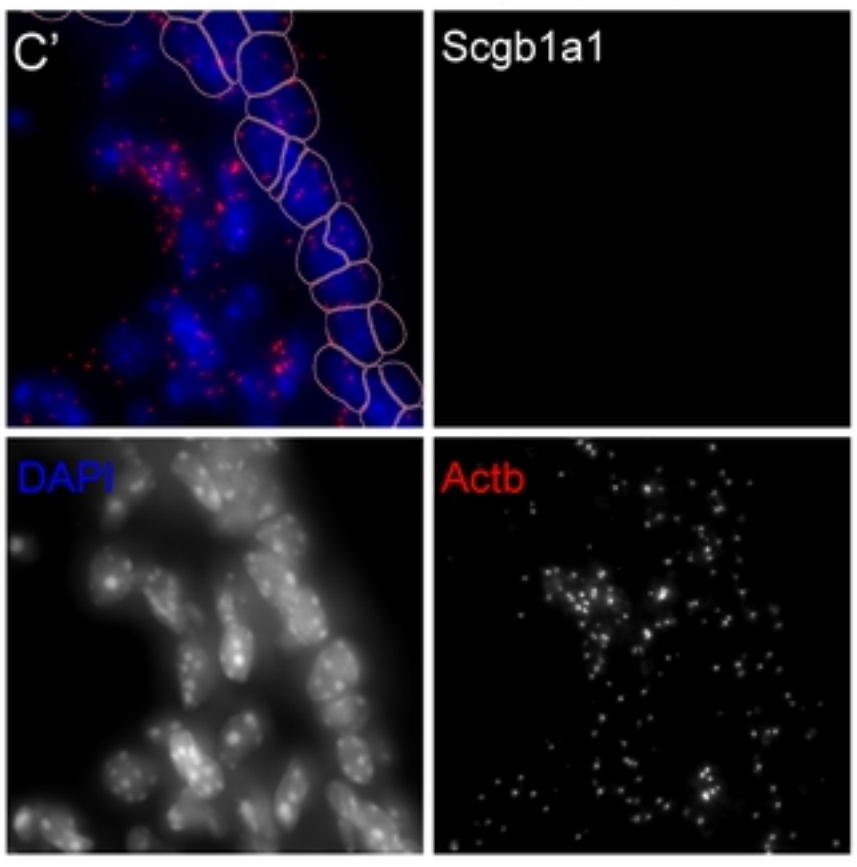
Mismatch at ligation site

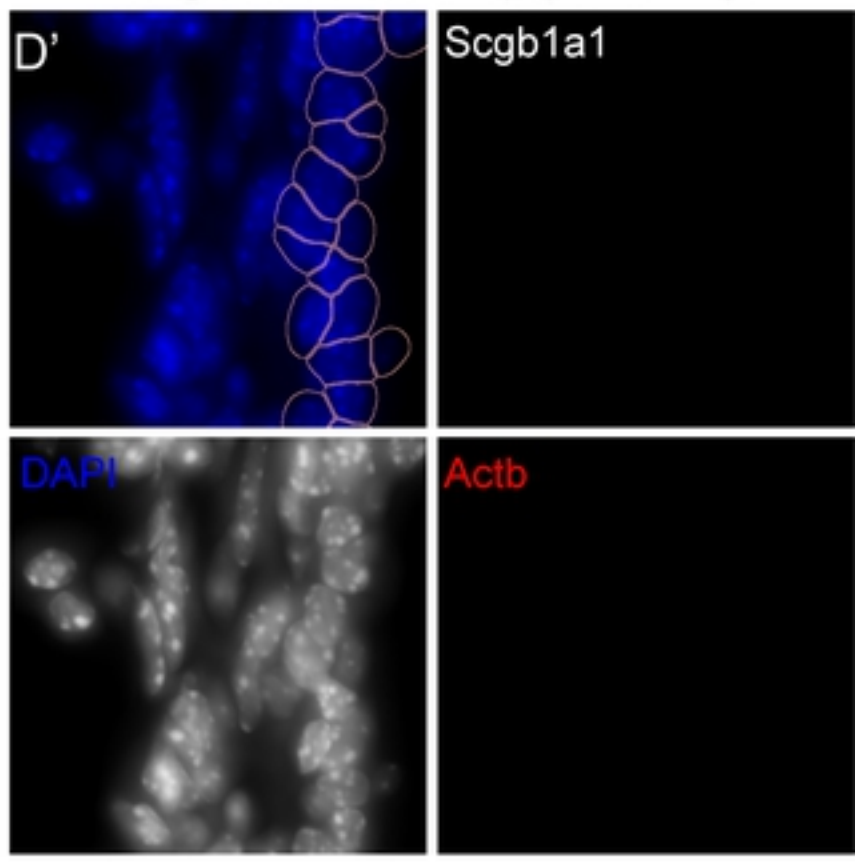


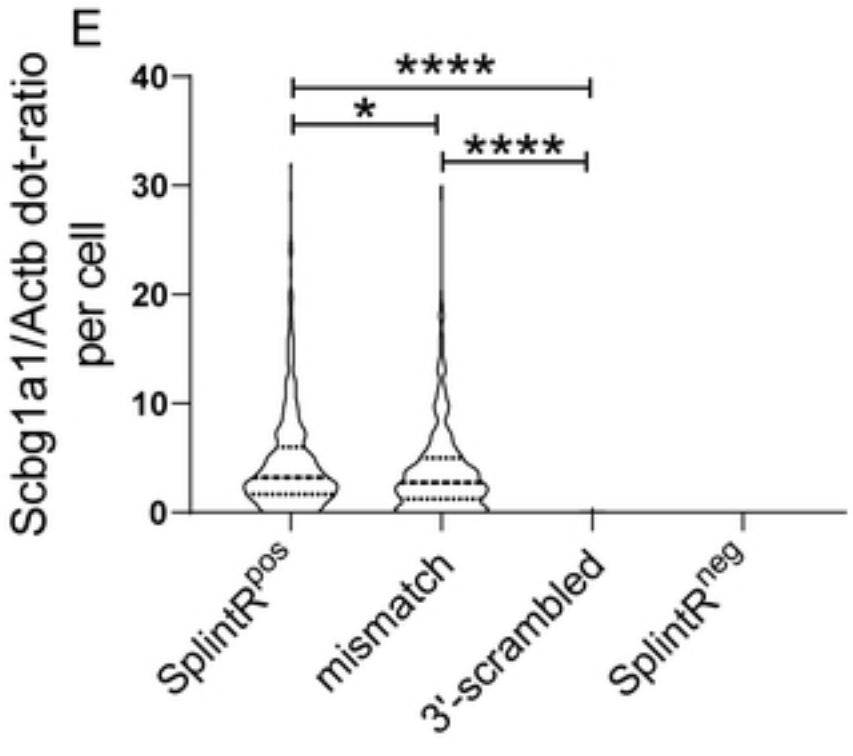


3'-scrambled padlock

Negative control (SplintR^neg)







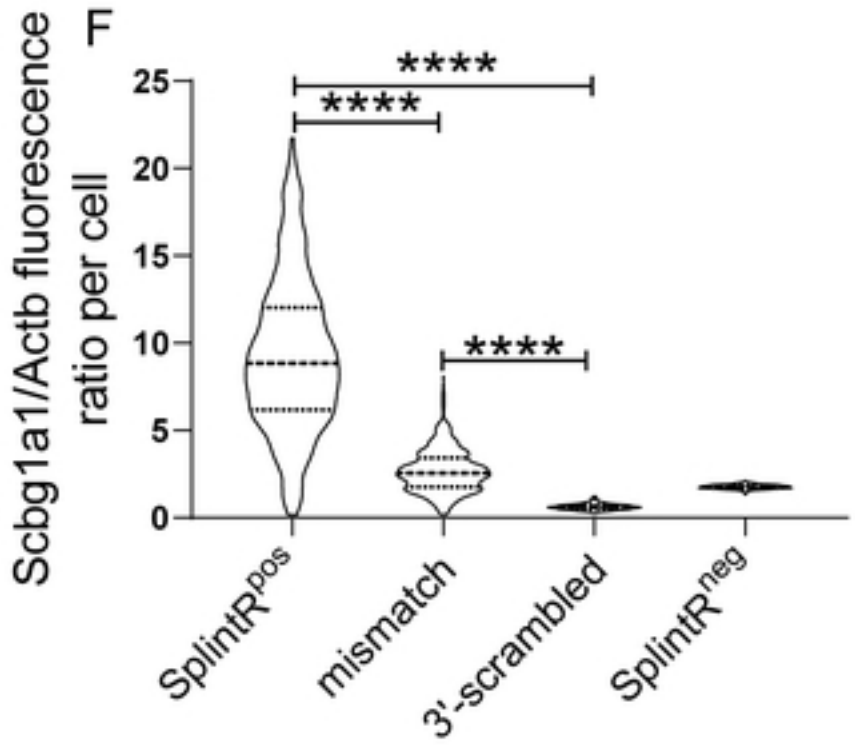
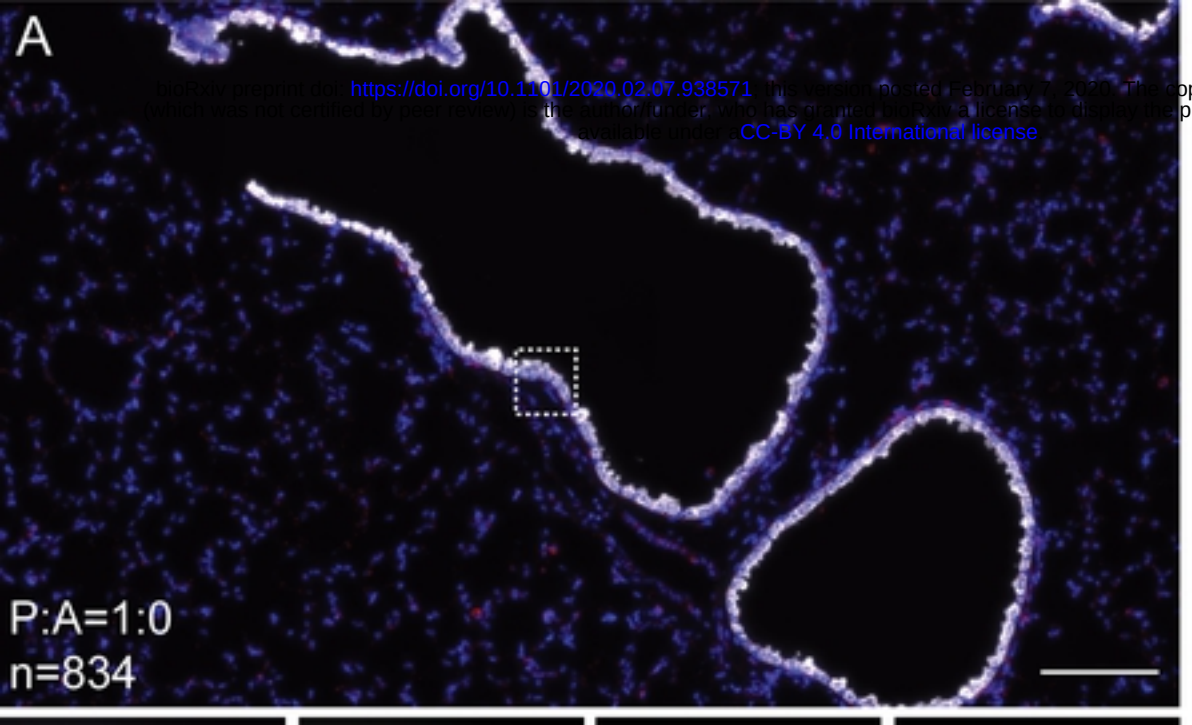
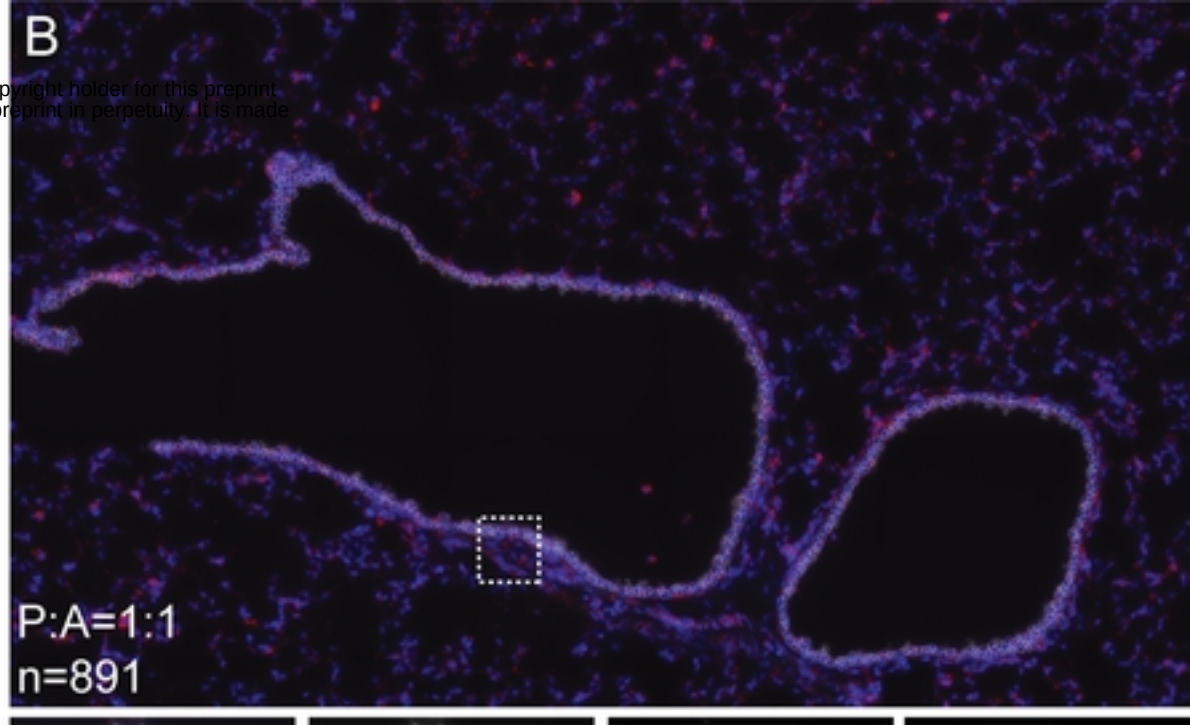
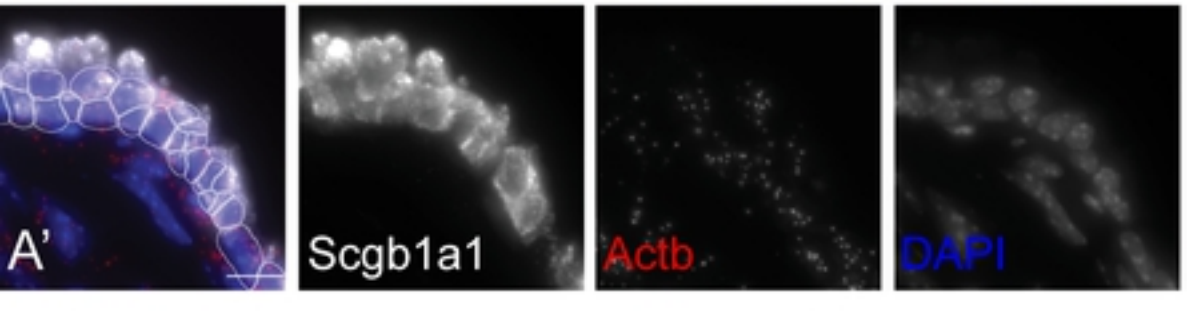


