View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Caltech Authors - Main

nature research

Corresponding author(s): Marianne E. Bronner

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	Zeiss Zen, Illumina HiSeq2500, WinList from Verity Software House
Data analysis	BLAT alignment software, R v3.6.1, Bowtie2, HTSeq-Count, DESeq2, TCoffee, Expresso, MegAlign Pro (DNAstar 15.0.0), MrBayes3.2.1, FigTree 1.4.2, Biotapestry, Adobe Illustrator CS5.1

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

All raw sequencing data for all RNAseq libraries (Figure 3) and merged reference transcriptomes are available online (NCBI BioProject# PRJNA497902). Sequences of in situ probe templates for Figures 1B, 1C, 2A, and 2C are available through GenBank accession codes found in the methods.

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	For chicken RNAseq replicates, a pre-determined number of a cells was collected per kit instructions for the SMART-seq v4 ultra low input RNA kit (Cranial Rep 1= 1534 cells, Cranial Rep 2= 1530 cells, Cranial Rep 3= 1527 cells; Trunk Rep 1= 1500, Trunk Rep 2= 721, Trunk Rep 3= 958). For lamprey dissections, approximately 100 embryos were dissected for each replicate of each axial level for RNAseq libraries. No statistical tests were used to determine sample size. The sample size provided enough RNA for subsequent library preparation.
Data exclusions	No datasets were excluded from this analysis.
Replication	Chicken RNAseq libraries were collected in triplicate for biological replicates. Lamprey RNAseq libraries were collected with replicates, as well, for biological replicates. All attempts at replication were successful. For in situ hybridization, embryos were pooled from different breeding pairs (fish), brooding stocks (skates), or embryo batches (lamprey) to ensure replication of results in multiple fixed collections. All in situ expression patterns were replicated 100% over multiple rounds of in situs.
Randomization	Chicken eggs obtained from a local chicken farm were incubated in multiple incubators over different days to account for inter-batch variability. Lamprey tissues were collected from different breeding events to account for inter-batch variability.
Blinding	Different batches of chicken embryos were separately incubated in different incubators and electroporated with fresh DNA reporter construct solution. Different breeding pairs for lamprey were used for tissue collections to blind for batch variability. For all trunk libraries, dissections were made at the same somitic levels and not based on reporter expression to ensure blinding in sample collection.

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	anti-Digoxigenin-Alkaline Phosphatase Fab Fragments (Roche ref#11093274910)
Validation	From the supplier: The polyclonal antibody from sheep is specific to digoxigenin and digoxin and shows no cross-reactivity with other steroids, such as human estrogens and androgens. Concentrations of antibody used for each animal has been previously reported and validated as cited in the methods section.

Animals and other organisms

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

Laboratory animals	Adult (3 months - 2 years of age) Danio rerio, zebrafish, were maintained in the Beckman Institute Fish Facility at Caltech, and all animal and embryo work was compliant under approved IACUC protocol 1346. Adult zebrafish are bred via natural spawning, which is triggered by light-dark cycle. The fish are kept in a specialized recirculating aquatic system with mechanical, chemical, and biological filters to maintain water quality. The water is kept at 28 degrees C and highly oxygenated.
Wild animals	Gravid male and female sea lamprey (Petromyzon marinus) were caught in the wild and provided by the Great Lakes Fishery

Wild animals	Commission, in cooperation with its partners at the USGS Hammond Bay Biological Station, USFWS Marquette Biological Station, and the Department of Fisheries and Oceans, Canada. They were sent overnight in chilled, oxygenated water to the lamprey facility at the California Institute of Technology, where they were maintained under the parameters set in accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health, with protocols approved by the Institutional Animal Care and Use Committees of the California Institute of Technology (lamprey, Protocol #1436-17). After spawning, the captive adult lamprey died of natural causes.
Field-collected samples	Lamprey embryos were produced by in vitro fertilization at the California Institute of Technology lamprey facility, using captive gravid lamprey (Petromyzon marinus) provided by the Great Lakes Fishery Commission, in cooperation with its partners at the USGS Hammond Bay Biological Station, USFWS Marquette Biological Station, and the Department of Fisheries and Oceans, Canada. Lamprey were maintained with a water tempature of 10-18 degrees C on a 15:9h light:dark cycle. All skate embryos were collected from wild-caught brood stock housed at the Marine Resources Centre of the Marine Biological Laboratory (MBL) in Woods Hole. Eggs were maintained in a flow-through seawater system with a water temp of 15 degrees C, on a 12h:12h light:dark cycle. Prior to fixation, all embryos were euthanized with an overdose of buffered MS-222 (1g/L in seawater). All embryos collection was performed in accordance with protocols approved by the MBL Institutional Animal Care and Use committee. Fertilized chicken eggs were obtained from a local farm in Sylmar, CA. Developing chicken embryos were maintained at a temperature of 37 degrees C.

Flow Cytometry

Plots

Confirm that:

- [x] The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- [x] 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).
- [x] All plots are contour plots with outliers or pseudocolor plots.
- [x] A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Chicken embryos were dissected in Ringers and washed thrice in chilled 1x PBS. The tissues were dissociated in Accumax (Innovative Cell Technologies, Inc.) for 15 minutes at 37ºC.
Instrument	Sony SY3200 Cell Sorter
Software	WinList from Verity Software House
Cell population abundance	Collected cells were obtained and analyzed on a hemocytometer for fluorescence and viability to ensure 100% purity.
Gating strategy	Gating was assigned according to standard protocols. Clear differentials between GFP+ and GFP- populations were observed.

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