Title

The TRIPLE PHD FINGERS proteins are required for SWI/SNF complex-mediated +1 nucleosome positioning and transcription start site determination in Arabidopsis.

Journal

Nucleic Acids Research, 50(18)

Authors

Diego-Martin, Borja
Pérez-Alemany, Jaime
Candela-Ferre, Joan
et al.

Publication Date

2022-10-14

DOI

10.1093/nar/gkac826

Peer reviewed

Published online 3 October 2022 *Nucleic Acids Research, 2022, Vol. 50, No. 18* ***10399–10417***
https://doi.org/10.1093/nar/gkac826

The TRIPLE PHD FINGERS proteins are required for SWI/SNF complex-mediated +1 nucleosome positioning and transcription start site determination in Arabidopsis

Borja Diego-Martin[1,†], Jaime Pérez-Alemany[1,†], Joan Candela-Ferre[1], Antonio Corbalán-Acedo[1], Juan Pereyra[1], David Alabadí[1], Yasaman Jami-Alahmadi[2], James Wohlschlegel[2] and Javier Gallego-Bartolomé[1,*]

[1]Instituto de Biología Molecular y Celular de Plantas (IBMCP), CSIC-Universitat Politècnica de València, Valencia, 46022, Spain and [2]Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, CA, 90095, USA

Received June 05, 2022; Revised September 08, 2022; Editorial Decision September 12, 2022; Accepted September 16, 2022

ABSTRACT

Eukaryotes have evolved multiple ATP-dependent chromatin remodelers to shape the nucleosome landscape. We recently uncovered an evolutionarily conserved SWItch/Sucrose Non-Fermentable (SWI/SNF) chromatin remodeler complex in plants reminiscent of the mammalian BAF subclass, which specifically incorporates the MINUSCULE (MINU) catalytic subunits and the TRIPLE PHD FINGERS (TPF) signature subunits. Here we report experimental evidence that establishes the functional relevance of TPF proteins for the complex activity. Our results show that depletion of TPF triggers similar pleiotropic phenotypes and molecular defects to those found in *minu* mutants. Moreover, we report the genomic location of MINU2 and TPF proteins as representative members of this SWI/SNF complex and their impact on nucleosome positioning and transcription. These analyses unravel the binding of the complex to thousands of genes where it modulates the position of the +1 nucleosome. These targets tend to produce 5´-shifted transcripts in the *tpf* and *minu* mutants pointing to the participation of the complex in alternative transcription start site usage. Interestingly, there is a remarkable correlation between +1 nucleosome shift and 5´ transcript length change suggesting their functional connection. In summary, this study unravels the function of a plant SWI/SNF complex involved in +1 nucleosome positioning and transcription start site determination.

INTRODUCTION

The genetic material in the eukaryotic nuclei is present in the form of chromatin. The nucleosome is the basic unit of chromatin and it is formed by pairs of histones H3, H4, H2A, and H2B wrapped with ∼147 base pairs (bp) of DNA (1). This structure represents the first layer of compaction of the genome and has a direct impact on the ability of nuclear proteins to access DNA. Moreover, it serves as a recruitment platform for multiple histone reader proteins (2). Thus, nucleosomes play an important role in the regulation of chromatin affecting diverse processes such as transcription, replication, and repair (3–5). Nucleosomes present a stereotypical distribution relative to the transcription start site (TSS) of genes transcribed by RNA Polymerase II (Pol II) where a strongly positioned nucleosome known as +1 is followed by an array of regularly spaced nucleosomes (+2, +3, etc.) and is preceded by a nucleosome depleted region (NDR) that facilitates the access of the transcriptional machinery to the promoter region (6–8). The +1 nucleosome exerts an important function in transcriptional regulation since it is a barrier to Pol II (9,10), promoting Pol II pausing in some species (11–14). Moreover, studies in fungi and animals have shown that the +1 nucleosome can have a negative effect on the ability of the transcriptional machinery to access the TSS and other promoter regulatory regions, leading to changes in gene expression and alternative TSS usage, which in turn can result in altered non-coding transcription and mRNA function (15–21).

Eukaryotes have evolved several ATP-dependent chromatin remodeler complexes to assemble, evict, slide, or restructure nucleosomes (22,23). Among these, the SWItch/sucrose non-fermenting (SWI/SNF) chromatin remodeler family has been extensively studied in diverse eu-

*To whom correspondence should be addressed. Tel: +34 963877856; Email: jagalbar@ibmcp.upv.es
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

karyotic model organisms (24–26). These complexes play a major role in the regulation of nucleosome positioning promoting an open NDR (16,27,28) and have a strong impact on cell differentiation processes being frequently mutated in different human cancers (29). The composition of SWI/SNF complexes is diverse and dynamic and is characterized by the presence of one catalytic ATPase subunit and a cohort of scaffold and regulatory subunits (30). There are different evolutionarily conserved SWI/SNF subclasses that are characterized by unique signature subunits and that show non-redundant functions (26,29,31). These distinct subclasses are known as SWI/SNF-BAF and RSC-PBAF in fungi and mammals, respectively, and a newer subclass was recently identified in mammals named non-canonical BAF (ncBAF) (32–35).

Plants conserve multiple SWI/SNF subunits and their misregulation can cause strong developmental defects and altered responses to the environment (36,37). The model plant Arabidopsis presents multiple paralogs of these subunits e.g. four ATPases (BRAHMA (BRM), SPLAYED (SYD), MINUSCULE1 (MINU1) and MINU2), four SWI3s (A, B, C, and D) and two SWP73s (A and B), which suggest the formation of multiple variants of the SWI/SNF complex (36). Through a comprehensive evolutionary study of the conservation of SWI/SNF subunits across eukaryotes, we recently reported a model for the evolution of SWI/SNF subclasses that predicts the presence of two major subclasses in plants (35). One, that we refer to as MINU-associated SWI/SNF (MAS) complex, is reminiscent of the mammalian BAF-PBAF complexes and incorporates the MINU ATPases while the other is similar to the metazoan ncBAF complex and incorporates the BRM ATPase (35). There is abundant information about the molecular function of the subunits of the plant ncBAF-like complex, like BRM (38–41), as well as the bromodomain-containing proteins BRD1,2,13 (BRDs), BRAHMA-INTERACTING PROTEINS (BRIPs) and GRF-INTERACTING FACTORS (GIFs) (42–45) signature subunits. However, there is a significant lack of information about the components and molecular function of the MAS complex. Through proteomics, we recently reported the composition of this complex in Arabidopsis, which incorporates conserved subunits as well as a set of plant-specific subunits of unknown function (35). Among them, we found two paralogs of an evolutionarily-conserved signature subunit –SMARCG- that we named TRIPLE PHD FINGERS (TPF1 and TFP2) and that have not been characterized in plants. These proteins are distant orthologs of the metazoan DOUBLE PHD FINGERS (DPF) / PHD FINGER PROTEIN 10 (PHF10) family and the fungal SWP82/Rsc7 proteins (46–48). Functional characterization of the DPF/PHF10 proteins in mammals has demonstrated their important function in the complex through dynamic incorporation upon different cellular contexts and their role in the recruitment of the complex to the chromatin mediated by their tandem PHD histone reader domains (46,49,50). The TPF proteins present three tandem PHD domains as well as a C-terminal Tudor domain, also involved in histone reading, suggesting a possible functional conservation with their animal counterparts (35).

In this study, we have explored the function of the MAS complex through a set of genomic studies that reveal a significant impact of the complex on the +1 nucleosome positioning and the TSS usage. Consistent with a prominent role in the function of the complex, *tpf* mutants present similar pleiotropic phenotypes and molecular defects to those found in mutants of the catalytic subunits of the MAS complex.

MATERIALS AND METHODS

Plant materials and growth conditions

Plants were grown in a growth chamber with a 16L/8D light cycle (LD) and 22°C. When grown on plates, plants were grown on $\frac{1}{2}$ MS medium with vitamins (Duchefa) 0.8% agar pH 5.7 under LD and 22°C. Transgenic lines were selected on $\frac{1}{2}$ MS medium supplemented with BASTA (6.67 μg/ml) or Hygromycin (35 μg/ml). For the root growth experiment, seeds were stratified for 7 days (darkness 4°C) and plants were grown on $\frac{1}{2}$ MS vertical plates for 12 days under LD photoperiod. The T-DNA insertion lines used in this study were the following: *minu1-2* (GK-146E09), *minu2-1* (SALK_057856C), *tpf1-1* (SALK_010411C), *tpf1-2* (WiscDsLox385F06), *tpf2-1* (SALK_141512C), and *tpf2-2* (SAIL_201_D01).

Transgenic lines

A genomic fragment of *TPF2*, including promoter until the codon before the stop, was amplified from genomic DNA and cloned into a pENTR/D plasmid by InFusion (Takara) to generate pENTR-gTPF2. The *TPF2* fragment was transferred by LR reaction into a modified pEarleyGate302 that contains a 3xFLAG tag downstream of the gateway cassette to generate pEG-gTPF2-3xFLAG. This construct was transformed into *Agrobacterium tumefaciens* strain *GV3101 C58C1*, which was used to transform the *tpf2-1* mutant by the floral dip method. A genomic fragment of *MINU2*, including promoter until the codon before the stop, and with an inserted XhoI site right upstream of the start codon was amplified from genomic DNA in two separate PCR reactions. These fragments were combined to amplify the full genomic sequence that was cloned into a pENTR/D plasmid by InFusion (Takara) to generate pENTR-gMINU2. A 3xFLAG tag was amplified and cloned into the unique XhoI site of pENTR-gMINU2. The 3xFLAG-gMINU2 fragment was transferred by LR reaction into a pMDC123 (51) to generate pMDC123-3xFLAG-gMINU2. This construct was transformed into *Agrobacterium tumefaciens* strain *GV3101 C58C1*, which was used to transform the *minu2-1* mutant by the floral dip method.

Root meristem microscopy

Roots from plants grown 12 days on vertical plates were incubated with 10 μg/ml Propidium Iodide and visualized in a confocal microscope Axio Observer 780 (Zeiss).

Yeast two-hybrid (Y2H)

Gateway-compatible vectors that incorporate the SWI/SNF subunits were either requested from ABRC as bacterial stabs (*SHH2* (U17329), *SWI3A* (U16949), *SWI3B* (G20908), *SWP73B* (G15375), *LFR* (G20071), *BSH* (G82285), *ARP4* (G23361), *ARP7* (U24661), *PSA2* (U25108), *BDH1* (G50477), *OPF2* (G68780), *TPF2* (TOPO-U19-C04), *MINU2* (TOPO-U19-F11)) or cloned from cDNA into a pENTR/D or pDONR207 plasmid (*PSA1* and *BDH2*, see Supplementary Data 1 for primer information). The pENTR-TPF1, pENTR-MINU1, and pENTR-BRD5 plasmids, containing their respective full-length CDS sequences, were synthesised from Genescript. The pENTR-TPF1 was mutated to generate the pENTR-N-TPF1 (TPF1 amino acids 1–422). The C-TPF1 fragment (TPF1 amino acids 423–697) was amplified from pENTR-TPF1 and cloned into a pDONR207 plasmid by BP reaction (Invitrogen). Plasmids were transferred to pGADT7 and pGBKT7 vectors (Clontech) by LR reaction (Invitrogen).

Yeasts were grown in SD media containing 6.7 g/l DifcoTM Yeast Nitrogen Base w/o amino acids (Becton Dickinson S.A.), and 1.4 g/l Yeast Synthetic Drop Out Medium Supplements (Sigma) pH 5.8, supplemented with 5% glucose, 10 μg/ml (Ura, Trp, His) and 50 μg/ml Leu. pGADT7 and pGBKT7 plasmids (Clontech) were transformed, following the lithium acetate/single-stranded carrier DNA/polyethylene glycol method, into the haploid strains *Y187* and *Y2HGold*, which were selected in SD media devoid of Leu and Trp, respectively. Diploid yeasts were obtained by mating, which was carried out overnight in YPD supplemented with 2% glucose, and the generated diploid cells were selected in SD/-Leu-Trp plates. Protein interaction was tested based on the complementation of the histidine auxotrophy, by dropping serial dilutions onto SD/-Leu-Trp-His plates.

Immunoprecipitation mass spectrometry (IP-MS)

Immunoprecipitation followed by mass spectrometry was done as previously described (35). Briefly, 8 g of inflorescences from one untransformed Col-0 replicate, two independent lines of TPF2-3xFLAG, and two independent lines of 3xFLAG-MINU2 were ground in liquid nitrogen and resuspended in 40 ml IP buffer (50 mM Tris pH 7.6, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.1% NP40, 0.5 mM DTT, 1 mM PMSF, 1 μg/μl pepstatin, and 1 × Complete EDTA-Free (Sigma)). Samples were filtered with one layer of Miracloth (Merck, cat#475855), and homogenized with a douncer (10 times soft, 10 times hard). Then, centrifuged at 4°C for 10 min at 10 000 × *g* and filtered again using a 40 μm cell strainer. 200 μl of Anti-FLAG M2 magnetic beads (Sigma, cat#M8823), previously blocked with 5% BSA, were added to the samples followed by 3 h rotating at 4°C. The samples were washed 4 times with IP buffer and two times with IP buffer without NP40. Samples were eluted with 300 μl of 250 μg/ml 3xFLAG peptide (Sigma, cat#F4799) in IP buffer without NP40 for 30 min at 25°C. This step was repeated one more time incubating samples for 15 min at 37°C. TCA was added to a final concentration of 20% and samples were incubated for 30 min on ice followed by a 30 min 4°C centrifugation at 12 000 × *g*. Subsequently, samples were washed three times with 250 μl of cold acetone and the pellet was air-dried.

