nature research

Corresponding author(s): Fabian J. Theis

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Experimental design

1. Sample size

   Describe how sample size was determined.

   Biological replicates were included at each time-point to establish reproducibility, but no formal power analysis was conducted. For each time-point, we included at least 2 biological replicates that were processed independently. This is enough to ensure sample-level variation was not completely aliased with temporal variation.

2. Data exclusions

   Describe any data exclusions.

   T-cell data: Replicates including female embryos were excluded (greater than 10 cells expressing Xist). Beta-cell data: None. The decision to exclude female embryos was made prior to analysis. The exact criterion (Xist < 10) was set in response to the data.

3. Replication

   Describe the measures taken to verify the reproducibility of the experimental findings.

   T-cell data: In silico separation of blood cells from others was broadly consistent across replicates. We collected multiple biological replicates of the wild-type thymus at each time point, as well as the Rag1 knock-out at day E16.5. For the Rag2 KO, only one replicate was included. We made no attempt to replicate the Rag2 KO results. We show in the thymus atlas paper that other results seem consistent across replicates, for instance, the overall distribution of cells in transcriptome space. There were no discarded unsuccessful attempts at replication; for example, no "outliers" or "bad samples" left out of the study.
   Beta-cell data: We collected replicates from 3 animals for each sampled time point.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   We grouped single-cell RNA-seq samples and total population size observations by developmental time. We only analyzed male embryos to avoid biological confounding by sex.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   The investigators were not blind to group allocation as mice were dissected and genotyped for sex at the predefined developmental time points. The analysts were not blinded to the group (the time point) allocation of the samples as these time labels are necessary to run the dynamics models.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ 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 any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ Test values indicating whether an effect is present *Provide confidence intervals or give results of significance tests (e.g. P values) as exact values whenever appropriate and with effect sizes noted.*
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Alignment, quantification, and in-silico isolation of the blood used the STAR aligner (v2.4.2), Picard tools (v1.96), Dropseq tools (v1.0), Seurat (1.4), and R 3.3.1. The extra TCR FASTA used for alignment will be available as a supplemental file on GEO. We used scanpy (v1.2) with python (v3.6.0). We used PESTO and AMICI and our pseudodynamics scripts in Matlab (2015a and 2017a). We used R (v3.4.1) to generate all figures and monocle (v2.5.7) to compute the monocle cell state ordering. We used Leica LAS AF lite 3.2.0, ImageJ 1.52a / Java 1.8.0_112 (64-bit) to process the stained pancreas sections.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

(Primary antibodies)
Name: Caspase-3 cleaved (Asp175)
Clone: 5A1E
Species: Rabbit
Dilution: 1 to 300
Company: Cell Signaling
Catalog number: 9664
Lot number: 20
Validation: https://media.cellsignal.com/pdf/9664.pdf#_ga=2.152620986.2027245922.1549289384-1498197464.1509104325

Name: Insulin
Clone: Polyclonal
Species: Guinea pig
Dilution: 1 to 300
Company: ABD Serotec
Catalog number: 5330-0104G
Lot number: 1707
Validation: https://www.labome.com/product/Bio-Rad/5330-0104G.html

Name: Ki67
Clone: SP6
Species: Rabbit
Dilution: 1 to 300
Company: Abcam
Catalog number: Ab16667
Lot number: GR322508-2
Validation: https://www.abcam.com/ki67-antibody-sp6-ab16667.html

Name: Urocortin
Clone: Polyclonal
Species: Rabbit
Dilution: 1 to 300
Company: Phoenix Pharmaceuticals
Catalog number: H-019-29
Lot number: 01459-1
Validation: https://www.biocompare.com/Product-Reviews/340931-Urocortin-III-in-mouse-cryosections/

(Secondary antibodies)
Name: Donkey anti-rabbit IgG Alexa 555
Clone: Polyclonal
Species: Donkey
Dilution: 1 to 800
Company: Invitrogen
Catalog number: A31572
Lot number: 1945911
Validation: https://www.thermofisher.com/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-31572

