Title

Single-cell transcriptomics of dynamic cell behaviors

Authors

Sheldon J.J. Kwok,[1,2,3,*] Daniel T. Montoro,[4,5*] Adam L. Haber,[6*] Seok-Hyun Yun,[1,2†] Vladimir Vinarsky[7,8,9†]

Affiliations

[1] Harvard Medical School and Wellman Center for Photomedicine, Massachusetts General Hospital, 65 Landsdowne St., Cambridge, MA 02139, USA.

[2] Harvard-MIT Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

[3] LASE Innovation Inc., Cambridge, MA 02140, USA.

[4] Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA.

[5] Department of Systems Biology, Harvard Medical School, Boston, MA, 02115, USA.

[6] Department of Environmental Health, Harvard T. H. Chan School of Public Health, Boston, MA, 02115 USA.

[7] Center for Regenerative Medicine, Massachusetts General Hospital, Boston, MA, 02114, USA.

[8] Division of Pulmonary and Critical Care Medicine, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA.

[9] Vertex Pharmaceuticals Incorporated, Boston, MA, USA.

* These authors contributed equally

† Correspondence to syun@hms.harvard.edu or vvinarsky@mgh.harvard.edu

Abstract

Despite advances in spatial transcriptomics, the molecular profiling of dynamic behaviors of cells in their native environment remains a major challenge. We present a method, termed behavioral transcriptomics, that allows us to couple physiological behaviors of single cells in an intact tissue to deep molecular profiling of individual cells. This method enabled us to establish a novel molecular signature for a striking migratory cellular behavior following tissue injury.

Introduction

Cells in a living organism are dynamic entities, changing their characteristics over space and time and constantly interacting with the host and pathogens. The ability to obtain such information and link it to detailed molecular phenotypes of the cells would be highly useful for biomedical investigations but has been underappreciated. Here, we present a method that allows us to characterize complex physiologic behaviors of single cells in an intact tissue and then perform live imaging-guided sequencing of the cells. We validate this approach using a regeneration model of airway tissues and demonstrate how this method leads to new biological findings.

There is pressing need for a comprehensive understanding of cellular behaviors in the lung, the site where aberrant cellular behavior has been linked to asthma (Kim et al., 2020; Park et al., 2015) pulmonary fibrosis (Fukumoto et al., 2016), and viral infections including influenza and coronaviruses (Kumar et al., 2011). Single-cell RNA-sequencing (scRNA-seq) has emerged as a precise way to define cell type and cell state, and new techniques are being developed to determine the spatial distribution of sequenced cells in tissues (Marx, 2021). However, the molecular pathways that drive the cellular behavior *in situ* continue to be inferred from time-lapse tissue sampling or transcriptional kinetics (La Manno et al., 2018). Moving beyond inference requires coupling visualized *in situ* cell behavior with deep molecular profiling of visualized cells.

Live cell imaging is an established technique for capturing morphology and cellular dynamics such as cellular migration during skin regeneration (Park et al., 2019, 2017), but imaging in the lung remains challenging due to difficult access and the constant motion of the respiratory system. Additionally, molecular information that accompanies live imaging is largely limited to a few fluorescent reporters. Prior attempts to link deep molecular profiling with live imaging have relied on imaging dissociated cells (Lane et al., 2017; Yuan et al., 2018), cell monolayers (Hu et al., 2020) or organoids (Konen et al., 2017) rather than cell behaviors in their native tissue environment.

Results

We describe a novel approach to linking live tissue imaging with single cell profiling (Figure 1a). In order to visualize the airway epithelium at high resolution over days, we explant a mouse trachea and secure it in a custom imaging platform, which minimizes sample movement during imaging and maintains a constant supply of nutrients from below the explant without disrupting the air-liquid interface (Figure 1a and Figure 1 – supplement figure 1a). This platform allows us to image common and rare cell types in the airway epithelium at high resolution in their native environment (Figure 1 – supplement figure 1b). The explant culture also allows an uninjured tracheal epithelium to survive with its native cellular anatomy for weeks with daily high-resolution imaging (Figure 1 – supplement figure 1c).

Discernable cell behaviors have a broad time scale, ranging from milliseconds to days. Thus, we imaged a wide range of cellular behaviors from rapid fluctuations of ciliary beating and directional mucociliary transport over milliseconds to wholescale regeneration of the airway epithelium, which occurs over days after the injury (Figure 1, c to f). Furthermore, this method allows single-cell level registration within tissues that are live-imaged and subsequently fixed and stained, which enables a unique

comparison between live fluorescence cellular patterns and immunostains that describe cell identity and function (Figure 1 – supplement figure 1d).

Remarkably, this airway imaging platform faithfully recapitulates and captures cellular dynamics of epithelial regeneration from native basal stem cells after an extensive epithelial injury (Figure 1f and Figure 2 – supplement figure 1) induced by sulfur dioxide (SO_2). In the first 5 days, the basal cells divide, increasing cellular density and reforming the pseudostratified epithelium. In the next 5-10 days, the epithelium differentiates, leading to restoration of the full epithelium, including the regeneration of ciliated cells. Complete regeneration requires an air-liquid interface (Figure 1f). Overall, this murine trachea explant ALI culture retains the nearly complete 3D organization and microenvironment of the basal progenitor cells and, therefore, offers a unique model to study organ physiology and regeneration outside of the body.

Continuous time-lapse imaging of the airway epithelium for up to 80 hours after injury (Figure 2a and Figure 2 – supplement figure 2a and Movie 1) demonstrated changes in cellular architecture over time, including an increase in the average cellular density and epithelial thickness, without apparent phototoxicity (Figure 2 – supplement figure 2a). We examined cell movement using single-cell tracking following segmentation of cell nuclei and particle image velocimetry (PIV) of non-segmented images (Figure 2 – supplement figure 2b and c). These analyses revealed a variety of regeneration cellular dynamics inaccessible without live imaging. For example, Hertwig's Rule predicts that a cell division plane is perpendicular to the long axis of the cell during the preceding interphase (Minc and Piel, 2012). This was established in plants (Besson and Dumais, 2011) and developing simple model organisms (Aigouy et al., 2010; Concha and Adams, 1998; Tsou et al., 2003) but has never been probed in an adult regenerating tissue. We found that the long axis in most cells predicts the cell division axis, while the axis of cellular movement prior to cell division does not (Figure 1e and Figure 2 – supplement figure 3a and b).

We also found a surprising degree of heterogeneity of collective cellular migration during regeneration throughout the injured airways. In regions that demonstrated rapid cellular movement after injury, the movement peaked at 26-38 hours after SO_2 injury and the speed declined significantly in most regions by 50 hours after injury (Figure 2c). There was a significant interaction between time and the mean speed, but no significant difference between mouse and mean speed (Figure 2c). Variable migratory behavior of airway epithelial cells has been observed in cell culture models (Kim et al., 2020; Park et al., 2015) but not previously in an intact regenerating airway tissue. We found that the frequency distribution of cellular speed in different regions at 26-38 hours demonstrated large variability ranging from "non-mover" regions (< 1.5 μm/hr) to "mover" regions (> 4 μm/hr) (Figure 2d). Furthermore, videos with higher temporal resolution revealed that the cells with slower movements in the "non-mover" regions had no directional preference, whereas the "mover" regions with faster cellular movements were more uni-directional (Figure 2 – supplement figure 3c and d and Movie 2).

Distinct subsets of cells have been theorized to contribute to the regeneration process (Pardo-Saganta et al., 2015a; Tadokoro et al., 2014), but it is unclear whether these heterogeneous transcriptional cell states reflect gene expression stochasticity or correlate with unique cell behaviors. To determine the molecular signatures of cells with

directional movement compared to regenerating non-moving cells, we marked epithelial regions by photoconversion at 24 hours after injury, imaged every 6 hours, screened for "movers" with >50 μm displacement (>3 μm/hr) at 18 hours after photoconversion, and then isolated photoconverted epithelial cells by FACS for plate-based scRNA-seq (Figure 3a and Materials and Methods). Dimensionality reduction revealed that cells from "moving" region (M) and a "non-moving" (NM) region cluster separately (Figure 3b). Using unsupervised clustering and cell identity signatures (Methods), we found that nearly all the cells in the M region are basal cells, whereas the NM region contains basal and club cells (Figure 3c). We defined the differences in gene expression between the M basal cells and NM basal cells and identified gene signatures that are enriched (FDR < 0.05, likelihood-ratio test) either in the M or the NM basal cells (Figure 3c).