For the purpose of proteomics, proteins were reduced and alkylated using 5 mM Tris (2-carboxyethyl) phosphine and 10 mM iodoacetamide, respectively. Protein digestion was achieved by sequential addition of endopeptidase Lys-C (BioLabs) and trypsin (Pierce™) at 1:100 enzyme/protein ratio followed by an incubation at 37°C overnight. Formic acid to 5% (v./v.) final concentration was added to quench the samples. Finally, desalting prior to LC-MS/MS analysis was done using C18 pipette tips (Thermo Scientific, cat # 87784) and reconstituted in 5% formic acid before analyzed by LC-MS/MS. A 25 cm long, 75 μm ID fused-silica capillary that was packed in-house with bulk ReproSil-Pur 120 C18-AQ particles as described elsewhere (52) was used to fractionate online the peptide mixtures. Peptides were subjected to a 140 min water-acetonitrile linear gradient in 6–28% buffer B (acetonitrile solution with 3% DMSO and 0.1% formic acid) at a flow rate of 200 nl min^{-1}, which was further increased to 35% followed by a rapid ramp-up to 85% using a Dionex Ultimate 3000 UHPLC (Thermo Fisher Scientific). The eluted peptides were then ionized via nanoelectrospray ionization. An *Orbitrap Fusion*™ Lumos™ *Tribrid*™ Mass Spectrometer (Thermo Fisher Scientific) was used to acquire the mass spectrometry data with an MS1 resolution of 120000 followed by sequential MS2 scans at a resolution of 15000. Raw data were searched against the TAIR Arabidopsis reference proteome. Default settings for LFQ analysis using MaxQuant 1.6.17.0 software were applied to calculate the Label-free quantitation (LFQ) intensities as described previously by (53). Shown interactors in Table 1 were previously identified in IP-MS experiments using as baits the MAS complex-specific subunits TPF1 and SHH2 (35) and follow the criteria (i) LFQ control < LFQ transgenic/10 in both transgenic lines compared to Col-0 control, and (ii) presence of unique peptides in both transgenic lines.

RNA extraction, qRT-PCR and RNA-seq libraries

Total RNA from inflorescences was extracted with the NucleoSpin RNA Plant kit (Macherey-Nagel) following the manusfacturer´s protocol and adding an on-column DNase treatment. For the RNA-seq experiments, RNA was extracted from inflorescences of Col-0, *minu1-2 minu2-1*, and *tpf1-1 tpf2-1* plants (three biological replicates per background) and sent to BGI to prepare strand-specific mRNA libraries that were sequenced by DNBSEQ high-throughput platform as PE100 reads. For the characterization of long/short full-length transcripts, 1 μg of the same RNA as the one used for RNA-seq experiments was used for cDNA synthesis using the Superscript IV kit (Invitrogen) and a modified oligodT (JG541) described in (54), followed by RNAse H treatment according to the manufacturer´s protocols. For the *TPF* expression in the T-DNA lines, 1 μg of total RNA extracted from 13-day-old seedlings was used for cDNA synthesis using the NZY First-Strand cDNA Synthesis kit (NZYTech) following the manufacturer´s protocol. Primer information can be found in Supplementary Data 1. The fold change was calculated

against a housekeeping gene (*PP2A*) following the AACt method (55).

5′ RACE

The 5′ RACE experiments were carried out by template switching reverse transcription using 200 ng of total RNA from two biological replicates employed in the RNA-seq experiment. First, total RNA (up to 2 μl) was incubated with 1.5 μl 3.3 M sorbitol/ 0.66 M trehalose, 1 μl oligodT 10 μM (Takara) and 0.5 μl 10 mM dNTPs at 65°C for 5 min. Then, the mix was added to 2 μl betaine 5 M, 2 μl Superscript II Buffer, 0.5 μl 0.1 M DTT, and 0.5 μl Superscript II (Invitrogen). After a 5 min incubation at 42°C, 0.25 μl of 400 μM Template Switching Oligo (TSO) recommended by NEB (JG656) was added and the reaction was kept at 42°C for 45 min more, followed by a final 10 min incubation at 70°C. A first PCR using High-Fidelity Phusion-Enzyme polymerase (Thermo Scientific) with a gene-specific oligo and an oligo binding to the TSO adapter (JG657), was followed by a nested PCR using MyTaq Red polymerase (Bioline) with a second gene-specific oligo and the oligo binding to the TSO adapter (JG657). After running the gel, bands from 350 to 650 bp corresponding to one biological replicate of Col-0 and *tpf* mutant were pooled-purified and cloned into pCR™4-TOPO™ TA vector (Invitrogen) following the manufacturer´s protocol. Twenty-four colonies were screened by colony PCR to identify the different fragments cloned. Plasmid DNA from at least 11 of these colonies that contained fragments representative of the different band sizes identified in the colony PCR experiment were Sanger sequenced for each gene and background tested.

Chromatin immunoprecipitation (ChIP-seq)

The chromatin immunoprecipitation protocol was done as previously described with minor changes (56). Briefly, 1 g of inflorescences were ground in liquid nitrogen and fixed in Nuclei Isolation Buffer containing 1% formaldehyde for 10 min at RT. In the case of 3xFLAG-MINU2 ChIPs, samples were first fixed in Nuclei Isolation Buffer containing 1.5 mM EGS for 20 min at RT followed by the addition of 1% formaldehyde and incubation for 10 additional minutes. Reactions were stopped with glycine followed by nuclei isolation and chromatin shearing using a Bioruptor Pico (Diagenode). Chromatin was immunoprecipitated overnight at 4°C with the following antibodies: Anti-FLAG M2 (5 μl/ChIP used, F1804, Sigma), H3K4me3 (5 μl/ChIP used, 04-745, Millipore), H3K36me3 (10 μl/ChIP used, Ab9050, Abcam), H3K27me3 (10 μl/ChIP used, 07-449, Millipore), H3 (5 μl/ChIP used, Ab1791, Abcam), H2A.Z (3 μl/ChIP used, (63)), and PanH3Ac (5 μl/ChIP used, 39140, active motif). Complexes were captured with a 1:1 mixture of magnetic Protein A and Protein G Dynabeads (Invitrogen) for 3 h at 4°C, washed with low salt, high salt, LiCl, and TE buffers for 10 min each at 4°C, and eluted for 2 × 20 min at 65°C with elution buffer. Reverse crosslink was done overnight at 65°C, followed by proteinase K treatment at 45°C for 5 h. DNA was purified using phenol:chloroform:isoamyl alcohol 25:24:1 (Fisher Scientific) and precipitated with GlycoBlue (Invitrogen) and NaAc/EtOH overnight at −20°C. DNA was resuspended in 75 μl of elution buffer. Libraries for sequencing were prepared using the Ovation Ultra Low System V2 1–16 kit (NuGEN) following the manufacturer's instructions. Libraries were sequenced in a HiSeq 4000 (Histone and TPF1 experiments) and HiSeq 2500 (TPF2 and MINU2 experiments) as SE50 reads.

Micrococcal nuclease assay (MNase-seq)

0.35 g of inflorescences from Col-0, *minu1-2 minu2-1*, and *tpf1-1 tpf2-1* (2 biological replicates per background) were ground and resuspended in 20 ml of Isolation Buffer 1 (20 mM Tris pH 8, 0.3 M sucrose, 5 mM $MgCl_2$, 0.2% Triton X-100, 5 mM BME, 35% glycerol, 0.1 μM PMSF, Complete mini –EDTA). After filtering through 1 layer of Miracloth (Merck, cat#475855), samples were spun at 3200 × *g* at 4°C for 20 min. The pellet was resuspended in 1 ml of HBB buffer (25 mM Tris–HCl, pH 7.6, 0.44 M sucrose, 10 mM $MgCl_2$, 0.1% Triton X-100 and 10 mM β-mercaptoethanol, 0.1 μM PMSF, Complete mini –EDTA) and spun at 3200 × *g* at 4°C for 10 min. Then, the pellet was resuspended in 1 ml of MNB buffer (10% sucrose, 50 mM Tris–HCl, pH 7.5, 4 mM $MgCl_2$ and 1 mM $CaCl_2$) and spun at 3200 × *g* at 4°C for 10 min. Finally, the pellet was resuspended in 900 μl of MNB buffer and 180 μl aliquots were digested with different amounts of Micrococcal Nuclease (N3755, Sigma) (0, 1, 3, 5 and 8U) followed by 10 min incubation at 37°C, quickly mixing samples every 3 min. Reactions were stopped by the addition of 20 μl of 0.5 mM EDTA. 4 μl of DNAse-Free RNAse (10 mg/ml) were added followed by incubation at 37°C for 1 h. 1 μl proteinase K (20 mg/ml) was added followed by incubation at 45°C for 1 h. DNA was extracted with Phenol:Chloroform:isoamyl followed by chloroform extraction and precipitation with NaOAc/glycogen/ethanol at −20°C overnight. Samples were centrifuged at 13 000 × *g* for 15 min at 4°C, washed with 70% ethanol and resuspended in 30 μl elution buffer. Samples were run in a 2% agarose gel to recover mononucleosomal DNA (∼150 bp) from digestions that presented 20% di-nucleosome and 80% of mononucleosome. Gel bands were cut and purified with Zymo gel extraction kit. Libraries for sequencing were prepared at BGI following the BGISEQ-500 ChIP-Seq library preparation protocol and were sequenced by DNBSEQ high-throughput platform as PE100 reads.

ChIP-seq data analysis

Read quality was checked with FastQC v0.11.9. Reads were then mapped to the TAIR10 genome using Bowtie2 v2.4.4 (57) with default parameters. Duplicates were marked with picard MarkDuplicates v2.26.6 (http://picard.sourceforge.net/). Reads that were either unmapped, duplicated or multimapped were discarded using the samtools (v1.14) (58) view command with options '-F 4 -F 1024 -q 5'. Chipseq-greylist v1.0.2 (https://github.com/roryk/chipseq-greylist) was used to detect anomalous enrichment in control samples. These regions were removed from a TAIR10 chromosome bedfile with bedtools v2.30.9 (59) complement to build a bedfile of included regions. Reads that

overlapped the included regions were counted to calculate Reads Per Million (RPM) normalization factors using samtools v.1.14. Fragment sizes of ChIP libraries were estimated with phantompeakqualtools v1.2.2. (60). ChIP coverage bedgraphs were built using the bedtools (v2.30.9) genomecov command, specifying '-scale < RPM factor> -fs < fragment size>' as parameters. These files were then converted to bigwig using bedGraphToBigWig from UCSC tools. ChIP bigwigs were $\log_2$ ratio-normalized against their controls with the bigwigCompare command from deepTools v3.5.1. (61). Coverage and log2 ratio bigwigs from replicates were averaged using wiggletools (v1.2.11) mean (62). Histone PTM ChIP-seqs were compared to a previously published dataset (63) to support the reproducibility of our data.

Peaks were called using MACS v2.7.1 (64) with options '–keep-dup all -g 119482427 –nomodel –extsize < fragment size> –call-summits'. Intersections between peaks were performed using the merge command from bedtools v2.30.9 (59). Reads overlapping common peaks were counted using the multiBamSummary command from deeptools v3.5.1. (61). Venn Diagrams were drawn using eulerr v6.1.1 and the correlation between replicates was calculated using the cor function in R v4.1.2. Peak summits were annotated with genomic features and genes by intersecting them against the Araport11 annotation with the intersect command from bedtools v2.30.9 (59). Genes where summits overlapped the gene body or were less than 200 bp away from the TSS were considered as targets. Conversely, genes that did not meet these criteria or overlapped enriched regions in ChIP controls were categorized as not targets.

MNase-seq data analysis

Read quality control and alignment were performed as described for ChIP-seq data analyses. Fragment size distributions were obtained with the samtools stats command from samtools v1.14. (58). As in ChIP-seq analyses, regions with anomalous enrichment were detected with chipseq-greylist v1.0.2 and removed from a TAIR10 chromosome bedfile using bedtools (v2.30.9) (59) complement. The resulting genomic intervals were converted to wig using a python script. DANPOS v2.2.2. (65) was used to obtain normalized coverage wigs and poisson difference wigs, as well as to carry out analyses of differential nucleosome dynamics. The dpos.py command of DANPOS was run with parameters '–save 1 –paired 1 –span 1 –clonalcut 1e-10 –nor_regions_file < included regions>'. Coverage wig files were converted to bedgraph format using the write_bg command from wiggletools v1.2.11 (62), and these were converted to bigwig with the bedGraphToBigWig command from UCSC tools. Additional MNase-seq raw files from (66) were downloaded and processed to evaluate the similarity of our dataset to previously published data.

Nucleosome bedfiles were built from the output excel files of DANPOS using python scripts. Dyad summits were annotated to their closest gene using the bedtools v2.30.9 (59) closest command and the Araport11 annotation. The +1 nucleosome was identified as the first nucleosome with significant occupancy (summit *P*-value < 0.05) downstream of the TSS position. To confidently detect differences in nucleosome occupancy and fuzziness, only nucleosomes with significant occupancy in at least one condition were considered. Furthermore, to evaluate position shifts, only those that displayed significant occupancy (summit *P*-value < 0.05) and fuzziness (fuzziness *P*-value < 0.05) in all conditions were considered. This is because nucleosome shift values between not well-positioned nucleosomes tend to be noisy.

RNA-seq data analysis

Quality control of reads was first performed with FastQC v0.11.9. Reads were then aligned to the TAIR10 genome with STAR v2.7.9a (67). Gene counts were obtained with featureCounts v2.0.1 (68) using the Araport11 annotation. Differential expression analyses were carried out with DESeq2 v1.32.0 (69). The PCA was computed with the plotPCA function over r-log normalized counts from DESeq2 (69). The lfcShrink function was used to obtain differential expression scores with default parameters (*P*-values were adjusted using the Benjamini–Hochberg method). Genes with adjusted *P*-values < 0.05 and $|\log_2\text{FoldChange}| > \log_2(1.5)$ were considered as differentially expressed. Gene Ontology enrichment analyses were performed over these genes using clusterProfiler v4.0.5 (70) and org.At.tair.db v3.13.0. Coverage bedgraph files were built using bedtools v2.30.9 (59) genomecov with parameters '-bga -split -scale < size factor>', in which size factors were normalization factors estimated by DESeq2. These were converted to bigwigs using the bedGraphToBigWig command from UCSC tools. The wiggletools v1.2.2 (62) mean command was used to average bigwigs of biological replicates. Finally, log2 ratio bigwigs were built using bigwigCompare from deepTools v3.5.1 (61).