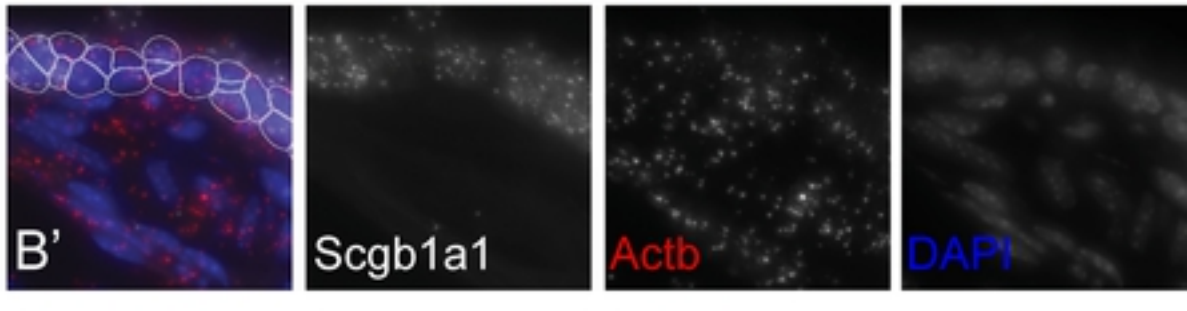
Figure_2

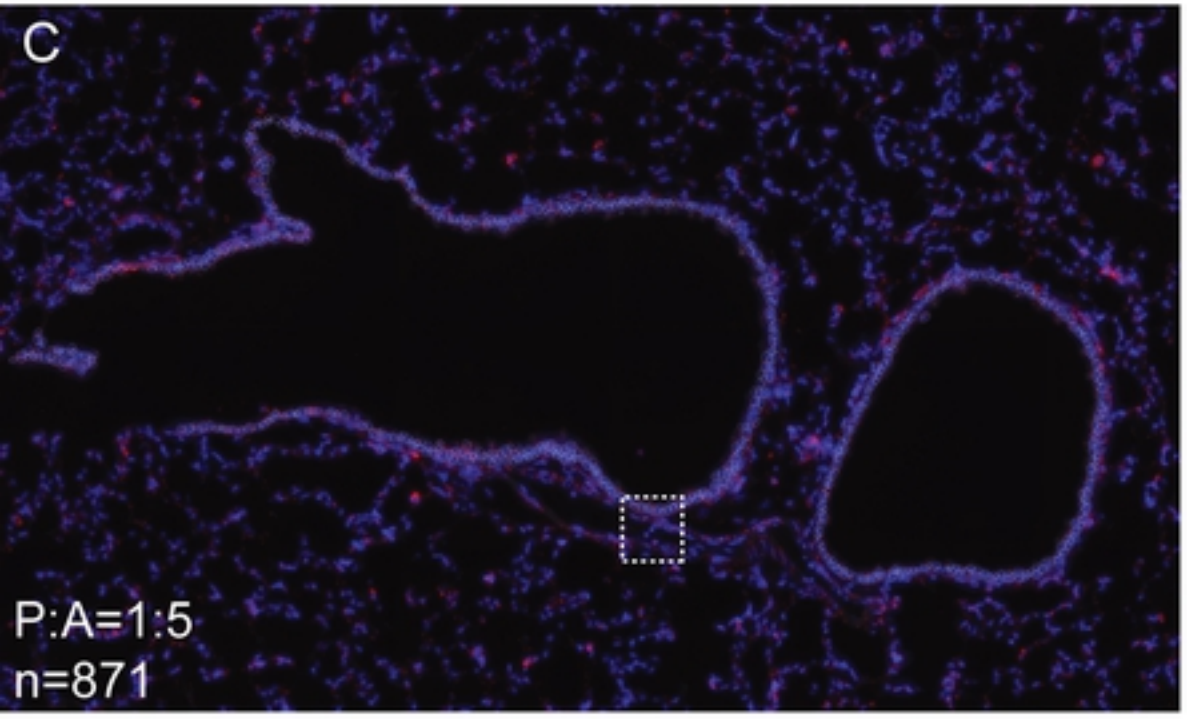
bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571 this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under CC-BY 4.0 International license

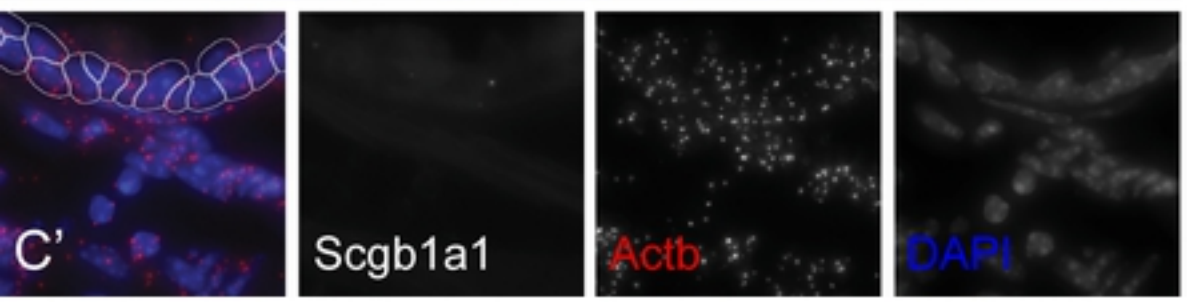












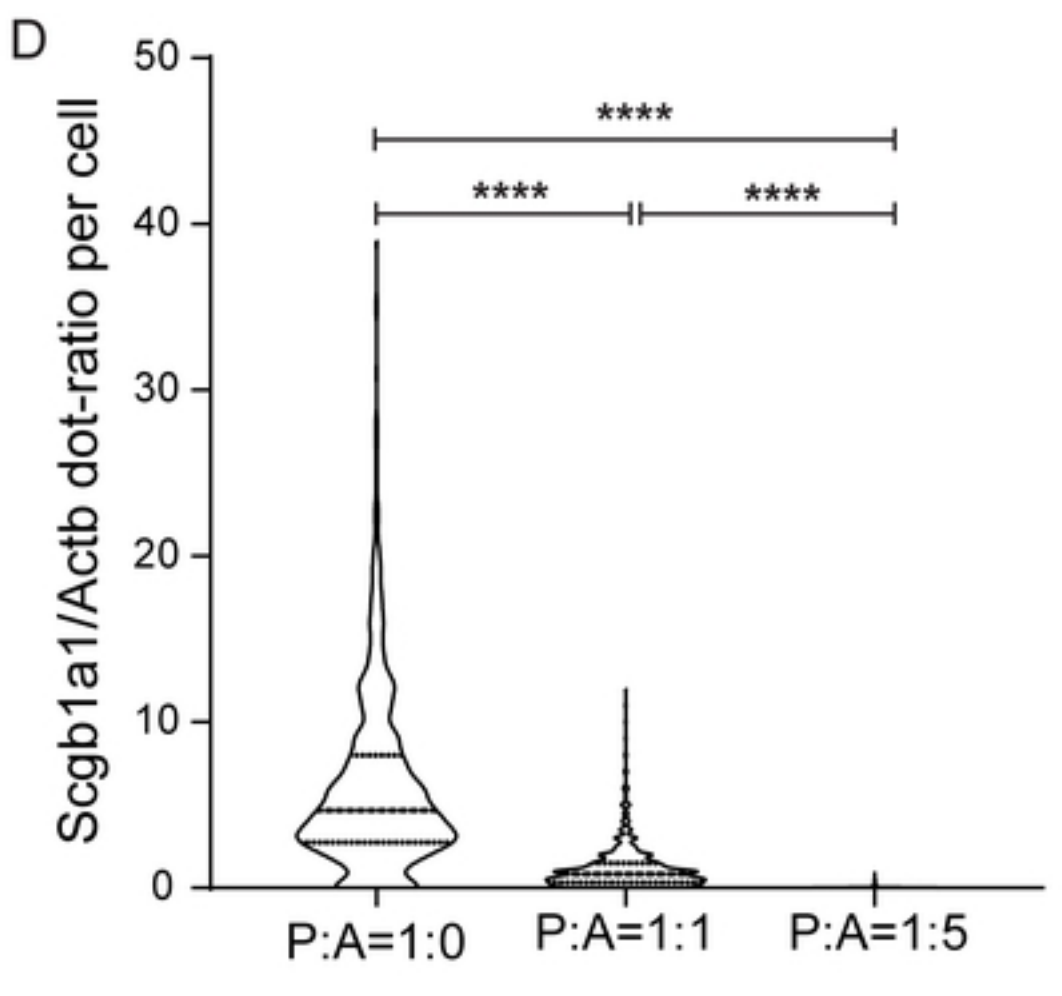
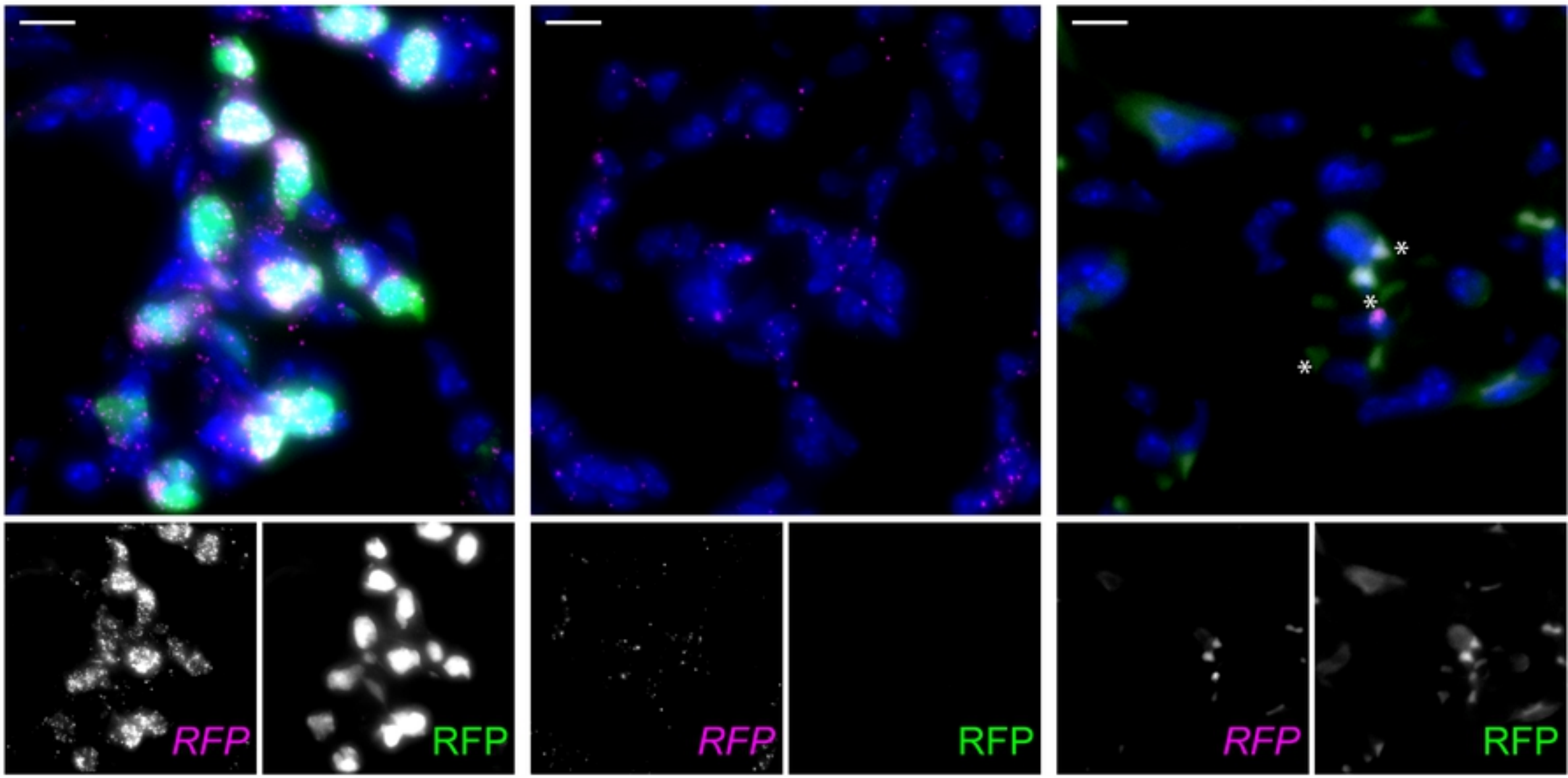


