nature research

Corresponding author(s): Mark S. Anderson

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 Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Confocal Microscopy: Leica LAS Flow Cytometry: BD FACSDiva NGS: Illumina HiSeq Software Suite
Data analysis	Statistics: GraphPad Prism v7.0 Image Analysis: Bitplane Imaris v8.3.1, Spatstat pcfcross (open source), Leica LAS RNA-Seq: RNA-STAR (open source), DESeq2 (open source), Subread (open source), indeXplorer (open source), SCDE (open source), R "stats" package (open source) Flow Cytometry: TreeStar FlowJo v10.4.2 General: Microsoft Excel v15.33

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/reviewers upon request. 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

RNA-seq data that support the findings of this study have been deposited in the GEO database. Source data for figures 1e, 2c, 3e,f,h, 4c-f and extended data figures 5d,f-i, 9b-d, and 10b-e,g are provided with the paper. All other data that support the findings of this study are available from the corresponding author upon reasonable request.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-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. Sample size was chosen based on the magnitude and consistency of measurable differences between groups and application of standard practices within the field. Whenever possible, experimental group sizes of at least 5 animals were used. In two cases, for Extended Data, experimental groups of 3 animals were used. For all findings, at least two independent experiments were conducted.
Data exclusions	Some cells were excluded from the single cell RNA sequencing analysis based on quality criteria. Cells were required to have a minimum of 40% total reads mapping to the genome, which all our cells met. Reads uniquely aligned to mRNA genes were counted with featureCounts, a part of the Subread software package 37. Genes present in a minimum of 5 cells and with at least 10 reads were retained in the analysis. Cells expressing fewer than 750 genes with 10 reads were removed from the analysis. All cells were required to have a minimum of 100,000 total reads mapping to genes, and less than 10% percent of reads mapping to mitochondrial genes. This resulted in 197 Flare25 cells and 95 Flare25.Aire-/- cells for further downstream analysis.
Replication	All experimental findings described in this manuscript were reliably reproduced.
Randomization	Animals were age-matched and were allocated to groups based on genotype. Whenever possible, animals were co-housed with littermate controls and sex-matched.
Blinding	Researchers were not blinded during experiments or data analysis since all of the findings are supported by quantitative measure (i.e. flow cytometry, quantitative image analysis, bionformatics, etc.).

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants

Methods

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

Antibodies

Antibodies used: Antibodies were chosen based on the available literature. All antibodies were validated using primary mouse or human tissues or single cell suspenstions. All staining reproduced previously published results or data available from the manufacturers. Lot

validation and quality control is available from the commercial source using the lot numbers listed below.

Mark Anderson Laboratory:
Anti-Mouse
Aire 5H12 eFluor660 eBiosciences 50593482 72423878
Aire 5H12 A488 eBiosciences 53593482 72306882
CD11c N418 Pe-Cy7 BioLegend 117318 B222652
CD3 17A2 APC-Cy7 BioLegend 100222 B239434
CD4 RM4-5 APC Tonbo 20-0042-u100 42010417203
CD45 30-F11 PercP BioLegend 103130 B214531
CD8 53-6.7 PercP-Cy5.5 BioLegend 100734B205157
DCLK1 Rab poly A488 Abcam ab203441 GR226078-1 (custom conj of ab31704)
Eomes Dan11mag A488 eBiosciences 53-4875-80 4314070
EpCAM G8.8 APC-Cy7 BioLegend 118218 B233587
I-Ab 25-9-17 eFluor450 eBiosciences 48-5320-82 E13383-105
Ly51 6C3 FITC BioLegend 108305 B147730
PLZF Mags.21F7 A488 eBiosciences 53-9320-82 4305496
RORg Q31-378 PE BD Pharm 568607 6223964
T-bet 4B10 A647 BioLegend 644804 B223600
TCRbeta H57-587 APC-eFluor780 eBiosciences 47-5961-82 4277886

Anti-Human
Epcam HEA-125 PE Miltenyi 130-091-253
CD45 HI30 Percp Biolegend 304026
HLA-DR L243 APC-Fire750 Biolegend 307658

