nature research

Corresponding author(s): Chia-Lin Wei

Last updated by author(s): Dec 20, 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	Base calling: Illumina RTA-1.12 to 1.18 software
Data analysis	ChIA-PET Utilities v0.0.1a-r1 (code available at https://github.com/cheehongsg/CPU) ChiaSigScaled (code available at https://github.com/cheehongsg/ChiaSigScaled) PhenStat_2.22.0 in R version 3.6.1. Trim Galore! v0.4.0 Hisat v2.1.0 HTSeq v0.6.1p1 Gorilla http://cbl-gorilla.cs.technion.ac.il/ MACS v2.1.0.20151222 DESeq2_1.16.1 with R version 3.4.0 Cufflinks v2.2.1

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

All data described in this study are deposited in NCBI's Gene Expression Omnibus GSE120393

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 calculation was performed.
Data exclusions	One replicate of KO si-chr9 F1 clones RNA-seq data was removed as the correlation was low (r<0.9).
Replication	Samples were generated in replicate or triplicate when possible. Only the above data point was excluded (Data exclusion)
Randomization	There was no randomization in this study
Blinding	There was no blinding in this study.

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	Antibodies used are anti-SUZ12 (ab12073, polyclonal, Abcam), anti-EED (ab4469, polyclonal, Abcam), EZH2 (#39875, AC22, Active Motif), H3K27me3 (ab6002, mAbcam 6002, Abcam), RNAPII (MMS126R, clone 8WG16, Covance) and CTCF (ab70303, polyclonal Abcam)
Validation	All antibodies were commercially available. Listed below are information available https://www.abcam.com/suz12-antibody-chip-grade-ab12073.html https://www.abcam.com/eed-antibody-ab4469.html http://www.activemotif.com/catalog/details/39875/ezh2-antibody-mab-clone-ac22 https://www.abcam.com/histone-h3-tri-methyl-k27-antibody-mabcam-6002-chip-grade-ab6002.html https://www.abcam.com/ctcf-antibody-chip-grade-ab70303.html

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	E14 cell line used was a courtesy from Edward Rubin Lab (Lawrence Berkeley National Lab); MEF was primary isolated expanded from mouse from the Genetic Engineering Technologies group of the Jackson Laboratory
Authentication	None of the cell lines used were authenticated.
Mycoplasma contamination	The cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used.

Animals and other organisms

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

Laboratory animals	Mouse C57BL/6NJ was used.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve samples collected from the field.
Ethics oversight	All procedures and protocols were approved by the Jackson Laboratory Animal Care and Use Committee and were conducted in compliance with the National Institute of Health Guideline for Care and Use of Laboratory Animals.

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.*	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120393
Files in database submission	ChIP-seq data for: RNAPII, H3K27me3, and CTCF. BAMs: RNAPolII_ChIP_rep1.unique.collapse.sorted.bam RNAPolII_ChIP_rep2.unique.collapse.sorted.bam RNAPolII_ChIP_rep3.unique.collapse.sorted.bam RNAPolII_ChIP_rep4.unique.collapse.sorted.bam RNAPolII_ChIP_rep1_input.unique.collapse.sorted.bam RNAPolII_ChIP_rep2_input.unique.collapse.sorted.bam RNAPolII_ChIP_rep3_input.unique.collapse.sorted.bam RNAPolII_ChIP_rep4_input.unique.collapse.sorted.bam H3K27me3_ChIP_rep1.unique.collapse.sorted.bam H3K27me3_ChIP_rep2.unique.collapse.sorted.bam H3K27me3_ChIP_rep1_input.unique.collapse.sorted.bam H3K27me3_ChIP_rep2_input.unique.collapse.sorted.bam CTCF_ChIP_rep1_input.unique.collapse.sorted.bam CTCF_ChIP_rep1.unique.collapse.sorted.bam CTCF_ChIP_rep2_input.unique.collapse.sorted.bam CTCF_ChIP_rep2.unique.collapse.sorted.bam Peaks: RNAPolII_ChIP.narrowPeak.gz H3K27me3_ChIP.broadPeak.gz CTCF_ChIP.narrowPeak.gz Signals: RNAPolII_ChIP_foldEnrich.bigWig RNAPolII_ChIP_treat.bigWig RNAPolII_ChIP_input.bigWig CTCF_ChIP_foldEnrich.bigWig CTCF_ChIP_treat.bigWig CTCF_ChIP_input.bigWig H3K27me3_ChIP_foldEnrich.bigWig H3K27me3_ChIP_treat.bigWig H3K27me3_ChIP_input.bigWig
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	Replicates were generated for all samples.
Sequencing depth	All reads are single-end 51 bp. Factor Total_reads Uniquely_mapped_PCRdedup RNA_PolII 43,381,959 32,741,261

H3K27me3 32,077,698 22,169,509
CTCF 66,418,692 45,777,229

Antibodies

Antibodies used are H3K27me3 (ab6002, Abcam), RNAPII (MMS126R, clone 8WG16, Covance) and CTCF (ab70303, Abcam)

Peak calling parameters

All factors are processed in the same manner, we provide an example for RNAPII as follows.
Mapping:
bwa aln -l 25 -f RNAPolII_ChIP_rep1.sai mm10.fa RNAPolII_ChIP_rep1.fastq.gz
bwa samse -f RNAPolII_ChIP_rep1.sam mm10.fa RNAPolII_ChIP_rep1.fastq.gz
samtools sort -O bam -o RNAPolII_ChIP_rep1.bam RNAPolII_ChIP_rep1.sam

Peak calling:
Peaks were called after merging replicates (RNAPolII_ChIP.rep*.unique.collapse.sorted.bam files), then the peak calling command was as follow.
macs2 callpeak --nomodel --extsize 250 -B --SPMR - RNAPolII_ChIP_treat.bam -c RNAPolII_ChIP_input.bam -f BAM -g mm

Data quality

All reported peaks have FDR 5%. The number of peaks:
RNAPolII: 20591 narrow peaks
H3K27me3: 4347 broad peaks
CTCF: 36992 narrow peaks

Software

bwa version 0.7.12
MACS version 2.1.1.20160309
samtools version 1.2 (merging and sorting BAM files).