



Programa de Doctorado en Bioingeniería
Universidad Miguel Hernández de Elche

**Análisis de la contribución de los genes
PRP8, *RPS24A* y *RPS24B* al metabolismo
del ARN en *Arabidopsis thaliana***

Adrián Cabezas Fuster

Directora de la tesis
María Rosa Ponce Molet

Elche, 2023

**Análisis de la contribución de los genes
PRP8, *RPS24A* y *RPS24B* al metabolismo
del ARN en *Arabidopsis thaliana***

Trabajo realizado por el Graduado Adrián Cabezas Fuster, en la Unidad de Genética del Instituto de Bioingeniería de la Universidad Miguel Hernández de Elche, para optar al grado de Doctor.

Elche, 1 de junio de 2023

La presente Tesis Doctoral, titulada “Análisis de la contribución de los genes *PRP8*, *RPS24A* y *RPS24B* al metabolismo del ARN en *Arabidopsis thaliana*”, se presenta bajo la modalidad de **tesis por compendio** de la siguiente **publicación**:

Cabezas-Fuster, A., Micol-Ponce, R., Fontcuberta-Cervera, S., y Ponce, M.R. (2022). Missplicing suppressor alleles of *Arabidopsis PRE-MRNA PROCESSING FACTOR 8* increase splicing fidelity by reducing the use of novel splice sites. *Nucleic Acids Research* **50**, 5513-5527.

MARÍA ROSA PONCE MOLET, Catedrática de Genética de la Universidad Miguel Hernández de Elche (UMH)

HAGO CONSTAR:

Que el presente trabajo ha sido realizado bajo mi dirección y recoge fielmente la labor desarrollada por el graduado Adrián Cabezas Fuster para optar al grado de Doctor. Las investigaciones reflejadas en esta memoria se han desarrollado íntegramente en la Unidad de Genética del Instituto de Bioingeniería de la UMH, según los términos y condiciones definidos en el Plan de Investigación del doctorando, y cumpliendo los objetivos inicialmente previstos de forma satisfactoria y lo establecido en el Código de Buenas Prácticas de la UMH.

María Rosa Ponce Molet

Elche, 1 de junio de 2023

PIEDAD NIEVES DE AZA MOYA, Coordinadora del Programa de Doctorado en Bioingeniería de la Universidad Miguel Hernández de Elche por Resolución Rectoral 0169/17, de 1 de febrero de 2017

HACE CONSTAR:

Que da su conformidad a la presentación de la Tesis Doctoral de Don Adrián Cabezas Fuster, titulada “Análisis de la contribución de los genes *PRP8*, *RPS24A* y *RPS24B* al metabolismo del ARN en *Arabidopsis thaliana*”, que se ha desarrollado en el Programa de Doctorado en Bioingeniería bajo la dirección de la profesora María Rosa Ponce Molet.

Lo que firmo en Elche, a instancias del interesado y a los efectos oportunos, a uno de junio de dos mil veintitrés.

Profesora PIEDAD NIEVES DE AZA MOYA
Coordinadora del Programa de Doctorado en Bioingeniería

A tot el que estime. A tot el que m'estima.
Als meus pares, a la meua germana i a la meua àvia.
A la Lluna.

ÍNDICE DE MATERIAS

ÍNDICE DE FIGURAS	II
ÍNDICE DE TABLAS	II
I.- PREFACIO	1
II.- RESUMEN	2
III.- SUMMARY	5
IV.- INTRODUCCIÓN	8
IV.1.- Diversidad del transcriptoma eucariótico	8
IV.2.- Metabolismo de los ARN eucarióticos	8
IV.2.1.- Maduración de los ARNm eucarióticos	8
IV.2.2.- El <i>splicing</i> de los pre-ARNm	9
IV.2.2.1.- Tipos de <i>splicing</i>	9
IV.2.2.2.- Estructura y función del espliceosoma	9
IV.2.2.3.- Actividad de PRP8 en el <i>splicing</i>	13
IV.2.2.4.- El <i>splicing</i> alternativo	15
IV.2.2.5.- Alteraciones del <i>splicing</i>	17
IV.2.3.- Mecanismos de control de calidad del ARNm	17
IV.3.- El ribosoma eucariótico	18
IV.3.1.- Función, estructura y composición del ribosoma	18
IV.3.2.- Biogénesis del ribosoma	19
IV.3.2.1.- Estructura, organización y expresión de los ADNr	20
IV.3.2.2.- Procesamiento de los pre-ARNr	21
IV.3.2.3.- Las proteínas ribosómicas	24
IV.3.2.3.1.- Genes que codifican proteínas ribosómicas	25
IV.3.2.3.2.- Funciones extrarribosómicas de las proteínas ribosómicas	25
IV.3.2.4.- Fenotipos causados por defectos en la biogénesis del ribosoma o en la traducción	27
IV.4.- Antecedentes y objetivos	28
IV.4.1.- Identificación de supresores del fenotipo morfológico de <i>ago1-52</i>	28
IV.4.2.- Identificación de interactores de MAS2	29
IV.4.3.- Objetivos de esta Tesis	30
V.- MATERIALES Y MÉTODOS	32
VI.- RESULTADOS Y DISCUSIÓN	34
VII.- CONCLUSIONES Y PERSPECTIVAS	37
VII.1.- Los alelos <i>mas5</i> del gen <i>PRP8</i> incrementan la fidelidad del <i>splicing</i>	37
VII.2.- RPS24A y RPS24B son factores de la biogénesis del ribosoma	38

VIII.- BIBLIOGRAFÍA DE LOS APARTADOS IV-VII.....	39
IX.- PUBLICACIONES.....	48
X.- AGRADECIMIENTOS.....	135

ÍNDICE DE FIGURAS

Figura 1.- Estructura de un intrón de tipo U2, con indicación de las secuencias implicadas en el <i>splicing</i>	10
Figura 2.- Diagrama del ensamblaje y desensamblaje de los complejos del espliceosoma de tipo U2 durante el <i>splicing</i>	12
Figura 3.- Detalles de las reacciones de transesterificación que ocurren durante el <i>splicing</i> de un intrón de tipo U2	13
Figura 4.- Tipos de eventos de <i>splicing</i> alternativo	15
Figura 5.- Estructura y localización de los genes del ADNr 45S en Arabidopsis.....	20
Figura 6.- Procesamiento del pre-ARNr 45S en Arabidopsis	22

I.- PREFACIO

I.- PREFACIO

Siguiendo la normativa de la Universidad Miguel Hernández de Elche para la “Presentación de Tesis Doctorales por compendio de publicaciones”, este documento se ha dividido en las partes siguientes:

I.- Este *Prefacio*.

II.- Un *Resumen* en español.

III.- Un *Summary* en inglés.

IV.- Una *Introducción*, en la que se presenta el tema de la Tesis y los antecedentes y objetivos del trabajo realizado.

V.- Un resumen de los *Materiales y métodos* de las publicaciones de la Tesis.

VI.- Un resumen de los *Resultados y discusión* de las publicaciones de la Tesis.

VII.- Un resumen de las *Conclusiones y perspectivas* del trabajo realizado.

VIII.- Una *Bibliografía de los apartados IV-VII*; algunas de las referencias que incluye se repiten en las bibliografías de los artículos incluidos en esta memoria.

IX.- Un apartado de *Publicaciones*, que incluye las dos siguientes, en una de las cuales se indica el último factor de impacto [FI] publicado.

Cabezas-Fuster, A., Micol-Ponce, R., Fontcuberta-Cervera, S., y Ponce, M.R. (2022). Missplicing suppressor alleles of Arabidopsis *PRE-MRNA PROCESSING FACTOR 8* increase splicing fidelity by reducing the use of novel splice sites. *Nucleic Acids Research* **10**, 5513-5527 [FI: 19,330; D1].

Cabezas-Fuster, A., Micol-Ponce, R., Sarmiento-Mañus, R., y Ponce, M.R. Cross-kingdom conservation of Arabidopsis RPS24 function in 18S rRNA maturation. Depositado en *bioRxiv* doi: 10.1101/2023.04.21.537868.

Una parte de los datos suplementarios de estos dos artículos no se ha incluido en esta memoria, por tratarse de tablas de gran longitud, que se han remitido a los miembros del tribunal en formato electrónico, en un archivo comprimido.

X.- Un apartado de *Agradecimientos*.

II.- RESUMEN

II.- RESUMEN

Casi todos los ARN eucarióticos maduran, sufriendo diferentes tipos de cortes exo y endonucleolíticos, la eliminación de algunos de sus segmentos y la sustitución y/o modificación química de determinados ribonucleótidos. Esto es así tanto para los que codifican proteínas (ARN mensajeros [ARNm]) como para los que no las codifican (ARN ribosómicos [ARNr] y otros). Se denomina metabolismo del ARN al conjunto de estos procesos, que se inicia con la síntesis de un transcrito primario y termina con la degradación del ARN maduro, usualmente posterior al cumplimiento de su función.

Se denomina pre-ARNm al producto primario de la transcripción de los genes eucarióticos que codifican proteínas, que sufre un complejo proceso de maduración antes de convertirse en un ARNm apto para su exportación desde el núcleo al citoplasma y su traducción por el ribosoma. El *splicing* de los pre-ARNm es uno de sus mecanismos de maduración mejor conocidos, en el que suceden dos transesterificaciones sucesivas que conllevan la eliminación de un intrón y la ligación de los exones adyacentes. El *splicing* es realizado por el espliceosoma, una maquinaria ribonucleoproteica compleja, dinámica y muy precisa, que reconoce secuencias conservadas en los extremos de los intrones y exones de los pre-ARNm: los sitios donante (5' splicing site; 5'SS) y aceptor (3'SS) del *splicing*. La eliminación de los intrones durante el *splicing* no siempre es idéntica para todas las moléculas de un mismo pre-ARNm, obteniéndose así diferentes variantes de ARNm maduro, que suelen ser específicas de tejido. Este fenómeno se denomina *splicing* alternativo y es una de las fuentes principales de la diversificación del proteoma de las plantas, y en particular, de los animales.

La proteína ARGONAUTE1 (AGO1) de *Arabidopsis thaliana* (en adelante, *Arabidopsis*) es el componente más importante de los RNA-induced silencing complexes (RISC), que juegan un papel central en la regulación postranscripcional de la expresión génica mediada por los microARN (miARN). El alelo *ago1-52* del gen *AGO1* es hipomorfo, viable y portador de una mutación puntual en su vigésimo intrón, que genera un 3'SS nuevo, que el espliceosoma usa más frecuentemente que el genuino, que está inalterado. Como resultado, el alelo *ago1-52* produce dos ARNm, uno silvestre y muy minoritario, y otro mutante; este último incluye 10 nucleótidos del vigésimo intrón, que desfasan su pauta de lectura. La traducción de este ARNm aberrante rinde una proteína AGO1-52 mutante, que además es mucho más abundante que la proteína AGO1 silvestre en el mutante *ago1-52*.

Los componentes del espliceosoma están muy conservados en todos los eucariotas. Uno de ellos es PRE-MRNA PROCESSING FACTOR 8 (PRP8), una de las proteínas centrales del espliceosoma, que reconoce los 5'SS y 3'SS de los intrones de los pre-ARNm.

En todas las especies en las que se han estudiado, los alelos nulos de *PRP8* son letales, y los hipomorfos causan alteraciones globales del *splicing*, que a su vez perturban el desarrollo. Hemos caracterizado en esta Tesis seis nuevos alelos mutantes del gen *PRP8* de Arabidopsis, que fueron aislados en una búsqueda de supresores extragénicos del fenotipo morfológico de *ago1-52*. Hemos llamado a estos alelos *morphology of argonaute1-52 suppressed 5-1 (mas5-1)* a *mas5-6*. Cuatro de los alelos *mas5* del gen *PRP8* de Arabidopsis causan sustituciones de uno de los aminoácidos de una región de PRP8 que forma una cavidad próxima al sitio activo de esta proteína.

Hemos establecido que la causa de la supresión del fenotipo morfológico del mutante *ago1-52* en los dobles mutantes *mas5 ago1-52* es la restauración parcial del uso por el espliceosoma del 3'SS genuino del vigésimo intrón del gen *AGO1*, que a su vez conlleva una mayor producción del ARNm de *AGO1* y la proteína AGO1 silvestres. Del análisis de las bases moleculares de la supresión del fenotipo de *ago1-52* y otros mutantes también afectados en el *splicing* de genes concretos hemos concluido que nuestros alelos *mas5* incrementan la fidelidad del *splicing*, favoreciendo el uso de los 5'SS y 3'SS genuinos tras la aparición por mutación de otros nuevos. En consecuencia, en los mutantes *mas5* no se altera globalmente el *splicing*, por lo que crecen como el tipo silvestre o aún mejor. Dado que actualmente es fácil inducir mutaciones mediante CRISPR/Cas, la obtención en líneas celulares de alelos del gen *PRP8* humano equivalentes a los *mas5* de Arabidopsis podría ser de gran utilidad para el estudio y la eventual terapia de enfermedades causadas por defectos en el *splicing* de genes concretos, ya que podrían suprimir sus efectos deletéreos sin perturbar globalmente el *splicing*.

El ribosoma citoplásmico 80S (en adelante, el ribosoma) es la maquinaria ribonucleoproteica que traduce a proteínas los ARNm de los genes nucleares. Está constituido por dos subunidades, que contienen cuatro ARNr y decenas de proteínas ribosómicas. La biogénesis del ribosoma es muy compleja y en ella participan tres ARN polimerasas y cientos de proteínas, conocidas colectivamente como factores de la biogénesis del ribosoma, que regulan de forma coordinada la transcripción de los genes del ADN ribosómico (ADNr), la maduración de los ARNr y el ensamblaje del ribosoma. En esta Tesis hemos caracterizado la función de RIBOSOMAL PROTEIN S24A (RPS24A) y RPS24B, una de las cuales está siempre presente, como componente estructural, en la subunidad menor del ribosoma de Arabidopsis. Los genes parálogos *RPS24A* y *RPS24B* son casi idénticos. El gen *RPS24* humano y el *Rps24* de la levadura *Saccharomyces cerevisiae* son de copia única y codifican una proteína con un papel dual, ya que no solo es un componente estructural del ribosoma sino también un factor de su biogénesis, que actúa en la maduración del ARNr 18S.

Hemos caracterizado funcionalmente los genes parálogos *RPS24A* y *RPS24B* de *Arabidopsis* mediante abordajes genéticos y moleculares. Hemos obtenido alelos mutantes de ambos genes, probablemente nullos, y comprobado que su fenotipo morfológico es similar al de los alelos mutantes de genes implicados en la maquinaria de la traducción. Hemos concluido, tras intentar obtener dobles mutantes *rps24a rps24b*, que *Arabidopsis* necesita al menos dos copias silvestres de alguno de sus dos genes *RPS24* para su viabilidad, y tres para su desarrollo normal. Estos resultados indican que *RPS24A* y *RPS24B* presentan haploinsuficiencia combinada; *RPS24*, su ortólogo humano, es de copia única y haploinsuficiente.

Nuestro estudio de los mutantes *rps24a* y *rps24b* ha revelado defectos en la maduración de sus ARNr: en estas estirpes se acumulan precursores del ARNr 18S y está incrementada la transcripción del ADNr 45S, que codifica los tres ARNr mayores (25S, 18S y 5,8S). Estos resultados sugieren funciones extrarribosómicas de *RPS24A* y *RPS24B*, en el procesamiento del pre-ARNr 45S y en la represión de la transcripción del ADNr 45S. Hemos observado fenotipos morfológicos y moleculares sinérgicos en las combinaciones dobles de *rps24b* con alelos mutantes de genes que codifican otros factores de la biogénesis del ribosoma. Hemos demostrado que *RPS24A* y *RPS24B* no solo juegan un papel estructural en el ribosoma, sino que actúan también como factores de la biogénesis del ribosoma, tal como hacen sus ortólogos humano y de la levadura.

III.- SUMMARY

III.- SUMMARY

Most eukaryotic RNAs, whether protein coding, such as messenger RNAs (mRNAs), or non-coding, such as ribosomal RNAs (rRNAs) and others, suffer maturation processes that involve exo- and endo-nucleolytic cleavages, removal of segments of mature RNA precursors (pre-RNAs), and ribonucleotide substitutions and chemical modifications. These processes are collectively referred to as RNA metabolism, which begins with the synthesis of a primary transcript and ends with the degradation of mature RNAs, usually after the execution of their functions.

Eukaryotic pre-mRNA, the primary product of the transcription of eukaryotic protein-coding genes, suffers a complex maturation to become a mature mRNA suitable for export from the nucleus to the cytoplasm and translation by the ribosome. Splicing of pre-mRNAs is a well-known mechanism of maturation in which two successive transesterifications remove an intron and ligate the adjacent exons. Splicing is catalysed by the spliceosome, a dynamic, complex, and precise ribonucleoprotein complex, which recognises conserved sequences at the ends of pre-mRNA introns and exons: the donor 5' splicing site (5'SS) and acceptor 3' splicing site (3'SS). Intron removal during splicing is not always identical for all the molecules of a given pre-mRNA, which generates different mature mRNA variants from a single gene, which are usually tissue specific. This phenomenon is known as alternative splicing and is one of the major sources of proteome diversity in plants and, in particular, in animals.

In *Arabidopsis thaliana* (hereafter, *Arabidopsis*), the ARGONAUTE1 (AGO1) protein is the most important component of the RNA-induced silencing complexes (RISC) that play a key role in the postranscriptional regulation of gene expression mediated by microRNAs (miRNAs). The *ago1-52* allele of the *AGO1* gene is hypomorphic, viable and carries a point mutation in its twentieth intron that generates a novel 3'SS, which is more frequently used by the spliceosome than the genuine one, which remains otherwise intact. Therefore, the *ago1-52* allele produces two mRNAs, one of which is wild-type and minority, while the other includes 10 nucleotides from its twentieth intron, which in turn causes a reading frame shift. Translation of this aberrant mRNA yields a mutant AGO1-52 protein, which is much more abundant than the wild-type AGO1 protein in the *ago1-52* mutant.

Spliceosome components are highly conserved in all eukaryotes. PRE-MRNA PROCESSING FACTOR 8 (PRP8) is a core component of the spliceosome, which recognizes the 5'SSs and 3'SSs. In all species where the *PRP8* gene has been studied, its null alleles are lethal, and its hypomorphic alleles cause global splicing alterations that disrupt development. In this Thesis, we characterized six new mutant alleles of *Arabidopsis PRP8*, which were isolated in a search for extragenic suppressors of the morphological phenotype of *ago1-52*.

We named these six mutant alleles *morphology of argonaute 1-52 suppressed 5-1 (mas5-1)* to *mas5-6*. Four out of the six *mas5* alleles of Arabidopsis *PRP8* cause single amino acid substitutions within a *PRP8* region that forms a cavity near to its active site.

We found that the suppression of the mutant morphological phenotype of *ago1-52* in the *mas5 ago1-52* double mutants is caused by the partial restoration of the use by the spliceosome of the genuine 3'SS of the twentieth intron of the *AGO1* gene, which in turn increases the levels of the wild-type *AGO1* mRNA and *AGO1* protein. From our analysis of the molecular basis of the suppression of the phenotypes of *ago1-52* and other mutants with alterations in the splicing of specific genes, we concluded that our *mas5* alleles increase splicing fidelity, favoring the use of the genuine 5'SSs and 3'SSs instead of the novel ones generated by point mutations. Therefore, in the *mas5* mutants splicing is not globally altered, and these mutants grow normally and even better than their wild types. Given that it is currently easy to induce mutations by CRISPR/Cas, obtaining human cell lines carrying alleles of the *PRP8* gene equivalent to the *mas5* alleles of Arabidopsis may be useful for the study and subsequent therapy of diseases caused by aberrations in the splicing of specific genes. Such alleles of human *PRP8* would be expected to suppress defects in the splicing of specific genes without globally perturbing splicing.

