nature research

Corresponding author(s): *Double-blind peer review submissions: write DBPR and your manuscript number here instead of author names.*

# 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<br>*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<br>*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<br>*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 | No software was used for data collection. |
| Data analysis | R/Bioconductor was used to analyze Affymetrix data, RNA-seq and WGBS data was analyzed applying custom algorithms. Fiji 2.0.0 and Adobe Photoshop CS5 were used to process confocal images. FlowJo vX.0.7 was used to process flow cytometry data. GraphPad Prism 6 and R/Bioconductor was used to make the plots. |

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

Mouse gene expression microarray data that support the findings have been deposited in www.ebi.ac.uk/arrayexpress with accession code E-MTAB-6612. Gene

expression data that support the findings of this study have been deposited in www.ebi.ac.uk/ena with the accession code PRJEB23111. WGBS data that support the findings of this study have been deposited in www.ebi.ac.uk/arrayexpress with the accession code E-MTAB-6202 and E-MTAB-556140.
Figures 2a,d,f,g,i,j, 3a,c,e,f,h, 4d, Supplementary Figures 1c, 3c, 6a,c,e,f have associated source data. Figures 5a,e,f, 6a, Supplementary Figure 7a have associated data sets. Data generated in immunofluorescence and flow cytometry experiments are available upon reasonable request.

# 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

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 determine the samples sizes. Samples sizes were determined by the amount of biological material available. For each experiment three or more biological replicates are presented. Sample sizes were sufficient where a marked contrast between the control and experimental samples was observed together with low variance between different replicates. |
| Data exclusions | Testicular weight of one testis from one experimental animal was excluded because tissue looked abnormal and atrophic and clearly different from the second testis in pair. Few data points from flow cytometry related experiments that showed sample preparation defects were excluded from final analysis. |
| Replication | Number of biological replicates and number of attempts to collect all the data is described in figure legends and Methods section. All attempts at replication were successful and all biological/experimental individual replicates were consistent. |
| Randomization | No randomization was employed in this study. In this study all samples/animals of precise genotype were analyzed and differences between different genotype groups were compared. |
| Blinding | No blind experiments were carried out in this study. All starting experimental material had a precise genotype assigned. |

## Materials & experimental systems

Policy information about availability of materials

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Unique materials |
| ☐ | ☒ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☐ | ☒ Research animals |
| ☒ | ☐ Human research participants |

### Antibodies

| | |
|---|---|
| Antibodies used | All suppliers and catalog numbers of antibodies used in this study are provided in section "Methods", paragraphs "Whole mount immunofluorescence on seminiferous tubules", "Spermatogonial cell isolation" and "Immunofluorescence". |
| Validation | Antibody validation was done by suppliers. Additionally, immunofluorescence antibodies were validated using wild-type controls where appropriate (eg. anti-LINE1 Orf1p, anti-IAP Gag, anti-V5, anti-HA antibodies) or by identifying a target population by specific nuclear morphology and position within seminiferous epithelium (eg. anti-PLZF, anti-c-Kit antibodies). FACS antibodies were additionally validated by using appropriate single-stain and FMO controls. |

### Research animals

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

| | |
|---|---|
| Animals/animal-derived materials | Described in section "Methods", paragraph "Mouse strains". |

# Method-specific reporting

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☐ | ☒ Flow cytometry |
| ☒ | ☐ Magnetic resonance imaging |

## 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 | Described in section "Methods", paragraph "Spermatogonial cell isolation". |
| Instrument | Described in section "Methods", paragraph "Spermatogonial cell isolation". |
| Software | Described in section "Methods", paragraph "Spermatogonial cell isolation". |
| Cell population abundance | Purity of the samples was 98%, which was determined by running sorted cells through the sorting procedure again and recording the percentage of a target population. |
| Gating strategy | Described in section "Methods", paragraph "Spermatogonial cell isolation" and shown in Supplementary Figure XXXX. |

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