nature research

Corresponding author(s): Timothy Bredy; Xiang Li

Last updated by author(s): Dec 19, 2018

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

Mass spectrometry detection was performed using an API5500 triple quadrupole and Analyst software 1.7 version (AB Sciex); Dot blot and western blot data were collected by by the Odyssey Fc system and Image studio version 5.0 (Licor) ; Quantitative PCR was performed on a RotorGeneQ (Qiagen) cycler and RotorGene Q software version 2.3.1.49; DpnI-seq data is collected by Hiseq4000 (illumina) and local run manager software 2.0 version. N6amt1 chIP-seq and RNA-seq is collected by Hiseq2500 (illumina) and local run manager software 2.0 version; Animal behavioral data was collected by Freezeframe (Colbourn) software version 4.

Data analysis

DpnI-seq Data analysis: Illumina pair-end sequencing data was aligned to the mouse reference genome (mm10) using BWA (v0.6.2) 6. Samtools (v0.1.17) 7was then used to convert "SAM" files to "BAM" files, sort and index the "BAM" files, and remove duplicate reads. Reads with low mapping quality (<20) or reads that were not properly paired-end aligned to the reference genome were excluded from the downstream analysis. These steps ensure that only high-quality alignments were used for the analysis of DpnI cleavage sites (Suppl. Table 2). After alignment, we applied a similar approach that infers potential DpnI cleavage sites based on the position of 5'ends as described in a previous study 5. Briefly, a binominal distribution model was assumed that each read could be randomly sheared and aligned to the genome with a probability p = 1/gs (gs = genome size) or cleaved by DpnI. For each individual sample, let n be the total number of reads. The P value of each genomic locus supported by x number of reads was calculated as C_n^x p^x 〖(1-p)〗^(n-x). Bonferroni correction was then applied for multiple testing correction. A genomic locus was determined as a real DpnI cleavage site if it satisfies the following criteria: i) the corrected P value < 0.01 in at least 2 of the 3 biological replicates in one condition or both conditions, and ii) the locus is not in the mm10 empirical blacklists identified by the ENCODE consortium 8.
N6amt1 ChIP-Seq data analysis: We performed the paired-end reads alignment and filter using the same analysis workflow as described in DpnI-Seq data analysis. After removed duplicate reads, low mapping quality reads, and not properly paired-end aligned reads, MACS2 (version 2.1.1.20160309) was used to call peaks for each sample with the parameter setting "callpeak -t SAMPLE -c INPUT -f BAMPE --keep-dup=all -g mm -p 0.05 -B". Peak summits identified by MACS2 from all samples were collected to generate a list of potential binding sites. Custom PERL script was then applied to parse the number of fragments (hereafter referred as counts) that cover the peak summit in each sample. Each pair of properly paired-end aligned reads covering the peak summit represents one count. The total counts in each sample were normalised to 20 million before comparison among samples. The potential binding sites were kept if they met all of the following conditions: i) the sites were not located in the mm10 empirical blacklists, and ii) the normalized counts in all three biological

replicates in one group were larger than that in its input sample, and the normalized counts in at least 2 replicates were more than 2-folds larger than its normalized input count.
RNA-Seq data analysis: Illumina paired-end reads were aligned to the mouse reference genome (mm10) using HISAT2 (version 2.0.5), with the parameter setting of "--no-unal --fr --known-splicesite-infile mm10_splicesites.txt". The "htseq-count" script in HTSeq package (v0.7.1) (http://www-huber.embl.de/HTSeq) was used to quantitate the gene expression level by generating a raw count table for each sample. Based on these raw count tables, edgeR (version 3.16.5) was adopted to perform the differential expression analysis between groups. EdgeR used a trimmed mean of M-values to compute scale factors for library size normalization. Genes with counts per million (CPM) > 1 in at least 3 samples were kept for downstream analysis. We applied the quantile-adjusted conditional maximum likelihood (qCML) method to estimate dispersions and the quasi-likelihood F-test to determine differential expression. Differentially expressed genes between two groups were identified when FDR < 0.05. Gene ontology enrichment analysis for differentially expressed genes was performed using the functional annotation tool in DAVID Bioinformatics Resources (version 6.8).
Other data analysis that are include in this study were used Prim 7 (GraphPad)

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

The raw data that support the findings of this study are available from the corresponding author upon reasonable request.