The 80S or cytoplasmic ribosome (hereafter, the ribosome) is a ribonucleoprotein machinery that translates into proteins the mRNAs from nuclear genes. It is composed of two ribosomal subunits that contain some rRNAs and tens of ribosomal proteins. Ribosome biogenesis is a complex process that involves three RNA polymerases and hundreds of proteins, collectively known as ribosome biogenesis factors, which regulate the transcription of rDNA genes, rRNA maturation, and ribosome assembly. In this Thesis, we characterized the function of the Arabidopsis RIBOSOMAL PROTEIN S24A (*RPS24A*) and *RPS24B* proteins, one of which is always present at the small subunit of the ribosome; the paralog *RPS24A* and *RPS24B* genes are almost identical. Human *RPS24* and *Saccharomyces cerevisiae Rps24* are single-copy genes and each encode a protein that plays a dual role as a structural component of the ribosome and as a ribosome biogenesis factor acting in 18S rRNA maturation.

We characterized the Arabidopsis *RPS24A* y *RPS24B* paralogs using classical and molecular genetic approaches. We obtained mutant alleles, likely null, for both genes, and found that their morphological phenotype is similar to that of other mutant alleles of genes involved in the translation machinery. After trying to obtain *rps24a rps24b* double mutants, we concluded that Arabidopsis needs at least two wild-type copies of either of these two genes to be viable, and three for its normal development. These results indicate that Arabidopsis

RPS24A and *RPS24B* show combined haploinsufficiency. It is of note that the single-copy human *RPS24* gene is haploinsufficient.

We found alterations in rRNA maturation in the *rps24a* and *rps24b* mutants, which accumulate 18S rRNA precursors, and that the transcription of the 45S rDNA, which encodes three rRNAs (25S, 18S and 5.8S), is increased. These results suggest that RPS24B has an extraribosomal function in 45S pre-rRNA processing and in the repression of 45S rDNA expression. We also found synergistic molecular and morphological phenotypes in the double mutant combinations of *rps24b* with alleles of genes encoding ribosome biogenesis factors. These results demonstrate that RPS24A and RPS24B not only are structural components of the ribosome, but also act as ribosome biogenesis factors, as their yeast and human orthologs do.

IV.- INTRODUCCIÓN

IV.- INTRODUCCIÓN

IV.1.- Diversidad del transcriptoma eucariótico

El transcriptoma de los procariotas y los eucariotas es heterogéneo, ya que está formado por miles de moléculas de ARN distintas, que difieren en sus secuencias, mecanismos de maduración, funciones y número de cada una de ellas en las células. Los ARN eucarióticos pueden clasificarse en dos grandes grupos: los mensajeros (ARNm), que codifican proteínas, están poliadenilados en su extremo 3' y son traducidos por los ribosomas, y los que no las codifican, no están poliadenilados ni se traducen, y son funcionales *per se* (Li y Liu, 2019).

Aunque las células eucarióticas contienen ribosomas citoplásmicos y organulares (mitocondriales, y en las plantas, también cloroplásticos), en esta Tesis nos referiremos exclusivamente a los primeros. En la traducción de los ARNm por el ribosoma intervienen dos tipos de ARN no codificantes: los ARN transferentes (ARNt), que descodifican la secuencia de los ARNm y aportan aminoácidos a las proteínas nacientes, y los ARN ribosómicos (ARNr), que son componentes estructurales del ribosoma [revisado en Perteza (2012)]. La ARN Polimerasa I (ARN Pol I) sintetiza el precursor de los ARNr de mayor tamaño (los 25S, 18S y 5,8S), y la ARN Pol III, los de los ARNt y del ARNr 5S. En la biogénesis del ribosoma también participan los ARN pequeños nucleolares (snoRNA), cuyos precursores son sintetizados por la ARN Pol II, que también cataliza la síntesis de los ARNm de todas las proteínas que intervienen en este proceso, a las que se denomina factores de la biogénesis del ribosoma [apartado IV.3.2, en la página 19; revisado en Barba-Aliaga *et al.* (2021)].

Existen otros tipos de ARN no codificantes, con función reguladora, que se clasifican en función de su longitud y biogénesis. Los microARN (miARN), cuyos precursores son sintetizados por la ARN Pol II o la ARN Pol III, según la especie, son los que mejor se han estudiado, por su gran importancia en la regulación postranscripcional de la expresión de muchos genes que codifican proteínas [revisado en O'Brien *et al.* (2018)]. En todos los eucariotas, las ribonucleasas de la familia ARGONAUTE (AGO) son las efectoras del silenciamiento génico postranscripcional que ejercen los miARN sobre sus ARNm diana, cuya traducción impiden. En *Arabidopsis thaliana* (en adelante, *Arabidopsis*), AGO1 juega un papel central en las rutas de regulación mediadas por los miARN y los alelos nulos del gen *AGO1* son letales [revisado en Zhang *et al.* (2015a); Li *et al.* (2022)].

IV.2.- Metabolismo de los ARN eucarióticos

IV.2.1.- Maduración de los ARNm eucarióticos

En los procariotas, la transcripción de los genes que codifican proteínas genera

moléculas de ARNm que requieren muy pocas o ninguna modificación, cuya traducción ocurre casi simultáneamente a su síntesis. Sin embargo, en los eucariotas la transcripción genera un transcrito primario (pre-ARNm) que es procesado cotranscripcionalmente en el núcleo, mediante la adición de una caperuza (cap) de guaninas a su extremo 5', de una cola de adenosinas (poli-A) al 3' y la eliminación de los intrones, en un proceso de corte y ligación denominado *splicing* [apartado IV.2.2; revisado en Bentley (2014)]. La caperuza del extremo 5' y la cola poli-A del 3' protegen a los ARNm de su degradación por exonucleasas, desde la transcripción hasta la traducción. La caperuza 5' facilita además la traducción del ARNm, al ser reconocida por factores de iniciación de la traducción [revisado en Ramanathan *et al.* (2016)]. La cola poli-A facilita la exportación del ARNm del núcleo al citoplasma, ya que a ella se unen proteínas de los complejos de exportación, y una vez en el citoplasma, otras que regulan el inicio de su traducción [revisado en Dreyfus y Régnier (2002); Rodríguez-Molina y Turtola (2022)].

IV.2.2.- El *splicing* de los pre-ARNm

IV.2.2.1.- Tipos de *splicing*

El *splicing* genera un ARNm en el que la región codificante no está interrumpida, que contiene una pauta de lectura abierta que puede ser traducida por el ribosoma. La existencia de más de un intrón en muchos genes eucarióticos posibilita su *splicing* alternativo, que enriquece el proteoma de los animales y las plantas [revisado en Wilkinson *et al.* (2020)]. La eliminación de los intrones y la unión subsiguiente de los exones adyacentes se ejecuta en el *splicing* mediante dos reacciones de transesterificación sucesivas, catalizadas por el spliceosoma. Este complejo ribonucleoproteico reconoce secuencias específicas situadas en los límites de los intrones y los exones de los pre-ARNm: los sitios donante o 5' (5' splice site; 5'SS) y aceptor o 3' (3'SS) del *splicing*, que contienen dinucleótidos muy conservados, en los extremos 5' y 3' de los intrones, respectivamente [revisado en Wilkinson *et al.* (2020)].

Existen dos tipos de intrones, que difieren en las secuencias de sus 5'SS y 3'SS, sobre los que actúan dos espliceosomas diferentes: los de tipo U2, que son los mayoritarios en todos los genomas eucarióticos y son procesados por el espliceosoma mayor, y los de tipo U12, minoritarios (unos 300 en *Arabidopsis*), por el menor [revisado en Turunen *et al.* (2013); Ding *et al.* (2022)]. En los siguientes apartados de esta Tesis se hará referencia únicamente al *splicing* de los de tipo U2 y, por tanto, al espliceosoma mayor, que se nombrará sin adjetivar.

IV.2.2.2.- Estructura y función del espliceosoma

El espliceosoma es el conjunto de proteínas y ribonucleoproteínas que intervienen de

forma directa o indirecta en el *splicing* de los pre-ARNm. Su composición es dinámica, ya que sus componentes se ensamblan y desensamblan formando nueve diferentes complejos a lo largo del *splicing*. Se conocen unas 170 proteínas humanas que participan en el *splicing*, como parte de alguno de los complejos del espliceosoma o como factores que intervienen en alguna etapa del proceso [revisado en Wahl *et al.* (2009); Turunen *et al.* (2013)]. Se han identificado en *Arabidopsis* 430 proteínas presuntamente homólogas de las que forman los espliceosomas de la levadura (unas 50-60) y humano (Koncz *et al.*, 2012).

Las U1, U2, U4, U5 y U6 snRNP (small nuclear ribonucleoproteins) contienen un snRNA (small nuclear RNA), que es rico en uridinas y da nombre a la partícula. Su núcleo central común incluye siete proteínas de los tipos Smith proteins (Sm) y Sm-like (Lsm), que forman una estructura anular a la que se unen los U snRNA [revisado en Will y Lührmann (2001)]. Al núcleo de las U snRNP (en adelante, partículas U) se suma un número variable de hasta 50 proteínas diferentes y específicas de cada partícula [revisado en Will y Lührmann (2011)]. Además de las proteínas que forman parte de las partículas U, otras 100 actúan como factores implicados en el *splicing* o en otros procesos relacionados con el metabolismo del ARNm, como la finalización de la transcripción, la exportación del ARNm del núcleo al citoplasma y el control de la calidad del ARNm (apartado IV.2.3, en la página 16; Fabrizio *et al.*, 2009).

Los 5'SS y 3'SS contienen invariablemente los dinucleótidos GU y AG en los extremos 5' y 3' de cada intrón, respectivamente (Figura 1). Los nucleótidos inmediatamente posteriores o anteriores a estos dinucleótidos también son parte de los 5'SS y 3'SS, pero su grado de conservación es menor; de hecho, solo son esenciales los dinucleótidos GU y AG para un *splicing* correcto (Brown *et al.*, 1996). En los intrones existen otras dos regiones conservadas y necesarias para la liberación del extremo 3' del exón 1 (según la nomenclatura de la Figura

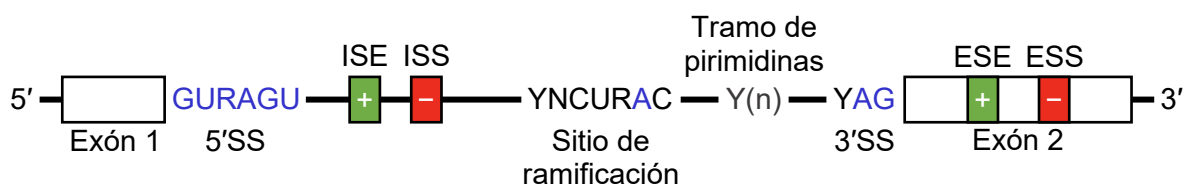


Figura 1.- Estructura de un intrón de tipo U2, con indicación de las secuencias implicadas en el *splicing*. Las líneas negras representan a los intrones, y los rectángulos blancos, a los exones. Se indican las secuencias de los 5'SS y 3'SS, el sitio de ramificación y el tramo de pirimidinas [Y(n)]. Se destacan con letras azules los nucleótidos conservados e invariantes. Y, R y N representan pirimidina, purina y cualquier base nucleotídica, respectivamente. Los rectángulos verdes con un símbolo + representan las señales potenciadoras del *splicing*, y los rojos con un símbolo -, las represoras. Modificada a partir de Will y Lührmann (2011) y Syed *et al.* (2012).

1): el sitio de ramificación (Branch Point; [BP]) con una A invariante, 15-50 nucleótidos aguas arriba del 3'SS y situado junto a un segmento de unas 10 pirimidinas, denominado tramo de pirimidinas [Polypyrimidine Tract (PT); Tolstrup *et al.* (1997)].

Otras secuencias intrónicas y exónicas actúan como elementos reguladores del *splicing*, ya que a ellas se unen factores de unión a ARN para estimularlo o dificultarlo: los Exonic Splicing Enhancer (ESE) e ISE (Intronic Splicing Enhancer) y Exonic Splicing Silencer (ESS) e Intronic Splicing Silencer (ISS), respectivamente [revisado en Smith y Valcárcel (2000); Figura 1, en la página 10].

Los complejos que se ensamblan y desensamblan sucesivamente durante el *splicing* difieren en su composición y/o conformación tridimensional. El complejo E es el primero en formarse, al reconocer la partícula U1 al 5'SS del intrón, al cual se une. El complejo E se convierte en complejo A (preespliceosoma) al unirse la partícula U2 al sitio de ramificación (Figura 2, en la página 12). El complejo B (precatalítico) es el primero al que se considera un espliceosoma completo; se forma tras la unión del complejo A a la partícula trimérica U4/U6.U5, ensamblada previamente, que formará parte del núcleo catalítico del espliceosoma. Tras su reclutamiento, las helicasas de ARN PRE-MRNA PROCESSING FACTOR 28 (PRP28) y BAD RESPONSE TO REFRIGERATION 2 (BRR2) retiran del complejo a las partículas U1 y U4, respectivamente, posibilitando la interacción de la partícula U6 con el 5'SS, el sitio de ramificación y la partícula U2, formándose así el complejo B^{act} (B activado). El primer complejo en el que el núcleo catalítico del espliceosoma está completamente formado es el B^{act}. En este complejo, el 5'SS del pre-ARNm se encuentra en el núcleo catalítico. La ATPasa PRP2 transloca el sitio de ramificación al núcleo catalítico, remodelando así la estructura de B^{act}, que se convierte en complejo B* (catalíticamente activo), que cataliza la primera de las dos reacciones de transesterificación del *splicing* (Figura 2, en la página 12).

En la primera reacción de transesterificación, el grupo hidroxilo 2' de la A del sitio de ramificación ataca al grupo fosfato que une al exón 1 (según la nomenclatura de la Figura 1) con la G del 5'SS (Figura 3A, en la página 13). Se obtiene así un intermediario formado por un exón 1 con su extremo 3' libre y un lazo intrónico unido al extremo 5' del exón 2 (Figura 3B). Se denomina complejo C a esta estructura, cuya composición es la misma que la del complejo B*, salvo por el lazo intrónico (Figura 2). Para la segunda reacción de transesterificación, la ATPasa PRP16 remodela la estructura del complejo B* para permitir su acceso al 3'SS, formándose así la denominada conformación de ligación de exones o complejo C*. En esta reacción, el grupo hidroxilo 3' libre del exón 1 ataca al grupo fosfato del extremo 5' del exón 2, propiciando la ligación de ambos exones (Figura 3B y C). La estructura

que incluye el lazo intrónico y los exones ligados se denomina complejo P (postesplíceosómico). A continuación, PRP22 libera al ARNm maduro y los factores de ligación unidos en los exones, a la vez que el intrón es degradado y se reciclan los ribonucleótidos resultantes y las partículas U del espliceosoma [Figura 2; revisado en Will y Lührmann (2011); Shi (2017); Wilkinson *et al.* (2020)].

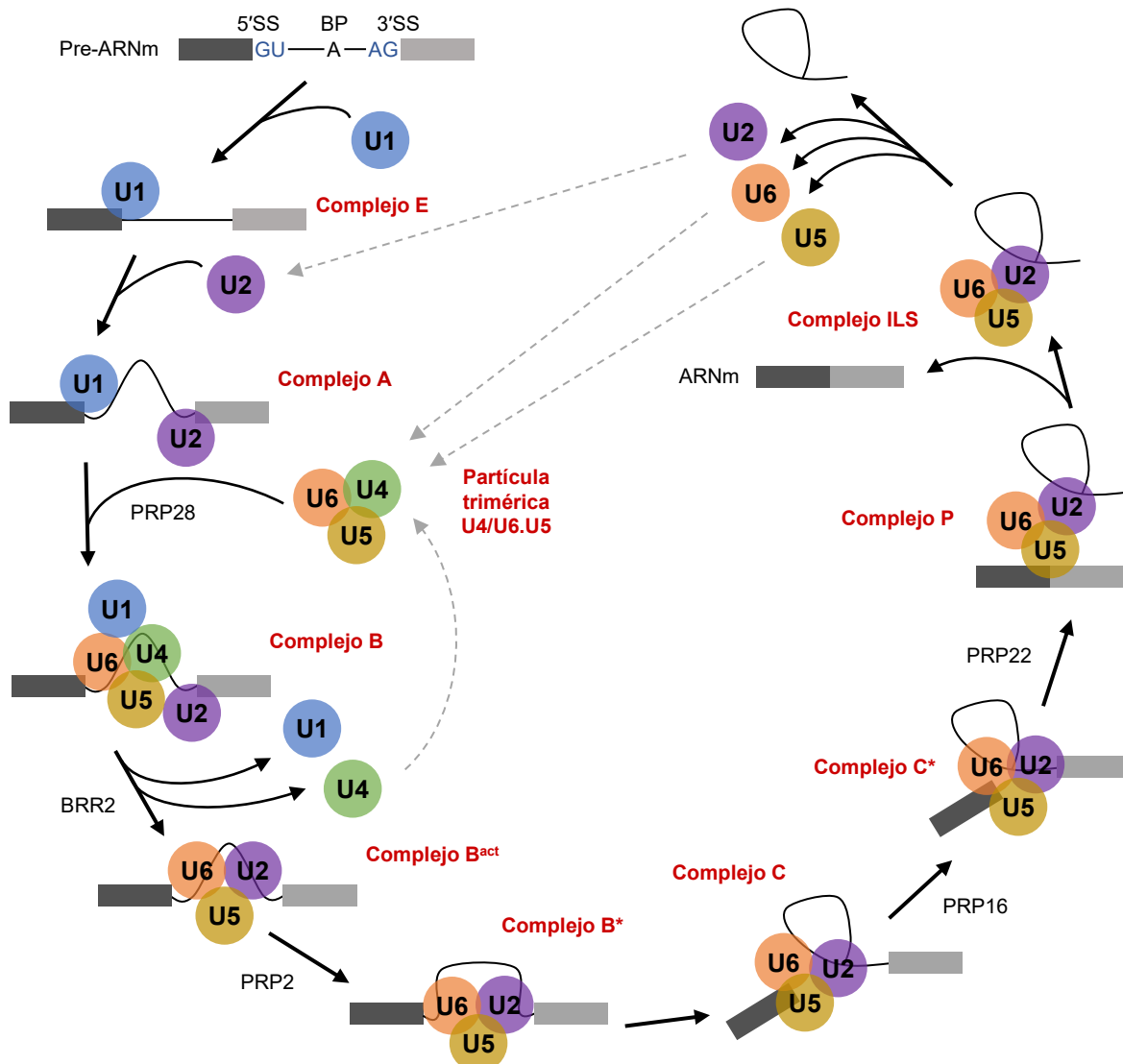


Figura 2.- Diagrama del ensamblaje y desensamblaje de los complejos del espliceosoma de tipo U2 durante el *splicing*. Los exones e intrones están representados por rectángulos y líneas, respectivamente. BP: sitio de ramificación. Las partículas U se representan con círculos. Las flechas negras continuas indican las etapas de ensamblaje y desensamblaje de los distintos complejos, y las grises discontinuas el reciclaje de las partículas U. Se indican con letras rojas y negras los nombres de los complejos y las principales enzimas implicadas en cada etapa del proceso, respectivamente. Modificado a partir de Will y Lührmann (2011).

