

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 (<i>n</i>) 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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ 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 <i>d</i> , Pearson's <i>r</i>), 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	Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.
Data analysis	All code used to generate the results and figures in this study are available in Zenodo (https://doi.org/10.5281/zenodo.12520577) scRNA-seq data analysis The generated fastq files were processed with Cell Ranger (version 7). Reads were aligned to the mm10-3.0.0 reference transcriptome annotation file. Sequences and annotations for tdTomato and Cre were manually added to the fasta and gtf files (Supplementary Data 5), respectively, using the function cellranger mkref. A total of 10,407 cells were sequenced, with a mean of 35,041 reads per cell and a median of 3,055 genes detected per cell. Demultiplexing was performed using Python according to the unique MULTI-seq barcodes given to each P1, P2, and P4 sample prior to pooling. Downstream bioinformatic analysis was performed on R (version 4.2.2) using the ‘Seurat’ R package (version 5.0.1). Briefly, pre-processing of the data was performed following quality assessment to include cells with >2,000 and <5,500 detected genes, and percentages of mitochondrial genes of >5% and <25% (Supplementary Fig. 3). Additional biological filtering was applied based on markers of common cell types known to exist in the heart. For instance, 1.03% of all cells (72 out of 6927 cells) that lacked expression of cardiomyocyte genes (Myh6, Myh7, Tnnt2, Tnni1, Ryr2) but expressed markers of fibroblasts (Postn, Col1a1, Col1a2, Col3a1) or smooth muscle cells (Cald1, Emilin1, Hexim1, Tagln and Fapb4), were excluded from downstream analysis (Supplementary Fig. 4). Following pre-processing of the data, feature counts were normalized by log-transformation using the function NormalizeData. The top 5,000 features were identified using “vst” as the selection method using FindVariableFeatures. Normalized counts for these variable genes were then scaled using ScaleData with regression of mitochondrial gene counts and the number of unique molecular identifiers (UMIs). Linear dimensional reduction was performed using principal component analysis (PCA) using RunPCA. For Uniform Manifold Approximation and Projection (UMAP) embedding using RunUMAP

analysis using all cells, including non-cardiomyocytes (Supplementary Fig. 4), CCS cells (Supplementary Fig. 5b), 'Nodes-HIS-BB' (Supplementary Fig. 5c), 'Nodes' (Supplementary Fig. 5d), the first 25 principal dimensions were used. Optimal clusters were identified by FindNeighbors and FindClusters using resolutions of 0.25, 0.19, 0.4, and 0.97, respectively, determined by the examination of known marker expression patterns. For the 'LNB-HIS-PBB' (Supplementary Fig. 5e), the first 10 principal components were used with a clustering resolution of 0.6. For 'Proliferative CCS' (Supplementary Fig. 8), 20 principal components were used with a clustering resolution of 0.6. Heatmaps were generated using the 'ComplexHeatmap' R package (version 2.15.4). E16.5 CCS data (GSE132658) was analyzed according to the author's described methods as well as using markers demonstrated in the study. Integration of E16.5 and Postnatal CCS was performed using Seurat's Integration pipeline. Briefly, anchors between the datasets were found using FindIntegrationAnchors followed by data integration using IntegrateData. For correlation between E16.5 and postnatal CCS components, averaged expressions of genes for each CCS component given by AverageExpression were used. Correlation analysis was then performed and visualized using the 'corrplot' R package (version 0.92). Direct comparison was performed on integrated data using FindMarkers.

Visualization of sequencing reads alignment to Scn10a was generated using the Integrative Genomics Viewer (IGV) (version 2.17.4), with the reference genome mm10 (Supplementary Fig. 12).

Gene regulatory network inference

Gene regulatory network inference was performed using 'pySCENIC' (version 0.12.1) following the SCENIC (single-cell regulatory network inference and clustering) pipeline. The count matrix was provided as input after filtering out genes with zero expression. Repressive regulons were excluded, and putative regulons lacking expression of the TF in at least 40% of any region were filtered out, resulting in 97 retained regulons. Regulon activity was quantified in individual cells using AUCell and averaged across each region.

To construct regional gene regulatory networks, regulons were scored based on regional specificity. The top 5 regulons specific to each region were retained after filtering for 40% expression of the TF in the respective region. Target genes differentially expressed in each region (adj. $P < 0.05$) were retained. The resulting regional networks with top 5 regulons were then visualized using the Fruchterman-Reingold layout ('iGraph', version 1.5.1), with node colours corresponding to the classification of the gene (transcription factors and CCS marker genes).