Kristin Hogquist Laboratory:
CD24 M1/69 PerCP-Cy5.5 Biolegend 101824 B229696
CD4 GK1.5 BV786 BD Bioscience 563331 7075503
CD45 30-F11 BV605 Biolegend 103155 B234402
CD8 53-6.7 BUV395 BD Bioscience 563786 7249930
DCLK1 Rab poly IgG Unconjugated Abcam ab31704 GR162298-2
Eomes Dan11mag A488 eBioscience 53-4875-82 4314070
EpCAM G8.8 A647 Biolegend 118212 B217174
Goat-anti-Rab Goat polyclonal IgG AF488 ThermoFisher Scientific A-11034 1751340
huCD2 TS1/8 BV421 Biolegend 309218 B212053
K10 EP1607IHCY A647 Abcam ab194231 GR201943-1
MHC-II M5/114.15.2 eFluor 450 eBioscience 48-5321-82 E08529-1634
PLZF R17-809 AF647 BD Bioscience 563490 7138685
ROR-rt Q31-378 PE-CF594 BD Bioscience 562684 7040792
Streptavidin PE-Cy7 eBioscience 25-4317-82 4290713
T-Bet 4B10 PE-Cy7 Biolegend 644824 B214294
UEAI lectin Biotinylated VECTOR Laboratory B-1065 X1020

Bruno Kyewski Laboratory:
DCLK1 Rab polyclonal IgG Unconjugated Abcam ab31704 GR162298-2
K10 EP1607IHCY AF647 Abcam ab194231 GR201943-1
MHC-II M5/114.15.2 eFluor 450 eBioscience 48-5321-82 E08529-1634
EpCAM G8.8 AF647 Biolegend 118212 B217174
CD45 30-F11 BV605 Biolegend 103155 B234402
UEAI lectin Biotinylated VECTOR Laboratory B-1065 X1020
CD4 GK1.5 BV786 BD Bioscience 563331 7075503
CD8 53-6.7 BUV395 BD Bioscience 563786 7249930
CD24 M1/69 PerCP/Cy5.5 Biolegend 101824 B229696
PLZF R17-809 AF647 BD Bioscience 563490 7138685
T-Bet 4B10 PE/Cy7 Biolegend 644824 B214294
ROR-rt Q31-378 PE-CF594 BD Bioscience 562684 7040792
Eomes Dan11mag AF488 eBioscience 53-4875-82 4314070
huCD2 TS1/8 BV421 Biolegend 309218 B212053
Goat-anti-Rab Goat polyclonal IgG AF488 ThermoFisher Scientific A-11034 1751340
Streptavidin PE-Cyanine7 eBioscience 25-4317-82 4290713

Validation

Antibodies were chosen based on the available literature. All antibodies were validated using primary mouse or human tissues or single cell suspenstions. All staining reproduced previously published results or data available from the manufacturers. Lot validation and quality control is available from the commercial source using the lot numbers listed.

Animals and other organisms

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

Laboratory animals

All mice were on the C57BL/6 or the mixed Balbc/ByJ x C57BL/6 F1 backgrounds. Both male and female mice were used. All mice were 4-8 weeks of age. Some animal strains used in this study are under specific MTA agreements.

Purchased animals were from The Jackson Laboratory:

C57BL/6 (Jax #000664)
Balbc/ByJ (Jax #001026)
B6.Cg-Gt(ROSA)26Sortm14(CAG-tdTomato)Hze/J (Jax #007914)
B6.Cg-Gt(ROSA)26Sortm6(CAG-ZsGreen1)Hze/J (Jax # #007906)

Provided by Mark S. Anderson (UCSF):
C57BL/6.Aire-/-
C57BL/6.Aire-CreERT2
C57BL/6.Aire-DTR
C57BL/6.iALT
C57BL/6.FoxN1-Cre

Provided by Richard M. Locksley (UCSF):
C57BL/6.Flare25
C57BL/6.Trpm5-/-

Provided by Bruno Kyewski (DKFZ):
C57BL/6.Hipk2fl/fl
C57BL/6.FoxN1-Cre

Provided by DTCC-KOMP2 (Toronto):
C57BL/6.Pou2f3-/-

Provided by Kristin Hogquist (U of Minnesota)
Balbc/ByJ.KN2
C57BL/6.Aire-DTR

Wild animals

Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

Field-collected samples

For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

Human research participants

Policy information about studies involving human research participants

Population characteristics

All human thymic tissue was collected at UCSF facilities in accordance with approved Internal Review Board Protocols and was de-identified at the point of collection. Fetal thymus was 22.3 gestational weeks and neonatal thymus was 21 days postpartum.

Recruitment

Thymus samples were collected from pediatric patients undergoing corrective cardiothoracic surgery.