Figure_3

A Sftpc-CreER^{pos}; Rosa-Ai14^{pos} Sftpc-CreER^{neg}; Rosa-Ai14^{pos} wildtype



B Sftpc-CreER^{pos}; Rosa-Ai14^{pos} Sftpc-CreER^{neg}; Rosa-Ai14^{pos} wildtype

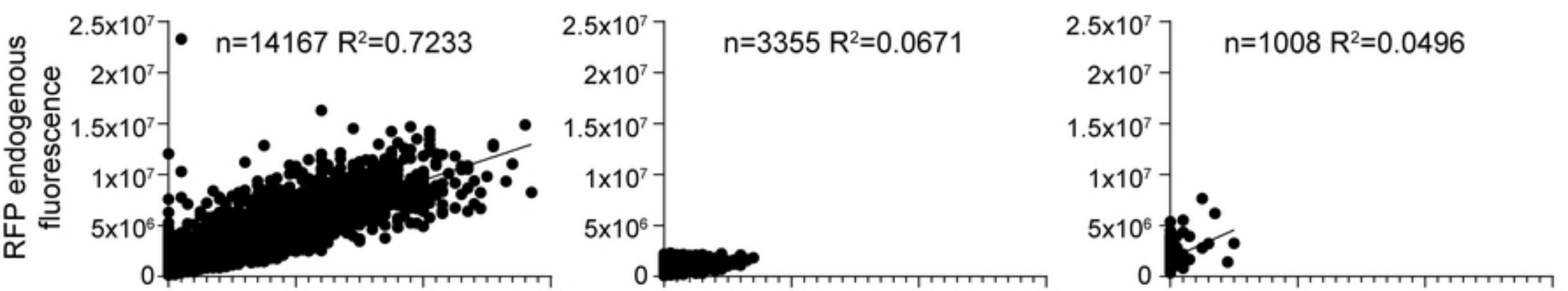


Figure_4

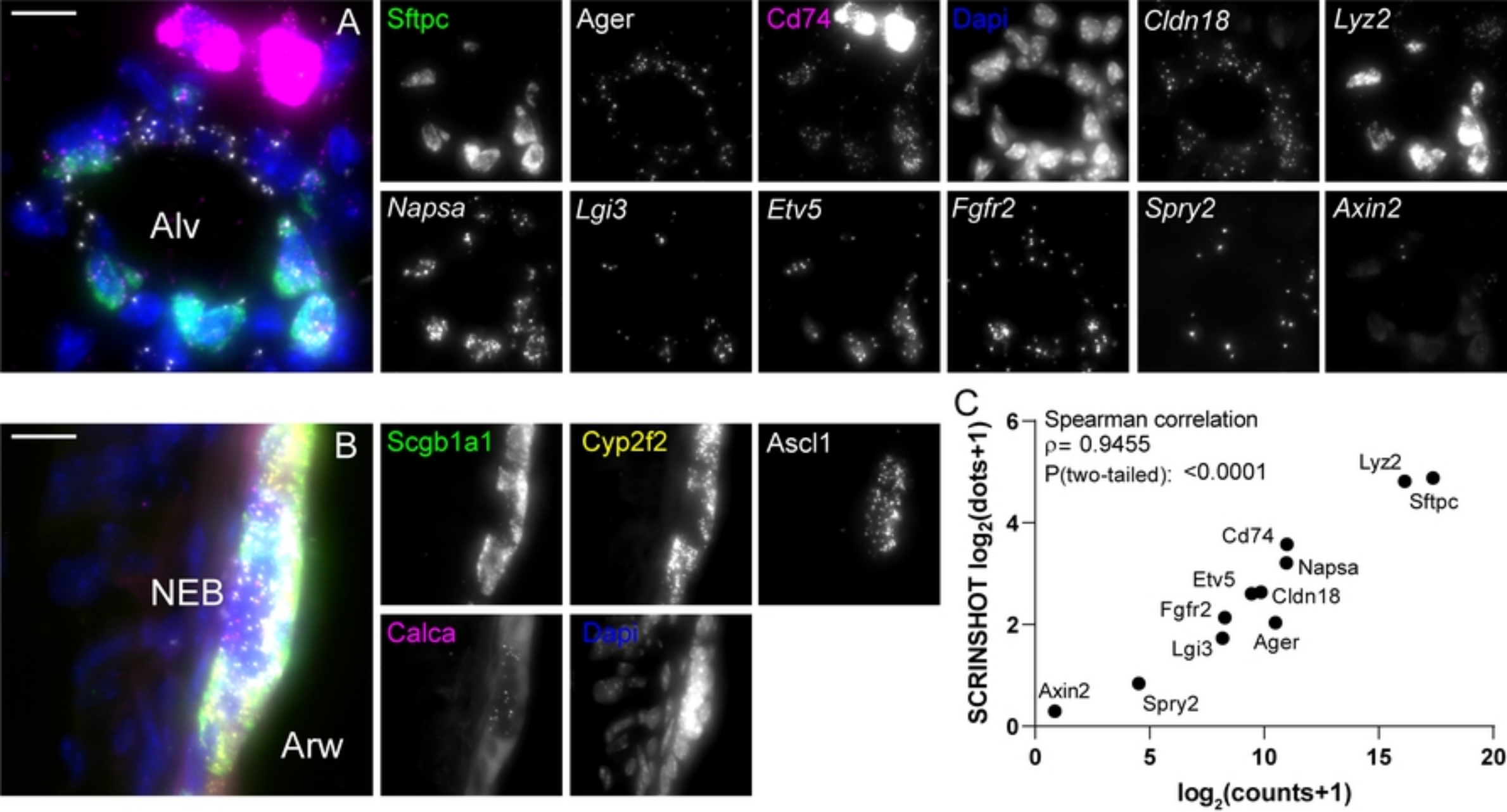
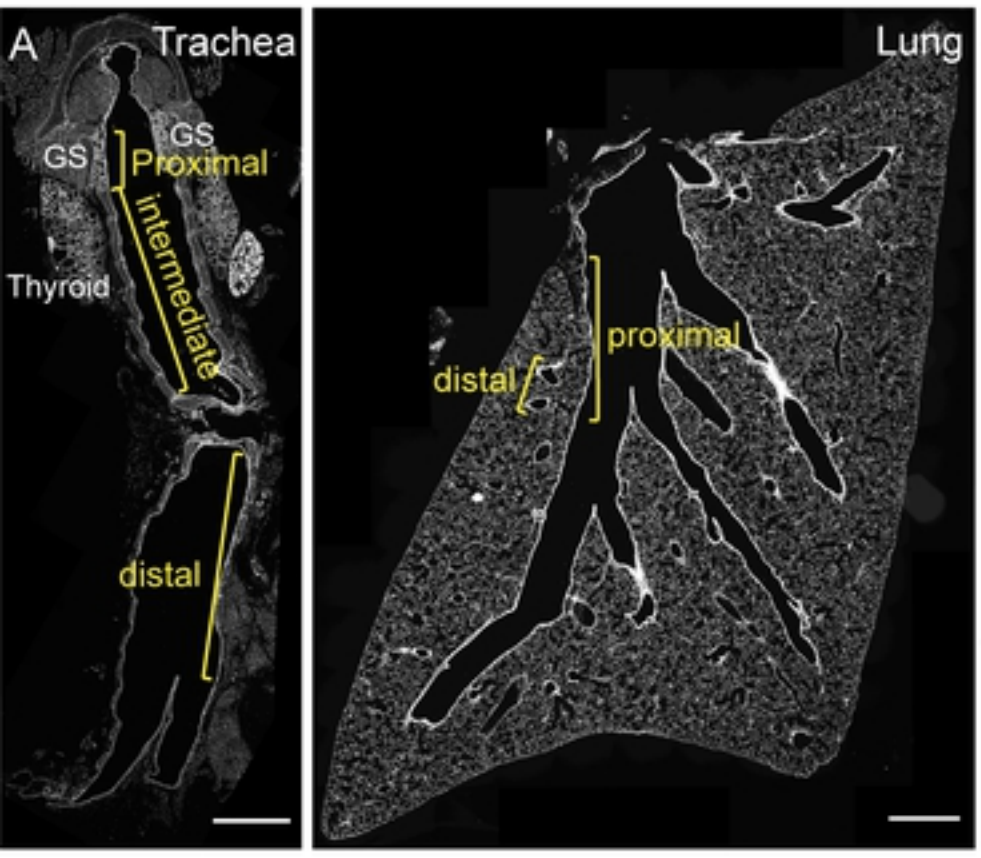
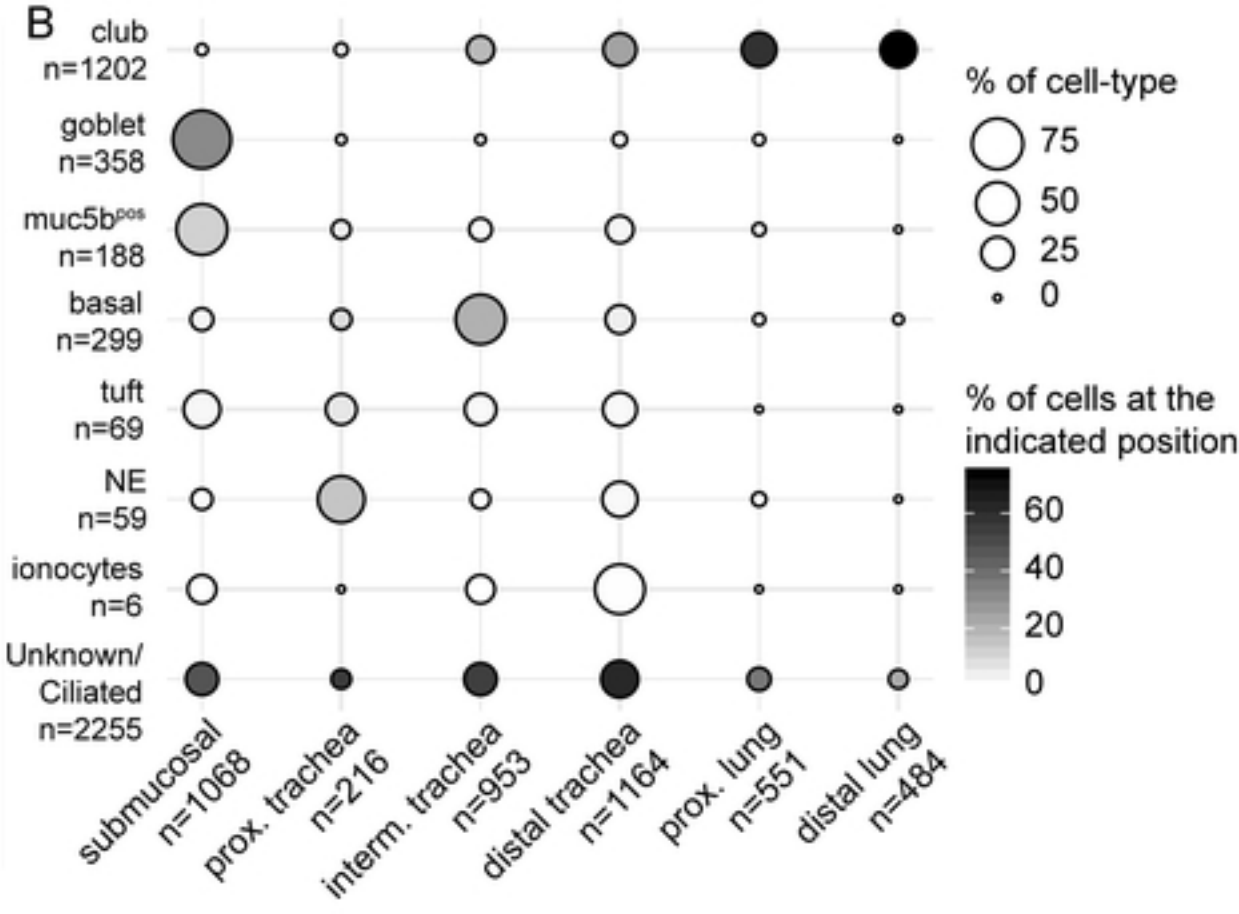
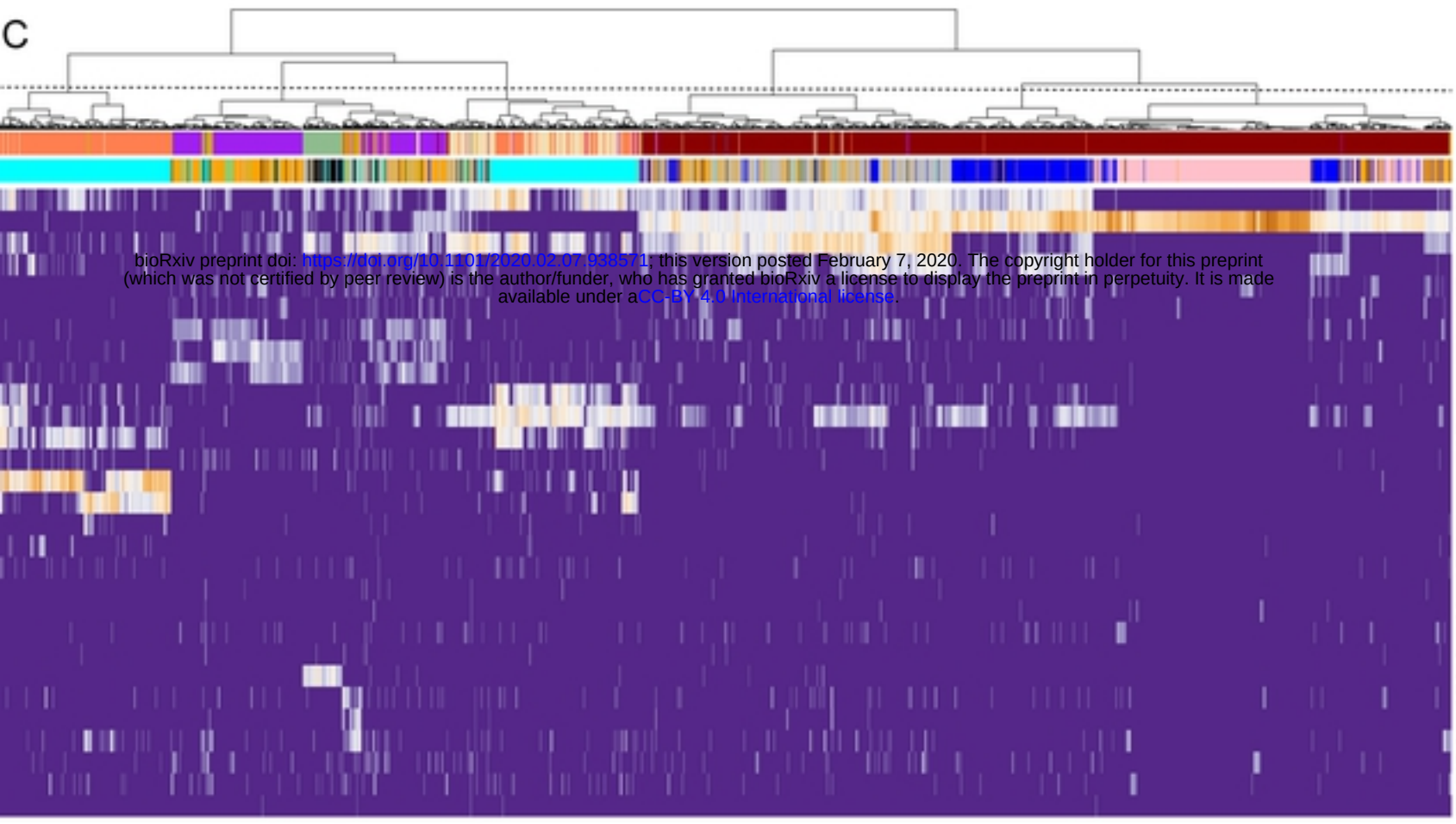

A
Alv
Sftpc
Ager
Cd74
Dapi
Cldn18
Lyz2
Napsa
Lgi3
Etv5
Fgfr2
Spry2
Axin2
B
NEB
Arw
Scgb1a1
Cyp2f2
Ascl1
Calca
Dapi
C
Spearman correlation
ρ= 0.9455
P(two-tailed): <0.0001
SCRINSHOT log2(dots+1)
log2(counts+1)
0
2
4
6
0
5
10
15
20
Lyz2
Sftpc
Cd74
Napsa
Etv5
Cldn18
Fgfr2
Lgi3
Ager
Axin2
Spry2

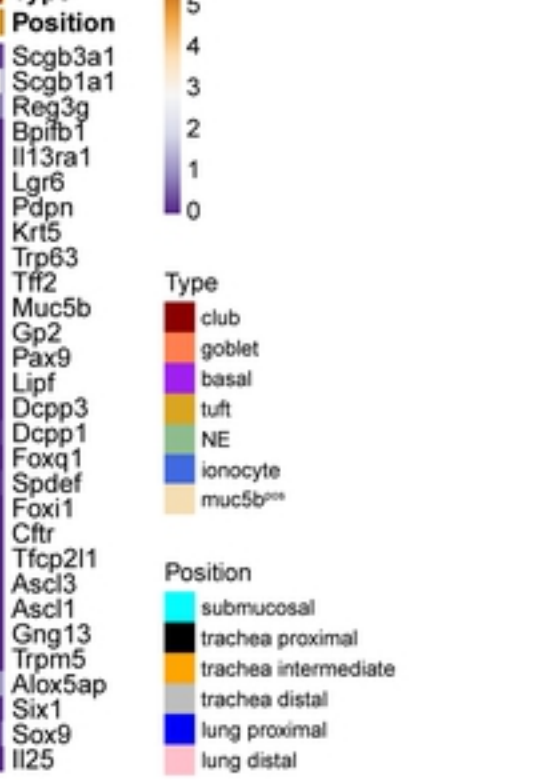
Figure_6

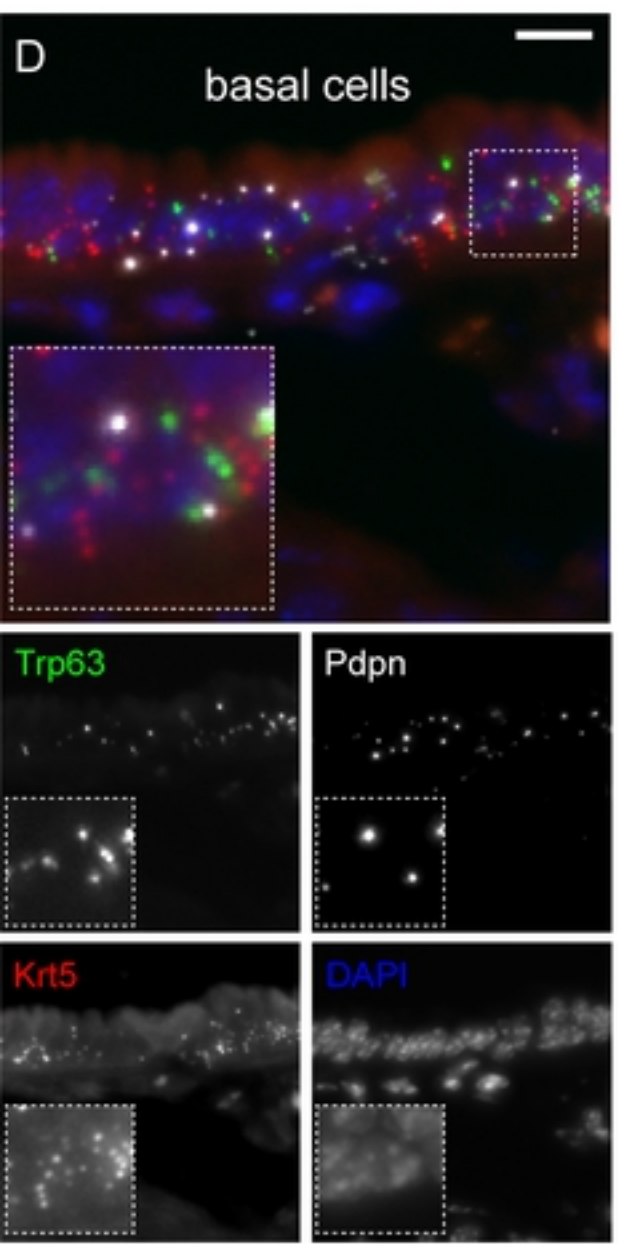


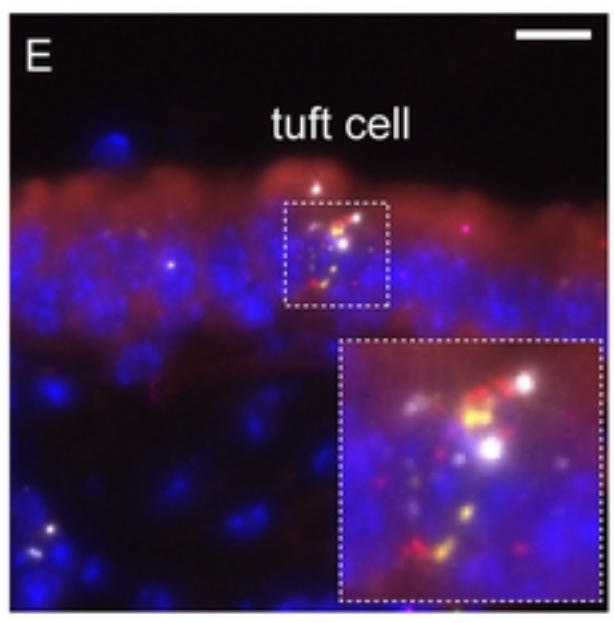


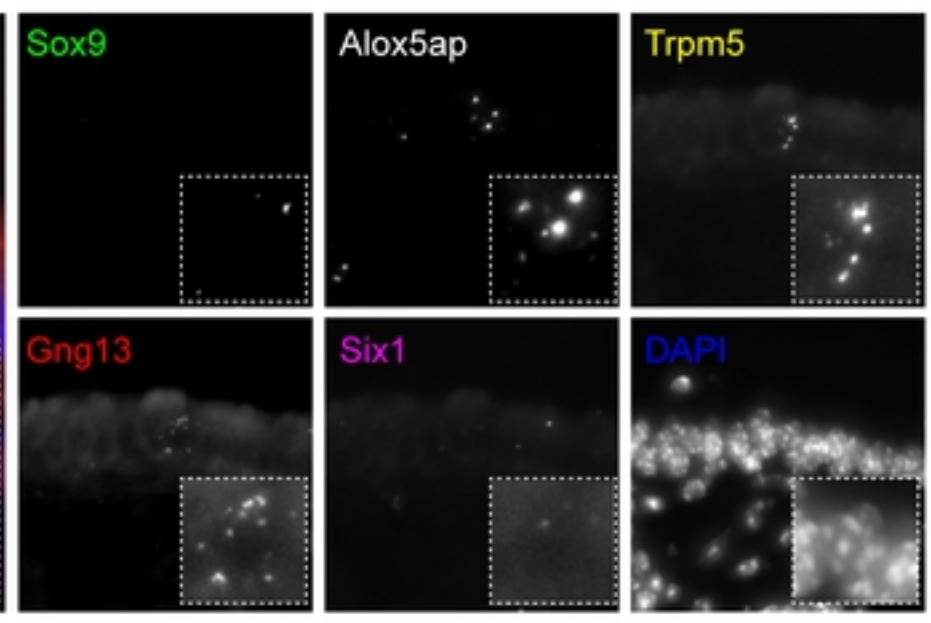


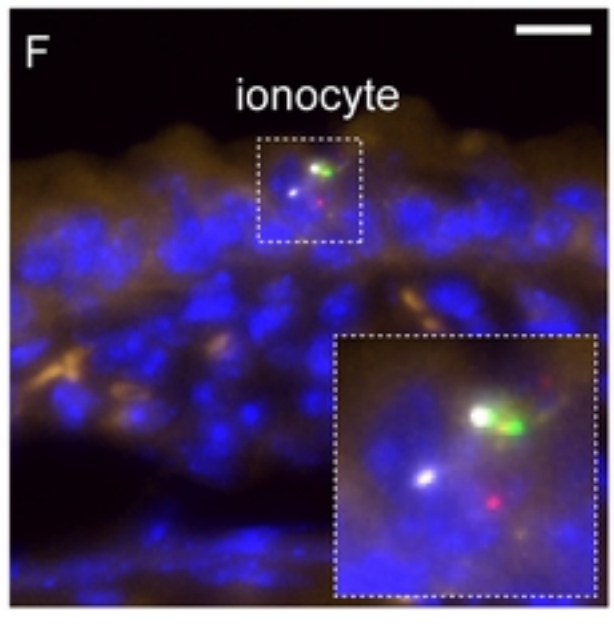
bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.











