*Published online 26 May 2022*

*Nucleic Acids Research, 2022, Vol. 50, No. 10* ***5513–5527***
*https://doi.org/10.1093/nar/gkac338*

# Missplicing suppressor alleles of Arabidopsis *PRE-MRNA PROCESSING FACTOR 8* increase splicing fidelity by reducing the use of novel splice sites

**Adrián Cabezas-Fuster, Rosa Micol-Ponce, Sara Fontcuberta-Cervera and María Rosa Ponce***

Instituto de Bioingeniería, Universidad Miguel Hernández, Campus de Elche, 03202 Elche, Alicante, Spain

Received July 27, 2021; Revised March 30, 2022; Editorial Decision April 21, 2022; Accepted April 25, 2022

## ABSTRACT

**Efficient splicing requires a balance between high-fidelity splice-site (SS) selection and speed. In *Saccharomyces cerevisiae*, Pre-mRNA processing factor 8 (Prp8) helps to balance precise SS selection and rapid, efficient intron excision and exon joining. *argonaute1-52* (*ago1-52*) and *incurvata13* (*icu13*) are hypomorphic alleles of the *Arabidopsis thaliana* genes *ARGONAUTE1* (*AGO1*) and *AUXIN RESISTANT6* (*AXR6*) that harbor point mutations creating a novel 3′SS and 5′SS, respectively. The spliceosome recognizes these novel SSs, as well as the intact genuine SSs, producing a mixture of wild-type and aberrant mature mRNAs. Here, we characterized five novel mutant alleles of *PRP8* (one of the two Arabidopsis co-orthologs of yeast *Prp8*), naming these alleles *morphology of ago1-52 suppressed5* (*mas5*). In the *mas5-1* background, the spliceosome preferentially recognizes the intact genuine 3′SS of *ago1-52* and 5′SS of *icu13*. Since point mutations that damage genuine SSs make the spliceosome prone to recognizing cryptic SSs, we also tested alleles of four genes carrying damaged genuine SSs, finding that *mas5-1* did not suppress their missplicing. The *mas5-1* and *mas5-3* mutations represent a novel class of missplicing suppressors that increase splicing fidelity by hampering the use of novel SSs, but do not alter general pre-mRNA splicing.**

## INTRODUCTION

Pre-mRNA splicing is a co-transcriptional, high-fidelity process consisting of two sequential transesterifications carried out by the spliceosome, resulting in the excision of an intron and the ligation of its flanking exons. These reactions depend on conserved intronic and exonic sequences: the branch-point sequence (BPS) and the 5′ splice site (5′SS) and 3′SS (reviewed in 1–3). Mutations that damage genuine (authentic) SSs or create novel SSs frequently produce a mixture of wild-type and aberrant mature mRNAs from a single pre-mRNA; these mutations are a major cause of several rare human diseases, including inherited mental disorders (4).

Pre-mRNA processing factor 8 (named Prp8 in the yeast *Saccharomyces cerevisiae* and PRPF8 in humans, and referred to across species as PRP8) is a core component of the spliceosome, as well as its largest (>2,000 amino acids) and most highly conserved protein. *PRP8* is an essential gene whose loss of function causes global splicing defects in all organisms studied. The *S. cerevisiae Prp8* gene was identified based on its conditional lethal alleles, and human *PRPF8* was identified based on the retinopathies caused by its loss-of-function alleles (1).

Efficient splicing requires a balance between high-fidelity sequence selection and speed. Structural analysis has revealed two alternative conformations of PRP8: one promotes fidelity over catalytic efficiency and the other promotes efficient, error-prone splicing (5). Indeed, some missense alleles of *PRP8* suppress missplicing of pre-mRNAs carrying mutations that damage SSs, which are not efficiently recognized by the spliceosome. Such extragenic suppressor alleles of *PRP8* reduce the frequency of selection of cryptic SSs by the spliceosome in *S. cerevisiae*, humans, *Caenorhabditis elegans* and *Arabidopsis thaliana* (hereafter Arabidopsis) (1,6).

ARGONAUTE1 (AGO1) is a key factor of microRNA (miRNA) pathways in Arabidopsis. The *ago1-52* mutation creates a novel 3′SS in the last intron of *AGO1* and the spliceosome uses this novel 3′SS more frequently than the genuine 3′SS, causing missplicing of the *AGO1* pre-mRNA (7,8). In a genetic screen for second-site suppressors, we previously mutagenized homozygous *ago1-52* plants using ethyl methanesulfonate (EMS) and identified 22 lines carrying extragenic suppressor mutations of its morphological phenotype (9). We named the suppressor genes *MORPHOLOGY OF AGO1-52 SUPPRESSED* (*MAS*). Eleven of the suppressed lines carried mutant alleles of the gene we

*To whom correspondence should be addressed. Email: mrponce@umh.es

named *MAS2*, which encodes the Arabidopsis ortholog of human NF-κB activating protein (NKAP) (10). The *mas2* mutations are predicted to cause amino acid substitutions, and the corresponding mutated MAS2 proteins act as dominant informational suppressors that partially suppress the missplicing of *ago1-52* by an unknown mechanism (8).

Five of the suppressor lines carried novel alleles of *PRP8* (AT1G80070), a gene that we initially dubbed *MAS5*. The Arabidopsis genome has two *PRP8* co-orthologs (11), but only AT1G80070 has been well studied, and traditionally named *PRP8* (its mutant alleles are named *prp8*). Due to the lethal effects of its loss-of-function mutations, other studies have named *PRP8* as *ABNORMAL SUSPENSOR 2* (*SUS2*), *EMBRYO DEFECTIVE 14* (*EMB14*), *EMB33* and *EMB177* (12). The other *PRP8* co-ortholog is AT4G38780, whose loss-of-function alleles do not seem to cause any phenotypic effect (11). The *mas5* alleles of *PRP8* may represent a novel class of missplicing suppressors that promote splicing fidelity by disfavoring the use of novel 3′SSs and 5′SSs created by mutation.

## MATERIALS AND METHODS

### Plant material and growth conditions

*Arabidopsis thaliana* Landsberg *erecta* (L*er*), Columbia-0 (Col-0), Wassilewskija (Ws-2), and Enkheim-2 (En-2) wild-type accessions were obtained from the Nottingham Arabidopsis Stock Center (NASC) and propagated in our laboratory for further analysis. Seeds of *prp8-6* (in the L*er* genetic background) and *prp8-7* (Col-0) mutants were provided by C. Dean and M. Matzke, respectively; *ago1-2* (Ws-2) by C. Benning; *ago1-25* and *ago1-27* (Col-0) by H. Vaucheret; and *sca3-1*, *anu4-1*, and *ang1-2* (L*er*) by J.L. Micol. Seeds of *icu13* (En-2; N349), *sar1-4* (Col-0; SALK_126801) and *atprmt5-1* (Col-0; SALK_065814) were provided by NASC. The *ago1-51*, *ago1-52* and *mas5* mutants were isolated in our laboratory (7,10). Seed sterilization and sowing, plant culture and crosses were performed as previously described (13,14).

### Positional cloning of *MAS5* and genotyping of single and double mutants

Genomic DNA extraction from *mas5-1* plants and mapping of the *mas5-1* mutation by iterative linkage analysis to molecular markers were performed as previously described (15,16). The PCR primers used for fine mapping are listed in Supplementary Table S1. The *mas5* point mutations were identified by Sanger sequencing using the primers described in Supplementary Table S2. At least two $M_3$ plants carrying each *mas5* allele were backcrossed twice. Plants that were phenotypically wild type but genotypically *AGO1*/*AGO1*;*mas5*/*mas5* (*AGO1* being the wild-type allele of the *AGO1* gene) were selected from the $F_2$ progeny. The *mas5* homozygous plants of the $F_2$ progeny derived from the second backcross were used for all further studies described here. The *ago1 mas5* double mutants studied in this work were also reconstructed from the *mas5* lines isolated after two backcrosses.

The single and double mutants carrying point mutations were genotyped by Sanger sequencing (*ago1-51*, *ago1-52*, *ago1-25*, *ago1-27*, *mas5-1*, *mas5-2*, *mas5-3*, *mas5-4*, *mas5-5*, *mas5-6*, *prp8-6*, *anu4-1*, *ang1-2*, *sca3-1* and *icu13*) or by restriction with *Mbo*II (*prp8-7*). The *sar1-4* and *atprmt5-1* insertional mutants were genotyped by PCR amplification. The primers used are listed in Supplementary Table S2. Most Sanger sequencing reactions and electrophoreses were carried out in our laboratory with ABI PRISM BigDye Terminator Cycle Sequencing kits and an ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems). Some Sanger sequencing reactions were carried out at Stab Vida (Caparica, Portugal).

### Plant morphometry and histology

Rosette pictures were taken with a Nikon SMZ1500 stereomicroscope equipped with a Nikon DXM1200F digital camera. For large rosettes, several pictures from the same plant were assembled with the Photomerge tool of Adobe Photoshop CS3 software. The backgrounds of the rosette pictures were homogenized using the Adobe Photoshop CS3 software without modifying the rosette images.

For cell size measurements, first-node leaves, collected 21 days after stratification (das), were cleared with ethanol and chloral hydrate, and mounted on slides. The samples were photographed under a Leica DMRB microscope equipped with a Nikon DXM1200 digital camera. The micrographs were transformed into diagrams by drawing the cell margins on a Cintiq 18SX Interactive Pen Display (Wacom) and using Adobe Photoshop CS3 software. Whole rosette area and palisade mesophyll cell size were measured with NIS Elements AR 3.1 (Nikon) software, as previously described (17).

### RT-qPCR analysis

RNA isolation was performed using TRIzol Reagent (Invitrogen). Following cDNA synthesis, RT-qPCR analyses were performed in a Step-One Real-Time PCR System (Applied Biosystems) as previously described (7). Three biological replicates were used, each consisting of a mixture of three rosettes collected 15 das. Three technical replicates were used per biological replicate. The *ACTIN2* (*ACT2*) housekeeping gene was used as an internal control for relative quantification.

### Immunoblot analysis

Immunoblot analyses were performed as previously described (18). The anti-AGO1 (AS09 527; Agrisera), anti-CUL1 (kindly provided by J.C. del Pozo), and anti-RbcL (AS03 037; Agrisera) primary antibodies were used at 1:10,000, 1:3,000 and 1:2,500 dilutions, respectively. WesternSure HRP Goat anti-Rabbit IgG (LI-COR) secondary antibody was used at 1:50,000 dilution. Detection was performed using the WesternSure PREMIUM Chemiluminescent Substrate (LI-COR) and a C-Digit Blot Scanner (LI-COR). The Image Studio Analysis (LI-COR) software was used for band quantification.

### RNA-seq and splicing analysis

Three biological replicates, each consisting of 5 μg of total RNA (isolated with TRIzol Reagent [Invitrogen]) from three rosettes collected 15 das, were sent to Novogene (Cambridge, United Kingdom) for high-throughput sequencing. cDNA libraries were produced with the NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs) and sequenced on a NovaSeq 6000 Illumina platform using a S4 Flow Cell and a 2 × 150 bp paired-end run. More than 98.5 M non-stranded 150 bp paired-end reads, equivalent to 14.8 Gbp of raw data, were generated from each library (Supplementary Table S3). All the FASTQ files were submitted to the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information (NCBI) under the BioProject accession PRJNA787038 (https://www.ncbi.nlm.nih.gov/sra/PRJNA787038).

Splicing analysis was carried out at the Bioinformatics for Genomics and Proteomics Unit of the Centro Nacional de Biotecnología (CNB, Madrid). Quality and purity of raw reads were assessed with FastQC 0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and FastQ Screen 0.14.1 (19), respectively. Reads were aligned against the Arabidopsis L*er* genome (NCBI accession GCA_001651475.1), using STAR 2.7.9a (20) with default parameters, except for –alignIntronMax and –alignMatesGapMax, which were set to 15,000. Potential optical duplicates and secondary alignments were identified and removed using the Picard Toolkit (https://broadinstitute.github.io/picard/) to get the effective reads from the aligned reads (Supplementary Table S3). Finally, differential splicing events were determined for each group pair (*mas5-1* or *mas5-3* versus L*er*) by applying the standard pipeline defined for the ASpli 2.4.0 R package (21), and indicating a minimum read length of 100 bp and a maximum intron size of 14,334 bp, which corresponds to that of the longest intron in the reference genome. Briefly, multiexonic genes were divided into bins, which were then classified as exclusively exonic (including the external exons, defined as the first or last exon of a transcript), exclusively intronic, original intron (Ios, defined as introns before splitting, resulting from the retention/inclusion of two or more subbins), or annotated alternative splicing bins. Bins (excluding the external exons and Ios) were subjected to differential splicing analysis if genes with which they were associated were expressed above a minimum threshold of 10 reads in both genotypes compared, and if bins had >5 reads in at least one genotype. Finally, reads at the bin level were normalized to the read counts of their corresponding gene, and the differential bin usage was estimated. Only those bin-based splicing events with a false discovery rate (FDR) <0.1 and an absolute Delta (percent spliced-in, PSI) or Delta (percent intron retention, PIR) >5% were considered statistically significant.