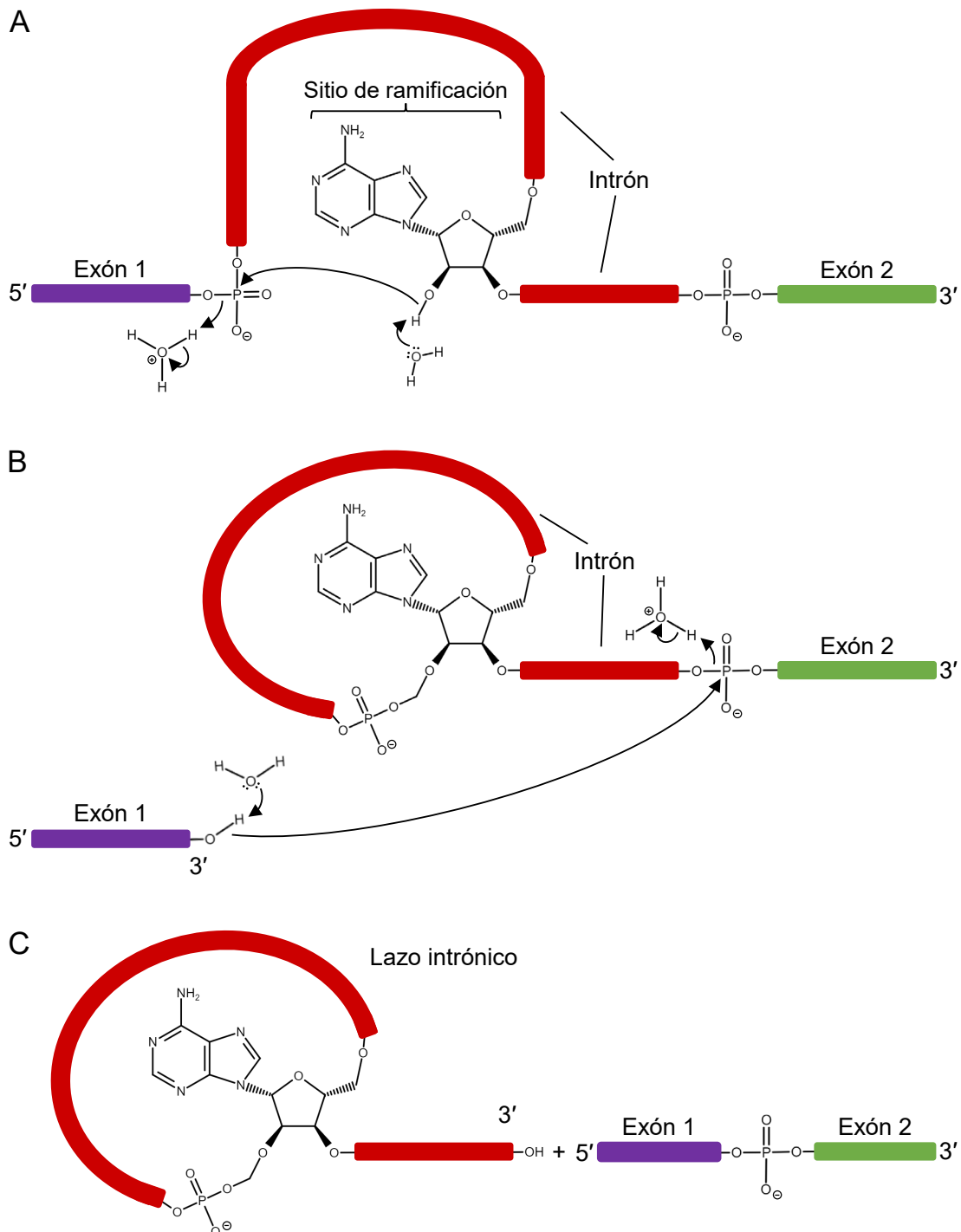


Figura 3.- Detalles de las reacciones de transesterificación que ocurren durante el *splicing* de un intrón de tipo U2. Modificado a partir de https://content.labxchange.org/learning-item-assets/LS1A_Ch+10_Transcription/LS1A_Ch+10_Transcription_Fig+17.png

IV.2.2.3.- Actividad de PRP8 en el *splicing*

PRP8 (PRPF8 en la especie humana, por Pre-mRNA processing factor 8) es una proteína grande, de 230-280 kDa, y muy conservada, que forma parte del núcleo catalítico del espliceosoma, participando en sus complejos precatalíticos (B y B^{act}), catalíticos (B*, C y C*)

y postcatalíticos (P e ILS) [revisado en Grainger y Beggs (2005)]. PRP8 ha sido muy estudiada en la levadura *Saccharomyces cerevisiae*, que paradójicamente es una especie con muy pocos genes con intrones (Spingola *et al.*, 1999). PRP8 presenta cuatro dominios altamente conservados en sus regiones central y carboxiterminal, denominados Reverse transcriptase-like, Endonuclease-like, RNaseH-like y Jab1/MPN, separados por regiones con estructuras desordenadas y flexibles, conocidas como Linkers. Dichos dominios forman una cavidad, cargada positivamente y próxima al núcleo catalítico del espliceosoma, que interacciona directamente con el 5'SS, el 3'SS y el sitio de ramificación [revisado en Galej *et al.* (2014)]. En la región aminoterminal de PRP8 se encuentra el dominio N, que interacciona con la partícula U5 (Turner *et al.*, 2006).

PRP8 interviene en las dos reacciones de transesterificación del *splicing*, formando parte de las partículas U5 y trimérica U4/U6.U5 [revisado en Grainger y Beggs (2005)]. El centro de la cavidad de PRP8 mantiene su conformación durante las diferentes etapas del *splicing*, pero su dominio N rota durante la activación del espliceosoma, posicionando adecuadamente al sitio activo de la proteína en cada uno de los complejos [revisado en Shi (2017)]. El dominio Jab1/MPN de PRP8 interacciona directamente con la helicasa BRR2, regulando la actividad de esta última durante la fase catalítica y la separación de las partículas U durante el desensamblaje del espliceosoma [revisado en Absmeier *et al.* (2016)].

La interacción entre PRP8 y los 5'SS y 3'SS resulta evidente en los análisis de algunos alelos del gen *Prp8* de la levadura, que suprimen las alteraciones del *splicing* de los pre-ARNm de genes cuyas funciones no están relacionadas con el metabolismo del ARN. El *splicing* aberrante de dichos pre-ARNm se debe a mutaciones en sus 5'SS o 3'SS genuinos (también llamados canónicos, auténticos o verdaderos) o a la selección preferente por el espliceosoma de 5'SS y 3'SS nuevos, creados por mutaciones. El efecto supresor de varios alelos mutantes de *Prp8* de la levadura se debe a que sus mutaciones modifican la cavidad que forma el sitio activo de PRP8, lo que evidencia la importancia de esta proteína como un elemento central del espliceosoma [revisado en Grainger y Beggs (2005)].

PRP8 fue identificado por primera vez en una búsqueda de mutantes letales termosensibles de la levadura, en los que se acumulaban moléculas de pre-ARNm en condiciones restrictivas, como consecuencia de un ensamblaje incorrecto del espliceosoma (Lustig *et al.*, 1986). Los alelos nulos del gen *PRP8* de *Arabidopsis* causan en homocigosis una letalidad muy temprana: se interrumpe el desarrollo embrionario al no formarse el suspensor, una estructura de soporte que es la primera que se diferencia durante la embriogénesis. Debido a este rasgo fenotípico de sus alelos mutantes, *PRP8* se denominó inicialmente *SUSPENSOR2* (*SUS2*; Schwartz *et al.*, 1994). Algunas mutaciones nulas que

dañan el extremo carboxilo de la PRP8 humana causan retinitis pigmentosa, una enfermedad que cursa con degeneración de la retina, que usualmente causa ceguera (McKie *et al.*, 2001). Se desconoce el motivo de que dichas mutaciones en *PRP8* y en genes que codifican otros componentes del espliceosoma, que se expresan ubicuamente, alteren específicamente un tejido como la retina. Se ha propuesto que la causa es la gran demanda de proteínas de la retina, uno de los tejidos de mayor tasa metabólica [revisado en Yang *et al.* (2021)].

IV.2.2.4.- El *splicing* alternativo

Se denomina *splicing* alternativo de un pre-ARNm al procesamiento que conlleva la retención o eliminación de alguno de sus intrones o exones, respectivamente, como consecuencia de la selección diferencial de determinados 5'SS o el 3'SS por el spliceosoma. Esta retención o eliminación puede ser total o parcial. Se llama exones constitutivos a los que están presentes en todas las variantes del ARNm de un gen que sufre *splicing* alternativo (Figura 4A). El *splicing* alternativo es un mecanismo de regulación postranscripcional que

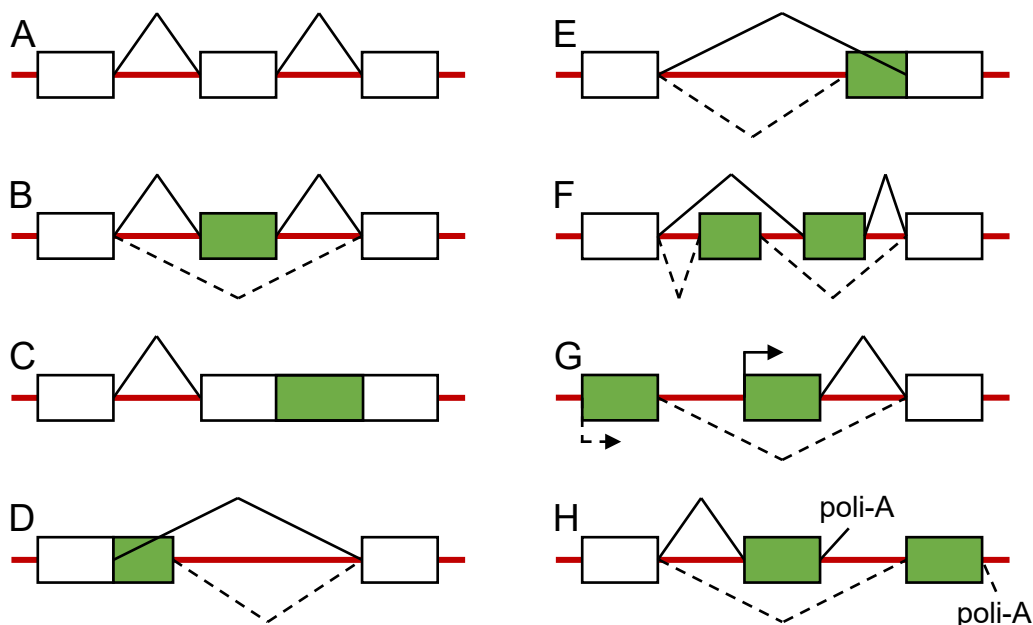


Figura 4.- Tipos de eventos de *splicing* alternativo. Los rectángulos blancos y verdes representan exones constitutivos y alternativos, respectivamente, y las líneas rojas, intrones. Los eventos de *splicing* constitutivo y alternativo se indican con líneas continuas y discontinuas, respectivamente. (A) *Splicing* constitutivo. (B) Exclusión de un exón. (C) Retención de un intrón. (D) Uso de un 5'SS alternativo presente en un intrón. (E) Uso de un 3'SS alternativo presente en un intrón. (F) Exclusión mutua de exones. (G) Uso de dos promotores alternativos (el símbolo \blacktriangleright señala los sitios de iniciación alternativa de la transcripción). (H) Uso de dos señales de poliadenilación (poli-A) alternativas. Modificado a partir de Wang y Burge (2008).

modula la expresión génica en diferentes tejidos, etapas del desarrollo o respuestas a estímulos, que puede producir múltiples variantes del ARNm de un gen, que a su vez pueden rendir diferentes isoformas de una proteína, contribuyendo así a la diversificación del proteoma [revisado en Graveley (2001); Lareau *et al.* (2004)].

La selección alternativa de los 5'SS y 3'SS es consecuencia de la existencia en los pre-ARNm de 5'SS y 3'SS alternativos, susceptibles de ser usados por el espliceosoma, o de los denominados 5'SS y 3'SS débiles, que requieren la actuación de factores activadores para ser reconocidos por el espliceosoma. Además de los 5'SS y 3'SS alternativos y los débiles, también intervienen factores de unión a ARN que estimulan o inhiben el *splicing* al unirse a las secuencias reguladoras ESE, ISE, ESS o ISS [revisado en Syed *et al.* (2012); Figura 1, en la página 10]. El 95% de los genes humanos que contienen intrones sufren *splicing* alternativo, mientras que en *Arabidopsis* es el 61% (Pan *et al.*, 2008; Marquez *et al.*, 2012).

Existen cuatro eventos de *splicing* alternativo frecuentes, que pueden darse en un mismo transcrito: la exclusión exónica, cuando un exón no es incluido en el ARNm (Figura 4B, en la página 15); la retención intrónica, cuando un intrón permanece en el ARNm maduro (Figura 4C), y el uso de 5'SS y 3'SS alternativos (Figura 4D y E), que incrementa la longitud de un exón a expensas de un intrón, o viceversa. Los tipos minoritarios son la exclusión mutua de exones (Figura 4F), que rinde transcritos alternativos que se diferencian en un exón, presente en una pero no en la otra variante del ARNm, y el uso alternativo de los exones inicial (5') o terminal (3') como consecuencia de la existencia de promotores y señales de poliadenilación alternativos, respectivamente (Figura 4G y H; revisado en Wang y Burge (2008)]. Los eventos de *splicing* alternativo más comunes en los genes humanos son los de eliminación de exones (40%), mientras que la retención intrónica supone menos del 5% [revisado en Keren *et al.* (2010); Reddy *et al.* (2013)]. Por el contrario, la retención intrónica es la más común en las plantas, suponiendo un 40% del total, y la eliminación de exones es solo del 8%, mientras que el resto de los eventos se deben principalmente al uso alternativo de los 5'SS y 3'SS (revisado en Reddy *et al.* (2013)].

El *splicing* alternativo contribuye principalmente a la regulación negativa de la expresión génica en las plantas, al alterar la pauta de lectura de un ARNm, cuya traducción puede por tanto rendir proteínas truncadas y usualmente no funcionales (Marquez *et al.*, 2012). Estos transcritos aberrantes pueden ser dianas de los mecanismos de control de la calidad del ARNm, que evitan su traducción, como la ruta Nonsense-Mediated Decay (NMD), que degrada en el citoplasma a los ARNm que contienen un codón de terminación prematuro [revisado en Rebbapragada y Lykke-Andersen (2009); Nicholson *et al.* (2010)].

Tanto en las plantas como en los animales se han identificado genes que poseen en

el extremo 3' de algunos de sus intrones tripletes NAG en tándem, que el espliceosoma usa como 3'SS alternativos. Las variantes de ARNm resultantes conservan su pauta de lectura abierta, pero a la vez incorporan o eliminan uno (cuando son dos los tripletes en tándem: NAGNAG) o más codones (Sinha *et al.*, 2010). En los animales y las plantas, este tipo de *splicing* alternativo en secuencias NAG en tándem puede generar diferencias tisulares en el proteoma (Bradley *et al.*, 2012). Se han descrito unos 5.000 genes de *Arabidopsis* que contienen secuencias NAG en tándem en alguno de sus intrones, que son usadas como 3'SS alternativos; algunos de estos genes codifican factores del *splicing* (Iida *et al.*, 2008; Schindler *et al.*, 2008).

IV.2.2.5.- Alteraciones del *splicing*

El *splicing* puede alterarse globalmente como consecuencia de mutaciones en los genes que codifican los componentes del espliceosoma o sus proteínas asociadas. Estas mutaciones causan un incremento generalizado de los eventos de retención intrónica o de las alteraciones en la fidelidad del *splicing*, al modificar las frecuencias relativas de selección de los 5'SS y 3'SS genuinos y alternativos [revisado en Dagueneit *et al.* (2015)]. Por el contrario, las mutaciones que dañan los 5'SS y 3'SS genuinos, el sitio de ramificación o los elementos activadores o silenciadores del *splicing* de genes concretos pueden causar alteraciones específicas en el patrón de *splicing* de estos últimos, pero no globales.

Las alteraciones del *splicing* de genes concretos dan cuenta de la tercera parte de las enfermedades genéticas hereditarias [revisado en Lim *et al.* (2011); Padgett (2012)]. En efecto, la eliminación o alteración de las secuencias necesarias para el *splicing*, principalmente los 5'SS y 3'SS, son la causa principal de generación de variantes de ARNm que provocan patologías. En el cáncer, sin embargo, prevalecen las mutaciones en los ESE y ESS, que inducen la exclusión exónica en algunos protooncogenes, que a su vez causa un incremento de su actividad (Sterne-Weiler *et al.*, 2011; Mort *et al.*, 2014; Supek *et al.*, 2014).

IV.2.3.- Mecanismos de control de calidad del ARNm

El metabolismo del ARNm, desde la síntesis del pre-ARNm hasta su exportación al citoplasma, no está exento de errores. La traducción de ARNm aberrantes producidos por un *splicing* alternativo o defectuoso puede generar proteínas truncadas y/o potencialmente tóxicas para la célula. Existen mecanismos de control de calidad de los ARNm que controlan su integridad durante su procesamiento, exportación y traducción.

El complejo TREX (TRanscription-Export) es el regulador principal de la exportación de los ARNm del núcleo al citoplasma. El complejo TREX se asocia cotranscripcionalmente a

los ARNm, propiciando su translocación a los poros nucleares al reclutar factores de exportación [revisado en Katahira (2012); Heath *et al.* (2016); Ehrnsberger *et al.* (2019)]. Los poros nucleares contienen unas 30 nucleoporinas, proteínas que forman una estructura en forma de cesta en la membrana nuclear, a través de la cual se exportan al citoplasma muchas macromoléculas y la mayoría de los ARNm. Dicha estructura presenta una parte externa en ambas caras de la membrana, con zonas de interacción con los factores de exportación, y una transmembrana, que forma el canal de traslocación [revisado en Björk y Wieslander (2017); Meier *et al.* (2017)].

El núcleo celular retiene moléculas de ARNm que sufren eventos de retención de intrones cortos y las que poseen 5'SS o 3'SS débiles, que el espliceosoma reconoce peor que los genuinos y, por tanto, usa poco o nada (Boutz *et al.*, 2015). Para la exportación nuclear del ARNm no solo es relevante el *splicing*, sino también la longitud de la cola poli-A, que debe ser de unos 50 nucleótidos; si es mayor se propicia la retención nuclear, que es completa cuando supera los 250 nucleótidos (Fuke y Ohno, 2008). La retención nuclear del ARNm también puede ser cotranscripcional, si se impide el reclutamiento de los factores de exportación, o postranscripcional, mediante su anclaje a diferentes estructuras nucleares, como la cromatina, que puede dificultar la separación de la ARN Pol II. También puede darse acumulación de ARNm en agregados nucleares o en la región interna de los poros nucleares [revisado en Wegener y Müller-McNicoll (2018)].

Los ARNm que contienen un codón de terminación prematura son reconocidos y degradados en la ruta NMD [revisado en Behm-Ansmant *et al.* (2007)], en la que intervienen tres factores muy conservados en todos los eucariotas, denominados UP-FRAMESHIFT1 (UPF1), UPF2 y UPF3 [revisado en Conti y Izaurralde (2005)]. Las uniones entre exones consecutivos que se producen tras el *splicing* quedan señalizadas por un Exon Junction Complex (EJC; Le Hir *et al.*, 2001). Dado que los codones de terminación normales suelen localizarse en el último exón de un gen, la posición de los EJC es crítica para el reconocimiento de los de terminación prematura: UPF2 y UPF3 activan a UPF1, que se une a los ARNm con algún codón de terminación situado aguas arriba de un EJC, promoviendo su degradación. Se impide así la síntesis de proteínas truncadas [revisado en Behm-Ansmant *et al.* (2007)].

IV.3.- El ribosoma eucariótico

IV.3.1.- Función, estructura y composición del ribosoma

El ribosoma citoplásmico eucariótico o 80S es un complejo ribonucleoproteico compuesto por dos subunidades, la 60S y la 40S, constituidas por cuatro ARNr y varias decenas de proteínas ribosómicas [revisado en Ramakrishnan (2002)]. La subunidad 40S

contiene el ARNr 18S y descodifica los ARNm que se incorporan al ribosoma. En la subunidad 60S radica el núcleo catalítico para la síntesis de la cadena polipeptídica y contiene los ARNr 25S, 5,8S y 5S [revisado en Merchante *et al.* (2017)]. Son 81 las proteínas ribosómicas que integran el ribosoma en *Arabidopsis*: 33 de la subunidad 40S, y 48 de la 60S (Barakat *et al.*, 2001; Carroll *et al.*, 2008; Carroll, 2013).

