nature portfolio

Corresponding author(s): Alejandro Vaquero, Jeannie T. Lee

Last updated by author(s): April 6, 2026

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

FastQC (version 0.11.3). Bowtie2 (version 2.2.9). SAMtools (version 0.1.19). deep Tools (version 2.4.2). Trimmomatic (version 0.36) HISAT2 (version 2.1.0). StringTie (version 1.3.3). DESeq2 (version 1.32.0). Novoalign (version 4.02.03). STAR (version 2.7.10a). ggplot2 (version 3.4.2). MACS (version 2.2.7.1). bedtools (version 2.25.0). SNPsplit (version 0.3.4). Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS, version 2023-08-02). HOMER (version 11.1). EdgeR (version 4.4.2; limma version 3.62.2). Trim Galore! (version v.0.6.6). ImageJ (version 1.53). GraphPad (Version 11.0.0). cutadapt (version 1.4.2). Juicebox (version 1.8.8). Python (version 3.10). NumPy (version 1.24.3). SciPy (version 1.10.1). scikit-learn (version 1.3.0). matplotlib (version 3.7.1). pandas (version 2.0.3).

Data analysis

Hi-C analysis
Hi-C reads were trimmed using cutadapt (version 1.4.2) with the options–adaptor = AAGCTTAAGCTT (HindIII ligation junction) and–minimum-length = 20. Reads of each pair were individually mapped to the mus (Xi) and cas (Xa) reference genomes using Bowtie2 (version 2.2.2) and merged into Hi-C summary files and filtered using HOMER (version 4.8) as previously described78. Hi-C contact maps were generated from reads in the filtered HOMER mus and cas tag directories using the 'pre' command of Juicer tools (version 1.13.02) with the genomeID mm1079. The resulting Hi-C contact maps in .hic format were visualized and normalized with the 'Coverage Square Root' option in Juicebox version 1.8.879. Two biological replicates for each cell line showed similar results. Principal component analysis (PCA) was done using the runHiCpca.pl command of HOMER with the filtered HOMER Hi-C tag directory as input and the options -res 250000 -superRes 250000 -genome mm10. Pearson correlation maps were also computed by HOMER using the analyzeHiC command with the filtered HOMER Hi-C tag directory as input and the options -chr chrX -res 250000 –corr.
4D-HiC: structural chromosome modeling
Simulations in this study were performed using the Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS)80 . In the initial configuration, the system was delineated into a series of sequentially connected 'monomers', the count of which was contingent upon the selected resolution. For instance, at a resolution of 200 kb, the entirety of the X-chromosome, spanning 166 Mb, was encapsulated by N=833

monomers, thereby emulating a linearly connected polymer chain. This chain was generated using a self-avoiding random walk (SARW) algorithm, ensuring that the beads did not occupy the same spatial coordinates. To mimic the experimental HiC constraints, harmonic bonds were introduced between interacting monomers based on HiC connectivity data. These bonds were established using the create_bonds command in LAMMPS, ensuring that monomers were in reasonable proximity prior to bond assignment. Once approximately 80% of bonding constraints were met, the structure underwent equilibration using Brownian Dynamics in an implicit solvent environment using the Langevin thermostat. The following pair interactions were defined between consecutively bonded monomers using FENE and Lennard-Jones potentials: The optimal parameter set for the maximum bond length is R0=20σ and the spring constant κ=30ε*/σ2. We choose the repulsive LJ strength ε*=1 in non-dimensional units for this bonded potential, which makes the equilibrium bond length rbond=0.99σ, yet allowing the bond to be stretched up to 20σ. For nonbonded monomers, only the repulsive part of the Lennard-Jones interaction potential was used.

Saddle plot analysis was performed to quantify the strength of A/B compartmentalization by measuring contact enrichment between genomic regions stratified by compartment identity. This approach reveals whether regions with similar compartment scores (A-A or B-B) preferentially interact, while regions with opposing scores (A-B) show depleted interactions. Raw Hi-C contact matrices were first transformed to observed-over-expected (O/E) matrices by normalizing each diagonal of the matrix by its mean contact frequency, thereby removing the distance-dependent decay signal:

O/E(i,j) = Observed(i,j) / E[contacts at distance |i–j|]

Genomic bins (250 kb resolution) were sorted by their PC1 compartment scores in ascending order (B compartment → A compartment). The sorted bins were then divided into n quantiles (n=40–50 bins). For each pair of quantiles (i, j), we computed the mean O/E contact frequency across all bin pairs where one bin belongs to quantile i and the other to quantile j.

Strong compartmentalization produces saddle plots with high A-A and B-B enrichment (O/E > 1) and low A-B values (O/E < 1). Weakened compartmentalization, as observed in Sirt7 KO samples, manifests as reduced dynamic range between same-compartment and cross-compartment interactions, indicating increased mixing between A and B compartments. Global PC1 variance analysis was performed to quantify the overall strength of A/B compartmentalization by measuring how much of the chromosomal contact pattern variance is explained by the first principal component. Higher variance indicates stronger, more distinct compartment segregation. Briefly, Hi-C contact matrices were first normalized to observed/expected (O/E) matrices by dividing each diagonal by its mean contact frequency. Pearson correlation matrices were then computed from the O/E matrices, generating n × n correlation matrices, where n is the number of genomic bins. Principal component analysis (PCA) was performed on the correlation matrices using singular value decomposition, and the variance explained by the first principal component was calculated as follows:

PC1 variance (%) = $(\lambda_1/\Sigma\lambda_i)\times100$, where λ_1 is the first eigenvalue and $\Sigma\lambda_i$ is the sum of all eigenvalues. In our analysis, PC1 variance remained stable between WT and Sirt7-/- (Xa: 22.9% → 24.0%; Xi: 38.4% → 37.3%), indicating that compartment identity is maintained despite loss of SIRT7.

