nature research

Corresponding author(s): Rafael Daniel Camerini-Otero
Galina Petukhova

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Sequencing data collection used Illumina pipelines for image processing and base calling.
Data analysis	Alignment of SSDS fragments to the genome and parsing of ss/dsDNA used the SSDS pipeline (Khil et al.; Genome Research 2012). Source code is available at https://github.com/kevbrick/SSDSpipeline. DSB hotspots were called and their strength calculated using the SSDS peak calling pipeline (Khil et al.; Genome Research 2012). Source code is available at https://github.com/kevbrick/callHotspotsSSDS. Algorithms used for ChIP-Seq analyses are specified below.

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 upon request. 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

The sequencing data reported in this paper are archived at the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) as accession no. GSE99921. A table of processed data is provided as a supplementary data file. This contains the data required to re-plot most figures. Data will be released upon publication with no restrictions.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-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. 9 individual male DSB maps were generated to estimate inter-individual variablity and technical noise in SSDS. Two female DSB maps were generated as replicates.
Data exclusions	No data were excluded from analyses.
Replication	The detailed protocols we use for DSB mapping and H3K4me3 ChIP-Seq have been used successfully by multiple research groups for DSB mapping in male mice. Applying these methods to female meiosis should be equally reproducible.
Randomization	Randomization was not employed in this study.
Blinding	Blinding was not employed in this study.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study	
☒	☐	Unique biological materials
☐	☒	Antibodies
☒	☐	Eukaryotic cell lines
☒	☐	Palaeontology
☐	☒	Animals and other organisms
☒	☐	Human research participants

Methods

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

Antibodies

Antibodies used	anti-DMC1 (Santa Cruz: c-20; sc-8973) anti-DMC1 (custom; polyclonal rabbit antibody to His-tagged full length human DMC1) anti-H3K4me3 (Millipore: #07-473) anti-H3K4me3 (Abcam 8580)
Validation	The c-20 anti-DMC1 antibody has been extensively validated in mice using both ChIP-Seq and immunofluorescence microscopy (Smagulova et al, Nature 2011). The custom anti-DMC1(human) polyclonal antibody was validated in mouse using ChIP-Seq in a male mouse. The hotspot signal correlated very well with SSDS using the established c-20 anti-DMC1 antibody (see manuscript). Both anti-H3K4me3 antibodies have been raised against peptides from human H3K4me3 that share extreme homology with

mouse H3K4me3. These have been extensively validated by the manufacturers (http://www.emdmillipore.com/US/en/product/Anti-trimethyl-Histone-H3-Lys4-Antibody,MM_NF-07-473# ; http://www.abcam.com/histone-h3-tri-methyl-k4-antibody-chip-grade-ab8580.html). Independent studies show that both antibodies recognize H3K4me3 from very distant species (fly, worm; http://compbio.med.harvard.edu/antibodies/targets/29). Furthermore, both antibodies yield a ChIP-Seq signal at meiosis-specific H3K4me3 sites specified by the binding of the PRDM9 protein.

Animals and other organisms

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

Laboratory animals	Both male and female wild-type adult mice were from C57Bl6/J background. Dnmt3L-/- adult mice (B6.129S6-Dnmt3ltm1Bes/J) were obtained from Jackson labs. Prdm9-/- adult mice (B6;129P2-Prdm9 tm1Ymat /J) were obtained from Jackson labs.
Wild animals	N/A
Field-collected samples	N/A

ChIP-seq

Data deposition

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

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

Data access links
May remain private before publication.

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE99921
Reviewer access token: wzgjyqaolzszfwv