Cis element analysis

To find cis elements occluded by the +1 nucleosome shift in *tpf* and *minu* mutants, we first generated bedfiles using the 5' ends of +1 nucleosomes of Col-0 and mutants as start and end coordinates respectively. The covered regions were merged so that they encompassed the maximum amount of shift observed in *tpf* and *minu* mutants. This was achieved using the bedtools v2.30.9 (59) merge command. The DNA sequences located in these regions were then obtained with the bedtools v2.30.9 (59) getfasta command using the TAIR10 genome as reference. The following analyses were performed over these sequences using the MEME suite v.5.4.1 (71) and specifying JASPAR CORE 2022 plants as reference database. MEME-ChIP was used for motif discovery and SEA was used to find known transcription factor (TF) cis elements enriched in the sequences. Both programs were run using the default parameters.

For core promoter elements analyses, PWM matrices for canonical TATA-box (5'-TATA[A/T]A[A/T]-3'), Y-patch (5'-CTCTTCTT-3') and GAGA (5'-GAGAGAGA-3') were first generated using the matrix2meme command from MEME suite v.5.4.1. (71). The GAGA and Y-patch sequences were taken from (72). Y-patch and GAGA elements were compared to the motifs that resulted from MEME-ChIP using TomTom from the MEME suite v.5.4.1 (71).

FIMO was then used to find instances of TATA boxes, Y-patches and GA repeats over TPF-occupied genes in a 400 bp window centred in the TSS. Motif frequency bedgraphs were obtained using bedtools v2.30 (59) genomecov. These files were converted to bigwig using the UCSC tool bedGraphToBigWig.

Data visualization

Statistical graphics such as bar plots, scatter plots, and boxplots were drawn using ggplot2 v3.3.6. Coverage bigwigs were visualized in the IGV browser and screenshots were rendered using SparK v2.6.2. Heatmaps and metaplots were obtained using the computeMatrix, plotHeatmap, and plotProfile commands from deeptools v3.5.1. (61). The gene ordering in heatmaps was controlled using the –sortUsingSamples and –sortRegions parameters from plotHeatmap. TSS reference positions for metaplots were obtained from Araport11 except for Figures 4E, 5B, and Supplementary Figures 12A, 12B, and 15C where they were obtained from an analysis performed in (73) (https://github.com/Maxim-Ivanov/Kindgren_et_al_2019/blob/master/Adjusted_Araport11/genes_araport_adj.bed.gz) which relies on TSS-seq and DR-seq experiments to precisely annotate TSSs.

RESULTS

tpf mutants have a strong impact on plant development and reproduction

The model plant Arabidopsis possesses two paralogs of the TPF proteins named TPF1 and TPF2 which were previously shown to belong to a plant BAF-like complex in Arabidopsis (35). To study the impact of TPF proteins on plant development, we identified T-DNA insertion lines for the *TPF1* and *TPF2* genes. For *TPF1*, we identified two lines with insertions located in the first exon (*tpf1-1*) and the first intron (*tpf1-2*) (Supplementary Figure 1A). For *TPF2*, we identified two lines, *tpf2-1* and *tpf2-2,* where the T-DNA was inserted in the 11th exon and the intron between exons 10 and 11, respectively (Supplementary Figure 1A). The expression of *TPF1* in the *tpf1* mutant lines was reduced compared to wild-type Col-0 plants (WT) (Supplementary Figure 1B, Supplementary Data 2). In the case of the *tpf2* alleles, both showed *TPF2* expression levels upstream of the T-DNA insertion similar to WT and reduced expression downstream of the insertion (Supplementary Figure 1C, Supplementary Data 2). According to a public database (www.genevestigator.com) the overall expression pattern of *TPF* genes is similar, showing medium-high expression levels and a stronger difference in senescent tissue (Supplementary Figure 1D). Single mutants of *TPF1* and *TPF2* did not show any apparent developmental differences compared to WT plants (Supplementary Figure 2A). However, combinations of the *tpf1* and *tpf2* alleles resulted in two different mutant phenotypes (Supplementary Figure 2B, C). While combinations of *tpf1-1* and *tpf1-2* with *tpf2-1* resulted in dwarf plants with strong developmental phenotypes, double mutants incorporating the *tpf2-2* allele showed milder defects, suggesting that *tpf2-2* is a weaker allele (Supplementary Figure 2B, C). Thus, further phenotypical characterizations were performed in the strong combination *tpf1-1 tpf2-1* and the weak combination *tpf1-2 tpf2-2*, hereafter named strong and weak *tpf1/2* mutants, respectively. Seedlings of the strong and weak *tpf1/2* mutants were smaller and developmentally delayed and presented shorter roots compared to WT plants (Figure 1A, B), which is probably the consequence of shorter meristems with fewer numbers of cells (Figure 1C, Supplementary Figure 3A, Supplementary Data 3). Interestingly, the niche of meristematic cells above the quiescent center incorporated high amounts of propidium iodide (PI) in the strong *tpf1/2* mutant (Figure 1C, Supplementary Figure 3B). This phenotype normally occurs in damaged or dead cells (74) thus suggesting that this pool of cells might have increased DNA damage response in the *tpf1/2* mutant. Adult plants of the strong and weak combinations were smaller and developmentally delayed (Figure 1D, E). The strong *tpf1/2* mutant showed pleiotropic defects with variable penetrance on flower development. Among the most frequent phenotypes were flowers with extra petals, open carpels, protuberances striking out of the stigma (Figure 1F), and defective anther dehiscence (Supplementary Figure 3C). Moreover, it produced small siliques with no seeds (Figure 1G), although very few seeds could be recovered in some mutant plants. The siliques from *tpf1-1* (+/–) *tpf2-1* (–/–) plants presented a high number of abortions and the segregation of double mutant plants was smaller than expected (3:0.38 +/- 0.07 observed; 3:1 expected; Average and standard error calculated from three independent *tpf1-1* (+/–) *tpf2-1* (–/–) populations (Figure 1H, Supplementary Figure 3D, Supplementary Data 4). In order to gain further insight into the sterility phenotype of the strong *tpf1/2* mutant, we performed reciprocal crosses with the WT (Supplementary Figure 3E, Supplementary Data 5). When WT pollen was used to pollinate mutant pistils, siliques did not elongate and just four seeds were recovered out of 28 crosses (Supplementary Figure 3E). On the contrary, when mutant pollen was used on WT pistils, siliques elongated and produced viable seeds (Supplementary Figure 3E), suggesting that the fertilization problem is mainly associated to the female gametophyte. This sterility phenotype is further enhanced by the dehiscence defects in the mutant anthers at the moment of anthesis when self-pollination takes place (Supplementary Figure 3C). In summary, this data shows that the TPF proteins are functionally redundant in the control of several aspects related to plant development and reproduction. Importantly, these phenotypes are reminiscent of the ones reported for double mutants of the functionally redundant MINU1 and MINU2 proteins (75) (Supplementary Figure 2D and 4). This is consistent with the idea that TPF proteins are important players in the overall function of the MAS complex.

TPF proteins are members of the MAS complex in plants

Through IP-MS experiments, we recently showed the composition of the MAS complex in plants (35). To further confirm the complex composition and the mutually exclusive presence of TPF paralogs, we performed additional IP-MS experiments in inflorescences of Arabidopsis using transgenic lines expressing 3xFLAG-tagged TPF2 and MINU2 proteins. Importantly, these transgenes were able to comple-

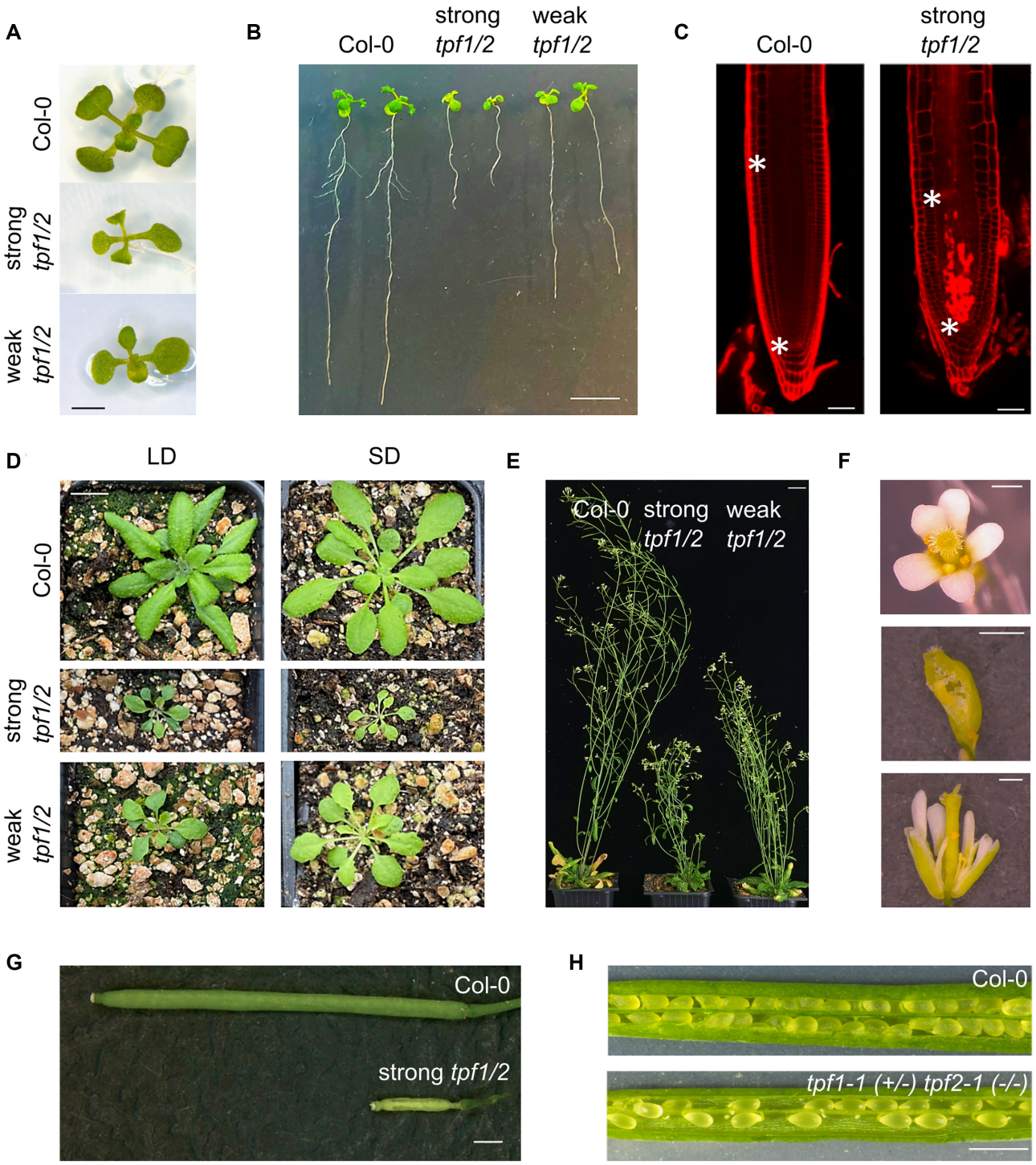

Figure 1. *tpf* mutants have a strong impact on plant development and reproduction. (**A**, **B**) *In vitro*-grown twelve-day-old seedlings of strong and weak *tpf1/2* mutants show reduced aerial and radicular growth, scale bars: 2 mm and 1 cm, respectively. (**C**) Confocal images of twelve-day-old root tips of Col-0 and the strong *tpf1/2* mutant stained with propidium iodide (PI). Asterisks indicate meristem length. A strong PI signal accumulates above the quiescent center in the *tpf1/2* mutant. Defects in the division plane can be observed in the strong *tpf1/2* mutant. scale bar: 50 μm. (**D**) Top view of three-week-old Col-0, strong, and weak *tpf1/2* mutants grown on soil in long-day (16L/8D) conditions and four-week-old plants of the same backgrounds grown on short-day (8L/16D) conditions, scale bar: 1 cm. (**E**) Seven-week-old Col-0, strong, and weak *tpf1/2* mutants grown on soil in long-day conditions, scale bar: 2 cm. (**F**) Representative images of the developmental defects found in flowers of the strong *tpf1/2* mutant. Extra petals (upper), open carpels (middle), and protuberances in the stigma (lower), scale bar: 0.5 mm. (**G**) Representative image of siliques of Col-0 and strong *tpf1/2* mutant plants, scale bar: 1 mm. (**H**) Representative images of open siliques of Col-0 and *tpf1-1 (+/–) tpf2-1 (–/–)* showing ovule abortion, scale bar: 1 mm.

