nature research

Corresponding author(s): Kristen Brennand

Last updated by author(s): Sep 25, 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	AxiS 2.4
Data analysis	Whole transcriptome RNAseq data was processed using publicly available code from Hoffman et. al 2017 Nature Communications and other published R packages. Code will be made available on github.com/nadschro/SZvariant-synergy. R version 3.4.0 R packages: 1] grid compiler stats4 parallel stats graphics grDevices [8] utils datasets methods base other attached packages: [1] sets_1.0-17 VennDiagram_1.6.17 [3] futile.logger_1.4.3 pheatmap_1.0.8 [5] GSEABase_1.38.1 graph_1.54.0 [7] annotate_1.54.0 XML_3.98-1.9 [9] AnnotationDbi_1.38.2 HTSanalyzeR_2.28.0 [11] igraph_1.1.2 gage_2.26.1 [13] GGally_1.3.2 knitr_1.18 [15] qvalue_2.8.0 reshape2_1.4.3 [17] xtable_1.8-2 colortools_0.1.5 [19] preprocessCore_1.38.1 DESeq2_1.16.1 [21] SummarizedExperiment_1.6.3 DelayedArray_0.2.7 [23] matrixStats_0.52.2 GenomicRanges_1.28.4 [25] GenomeInfoDb_1.12.2 IRanges_2.10.2 [27] S4Vectors_0.14.3 readr_1.1.1 [29] latex2exp_0.4.0 mixtools_1.1.0 [31] ape_4.1 ggrepel_0.6.5

[33] variancePartition_1.6.0 Biobase_2.36.2
[35] BiocGenerics_0.22.0 ggplot2_2.2.1
[37] gridExtra_2.3 edgeR_3.18.1
[39] limma_3.32.5 doParallel_1.0.10
[41] iterators_1.0.9 foreach_1.4.4
[43] synapseClient_1.15-1
loaded via a namespace (and not attached):
[1] backports_1.1.2 Hmisc_4.1-0
[3] splots_1.42.0 plyr_1.8.4
[5] lazyeval_0.2.0 splines_3.4.0
[7] gmp_0.5-13.1 BiocParallel_1.10.1
[9] digest_0.6.13 BiocInstaller_1.26.1
[11] htmltools_0.3.6 gdata_2.18.0
[13] magrittr_1.5 checkmate_1.8.5
[15] memoise_1.1.0 cluster_2.0.6
[17] prada_1.52.0 Biostrings_2.44.2
[19] colorspace_1.3-2 blob_1.1.0
[21] rrcov_1.4-3 RCurl_1.95-4.8
[23] genefilter_1.58.1 lme4_1.1-13
[25] survival_2.41-3 gtable_0.2.0
[27] zlibbioc_1.22.0 XVector_0.16.0
[29] RankProd_3.2.0 Rmpfr_0.6-1
[31] cellHTS2_2.40.0 DEoptimR_1.0-8
[33] scales_0.5.0 futile.options_1.0.0
[35] vsn_3.44.0 mvtnorm_1.0-6
[37] DBI_0.7 Rcpp_0.12.14
[39] htmlTable_1.11.1 foreign_0.8-69
[41] bit_1.1-12 Formula_1.2-2
[43] htmlwidgets_0.9 httr_1.3.1
[45] gplots_3.0.1 RColorBrewer_1.1-2
[47] acepack_1.4.1 pkgconfig_2.0.1
[49] reshape_0.8.7 nnet_7.3-12
[51] locfit_1.5-9.1 labeling_0.3
[53] rlang_0.1.6 munsell_0.4.3
[55] tools_3.4.0 RSQLite_2.0
[57] evaluate_0.10.1 stringr_1.2.0
[59] bit64_0.9-7 robustbase_0.92-7
[61] caTools_1.17.1 KEGGREST_1.16.1
[63] RBGL_1.52.0 nlme_3.1-131
[65] BioNet_1.36.0 biomaRt_2.32.1
[67] pbkrtest_0.4-7 rstudioapi_0.7
[69] png_0.1-7 affyio_1.46.0
[71] statmod_1.4.30 tibble_1.3.4
[73] geneplotter_1.54.0 pcaPP_1.9-72
[75] stringi_1.1.5 highr_0.6
[77] lattice_0.20-35 Matrix_1.2-11
[79] nloptr_1.0.4 data.table_1.10.4-3
[81] bitops_1.0-6 colorRamps_2.3
[83] R6_2.2.2 latticeExtra_0.6-28
[85] affy_1.54.0 hwriter_1.3.2
[87] KernSmooth_2.23-15 codetools_0.2-15
[89] lambda.r_1.1.9 MASS_7.3-47
[91] gtools_3.5.0 Category_2.42.1
[93] rprojroot_1.2 rjson_0.2.15
[95] GenomeInfoDbData_0.99.0 hms_0.3
[97] rpart_4.1-11 minqa_1.2.4
[99] rmarkdown_1.6 segmented_0.5-2.1
[101] base64enc_0.1-3Biobase_2.42.0
[102] DEGseq_1.38.0
BiocGenerics_0.28.0
limma_3.38.3
WGCNA_1.67
Other software
GraphPad PRISM 9
ImageJ
STAR
verifyBamID
featureCounts
Kallisto
Samtools
Bedtools

