nature research

Corresponding author(s): Frederic Berger

Last updated by author(s): Mar 27, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	TrimGalore v0.4.1 STAR aligner v2.5.2a Bowtie2 v2.1.0 SAMtools v1.3 Picard tools v1.141 deepTools v2.5.0.1 MACS2 v2.1.0 Kallisto version 0.43.1 Zeiss ZEN 2012 SP1 (black edition)
Data analysis	pheatmap R package v1.0.12 DESeq2 v1.22.2 GeneOverlap R package v1.18.0 EnrichedHeatmap R package v1.12.0 Preseq v2.0 ChIPpeakAnno R package v3.16.1 gProfileR R package v0.6.7 ImageJ v2.0.0-rc-69/1.52p

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The deep sequencing data generated in this study has been deposited at the Gene Expression Omnibus (GSE120669).
Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120669

Egg cell transcriptomic data has been deposited at the DNA Data Bank of Japan (BioProject: PRJDB8211).
Go to https://ddbj.nig.ac.jp/DRASearch/query?keyword=PRJDB8211&show=20

Previously published RNA-seq and ChIP-seq datasets re-analysed in this study are detailed in Supplementary Table 6. The resulting normalized TPM data and bigwig files has been deposited at the Gene Expression Omnibus (GSE120669).

All statistical source data and unprocessed gel and blot images have been submitted. All other data supporting the findings in this study are available from the corresponding author on reasonable request.

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	For quantification of H3K27me1 and H3K27me3 levels in sperm nuclei, a pooled pollen population was isolated for each genotype from multiple plants. The resulting nuclei from this pooled pollen population was processed and images of at least 40 nuclei were taken for each genotype. No statistical method was used to predetermine the sample size. The sample size was based on similar studies in field and the high sample size should meet the criteria for sound statistical results.
Data exclusions	a) RNA-seq: As detailed in the methods, for differential gene expression analysis of mutant transcriptomes, only transcripts that had 10 counts or more in at least one sample were included such that only genes with detectable expression were analysed. Such quality control steps are customary in the field and were pre-established. b) ChIP-seq: As detailed in the methods, duplicate reads were removed while reads with a poor mapping score (mapQ<10) were removed prior to merging replicates and performing downstream analyses. Such quality control steps are customary in the field and were pre-established.
Replication	a) RNA-seq: Three or more replicates were used for RNA-seq analysis of mutant pollen, as detailed in the legend for Extended Data Figure 5a. b) ChIP-seq: Two or more replicates were used for ChIP-seq analysis of each histone mark. c) Microscopy & blots: Immunostaining, marker line analysis and western blots were performed at least twice with reproducible results. This is detailed in each relevant figure legend. Histone methyltransferase assays were performed once for ATXR5/6 and twice for PRC2.
Randomization	Not relevant - no treatment groups.
Blinding	Not relevant - no treatment groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used

Antibody, Clonality, Manufacturer, Catalogue number, Dilutions used

anti-H3, mouse monoclonal, Abcam, ab1791, 1:500 for immunostaining
anti-H3K4me3, rabbit polyclonal, Abcam, ab8580, 1:1000 for western blot
anti-H3K27ac, rabbit polyclonal, Abcam, ab4729, 1:100 for immunostaining, 1:1000 for western blot
anti-H3K27me1, rabbit polyclonal, Millipore, 17-643, 1:100 for immunostaining, 1:1000 for western blot
anti-H3K27me3, rabbit polyclonal, Millipore, 07-449, 1:100 for immunostaining, 1:1000 for western blot
anti-HA, rat monoclonal, Roche, 11867423001, 1:2000 for western blot

anti-mouse IgG-AlexaFluor488 conjugate, goat polyclonal, Invitrogen, A-11034, 1:500 for immunostaining
anti-rabbit IgG-AlexaFluor555 conjugate, goat polyclonal, Invitrogen, A-21422, 1:500 for immunostaining

anti-rabbit IgG-HRP conjugate, goat polyclonal, Bio-Rad, 170-6515, 1:10000 for western blot
anti-rat IgG-HRP conjugate, rabbit polyclonal, Sigma, A5795, 1:10000 for western blot

