Science
AAAS

Supplementary Materials for

The histone H3.1 variant regulates TONSOKU-mediated DNA repair during replication

Hossein Davarinejad[1]†, Yi-Chun Huang[2]†, Benoit Mermaz[2], Chantal LeBlanc[2], Axel Poulet[2], Geoffrey Thomson[2], Valentin Joly[2], Marcelo Muñoz[1], Alexis Arvanitis-Vigneault[1], Devisree Valsakumar[3,4], Gonzalo Villarino[2], Alex Ross[5,6], Benjamin H. Rotstein[6,7], Emilio I. Alarcon[5,6], Joseph S. Brunzelle[8], Philipp Voigt[3,4], Jie Dong[2,9], Jean-François Couture[1*] and Yannick Jacob[2*]

Correspondence to: E-mails: yannick.jacob@yale.edu and jean-francois.couture@uottawa.ca

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Materials and Methods

Plant materials

A. thaliana plants were grown under cool-white fluorescent lights (~100 μmol m^{-2} s^{-1}) in long-day conditions (16 h light/8 h dark). The *atxr5/6* double mutant was described previously (*17*). *tsk/bru1-4* (*At3g18730*, SALK_034207 (*29*)), *ku70-2* (*At1g16970*, SALK_123114c (*30*)), *ku80-7* (*At1g48050*, SALK_112921 (*30*)), *lig4-4* (*At5g57160*, SALK_044027 (*31*)), *rad17-2* (*At5g66130*, SALK_009384 (*32*)), *brca2a* (*At4g00020*, 13F-1 allele (*33*)), *brca2b* (*At5g01630*, SALK_037617 (*33*)), *rad54-1* (*At3g19210*, SALK_038057 (*34*)), and *teb-5* (*At4g32700*, SALK_018851 (*35*)) are in the Col-0 genetic background. They were obtained from the Arabidopsis Biological Resource Center (Columbus, Ohio), and described in previous publications. CRISPR/Cas9 was used to mutate *RAD51* (*At5g20850*) in Col-0 and in the *atxr5/6* mutant background. The *h3.1* quadruple mutant (*htr1 htr2 htr3 htr9*) used for transformation of the *H3.1* transgenes (WT and point mutants) was described previously (*16*). The *h3.1* pentuple mutant (*htr1 htr2 htr3 htr9 htr13*) was created by performing temperature-optimized CRISPR/Cas9 at the *HTR1*, *HTR2* and *HTR13* genes in the *h3.1* quadruple mutant (*36*). *htr1* and *htr2* have T-DNA insertions just outside the coding sequence in the *h3.1* quadruple mutant (*16*). Transgenic plants expressing H3.1 WT or H3.1A31T (in the *h3.1* quadruple mutant) used in the MMS genotoxic assay were described previously (*16*). The septuple mutant *atxr5 atxr6 htr1 htr2 htr3 htr9 htr13* was generated by crossing the *atxr5/6* double mutant with the *h3.1* pentuple mutant, followed by multiplex CRISPR/Cas9 editing at *HTR9* and *HTR13*.

Plasmid constructs

The coding sequences of the TPR domains of *A. thaliana* TSK (AtTSK; a.a. 1-525 followed by a stop codon) and of mouse TONSL (a.a. 1-515 followed by a stop codon) were cloned into the pET32a vector using BamHI and SalI, yielding pET32a-TSK and pET32a-TONSL, respectively. For the *C. unishu* TSK (CuTSK) construct, a.a. sequences ENLYFQG (TEV cleavage site) followed by CuTSK a.a. 1-530 (accession ID: GAY58445.1) were cloned into the pET22a(+) vector using BamHI and XhoI sites (GeneScript, Piscataway, NJ). CuTSK 1-490 construct was generated by placing a stop codon after residue 490 using site-directed mutagenesis (QuickChange, Agilent Technologies, Santa Clara, CA).

The coding sequences of the N-terminal tails of *A. thaliana* H2A.Z (a.a. 1-16), H2A.X (a.a. 1-16), H2B (a.a. 1-35), H3.1 (a.a. 1-58), H3.3 (a.a. 1-58), and H4 (a.a. 1-30) were fused with a C-terminal GST tag by cloning into pET28-Mff(1-61)-PP-GST (Addgene plasmid #73042; gifted by D. Chan) using the NdeI and BamHI sites (*37*). The H3.1A31T and H3.1F41Y mutations were engineered by site-directed mutagenesis (QuikChange II XL, Agilent Technologies).

HTR1 (H3.1, *At5g65360*) and its promoter (1167 bp upstream of the start codon) were cloned into pENTR/D-TOPO (Thermo Fisher Scientific, Waltham, MA), subcloned using Gateway Technology into pB7WG (*38*), and modified using site-directed mutagenesis to generate the following H3.1 point mutant constructs: H3.1, H3.1S28A, H3.1K4A, H3.1S28A K4A, H3.1K9A, H3.1S28A K9A, H3.1K36A, H3.1S28A K36A, H3.1A31T and H3.1S28A A31T.

Protein expression

A Rosetta (DE3) *E. coli* strain (#70954, Sigma, St. Louis, MO) was used for the expression of the following recombinant proteins: AtTSK, mouse TONSL and the histone-GST fusion proteins. The bacteria were cultured in LB, and 1 mM IPTG was used to induce protein expression. For Selenium Methionine (SelMet)-CuTSK, the plasmids were transformed into B843 *E. coli* and grown in M9 minimal medium supplemented with SelMet (Complete kit MD12-500, Molecular Dimensions, Holland, OH). CuTSK constructs were expressed at 18°C and induced with IPTG (0.2 mM) at $OD_{600nm} = 0.6$.

For purification of AtTSK and TONSL (containing an N-terminal Trx-His-S tag), the cell pellets were resuspended in NPI-10 buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM Imidazole, pH 8) containing 1 mM PMSF and sonicated. After centrifugation, the supernatant was passed through a Ni-NTA agarose column. The column was then washed with NPI-20 buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8), and the recombinant proteins were eluted with NPI-250 buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8). The eluted proteins were further purified by size exclusion chromatography (SEC), aliquoted, and stored at -80°C.

For purification of the histone-GST fusion proteins, the cell pellets were resuspended in 1X PBS (137 mM NaCl, 10 mM phosphate, 2.7 mM KCl, pH 7.4) containing 1 mM PMSF before sonification and centrifugation. The supernatant was passed through a Glutathione Sepharose 4B column, and bound proteins were washed with 1X PBS and eluted using EB buffer (50 mM Tris, 50 mM NaCl, 30 mM reduced L-Glutathione, 10% glycerol, pH 8.0). Proteins were aliquoted and stored at -80°C.

For purification of CuTSK, cell pellets were resuspended in NaP_i buffer (50 mM NaP_i pH 7.5, 1 M NaCl, 10% glycerol, 5 mM βME) and purified at 4°C using cobalt resin (Talon) (TaKaRa, Japan). Proteins were TEV-cleaved on beads and purified using SEC (Superdex75, GE Healthcare, Chicago, IL) columns equilibrated with NaP_i buffer (250 mM NaCl for pulldowns; 350 mM NaCl for ITC) or Tris buffer (20 mM Tris pH 7.5, 200 mM NaCl, 5% glycerol, 5 mM βME) for crystallography.

In vitro binding assays

For the binding assays involving AtTSK and the GST-tagged histone N-terminal tail proteins, 2 µg of AtTSK was mixed with 2 µg of GST or GST-tagged histone proteins in 400 µl of binding buffer (25 mM Tris, 250 mM NaCl, 0.05% NP-40, pH 8.0), and the mixture was incubated at 4°C overnight. 15 µl pre-washed Glutathione Sepharose 4B agarose beads were added to each tube and incubated for 30 min to pull down the GST-tagged histone proteins. The beads were washed four times with 1 mL of binding buffer, with each wash performed by rotating for 5 min at 4°C. After the final wash, 15 µl of 2X SDS loading buffer was added to each tube, and the proteins were eluted by boiling at 95°C for 5 minutes. The samples were separated on a 10% SDS-PAGE gel. The lower part of the gel was subjected to Coomassie staining to visualize the GST or GST-tagged histone N-terminal tail proteins, and the upper part of the gel was subjected to Western blot analyses using an anti-His antibody (H1029) (Sigma). Each pulldown assay was performed at least three times.

For the binding assays using TONSL and the biotin-tagged full-length histone proteins (1-135 aa, Active Motif, Carlsbad, CA), 1 μg of TONSL was mixed with 1 μg of biotin or biotin-tagged histone proteins in 400 μl of binding buffer (25 mM Tris, 450 mM NaCl, 0.05% NP-40, pH 8.0), and incubated at 4°C for 30 min. Pre-washed MyOne Streptavidin beads (20 μl) (Invitrogen, Waltham, MA) were added to each tube and incubated for 30 min to pull down the biotin-tagged histone proteins. The beads were washed four times with 1 mL of binding buffer for 5 min at 4°C. After the final wash, 15 μl of 2× SDS loading buffer was added to each tube, and the proteins were eluted by boiling at 95°C for 5 min. The samples were separated on a Bio-Rad 4–20% Mini-PROTEAN® TGX™ Precast Protein Gel. The lower part of the gel was subjected to Coomassie staining to visualize the biotin or biotin-tagged histone proteins, and the upper part of the gel was subjected to Western blot analyses using an anti-His antibody (H1029) (Sigma). Pulldown assays were performed three times.

For binding assays involving CuTSK and biotinylated H3.1 peptides, 50 μL of streptavidin agarose (Millipore, Burlington, MA, #69203-3) resin was washed with pulldown buffer (200 mM NaCl, 150 mM NaP_i pH 7.5, 10% glycerol, 5 mM βME) and saturated with $H3.1_{(1\text{-}45)}$ peptides (unmodified/modified) while incubating at 4°C for 30 min. Peptide-bound resin was incubated with 50 μg $CuTSK_{(1\text{-}530)}$ for 30 min in 200 μL of pulldown buffer, washed, boiled, and loaded onto SDS 4-20% acrylamide gels. Pulldown assays were performed three times with two separate purifications of CuTSK.