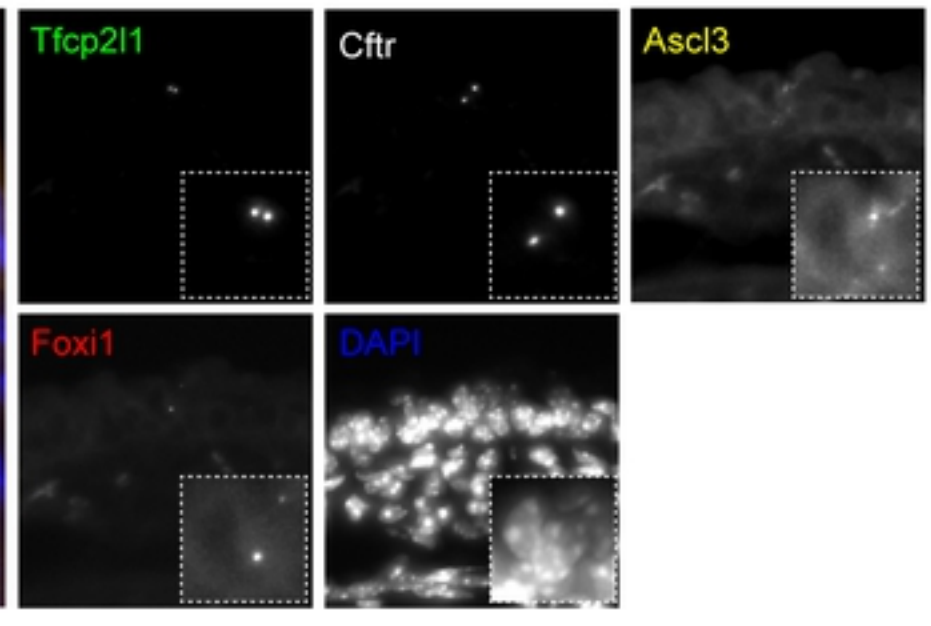
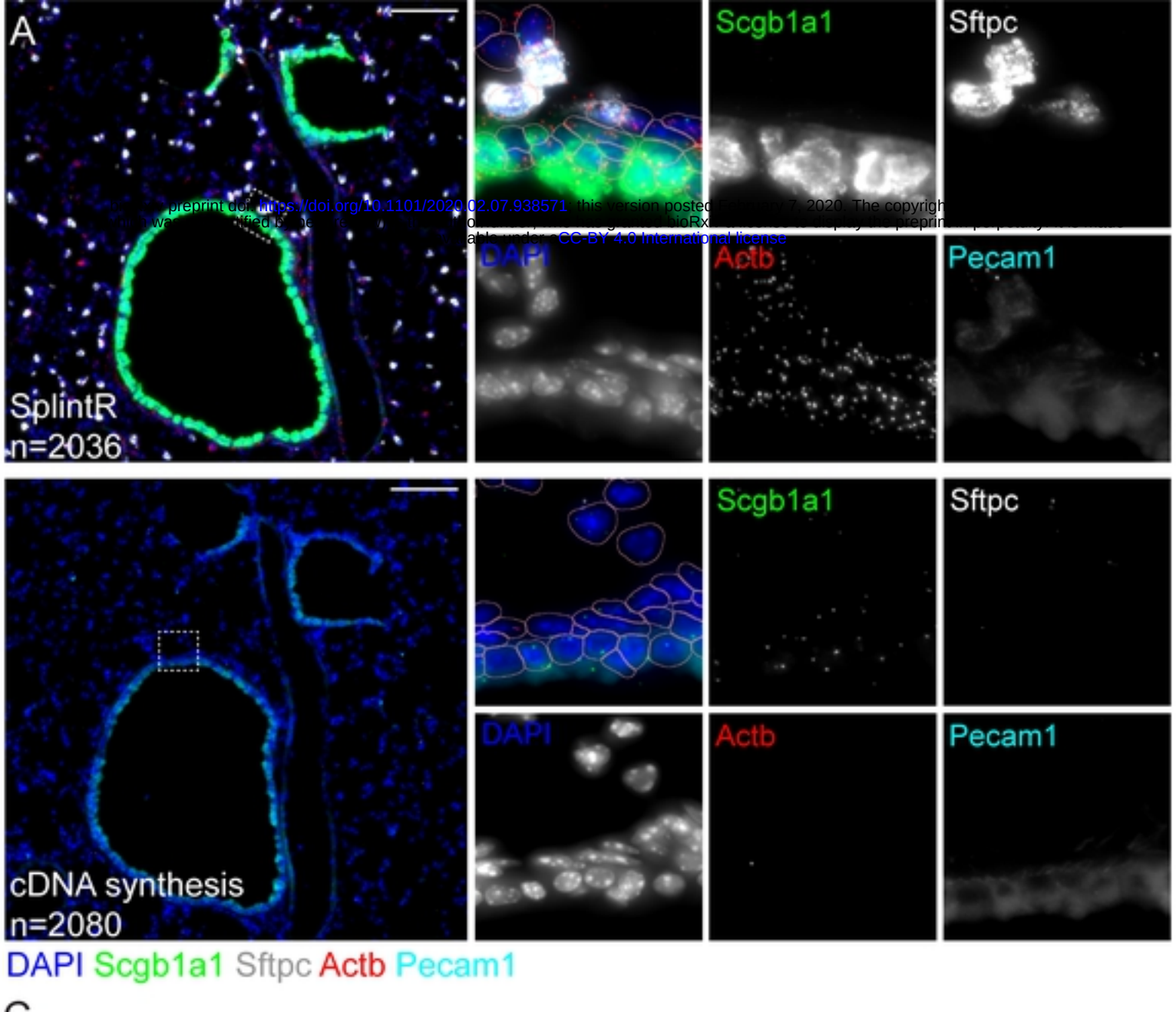
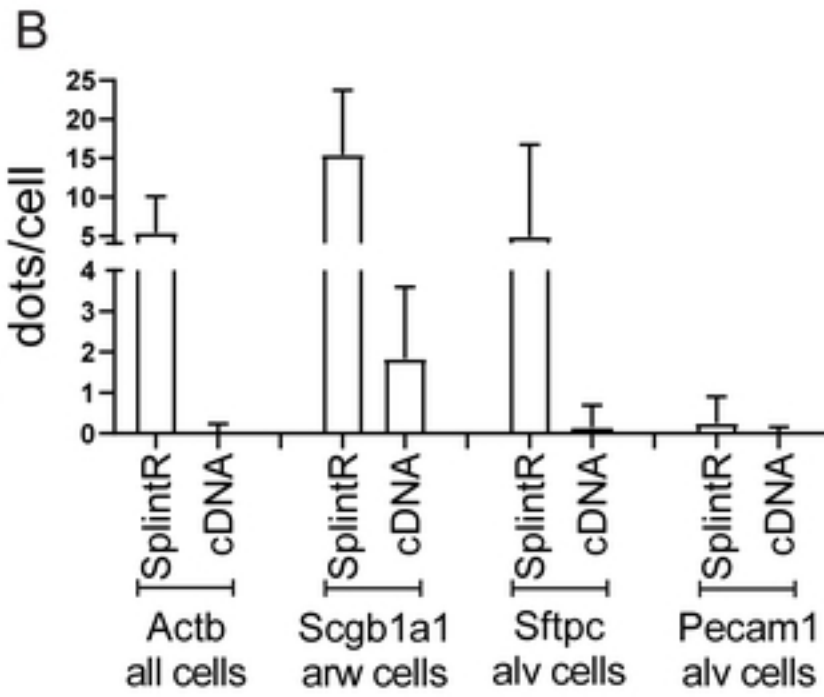
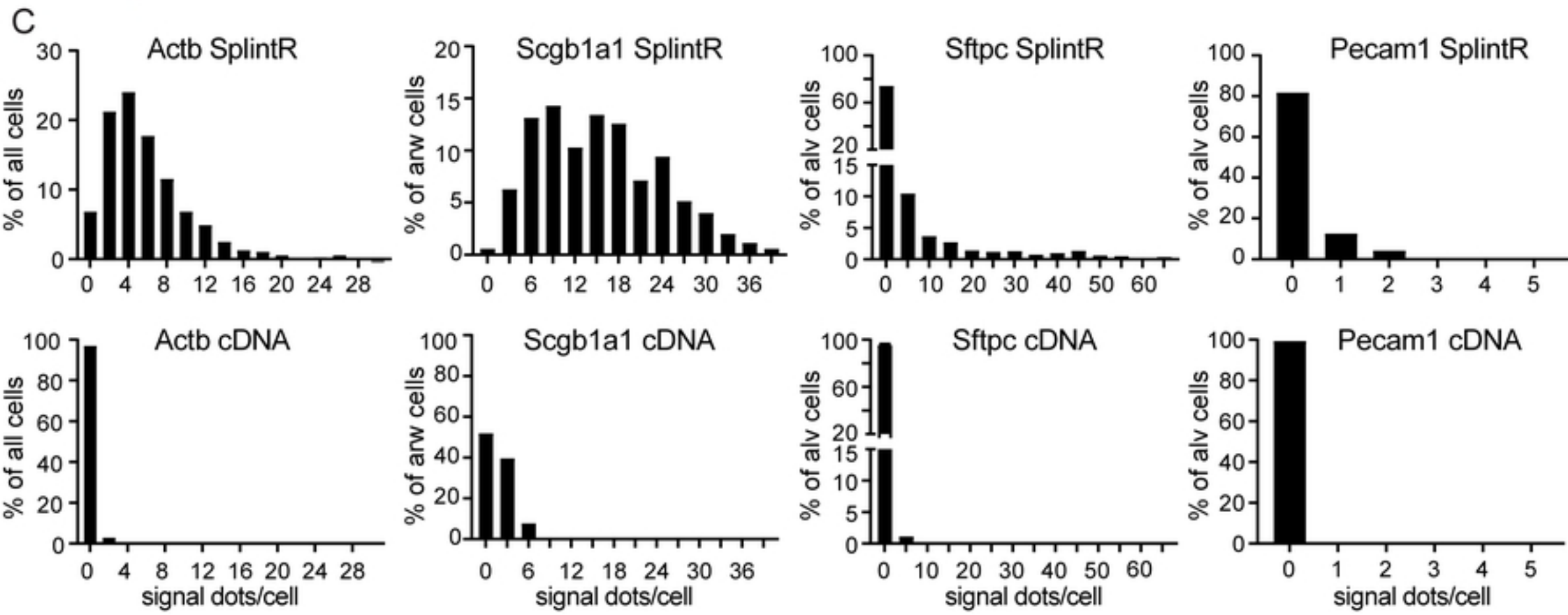


Figure_7

preprint doi: https://doi.org/10.1101/2020.02.07.938571 this version posted February 7, 2020. The copyright holder was certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under CC-BY 4.0 International license



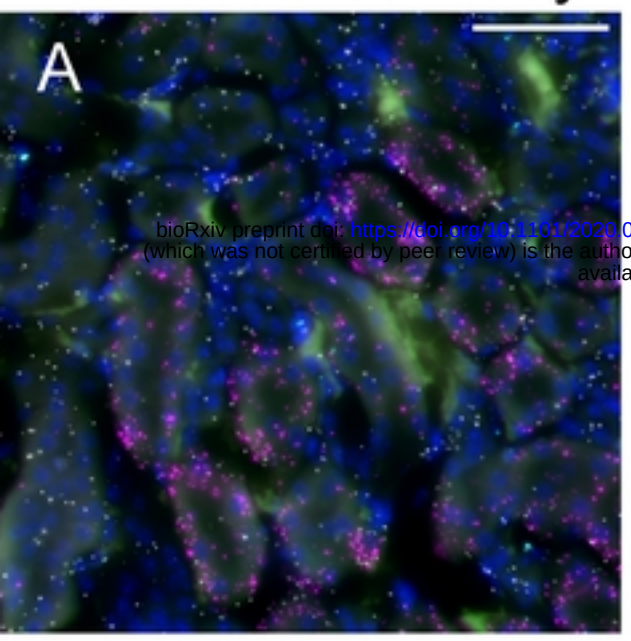


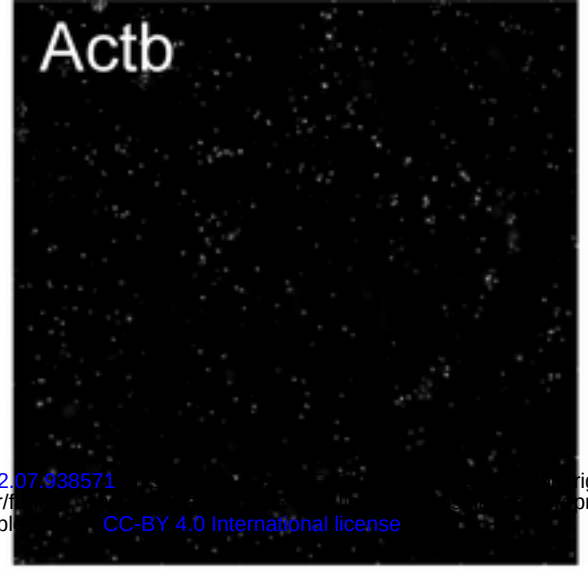


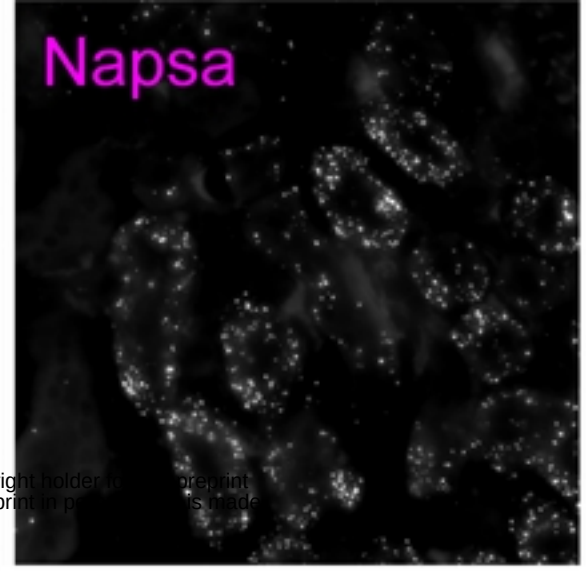
Suppl_Figure_1

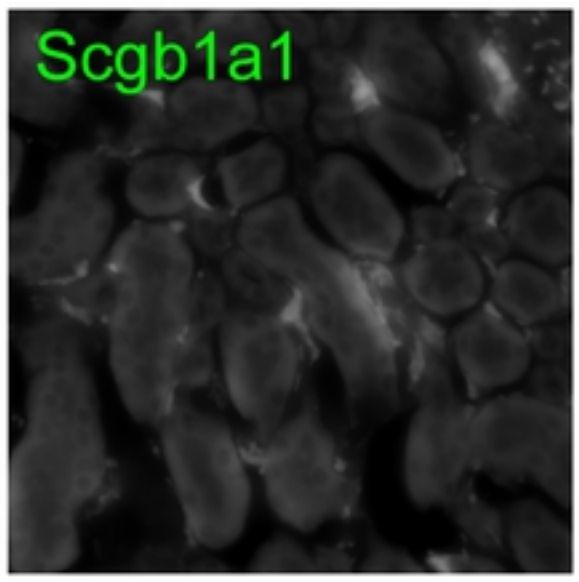
bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.038571
(which was not certified by peer review) is the author/f
availabl
right holder
preprint
print in pe
is made
CC-BY 4.0 International license