The local mixing index (LMI) was used to quantify the degree of spatial segregation between A and B compartments at local genomic scales. Unlike global metrics, this captures fine-scale mixing that may occur even when compartment identity is maintained. Each genomic bin was classified as either A compartment (PC1>0) or B compartment (PC1<0). For each bin i, a local window spanning ±w bins (typically w = 5–10 bins, or 2.5–5 Mb at 250 kb resolution) was defined. The LMI was then calculated for each bin using the formula: LMIi = |NA − NB| / (NA + NB), where NA is the number of A-compartment bins in the window and NB is the number of B-compartment bins. The overall Local Mixing Index was determined as the genome-wide mean of LMIi. Lower LMI values indicate more mixing (A and B compartments are interspersed locally), while higher values reflect greater segregation (A and B compartments are separated). In our analysis, the LMI on Xa decreased from 4.33 to 3.66 in Sirt7-/- cells, corresponding to a 15.5% increase in local mixing despite maintained compartment identity.

Boundary sharpness was assessed to quantify the abruptness in PC1 scores at compartment boundaries, Sharp boundaries indicate strong physical segregation between A and B compartments, while gradual transitions suggest permeable or weakened boundaries. Discrete derivatives of the PC1 signal were calculated as ∇PC1i = |PC1i+1 − PC1i|. Compartment boundaries were defined as positions where PC1 values crossed zero (sign changes from positive to negative or vice versa). For each boundary b, the local sharpness was defined as the maximum PC1 gradient within a ±3 bin window centered on the boundary (Sharpnessb = max(∇PC1b−3:b+3)). The mean boundary sharpness across all compartment boundaries was calculated as Boundary Sharpness = mean(Sharpnessb). Higher values indicate sharper, more distinct boundaries. The observed decrease from 1.98 to 1.78 on Xa (10.1% reduction) demonstrates that compartment boundaries become more gradual and permeable after SIRT7 loss.

The A-B Mixing Score quantifies the proportion of chromatin contacts that occur between different compartment types (A-B contacts) relative to the total number of contacts, providing a direct measure of the degree of compartment mixing in three-dimensional space. Each genomic bin was first classified as belonging to either the A compartment (PC1>0) or the B compartment (PC1< 0). For each contact Cij, in the Hi-C matrix, contacts were categorized as A-A if both bins were in the A compartment, B-B if both bins were in the B compartment, or A-B if the bins were in different compartments. The mixing score was calculated as follows: A-B Mixing Score = ΣABCij / (ΣAACij + ΣBBCij + ΣABCij). Values range from 0 (perfect segregation) to 0.5 (complete mixing), with typical scores for well-compartmentalized chromosomes falling between 0.15 and 0.25. In our data, the unchanged A-B mixing score (0.22 for both WT and Sirt7-/-) indicates that the global ratio of inter- versus intra-compartment contacts remains constant, even though local boundary integrity decreases.

PC1 Validation Metrics were used to confirm that PC1 scores accurately reflect compartment properties by comparing them with established genomic features. The A compartment, associated with active chromatin, is expected to form longer-range contacts owing to its open and accessible structure. To assess this, the total contact count for each bin (i) was calculated as contact rangei = ΣjCij, and the Pearson correlation between PC1 values and Contact Range was computed across all bins. A positive correlation supports correct compartment assignment. Additionally, the B compartment, corresponding to heterochromatin, typically exhibits enriched short-range contacts due to its compact, self-interacting nature. For each bin, local contacts within ±10 bins were summed (Diagonal Enrichmenti = Σj=i−10i+10Cij), and the Pearson correlation was calculated between PC1 and Diagonal Enrichment. Correlation can be positive or negative depending on whether A or B compartment dominates diagonal contacts.

The TAD Insulation Score was used to quantify the strength of topologically associating domain (TAD) boundaries by measuring the degree to which genomic positions block chromatin contacts from crossing them. Strong boundaries are characterized by low contact frequency across the boundary, while weak boundaries allow contacts to “pass through.” For each genomic position i, a square window of size ±w (typically w = 5 bins, or 2.5 Mb at 500 kb resolution) was defined. The insulation score was calculated as the sum of contacts crossing position Insulationi = Σj=i−wiΣk=ii+wCjk. A log transformation was applied to the insulation vector, and values were smoothed with a Gaussian filter (σ = 2 bins): Insulationi' = Gaussianσ=2(log2(Insulationi + 1)). TAD boundaries were then identified as local minima in the insulation profile using peak detection (prominence threshold = 0.3–0.5). The mean insulation score at all detected boundaries was reported as the TAD Insulation Score. Lower values indicate stronger boundaries, with fewer contacts crossing. In our analysis, the 13.4% decrease on Xa (0.82 → 0.71) demonstrates substantial weakening of TAD boundaries after SIRT7 loss, while Xi shows minimal change (−2.6%) as its boundaries were already disrupted.

All Hi-C analyses were implemented in Python 3.10 using NumPy (v1.24.3), SciPy (v1.10.1), scikit-learn (v1.3.0), matplotlib (v3.7.1), and pandas (v2.0.3). Contact matrices were processed using sparse matrix representations (scipy.sparse.csr_matrix) for computational efficiency. Molecular dynamics simulations were performed using LAMMPS80 daily version (2023-08-02). Statistical analyses and bootstrap procedures

were conducted using Python.

Volume calculation method

Chromosome volumes were calculated from coarse-grained polymer structures, as described previously51. They are computed by generating a simulated Gaussian density map from the structure, then measuring the volume enclosed within a contour threshold. Each coarse grained 'bead' is represented as a 3D Gaussian:

where ri is the position of bead i and σ is the Gaussian width determined by the resolution parameter, which integrates the number of grid voxels exceeding a specified density threshold.

RNA-seq analysis

For RNA-seq analysis from mESC and primary MEF (female and male), raw reads were prefiltered (Novocraft software fastqfilter) for rDNA (mm10 and DM6) using a quick alignment to hybrid mm10 and dm6 genome (novoalign 4.02.03). In all cases (including TTF cells analysis), rDNA and mtDNA were removed by aligning to a library (STAR_2.7.10a) containing rDNA and mtDNA sequences. Adaptor sequences were trimmed with Trim Galore! (version v.0.6.6), and PCR duplicates were removed using adhoc methods based on exact sequence matches between reads utilizing the Linux uniq function or for MEFs after alignment (markdup -r)84. MEFs utilized DM6 spike-in reads for normalization. These were counted commensurate with alignment to a hybrid genome. Alignments were a two-pass alignment approach using STAR. Of note for TST cells, independent alignments were performed to an mm10 genome modified to Mus musculus and a mm10 genome modified to Cas castaneous/EiJ genome, scoring as cas/mus/neutral (equal). The processing of allelic-specific alignments was performed similarly to that described before85.