We wondered whether the identified phenotypes may be a common feature of injury-induced epithelial regeneration. We examined published data of an independent injury model (Borthwick et al., 2001) and analyzed the prevalence of these signatures during repair after polidocanol injury. As predicted, the M basal cell signature is strongly enriched 24 hours post injury (hpi), declines at 48 hpi and 72 hpi, and returns to baseline at 1 week after injury (all $p < 10^{-16}$, Mann-Whitney U test, Figure 4a). Similarly, the NM signature is decreased at 24 hpi when cell migration is presumed to be active, increases at 48 hpi and 72 hpi when cell migration is presumed to be diminished, and returns to baseline at 1-week post-injury when regeneration is complete (Figure 4a). Furthermore, at 24 hpi we found that scoring basal cells using M and NM signatures segregated basal cells into two statistically distinct cell populations (Figure 4b), indicating that polidocanol regeneration is likely also characterized by these cell phenotypes. To test this possibility, we used unsupervised clustering (Methods) to define two groups of basal cells at 24 hpi and found that these two populations were indeed separately enriched for the M and NM basal cell signatures (Figure 4c), confirming the presence of distinct M and NM basal cells during polidocanol regeneration. Taken together, these findings suggest that distinct M and NM cell behaviors are conserved features of early epithelial regeneration and demonstrate that our live imaging-guided single-cell profiling approach can discover generalizable principles of tissue biology.

Discussion

The rapid progress in spatially resolved transcriptomics is enabling the discovery and characterization of transcriptionally heterogenous cells in diverse tissue contexts (Lee et al., 2021; Ståhl et al., 2016). However, these methods do not capture the dynamics of cell behaviors that often define the unique biological processes that occur in the tissues. To address this gap, we developed an approach to examine the association of molecular and behavioral phenotypes of single cells in their native tissues. We first established a respiratory organ explant culture that maintains tissue dynamics for an extended length of time, and subsequently combined this platform with live imaging in

order to observe distinct lung cellular behaviors at a broad time scale, spanning cell migration, cell division, and ciliary beating.

To link cell behavior to molecular analysis, we used photoconversion to mark cells that display distinct cell behaviors for subsequent single-cell genomics analysis. We found that a subpopulation of basal stem cells migrates within the lung during early regeneration. We used recently developed single-cell RNA-sequence approaches to establish molecular signatures for moving and nonmoving basal cells. Furthermore, we found these distinct cell signatures across independent lung regeneration models, suggesting that M and NM cell behaviors are likely not only conserved cellular features of early epithelial regeneration, and but also that live imaging-guided single-cell profiling approach can discover general principles of tissue biology.

Materials and Methods

Mice

mT-mG (stock no. 007676), nT-nG (stock no. 023035), CAGs-LSL-rtTA3 (stock no. 029617), and Col1a1-tetO-H2B-mCherry (stock no. 014602), CD11cCre (stock no. 007567), and Ascl1nGFP (stock no 012881) mice were purchased from the Jackson Laboratory. Foxj1Cre (Zhang et al., 2007), KRT5rtTA(Diamond et al., 2000), B1EGFP (Miller et al., 2005), Foxj1CreER (Rawlins and Hogan, 2008), CC10CreER (Rawlins et al., 2009), and Kaede (Tomura et al., 2008) lines were previously described. A line of Membrane-GFP (mG) mice was generated by selecting GFP-positive pups of a Foxj1Cre-mTmG male parent (with mT to mG recombination in the sperm) and backcrossing to WT background to eliminate the Cre allele. The mG line without Cre was crossed to nT-nG to generate the “nT-mG” strain. Mice were maintained in an Association for Assessment and Accreditation of Laboratory Animal Care-accredited animal facility at the Massachusetts General Hospital, and procedures were performed with Institutional Animal Care and Use Committee (IACUC)-approved protocols. Mice of all strains were housed in an environment with controlled temperature and humidity, on 12-hour light-dark cycles, and fed with regular rodent’s chow.

Sulfur Dioxide Injury

Sulfur dioxide (SO_2) injury model was performed as previously described (Kim et al., 2012; Pardo-Saganta et al., 2015b). In brief, mice were exposed to 500 p.p.m. of SO_2 for 3 h 40 min and the trachea was collected 16-24 hours after injury for imaging and explant culture.

Tracheal Explant

Tracheas were dissected, cleared of connective tissue and adjacent organs, and opened longitudinally along the anterior tracheal wall. The tracheas were placed on ice in DMEM/F-12 Media with Primocin (InVivoGen) and 15 mM HEPES until culture. Trachea explants were then sutured onto a silicone o-ring and placed in a custom-made

tissue culture dish over an inverted air-liquid interface (ALI) insert secured in a 60 mm tissue culture dish by PDMS. This approach ensured stability during high resolution imaging. The media contacted the explant from below through the ALI membrane. The dish was placed in a physiological live imaging chamber (CO_2 and temperature-controlled, TokaiHit) on the stage of the 2-photon microscope.

Physiological 2-Photon Imaging

Trachea explant imaging was performed on an Olympus FVMPE-RS multiphoton laser scanning microscope equipped with a MaiTai HPDS-O IR pulsed laser (900 nm for GFP and SHG) and INSIGHT X3-OL IR pulsed laser (1100 nm for tdTomato), using a 25X water immersion lens (NA 1.05). Explants were imaged at time points as indicated in the Figures. For orthogonal view reconstruction, we scanned the trachea with 0.75 µm Z steps. To reimage the same trachea at high resolution at different time points, landmarks such as cartilage rings and vascularity patterns were used as fiducial marks. These fiducial marks were also used for 2D and 3D registration of different time points.

Image Analysis

4D images (x,y,z,t) were imported into MATLAB and/or ImageJ for image processing and analysis. Because the curvature of the tissue changes over time, we first normalized each 3D image to generate a flat basement membrane. As the SHG signal is maximal at the basement membrane, we computed the *z* height of the basement membrane across the image after applying a Gaussian blur (typical σ values: 10-25 µm in xy, 1-4 µm in z). This height was subtracted from the original 3D data to level the basement membrane. MATLAB code for flattening the 3D images is available upon request. Other image processing steps including brightness and contrast adjustments, background subtraction, photobleaching correction, pseudocoloring, 3D time-lapse registration, and stitching were performed using built-in functions in ImageJ.

Cilia beating was recorded by acquiring time-lapse two-photon images of the epithelial surface at 150 Hz over 200 frames using a resonant galvanometer scanner. To estimate the cilia beat frequency (CBF), we estimated the power spectral density of the fluorescence intensity fluctuations across the image using Welch's method in MATLAB. The peak fluctuation frequency was computed for each pixel across the image corresponding to bright cilia. Mucociliary transport was measured by applying 1 µm fluorescent spherical beads to the epithelial surface and recording their displacement over time after equilibration.

To track individual cells over time-lapse imaging, images were imported into ilastik for segmentation and cell tracking (Berg et al., 2019). Briefly, pixels corresponding to nuclei were first classified using manual training and machine learning. Next, individual cells were similarly identified through manual training and machine learning algorithms to classify objects. Finally, classified cells were tracked over time using a conservation tracking algorithm. Segmented and tracked cells were then imported into MATLAB for quantitative analysis, including computation of individual cell speeds over time.

For automated estimation of cell speed from time-lapse imaging, we performed automated particle imaging velocimetry (PIV). Image sequences were imported into MATLAB and analyzed using the PIVlab plugin (Thielicke and Stamhuis, 2014). A direct Fourier transform correlation with multiple passes of sizes consisting of 24 µm, 16 µm and 10 µm was used. This generated displacement vectors arranged in a grid with 10 µm spacing for each sequential pair of images. The average cell speed for each time-point was estimated by computing the average absolute displacement estimated by PIV in each (10 x 10) μm^2 region divided by the time between images. To quantify the directionality of cell movement, we computed the circular variance of the displacement vectors generated by PIV analysis.

Kaede Photoconversion

Trachea from Kaede mice were explanted 20 hours after SO2 inhalation injury, sutured onto a silicone O-ring, and secured on an inverted ALI insert in media on ice, and placed on the imaging platform of a FV3000 Olympus Laser Scanning confocal microscope. Selected regions were outlined and photoconverted using the 405 nm laser for 2 minutes, while both disappearance of KaedeGreen and appearance of KaedeRed were simultaneously visualized using the 488 nm and the 561 nm lasers, respectively.