IV.3.2.- Biogénesis del ribosoma

La biogénesis del ribosoma es un proceso clave para cualquier célula, que conlleva un elevado gasto energético [revisado en Ni y Buszczak (2023)]. Son partes del proceso de generación de nuevos ribosomas la transcripción de los genes que codifican sus componentes estructurales y su maduración, su ensamblaje independiente en las subunidades 60S y 40S, la maduración de las propias subunidades y su exportación al citoplasma, y su ensamblaje final en un ribosoma 80S competente para llevar a cabo la traducción de los ARNm [revisado en Kressler *et al.* (2017); Sáez-Vásquez y Delseny (2019)].

La biogénesis del ribosoma se inicia en el nucleolo con la transcripción por la ARN Pol I del ADNr, denominado 45S en las plantas, 47S en la especie humana y 35S en los hongos. El ADNr 45S (47S o 35S) codifica los tres ARNr mayores que son componentes estructurales del ribosoma: el 25S, el 18S y el 5,8S. La ARN Pol II transcribe en el nucleoplasma los genes que codifican las proteínas ribosómicas, los factores de la biogénesis del ribosoma y los snoRNA. El ADNr 45S codifica el ARNr 45S, que es sintetizado por la ARN Pol III. El procesamiento del pre-ARNr 45S se inicia cotranscripcionalmente para generar finalmente los ARNr maduros 25S, 18S y 5,8S, mediante una serie de cortes endo y exonucleolíticos catalizados por decenas de factores de la biogénesis del ribosoma; ocurre en la región del componente fibrilar denso, uno de los tres compartimentos funcionales del nucleolo [revisado en Sáez-Vásquez y Delseny (2019)].

El procesamiento de los pre-ARNr ocurre coordinadamente con su ensamblaje con las proteínas ribosómicas en la región del componente granular del nucleolo, generándose así las subunidades prerribosómicas. Las proteínas ribosómicas estabilizan la estructura tridimensional de los ARNr, evitando plegamientos indebidos y facilitando la unión sucesiva de las siguientes. Las subunidades ribosómicas 40S y 60S, ensambladas pero inmaduras, se exportan al citoplasma, donde se producen las últimas etapas de la maduración de los ARNr y la incorporación de las últimas proteínas ribosómicas. Las subunidades ribosómicas maduras se ensamblan formando el ribosoma 80S, a la vez que se eliminan algunos factores de ensamblaje que no formarán parte del ribosoma maduro [revisado en Kressler *et al.* (2017)].

IV.3.2.1.- Estructura, organización y expresión de los ADNr

Los genes ADNr 45S (47S o 35S) y ADNr 5S están organizados en tándem y su número varía entre especies. *Arabidopsis* cuenta con unas 750 copias del ADNr 45S por genoma haploide, situadas en los brazos cortos de los cromosomas acrocéntricos 2 y 4 (Figura 5; Copenhaver y Pikaard, 1996a; b; Douet y Tourmente, 2007; Robledo *et al.*, 2008).

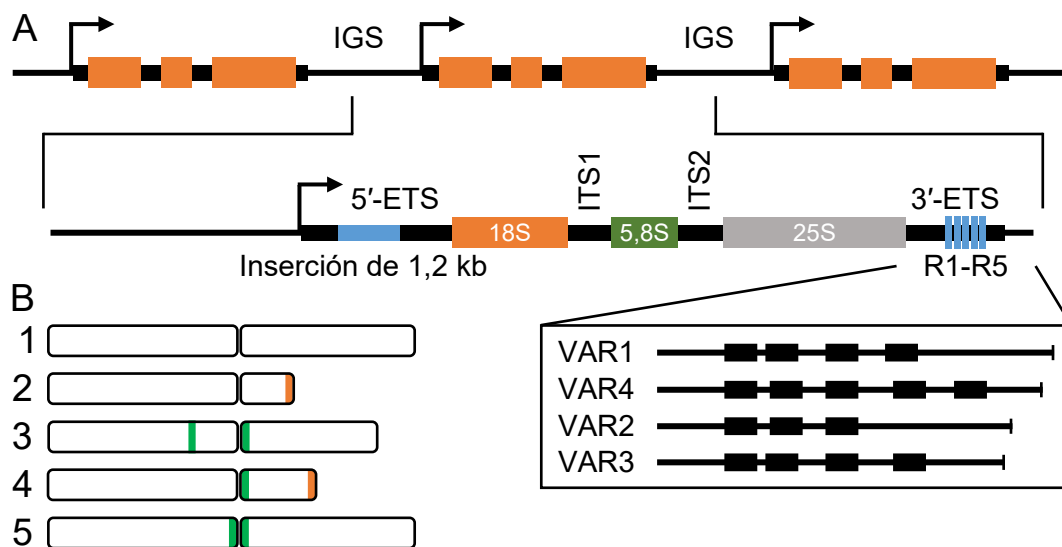


Figura 5.- Estructura y localización de los genes del ADNr 45S en *Arabidopsis*. (A) Estructura de una unidad del ADNr 45S, en la que los rectángulos naranjas, verdes y grises representan las regiones que codifican los ARNr 18S, 5,8S y 25S, respectivamente. Las regiones intergénicas (IGS) que separan las repeticiones se representan con una línea negra fina, y las espaciadoras internas (ITS1 e ITS2) y externas (5'-ETS y 3'-ETS) con una línea negra gruesa. El símbolo \rightarrow señala el sitio de inicio de la transcripción. Se representa en azul la inserción de 1,2 kb del 5'-ETS y la región polimórfica de las *VAR* en el acceso Col-0 (R1-R5), que se ha ampliado para destacar sus repeticiones internas (rectángulos negros). (B) Localización de los ADNr 45S (en naranja) y 5S (en verde) en los cromosomas de Col-0. Modificado a partir de Sáez-Vásquez y Delseny (2019).

Las regiones codificantes de los tres ARNr están separadas por dos espaciadores internos (Internal Transcribed Spacer [ITS]) en cada unidad del ADNr 45S: el ITS1 entre las regiones de los ARNr 18S y 5,8S, y el ITS2, entre las del 5,8S y el 25S. Dos espaciadores externos (External Transcribed Spacer [ETS]) flanquean las regiones codificantes: el 5'-ETS se sitúa aguas arriba de la región del ARNr 18S, y el 3'-ETS, aguas abajo de la del 25S (Figura 5). Cada unidad del ADNr 45S está separada de sus vecinas por un espaciador intergénico (Intergenic Spacer [IGS]), que contiene el promotor del gen ADNr 45S [revisado en Sáez-Vásquez y Delseny (2019)]. La transcripción de los ADNr 45S, 47S y 35S por la ARN Pol I

rinde un transcrito primario policistrónico, denominado pre-ARNr 45S, 47S y 35S (Gruendler *et al.*, 1991; Doelling y Pikaard, 1995).

Las agrupaciones en tándem de los ADNr 45S, 47S y 35S se encuentran en las regiones de los organizadores nucleolares (Nucleolus Organizer Regions [NOR]), cuyo nombre deriva de que el nucleolo se organiza en torno a las repeticiones de ADNr cuando la ARN pol I comienza a transcribirlas. Los 3'-ETS de los ADNr 45S de *Arabidopsis* son polimórficos y su transcripción rinde diferentes isoformas del pre-ARNr 45S, a las que se ha denominado variantes *VAR*. En el acceso Columbia-0 (Col-0) se han detectado cuatro *VAR* (*VAR1-VAR4*), asociándose mayoritariamente la *VAR1* a la expresión de los genes del NOR2, y las otras tres, a los del NOR4 (Figura 5, en la página 20; Earley *et al.*, 2010; Pontvianne *et al.*, 2010; Chandrasekhara *et al.*, 2016). La existencia de estos polimorfismos ha permitido establecer que no todas las copias del ADNr 45S se expresan. Los ADNr 45S del NOR2 de Col-0 están epigenéticamente silenciados, expresándose solo durante los primeros días tras la germinación (Earley *et al.*, 2006; Pontvianne *et al.*, 2010); los del NOR4 están transcripcionalmente activos durante el resto del ciclo de vida de la planta (Chen y Pikaard, 1997; Fransz *et al.*, 2002).

Se han identificado reguladores de la expresión de los ADNr 45S del NOR2. Las chaperonas de histonas NUCLEOLIN1 (NUC1) y NUC2 actúan de forma antagónica, activando y reprimiendo la transcripción de dichos ADNr 45S, respectivamente. CELL DIVISION CYCLE 48A (CDC48A), HISTONE DESACETYLASE 6 (HDA6) y MORPHOLOGY OF ARGONAUTE1-52 SUPPRESSED2 (MAS2) reprimen la expresión de estos ADNr 45S [revisado en Sáez-Vásquez y Delseny (2019)].

Existen en *Arabidopsis* unas 1.000 copias del ADNr 5S por genoma haploide (Campbell *et al.*, 1992), situadas en las regiones pericentroméricas de los cromosomas 3, 4 y 5 (Figura 5; Murata *et al.*, 1997; Fransz *et al.*, 1998; Tutois *et al.*, 1999). Cada gen del ADNr 5S tiene unos 500 pb, de las que 120 corresponden a la unidad de transcripción y el resto constituyen un espaciador intergénico (Cloix *et al.*, 2003; Layat *et al.*, 2012). En Col-0 solo se transcriben las copias del ADNr 5S del cromosoma 4 y las del brazo largo del 5 (Douet y Tourmente, 2007).

IV.3.2.2.- Procesamiento de los pre-ARNr

Durante el procesamiento de los pre-ARNr 45S, 47S o 35S se eliminan los 5'-ETS, 3'-ETS, ITS1 e ITS2 (Figura 6, en la página 22). Este proceso está muy conservado en los eucariotas, habiéndose estudiado principalmente en la levadura y la especie humana. La gran conservación estructural y funcional de los factores de la biogénesis del ribosoma y de las

rutas de maduración de los ARNr facilita su estudio en otras especies, como *Arabidopsis* (Tomecki *et al.*, 2017).

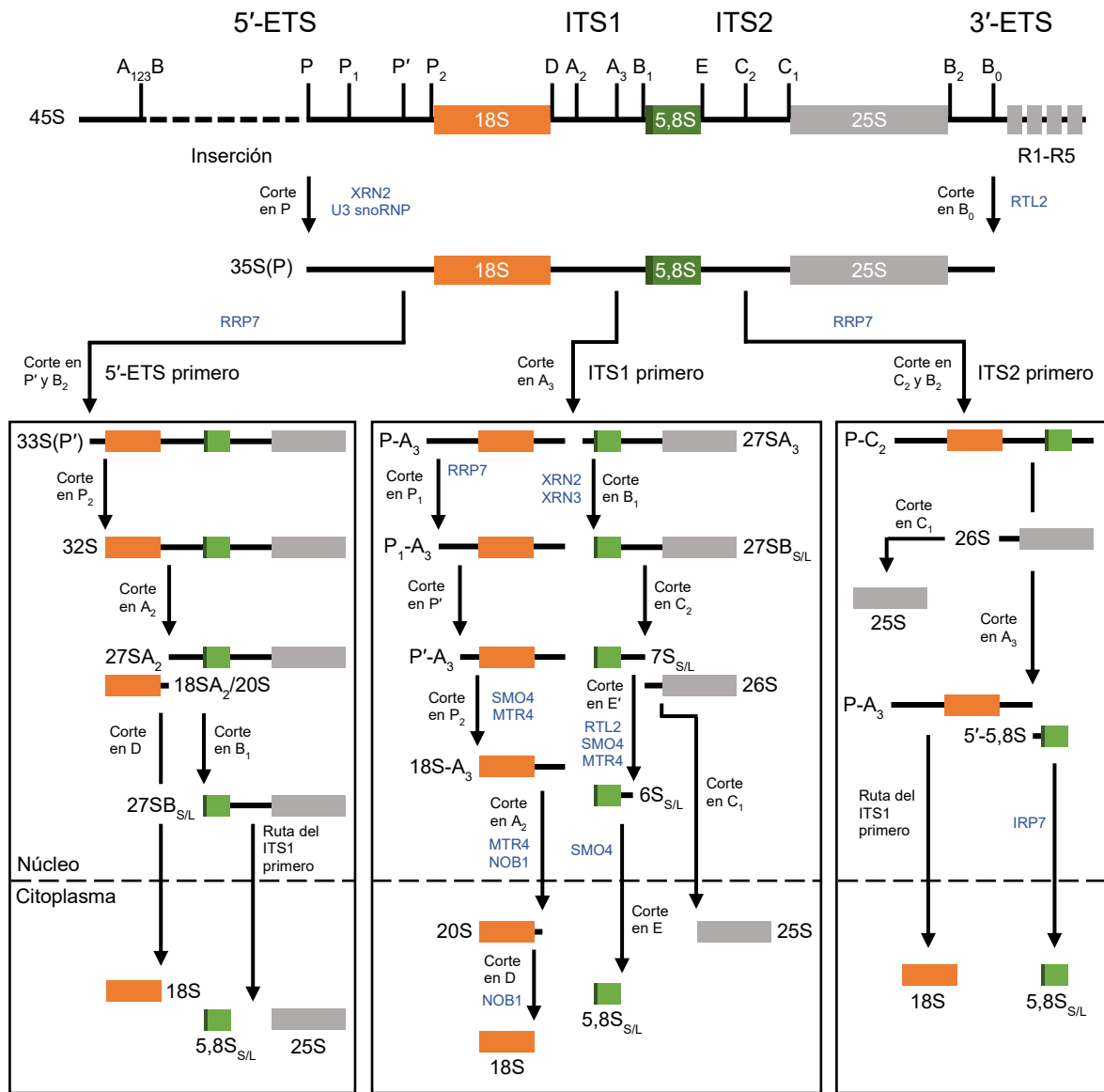


Figura 6.- Procesamiento del pre-ARNr 45S en *Arabidopsis*. Se muestran sus tres rutas de maduración de los ARNr. Las líneas verticales en el pre-ARNr 45S señalan los sitios de corte endonucleolítico; las flechas, los cortes que ocurren en cada etapa, en las que también se indican en azul los factores implicados; la línea horizontal discontinua, la membrana nuclear, y las flechas que la atraviesan, la salida de las distintas especies de pre-ARNr y ARNr al citoplasma. Modificado a partir de Sáez-Vásquez y Delseny (2019).

La maduración de los ARNr comienza en el 5'-ETS, que posee varios sitios de corte endonucleolítico, que reciben diferentes nombres según la especie; en *Arabidopsis* se denominan A₁₂₃B, P, P₁, P' y P₂ (Figura 6). El 5'-ETS es más largo en *Arabidopsis* que en la

levadura o los animales, como consecuencia de una inserción de 1,2 kb (Figura 5, en la página 20; Tomecki *et al.*, 2017). La primera etapa de la eliminación del 5'-ETS humano y de la levadura es un corte endonucleolítico en el sitio P. En Arabidopsis, sin embargo, le precede la degradación exonucleolítica de la antes mencionada inserción de 1,2 kb hasta el sitio de corte A₁₂₃B, ubicado aguas arriba de P (Figura 6, en la página 22). Esta degradación es catalizada por la 5'-3' EXORIBONUCLEASE2 (XRN2; Sáez-Vásquez *et al.*, 2004; Zakrzewska-Placzek *et al.*, 2010).

La eliminación del 3'-ETS comienza con un corte endonucleolítico en el pre-ARNr que se está transcribiendo, que causa su separación de los complejos de elongación de la transcripción, que contienen la ARN Pol I, finalizando así la transcripción. En la levadura, dicho corte lo realiza la RNase three1 (Rnt1), y en Arabidopsis, su ortóloga RNASE THREE LIKE 2 (RTL2; Comella *et al.*, 2008). No se ha establecido como actúa RTL2 ni se ha determinado su sitio de corte [revisado en Tomecki *et al.* (2017); Sáez-Vásquez y Delseny (2019)]. La molécula resultante, denominada pre-ARNr 35S(P) en Arabidopsis (35S en la levadura y 45S en la especie humana), es procesada mediante dos rutas alternativas (Henras *et al.*, 2015): la denominada del ITS1 primero (ITS1 first), que es la mayoritaria y se inicia con un corte en el ITS1 (Zakrzewska-Placzek *et al.*, 2010), y la del 5'-ETS primero (5'-ETS first), que es minoritaria y conlleva la degradación total del 5'-ETS antes de que sucedan cortes en los ITS (Figura 6; Weis *et al.*, 2015b).

La ruta del 5'-ETS primero se inicia con dos cortes sucesivos en los sitios P' y P₂ del precursor 35S(P), que dan lugar a los pre-ARNr 33S(P') y 32S, respectivamente. Se producen a continuación otros dos cortes en los sitios A₂ y B_{1L} del ITS1, que rinden el pre-ARNr 20S, precursor del ARNr 18S, y el pre-ARNr 27B_{S/L}, precursor de los ARNr 5,8S y 25S. El ITS1 está totalmente ausente del pre-ARNr 27B_{S/L} (Figura 6; Zakrzewska-Placzek *et al.*, 2010; Weis *et al.*, 2015a). Un corte en C₂ separa a continuación el precursor 7S del ARNr 5,8S, y el 26S del ARNr 25S. El pre-ARNr 7S sufre una degradación exonucleolítica hasta el sitio E', se exporta al citoplasma y allí completa su procesamiento, tras el corte en E, generando el ARNr 5,8S maduro. Por su parte, el pre-ARNr 26S sufre un corte en C₁ que genera el ARNr 25S, que se exporta maduro al citoplasma (Figura 6).

En la ruta del ITS1 primero, un corte en el sitio A₃ fragmenta el pre-ARNr 35S(P), rindiendo los precursores P-A₃ y 27SA₃ (Weis *et al.*, 2015a). Se elimina a continuación el 5'-ETS del pre-ARNr P-A₃, mediante cortes en P' y P₂, generándose así el pre-ARNr 18S-A₃, que con un último corte en D, realizado en el citoplasma por NIN1 (ONE) BINDING PROTEIN 1 (NOB1), produce el ARNr 18S maduro (Veith *et al.*, 2012; Missbach *et al.*, 2013). Por su parte, el pre-ARNr 27SA₃, sufre un corte en C₂ que produce el pre-ARNr A₃-C₂, precursor del ARNr

5,8S, y el pre-ARNr 27SB_{S/L}, precursor del ARNr 25S. Estos dos fragmentos sufren un procesamiento muy similar al descrito para los pre-ARNr 7S y 26S para dar lugar a los ARNr maduros correspondientes (Figura 6, en la página 22).

A pesar de la gran conservación entre estas rutas de procesamiento, se siguen identificando sitios de corte distintos entre diferentes eucariotas. Por ejemplo, los pre-ARNr equivalentes a los 35S, 33S y 32S de *Arabidopsis* se han detectado en la especie humana, pero no en la levadura (Zakrzewska-Placzek *et al.*, 2010; Missbach *et al.*, 2013; Henras *et al.*, 2015; Weis *et al.*, 2015b; Maekawa *et al.*, 2018). Esta diferencia sugiere que, mientras que en *Arabidopsis* (sitios P, P₁, P' y P₂) y la especie humana (A₀, A₁ y A₂) los cortes son postranscripcionales, sus equivalentes en la levadura (A₀, A₁ y A₂) son cotranscripcionales [revisado en Sáez-Vásquez y Delseny (2019)].

Se ha identificado en *Arabidopsis* una tercera ruta de procesamiento que parece ser específica de las plantas (Palm *et al.*, 2019). En esta ruta, llamada del ITS2 primero (ITS2 first), se produce un corte inicial en el sitio C₂, que fragmenta el pre-ARNr 35S(P) en dos precursores: el pre-ARNr P-C₂, que contiene las secuencias de los ARNr 18S y 5,8S, y el pre-ARNr 26S, que contiene la del ARNr 25S (Figura 6). Esta ruta solo se ha detectado en mutantes portadores de alelos del gen *INVOLVED IN RNA PROCESSING 7 (IRP7)*, en los que se observa una acumulación de estos precursores en tejidos de rápida proliferación, tras un tratamiento con auxinas. En esta ruta, IRP8 e IRP9 son necesarios para el procesamiento del pre-ARNr 5'-5,8S, que también parece ser exclusivo de las plantas (Palm *et al.*, 2019).

