nature portfolio

Corresponding author(s): Cornelis Murre

Last updated by author(s): Oct 8, 2023

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies 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 (NGS) data was collected using NovaSeq 6000, Flowcytometry data was collected using FACSDiva V=v9.0 (BD), Western blot images were collected using BioRad ChemiDoc Imaging System, Fluorescence microscopy images of nuclear volume measurements were collected using Zeiss LSM 980 confocal microscope, Wright-Giemsa staining images were collected using Keyence BZX microscope.
Data analysis	1.RNA-Seq analysis: Quality control of the raw fastq files was performed using FastQC v0.11.9. Reads were aligned to the mouse genome (GRCm39.primary_assembly.genome.fa) using the STAR v2.7.9a aligner and reads were quantified using RSEM v1.3.3. Differential expression analyses were performed using edgeR (R Bioconductor package). Functional enrichment analysis was performed using GProfiler. 2.ATAC-Seq analysis: Quality control of raw reads was performed using FastQC v0.11.8, MultiQC v1.10.1 and ataqv v1.2.1. Reads were aligned to mouse genome using BWA v0.7.17. Peaks were called using MACS2 v2.2.7. deep Tools were used to analyze enrichment signals and PyRanges were used for quantification of signal strength. Motif enrichment analysis was performed with Homer v4.11. 3. scRNA-Seq analysis: Raw reads were processed using CellRanger version 6.0.0 10X Genomics. Filtering, normalization, scaling, and clustering of data was performed using Seurat version 4.0.6. 4. Hi-C Analysis: We used the runHiC pipeline (https://github.com/XiaoTaoWang/HiC_pipeline) to process Hi-C data. Specifically, we first used Chromap (PMID: 34772935) to map the fastq file from Hi-C experiment to the reference genome mm10. We then kept the uniquely mapped reads, removed PCR duplicates, and output the valid .pairs files defined by the 4DN consortium (https://github.com/4dn-dcic/pairix/blob/master/pairs_format_specification.md). In addition, we used Juicer (PMID: 27467249) to convert the processed Hi-C data into the .hic format, for visualization in the software Juicebox (https://www.aidenlab.org/juicebox/). We first performed such an analysis for each biological replicate separately, and applied HiCrep to evaluate the reproducibility of Hi-C contact frequency between two biological replicates. After confirming the high reproducibility, we pooled the two biological replicates from the same experimental condition and used the pooled Hi-C data for all the downstream analysis.

5.ChIP-Seq Analysis: FASTQC v0.12.1 is used for checking the sequencing quality. ChIP-seq reads were aligned to the reference genome mm10 using Bowtie2 v2.5.1. Samtools v1.17 was used to remove PCR duplicates. Only the uniquely mapped reads were included for further processing. MACS2 v2.2.7.1 was used to call RAD21 ChIP-Seq peaks in progenitors and NIPBL-depleted progenitors. To ensure high data quality, we call Rad21 peaks with FDR<1e-25 and fold enrichment>5. Based on the binary peak status in progenitors and NIPBL depleted progenitors, we defined three groups of RAD21 peaks: shared RAD21 peaks, progenitor-specific RAD21 peaks and NIPBL depleted progenitor-specific RAD21 peaks. Last, we selected promoters and enhancers within progenitor-specific RAD21 peaks and NIPBL depleted progenitor-specific RAD21 peaks, and applied HOMER (PMID: 20513432) with the default settings to perform the motif enrichment analysis.
6. Flow cytometry data was analyzed using FlowJo v10.8.1 (BD).

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio 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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

RNA-Seq, ChIP-seq, ATAC–seq, scRNA-seq and Hi-C datasets generated in the current study have been deposited in the Gene Expression Omnibus (GEO) GSE211817, GSE211818, GSE211819, GSE211820, GSE232064 and GSE235645.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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

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	Sample size was determined based on similar studies in this field.
Data exclusions	No data were excluded other than the part of QC (NGS) that have been detailed in Methods.
Replication	All the data were reproduced. The vast majority of the experiments and analysis presented were the result of two or three independent biological replicates. For non-quantitative data (Western blots and so on), the findings were validated and representative results are shown.
Randomization	This is not relevant to this study. The samples form defined groups.
Blinding	No blinding was performed. The groups were defined and studied using standard protocols.