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

To facilitate improved sharing between stem cell laboratories, all hiPSCs have already been deposited at the Rutgers University Cell and DNA Repository (study 160; http://www.nimhstemcells.org/) and all sequencing data has been deposited to GEO (GSE137101 whole short-read RNA-seq and scRNA-seq and GSE137127 targeted short-read RNA-seq) and SRA (PRJNA563972 Isoseq). To facilitate improved reproducibility of our data analyses, all code has been deposited at github.com/zhushijia/STAR2bSMRT.

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	Sample size was determined by availability of hiPCS lines from donors carrying 2p16.3 deletions. Sample size for RNA-seq was > 30, which allows an adequate observation to take the benefits of the Central limit Theorem, i.e. normally distributed data.
Data exclusions	Three samples were excluded from whole-transcriptome RNA-seq analysis due incompatability of verifyBamID results and check on sample gender with the sample ID. These criteria were pre-determined to establish proper interpretation of resulting RNAseq data. Iso-seq samples which failed QC (as discussed in Supplementary Text 1) were also not included in the final data and this criteria was empirically determined and validated within the manuscript.
Replication	hiPSC-NPCs and hiPSC-neurons underwent multiple differentiations and multiple wells for each phenotypic assay. Whole-transcriptome RNAseq results were compared to post-mortem brain samples to assess replication of findings. All attempts at replication were successful.
Randomization	Samples without 2p16.3 deletions were allocated to control groups while those with 2p16.3 deletions were allocated according to the position of the deletion.
Blinding	Neuronal tracing was done blinded to genotype. Other experiments were not performed blinded given that the measurements were obtained through automated software following cell culture.

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	Rabbit anti-NESTIN, Millipore, ABD69, 1:500; Goat anti-SOX2, Santa Cruz, sc-17320, 1:500; Chicken anti-MAP2, ABcam, ab5392, 1:500; Mouse anti-SYN1 Millipore, 574778, 1:500;

Mouse anti-GFP ThermoFisher, MA1-952, 1:1000;
Rabbit anti-RFP Life Technologies,R10367, 1:1000;
Rabbit-anti-GABA, Sigma, A2052, 1:500;
Mouse-anti-GAD67, Millipore, MAB5406, 1:1000;
Alexa 488 Donkey-anti-Rabbit, Jackson immunoResearch 711-545-152, 1:500;
Alexa 568 Goat-anti-Chicken, Thermo FIsher Scientific A-11041, 1:500;
Alexa 647 Donkey-anti-mouse, Jackson immunoResearch 715-605-150, 1:500;

Validation

All antibodies were previously validated (in Brennand et al 2011, Brennand et al 2015, Ho et al 2017, Yang et al 2017) for immunocytochemistry in human cells.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

Skin biopsies were obtained from four cases, four unrelated controls and one non-carrier relative from an NIMH study into childhood-onset schizophrenia.

Authentication

Fibroblasts were re-genotyped using PsychChip and exome sequencing. hiPSCs were reprogrammed via sendal viral reprogramming (Life Technologies) and validated by karyotyping, gene expression and protein immunocytochemistry.

Mycoplasma contamination

Cells were tested for mycoplasma monthly with all test being negative.

Commonly misidentified lines
(See ICLAC register)

none

Human research participants

Policy information about studies involving human research participants

Population characteristics

No subjects were consented or recruited by the stem cell investigators. Only de-identified fibroblast samples, with limited clinical information (noted in Table 1) were provided by the team at NIMH. There were four controls, three males and one female as well as two male MZ twins with 3'NRXN1 deletion diagnosed with SA/SZ and a mother and son pair with 5' NRXN1 deletion diagnosed with BD+psychosis. All NRXN1 deletion cases had onset before the age of 17.

Recruitment

Subjects were recruited as part of a longitudinal study of childhood onset schizophrenia, since closed, conducted at the NIMH and led by Judith Rapoport.

Ethics oversight

Informed consent was obtained from all fibroblast donors at the National Institute of Mental Health, under the review of the Internal Review Board of the NIMH. All hiPSC work was reviewed by the Internal Review Board of the Icahn School of Medicine at Mount Sinai. This work was also reviewed by the Embryonic Stem Cell Research Oversight Committee at the Icahn School of Medicine at Mount Sinai.

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