To identify regions of movement and no movement, we explanted Kaede (Tomura et al., 2008) mouse tracheas 20 hours after SO_2 injury, photoconverted distinct regions with a specific shape, and proceeded with timelapse live imaging, screening for regions with significant shape displacement over time (from epithelial movement) (Figure 3a). After defining whether a region moved or remained non-moving, we excised a trachea fragment, dissociated the fragment into single cells, and used flow activated cell sorting (FACS) to isolate photoconverted (KaedeRed) epithelial cells. We then proceeded to single-cell RNA sequencing of cells isolated from moving and non-moving regions.

Cell Dissociation and FACS

Airway epithelial cells were dissociated using papain solution. Tracheal fragments with photoconverted regions were trimmed and incubated in papain dissociation solution and incubated at 37 °C for 2 hours. After incubation, dissociated tissues were passed through a cell strainer and centrifuged and pelleted at 500g for 5 min. Cell pellets were dispersed and incubated with Ovo-mucoid protease inhibitor (Worthington Biochemical, cat. no. LK003182) to inactivate residual papain activity by incubating on a rocker at 4 °C for 20 min. Cells were then pelleted and stained with EpCAM–BV421 (1:50; BD Bioscience, #563214) for 30 min in 2.5% FBS in PBS on ice. After washing, cells were sorted by fluorescence (antibody staining, Kaede-Green and Kaede-Red) on a BD FACS Aria (BD Biosciences) using FACS Diva software and analysis was performed using FlowJo (version 10) software.

Single cells were sorted into each well of a 96-well PCR plate containing 5 µl buffer. After sorting, the plate was sealed with a Microseal F, centrifuged at 800g for 1 min and

immediately frozen on dry ice. Plates were stored at −80 °C and submitted to a core facility for cDNA library generation, amplification, and sequencing.

Single Cell Sequencing and Sequence Analysis

cDNA was generated from single cells in the 96-well plate using the SmartSeq v4 kits (Takara Bio) using 1/4th volume reactions dispensed using a Mantis dispenser (Formulatrix). Samples were amplified using 18 cycles of PCR. Resulting cDNA was then made into Illumina-compatible libraries using the Nextera XT kit (Illumina Inc). Libraries were sequenced on a NextSeq using a Mid Output 150 cycle kit (Illumina Inc.) using 75bp paired end reads.

Pre-processing of plate-based scRNA-seq data

BCL files were converted to merged, de-multiplexed FASTQ files using the Illumina Bcl2Fastq software package v.2.17.1.14. Paired-end reads were mapped to the UCSC mm10 mouse transcriptome using Bowtie (Langmead et al., 2009) with parameters '-q–phred33-quals -n 1 -e 99999999 -l 25 -I 1 -X 2000 -a -m 15 -S -p 6', which allows alignment of sequences with one mismatch. Expression levels of genes were quantified as transcript-per-million (TPM) values by RSEM (Li and Dewey, 2011) v.1.2.3 in paired-end mode. For each cell, we determined the number of genes for which at least one read was mapped, and then excluded all cells with fewer than 1,000 or more than 10,000 detected genes, or less than 25% of reads mapping to the transcriptome.

To identify variable genes a logistic regression was fit to the cellular detection fraction, using the total number of transcripts per cell as a predictor. Outliers from this curve are genes that are expressed in a lower fraction of cells than would be expected given the total number of transcripts mapping to that gene, that is, cell-type or state-specific genes. We used a threshold of deviance <−0.15, producing a set of 1910 variable genes.

Dimensionality reduction by PCA and t-SNE

We restricted the expression matrix to the subsets of variable genes and high-quality cells noted above, and values were $\log_2$-transformed, and then centered and scaled before input to PCA, which was implemented using the R function 'prcomp' from the 'stats' package. After PCA, significant principal components were identified by inspection of the scree plot. Only scores from the first 20 PCs were used as the input to further analysis.

For visualization purposes only (and not for clustering), dimensionality was further reduced using the Barnes–Hut approximate version of t-SNE (Van Der Maaten, 2014; van der Maaten, 2008) (Figure 3b). This was implemented using the 'Rtsne' function from the 'Rtsne' R package.

To identify cell types within the data, unsupervised hierarchical clustering was used using the 'Ward.D2' metric in the 'hclust' R package. Pearson's correlation was used as a distance metric. This produced 3 clusters, 2 were clearly identifiable as Basal and Club cells, based on disjoint expression of known markers *Krt5* and *Scgb1a1*, respectively, while the third was distinguished by much lower technical quality (an average of 2373 genes detected per cell compared to 5193 for the Basal and 5480 for the club clusters respectively, p=0.0004, Mann-Whitney U-test). These low-quality cells were not used for DE testing.

To identify the signature of moving vs non-moving basal cells (Figure 3c) we ran differential expression tests between cells in the Basal cluster between the two conditions (moving and non-moving), and selected genes that were differentially expressed (FDR<0.05). Differential expression tests were carried using a two part 'hurdle' model to control for both technical quality and mouse-to-mouse variation. This was implemented using the R package MAST (Finak et al., 2015), and *P* values for differential expression were computed using the likelihood-ratio test. Multiple hypothesis testing correction was performed by controlling the false discovery rate using the R function 'p.adjust'.

Re-analysis of polidocanol injury dataset

Previously published single-cell RNA sequencing data from mouse trachea injured using polidocanol(Plasschaert et al., 2018) was downloaded from the NCBI GEO (GSE102580). All available unique molecular identifier (UMI) counts tables from mice at 24, 48, 72 and 168 hours after injury along with uninjured controls were downloaded. Cell-types were determined using the authors provided annotations. To determine the expression of migration-associated genes in the injury response, we scored the Basal cells for the set of genes (Figure 3c) both significantly up- ('mover') and down-regulated ('non-mover') (Figure 3d). Scoring cells was computed as described previously(Montoro et al., 2018). To obtain a score for a specific set of n genes in a given cell, a 'background' gene set was defined to control for differences in sequencing coverage and library complexity. The background gene set was selected for similarity to the genes of interest in terms of expression level. Specifically, the 10n nearest neighbors in the 2D space defined by mean expression and detection frequency across all cells were selected. The signature score for that cell was then defined as the mean expression of the n signature genes in that cell, minus the mean expression of the 10n background genes in that cell.

Unsupervised cluster analysis of polidocanol-injured basal cells 24 hours after injury was computed using default settings in Seurat. Briefly, variable genes were selected using the method 'vst', and then PCA was computed using only these genes. Shared-nearest neighbor (SNN)-based clustering was implemented using the 'FindClusters' function (resolution parameter = 0.25) using the first 25 principal components as input, resulting in two clusters (Figure 3e).

Statistical Analysis

Data was compared among groups using the Student's *t*-test (unpaired, two-tailed) unless otherwise specified in the Figure legends. Analysis was performed with Graphpad Prism software (version 9.1.0).

Acknowledgements

We thank Drs. Jin-ah Park, Jeffrey Fredberg, and Jeffrey Drazen for stimulating discussion about the migratory behaviors of cells. We thank all of the laboratory of Dr. Jayaraj Rajagopal for providing the cell type-specific fluorescent reporter mice, access to advanced microscopy, and critical appraisal of the project. We thank the HSCI-CRM Flow Cytometry and Microscopy Core Facility at the Massachusetts General Hospital, and The Bauer Core Facility at Harvard University for facilitating this project.

Competing interests

SJJK is currently an employee of and has financial interests in LASE Innovation Inc. SHY has financial interests in LASE Innovation Inc. that were reviewed and are managed by Massachusetts General Hospital and Mass General Brigham in accordance with their conflict-of-interest policies. VV is currently an employee and has financial interest in Vertex Pharmaceuticals, Inc. All other authors declare they have no competing interests.

References

Aigouy, B., Farhadifar, R., Staple, D.B., Sagner, A., Röper, J.-C., Jülicher, F., Eaton, S., 2010. Cell flow reorients the axis of planar polarity in the wing epithelium of Drosophila. Cell 142, 773–786.

Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J.I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F.A., Kreshuk, A., 2019. ilastik: interactive machine learning for (bio)image analysis. Nat. Methods 16, 1226–1232.

Besson, S., Dumais, J., 2011. Universal rule for the symmetric division of plant cells. Proc. Natl. Acad. Sci. U. S. A. 108, 6294–6299.

Borthwick, D.W., Shahbazian, M., Krantz, Q.T., Dorin, J.R., Randell, S.H., 2001. Evidence for stem-cell niches in the tracheal epithelium. Am. J. Respir. Cell Mol. Biol. 24, 662–670.

