nature research

Corresponding author(s): Fred H. Gage

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	For data collection, the following software was used: Zeiss ZEN (black edition; Carl Zeiss Microscopy), Zeiss ZEN (blue edition; Carl Zeiss Microscopy), Neurolucida v11 (MBF Bioscience), StereoInvestigator v11 (MBF Bioscience) and the BD FACSDiva acquisition software (Becton-Dickinson).
Data analysis	The following R packages were used for data analysis: DESeq2 v1.20.0, WGCNA v1.51, dtw v1.20-1, DiffBind v2.4.8, edgeR v3.18.1, Gviz v1.20.0. Other packages and software used in this study included: STAR v2.5.2a, Cudadapt v1.6 with Python 3.4.1, FastQC v0.10.1, Bowtie2 v2.2.8, PicardTools MarkDuplicates v2.18.3-SNAPSHOT, samtools v1.6, deeptools2 v2.5.4, macs2 v2.1.1.20160309, Homer, GAT v1.3.5. and GraphPad Prism v6.

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 datasets are available at EMBL-EBI ArrayExpress with the accession code: E-MTAB-6018. This raw data is associated with the following figures: 1, 2a-c, 4i-k, S3c, S4, S7a-e, S8f-g and S10. Additional data generated or analyzed during this study is included in this published article and its supplementary information files. Supplementary tables are available for the following figures: 1, 2a-c, 4i-k, S3c, S4a, S7a-d, S8f-g and S10 as well as for additional differential expression analysis (Supplementary table 8). Supplementary table 1 (Supplementary Data and Notes) is providing additional information about all patient-derived iPSC lines used in this study. ATAC-seq datasets are available from the corresponding author upon 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 but our sample sizes are similar to those reported in previous publications and were estimated based on previous experience with a similar setup that showed significance (Marchetto et al., 2016).
Data exclusions	All biological assays were performed with iPSCs derived from eight ASD patients and five healthy controls (nASD=8, ncontrol=5; all male), with the exception of the organoid validation experiments, for which we randomly selected a subset of three ASD patients and three controls (nASD=3, ncontrol=3; Fig.3 and Extended Data Fig.6f-l). No other collected samples were excluded from the analysis.
Replication	Biological assays were performed with iPSCs derived from eight ASD patients and five healthy controls (nASD=8, ncontrol=5; all male), with the exception of the organoid validation experiments, for which we randomly selected a subset of three ASD patients and three controls (nASD=3, ncontrol=3; Fig.3 and Extended Data Fig.6f-l). In all experiments, two lines from each patient were prepared, and one or two lines were eventually used for the experiment. Replications were defined as such in the associated figure legends. All attempts of replication were successful.
Randomization	Cell culture experiments were not completely randomized, however all biological assays and sequencing samples were randomly assigned to different experimental groups and batches to control for covariates.
Blinding	For data collection, every cell line had a unique code that could not reveal the identity of the subject but could tell which lines belonged to the same subject, so that the investigators were blinded with regard to the group category. For data analysis (counting, tracing, imaging and FACS analysis), technicians and other investigators were blinded for group categories.

