nature portfolio

Corresponding author(s): Howard Lipshitz, Craig Smibert, Shreejoy Tripathy

Last updated by author(s): Jul 27, 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	Proteomics Data Collection: * Orbitrap Lumos Mass Spectrometer (Thermo Fisher Scientific) * Sequest/COMET algorithm — real-time peptide-spectrum matching during acquisition Long-read RNAseq Data Collection: * PacBio basecaller — default Revio parameters * ccs v8.0.1 — HiFi/CCS read generation
Data analysis	Long-read RNAseq Data Processing: * skera v1.2.0 — Kinnex S-read segmentation * lima v2.9.0 — demultiplexing and trimming * isoseq refine v4.2.0 — FLNC read generation * pbmm2 v1.13.1 — alignment to GRCh38 * SMRT Link v13.0.0.207600 — end-to-end workflow manager * samtools v1.20 — BAM merging and sorting * isoseq collapse — default settings * pigeon classify v1.3.0 — isoform annotation/classification * pigeon filter v1.3.0 — artifact filtering Short-read RNAseq Processing: * STAR v2.4.2a — alignment to GRCh38.p13/hg38 with GENCODE V47 annotations

* samtools v1.20 — BAM sorting
Proteomics Data Analyses:
* Comet v2024.01 rev. 1 — offline MS2 spectra searching
* Percolator v3.05.0 — PSM rescoring
Ribosome Profiling Reanalysis:
* cutadapt — adapter and poly(A) trimming (no version specified)
* STAR — alignment (no version specified in this section, though v2.4.2a used elsewhere)
* HTSeq-count — read quantification (no version specified)
Downstream Bioinformatic Analyses:
* ORFanage v1.1.0 — ORF prediction
* SQANTI3/pigeon — isoform and protein classification framework (version covered under pigeon above)
* edgeR v4 — CPM normalization
* DESeq2 — differential gene and transcript expression
* IsoformSwitchAnalyzeR — DTU analysis
* spliceR — alternative splicing event identification
* Pfam — protein domain annotation
* IUPred2A — intrinsic disorder prediction
* MSstatsTMT — differential protein expression
* gprofiler2 — GO term enrichment
* SUPPA2 v2.4 — alternatively spliced exon identification
* emmeans R package v2.0.2 — estimated marginal means for GLMs
* QAPA v1.3.3 — alternative polyadenylation analysis
* VEP v114 — variant effect prediction for gnomAD variants
* ANNOVAR (2019Oct24 release) — variant annotation for ASD cohort DNMs

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 following secure token has been created to allow review of record GSE311240
while it remains in private status:
ulcvieyiblqnvax

The following secure token has been created to allow review of record GSE310131
while it remains in private status:
szmjegiodharrmd

Reviewer access details for the proteomics data is below:
Log in to the PRIDE website using the following details:
Project accession: PXD069736
Token: Wf7UVmoQX92s
Alternatively, reviewer can access the dataset by logging in to the PRIDE website
using the following account details:
Powered by Editorial Manager® and ProduXion Manager® from Aries Systems CorporationUsername: reviewer_pxd069736@ebi.ac.uk
Password: M2HwXlucSlQx

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	We employed an engineered human induced pluripotent stem cell (iPSC) line from a neurotypical male donor harboring a doxycycline-inducible NGN2 construct for targeted differentiation into excitatory cortical neurons
Reporting on race, ethnicity, or other socially relevant groupings	The population characteristics of the donor is available from Hilldebrandt et al., 2019 (10.1016/j.stemcr.2019.11.003) Table S1.
Population characteristics	The population characteristics of the donor is available from Hilldebrandt et al., 2019 (10.1016/j.stemcr.2019.11.003) Table S1.
Recruitment	No human participants were directly recruited for the iPSC-based experimental component of this study. The PGPC17_11 iPSC line was obtained from the Personal Genome Project Canada biobank, where the original donor provided informed

	consent for research use of their biological material under the ethics protocols governing that resource (Hildebrandt et al., 2019).
Ethics oversight	The use of the PGPC17_11 iPSC line was conducted in accordance with the ethical guidelines of the Personal Genome Project Canada, under which informed consent was obtained from the original donor for research use (Hildebrandt et al., 2019). All iPSC culture and differentiation experiments were conducted under ethics approval from the Research Ethics Board of the Hospital for Sick Children, Toronto, Ontario, Canada.

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	No sample size calculations were performed. n = 3–6 biological replicates per timepoint is consistent with, and in some cases exceeds, what has been used in comparable long-read proteogenomic studies of iPSC-derived neurons.
Data exclusions	No data were excluded from the analysis.
Replication	Three independent iPSC stocks were used as biological replicates, representing genuine biological independence rather than technical replication. Short-read RNAseq was used to independently validate 78.9% of novel splice junctions identified by long-read sequencing. Proteomics independently validated translation of novel isoforms.
Randomization	Randomization was not relevant as the grouping variable (developmental time point) is an intrinsic biological property that cannot be randomly assigned.
Blinding	Immunofluorescence images were coded prior to analysis and decoded only after quantification was complete, ensuring that cell counting was performed without knowledge of which differentiation stage each sample represented.

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	Primary Antibodies: Chicken anti-βIII tubulin | Sigma Aldrich | Catalog #AB9354 | 0.3 µg/mL Rabbit anti-OCT4 | Abcam | Catalog #ab19857 | 1 µg/mL Mouse anti-CTIP2 (BCL11B) | Sigma Aldrich | Catalog #MABE882 | 1 µg/mL Secondary Antibodies: Goat anti-Mouse IgG Alexa Fluor™-488 | Thermo Fisher | Catalog #A11029 | 2 µg/mL Goat anti-Chicken IgG Alexa Fluor™-555 | Thermo Fisher | Catalog #A21437 | 2 µg/mL Goat anti-Rabbit IgG Alexa Fluor™-647 | Thermo Fisher | Catalog #A21245 | 2 µg/mL
Validation	All primary antibodies were validated by the manufacturer for use in immunofluorescence applications on human cells.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	We employed an engineered human induced pluripotent stem cell (iPSC) line from a neurotypical male donor harboring a doxycycline-inducible NGN2 construct for targeted differentiation into excitatory cortical neurons.
Authentication	We verified correct integration by PCR and confirmed cellular identity at each differentiation stage using immunofluorescence markers (OCT4, CTIP2, βIII-tubulin).
Mycoplasma contamination	The cell line was tested for mycoplasma contamination and was not contaminated.
Commonly misidentified lines (See ICLAC register)	The study uses the PGPC17_11 iPSC line, which is not itself one of the commonly misidentified cell lines flagged by organizations like the International Cell Line Authentication Committee (ICLAC).

Plants

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