Concha, M.L., Adams, R.J., 1998. Oriented cell divisions and cellular morphogenesis in the zebrafish gastrula and neurula: a time-lapse analysis. Development 125, 983–994.

Diamond, I., Owolabi, T., Marco, M., Lam, C., Glick, A., 2000. Conditional gene expression in the epidermis of transgenic mice using the tetracycline-regulated

transactivators tTA and rTA linked to the keratin 5 promoter. J. Invest. Dermatol. 115, 788–794.

Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., Linsley, P.S., Gottardo, R., 2015. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol. 16, 278.

Fukumoto, J., Soundararajan, R., Leung, J., Cox, R., Mahendrasah, S., Muthavarapu, N., Herrin, T., Czachor, A., Tan, L.C., Hosseinian, N., Patel, P., Gone, J., Breitzig, M.T., Cho, Y., Cooke, A.J., Galam, L., Narala, V.R., Pathak, Y., Lockey, R.F., Kolliputi, N., 2016. The role of club cell phenoconversion and migration in idiopathic pulmonary fibrosis. Aging 8, 3091–3109.

Hu, K.H., Eichorst, J.P., McGinnis, C.S., Patterson, D.M., Chow, E.D., Kersten, K., Jameson, S.C., Gartner, Z.J., Rao, A.A., Krummel, M.F., 2020. ZipSeq: barcoding for real-time mapping of single cell transcriptomes. Nat. Methods 17, 833–843.

Kim, J.H., Pegoraro, A.F., Das, A., Koehler, S.A., Ujwary, S.A., Lan, B., Mitchel, J.A., Atia, L., He, S., Wang, K., Bi, D., Zaman, M.H., Park, J.-A., Butler, J.P., Lee, K.H., Starr, J.R., Fredberg, J.J., 2020. Unjamming and collective migration in MCF10A breast cancer cell lines. Biochem. Biophys. Res. Commun. 521, 706–715.

Kim, J.K., Vinarsky, V., Wain, J., Zhao, R., Jung, K., Choi, J., Lam, A., Pardo-Saganta, A., Breton, S., Rajagopal, J., Yun, S.H., 2012. In VivoImaging of Tracheal Epithelial Cells in Mice during Airway Regeneration. American Journal of Respiratory Cell and Molecular Biology. https://doi.org/10.1165/rcmb.2012-0164oc

Konen, J., Summerbell, E., Dwivedi, B., Galior, K., Hou, Y., Rusnak, L., Chen, A., Saltz, J., Zhou, W., Boise, L.H., Vertino, P., Cooper, L., Salaita, K., Kowalski, J., Marcus, A.I., 2017. Image-guided genomics of phenotypically heterogeneous populations reveals vascular signalling during symbiotic collective cancer invasion. Nat. Commun. 8, 15078.

Kumar, P.A., Hu, Y., Yamamoto, Y., Hoe, N.B., Wei, T.S., Mu, D., Sun, Y., Joo, L.S., Dagher, R., Zielonka, E.M., Wang, D.Y., Lim, B., Chow, V.T., Crum, C.P., Xian, W., McKeon, F., 2011. Distal Airway Stem Cells Yield Alveoli In Vitro and during Lung Regeneration following H1N1 Influenza Infection. Cell. https://doi.org/10.1016/j.cell.2011.10.001

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L.E., Liu, Z., van Bruggen, D., Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco, G., Cramer, P., Adameyko, I., Linnarsson, S., Kharchenko, P.V., 2018. RNA velocity of single cells. Nature 560, 494–498.

Lane, K., Van Valen, D., DeFelice, M.M., Macklin, D.N., Kudo, T., Jaimovich, A., Carr, A., Meyer, T., Pe'er, D., Boutet, S.C., Covert, M.W., 2017. Measuring Signaling and RNA-Seq in the Same Cell Links Gene Expression to Dynamic Patterns of NF-κB Activation. Cell Syst 4, 458-469.e5.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25.

Lee, Y., Bogdanoff, D., Wang, Y., Hartoularos, G.C., Woo, J.M., Mowery, C.T., Nisonoff, H.M., Lee, D.S., Sun, Y., Lee, J., Mehdizadeh, S., Cantlon, J., Shifrut, E., Ngyuen, D.N., Roth, T.L., Song, Y.S., Marson, A., Chow, E.D., Ye, C.J., 2021. XYZeq: Spatially resolved single-cell RNA sequencing reveals expression heterogeneity in the tumor microenvironment. Sci Adv 7. https://doi.org/10.1126/sciadv.abg4755

Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323.

Marx, V., 2021. Method of the Year: spatially resolved transcriptomics. Nat. Methods 18, 9–14.

Miller, R.L., Zhang, P., Smith, M., Beaulieu, V., Paunescu, T.G., Brown, D., Breton, S., Nelson, R.D., 2005. V-ATPase B1-subunit promoter drives expression of EGFP in intercalated cells of kidney, clear cells of epididymis and airway cells of lung in transgenic mice. Am. J. Physiol. Cell Physiol. 288, C1134-44.

Minc, N., Piel, M., 2012. Predicting division plane position and orientation. Trends Cell Biol. 22, 193–200.

Montoro, D.T., Haber, A.L., Biton, M., Vinarsky, V., Lin, B., Birket, S.E., Yuan, F., Chen, S., Leung, H.M., Villoria, J., Rogel, N., Burgin, G., Tsankov, A.M., Waghray, A., Slyper, M., Waldman, J., Nguyen, L., Dionne, D., Rozenblatt-Rosen, O., Tata, P.R., Mou, H., Shivaraju, M., Bihler, H., Mense, M., Tearney, G.J., Rowe, S.M., Engelhardt, J.F., Regev, A., Rajagopal, J., 2018. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature 560, 319–324.

Pardo-Saganta, A., Law, B.M., Tata, P.R., Villoria, J., Saez, B., Mou, H., Zhao, R., Rajagopal, J., 2015a. Injury induces direct lineage segregation of functionally distinct airway basal stem/progenitor cell subpopulations. Cell Stem Cell 16, 184–197.

Pardo-Saganta, A., Tata, P.R., Law, B.M., Saez, B., Chow, R.D.-W., Prabhu, M., Gridley, T., Rajagopal, J., 2015b. Parent stem cells can serve as niches for their daughter cells. Nature 523, 597–601.

Park, J.-A., Kim, J.H., Bi, D., Mitchel, J.A., Qazvini, N.T., Tantisira, K., Park, C.Y., McGill, M., Kim, S.-H., Gweon, B., Notbohm, J., Steward, R., Jr, Burger, S., Randell, S.H., Kho, A.T., Tambe, D.T., Hardin, C., Shore, S.A., Israel, E., Weitz, D.A., Tschumperlin, D.J., Henske, E.P., Weiss, S.T., Manning, M.L., Butler, J.P., Drazen, J.M., Fredberg, J.J., 2015. Unjamming and cell shape in the asthmatic airway epithelium. Nat. Mater. 14, 1040–1048.

Park, M., Richardson, A., Pandzic, E., Lobo, E.P., Whan, R., Watson, S.L., Lyons, J.G., Wakefield, D., Di Girolamo, N., 2019. Visualizing the Contribution of Keratin-14+ Limbal Epithelial Precursors in Corneal Wound Healing. Stem Cell Reports 12, 14–28.

Park, S., Gonzalez, D.G., Guirao, B., Boucher, J.D., Cockburn, K., Marsh, E.D., Mesa, K.R., Brown, S., Rompolas, P., Haberman, A.M., Bellaïche, Y., Greco, V., 2017. Tissue-scale coordination of cellular behaviour promotes epidermal wound repair in live mice. Nat. Cell Biol. 19, 155–163.

Plasschaert, L.W., Žilionis, R., Choo-Wing, R., Savova, V., Knehr, J., Roma, G., Klein, A.M., Jaffe, A.B., 2018. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature 560, 377–381.

Rawlins, E.L., Hogan, B.L.M., 2008. Ciliated epithelial cell lifespan in the mouse trachea and lung. Am. J. Physiol. Lung Cell. Mol. Physiol. 295, L231-4.

Rawlins, E.L., Okubo, T., Xue, Y., Brass, D.M., Auten, R.L., Hasegawa, H., Wang, F., Hogan, B.L.M., 2009. The role of Scgb1a1+ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. Cell Stem Cell 4, 525–534.

Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, Å., Pontén, F., Costea, P.I., Sahlén, P., Mulder, J., Bergmann, O., Lundeberg, J., Frisén, J., 2016. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 353, 78–82.

Tadokoro, T., Wang, Y., Barak, L.S., Bai, Y., Randell, S.H., Hogan, B.L.M., 2014. IL-6/STAT3 promotes regeneration of airway ciliated cells from basal stem cells. Proc. Natl. Acad. Sci. U. S. A. 111, E3641-3649.

Thielicke, W., Stamhuis, E.J., 2014. PIVlab – Towards User-friendly, Affordable and Accurate Digital Particle Image Velocimetry in MATLAB. Journal of Open Research Software. https://doi.org/10.5334/jors.bl

Tomura, M., Yoshida, N., Tanaka, J., Karasawa, S., Miwa, Y., Miyawaki, A., Kanagawa, O., 2008. Monitoring cellular movement in vivo with photoconvertible fluorescence protein “Kaede” transgenic mice. Proc. Natl. Acad. Sci. U. S. A. 105, 10871–10876.

Tsou, M.-F.B., Ku, W., Hayashi, A., Rose, L.S., 2003. PAR-dependent and geometry-dependent mechanisms of spindle positioning. J. Cell Biol. 160, 845–855.

Van Der Maaten, L., 2014. Accelerating t-SNE using tree-based algorithms. J. Mach. Learn. Res. 15, 3221–3245.

van der Maaten, L., 2008. Visualizing Data using t-SNE [WWW Document]. URL https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf?fbclid=IwAR0Bgg1eA5TFmqOZeCQXsIoL6PKrVXUFaskUKtg6yBhVXAFFvZA6yQiYx-M (accessed 2.23.21).

Yuan, J., Sheng, J., Sims, P.A., 2018. SCOPE-Seq: a scalable technology for linking live cell imaging and single-cell RNA sequencing. Genome Biol. 19, 227.

Zhang, Y., Huang, G., Shornick, L.P., Roswit, W.T., Shipley, J.M., Brody, S.L., Holtzman, M.J., 2007. A transgenic FOXJ1-Cre system for gene inactivation in ciliated epithelial cells. Am. J. Respir. Cell Mol. Biol. 36, 515–519.

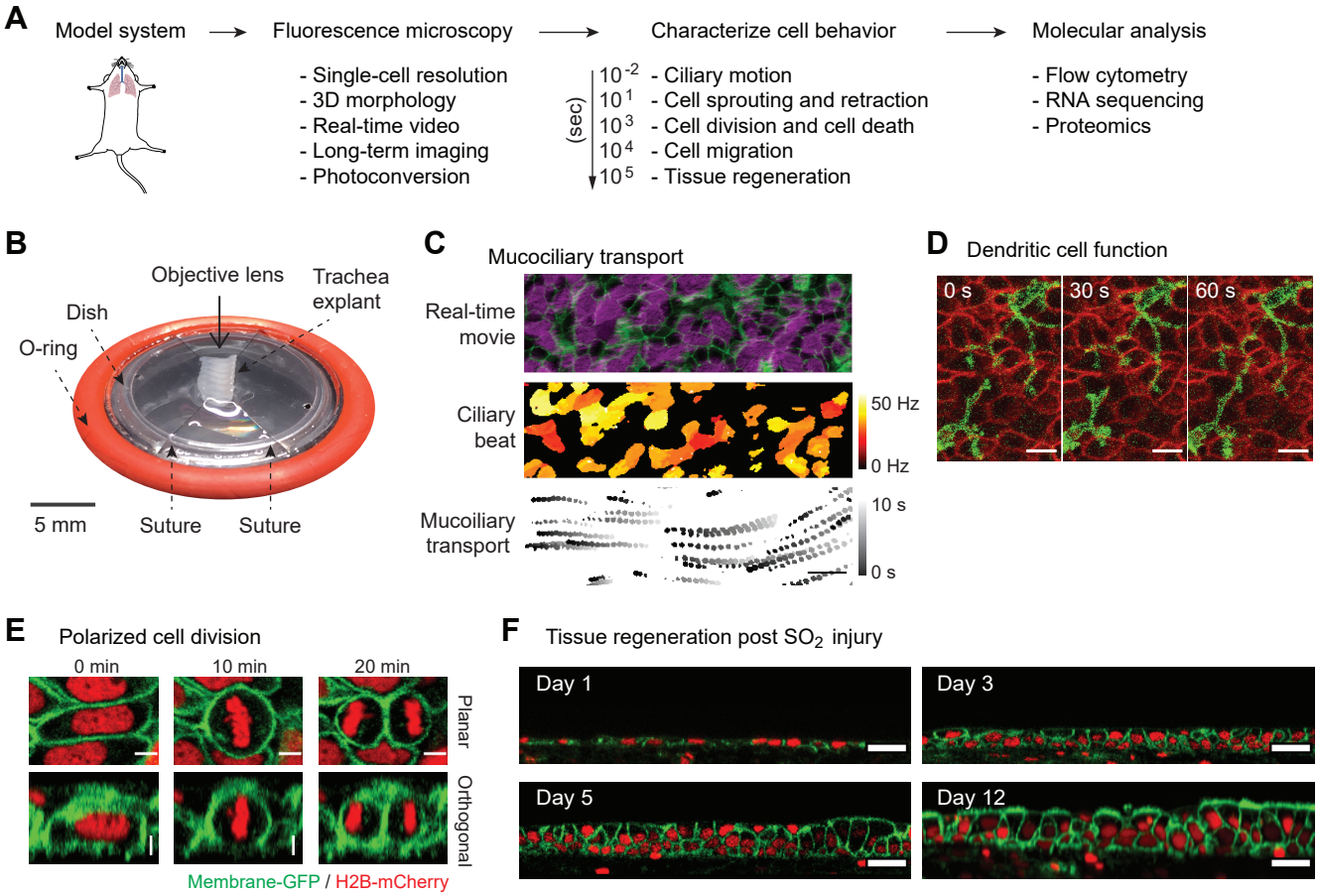

Figure 1. Platform for live imaging of airway tissue. (A) Behavioral transcriptomics workflow, starting with imaging, followed by image analysis to characterize cellular behavior over different time frames, leading to single cell applications. **(B)** Airway tissue (trachea) is explanted from a mouse and affixed to a custom platform for long-term air-liquid-interface (ALI) culture and imaging. The platform enables both time-lapse microscopy and downstream single-cell applications. **(C)** Imaging and image analysis of ciliary beating and mucociliary transport 1 µm spherical beads. **(D)** Intraepithelial dendritic cells (CD11cCre-MTMG) grow and retract dendrites in real time; scale bar = 5 µm **(E)** Selected snapshots of cell division during regeneration post SO2 injury. Epithelial cell divides along its long axis during regeneration (Hertwig's rule); scale bar = 5 µm. **(F)** Long-term ALI culture enables imaging of tissue regeneration post SO2 injury over > 12 days. ALI culture enables regeneration of entire epithelial thickness; scale bar = 20 µm. Green = membrane-GFP; red = nuclear-tdTomato.