For nucleosome pulldown assays, histone octamers were reconstituted using *Xenopus laevis* H2A, H2B, H4, and *A. thaliana* H3.1 or H3.3 and purified by gel filtration on a S200 size exclusion column (GE Healthcare). A biotinylated 209-bp DNA fragment containing the 601 nucleosome positioning sequence was generated by PCR with a biotinylated forward primer and purified by ion exchange chromatography on a HiTrap Q column followed by ethanol precipitation. Mononucleosomes were assembled from histone octamers and 601 DNA by gradient dialysis as described previously (*39*). Nucleosome assembly was verified by native gel electrophoresis on 6% acrylamide gels in 0.5× TGE buffer (12.5 mM Tris pH 8.0, 95 mM glycine, 0.5 mM EDTA). To carry out the binding assays, all incubation and wash steps were performed at 4°C with end-over-end rotation. Centrifugation of beads before washes was done at 1,500 *g* for 2 min at 4°C. 3 μg (23 pmol) of assembled nucleosomes were immobilized on streptavidin sepharose high performance beads (GE Healthcare) that were blocked with 1 mg/ml BSA in pulldown buffer (20 mM HEPES pH 7.9, 175 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM DTT, 0.1% NP-40, 0.1 mg/ml BSA) by overnight incubation. After three 5-min washes with pulldown buffer, nucleosome-bound beads were incubated with TPR_{TSK}-GST for 2 h. Beads were then washed 5 times for 10 min each with high salt pulldown buffer (as above but with 350 mM NaCl) before elution of bound proteins by boiling in 1.5× SDS sample buffer (95 mM Tris HCl pH 6.8, 15% glycerol, 3% SDS, 75 mM DTT, 0.15% bromophenol blue). Protein binding was analysed by Western Blotting with anti-GST-HRP antibody (ab3416, Abcam, Waltham, MA). Nucleosome loading was confirmed by Western Blotting with anti-H3 antibody (ab176842, Abcam). The nucleosome pulldown assays were repeated three times, using three different preparations of nucleosomes.

ITC assay

ITC experiments were performed using a VP-ITC calorimeter (MicroCal, Northampton, MA) by injecting peptides (750 μM) into a solution of CuTSK (50 μM) in 50 mM NaP_i pH 7.5, 350 mM NaCl, 10% glycerol, and 5 mM βME. The experiment was performed at 19°C, and the titration data were analyzed using Origin software (OriginLab Corporation, Northampton, MA). ITC experiments were replicated three times using two separate batches of synthesized peptides and three separate purifications of CuTSK.

Histone H3 peptide synthesis

Fmoc-protected amino acids and rink amide low loading resin were purchased from CEM (Matthews, NC). Fmoc-Lysine(Boc)(Me)-OH, Fmoc-Lysine$(Me)_2$-OH, and Fmoc-Lysine$(Me)_3$-OH building blocks were purchased from Bachem (Bubendorf, Switzerland). All peptides were synthesized using microwave-assisted Fmoc solid phase peptide synthesis in a Liberty Blue automated system. Briefly, the required amount of resin was swelled in DMF for 5 min. Next, Fmoc deprotection was carried out with 20% piperidine at 90°C for 60s. Standard coupling cycles using DIC/Oxyma Pure were run at 90°C for 240s in each amino acid. Peptides were cleaved from the resin and deprotected with TFA/TIS/EDT/H_2O (92.5/2.5/2.5/2.5 % v/v) at 42°C for 30 min, and then precipitated in -20°C diethyl ether. Peptide crude products were then dried under vacuum overnight and purified by RP-HLPC in a Waters 1525EF semi-preparative system with a 21.6 × 250 mm C18 column at 20 mL/min. Peptide purity and identity were confirmed via RP-UPLC-UV/MS in a Waters Acquity UPLC Xevo TQD using a 2.1 × 100 mm UPLC BEH C8 column. A purity of >95% was determined through HPLC peak analysis. The molecular ions found for each peptide are described in Table S3. Lyophilized peptides were resuspended in water.

Crystallography

$CuTSK_{(1\text{-}490)}$ (20 mg/ml) was incubated with $H3.1_{(1\text{-}45)}$ (5:1 peptide:TSK molar ratio) and crystalized in 25% 1,2-propanediol, 5% glycerol, 0.1 M Na/K phosphate pH 6.0. A single-wavelength anomalous dispersion (SAD) data set was collected at the 21-ID-D beamline of the Life Science-Collaborative Access Team at the Advanced Photon Source Synchrotron. The structure of $CuTSK_{(1\text{-}490)}$ was determined by SAD at the selenium peak wavelength. The reflections were processed and scaled using HKL2000 (*40*) and 23 selenium atom were identified and refined using the SHELX C/D programs (*41*). Phases were calculated using SHELX-E and the Arp/Warp program was used to generate the initial model. One chain was traced and used as a search model for molecular replacement of two TSK chains in the asymmetric unit using Phaser. Missing residues were modeled in the calculated phases using Coot (*42*) and the structure was further refined using phenix.refine (*43*). Clash scores were determined by MolProbity (*44*). The final model includes 17 selenium atoms, one water molecule, CuTSK residues 1-124, 128-150 and 159-483, and H3.1 residues 4-9 and 18-40. Missing residues were not modeled due to lack of electron density.

BiFC and confocal microscopy

The TPR domain (a.a. 1-524) of TSK with a nuclear localization signal (NLS) was cloned into the Gateway destination vector pUC-DEST-VYCE®GW. Histones H3.1 and H3.3 (a.a.1-136) were

cloned into the Gateway destination vector pUC-DEST-VYNE®GW (*45*). pSAT6-mCherry-VirD2NLS was used as a nuclear marker. The protoplasts were isolated from 3- to 4-week-old *A. thaliana* plants (*atxr5/6)* and transfected following the tape-Arabidopsis sandwich method (*46*). After 14–18 h incubation in low-light conditions, protoplast images were acquired using a confocal spinning disk unit (Yokogawa CSU-W1), mounted on a Nikon Eclipse Ti2 microscope body (Nikon, Minato City, Tokyo, Japan). A 60× water objective (N.A. = 1.2) and a 1.5× post magnification along with 514 nm and 561 nm lasers were used for imaging as described (*47*). The images were processed with FIJI (*48*). Assays were repeated three times with similar results.

Plant nuclei microscopy

Leaves from four-week-old plants were fixed in 3.7% formaldehyde in cold Tris buffer (10 mM Tris-HCl pH 7.5, 10 mM NaEDTA, 100 mM NaCl) for 20 min, then washed for 10 min in Tris buffer. The leaves were finely chopped with a razor blade in 500 µl LB01 buffer (15 mM Tris-HCl pH7.5, 2 mM NaEDTA, 0.5 mM spermine-4HCl, 80 mM KCl, 20 mM NaCl and 0.1% Triton X-100) and filtered through a 30 µm mesh (Sysmex Partec, Gorlitz, Germany). 10 µl of lysate was mixed to 10 µl of sorting buffer (100 mM Tris-HCl pH 7.5, 50 mM KCl, 2mM $MgCl_2$, 0.05% Tween-20 and 5% sucrose) and spread onto a coverslip. After drying for 30 min, cold methanol was added onto each coverslip for 3 min. Methanol was removed and TBS-Tx (20 mM Tris pH 7.5, 100 mM NaCl, 0.1% Triton X-100) was added for 5 min. The coverslips were mounted onto slides with Vectashield mounting medium containing DAPI (Vector Laboratories, Burlingame, CA). Imaging was done using a Nikon Eclipse Ni-E microscope with a 100× CFI PlanApo Lamda objective (Nikon) and equipped with an Andor Clara camera. Z-series optical sections of each nucleus were obtained at 0.3-µm steps. Images were deconvolved by ImageJ using the deconvolution plugin. Three biological samples per genotype were assessed for each experiment. Twenty-five nuclei were analyzed for each sample.

RT-qPCR

RNA extraction from three-week-old leaf tissue was performed using TRIzol (Invitrogen). RNA samples were treated with RQ1 RNase-free DNase (Promega, Madison, WI) at 37°C for 30 min. 1 µg of total RNA was used to produce cDNA with SuperScript II Reverse Transcriptase (Invitrogen) using oligo-dT primers. Real-time PCR was done using a CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with KAPA SYBR FAST qPCR Master Mix (2×) Kit (Kapa Biosystems, Wilmington, MA). Relative quantities were determined by the C_t method (*49*) with *ACTIN* as the normalizer. At least three biological samples were used for each experiment.

Primer name	Sequence
ACTIN-F	TCGTGGTGGTGAGTTTGTTAC
ACTIN-R	CAGCATCATCACAAGCATCC
TSI-F	ATCCAGTCCGAAGAACGCGAACTA
TSI-R	TCACTTGTGAGTGTTCGTGAGGTC
BRCA1-F	CATGTGCCTTTTGTCAGTGTTC
BRCA1-R	TGGAGCCCATTCAGCACAGTTT
H3.1 transgene-F	GCAGCGCCGTCGCAGCACTTCAGG
H3.1 transgene-R	ACTCTAGCATGGCCGCGGGATATC

Flow cytometry

To generate flow cytometry profiles, rosette leaves from three-week-old plants were finely chopped in 0.5 ml Galbraith buffer (45 mM $MgCl_2$, 20 mM MOPS, 30 mM sodium citrate, 0.1% Triton X-100, 40 µg/µl RNase A) and filtered through a 30 µm mesh (Sysmex Partec). Nuclei were stained by adding 20 µg/ml propidium iodide (Sigma) to each sample, followed by vortexing. The samples were analyzed using a BD FACS LSR Fortessa X20 or BD FACSAria II sorter (Becton Dickinson, Franklin Lakes, NJ). FlowJo 10.0.6 (Tree Star, Ashland, Oregon) was used to generate profiles and for quantification (nuclei counts and rCV values). Each biological replicate consisted of a leaf from one plant. To sort 16C nuclei for DNA sequencing, samples were prepared by chopping rosette leaves from four-week-old plants as described above. 100,000 nuclei for each sample (two biological replicates per genotype) were sorted using a BD FACSAria II sorter with a 100-µm nozzle.

DNA extraction and sequencing

Genomic DNA was extracted from sorted 16C nuclei using the Arcturus™ PicoPure™ DNA Extraction Kit (ThermoFisher Scientific, Waltham, MA). Samples were incubated at 65°C overnight, and then at 95°C for 10 min. The DNA samples were purified using a genomic DNA Clean and Concentrator kit (Zymo Research, Irvine, CA). Sequencing libraries were generated at the Yale Center for Genome Analysis (YCGA) using the xGen Prism DNA library prep kit for NGS (Integrated DNA Technologies, Coralville, IA). Sequencing was performed on an Illumina NovaSeq 6000 using the S4 XP workflow (Illumina, San Diego, CA). Paired-end reads were filtered and trimmed using fastp (version 0.21.0 with default parameters) (*50*). Reads with quality inferior to 20 were removed from the datasets (Table S4). Sequencing datasets were aligned against the *A. thaliana* genome (TAIR10) using bowtie2 with default parameters (*51*). Duplicate reads were removed using the Picard toolkit (https//broadinstitute.github.io/picard) (MarkDuplicates with *REMOVE_DUPLICATES=true*). The mapped reads were filtered based on mapping quality using samtools (-q 30) (*52*) (Table S4). Biological replicates were analyzed for consistency with deepTools2 (fig. S12A) (*53*). For generating the chromosomal representations, the program featureCounts (version 1.6.4 (*54*)) was used to count the paired-end fragments present in each 200-kb region of the *A. thaliana* genome. As previously described (*55*), the log2 ratio was centered on the average ratio of any two compared libraries (i.e. mutant vs Col) on the first 5 Mbp of chromosome 1 for normalization. Plot profiles were done using R (version 3.6.2) (*56*) and Gviz (*57*).

