nature research

Corresponding author(s): Francois Spitz, Leonid Mirny

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   Sample size was not pre-determined, We used sample sizes commonly used and accepted for the type of experiments. For animal experiments, we used at least 3 animals per group (range 3 to 17) to allow basic statistical inference while using a justifiable number of mutant mice. We used 4 biological replicates per conditions for RNA-Seq, 2 for Hi-C, and 2 for ChiP-Seq. Hi-C data was subsequently pooled. For ChIP-Seq, representative tracks from one of the replicates are shown.

2. Data exclusions

   Describe any data exclusions.

   No data were excluded from the analysis, except for Hi-C. As recommended for the analysis of Hi-C data, we removed bins with low-coverage as well as contact signals between same of adjacent genomic bins.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   all replication attempts were successful

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   allocation of mice to one or the other group was not randomized. However, samples were treated in parallel, and whenever possible with internal controls, so as to equally match potential confounding effects

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   Experiments were not blinded

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Statistical analyses were performed with R (R studio) and Prism 6. Genomic analyses used Galaxy and ENCODE tools, Hiclib, STAR V2.5.0a, HTSeq, DESeq2, HISAT2. Details are included in the Method section. Custom codes used are described in Method section and Code availability section.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	No unique material have been used. Mouse strains - as cryopreserved sperm - are available for distribution by the authors, upon request.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	αH3k4me3: C15410003-50, Diagenode (valided by the manufacturer) αH3K27Ac: ab4729, Abcam; lot GR7675-1 (valided by the manufacturer) Anti-Rad21: Abcam ab992; lot GR214359-7 (valided by the manufacturer) Anti-CTCF: Millipore 07-729; lot 2887267 (valided by the manufacturer) anti-SMC3: Abcam ab9263, lot GR290533-3 (valided by the manufacturer) anti SMC1: provided by Ana Losada (described/validated in 10.1038/emboj.2012.11) anti SA1: provided by Ana Losada (described/validated in 10.1038/emboj.2012.11)

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	HEK293 cells (originally from the Trono lab - EPFL - Lausanne - Switzerland)
b. Describe the method of cell line authentication used.	HEK293 cells were not authenticated
c. Report whether the cell lines were tested for mycoplasma contamination.	HEK293 cells were not tested for mycoplasma
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	cell line used (HEK-293) is not in the ICLAC database

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	we used animals from the Nipbl [flox] allele made in the lab (internal reference clone 3H6), crossed with Ttr-cre/Esr1 strain (MGI:3046546 Tg(Ttr-cre/Esr1*)1Vco). Mice used were adult (10-20 week-old, usually 12 week-old) of both sexes, except for Hi-C experiments for which only males were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	the study did not involve human participants

nature research

Corresponding author(s): Francois Spitz, Leonid Mirny

☐ Initial submission ☐ Revised version ☒ Final submission

ChIP-seq Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data deposition

1. For all ChIP-seq data:

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

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

2. Provide all necessary reviewer access links.
The entry may remain private before publication.

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE93431

WT vs TAM vs ΔNipbl, locked by zoom and location
• http://higlass.io/app/?config=fLYyoNpETSe-QbXg50_DOA
Same, with overlaid eigenvector tracks:
• http://higlass.io/app/?config=RNj34_97T3SWJ5RcSDonjQ
WT vs ΔNipbl with zoom-ins connected by view projections (grey squares)
• http://higlass.io/app/?config=Tf2-ublRTey9hiBKMlgzwg

UCSC tracks (ChIP-Seq and RNA-Seq)
http://genome-euro.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=spitzfr&hgS_otherUserSessionName=Ko_Nipbl_Public0

3. Provide a list of all files available in the database submission.

4. If available, provide a link to an anonymized genome browser session (e.g. UCSC).

• http://higlass.io/app/?config=fLYyoNpETSe-QbXg50_DOA
• http://higlass.io/app/?config=RNj34_97T3SWJ5RcSDonjQ
• http://higlass.io/app/?config=Tf2-ublRTey9hiBKMlgzwg

UCSC tracks (ChIP-Seq and RNA-Seq)
http://genome-euro.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=spitzfr&hgS_otherUserSessionName=Ko_Nipbl_Public0

Methodological details

5. Describe the experimental replicates.

See Method section.