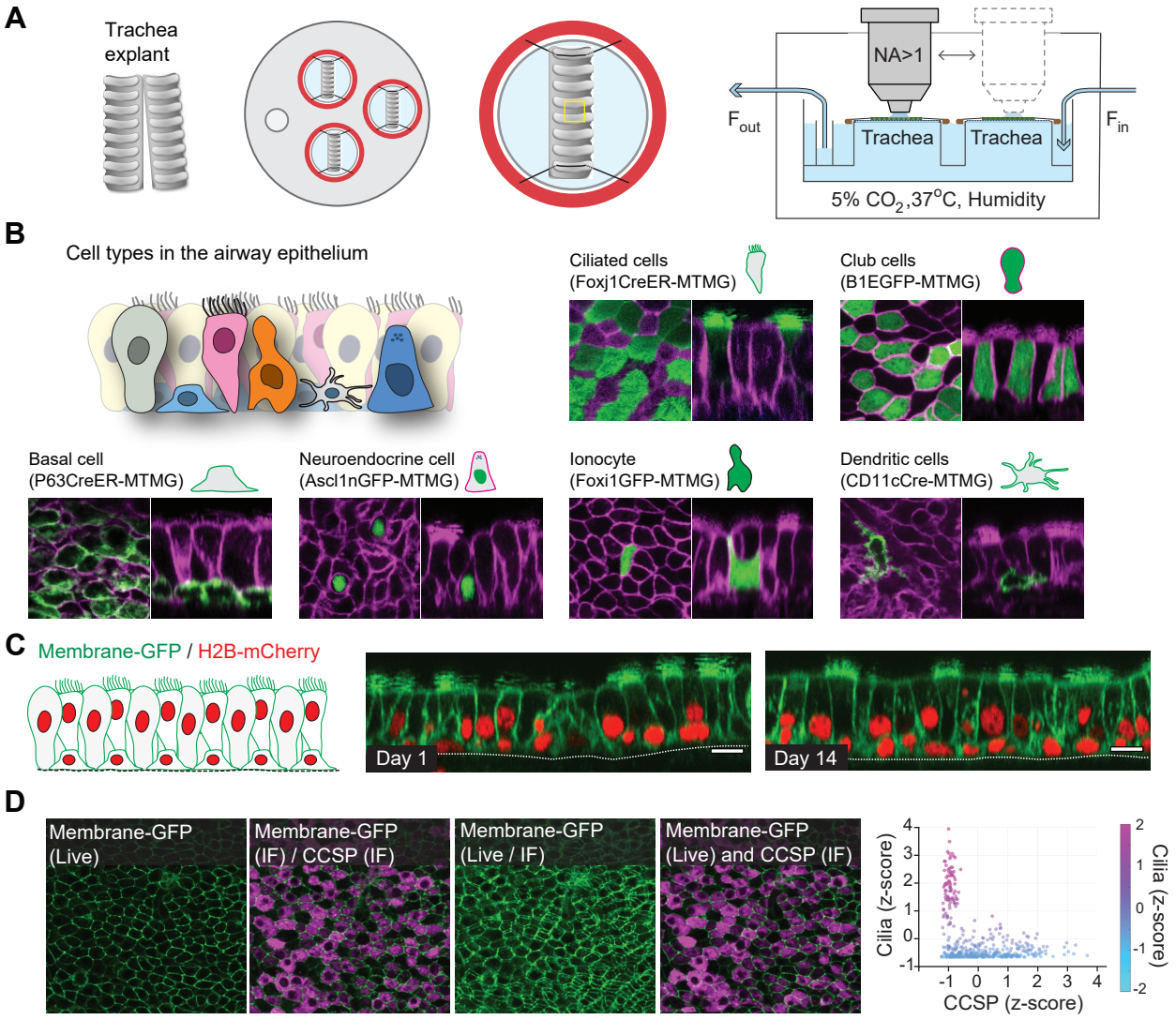

Figure 1 – figure supplement 1. Platform for live imaging of airway tissue. (A) Detailed schematic of ALI platform for airway tissue imaging. The platform enables imaging of multiple tissue explants at a time in a controlled environment maintaining ALI, temperature, humidity, and CO2 content. (B) Representative images of airway epithelial cells using different transgenic mouse models including fluorescent reporters for ciliated, club, basal, neuroendocrine, ionocyte, and dendritic cells. Green = cell type-specific GFP reporter; magenta = membrane-tdTomato. Scale bar = 10 μm. (C) Representative two-photon images of the same airway tissue explanted from a membrane-GFP/H2B-mCherry transgenic mouse at day 1 and day 14 in ALI culture. Scale bar = 10 μm. (D) Registration of live imaging with post-fixation-staining imaging using in silico tissue flattening followed by non-rigid 3D registration. Ciliated cells identified by live imaging were found to have low CCSP expression, while cells with high CCSP expression tended to have no cilia.

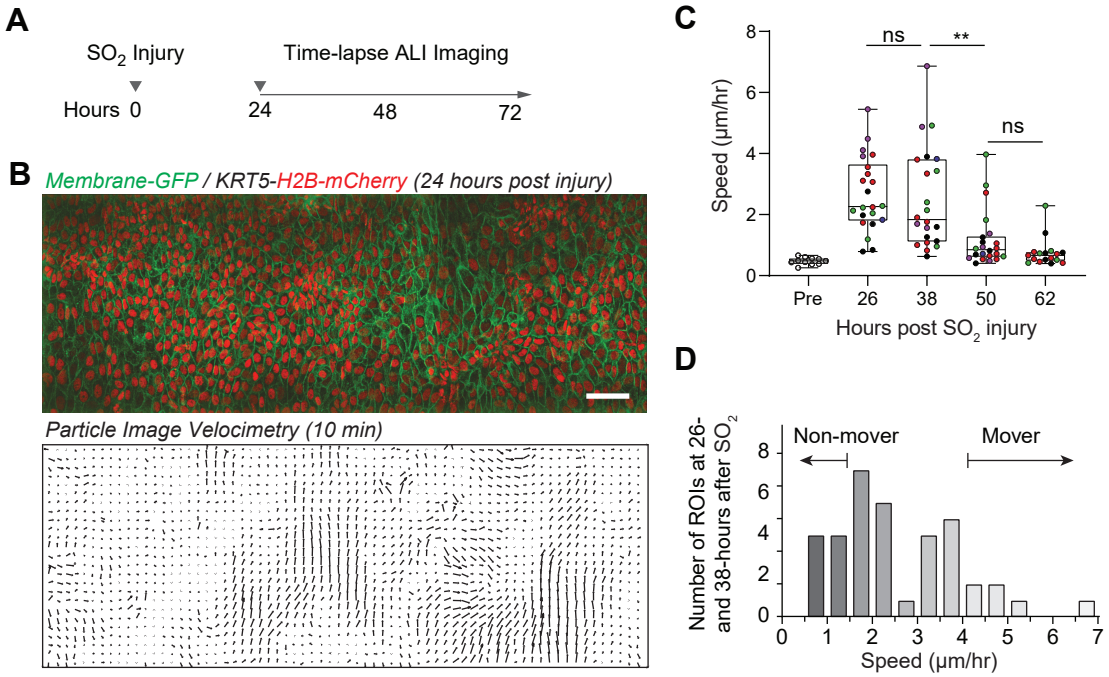

Figure 2. Live imaging enables observation of movement of regenerating airway epithelial cells. (A) Experimental design: tracheas are explanted 24 hours post SO2 injury for continuous time-lapse imaging. **(B)** Two-photon imaging of trachea epithelium from membrane-GFP, KRT5-H2B-mCherry transgenic mouse. Top image is a stitch of three areas. Bottom image shows displacement vectors over 10 minutes computed using particle imaging velocimetry (PIV). Scale bar = 50 μm. **(C)** Computed speed of epithelial cells measured at different time points post SO2 injury in 22 independent regions from a total of 5 mice at 4 different time points (mouse origin is color-coded). A two-way ANOVA was run to examine the effect of time post SO2 injury and different mice on the mean speed determined by PIV. There were 22 ROIs analyzed from 5 mice over 4 time-points. There was a significant interaction between time and the mean speed, $F(2.219,42.91)=16.12$, $p<0.0001$, but no significant difference between mouse and mean speed, $F(4,17)=2.193$, $p=0.113$. A Tukey post-hoc test revealed significant pairwise differences between 26 and 50 hr, 26 and 62 hr, 38 and 50 hr, as well as 38 and 62 hr. ** $p<0.01$. **(D)** Frequency distribution of injury-induced cell movements measured at 26- and 38-hours after injury identifies “mover” and “non-mover” regions.