### RNA fluorescence *in situ* hybridization

Tissue preparation and RNA fluorescence *in situ* hybridization (RNA-FISH) were performed as previously described (22,23), using a 40-mer fluorescein-labeled oligo(dT) probe (Eurofins Genomics) at a concentration of 0.5 μg/ml in PerfectHyb Plus Hybridization Buffer (Sigma-Aldrich). Fluorescein was excited at 488 nm and its emission collected at 515/30 nm, maintaining the same detector gain to allow direct comparisons of fluorescence intensity between samples.

### Accession numbers

Sequence data from this article can be found at The Arabidopsis Information Resource (TAIR; https://www.arabidopsis.org) under the following accession numbers: *PRP8* (AT1G80070), *AGO1* (AT1G48410), *SCA3* (AT2G24120), *ANU4* (AT1G02280), *ANG1* (AT2G27530), *AXR6* (AT4G02570; also called *ICU13*), *SAR1* (AT1G33410) and *ATPRMT5* (AT4G31120).

## RESULTS

### Isolation of dominant mutant alleles of *PRP8* that suppress the morphological phenotype of *ago1-52*

*ago1-52*, a recessive and hypomorphic allele of *AGO1*, causes a pleiotropic phenotype (7) that is easily distinguishable from that of its wild type L*er* at any developmental stage (Figure 1A and B). We performed a second-site mutagenesis screen for suppressors of the morphological phenotype of *ago1-52* (10). The suppressor mutation carried by the $M_3$ progeny of an $M_2$ plant (P8 25.1) was named *mas5-1* (Figure 1C) and crossed to Col-0 to obtain an $F_2$ mapping population. Two phenotypic classes were defined: plants exhibiting the phenotype of *ago1-52* and plants similar to a wild-type Col-0/L*er* hybrid. We genotyped plants from each class for 32 molecular markers known to be polymorphic between Col-0 and L*er*, as well as for the presence of *ago1-52* and its *AGO1* wild-type allele.

Iterative linkage analysis of these $F_2$ plants with additional molecular markers, allowed us to define a 446-kb candidate interval flanked by cer461530 and cer470312 (Supplementary Table S1) that harbored the *mas5-1* suppressor mutation. Among the genes within the above interval, we considered AT1G80070 (*PRP8*) to be the best candidate gene for *MAS5*, since it is one of the two co-orthologs encoding PRP8, the largest factor of the core spliceosome. Indeed, Sanger sequencing of AT1G80070 in *ago1-52 mas5-1* plants revealed a G→A mutation that is predicted to cause an E1769K substitution (Figure 1I). We sequenced AT1G80070 in all lines carrying putative extragenic suppressors of *ago1-52* that we had isolated in our screen and identified 10 that carried *mas5* alleles with mutations in *PRP8* (Supplementary Figure S1). Six of these *mas5* alleles are unequivocally of different origins and some alleles were isolated twice in lines originating from the same parental group subjected to mutagenesis.

The *mas5-1* to *mas5-6* alleles of *PRP8* caused different degrees of suppression in the $M_3$ generation (Figure 1C–H). The *mas5-1* and *mas5-2* mutations are identical even though they were derived from different parental groups and therefore originated from independent mutational events (Supplementary Figure S1). However, the *ago1-52 mas5-1* and *ago1-52 mas5-2* plants of the $M_3$ generation had different morphological phenotypes (Figure 1C and D), likely due to the presence of other mutations resulting from EMS mutagenesis. For example, the *ago1-52 mas5-2* mutant also car-

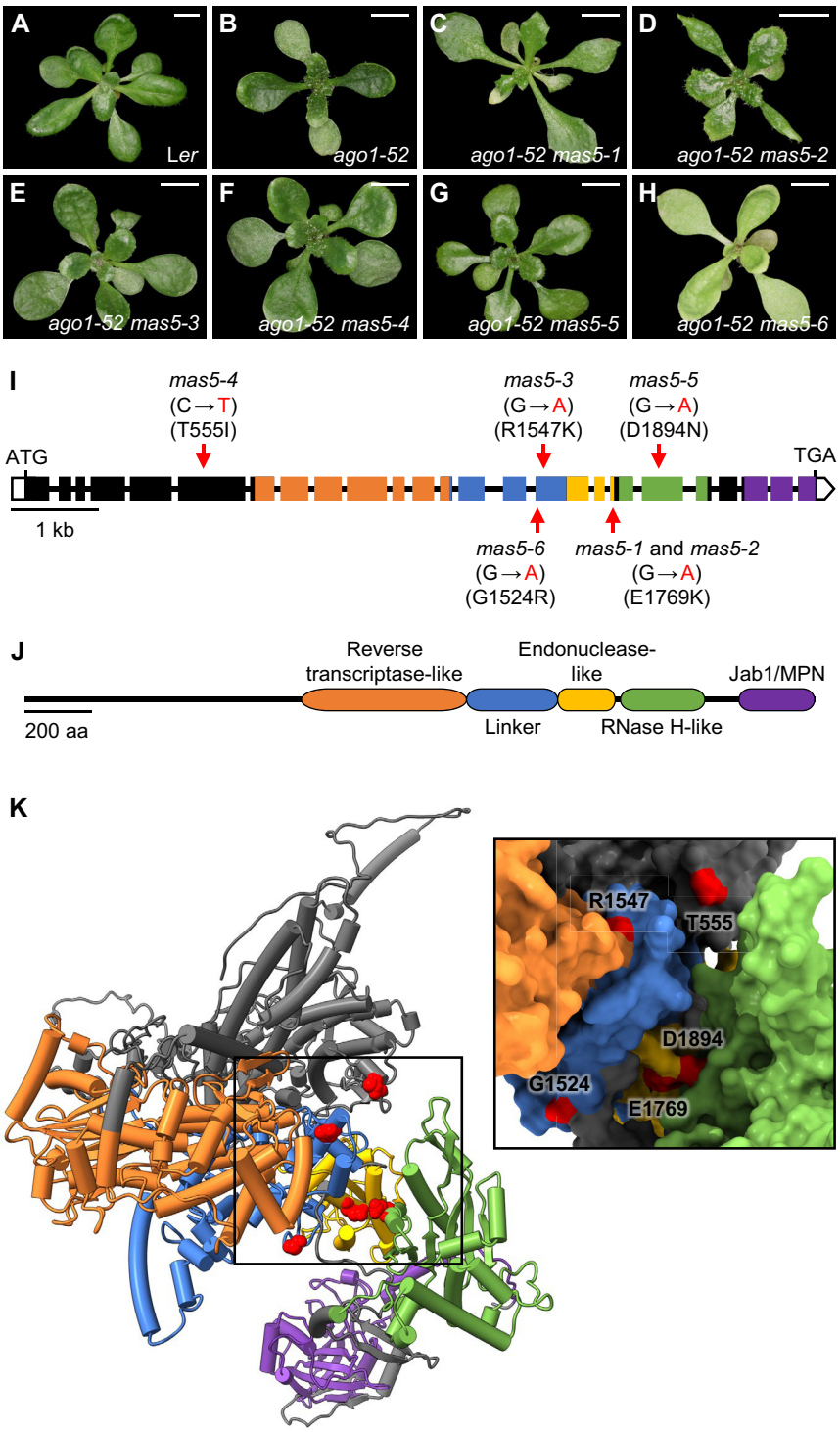


**Figure 1.** Molecular nature and effects of the *mas5* alleles of *PRP8* examined in this study. (**A–H**) Suppression of the morphological phenotype of *ago1-52* by the *mas5* mutations. Rosettes of (A) the wild-type L*er*, (B) the *ago1-52* single mutant, and the (C) *ago1-52 mas5-1*, (D) *ago1-52 mas5-2*, (E) *ago1-52 mas5-3*, (F) *ago1-52 mas5-4*, (G) *ago1-52 mas5-5*, and (H) *ago1-52 mas5-6* double mutants. The plants shown in (C–H) belong to the $M_3$ generation of the genetic screen described in (10) and still had not been backcrossed to L*er*. Photographs were taken 21 days after stratification (das). Scale bars: 4 mm. (**I**) Schematic representation of the *PRP8* gene, indicating the nature and positions of the *mas5* mutations and their predicted effects on the PRP8 protein. Empty and filled boxes represent untranslated and coding exonic regions, respectively. Lines between boxes represent introns, and red arrows indicate the positions of point mutations. Mutated nucleotides are shown in red. (**J**) Schematic representation of the PRP8 domains. The same colors have been used to highlight the regions of the *PRP8* gene encoding the corresponding domains (in I), and those domains in the PRP8 protein (in J). Sequence and domain information about PRP8 was obtained from TAIR10 (https://www.arabidopsis.org/) and (3). (**K**) Prediction of the Arabidopsis PRP8 3D structure with indication of the residues altered by the *mas5* mutations. The structure was downloaded from AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/; PDB: AF-Q9SSD2-F1) and visualized with the ChimeraX 1.2.5 software (https://www.rbvi.ucsf.edu/chimerax/). PRP8 domain colors are the same than those used in (J), and residues altered by the *mas5* mutations are shown in red. The close-up view of PRP8 surface, with focus on the region containing the *mas5* mutations, has been shaded to highlight the protein cavities and pockets.

ried *mas2-7*, an allele of *MAS2* that also suppresses *ago1-52* (8). The mutational burden caused by EMS is also clearly evidenced by the chlorotic phenotype of the *ago1-52 mas5-6* $M_3$ plant shown in Figure 1H, as also observed in all *ago1-52 mas5* lines shown in Figure 1. *mas5-1* and *mas5-3* $F_2$ plants, from second backcrosses to L*er*, were used for all further studies in this work.

The finding of six allelic *mas5* mutations of independent origin in a single genetic screen strongly supports the hypothesis that *PRP8* is the causal gene for the suppression of the *ago1-52* phenotype in the *ago1-52 mas5-1* to *ago1-52 mas5-6* double mutants. The known functional role of PRP8 as a core component of the spliceosome also suggests that the *mas5* alleles act as informational suppressors of the aberrant splicing of *ago1-52*.

### Mutations in yeast *Prp8* that act as missplicing suppressors map to the same regions that harbor *mas5* mutations in Arabidopsis *PRP8*

Crystallographic structural analyses revealed the existence of five functional domains in yeast Prp8 (Figure 1J and K, and Supplementary Figure S2): the Reverse transcriptase-like, Linker, Endonuclease-like (24), RNase H-like (25) and C-terminal Jab1/MPN (26) domains. Crystallography also revealed the existence of a cavity formed by the Reverse transcriptase thumb (one of the three subdomains of the Reverse transcriptase-like domain [amino acids 1257–1375]), the Endonuclease-like domain (amino acids 1652–1821), and the RNase H-like domain (amino acids 1836–2091; Figure 1J and K, and Supplementary Figure S2). This cavity is involved in the interaction of PRP8 with the 5′SS, 3′SS and BPS of any intron (Supplementary Figures S3 and S4), contributing to the fidelity of the two splicing steps, and is where most missplicing suppressor mutations that have been mapped in yeast and humans are located (3,24,27,28).

To find the residues in yeast Prp8 that are homologous to those that the *mas5-1* to *mas5-6* mutations affect in Arabidopsis PRP8, we performed a multiple alignment of the amino acid sequences of PRP8 orthologs. We focused on species with the highest number of previously described mutations, including informational missplicing suppressor alleles from humans, *S. cerevisiae*, and *C. elegans* (Supplementary Figure S2). Then, using the cryo-electron microscopy (cryo-EM) structures of the yeast spliceosomal and post-spliceosomal complexes assembled on a single-intron pre-mRNA from the Protein Data Bank (https://www.rcsb.org) and the ChimeraX 1.2.5 software for their visualization, we located the homologous residues in Prp8 (Supplementary Figures S3 and S4). Specifically, we located the homologous residues in the most recently determined cryo-EM structures for five of the eight major functional states of the spliceosome: activated complex ($B^{act}$; PDB: 7DCO, at 2.5 Å resolution; 29), catalytically activated complex (B*; PDB: 6J6Q, at 3.7 Å resolution; 30), catalytic step I complex (C; PDB: 7B9V, at 2.8 Å resolution; 31), catalytically activated step II complex (C*; PDB: 5WSG, at 4.0 Å resolution; 32), and post-catalytic complex (P; PDB: 6BK8, at 3.3 Å resolution; 33).