IV.3.2.3.- Las proteínas ribosómicas

Las proteínas ribosómicas son componentes estructurales del ribosoma maduro. Existen unas 80 en los ribosomas eucarióticos 80S, que están muy conservadas entre los hongos, los animales y las plantas (Doudna y Rath, 2002). Los genes que codifican proteínas ribosómicas son transcritos en el núcleo y sus ARNm se traducen en el citoplasma. Las proteínas ribosómicas son posteriormente importadas al núcleo.

Las proteínas ribosómicas no solo son componentes estructurales del ribosoma, ya que también son fundamentales en las diferentes etapas de la maduración de los ARNr: la maduración de los pre-ARNr ocurre a la vez que se incorporan a las partículas prerribosómicas. De hecho, las proteínas ribosómicas participan en el reclutamiento de algunos factores de la biogénesis del ribosoma y guían el plegamiento adecuado de los pre-ARNr en su incorporación a la partícula prerribosómica en la levadura [revisado en de la Cruz *et al.* (2015)].

IV.3.2.3.1.- Genes que codifican proteínas ribosómicas

Las 80 proteínas que conforman el ribosoma 80S humano están codificadas por 85 genes [revisado en Uechi *et al.* (2001); Melnikov *et al.* (2012)], mientras que en la levadura son 137 los genes que codifican sus 78 proteínas ribosómicas (Planta y Mager, 1998). Como consecuencia de las duplicaciones genómicas que han sufrido las plantas a lo largo de su evolución, son dos o más genes parálogos los que codifican cada proteína ribosómica (Barakat *et al.*, 2001). Se han identificado 255 genes de *Arabidopsis* que codifican proteínas ribosómicas, que se agrupan en 81 familias génicas (Barakat *et al.*, 2001; Carroll *et al.*, 2008; Hummel *et al.*, 2012; Ding *et al.*, 2022); 103 de estos genes codifican las 33 proteínas ribosómicas de la subunidad 40S, y los restantes 152, las 48 de la subunidad 60S. Una de estas familias génicas de *Arabidopsis* cuenta con 7 miembros, que producen 7 proteínas RPS15 casi idénticas (Carroll *et al.*, 2008). En otras especies vegetales el número de genes es mucho mayor: *Brassica napus* cuenta con 996 genes de proteínas ribosómicas (Lysak *et al.*, 2005; Whittle y Krochko, 2009).

Los análisis transcriptómicos de los genes parálogos de *Arabidopsis* que codifican proteínas ribosómicas han revelado que casi todos se expresan, muchos de ellos con distintos patrones espaciales y/o temporales (Schmid *et al.*, 2005; Savada y Bonham-Smith, 2014). En la levadura, sin embargo, muchos de estos genes están pseudogenizados (Wolfe y Shields, 1997). Los análisis proteómicos del ribosoma 80S han confirmado que sus proteínas suelen estar codificadas por un único gen en los mamíferos, y entre uno y dos en la levadura, mientras que en las plantas ninguna de las proteínas ribosómicas estudiadas está codificada por un único gen [revisado en Carroll (2013)].

IV.3.2.3.2.- Funciones extrarribosómicas de las proteínas ribosómicas

Se denomina haploinsuficiencia a una condición genética en la que la presencia de un único alelo funcional de un gen no es suficiente para rendir un fenotipo silvestre y, en consecuencia, la heterocigosis para un alelo silvestre y otro hipomorfo o nulo rinde un fenotipo mutante [revisado en Veitia (2002); Navarro-Quiles *et al.* (2023)]. Aunque en *Arabidopsis* se han identificado pocos casos de haploinsuficiencia, en la especie humana son muchas las enfermedades neurodegenerativas y alteraciones del desarrollo embrionario que se deben a loci haploinsuficientes [revisado en Zug (2022)]. Esto puede explicarse por la mayor tolerancia a los cambios en los niveles de ploidía de las plantas respecto a los animales [revisado en Meinke (2013)].

Se denomina no complementación no alélica al fenómeno que manifiestan algunos diheterocigotos para los alelos silvestres y mutantes recesivos de dos genes no ligados, cuyo

fenotipo es mutante a pesar de que los correspondientes heterocigotos simples son fenotípicamente silvestres [revisado en Hawley y Gilliland (2006)]. Se denomina haploinsuficiencia combinada a la no complementación no alélica que no muestra especificidad de alelo y es dependiente de dosis, que se ha descrito para las familias génicas de *Arabidopsis* que codifican las proteínas ribosómicas RPS6, RPL5, RPL23 y RPL36. Estas familias génicas incluyen genes parálogos muy redundantes, y sus diheterocigotos presentan un fenotipo similar al de los homocigotos simples (Degenhardt y Bonham-Smith, 2008; Fujikura *et al.*, 2009; Creff *et al.*, 2010; Casanova-Sáez *et al.*, 2014).

La existencia de genes parálogos que codifican proteínas ribosómicas total o parcialmente redundantes ha posibilitado la adquisición, en algunos casos, de funciones adicionales a las que realizan en el ribosoma, que pueden estar relacionadas o no con la traducción [revisado en Xiong *et al.* (2021)]. Son ejemplos de ello en *Arabidopsis* las proteínas RPS2, RPS6, RPL10 y RPL24; tal como a continuación se detalla, mientras que las dos primeras tienen funciones extrarribosómicas, ya que actúan como factores de la biogénesis del ribosoma, las dos últimas participan en procesos no relacionados con la traducción.

La familia RPS2 incluye seis genes, cuyos mutantes simples tienen el fenotipo característico de hojas apuntadas y crecimiento lento que causan los alelos de otros genes de la maquinaria de la traducción. Los cuádruples mutantes obtenidos combinando alelos de pérdida de función de genes *RPS2* presentan defectos en el procesamiento del 5'-ETS y un retraso en el de los precursores de los ARNr 25S y 18S (Hang *et al.*, 2021). A su vez, la familia RPS6 incluye dos genes redundantes, que manifiestan no complementación no alélica y son esenciales (Creff *et al.*, 2010). Las proteínas RPS6 se unen al promotor del ADNr 45S, cuya transcripción reprimen (Kim *et al.*, 2014).

La familia RPL10 incluye tres genes, que codifican proteínas que son sustratos de la NUCLEAR SHUTTLE PROTEIN (NSP)-INTERACTING KINASE (NIK). La fosforilación de las proteínas RPL10 por NIK propicia su translocación al núcleo, en donde activa la respuesta antiviral (Carvalho *et al.*, 2008). Los genes *RPL10* también están relacionados con la respuesta al estrés causado por la radiación ultravioleta; sus alelos nulos son deficientes en dicha respuesta, en un grado variable (Ferreira *et al.*, 2010a; Ferreira *et al.*, 2010b). La familia RPL24 incluye dos miembros, que son esenciales para el inicio de la traducción de muchas proteínas, como algunas de respuesta a las auxinas. Las proteínas RPL24 se translocan al núcleo, en donde interactúan con los transcritos primarios de los genes que codifican miARN, para facilitar su procesamiento (Li *et al.*, 2017).

Las fosfoproteínas ácidas ribosómicas (RRP) son proteínas ribosómicas que forman parte de una protuberancia lateral de la subunidad 60S (Szick *et al.*, 1998). La familia RRP3

es exclusiva de las plantas, y en *Arabidopsis* está formada por dos parálogos. Se ha encontrado a RRP3 en complejos que se forman bajo condiciones de estrés causado por temperaturas altas, que contienen principalmente proteínas de respuesta a choque térmico. RRP3 actúa como chaperona de otras proteínas o de ARN, protegiendo a las plantas ante el estrés térmico causado por el calor o el frío, respectivamente (Kang *et al.*, 2016).

IV.3.2.4.- Fenotipos causados por defectos en la biogénesis del ribosoma o en la traducción

La biogénesis del ribosoma requiere un equilibrio entre las concentraciones de los componentes del ribosoma, razón por la que la transcripción de los genes implicados y la síntesis de las proteínas ribosómicas y los factores de la biogénesis del ribosoma debe estar estrictamente regulada (Laferté *et al.*, 2006). La depleción de solo uno de los componentes de la maquinaria de biogénesis del ribosoma puede alterar la producción de los restantes y causar la formación de ribosomas defectuosos. La disfunción de alguno de los genes que codifican proteínas ribosómicas o factores de la biogénesis del ribosoma causa un fenotipo característico. Una de las especies en las que se ha observado este fenómeno es *Arabidopsis*, en la que la homocigosis para alelos de insuficiencia de función de genes de la maquinaria de la traducción ralentiza el desarrollo de las plantas, que tienen hojas apuntadas con márgenes dentados (Horiguchi *et al.*, 2011).

Las enfermedades humanas causadas por mutaciones en genes que codifican proteínas ribosómicas o factores de la biogénesis del ribosoma se denominan ribosomopatías [revisado en Narla y Ebert (2010)]. Aunque los defectos producidos por la pérdida de función de un componente de la maquinaria de la traducción deberían alterar el desarrollo en cualquier histotipo, se manifiestan especialmente en la hematopoyesis. Por ejemplo, la anemia de Diamond-Blackfan es una enfermedad congénita rara, asociada a insuficiencias en la médula ósea, causadas por mutaciones en genes que codifican factores de la biogénesis del ribosoma o proteínas ribosómicas, entre ellos los de la familia RPS24 (revisado en Kampen *et al.* (2020)]. El síndrome de Shwachmann-Diamond es otra ribosomopatía, que altera las funciones del páncreas y la médula ósea como consecuencia de la incorrecta maduración de la subunidad 60S (Woloszynek *et al.*, 2004; Weis *et al.*, 2015c). Por último, el síndrome de hipoplasia cartílago-cabello, que se caracteriza por la escasa actividad del cartílago de crecimiento de los huesos largos, se debe a mutaciones en el gen del ARN nuclear no codificante *Ribonuclease Mitochondrial RNA Processing (RMRP)* que inhiben el corte del ITS1 del pre-ARNr 47S y, en consecuencia, dificultan la maduración de los ARNr 18S y 5,8S (Thiel *et al.*, 2005; Goldfarb y Cech, 2017).

Las ribosomopatías, además, incrementan el riesgo de sufrir cáncer (Taskinen *et al.*, 2008; Alter *et al.*, 2018; Vlachos *et al.*, 2018), ya que pueden propiciar la aparición de ribosomas especializados en la traducción de determinados ARNm [revisado en Elhamamsy *et al.* (2022)], reduciendo la expresión de genes supresores de tumores como *TUMOR PROTEIN 53 (TP53)*, o incrementando la de protooncogenes como *B-CELL LYMPHOMA 2 (BCL-2)*. En estos dos últimos casos se favorecería la formación de tumores [revisado en Kampen *et al.* (2020)]. Además, la presencia de ribosomas defectuosos también induce estrés oxidativo, por la generación de especies reactivas de oxígeno. Aunque estas últimas reducen la proliferación celular, también dañan moléculas como el ADN. Las mutaciones en los protooncogenes y genes supresores de tumores propician procesos tumorales [revisado en De Keersmaecker *et al.* (2015); Kampen *et al.* (2020)].

IV.4.- Antecedentes y objetivos

IV.4.1.- Identificación de supresores del fenotipo morfológico de *ago1-52*

El silenciamiento génico mediado por miARN es uno de los principales mecanismos de regulación postranscripcional de la expresión génica en los animales y las plantas. El principal efector de esta ruta en *Arabidopsis* es la proteína ARGONAUTE1 (AGO1), una endonucleasa de ARN que impide la traducción de los ARNm a los que se ha unido por complementariedad un miARN (Bartel, 2009). El mutante *ago1-52* fue identificado en una búsqueda de mutantes foliares realizada en el laboratorio de José Luis Micol. Dichos mutantes se indujeron mediante tratamiento de semillas del acceso Landsberg *erecta* (Ler) con metanosulfonato de etilo (EMS), un mutágeno que genera transiciones G→A, y en consecuencia, también C→T (Berná *et al.*, 1999; Jover-Gil *et al.*, 2012). El alelo *ago1-52* presenta una transición G→A en su vigésimo intrón, que provoca la aparición de un 3'SS nuevo, que el espliceosoma utiliza más frecuentemente que el genuino, que está inalterado. Como resultado, *ago1-52* produce dos ARNm, uno silvestre y muy minoritario, y otro mutante, que incluye 10 nucleótidos del vigésimo intrón, que además desfasan su pauta de lectura. La traducción de este ARNm aberrante genera una proteína AGO1-52 mutante, con 55 aminoácidos menos y 15 diferentes de los de la proteína silvestre AGO1 en su extremo carboxilo. Además, la proteína mutante AGO1-52 es mayoritaria frente a la silvestre AGO1 silvestre en el mutante *ago1-52* (Jover-Gil *et al.*, 2012; Sánchez-García *et al.*, 2015).

Un escrutinio posterior, realizado en el laboratorio de María Rosa Ponce tras una mutagénesis de segundos sitios del mutante *ago1-52*, permitió la identificación de 23 líneas portadoras de mutaciones supresoras de su fenotipo morfológico. A los correspondientes genes supresores se les denominó *MORPHOLOGY OF ARGONAUTE1-52 SUPPRESSED*

(MAS; Aguilera-Díaz, 2009; Micol-Ponce *et al.*, 2014). Diez de las mutaciones *mas* dañaban el gen AT4G02720, que codifica el presunto ortólogo de la NFκB activating protein (NKAP) humana, y fue denominado MAS2 (Sánchez-García *et al.*, 2015). Otras once líneas eran portadoras de mutaciones en el gen *PRE-MRNA PROCESSING FACTOR 8 (PRP8)*, que codifica la proteína PRP8, un factor central del espliceosoma, que está ampliamente conservado en los eucariotas (apartado IV.2.2.3, en la página 13). Todas las mutaciones *mas2* y *mas5* son de cambio de sentido, y las que se estudiaron con más detalle resultaron ser dominantes frente a su alelo silvestre.

IV.4.2.- Identificación de interactores de MAS2

Las mutaciones puntuales *mas2* son supresores extragénicos del fenotipo morfológico del mutante *ago1-52* porque corrigen parcialmente el *splicing* aberrante del alelo *ago1-52*: reducen el uso preferente por el espliceosoma del 3'SS nuevo generado por una mutación G→A en el vigésimo intrón del gen *AGO1* (Sánchez-García *et al.*, 2015). En efecto, mientras que los niveles del ARNm silvestre de *AGO1* fueron casi indetectables en el mutante simple *ago1-52*, su nivel respecto al del ARNm *AGO1-52* mutante se incrementó ligeramente en los dobles mutantes *ago1-52 mas2*. La causa molecular de la supresión fue aún más evidente al analizar las concentraciones relativas de la proteína silvestre *AGO1* y la mutante *AGO1-52*: se igualaron en el doble mutante, mientras que *AGO1* fue muy minoritaria en el mutante simple *ago1-52* (Sánchez-García *et al.*, 2015). Estos resultados sugirieron la relación directa o indirecta de MAS2 con el *splicing* de los pre-ARNm del gen *AGO1*, la exportación al citoplasma de sus ARNm defectuosos y/o su traducción.

El estudio de MAS2 reveló que es un gen esencial en Arabidopsis. Su silenciamiento parcial en estirpes transgénicas productoras de miARN artificiales produjo plantas con un fenotipo pleiotrópico, más pequeñas que las silvestres, ligeramente cloróticas y con hojas apuntadas y con margen dentado; esta observación sugirió la implicación de MAS2 en la traducción. Además, el análisis de la biogénesis del ribosoma en una de dichas estirpes reveló la implicación de MAS2 en la represión de la transcripción del ADNr 45S (Sánchez-García *et al.*, 2015).

La identificación de interactores de la proteína nuclear MAS2 mediante ensayos del doble híbrido de la levadura confirmó su naturaleza multifuncional, ya que parecía actuar en diferentes rutas del metabolismo de los ARN. Trece de las 14 presuntas interactoras de MAS2 no se habían estudiado previamente en las plantas. Tres de ellas eran ortólogas de factores que se habían encontrado asociados a alguno de los complejos del espliceosoma en otras especies. Otras tres resultaron ser ortólogas de proteínas que actúan como factores de la

biogénesis del ribosoma en otras especies. Dos de estas últimas tenían ortólogas humanas y de la levadura que eran factores de la biogénesis del ribosoma: RIBOSOMAL RNA PROCESSING 7 (RRP7), que participa en la maduración del ARNr 18S en ambas especies (Baudin-Baillieu *et al.*, 1997; Tafforeau *et al.*, 2013) y NUCLEOLAR PROTEIN 53 (NOP53), que actúa en la maduración de los ARNr 5,8S y 25S de la levadura y los 5,8S y 18S humanos (Granato *et al.*, 2005; Sydorsky *et al.*, 2005; Thomson y Tollervey, 2005; Tafforeau *et al.*, 2013). Poco después de publicarse el artículo del laboratorio de María Rosa Ponce sobre MAS2 (Sánchez-García *et al.*, 2005), otros autores llamaron *SMALL ORGAN 4 (SMO4)* al gen *NOP53* de *Arabidopsis*, por los defectos en la proliferación celular causados por sus alelos mutantes (Zhang *et al.*, 2015b). Los estudios de RRP7 y SMO4 (NOP53) realizados en el laboratorio de María Rosa Ponce demostraron su conservación funcional como factores de la biogénesis del ribosoma, ya que participan en la maduración de los ARNr 18S y 5,8S, respectivamente (Micol-Ponce *et al.*, 2018; Micol-Ponce *et al.*, 2020). RRP7 también parecía jugar un papel represor de la transcripción del ADNr 45S, como MAS2 (Micol-Ponce *et al.*, 2018). La tercera interactora de MAS2 potencialmente relacionada con la biogénesis del ribosoma fue RPS24B, una de las dos coortólogas de la proteína ribosómica RPS24 humana y de la levadura, que no solo son componentes estructurales del ribosoma 80S, sino que también presentan funciones extrarribosómicas, actuando como factores de la biogénesis del ribosoma en la maduración del ARNr 18S (Ferreira-Cerca *et al.*, 2005; Choessel *et al.*, 2008; Robledo *et al.*, 2008).

IV.4.3.- Objetivos de esta Tesis

Los principales objetivos iniciales de esta Tesis Doctoral fueron establecer (1) la naturaleza molecular del efecto de las mutaciones *mas5* sobre la función de PRP8, así como la de su supresión de los fenotipos de mutantes en los que está alterado el *splicing* de genes concretos, y (2) si RPS24B y su paróloga RPS24A tienen funciones extrarribosómicas, tal como ocurre con sus ortólogas humana y de la levadura, que actúan en la biogénesis del ribosoma.

En cuanto al primer objetivo inicial, nos propusimos: (1) analizar los efectos de las mutaciones *mas5* sobre la estructura primaria de la proteína PRP8, (2) deducir el impacto de las mutaciones *mas5* sobre la estructura secundaria y actividad de PRP8, (3) establecer el mecanismo molecular de la supresión del fenotipo del mutante *ago1-52* por las mutaciones *mas5*, (4) analizar los eventuales efectos supresores de las mutaciones *mas5* sobre mutantes con diferentes alteraciones en el *splicing* de genes concretos, y (5) analizar el impacto global de las mutaciones *mas5* en el *splicing* de los pre-ARNm de *Arabidopsis*.

En cuanto al segundo objetivo inicial, nos propusimos: (6) caracterizar alelos mutantes insercionales y puntuales de *RPS24B* y *RPS24A*, y estudiar su eventual redundancia funcional, (7) determinar la localización subcelular de las proteínas RPS24A y RP24B, (8) inferir mediante abordajes genéticos las eventuales relaciones funcionales de *RPS24A* y *RPS24B* con *RRP7*, *SMO4* (*NOP53*) y otros genes que codifican factores de la biogénesis del ribosoma, (9) estudiar la maduración de los ARNr en los mutantes simples *rps24a* y *rps24b* y en sus combinaciones dobles con alelos mutantes de genes que codifican factores de la biogénesis del ribosoma, (10) analizar la expresión del ADNr 45S en los mutantes *rps24a* y *rps24b*, y (11) identificar interactores físicos de RPS24B.