Normalization of MEFs followed a process similar to that of calibrated ChIP. The total number of DM6 fragments was matched between samples by scaling to the smallest DM6 count, subsampling factors were further adjusted based on a ratio of DM6 to total fragments (mm10 +DM6) to account for variability in spike-in to library size (scaled to the largest ratio: (dm6/(mm10+DM6))/(largest dm6/(mm10+DM6)). The derived scale factors follow: MEF Male Sirt7-/-_1 (0.54), MEF Male Sirt7-/-_2 (0.50), MEF Male Sirt7-/-_3 (0.50), MEF Male WT_1 (0.40), MEF Male WT_2 (0.34), MEF Male WT_3 (0.40), MEF Female Sirt7-/-_1 (0.58), MEF Female Sirt7-/-_2 (0.47), MEF Female Sirt7-/-_3 (0.49), MEF Female WT_1 (0.45), MEF Female WT_2 (0.56), MEF Female WT_3 (0.50). To obtain a counts table for DE RNA analysis, MEF sampled bam files were subjected to the software FeatureCounts86 with largestOverlap, whereas mESC and TTF used bedtools (2.31.0 or lesser) to map over reference genes. Differential expression was analyzed with DESeq2 (1.50.0 or higher). Where possible, we used models that incorporated batch effect mitigation. For instance, when performing DE analysis of rescue vs. WT, Sirt7H187Y vs. WT, we used matched Sirt7-/- sample cohorts for potential batch effects. Gene counts were prefiltered such that each gene (WT and treatment) had as sum of reads > 10 and for day 0, day 7 (mESC) and TTF, FPKM > 0.5. MEFs used EdgeR87 (v4.4.2; limma v3.62.2) trimmed mean of M-values (TMM) normalization in addition to sampling. We only perform contrasts on pairs of experiments. After creating a Digital Gene Expression (DGE) list from raw counts, we keep only genes where at least one condition has a FPM > 1. After which, we normalize the full complement of genes and keep only those genes for which at least one condition has a mean normalized count > 50. We then calculate dispersion estimates, which account for biological variability across replicates, as calculated in three steps: Common Dispersion, Trended Dispersion, and Tagwise Dispersion. Finally, a contrast is performed between the treatment group and WT (control) or Female vs. Male (control) using glmQLFTest and FDR. The function ggplot2 (v3.4.2) with stat ecdf creates cumulative distribution plots on mean values of fpm (normalized.lib.sizes = TRUE) counts, normalized data. data. Similarly, MA plots were generated with ggplot2 (v3.4.2) comparing log2 FC vs. log2 CPM. Batch 1 biological replicate 1 TTF Sirt7-/- was removed from the analysis. RNA-seq analysis from primary MEF was done using the previous dataset GSE128893 and the new datasets GSE256287.

Vector analysis

Vector analysis was used to identify coordinated changes in gene expression from RNA-seq and ChIP-seq experiments resulting from SIRT7 loss. For RNA, gene-level expression was quantified in FPKM units and low-noise features were retained by requiring (i) at least 11 "comp" reads aggregated across WT and perturbed samples (Sirt7-/-, Sirt7H187Y) and (ii) WT FPKM greater than 0.5. Allele-specific RNA values were reconstructed by multiplying total "comp" counts by the allele-specific fraction. For ChIP (H3K36ac), we quantified transcription start site (TSS)–proximal signal. SES-normalized 50 bp bigWig coverage files were summarized using deepTools (multiBigwigSummary, v3.5.2) and converted to bed-like tables. Coverage was extracted over the union of significant MACS (v2.2.7.1) peaks ($q < 0.05$) expanded by ±500 nt genome-wide, across all alleles and both WT and perturbed conditions. TSS-level ChIP coverage was then computed using bedtools (v2.25.0) map over both Watson and Crick strands within ±3 kb of annotated TSSs. These data were used to construct a chrX counts table containing RNA and ChIP measurements for each condition (WT and Sirt7-/-, WT and Sirt7H187Y) and each allele. To quantify the combined RNA-ChIP response, we computed for each gene a two-dimensional effect vector $v = (dX, dY)$, where $dY = \log2(RNA\_KO + 1) - \log2(RNA\_WT + 1)$ and $dX = \log2(ChIP\_KO + 1) - \log2(ChIP\_WT + 1)$. A pseudocount of 1 stabilized log-fold changes for zero values, although most very low-expression genes were removed by the RNA prefilter. For each gene we calculated the polar angle theta = atan2(dY, dX) converted to degrees, and the magnitude $|v| = \text{sqrt}(dX2 + dY2)$. Genes were assigned qualitative directional labels based on the angle. We classified genes as RNA-Up or RNA-Down only when the RNA response was substantial relative to the corresponding ChIP response. Genes with angles satisfying $|\theta| < 10°$ were considered uninformative, reflecting an RNA change less than approximately 20 percent of the ChIP change (i.e., $|dY/dX| < \tan 10° \approx 0.176$). This threshold removes cases in which RNA changes are dominated by noise or are only weakly responsive relative to H3K36ac ChIP. In addition, a gene was considered RNA-responsive only if the RNA change exceeded ≈0.05 in log2(x+1) space. To assess significance of coordinated RNA–ChIP responses, we computed a synergy score equal to the projection of v onto the (1,1) axis: $s = (dX + dY)$ divided by sqrt(2). Positive values indicate RNA and H3K36ac ChIP increasing together, whereas negative values indicate coordinated decreases. The null distribution scale was estimated using the median absolute deviation (MAD), and MAD-scaled synergy scores were converted to p-values under a normal approximation. False discovery rates were computed using Benjamini–Hochberg correction.