Files in database submission

B6wt_30dpp_H3K4me3.fastq.gz
B6wt_30dpp_H3K4me3.peaks.bed
Dnmt3Lko_30dpp_H3K4me3.fastq.gz
Dnmt3Lko_30dpp_H3K4me3.peaks.bed
Dnmt3Lko_30dpp_input.fastq.gz
Ovary_H3K4me3.R1.fastq.gz
Ovary_H3K4me3.R2.fastq.gz
Ovary_H3K4me3.peaks.bedgraph.gz
Ovary_SSDS_O1.DSBhotspots.bedgraph
Ovary_SSDS_O1.R1.fastq.gz
Ovary_SSDS_O1.R2.fastq.gz
Ovary_SSDS_O1.ssDNA_type1.bam
Ovary_SSDS_O1.ssDNA_type1.bam.bai
Ovary_SSDS_O1.ssDNA_type1.bed
Ovary_SSDS_O2.DSBhotspots.bedgraph
Ovary_SSDS_O2.R1.fastq.gz
Ovary_SSDS_O2.R2.fastq.gz
Ovary_SSDS_O2.ssDNA_type1.bam
Ovary_SSDS_O2.ssDNA_type1.bam.bai
Ovary_SSDS_O2.ssDNA_type1.bed
Ovary_input_SSDS_O1.R1.fastq.gz
Ovary_input_SSDS_O1.R2.fastq.gz
Testis_DMC1_200k_N1.fastq.gz
Testis_SSDS_DN1.DSBhotspots.bedgraph
Testis_SSDS_DN1.R1.fastq.gz
Testis_SSDS_DN1.R2.fastq.gz
Testis_SSDS_DN1.ssDNA_type1.bam
Testis_SSDS_DN1.ssDNA_type1.bam.bai
Testis_SSDS_DN1.ssDNA_type1.bed
Testis_SSDS_Prdm9ko.DSBhotspots.bedgraph
Testis_SSDS_Prdm9ko.R1.fastq.gz
Testis_SSDS_Prdm9ko.R2.fastq.gz
Testis_SSDS_Prdm9ko.mm10.ssDNA_type1.bam
Testis_SSDS_Prdm9ko.mm10.ssDNA_type1.bam.bai
Testis_SSDS_Prdm9ko.mm10.ssDNA_type1.bed
Testis_SSDS_Hop2ko.R1.fastq.gz
Testis_SSDS_Hop2ko.R2.fastq.gz
Testis_SSDS_Hop2ko.mm10.ssDNA_type1.bam
Testis_SSDS_Hop2ko.mm10.ssDNA_type1.bam.bai
Testis_SSDS_Hop2ko.mm10.ssDNA_type1.bed
Testis_SSDS_T1.DSBhotspots.bedgraph
Testis_SSDS_T1.R1.fastq.gz
Testis_SSDS_T1.R2.fastq.gz
Testis_SSDS_T1.ssDNA_type1.bam

Testis_SSDS_T1.ssDNA_type1.bam.bai
Testis_SSDS_T1.ssDNA_type1.bed
Testis_SSDS_T2.DSBhotspots.bedgraph
Testis_SSDS_T2.R1.fastq.gz
Testis_SSDS_T2.R2.fastq.gz
Testis_SSDS_T2.ssDNA_type1.bam
Testis_SSDS_T2.ssDNA_type1.bam.bai
Testis_SSDS_T2.ssDNA_type1.bed
Testis_SSDS_T3.DSBhotspots.bedgraph
Testis_SSDS_T3.R1.fastq.gz
Testis_SSDS_T3.R2.fastq.gz
Testis_SSDS_T3.ssDNA_type1.bam
Testis_SSDS_T3.ssDNA_type1.bam.bai
Testis_SSDS_T3.ssDNA_type1.bed
Testis_SSDS_T4.DSBhotspots.bedgraph
Testis_SSDS_T4.R1.fastq.gz
Testis_SSDS_T4.R2.fastq.gz
Testis_SSDS_T4.ssDNA_type1.bam
Testis_SSDS_T4.ssDNA_type1.bam.bai
Testis_SSDS_T4.ssDNA_type1.bed
Testis_SSDS_T5.DSBhotspots.bedgraph
Testis_SSDS_T5.R1.fastq.gz
Testis_SSDS_T5.R2.fastq.gz
Testis_SSDS_T5.ssDNA_type1.bam
Testis_SSDS_T5.ssDNA_type1.bam.bai
Testis_SSDS_T5.ssDNA_type1.bed
Testis_SSDS_T6.DSBhotspots.bedgraph
Testis_SSDS_T6.R1.fastq.gz
Testis_SSDS_T6.R2.fastq.gz
Testis_SSDS_T6.ssDNA_type1.bam
Testis_SSDS_T6.ssDNA_type1.bam.bai
Testis_SSDS_T6.ssDNA_type1.bed
Testis_SSDS_T7.DSBhotspots.bedgraph
Testis_SSDS_T7.R1.fastq.gz
Testis_SSDS_T7.R2.fastq.gz
Testis_SSDS_T7.ssDNA_type1.bam
Testis_SSDS_T7.ssDNA_type1.bam.bai
Testis_SSDS_T7.ssDNA_type1.bed
Testis_SSDS_T8.DSBhotspots.bedgraph
Testis_SSDS_T8.R1.fastq.gz
Testis_SSDS_T8.R2.fastq.gz
Testis_SSDS_T8.ssDNA_type1.bam
Testis_SSDS_T8.ssDNA_type1.bam.bai
Testis_SSDS_T8.ssDNA_type1.bed
Testis_SSDS_T9.DSBhotspots.bedgraph
Testis_SSDS_T9.R1.fastq.gz
Testis_SSDS_T9.R2.fastq.gz
Testis_SSDS_T9.ssDNA_type1.bam
Testis_SSDS_T9.ssDNA_type1.bam.bai
Testis_SSDS_T9.ssDNA_type1.bed
Testis_input_SSDS.R1.fastq.gz
Testis_input_SSDS.R2.fastq.gz
Testis_input_SSDS.ssDNA_type1.bam
Testis_input_SSDS.ssDNA_type1.bam.bai
Testis_input_SSDS.ssDNA_type1.bed
Testis_H3K4me3_BSSeq.run1.mm10.R1.fastq.gz
Testis_H3K4me3_BSSeq.run1.mm10.R2.fastq.gz
Testis_H3K4me3_BSSeq.run2.mm10.R1.fastq.gz
Testis_H3K4me3_BSSeq.run2.mm10.R2.fastq.gz
Testis_H3K4me3_BSSeq.bismarkMethylationCalls.bedgraph.gz