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 and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

1. APC anti-mouse Ly-6G Antibody, BioLegend, Cat # 127613, Clone 1A8,
2. APC Ly-6G/Ly-6C Monoclonal Antibody, Invitrogen, Cat # 17-5931-81, Clone RB6-8C5,
3. FITC anti-mouse CD11c Antibody, BioLegend, Cat # 117305, Clone N418
4. FITC anti-mouse CD8a Antibody, Cat # 100705, BioLegend, Clone 53-6.7
5. FITC anti-mouse CD3 Antibody, BioLegend, Cat # 100203, Clone 17A2,
6. FITC anti-mouse CD4 Antibody, BioLegend, Cat # 100405, Clone GK1.5
7. FITC anti-mouse NK-1.1 Antibody, BioLegend, Cat # 108705, Clone PK136
8. FITC CD19 Monoclonal Antibody, Invitrogen, Cat # 11-0191-82, Clone MB19-1
9. FITC anti-mouse TER-119/Erythroid Cells Antibody, BioLegend, Cat # 116205, Clone TER-119
10. FITC CD45R (B220) Monoclonal Antibody, Invitrogen, Cat # 11-0452-82, Clone RA3-6B2
11. APC-Cy7 Rat Anti-CD11b, BD Biosciences, Cat # 557657, Clone M1/70
12. Brilliant Violet 421 anti-mouse/human CD11b Antibody, BioLegend, Cat # 101235, Clone M1/70
13. Biotin anti-mouse TER-119/Erythroid Cells Antibody, BioLegend, Cat # 116204, Clone TER-119
14. Biotin anti-mouse CD19, BioLegend, Cat # 115504, Clone 6D5,
15. Biotin anti-mouse Ly-6G, BioLegend, Cat # 127604, Clone 1A8,
16. Biotin anti-CD11c Armenian Hamster Monoclonal Antibody, BioLegend, Cat # 117304, Clone N418
17. Biotin anti-CD45R Rat Monoclonal Antibody, BioLegend, Cat # 103204, Clone RA3-6B2,
18. Biotin Anti-CD3 Rat Monoclonal Antibody, BioLegend, Cat # 100244, Clone 17A2,
19. Biotin anti-CD115 Rat Monoclonal Antibody, BioLegend, Cat # 135507, Clone AFS98
20. Biotin anti-mouse CD170 (Siglec-F) Antibody, BioLegend, Cat # 155512, Clone S17007L,
21. Biotin anti-mouse F4/80 Antibody, BioLegend, Cat # 123105, Clone BM8
22.Anti-Rad21 antibody, Abcam, Cat # ab992
23.NIPBL Antibody, Santa Cruz, Cat # sc-374625, Clone C-9
24. β-Actin Antibody, Cell Signaling Technology, Cat # 4970, Clone 13E5
25. Ki-67 Monoclonal Antibody, eBioscience™, Cat # 17-5698-80, Clone SolA15
26. SMC3 Antibody, Abcam, Cat # ab9263
27. SMC1 Antibody, Bethyl Laboratories, Cat # A300-055A