RNA sequencing

For each biological replicate, leaves from three individual plants growing in the same flat were pooled. Two biological replicates per genotype were sequenced. RNA was extracted from three-week-old leaf tissue using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). RNA quality was verified using the Agilent 2100 Bioanalyzer Nano RNA Assay. Libraries were prepared at the YCGA with 1 µg of total RNA using Illumina's TruSeq Stranded Total RNA with Ribo-Zero Plant (Illumina). The libraries were amplified with eight PCR cycles, validated using Agilent Bioanalyzer 2100 High sensitivity DNA assay and quantified using the KAPA Library Quantification Kit (Illumina® Platforms). Sequencing was done on an Illumina NovaSeq 6000 using the S4 XP workflow. Paired-end reads were filtered and trimmed using fastp (version 0.21.0

with default parameters) (*50*). Reads with quality inferior to 20 were removed from the data sets (Table S5). Biological replicates were analyzed for consistency with deepTools2 (fig. S12B) (*53*). Data sets were aligned against the *A. thaliana* genome (TAIR10) using STAR (version 2.7.2a) allowing two mismatches (--outFilterMismatchNmax 2) (*58*). Transposable elements (TEs) were defined according to Panda *et al*, 2016 (*59*). featureCounts (version 1.6.4) (*23*) was used to count the paired-end fragments overlapping with TEs. TPM (transcripts per million) values were calculated for TEs. TEs were considered to be upregulated in mutant lines if they showed a ≥2-fold up-regulation as compared to Col in both biological replicates, and had a value of TPM ≥ 5. The heatmap was drawn with the R built-in function (version 3.6.2) (*25*).

Somatic recombination assay

The inverted repeat GUS reporter line used in this study was described previously (*23*). This reporter line was crossed with the following mutants: *atxr5 atxr6, tsk,* and *atxr5 atxr6 tsk*. Plants in the F3 generation homozygous for the GUS reporter gene, and either WT (control) or homozygous mutant for *atxr5*/*6, tsk*, or *atxr5/6 tsk* were identified. Experiments were performed at least three times in four-week-old F3 plants as previously described (*23*).

MMS genotoxic assay

Seeds were germinated and grown on ½ MS plates with or without 100 µg/ml methyl methansulfonate (MMS) (Thermo Fisher Scientific) under cool-white fluorescent lights (~100 µmol m^{-2} s^{-1}) in long-day conditions (16 h light/8 h dark). Seedlings were grown on vertically oriented plates for root length measurements. Measurements were done 14 days after germination. The experiments were repeated tree times with similar results.

CRISPR

The *rad51* mutant was obtained by multiplex CRISPR/Cas9-mediated deletion of the full *RAD51* gene. Two guide RNAs (*RAD51-F:* GTAGTGTTGTATAAACCACG and *RAD51-R:* AACACCTAGGTATCACTCGG) were designed with CHOPCHOP v3 (*24*) and cloned into an entry vector as described previously (*10*). The resulting AtU6.26:gRNA cassettes were amplified using the PhusionFlash polymerase (ThermoFisher Scientific). A modular cloning (MoClo) reaction (*25*) was then used to clone the F and R amplicons at positions #3 and #4, respectively, of the pAGM65879 acceptor vector (Addgene plasmid #153214; gifted by S. Marillonnet), which provides an *RPS5a*-driven, intron-optimized SpCas9 variant at position #2 (*26*). An *OLE1p:OLE1-RFP* reporter construct was used at position #5 for selection of transformants. Final constructs were agroinfiltrated into Col and *atxr5/6* T0 plants. In both backgrounds, one transgene-free T2 plant heterozygous for the *rad51* deletion was selected by PCR and selfed. The resulting T3 populations were screened for homozygous *rad51* mutants.

CRISPR/Cas9 editing of *HTR9* and *HTR13* used a Level2 MoClo vector. This was constructed using the pAGM4723 acceptor vector containing a YAOp:SaCas9:E9 cassette at position #2 (reverse orientation), AtU6.26:gRNA cassettes targeting *HTR9* (CTCAACGCCACCGTTCCTGG A) and *HTR13 (*CTCAAGGCAACAGTTCCTGGA) at positions #3 and #4, an *OLE1p:OLE-RFP* cassette at position #5 and a *Nos*:*Hyg*:*Ocs* at position #6. The YAOp:SaCas9::E9 cassette was

cloned from plasmids LBJJ491 (Addgene plasmid #117513), pEPOR0SP0020 (Addgene plasmid #117531) and pICSL60004 (Addgene plasmid #117519), which were gifts from Jonathan D. Jones (*60*) and Nicola Patron (*61*). The AtU6.26:gRNA cassettes used pICSL90002 (Addgene plasmid #68261), also a gift from Nicola Patron.

All CRISPR/Cas9 editing vectors used components from The MoClo Toolkit (Addgene kit #1000000044) (*62*) and The MoClo Plant Parts Kit (Addgene kit #1000000047) (*63*). Transformations were done by floral dip using *Agrobacterium tumefaciens* GV3101. Transformant T1 seeds were selected using the *OLE1p:OLE-RFP* reporter.

Amplicon sequencing

DNA extracted from T1 plants was amplified and pooled using Custom rhAmpSeq Panels and the rhAmpSeq Library Kit (Integrated DNA Technologies, Coralville, IA). Multiplexed libraries were then sequenced on a NovaSeq 6000 (Illumina) producing paired-end 2×150bp reads. Reads were analyzed using the CRISPResso2 pipeline (*64*).

rhAMpSeq primers	Sequence
HTR9-F	AACTCCTAAAATGGCTCGTAC**rC**AAGC
HTR9-R	AAGCTCAGTACTCTTCTGATACTT**rC**CTGA
HTR13-F	GTTTGATTTCGAAATGGCTCGTA**rC**TAAG
HTR13-R	CAGTGCTCTTCTGATACTTCCT**rG**ATCT

* r indicates RNA bases

Graphic design

The model depicted in Fig. 4G was created with BioRender.com.

Figure S1: Alignment of plant TSK proteins. The alignment was generated with Clustal Omega and represented using Jalview. NCBI reference sequences: GAY58445.1 (*Citrus unshiu*), NP_188503.2 (*A. thaliana*), XP_006585323.1 (*Glycine max*), XP_015624059.1 (*Oryza sativa*), XP_024378964.1 (*Physcomitrium patens*), and XP_024518191.1 (*Selaginella moellendorffii*). Dark blue, blue and light blue residues indicate 100%, 80%, and 60% identity, respectively, across all six protein sequences. The sequences corresponding to the TPR domains are indicated.

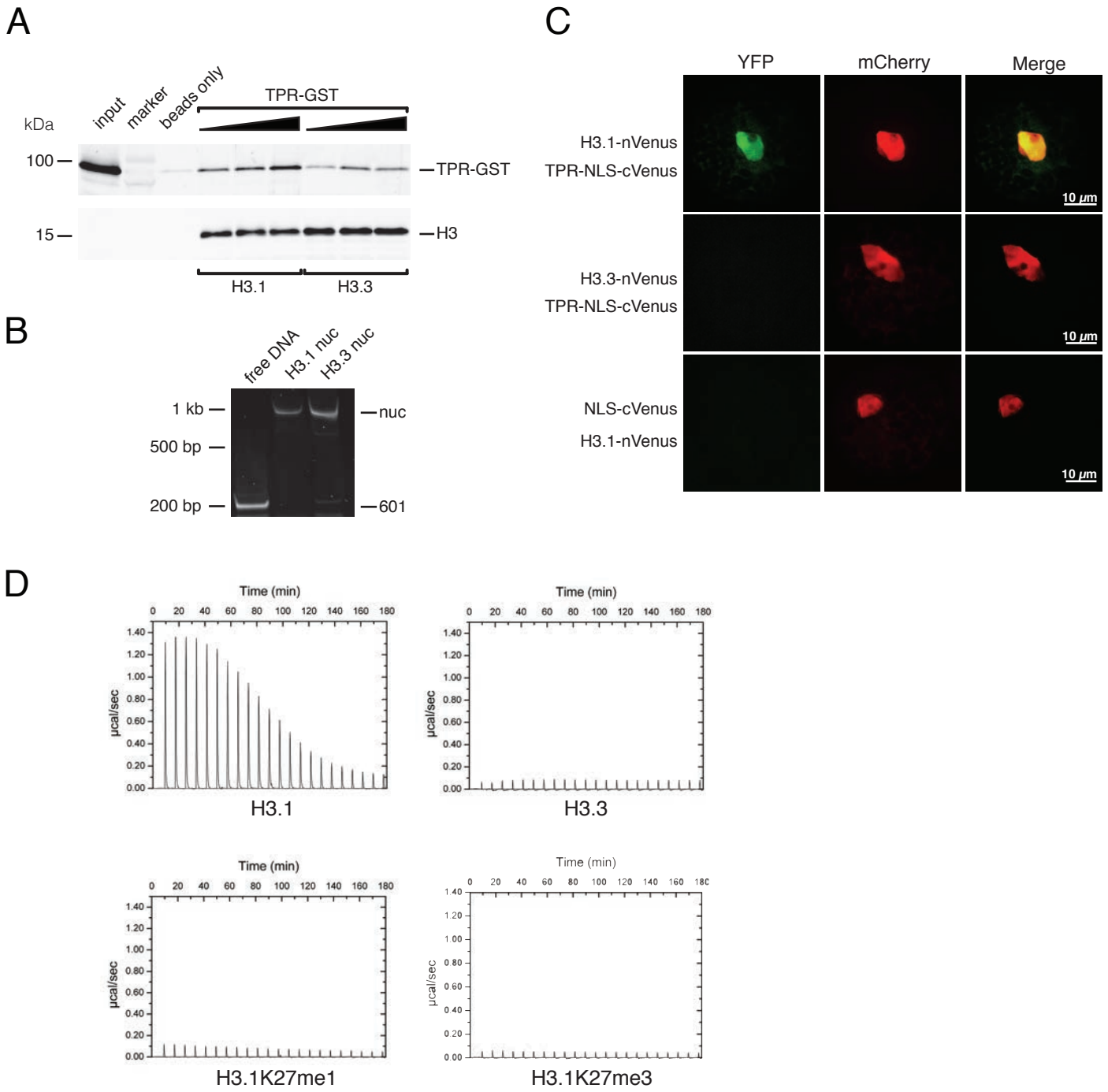
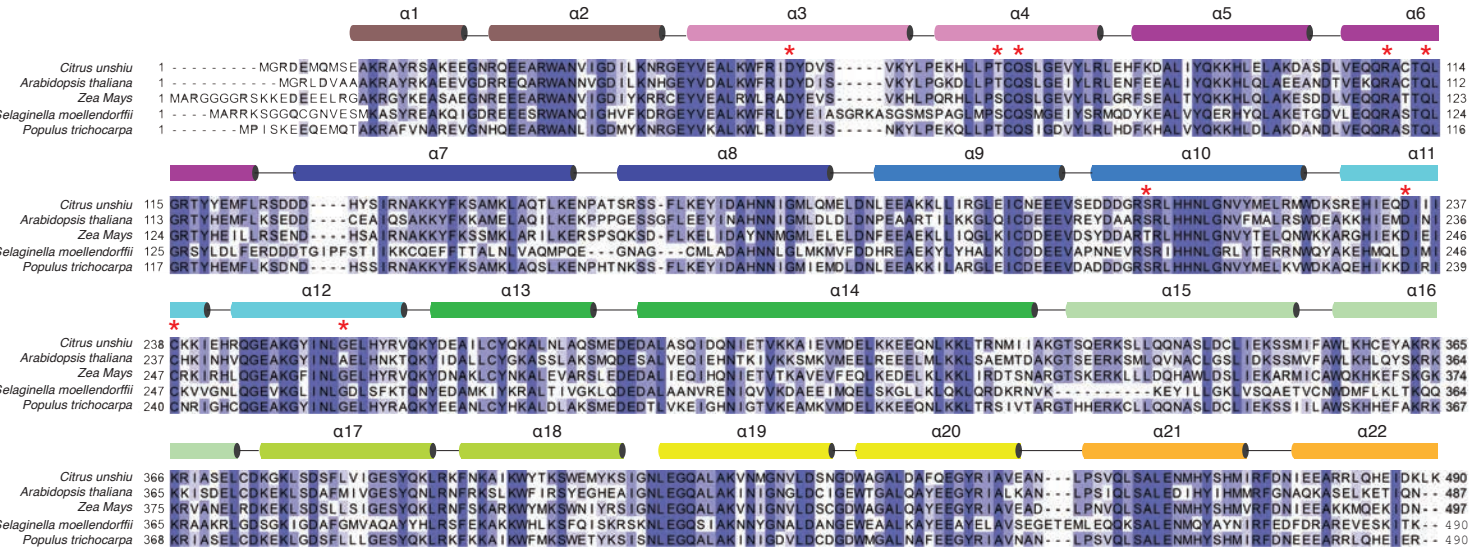
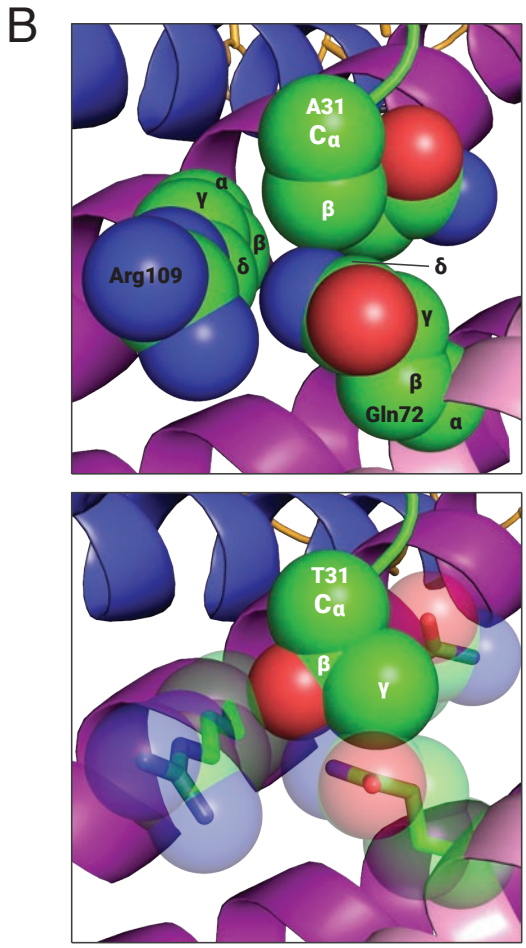