ChIP-seq and CHART-seq Analysis

ChIP-seq was conducted as described before4,5,29. Raw read quality was first assessed with FastQC88 version 0.11.3, followed by adapter clipping and quality trimming with Trimmomatic89 version 0.36. Reads were then aligned to the mouse (UCSC mm10) reference genome with Bowtie290 version 2.2.9 with options --n 2 --k 10 and converted to BAM format with SAMtools version 0.1.19. ChIP, library, and alignment quality was evaluated by PCR bottleneck coefficient (NRF, PBC1, PBC2), insert size distribution, and strand cross-correlation analysis. All samples passed the ENCODE quality standards for ChIP-seq studies. BigWig ChIP signal files were generated from BAM data using deepTools version 2.4.2 using a bin size of 50 bp, scaled with SES, and normalized by INPUT. Where available, data were scaled based on spike-in controls (EpiCypher K-MetStat, K-AcylStat, or exogenous chromatin). CHART-seq was analyzed in similar fashion as ChIP-seq.

Peak calling was performed using MACS version 2.2.7.1 by providing ChIP as --treatment and matched Input as --control. For H3K27me3 and H3K36Ac, peak calling was performed in --broad mode with --broad-cutoff 0.1 --qvalue 0.1 --scale-to small. Peak intersections were computed using bedtools version 2.25.0, and the number of peaks was tabulated and plotted as Venn diagrams. IGV was used to generate visualizations of ChIP signals at representative loci. Average profiles of ChIP signals at genes were generated using custom scripts. For H3K27me3 average profiles at DEGs, DEGs were determined from RNA-seq data in female MEF cells (see RNA-seq methods for details) as genes with adjusted $P < 0.05$ with the study design "~ genotype" and the sign of the log2 FC determining "Up" or "Down" classification, while all other genes were

classified as “No-change”. CEAS was used to generate plots of peak frequencies per chromosome and average profiles at TSS, TTS, and metagenes, and replicates combined and plotted as a mean profile with SEM bands using Python and seaborn. GREAT36 analysis to predict functions of cis-regulatory regions of H3K36ac peaks was performed using significant region-based binomial statistical tests using the basal plus extension rule. Informed ontology terms were significant by both binomial and hypergeometric tests, with an FDR of 0.05. Genomic regions associated with their respective genes are listed in tables according to the basal plus extension rule.
For γH2AX, the genome was divided into a set of 100KB non-overlapping bins using bedtools `makewindows` command. The number of reads in each of the bins was counted for γH2AX-ChIP and H2AX-ChIP samples using bedtools (v2.25.0) `coverage` command. The resulting bed files were read into a pandas dataframe, and γH2AX-ChIP read counts per bin were adjusted for library size and then normalized by corresponding H2AX-ChIP read counts per bin. Replicates were then averaged on a per-bin basis. Finally, for each bin, the log2 ratio of Sirt7-/- and WT was calculated, and these bin values were plotted as boxplots. Significance testing was performed using a t-test for independent samples from scipy (v 1.10.1).
ENCODE current standards for SIRT7 ChIP-seq were obtained in MEF and TTF clones in WT and Sirt7-/- cells. Library complexity was measured using the non-redundant fraction (NRF) and PCR bottlenecking coefficients 1 and 2 (PBC1 and PBC2). SIRT7 ChIP-seq in MEF samples fall into the “Moderate/Acceptable” range. SIRT7 ChIP-seq in TTF samples fall into the “ideal” range (WT) or “Mild/Compliant” for negative sample controls (Sirt7-/- cells) as detailed in Supplementary Table 16.

Allele-specific analysis
SNPsplit version 0.3.4 was used. We first used the command `SNPsplit_genome_preparation` in single strain mode to generate new reference genome artifacts where high-confidence SNPs from the Mouse Genomics Project between the UCSC mm10 reference genome and strain “CAST_EiJ” (mgp.v5.merged.snps_all.dbSNP142.ucsc.vcf) were N-masked. The N-masked genomes were then used to generate indexes for the Bowtie2 and HISAT2 aligners using default options.
Aligned reads were sorted into three categories based on informative SNPs using the command `SNPsplit`: A) reads from mus musculus [genome1] B) reads from mus Castaneus [genome2] or C) unassignable reads [unassigned]. The resulting BAM files were then sorted by position and indexed using samtools version 0.1.19.
For allele-specific SIRT7 ChIP-seq, the genome was divided into a set of 1MB non-overlapping bins using bedtools `makewindows` command. The number of reads in each of the 1MB bins was counted for ChIP and INPUT samples using bedtools `coverage` command. The resulting bed files were read into a pandas dataframe, and ChIP read counts per bin were normalized by corresponding Input read counts per bin. Allele-specific read counts for genome2 (cas) were plotted against read counts for genome1 (mus) in chromosomes 13 and X using matplotlib, and a linear regression was performed for each using scipy.stats.linregress. For SNP distribution, there are 636,442 informative SNPs on chromosome X between mus and cas. In our 1Mb bins on chromosome X, the median bin contains 4,750 informative SNPs (mean=3,721.9, max=7,510). Overall, on chromosome X, the average spacing between adjacent SNPs is 263.6 bp. We typically find that 40-60% of reads are assignable to mus or cas genome wide.
For H3K36Ac fold-change plots, average signal profiles at chrX genes on the cas (Xa) and chr13 genes on the cas allele were calculated. The ratio between these two average signals was calculated and plotted. KS test was used for significance testing of the TSS region, as indicated by the shaded region, the TSS to the TSS + 1.25Kb.

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

All new datasets generated from this work have been deposited in GEO under accession GSE256287. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD057198.
The custom code for RNA-seq and related ChIP-seq presented in this study is archived on Zenodo https://doi.org/10.5281/zenodo.19464181.