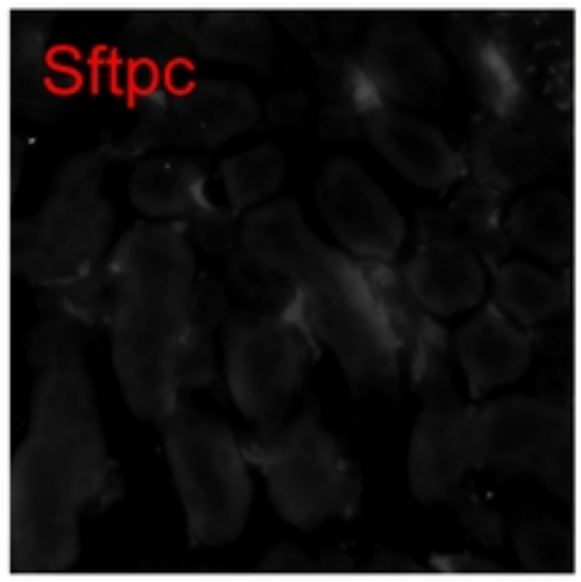
P60 mouse kidney

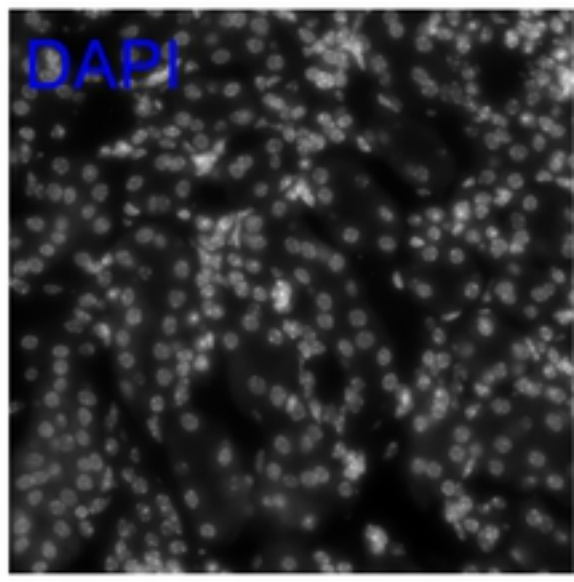




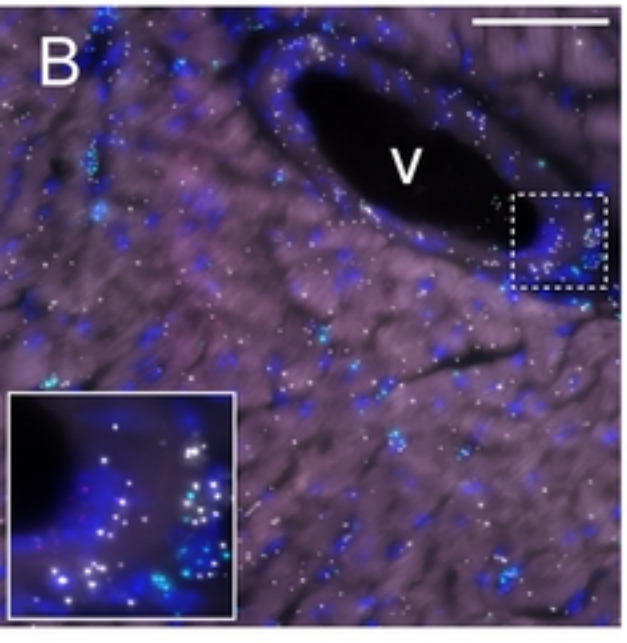


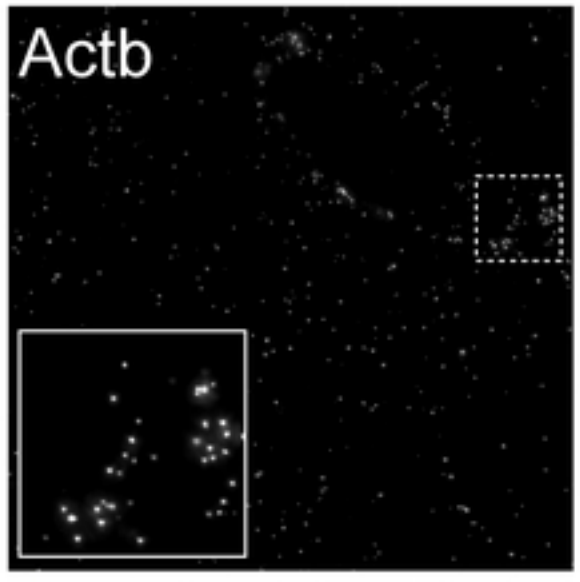


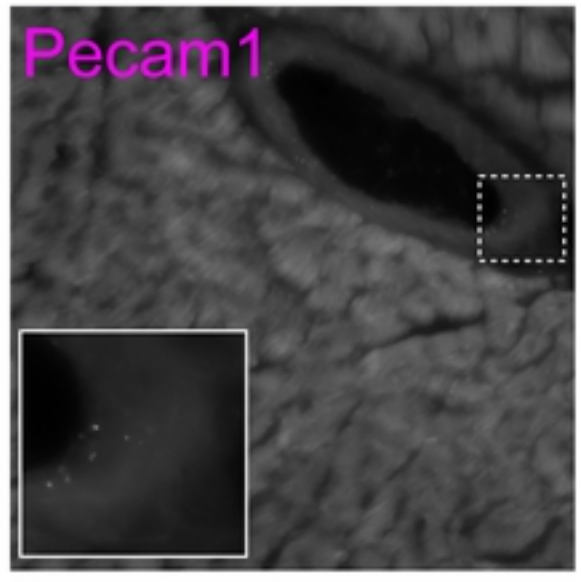


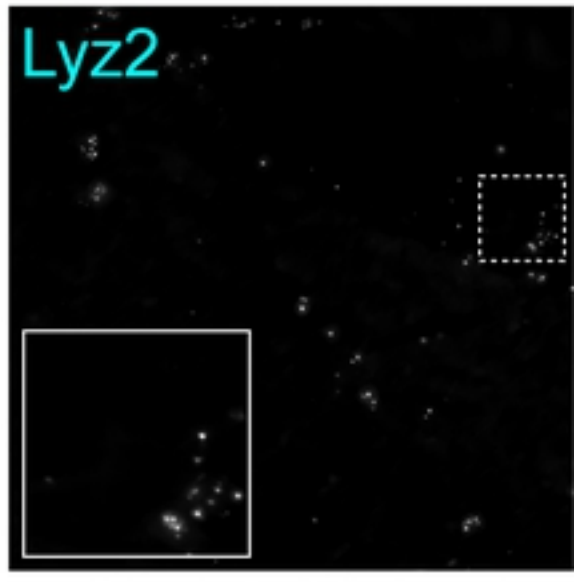


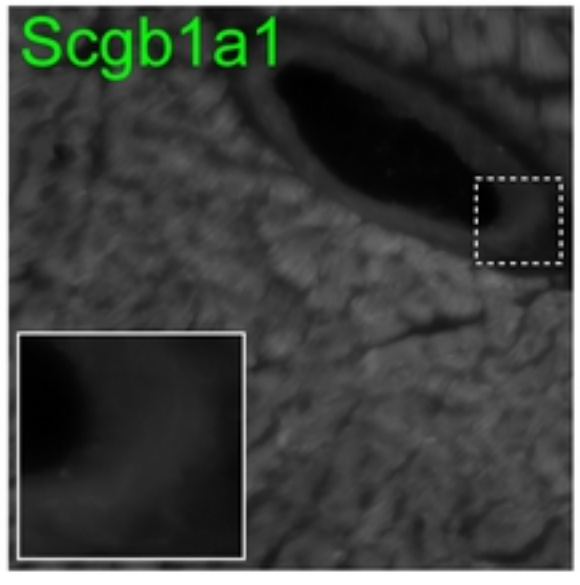
P60 mouse heart

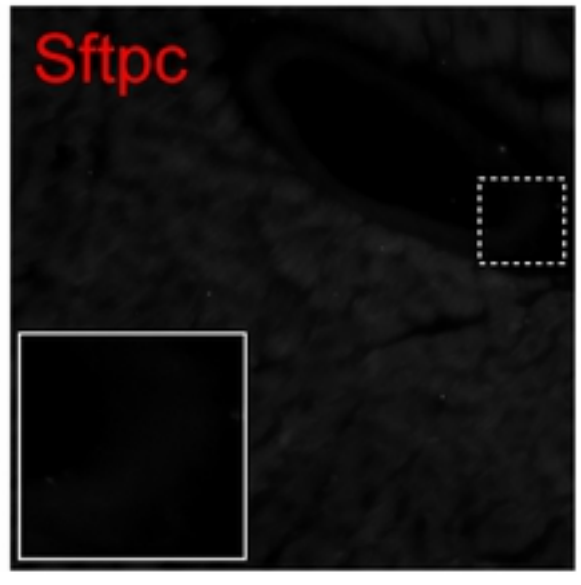


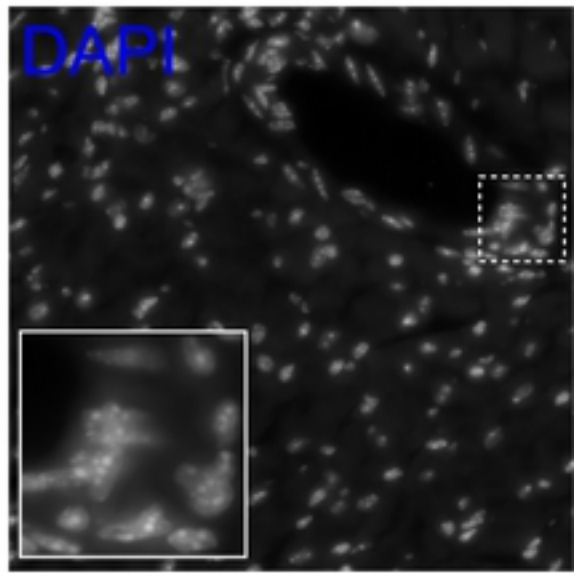




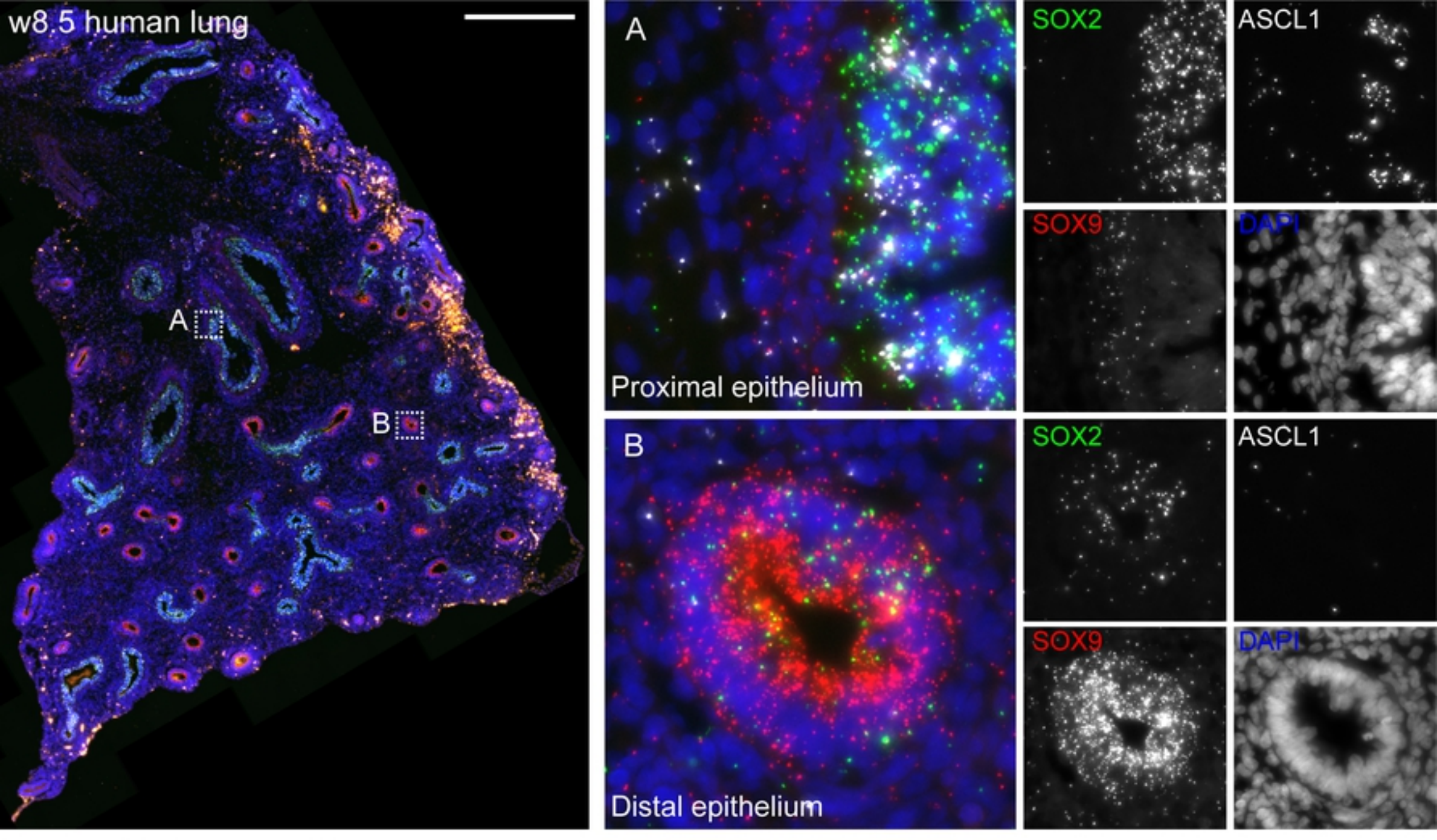