Figure S2. The TPR domain of TSK specifically interacts with the N-terminal tail of the H3.1 variant. (A) Nucleosome pulldown assays with recombinant mononucleosomes containing either *A. thaliana* H3.1 or H3.3 and increasing amounts of GST-tagged TPR_{TSK} (0.5, 1, and 2-fold molar ratio to nucleosomes). Nucleosome loading and binding of TPR was assessed by Western Blot with anti-H3 and anti-GST antibodies, respectively. (B) Native gel electrophoresis was performed to verify assembly of recombinant nucleosomes used for pulldown experiments shown in panel A. Mononucleosomes (nuc) were separated from free 601 DNA (601) on native 6% polyacrylamide gels and visualized with SYBR safe stain. (C) Bimolecular fluorescence complementation (BiFC) assay in *A. thaliana* protoplasts. H3.1 or H3.3 fused to the N-terminus of Venus (YFP) and TSKTPR-NLS (nuclear localization signal) fused to the C-terminus of Venus were co-transformed into protoplasts. mCherry-VirD2NLS was co-expressed as a nuclear marker. H3.1-nVenus and NLS-cVenus were co-transformed into protoplasts as a negative control. (D) Thermograms of ITC assays using plant TPR_{TSK} and H3.1, H3.3, H3.1K27me1, and H3.1K27me3 peptides.





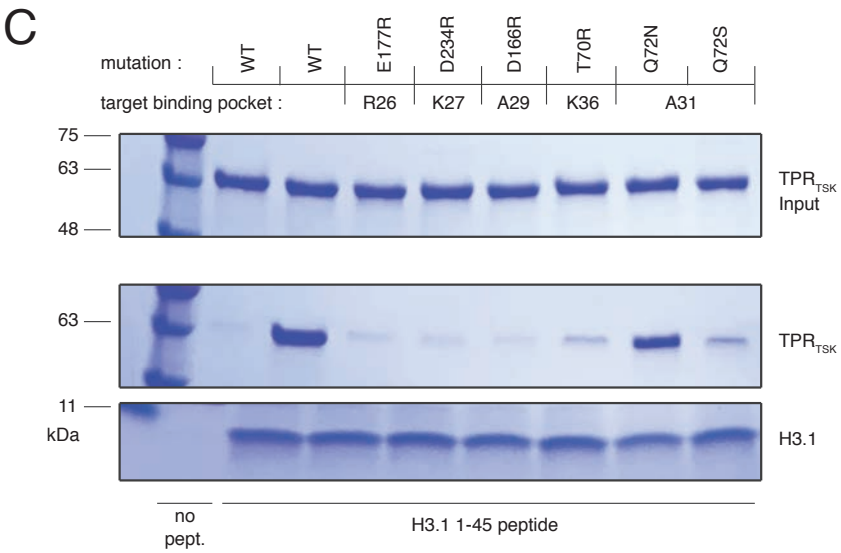


Figure S3: Analysis of amino acid residues of TPR$_{TSK}$ interacting with H3.1 (A) Sequence alignment of the TPR domain of TSK from multiple plant species. Red asterisks indicate the CuTPR$_{TSK}$ residues which interact with H3.1 K27, K36, and A31. Cylinders atop the sequences mark boundaries of CuTPR$_{TSK}$ sequences of individual TPR folds. (B) Spherical representation of the atomic radii of H3.1A31 in its binding pocket on CuTPR$_{TSK}$. (Top panel) A31 is shown with its Cβ bordering the amine group of Gln72. (Bottom panel) Threonine replacing alanine at position 31 (as in H3.3) produces van der Waals clashes between T31-C γ and Gln72-Cγ/Cδ/amine group, between T31-OH and Arg109-Cβ/Cγ, and between T31-OH and Gln72-amine group. (C) Streptavidin pulldowns with biotin-H3.1$_{(1-45)}$ and various TPR single mutants targeting different H3.1 binding pockets.