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	From many previous experiments, we know that groups of 8 animals are sufficient to obtain statistically reliable results under the conditions of our experiments and typical effect sizes. This was calculated from the average of 13 experiments where the average freezing score during the last CS of a standard 3 CS-US fear conditioning protocol was 62.12% and the average freezing score on a subsequent behavioural task was 48.04%. Assuming a common standard deviation of 10, and performing a 2-sided test with alpha .05 and power of .80, we can obtain sufficient statistical power from 8 animals. In some cases, we present results from n=7 due to the loss of animals from the experiment prior to data collection
Data exclusions	There are data points excluded from the ChIP experiment because the binding efficiency of a few samples was very low. This data is not valid and was excluded. For behavioural experiments, we excluded data points from animals that had misplaced cannulas or inefficient viral spread.
Replication	All technical replicates were successful reproduced.
Randomization	For molecular profilling , naive mice of 9-12 weeks old C56/BL6 male mice were assigned randomly to experimental groups. For viral manipulation animal behavioral experiments in which 9-12 weeks old C56/BL6 cannulated male mice underwent fear conditioning prior to injection of a lentiviral vector, mice were assigned to receive a virus containing either control or active shRNA based on their freezing score from the last CS of a 3 CS-US training protocol. There is natural variation in the level of freezing behaviour displayed, which is increased in animals which have undergone surgery; group assignments were made so that the average freezing score for each treatment group was as similar as possible, to avoid a risk of random variation in learning confounding the effect of the shRNA injection.
Blinding	For behavioral studies, no blinding to group allocation since the groups need to counter balanced based on fear condition score. During behavioral test, data was blinded captured and analyzed by full automatic analysis software. For IHC, FACS and Sequencing, the investigators were blind to group allocation during data collection and analysis.

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

H3K4me3 (active motif Cat# 39915, lot:14013006); YY1 (Abcam Cat# ab38422, lot:GR118077-26) ; TFIIB (Santa Cruz Cat# SC-274X, lot: J0215); Pol II (Santa Cruz Cat# sc-899X, clone: N-20, lot:F1715); m6A (Active motif Cat# 61495, clone:17-3-4-1, lot:06116001); Mouse IgG (Santa Cruz Cat# SC-2025; lot:K1115); N6amt1 (Santa Cruz Cat# SC-83304, lot:H0384); N6amt2 (Santa Cruz Cat# SC-390240, clone: H-3, lot:B0813); NeuN-Alex 488 (Abcam Cat# ab190195, lot:316293-3); Arc-Alex 647 (Bioss Cat# BS-0385R-A647, lot:AG11207885); Beta-actin (Santa Cruz Cat#SC-69879 Clone:AC-15 Lot#K1715); Beta tubulin (Santa Cruz Cat# SC-55529 clone:G8 lot# A1011); Mouse IgG (Active motif Cat# 103533 lot# 31111003); Rabbit IgG (cell signaling Cat# 29293 lot# 7).

Validation

N6amt1 antibody were selected from Santa Cruz. antibody has been validated by using IgG for chIP experiment, and over expression of N6amt1 in HEK293t cells was used to verify the antibody (Suppl.Fig.15 F). Previous validated antibody has been selected for H3K4me3 (Jun et al., Nucl. Acids Res. 2015), YY1 (Song et al., PLOS one, 2009), TFIIB (Pan et al., Cell report, 2014) and RNA Pol II (Chaudhary et al., Nucl. Acids Res. 2016). To re-validate the antibody, we used equivalent amount of control normal rabbit IgG (Santa Cruz), Rabbit IgG (Cell signalling) and mouse IgG (Active motif) was used for non-specificity control. For NeuN-Alex 488 and Arc-Alex 647, we run non labeling control to selected cells that co-express NeuN and Arc.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

N2A cell lines: ATCC® CCL-131™

Authentication

since we only use the cell line to validate the knock down constructs, we have not authenticated the cell line.

Mycoplasma contamination

The cell line were tested for mycoplasma contamination, and no mycoplasma contamination was observed

Commonly misidentified lines
(See ICLAC register)

no commonly misidentified cell lines were used in this study

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

C57/Bl6J male age between 10-14 weeks

Wild animals

no wild animal used in this study

Field-collected samples

no Field-collected samples used in this study

Ethics oversight

All testing took place during the light phase in red-light-illuminated testing rooms following protocols approved by the Institutional Animal Care and Use Committee of the University of California, Irvine and by the Animal Ethics Committee of The University of Queensland. Animal experiments were carried out in accordance with the Australian Code for the Care and Use of Animals for Scientific Purposes (8th edition, revised 2013).