GSE256287 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome (the SuperSeries)

GSE256277 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [CHART-seq]
GSE256278 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [H3K27me3_ES]
GSE256279 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [H3K27me3_TTF]
GSE256281 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [H3K36Ac_ES]
GSE256282 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [H3K36Ac_TTF]
GSE256283 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [ChIP_SIRT7_TTF]
GSE256284 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [RNA-seq ES]
GSE256285 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [RNA-seq TTF]
GSE262651 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [Hi-C]
GSE295582 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [ChIP_H2AX_gH2AX_MEF]
GSE295583 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [ChIP_H2AX_gH2AX_TTF]
GSE295584 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [ChIP_SIRT7_MEF]
GSE295585 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [ChIP_SIRT7_primary_TTF]
GSE295586 Sirtuin 7 regulates dosage compensation and safeguards the female X-chromosome [RNA-seq_male_female_MEF]

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

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	No statistical methods were used to predetermine sample size. Sample sizes were chosen based on established precedents in the field and prior experimental experience to ensure sufficient power for detecting biologically meaningful differences.For the longitudinal survival study, a large-scale cohort of 2,132 mice was monitored to provide high-resolution Kaplan-Meier curves. For disease onset phenotyping, cohorts of n=21 23 per group were used. For oocyte studies, we used biologically independent mice (WT, n=5; Sirt7-/-, n=4). In vivo experiments (including EZH2 RIP and RT-qPCR) were typically performed with n=3-5 biological replicates. High-throughput sequencing (RNA-seq, ChIP-seq, and Hi-C) utilized n 2 or n=3 biological replicates per condition, consistent with ENCODE guidelines. In all cases, the adequacy of the sample size was confirmed post-hoc by the achievement of statistical significance and the high level of reproducibility across independent experiments. Below published precedents in epigenetics for NGS: ChIP-seq: https://doi.org/10.15252/embj.201593499 https://doi.org/10.1126/sciadv.aaz2590 HiC: https://doi.org/10.1038/s41467-019-10755-3 https://doi.org/10.1073/pnas.2107092118 https://doi.org/10.1371/journal.pcbi.1005665 RNA-seq: https://doi.org/10.1038/s41586-022-04537-z
Data exclusions	no data excluded
Replication	All experiments were performed independently and yielded similar results. For high-throughput sequencing and genomic studies (RNA-seq, Hi-C, ChIP-seq, and CHART), experiments were performed in at least two or three biological replicates initiated from different cell passages on different days. For molecular and cytogenetic assays, including RNA-FISH, DNA-FISH, and immunostaining, experiments were performed using at least three to five biologically independent samples. RT-qPCR and Western blot experiments were replicated at least three times independently using different biological replicates.
Randomization	Randomization was not applicable to these studies because the experimental groups were defined by fixed genetic variables (genotype: WT, Sirt7-/-) and/or sex (male, female). In cell culture experiments, comparisons were made between established WT and Sirt7-/- cell lines from our mouse colony or by CRISPR/Cas9 (properly genotyped for genotype and sex); therefore, samples could not be randomly assigned to groups. For in vivo studies, including the Kaplan-Meier survival curves and pachytene oocyte/ spermatocytes analysis, mice were assigned to groups based on their predetermined genotype and sex. To minimize bias, we ensured that all groups were processed under identical conditions, using age-matched and littermate-controlled cohorts wherever possible.
Blinding	Blinding was not applicable to the experimental setup and sample collection phases, as the identification of specific genotypes and sex variables is a technical requirement for the precise timing of embryonic tissue harvesting and for the management of long-term cohorts for mouse survival analysis. To ensure maximum objectivity and mitigate bias during data acquisition and data analysis, the following measures

were implemented: 1) Automated Data Processing: All high-throughput sequencing data (RNA-seq, ChIP-seq, Hi-C, and Xist CHART) were processed using standardized, automated computational pipelines. These pipelines treat all raw data files identically and autonomously, ensuring that the genotype of the sample does not influence the bioinformatic output. 2) Independent Reproducibility: The consistency of results across biologically independent replicates, performed on different days by different researchers where applicable, and further validates the robustness of the findings in the absence of formal 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 and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Antibodies

Antibodies used

Anti-SIRT7 (Cell Signaling, #5360, lot number 1), anti-Histone H3 (#9715, Cell Signaling), anti-H3K27me3 (Cell Signaling, #9733, anti-H3K36ac (#27683, Cell Signaling, EpiCypher, #13-0035), anti-tubulin (ab-6160, Abcam), anti-EZH2 antibody (#5246, Cell Signaling), anti- phospho-EZH2 Thr345 (PA5-114574, Thermo Fisher), anti-H2A.X antibody (ab20669, Abcam), anti-H2A.X phospho S139 (ab2893, Abcam), rabbit anti-SYCP3 (Novus biologicals, NB300-232), mouse anti-ɣH2AX (Sigma, 05-636-I), rabbit anti-GAPDH (#5174 Cell Signaling), Alexa Fluor-488 anti-rabbit IgG(Thermo Fisher, A-11034), Alexa Fluor-555 anti-rabbit IgG (Thermo Fisher, A-32727), rabbit anti-HORMAD1 (Proteintech 13917-1-AP), Alexa Fluor 555 goat anti-mouse IgG (ThermoFisher A-32727), AlexaFluor 647 goat anti-rabbit IgG (TermoFisher A-21245), Cy3-conjugated polyclonal goat anti-mouse antibody (#115-165-003, Jackson ImmunoResearch Laboratories).

Validation

Antibodies used in this study were validated for species specificity (Mus musculus) and application through manufacturer testing, established literature, and rigorous in-house controls. For instance, the anti-SIRT7 antibody (Cell Signaling, #5360) was validated in-house using CRISPR/Cas9-mediated knockout of Sirt7 and confirmed via Western blot and ChIP-seq in MEF cell lines derived from Sirt7 KO mice, where the SIRT7-specific signal was entirely abolished. Validation for chromatin marks anti-H3K27me3 (Cell Signaling, #9733) and anti-H3K36ac (EpiCypher, #13-0035, Cell Signaling, #27683) was performed using SNAP-ChIP spike-ins for K-AcylStat and K-MetStat (EpiCypher, #19-1001, #19-3001), followed by next-generation sequencing and quantitative recovery of barcoded recombinant dNucs to ensure minimal cross-reactivity and high mapping specificity.
Meiotic markers anti-SYCP3 (Novus, NB300-232) and anti-HORMAD1 (Proteintech, 13917-1-AP) were validated by their characteristic localization to the synaptonemal complex axial elements and unsynapsed regions in pachytene spreads, respectively. Loading controls including anti-H3 (CST, #9715), anti-tubulin (Abcam, ab-6160), and anti-GAPDH (CST, #5174) were validated by the presence of a single band at the predicted molecular weight in mouse protein lysates. All antibodies cited have been extensively used in peer-reviewed literature for the applications described herein.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