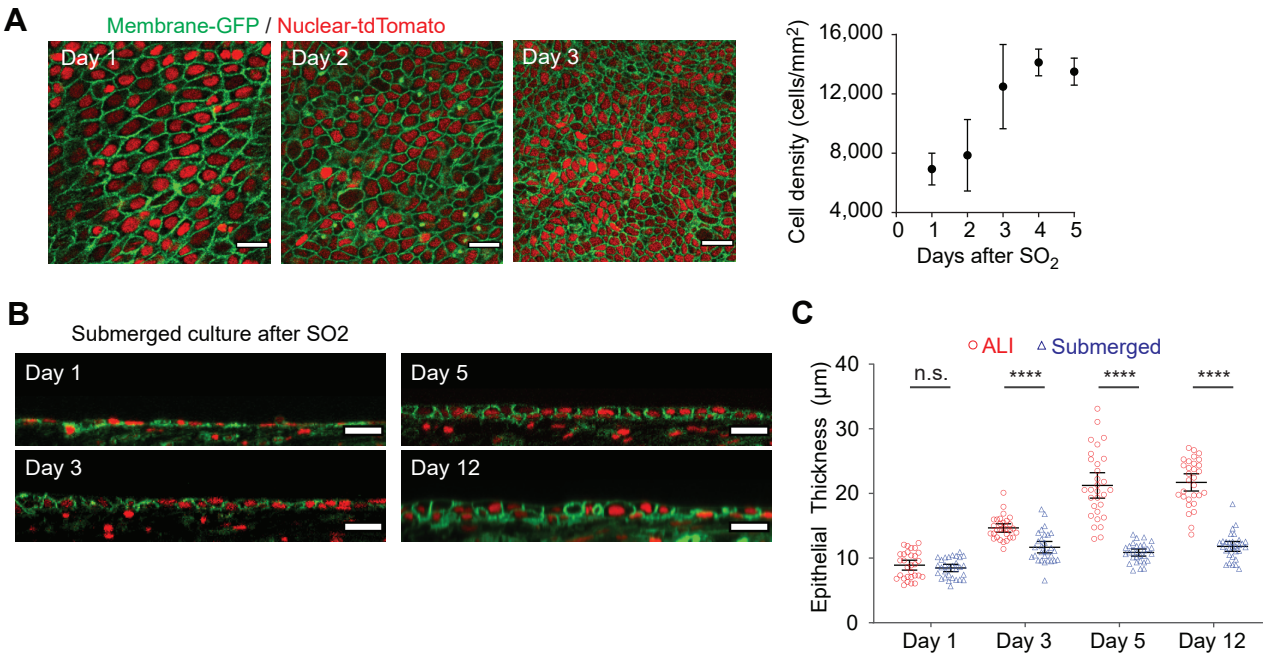

Figure 2 – figure supplement 1. Epithelial regeneration after sulfur dioxide injury *ex vivo*. (A) Quantification of cell density measured post SO2 injury using membrane-GFP/nuclear-tdTomato mice. Scale bar = 20 µm. n=5, error bars indicate the standard deviation. (B) Long - term imaging of tissue regeneration post SO2 injury in a submerged culture condition. (C) Comparison of tissue regeneration between ALI and submerged culture. Compared to submerged culture, ALI culture enables regeneration of the full epithelial thickness. Green = membrane-GFP; red = nuclear tdTomato. Error bars indicate the 95% confidence intervals. **** p< 0.0001.

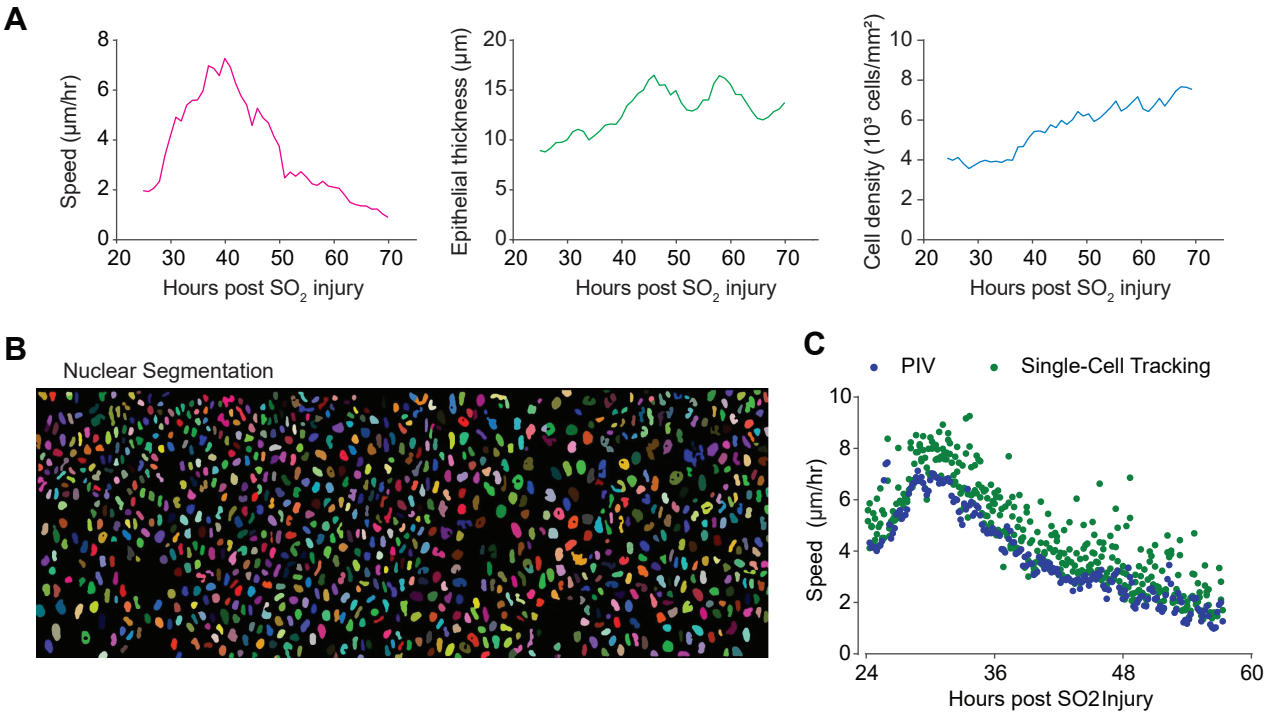

Figure 2 – figure supplement 2. Live imaging enables quantitative analysis of epithelial cell movement over time. (A) Quantification of key parameters quantified by time-lapse imaging of the regenerating epithelium, including speed, epithelial thickness, and cell density. (B) Nuclear segmentation of epithelial cells from KRT5rtTA-H2BmCherry transgenic mouse. (C) Speed of moving cells from time-lapse microscopy computed by particle-imaging-velocimetry (PIV) and single-cell tracking of segmented nuclei. There is no significant difference in the computed speed using different methods.

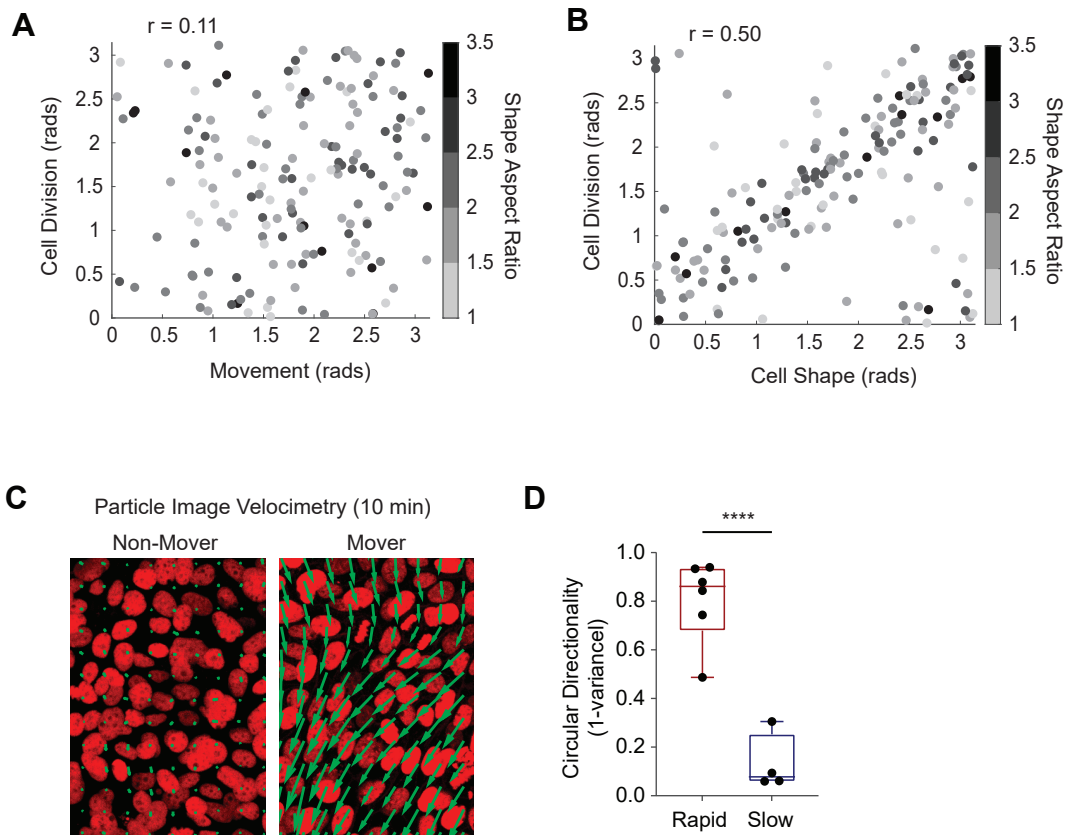

Figure 2 – figure supplement 3. Live imaging with high temporal resolution.
(A) Quantification of cell division behaviors during regeneration. The axis of cell division was found to have minimal correlation (r=0.11) with local movement, (B) but a moderate correlation with the long axis of the cell (r=0.50). (C) Representative movement vectors computed from PIV analysis for a rapid “mover” and a slow “non-mover” region. (D) Cell movement vectors computed from PIV analysis were fit to a von Mises distribution to compute the circular variance. Fast moving cells had a low circular variance (directional), while slow moving cells had a high circular variance (no directional preference). **** $p<0.0001$.