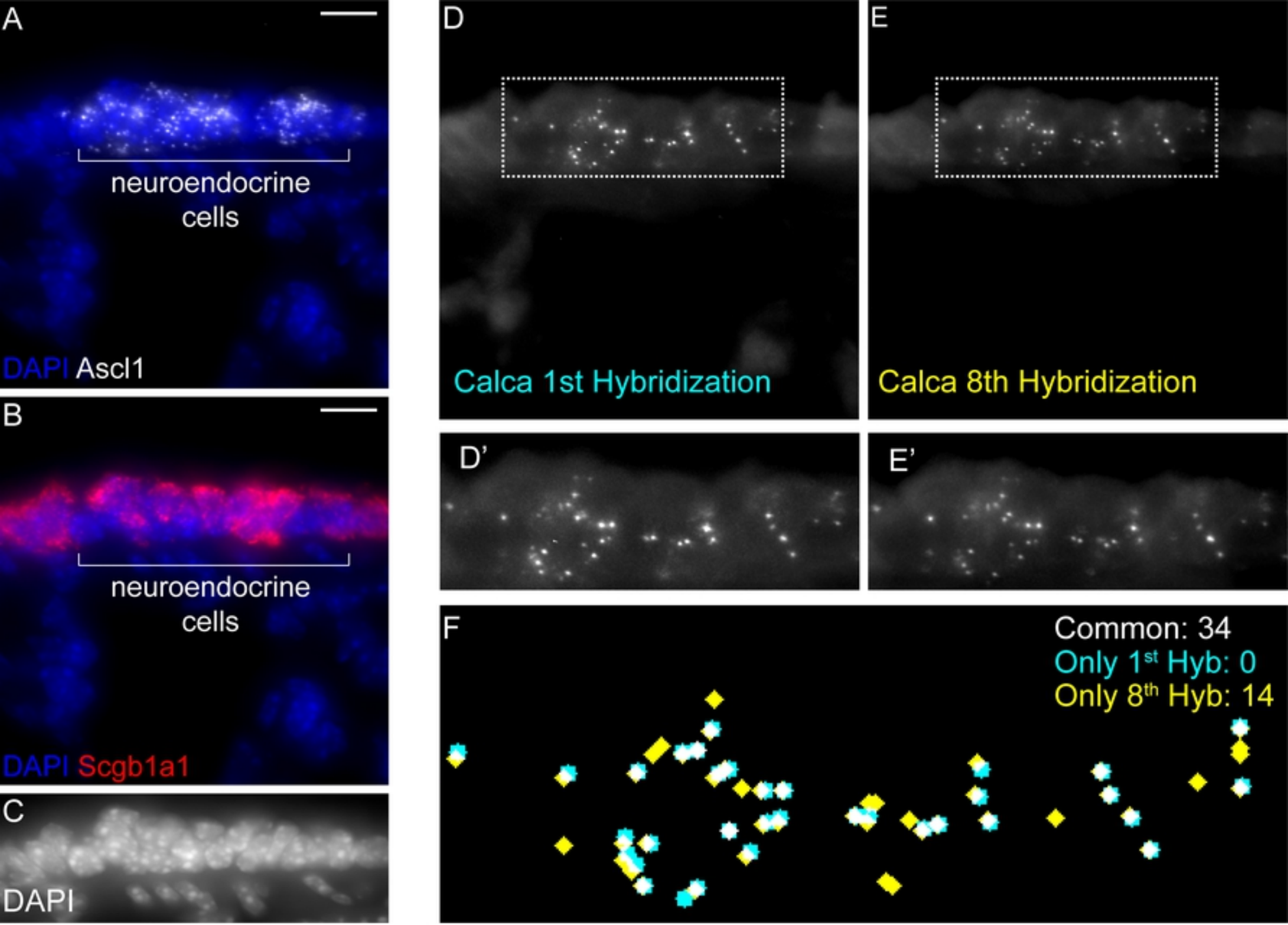




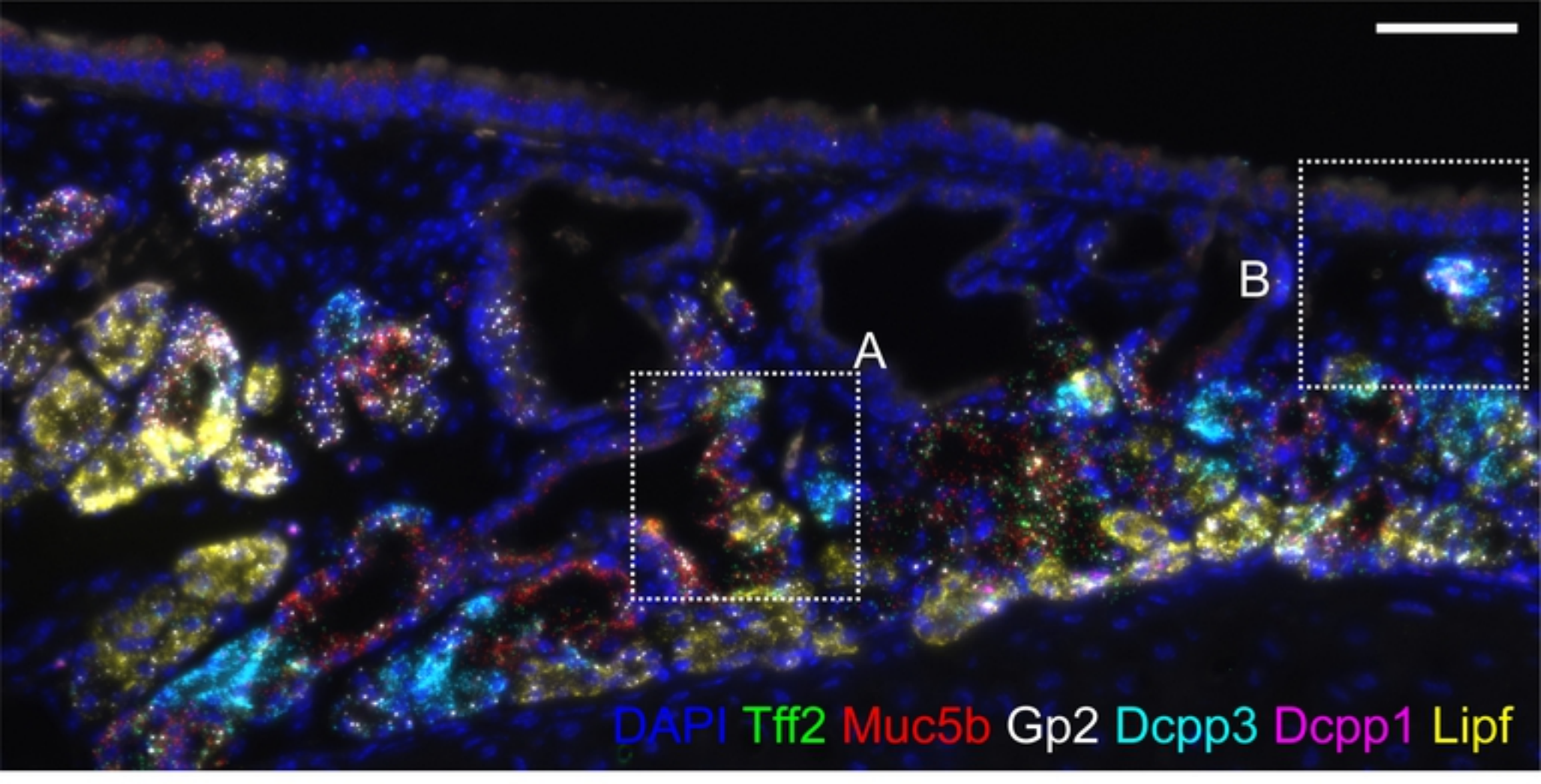
Suppl_Figure_2

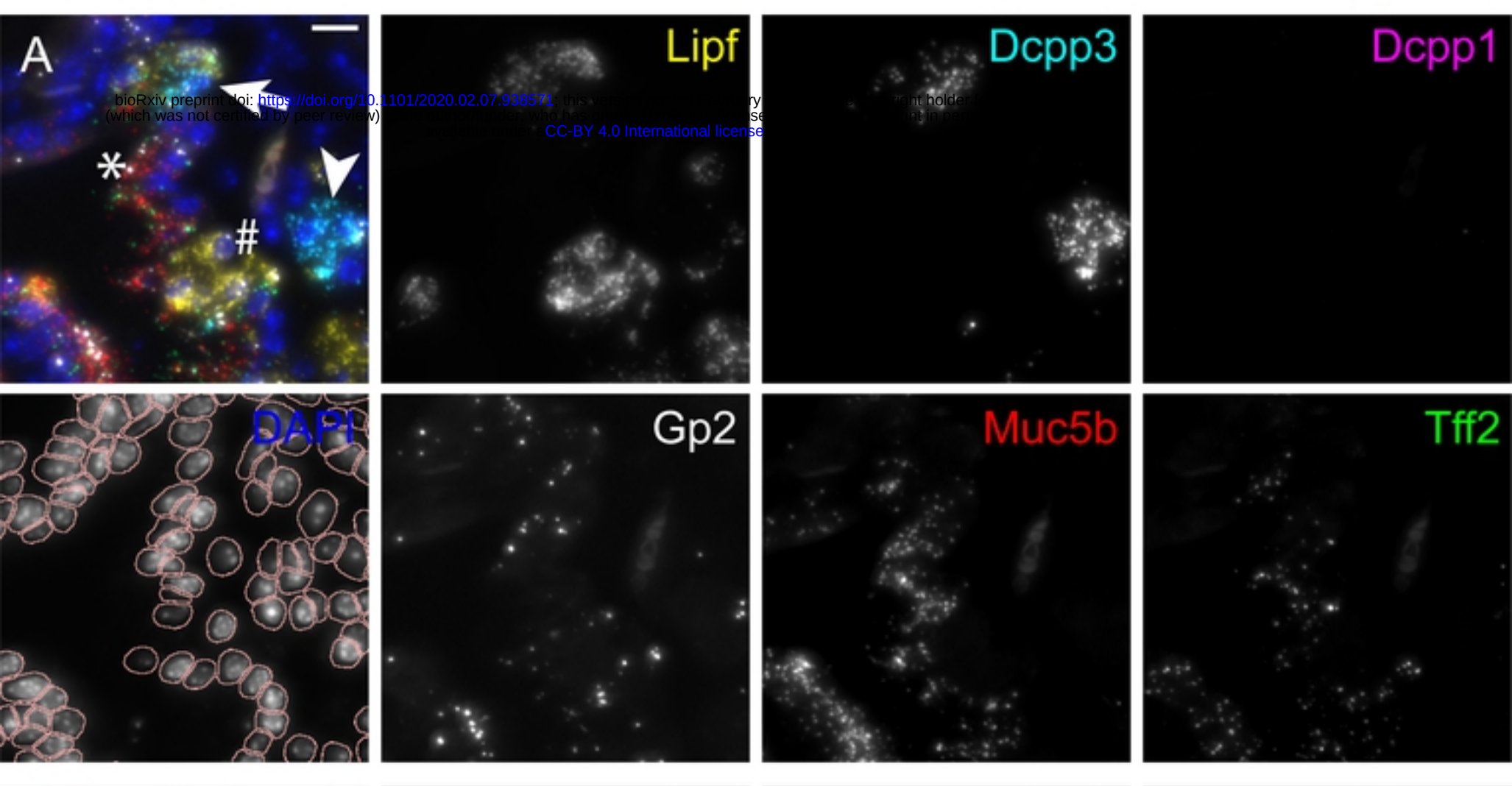


Suppl_Figure_3

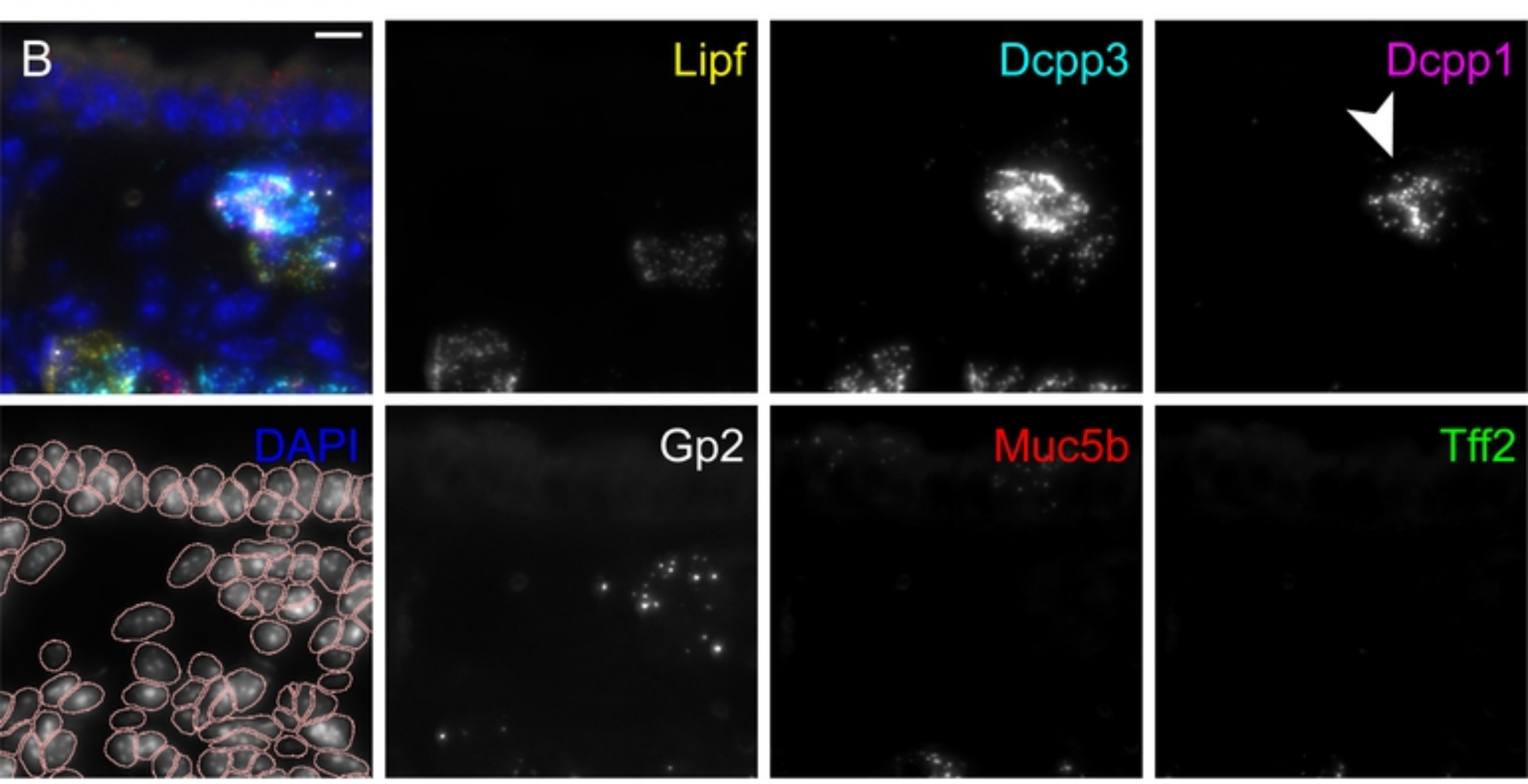


Suppl_Figure_4



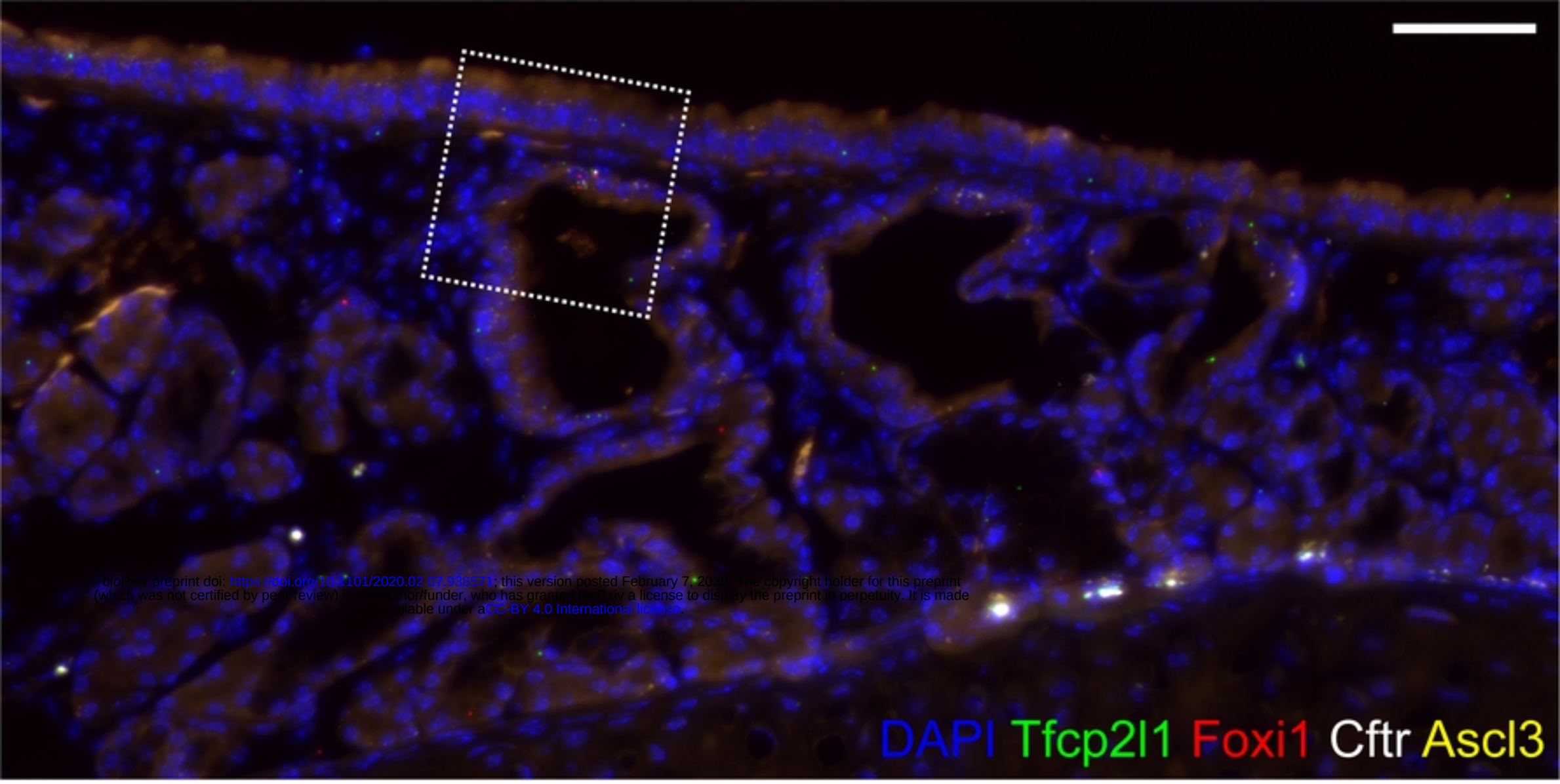


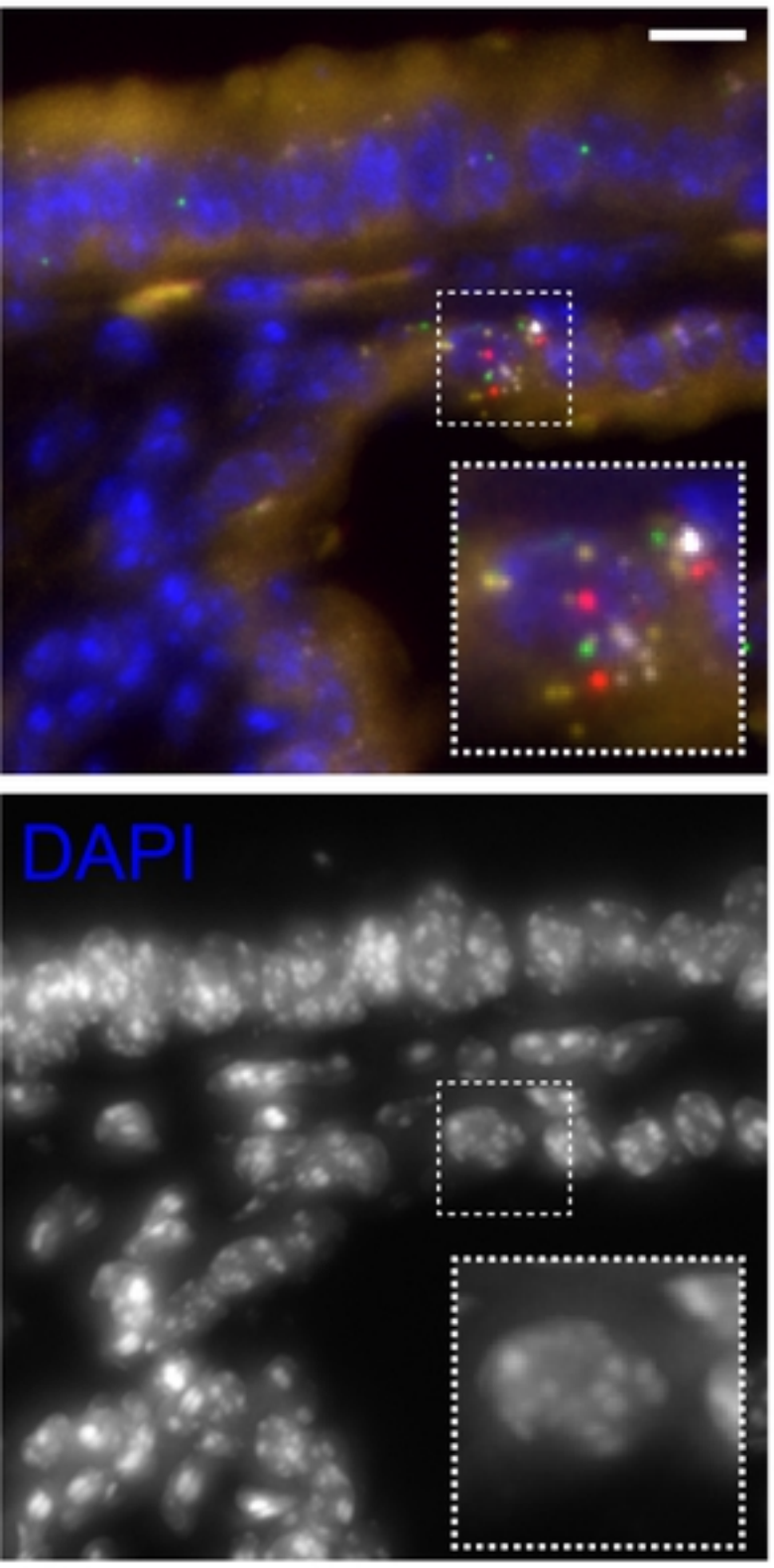
bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571 this version
(which was not certified by peer review)
CC-BY 4.0 International license