Male and Female primary MEF (E14.5) obtained from 129sv Crossed heterozygous mice for one SIRT7 knockout allele. tail tip female fibroblasts (TTF) isolated from hybrid post-natal mice. The hybrid TTFs carry one inactive X (Xi) chromosome of Mus musculus (mus) and one active X (Xa) of Mus castaneus (cas) origin, thereby enabling allelic discrimination. mESC line (TST mouse ES cells) also carries one inactive X (Xi) chromosome of Mus musculus (mus) and one active X (Xa) of Mus castaneus (cas) origin for allelic discrimination. Pachytene stage oocytes at embryonic day 17.5 from WT and Sirt7-/- female mouse embryos. WT and Sirt7-/- female non-clonal primary fibroblast F1 hybrid with random XCI (Mus × Cas). WT and Sirt7-/- Male primary spermatocytes.

Authentication

hybrid (mus and cas) cell lines TTF and mESC (TST) were analyzed and observed an Xi-specific regulation. In the case of primary MEFs, cell preparations from embryos were expected to reach 80% to 90% confluence and harvested between passages 0 and 3/4, and their MEF morphology remained as mesenchymal fibroblasts

Mycoplasma contamination

primary MEF cells were tested

Commonly misidentified lines
(See ICLAC register)

the study did not use misidentified cell lines

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	Species & Strain: Mus musculus, 129sv background. Age and Sex: The study utilized both male and female mice across a broad developmental and chronological range. Survival analysis and disease-onset phenotyping were performed on a cohort of 2,132 mice monitored from birth up to 26 months of age. Embryos were harvested at specific developmental stages, including E14.5 (for RNA expression and Mendelian ratio analysis) and E17.5 (for oocyte pachytene chromosome spreads). Genotypes: Experimental groups consisted of wild-type (WT), heterozygous (Sirt7+/-), and homozygous knockout (Sirt7-/-) littermates. All mice were housed in a centralized, pathogen-free barrier facility. Animals were maintained under a standardized 14:10 hour light/dark cycle with the ambient temperature strictly regulated at 22 ± 1°C and relative humidity maintained between 40% and 60%. Mice had ad libitum access to standard irradiated rodent chow and water. All animal husbandry and experimental procedures were performed in accordance with the protocols approved by the Rutgers University IACUC and CMCiB policies.
Wild animals	no wild animals were used in the study
Reporting on sex	Sex was considered in the study design. A PCR method for assessing sex was used for primary MEF and fibroblasts, while male and female mice were assessed by measuring the anogenital distance.
Field-collected samples	no field collected samples were used in the study.
Ethics oversight	Animal studies were conducted in accordance with Rutgers University IACUC and CMCiB policies.

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

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA

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 All data sets generated from this work have been deposited to GEO under the accession number GSE256287.
Files in database submission	GSM8092450 ChIP_H3K27me3_S7KO_ES_d0_Rep1 GSM8092451 ChIP_H3K27me3_S7KO_ES_d0_Rep2 GSM8092452 ChIP_H3K27me3_S7KO_ES_d7_Rep1 GSM8092453 ChIP_H3K27me3_S7KO_ES_d7_Rep2 GSM8092454 ChIP_H3K27me3_S7WT_ES_d0_Rep1 GSM8092455 ChIP_H3K27me3_S7WT_ES_d0_Rep2 GSM8092456 ChIP_H3K27me3_S7WT_ES_d7_Rep1 GSM8092457 ChIP_H3K27me3_S7WT_ES_d7_Rep2 GSM8092458 INPUT_H3K27me3_S7KO_ES_d0_Rep1 GSM8092459 INPUT_H3K27me3_S7KO_ES_d0_Rep2 GSM8092460 INPUT_H3K27me3_S7KO_ES_d7_Rep1 GSM8092461 INPUT_H3K27me3_S7KO_ES_d7_Rep2 GSM8092462 INPUT_H3K27me3_S7WT_ES_d0_Rep1 GSM8092463 INPUT_H3K27me3_S7WT_ES_d0_Rep2 GSM8092464 INPUT_H3K27me3_S7WT_ES_d7_Rep1 GSM8092465 INPUT_H3K27me3_S7WT_ES_d7_Rep2 GSM8092466 ChIP_H3K27me3_S7KO_TTF_Rep1 GSM8092467 ChIP_H3K27me3_S7KO_TTF_Rep2 GSM8092468 ChIP_H3K27me3_S7WT_TTF_Rep1 GSM8092469 ChIP_H3K27me3_S7WT_TTF_Rep2

