nature research

Corresponding author(s): Jia Chen, Xingxu Huang, Li Yang

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	The single nucleotide variants (SNV) information was downloaded from NCBI ClinVar database (ftp://ftp.ncbi.nlm.nih.gov/pub/clinvar/tab_delimited/variant_summary.txt.gz);The pathogenic human allele sequences were downloaded from NCBI dbSNP database (https://www.ncbi.nlm.nih.gov/snp/?term=((%22pathogenic%22%5BClinical+Significance%5D+%20AND+%22snp%22%5BSNP+Class%5D+AND+homo+sapiens%5BOrganism%5D+)+AND(%22y%22%5BAllele%5D%20).
Data analysis	High-throught sequencing read quality was evaluated by FastQC (v0.11.4) and then aligned to reference sequences by BWA (v0.7.9a). Substitution calling was performed by SAMtools (v0.1.18) mpileup function and indel calling was based on the CIGAR values of the SAM/BAM files (Li et al, Nat Biotechnol, 2018). BE-targetable genetic variants were selected by judging whether a nearby PAM sequence exists with a Python (v.2.7.13) script (Kim et al, Nat Biotechnol, 2017).

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.

Nature Biotechnology: doi:10.1038/nbt.4198

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

The deep-sequencing data from this study have been deposited in the NCBI Gene Expression Omnibus (accession no. GSE114999) and the National Omics Data Encyclopedia (accession no. OEP000030). The dataset will be released prior to publication.

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. Experiments were performed three times independently unless indicated. In previous studies using related experiments, the sample size has been determined to be sufficient to ensure reproducibility.
Data exclusions	No data were excluded
Replication	The experimental findings in all figures were reproduced successfully
Randomization	Samples were not randomized
Blinding	The investigators were not blinded to group allocation

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	Antibodies were purchased from the following sources: against alpha-tubulin (T6199) - Sigma; against Cas9 (ab204448) - Abcam.
Validation	The alpha-tubulin (T6199, Sigma) antibody has been validated by western blot in human HeLa cell lysate (https://www.sigmaaldrich.com/catalog/product/sigma/t6199?lang=en&region=US); the Cas9 (ab204448, Abcam) antibody has been validated by western blot in human 293T cells transfected with DDDDK tag-CRISPR-Cas9 lysate (http://www.abcam.com/crispr-cas9-antibody-ab204448.html#description_images_1).

Nature Biotechnology: doi:10.1038/nbt.4198

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T cell line was from ATCC
Authentication	No cell lines were authenticated
Mycoplasma contamination	HEK293T cell line has been tested negative for mycoplasma contamination by PCR methods
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used

Nature Biotechnology: doi:10.1038/nbt.4198