Table 1. Interactors of TPF2 and MINU2 identified by mass spectrometry

		log_2 LFQ intensities			Peptides			Unique peptides		
	Protein IDs	Col-0	TPF2–15	TPF2–9	Col-0	TPF2–15	TPF2–9	Col-0	TPF2–15	TPF2–9
ARP4	AT1G18450	25	32	33	6	22	25	6	22	25
ARP7	AT3G60830	-	32	33	0	13	17	0	13	17
BDH1	AT4G22320	-	30	31	0	7	9	0	7	9
BDH2	AT5G55210	-	31	32	0	5	7	0	5	7
BRD5	AT1G58025	-	31	32	0	13	19	0	13	19
BSH	AT3G17590	-	30	32	0	6	9	0	6	9
SWP73B	AT5G14170	-	32	33	0	25	27	0	25	27
LFR	AT3G22990	-	32	33	0	12	14	0	12	14
MINU1	AT3G06010	-	32	33	0	45	53	0	36	43
MINU2	AT5G19310	-	31	31	0	33	39	0	24	29
OPF1	AT1G50620	-	31	32	0	18	22	0	18	21
OPF2	AT3G20280	-	28	30	0	4	13	0	4	12
PSA1	AT1G32730	-	32	33	0	17	20	0	17	20
PSA2	AT1G06500	-	31	32	0	7	9	0	7	9
SHH2	AT3G18380	-	30	30	0	14	16	0	14	16
SWI3A	AT2G47620	-	32	33	0	28	42	0	28	42
SWI3B	AT2G33610	-	33	33	0	23	30	0	23	30
TPF2	AT3G08020	-	36	38	0	45	60	0	42	57
	Protein IDs	Col-0	MINU2–10	MINU2–7	Col-0	MINU2–10	MINU2–7	Col-0	MINU2–10	MINU2–7
ARP4	AT1G18450	25	33	32	6	22	22	6	22	22
ARP7	AT3G60830	-	34	32	0	16	15	0	16	15
BDH1	AT4G22320	-	32	31	0	7	6	0	7	6
BDH2	AT5G55210	-	31	32	0	5	5	0	5	5
BRD5	AT1G58025	-	32	31	0	19	19	0	19	19
BSH	AT3G17590	-	32	31	0	7	7	0	7	7
SWP73B	AT5G14170	-	33	32	0	27	25	0	27	25
LFR	AT3G22990	-	33	32	0	14	12	0	14	12
MINU2	AT5G19310	-	34	33	0	44	44	0	34	34
OPF1	AT1G50620	-	32	31	0	20	15	0	19	15
OPF2	AT3G20280	-	31	30	0	16	15	0	15	15
PSA1	AT1G32730	-	34	32	0	18	16	0	18	16
PSA2	AT1G06500	-	32	31	0	9	7	0	9	7
SHH2	AT3G18380	-	31	29	0	19	16	0	19	16
SWI3A	AT2G47620	-	33	32	0	34	35	0	34	35
SWI3B	AT2G33610	-	33	33	0	25	25	0	25	25
TPF1	AT3G52100	-	35	34	0	35	30	0	32	27
TPF2	AT3G08020	-	33	33	0	37	37	0	34	34

Data represents values of log_2LFQ intensity, peptides and unique peptides from an IP-MS experiment using samples from two independent transgenic lines of TPF2-3xFLAG (TPF2–15 and –9) and 3xFLAG-MINU2 (MINU2–10 and –7) and one untransformed Col-0 control. Shown interactors were selected based on the previous identification of MAS complex subunits using TPF1 and SHH2 as baits (35) and followed the criteria: (i) LFQ control < LFQ transgenic/10 in both transgenic lines compared to Col-0 control, (ii) presence of unique peptides in both transgenic lines. Raw data can be found in Supplementary Data 14.

ment the strong *tpf1/2* and the *minu1-2 minu2-1* (*minu1/2*) double mutant backgrounds, respectively (Supplementary Figure 4). These experiments confirmed the presence of a well-defined MAS complex that incorporates one TPF subunit and specifically the MINU1 and MINU2 ATPases, as well as a set of known and several uncharacterized subunits (Table 1). The mammalian TPF distant ortholog DPF2 directly interacts with several subunits of the SWI/SNF complex including ARID1 (plant LFR) and SMARCD (plant SWP73) (31). We tested the direct interaction between TPF1 and all the identified subunits in the complex by Y2H. For this experiment we split TPF1 into an N-terminal (N-TPF1) fragment that includes the 3xPHD domains and a C-terminal fragment that includes the Tudor domain (C-TPF1) (Supplementary Figure 5A). The N-TPF1 fragment was able to interact with SWP73B and LFR (Supplementary Figure 5B), suggesting conservation of the TPF1 position within the complex compared to its mammalian counterpart. Surprisingly, the same experiment using the TPF1 full-length and C-TPF1 fragments did not result in any positive interactions, probably due to steric impediments or misfolding of the fusion proteins.

TPF and MINU2 proteins locate over the 5′ regions of thousands of genes

The genome-wide location of the MAS complex has not been reported in plants. We profiled the location of the MINU2 ATPase, as a core member of the complex, as well as the location of the TPF1 and TPF2 proteins to study whether chromatin-bound MAS complex incorporates TPF proteins at all targets. For this purpose, we performed ChIP-seq experiments with two independent transgenic lines expressing 3xFLAG-tagged TPF1, TPF2, and MINU2 (Supplementary Data 6). The two lines tested showed a highly reproducible profile (Supplementary Figure 6A, B). We found

14419 peaks for TPF1, 15337 peaks for TPF2, and 12169 peaks for MINU2 (Figure 2A). Consistent with their redundant function at the phenotypical level, the overlap of TPF peaks was very high (Figure 2A,B,C). These peaks were mostly located over the region downstream of the TSS of thousands of genes (Figure 2B–D). The overlap between MINU2 and TPF peaks was also very high, although the position of MINU2 position was slightly shifted towards the promoter region, perhaps reflecting the different positions occupied by TPF and the ATPase within the complex (Figure 2A–D). The target genes were enriched with epigenetic marks related to transcription activation such as H3K4me3, H3 pan acetylation, H3K36me3, and H2A.Z (over the TSS regions) while non-target genes were enriched in silencing marks such as H3K27me3 and H2A.Z (over gene body regions) (Figure 2E, Supplementary Figure 7A, B, Supplementary Data 7). Consistently, direct targets were more expressed than non-targets (Figure 2F) and the accumulation of the complex positively correlated with gene expression (Figure 2G). In summary, the MAS complex locates over the 5′ regions of euchromatic expressed genes with an epigenome favourable for transcription. Moreover, TPF1 and TPF2 share a large number of targets with MINU2 suggesting they are core members of the chromatin-bound MAS complex.

The MAS complex controls the +1 nucleosome positioning of thousands of genes

In order to investigate the role of the MAS complex on the genome-wide nucleosome landscape and study the contribution of TPF subunits to the complex′s activity, we performed MNase-seq experiments comparing the *minu1/2,* and the strong *tpf1/2* mutants (from now on *minu* and *tpf* mutants, respectively) with WT plants (Supplementary Data 6). This technique relies on the function of the Micrococcal Nuclease (MNase), which has endo-exonuclease activity and degrades faster the DNA that is not associated with nucleosomes, thus allowing the inference of the nucleosome positioning and the differences between nucleosomes in mutant and WT backgrounds. The MNase-seq replicates were highly reproducible (Supplementary Figure 8A, B). Moreover, the WT samples were very similar to a previously published MNase-seq experiment using chromatin from the same tissue (Supplementary Figure 8B, C) (66). The changes observed in nucleosome position can be separated into changes in position, fuzziness, and occupancy (Figure 3A). Both the *tpf* and *minu* mutants presented changes in all three events but there was a clear preference for the control of positioning (Figure 3A, Supplementary Figure 9A, B). A visual inspection of the MNase-seq tracks revealed an overall shift of the +1 nucleosome towards the NDR, which was further confirmed with metaplot and heatmap representations over TPF target genes (Figure 3B–D, Supplementary Figure 9A, Supplementary Data 8). Importantly, this trend was not observed over non-target genes suggesting a direct implication of the MAS complex (Figure 3C, E). The distance between nucleosomes over target loci was not affected suggesting that this complex does not play a major role in nucleosome phasing. The shift observed was variable among genes and, on average, represented a 10–15 bp shift although some loci showed a consistent shift of >40 bp towards the NDR (Figure 3D, E). In summary, this data shows that the MAS complex is mainly involved in the control of the nucleosome positioning and prevents +1 nucleosome shift towards upstream regions.

The MAS complex affects the transcription start site determination of thousands of genes with no major impact on differential gene expression

Previous studies in yeast have shown that +1 nucleosome shift results in overall gene repression due to interference with the accessibility of the transcriptional machinery to the promoter (15,16,20). Thus, we performed RNA-seq experiments comparing the *tpf* and *minu* mutants with WT plants (Supplementary Figure 10A, Supplementary Data 6). Consistent with the strong developmental phenotype of these mutants, we found thousands of differentially expressed genes (DEGs), of which 2436 and 1225 were upregulated and 4081 and 2663 were downregulated in the *tpf* and *minu* mutants, respectively (Figure 4A, B, Supplementary Data 9). As expected, the DEGs from the two mutants showed a high degree of overlap although the *tpf* mutant showed a larger number of up- and down-DEGs (Figure 4B, Supplementary Figure 10B). This is consistent with the stronger developmental phenotype of the *tpf* mutant compared to *minu*, which is a weak combination of *minu1* and *minu2* alleles (75) (Supplementary Figure 2D and 4). Importantly, the combination of strong *minu1 minu2* alleles causes embryo lethality, which hampers the ability to work with such material (75). In line with these results, both mutants showed similar enriched gene ontology categories, which were especially related to reproductive processes and response to biotic and abiotic stresses (Supplementary Figure 10C). Interestingly, we found that multiple subunits that are specific to the MAS complex were upregulated in the *tpf* and *minu* mutants while subunits specific to the plant ncBAF-like complex were not (Supplementary Figure 10D). This suggests that there is a negative feedback mechanism to upregulate the expression of the MAS-specific subunits in situations where the complex is not functional or is highly demanded.

In order to identify genes that are potentially regulated by the MAS complex, we selected TPF-occupied genes (which result from the intersection between TPF1 and TPF2 target genes) and looked for their overlap with DEGs. To work with highly confident MAS complex-regulated genes, we used the list of DEGs commonly misregulated in *tpf* and *minu* mutants (Figure 4B). We identified 427 direct targets that were upregulated and 531 targets that were downregulated, which together represented 6% of all direct targets (Figure 4C, Supplementary Data 10). This result indicates that most complex targets are not misregulated under the conditions where samples were taken. Perhaps sampling under more dynamic conditions would reveal a stronger effect. The number of DEGs that showed a +1 nucleosome shift was around one-third of the total number of direct DEGs (155/427 for up-DEGs and 152/531 for down-DEGs) but this trend was also shared by targets that were not differentially expressed, suggesting a small correlation between the +1 nucleosome shift and differential expression (Figure 4D). We compared the nucleosome profile over direct-

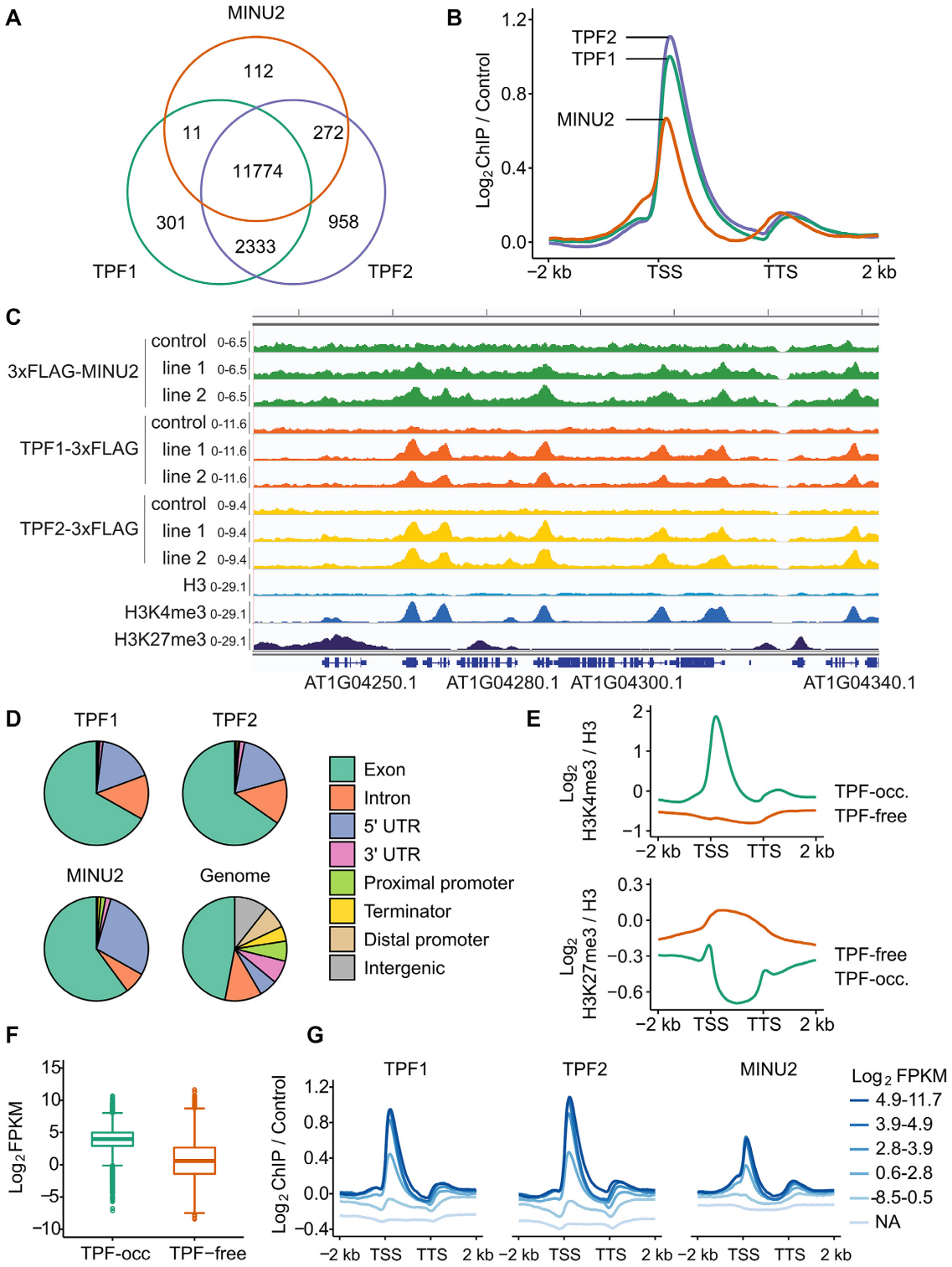

Figure 2. TPF and MINU2 proteins locate over the 5′ regions of thousands of genes. (**A**) Overlap between TPF1, TPF2, and MINU2 peaks. The peaks of each subunit are the union set of peaks detected (MACS2 $q < 0.05$) in ChIP-seqs from two independent transgenic lines. (**B**) Average occupancy of TPF1, TPF2, and MINU2 over the intersection of TPF1 and TPF2 target genes ($n = 15859$), hereafter referred to as TPF-occupied genes. (**C**) Browser screenshot showing the reproducibility of TPF and MINU2 peaks between lines. Notice these peaks co-localize with H3K4me3 and are absent in H3K27me3-enriched regions. Coverage values are Reads Per Million (RPM) mapped reads. (**D**) Percentage of TPF1, TPF2, and MINU2 peak summits overlapping genomic features. Proximal promoters and terminators were defined as regions 500 bp upstream or downstream of the TSS or TTS, respectively. Distal promoters ranged from 500 bp to 2000 bp upstream of the TSS. Other non-genic regions were classified as intergenic. (**E**, **F**) Average occupancy of H3K4me3 and H3K27me3 (E), and gene expression (F) over genes occupied by TPF (n = 15859) and genes lacking TPF peaks (hereafter TPF-free genes, $n = 21574$). (**G**) Average occupancy of TPF and MINU2 over genes ranked by their expression level. The lowest group consists of non-expressed genes (0 FPKM), while expressed genes were grouped into five quintiles of ascending expression.

