nature research

Corresponding author(s): Jacob Hanna, Alejandro Aguilera, Itay Maza

Last updated by author(s): 17/02/21

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	Single cell sequencing data were collected using Novaseq platform. Bulk RNA seq were collected with TrueSeq Illumina. qPCR data were obtained with the Viia7 platform (Applied Biosystems). Microscopy images were acquired with a Zeiss LSM 700 inverted confocal microscope (Carl Zeiss). 3D images of embryos were acquired with a light-sheet microscope (Ultramicroscope II, LaVision Biotec)
Data analysis	All statistical analysis besides single cell and bulk RNA-seq analysis were performed using the GraphPad Prism 8 software (La Joya, California). Fiji/Image J (version 1.52p) was used to count cell numbers. 10X Genomics data analysis was performed with the Cell Ranger 3.1.0 software (10x Genomics) and Seurat 3.0. The annotations were performed using the R package AUCELL 1.10.041, using parameters: aucMaxRank =100 (5% of the total gene count) under the AUCell_calcAUC function. To evaluate the probability of a certain cluster to be enriched to a certain tissue, we utilized the annotated AUC predictions of each cell to a tissue to compare to our observed cluster annotation of each cell, thus producing a p-value based on Mann-Whitney U statistics, this was performed using the R package roc.area v1.42 (CRAN.R-project.org). Differential expression (DEGs) of compatible clusters between in utero and ex utero was performed using parameters : fold-change-threshold of log(0.5) and with min.pct=0.25. DEGs with significant values were also enriched using the gene ontology database via R package limma 3.42.2 using function "goana". For bulk RNA seq, reads were trimmed with TrimGalore 0.6.5 (flags --stringency 3 --paired) and aligned to GRCm38 genome using STAR aligner (flags --runThreadN 64 --genomeLoad). Counts were estimated using HTSeq-count 0.7.2 (flags -q -f bam -r pos -s no -t exon -i gene_name). Normalization and differentially expressed genes were calculated using DESeq2 R package, with default parameters. External gene signatures, based on mouse microarray data, were calculated from GSE60603 (PMID 25945737). Flow cytometry was analyzed using FlowJo v10.7. Morphometric measurements were performed using the CellSens Entry 1.18 software (Olympus).

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

The single cell RNA-seq and bulk RNA-seq data sets generated in this study have been deposited and publicly-available in the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo) under accession number GSE149372. Source data are provided with this paper.

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	No statistical methods were used to predetermine sample size. The number of embryos used in each experiment was determined taking into account dat consistency, resources available and ethical reduction of animal use. For single cell RNA-Seq, the sample size was determined when the main cell lineages at each developmental stages were captured.
Data exclusions	For scRNA-seq, to filter out low expressing single cells, possible doublets produced during the 10X sample processing or single cells with extensive mitochondrial expression, we filtered out cells with under 200 expressing genes, over 4000 expressing genes or over 10% mitochondrial gene expression.
Replication	The exact number of replicate independent experiments and embryos is indicated in the figure or in the figure legend. All data refer to biological replicates. Single cell RNA-seq was performed in 2 biological replicates for E8.5/Day 2 and 5 &7 biological replicates for E10.5/+Day 4. All the attempts at replication were successful.
Randomization	Embryos were chosen randomly when placed in different culture conditions. Other experiments were not randomized.
Blinding	The investigators were not blinded to allocation during experiments and outcome assessment. We had no relevant scientific reasons to conduct 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	Rabbit monoclonal anti-Brachyury (D2Z3J) (Cell Signaling Cat# 81694); Rabbit polyclonal anti-Cdx2 (Cell Signaling Cat# 3977); Mouse monoclonal anti-Cdx2 (Biogenex Cat# MU392A-UC); Goat polyclonal anti-Gata4 (Santa Cruz Cat# SC-1237); Rabbit polyclonal anti-Gata4 (Abcam Cat# Ab84593); Rabbit monoclonal anti-Foxa2 (Abcam Cat# Ab108422); Mouse monoclonal anti-Myosin Heavy Chain II (clone MF-20) (R&D Cat# MAB4470); Goat polyclonal anti-Otx2 (R&D Cat# AF1979); Rabbit polyclonal anti-Pax6 (Covance; Cat# PBR-278P); Goat polyclonal anti-Sox2 (R&D Cat# AF2018); Rabbit polyclonal anti-Sox9 (Millipore Cat# AB5535); Goat polyclonal anti-Sox17 (R&D Cat# AF1924); Mouse monoclonal anti-TUBB3 (Tuj1) (Covance Cat# MMS-435P); Chicken polyclonal anti-GFP (Abcam

Cat# Ab13970); Mouse monoclonal anti-Oct4 (clone C-10) (Santa Cruz Cat# SC-5279); Goat polyclonal anti-Lefty 1 (R&D Cat# AF746); Goat polyclonal anti-mCherry/Tomato (SiCGEN Cat# AB0040-200); Rabbit anti-hNUMA (Abcam Cat# ab84680); Rabbit polyclonal anti-human TEMEM119 (Invitrogen Cat# PA562505); Mouse monoclonal anti-human CD34-PE (BioLegend, Cat# 561), Mouse monoclonal anti-human CD43-APC (ebioscience, Cat# 84-3C1).

