nature research

Corresponding author(s): Frederick W. Alt

Last updated by author(s): Aug 4, 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	Next generation sequencing data were collected via Illumina sequencing platforms (MiSeq, NextSeq 550 or HiSeq 2500). Data generated from Hi-Seq 2500 were demultiplexed via bcl2fastq-1.8.3. Data generated from MiSeq and NextSeq 550 were demultiplexed via TranslocPreprocess.pl, a published pipeline available at (http://robinmeyers.github.io/transloc_pipeline/).
Data analysis	HTGTS V(D)J-seq and 3C-HTGTS data were processed through a published pipeline available at (http://robinmeyers.github.io/transloc_pipeline/). Code for Hi-C data process is available at (github.com/aidenlab). GRO-Seq and ChIP-Seq were aligned to mm9 genome with bowtie2 v2.2.8, processed by samtools v1.8 and generated graph files via RSeQC tool v2.6. Statistical analysis in all assays were performed via GraphPad Prism6.

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

HTGTS V(D)J-seq, Hi-C, 3C-HTGTS, GRO-seq and ChIP–seq sequencing data reported in this study have been deposited in the GEO database under the accession number GSE130224. Specifically, HTGTS V(D)J-seq data are deposited in the GEO database under the accession number GSE130216 and are related to Figs. 1e–h, 2b-d, 3a–c, e, 4c, Extended Data Figs. 2d, e, g, h, 3e, f, 4a, c, d, 5a–c, 6c, 7d, e, 10b, and Supplementary Tables 1, 2. Hi-C data are deposited in the GEO database under accession number GSE134543 and are related to Extended Data Fig. 8a. 3C-HTGTS data are deposited in the GEO database under accession number GSE130214 and are related to Figs. 3f, 4d and Extended Data Figs. 8b, 9q, r,10c. GRO-seq data are deposited in the GEO database under accession number

GSE130215 and are related to Figs. 3d, 4b and Extended Data Figs. 4e, 6d, 7f, 10a. ChIP–seq data are deposited in the GEO database under the accession number GSE130213 and are related to Extended Data Figs. 8c, d, 10d.

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 statistical methods were used to predetermine sample size for all experiments. Sample sizes were chosen based on previous studies in this field that used similar sample sizes to generate reproducible results.
Data exclusions	No experiments were excluded from analysis.
Replication	Samples were analyzed with biological repeats. All attempts for replication were successful.
Randomization	Experiments were not randomized. Each experiment was performed with identified controls and mutant strains. Randomization was not relevant to the study as the study does not involve participant groups.
Blinding	Investigators were not blinded to allocation during experiments and outcome assessment. Blinding was not possible as investigators need to verify the control and matched mutant strains before each experiment. Also, based on previous studies in this field, these assays do not require blinding.

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	Antibody (clone), Host, Supplier, Cat. number, Technique, Concentration/Dilution Rad21, Rabbit, Abcam, ab992, ChIP-Seq, 5 ug/ChIP Nipbl, Rabbit, Bethyl, A301-779A, ChIP-Seq, 5 ug/ChIP Cas9, Rabbit, Diagenode, C15310258-100, WB, 1:5,000 β-Actin (8H10D10), Mouse, Cell Signaling Technology, 3700S, WB, 1:2,000
Validation	Rad21 antibody (Abcam, ab992), manufacturer's validation: https://www.abcam.com/rad21-antibody-chip-grade-ab992.html This antibody was validated for ChIP-Seq by published papers including: Vian, L. et al. The Energetics and Physiological Impact of Cohesin Extrusion. Cell 173, 1165-1178 e1120, doi:10.1016/j.cell.2018.03.072 (2018). Nipbl antibody (Bethyl, A301-779A), manufacturer's validation: https://www.bethyl.com/product/A301-779A/NIPBL+Antibody This antibody was validated for ChIP-Seq by published papers including: Vian, L. et al. The Energetics and Physiological Impact of Cohesin Extrusion. Cell 173, 1165-1178 e1120, doi:10.1016/j.cell.2018.03.072 (2018). Cas9 antibody (Diagenode, C15310258), manufacturer's validation: https://www.diagenode.com/en/p/crispr-cas9-polyclonal-antibody This antibody was validated for ChIP in the following paper:

Howe, FS. et al. CRISPRi is not strand-specific at all loci and redefines the transcriptional landscape. Elife. Oct 23;6. pii: e29878. doi: 10.7554/eLife.29878 (2017).