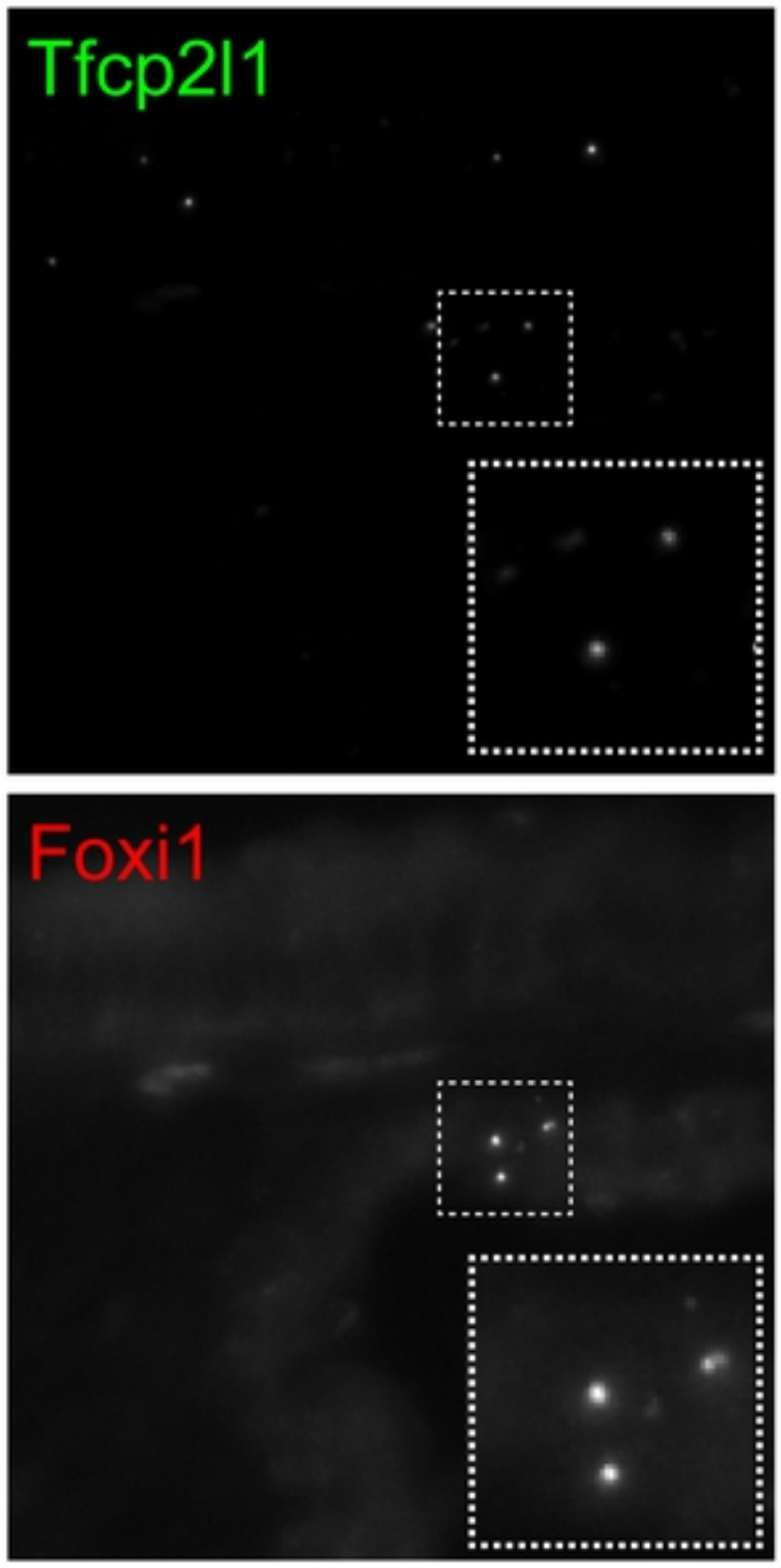


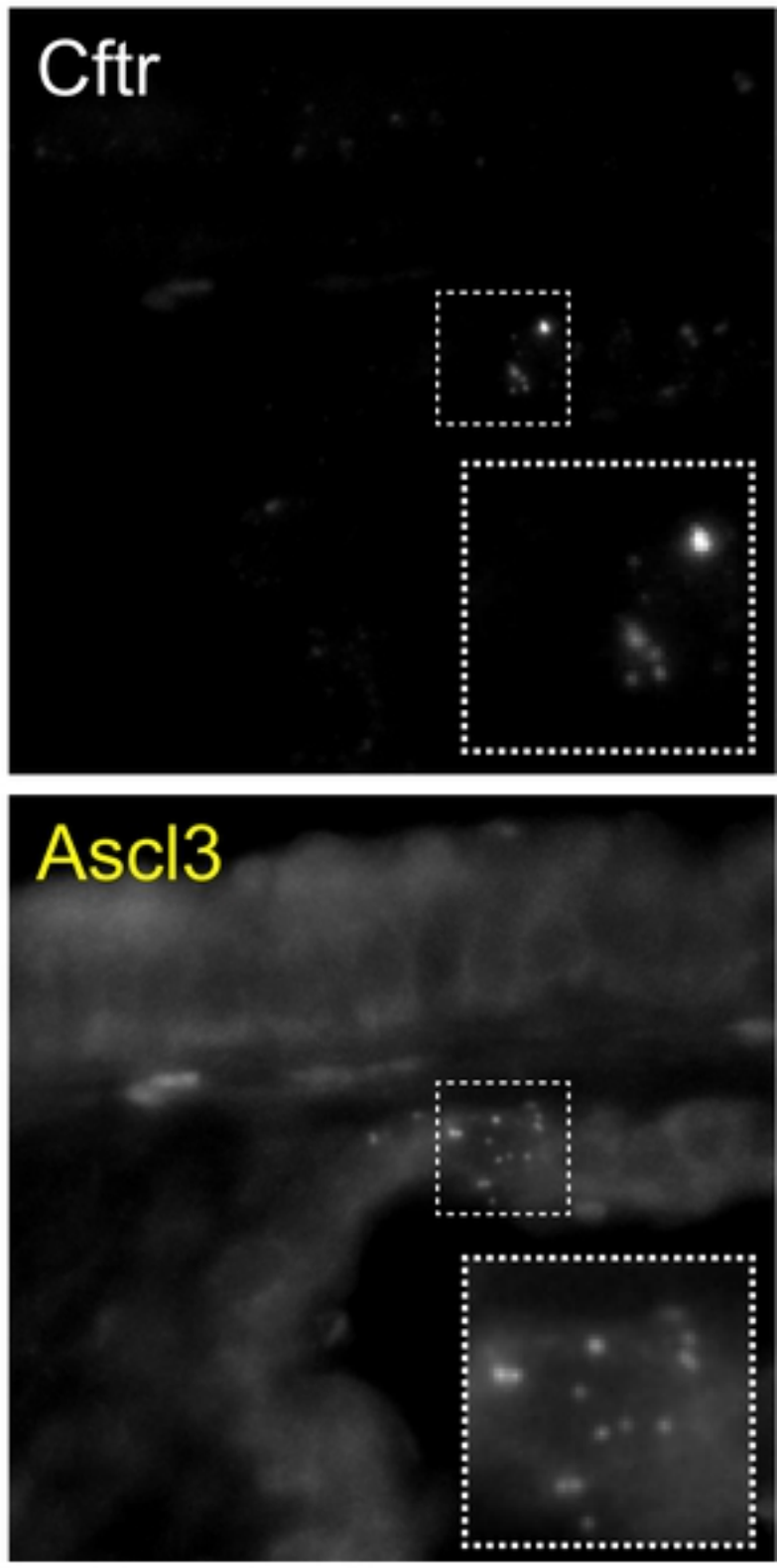
Suppl_Figure_5

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this version posted February 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

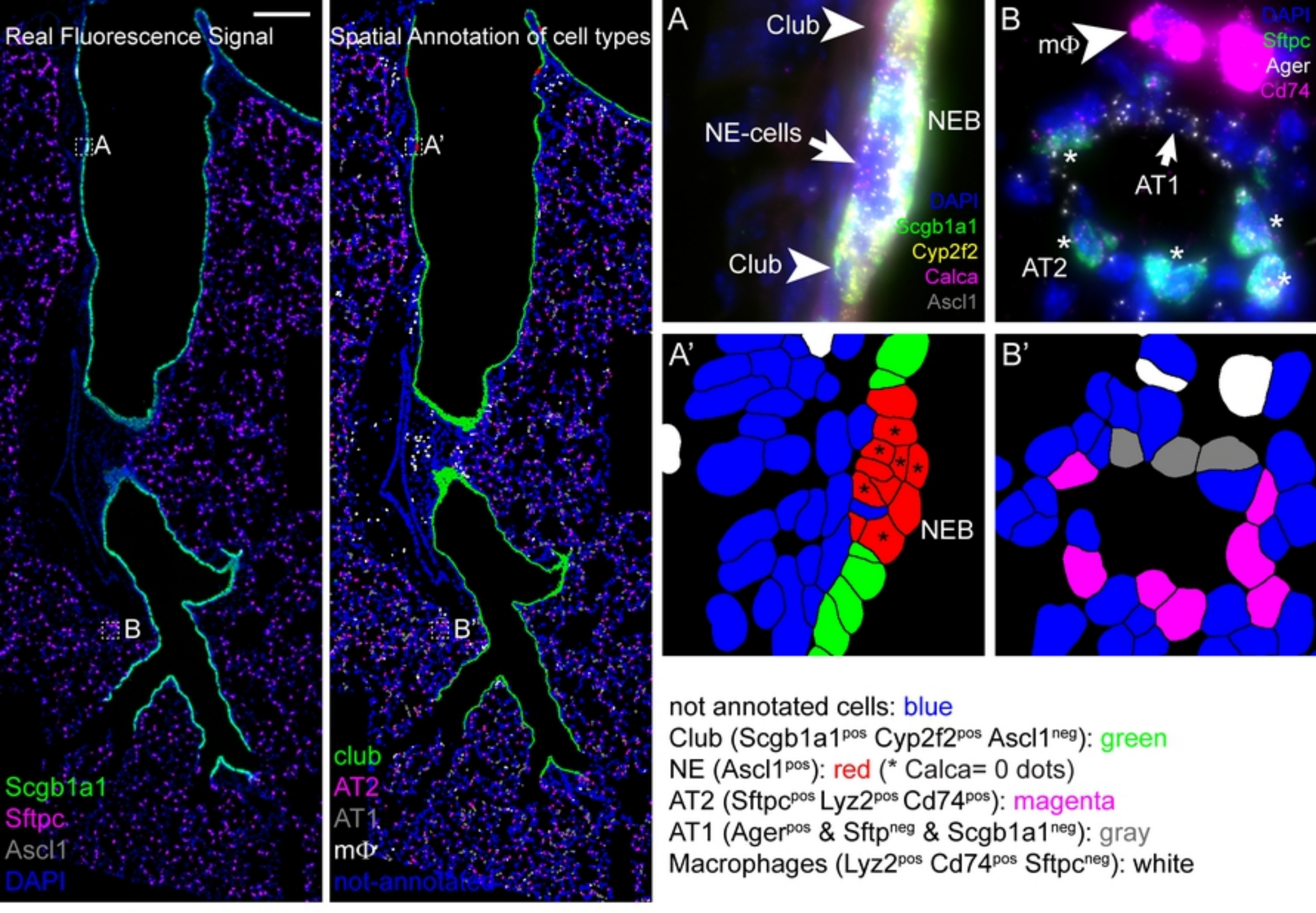








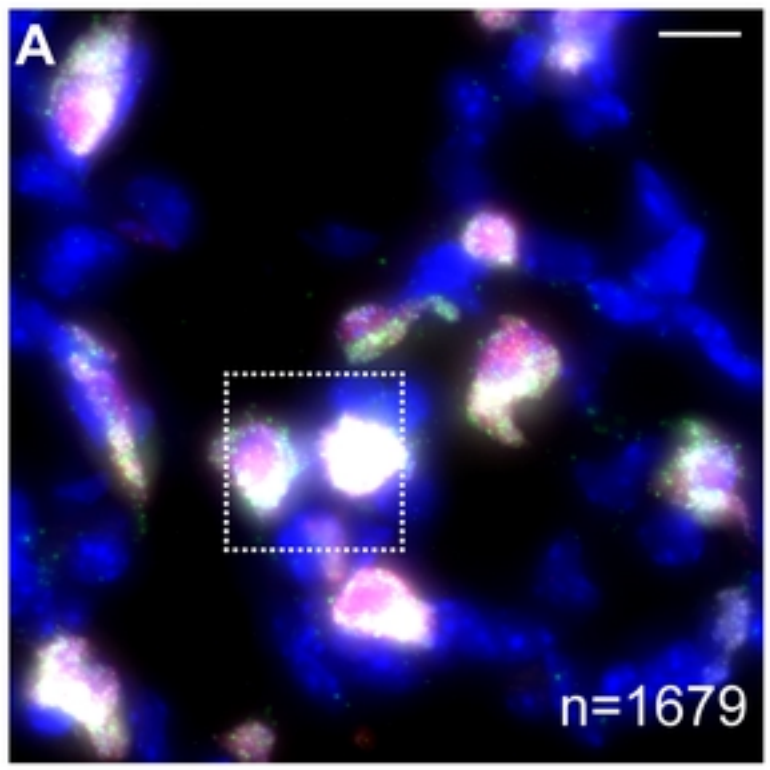
Suppl_Figure_6

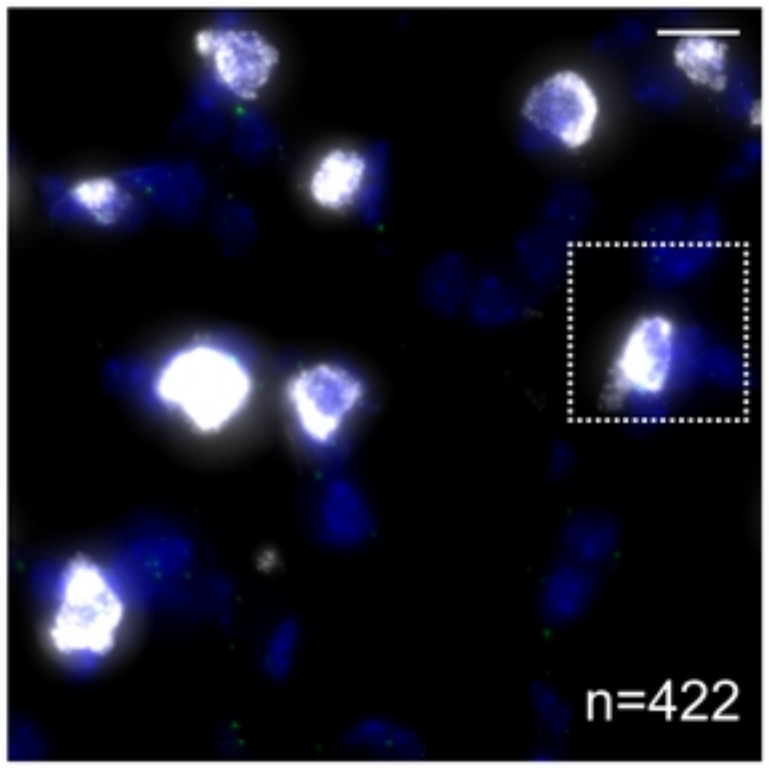


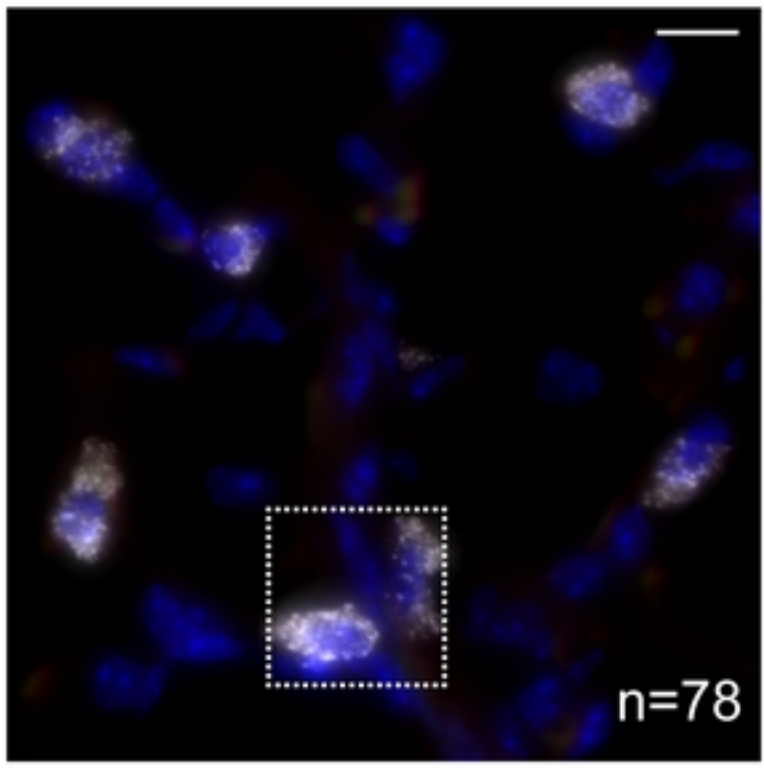
not annotated cells: blue
Club (Scgb1a1^{pos} Cyp2f2^{pos} Ascl1^{neg}): green
NE (Ascl1^{pos}): red (* Calca= 0 dots)
AT2 (Sftpc^{pos} Lyz2^{pos} Cd74^{pos}): magenta
AT1 (Ager^{pos} & Sftp^{neg} & Scgb1a1^{neg}): gray
Macrophages (Lyz2^{pos} Cd74^{pos} Sftpc^{neg}): white