V.- MATERIALES Y MÉTODOS

V.- MATERIALES Y MÉTODOS

Para la redacción de los apartados I a VII de esta memoria se han seguido las mismas pautas que en Tesis anteriores de los laboratorios de María Rosa Ponce y José Luis Micol. En este apartado de Materiales y métodos se reproducen literalmente algunas frases procedentes de dichas Tesis. Se ha preferido usar los acrónimos ADN y ARN para los ácidos desoxirribonucleico y ribonucleico, respectivamente, ya que son de uso común en los medios de comunicación españoles. Sin embargo, hemos mantenido el término “*splicing*”, y adoptado “*espliceosoma*” como versión castellanizada de *spliceosome*, ya que ambos se recogen en “Enclave de Ciencia” (<https://enclavedeciencia.rae.es>), una plataforma que pretende dar soporte a la comunicación científica y tecnológica, que ha desarrollado la Real Academia Española (RAE) en colaboración con la Fundación Española para la Ciencia y la Tecnología (FECYT). Además, *splicing* está ampliamente aceptado en la mayoría de los textos docentes de genética y biología molecular en español; sus alternativas, como “ajuste” o “corte y empalme” nos parecen poco adecuadas, al ser “ajuste” un término inusual, y “corte y empalme”, poco riguroso. También hemos usado repetidas veces el acrónimo SS (de “splice site”). Tal como recomienda la RAE, en esta memoria no se pluralizan las siglas, y por tanto se escribe “el ARN” y también “los ARN”.

La nomenclatura que se aplica en esta memoria a genes y mutaciones se atiene a las pautas propuestas para Arabidopsis por Meinke y Koornneef (1997). Hemos empleado la tipografía cursiva exclusivamente para los genes, alelos, mutaciones y mutantes. También hemos escrito *splicing* en cursiva para destacar su carácter foráneo y que conserva su pronunciación original. No hemos traducido al español la mayoría de los nombres de genes y proteínas que se mencionan en esta memoria. Los transgenes se denotan según lo establecido en las instrucciones a los autores de la revista *Plant Cell*. Los genotipos completos, como *api6-1/rps24b-2*, en los que los alelos de un gen en cromosomas homólogos se separan con una barra, se han utilizado únicamente cuando fue imprescindible. Salvo que se indique lo contrario, los individuos que se describen en este trabajo son homocigóticos para la mutación que se menciona en cada caso. Hemos utilizado en algunos casos un punto y coma como separador entre mutaciones no alélicas.

Las estirpes de Arabidopsis, los procedimientos para su manipulación y las condiciones de cultivo usados en esta Tesis se describen en las páginas 49 y 109. Hemos identificado mutaciones puntuales mediante análisis del ligamiento a marcadores moleculares (en las páginas 49 y 109). Hemos realizado análisis histológicos y morfométricos y de microscopía confocal de los mutantes a estudio (en las páginas 49 y 110). Hemos aislado ARN para su retrotranscripción seguida de PCR cuantitativa (RT-qPCR), semicuantitativa y

para su secuenciación masiva (en las páginas 49, 50 y 110). Hemos construido fusiones traduccionales con el gen de la proteína fluorescente verde (en la página 108). Hemos realizado análisis de tipo *western* y *northern* (en las páginas 49 y 110), así como de hibridación *in situ* fluorescente de ARN e inmunolocalización (en las páginas 50 y 111). Hemos identificado interactores físicos mediante coimmunoprecipitación y su posterior análisis por cromatografía líquida con espectrometría de masas (en la página 111). También hemos realizado alineamientos de secuencias aminoacídicas, así como análisis bioinformáticos del *splicing* alternativo y la visualización tridimensional de proteínas cristalografiadas (en la página 50).

VI.- RESULTADOS Y DISCUSIÓN

VI.- RESULTADOS Y DISCUSIÓN

El alelo *ago1-52* del gen *AGO1* de *Arabidopsis* es hipomorfo y viable y fue aislado en el laboratorio de José Luis Micol y caracterizado en el de María Rosa Ponce. El mutante *ago1-52* es portador de una mutación puntual que crea un 3'SS nuevo, que es elegido por el espliceosoma más frecuentemente que el genuino, del que dista 10 nucleótidos. Como consecuencia, en el mutante *ago1-52* coexisten dos variantes del ARNm del gen *AGO1*, una de ellas minoritaria y de secuencia silvestre, y otra mayoritaria y mutante, que es 10 nucleótidos más larga que la silvestre. La traducción de estos dos ARNm rinde dos proteínas: una silvestre y otra mutante y mayoritaria.

Las mutaciones *mas5* son supresores extragénicos del fenotipo morfológico del mutante *ago1-52*. Se identificaron en una mutagénesis de segundos sitios realizada en el laboratorio de María Rosa Ponce, anterior al comienzo de esta Tesis Doctoral. Se estableció mediante análisis iterativo del ligamiento a marcadores moleculares que las mutaciones *mas5* son alelos del gen *PRP8*, que codifica el factor central del espliceosoma (página 50). Todas las mutaciones *mas5* resultaron ser de cambio de sentido, y al menos dos de ellas, dominantes en cuanto a su efecto supresor (Micol-Ponce *et al.*, 2014).

Hemos establecido que todas las mutaciones *mas5* causan la sustitución de un aminoácido por otro cargado positivamente, o de uno con carga negativa por otro sin ella (página 52); mediante su reconstrucción tridimensional *in silico*, hemos establecido que estas sustituciones afectan a algunos de los dominios que forman parte de una cavidad de la proteína PRP8 implicada directamente en el reconocimiento de los 5'SS y 3'SS y el sitio de ramificación de los pre-ARNm. También dañan esta cavidad casi todas las mutaciones del gen *PRP8* con efectos supresores descritas en otras especies, como la levadura (Grainger y Beggs, 2005).

Hemos obtenido combinaciones dobles mutantes de *mas5-1* con cuatro mutaciones puntuales e hipomorfias del gen *AGO1*, una de las cuales (la antes mencionada *ago1-52*) crea un 3'SS nuevo, otras dos (*ago1-25* y *ago1-27*) causan sustituciones de aminoácidos sin alterar el *splicing* del pre-ARNm del gen *AGO1*, y la cuarta (*ago1-51*) elimina un 5'SS genuino. Solo hemos observado supresión del fenotipo morfológico de *ago1-52* en los dobles mutante *mas5 ago1-52* (página 53). Hemos comprobado que en el doble mutante *mas5-1 ago1-52* se incrementa la concentración relativa del ARNm *AGO1* y la proteína *AGO1* silvestres, respecto a lo que se observa en el mutante simple *ago1-52*. Hemos establecido que este fenotipo molecular del doble mutante se debe al incremento del uso por el espliceosoma del 3'SS genuino frente al nuevo creado por la mutación *ago1-52* (página 53).

También hemos combinado *mas5-1* con alelos mutantes de genes no relacionados ni con *AGO1* ni con el *splicing*, cuyas mutaciones eliminan un 5'SS (*ago1-51* y *sca3-1*) o un 3'SS (*anu4-1* y *ang1-2*) genuinos, o crean un nuevo 5'SS (*icu13-1*). Su estudio indica que la mutación *mas5-1* suprime el fenotipo morfológico y molecular de *icu13-1*, al incrementar el uso por el espliceosoma del 5'SS genuino en las plantas *mas5-1 icu13-1* (página 53). Sin embargo, *mas5-1* no suprime el fenotipo morfológico de los mutantes *ago1-51*, *ang1-2*, *anu4-1* y *sca3-1*, cuyas mutaciones dañan los 5'SS o 3'SS (página 56). Considerados en conjunto, estos resultados indican que las mutaciones *mas5* incrementan la fidelidad del *splicing* al reducir parcialmente el uso por el espliceosoma de los 5'SS y 3'SS nuevos generados por una mutación puntual.

Hemos demostrado también que *mas5-1* y *mas5-3* no alteran globalmente el *splicing*, a diferencia de su alelo hipomorfo *prp8-7* (Sasaki et al., 2015). En efecto, *mas5-1* y *mas5-3* solo afectan al 0,1% del total de los transcritos que hemos analizado mediante secuenciación masiva de ARN, principalmente relacionados con eventos de retención intrónica (página 56), mientras que en el mutante *prp8-7* este tipo de alteraciones del *splicing* sucede en el 6,7% de los intrones.

Concluimos que nuestros alelos *mas5* del gen *PRP8* de Arabidopsis incrementan la fidelidad del *splicing* al favorecer específicamente el uso de los 5'SS y 3'SS genuinos frente a los alternativos en secuencias NAG en tándem, o a los nuevos, generados por mutaciones puntuales. Los alelos *mas5* de *PRP8* no alteran globalmente el *splicing*, a diferencia de los alelos hipomorfos de este gen (Sasaki et al., 2015).

También hemos caracterizado en esta Tesis los genes parálogos *RPS24A* y *RPS24B* de Arabidopsis, empleando para ellos dos mutantes insercionales (*rps24b-2* y *rps24a-1*) y uno portador de una mutación puntual inducida mediante EMS (*api6*, otro alelo de *RPS24B*; en la página 88). Hemos intentado combinar *rps24b-2* y *rps24a-1* y solo hemos obtenido diheterocigotos; la ausencia de dobles mutantes y sesquimutantes revela la existencia de no complementación no alélica entre estos dos genes e indica no solo su alto grado de redundancia funcional sino también su dependencia de dosis. En otras palabras, *RPS24A* y *RPS24B* manifiestan haploinsuficiencia combinada, ya que se necesitan al menos dos copias silvestres de alguno de los dos para la viabilidad de la planta, y tres para que su fenotipo sea silvestre (página 90). Hemos localizado *RPS24B* principalmente en el nucleolo, pero también en el nucleoplasma y el citoplasma, mediante la obtención de una fusión traduccional *RPS24B:GFP* y su transferencia a plantas Col-0 (página 90), confirmando así los resultados previamente obtenidos en diferentes análisis de los proteomas nuclear y nucleolar (Pendle et al., 2005; Montacié et al., 2017; Palm et al., 2019; Ayash et al., 2021).

La RPS24 humana y la de *Saccharomyces cerevisiae* son factores de la biogénesis del ribosoma, cuya pérdida de función causa la acumulación de precursores del ARNr 18S y la disminución de los niveles del ARNr 18S maduro (Ferreira-Cerca *et al.*, 2005; Choismel *et al.*, 2008; Robledo *et al.*, 2008). Hemos constatado que los mutantes *rps24a-1*, *api6* y *rps24b-2* acumulan el pre-ARNr P-A₃, un precursor del ARNr 18S, y que sus niveles del ARNr 18S maduro son inferiores a los del tipo silvestre (página 92).

Hemos obtenido dobles mutantes combinando *rps24b-2* y alelos de genes que codifican otros factores de la biogénesis del ribosoma. Los fenotipos morfológicos de los dobles mutantes de *rps24b-2* con *smo4-3*, *mtr4-2*, *par11-2* y *rrp7-1* resultaron sinérgicos (página 96). Los niveles de los precursores del ARNr 5,8S fueron más altos en los dobles mutantes *rps24b-2 smo4-3* y *rps24b-2 mtr4-2* (página 96) que en los mutantes simples *smo4-3* y *mtr4-2*. Este fue un resultado inesperado, ya que en el mutante simple *rps24b-2* no se acumula ningún pre-ARNr 5,8S. Hemos observado hipertrofia del nucleolo en el mutante simple *rps24b-2* y los dobles mutantes *rps24b-2 mtr4-2* y *rps24b-2 smo4-3*, que atribuimos al estrés nucleolar que deben causar las alteraciones de la maduración de los ARNr (página 98). Hemos constatado mediante RT-PCR y RT-qPCR (página 98) un aumento en los niveles de transcripción del ADNr 45S en los mutantes *rps24*; esta observación refuerza la hipótesis de que la acumulación del pre-ARNr 5,8S en los dobles mutantes *rps24b-2 smo4-3* y *rps24b-2 mtr4-2* es consecuencia del incremento de la transcripción de los genes ADNr 45S, y además demuestra que RPS24A y RPS24B regulan la transcripción del ADNr 45S. Por último, hemos constatado que varios factores de la biogénesis del ribosoma coimmunoprecipitan con la proteína de fusión RPS24B-GFP (página 100), observación que también apoya la hipótesis de la actuación de RPS24A y RPS24B como factores de la biogénesis del ribosoma en *Arabidopsis*.

En conclusión, hemos obtenido resultados que indican que las proteínas redundantes RPS24A y RPS24B de *Arabidopsis* presentan un alto grado de conservación funcional con sus ortólogas humana y de la levadura, ya que no solo actúan en estas tres especies como componentes estructurales de la subunidad 40S del ribosoma, sino también como factores de la biogénesis del ribosoma, interviniendo en la maduración del ARNr 18S. La participación de RPS24A y RPS24B en la regulación de la transcripción del ADNr 45S, sin embargo, no se ha descrito para sus ortólogas humana y de la levadura, lo que supondría una divergencia evolutiva.

VII.- CONCLUSIONES Y PERSPECTIVAS

VII.- CONCLUSIONES Y PERSPECTIVAS

VII.1.- Los alelos *mas5* del gen *PRP8* incrementan la fidelidad del *splicing*

PRP8 es el componente central del espliceosoma y está muy conservado en los eucariotas. Esta proteína reconoce los 5'SS, los 3'SS y los sitios de ramificación de los intrones, e interviene en el ensamblaje y el reciclaje del espliceosoma. La mayoría de los alelos mutantes viables del gen *PRP8* previamente identificados en los eucariotas son hipomorfos y alteran globalmente el *splicing*, causando fundamentalmente un incremento en los eventos de retención intrónica. Hemos caracterizado en esta Tesis seis alelos *mas5* del gen *PRP8* de *Arabidopsis*, que incrementan la fidelidad del *splicing*, reduciendo el uso por el espliceosoma de los 5'SS o 3'SS nuevos creados por mutaciones puntuales, como los de los mutantes *ago1-52* e *icu13-1*. También hemos constatado que en los mutantes *mas5* se reduce el uso por el espliceosoma de los 3'SS alternativos, que es poco frecuente en la estirpe silvestre Col-0; estos 3'SS alternativos forman parte de grupos de trinucleótidos NAG dispuestos en tándem en el extremo 3' de muchos intrones. Hemos demostrado que los alelos *mas5* de *PRP8* favorecen el uso del 3'SS proximal (el más cercano al extremo 5' del grupo de trinucleótidos NAG en tándem), que es el más usado por la estirpe silvestre.

Desconocemos el mecanismo molecular por el que los alelos *mas5* incrementan la fidelidad del *splicing*. Nuestro análisis *in silico* de la estructura tridimensional de PRP8 indica que los aminoácidos afectados por las mutaciones *mas5* se encuentran en regiones muy cercanas al sitio activo de esta proteína, que interacciona directamente con el pre-ARNm. Se requerirán análisis adicionales para interpretar los efectos de las mutaciones *mas5* sobre la estructura secundaria de las proteínas PRP8 mutantes. Consideramos verosímil que estas mutaciones modifiquen la estructura secundaria de la cavidad de PRP8 y, en consecuencia, sus interacciones con los pre-ARNm y/o con las proteínas que conforman los distintos complejos del espliceosoma.

Se conocen mutaciones puntuales que dañan intrones y están asociadas a enfermedades humanas, como el cáncer. Se sabe que muchas de ellas, que habían pasado desapercibidas al no afectar a regiones codificantes, generan 5'SS o 3'SS nuevos, cuyo uso por el espliceosoma rinde proteínas aberrantes, usualmente truncadas. *Arabidopsis* podría ser muy útil como sistema modelo para el estudio de la base molecular de estas enfermedades. *Saccharomyces cerevisiae* dista mucho de ser el mejor sistema para este fin, dado que la mayoría de sus genes carecen de intrones y su *splicing* alternativo es casi inexistente. Dado que actualmente es fácil inducir mutaciones mediante CRISPR/Cas, la obtención en líneas celulares de alelos del gen *PRP8* humano equivalentes a los *mas5* de *Arabidopsis* podría ser de gran utilidad para el estudio y la eventual terapia de enfermedades

causadas por defectos en el *splicing* de genes concretos, ya que podrían suprimir sus efectos deletéreos sin perturbar globalmente el *splicing*.

VII.2.- RPS24A y RPS24B son factores de la biogénesis del ribosoma

El análisis de la composición de los ribosomas eucarióticos 80S revela su heterogeneidad histotípica, que se ha relacionado con su especialización para traducir determinadas subpoblaciones de ARNm. Esta heterogeneidad estructural es fruto de la existencia de familias génicas de parálogos que codifican proteínas ribosómicas, un fenómeno que se manifiesta particularmente en el reino vegetal, como consecuencia de la tolerancia de las plantas a las duplicaciones génicas y genómicas. Los genes *RPS24A* y *RPS24B* de *Arabidopsis* presentan patrones de expresión espaciotemporales similares y codifican proteínas casi idénticas, lo que sugiere su redundancia funcional. Nuestro análisis de estos genes ha confirmado experimentalmente esta hipótesis y ha revelado su dependencia de dosis: el requerimiento de al menos dos copias silvestres de alguno de los dos parálogos para la viabilidad de *Arabidopsis* y de al menos tres para su desarrollo normal. Casos de haploinsuficiencia combinada como el de *RPS24A* y *RPS24B* se habían descrito previamente para otras familias de proteínas ribosómicas de *Arabidopsis*, como las de *RPS6*, *RPL5*, *RPL23* y *RPL36*.

Las duplicaciones génicas favorecen la diversificación funcional del proteoma, y en el caso de los genes que codifican proteínas ribosómicas, que algunas de ellas adquieran funciones extrarribosómicas, una de las cuales es actuar como factores de la biogénesis del ribosoma. Nuestros análisis moleculares de los mutantes simples *rps24a* y *rps24b* y sus dobles mutantes con alelos de otros genes implicados en el procesamiento del pre-ARNr 45S han demostrado que *RPS24A* y *RPS24B* actúan como factores de la biogénesis del ribosoma, concretamente en la maduración del ARNr 18S, tal como hacen sus ortólogos humano y de la levadura. Hemos constatado también que *RPS24A* y *RPS24B* regulan negativamente la expresión del ADNr 45S. Esta observación se ve confirmada por los fenotipos sinérgicos de los dobles mutantes de *rps24a* o *rps24b* con alelos de pérdida de función de *RRP7* y *PARL1*, que codifican represores de la transcripción del ADNr 45S previamente conocidos. En consecuencia, *RPS24B* tendría un doble papel en la biogénesis del ribosoma, actuando primero en la regulación transcripcional de los genes del ADNr 45S, y posteriormente, en la maduración de los ARNr. La implicación de *RPS24A* y *RPS24B* en la biogénesis del ARNr 18S revela un alto grado de conservación funcional con sus ortólogos humano y de la levadura. El papel que parecen jugar *RPS24A* y *RPS24B* en la regulación de la transcripción del ADNr 45S no ha sido descrito para ninguno de sus ortólogos.

