nature research

Corresponding author(s): Dr. Jonghwan Kim & Dr. Joo-Hyeon Lee

Last updated by author(s): Jul 15, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Leica Application Suite X (Confocal microscopy, version 2.0.0) BD FACS Diva (Flow cytometry, version 8.0.1) StepOnePlusTM Software (version 2.3) Hiseq Software Suite (version 3.4.0)
Data analysis	FlowJo software (version: 8.72, Tree Star) Fiji software (ImageJ, version: 2.00-rc-69/1/52i) GraphPad Prism 7.0 R package DESeq2 (version 1.28.0) Java Treeview (version 1.1.6r4) Integrative Genomics Viewer (IGV, version 2.7.2) MAnorm (version 1.1.4) Bowtie2 (version 2.3.4.1) MACS2 (version 2.1.2) HOMER (version 4.11) GREAT (version 4.0.4) STAR (version 2.5.2b) Cell Ranger Single-Cell Software Suite (version 2.0.2, 10x Genomics Inc) python package scanpy (version 1.5.1)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

ATAC-seq, RNA-seq and scRNA-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) under accession codes GSE153677 (RNA-seq and ATACC-seq for organoids) and GSE154218 (scRNA-seq). Previous published sequencing data that were re-analysed here are available under accession code GAS0000100434445, GSE13589344, and GSE14455330. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size for animal experiments was made as large as possible, but usually determined based upon pilot experiments. For the injury experiments, the baseline mouse number is N=3 by considering the variation in response to injury from a series of prelimary studies. All replicates reported in the manuscript are biological replicates. Cohort size of mice was designed to be sufficient to enable accurate determination of statistical significance. For in vitro experiments including organoids culture, sample size was determined based on our previous experiences an all experiments used more than duplicated biological replicates.
Data exclusions	No animals were excluded from the statistical analysis
Replication	All experiments for this study are from multiple biological and technical replicates and are reported in the legend of each Figure.
Randomization	All experiments include age- and sex-matched mice assigned randomly to all experimental and control groups. For in vitro assays including organoids culture, all groups of control and treatment were allocated randomly as well.
Blinding	Blinding to experimental condition during injury experiments using bleomycin was challenging and difficult as only few investigator was allowed to do the surgery. However, information about the mice used for experiments is known to all investigators through the online system. Blinding to condition for quantitative counting and scoring was attempted but is not possible given the severity/heterogeinity of injury and the morphological differences in injured tissues. For flow cytometry and computational analysis in sequencing experiments, as only few investigators were able to access to machines and analyze acquired/processed files, those experiments were not suitable for blinding.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For flow cytometry, the antibodies used are as follows:

Antibodies used

For mouse: CD45 (30-F11)-APC or -APC-Cy7 (BD Biosciences), CD31 (MEC13.3)-APC (BD Biosciences), EpCAM (G8.8)-PE-Cy7 or FITC (BioLegend), and CD24 (M1/69)-APC (eBioscience), MHC-II (I-A/I-E, M5)-FITC or -APC-Cy7 (eBioscience), and CD309/Kdr/Flk-1 (7D4-6)-APC (BioLegend). The titer of antibodies used for sorting was 1:200.

For human: CD45 (2D1)-APC (BioLegend), CD31 (VM59)-APC (BioLegend), EpCAM (9C4)- FITC (BioLegend), HTII-280-mouse IgM (Terrace Biotech), Purified CD309/KDR (A16085H), anti-mouse IgG1 (RMG1-1)-APC-Cy7 (BioLegend), and anti-mouse IgM (II/41)-PE (eBioscience). All antibodies was diluted in 1:200.