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

All antibodies used in this study were commercially available antibodies. FACS and flow cytometry were performed with the following antibodies: APC-conjugated PSA-NCAM (Miltenyi Biotec, cat.no. 130-093-273, 1:50) and APC-conjugated IgM Isotype control (Miltenyi Biotec, cat.no. 130-093-176, 1:50), all antibodies supplied in the BD Stemflow™ Human Neural Cell Sorting Kit (BD Biosciences, cat.no. 562271, 1:20), FBXO2 (abcam, cat.no. ab133717, 1:100), 647-conjugated SOX2 (BD Biosciences, cat.no. 562139, 1:25) and 647-conjugated IgG1, k Isotype Control (BD Biosciences, cat.no. 562139, 1:25), 488-conjugated SSEA-4 (BioLegend, cat.no. 330411, 1:20) and 488-conjugated IgG3, k Isotype Control (BioLegend, cat.no. 401323, 1:25), 647-conjugated TRA1-81 (BioLegend, cat.no. 330705, 1:20) and 647-conjugated IgM, k Isotype Control (BioLegend, cat.no. 401618, 1:200).
The following antibodies were used for immunofluorescence: MAP2AB (Abcam, cat.non. ab5392, RRID: AB_2138153, 1:500), PSA-NCAM (EMD Millipore, cat.non. MAB5324, RRID: AB_95211, 1:500), Synapsin I (EMD Millipore, cat.non. 574778, RRID: AB_2200121, 1:200), Mouse anti-TUJ1/TUBB3 (BioLegend, cat.non. 801201, RRID: AB_2313773, 1:1000), Rabbit anti-TUJ1 (BioLegend, cat.no. 802001, RRID: AB_2564645, 1:2000), VGLUT1 (Synaptic Systems, cat.no. 135302, RRID: AB_887877, 1:1000), Chicken anti-GFP (Aves Labs, cat.no. GFP-1020, RRID: AB_10000240, 1:1000), Rabbit anti-GFP (Thermo Fisher Scientific, cat.no. A-6455, RRID: AB_221570, 1:1000), SMI312 (BioLegend, cat.no. 837904, RRID: AB_2566782, 1:1000), DCX (Santa Cruz, cat.no. sc-8066, RRID: AB_2088494, 1:500), FOXG1 (Abcam, cat.no. ab18259, RRID: AB_732415, 1:500), PAX6 (BioLegend, cat.no.901301, RRID: AB_2565003, 1:400), NESTIN (MilliporeSigma, cat.no. MAB5326, RRID: AB_2251134, 1:500), TBR1 (Abcam, cat.no. ab31940, RRID:AB_2200219, 1:1000), SOX2 (Cell Signaling, cat.no. 3579, RRID: AB_2195767, 1:500), BrdU (Accurate Chemical, cat.no. OBT0030, RRID: AB_2313756, 1:250) and PHF1 (generously provided by Dr. Peter Davies, RRID: AB_2315150, 1: 250).