β-Actin antibody (Cell Signaling Technology, 3700S), manufacturer's validation: https://www.cellsignal.com/products/primary-antibodies/b-actin-8h10d10-mouse-mab/3700
This antibody was validated for western blotting in publications including:
Lahrouchi, N. et al., Homozygous frameshift mutations in FAT1 cause a syndrome characterized by colobomatous-microphthalmia, ptosis, nephropathy and syndactyly. Nat Commun. Mar 12;10(1):1180. doi: 10.1038/s41467-019-08547-w. (2019)

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	The DH-JH+/- line was derived from a previously described RAG2-/- Eμ-bcl2+ v-Abl pro-B cell line from Dr. Barry P. Sleckman's lab. Gapud, E.J., Lee, B.S., Mahowald, G.K., Bassing, C.H. & Sleckman, B.P. Repair of chromosomal RAG-mediated DNA breaks by mutant RAG proteins lacking phosphatidylinositol 3-like kinase consensus phosphorylation sites. J Immunol 187, 1826-1834 (2011). The DH-JH+/- line was used as the parental line to generate mutant derivatives including DFL16.1-RSS-DN-inv, DFL16.1-RSS-UP-inv, DFL16.1-inv, JHΔ and DH-JH1+/- lines. The DH-JH1+/- lines were used to generate mutant derivatives including DQ52-inv, DQ52Δ DFL16.1-to-DQ52, DQ52Δ DFL16.1-to-DQ52-inv, DFL16.1Δ DQ52-to-DFL16.1, DFL16.1Δ DQ52-to-DFL16.1-inv and intervening DHs inversion lines. The JHΔ lines were used to generate mutant derivatives including JHΔ-dCas9 and JHΔ-dCas9-sgRNA-Sγ1 lines. The JHΔ-dCas9 lines were used to generate JHΔ-dCas9-Iγ2b-/- lines. The DFL16.1-JH4 line was a published line from this lab. Jain, S., Ba, Z., Zhang, Y., Dai, H. & Alt, F.W. CTCF-Binding Elements Mediate Accessibility of RAG Substrates During Chromatin Scanning. Cell 174, 1-15 (2018). The DFL16.1-JH4 line was used as the parental line to generate the DFL16.1-JH4-inv lines, which was used to further generate the DFL16.1-JH4-inv 3'CBE-/- and DFL16.1-JH4-inv RAG2-/- lines.
Authentication	The DH-JH+/-, JHΔ, DH-JH1+/-, intervening DHs inversion, DFL16.1-JH4-inv and DFL16.1-JH4-inv 3'CBE-/- lines were verified via Southern blotting. The DFL16.1-RSS-DN-inv, DFL16.1-RSS-UP-inv, DFL16.1-inv, DQ52-inv, DQ52Δ DFL16.1-to-DQ52, DQ52Δ DFL16.1-to-DQ52-inv, DFL16.1Δ DQ52-to-DFL16.1, DFL16.1Δ DQ52-to-DFL16.1-inv, JHΔ-dCas9-Iγ2b-/- and DFL16.1-JH4-inv RAG2-/- lines were verified via PCR genotyping and Sanger sequencing. The JHΔ-dCas9 lines were verified via Western blotting. The JHΔ-dCas9-sgRNA-Sγ1 lines were verified via RT-PCR.
Mycoplasma contamination	The cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

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.*	To review GEO accession GSE130224: Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE130224 Enter token erozikgorhctbil into the box
Files in database submission	Rad21_dCas9Control_Input.bw Rad21_dCas9Block_Input.bw Rad21_dCas9Control_rep1.bw Rad21_dCas9Control_rep2.bw Rad21_dCas9Block_rep1.bw Rad21_dCas9Block_rep2.bw Rad21_Ig2bwt_rep1.bw Rad21_Ig2bwt_rep2.bw Rad21_Ig2bwt_rep3.bw Rad21_Ig2bdel_rep1.bw Rad21_Ig2bdel_rep2.bw Rad21_Ig2bdel_rep3.bw Nipbl_dCas9Control_Input.bw Nipbl_dCas9Block_Input.bw Nipbl_dCas9Control_rep1.bw Nipbl_dCas9Control_rep2.bw Nipbl_dCas9Block_rep1.bw Nipbl_dCas9Block_rep2.bw Rad21_dCas9Control_rep1.broadpeak.bed

