nature portfolio

Corresponding author(s): Scherer, Stephen W. and Bradley, Clarrisa A.

Last updated by author(s): Mar 26, 2026

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 (*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	Benchling (http://benchling.com/) UltraSoundGate system (Avisoft Bioacoustics, DE) Limelight software (Actimetrics) FreezeFrame scoring system (Actimetrics, IL) Bussey-Saksida Touchscreen System (Lafayette Instrument, IN) Multiclamp 700B (Molecular Devices) Bio-Rad QX200 Droplet Reader
Data analysis	ERDS CNVnator Mutect2 (GATK3.7) MUPET MATLAB R2022a (MathWorks, MA) LimeLight software (Actimetrics, Wilmette, IL) (v3) WinLTP software Quantasoft (version 1.7.4) FastQC (v.0.11.5) Trim Galore (v. 0.5.0) FastQ-Screen26 (v.0.10.0) RseQC package27 (v.2.6.2) STAR aligner28 (v.2.6.0c)

DESeq2 package30, (v.1.26.0)
R (v3.6.1 (R Core Team (2019))
DoubletFinder (v2.0.3)
SoupX (v1.6.1)
Seurat (v4.3.0)
Signac (v1.8.0)
MACS2 (v2.2.7.1)
10x Genomics Cell Ranger (cellranger-arc-2.0.1)
SCTransform (v0.3.5)
ChatGPT 4.0 (Data analyst)
Cytoscape (v3.10.1)
Enrichment map plug-in
Proteomics Analysis workflow (PAW) (https://github.com/pwilmart/PAW_pipeline)
Bioconductor package edgeR
Prism built-in analysis (version 10 or earlier; GraphPad Software Inc., MA)
GSEA v4.0.3 with gene-sets from Broad (https://www.gsea-msigdb.org/gsea/msigdb/) downloaded in 2023

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

WGS data from MSSNG and SFARI datasets can be obtained at https://research.mss.ng and https://www.sfari.org/resource/sfari-base/ as was done for this study. Public databases used in the study can be accessed through the following sites: 1000 Genomes Project (https://www.internationalgenome.org/); Medical Genome Reference Bank (MGRB) (https://doi.org/10.1038/s41467-019-14079-0). Additional control datasets can be requested at CHILD (https://ega-archive.org/dacs/EGAC00001002953); HostSeq (https://www.cgen.ca/hostseq-databank-access-request) and Inova (www.inova.org; VA, USA).. Data from the cardiology datasets are available through (https://ega-archive.org/studies/EGAS50000000586) and (https://ega-archive.org/studies/EGAS00001004929)snRNA-seq data can be found at NCBI Bioproject, RNA Accession: PRJNA1195946, and the proteomic data at MassIVE (Mass Spectrometry Interactive Virtual Environment @ UCSD), Proteomics Accession: MSV000096677. Source data supporting the findings of this study are provided with this paper and can be requested from authors S.W.S and C.A.B.

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	There were no restrictions on the sex of the individual to enter the study however, autism does present in a four-fold excess in males. By examining copy number variants in individuals with autism we noted that there was a locus at Xp22.11 that appeared to be more frequently deleted in individuals with autism as compared to controls. Our human genomic data presented in this manuscript examines deletions found in males only as they are expected to be more deleterious.
Reporting on race, ethnicity, or other socially relevant groupings	We used only computed ancestry based on genotype. All ancestry groups available were included in the study without restrictions. Details on the computed ancestry can be found from the autism controlled-access databases described in the data access statement.
Population characteristics	Demographics of all participants including ancestry and sex are provided in the supplementary table and is available from the respective databases. All participants were between 2-18 year old at the time of analysis.
Recruitment	Data utilized from controlled-access databases as described in the data access statement.
Ethics oversight	This research involving human participants complies with all relevant ethical regulations including obtaining informed consent from all participants through the recruitment sites. Ethical review and approval was obtained from The Hospital for Sick Children Research Ethics Board (REB# 1000080561).