Validation

FACS and flow cytometry antibodies: APC-conjugated PSA-NCAM (Miltenyi Biotec, cat.no. 130-093-273, 1:50) and APC-conjugated IgM Isotype control (Miltenyi Biotec, cat.no. 130-093-176, 1:50). Validations and references at https://www.miltenyibiotec.com/US-en/products/macs-flow-cytometry/antibodies/primary-antibodies/anti-psa-ncam-antibodies-human-mouse-rat-2-2b-1-11.html#apc:for-100-tests. All antibodies supplied in the BD Stemflow™ Human Neural Cell Sorting Kit (BD Biosciences, cat.no. 562271, 1:20). Validations and references at http://www.bdbiosciences.com/us/applications/research/stem-cell-research/stem-cell-kits-and-cocktails/human/human-neural-cell-sorting-kit/p/562271. FBXO2 (abcam, cat.no. ab133717, 1:100). Validations at https://www.abcam.com/fbxo2-antibody-epr73282-ab133717-references.html. 647-conjugated SOX2 (BD Biosciences, cat.no. 562139, 1:25) and 647-conjugated IgG1, k Isotype Control (BD Biosciences, cat.no. 562139, 1:25). Validations and references at https://www.bdbiosciences.com/eu/applications/research/intracellular-flow/intracellular-antibodies-and-isotype-controls/anti-human-antibodies/alexa-fluor-647-mouse-anti-sox2-o30-678/p/562139. 488-conjugated SSEA-4 (BioLegend, cat.no. 330411, 1:20) and 488-conjugated IgG3, k Isotype Control (BioLegend, cat.no. 401323, 1:25). Validations and references at https://www.biolegend.com/en-us/products/alexa-fluor-488-anti-human-ssea-4-antibody-4825. 647-conjugated TRA1-81 (BioLegend, cat.no. 330705, 1:20) and 647-conjugated IgM, k Isotype Control (BioLegend, cat.no. 401618, 1:200). Validations and references at https://www.biolegend.com/fr-ch/products/alexa-fluor-647-anti-human-tra-1-81-antibody-4827.
The following antibodies were used for immunofluorescence: MAP2AB (Abcam, cat.non. ab5392, RRID: AB_2138153, 1:500), Validations and references at https://www.abcam.com/map2-antibody-ab5392.html. PSA-NCAM (EMD Millipore, cat.non. MAB5324, RRID: AB_95211, 1:500). Validations and references at http://www.emdmillipore.com/US/en/product/Anti-Polysialic-Acid-NCAM-Antibody-clone-2-2B,MM_NF-MAB5324. Synapsin I (EMD Millipore, cat.non. 574778, RRID: AB_2200121, 1:200). Validation and references at http://www.emdmillipore.com/US/en/product/Anti-Synapsin-I-Rabbit-pAb,EMD_BIO-574778. Mouse anti-TUJ1/TUBB3 (BioLegend, cat.non. 801201, RRID: AB_2313773, 1:1000). Validations and references at https://www.biolegend.com/fr-ch/products/purified-anti-tubulin-beta-3-tubb3-antibody-11580. Rabbit anti-TUJ1 (BioLegend, cat.no. 802001, RRID: AB_2564645, 1:2000). Validations and references at https://www.biolegend.com/fr-ch/products/purified-anti-tubulin-beta-3-tubb3-antibody-11579. VGLUT1 (Synaptic Systems, cat.no. 135302, RRID: AB_887877, 1:1000). Validations and references at https://www.sysy.com/products/vglut1/facts-135302.php. Chicken anti-GFP (Aves Labs, cat.no. GFP-1020, RRID: AB_10000240, 1:1000). Validations and references at http://www.aveslab.com/products/epitope-tag-and-gfp-antibodies/anti-gfp-green-fluorescent-protein-antibodies-2/. Rabbit anti-GFP (Thermo Fisher Scientific, cat.no. A-6455, RRID: AB_221570, 1:1000). Validations and references at https://www.thermofisher.com/antibody/product/GFP-Tag-Antibody-Polyclonal/A-6455. SMI312 (BioLegend, cat.no. 837904, RRID: AB_2566782, 1:1000). Validations and references at https://www.biolegend.com/en-us/products/purified-anti-neurofilament-marker-pan-axonal-cocktail-12811. DCX (Santa Cruz, cat.no. sc-8066, RRID: AB_2088494, 1:500). Validations and references at https://www.scbt.com/scbt/product/doublecortin-antibody-c-18. FOXG1 (Abcam, cat.no. ab18259, RRID: AB_732415, 1:500). Validations and references at https://www.abcam.com/foxg1-antibody-chip-grade-ab18259-references.html. PAX6 (BioLegend, cat.no.901301, RRID: AB_2565003, 1:400). Validations and references at https://www.biolegend.com/en-us/products/purified-anti-pax-6-antibody-11511. NESTIN (MilliporeSigma, cat.no. MAB5326, RRID: AB_2251134, 1:500). Validations and references at http://www.emdmillipore.com/US/en/product/Anti-Nestin-Antibody-clone-10C2,MM_NF-MAB5326. TBR1 (Abcam, cat.no. ab31940, RRID:AB_2200219, 1:1000). Validations and references at https://www.abcam.com/tbr1-antibody-ab31940.html. SOX2 (Cell Signaling, cat.no. 3579, RRID: AB_2195767, 1:500). Validations and references at https://www.cellsignal.com/products/primary-antibodies/sox2-d6d9-xp-rabbit-mab/3579. BrdU (Accurate Chemical, cat.no. OBT0030, RRID: AB_2313756, 1:250). Validations and references at http://antibodyregistry.org/search.php?q=AB_2313756. PHF1 (generously provided by Dr. Peter Davies, RRID: AB_2315150, 1: 250) is specific towards phosphorylated human tau (hTau) and was used as published previously (Mertens et al., 2015).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