For immunostaining, the antibodies used are as follows:
Goat anti-CCSP/CC10 (T-18) (1:200, Santa Cruz Biotechnology Inc., sc-9772), Mouse anti-CCSP/CC10 (E11) (1:200, Santa Cruz Biotechnology Inc., sc-365992), Goat anti-SP-C (1:200, Santa Cruz Biotechnology Inc., sc-7706), Rabbit anti-pro-SP-C (1:300, Millipore, AB3786), Mouse anti-Acetylated Tubulin (6-11B-1) (1:300, Sigma-Aldrich, T7451), Hamster anti-PDPN (1:300, DSHB, 8.1.1), Rabbit anti-KRT5 (1:300, BioLegend, 905501), Mouse anti-P63 (1:200, Abcam, ab735), Rabbit anti-PORCN (1:100, Abcam, ab201793), Rabbit anti-P57/KIP2 (1:100, Abcam, ab75974), Rabbit anti-RFP (1:250, Rockland, 600–401379), Rabbit anti-HOPX (1:100, Santa Cruz Biotechnology Inc., sc-30216), Rabbit anti-HES1 (1:200, D6P2U, #11988, Cell Signaling), Rabbit anti-NOTCH1 (1:50, Abcam, ab52627), Mouse IgM anti-HT2-280 (1:300, Terrace Biotech, TB-27AHT2-280), Rat anti-human SCGB1A1/CC10 (1:200, R&D system, MAB4218), Rabbit anti-KDR/VEGFR2 (1:100, Cell Signaling, 2479). Alexa Fluor-coupled secondary antibodies (1:500, Invitrogen). Anti-human Jag1 antibody (1:100, Santa Cruz Biotechnology Inc., sc-390177).

Validation

Antibodies were well validated by the manufacturer and previous our studies published in other journalsas well.
Antibodies used for mouse cell sorting:
CD45 (30-F11, BD Bioscience) was validated from FACS staining using mouse splenocytes (https://www.biolegend.com/en-us/search-results/apc-anti-mouse-cd45-antibody-97).
CD31 (MEC13.3, BD Bioscience) was validated from staining using FACS staining using mouse splenocytes (https://www.biolegend.com/de-de/products/apc-anti-mouse-cd31-antibody-375).
EpCAM (G8.8, BioLegend) was validated from FACS staining using TE-71 (mouse thymic epithelial stromal cell line) cells (https://www.biolegend.com/en-us/products/pe-anti-mouse-cd326-ep-cam-antibody-4726?GroupID=BLG5748).
CD24 (M1/69, eBioscience) was validated from FACS staining using mouse splenocytes (https://www.thermofisher.com/antibody/product/CD24-Antibody-clone-M1-69-Monoclonal/17-0242-82).
MHC-II (I-A/I-E, eBioscience) was validated from FACS staining using mouse splenocytes (https://www.thermofisher.com/antibody/product/MHC-Class-II-I-A-I-E-Antibody-clone-M5-114-15-2-Monoclonal/14-5321-82).
CD309 (7D4-6, BioLegend) was validated from FACS staining using bEnd.3 endothelial cells (https://www.biolegend.com/en-us/products/apc-anti-mouse-cd309-vegfr2-flk-1-antibody-6472?GroupID=BLG8270).

Antibodies used for human cell sorting:
CD45 (2D1, BioLegend) was validated from FACS staining using human peripheral blood lymphocytes (https://www.biolegend.com/en-us/products/apc-anti-human-cd45-antibody-12397?GroupID=BLG14850).
CD31 (VM59, BioLegend) was validated from FACS staining using human peripheral blood lymphocytes (https://www.biolegend.com/en-us/products/apc-anti-human-cd31-antibody-6123).
EpCAM (9C4, BioLegend) was validated from FACS staining using HT29 human colon carcinoma cell line (https://www.biolegend.com/en-us/search-results/fitc-anti-human-cd326-epcam-antibody-3756?GroupID=BLG5134).
HTII-280-mouse IgM (Terrace Biotech) was validated using human lung alveolar type 2 cells (https://www.terracebiotech.com/product-page/anti-ht2-280-1ml).
Purified CD309/KDR (A16085H, BioLegend) was validated using HUVEC human endothelial cells (https://www.biolegend.com/en-gb/products/purified-anti-human-cd309-antibody-15471?GroupID=GROUP524).
Anti-mouse IgG1 (RMG1-1, BioLegend) was validated using human peripheral lymphocytes (https://www.biolegend.com/en-us/products/apc-anti-mouse-igg1-7022?GroupID=BLG3729).
Anti-mouse IgM (II/41, eBioscience) was validated using mouse splenocytes (thermofisher.com/antibody/product/IgM-Antibody-clone-II-41-Monoclonal/14-5790-82).