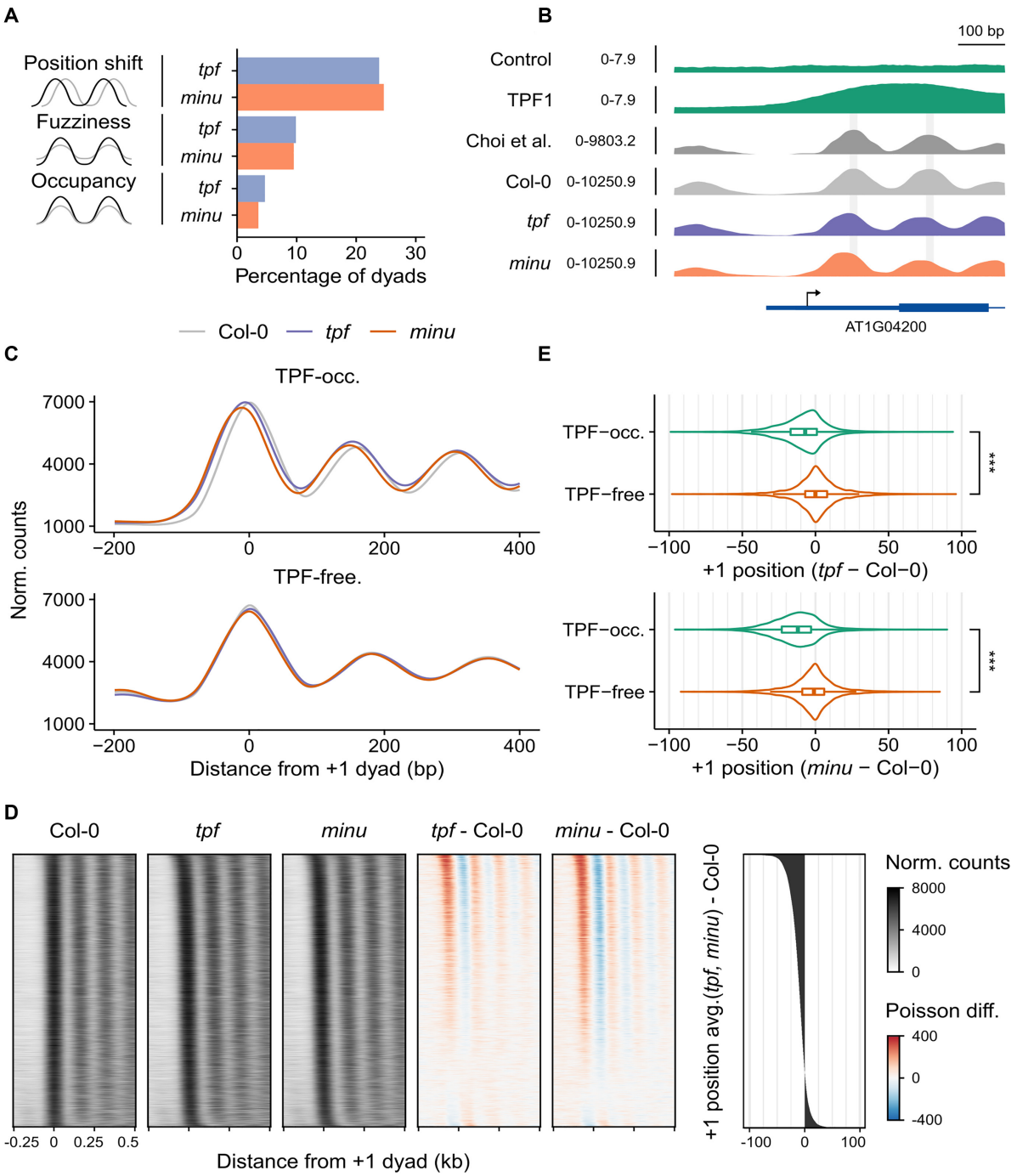

Figure 3. TPF and MINU control the position of the +1 nucleosome. (**A**) Cartoon depicting the three different nucleosome features analyzed. Percentage of dyads showing significant position shift (|mutant – Col-0 position| > 5 bp), fuzziness (FDR < 0.05) and occupancy (FDR < 0.05) differences in *tpf* and *minu* mutants compared to Col-0. Position shifts are differences between nucleosome summits of mutants and Col-0 samples, while occupancy and fuzziness differences are poisson tests comparing read counts and read position deviations between mutants and Col-0 samples, respectively. (**B**) Representative TPF-occupied locus in which dyads are shifted towards the TSS in *tpf* and *minu* compared to Col-0. In green, ChIP-seq signal from Col-0 control and the mean Reads Per Million (RPM) of TPF1 ChIPs performed with two independent transgenic lines. In dark grey, nucleosome coverage from an independent study (66). Light grey, blue and orange represent normalized nucleosome coverages of Col-0, *tpf* and *minu* mutants, respectively, from the mean of two biological replicates. In this locus, the +1 nucleosome is found 14 bp (*tpf*) and 23 bp (*minu*) upstream to the position of the Col-0 +1 nucleosome, scale bar 100 bp. The dominant TSS position derived from (74) is depicted as an horizontal arrow. (**C**) Average nucleosome occupancy in Col-0, *tpf,* and *minu* mutants over TPF-occupied ($n = 11768$) and TPF-free ($n = 8991$) genes that display well-positioned +1 nucleosomes. (**D**) Heatmaps representing nucleosome positions over genes that display well-positioned +1 nucleosomes in Col-0, *tpf,* and *minu* mutants ranked by the average +1 dyad shift observed in mutants. The first three heatmaps show normalized nucleosome occupancies of Col-0, *tpf,* and *minu* (average from two replicates), while the last two represent the poisson difference between Col-0 and mutant occupancies. The far-right plot shows the average +1 position shift (bp) between *tpf* and *minu* mutants ordered from most upstream to most downstream. (**E**) Distribution of the +1 dyad shift values in *tpf* and *minu* mutants compared to Col-0 over TPF-occupied and TPF-free genes with well-positioned nucleosomes. ***Welch two-sample *t*-test *P*-value < 2.2e–16

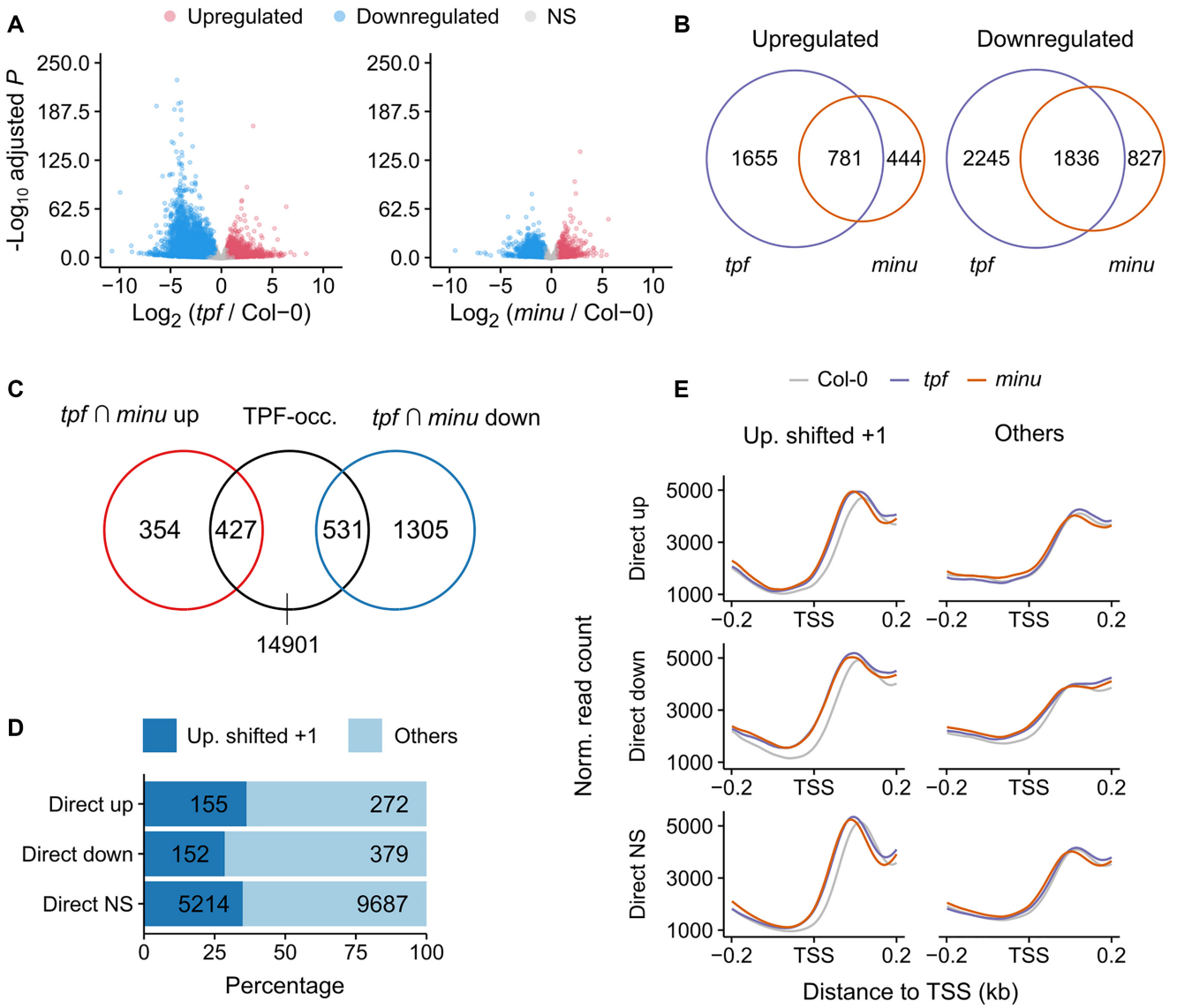

Figure 4. Changes in nucleosome positioning have a small impact on the differential expression of MAS complex targets. (**A**) Volcano plots showing $\log_2$ differential expression of *tpf* and *minu* mutants compared to Col-0. Genes with adjusted $P < 0.05$ and fold change >1.5 were considered differentially expressed. NS, not significant. (**B**) Overlap between upregulated and downregulated DEGs detected in *tpf* and *minu* mutants. (**C**) Overlap between TPF-occupied genes and shared DEGs of *tpf* and *minu* mutants. DEGs overlapping TPF-occupied genes were classified as either 'direct up' or 'direct down' depending on their misregulation direction. The rest were considered direct not significant (direct NS). (**D**) Percentage of direct up, direct down or direct NS genes showing an upstream shift at the +1 dyad of at least 5 bp in both *tpf* and *minu* mutants compared to Col-0. (**E**) Nucleosome occupancy profiles of direct up, direct down and direct NS genes in Col-0, *tpf* and *minu* mutants, centered on the TSS. Plots represent occupancy in loci that show a +1 nucleosome shift of at least 5 bp in both *tpf* and *minu* mutants compared to Col-0, or loci that do not show a shift (others). Occupancy values are normalized read counts averaged from two replicates.

up-DEGs, direct-down-DEGs, and direct-non-DEGs (Figure 4E). Interestingly, we found that the nucleosomal signal over the NDR region of the direct-down-DEGs was higher compared to the other two groups. This suggests that nucleosome invasion of the NDR over this set of genes could be responsible for the observed transcriptional downregulation in the *tpf* and *minu* mutants probably by interfering with the recruitment of the transcriptional machinery (15) (Figure 4E, Supplementary Figure 11).