GSM8092470 Input_S7KO_TTF_Rep1
GSM8092471 Input_S7KO_TTF_Rep2
GSM8092472 Input_S7WT_TTF_Rep1
GSM8092473 Input_S7WT_TTF_Rep2
GSM8092476 ChIP_H3K36Ac_S7KO_d0_Rep1
GSM8092477 ChIP_H3K36Ac_S7KO_d0_Rep2
GSM8092478 ChIP_H3K36Ac_S7KO_d7_Rep1
GSM8092479 ChIP_H3K36Ac_S7KO_d7_Rep2
GSM8092480 ChIP_H3K36Ac_S7WT_d0_Rep1
GSM8092481 ChIP_H3K36Ac_S7WT_d0_Rep2
GSM8092482 ChIP_H3K36Ac_S7WT_d7_Rep1
GSM8092483 ChIP_H3K36Ac_S7WT_d7_Rep2
GSM8092484 INPUT_H3K36Ac_S7KO_d0_Rep1
GSM8092485 INPUT_H3K36Ac_S7KO_d0_Rep2
GSM8092486 INPUT_H3K36Ac_S7KO_d7_Rep1
GSM8092487 INPUT_H3K36Ac_S7KO_d7_Rep2
GSM8092488 INPUT_H3K36Ac_S7WT_d0_Rep1
GSM8092489 INPUT_H3K36Ac_S7WT_d0_Rep2
GSM8092490 INPUT_H3K36Ac_S7WT_d7_Rep1
GSM8092491 INPUT_H3K36Ac_S7WT_d7_Rep2
GSM8092492 ChIP_H3K36ac_S7KO_TTF_Rep1
GSM8092493 ChIP_H3K36ac_S7KO_TTF_Rep2
GSM8092494 ChIP_H3K36ac_S7WT_TTF_Rep1
GSM8092495 ChIP_H3K36ac_S7WT_TTF_Rep2
GSM8092496 Input_S7KO_TTF_Rep1
GSM8092497 Input_S7KO_TTF_Rep2
GSM8092498 Input_S7WT_TTF_Rep1
GSM8092499 Input_S7WT_TTF_Rep2
GSM8092500 ChIP_SIRT7_S7KO
GSM8092501 ChIP_SIRT7_S7WT_Rep1
GSM8092502 ChIP_SIRT7_S7WT_Rep2
GSM8092503 INPUT_S7KO
GSM8092504 INPUT_S7WT
GSM8953394 ChIP_H2AX_Female_KO_XRAY_rep1
GSM8953395 ChIP_H2AX_Female_KO_XRAY_rep2
GSM8953396 ChIP_H2AX_Female_WT_XRAY_rep1
GSM8953397 ChIP_H2AX_Female_WT_XRAY_rep2
GSM8953398 ChIP_H2AX_Male_KO_XRAY_rep1
GSM8953399 ChIP_H2AX_Male_KO_XRAY_rep2
GSM8953400 ChIP_H2AX_Male_WT_XRAY_rep1
GSM8953401 ChIP_H2AX_Male_WT_XRAY_rep2
GSM8953402 ChIP_gH2AX_Female_KO_XRAY_rep1
GSM8953403 ChIP_gH2AX_Female_KO_XRAY_rep2
GSM8953404 ChIP_gH2AX_Female_WT_XRAY_rep1
GSM8953405 ChIP_gH2AX_Female_WT_XRAY_rep2
GSM8953406 ChIP_gH2AX_Male_KO_XRAY_rep1
GSM8953407 ChIP_gH2AX_Male_KO_XRAY_rep2
GSM8953408 ChIP_gH2AX_Male_WT_XRAY_rep1
GSM8953409 ChIP_gH2AX_Male_WT_XRAY_rep2
GSM8953410 INPUT_Female_KO_XRAY_rep1
GSM8953411 INPUT_Female_KO_XRAY_rep2
GSM8953412 INPUT_Female_WT_XRAY_rep1
GSM8953413 INPUT_Female_WT_XRAY_rep2
GSM8953414 INPUT_Male_KO_XRAY_rep1
GSM8953415 INPUT_Male_KO_XRAY_rep2
GSM8953416 INPUT_Male_WT_XRAY_rep1
GSM8953417 INPUT_Male_WT_XRAY_rep2
GSM8953418 ChIP_H2AX_KO_XRAY_rep1
GSM8953419 ChIP_H2AX_KO_XRAY_rep2
GSM8953420 ChIP_H2AX_WT_XRAY_rep1
GSM8953421 ChIP_H2AX_WT_XRAY_rep2
GSM8953422 ChIP_gH2AX_KO_XRAY_rep1
GSM8953423 ChIP_gH2AX_KO_XRAY_rep2
GSM8953424 ChIP_gH2AX_WT_XRAY_rep1
GSM8953425 ChIP_gH2AX_WT_XRAY_rep2
GSM8953426 INPUT_KO_XRAY_rep1
GSM8953427 INPUT_KO_XRAY_rep2
GSM8953428 INPUT_WT_XRAY_rep1
GSM8953429 INPUT_WT_XRAY_rep2
GSM8953430 ChIP_SIRT7_KO_rep1_Female
GSM8953431 ChIP_SIRT7_KO_rep1_Male
GSM8953432 ChIP_SIRT7_KO_rep2_Female
GSM8953433 ChIP_SIRT7_KO_rep2_Male
GSM8953434 ChIP_SIRT7_WT_rep1_Female
GSM8953435 ChIP_SIRT7_WT_rep1_Male
GSM8953436 ChIP_SIRT7_WT_rep2_Female

GSM8953437 ChIP_SIRT7_WT_rep2_Male
GSM8953438 INPUT_KO_Female
GSM8953439 INPUT_KO_Male
GSM8953440 INPUT_WT_Female
GSM8953441 INPUT_WT_Male
GSM8953442 ChIP_SIRT7_KO_rep1
GSM8953443 ChIP_SIRT7_KO_rep2
GSM8953444 ChIP_SIRT7_WT_rep1
GSM8953445 ChIP_SIRT7_WT_rep2
GSM8953446 INPUT_KO_rep1
GSM8953447 INPUT_KO_rep2
GSM8953448 INPUT_WT_rep1
GSM8953449 INPUT_WT_rep2
GSM9499520 ChIP-gH2AX_S7HY-c5
GSM9499521 ChIP-gH2AX_S7HY-c7
GSM9499522 ChIP-gH2AX_S7KO-c22
GSM9499523 ChIP-gH2AX_S7KO-c22+S7WT
GSM9499524 ChIP-gH2AX_S7KO-c9
GSM9499525 ChIP-gH2AX_S7KO-c9+S7WT
GSM9499526 ChIP-H2AX_S7HY-c5
GSM9499527 ChIP-H2AX_S7HY-c7
GSM9499528 ChIP-H2AX_S7KO-c22
GSM9499529 ChIP-H2AX_S7KO-c22+S7WT
GSM9499530 ChIP-H2AX_S7KO-c9
GSM9499531 ChIP-H2AX_S7KO-c9+S7WT
GSM9499540 ChIP-H3K36Ac_S7HY-c5
GSM9499541 ChIP-H3K36Ac_S7HY-c7
GSM9499542 ChIP-H3K36Ac_S7KO-c9
GSM9499543 ChIP-H3K36Ac_S7KO-c9+S7WT
GSM9499544 ChIP-H3K36Ac_S7KO-c22
GSM9499545 ChIP-H3K36Ac_S7KO-c22+S7WT
GSM9499546 INPUT_S7HY-c5
GSM9499547 INPUT_S7HY-c7
GSM9499548 INPUT_S7KO-c9
GSM9499549 INPUT_S7KO-c9+S7WT
GSM9499550 INPUT_S7KO-c22
GSM9499551 INPUT_S7KO-c22+S7WT
GSM9499552 SIRT7-HY-c4
GSM9499553 SIRT7-HY-c5
GSM9499554 SIRT7-HY-c7
GSM9499555 SIRT7-KO-c2
GSM9499556 SIRT7-KO-c2+S7WT
GSM9499557 SIRT7-KO-c22
GSM9499558 SIRT7-KO-c22+S7WT
GSM9499559 SIRT7-KO-c9
GSM9499560 SIRT7-KO-c9+S7WT