Validation for Antibodies used for immunostaining:
Goat anti-CCSP/CC10 (T-18,Santa Cruz Biotechnology Inc., sc-9772) was validated from manufacturer (https://www.scbt.com/p/cc10-antibody-t-18).
Mouse anti-CCSP/CC10 (E11,Santa Cruz Biotechnology Inc., sc-365992) was validated from manufacturer (https://www.scbt.com/p/cc10-antibody-e-11).
Goat anti-SP-C (Santa Cruz Biotechnology Inc., sc-7706),Rabbit anti-pro-SP-C (Millipore, AB3786), Mouse anti-Acetylated Tubulin (6-11B-1,Sigma-Aldrich, T7451), Hamster anti-PDPN (1:300, DSHB, 8.1.1), Rabbit anti-HOPX (1:100, Santa Cruz Biotechnology Inc., sc-30216) were validated from our previous study (https://www.cell.com/cell-stem-cell/fulltext/S1934-5909(20)30287-3?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS1934590920302873%3Fshowall%3Dtrue).
Rabbit anti-KRT5 (1:300, BioLegend, 905501) was validated from manufacturer (https://www.biolegend.com/en-us/products/keratin-5-polyclonal-antibody-purified-10956?GroupID=GROUP26).
Mouse anti-P63 (1:200, Abcam, ab735) was validated from manufacturer (https://www.abcam.com/p63-antibody-4a4-ab735.html).
Rabbit anti-PORCN (1:100, Abcam, ab201793) was validated from manufacturer (https://www.abcam.com/porcnppn-antibody-epr19741-ab201793.html).
Rabbit anti-P57/KIP2 (1:100, Abcam, ab75974) was validated from manufacturer (https://www.abcam.com/p57-kip2-antibody-ep2515y-ab75974.html).
Rabbit anti-RFP (1:250, Rockland, 600–401379) was validated from manufacturer (https://rockland-inc.com/store/Antibodies-to-GFP-and-Antibodies-to-RFP-600-401-379-O4L_24299.aspx).
Rabbit anti-HES1 (1:200, D6P2U, #11988, Cell Signaling) was validated from manufacturer (https://www.cellsignal.co.uk/products/primary-antibodies/hes1-d6p2u-rabbit-mab/11988).
Rabbit anti-NOTCH1 (1:50, Abcam, ab52627) was validated from manufacturer (https://www.abcam.com/notch1-antibody-ab27526.html).
Mouse IgM anti-HT2-280 (1:300, Terrace Biotech, TB-27AHT2-280) was validated from manufacturer (https://www.terracebiotech.com/product-page/anti-ht2-280-1ml).
Rat anti-human SCGB1A1/CC10 (1:200, R&D system, MAB4218) was validated from manufacturer (https://www.rndsystems.com/products/human-uteroglobin-scgb1a1-antibody-394324_mab4218).