**VIII.- BIBLIOGRAFÍA
DE LOS APARTADOS IV-VII**

VIII.- BIBLIOGRAFÍA DE LOS APARTADOS IV-VII

- Absmeier, E., Santos, K.F., y Wahl, M.C. (2016). Functions and regulation of the Brr2 RNA helicase during splicing. *Cell Cycle* **15**, 3362-3377.
- Alter, B.P., Giri, N., Savage, S.A., y Rosenberg, P.S. (2018). Cancer in the National Cancer Institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. *Haematologica* **103**, 30-39.
- Ayash, M., Abukhalaf, M., Thieme, D., Proksch, C., Heilmann, M., Schattat, M.H., y Hoehenwarter, W. (2021). LC-MS based draft map of the *Arabidopsis thaliana* nuclear proteome and protein import in pattern triggered immunity. *Frontiers in Plant Science* **12**, 744103.
- Barakat, A., Szick-Miranda, K., Chang, I.F., Guyot, R., Blanc, G., Cooke, R., Delseny, M., y Bailey-Serres, J. (2001). The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. *Plant Physiology* **127**, 398-415.
- Barba-Aliaga, M., Alepuz, P., y Pérez-Ortín, J.E. (2021). Eukaryotic RNA polymerases: the many ways to transcribe a gene. *Frontiers in Molecular Biosciences* **8**, 663209.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233.
- Baudin-Baillieu, A., Tollervey, D., Cullin, C., y Lacroute, F. (1997). Functional analysis of Rrp7p, an essential yeast protein involved in pre-rRNA processing and ribosome assembly. *Molecular and Cellular Biology* **17**, 5023-5032.
- Behm-Ansmant, I., Kashima, I., Rehwinkel, J., Sauliere, J., Wittkopp, N., y Izaurralde, E. (2007). mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons. *FEBS Letters* **581**, 2845-2853.
- Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. *Nature Reviews Genetics* **15**, 163-175.
- Berná, G., Robles, P., y Micol, J.L. (1999). A mutational analysis of leaf morphogenesis in *Arabidopsis thaliana*. *Genetics* **152**, 729-742.
- Björk, P., y Wieslander, L. (2017). Integration of mRNP formation and export. *Cellular and Molecular Life Sciences* **74**, 2875-2897.
- Boutz, P.L., Bhutkar, A., y Sharp, P.A. (2015). Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes and Development* **29**, 63-80.
- Bradley, R.K., Merkin, J., Lambert, N.J., y Burge, C.B. (2012). Alternative splicing of RNA triplets is often regulated and accelerates proteome evolution. *PLOS Biology* **10**, e1001229.
- Brown, J.W., Smith, P., y Simpson, C.G. (1996). Arabidopsis consensus intron sequences. *Plant Molecular Biology* **32**, 531-535.
- Campell, B.R., Song, Y., Posch, T.E., Cullis, C.A., y Town, C.D. (1992). Sequence and organization of 5S ribosomal RNA-encoding genes of *Arabidopsis thaliana*. *Gene* **112**, 225-228.
- Carroll, A.J., Heazlewood, J.L., Ito, J., y Millar, A.H. (2008). Analysis of the *Arabidopsis* cytosolic ribosome proteome provides detailed insights into its components and their post-translational modification. *Molecular and Cellular Proteomics* **7**, 347-369.
- Carroll, A.J. (2013). The *Arabidopsis* cytosolic ribosomal proteome: from form to function. *Frontiers in Plant Science* **4**, 32.
- Carvalho, C.M., Santos, A.A., Pires, S.R., Rocha, C.S., Saraiva, D.I., Machado, J.P., Mattos, E.C., Fietto, L.G., y Fontes, E.P. (2008). Regulated nuclear trafficking of rpL10A mediated by NIK1 represents a defense strategy of plant cells against virus. *PLOS Pathogens* **4**, e1000247.
- Casanova-Sáez, R., Candela, H., y Micol, J.L. (2014). Combined haploinsufficiency and purifying selection drive retention of *RPL36a* paralogs in Arabidopsis. *Scientific Reports* **4**, 4122.
- Chandrasekhara, C., Mohannath, G., Blevins, T., Pontvianne, F., y Pikaard, C.S. (2016). Chromosome-specific NOR inactivation explains selective rRNA gene silencing and dosage control in *Arabidopsis*. *Genes and Development* **30**, 177-190.

- Chen, Z.J., y Pikaard, C.S. (1997). Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation and histone modification in nucleolar dominance. *Genes and Development* **11**, 2124-2136.
- Choesmel, V., Fribourg, S., Aguisa-Touré, A.H., Pinaud, N., Legrand, P., Gazda, H.T., y Gleizes, P.E. (2008). Mutation of ribosomal protein RPS24 in Diamond-Blackfan anemia results in a ribosome biogenesis disorder. *Human Molecular Genetics* **17**, 1253-1263.
- Cloix, C., Yukawa, Y., Tutois, S., Sugiura, M., y Tourmente, S. (2003). *In vitro* analysis of the sequences required for transcription of the *Arabidopsis thaliana* 5S rRNA genes. *Plant Journal* **35**, 251-261.
- Comella, P., Pontvianne, F., Lahmy, S., Vignols, F., Barbezier, N., Debures, A., Jobet, E., Brugidou, E., Echeverria, M., y Sáez-Vásquez, J. (2008). Characterization of a ribonuclease III-like protein required for cleavage of the pre-rRNA in the 32ETS in *Arabidopsis*. *Nucleic Acids Research* **36**, 1163-1175.
- Conti, E., y Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Current Opinion in Cell Biology* **17**, 316-325.
- Copenhaver, G.P., y Pikaard, C.S. (1996a). Two-dimensional RFLP analyses reveal megabase-sized clusters of rRNA gene variants in *Arabidopsis thaliana*, suggesting local spreading of variants as the mode for gene homogenization during concerted evolution. *Plant Journal* **9**, 273-282.
- Copenhaver, G.P., y Pikaard, C.S. (1996b). RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of *Arabidopsis thaliana* adjoin the telomeres on chromosomes 2 and 4. *Plant Journal* **9**, 259-272.
- Creff, A., Sormani, R., y Desnos, T. (2010). The two *Arabidopsis RPS6* genes, encoding for cytoplasmic ribosomal proteins S6, are functionally equivalent. *Plant Molecular Biology* **73**, 533-546.
- Daguenet, E., Dujardin, G., y Valcárcel, J. (2015). The pathogenicity of splicing defects: mechanistic insights into pre-mRNA processing inform novel therapeutic approaches. *EMBO Reports* **16**, 1640-1655.
- De Keersmaecker, K., Sulima, S.O., y Dinman, J.D. (2015). Ribosomopathies and the paradox of cellular hypo- to hyperproliferation. *Blood* **125**, 1377-1382.
- de la Cruz, J., Karbstein, K., y Woolford, J.L., Jr. (2015). Functions of ribosomal proteins in assembly of eukaryotic ribosomes in vivo. *Annual Review of Biochemistry* **84**, 93-129.
- Degenhardt, R.F., y Bonham-Smith, P.C. (2008). *Arabidopsis* ribosomal proteins RPL23aA and RPL23aB are differentially targeted to the nucleolus and are disparately required for normal development. *Plant Physiology* **147**, 128-142.
- Ding, Z., Meng, Y.R., Fan, Y.J., y Xu, Y.Z. (2022). Roles of minor spliceosome in intron recognition and the convergence with the better understood major spliceosome. *Wiley Interdisciplinary Reviews RNA* **14**, e1761.
- Doelling, J.H., y Pikaard, C.S. (1995). The minimal ribosomal RNA gene promoter of *Arabidopsis thaliana* includes a critical element at the transcription initiation site. *Plant Journal* **8**, 683-692.
- Doudna, J.A., y Rath, V.L. (2002). Structure and function of the eukaryotic ribosome: the next frontier. *Cell* **109**, 153-156.
- Douet, J., y Tourmente, S. (2007). Transcription of the 5S rRNA heterochromatic genes is epigenetically controlled in *Arabidopsis thaliana* and *Xenopus laevis*. *Heredity* **99**, 5-13.
- Dreyfus, M., y Régnier, P. (2002). The poly(A) tail of mRNAs: bodyguard in eukaryotes, scavenger in bacteria. *Cell* **111**, 611-613.
- Earley, K., Lawrence, R.J., Pontes, O., Reuther, R., Enciso, A.J., Silva, M., Neves, N., Gross, M., Viegas, W., y Pikaard, C.S. (2006). Erasure of histone acetylation by *Arabidopsis* HDA6 mediates large-scale gene silencing in nucleolar dominance. *Genes and Development* **20**, 1283-1293.
- Earley, K.W., Pontvianne, F., Wierzbicki, A.T., Blevins, T., Tucker, S., Costa-Nunes, P., Pontes, O., y Pikaard, C.S. (2010). Mechanisms of HDA6-mediated rRNA gene silencing: suppression of intergenic Pol II transcription and differential effects on maintenance versus siRNA-directed cytosine methylation. *Genes and Development* **24**, 1119-1132.

- Ehrnsberger, H.F., Grasser, M., y Grasser, K.D. (2019). Nucleocytoplasmic mRNA transport in plants: export factors and their influence on growth and development. *Journal of Experimental Botany* **70**, 3757-3763.
- Elhamamsy, A.R., Metge, B.J., Alsheikh, H.A., Shevde, L.A., y Samant, R.S. (2022). Ribosome biogenesis: a central player in cancer metastasis and therapeutic resistance. *Cancer Research* **82**, 2344-2353.
- Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., y Luhrmann, R. (2009). The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Molecular Cell* **36**, 593-608.
- Ferreira-Cerca, S., Poll, G., Gleizes, P.E., Tschochner, H., y Milkereit, P. (2005). Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. *Molecular Cell* **20**, 263-275.
- Ferreira, M.L.F., Biarc, J., Burlingame, A.L., y Casati, P. (2010a). Arabidopsis L10 ribosomal proteins in UV-B responses. *Plant Signaling and Behavior* **5**, 1222-1225.
- Ferreira, M.L.F., Pezza, A., Biarc, J., Burlingame, A.L., y Casati, P. (2010b). Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B Stress. *Plant Physiology* **153**, 1878-1894.
- Franz, P., Armstrong, S., Alonso-Blanco, C., Fischer, T.C., Torres-Ruiz, R.A., y Jones, G. (1998). Cytogenetics for the model system *Arabidopsis thaliana*. *Plant Journal* **13**, 867-876.
- Franz, P., De Jong, J.H., Lysak, M., Castiglione, M.R., y Schubert, I. (2002). Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. *Proceedings of the National Academy of Sciences USA* **99**, 14584-14589.
- Fujikura, U., Horiguchi, G., Ponce, M.R., Micol, J.L., y Tsukaya, H. (2009). Coordination of cell proliferation and cell expansion mediated by ribosome-related processes in the leaves of *Arabidopsis thaliana*. *Plant Journal* **59**, 499-508.
- Fuke, H., y Ohno, M. (2008). Role of poly (A) tail as an identity element for mRNA nuclear export. *Nucleic Acids Research* **36**, 1037-1049.
- Galej, W.P., Nguyen, T.H., Newman, A.J., y Nagai, K. (2014). Structural studies of the spliceosome: zooming into the heart of the machine. *Current Opinion in Structural Biology* **25**, 57-66.
- Goldfarb, K.C., y Cech, T.R. (2017). Targeted CRISPR disruption reveals a role for RNase MRP RNA in human preribosomal RNA processing. *Genes and Development* **31**, 59-71.
- Grainger, R.J., y Beggs, J.D. (2005). Prp8 protein: at the heart of the spliceosome. *RNA* **11**, 533-557.
- Granato, D.C., Gonzales, F.A., Luz, J.S., Cassiola, F., Machado-Santelli, G.M., y Oliveira, C.C. (2005). Nop53p, an essential nucleolar protein that interacts with Nop17p and Nip7p, is required for pre-rRNA processing in *Saccharomyces cerevisiae*. *FEBS Journal* **272**, 4450-4463.
- Graveley, B.R. (2001). Alternative splicing: increasing diversity in the proteomic world. *Trends in Genetics* **17**, 100-107.
- Gruendler, P., Unfried, I., Pascher, K., y Schweizer, D. (1991). rDNA intergenic region from *Arabidopsis thaliana*. Structural analysis, intraspecific variation and functional implications. *Journal of Molecular Biology* **221**, 1209-1222.
- Hang, R., Wang, Z., Yang, C., Luo, L., Mo, B., Chen, X., Sun, J., Liu, C., y Cao, X. (2021). Protein arginine methyltransferase 3 fine-tunes the assembly/disassembly of pre-ribosomes to repress nucleolar stress by interacting with RPS2B in *arabidopsis*. *Molecular Plant* **14**, 223-236.
- Hawley, R.S., y Gilliland, W.D. (2006). Sometimes the result is not the answer: the truths and the lies that come from using the complementation test. *Genetics* **174**, 5-15.
- Heath, C.G., Viphakone, N., y Wilson, S.A. (2016). The role of TREX in gene expression and disease. *Biochemical Journal* **473**, 2911-2935.
- Henras, A.K., Plisson-Chastang, C., O'Donohue, M.F., Chakraborty, A., y Gleizes, P.E. (2015). An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdisciplinary Reviews RNA* **6**, 225-242.

- Horiguchi, G., Mollá-Morales, A., Pérez-Pérez, J.M., Kojima, K., Robles, P., Ponce, M.R., Micol, J.L., y Tsukaya, H. (2011). Differential contributions of ribosomal protein genes to *Arabidopsis thaliana* leaf development. *Plant Journal* **65**, 724-736.
- Hummel, M., Cordewener, J.H., de Groot, J.C., Smeeckens, S., America, A.H., y Hanson, J. (2012). Dynamic protein composition of *Arabidopsis thaliana* cytosolic ribosomes in response to sucrose feeding as revealed by label free MS^E proteomics. *Proteomics* **12**, 1024-1038.
- Iida, K., Shionyu, M., y Suso, Y. (2008). Alternative splicing at NAGNAG acceptor sites shares common properties in land plants and mammals. *Molecular Biology and Evolution* **25**, 709-718.
- Jover-Gil, S., Candela, H., Robles, P., Aguilera, V., Barrero, J.M., Micol, J.L., y Ponce, M.R. (2012). The microRNA pathway genes *AGO1*, *HEN1* and *HYL1* participate in leaf proximal-distal, venation and stomatal patterning in *Arabidopsis*. *Plant and Cell Physiology* **53**, 1322-1333.
- Kampen, K.R., Sulima, S.O., Vereecke, S., y De Keersmaecker, K. (2020). Hallmarks of ribosomopathies. *Nucleic Acids Research* **48**, 1013-1028.
- Kang, C.H., Lee, Y.M., Park, J.H., Nawkar, G.M., Oh, H.T., Kim, M.G., Lee, S.I., Kim, W.Y., Yun, D.J., y Lee, S.Y. (2016). Ribosomal P3 protein AtP3B of *Arabidopsis* acts as both protein and RNA chaperone to increase tolerance of heat and cold stresses. *Plant, Cell and Environment* **39**, 1631-1642.
- Katahira, J. (2012). mRNA export and the TREX complex. *Biochimica et Biophysica Acta* **1819**, 507-513.
- Keren, H., Lev-Maor, G., y Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. *Nature Reviews Genetics* **11**, 345-355.
- Kim, Y.K., Kim, S., Shin, Y.J., Hur, Y.S., Kim, W.Y., Lee, M.S., Cheon, C.I., y Verma, D.P. (2014). Ribosomal protein S6, a target of rapamycin, is involved in the regulation of rRNA genes by possible epigenetic changes in *Arabidopsis*. *Journal of Biological Chemistry* **289**, 3901-3912.
- Koncz, C., Dejong, F., Villacorta, N., Szakonyi, D., y Koncz, Z. (2012). The spliceosome-activating complex: molecular mechanisms underlying the function of a pleiotropic regulator. *Frontiers in Plant Science* **3**, 9.
- Kressler, D., Hurt, E., y Baßler, J. (2017). A puzzle of life: crafting ribosomal subunits. *Trends in Biochemical Sciences* **42**, 640-654.
- Laferté, A., Favry, E., Sentenac, A., Riva, M., Carles, C., y Chédin, S. (2006). The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. *Genes and Development* **20**, 2030-2040.
- Lareau, L.F., Green, R.E., Bhatnagar, R.S., y Brenner, S.E. (2004). The evolving roles of alternative splicing. *Current Opinion in Structural Biology* **14**, 273-282.
- Layat, E., Sáez-Vásquez, J., y Tourmente, S. (2012). Regulation of Pol I-transcribed 45S rDNA and Pol III-transcribed 5S rDNA in *Arabidopsis*. *Plant and Cell Physiology* **53**, 267-276.
- Le Hir, H., Gatfield, D., Izaurralde, E., y Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO Journal* **20**, 4987-4997.
- Li, J., y Liu, C. (2019). Coding or noncoding, the converging concepts of RNAs. *Frontiers in Genetics* **10**, 496.
- Li, S., Liu, K., Zhang, S., Wang, X., Rogers, K., Ren, G., Zhang, C., y Yu, B. (2017). STV1, a ribosomal protein, binds primary microRNA transcripts to promote their interaction with the processing complex in *Arabidopsis*. *Proceedings of the National Academy of Sciences USA* **114**, 1424-1429.
- Li, Z., Li, W., Guo, M., Liu, S., Liu, L., Yu, Y., Mo, B., Chen, X., y Gao, L. (2022). Origin, evolution and diversification of plant ARGONAUTE proteins. *Plant Journal* **109**, 1086-1097.
- Lim, K.H., Ferraris, L., Filloux, M.E., Raphael, B.J., y Fairbrother, W.G. (2011). Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. *Proceedings of the National Academy of Sciences USA* **108**, 11093-11098.

- Lustig, A.J., Lin, R.J., y Abelson, J. (1986). The yeast RNA gene products are essential for mRNA splicing in vitro. *Cell* **47**, 953-963.
- Lysak, M.A., Koch, M.A., Pecinka, A., y Schubert, I. (2005). Chromosome triplication found across the tribe *Brassicaceae*. *Genome Research* **15**, 516-525.
- Maekawa, S., Ishida, T., y Yanagisawa, S. (2018). Reduced expression of *APUM24*, encoding a novel rRNA processing factor, induces sugar-dependent nucleolar stress and altered sugar responses in *Arabidopsis thaliana*. *Plant Cell* **30**, 209-227.
- Marquez, Y., Brown, J.W., Simpson, C., Barta, A., y Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in *Arabidopsis*. *Genome Research* **22**, 1184-1195.
- McKie, A.B., McHale, J.C., Keen, T.J., Tarttelin, E.E., Goliath, R., van Lith-Verhoeven, J.J., Greenberg, J., Ramesar, R.S., Hoyng, C.B., Cremers, F.P., Mackey, D.A., Bhattacharya, S.S., Bird, A.C., Markham, A.F., y Inglehearn, C.F. (2001). Mutations in the pre-mRNA splicing factor gene *PRPC8* in autosomal dominant retinitis pigmentosa (RP13). *Human Molecular Genetics* **10**, 1555-1562.
- Meier, I., Richards, E.J., y Evans, D.E. (2017). Cell biology of the plant nucleus. *Annual Review of Plant Biology* **68**, 139-172.
- Meinke, D.W., y Koornneef, M. (1997). Community standards for *Arabidopsis* genetics. *Plant Journal* **12**, 247-253.
- Meinke, D.W. (2013). A survey of dominant mutations in *Arabidopsis thaliana*. *Trends in Plant Science* **18**, 84-91.
- Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., y Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. *Nature Structural and Molecular Biology* **19**, 560-567.
- Merchante, C., Stepanova, A.N., y Alonso, J.M. (2017). Translation regulation in plants: an interesting past, an exciting present and a promising future. *Plant Journal* **90**, 628-653.
- Micol-Ponce, R., Aguilera, V., y Ponce, M.R. (2014). A genetic screen for suppressors of a hypomorphic allele of *Arabidopsis ARGONAUTE1*. *Scientific Reports* **4**, 5533.
- Micol-Ponce, R., Sarmiento-Mañús, R., Ruiz-Bayón, A., Montacie, C., Sáez-Vásquez, J., y Ponce, M.R. (2018). *Arabidopsis* RIBOSOMAL RNA PROCESSING7 is required for 18S rRNA maturation. *Plant Cell* **30**, 2855-2872.
- Micol-Ponce, R., Sarmiento-Mañús, R., Fontcuberta-Cervera, S., Cabezas-Fuster, A., de Bures, A., Sáez-Vásquez, J., y Ponce, M.R. (2020). SMALL ORGAN4 is a ribosome biogenesis factor involved in 5.8S ribosomal RNA maturation. *Plant Physiology* **184**, 2022-2039.
- Missbach, S., Weis, B.L., Martin, R., Simm, S., Bohnsack, M.T., y Schleiff, E. (2013). 40S ribosome biogenesis co-factors are essential for gametophyte and embryo development. *PLoS One* **8**, e54084.
- Montacié, C., Durut, N., Opsomer, A., Palm, D., Comella, P., Picart, C., Carpentier, M.C., Pontvianne, F., Carapito, C., Schleiff, E., y Sáez-Vásquez, J. (2017). Nucleolar proteome analysis and proteasomal activity assays reveal a link between nucleolus and 26S proteasome in *A. thaliana*. *Frontiers in Plant Science* **8**, 1815.
- Mort, M., Sterne-Weiler, T., Li, B., Ball, E.V., Cooper, D.N., Radivojac, P., Sanford, J.R., y Mooney, S.D. (2014). MutPred Splice: machine learning-based prediction of exonic variants that disrupt splicing. *Genome Biology* **15**, R19.
- Murata, M., Heslop-Harrison, J.S., y Motoyoshi, F. (1997). Physical mapping of the 5S ribosomal RNA genes in *Arabidopsis thaliana* by multi-color fluorescence *in situ* hybridization with cosmid clones. *Plant Journal* **12**, 31-37.
- Narla, A., y Ebert, B.L. (2010). Ribosomopathies: human disorders of ribosome dysfunction. *Blood* **115**, 3196-3205.
- Navarro-Quiles, C., Lup, S.D., Muñoz-Nortes, T., Candela, H., y Micol, J.L. (2023). The genetic and molecular basis of haploinsufficiency in flowering plants *Trends in Plant Science* en prensa.