Genome browser session
(e.g. UCSC)

Not provided

Methodology

Replicates

Male DSB maps: 9 replicates from individual males; Maximum reciprocal overlap between hotspots > 90%
Female DSB maps: 2 replicates from pools of 230 and 90 ovaries; Maximum reciprocal overlap between hotspots = 94%
H3K4me3 experiments; no replicates.

Sequencing depth

Sample Total_reads* Aligned_reads** Unique_aligned_reads*** Single/Paired-End
Dnmt3Lko_30dpp_input 33,036,289 29,394,228 26,801,138 Single-end
B6wt_30dpp_H3K4me3 147,013,281 135,945,673 101,088,245 Single-end
Ovary_H3K4me3 40,910,457 21,695,834 21,481,579 Paired-end
Dnmt3Lko_30dpp_H3K4me3 39,073,925 36,312,967 32,034,689 Single-end

Ovary_SSDS_O1 132,726,861 38,270,223 6,912,500 Paired-end(SSDS)
Ovary_SSDS_O2 184,101,322 66,594,381 14,562,130 Paired-end(SSDS)
Testis_SSDS_T1 173,499,622 41,448,730 19,858,588 Paired-end(SSDS)
Testis_SSDS_T2 133,369,383 40,035,209 16,197,173 Paired-end(SSDS)
Testis_SSDS_T3 103,583,213 40,287,172 14,690,253 Paired-end(SSDS)
Testis_SSDS_T4 121,949,056 33,525,473 8,505,069 Paired-end(SSDS)
Testis_SSDS_T5 65,228,101 25,756,338 7,871,727 Paired-end(SSDS)
Testis_SSDS_T6 83,861,621 45,871,302 24,957,563 Paired-end(SSDS)
Testis_SSDS_T7 111,171,893 49,394,237 21,290,033 Paired-end(SSDS)
Testis_SSDS_T8 183,443,278 32,640,277 11,999,412 Paired-end(SSDS)
Testis_SSDS_T9 26,452,649 7,936,293 6,161,585 Paired-end(SSDS)
Testis_SSDS_DN1 39,592,819 12,440,774 7,986,799 Paired-end(SSDS)
Testis_SSDS_Prdm9ko 253,731,930 67,111,851 49,802,403 Paired-end(SSDS)
Testis_SSDS_Hop2ko 71,946,748 2,464,767 2,092,085 Paired-end(SSDS)
Testis_input_SSDS 1,045,039,445 95,673,845 61,744,518 Paired-end(SSDS)
Ovary_input_SSDS_O1 122,141,174 49,622,950 34,083,176 Paired-end(SSDS)
Testis_DMC1_200k_N1 60,071,689 4,397,658 900,574 Single-end
Testis_DMC1_200k_N1 60,071,689 4,397,658 900,574 Single-end
Testis_H3K4me3_BSSeq 149,850,372 98,065,992 98,065,992 Paired-end
* For Paired-End samples, only the number of first end reads are shown
** For SSDS samples, the aligned reads are given as the number of aligned ssDNA-type1 fragments.
For other samples, this shows the number of aligned reads. For paired-end samples, it is the number of aligned first-end reads.
*** For SSDS samples, this shows the number of unique type-1 ssDNA fragments (determined from position and ITR structure; see Khil et al. Genome Res. 2012). For the H3K4me3_BSSeq sample, the total and uniquely aligning read counts are equal because the reads are deduplicated during the bismark alignment pipeline. For other samples, duplicates flagged by picard are removed.