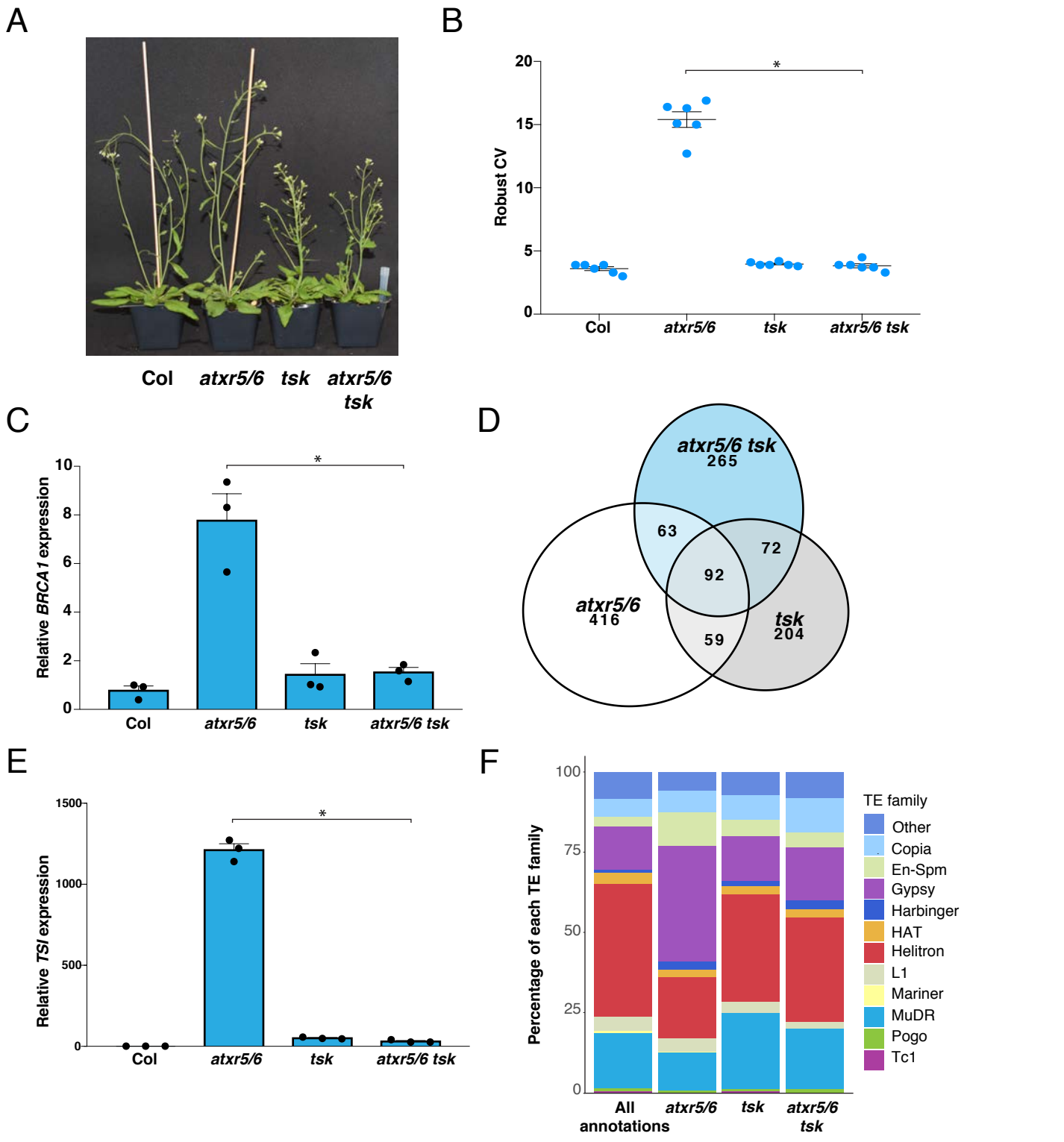


Figure S4: Effect of *TSK* on genome stability and transcriptional de-repression in *atxr5/6* mutants. (A) Morphological phenotypes of *atxr5/6*, *tsk* and *atxr5/6 tsk*. (B) Robust CV values for 16C nuclei obtained by flow cytometry analysis. Each dot represents a biological replicate. Horizontal bars indicate the mean. SEM are shown. The asterisk indicates a significant difference as determined by a Brown-Forsythe and Welch ANOVA test followed by the Dunnett T3 test for multiple comparisons: * $p < 0.0001$ (C) RT-qPCR analysis of the genome stability marker *BRCA1* in Col, *atxr5/6*, *tsk* and *atxr5/6 tsk*. The average of three biological replicates and SEM are shown. Unpaired *t*-test: * $p < 0.01$. (D) Venn diagram showing the number of upregulated and downregulated TEs (≥2-fold change) in *atxr5/6*, *tsk* and *atxr5/6 tsk* compared to Col plants ($p_{adj} < 0.05$). (E) RT-qPCR analysis of the DNA repeat *TSI* in Col, *atxr5/6*, *tsk* and *atxr5/6 tsk*. The average of three biological replicates and SEM are shown. Unpaired *t*-test: * $p < 0.0001$. (F) Distribution of reactivated TEs in *atxr5/6*, *tsk*, and *atxr5/6 tsk* among the different TE families.

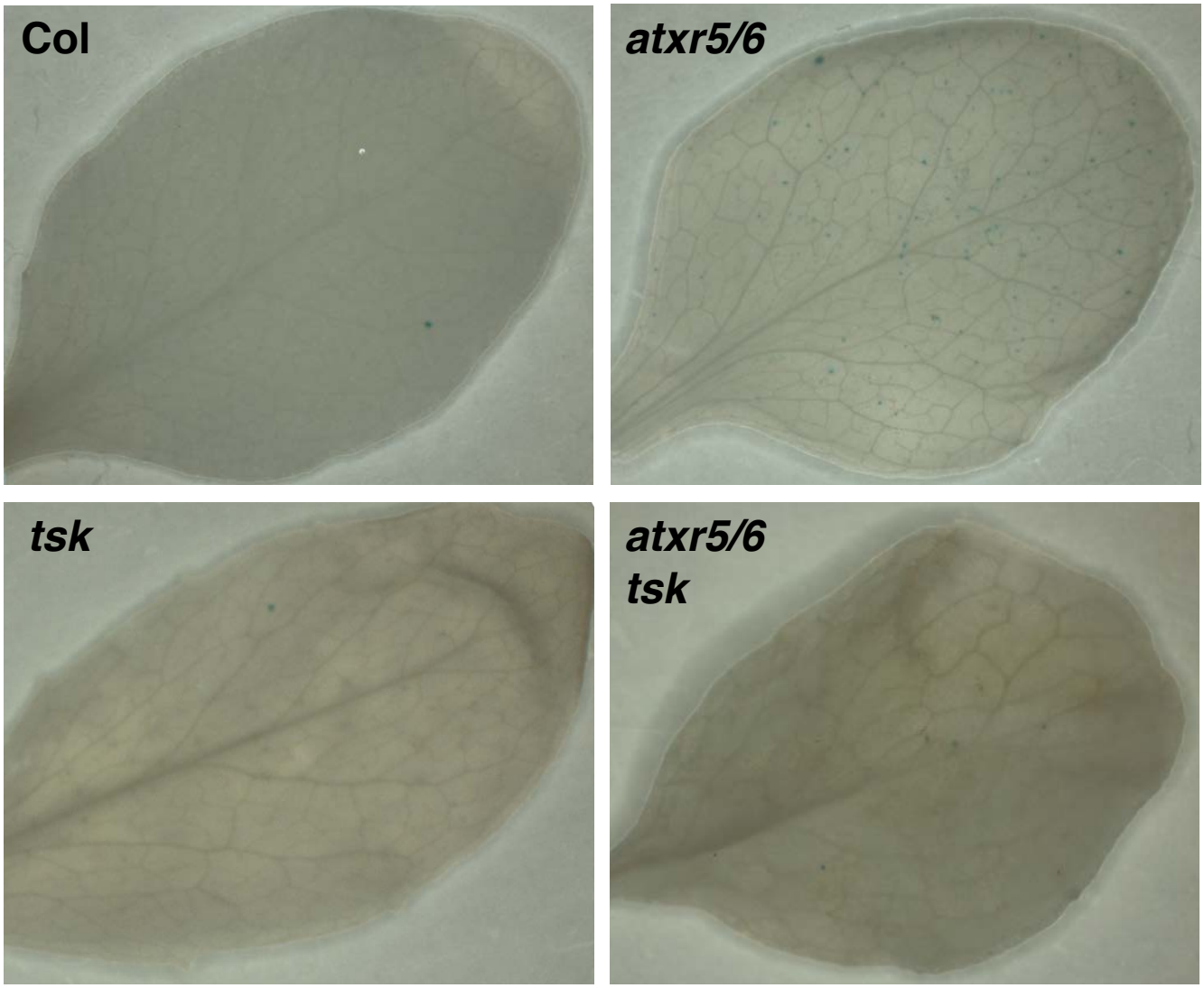


Figure S5: Increased levels of homologous recombination in *atxr5/6* mutants. Representative images of GUS activity in the leaves of Col, *atxr5/6*, *tsk* and *atxr5/6 tsk*.

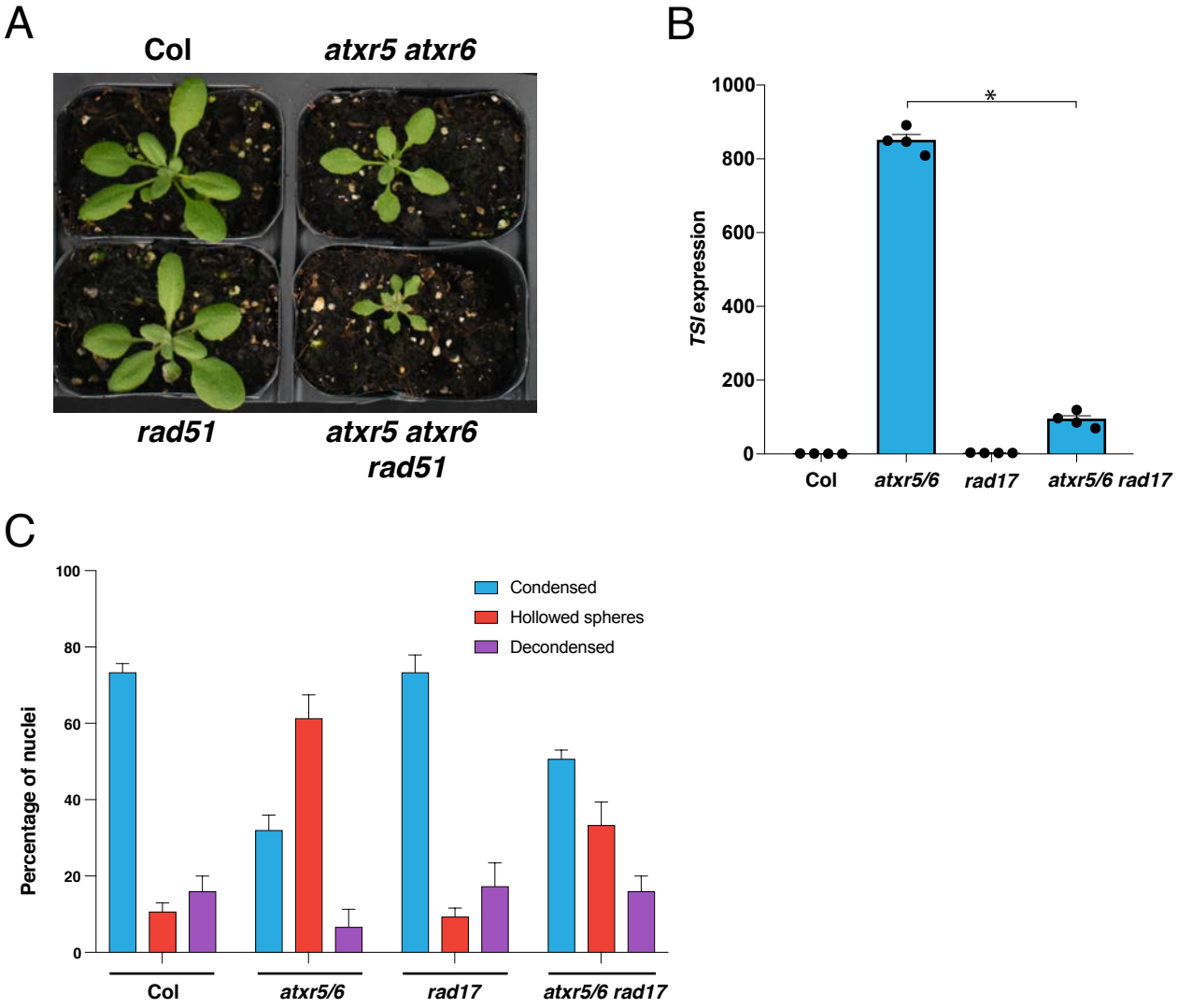


Figure S6: Inactivating *RAD17* in *atxr5/6* mutants suppresses transcriptional de-repression and the organization of heterochromatin. (A) Morphological phenotypes of *atxr5/6*, *rad51* and *atxr5/6 rad51*. (B) RT-qPCR analyses of the repetitive element *TSI*. Each dot represents an independent biological replicate. The average of three biological replicates and SEM are shown. Unpaired *t*-test: * $p < 0.0001$. (C) Quantification of chromocenter appearance from DAPI-stained nuclei. Shown is the percentage of nuclei that are fully condensed, displaying a hollowed sphere conformation and irregularly/partially decondensed. Twenty-five nuclei for three biological replicates of each genotype were assessed. Error bars indicate SEM.