ATAC-seq and ChIP-seq analyses

For chromatin accessibility and binding analyses, raw data was downloaded in fastq format using the fastq-dump function in the SRA toolkit. Fastq files were then uploaded onto the UseGalaxy server. Quality control assessments were performed using FastQC (version 0.11.7). Reads were then aligned to the mm9 mouse reference genome using Bowtie2 (Galaxy Version 2.4.2+galaxy0) with the '--very-sensitive' parameter applied for optimal mapping. Mitochondrial genome reads and low mapping quality (mapQuality <30) were considered uninformative reads and filtered out using the Filter BAM datasets on a variety of attributes function (Galaxy Version 2.4.1). Duplicate reads were then removed using 'Picard MarkDuplicates' (Galaxy Version 2.18.2.2) (broadinstitute.github.io/picard). MACS2 callpeak function (Galaxy Version 2.1.1.20160309.6) was used to call peaks with the following parameters: 'Single-end BED', '--nomodel', '--shift -100', '--extend 200', and FDR < 0.05. The following datasets were analysed: E13.5 Tbx3+ AVJ ATAC-seq (GSE121464), adult ventricle ATAC-seq (GSE110209), P0 sinoatrial pacemaker ATAC-seq (GSE148515), P0 right atrial cardiomyocyte ATAC-seq (GSE148515), and adult heart Tbx3 ChIP-seq (GSE35151). Moreover, two Iroquois (Irx) motifs, ACATGT (Motif ID: UP00223_2) and ACAnnTGT were used for predicting genome-wide motif occurrences. Each motif sequence was uploaded onto FIMO (Find Individual Motif Occurrences), part of the MEME software package (tools.genouest.org/tools/meme/), and motifs were scanned in the UCSC mm9 mouse genome. Only motif occurrences with P values $\leq 1.0 \times 10^{-4}$ were considered significant, which are automatically calculated by FIMO based on a log-likelihood ratio score conversion. Peaks were visualized in bigWig format by converting MACS2 Bedgraph output to bigWig via the Wig/BedGraph-to-bigWig function (Galaxy version 1.1.1) and uploading it onto the Integrative Genomics Viewer (IGV) genome browser (version 2.16.2). For datasets where peaks were readily available in bigWig format, peak files were directly downloaded from the NCBI Gene Expression Omnibus (GEO) and uploaded onto IGV. For datasets with multiple replicates, a mean bigWig file was generated using the deepTools2 'bigwigCompare' function (Galaxy version 3.5.1.0.0).

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](#)

The scRNA-seq data and spatial transcriptomics data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under accession code, GSE231547 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE231547]. Source data for Fig. 6 and Supplementary Fig. 25 and 26 are provided with this paper.

All code used to generate the results and figures in this study are available in Zenodo (https://doi.org/10.5281/zenodo.12520577)

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 N/A

Reporting on race, ethnicity, or N/A

other socially relevant groupings

Population characteristics

N/A

Recruitment

N/A

Ethics oversight

N/A

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](https://www.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 method was used to predetermine sample size.
Data exclusions	No data were excluded from the analyses
Replication	Fluorescence stereomicroscope imaging of the CCS-tdT heart (Fig. 1a) was performed on more than five hearts. All immunofluorescence experiments (Fig. 3a, c, d, e, f, and Supplementary Fig. 12, 13, 15, 16, 17, 19, 20) were independently performed at least two times with varying antibody concentrations. All fluorescence in situ hybridization (FISH) experiments (Supplementary Fig. 14, 18, 21) were conducted at least twice.
Randomization	Randomization is not relevant to this study
Blinding	Blinding is not relevant to this study

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

chicken anti-RFP (Rockland Immunochemicals, 600-901-379, lot #42649, 1:100)
<https://www.rockland.com/categories/primary-antibodies/rfp-antibody-600-901-379/>

goat anti-Cntn2 (R&D Systems, AF4439, lot #CDSO0219101, 1:50)
https://www.rndsystems.com/products/human-mouse-rat-contactin-2-tag1-antibody_af4439

rabbit anti-Gnao1 (Thermo Fisher Scientific, PA5-30044, lot #YH4027897A, 1:100)
<https://www.thermofisher.com/antibody/product/GNAO1-Antibody-Polyclonal/PA5-30044>

rabbit anti-Scn10a (Alomone Labs, ASC-016, lot #, ASC016AN2550, 1:50)
https://www.alomone.com/p/anti-nav1-8-2/ASC-016?gad_source=1&gclid=Cj0KCQjwsuSzBhCLARIsAlcdLm4B0rNv-sjCzZ9isXss6nWHfQAQKkEhlScsdY27TcmBALbtoXzta4aAp-zEALw_wcB

rabbit anti-Igfbp7 (Abcam, ab74169, lot #GR242021-1, 1:50)
<https://www.abcam.com/products/primary-antibodies/igfbp7-antibody-ab74169.html>