Validation

All the antibodies have been validated by the companies from which they were offered. Details of the validation statements, antibody profiles and relevant citations can be found on the manufacturer's website.
Antibodies for immunofluorescence: All antibodies were validated herein by immunofluorescence on mouse embryos, based on their well described expression patterns.
Rabbit monoclonal anti-Brachyury (Cell Signaling Cat# 81694). The antibody has been referenced in 5 publications: https://www.cellsignal.com/products/primary-antibodies/brachyury-d2z3j-rabbit-mab/81694
Rabbit polyclonal anti-Cdx2 (Cell Signaling Cat# 3977). The antibody has been referenced in 13 publications: https://www.cellsignal.com/products/primary-antibodies/cdx2-antibody/3977
Mouse monoclonal anti-Cdx2 (Biogenex Cat# MU392A-UC). The antibody has been referenced in more than 49 publications: https://store.biogenex.com/us/anti-cdx-2-clone-cdx2-130.html
Goat polyclonal anti-Gata4 (Santa Cruz Cat# SC-1237). The antibody has been referenced in 176 publications: https://www.scbt.com/p/gata-4-antibody-c-20
Rabbit polyclonal anti-Gata4 (Abcam Cat# Ab84593). The antibody has been referenced in 45 publications: https://www.abcam.com/gata4-antibody-ab84593.html
Rabbit monoclonal anti-Foxa2 (Abcam Cat# Ab108422). The antibody has been referenced in 42 publications: https://www.abcam.com/foxa2-antibody-epr4466-ab108422.html
Mouse monoclonal anti-Myosin Heavy Chain II (clone MF-20) (R&D Cat# MAB4470). The antibody has been referenced in 50 publications: https://www.rndsystems.com/products/myosin-heavy-chain-antibody-mf20_mab4470
Goat polyclonal anti-Otx2 (R&D Cat# AF1979). The antibody has been referenced in 37 publications: https://www.rndsystems.com/products/human-otx2-antibody_af1979
Rabbit polyclonal anti-Pax6 (Covance; Cat# PBR-278P). The antibody has been referenced in 169 publications: https://www.biolegend.com/en-us/products/purified-anti-pax-6-antibody-11511
Goat polyclonal anti-Sox2 (R&D Cat# AF2018). The antibody has been referenced in 93 publications: https://www.rndsystems.com/products/human-mouse-rat-sox2-antibody_af2018
Rabbit polyclonal anti-Sox9 (Millipore Cat# AB5535). Anti-Sox9 Antibody is a well characterized affinity purified Rabbit Polyclonal Antibody that reliably detects Transcription Factor Sox-9. This highly published antibody has been validated in IHC & WB: https://www.merckmillipore.com/INTL/en/product/Anti-Sox9-Antibody,MM_NF-AB5535?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1
Goat polyclonal anti-Sox17 (R&D Cat# AF1924). The antibody has been referenced in 163 publications: https://www.rndsystems.com/products/human-sox17-antibody_af1924
Mouse anti-TUBB3 (Tuj1) (Covance Cat# MMS-435P). The antibody has been referenced in 436 publications. https://www.biolegend.com/en-us/products/purified-anti-tubulin-beta-3-tubb3-antibody-11580
Chicken polyclonal anti-GFP (Abcam Cat# Ab13970). The antibody has been referenced in 2036 publications: https://www.abcam.com/gfp-antibody-ab13970.html
Mouse monoclonal anti-Oct4 (clone C-10) (Santa Cruz Cat# SC-5279). The antibody has been referenced in 1905 publications: https://www.scbt.com/p/oct-3-4-antibody-c-10
Goat polyclonal anti-Lefty 1 (R&D Cat# AF746). Validated in for IHC in mouse tissues by a previous study. The antibody has been referenced in 1 publication: https://www.rndsystems.com/products/human-mouse-lefty-antibody_af746
Goat polyclonal anti-mCherry/Tomato (SiCGEN Cat# AB0040-200). This antibody (AB0040) recognizes very well tdTomato and does not recognize GFP (green fluorescent protein). The antibody has been referenced in 15 publications: https://www.origene.com/catalog/antibodies/primary-antibodies/ab0040-200/mcherry
Rabbit anti-hNUMA (Abcam Cat# ab84680). The antibody has been referenced in 6 publications: https://www.abcam.com/numa-antibody-ab84680.html
Rabbit polyclonal anti-human TEMEM119 (Invitrogen Cat# PA562505). This antibody has been validated by the manufacturer to react with human human cerebral cortex microglia. https://www.thermofisher.com/antibody/product/TMEM119-Antibody-Polyclonal/PA5-62505
Antibodies for immunohistochemistry:
Rabbit monoclonal anti-Foxa2 (Abcam Cat# Ab108422). Validated by the manufacturer using Formalin/PFA-fixed paraffin-embedded sections of mouse liver tissue. The antibody has been referenced in 42 publications: https://www.abcam.com/foxa2-antibody-epr4466-ab108422.html
Goat polyclonal anti-Gata4 (Santa Cruz Cat# SC-1237). Validated by the manufacturer for IHC in mouse tissues. The antibody has been referenced in 176 publications: https://www.scbt.com/p/gata-4-antibody-c-20
Antibodies for flow cytometry: All the antibodies guarantee covers the use of the antibody for flow cytometry applications.
Mouse monoclonal anti-human CD34-PE (BioLegend, Cat# 561). Each lot of this antibody is quality control tested by immunofluorescent staining with flow cytometric analysis. The antibody has been referenced in 8 publications: https://www.biolegend.com/en-us/products/pe-anti-human-cd34-antibody-6036
Mouse monoclonal anti-human CD43-APC (ebioscience, Cat# 84-3C1). This eBio84-3C1 (84-3C1) antibody has been pre-titrated and tested by flow cytometric analysis of human peripheral blood monocytes. The antibody has been referenced in 7 publications: https://www.thermofisher.com/antibody/product/CD43-Antibody-clone-eBio84-3C1-84-3C1-Monoclonal/17-0439-42