Most amino acids affected by the *mas5* mutations in Arabidopsis PRP8 are conserved across all PRP8 orthologs and five of these mutations affect residues of the cavity mentioned above, including the Linker region (Figure 1I-K and Supplementary Figures S2-S4). Three *mas5* mutations causing the highest levels of suppression are located close to each other, affecting the end of the Endonuclease-like domain (*mas5-1* and *mas5-2*) and the start of the RNase H-like domain (*mas5-5*). The *mas5-3* and *mas5-6* mutations damage the Linker region (amino acids 1375–1648), where several missplicing suppressor mutations also occur in yeast Prp8 (Figure 1I-K, Supplementary Figures S2-S4, and Supplementary Table S4).

The identical *mas5-1* and *mas5-2* mutations (Figure 1I–K) are predicted to cause an E1769K change, which substitutes a basic amino acid by an acidic one; their most similar mutation affecting the homologous residue of yeast is *D-135*, which causes an E1817G change, substituting a non-polar amino acid by an acidic one (Supplementary Figure S2 and Supplementary Table S4). This *D-135* mutation suppresses missplicing caused by mutations in position 2 of the 5′SS of a reporter gene, which is used as a cryptic 5′SS (34). The *mas5-5* mutation causes a D1894N change, which corresponds to D1942 of yeast Prp8. D and N are polar amino acids, but the substitution caused by *mas5-5* adds an amino group and removes the negative charge of D (Supplementary Figure S2). In the B* to P yeast spliceosomal complexes, the homologous residues to those of Arabidopsis PRP8 affected by the *mas5-1* (*mas5-2*) and *mas5-5* mutations are very close to each other (Supplementary Figures S3 and S4).

The amino acids affected by *mas5-3* (R1547K) and *mas5-6* (G1524R) do not correspond to those of yeast suppressor mutations identified in the same region (W1609R, W1575R, E1576V, and T1565A; Supplementary Figure S2) (35). This region forms a disordered loop that interacts with the BPS (24). However, the *mas5-3* and *mas5-6* missense mutations cause opposite changes. On the one hand, R and K are similar positively charged amino acids, and the R1547K change only eliminates two amino groups. On the other hand, the G1524R substitution in the *mas5-6* mutant adds three amino groups and a positive charge. We found that the yeast Prp8 residue homologous to the PRP8 residue affected by the Arabidopsis *mas5-3* mutation (R1595) forms part of the so-called 1585-loop of this protein (amino acids 1585–1598 [32] or 1576–1599 [33]), which interacts directly with the intron lariat–3′ exon in the C* spliceosomal complex, stabilizing the 3′SS for the second transesterification (Supplementary Figure S4). However, the yeast Prp8 residue homologous to the PRP8 residue affected by the Arabidopsis *mas5-6* mutation (G1572) was not located near the pre-mRNA, snRNAs, of any of the conserved residues affected by the other *mas5* mutations, in all yeast spliceosomal complexes (Supplementary Figures S3 and S4).

The only *mas5* mutation that affects the N-terminal domain of PRP8 is *mas5-4* (T555I), but we observed a clear interaction of the homologous residue in yeast Prp8 with the 5′SS in all spliceosomal complexes (Supplementary Figures S3 and S4). An *az50* semidominant mutation affecting the corresponding residue in *C. elegans* (T524S) was previously found in a screen for factors that modify the frequency of cryptic splicing, but does not produce overall changes in splicing (6).

### *mas5-1* does not suppress the morphological phenotypes of *ago1-25* or *ago1-27*

For further analysis, we selected the *mas5-1* and *mas5-3* mutations, which affect two different regions: the Endonuclease-like domain and the Linker region of PRP8, respectively (Figure 1I-K). We backcrossed the *ago1-52 mas5-1* (P8 25.1; Figure 1C) and *ago1-52 mas5-3* (P7 24.1; Figure 1E) $M_3$ lines twice to L*er*. After these backcrosses, *mas5-3* plants exhibited moderately larger rosettes compared to L*er* and *mas5-1* but were otherwise very similar to L*er* throughout vegetative and reproductive development (Figure 2A, C, E, G and H). We then crossed the backcrossed *mas5* mutants to *ago1-52*, to confirm the suppression, and to *ago1-25* and *ago1-27*, to determine the specificity of the suppression. *ago1-25* and *ago1-27* carry EMS-induced point mutations in the Col-0 genetic background that cause single amino acid substitutions but do not alter their own pre-mRNA splicing (36).

In *ago1-52 mas5-1* plants, and to a lesser extent in *ago1-52 mas5-3,* rosette size and whole plant height were partially restored to the L*er* values (Figure 2A-H). The *ago1-52 mas5-1* plants were similar to L*er*, but *ago1-52 mas5-3* resembled *ago1-52*, with dark green rosettes harboring two large, spatulate first leaves, like those of *ago1-52* (Figure 2A-F). However, the main stem heights of *ago1-52 mas5-1* and *ago1-52 mas5-3* were closer to those of L*er, mas5-1* and *mas5-3* (Figure 2G). We also obtained histological evidence for suppression: the small palisade mesophyll cell size of *ago1-52* was normalized in *ago1-52 mas5-1* plants, and to a lesser extent in *ago1-52 mas5-3* (Supplementary Figure S5). We did not find any evidence of morphological suppression during vegetative or reproductive development in *ago1-25 mas5-1*, *ago1-27 mas5-1*, *ago1-25 mas5-3* or *ago1-27 mas5-3* plants (Supplementary Figure S6).

Taken together, these results indicate that *mas5-1*, and to a lesser extent *mas5-3*, appears to specifically suppress the *ago1-52* allele. It is likely that the remaining *mas5* alleles also act specifically on the *ago1-52* mutation, but this has yet to be demonstrated.

### The *mas5* mutations modify the ratios of mRNA splice variants and protein isoforms produced by *ago1-52*

The novel 3′SS of *ago1-52* seems to be used by the spliceosome more frequently than the genuine one, as shown by the level of the mRNA variant containing 10 nt of the 21$^{st}$ intron, which is more abundant than the wild-type variant. Translation of the misspliced *ago1-52* mRNA variant produces a truncated protein (AGO1-52), which is 55 amino acids shorter than the wild-type AGO1 (wAGO1) and includes 15 amino acids not present in wAGO1 at its C-terminus (Figure 2I) (7).

To study the suppression of *ago1-52* by *mas5-1* and *mas5-3* at the molecular level, we amplified total (*tAGO1*) and wild-type (*wAGO1*) mRNA splicing variants by RT-qPCR using specific primers (Supplementary Table S2) (8). In agreement with the different levels of morphological suppression in both double mutants, the ratio of *wAGO1*/*ago1-52* mRNAs was higher in *ago1-52 mas5-1* than in *ago1-52 mas5-3* (Figure 2I-K). Therefore, the suppression of *ago1-52* by *mas5-1* could be due to the almost 10-fold increase in *wAGO1* mRNA levels in *ago1-52 mas5-1* plants, the reduced levels of aberrant *ago1-52* mRNA in the double mutant, or both.

As a control, we performed immunoblot analysis using the *ago1-2* null mutant (37), which does not produce any AGO1 protein (Figure 2L). The wAGO1 protein (∼130 kDa) was the only AGO1 protein detected in the L*er*, *mas5-1*, and *mas5-3* extracts. In agreement with our RT-qPCR results, we detected high levels of the mutant AGO1-52 protein (∼125 kDa) in *ago1-52*, along with low levels of wAGO1. We also detected two bands in *ago1-52 mas5-1* and *ago1-52 mas5-3*, corresponding to the wAGO1 and AGO1-52 protein isoforms, as previously shown in *ago1-52 mas2-1* plants (8). Therefore, the level of wAGO1 was higher than that of AGO1-52 in both *ago1-52 mas5-1* and *ago1-52 mas5-3* (Figure 2L and M). These results are in agreement with the stronger suppression of morphological defects in *ago1-52 mas5-1* compared to *ago1-52 mas5-3* (Figure 2A–H).

We repeated the RT-qPCR and immunoblot analyses with the original $M_4$ lines harboring the *mas5-4*, *mas5-5* and *mas5-6* alleles (Figure 1F–H). We obtained similar results to those found with the *ago1-52 mas5-1* double mutant with the *mas5-6* allele, and less suppression with *mas5-5*, and in particular with *mas5-4*, as we observed with the *mas5-3* allele in *ago1-52 mas5-3* (Supplementary Figure S7). Therefore, the suppression of *ago1-52* by the *mas5* mutations may involve effects at the translational level; for example, the *wAGO1* splice variant may be more translatable than the *ago1-52* mRNA.

### *mas5-1* increases splicing fidelity in the *icu13* allele of *AXR6*, which contains a novel 5′SS

To determine whether the suppression by *mas5* alleles is specific to the *AGO1* gene or whether it also occurs in other genes whose mutations eliminate or create novel SSs, we crossed *mas5-1* to the *incurvata13* (*icu13*), *scabra3-1* (*sca3-1*), *angulata4-1* (*anu4-1*) and *angusta1-2* (*ang1-2*) mutants. We selected these four additional mutants because 1) they harbor the same type of transitions (G→A or C→T) but exhibit different types of missplicing, 2) they do not appear to be functionally related to each other or to *AGO1* or *PRP8*, and 3) their morphological phenotypes are easily distinguishable by eye (Figure 3 and Supplementary Figure S8). We also crossed *mas5-1* to the *ago1-51* mutant, which is in a L*er* background, like *ago1-52*. Unlike *ago1-52* and *icu13*, the *ago1-51*, *sca3-1*, *anu4-1* and *ang1-2* point mutations damage genuine SSs, favorizing the recognition of nearby cryptic SSs by the spliceosome (Supplementary Figure S8).

*icu13* is a recessive hypomorphic allele of the *AUXIN RESISTANT6* (*AXR6*) gene, which encodes CULLIN1 (CUL1), a component of the core SCF complex that catalyzes the ubiquitination of proteins for their degradation by the proteasome (38). A C→T transition in *icu13* creates a novel 5′SS upstream of the genuine 5′SS of its 15$^{th}$ intron. The alternative use of both 5′SSs by the spliceosome generates two different splice variants from *icu13*. One of these mRNA variants (which we named *icu13.1*) has a synonymous mutation (GGC→GGU, both codons encoding glycine) and is produced when the genuine 5′SS is used by