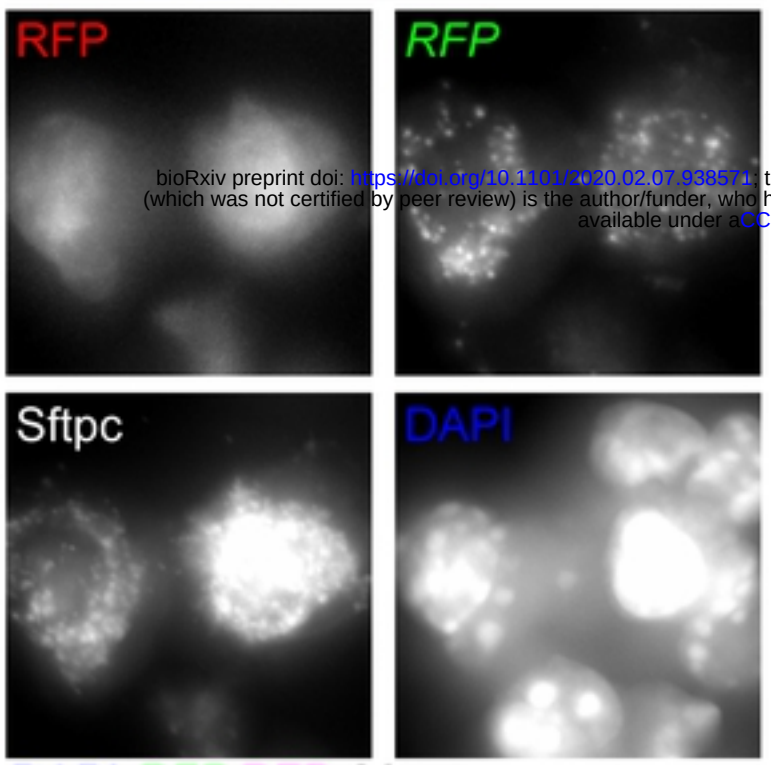
Suppl_Figure_7

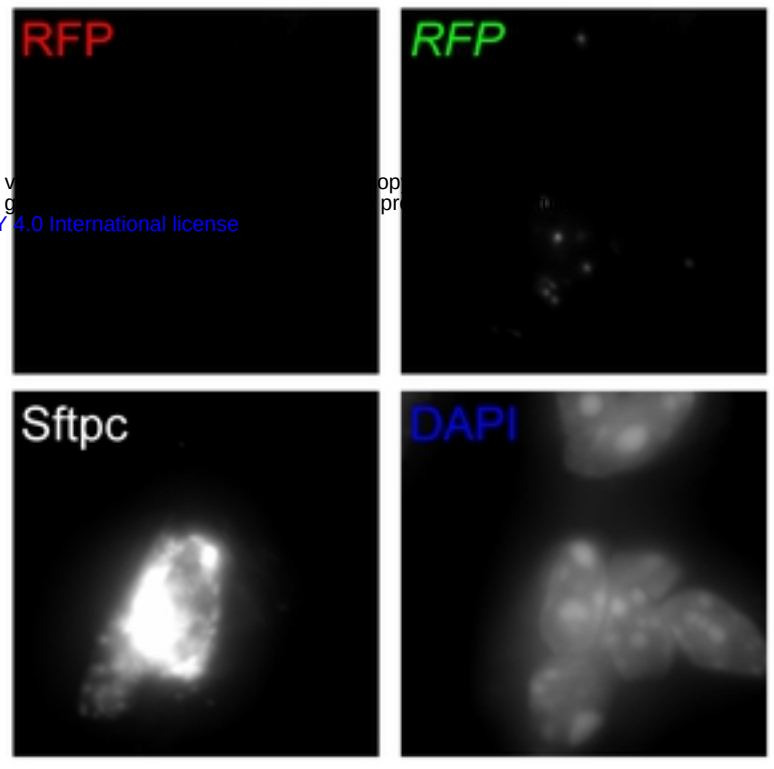
bioRxiv preprint doi: https://doi.org/10.1101/2020.02.07.938571; this v... op...
(which was not certified by peer review) is the author/funder, who has g... pr...
available under aCC-BY 4.0 International license

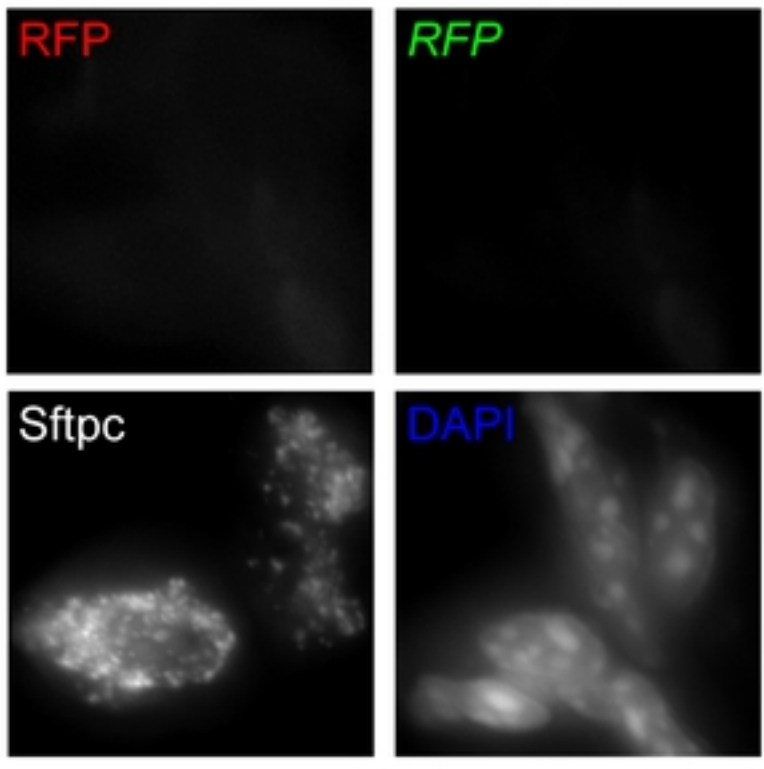




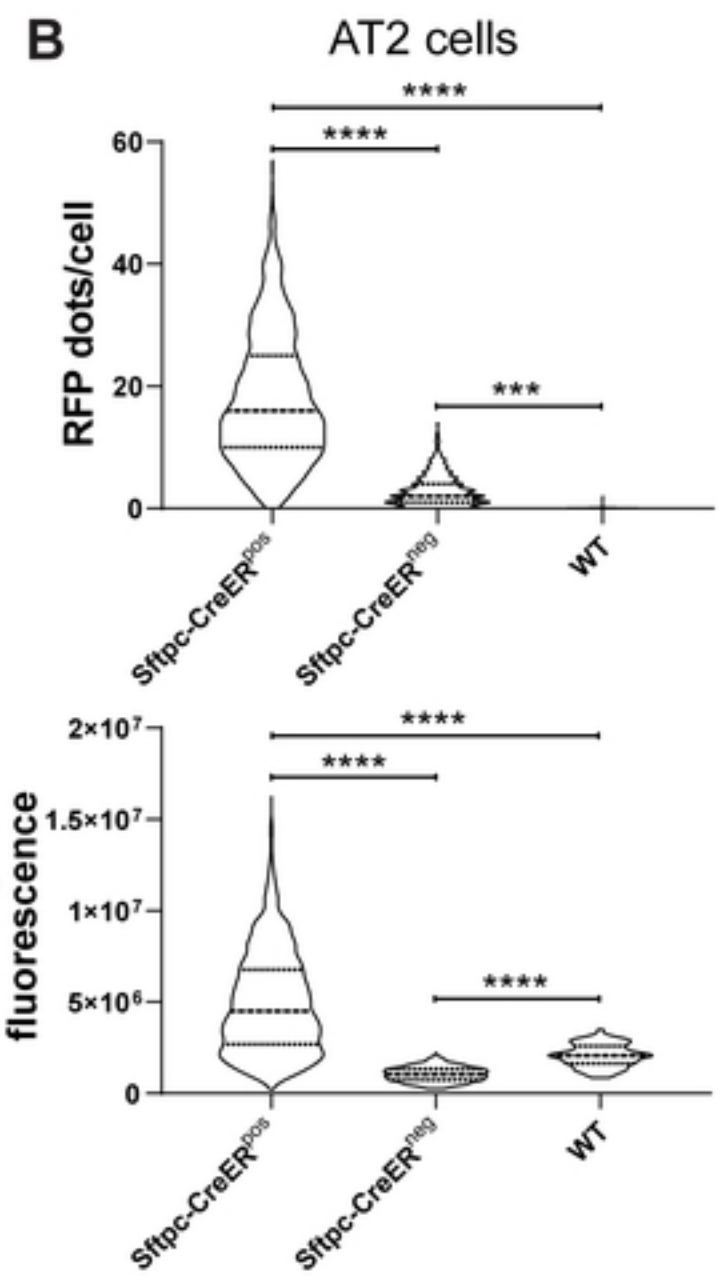


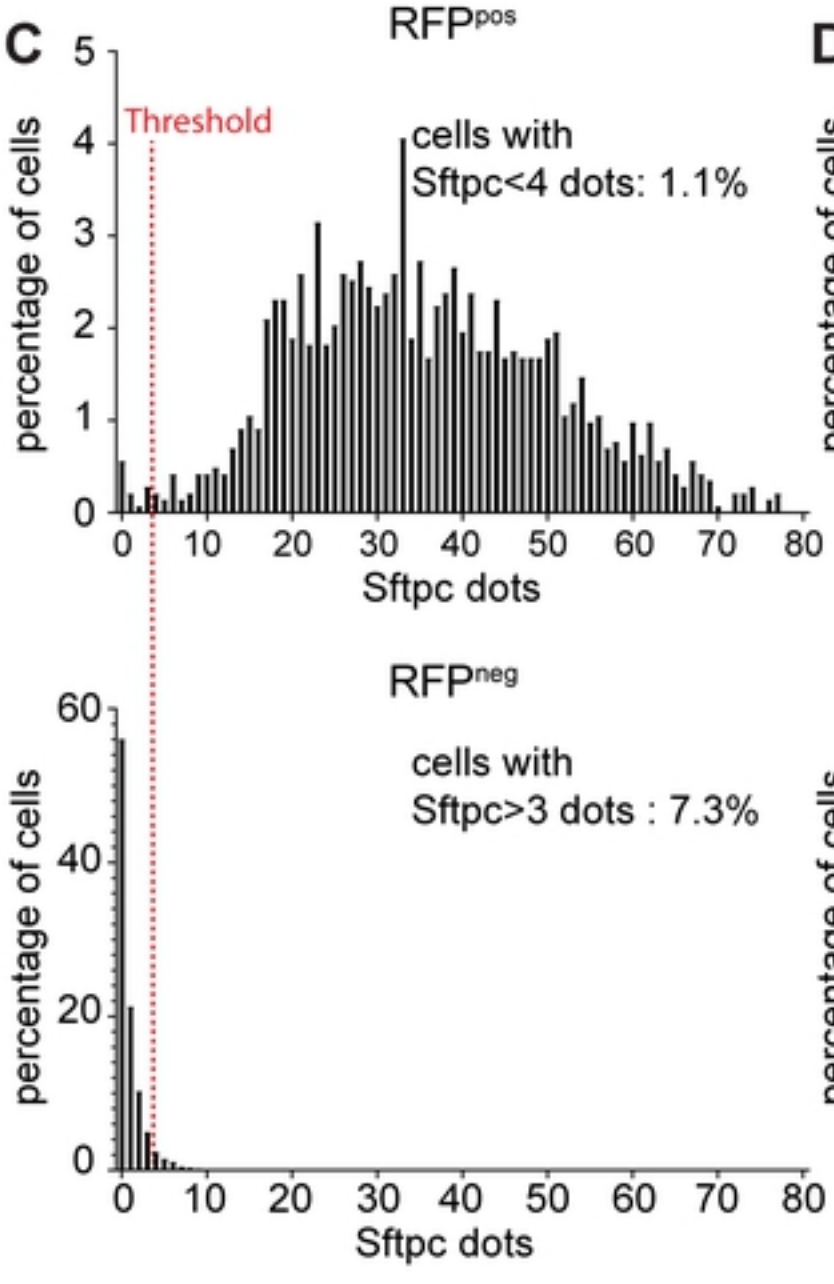






DAPI *RFP* RFP Sftpc





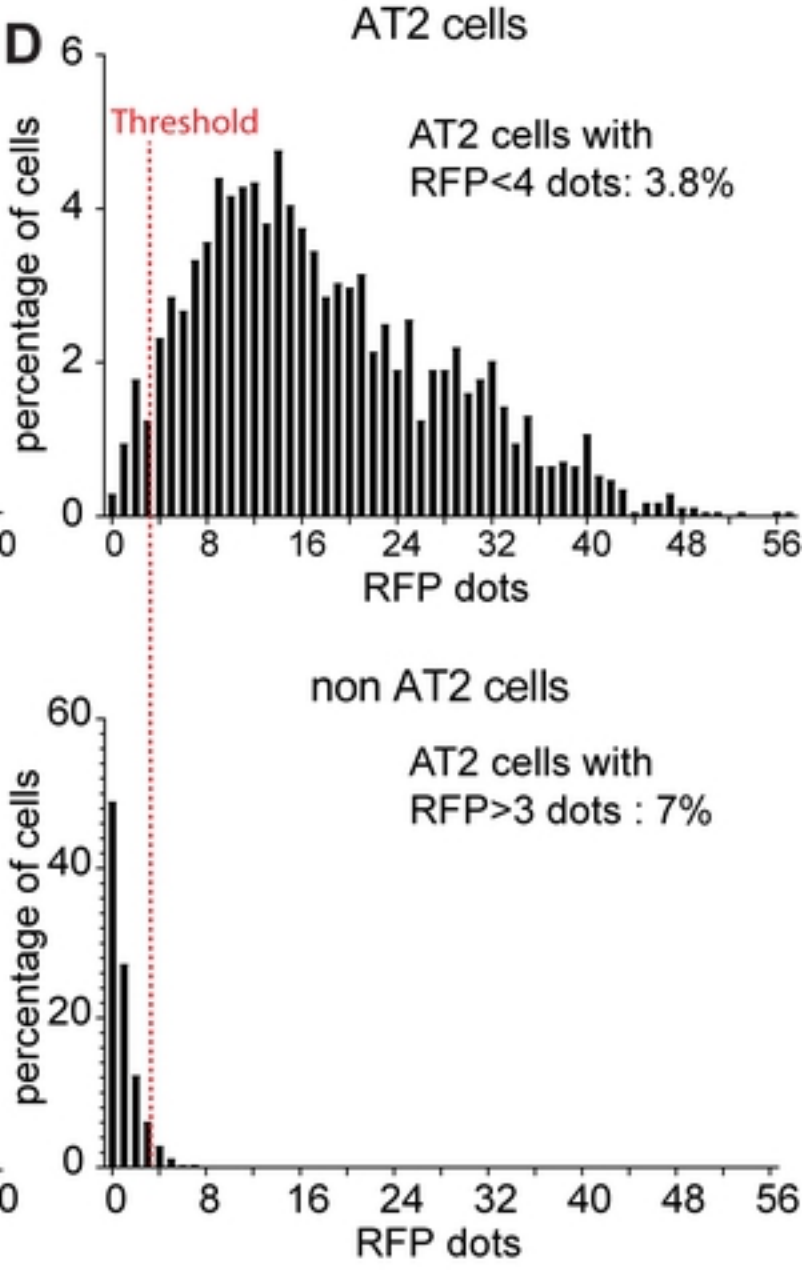


Figure_5