nature research

Corresponding author(s): Yin Shen

Last updated by author(s): Jun 17, 2019

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.

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	Sequencing data was obtained from the HiSeq 4000, HiSeq X Ten, and NovaSeq 6000 systems (Illumina).
Data analysis	All software and code used in the text are now reported in the text along with their version numbers. We used the following software: fastp 0.20.0, HiCUP 0.71, bowtie 2.3.2, CHiCAGO 1.1.8, HOMER 4.10, HaploReg 4.1, HiC-Pro 2.11.0, bowtie 1.2.1.1, https://github.com/kundajelab/atac_dnase_pipelines (June 2018), STAR 2.7.0f, RSEM 1.3.1, TrimGalore 0.4.5, edgeR 3.20.9, Picard Tools 1.141, MACS2 2.1.1, HiCRep 1.4.0, DiffBind 2.6.6, DESeq2 1.18.1. A copy of the custom code used for all the data analysis and figure generation in this study can be viewed and downloaded at the following GitHub repository: https://github.com/stayingsong/brain_pchic.

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. 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 datasets used in this study (pcHi-C, ATAC-seq, RNA-seq, CUT&RUN, and chromosome-wide SNP phasing data) are available at the Gene Expression Omnibus under the accession number GSE113483. Open chromatin peaks and gene expression results for each cell type are also available on Zenodo through the following link: https://zenodo.org/record/3243977.
Data can be visualized on the WashU Epigenome Browser using the session bundle ID (session ID in parentheses): 6e375740-8e71-11e9-be37-cb77c4bbb5fc (brain_pchic_nature_genetics_00).
Alternatively, the data can also be visualized on the legacy WashU Epigenome Browser at the following link (session ID in parentheses): http://

epigenomegateway.wustl.edu/legacy/?genome=hg19&session=8OCs2rkpEA (brain_pchic_nature_genetics_00).
Tracks include ATAC-seq signal, chromatin interactions with score >= 5, and RNA-seq plus and minus strand signal for each cell type. HindIII fragments, in vivo-validated enhancer elements, GENCODE 19 genes, and GWAS SNPs are also displayed.

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
Our study design, which ranges from between two to two and four to four pairwise comparisons across the various assays, has approximately 80% power to detect a mean difference of 2.39 to 5.66 standard deviations at a nominal significance threshold of 0.05. We chose this study design because effect sizes of this magnitude are compatible with our research goals.

Data exclusions
No data were excluded from the analyses.

Replication
All attempts at replication were successful and are described in the text.

Randomization
Randomization and blinding is not relevant to our study because we do not apply any differential treatment, intervention, or perturbation to our samples. Instead, we compare the epigenomic profiles for different cell types.

Blinding
Randomization and blinding is not relevant to our study because we do not apply any differential treatment, intervention, or perturbation to our samples. Instead, we compare the epigenomic profiles for different cell types.

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
☒	☐	Animals and other organisms
☒	☐	Human research participants
☒	☐	Clinical data

Methods

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

Antibodies

Antibodies used
We used the following primary antibodies for immunofluorescence: CUX1 (CASP) (Abcam, ab54583, lot: GR3224721-2, 1:500 dilution), MAP2 (Abcam, ab5392, lot: GR3242762-1, 1:1000 dilution), PROX1 (Millipore, MAB5654, lot: 3075604, 1:500 dilution), HB9 (Millipore, ABN174, lot: 3050643, 1:500 dilution), SMI32 (Abcam, ab7795, lot: GR299862-23, 1:1000 dilution), and GFAP (Abcam, ab7260, lot: GR3240356-1, 1:1000 dilution).

Secondary antibodies for immunofluorescence included: Alexa Fluor 594 goat anti-chicken IgG (Thermo Fisher Scientific, A11042, lot: 1977707, 1:500 dilution), Alexa Fluor 568 donkey anti-mouse IgG (Thermo Fisher Scientific, A10037, lot: 1917938, 1:500 dilution), Alexa Fluor 488 donkey anti-rabbit IgG (Thermo Fisher Scientific, A21206, lot: 1981155, 1:500 dilution), and Alexa Fluor 488 donkey anti-mouse IgG (Thermo Fisher Scientific, A21202, lot: 2018296, 1:500 dilution).

We used the following antibodies for CUT&RUN: H3K27ac from Active Motif, 39122, lot: 22618011, 1:100 dilution, and CTCF from Millipore, 07-729, lot: 305960, 1:100 dilution.

Validation
All primary antibodies for immunofluorescence recognize human proteins and are verified for staining on the manufacturers' websites.

CUX1 (Abcam, ab54583, lot: GR3224721-2)
Mouse monoclonal [2A10] to Protein CASP; Recombinant fragment (GST-tag) corresponding to Human Protein CASP aa 521-621; Suitable for: IHC-P, ICC, WB, IHC-FoFr, ICC/IF, Sandwich ELISA, Flow Cyt; Reacts with: Mouse, Human; This product has been

referenced in 10 publications.