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

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

ftp://ftp-trace.ncbi.nlm.nih.gov/sra/review/SRP110529_20170705_073201_94e0b7097bece407207267f7787e3012

Files in database submission

Raw fastq files have been deposited at the Sequence Read Archive (accession SRP110529) and BioProject (accession PRJNA391201).

Genome browser session (e.g. UCSC)

http://software.broadinstitute.org/software/igv/

Methodology

Replicates

Three biological replicates of each group were used

Sequencing depth

We performed the paired-end reads alignment for N6amt1 chIP-seq, and The total counts in each sample were normalised to 20 million before comparison among samples.

Antibodies

N6amt1 antibody purchased from Santa Cruz. The catalog is SC-83304 and host is rabbit.

Peak calling parameters

The potential binding sites were kept if they met all of the following conditions: i) the sites were not located in the mm10 empirical blacklists, and ii) the normalized counts in all three biological replicates in one group were larger than that in its input sample, and the normalized counts in at least 2 replicates were more than 2-folds larger than its normalized input count.

Data quality

We performed the paired-end reads alignment against mouse reference genome (mm10) using BWA (v0.6.2). Samtools (v0.1.17) was then used to convert “SAM” files to “BAM” files, sort and index the “BAM” files, and remove duplicate reads (i.e. potential PCR duplicates). Reads with low mapping quality (<20) or reads that were not properly paired-end aligned to the mouse reference genome were excluded from the downstream analysis. These steps ensure that only high-quality alignments were used for the downstream peak calling step. MACS2 (version 2.1.1.20160309) was used to call peaks for each sample with the parameter setting “callpeak -t SAMPLE -c INPUT -f BAMPE --keep-dup=all -g mm -p 0.05 -B”. Here, we used the parameter "--keep-dup=all" for peak calling as we had removed the duplicate reads in previous steps. From the MACS output, 66 peaks are at FDR 5% and above 5-fold enrichment. Peak summits identified by MACS2 from all samples were collected to generate a list of potential n6amt1 binding sites. Custom PERL script was then applied to parse the number of fragments (hereafter referred as counts) that cover the peak summit in each sample. Each pair of properly paired-end aligned reads covering the peak summit represents one count. The total counts in each sample were normalized to 20 million before comparison among samples. The potential binding sites were kept if they met all of the following conditions: i) the sites were not located in the mm10 empirical blacklists, and ii) the normalized counts in all three biological replicates in one group were larger than that in its input sample, and the normalized counts in at least 2 replicates were more than 2-folds larger than its normalized input count. These steps ensure that the peaks we identified are consistently enriched in IP samples compared to input samples across different biological replicates.

Software

MACS2 (version 2.1.1.20160309) was used to call peaks for each sample with the parameter setting “callpeak -t SAMPLE -c INPUT -f BAMPE --keep-dup=all -g mm -p 0.05 -B”. Peak summits identified by MACS2 from all samples were collected to generate a list of potential binding sites. Custom PERL script was then applied to parse the number of fragments (hereafter referred as counts) that cover the peak summit in each sample.

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

FACS sorted RNA-seq: Tissue was dissociated with FACS lysis buffer (final concentration: 0.32M Sucrose, 10mM pH8.0, 5mM , 3mM , 0.1mM EDTA, 1mM DTT, 0.3% Triton-X -100 and 100x PIC) into single cell suspension, then fixed with 1% formaldehyde for 5 mins, and stop by 0.125M Glycine. Then, cells were washed twice with cold 1xPBS to remove excessed formaldehyde and glycine. After incubating with blocking buffer (Final concentration: 10% normal goat serum, 5% BSA, 0.1% Triton-X-100 and 1XPIC) for half-hour, cells were double-labeled with Arc antibody (Bioss) in 1:20000 dilution per million cell and NeuN antibody (Abcam) in 1:20000 per million cells, together with DAPI (Thermofisher) in 1:2000. PBPT buffer was used for washing (twice each time) and resuspend into 500ul 1x PBS for FACS sorting. FACS was performed on a BD FACSArial (BD Science)

Instrument

BD FACSAria Cell Sorter and BD FACSAriaII Cell Sorter

Software

BD FACSAria software

Cell population abundance

The sorting is for enriching the activated neurons. 518 positive cells were selected from 50,000 events

Gating strategy

The conjugated antibodies were used in FACS experiment. To set up sorting gate, we used non-neuronal cells to step up the

Gating strategy

gating. Then, the sorted cells were selected by double labeling of Arc-647 and NeuN-488 antibody

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