Validation

Both anti-H3K27me1 (Millipore, 17-643) and anti-H3K27me3 (Millipore, 07-449) antibodies were tested in a dot blot assay to assess cross reactivity with methylated histone H3.1, H3.3 and H3.10 peptides. This analysis is shown in Extended Data Figure 2.

anti-H3 (Abcam, ab179) is validated for use in Arabidopsis thaliana and for immunostaining on the manufacturer's website.

anti-H3K4me3 (Abcam, ab8580) is validated for use in Arabidopsis thaliana and for ChIP-seq and western blot on the manufacturer's website.

anti-H3K27ac (Abcam, ab4729) is validated for use in Arabidopsis thaliana and for ChIP-seq, western blot and immunostaining on the manufacturer's website.

anti-H3K27me1 (Millipore, 17-643) is validated for ChIP-seq, western blot and immunostaining on the manufacturer's website. See Extended Data Figure 2 for validation in Arabidopsis thaliana.

anti-H3K27me3 (Millipore, 07-449) is validated for ChIP-seq, western blot and immunostaining on the manufacturer's website. See Extended Data Figure 2 for validation in Arabidopsis thaliana.

anti-GFP (Thermo Scientific, A-11122) is validated for ChIP-seq and western blot on the manufacturer's website.

ChIP-seq

Data deposition

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

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

Data access links
May remain private before publication.

To view GEO accession GSE120669:
Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120669

Files in database submission

GSM3407988 Input_A_rep1
GSM3407989 Input_A_rep2
GSM3407990 Input_A_rep3
GSM3407991 H3K27ac_rep1
GSM3407992 H3K27ac_rep2
GSM3407993 H3K27ac_rep3
GSM3407994 H3K27me1_rep1
GSM3407995 H3K27me1_rep2
GSM3407996 Input_B_rep1
GSM3407997 Input_B_rep2
GSM3407998 Input_B_rep3
GSM3407999 H3.10_rep1

GSM3408000 H3.10_rep2
GSM3408001 H3.10_rep3
GSM3408002 H3K27me3_rep1
GSM3408003 H3K27me3_rep2
GSM3408004 H3K4me3_rep1
GSM3408005 H3K4me3_rep2
GSM3408006 H3K4me3_rep3
GSM4300539 Input_htr10_rep1
GSM4300540 Input_htr10_rep2
GSM4300541 Input_htr10_rep3
GSM4300542 H3K27me1_htr10_rep1
GSM4300543 H3K27me1_htr10_rep2
GSM4300544 H3K27me1_htr10_rep3
GSM4300545 H3K27me3_htr10_rep1
GSM4300546 H3K27me3_htr10_rep2
GSM4300547 H3K27me3_htr10_rep3

Genome browser session
(e.g. UCSC)

No longer applicable

Methodology

Replicates

All ChIP-seq datatsets had 2 or more biological replicates.

Sequencing depth

All samples were sequenced paired-end 50 with depth ranging between ~8-40 million reads. Alignment rate ranged between 10-90%, with the total number of reads resulting in >3x coverage of the Arabidopsis genome. Mapping statistics are detailed below:

Sample Total Aligned Unique.mapQ>10 Run.Type Length
Rep1_H3K27Ac_SN 31175898 28143419 19758937 PE50 1407170950
Rep2_H3K27Ac_SN 27999923 25817265 18154265 PE50 1290863250
Rep3_H3K27Ac_SN 28070563 26477778 17016673 PE50 1323888900
Rep1_H3K27me1_SN 27290397 21930981 6661114 PE50 1096549050
Rep3_H3K27me1_SN 27987690 13965774 5610987 PE50 698288700
Rep1_Input_SN 9681462 9144296 4995492 PE50 457214800
Rep2_Input_SN 9956654 9037727 5343050 PE50 451886350
Rep3_Input_SN 9568530 9117092 4667164 PE50 455854600
Rep1_Input 8167655 7837330 6173539 PE50 391866500
Rep2_Input 7604002 6759853 5243518 PE50 337992650
Rep3_Input 7647152 6903507 5325469 PE50 345175350
Rep1_GFP 42700518 18978776 7312229 PE50 948938800
Rep2_GFP 42316848 30511855 7074720 PE50 1525592750
Rep3_GFP 43161333 26613706 5153101 PE50 1330685300
Rep2_H3K27me3 46612478 3383699 337383 PE50 169184950
Rep3_H3K27me3 45306687 4337847 509360 PE50 216892350
Rep1_H3K4me3 44105345 30369320 8462007 PE50 1518466000
Rep2_H3K4me3 44298328 30893269 8185304 PE50 1544663450
Rep3_H3K4me3 49646310 38200481 5298260 PE50 1910024050
Rep1_htr10_Input 18615458 16000364 2237983 PE75 1200027300
Rep2_htr10_Input 15937698 15229023 4426464 PE75 1142176725
Rep3_htr10_Input 13774187 11785966 1750123 PE75 883947450
Rep1_H3K27me1_htr10 15622696 13986499 9669024 PE75 1048987425
Rep2_H3K27me1_htr10 16140561 14991041 10135381 PE75 1124328075
Rep3_H3K27me1_htr10 17504709 8630613 5023772 PE75 647295975
Rep1_H3K27me3_htr10 14192324 7824067 5265557 PE75 586805025
Rep2_H3K27me3_htr10 11469439 7531574 4782677 PE75 564868050
Rep3_H3K27me3_htr10 15556148 5278243 2131720 PE75 395868225

Antibodies

Antibody, Clonality, Manufacturer, Catalogue number, Amount used

anti-H3K4me3, rabbit polyclonal, Abcam, ab8580, 1 µg for ChIPseq
anti-H3K27ac, rabbit polyclonal, Abcam, ab4729, 1 µg for ChIPseq
anti-H3K27me1, rabbit polyclonal, Millipore, 17-643, 1 µg for ChIPseq
anti-H3K27me3, rabbit polyclonal, Millipore, 07-449, 2 µl serum for ChIPseq
anti-GFP, rabbit polyclonal, Thermo Scientific, A-11122), 1 µg for ChIPseq

Peak calling parameters

Read were mapped using the TAIR10 genome with Bowtie2 version 2.1.0.
Reads were filtered for a MAPQ score > 10 using SAMtools version 1.3.
Duplicates were filtered out using Picard tools MarkDuplicates version 1.141.
Broad peaks were called using MACS2 version 2.1.0 callpeak function -f BAMPE --broad --broad-cutoff 0.1 -g 1.2e8.
Narrow peaks were called using MACS2 version 2.1.0 callpeak function -f BAMPE -q 0.1 -g 1.2e8.

Data quality

See Extended Data Figure 3 for validation of the ChIP-seq data. Unique reads filtered for quality (mapQ >10) from each biological replicate were merged for downstream analysis after confirming high correlation among replicates, resulting in at least x6 coverage of the Arabidopsis genome per group of replicates (except for H3K27me3 in sperm, see below).

Consistent with its erasure from sperm chromatin, library complexity for sperm H3K27me3 ChIP-seq was poor compared to other histone marks. Nonetheless, the sperm H3K27me3 replicates were highly correlated (Extended Data Fig. 3i) and confirmed to have been sequenced to a saturating depth (Extended Data Fig. 3h) using Preseq v2.0. Read depth discrepancy was accounted for prior to comparing the number of peaks between sperm and somatic tissues by subsampling to the same read depth using SAMtools version 1.3.

Sperm H3K27me1 narrow peaks FDR 1% and >3-fold enrichment: 1091
Sperm H3K27me3 narrow peaks FDR 1% and >2-fold enrichment: 478
Sperm H3K4me3 narrow peaks FDR 1% and >3-fold enrichment: 14571
Sperm H3K27ac narrow peaks FDR 1% and >2-fold enrichment: 3450
htr10 sperm H3K27me1 narrow peaks FDR 1% and >2-fold enrichment: 4193
htr10 sperm H3K27me3 narrow peaks FDR 1% and >2-fold enrichment: 5078

Software

TrimGalore v0.4.1
Bowtie2 v2.1.0
SAMtools v1.3
Picard tools v1.141
deepTools v2.5.0.1
MACS2 v2.1.0
GeneOverlap R package v1.18.0
EnrichedHeatmap R package v1.12.0
Preseq v2.0
ChIPpeakAnno R package v3.16.1