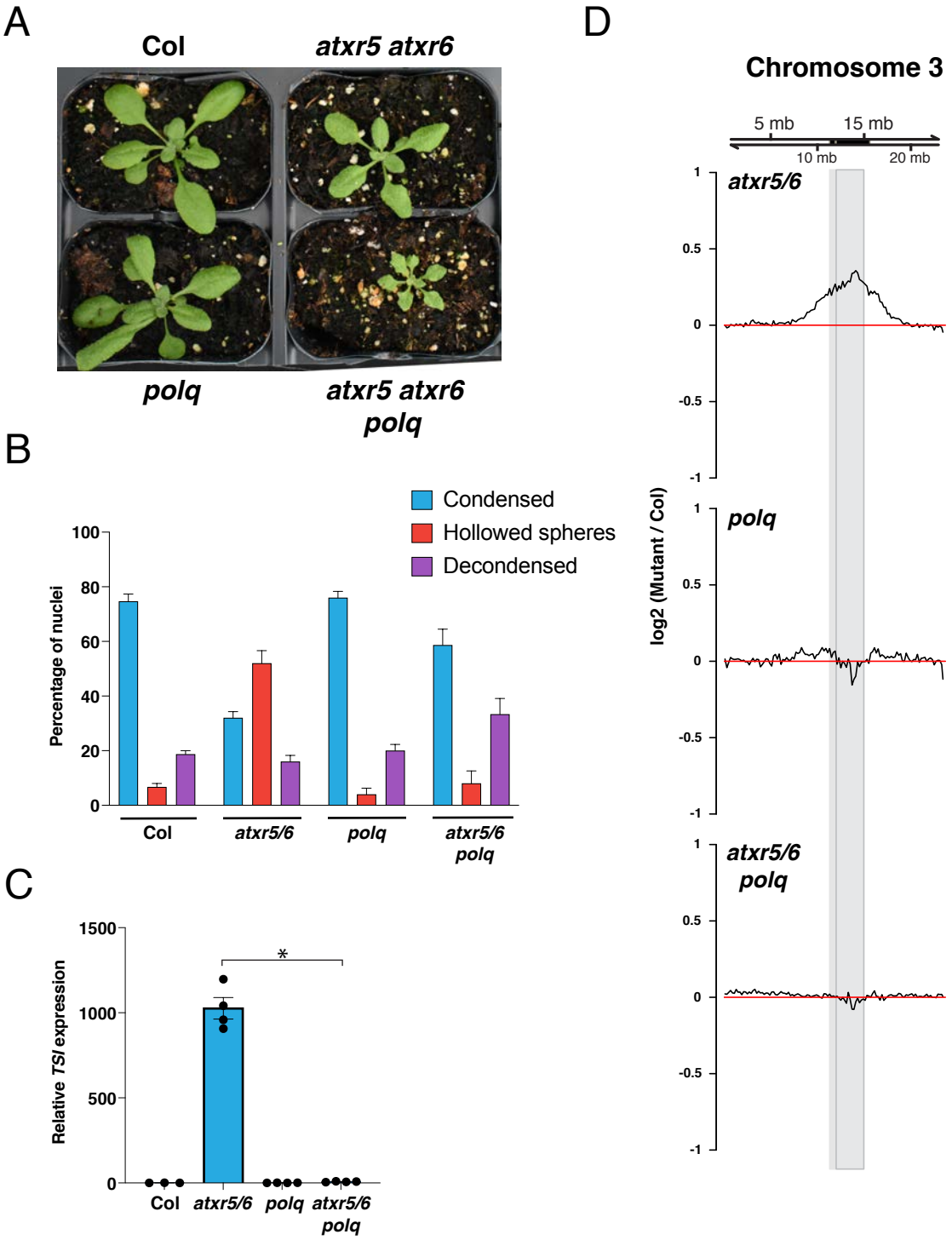


Figure S7: Mutations in *POLQ* suppress the heterochromatin phenotypes of *atxr5/6* mutants. (A) Morphological phenotypes of *atxr5/6*, *polq* and *atxr5/6 polq*. (B) Quantification of chromocenter appearance from DAPI-stained nuclei. Shown is the percentage of nuclei that are fully condensed, displaying a hollowed sphere conformation and irregularly/partially decondensed. Twenty-five nuclei for three biological replicates of each genotype were assessed. Error bars indicate SEM. (C) RT-qPCR analyses of the repetitive element *TSI*. Each dot represents an independent biological replicate. The average of three biological replicates and SEM are shown. Unpaired *t*-test: * $p < 0.0001$. (D) Chromosomal view (Chromosome 3 of *A. thaliana*) of DNA sequencing reads from *atxr5/6*, *polq* and *atxr5/6 polq* sorted 16C nuclei, normalized to reads from Col. The centromeric region is highlighted in gray.

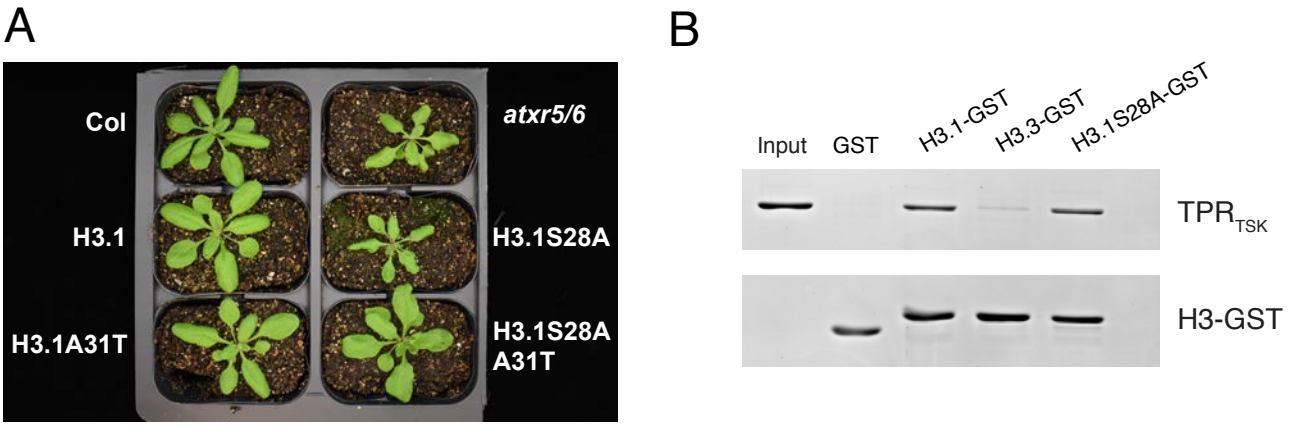


Figure S8: Expression of the histone point mutant H3.1S28A in *A. thaliana* does not interfere with binding of TSK to H3.1. (A) Morphological phenotypes of T1 plants expressing different H3.1 transgenes. (B) Pull-down assay using the TPR domain of TSK and GST-tagged histones H3.1, H3.3 and H3.1S28A.

A

Col	*tsk* mutants	*h3.1* mutants
WT H3.1 Line 1	WT H3.1 Line 2	WT H3.1 Line 3
H3.1A31T Line 1	H3.1A31T Line 2	H3.1A31T Line 3

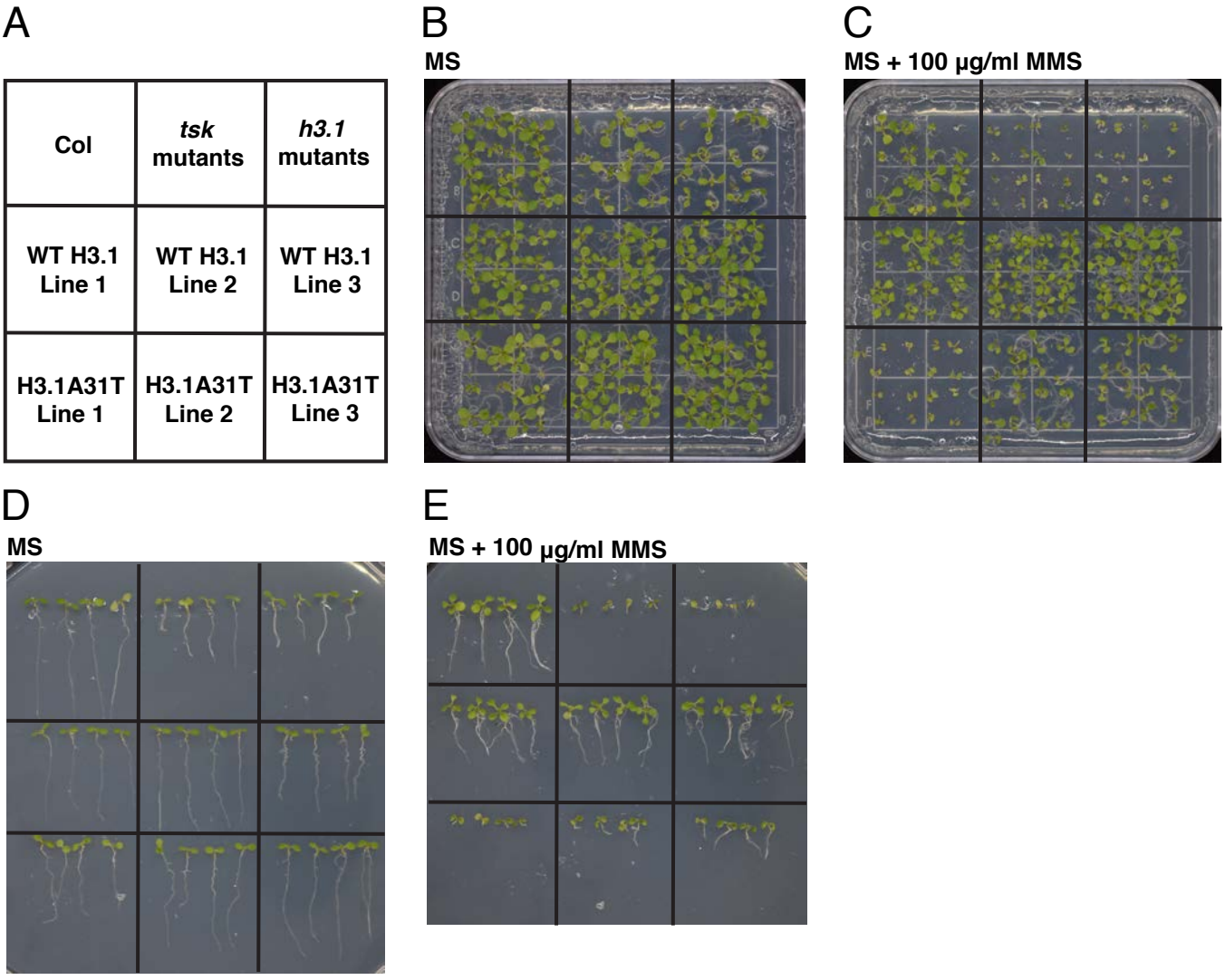


Figure S9. Sensitivity of *tsk* and *h3.1* mutants, and H3.1A31T-expressing plants, to genotoxic stress. (A) Layout of plant genotypes grown on plates shown in B-E. H3.1WT lines 1-3 and H3.1A31T lines 1-3 are T4 transgenic plants from independent T1 parents. (B-C) Representative seedlings grown on horizontally-oriented ½ MS plates in the absence (B) or the presence (C) of 100 µg/ml MMS. (D-E) Representative seedlings grown on vertically-oriented plates in the absence (D) or the presence (E) of MMS. (F) Root length of seedlings grown with or without MMS. SEM is shown. Eight seedlings were measured for each genotype.