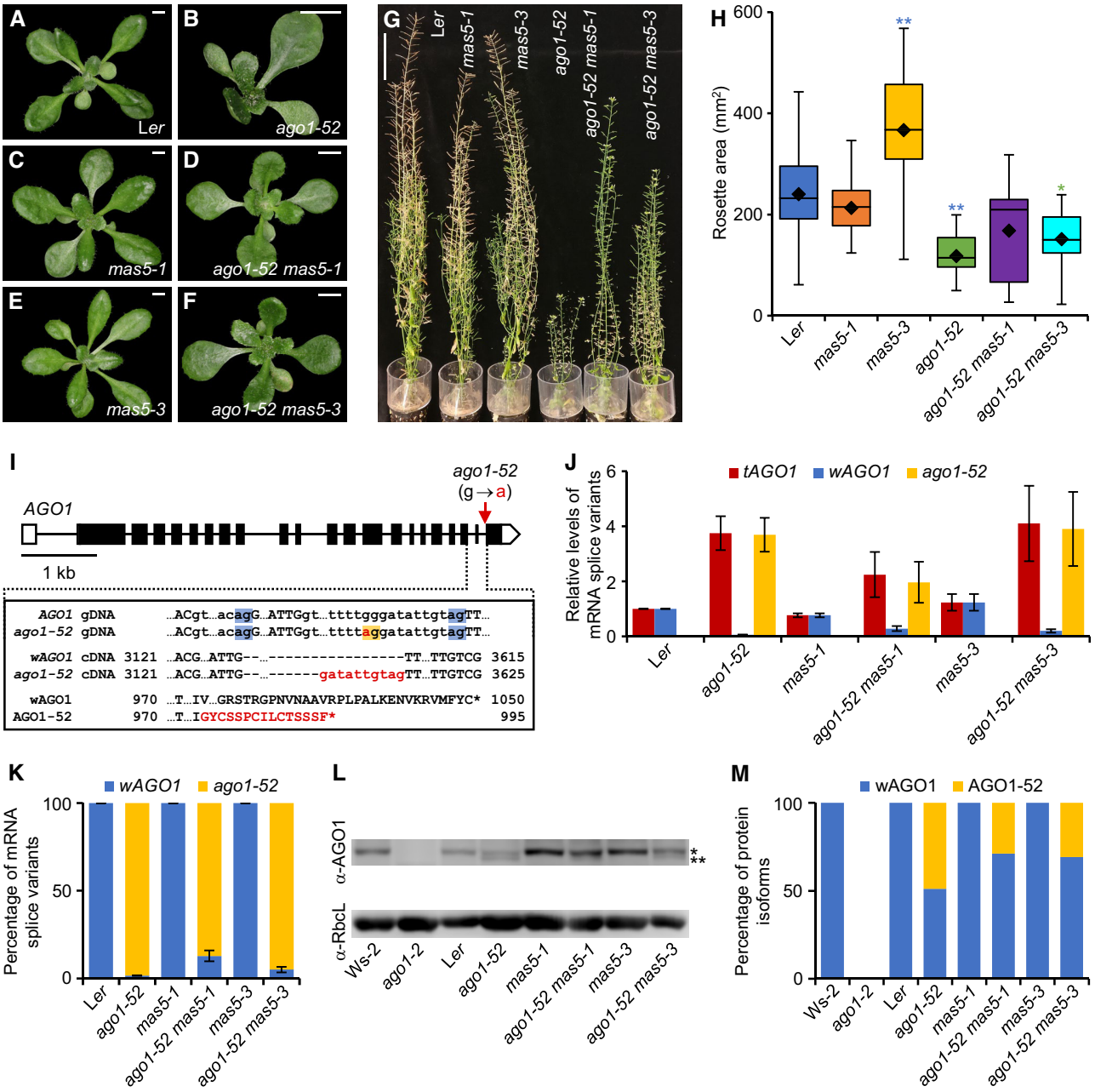


**Figure 2.** Suppression of the morphological and molecular phenotypes of *ago1-52* by *mas5-1* and *mas5-3*. (**A–F**) Rosettes of (A) L*er*, (B) *ago1-52*, (C) *mas5-1*, (D) *ago1-52 mas5-1*, (E) *mas5-3* and (F) *ago1-52 mas5-3* plants. (**G**) From left to right, adult plants of L*er*, *mas5-1*, *mas5-3*, *ago1-52*, *ago1-52 mas5-1*, and *ago1-52 mas5-3*. Photographs were taken (A–F) 21 and (G) 52 das. Scale bars: (A–F) 4 mm, and (G) 5 cm. (**H**) Boxplot showing the distribution of rosette areas in plants of the genotypes shown on the X-axis. Boxes are delimited by the first (Q1, lower hinge) and third (Q3, upper hinge) quartiles. Whiskers represent the most extreme data points that are no more than Q3 + 1.5 × IQR or no less than Q1 − 1.5 × IQR, where the interquartile range (IQR) is Q3 − Q1. ♦: Mean. —: Median. Asterisks indicate significant differences from the corresponding parental lines (indicated by color) in a Student's *t*-test (*$P$ <0.05 and **$P$ <0.0001). At least 15 rosettes per genotype were measured from plants collected 21 das. (**I**) Schematic representation of the *AGO1* gene and molecular effects of the *ago1-52* mutation. Gene structure is represented as described in the legend of Figure 1. gDNA and cDNA indicate genomic and complementary DNA, respectively. The molecular changes in mutant cDNAs and proteins are shown as red letters. The genuine and novel 3′SSs are boxed in blue and yellow, respectively. (**J**) RT-qPCR analysis of the expression of total (*tAGO1*), wild-type (*wAGO1*), and mutant (*ago1-52*) mRNA splice variants in plants of the genotypes shown. (**K**) Percentage of *wAGO1* and *ago1-52* mRNA splice variants. Error bars in (J, K) indicate standard deviation. (**L**) Detection of AGO1 protein isoforms by immunoblot analysis using a primary antibody against AGO1 (α-AGO1). Asterisks indicate the wild-type AGO1 (*) and mutant AGO1-52 (**) proteins. Detection of the RuBisCO large subunit with α-RbcL was used as a loading control. (**M**) Relative quantification of the wAGO1 and AGO1-52 proteins shown in (L), using the Image Studio Analysis software (LI-COR). Total RNA and proteins were extracted from plants collected 15 das.

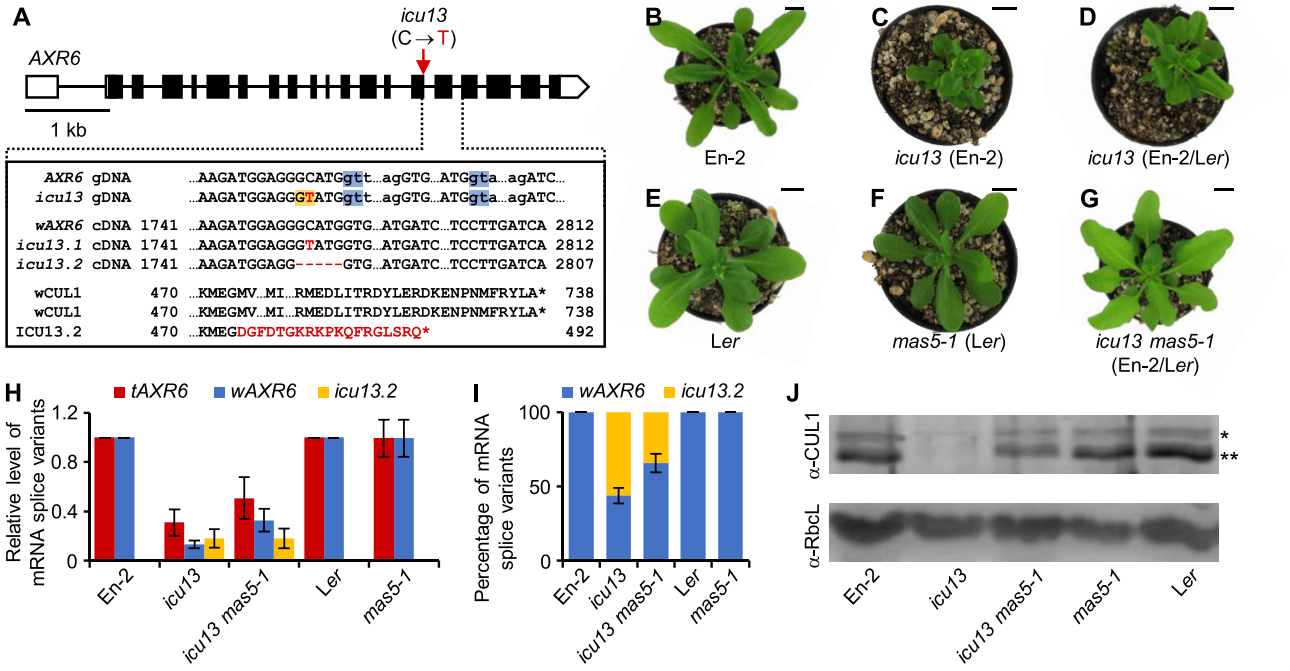


**Figure 3.** Suppression of the morphological and molecular phenotypes of *icu13* by *mas5-1*. (**A**) *AXR6* gene structure, mRNA splice variants, and CUL1 isoforms from the translation of *icu13* mRNA transcripts, represented as described in the legend of Figure 2. A red arrow indicates the position of the *icu13* mutation. (**B–G**) Rosettes of (B) En-2, (C) *icu13* (in the En-2 genetic background), (D) *icu13* (in the En-2/L*er* hybrid genetic background), (E) L*er*, (F) *mas5-1*, and (G) *icu13 mas5-1*. Photographs were taken 28 das. Scale bars: 1 cm. (**H**) RT-qPCR analysis of the expression of the total (*tAXR6*), wild-type (*wAXR6*) and mutant (*icu13.2*) mRNA splice variants in En-2, *icu13*, *icu13 mas5-1*, L*er*, and *mas5-1* plants. (**I**) Percentage of *wAXR6* and *icu13.2* mRNA splice variants. Error bars in (H, I) indicate standard deviations. (**J**) Immunoblot analysis of CUL1 proteins using a primary antibody against CUL1 (α-CUL1). Asterisks indicate RUB-modified CUL1 (*) and CUL1 (**). Detection of the RuBisCO large subunit with α-RbcL was used as a loading control. Total RNA and proteins were extracted from plants collected 15 das.

the spliceosome. The other mRNA variant (*icu13.2*) lacks the last 5 nt of the 15$^{th}$ exon, which causes a frameshift that generates a premature termination codon (PTC) due to the use of the novel 5′SS. Translation of the latter mRNA is predicted to produce a truncated protein (ICU13.2) with only 492 amino acids, instead of the 738 amino acids of the wild-type CUL1 (wCUL1) (Figure 3A).

We genotyped plants from all the phenotypic classes found in the $F_2$ progeny of a *mas5-1* × *icu13* cross. The *icu13/icu13;PRP8/PRP8* plants were identical to their *icu13/icu13* parent, whereas *icu13/icu13;PRP8/mas5-1* and *icu13/icu13;mas5-1/mas5-1* plants were similar to En-2 (Figure 3B-G). These results indicate that *mas5-1* acts as a dominant suppressor of the *icu13* mutant phenotype, as it does for *ago1-52*. Accordingly, we analyzed the relative levels of the mRNA variants known to be produced by *icu13* (18): *wAXR6* (including the completely wild-type *AXR6* variant and the *icu13.1* variant, which carries a synonymous mutation) and *icu13.2* (Figure 3A). Similar to previous findings, the levels of mature mRNAs produced by the *icu13* allele of *AXR6* were reduced 0.3-fold compared to wild type and less than 50% of these mRNAs were *wAXR6* (including the *icu13.1* variant). In the *icu13 mas5-1* double mutant, however, the mRNA levels were higher, and *wAXR6* became the major variant (Figure 3H and I).

*icu13.2* might be targeted by the nonsense-mediated mRNA decay (NMD) pathway, as its mutation maps to the 15$^{th}$ of its 20 exons (Figure 3A) and produces a PTC at the 16$^{th}$ exon. NMD is the major RNA surveillance pathway and is universal among eukaryotes; NMD recognizes and elicits the degradation of unproductive mRNA variants with PTCs, thereby preventing their translation (39). This would explain the low levels of mRNAs produced by the *icu13* allele, since its major variant *icu13.2* is likely to be degraded by NMD. However, *mas5-1* partially restored the use of the novel 5′SS of the *icu13* pre-mRNA, thereby decreasing the *icu13.2*/*wAXR6* ratio (Figure 3H and I). These findings explain why the total amounts of mature RNAs produced by *icu13* in the *icu13 mas5-1* double mutant were higher than those of the *icu13* single mutant.

We also examined the protein products of *icu13* in *icu13 mas5-1* plants by performing an immunoblot assay with a polyclonal antibody against CUL1. CUL1 was more abundant in *icu13 mas5-1* plants than in *icu13* (Figure 3J). These findings confirm (at the protein level) the suppression of *icu13* by *mas5-1* that we observed at the morphological and mRNA levels. Similar to a previous report (18), we did not detect the predicted truncated CUL1 isoform (ICU13.2) in *icu13* or *icu13 mas5-1* plants, reinforcing the notion that the NMD pathway degrades the *icu13.2* mRNA.

*ago1-51* and *sca3-1* carry transitions that damage a 5′SS of the *AGO1* and *SCA3* genes, respectively. In the case of *ago1-51*, three detectable mature mRNAs were produced, which include very low amounts of the wild-type variant (dubbed here as *wAGO1*) (Supplementary Figure S8A) (7,8). As in *ago1-51*, the cryptic 5′SS in *sca3-1* appears to be stronger than the damaged genuine one, as shown by the very low levels of wild-type *SCA3* mRNA (*sca3-1.1* in Supplementary Figure S8B) compared to those of the *sca3-1.2* variant, whose translation should result in a wild-type and a truncated protein, respectively (40). The *anu4-1* and *ang1-2* mutations damage a 3′SS that the spliceosome does not seem to recognize. Splicing of the *anu4-1* and *ang1-2* pre-mRNAs generates three different mRNA variants that suffer frameshifts (Supplementary Figure S8C and S8D), and in consequence do not produce detectable wild-type ANU4 and RPL10aB proteins, respectively (41). We performed Sanger sequencing to genotype plants from all the phenotypic classes in the different $F_2$ populations that we obtained. The *ago1-51 mas5-1*, *sca3-1 mas5-1*, *anu4-1 mas5-1* and *ang1-2 mas5-1* double mutant plants were indistinguishable from their respective *ago1-51*, *sca3-1*, *anu4-1* and *ang1-2* single mutant $F_2$ siblings (Supplementary Figure S8E-N), suggesting that in these mutants, *mas5-1* does not reduce the frequency of the selection of cryptic SSs by the spliceosome.