rabbit anti-S100a6 (Abcam, ab181975, lot #GR3438351-1, 1:100)
<https://www.abcam.com/products/primary-antibodies/s100-alpha-6pra-antibody-epr13084-69-ab181975.html>

rabbit anti-Nppa (EMD Millipore, ab5490, lot #Q3015715, 1:250)
https://www.emdmillipore.com/US/en/product/Anti-Atrial-Natriuretic-Peptide-Antibody,MM_NF-AB5490

mouse anti-Shox2 IgG2a (Santa Cruz Biotechnology, sc-81955, lot #B2822, 1:50)
<https://www.scbt.com/p/shox2-antibody-jk-6e>

rabbit anti-Vsn1 (GeneTex, GTX115039, lot #40247, 1:100)
<https://www.genetex.com/Product/Detail/Visinin-like-1-antibody/GTX115039>

rabbit anti-Ppp1r17 (Thermo Fisher Scientific, PA5-61599, lot #YH4028316B, 1:200)
<https://www.thermofisher.com/antibody/product/PPP1R17-Antibody-Polyclonal/PA5-61599>

goat anti-chicken IgY Alexa FluorTM 568 (Thermo Fisher Scientific, A11041, lot #2482953, 1:200)
<https://www.thermofisher.com/antibody/product/Goat-anti-Chicken-IgY-H-L-Secondary-Antibody-Polyclonal/A-11041>

donkey anti-chicken IgY Alexa FluorTM 568 (Thermo Fisher Scientific, A78950, lot # 2622381, 1:200)
<https://www.thermofisher.com/antibody/product/Donkey-anti-Chicken-IgY-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A78950>

goat anti-rabbit Alexa FluorTM 488 (Thermo Fisher Scientific, A-11008, lot #2420730, 1:500)
<https://www.thermofisher.com/antibody/product/Goat-anti-Rabbit-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11008>

goat anti-mouse IgG2a Alexa FluorTM 488 (Thermo Fisher Scientific, A21131, lot #2420706, 1:500)
<https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG2a-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21131>

donkey anti-rabbit Alexa Fluor® 647 (Jackson ImmunoResearch Laboratories Inc., 711-605-152, 1:200)
<https://www.jacksonimmuno.com/catalog/products/711-605-152>

goat anti-mouse IgG2b Alexa FluorTM 488 (Thermo Fisher Scientific, A21141, lot #2432059, 1:500)
<https://www.thermofisher.com/antibody/product/Goat-anti-Mouse-IgG2b-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21141>

Validation

Manufacturers' websites are listed above

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

Early postnatal mice (1 to 4 days old) were used for experiments. All mice had free access to standard chow (2019 Teklad Global Diet; Envigo) and water and were housed in a controlled environment with a 12h light/12h dark cycle, ambient temperature of 22-24°C and humidity levels of 30-60%. Mice were humanely euthanized via CO₂ inhalation and/or cervical dislocation based on CCAC guidelines on euthanasia of animals used in science.

For scRNA-seq, ST, and histological validation experiments, homozygous Cntn2tm1.1(cre/EGFP)Nmun (Cntn2Cre) and homozygous Rosa26tdTomato mice³², which were maintained on a CD-1 background, were crossed to generate CCS-fluorescent reporter mice (Cntn2Cre/+;Rosa26tdTomato/+, CCS-tdT). To achieve the maximum number of age-matched pups simultaneously, 8-15 weeks old female mice were housed together without exposure to male mice for 2-3 weeks prior to mating to synchronize their estrous cycles. Following the mating setup, vaginal plugs were monitored daily. Neonatal mice on postnatal day 1, 2, and 4 were euthanized by cervical dislocation immediately before heart collection for cardiomyocyte isolation, ST or histological validation. A total of 103 CCS-tdT pups were used for scRNA-seq. For in vitro adenovirus-based overexpression studies, 2 to 3 P0 neonatal litters from CD-1 male and female breeders were used to isolate atrial and ventricular cardiomyocytes.

Wild animals

Provide details on animals observed in or captured in the field; report species 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.

Reporting on sex

Both males and females were included in all experiments, and sex differences were not considered.

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.

Ethics oversight

All experimental protocols were approved by the Animal Care Committees of the University of Ottawa (HI-2950 and 3060), in accordance with the guidelines of the Canadian Council on Animal Care (CCAC).

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

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A

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	Neonatal cardiomyocytes were enzymatically isolated from CCS-tdT report mice and subjected to Sytox staining.
Instrument	MoFlo XDP
Software	FlowJo v. 10
Cell population abundance	Percentage of purified cells were provided.
Gating strategy	Sytox-negative and tdTomato-positive

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