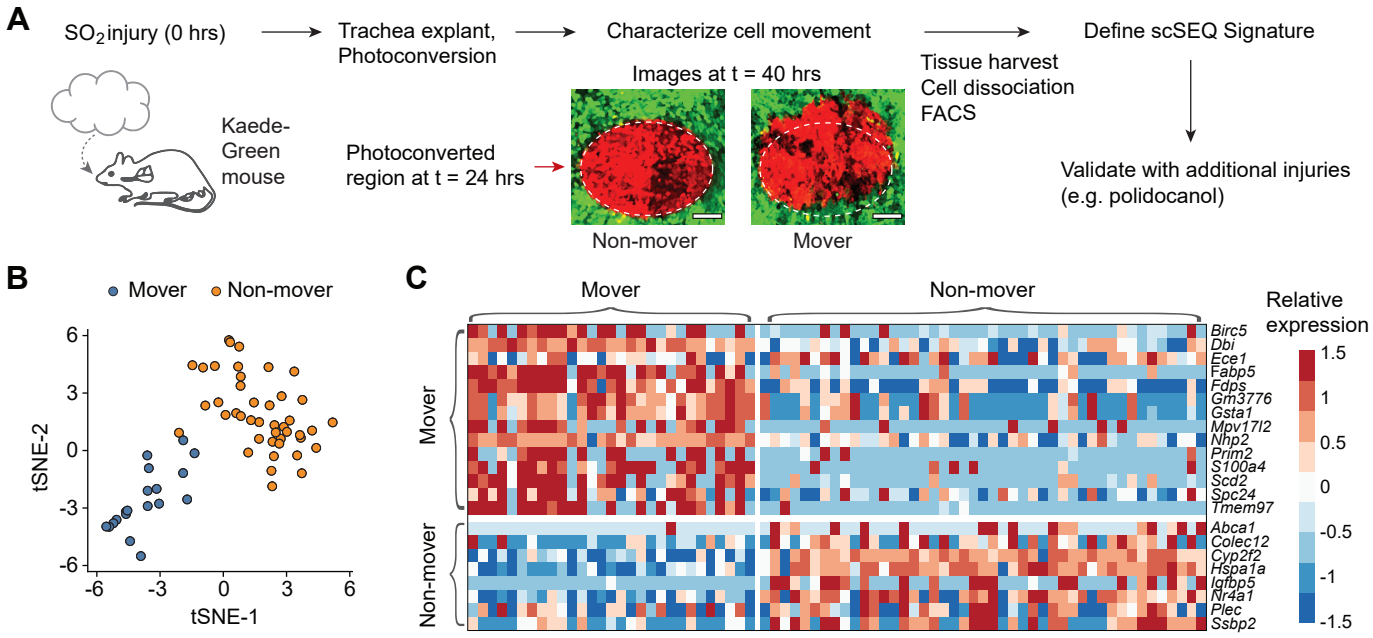

Figure 3. Transcriptionally distinct moving (M) and non-moving (NM) cells coordinate early airway epithelial regeneration across multiple injury types. (a) Experimental design: tracheas are explanted 24 hours post SO_2 injury (24 hpi) for continuous time-lapse two-photon imaging. Distinct cellular phenotypes are observed and labeled by photo-conversion for subsequent isolation and transcriptional analysis by full-length single-cell RNA-sequencing. Scale bar = 100 μm **(b)** Unsupervised clustering of regenerating cells partitions mover and non-mover cell phenotypes. **(c)** Heatmap of transcriptional signatures of mover and non-mover cells.

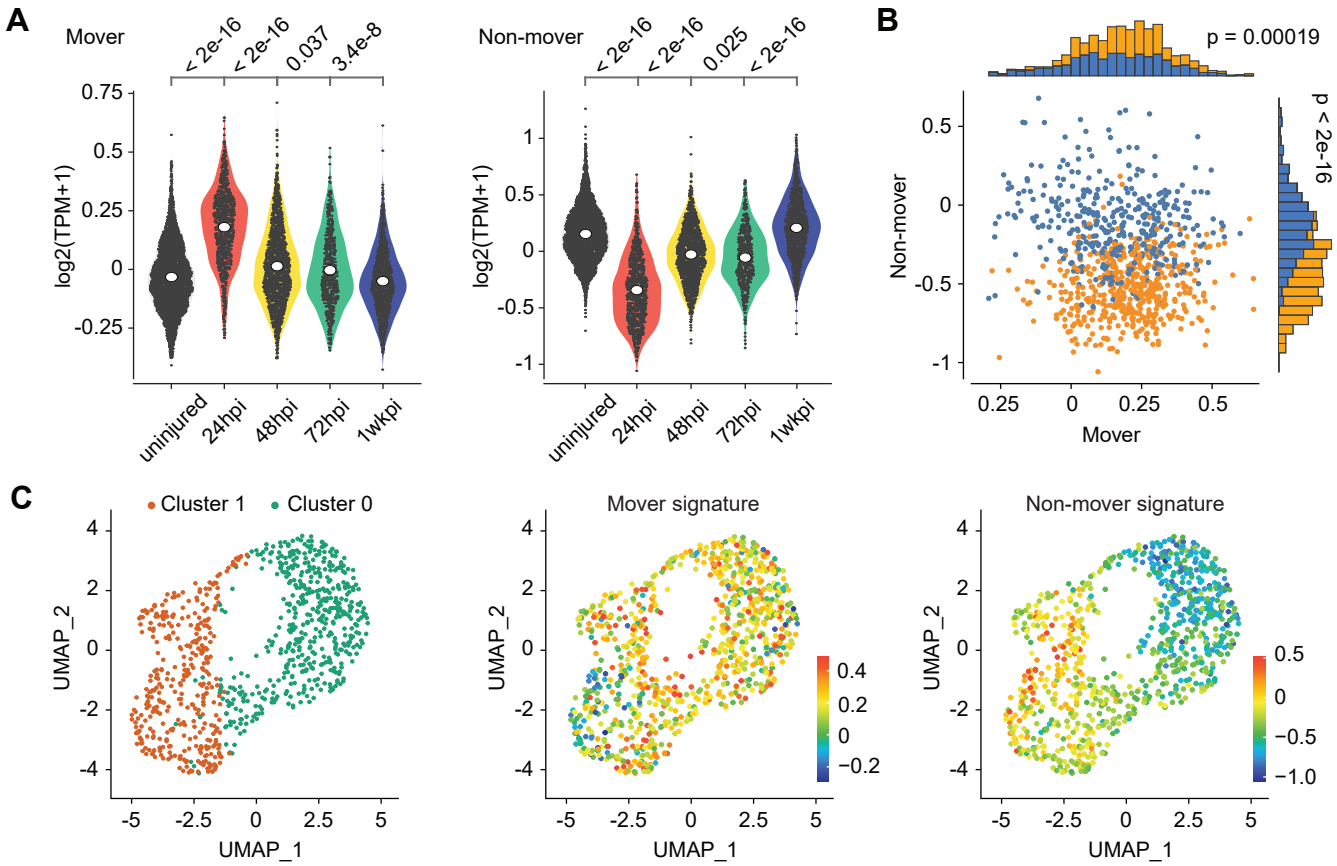

Figure 4. Transcriptionally distinct moving (M) and non-moving (NM) cells coordinate early airway epithelial regeneration across multiple injury types. (A) Mover and non-mover transcriptional signatures are also enriched in early airway epithelial regeneration 24hours post-injury of an independent murine airway injury induced by polidocanol administration. **(B)** Scoring for mover and non-mover transcriptional signatures in 24hpi regenerating cells following polidocanol treatment partitions cells into two populations. **(C)** Unsupervised clustering of 24 hpi regenerating cells yields two cell populations enriched in the expression of mover (cluster 0) or non-mover (cluster 1) signatures. *P* values from Mann-Whitney U test.