Our results indicate that *mas5-1* partially suppresses the missplicing caused by the preferential use of novel 5′SS (in *icu13*) or 3′SS (in *ago1-52*). Our results also suggest that *mas5-1* (and probably the other *mas5* mutations) does not have global effects on splicing. This would explain why the *mas5-1* single mutant is similar to the wild type, as has been shown for several missplicing suppressor alleles of *S. cerevisiae Prp8* and *C. elegans prp-8* (1,6).

### *ago1-52* synergistically interacts with hypomorphic alleles of *PRP8*

To compare the functional nature of the *mas5* alleles with other *prp8* alleles previously studied, we crossed *ago1* plants to *prp8-6* and *prp8-7* plants, which carry hypomorphic alleles of *PRP8* (Figure 4 and Supplementary Figure S9). The *prp8-6* mutant is in the L*er* genetic background (42), as are *ago1-51* and *ago1-52*, whereas *prp8-7* is in the Col-0 genetic background (43), as are *ago1-25* and *ago1-27*. Under our growth conditions, the rosette leaves of *prp8-6* and L*er* were very similar (Figure 4B), whereas those of *prp8-7* were slightly pointed, serrated, and pale (Figure 4C).

The residue altered by the *prp8-6* missense mutation (G1891E) in Arabidopsis is conserved with human PRPF8 (G1867) but not with yeast Prp8 (A1939; Supplementary Figure S2); PRP8 protein levels are similar in *prp8-6* and the wild type (42). The Arabidopsis hypomorphic *prp8-7* mutation causes a G1820E substitution in a 17-amino-acid extension within the RNase H-like domain of PRP8 (residues 1860–1875 in yeast, which correspond to residues 1812–1827 in Arabidopsis; Supplementary Figure S2). Cryo-EM analyses of yeast Prp8 revealed that this protein undergoes conformational rearrangements during pre-mRNA splicing and that the 17-amino-acid region can exist as a β-hairpin or a disordered loop, depending of the splicing step (44). Some missense *prp8* alleles affecting residues of this 17-amino-acid region stabilize the disordered loop conformation, which in turn provides high efficiency but low fidelity to the splicing of pre-mRNAs from reporter constructs. The growth of these yeast mutants with error-prone splicing resembles that of the wild type. By contrast, other *prp8* alleles harbor missense mutations affecting the same 17-amino-acid region that stabilize the β-hairpin conformation and cause low efficiency but high-fidelity splicing; therefore, the growth of yeast harboring these alleles is worse than the wild type (5). The global effect of the highly efficient but error-prone splicing caused by *prp8-7* is the retention of a low amount (6.7%) of introns (43).

We believe that the different combinations of Col-0 and L*er* genetic backgrounds that are present in the *ago1-25 prp8-6*, *ago1-27 prp8-6*, *ago1-25 mas5-1*, *ago1-27 mas5-1*, *ago1-25 mas5-3* and *ago1-27 mas5-3* double mutants contribute to their phenotypes (Supplementary Figures S6 and S9), making difficult to interpret the phenotypes of these double mutants. However, the *ago1-52 prp8-6* and *ago1-52 prp8-7* plants displayed a more severe morphological phenotype than *ago1-52* and were completely sterile; in addition, the phyllotaxy of *ago1-52 prp8-7* was strongly affected (Figure 4D–F). However, RT-qPCR and immunoblot analyses allowed us to conclude that the *prp8-6* and *prp8-7* mutations do not seem to modify *ago1-52* pre-mRNA splicing, and that the synergistic morphological phenotypes of *ago1-52 prp8-6* and *ago1-52 prp8-7* plants cannot be explained by an increase in *ago1-52* missplicing (Figure 4G-J). Nevertheless, these results clearly reveal that the *mas5* alleles are not hypomorphic.

### The *mas5* mutations do not alter global pre-mRNA splicing, but modify the ratio of the proximal/distal 3′SS use in NAG-NAG motifs

Some *prp8* suppressor mutations modify the genetic interactions among mutant alleles of the genes encoding different spliceosome factors and cofactors, alleles that cause global missplicing (1). This is the case for *prp8-8* and *prp8-9*; these alleles were isolated in a genetic screen for suppressors of the phenotype of the Arabidopsis *atprmt5-1* mutant, which carries a T-DNA insertion in the 21$^{st}$ exon of *PROTEIN ARGININE METHYLTRANSFERASE 5* (*PRMT5*). Arabidopsis PRMT5 regulates constitutive and alternative pre-mRNA splicing by promoting spliceosome assembly and activation (45–47). The *atprmt5-1* mutation causes an increase in intron retention (IR) events, a global splicing alteration that is suppressed by the *prp8-8* (P347S) and *prp8-9* (P1141S) mutations, which are both considered neomorphic, since the loss of function of Arabidopsis *PRP8* did not suppress the splicing defects of *atprmt5-1* (48).

To shed further light on the functional nature of the *mas5* alleles, we generated the *atprmt5-1 mas5-1* double mutant, which was indistinguishable from *atprmt5-1* (Supplementary Figure S10). Therefore, whereas *prp8-8* and *prp8-9* (which are dominant, like *mas5-1* and *mas5-3*) suppressed *atprmt5-1, mas5-1* did not. These results strongly suggest that these different alleles of *PRP8* alter different PRP8 protein activities, as expected from the different locations of the amino acids changed by the *mas5-1*, *prp8-8* and *prp8-9* mutations in PRP8 (Supplementary Figure S2).

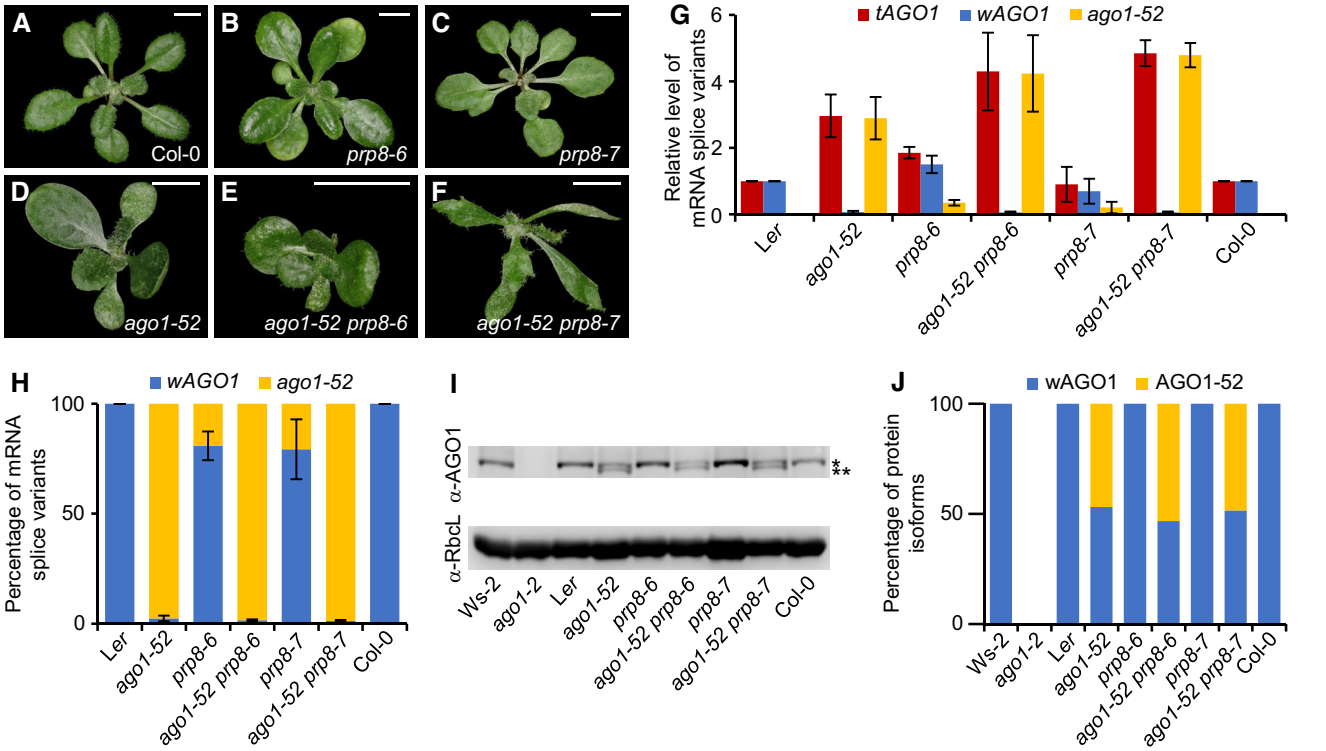


**Figure 4.** Genetic interactions between *prp8* hypomorphic alleles and *ago1-52*. (**A–F**) Rosettes of (A) Col-0, (B) *prp8-6*, (C) *prp8-7*, (D) *ago1-52*, (E) *ago1-52 prp8-6*, and (F) *ago1-52 prp8-7*. Photographs were taken 21 das. Scale bars: 4 mm. (**G**) RT-qPCR analysis of the expression of total (*tAGO1*), wild-type (*wAGO1*), and mutant (*ago1-52*) *AGO1* mRNA splice variants. (**H**) Percentage of *wAGO1* and *ago1-52* splice variants. Error bars in (G, H) indicate standard deviation. (**I**) Detection of AGO1 protein isoforms by immunoblot analysis using a primary antibody against AGO1 (α-AGO1). Asterisks indicate wild-type AGO1 (*) and mutant AGO1-52 (**). Detection of the RuBisCO large subunit with α-RbcL was used as a loading control. (**J**) Relative quantification of the wAGO1 and AGO1-52 proteins shown in (I), using the Image Studio Analysis software (LI-COR). Total RNA and proteins were extracted from plants collected 15 das.

To test whether the *mas5* mutations cause global alterations in pre-mRNA splicing, we carried out RNA-seq analyses of RNA extracted from L*er*, *mas5-1*, and *mas5-3*. Using the ASpli software (21), 98,488 exons and 118,974 introns from 21,790 multiexonic genes were evaluated for each sample (Supplementary Dataset S1), which excluded external exons and Ios (see Materials and Methods). We filtered bin-based splicing events using an FDR <0.1 and an absolute Delta PSI or Delta PIR >5%. We only found 251 and 164 differential splicing events in *mas5-1* and *mas5-3*, respectively, compared to the wild type, 33 of which were common to both mutants (Figure 5A and Supplementary Datasets S2–S4).

Increased IR events were the most frequent: 201 (80.1% of the total missplicing events) and 117 (71.3%) in *mas5-1* and *mas5-3*, respectively, with 25 common to both mutants (Figure 5A and Supplementary Datasets S2-S4). The second most frequent event found in both mutants was decreased Alt 3′SS: 31 (12.3%) and 25 (21.4%) in *mas5-1* and *mas5-3*, respectively, with only 7 common to both mutants (Figure 5A and Supplementary Datasets S2-S4).

Using the IGV software, we confirmed the increased IR events in both mutants (Figure 5B) and found that 27 out of 31 (in *mas5-1*) and 21 out of 25 (in *mas5-3*) of the decreased Alt 3′SS events affected tandem 3′SSs, which were exactly 3 nt apart (NAGNAG). We also confirmed that both mutants used the proximal 3′SS more frequently than the distal one, compared to the wild type, which also uses both 3′SSs (Figures 5C and D, and Supplementary Figure S11). The presence of NAGNAG motifs in the 3′SS occurs widely in eukaryotic genomes, including the human and Arabidopsis genomes, in which 1,890 have been found by analyzing 435 RNA-seq datasets, with a mean number of 201 NAGNAG motifs with confirmed alternative use per sample (49). Because both 3′SSs are exactly 3 nt apart (in-frame), their alternative choice for the spliceosome would produce proteins differing in a single amino acid, which might not affect its function. Interestingly, in all cases the proximal 3′SS, which seems to be the strongest one because is more frequently chosen by the wild type, seemed to be more favored over the distal 3′SS in both *mas5* mutants (Supplementary Datasets S2 and S3).