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Mouse V6.5 and WIS2 human embryonic stem cell lines were previously reported in Gafni et al. Nature 2013 and Geula et al. Science 2015. HEK293 was obtained from ATCC.

Authentication

Karyotype and sequencing data cofirmed expected sex, karyotype, gene reporters and SNPs.

Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination by using the MycoAlert plasma Detection Kit (Lonza, Cat# LT07-318) and were routinely screened every 2 months.
Commonly misidentified lines (See ICLAC register)	None were used herein

Animals and other organisms

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

Laboratory animals	All mice were housed in a standard 12-hour light/12-hour dark cycle conditions in a specialized and certified animal facility. 5-8 weeks old female mice (mus musculus) of the strains ICR, C57BL/6, 129 and BDF1 were mated with male (5-40 weeks old) ICR, BDF1 or C57/B6 mice. Established transgenic mouse strains used were: Jackson #007576, Jackson #022137, Jackson #024242, Jackson #007914, Jackson #030539
Wild animals	The study did not involve wild animals
Field-collected samples	The study did not involve samples collected from the field
Ethics oversight	All animal experiments were performed according to the Animal Protection Guidelines of Weizmann Institute of Science, Rehovot, Israel. All animal experiments described herein were approved by relevant Weizmann Institute IACUC (#01390120-1, 01330120-2, 33520117-2).

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	Healthy pregnant women who were asked to give their informed consent to have blood collected from their umbilical cord, as approved by a Rambam Medical Center Helsinki committee (#RMB-0452-15). The source of each collection underwent full anonymization and was not identified by name or other designation, and the extracted serum was only used as described herein. Healthy women over the age of 18 and under 40, who gave their consent and were scheduled for caesarian section delivery by their obstetrician following a prenatal clinic visit, were eligible for cord blood collection. We excluded women who gave vaginal birth as well as women with any chronic illness or active medical conditions, including gestational diabetes or hypertension. All adult blood samples were collected from healthy donors.
Recruitment	Research donors in the study were recruited from the prenatal clinic at the Rambam Medical Center were asked to give their informed consent to have blood collected from their umbilical cord. Adult blood samples were collected from the authors of the study.
Ethics oversight	Rambam Medical Center Helsinki committee (#RMB-0452-15)

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	Cells were incubated for half an hour with CD34-PE (BioLegend, 561) and CD43-APC (ebioscience, 84-3C1) antibodies (1:50) on PBS/0.5% BSA.
Instrument	BD FACS-Aria III
Software	FlowJo v10.7
Cell population abundance	Only one cell population was analyzed and the purity was verified
Gating strategy	FSC and SSC singlets were gated to remove debris and aggregated cells, and only single cells were considering for all analyses. To determine the gating for positive of negative populations, an unstained control was employed, making sure that 100% of the unstained population was allocated on the negative area of the histogram/dot plot. Gating strategies are shown in

Extended Data Fig. 12b.

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