Remarkably, looking at the RNA-seq tracks we found a consistent increase in the accumulation of reads over the 5′ regions of many genes, suggesting an upstream change in TSS usage in the *tpf* and *minu* mutants (Figure 5A–C). Importantly, the accumulation of these additional reads was only observed in genes targeted by the complex, suggesting that it may be participating in the TSS change (Figure 5B,C, Supplementary Figure 12A). This change in TSS usage did not have an overall impact on gene expression since a similar 5′ differential enrichment was found in all TPF-occupied genes (Supplementary Figure 12B), of which only a fraction was differentially expressed (Figure 4C, Supplementary Figure 12B). To confirm that these extended 5′ reads correspond to longer full-length transcripts, we analysed in Col-0 and the *tpf* and *minu* mutants the transcripts of five

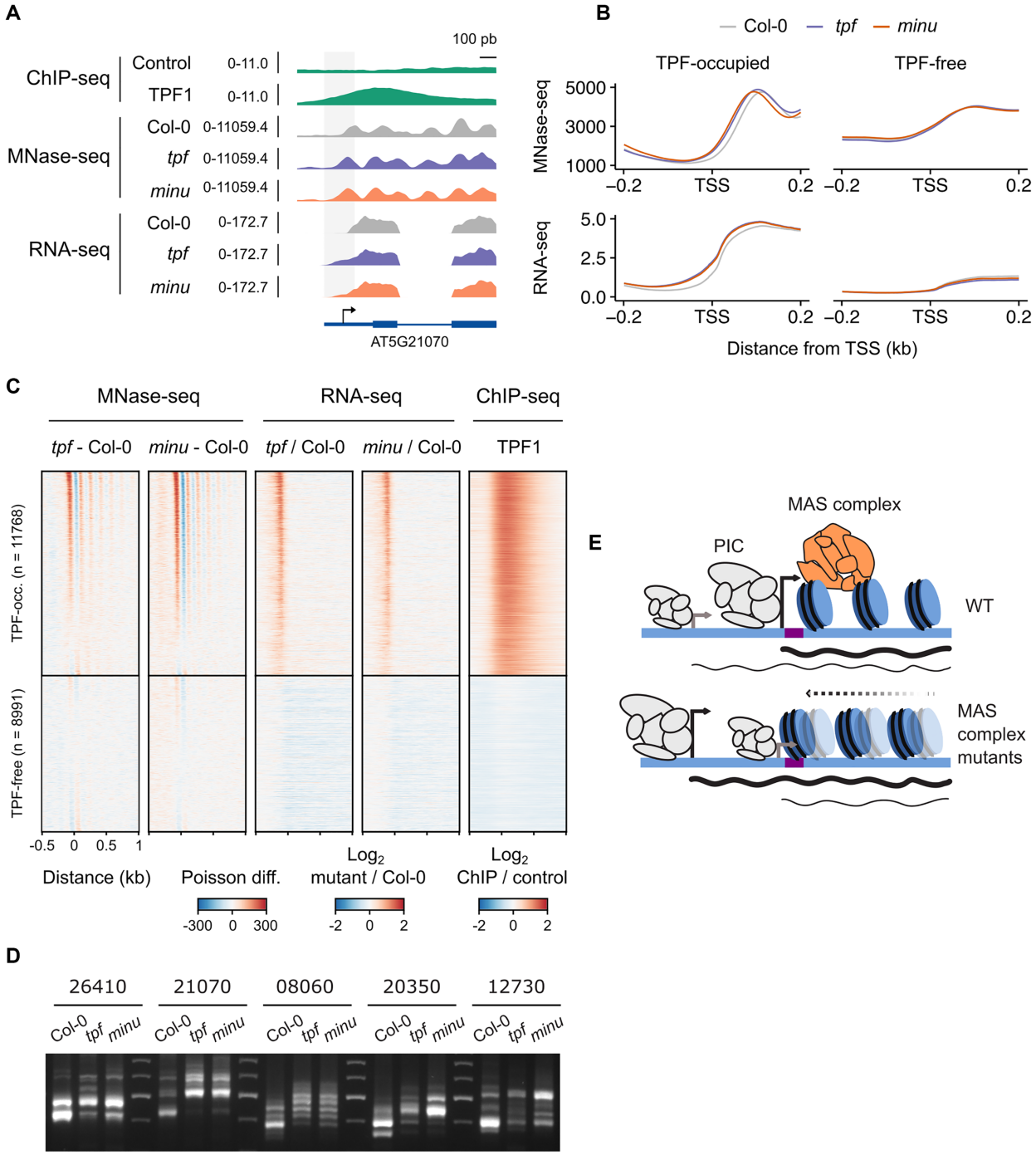

Figure 5. TPF and MINU affect the TSS determination of thousands of genes. (**A**) Example of a TPF-bound locus showing longer 5′ transcript length and +1 nucleosome shift towards the promoter in *tpf* and *minu* mutants compared to Col-0. ChIP-seq values are the mean reads per million (RPM) of ChIPs performed with two independent transgenic lines. MNase-seq and RNA-seq values are normalized counts averaged from two and three biological replicates, respectively. The dominant TSS position derived from (74) is depicted as an horizontal arrow. (**B**) Normalized nucleosome (upper part) and transcript coverage (lower part, shown in $\log_2$ scale) in Col-0, *tpf* and *minu* mutants over TPF-occupied and TPF-free genes that display well-positioned +1 nucleosomes. (**C**) Heatmap showing MNase-seq poisson differences between genes with a properly positioned +1 nucleosome, RNA-seq $\log_2$ fold changes, and TPF1 ChIP-seq $\log_2$ ratio values between transgenic lines and non-transgenic Col-0 control. Genes were sorted by the average +1 dyad position shift observed in *tpf* and *minu* mutants. (**D**) 5′ RACE analysis of five targets of the MAS complex displaying +1 nucleosome shift and 5′ extended transcripts from Col-0, *tpf* and *minu* samples. Results from one of two biological replicates is shown. A lane depicting a DNA ladder (400, 500, 600 and 700 bp bands) separates each gene. 26410: AT4G26410; 21070: AT5G21070; 08060: AT5G08060; 20350: AT5G20350; 12730: AT1G12730. (**E**) Proposed model of the role of the MAS complex on +1 nucleosome position and TSS usage. An upstream shift in the position of the +1 nucleosome in the *tpf* and *minu* mutants negatively impacts the function of promoter elements or the transcriptional machinery, triggering the preferential use of upstream alternative TSSs. The MAS complex is depicted in orange, the promoter elements in purple, and the transcriptional machinery in grey. The size of the transcriptional machinery cartoon represents its relative activity on alternative TSSs. The size of the arrows depicting alternative TSSs represents their preferential usage in each background (WT and MAS complex mutants).

genes that showed a consistent +1 nucleosome shift and extended 5′ RNA-seq reads (Supplementary Figure 13A, B). Results indicate an increased accumulation of longer transcripts in the *tpf* and *minu* mutants compared to Col-0 while overall expression levels were not affected (Supplementary Figure 13B, Supplementary Data 11). To further confirm the preferential use of upstream alternative TSSs in the *tpf* and *minu* mutant backgrounds, we performed 5′ RACE of the five selected genes. Samples from Col-0 presented multiple bands indicative of diverse TSS usage although a predominant band, most likely corresponding to the dominant TSS, was observed in all cases (Figure 5D, Supplementary Data 12). Importantly, a concomitant decrease in the intensity of this band and the accumulation of larger bands were observed in all five genes tested in the *tpf* and *minu* mutants, which is indicative of preferential use of alternative upstream TSSs (Figure 5D, Supplementary Data 12). This was confirmed by sequencing the DNA in these bands. Results show that both mutants and WT can transcribe from diverse TSSs, although mutants present a preference for upstream sites (Supplementary Figure 14).

The accumulation of 5′ reads in *tpf* and *minu* mutants showed a remarkable correlation with their +1 nucleosome shift (Figure 5B, C, Supplementary Figure 12C), thus pointing to a functional connection between the TSS determination and the +1 nucleosome repositioning. A shift in nucleosome position could interfere with the recruitment or the activity of the transcriptional machinery forcing the use of alternative upstream TSSs. To explore this possibility, we searched for enriched cis elements in the regions occluded by the shifted +1 nucleosomes in *tpf* and *minu* mutants (Supplementary Figure 15A). We identified two *de novo* motifs that highly resembled GA and Y-patch boxes (Supplementary Figure 15B). Importantly, these motifs peak in the same region where nucleosome occupancies are most altered (Supplementary Figure 15C). Moreover, several known TF binding sites were also significantly enriched in the occluded regions, including the Telobox, the GCC-box, and the dehydration response element (DRE) (Supplementary Data 13). Thus, TFs that recognize these promoter elements represent candidates whose binding to DNA could be affected upon the nucleosome shift. By contrast, TATA boxes were located in a more upstream position and did not overlap with the shifted +1 nucleosomes suggesting no effect by +1 nucleosome shifts (Supplementary Figure 15C). In summary, these results indicate that the MAS complex plays an important role in determining the TSS of its targets without affecting overall expression levels and also suggest that the +1 nucleosome repositioning could be functionally involved in the TSS change through steric impediments with the transcriptional machinery.

DISCUSSION

We recently identified an evolutionarily conserved plant SWI/SNF complex reminiscent of the mammalian BAF subclass (35). In the present study, we report the function of the MAS complex, which regulates the position of the +1 nucleosome of thousands of genes and the TSS determination. Moreover, we report experimental evidence that establishes the functional relevance of TPF proteins for the complex activity based on the similarities between the *tpf* and *minu* mutants.

TPF proteins play an important role in the function of the MAS complex

Our results show that double *tpf* mutants present developmental phenotypes similar to the ones reported for *minu* mutants (75). Moreover, *tpf* and *minu* mutants show similar molecular phenotypes, including changes in gene expression, +1 nucleosome shift, and change in TSS usage. This evidence strongly suggests that TPF proteins play a prominent role in the overall complex function. The characterization of the genome-wide TPF location shows a very high overlap with MINU2, reinforcing the idea that they are core components of the same complex. The fact that strong *minu* alleles cause more severe developmental defects than the strong combination of *tpf* alleles identified in our study can be interpreted in two ways: either the MAS complex retains some activity in the absence of TPF, or the *tpf* alleles used in this study are not null and stronger alleles could be identified that would mimic strong *minu* mutants. This is most likely the case for the *tpf2* alleles that present a T-DNA insertion in the 3′ portion of the gene body. The TPF orthologs in mammals participate in the recruitment of the complex to the chromatin through their tandem PHD domains that can recognize acetylated histones (49,50). The strong similarities between *tpf* and *minu* mutants suggest that TPF could be playing a significant role in the recruitment of the complex to their targets through their PHD and Tudor histone-reader domains. Future functional studies will shed light on these aspects.

The MAS complex partially overlaps with ncBAF-like components

Our results show that the MAS complex is found over the 5′ regions of thousands of expressed genes. This location is shared by other SWI/SNF chromatin remodeling proteins in animals and fungi (76,77). Studies of the genomic location of the BRM and SYD ATPases revealed their location over TSS regions of many genes (38,39,78), pointing to their co-localization with the MAS complex. An open question is whether there is functional crosstalk between the MAS and ncBAF-like complexes over shared targets. Importantly, it remains to be shown whether the SYD ATPase is incorporated into a ncBAF-like complex or into an uncharacterized plant SWI/SNF subclass. However, previously reported functional redundancy between BRM and SYD (40) supports its presence in ncBAF-like. BRM and SYD ATPases were also found over more distal promoter regions and downstream genic regions (39,42,78). Indeed, BRM participates over the 3′ regions of genes to regulate antisense transcription (79). Our results did not show significant enrichment of the MAS complex over distal promoter and terminator regions indicating that the MAS and ncBAF-like complexes can bind to distinct targets to perform non-redundant functions. We found a strong correlation in the position of the MAS complex with permissive marks like H3K4me3 and H3panAc marks. The fact that TPF contains 3xPHD and Tudor domains, previously shown to interact with methylated and acetylated histones (49,50,80),

makes these two histone marks likely direct targets of TPF proteins to promote recruitment or stability of the complex on target sites. On the contrary, the complex was not found over facultative or constitutive heterochromatic regions, similar to what was observed for BRM and SYD targets, indicating that plant SWI/SNF complexes do not normally localize to these regions (38,39,78,79). Interestingly, BRM and SYD have been shown to play antagonistic and synergistic interactions with the Polycomb complex, responsible for the deposition of the facultative heterochromatic mark H3K27me3 (81–83). Future studies will reveal if the MAS complex plays similar roles. The MAS complex contains the plant-specific SHH2 subunit able to interact with the constitutive heterochromatic mark H3K9me1 in *in vitro* assays (84,85). Where and when the H3K9me-reader ability of SHH2 becomes required in the MAS complex remains to be explored.

The MAS complex controls the positioning of the +1 nucleosome at thousands of loci

Depletion of the MAS complex showed a genome-wide effect on the positioning of the +1 nucleosome, with smaller effects on the fuzziness and occupancy of nucleosomes. A similar +1 nucleosome shift has been reported in yeast SWI/SNF mutants indicating a functional conservation of their activity (15,20). However, the distance between nucleosomes was not altered indicating that the MAS complex is not required to control nucleosome phasing, which has been attributed to the ISWI remodeler in plants (86). This is consistent with results in yeast where RSC depletion triggered +1 nucleosome shifts with no effect on internucleosomal spacing (87,88). A recent study on the genome-wide impact of BRM depletion on nucleosomes reported defects mostly related to the control of the occupancy of nucleosomes flanking the BRM targets, especially those modified with the H2A.Z variant (38). This indicates that MAS and ncBAF-like complexes can play unique roles in setting up the nucleosomal landscape. Importantly, while there is information about BRM (38) and MINU (this study), no study so far has reported the genome-wide impact on nucleosomes of the remaining ATPase in plants –SYD-, which would provide a complete picture of the effect of the different plant SWI/SNF ATPases on the nucleosome landscape.

A genome-wide analysis of the effect of MAS complex depletion on the transcriptome revealed that only a small number of direct targets were misregulated. This is despite the large number of direct targets that show a +1 nucleosome shift, suggesting that this change in nucleosome positioning does not play a prominent role in the regulation of transcript levels. There are examples in fungi and animals where an upstream nucleosome shift towards the NDR occludes the dominant TSS and triggers a global downregulation of expression (16,19,20). While we did not see such global effect, we did observe a correlation between downregulated direct targets and a stronger nucleosome invasion of the NDR. Perhaps more dynamic situations such as during the response to environmental or developmental cues would reveal a larger number of genes whose expression is affected by the upstream repositioning of the +1 nucleosome. Importantly, the disconnect between bound target genes and DEGs could be due to functional redundancy with other remodelers, in particular BRM and SYD complexes that also localize to many promoter regions.

The MAS complex affects the TSS determination of thousands of genes

Interestingly, we observed that thousands of MAS complex targets presented extended 5′ transcript regions in the *tpf* and *minu* mutants pointing to the participation of the complex in the selection of alternative TSSs. This change in TSS did not affect the overall expression level of most target genes suggesting that the promoter of these genes can still drive the same amount of expression regardless of the TSS used. Recent studies have reported that approximately 75% of genes are estimated to use multiple TSSs in Arabidopsis (89,90). Thus, the new TSSs used in the MAS mutants most likely correspond to previously found TSSs that are not preferentially used in unchallenged conditions but might be chosen in certain developmental or environmental situations. Importantly, changes in TSS selection have been reported in different tissues or upon environmental changes (89,91–93). The extended 5′ transcript regions observed in the *tpf* and *minu* mutants could influence mRNA function through, for example, the incorporation of upstream ORFs and localisation signals that can impact translation efficiency and final localisation of proteins, respectively (89,91). Perhaps these events could be responsible for some of the pleiotropic defects observed in the *tpf* and *minu* mutants.