Comparing the number of IR events previously detected in *prp8-7* (8,124 events affecting 6.7% of total introns; 43), we conclude that the *mas5-1* and *mas5-3* suppressor mutations do not alter global pre-mRNA splicing.

Misspliced mRNAs prevent the recruitment of mRNA export factors, causing nuclear accumulation of poly(A)+ RNAs (50), which should be evident in *prp8-7*, but not in *mas5-1* or *mas5-3*, according to the RNA-seq results. To test this hypothesis, we carried out RNA-FISH assays with a fluorescently labeled oligo-dT probe against

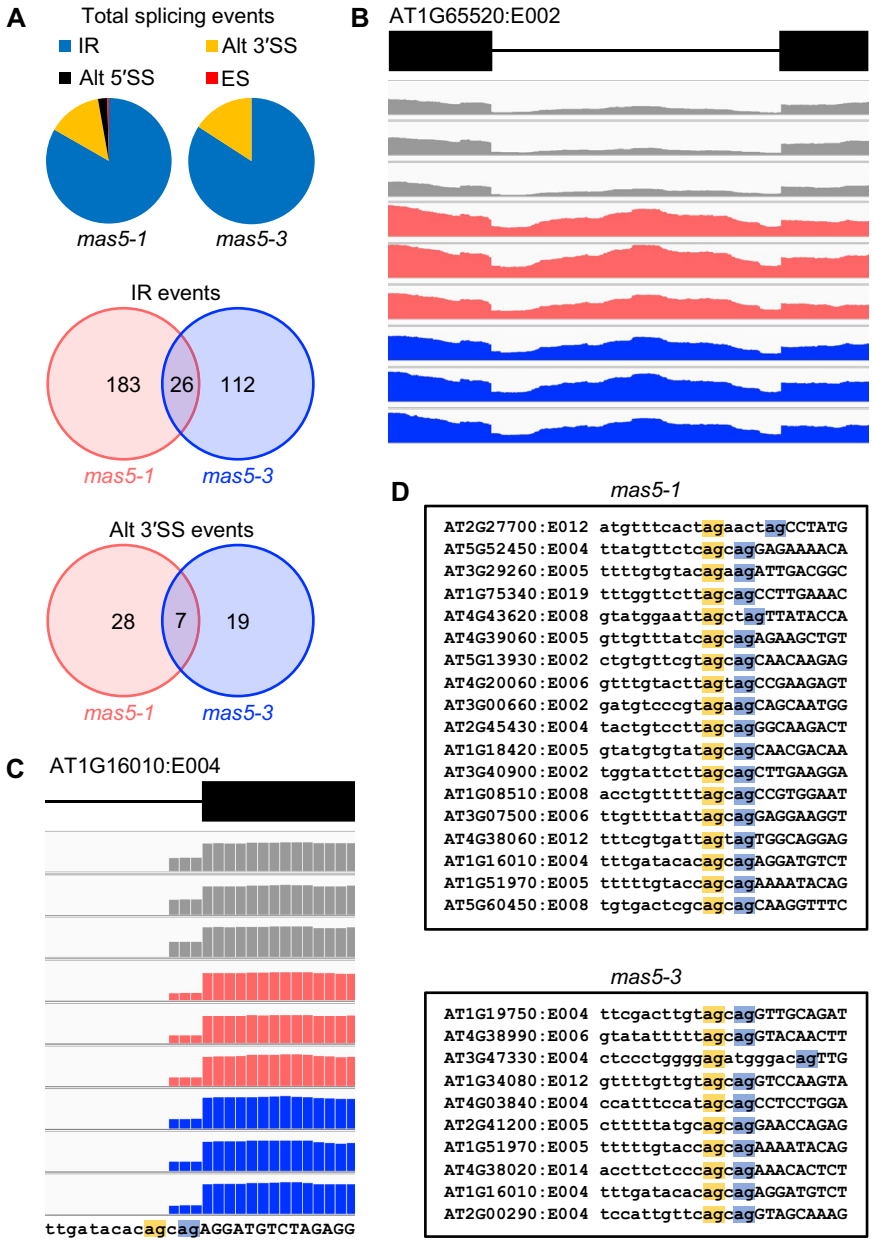


**Figure 5.** Genome-wide analysis of pre-mRNA splicing in the *mas5* mutants. (**A**) Percentage of differential splicing events identified in *mas5-1* and *mas5-3*, and Venn diagrams showing the IR and Alt 3′SS events. IR: intron retention; ES: exon skipping; Alt 3′SS/5′SS: alternative 3′/5′ splicing site. (**B, C**) Plots of AT1G65520 and AT1G16010 aligned reads taken as representative examples of (B) IR and (C) Alt 3′SS events that were statistically different in the three biological samples of L*er* (in grey), *mas5-1* (in red), and *mas5-3* (in blue). Only the intron and flanking exons (represented in gene structures by black lines and boxes, respectively) corresponding to the sites of the events are shown in plots obtained with the IGV software (http://software.broadinstitute.org/software/igv/). (**D**) DNA sequences corresponding to the statistically significant Alt 3′SS events identified in *mas5-1* and *mas5-3*, with an absolute Delta PSI >20%. Intronic and exonic sequences are shown in lowercase and uppercase, respectively. The proximal and distal 3′SSs are boxed in blue and yellow, respectively. E0XX indicates the exon number of each gene.

poly(A)+ RNAs. We used as a positive control the *sar1-4* mutant, which carries a null allele of *SUPPRESSOR OF AUXIN RESISTANCE1* (*SAR1*), encoding NUCLEOPORIN160 (NUP160), and shows elevated nuclear retention of poly(A)+ RNAs (22,50). We found nuclear accumulation of poly(A)+ RNAs within the nucleus of *prp8-7* and *sar1-4* leaf palisade mesophyll cells, but not in *mas5-1* (Figure 6). These results further support our RNA-seq results.

## DISCUSSION

### The *mas5* mutations belong to an unusual class of missplicing suppressors that increase splicing fidelity without causing a dramatic global increase in missplicing

Point mutations at the 5′SSs or 3′SSs may abolish or reduce their recognition by the spliceosome, which then recognizes a cryptic site close to the genuine, mutated SS. Moreover, point mutations in exonic or intronic sequences may cre-

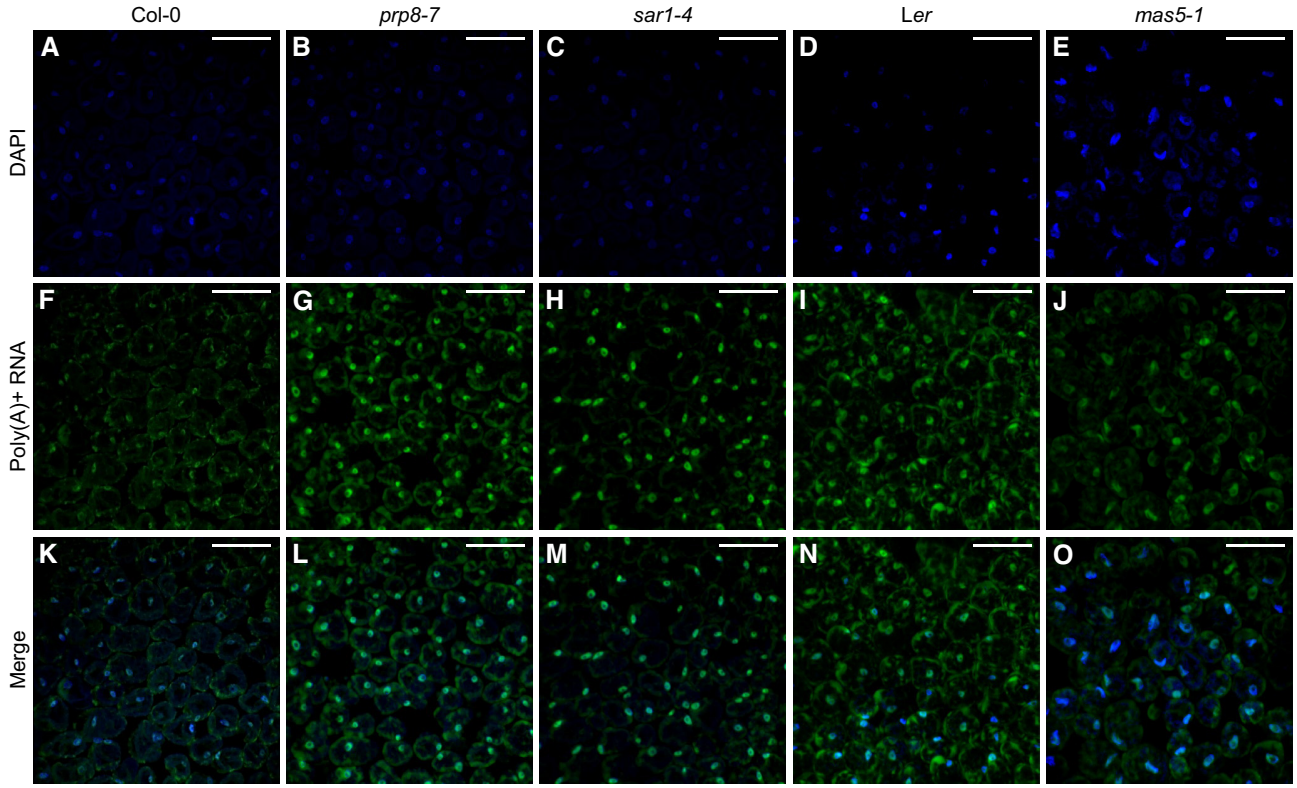


**Figure 6.** Detection of poly(A)+ RNAs in *prp8-7* and *mas5-1* leaf cells. (**A–O**) Poly(A)+ RNA-FISH assays in palisade mesophyll cells of (A, F, K) Col-0, (B, G, L) *prp8-7*, (C, H, M) *sar1-4*, (D, I, N) L*er* and (E, J, O) *mas5-1* leaves. Fluorescent signals correspond to (A–E) nuclear 4′,6-diamidino-2-phenylindole (DAPI) staining, (F–J) fluorescein from an oligo(dT) probe, and (K–O) their overlay. Confocal laser-scanning micrographs were taken from 10 leaves per genotype of plants collected 14 das. Scale bars: 50 μm.

ate novel SSs. *PRP8* encodes the core component of the spliceosome and is conserved across all eukaryotes. Numerous *prp8* alleles have been described to act as missplicing suppressors, mainly in *S. cerevisiae*. Most of these mutations promote changes in SS choice by the spliceosome, increasing the frequency of proper, in-frame splicing of pre-mRNAs. These suppressor mutations reduce the frequency of use of cryptic SSs, which are already close to a mutated genuine SS (1).

We identified five different, allelic *mas5* mutations in the Arabidopsis *PRP8* gene in a second-site mutagenesis screen for extragenic suppressors of the morphological phenotype of *ago1-52*, a mutant allele of *AGO1* that undergoes missplicing. Except *mas5-4*, all *mas5* mutations affect residues of the cavity and the Linker of PRP8, where most yeast suppressor mutations are also located (Figure 1J and K, and Supplementary Figure S2). In the five *mas5* alleles, the wild-type residue is predicted to be replaced by a positively charged amino acid (K in the identical *mas5-1* and *mas5-2* mutations, and *mas5-3*, and R in *mas5-6*), or a negatively charged amino acid is removed (D in *mas5-5*). The cavity interacts with the 5′SS, 3′SS and BPS, and is thought to play an important role in splicing fidelity. Our results suggest that regions that make up the cavity are also hotspots for missplicing suppressor mutations in Arabidopsis PRP8 that are likely involved in splicing fidelity. However, when we looked at the yeast residues homologous to those affected by the Arabidopsis *mas5* mutations in five cryo-EM structures of spliceosomal complexes, we did not find interactions with mRNA, snRNAs, or other spliceosomal factors, except for the residues homologous to those affected by the *mas5-3* and *mas5-4* mutations; these residues could interact with the intron lariat–3′ exon (stabilizing the 3′SS for the second transesterification) and 5′SS of the single-intron pre-mRNA used in the models, respectively (Supplementary Figures S3 and S4). It is possible that the *mas5* mutations modify the conformation of these spliceosomal complexes, but this remains to be tested.