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

Murine Thymocytes:
Thymi were isolated, cleaned of fat and transferred to DMEM (UCSF Cell Culture Facility) containing 2% FBS (Atlanta Biologics) on ice. Single cell suspensions of thymocytes were prepared by mashing through a 40 micron filter (Falcon).

Murine Thymic Epithelial Cells:
Thymi were minced with a razor blade. Up to five thymi were pooled into a single digestion. Tissue pieces were moved with a glass Pasteur pipette to 15 ml tubes and vortexed briefly in 10 ml of media. Fragments were allowed to settle before removing the media and replacing it with 4 ml of digestion media containing 2% FBS, 100 ug/ml DNase I (Roche), and 100 ug/ml Liberase TM (Sigma Aldrich) in DMEM. Tubes were moved to a 37 °C water bath and fragments were titurated through a glass Pasteur pipette at 0 min and 6 min to mechanically aid digestion. At 12 min tubes were spun briefly to pellet undigested fragments and the supernatant was moved to 20 ml of 0.5% BSA (Sigma Aldrich), 2 mM EDTA (TekNova), in PBS (MACS buffer) on ice to stop the enzymatic digestion. This was repeated twice for a total of three 12 min digestion cycles, or until there were no remaining tissue fragments. The single cell suspension was then pelleted and washed once in MACS Buffer. Density-gradient centrifugation using a three-layer Percoll gradient (GE Healthcare) with specific gravities of 1.115, 1.065, and 1.0 was used to enrich for stromal cells. Cells isolated from the Percoll-light fraction, between the 1.065 and 1.0 layers, were then resuspended in 0.5% BSA (Sigma Aldrich), 2 mM EDTA (TekNova) (FACS buffer) and counted.

Human Thymic Epithelial Cells:
Thymi were transferred to RPMI (ThermoFisher) and cut into small pieces using scissors. Tissue pieces were mashed gently using the back of a sterile syringe to extract most of the thymocytes. The supernatent was removed and replaced with fresh RPMI. Tissue pieces were moved to a 15 ml tube using a 5 ml pipette. Fragments were allowed to settle before discarding the media and replacing it with 5 ml of digestion media containing 100 ug/ml DNase I (Roche), and 100 ug/ml Liberase TM (Sigma Aldrich) in RPMI. Tubes were moved to a 37 C water bath and fragments were triturated through a 5 ml pipette at 0 min and 6 min to mechanically aid digestion. At 12 min tubes were spun briefly to pellet undigested fragments and the supernatant was discarded. Fresh digestion media was added to remaining fragments and the digestion was repeated using a glass Pasteur pipette for trituration. Supernatant from the second round of digestion was also discarded. A third round of enzymatic was performed using 5 ml of digestion media supplemented with Trypsin/EDTA (ThermoFisher) for a final concentration of 0.05%. Remaining thymic fragments were digested for an additional hour with trituration every 15 minutes. The cells were moved to cold MACS buffer to stop the enzymatic digestion. TECs were enriched using the same Percoll gradient as for murine mTECs.

Murine Small Intestinal Epithelial Cells
Small intestines were flushed with PBS, opened, and rinsed with PBS to remove luminal contents. Two-and-a-half- to five-cm-long segments of jejunum were incubated with rocking for 20 min at 37 °C in 5 ml PBS containing 2.5 mM EDTA (Sigma-Aldrich), 0.75 mM dithiothreitol (DTT; Sigma-Aldrich), and 10 µg ml−1 DNaseI (Sigma-Aldrich). Tissues were shaken vigorously for 30 s and released cells were incubated with rocking for 10 min at 37 °C in 5 ml HBSS (Ca2+/Mg2+ free) containing 1.0 U ml−1 Dispase (Gibco) and 10 µg ml−1 DNaseI. Digested cells were passed through a 70 µm filter and washed once before staining.

Instrument

Data were collected using the BD LSR2, BD Fortessa, and BD Aria Fusion instruments.

Software

Flow cytometry data were collected and analyzed using the BD FACSDiva v8.0.1,TreeStar FlowJo v10.3 software.

Cell population abundance

All cells were sorted to a purity of > 95%

Gating strategy

All cells were gated on singlets, live cells, and FSC/SSC.

iNKT:
CD1d+ TCRbint

EOMES+ Thymocytes:
TCRb+ CD4- CD8+

mTECs:
CD11c- CD45- EPCAM+ (+/- Ly51-)

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