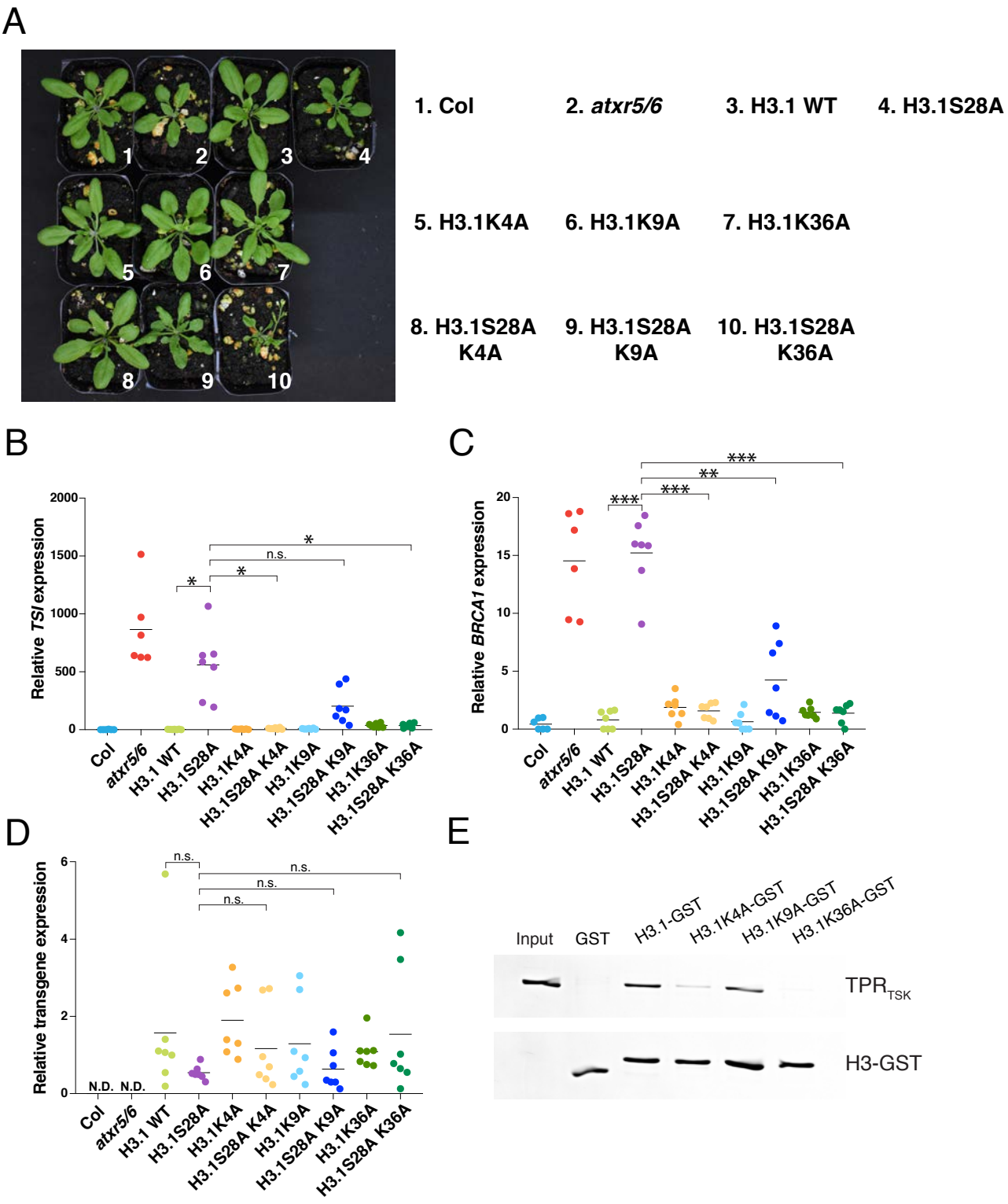


Figure S10: Effects of alanine replacement at K4, K9 and K36 of H3.1 on genomic stability. (A) Morphological phenotypes of T1 plants expressing different *H3.1* transgenes. (B-D) RT-qPCR analyses of the repetitive element *TSI* (B), *BRCA1* (C) and the *H3.1* transgene (D) in Col, *atxr5/6* mutants and the *H3.1* replacement lines. For Col and *atxr5/6*, each dot represents an independent biological replicate. For the *H3.1* lines, each dot represents one T1 plant. Horizontal bars indicate the mean. N.D. = not detected. The asterisks indicate a significant difference as determined by a Brown-Forsythe and Welch ANOVA test followed by the Dunnett T3 test for multiple comparisons: * $p < 0.05$, ** $p < 0.001$, $p < 0.0005$ and n.s. = not significantly different. (E) Pull-down assay using the TPR domain of TSK and GST-tagged histones H3.1, H3.1K4A, H3.1K9A and H3.1K36A.

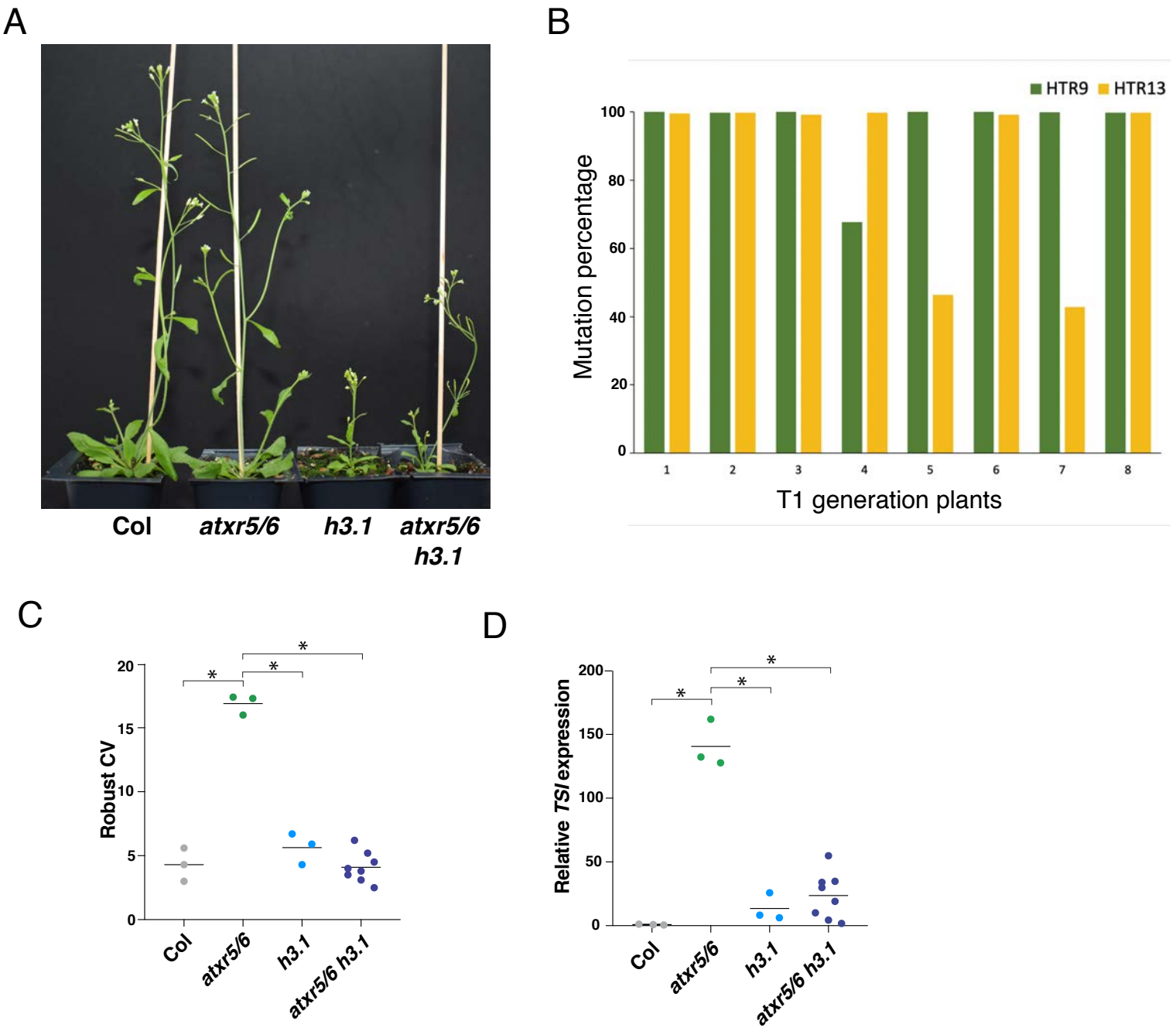


Figure S11: H3.1 is required to induce heterochromatin amplification and transcriptional reactivation in the absence of *ATXR5/6*. (A) Morphological phenotype of *atxr5/6*, *h3.1* and *atxr5/6 h3.1* mutants. (B) Percentage of mutated *HTR9* and *HTR13* alleles in individual T1 plants (*atxr5/6 h3.1*) as determined by amplicon sequencing. (C) rCV values for 16C nuclei obtained by flow cytometry analysis. For Col, *atxr5/6*, and *h3.1,* each dot represents a biological replicate. For the *atxr5/6 h3.1* CRISPR lines, each dot represents one T1 plant. Horizontal bars indicate the mean. The asterisks indicate a significant difference as determined by a Brown-Forsythe and Welch ANOVA test followed by the Dunnett T3 test for multiple comparisons: * $p < 0.005$. (D) RT-qPCR analyses of the repetitive element *TSI*. Each dot represents an independent biological replicate. The average of three biological replicates and SEM are shown.The asterisks indicate a significant difference as determined by a Brown-Forsythe and Welch ANOVA test followed by the Dunnett T3 test for multiple comparisons: * $p < 0.05$.