We found a remarkable correlation between the changes in +1 nucleosome positioning and 5′ transcript length observed in the *tpf* and *minu* mutants. Previous reports in yeast have proposed a role for the +1 nucleosome in determining the usage of the TSS (16,94,95). In yeast SWI/SNF mutants, an upstream +1 nucleosome shift towards the NDR can position the TSS closer to the midpoint of the nucleosome, which is a more repressive location (76). Moreover, +1 nucleosome shifts can directly impact the ability of TATA-binding protein (TBP) to bind to the promoter hampering the use of the dominant TSS (15). Thus, in promoters with multiple TSSs, those not occluded by the shifted nucleosome would remain active and allow gene expression (16). Our results indicate that shifted nucleosomes in the *tpf* and *minu* mutants can occupy the position of the Y patch and GA core promoter elements, as well as the regions of several known TF binding sites, including the Telobox and CGG-box. The Y patch is a plant-specific element of unknown function that has been associated with higher promoter strength (96). The GA box and Telobox have been associated with the recruitment of polycomb through TFs like the BPC1 and AZF1, respectively (97). Future studies will reveal the direct impact of MAS complex-dependent nucleosome positioning on the binding of these TFs to their targets and the downstream consequences on polycomb activity. In conclusion, occupancy of these regions by the shifted nucleosomes could interfere with diverse TFs or transcriptional machinery which, in turn, could alter the normal function of the dominant TSSs found in WT plants and trigger the use of alternative upstream TSSs, resulting in the longer mRNA isoforms found in the *tpf* and *minu* mutants

(Figure 5E). Interestingly, we observed a positive correlation between the upstream displacement of the +1 nucleosome shift and the amount of reads differentially accumulated in the 5′ transcript ends, suggesting that the new +1 nucleosome position could be dictating the position of the alternative TSS, as it has been previously suggested in yeast studies (16,94,95).

In summary, this study reports the function of the MAS complex as an important regulator of the +1 nucleosome positioning and the TSS determination and provides experimental evidence of the functional relevance of the TPF signature subunits. We envision a scenario where internal and external stimuli could modulate the activity of the MAS complex to influence the TSS determination. This would result in altered mRNA functionality that could contribute to the adaptive response of the plant to the triggering stimuli. Future work should focus on several open questions such as which factors influence the final position of the +1 nucleosomes in the MAS mutants or what is the mechanistic connection between the +1 nucleosome shift and the selection of specific upstream TSSs.

DATA AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (98) partner repository with the dataset identifier PXD035892. The genomics data have been deposited to the GEO repository with accession number GSE205112. The code used to perform the bioinformatic analyses is available at https://github.com/jperezalemany/DiegoMartin2022. All the ChIP-seq, MNase-seq, and RNA-seq tracks can be visualized in the UCSC browser using the following link: https://genome-euro.ucsc.edu/s/jaipea/DiegoMartin2022.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Drs Miguel Blázquez and Maria A. Nohales for critical reading and discussion of the manuscript.
Author contributions: B.D.-M. performed all the *tpf* phenotyping experiments. B.D.-M., and J.G.-B. performed ChIP-seq, IP-MS and MNase-seq experiments. J.P.-A. performed all bioinformatics analyses. B.D.-M., J.C.-F., J.P. performed the Y2H experiments. B.D.-M. and A.C.-A. performed the TSS characterization experiments. D.A. supervised the work and edited the manuscript. Y.J-A. and J.W. performed the IP-MS analyses. B.D.-M., J.P.-A. and J.G.-B. conceived this study and wrote the manuscript.

FUNDING

MCIN/AEI/10.13039/501100011033 [RYC2018-024108-I, PID2019-108577GA-I00 to J.G.B.]. Funding for open access charge: CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

Conflict of interest statement. None declared.

REFERENCES

1. Luger,K., Mäder,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 a resolution. *Nature*, **389**, 251–260.
2. Jenuwein,T. and Allis,C.D. (2001) Translating the histone code. *Science*, **293**, 1074–1080.
3. Azmi,I.F., Watanabe,S., Maloney,M.F., Kang,S., Belsky,J.A., MacAlpine,D.M., Peterson,C.L. and Bell,S.P. (2017) Nucleosomes influence multiple steps during replication initiation. *Elife*, **6**, e22512.
4. Davarinejad,H., Huang,Y.C., Mermaz,B., LeBlanc,C., Poulet,A., Thomson,G., Joly,V., Muñoz,M., Arvanitis-Vigneault,A., Valsakumar,D. *et al.* (2022) The histone H3.1 variant regulates TONSOKU-mediated DNA repair during replication. *Science*, **375**, 1281–1286.
5. Zhu,F., Farnung,L., Kaasinen,E., Sahu,B., Yin,Y., Wei,B., Dodonova,S.O., Nitta,K.R., Morgunova,E., Taipale,M. *et al.* (2018) The interaction landscape between transcription factors and the nucleosome. *Nature*, **562**, 76–81.
6. Lee,C.K., Shibata,Y., Rao,B., Strahl,B.D. and Lieb,J.D. (2004) Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat. Genet.*, **36**, 900–905.
7. Yuan,G.C., Liu,Y.J., Dion,M.F., Slack,M.D., Wu,L.F., Altschuler,S.J. and Rando,O.J. (2005) Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science*, **309**, 626–630.
8. Jiang,C. and Pugh,B.F. (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.*, **10**, 161–172.
9. Nock,A., Ascano,J.M., Barrero,M.J. and Malik,S. (2012) Mediator-regulated transcription through the +1 nucleosome. *Mol. Cell*, **48**, 837–848.
10. Weber,C.M., Ramachandran,S. and Henikoff,S. (2014) Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell*, **53**, 819–830.
11. Mavrich,T.N., Jiang,C., Ioshikhes,I.P., Li,X., Venters,B.J., Zanton,S.J., Tomsho,L.P., Qi,J., Glaser,R.L., Schuster,S.C. *et al.* (2008) Nucleosome organization in the Drosophila genome. *Nature*, **453**, 358–362.
12. Muse,G.W., Gilchrist,D.A., Nechaev,S., Shah,R., Parker,J.S., Grissom,S.F., Zeitlinger,J. and Adelman,K. (2007) RNA polymerase is poised for activation across the genome. *Nat. Genet.*, **39**, 1507–1511.
13. Zeitlinger,J., Stark,A., Kellis,M., Hong,J.W., Nechaev,S., Adelman,K., Levine,M. and Young,R.A. (2007) RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat. Genet.*, **39**, 1512–1516.
14. Jimeno-González,S., Ceballos-Chávez,M. and Reyes,J.C. (2015) A positioned +1 nucleosome enhances promoter-proximal pausing. *Nucleic Acids Res.*, **43**, 3068–3078.
15. Kubik,S., O'Duibhir,E., de Jonge,W.J., Mattarocci,S., Albert,B., Falcone,J.L., Bruzzone,M.J., Holstege,F.C.P. and Shore,D. (2018) Sequence-Directed action of RSC remodeler and general regulatory factors modulates +1 nucleosome position to facilitate transcription. *Mol. Cell*, **71**, 89–102.
16. Klein-Brill,A., Joseph-Strauss,D., Appleboim,A. and Friedman,N. (2019) Dynamics of chromatin and transcription during transient depletion of the RSC chromatin remodeling complex. *Cell Rep.*, **26**, 279–292.
17. Oldfield,A.J., Henriques,T., Kumar,D., Burkholder,A.B., Cinghu,S., Paulet,D., Bennett,B.D., Yang,P., Scruggs,B.S., Lavender,C.A. *et al.* (2019) NF-Y controls fidelity of transcription initiation at gene promoters through maintenance of the nucleosome-depleted region. *Nat. Commun.*, **10**, 3072.
18. Haberle,V., Li,N., Hadzhiev,Y., Plessy,C., Previti,C., Nepal,C., Gehrig,J., Dong,X., Akalin,A., Suzuki,A.M. *et al.* (2014) Two independent transcription initiation codes overlap on vertebrate core promoters. *Nature*, **507**, 381–385.
19. Lam,K.C., Chung,H.R., Semplicio,G., Iyer,S.S., Gaub,A., Bhardwaj,V., Holz,H., Georgiev,P. and Akhtar,A. (2019) The NSL complex-mediated nucleosome landscape is required to maintain transcription fidelity and suppression of transcription noise. *Genes Dev.*, **33**, 452–465.

20. Kubik,S., Bruzzone,M.J., Challal,D., Dreos,R., Mattarocci,S., Bucher,P., Libri,D. and Shore,D. (2019) Opposing chromatin remodelers control transcription initiation frequency and start site selection. *Nat. Struct. Mol. Biol.*, **26**, 744–754.
21. Whitehouse,I., Rando,O.J., Delrow,J. and Tsukiyama,T. (2007) Chromatin remodelling at promoters suppresses antisense transcription. *Nature*, **450**, 1031–1035.
22. Narlikar,G.J., Sundaramoorthy,R. and Owen-Hughes,T. (2013) Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell*, **154**, 490–503.
23. Clapier,C.R. and Cairns,B.R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.*, **78**, 273–304.
24. Monahan,B.J., Villén,J., Marguerat,S., Bähler,J., Gygi,S.P. and Winston,F. (2008) Fission yeast SWI/SNF and RSC complexes show compositional and functional differences from budding yeast. *Nat. Struct. Mol. Biol.*, **15**, 873–880.
25. Shi,J., Zheng,M., Ye,Y., Li,M., Chen,X., Hu,X., Sun,J., Zhang,X. and Jiang,C. (2014) Drosophila brahma complex remodels nucleosome organizations in multiple aspects. *Nucleic Acids Res.*, **42**, 9730–9739.
26. Mashtalir,N., Dao,H.T., Sankar,A., Liu,H., Corin,A.J., Bagert,J.D., Ge,E.J., D'Avino,A.R., Filipovski,M., Michel,B.C. *et al.* (2021) Chromatin landscape signals differentially dictate the activities of mSWI/SNF family complexes. *Science*, **373**, 306–315.
27. Krietenstein,N., Wal,M., Watanabe,S., Park,B., Peterson,C.L., Pugh,B.F. and Korber,P. (2016) Genomic nucleosome organization reconstituted with pure proteins. *Cell*, **167**, 709–721.
28. Brahma,S. and Henikoff,S. (2019) RSC-associated subnucleosomes define MNase-sensitive promoters in yeast. *Mol. Cell*, **73**, 238–249.
29. Hodges,C., Kirkland,J.G. and Crabtree,G.R. (2016) The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harb. Perspect. Med.*, **6**, a026930.
30. Middeljans,E., Wan,X., Jansen,P.W., Sharma,V., Stunnenberg,H.G. and Logie,C. (2012) SS18 together with animal-specific factors defines human BAF-type SWI/SNF complexes. *PLoS One*, **7**, e33834.
31. Mashtalir,N., D'Avino,A.R., Michel,B.C., Luo,J., Pan,J., Otto,J.E., Zullow,H.J., McKenzie,Z.M., Kubiak,R.L., St Pierre,R. *et al.* (2018) Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell*, **175**, 1272–1288.
32. Kadoch,C., Hargreaves,D.C., Hodges,C., Elias,L., Ho,L., Ranish,J. and Crabtree,G.R. (2013) Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.*, **45**, 592–601.
33. Alpsoy,A. and Dykhuizen,E.C. (2018) Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. *J. Biol. Chem.*, **293**, 3892–3903.
34. Wang,X., Wang,S., Troisi,E.C., Howard,T.P., Haswell,J.R., Wolf,B.K., Hawk,W.H., Ramos,P., Oberlick,E.M., Tzvetkov,E.P. *et al.* (2019) BRD9 defines a SWI/SNF sub-complex and constitutes a specific vulnerability in malignant rhabdoid tumors. *Nat. Commun.*, **10**, 1881.
35. Hernández-García,J., Diego-Martin,B., Kuo,P.H., Jami-Alahmadi,Y., Vashisht,A.A., Wohlschlegel,J., Jacobsen,S.E., Blázquez,M.A. and Gallego-Bartolomé,J. (2022) Comprehensive identification of SWI/SNF complex subunits underpins deep eukaryotic ancestry and reveals new plant components. *Commun. Biol.*, **5**, 549.
36. Shang,J.Y. and He,X.J. (2022) Chromatin-remodeling complexes: conserved and plant-specific subunits in arabidopsis. *J. Integr. Plant Biol.*, **64**, 499–515.
37. Sarnowska,E., Gratkowska,D.M., Sacharowski,S.P., Cwiek,P., Tohge,T., Fernie,A.R., Siedlecki,J.A., Koncz,C. and Sarnowski,T.J. (2016) The role of SWI/SNF chromatin remodeling complexes in hormone crosstalk. *Trends Plant Sci.*, **21**, 594–608.
38. Torres,E.S. and Deal,R.B. (2019) The histone variant H2A.Z and chromatin remodeler BRAHMA act coordinately and antagonistically to regulate transcription and nucleosome dynamics in arabidopsis. *Plant J.*, **99**, 144–162.
39. Li,C.L., Gu,L.F., Gao,L., Chen,C., Wei,C.Q., Qiu,Q., Chien,C.W., Wang,S.K., Jiang,L.H., Ai,L.F. *et al.* (2016) Concerted genomic targeting of H3K27 demethylase REF6 and chromatin-remodeling ATPase BRM in Arabidopsis. *Nat. Genet.*, **48**, 687–693.
40. Bezhani,S., Winter,C., Hershman,S., Wagner,J.D., Kennedy,J.F., Kwon,C.S., Pfluger,J., Su,Y. and Wagner,D. (2007) Unique, shared, and redundant roles for the arabidopsis SWI/SNF chromatin remodeling ATPases BRAHMA and SPLAYED. *Plant Cell*, **19**, 403–416.
41. Wang,Z., Ma,Z., Castillo-González,C., Sun,D., Li,Y., Yu,B., Zhao,B., Li,P. and Zhang,X. (2018) SWI2/SNF2 ATPase CHR2 remodels pri-miRNAs via Serrate to impede miRNA production. *Nature*, **557**, 516–521.
42. Yu,Y., Liang,Z., Song,X., Fu,W., Xu,J., Lei,Y., Yuan,L., Ruan,J., Chen,C., Fu,W. *et al.* (2020) BRAHMA-interacting proteins BRIP1 and BRIP2 are core subunits of arabidopsis SWI/SNF complexes. *Nat. Plants*, **6**, 996–1007.
43. Yu,Y., Fu,W., Xu,J., Lei,Y., Song,X., Liang,Z., Zhu,T., Liang,Y., Hao,Y., Yuan,L. *et al.* (2021) Bromodomain-containing proteins BRD1, BRD2, and BRD13 are core subunits of SWI/SNF complexes and vital for their genomic targeting in arabidopsis. *Mol. Plant*, **14**, 888–904.
44. Jarończyk,K., Sosnowska,K., Zaborowski,A., Pupel,P., Bucholc,M., Małecka,E., Siwirykow,N., Stachula,P., Iwanicka-Nowicka,R., Koblowska,M. *et al.* (2021) Bromodomain-containing subunits BRD1, BRD2, and BRD13 are required for proper functioning of SWI/SNF complexes in arabidopsis. *Plant Commun.*, **2**, 100174.
45. Vercruyssen,L., Verkest,A., Gonzalez,N., Heyndrickx,K.S., Eeckhout,D., Han,S.K., Jégu,T., Archacki,R., Van Leene,J., Andriankaja,M. *et al.* (2014) ANGUSTIFOLIA3 binds to SWI/SNF chromatin remodeling complexes to regulate transcription during arabidopsis leaf development. *Plant Cell*, **26**, 210–229.
46. Lessard,J., Wu,J.I., Ranish,J.A., Wan,M., Winslow,M.M., Staahl,B.T., Wu,H., Aebersold,R., Graef,I.A. and Crabtree,G.R. (2007) An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron*, **55**, 201–215.
47. Chugunov,A.O., Potapova,N.A., Klimenko,N.S., Tatarskiy,V.V., Georgieva,S.G. and Soshnikova,N.V. (2021) Conserved structure and evolution of DPF domain of PHF10-The specific subunit of PBAF chromatin remodeling complex. *Int. J. Mol. Sci.*, **22**, 11134.
48. Wilson,B., Erdjument-Bromage,H., Tempst,P. and Cairns,B.R. (2006) The RSC chromatin remodeling complex bears an essential fungal-specific protein module with broad functional roles. *Genetics*, **172**, 795–809.
49. Zeng,L., Zhang,Q., Li,S., Plotnikov,A.N., Walsh,M.J. and Zhou,M.M. (2010) Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature*, **466**, 258–262.
50. Lange,M., Kaynak,B., Forster,U.B., Tönjes,M., Fischer,J.J., Grimm,C., Schlesinger,J., Just,S., Dunkel,I., Krueger,T. *et al.* (2008) Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev.*, **22**, 2370–2384.
51. Curtis,M.D. and Grossniklaus,U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.*, **133**, 462–469.
52. Jami-Alahmadi,Y., Pandey,V., Mayank,A.K. and Wohlschlegel,J.A. (2021) A robust method for packing high resolution C18 RP-nano-HPLC columns. *J. Vis. Exp.*, https://doi.org/10.3791/62380.
53. Cox,J. and Mann,M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, **26**, 1367–1372.
54. Thomas,Q.A., Ard,R., Liu,J., Li,B., Wang,J., Pelechano,V. and Marquardt,S. (2020) Transcript isoform sequencing reveals widespread promoter-proximal transcriptional termination in arabidopsis. *Nat. Commun.*, **11**, 2589.
55. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta delta C(T)) method. *Methods*, **25**, 402–408.
56. Gallego-Bartolome,J., Liu,W., Kuo,P.H., Feng,S., Ghoshal,B., Gardiner,J., Zhao,J.M., Park,S.Y., Chory,J. and Jacobsen,S.E. (2019) Co-targeting RNA polymerases IV and v promotes efficient de novo DNA methylation in Arabidopsis. *Cell*, **176**, 1068–1082.
57. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with bowtie 2. *Nat. Methods*, **9**, 357–359.
58. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

59. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
60. Landt,S.G., Marinov,G.K., Kundaje,A., Kheradpour,P., Pauli,F., Batzoglou,S., Bernstein,B.E., Bickel,P., Brown,J.B., Cayting,P. *et al.* (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.*, **22**, 1813–1831.
61. Ramírez,F., Ryan,D.P., Grüning,B., Bhardwaj,V., Kilpert,F., Richter,A.S., Heyne,S., Dündar,F. and Manke,T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165.
62. Zerbino,D.R., Johnson,N., Juettemann,T., Wilder,S.P. and Flicek,P. (2014) WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics*, **30**, 1008–1009.
63. Liu,W., Gallego-Bartolomé,J., Zhou,Y., Zhong,Z., Wang,M., Wongpalee,S.P., Gardiner,J., Feng,S., Kuo,P.H. and Jacobsen,S.E. (2021) Ectopic targeting of CG DNA methylation in arabidopsis with the bacterial SssI methyltransferase. *Nat. Commun.*, **12**, 3130.
64. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of chip-Seq (MACS). *Genome Biol.*, **9**, R137.
65. Chen,K., Xi,Y., Pan,X., Li,Z., Kaestner,K., Tyler,J., Dent,S., He,X. and Li,W. (2013) DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res.*, **23**, 341–351.
66. Choi,K., Zhao,X., Tock,A.J., Lambing,C., Underwood,C.J., Hardcastle,T.J., Serra,H., Kim,J., Cho,H.S., Kim,J. *et al.* (2018) Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana* transposons and gene regulatory regions. *Genome Res.*, **28**, 532–546.
67. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
68. Liao,Y., Smyth,G.K. and Shi,W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.
69. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
70. Wu,T., Hu,E., Xu,S., Chen,M., Guo,P., Dai,Z., Feng,T., Zhou,L., Tang,W., Zhan,L. *et al.* (2021) clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Cambridge (Mass.))*, **2**, 100141.
71. Bailey,T.L., Boden,M., Buske,F.A., Frith,M., Grant,C.E., Clementi,L., Ren,J., Li,W.W. and Noble,W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208.
72. Hiratsuka,T., Makita,Y. and Yamamoto,Y.Y. (2022) Sequence-based evaluation of promoter context for prediction of transcription start sites in arabidopsis and rice. *Sci. Rep.*, **12**, 6976.
73. Kindgren,P., Ivanov,M. and Marquardt,S. (2020) Native elongation transcript sequencing reveals temperature dependent dynamics of nascent RNAPII transcription in arabidopsis. *Nucleic Acids Res.*, **48**, 2332–2347.
74. Fulcher,N. and Sablowski,R. (2009) Hypersensitivity to DNA damage in plant stem cell niches. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 20984–20988.
75. Sang,Y., Silva-Ortega,C.O., Wu,S., Yamaguchi,N., Wu,M.F., Pfluger,J., Gillmor,C.S., Gallagher,K.L. and Wagner,D. (2012) Mutations in two non-canonical arabidopsis SWI2/SNF2 chromatin remodeling ATPases cause embryogenesis and stem cell maintenance defects. *Plant J. Cell Mol. Biol.*, **72**, 1000–1014.
76. Yen,K., Vinayachandran,V., Batta,K., Koerber,R.T. and Pugh,B.F. (2012) Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell*, **149**, 1461–1473.
77. Tolstorukov,M.Y., Sansam,C.G., Lu,P., Koellhoffer,E.C., Helming,K.C., Alver,B.H., Tillman,E.J., Evans,J.A., Wilson,B.G., Park,P.J. *et al.* (2013) Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 10165–10170.
78. Shu,J., Chen,C., Li,C., Thapa,R.K., Song,J., Xie,X., Nguyen,V., Bian,S., Liu,J., Kohalmi,S.E. *et al.* (2021) Genome-wide occupancy of arabidopsis SWI/SNF chromatin remodeler SPLAYED provides insights into its interplay with its close homolog BRAHMA and polycomb proteins. *Plant J. Cell Mol. Biol.*, **106**, 200–213.
79. Archacki,R., Yatusevich,R., Buszewicz,D., Krzyczmonik,K., Patryn,J., Iwanicka-Nowicka,R., Biecek,P., Wilczynski,B., Koblowska,M., Jerzmanowski,A. *et al.* (2017) Arabidopsis SWI/SNF chromatin remodeling complex binds both promoters and terminators to regulate gene expression. *Nucleic Acids Res.*, **45**, 3116–3129.
80. Lu,R. and Wang,G.G. (2013) Tudor: a versatile family of histone methylation 'readers'. *Trends Biochem. Sci*, **38**, 546–555.
81. Xu,Y., Guo,C., Zhou,B., Li,C., Wang,H., Zheng,B., Ding,H., Zhu,Z., Peragine,A., Cui,Y. *et al.* (2016) Regulation of vegetative phase change by SWI2/SNF2 chromatin remodeling ATPase BRAHMA. *Plant Physiol.*, **172**, 2416–2428.
82. Wu,M.F., Sang,Y., Bezhani,S., Yamaguchi,N., Han,S.K., Li,Z., Su,Y., Slewinski,T.L. and Wagner,D. (2012) SWI2/SNF2 chromatin remodeling ATPases overcome polycomb repression and control floral organ identity with the LEAFY and SEPALLATA3 transcription factors. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 3576–3581.
83. Li,C., Chen,C., Gao,L., Yang,S., Nguyen,V., Shi,X., Siminovitch,K., Kohalmi,S.E., Huang,S., Wu,K. *et al.* (2015) The arabidopsis SWI2/SNF2 chromatin remodeler BRAHMA regulates polycomb function during vegetative development and directly activates the flowering repressor gene SVP. *PLoS Genet.*, **11**, e1004944.
84. Zhang,H., Ma,Z.Y., Zeng,L., Tanaka,K., Zhang,C.J., Ma,J., Bai,G., Wang,P., Zhang,S.W., Liu,Z.W. *et al.* (2013) DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of pol iV. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 8290–8295.
85. Wang,Y., Zhou,X., Luo,J., Lv,S., Liu,R., Du,X., Jia,B., Yuan,F., Zhang,H. and Du,J. (2021) Recognition of H3K9me1 by maize RNA-directed DNA methylation factor SHH2. *J. Integr. Plant Biol.*, **63**, 1091–1096.
86. Li,G., Liu,S., Wang,J., He,J., Huang,H., Zhang,Y. and Xu,L. (2014) ISWI proteins participate in the genome-wide nucleosome distribution in arabidopsis. *Plant J. Cell Mol. Biol.*, **78**, 706–714.
87. Ganguli,D., Chereji,R.V., Iben,J.R., Cole,H.A. and Clark,D.J. (2014) RSC-dependent constructive and destructive interference between opposing arrays of phased nucleosomes in yeast. *Genome Res.*, **24**, 1637–1649.
88. Parnell,T.J., Schlichter,A., Wilson,B.G. and Cairns,B.R. (2015) The chromatin remodelers RSC and ISW1 display functional and chromatin-based promoter antagonism. *Elife*, **4**, e06073.
89. Kurihara,Y., Makita,Y., Kawashima,M., Fujita,T., Iwasaki,S. and Matsui,M. (2018) Transcripts from downstream alternative transcription start sites evade uORF-mediated inhibition of gene expression in arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 7831–7836.
90. Tokizawa,M., Kusunoki,K., Koyama,H., Kurotani,A., Sakurai,T., Suzuki,Y., Sakamoto,T., Kurata,T. and Yamamoto,Y.Y. (2017) Identification of arabidopsis genic and non-genic promoters by paired-end sequencing of TSS tags. *Plant J. Cell Mol. Biol.*, **90**, 587–605.
91. Ushijima,T., Hanada,K., Gotoh,E., Yamori,W., Kodama,Y., Tanaka,H., Kusano,M., Fukushima,A., Tokizawa,M., Yamamoto,Y.Y. *et al.* (2017) Light controls protein localization through phytochrome-mediated alternative promoter selection. *Cell*, **171**, 1316–1325.
92. Thieffry,A., López-Márquez,D., Bornholdt,J., Malekroudi,M.G., Bressendorff,S., Barghetti,A., Sandelin,A. and Brodersen,P. (2022) PAMP-triggered genetic reprogramming involves widespread alternative transcription initiation and an immediate transcription factor wave. *Plant Cell*, **34**, 2615–2637.
93. Mejía-Guerra,M.K., Li,W., Galeano,N.F., Vidal,M., Gray,J., Doseff,A.I. and Grotewold,E. (2015) Core promoter plasticity between maize tissues and genotypes contrasts with predominance of sharp transcription initiation sites. *Plant Cell*, **27**, 3309–3320.
94. Hughes,A.L., Jin,Y., Rando,O.J. and Struhl,K. (2012) A functional evolutionary approach to identify determinants of nucleosome positioning: a unifying model for establishing the genome-wide pattern. *Mol. Cell*, **48**, 5–15.
95. Rhee,H.S. and Pugh,B.F. (2012) Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature*, **483**, 295–301.

96. Jores,T., Tonnies,J., Wrightsman,T., Buckler,E.S., Cuperus,J.T., Fields,S. and Queitsch,C. (2021) Synthetic promoter designs enabled by a comprehensive analysis of plant core promoters. *Nat Plants*, **7**, 842–855.
97. Xiao,J., Jin,R., Yu,X., Shen,M., Wagner,J.D., Pai,A., Song,C., Zhuang,M., Klasfeld,S., He,C. *et al.* (2017) Cis and trans determinants of epigenetic silencing by polycomb repressive complex 2 in arabidopsis. *Nat. Genet.*, **49**, 1546–1552.
98. Perez-Riverol,Y., Bai,J., Bandla,C., García-Seisdedos,D., Hewapathirana,S., Kamatchinathan,S., Kundu,D.J., Prakash,A., Frericks-Zipper,A., Eisenacher,M. *et al.* (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.*, **50**, D543–D552.