nature portfolio

Corresponding author(s): Robbert Rottier; NCOMMS-24-27840-T

Last updated by author(s): May 24, 2024

Reporting Summary

Nature Portfolio 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 Portfolio 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	N/A, data collected via website based databases as described in manuscript
Data analysis	In general custom Python scripts are available upon request with exceptions to scripts concerning processing of raw MeD-seq data which are limited under a licensing agreements between Erasmus MC and commercial partners. However scripts regarding downstream analysis of MeD-seq data are available.

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 Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All raw and processed high-throughput sequencing data (MeD-seq, scRNA-seq) generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO) under accession number GSE268414.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status). Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.) Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	*Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."*
Recruitment	*Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.*
Ethics oversight	*Identify the organization(s) that approved the study protocol.*

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

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

Ecological, evolutionary & environmental sciences study design

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

Study description	Embryonics lungs were isolated at different time points of wild type C57/Bl6 mice or transgenic mice containing DCM-Pol2B that were induced with doxycycline. Genomic DNA and RNA was isolated from total lungs or from FACS-sorted cells and processed for RNA-Seq, Med-Seq or histology/IF analysis. For every condition and time point triplicates were generated.
Research sample	Embryonic lungs originate from Mus musculus (C57/Bl6) and genomic DNA or RNA were isolated and analyzed. Also, some lungs were processed for paraffin embedding en analyzed by IF stainings.
Sampling strategy	No sample size calculations were made, variance of data points showed that aiming at commonly used triplicates was sufficient.
Data collection	MeD-seq and RNA-seq data was generated by in house sequencing facility at Erasmus MC which is led by co-author Wilfred van IJcken; IF data were generated and analyzed in house (Prof Rottier lab)
Timing and spatial scale	Timing of doxycycline treatment of the DCM-Pol2B transgenic mouse line was based on known key steps in mouse lung development. Isolation of embryonic lungs of wild type C57/bl6 mice for RNA-Seq and IF was also based on critical points in development.
Data exclusions	Data was only excluded if during sequencing insufficient reads could be generated and/or data complexity was too low. Both are hallmarks of insufficient amounts of DNA or RNA provided for sequencing
Reproducibility	For reproducibility we choose to use standard triplicates. Furthermore multiple time points provided extra power to detect patterns during cell state tracing
Randomization	Randomization was not applicable since we worked with a Doxycycline induced mouse model where the difference between experimental group and control group is determined by administration of Doxycycline (Med-Seq), or by isolation of embryonic lungs at different gestational ages (scRNA-Seq/IF analysis).

Blinding	Blinding was not possible because experimental and control groups were based on Doxycycline induction.

Did the study involve field work? ☐ Yes ☒ No

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
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Antibodies

Antibodies used	Notch3 R&D systems, AF1308; Ebf1 Merck, AB10523; Mcam BioLegend, 134712; Cspg4 Merck, AB5320; Pdgfrb Cell signaling; Technology, 3169; Donkey anti-rabbit HRP Jackson ImmunoResearch, 711-035-152; Donkey anti-rabbit Alexa 488 Jackson; ImmunoResearch, 711-545-152; Donkey anti-rabbit Alexa 647 Jackson ImmunoResearch 711-605-152; Donkey anti-goat Alexa 594 Jackson ImmunoResearch, 705-585-147; Donkey anti-goat Alexa 647 Jackson ImmunoResearch, 705-605-147; Donkey anti-rat Alexa 647 Jackson ImmunoResearch, 712-605-153;
Validation	Validation was performed by Manufacturers, which are provided after the name of the antibody, and in the manuscript

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	Mu musculus; C57/bl6 and DCM-Pol2B transgenic mouse; age 9.5 until 18.5 gestational age;
Wild animals	none
Reporting on sex	For scRNA-Seq and Med-Seq, male pups were selected to avoid sex-bias. For IF analysis, lungs of male and female pups were used
Field-collected samples	none
Ethics oversight	Erasmus MC

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

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*

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	Embryos were harvested at E18.5 and their lungs were isolated. For each induction, three independent cell preparations were processed. Three to four fetal lungs of male embryos were pooled in 750 µl of DMEM (Lonza) and mechanically dissociated them using dissection scissors. Following the addition of collagenase I, II, and IV (Sigma-Aldrich) at a final concentration of 5 µg/ml, the samples were incubated at 37°C in an Eppendorf shaker at 1000 rpm for 10 minutes. Subsequent mechanical dissociation involved vigorous resuspension, after which the cell suspensions were filtered through a 40-µm cell strainer. The filters were rinsed with 750 µl of DMEM with 10% FCS, and the resulting cell suspensions were centrifuged at 4°C for 10 minutes at 1000 rpm. After removing the supernatant, the cells were resuspended in 300 µl of 1x red blood cell lysis buffer at 4°C for 1 minute, diluted them in 900 µl of PBS with 10% FCS, and centrifuged them (10 minutes at 4°C, 1000 rpm). The cells were then resuspended in 300 µl of PBS with 5% FCS. Cells were counted and incubated at 4°C in the dark for 40 minutes with the following antibodies: CD31-PE-Cy7 (1:100, ThermoFisher), CD45-PE-Texas Red (1:100, ThermoFisher), CSPG4-Alexa-488 (1:100, Merck), and PDGFR-β-APC (1:100, ThermoFisher). Prior to sorting, the cells were concentrated and resuspended in 800 µl of PBS with 5% FCS. DAPI (1:10000, BD Biosciences). CD45- cells were sorted for CD31 and for CSPG4/PDGFRB using a BD FACS Aria III and BD FACSDiva software version 8.0.1. DNA was isolated from the sorted cells using QIAamp DNA micro kit according to the manufacturer's protocol, and RNA was isolated by Trizol extraction. At the E8.5-10.10 induction time point, cells were sorted using an unconjugated CSPG4 antibody (1:100, Millipore) for 40 minutes, and subsequently incubated with goat anti-rabbit Alexa 488 secondary antibody (1:200, Jackson) for 30 minutes
Instrument	BD FACS Aria III
Software	BD FACSDiva software version 8.0.1
Cell population abundance	Post-sort abundance was determined using RNA-seq and validated using published scRNA-seq
Gating strategy	Unstained and single stained samples were used to set the gates; strategy was previously published

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