6. Describe the sequencing depth for each experiment.

Condition Total Mapped Unique
Nipbl_8094_ChIPK27ac 28888231 28138779 22336333
Nipbl_8094_ChIPK4me3 19322681 18765579 13559526
Nipbl_8094_Input 20708221 20033518 14857648
WT_8097_ChIPK27ac 23893468 23425426 18952812
WT_8097_ChIPK4me3 21546139 21033793 15429627
WT_8097_Input 22718479 22040698 17235134
Rad21Ctrl 55861956 52702007 39862949
Rad21Delta 50724880 48735141 36299305
Smc3Ctrl 53349512 50898271 38132996
Smc3Delta 59299518 56890933 43428069
InputCtrl 55442998 53763091 36651837

InputDelta 64217332 60983192 43185393
CtcfCtrl 33078713 32168242 22805555
CtcfDelta 43884331 42740092 32374478

7. Describe the antibodies used for the ChIP-seq experiments.

αH3k4me3: C15410003-50, Diagenode
αH3K27Ac: ab4729, Abcam; lot GR7675-1
Anti-Rad21: Abcam ab992; lot GR214359-7
Anti-CTCF: Millipore 07-729; lot 2887267
anti-SMC3: Abcam ab9263, lot GR290533-3

all those antibodies have been used for ChIP-Seq in multiple publications

8. Describe the peak calling parameters.

We processed ChIP-seq data following the steps of the ENCODE pipeline [https://github.com/ENCODE-DCC/chip-seq-pipeline]. Alignment: we used bwa 0.7.12 (program bwa aln with parameters: -q 5 -l 32 -k 2). Filtering: PCR duplicates were marked using picardtools 2.7.1. Unmapped reads, non-primary alignments, and low quality alignments (mapQ < 30) were removed using samtools 1.3. Cross-correlation analysis was performed using phantompeakqualtools. Peaks and signal tracks were generated using MACS2. Peak calls were generated using MACS2 with parameters --p 1e-2, --nomodel, --shift 0, --keep-dup all. For Rad21, Smc3 and CTCF ChIP-seq, we followed the same steps with the following alterations: reads from pooled mouse hepatocyte chromatin and HEK human chromatin (internal control and calibration) were mapped to the combined mm9 and hg19 assemblies using the bwa mem program with default parameters. After filtering, reads were divided into those that mapped uniquely to either mm9 or hg19.

9. Describe the methods used to ensure data quality.

Raw read files were assessed using fastqc prior to processing. Cross-correlation and phantom peak analysis was used to ensure data quality and estimate fragment lengths. All datasets used for peak calling received a quality tag of 1 (High) or 2 (veryHigh) from cross-correlation analysis using phantompeakqualtools. Peak statistics are provided in the following table.
name treatment control total_peaks FDR_5_percent gt_5_fold_fc
CtcfCtrl CtcfCtrl InputCtrl 71347 53920 32488
CtcfDelta CtcfDelta InputDelta 43277 23347 9805
Rad21Ctrl Rad21Ctrl InputCtrl 44306 37094 20782
Rad21Delta Rad21Delta InputDelta 20730 7216 2577
Smc3Ctrl Smc3Ctrl InputCtrl 44306 37094 20782
Smc3Delta Smc3Delta InputDelta 31368 9683 2457

10. Describe the software used to collect and analyze the ChIP-seq data.

We followed and used the same software tools used in the ENCODE ChIP-seq pipeline (https://github.com/ENCODE-DCC/chip-seq-pipeline): bwa, picard, phantompeakqualtools, MACS2, bedtools, kent UCSC binaries.