Validation

1.https://www.biolegend.com/en-ie/products/apc-anti-mouse-ly-6g-antibody-6115
2. https://www.thermofisher.com/antibody/product/Ly-6G-Ly-6C-Antibody-clone-RB6-8C5-Monoclonal/17-5931-82.
3. https://www.biolegend.com/en-ie/products/fitc-anti-mouse-cd11c-antibody-1815
4. https://www.biolegend.com/en-ie/products/fitc-anti-mouse-cd8a-antibody-153
5. https://www.biolegend.com/en-ie/products/fitc-anti-mouse-cd3-antibody-45
6. https://www.biolegend.com/en-ie/products/fitc-anti-mouse-cd4-antibody-248
7. https://www.biolegend.com/en-ie/products/fitc-anti-mouse-nk-1-1-antibody-429
8. https://www.thermofisher.com/antibody/product/CD19-Antibody-clone-MB19-1-Monoclonal/11-0191-82
9. https://www.biolegend.com/en-ie/products/fitc-anti-mouse-ter-119-erythroid-cells-antibody-1865
10. https://www.thermofisher.com/antibody/product/CD45R-B220-Antibody-clone-RA3-6B2-Monoclonal/11-0452-82
11. https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/apc-cy-7-rat-anti-cd11b.557657
12. https://www.biolegend.com/en-ie/products/brilliant-violet-421-anti-mouse-human-cd11b-antibody-7163
13. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-ter-119-erythroid-cells-antibody-1864
14. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-cd19-antibody-1527
15. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-ly-6g-antibody-4772
16. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-cd11c-antibody-1814
17. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-human-cd45r-b220-antibody-444
18. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-cd3-antibody-10023
19. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-cd115-csf-1r-antibody-6335
20. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-cd170-siglec-f-antibody-18059
21. https://www.biolegend.com/en-ie/products/biotin-anti-mouse-f4-80-antibody-4066
22. https://www.abcam.com/products/primary-antibodies/rad21-antibody-ab992.html
23. https://www.scbt.com/p/nipbl-antibody-c-9
24. https://www.cellsignal.com/products/primary-antibodies/b-actin-13e5-rabbit-mab/4970
25. https://www.thermofisher.com/antibody/product/Ki-67-Antibody-clone-SolA15-Monoclonal/17-5698-82
26. https://www.abcam.com/products/primary-antibodies/smc3-antibody-ab9263.html
27. https://www.fortislife.com/products/primary-antibodies/rabbit-anti-smc1-antibody/BETHYL-A300-055

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	1.ECOMG myeloid cell line was established by conditional immortalization of murine primary bone marrow progenitors with an estrogen-regulated E2a/Pbx1-estrogen receptor fusion protein (Sykes et al., 2003). 2. ECOMG progenitors carrying NIPBL-FKBP12F36V-EYFP or MAU2-FKBP12F36V-mScarlet or both NIPBL-FKBP12F36V-EYFP and MAU2-FKBP12F36V-mScarlet were generated (Zhu et al., 2021).
Authentication	Authentication of ECOMG progenitors carrying NIPBL-FKBP12F36V-EYFP or MAU2-FKBP12F36V-mScarlet or both NIPBL-FKBP12F36V-EYFP and MAU2-FKBP12F36V-mScarlet were performed by screening for the presence of inserted degrons at endogenous NIPBL and MAU2 loci by PCR and flowcytometry.
Mycoplasma contamination	Cell lines tested negative for mycoplasma when obtained from the Kamps laboratory.
Commonly misidentified lines (See ICLAC register)	Not applicable.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	6-10 weeks old wild type C57BL/6 mice were used in this study.
Wild animals	This study did not involve wild animals
Reporting on sex	Male mice were used for majority of the experiments.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All mice were maintained in breeding facility in accordance with protocols approved by Institutional Animal Care and Use Committee of the University of California San Diego (La Jolla, CA).

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.*	GSE235645 https://urldefense.com/v3/__https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE235645__;!!Mih3wA!GomXer6QzAyVIGlh8zWYiG2N9HhMe3q2W7h6PkH08LVtS4iX4d2kgHZvnuaSFfY2_gzs_ATJa1jYVremVA$
Files in database submission	GSM7506769 NIPBL Untreated +estradiol 72hrs_RAD21_ChIPSeq approved BW NARROWPEAK GSM7506770 NIPBL Untreated +estradiol 72hrs_Input approved BW GSM7506771 NIPBL dTAG treated +estradiol 72hrs_RAD21_ChIPSeq approved BW NARROWPEAK GSM7506772 NIPBL dTAG treated +estradiol 72hrs_Input approved BW
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	1.NIPBL Untreated +estradiol 72hrs_RAD21_ChIPSeq NIPBL Untreated +estradiol 72hrs_Input 2.NIPBL dTAG treated +estradiol 72hrs_RAD21_ChIPSeq NIPBL dTAG treated +estradiol 72hrs_Input
Sequencing depth	1.NIPBL Untreated +estradiol 72hrs_RAD21_ChIPSeq (Progenitors), Raw reads-78,356,385, % of unique reads - 62.03%, single end sequencing. 2.NIPBL dTAG treated +estradiol 72hrs_RAD21_ChIPSeq (NIPBL depleted Progenitors), Raw reads - 67,945,293, % of unique reads -