Antibodies

anti-DMC1 (Santa Cruz: c-20; sc-8973)
anti-DMC1 (custom; polyclonal rabbit antibody to His-tagged full length human DMC1)
anti-H3K4me3 (Millipore: #07-473)
anti-H3K4me3 (Abcam 8580)

Peak calling parameters

For SSDS:
Uniquely mapping fragments unambiguously derived from ssDNA (ssDNA type 1) and having both reads with a mapping quality score ≥ 30 were used for identifying hotspot locations (peak calling). NCIS was used to estimate the background fraction for each library. Peak calling was performed using MACS (v.2.1.0.20150420) with the following parameters :

--ratio [output from NCIS]
-g mm
--bw 1000
--keep-dup all
--slocal 5000

For H3K4me3:
Uniquely mapping reads with a mapping quality score ≥ 30 were used for peak calling. NCIS was used to estimate the background fraction relative to an input DNA library. Peak calling was performed using MACS (v.2.1.0.20150420) with the following parameters :

--ratio [output from NCIS]
-g mm
--bw 1000
--keep-dup all
--slocal 5000.

Peak strength was subsequently calculated by subtracting the NCIS normalized input read count from the ChIP-Seq read count.

Data quality

For SSDS, we use a mixture-model-based approach that accounts for GC-biases to calculate a corrected p-value for each hotspot (model = negative binomial; num. iterations for refinement = 100) (Teng & Irizarry; Genome Research 2017). P-values were adjusted for multiple testing using the Benjamini-Hochberg method. Hotspots with a GC-corrected P-value > 0.05 and DSB hotspots within regions previously blacklisted (Smagulova, Brick et al.; Genes & Dev 2016) were discarded.

Software

Alignment of SSDS fragments to the genome and parsing of ss/dsDNA used the SSDS pipeline (Khil et al.; Genome Research 2012). Source code is available at https://github.com/kevbrick/SSDSpipeline. DSB hotspots were called and their strength calculated using the SSDS peak calling pipeline (Khil et al.; Genome Research 2012). Source code is available at https://github.com/kevbrick/callHotspotsSSDS. MAnorm was used to determine DSB hotspots with differential usage between males and females.

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	Fetal gonads were dissected from 15.5 dpc C57Bl/6J mice. Nuclei were prepared using the BITS-ChIP protocol (Bonn, Zinzen et al; Nature Genetics, 2012) and an antibody against a component of the meiotic synaptonemal complex, SCP3 (Santa Cruz: sc-74569).
Instrument	BD FACSAria Fusion
Software	BD FACS Diva 8.0.1, FlowCore package (R)
Cell population abundance	>90% of the post-sort fraction was validated as SCP3 positive using immunofluorescence microscopy.
Gating strategy	We first gated to retain only single nuclei (not shown). We then gated 4C nuclei using DAPI (92,000 < DAPI signal <170,000), and using an oocyte-specific marker, SCP3 (SCP signal >10,000) . We defined the SCP3 gate using an aliquot of our sample to which the primary antibody was not added (Secondary only). This estimates the background fluorescence from the secondary antibody. A figure describing the sorting is provided in supplementary information.

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