A detailed description of all reprogrammed iPSCs is listed in Supplementary Table 1. Additional information was published in Marchetto et al., Mol. Psychiatry (2016). The iPSC lines obtained from this cohort were reprogrammed in the same facility (Salk Institute for Biological Studies, Laboratory of Genetics) and under the same conditions. Briefly, fibroblasts were transduced with retroviruses containing SOX2, OCT4, KLF4 and MYC to induce overexpression of these genes and were transferred to a co-culture system with murine embryonic fibroblasts. IPSC colonies were identified after around two weeks

	in this culture system, plated onto Matrigel-coated plates (BD Biosciences) and maintained in mTeSR1 media (Stem Cell Technologies). Experiments were performed with 8 independent ASD patient lines (referred to as: Able, Avid, Aqua, Arch, Ahoy, Aero, Acai, Apex) and 5 independent control lines (Clue, Cent, Cove, Chap and Cent) with one or two clones per line. The above mentioned identifiers, accompanied by their respective clone ID, have also been used to allow sample identification in the deposited raw sequencing data.
Authentication	The characterization and authentication of these lines was recently performed in Marchetto et al., Mol. Psychiatry (2016). Here, analyses of patient blood cells, fibroblasts and iPSCs were performed using CNV analysis, whole-exome sequencing and G-banding analysis to authenticate and characterize patient-derived cell lines. For subsequent experiments, iPSC lines were grown separately and authentication was performed through the use of standardized names and clone-specific unique identifiers.
Mycoplasma contamination	Cell lines were regularly tested for mycoplasma contamination. All cell lines used for this study were tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Human research participants

Policy information about studies involving human research participants

Population characteristics	IPSCs were generated from eight male subjects diagnosed with ASD according to the DSM-IV (average age 13.25 ± 5.67 years) and five male control participants (average age 11.4 ± 5.22 years), which had no history of psychological, genetic or other disorders (also shown in Supplementary Table 1). ASD subjects were selected from lists of ASD subjects with larger than normal average total brain volume (at 3.56 ± 0.67 years) as measured by MRI and compared with typically developing toddlers.
Recruitment	Subjects were recruited through the University of California San Diego, Autism Center of Excellence from a pool of volunteers formerly included in previous brain imaging studies. Control subjects were selected randomly from lists of typically developing individuals who had had the magnetic resonance imaging (MRI) scan when they were toddlers. ASD subjects were selected from lists of ASD subjects with larger than normal average total brain volume as compared with typically developing toddlers. The ASD subjects demonstrated a behavioral presentation consistent with autism as defined by the criteria set forth in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; APA). Assessment of the history and presence of the disorder was achieved via standardized behavioral, cognitive and functional assessments, including the appropriate Wechsler Intelligence Scale, the Autism Diagnostic Observation Schedule, the Autism Diagnostic Interview, Revised and Vineland Adaptive Behavior Scales. Participants in the control group had no history of psychological, genetic or other disorders. The recruitment criteria thus was specific to a subset of autistic patients, which show early signs of macrocephaly. We have highlighted this in the manuscript and abstract to specify the sampling of a highly stratified subset of macro-cephalic ASD patients.

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	NSCs and maturing neuron cultures as well as iPSCs and iPSC-iNs were detached using accutase (Innovative Cell Technologies) and then washed and stained with fluorophore-coupled antibodies for 45 min at 4ºC in phosphate buffered saline (PBS). Cells were washed, suspended in PBS and filtered using a 40-µm cell strainer.
Instrument	Flow cytometry analysis was performed with a BD LSR II (Becton-Dickinson) and the BD InFlux Cytometer (Becton-Dickinson) was used for FACS-based purification of neuronal cell types.
Software	Flow cytometry data was analyzed using the FlowJo 10.4.2 software.
Cell population abundance	For flow cytometry analysis (measuring MFIs for reporter assays), a similar number of cells were analyzed for each replicate as indicated in the Supplementary Information (50,000 cells). Gates and histograms are shown for the populations analyzed and percentages for cell populations were indicated.
Gating strategy	Gating strategies for flow cytometry analysis are shown in the figures, including the respective isotype controls to determine the gates. Briefly, for FACS sorting we gated on APC-conjugated PSA-NCAM+ and GFP+ populations, which we determined in every experiment by using the same non RV-infected cell line stained with the appropriate APC-conjugated Isotype control antibody. Gating for the flow cytometry analysis of reporter activity (hDCX and TCF/LEF) was performed similar by gating on APC-conjugated PSA-NCAM+ and GFP+ populations.

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