Some dominant alleles are antimorphic (with a dominant negative effect); when they are heterozygous with a wild-type allele, they antagonize the function of the wild-type protein, thus leading to a loss of function. This mainly occurs in genes encoding subunits of multimeric complexes (51), as is the case of *PRP8*. In genetic screens, performed in *S. cerevisiae*, several dominant alleles of *Prp8* have been identified that alter both SS choice by the spliceosome and alternative splicing efficiency (1). Many of these mutations do not have detrimental effects or visible phenotypes.

However, null alleles of genes encoding components of the spliceosome, or its associated factors, can cause global missplicing and lethality, as is the case of *PRP8*. Our RNA-seq analysis showed that the *mas5-1* and *mas5-3* mutations, and probably the other *mas5* mutations, do not cause global defects in splicing, since only a few missplicing events were detected. Most of these missplicing events are increased IR events, corresponding to 0.17% (201 events) and

0.1% (117 events) of the introns analyzed in *mas5-1* and *mas5-3*, respectively (Figure 5A and Supplementary Figures S2 and S3), which are minimal compared to those found in the *prp8-7* mutant (8,124 events, corresponding to 6.7% of total introns), which exhibits a very weak morphological phenotype (Figure 4C) (43). Our results suggest that the *mas5-1* and *mas5-3* mutations improve the choice of the strongest 3′SS (the proximal one) compared with the wild type, at least in cases where there are NAGNAG sequences. These findings are in line with our previous results that suggested an increase in splicing fidelity in the *ago1-52 mas5-1*, *ago1-52 mas5-3* and *icu13 mas5-1* double mutants, and explain why *mas5-1* and *mas5-3* plants exhibit a wild-type phenotype. Based on their suppression of the missplicing of *ago1-52* and *icu13*, we propose that the *mas5* mutations represent a class of novel and uncommon *PRP8* alleles whose behavior differs from that of alleles that increase splicing fidelity by suppressing cryptic splicing.

In animals and land plants, around 25% of the alternative splicing events are due to the use of alternative 3′SSs and 5′SSs, and about half of these 3′SSs are present in a NAGNAG motif and thus are separated by only 3 nt (52–55). In most cases, NAG tandem repeats are in phase and their differential splicing events give rise to a protein with an insertion or a deletion of a single amino acid (52,53,56). There is evidence that both protein isoforms from hundreds of genes with NAGNAG 3′SSs exist in Arabidopsis, rice (*Oryza sativa*), and the moss *Physcomitrella patens* (53–55). For example, the alternative splicing of the 3′SS of intron 14 in the Arabidopsis *ZINC-INDUCED FACILITATOR-LIKE1* gene produces two mRNA variants that differ by 2 nt. One of these mRNAs codes for a full-length protein that localizes to the plasma membrane and functions in auxin-regulated processes, whereas the second variant codes for a truncated protein that localizes to the tonoplast membrane and functions in drought tolerance (57). The use of an alternative 3′SS in a NAGNAG sequence also produces the two isoforms of Arabidopsis U1-35K, a factor involved in splicing of rare U12-type introns. The shorter isoform, which lacks a glutamine, exhibits altered binding affinity to different components of the spliceosome complex (58). These studies suggest the functional significance of alternative splicing, as a result of the presence of tandem 3′SSs in plants.

In addition, our *mas5* alleles appear to differ from other Arabidopsis *prp8* dominant alleles, such as *prp8-8* and *prp8-9* (48), because they do not suppress the morphological phenotype caused by mutations in *ATPRMT5*, which encodes another spliceosome-related factor (Supplementary Figure S10). These findings strongly suggest that these different alleles of *PRP8* alter different activities of PRP8, as expected based on the different localizations of the *mas5-1*, *prp8-8* and *prp8-9* mutations (Supplementary Figure S2).

### Effect of mutations that create novel SSs but do not alter genuine SSs in model species and humans

Base substitutions are the most frequent type of mutations induced by the chemical mutagens most widely used to study model organisms, and they represent the major form of spontaneous genetic polymorphisms found in many species, including humans (59). The identification of mutated genes that cause a phenotype of interest has traditionally relied on the use of iterative linkage analysis to identify candidate mutations. Such candidate mutations are commonly chosen by focusing mostly on nonsynonymous substitutions in exons or, to a lesser extent, on substitutions that disturb SSs (60). However, in not few cases, none of the candidate genes was ultimately found to be the causal gene for the phenotype under study, despite recent progress in whole-genome sequencing technologies. Some of these cloning failures could be due to mutations that remain unnoticed because they create a synonymous codon or occur in a deep intronic region that does not form part of a genuine SS. Nevertheless, the effects of these apparently silent mutations can be strong, since some create novel SSs that are favored by the spliceosome compared to the genuine SSs, even though these SSs are otherwise intact. *ago1-52* and *icu13* belong to this class of mutations. The morphological and molecular phenotypes of *ago1-52* are caused by a point mutation in an intronic region that has no obvious functional role, whereas in *icu13*, these phenotypes appear to be caused by a synonymous change at the end of an exon. In both cases, however, the mutation creates a novel SS that causes missplicing.

Recent studies integrating DNA and RNA data from whole-genome exon sequencing and transcriptomic analysis revealed that human mutations in deep intronic regions or those that yield synonymous codons in coding regions are the causes of several hereditary disorders and have been associated with cancer. A computational genomic analysis of 235 individuals of the 1000 Genomes Project estimated that each genome contains an average of 10 intronic mutations in sequences other than SSs or BPS. These mutations are associated with disorders, since they generate novel SSs without damaging genuine SSs, which in turn cause missplicing and often introduce a PTC in the misspliced mRNA (60). In addition, computational analysis of 8,656 tumors from The Cancer Genome Atlas project discovered several hundred novel mutations in intronic sequences, which cause missplicing and might have an impact on cancer; some of these mutations damage key tumor suppressor genes, such as *TP53*, the key tumor suppressor gene that encodes P53, the so-called guardian of the genome (61,62). These mutations cannot be detected by sequencing exomes, which is the most frequently used method to identify mutations associated with human genetic disorders.

### The study of missplicing suppressors may be useful for a better understanding of splicing, as well as for engineering SS selection by the spliceosome

Due to its relative simplicity and rapid growth, *S. cerevisiae* has traditionally been recognized as the best model organism to study several cross-kingdom conserved processes, including splicing. However, 97% of protein-coding genes of *S. cerevisiae* lack introns, and several splicing factors and cofactors that are present in multicellular organisms are not encoded by its genome, including those that participate in alternative splicing, an event that is rare in this yeast but common in plants and animals (63). Several animal species are used as models to better understand missplicing caus-

ing human diseases and to design strategies for suppressing missplicing (64). Our findings indicate that Arabidopsis, like other multicellular organisms, could be useful for analyzing human disorders involving highly conserved genes, such as *PRP8*. It might be possible to suppress the effects of some mutations that cause missplicing in mammalian and particularly human cells by obtaining mutations equivalent to the *mas5* mutations that mutate amino acids that exhibit cross-kingdom conservation and do not impair Arabidopsis growth or development.

Our findings also suggest that mutants that show missplicing may be good candidates for investigating both missplicing suppression and splicing itself. Indeed, such an approach might be a better choice than using minigenes to recapitulate artificial exon skipping events, because mutations such as those in the *mas5* lines, are present in their natural cellular and chromosomal context.

## DATA AVAILABILITY

Sequence data from this article can be found at The Arabidopsis Information Resource (TAIR; https://www.arabidopsis.org) under the following accession numbers: PRP8 (AT1G80070), AGO1 (AT1G48410), SCA3 (AT2G24120), ANU4 (AT1G02280), ANG1 (AT2G27530), AXR6 (AT4G02570), SAR1 (AT1G33410), and ATPRMT5 (AT4G31120). All the FASTQ files were submitted to the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information (NCBI) under the BioProject accession PRJNA787038 (https://www.ncbi.nlm.nih.gov/sra/PRJNA787038).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors thank J.A. García-Martín for the global splicing analysis, J.M. Serrano, T. Trujillo, J. Castelló and D. Navarro for their excellent technical assistance, J.C. del Pozo for providing the anti-CUL1 antibody, and J.L. Micol for useful discussions and comments on the manuscript, as well as for the use of his facilities.

*Author Contributions*: M.R.P. obtained funding and conceived, designed, and supervised research; all authors performed research, interpreted the results and wrote the manuscript.

## FUNDING

Ministerio de Ciencia e Innovación of Spain [BIO2017-89728-R and PID2020-117125RB-I00 (MCI/AEI/FEDER, UE)]; Generalitat Valenciana [PROMETEO/2019/117 to M.R.P.]. R.M.-P. held a postdoctoral fellowship from the Generalitat Valenciana [APOSTD/2019/001].

*Conflict of interest statement.* None declared.

## REFERENCES

1. Grainger,R.J. and Beggs,J.D. (2005) Prp8 protein: At the heart of the spliceosome. *RNA*, **11**, 533–557.
2. Hoskins,A.A. and Moore,M.J. (2012) The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.*, **37**, 179–188.
3. Galej,W.P., Nguyen,T.H., Newman,A.J. and Nagai,K. (2014) Structural studies of the spliceosome: zooming into the heart of the machine. *Curr. Opin. Struct. Biol.*, **25**, 57–66.
4. Joynt,A.T., Evans,T.A., Pellicore,M.J., Davis-Marcisak,E.F., Aksit,M.A., Eastman,A.C., Patel,S.U., Paul,K.C., Osorio,D.L., Bowling,A.D. *et al.* (2020) Evaluation of both exonic and intronic variants for effects on RNA splicing allows for accurate assessment of the effectiveness of precision therapies. *PLoS Genet.*, **16**, e1009100.
5. Mayerle,M., Raghavan,M., Ledoux,S., Price,A., Stepankiw,N., Hadjivassiliou,H., Moehle,E.A., Mendoza,S.D., Pleiss,J.A., Guthrie,C. *et al.* (2017) Structural toggle in the RNaseH domain of Prp8 helps balance splicing fidelity and catalytic efficiency. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 4739–4744.
6. Mayerle,M., Yitiz,S., Soulette,C., Rogel,L.E., Ramirez,A., Ragle,J.M., Katzman,S., Guthrie,C. and Zahler,A.M. (2019) Prp8 impacts cryptic but not alternative splicing frequency. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 2193–2199.
7. Jover-Gil,S., Candela,H., Robles,P., Aguilera,V., Barrero,J.M., Micol,J.L. and Ponce,M.R. (2012) The microRNA pathway genes *AGO1*, *HEN1* and *HYL1* participate in leaf proximal-distal, venation and stomatal patterning in Arabidopsis. *Plant Cell Physiol.*, **53**, 1322–1333.
8. Sánchez-García,A.B., Aguilera,V., Micol-Ponce,R., Jover-Gil,S. and Ponce,M.R. (2015) Arabidopsis *MAS2*, an essential gene that encodes a homolog of animal NF-kappa B activating protein, is involved in 45S ribosomal DNA silencing. *Plant Cell*, **27**, 1999–2015.
9. Weigel,D. and Glazebrook,J. (2006) EMS mutagenesis of *Arabidopsis* seed. *Cold Spring Harb. Protoc.*, **2006**, pdb.prot4621.
10. Micol-Ponce,R., Aguilera,V. and Ponce,M.R. (2014) A genetic screen for suppressors of a hypomorphic allele of Arabidopsis *ARGONAUTE1*. *Sci. Rep.*, **4**, 5533.
11. Kulichová,K., Kumar,V., Steinbachová,L., Klodová,B., Timofejeva,L., Juříček,M., Honys,D. and Hafidh,S.S. (2020) PRP8A and PRP8B spliceosome subunits act coordinately to control pollen tube attraction in *Arabidopsis thaliana*. *Development*, **147**, dev186742.
12. Schwartz,B., Yeung,C. and Meinke,W. (1994) Disruption of morphogenesis and transformation of the suspensor in abnormal suspensor mutants of Arabidopsis. *Development*, **120**, 3235–3245.
13. Ponce,M.R., Quesada,V. and Micol,J.L. (1998) Rapid discrimination of sequences flanking and within T-DNA insertions in the Arabidopsis genome. *Plant J.*, **14**, 497–501.
14. Berná,G., Robles,P. and Micol,J.L. (1999) A mutational analysis of leaf morphogenesis in *Arabidopsis thaliana*. *Genetics*, **152**, 729–742.
15. Ponce,M.R., Robles,P., Lozano,F.M., Brotóns,M.A. and Micol,J.L. (2006) Low-resolution mapping of untagged mutations. *Methods. Mol. Biol.*, **323**, 105–113.
16. Ponce,M.R., Robles,P. and Micol,J.L. (1999) High-throughput genetic mapping in *Arabidopsis thaliana*. *Mol. Gen. Genet.*, **261**, 408–415.
17. Pérez-Pérez,J.M., Rubio-Díaz,S., Dhondt,S., Hernández-Romero,D., Sánchez-Soriano,J., Beemster,G.T., Ponce,M.R. and Micol,J.L. (2011) Whole organ, venation and epidermal cell morphological variations are correlated in the leaves of *Arabidopsis* mutants. *Plant Cell Environ.*, **34**, 2200–2211.
18. Esteve-Bruna,D., Pérez-Pérez,J.M., Ponce,M.R. and Micol,J.L. (2013) *incurvata13*, a novel allele of *AUXIN RESISTANT6*, reveals a specific role for auxin and the SCF complex in Arabidopsis embryogenesis, vascular specification, and leaf flatness. *Plant Physiol.*, **161**, 1303–1320.
19. Wingett,S.W. and Andrews,S. (2018) FastQ screen: a tool for multi-genome mapping and quality control. *F1000Research*, **7**, 1338.
20. 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.
21. Mancini,E., Rabinovich,A., Iserte,J., Yanovsky,M. and Chernomoretz,A. (2021) ASpli: integrative analysis of splicing landscapes through RNA-Seq assays. *Bioinformatics*, **37**, 2609–2616.
22. Parry,G., Ward,S., Cernac,A., Dharmasiri,S. and Estelle,M. (2006) The Arabidopsis SUPPRESSOR OF AUXIN RESISTANCE proteins are nucleoporins with an important role in hormone signaling and development. *Plant Cell*, **18**, 1590–1603.