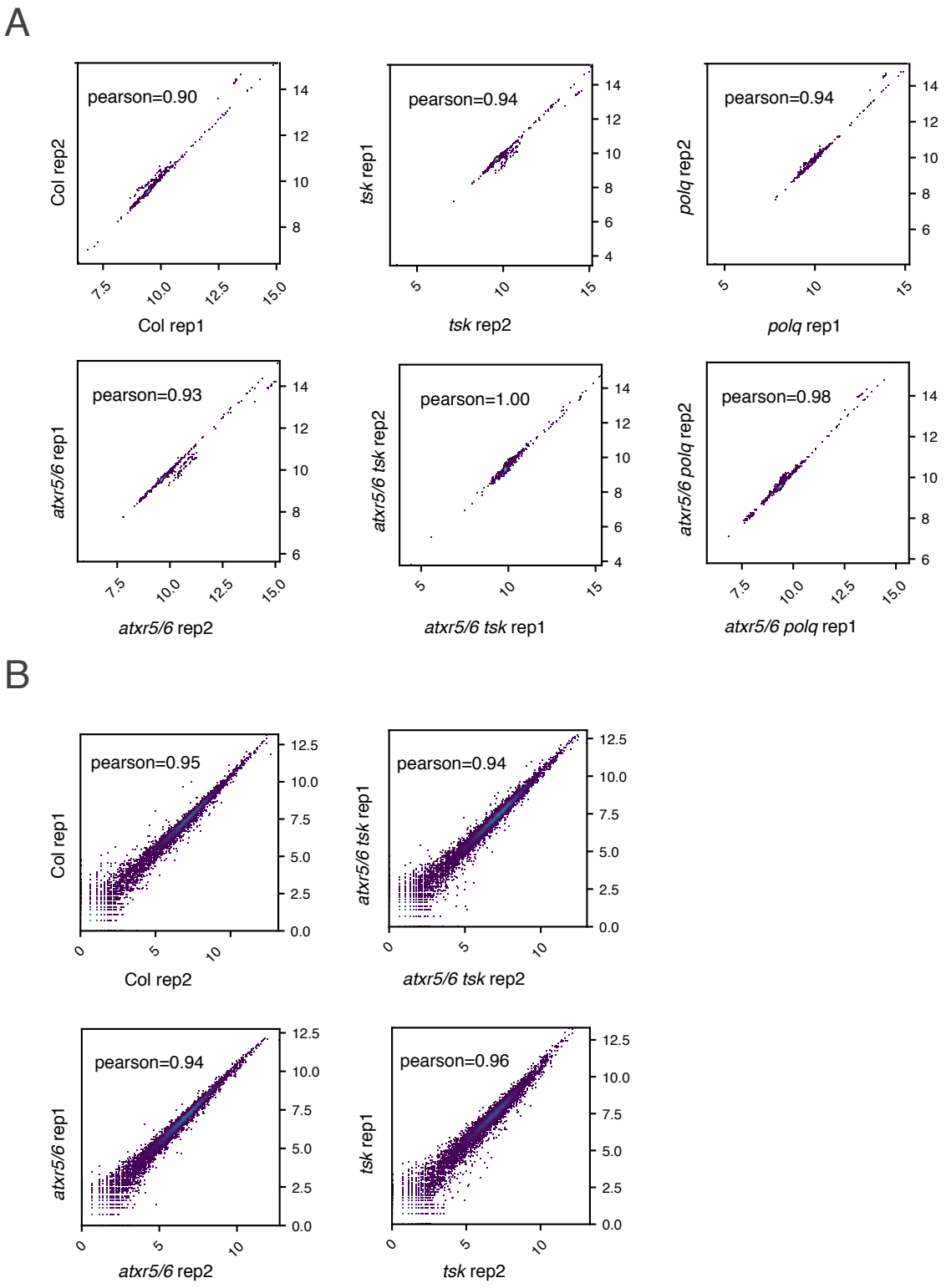


Figure S12: Scatterplots and Pearson correlation coefficients for the sequencing biological replicates. (A) DNA-seq replicates of Col, *atxr5/6, tsk*, *atxr5/6 tsk*, *polq* and *atxr5/6 polq*. (B) RNA-seq replicates of Col, *atxr5/6*, *tsk* and *atxr5/6 tsk*.

Table S1. Data collection and refinement statistics for TPR_{TSK} and H3.1 complex.

PDB accession number	7T7T
Data Collection	
Space group	P 2_1 2_1 2_1
Cell dimensions	
a, b, c (Å)	82.89, 92.33, 209.91
α, β, γ (°)	90, 90, 90
Resolution	37.21 - 3.17 (3.28 - 3.17)
R_{meas}	0.05 (0.33)
R_{pim}	0.01
I / σI	19.5 (2.6)
No. unique reflections	27369 (2515)
Completeness (%)	99.9 (99.9)
Redundancy	26.1 (26.1)
CC1/2	0.99 (0.92)
Wilson *B*-Factor	31
Refinement	
Resolution (Å)	37.21 - 3.17
No. reflections	27365 (2515)
R_{work} / R_{free}	0.27 / 0.31
Num. atoms	
TPR^{TSK}	6832
$H3.1^{1-45}$	343
Water	1
B-factors (Å^2)	
TPR^{TSK}	29
$H3.1^{1-45}$	28
R.m.s. deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.74
Molprobity score	1.76
Clashscore	8.67
Ramachandran favored (%)	95.74
Ramachandran allowed (%)	4.26
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00

* Highest-resolution shell is shown in parentheses.

Table S3: Summary of synthesized histone H3 peptides. Peptide synthesis based on the main backbone peptide, and further modifications with different amino acids in positions "1", "2", and "3". Most relevant mass spectra molecular ions identified are presented.

Main Peptide Backbone					
ARTKQTARKSTGGKAPRKQLATKAAR**-1-**SAP**-2-**TGGVKKPHR**-3-**RPGTY-NH_2					
Peptide	**Amino Acid Modification**			**Theorical Molecular Ions**	**Identified Molecular Ions**
	1	**2**	**3**		
H3.1	Lys	Ala	Phe	[M+7H]7+: 703.7 [M+8H]8+: 615.7 [M+9H]9+: 547.4	[M+7H]7+: 703.6 [M+8H]8+: 615.7 [M+9H]9+: 547.5
H3.3	Lys	Thr	Tyr	[M+7H]7+: 710.1 [M+8H]8+: 621.5 [M+9H]9+: 552.7	[M+7H]7+: 710.1 [M+8H]8+: 621.5 [M+9H]9+: 552.5
H3.1-K27Me1	Lys(Me)	Ala	Phe	[M+7H]7+: 705.6 [M+8H]8+: 617.5 [M+9H]9+: 549.0	[M+7H]7+: 705.7 [M+8H]8+: 617.5 [M+9H]9+: 549.1
H3.1-K27Me3	$Lys(Me)_3$	Ala	Phe	[M+7H]7+: 709.6 [M+8H]8+: 621.0 [M+9H]9+: 552.1	[M+7H]7+: 709.8 [M+8H]8+: 621.1 [M+9H]9+: 552.3

Additional data (separate files)

Table S2: TEs de-repressed in *atxr5/6*, *tsk* and *atxr5/6 tsk*.

Table S4: Sequencing summary statistics for the 16C nuclei analysis.

Table S5: Sequencing summary statistics for the RNA-sequencing experiment.

References

29. Y. Jacob *et al.*, ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nat Struct Mol Biol* **16**, 763-768 (2009).
30. K. Brzezinka, S. Altmann, I. Baurle, BRUSHY1/TONSOKU/MGOUN3 is required for heat stress memory. *Plant Cell Environ* **42**, 771-781 (2019).
31. S. Valuchova *et al.*, Protection of Arabidopsis Blunt-Ended Telomeres Is Mediated by a Physical Association with the Ku Heterodimer. *Plant Cell* **29**, 1533-1545 (2017).
32. M. L. Heacock, R. A. Idol, J. D. Friesner, A. B. Britt, D. E. Shippen, Telomere dynamics and fusion of critically shortened telomeres in plants lacking DNA ligase IV. *Nucleic Acids Res* **35**, 6490-6500 (2007).
33. F. Heitzeberg *et al.*, The Rad17 homologue of Arabidopsis is involved in the regulation of DNA damage repair and homologous recombination. *Plant J* **38**, 954-968 (2004).
34. S. Wang, W. E. Durrant, J. Song, N. W. Spivey, X. Dong, Arabidopsis BRCA2 and RAD51 proteins are specifically involved in defense gene transcription during plant immune responses. *Proc Natl Acad Sci U S A* **107**, 22716-22721 (2010).
35. K. Osakabe *et al.*, Isolation and characterization of the RAD54 gene from Arabidopsis thaliana. *Plant J* **48**, 827-842 (2006).
36. S. Inagaki *et al.*, Arabidopsis TEBICHI, with helicase and DNA polymerase domains, is required for regulated cell division and differentiation in meristems. *Plant Cell* **18**, 879-892 (2006).
37. Y. Jacob *et al.*, Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* **343**, 1249-1253 (2014).
38. C. LeBlanc *et al.*, Increased efficiency of targeted mutagenesis by CRISPR/Cas9 in plants using heat stress. *Plant J* **93**, 377-386 (2018).
39. R. Liu, D. C. Chan, The mitochondrial fission receptor Mff selectively recruits oligomerized Drp1. *Mol Biol Cell* **26**, 4466-4477 (2015).
40. M. Karimi, D. Inze, A. Depicker, GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci* **7**, 193-195 (2002).
41. P. Voigt *et al.*, Asymmetrically modified nucleosomes. *Cell* **151**, 181-193 (2012).
42. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326 (1997).
43. G. M. Sheldrick, Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr* **66**, 479-485 (2010).
44. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
45. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
46. V. B. Chen *et al.*, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
47. C. M. Lee *et al.*, GIGANTEA recruits the UBP12 and UBP13 deubiquitylases to regulate accumulation of the ZTL photoreceptor complex. *Nat Commun* **10**, 3750 (2019).
48. F. H. Wu *et al.*, Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method. *Plant Methods* **5**, 16 (2009).
49. J. Nijjer *et al.*, Mechanical forces drive a reorientation cascade leading to biofilm self-patterning. *Nat Commun* **12**, 6632 (2021).

50. J. Schindelin *et al.*, Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).
51. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408 (2001).
52. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890 (2018).
53. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).
54. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
55. F. Ramirez *et al.*, deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165 (2016).
56. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
57. C. J. Hale *et al.*, Identification of Multiple Proteins Coupling Transcriptional Gene Silencing to Genome Stability in Arabidopsis thaliana. *PLoS Genet* **12**, e1006092 (2016).
58. R. C. Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria., (2018).
59. F. Hahne, R. Ivanek, Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol* **1418**, 335-351 (2016).
60. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
61. K. Panda *et al.*, Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biol* **17**, 170 (2016).
62. J. M. Lucht *et al.*, Pathogen stress increases somatic recombination frequency in Arabidopsis. *Nat Genet* **30**, 311-314 (2002).
63. B. Castel, L. Tomlinson, F. Locci, Y. Yang, J. D. G. Jones, Optimization of T-DNA architecture for Cas9-mediated mutagenesis in Arabidopsis. *PLoS One* **14**, e0204778 (2019).
64. O. Raitskin, C. Schudoma, A. West, N. J. Patron, Comparison of efficiency and specificity of CRISPR-associated (Cas) nucleases in plants: An expanded toolkit for precision genome engineering. *PLoS One* **14**, e0211598 (2019).
65. E. Weber, C. Engler, R. Gruetzner, S. Werner, S. Marillonnet, A modular cloning system for standardized assembly of multigene constructs. *PLoS One* **6**, e16765 (2011).
66. C. Engler *et al.*, A golden gate modular cloning toolbox for plants. *ACS Synth Biol* **3**, 839-843 (2014).
67. K. Clement *et al.*, CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol* **37**, 224-226 (2019).