Rad21_dCas9Control_rep2.broadpeak.bed
Rad21_dCas9Block_rep1.broadpeak.bed
Rad21_dCas9Block_rep2.broadpeak.bed
Rad21_Ig2bwt_rep1.broadpeak.bed
Rad21_Ig2bwt_rep2.broadpeak.bed
Rad21_Ig2bwt_rep3.broadpeak.bed
Rad21_Ig2bdel_rep1.broadpeak.bed
Rad21_Ig2bdel_rep2.broadpeak.bed
Rad21_Ig2bdel_rep3.broadpeak.bed
Nipbl_dCas9Control_rep1.broadpeak.bed
Nipbl_dCas9Control_rep2.broadpeak.bed
Nipbl_dCas9Block_rep1.broadpeak.bed
Nipbl_dCas9Block_rep2.broadpeak.bed
Rad21_dCas9Control_Input_R1.fastq.gz
Rad21_dCas9Block_Input_R1.fastq.gz
Rad21_dCas9Control_rep1_R1.fastq.gz
Rad21_dCas9Control_rep2_R1.fastq.gz
Rad21_dCas9Block_rep1_R1.fastq.gz
Rad21_dCas9Block_rep2_R1.fastq.gz
Rad21_Ig2bwt_rep1_R1.fastq.gz
Rad21_Ig2bwt_rep2_R1.fastq.gz
Rad21_Ig2bwt_rep3_R1.fastq.gz
Rad21_Ig2bdel_rep1_R1.fastq.gz
Rad21_Ig2bdel_rep2_R1.fastq.gz
Rad21_Ig2bdel_rep3_R1.fastq.gz
Nipbl_dCas9Control_Input_R1.fastq.gz
Nipbl_dCas9Block_Input_R1.fastq.gz
Nipbl_dCas9Control_rep1_R1.fastq.gz
Nipbl_dCas9Control_rep2_R1.fastq.gz
Nipbl_dCas9Block_rep1_R1.fastq.gz
Nipbl_dCas9Block_rep2_R1.fastq.gz
Rad21_dCas9Control_Input_R2.fastq.gz
Rad21_dCas9Block_Input_R2.fastq.gz
Rad21_dCas9Control_rep1_R2.fastq.gz
Rad21_dCas9Control_rep2_R2.fastq.gz
Rad21_dCas9Block_rep1_R2.fastq.gz
Rad21_dCas9Block_rep2_R2.fastq.gz
Rad21_Ig2bwt_rep1_R2.fastq.gz
Rad21_Ig2bwt_rep2_R2.fastq.gz
Rad21_Ig2bwt_rep3_R2.fastq.gz
Rad21_Ig2bdel_rep1_R2.fastq.gz
Rad21_Ig2bdel_rep2_R2.fastq.gz
Rad21_Ig2bdel_rep3_R2.fastq.gz
Nipbl_dCas9Control_Input_R2.fastq.gz
Nipbl_dCas9Block_Input_R2.fastq.gz
Nipbl_dCas9Control_rep1_R2.fastq.gz
Nipbl_dCas9Control_rep2_R2.fastq.gz
Nipbl_dCas9Block_rep1_R2.fastq.gz
Nipbl_dCas9Block_rep2_R2.fastq.gz

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

Two or three independent libraries were performed from at least two independent clones for each genotype analyzed.

Sequencing depth

Read depth is > 50 million/sample

Antibodies

Rad21 antibody (Abcam, ab992); Nipbl antibody (Bethyl, A301-779A)

Peak calling parameters

Peak calling was performed using MACS2 with the following commands:
macs2 callpeak -t IP.bam -c Input.bam -n Result_directory --nomodel --extsize 51 --nolambda -B --SPMR -g mm --verbose 0 --broad
or
macs2 callpeak -t IP.bam -n Result_directory --nomodel --extsize 51 --nolambda -B --SPMR -g mm --verbose 0 --broad

Data quality

We relied on MACS2 algorithm to detect significant peaks, obtaining 15-30 thousand peaks at FRD5% and above 5 fold enrichment for Rad21 ChIP-Seq libraries and obtaining 4-12 thousand peaks at FRD5% and above 5 fold enrichment for Nipbl ChIP-Seq libraries.

Software

ChIP-seq data was aligned to mm9 by bowtie 2.2.8., processed by samtools v1.8 and generated graph files via RSeQC tool v2.6. MACS 2.1.2 was used to call peaks.