MAP2 (Abcam, ab5392, lot: GR3242762-1)
Chicken polyclonal to MAP2; Recombinant fragment corresponding to Human MAP2. Mix of recombinant human constructs of projection domain sequences, amino acids 235-1588; Suitable for: ELISA, IHC-Fr, IHC-FoFr, IHC-P, WB, ICC/IF, IHC (PFA fixed); Reacts with: Mouse, Rat, Sheep, Cow, Dog, Human, Cynomolgus monkey, Common marmoset, Aplysia; This product has been referenced in 274 publications.

PROX1 (Millipore, MAB5654, lot: 3075604)
Monoclonal Antibody; Protein A Purfied; Recombinant human Prox1 protein; Anti-Prox1 Antibody, clone 4G10 is an antibody against Prox1 for use in WB, IH; Validated using positive control (mouse dentate granule neurons) and negative control (secondary antibody only). Also validated in fresh formaldehyde-fixed human hippocampal tissue with specific staining in the dentate granule layer.

HB9 (Millipore, ABN174, lot: 3050643)
Polyclonal Antibody; Affinity Purified; KLH-conjugated linear peptide corresponding to human MNX1; Anti-MNX1 (HB9) Antibody detects level of MNX1 (HB9) & has been published & validated for use in Western Blotting & IHC.

SMI32 (Abcam, ab7795, lot: GR299862-23)
Mouse monoclonal [NF-01] to Neurofilament heavy polypeptide; This antibody recognizes a phosphorylated epitope on heavy neurofilament protein (210 kDa) of various species; Suitable for: ELISA, IHC-Fr, ICC, IHC-P, WB, IHC - Wholemount, ICC/IF, Flow Cyt; Reacts with: Mouse, Rat, Cow, Human, Pig; Predicted to work with: a wide range of other species, Mammals; This product has been referenced in 11 publications.

GFAP (Abcam, ab7260, lot: GR3240356-1)
Rabbit polyclonal to GFAP; Specifically recognizes mammalian GFAP on western blots and immunocytochemically. Detects a band of 55kDa corresponding to GFAP and also a GFAP derived 48kDa band; Suitable for: IHC-FoFr, IHC-Fr, IHC-FrFl, ICC/IF, WB, IHC-P, IHC - Wholemount, ICC; Reacts with: Mouse, Rat, Cat, Dog, Human, Common marmoset; Predicted to work with: Cow, Pig, Mammals; This product has been referenced in 343 publications.

Alexa Fluor 568 goat anti-chicken IgG (Thermo Fisher Scientific, A11042, lot: 1977707)
Goat / IgG Polyclonal to Chicken; This product has been referenced in 26 publications.

Alexa Fluor 568 donkey anti-mouse IgG (Thermo Fisher Scientific, A10037, lot: 1917938)
Donkey / IgG Polyclonal to Mouse; This product has been referenced in 32 publications.

Alexa Fluor 488 donkey anti-rabbit IgG (Thermo Fisher Scientific, A21206, lot: 1981155)
Donkey / IgG Polyclonal to Rabbit; This product has been referenced in 70 publications.

Alexa Fluor 488 donkey anti-mouse IgG (Thermo Fisher Scientific, A21202, lot: 2018296)
Donkey / IgG Polyclonal to Mouse; This product has been referenced in 68 publications.

H3K27ac and CTCF antibodies were validated by the ENCODE project and have been used in many publications.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	We used excitatory neurons (i3N iPSCs) from co-author Dr. Li Gan's lab (Mertens et al., 2016 Nature and Wang et al., 2017 Stem Cell Reports), hippocampal DG-like neurons from co-author Dr. Hongjun Song's lab, and lower motor neurons (i3LMN iPSCs) from co-author Dr. Bruce Conklin's lab (Fernandopulle et al., 2018 Curr Protoc Cell Biol). Two batches of primary astrocytes derived from two individuals were purchased from ScienCell. The mouse embryonic fibroblasts (MEF) were isolated in Dr. Hongjun Song's lab. They were derived from E13 embryos extracted from pregnant CF1 mice (Charles River, Strain Code: 023). The MEFs were irrated with 3000 rads before plating them as feeder cells.
Authentication	We checked the expression of key marker genes for each cell type using immunofluorescence and RNA-seq. For excitatory neurons, we used VGLUT1 and CUX1, for hippocampal DG-like neurons, we used SOX2 and PROX1, for lower motor neurons, we used HB9 and SMI32, and for astrocytes, we used GFAP.
Mycoplasma contamination	All cells used in the present study were verified as mycoplasma contamination free.
Commonly misidentified lines (See ICLAC register)	None of the cell lines used are commonly misidentified lines.