Genome browser session
(e.g. UCSC)

NA

Methodology

Replicates

RNA-seq experiments: 3 replicates per genotype (WT and Sirt7-/-, rescue with WT, rescue with Sirt7H187Y). XIST CHART-seq: 2 replicates. SIRT7, H3K36ac, H3K27me3, H2AX, gH2AX ChIP-seq experiments: 2 replicates per genotype (WT and SIRT7 KO). Hi-C: 2 replicates per genotype (WT and Sirt7-/-)

Sequencing depth

All experiments were performed using paired-end sequencing with a mixture of 100 bp and 150 bp read lengths, yielding an average sequencing depth of 25–80 million reads per sample.

Antibodies

Anti-SIRT7 (Cell Signaling, #5360, lot number 1), anti-H3K27me3 (Cell Signaling, #9733, anti-H3K36ac (#27683, Cell Signaling, EpiCypher, #13-0035), anti-H2A.X antibody (ab20669, Abcam), mouse anti-γH2AX (Sigma, 05-636-I).

Peak calling parameters

Peak calling was performed using MACS version 2.2.7.1 by providing ChIP as --treatment and matched Input as --control. For H3K27me3 and H3K36Ac, peak calling was performed in --broad mode with --broad-cutoff 0.1 --qvalue 0.1 --scale-to small. Peak intersections were computed using bedtools version 2.25.0, and the number of peaks were tabulated and plotted as venn diagrams. Average profiles of ChIP signal at genes were generated using custom scripts. CEAS was used to generate plots of peak frequencies per chromosome and average profiles.

Data quality

Raw read quality was first assessed with FastQC67 version 0.11.3, followed by adapter clipping and quality trimming with Trimmomatic 68 version 0.36. Reads were then aligned to the mouse (UCSC mm10) reference genome with Bowtie269 version 2.2.9 with options --n 2 --k 10 and converted to BAM format with SAMtools version 0.1.19. ChIP, library, and alignment quality was evaluated by PCR bottleneck coefficient (PBC), insert size distribution, and strand cross-correlation analysis. All samples passed the ENCODE quality standards for ChIP-seq studies. BigWig ChIP signal files were generated from BAM data using deepTools version 2.4.2 using a bin size of 50 bp, scaled with SES, and normalized by INPUT. Where available, data were scaled based on spike-in controls (EpiCypher K-MetStat, K-AcylStat, or exogenous chromatin).

Software	Quality control: FastQC version 0.11.3, followed by adapter clipping and quality trimming with Trimmomatic. Aligment to genome: Bowtie2 version 2.2.9 with options --n 2 --k 10 and converted to BAM format with SAMtools version 0.1.19. BigWig ChIP signal files were generated from BAM data using deepTools version 2.4.2

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	Tail tip fibroblasts (TTFs) were isolated from 129sv mice and cultured in DMEM, high glucose, GlutaMAX™ supplement, pyruvate (Thermo Fisher Scientific, #10569044), supplemented with 20% FBS, 25 mM HEPES pH 7.2-7.5, 500 U/ml), 1x MEM Non-Essential Amino Acids (Thermo Fisher Scientific, #11140076), 100U/ml Penicillin-Streptomycin (Thermo Fisher Scientific, # 15140163), and 0.1 mM 2-mercaptoethanol (Thermo Fisher Scientific, #21985023) at 37°C with 5% CO2. For gene editing experiments, 1 million cells were nucleofected using the Alt-R CRISPR-Cas9 system (Integrated DNA Technologies). Following nucleofection, cells were harvested and resuspended as a single-cell suspension in PBS supplemented with 1% BSA.
Instrument	Data collection was performed using a BD FACSAria flow cytometer
Software	Data were acquired and analyzed using BD FACSDiva software
Cell population abundance	Cell population abundance was determined through a hierarchical gating strategy. From an initial pool of 1,000,000 events, the primary scatter gate identified 907,538 events (90.8%). Subsequent doublet exclusion yielded 782,514 events in the first singlet gate (Singlets f) and 776,426 events in the final singlet gate (Singlets s), representing 77.6% of the total starting population. Within this highly purified singlet fraction, the target GFP-positive (GFP+) population consisted of 333 cells.
Gating strategy	The gating hierarchy was designed to systematically exclude debris and cell aggregates to ensure the purity of the analyzed population. Initial events were gated on a Forward Scatter Area (FSC-A) vs. Side Scatter Area (SSC-A) plot to define the primary fibroblast population (Scatter). To exclude doublets and higher-order aggregates, a sequential two-step gating strategy was applied: first, using FSC-A vs. Forward Scatter Height (FSC-H) to identify the "Singlets f" population, followed by Side Scatter Area (SSC-A) vs. Side Scatter Height (SSC-H) to identify the "Singlets s" population. Finally, GFP-positive cells were identified on a FSC-A vs. FITC-A (GFP) plot. The boundary between the GFP-negative and GFP-positive populations was defined based on the baseline autofluorescence of non-nucleofected control cells, allowing for the precise isolation of the GFP+ fraction. A figure exemplifying this hierarchical gating strategy and the corresponding population statistics is provided in Supplementary Figure 2.

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