23. Micol-Ponce,R., Sarmiento-Mañús,R., Fontcuberta-Cervera,S., Cabezas-Fuster,A., de Bures,A., Sáez-Vásquez,J. and Ponce,M.R. (2020) SMALL ORGAN4 is a ribosome biogenesis factor involved in 5.8S ribosomal RNA maturation. *Plant Physiol.*, **184**, 2022–2039.
24. Galej,W.P., Oubridge,C., Newman,A.J. and Nagai,K. (2013) Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature*, **493**, 638–643.
25. Pena,V., Rozov,A., Fabrizio,P., Lührmann,R. and Wahl,M.C. (2008) Structure and function of an RNase H domain at the heart of the spliceosome. *EMBO J.*, **27**, 2929–2940.
26. Pena,V., Liu,S., Bujnicki,J.M., Lührmann,R. and Wahl,M.C. (2007) Structure of a multipartite protein-protein interaction domain in splicing factor Prp8 and its link to *Retinitis pigmentosa*. *Mol. Cell*, **25**, 615–624.
27. Fica,S.M., Oubridge,C., Galej,W.P., Wilkinson,M.E., Bai,X.C., Newman,A.J. and Nagai,K. (2017) Structure of a spliceosome remodelled for exon ligation. *Nature*, **542**, 377–380.
28. Brow,D.A. (2019) An allosteric network for spliceosome activation revealed by high-throughput suppressor analysis in *Saccharomyces cerevisiae*. *Genetics*, **212**, 111–124.
29. Bai,R., Wan,R., Yan,C., Jia,Q., Lei,J. and Shi,Y. (2021) Mechanism of spliceosome remodeling by the ATPase/helicase Prp2 and its coactivator Spp2. *Science*, **371**, eabe8863.
30. Wan,R., Bai,R., Yan,C., Lei,J. and Shi,Y. (2019) Structures of the catalytically activated yeast spliceosome reveal the mechanism of branching. *Cell*, **177**, 339–351.
31. Wilkinson,M.E., Fica,S.M., Galej,W.P. and Nagai,K. (2021) Structural basis for conformational equilibrium of the catalytic spliceosome. *Mol. Cell*, **81**, 1439–1452.
32. Yan,C., Wan,R., Bai,R., Huang,G. and Shi,Y. (2017) Structure of a yeast step II catalytically activated spliceosome. *Science*, **355**, 149–155.
33. Liu,S., Li,X., Zhang,L., Jiang,J., Hill,R.C., Cui,Y., Hansen,K.C., Zhou,Z.H. and Zhao,R. (2017) Structure of the yeast spliceosomal postcatalytic P complex. *Science*, **358**, 1278–1283.
34. Collins,C.A. and Guthrie,C. (1999) Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. *Genes Dev.*, **13**, 1970–1982.
35. Umen,J.G. and Guthrie,C. (1996) Mutagenesis of the yeast gene *PRP8* reveals domains governing the specificity and fidelity of 3′ splice site selection. *Genetics*, **143**, 723–739.
36. Morel,J.B., Godon,C., Mourrain,P., Béclin,C., Boutet,S., Feuerbach,F., Proux,F. and Vaucheret,H. (2002) Fertile hypomorphic *ARGONAUTE* (*ago1*) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell*, **14**, 629–639.
37. Bohmert,K., Camus,I., Bellini,C., Bouchez,D., Caboche,M. and Benning,C. (1998) *AGO1* defines a novel locus of *Arabidopsis* controlling leaf development. *EMBO J.*, **17**, 170–180.
38. Gray,W.M., del Pozo,J.C., Walker,L., Hobbie,L., Risseeuw,E., Banks,T., Crosby,W.L., Yang,M., Ma,H. and Estelle,M. (1999) Identification of an SCF ubiquitin-ligase complex required for auxin response in *Arabidopsis thaliana*. *Genes Dev.*, **13**, 1678–1691.
39. Isken,O. and Maquat,L.E. (2008) The multiple lives of NMD factors: balancing roles in gene and genome regulation. *Nat. Rev. Genet.*, **9**, 699–712.
40. Hricová,A., Quesada,V. and Micol,J.L. (2006) The *SCABRA3* nuclear gene encodes the plastid RpoTp RNA polymerase, which is required for chloroplast biogenesis and mesophyll cell proliferation in Arabidopsis. *Plant Physiol.*, **141**, 942–956.
41. Mateo-Bonmatí,E., Casanova-Sáez,R., Candela,H. and Micol,J.L. (2014) Rapid identification of *angulata* leaf mutations using next-generation sequencing. *Planta*, **240**, 1113–1122.
42. Marquardt,S., Raitskin,O., Wu,Z., Liu,F., Sun,Q. and Dean,C. (2014) Functional consequences of splicing of the antisense transcript *COOLAIR* on *FLC* transcription. *Mol. Cell*, **54**, 156–165.
43. Sasaki,T., Kanno,T., Liang,S.C., Chen,P.Y., Liao,W.W., Lin,W.D., Matzke,A.J. and Matzke,M. (2015) An Rtf2 domain-containing protein influences pre-mRNA splicing and is essential for embryonic development in *Arabidopsis thaliana*. *Genetics*, **200**, 523–535.
44. Garside,E.L., Whelan,T.A., Stark,M.R., Rader,S.D., Fast,N.M. and MacMillan,A.M. (2019) Prp8 in a reduced spliceosome lacks a conserved toggle that correlates with splicing complexity across diverse taxa. *J. Mol. Biol.*, **431**, 2543–2553.
45. Sanchez,S.E., Petrillo,E., Beckwith,E.J., Zhang,X., Rugnone,M.L., Hernando,C.E., Cuevas,J.C., Godoy Herz,M.A., Depetris-Chauvin,A., Simpson,C.G. *et al.* (2010) A methyl transferase links the circadian clock to the regulation of alternative splicing. *Nature*, **468**, 112–116.
46. Hernando,C.E., Sanchez,S.E., Mancini,E. and Yanovsky,M.J. (2015) Genome wide comparative analysis of the effects of PRMT5 and PRMT4/CARM1 arginine methyltransferases on the *Arabidopsis thaliana* transcriptome. *BMC Genomics*, **16**, 192.
47. Deng,X., Gu,L., Liu,C., Lu,T., Lu,F., Lu,Z., Cui,P., Pei,Y., Wang,B., Hu,S. *et al.* (2010) Arginine methylation mediated by the *Arabidopsis* homolog of PRMT5 is essential for proper pre-mRNA splicing. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 19114–19119.
48. Deng,X., Lu,T., Wang,L., Gu,L., Sun,J., Kong,X., Liu,C. and Cao,X. (2016) Recruitment of the NineTeen Complex to the activated spliceosome requires AtPRMT5. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 5447–5452.
49. Zhang,Q., Zhang,Q., Li,S., Ye,J., Tang,W., Yin,M., Wang,K., Wang,K., Shi,C., Wang,C. *et al.* (2018) AtNAGNAG: a comprehensive database for NAGNAG alternative splicing in *Arabidopsis thaliana*. In: *Proceedings of the 2nd International Conference on Big Data Research*. pp. 33–37.
50. Dong,C.H., Hu,X., Tang,W., Zheng,X., Kim,Y.S., Lee,B.H. and Zhu,J.K. (2006) A putative *Arabidopsis* nucleoporin, AtNUP160, is critical for RNA export and required for plant tolerance to cold stress. *Mol. Cell. Biol.*, **26**, 9533–9543.
51. Wilkie,A.O. (1994) The molecular basis of genetic dominance. *J. Med. Genet.*, **31**, 89–98.
52. Hiller,M., Huse,K., Szafranski,K., Jahn,N., Hampe,J., Schreiber,S., Backofen,R. and Platzer,M. (2004) Widespread occurrence of alternative splicing at NAGNAG acceptors contributes to proteome plasticity. *Nat. Genet.*, **36**, 1255–1257.
53. Iida,K., Shionyu,M. and Suso,Y. (2008) Alternative splicing at NAGNAG acceptor sites shares common properties in land plants and mammals. *Mol. Biol. Evol.*, **25**, 709–718.
54. Schindler,S., Szafranski,K., Hiller,M., Ali,G.S., Palusa,S.G., Backofen,R., Platzer,M. and Reddy,A.S. (2008) Alternative splicing at NAGNAG acceptors in *Arabidopsis thaliana* SR and SR-related protein-coding genes. *BMC Genomics*, **9**, 159.
55. Sinha,R., Zimmer,A.D., Bolte,K., Lang,D., Reski,R., Platzer,M., Rensing,S.A. and Backofen,R. (2010) Identification and characterization of NAGNAG alternative splicing in the moss *Physcomitrella patens*. *BMC Plant Biol.*, **10**, 76.
56. Bradley,R.K., Merkin,J., Lambert,N.J. and Burge,C.B. (2012) Alternative splicing of RNA triplets is often regulated and accelerates proteome evolution. *PLoS Biol.*, **10**, e1001229.
57. Remy,E., Cabrito,T.R., Baster,P., Batista,R.A., Teixeira,M.C., Friml,J., Sá-Correia,I. and Duque,P. (2013) A major facilitator superfamily transporter plays a dual role in polar auxin transport and drought stress tolerance in *Arabidopsis*. *Plant Cell*, **25**, 901–926.
58. Lorković,Z.J., Lehner,R., Forstner,C. and Barta,A. (2005) Evolutionary conservation of minor U12-type spliceosome between plants and humans. *RNA*, **11**, 1095–1107.
59. Collins,F.S., Brooks,L.D. and Chakravarti,A. (1998) A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res.*, **8**, 1229–1231.
60. Sakaguchi,N. and Suyama,M. (2021) In silico identification of pseudo-exon activation events in personal genome and transcriptome data. *RNA Biol.*, **18**, 382–390.
61. Jayasinghe,R.G., Cao,S., Gao,Q., Wendl,M.C., Vo,N.S., Reynolds,S.M., Zhao,Y., Climente-González,H., Chai,S., Wang,F. *et al.* (2018) Systematic analysis of splice-site-creating mutations in cancer. *Cell. Rep.*, **23**, 270–281.
62. Cao,S., Zhou,D.C., Oh,C., Jayasinghe,R.G., Zhao,Y., Yoon,C.J., Wyczalkowski,M.A., Bailey,M.H., Tsou,T., Gao,Q. *et al.* (2020) Discovery of driver non-coding splice-site-creating mutations in cancer. *Nat. Commun.*, **11**, 5573.
63. Bon,E., Casaregola,S., Blandin,G., Llorente,B., Neuvéglise,C., Munsterkotter,M., Guldener,U., Mewes,H.W., Van Helden,J., Dujon,B. *et al.* (2003) Molecular evolution of eukaryotic genomes: hemiascomycetous yeast spliceosomal introns. *Nucleic Acids Res.*, **31**, 1121–1135.
64. Montes,M., Sanford,B.L., Comiskey,D.F. and Chandler,D.S. (2019) RNA splicing and disease: animal models to therapies. *Trends Genet.*, **35**, 68–87.