Name: Donkey anti-guinea pig Alexa 647
Clone: Polyclonal
Species: Donkey
Dilution: 1 to 800
Company: Dianova
Catalog number: 706-495-148
Lot number: 95544
Validation: https://www.dianova.com/en/produkte/706-605-148-donkey-igg-anti-guinea-pig-igg-hl-alexa-fluor-647-minx-bockgohshohumsrbrtsh-500-%C2%B5g/

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

We did not use eukaryotic cell lines.

b. Describe the method of cell line authentication used.

We did not use eukaryotic cell lines.

c. Report whether the cell lines were tested for mycoplasma contamination.

We did not use eukaryotic cell lines.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

We did not use eukaryotic cell lines.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

Thymus study (Maehr lab):
Mouse (Mus musculus) C57BL6/J "wild type": , male, age E13.5-P0 from The Jackson Laboratory
Mouse (Mus musculus) Rag1-/- "Rag1 KO": , male, age E16.5 from The Jackson Laboratory
Mouse (Mus musculus) Rag2-/- "Rag2 KO": , male, age E14.5 from The Jackson Laboratory
E12.5 embryos were generated from the Pax9Venus reporter mouse strain for easier identification of the thymic primordium at this early time-point (Stem Cell Res. 2013 Nov;11(3):1003-12. doi: 10.1016/j.scr.2013.06.007)
Pancreas study (Lickert lab):
Mouse (Mus musculus) C57BL6/J "wild type", male and female, age: E17.5, P0, P4, P9, P14, P25, P45

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

There were no human participants in this study.

nature research

Corresponding author(s): Fabian J. Theis

Code and Software Submission Checklist

Prior to submitting your work to Nature Research, we strongly recommend that you ask at least one colleague who is unfamiliar with your software to install the tool(s), follow the instructions, and provide feedback. This process will help ensure that reviewers will also be able to run your software.

You must submit all required content as a single zip file prior to peer review or provide a link where editors and reviewers can access all required content.

Required content

- [x] Compiled standalone software and/or source code
- [x] A small (simulated or real) dataset to demo the software/code

A README file that includes:

1. System requirements
   - [x] All software dependencies and operating systems (including version numbers)
   - [x] Versions the software has been tested on
   - [x] Any required non-standard hardware
2. Installation guide
   - [x] Instructions
   - [x] Typical install time on a "normal" desktop computer
3. Demo
   - [x] Instructions to run on data
   - [x] Expected output
   - [x] Expected run time for demo on a "normal" desktop computer
4. Instructions for use
   - [x] How to run the software on your data
   - [x] (OPTIONAL) Reproduction instructions

   We encourage you to include instructions for reproducing all the quantitative results in the manuscript.

Additional information

Describe your software's license for use. We strongly recommend using a license approved by the Open Source Initiative.

GPL-3.0

Provide a link to the code in an open source repository (when available).

https://github.com/theislab/pseudodynamics

Your manuscript should include a complete, detailed description of the code's functionality (i.e. pseudocode).

Please indicate where this is found:

- [ ] Main text
- [ ] Methods section
- [x] Elsewhere (specify):

  Supplementary Note 1 contain a detailed description of the algorithm. Supp. Note 1 Fig. 5 and 6 contain a conceptual description of the underlying algorithms used (published elsewhere).

Examples of well-structured software packages

1. https://github.com/neurodata-papers/MGC
2. https://github.com/neurodata-papers/LOL
3. https://www.nature.com/nbt/journal/v34/n6/abs/nbt.3569.html#supplementary-information
4. https://www.nature.com/nature/journal/v548/n7669/full/nature23463.html#extended-data
   https://github.com/yasharhezaveh/Ensai
5. https://www.nature.com/nbt/journal/v34/n11/full/nbt.3685.html#supplementary-information
   https://github.com/IFIproteomics/LFQbench