	62.78%, Single end sequencing.
Antibodies	Anti-Rad21 antibody, Abcam, Cat # ab992 https://www.abcam.com/products/primary-antibodies/rad21-antibody-ab992.html
Peak calling parameters	Peak calling parameters: A publicly available ChIP-seq data analysis pipeline (https://hbctraining.github.io/Intro-to-ChIPseq-flipped/) was used to process the Rad21 ChIP-seq data. Briefly, Bowtie2 v2.5.1 was used to map the fastq files to the reference genome mm10 with the following parameters: bowtie2 -p 10 -q (query input files are FASTQ .fq/.fastq (default)) --local -x mm10(ReferenceGenome) -U Inputfastq -S SAMfile. Samtools v1.17 was used to remove PCR duplicates. MACS2 v2.2.7.1 was used to call Rad21 peaks with the following parameters: macs2 callpeak -t TAGbam -c Inputbam -f BAM -n NO.TAG --outdir outputdirectory -B.
Data quality	FASTQC v0.12.1 is run to check the sequencing quality. To ensure high data quality, we call Rad21 peaks with FDR<1e-25 and fold enrichment>5, results in 17,066 peaks in progenitors and 8,304 peaks in NIPBL depleted progenitors.
Software	FASTQC v0.12.1, Bowtie v2.5.1, Samtools v1.17, MACS v2.2.7.1.

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	Single cell suspensions of ECOMG cells and mouse primary bone marrow cells were prepared. Cells were counted and adjusted the concentration of cells to 1-5 x 106 cells/ml using ice cold FACS buffer (1X PBS with 2% FCS, 1 mM EDTA and 0.1% sodium azide). Approximately, 100 µl of cell suspension was used for staining with optimal dilutions of the indicated antibodies. The following fluorescently labelled antibodies were used for surface staining of cells: APC anti-Ly-6G/Ly-6C (RB6-8C5), APC anti-Ly-6G (1A8), APC-Cy7 anti-CD11b (M1/70), Brilliant Violet 421 anti-CD11b (M1/70), FITC anti-CD11c (N418), FITC anti-CD3 (17A2), FITC anti-CD4 (GK1.5), FITC anti-CD8a (53-6.7), FITC anti-NK-1.1 (PK136), FITC anti-CD19 (MB19-1), FITC anti-TER-119, FITC anti-CD45R (B220) (RA3-6B2), and Fixable Viability Dye eFluor 780. After staining, cells were washed with 3 ml of ice cold FACS buffer and resuspended in 500-800 µl FACS buffer for flow cytometry or 1-3 ml for sorting. Intra-cellular staining of Ki-67 (APC anti-Ki67, SolA15) was performed using Foxp3/Transcription Factor Staining Buffer Kit (Tonbo Biosciences). The profiling of cell cycle was performed by fixing cells in 70% ethanol and further incubating with 5µL of propidium iodide (1mg/mL). Flow cytometry was performed using BD LSRFortessa X-20 and FACS analyses were performed using FlowJo v10.8.1. Sorting of cells were performed using FACS Aria Fusion.
Instrument	LSRFortessa X-20, FACS Aria Fusion.
Software	The flowcytometry data was collected using FACSDiva v9.0 (BD) and analyzed with FlowJo v10.8.1(BD).
Cell population abundance	The purity of sorted cells was >90% as assessed by flow cytometry.
Gating strategy	Live cells were gated using FSC-A/SSC-A. Doublets were removed and singlets were selected using FSC-A/FSC-W and SSC-A/SSC-W. Live cells were further gated using Fixable Viability Dye eFluor 780. Live cell population was used for further downstream gating strategies as shown in the manuscript Figures.

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