Rabbit anti-KDR/VEGFR2 (1:100, Cell Signaling, 2479) was validated from manufacturer (https://www.cellsignal.co.uk/products/primary-antibodies/vegf-receptor-2-55b11-rabbit-mab/2479).
Anti-human Jag1 antibody (1:100, Santa Cruz Biotechnology Inc., sc-390177) was validated from manufacturer (https://www.scbt.com/p/jagged1-antibody-e-12).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	293T cells (ATCC: CRL-11268), Human kidney cells
Authentication	No further in-house authentication of the cells was further performed.
Mycoplasma contamination	The cell lines were tested fro mycoplsama contamination using PCR kit before experiments.
Commonly misidentified lines (See ICLAC register)	Our study did not involve any misidentified cell lines.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Scgb1a1-CreERTM, Sftpc-CreERT, Foxj1-CreERT2, Red2-NotchN1ICD, Rosa26R-CAG-fGFP, Rosa26R-lox-stop-lox-tdTomato, and Il1r1flox/flox mice have been described and are available from Jackson Laboratory. Il1r1-P2A-eGFP-IRES-CreERT2 (Il1r1-CreERT2) and Sftpc-IRES-DTR-P2A-dsRed (Sftpc-dsRedIRES-DTR) mice were generated in our laboratory. All mice were bred and maintained under specific-pathogen-free conditions in individually ventilated cages. All animals were placed under a 12hr light-dark cycle. Room temperature was maintained at 22 °C ± 1 °C with 30–70% humidity. Mice for the lineage tracing and injury experiments were on a C57BL/6 background and 6-10 weeks old mice and both sexes were used for most of the experiments. None of the mice were involved in any previous procedures before the study. To label the cells, Tamoxifen (Sigma) was dissolved in Mazola corn oil (Sigma) in a 20mg/ml stock solution. 0.2mg/g body weight tamoxifen was given via intraperitoneal (IP) injection. At indicated time points lungs were collected for analysis. For bleomycin injury, 6-10 weeks mice were anesthetised via inhalation of isoflurane for approximately 3 mins. The mice were positioned on the intratracheal intubation stand, and 1.25U/kg of bleomycin, or PBS control, were delivered intratracheally by a catheter (22G). During the procedure anaesthesia was maintained by isoflurane and oxygen delivery.
Wild animals	No wild animals were used in the study.
Field-collected samples	No field-collected samples were used in the study.
Ethics oversight	Experiments were approved by local ethical review committees and conducted according to UK Home Office project license PC7F8AE82.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	There is no specific ages and gender for human sample collection. Samples were collected from all ethnic groups.
Recruitment	We didn't recruit patients specifically for this study. All samples were collected from the healthy region in the lungs after surgery or transplantation.
Ethics oversight	Human tissue samples were procured and processed under the protocol (REC reference 06/Q0505/12; An investigation into the molecular pathogenesis of lung cancer) approved from University of College London (UCL) and Papworth Hospital Research Tissue Bank (T02233), and Adddenbrookes Hospital (Cambridge University NHS Foundations trust) under the collaboration of Cambridge Biorepository for Translational Medicine (CBTM).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Lung tissues were cleared by perfusion with cold PBS through the right ventricle, and then 2ml of dispase (BD Bioscience, 50U/ml) was instilled into the lung through the trachea until the lungs inflated. The lung tissues were digested with the mixture of collagenase/dispase (Roche), and 1% DNase I, followed by rotating incubation as described in the method section. Cells were passed through sequential filtering with 100um and 40um filter and then red blood cells were removed by ACK lysis buffer. Pelleted cells were resuspended in PF10 buffer (PBS with 10% FBS) for staining.
Instrument	Flexible BD InfluxTM cell sorter (BD Bioscience) for cell sorting and LSRII analyser (BD Biosceince) for analysis.
Software	Raw FCS files were obtained from each experiment and analysed in FlowJo (version: 8.72, Tree Star).
Cell population abundance	In the lung tissue, EPCAM+ lung epithelial cells comprise approximately 15% of single cell suspensions prepared as above and gated as below. Percentages for each fractions including epithelial cells and immune cells are noted in each figure and legend.
Gating strategy	Based on the pattern of FSC-A/SSC-A intact cells were gated and then singlets were gated according to the pattern of FSC-H vs. FSC-A. Positive populations were determined by the specific antibodies, which were distinct from negative populations. Example gating strategies for each sorting experiments are shown in Extended Data Fig. 5a,b; Extended Data Fig. 6a; Extended Data Fig. 7d; Extended Data Fig. 9b,l; Extended Data Fig. 10d).

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.