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

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For the human genetics, there was no sample size determined. We used all available datasets to search the Xp11.2 locus for variants. Mouse experiments: Two mutant mice were generated using different guide RNAs to create the deletion required for experiments. The number of mice and sample sizes were determined based on established standards in the field and previous experience. No statistical methods were used to predetermine sample sizes. The number of mice and sample sizes were determined based on established standards in the field and previous experience (e. g. Frankland et al. Molecular Psychiatry 2004).
Data exclusions	Human genetics: our analysis exclusively considered high-quality deletions based on the following criteria: i. Length of at least 1kb, ii. Identification by both ERDS and CNVnator, with at least 50% reciprocal overlap between the two methods; iii. Less than 70% overlap with repetitive or low complexity genomic regions (such as telomeres, centromeres, and segmental duplications); iv. Exclusion of CNVs in PAR for X chromosomal calls in males. Mouse experiments: we used genotype analysis to confirm successful gene-editing with no exonic off-target effects by WGS. Experiments were performed on the genotype confirmed mice and no high quality data generated from these experiments was excluded.
Replication	Mouse experiments: Sample sizes for replication were determined based on established standards in the filed and previous experience. General measures taken to verify the reproductivity include multiple cohorts across several years of experiments and all attempts at replication were backcrossing was controlled were successful. We noted experimental means due to the length of time generations were inbreed. Outbreeding methods were increased to every three generations.
Randomization	Mouse experiments: Male hemizygous mutant (KO) and WT mice were housed. Researchers were blinded to genotype and given group codes that were also randomized for each cohort/batch. These methods were used for all electrophysiology, western blotting and behavior experiments.
Blinding	Mouse experiments: All mouse experiments were conducted and analyzed in a blinded manner, with the experimenter unaware of the experimental conditions. The researcher was unaware of what group the mice were in (WT or KO).

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

Please also refer to Supplementary Information (Supplementary Methods Table 3)
GFAP, Cat# 5-240, Neuromab, Ms
FMRP, Cat#4317, Cell Signaling Technology, Rb
GluN2A, Cat# MA5-27692, Invitrogen, Ms
pGluN2A Y1426, Cat# 4206, Cell Signaling Technology, Rb
GluN2B Cat#66565-1, Proteintech, Ms
pGluN2B Y1472 Cat#4208 Cell Signaling Technology, Rb
PKC alpha Cat#5578-MSM2-P0 Thermo Fisher Scientific, Ms
Src, Cat #2110, Cell Signaling Technology, Ms
pSrc Y416, Cat #6943, Cell Signaling Technology, Rb
pGSK-3α/β S21/9 Cat#9331, Cell Signaling Technology, Rb
GSK-3α/β, Cat #44-610, Invitrogen, Ms
StarBright Blue 700 Anti-Rb, Cat#12004161, BioRad, Goat
StarBright Blue 520 Anti-Ms Cat#12005866, BioRad, Goat
StarBright Blue 700 Anti-Ms Cat#12004158, BioRad, Goat
StarBright Blue 520 Anti-Rb Cat#12005869 BioRad Goat

Validation	The Gfap KO was validated.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	Human fetal neural stem cells HF6562 were obtained from Dr. Peter Dirks lab at Hospital for Sick Children. This cell line is derived from human fetal brain tissue and has been previously used as a normal neural stem cell control in published studies (Refs. 1-3). This cell line corresponds to HF6562 as annotated in Cellosaurus (RRID:CVCL_C8ZT)
Authentication	HF6562 cells were not authenticated in our laboratory. Cell identity was based on the source laboratory.
Mycoplasma contamination	HF6562 cells were not tested for mycoplasma contamination in our laboratory. Cells were directly transferred from the originating laboratory internally and maintained under sterile conditions. Cells grew normally and no signs of contamination were observed during the course of the experiments.
Commonly misidentified lines (See ICLAC register)	HF6562 is not known to be among those listed in the ICLAC database of commonly misidentified cell lines.

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	Laboratory mice from the C57BL/6J strain were used in the experiments. Please see lines 845-848 for housing conditions "The mice were bred at the Hospital for Sick Children or The Centre for Phenogenomics, Neurobehavior Core and group-housed in cages with 3-5 mice per cage. The housing conditions maintained a constant temperature of 22 °C and a 12-hour light/dark cycle, with food and water available ad libitum."
Wild animals	This study did not involve wild animals.
Reporting on sex	We used all male mice for the experiments as the region that we are investigating is on human chromosome X and therefore we were most interested in the phenotype of the male mice.
Field-collected samples	This study did not involve samples collected in the field.
Ethics oversight	All procedures were approved by the Hospital for Sick Children Animal Care and Use Committee (AUP# 49151) and The Centre for Phenogenomics Animal Care and Use Committee (AUP# 25-0307; #20-0307H) and conducted in accordance with Canadian Council on Animal Care guidelines. Electrophysiological recordings were carried out at the Lunenfeld Tanenbaum Research Institute in Mount Sinai Hospital (Toronto) (AUP #24-0292H) and at the University of Toronto's, Tanz Centre for Research in Neurodegenerative Diseases, University Hospital Network (AUP #6668.8).

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

Plants

Seed stocks	Plants were not used as part of this study.
Novel plant genotypes	Plants were not used as part of this study.
Authentication	Plants were not used as part of this study.