- Ni, C., y Buszczak, M. (2023). The homeostatic regulation of ribosome biogenesis. *Seminars in Cell and Developmental Biology* **136**, 13-26.
- Nicholson, P., Yepiskoposyan, H., Metze, S., Zamudio Orozco, R., Kleinschmidt, N., y Muhlemann, O. (2010). Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cellular and Molecular Life Sciences* **67**, 677-700.
- O'Brien, J., Hayder, H., Zayed, Y., y Peng, C. (2018). Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in Endocrinology* **9**, 402.
- Padgett, R.A. (2012). New connections between splicing and human disease. *Trends in Genetics* **28**, 147-154.
- Palm, D., Streit, D., Shanmugam, T., Weis, B.L., Ruprecht, M., Simm, S., y Schleiff, E. (2019). Plant-specific ribosome biogenesis factors in *Arabidopsis thaliana* with essential function in rRNA processing. *Nucleic Acids Research* **47**, 1880-1895.
- Pan, Q., Shai, O., Lee, L.J., Frey, B.J., y Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics* **40**, 1413-1415.
- Pendle, A.F., Clark, G.P., Boon, R., Lewandowska, D., Lam, Y.W., Andersen, J., Mann, M., Lamond, A.I., Brown, J.W., y Shaw, P.J. (2005). Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. *Molecular Biology of the Cell* **16**, 260-269.
- Peretea, M. (2012). The human transcriptome: an unfinished story. *Genes* **3**, 344-360.
- Planta, R.J., y Mager, W.H. (1998). The list of cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. *Yeast* **14**, 471-477.
- Pontvianne, F., Abou-Elail, M., Douet, J., Comella, P., Matia, I., Chandrasekhara, C., DeBures, A., Blevins, T., Cooke, R., Medina, F.J., Tourmente, S., Pikaard, C.S., y Sáez-Vásquez, J. (2010). Nucleolin is required for DNA methylation state and the expression of rRNA gene variants in *Arabidopsis thaliana*. *PLOS Genetics* **6**, e1001225.
- Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. *Cell* **108**, 557-572.
- Ramanathan, A., Robb, G.B., y Chan, S.H. (2016). mRNA capping: biological functions and applications. *Nucleic Acids Research* **44**, 7511-7526.
- Rebbapragada, I., y Lykke-Andersen, J. (2009). Execution of nonsense-mediated mRNA decay: what defines a substrate? *Current Opinion in Cell Biology* **21**, 394-402.
- Reddy, A.S., Marquez, Y., Kalyna, M., y Barta, A. (2013). Complexity of the alternative splicing landscape in plants. *Plant Cell* **25**, 3657-3683.
- Robledo, S., Idol, R.A., Crimmins, D.L., Ladenson, J.H., Mason, P.J., y Bessler, M. (2008). The role of human ribosomal proteins in the maturation of rRNA and ribosome production. *RNA* **14**, 1918-1929.
- Rodríguez-Molina, J.B., y Turtola, M. (2022). Birth of a poly(A) tail: mechanisms and control of mRNA polyadenylation. *FEBS Open Bio* doi: 10.1002/2211-5463.13528.
- Sáez-Vásquez, J., Caparros-Ruiz, D., Barneche, F., y Echeverría, M. (2004). A plant snoRNP complex containing snoRNAs, fibrillarin, and nucleolin-like proteins is competent for both rRNA gene binding and pre-rRNA processing in vitro. *Molecular and Cellular Biology* **24**, 7284-7297.
- Sáez-Vásquez, J., y Delseny, M. (2019). Ribosome biogenesis in plants: from functional 45S ribosomal DNA organization to ribosome assembly factors. *Plant Cell* **31**, 1945-1967.
- Sánchez-García, A.B., Aguilera, V., Micol-Ponce, R., Jover-Gil, S., y Ponce, M.R. (2015). *Arabidopsis MAS2*, an essential gene that encodes a homolog of animal NF- κ B activating protein, is involved in 45S ribosomal DNA silencing. *Plant Cell* **27**, 1999-2015.
- Sasaki, T., Kanno, T., Liang, S.C., Chen, P.Y., Liao, W.W., Lin, W.D., Matzke, A.J., y 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.
- Savada, R.P., y Bonham-Smith, P.C. (2014). Differential transcript accumulation and subcellular localization of *Arabidopsis* ribosomal proteins. *Plant Science* **223**, 134-145.

- Schindler, S., Szafranski, K., Hiller, M., Ali, G.S., Palusa, S.G., Backofen, R., Platzer, M., y Reddy, A.S. (2008). Alternative splicing at NAGNAG acceptors in *Arabidopsis thaliana* SR and SR-related protein-coding genes. *BMC Genomics* **9**, 159.
- Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., y Lohmann, J.U. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics* **37**, 501-506.
- Schwartz, B., Yeung, C., y Meinke, W. (1994). Disruption of morphogenesis and transformation of the suspensor in abnormal suspensor mutants of *Arabidopsis*. *Development* **120**, 3235-3245.
- Shi, Y. (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nature Reviews Molecular Cell Biology* **18**, 655-670.
- Sinha, R., Zimmer, A.D., Bolte, K., Lang, D., Reski, R., Platzer, M., Rensing, S.A., y Backofen, R. (2010). Identification and characterization of NAGNAG alternative splicing in the moss *Physcomitrella patens*. *BMC Plant Biology* **10**, 76.
- Smith, C.W., y Valcárcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends in Biochemical Sciences* **25**, 381-388.
- Spingola, M., Grate, L., Haussler, D., y Ares, M. (1999). Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. *RNA* **5**, 221-234.
- Sterne-Weiler, T., Howard, J., Mort, M., Cooper, D.N., y Sanford, J.R. (2011). Loss of exon identity is a common mechanism of human inherited disease. *Genome Research* **21**, 1563-1571.
- Supek, F., Miñana, B., Valcárcel, J., Gabaldón, T., y Lehner, B. (2014). Synonymous mutations frequently act as driver mutations in human cancers. *Cell* **156**, 1324-1335.
- Sydorsky, Y., Dilworth, D.J., Halloran, B., Yi, E.C., Makhnevych, T., Wozniak, R.W., y Aitchison, J.D. (2005). Nop53p is a novel nucleolar 60S ribosomal subunit biogenesis protein. *Biochemical Journal* **388**, 819-826.
- Syed, N.H., Kalyna, M., Marquez, Y., Barta, A., y Brown, J.W. (2012). Alternative splicing in plants – coming of age. *Trends in Plant Science* **17**, 616-623.
- Szick, K., Springer, M., y Bailey-Serres, J. (1998). Evolutionary analyses of the 12-kDa acidic ribosomal P-proteins reveal a distinct protein of higher plant ribosomes. *Proceedings of the National Academy of Sciences USA* **95**, 2378-2383.
- Tafforeau, L., Zorbas, C., Langhendries, J.L., Mullineux, S.T., Stamatopoulou, V., Mullier, R., Wacheul, L., y Lafontaine, D.L. (2013). The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of pre-rRNA processing factors. *Molecular Cell* **51**, 539-551.
- Taskinen, M., Ranki, A., Pukkala, E., Jeskanen, L., Kaitila, I., y Mäkitie, O. (2008). Extended follow-up of the finnish cartilage-hair hypoplasia cohort confirms high incidence of non-Hodgkin lymphoma and basal cell carcinoma. *American Journal of Medical Genetics Part A* **146A**, 2370-2375.
- Thiel, C.T., Horn, D., Zabel, B., Ekici, A.B., Salinas, K., Gebhart, E., Ruschendorf, F., Sticht, H., Spranger, J., Müller, D., Zweier, C., Schmitt, M.E., Reis, A., y Rauch, A. (2005). Severely incapacitating mutations in patients with extreme short stature identify RNA-processing endoribonuclease *RMRP* as an essential cell growth regulator. *American Journal of Human Genetics* **77**, 795-806.
- Thomson, E., y Tollervey, D. (2005). Nop53p is required for late 60S ribosome subunit maturation and nuclear export in yeast. *RNA* **11**, 1215-1224.
- Tolstrup, N., Rouzé, P., y Brunak, S. (1997). A branch point consensus from *Arabidopsis* found by non-circular analysis allows for better prediction of acceptor sites. *Nucleic Acids Research* **25**, 3159-3163.
- Tomecki, R., Sikorski, P.J., y Zakrzewska-Placzek, M. (2017). Comparison of preribosomal RNA processing pathways in yeast, plant and human cells - focus on coordinated action of endo- and exoribonucleases. *FEBS Letters* **591**, 1801-1850.
- Turner, I.A., Norman, C.M., Churcher, M.J., y Newman, A.J. (2006). Dissection of Prp8 protein defines multiple interactions with crucial RNA sequences in the catalytic core of the spliceosome. *RNA* **12**, 375-386.

- Turunen, J.J., Niemelä, E.H., Verma, B., y Frilander, M.J. (2013). The significant other: splicing by the minor spliceosome. *Wiley Interdisciplinary Reviews RNA* **4**, 61-76.
- Tutois, S., Cloix, C., Cuvillier, C., Espagnol, M.C., Lafleurriel, J., Picard, G., y Tourmente, S. (1999). Structural analysis and physical mapping of a pericentromeric region of chromosome 5 of *Arabidopsis thaliana*. *Chromosome Research* **7**, 143-156.
- Uechi, T., Tanaka, T., y Kenmochi, N. (2001). A complete map of the human ribosomal protein genes: assignment of 80 genes to the cytogenetic map and implications for human disorders. *Genomics* **72**, 223-230.
- Veith, T., Martin, R., Wurm, J.P., Weis, B.L., Duchardt-Ferner, E., Safferthal, C., Hennig, R., Mirus, O., Bohnsack, M.T., Wöhnert, J., y Schleiff, E. (2012). Structural and functional analysis of the archaeal endonuclease Nob1. *Nucleic Acids Research* **40**, 3259-3274.
- Veitia, R.A. (2002). Exploring the etiology of haploinsufficiency. *Bioessays* **24**, 175-184.
- Vlachos, A., Rosenberg, P.S., Atsidaftos, E., Kang, J., Onel, K., Sharaf, R.N., Alter, B.P., y Lipton, J.M. (2018). Increased risk of colon cancer and osteogenic sarcoma in Diamond-Blackfan anemia. *Blood* **132**, 2205-2208.
- Wahl, M.C., Will, C.L., y Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701-718.
- Wang, Z., y Burge, C.B. (2008). Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA* **14**, 802-813.
- Wegener, M., y Müller-McNicoll, M. (2018). Nuclear retention of mRNAs - quality control, gene regulation and human disease. *Seminars in Cell and Developmental Biology* **79**, 131-142.
- Weis, B.L., Kovacevic, J., Missbach, S., y Schleiff, E. (2015a). Plant-specific features of ribosome biogenesis. *Trends in Plant Science* **20**, 729-740.
- Weis, B.L., Palm, D., Missbach, S., Bohnsack, M.T., y Schleiff, E. (2015b). atBRX1-1 and atBRX1-2 are involved in an alternative rRNA processing pathway in *Arabidopsis thaliana*. *RNA* **21**, 415-425.
- Weis, F., Giudice, E., Churcher, M., Jin, L., Hilcenko, C., Wong, C.C., Traynor, D., Kay, R.R., y Warren, A.J. (2015c). Mechanism of eIF6 release from the nascent 60S ribosomal subunit. *Nature Structural and Molecular Biology* **22**, 914-919.
- Whittle, C.A., y Krochko, J.E. (2009). Transcript profiling provides evidence of functional divergence and expression networks among ribosomal protein gene paralogs in *Brassica napus*. *Plant Cell* **21**, 2203-2219.
- Wilkinson, M.E., Charenton, C., y Nagai, K. (2020). RNA splicing by the spliceosome. *Annual Review of Biochemistry* **89**, 359-388.
- Will, C.L., y Lührmann, R. (2001). Spliceosomal UsnRNP biogenesis, structure and function. *Current Opinion in Cell Biology* **13**, 290-301.
- Will, C.L., y Lührmann, R. (2011). Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology* **3**, a003707.
- Wolfe, K.H., y Shields, D.C. (1997). Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**, 708-713.
- Woloszynek, J.R., Rothbaum, R.J., Rawls, A.S., Minx, P.J., Wilson, R.K., Mason, P.J., Bessler, M., y Link, D.C. (2004). Mutations of the *SBDS* gene are present in most patients with Shwachman-Diamond syndrome. *Blood* **104**, 3588-3590.
- Xiong, W., Lan, T., y Mo, B. (2021). Extraribosomal functions of cytosolic ribosomal proteins in plants. *Frontiers in Plant Science* **12**, 607157.
- Yang, C., Georgiou, M., Atkinson, R., Collin, J., Al-Aama, J., Nagaraja-Grellscheid, S., Johnson, C., Ali, R., Armstrong, L., Mozaffari-Jovin, S., y Lako, M. (2021). Pre-mRNA processing factors and retinitis pigmentosa: RNA splicing and beyond. *Frontiers in Cell and Developmental Biology* **9**, 700276.

- Zakrzewska-Placzek, M., Souret, F.F., Sobczyk, G.J., Green, P.J., y Kufel, J. (2010). *Arabidopsis thaliana* XRN2 is required for primary cleavage in the pre-ribosomal RNA. *Nucleic Acids Research* **38**, 4487-4502.
- Zhang, H., Xia, R., Meyers, B.C., y Walbot, V. (2015a). Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Current Opinion in Plant Biology* **27**, 84-90.
- Zhang, X.R., Qin, Z., Zhang, X., y Hu, Y. (2015b). *Arabidopsis* SMALL ORGAN 4, a homolog of yeast NOP53, regulates cell proliferation rate during organ growth. *Journal of Integrative Plant Biology* **57**, 810-818.
- Zug, R. (2022). Developmental disorders caused by haploinsufficiency of transcriptional regulators: a perspective based on cell fate determination. *Biology Open* **11**, bio058896.

IX.- PUBLICACIONES

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

Adrián Cabezas-Fuster¹, Rosa Micol-Ponce¹, Sara Fontcuberta-Cervera¹ and María Rosa Ponce^{1*}

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- κ B 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* (*Ler*), 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 *Ler* 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* (*Ler*) 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 *Ler* 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 *Ler*) 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 sub-bins), 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). Flu-

orescein 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*), *SARI* (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 *Ler* 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 M₃ progeny of an M₂ plant (P8 25.1) was named *mas5-1* (Figure 1C) and crossed to Col-0 to obtain an F₂ mapping population. Two phenotypic classes were defined: plants exhibiting the phenotype of *ago1-52* and plants similar to a wild-type Col-0/*Ler* hybrid. We genotyped plants from each class for 32 molecular markers known to be polymorphic between Col-0 and *Ler*, as well as for the presence of *ago1-52* and its *AGO1* wild-type allele.

Iterative linkage analysis of these F₂ 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 M₃ 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 M₃ 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-

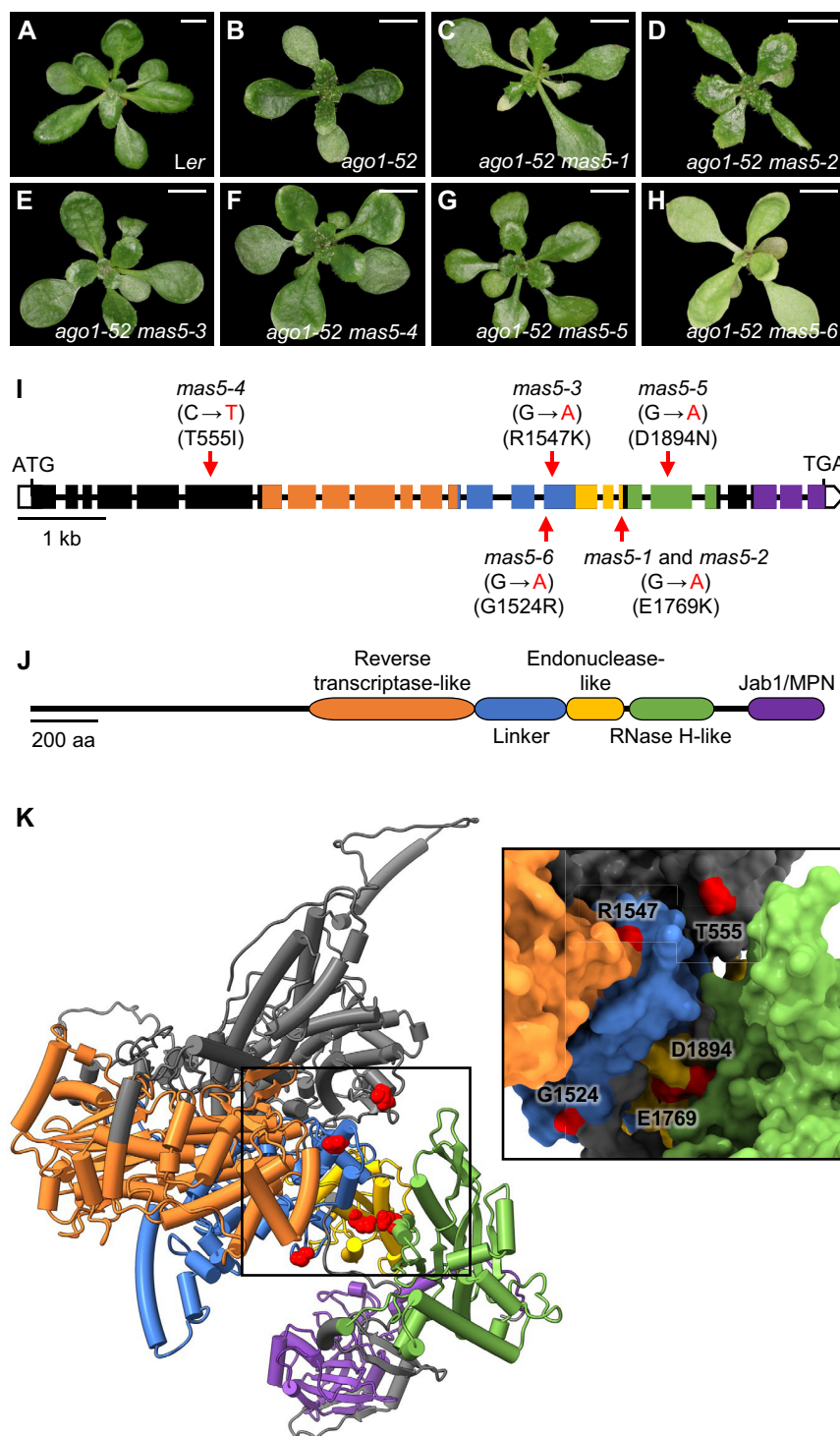


Figure 1. Molecular nature and effects of the *mas5* alleles of *PRP8* examined in this study. (A–H) Suppression of the morphological phenotype of *ago1-52* by the *mas5* mutations. Rosettes of (A) the wild-type *Ler*, (B) the *ago1-52* single mutant, and the (C) *ago1-52 mas5-1*, (D) *ago1-52 mas5-2*, (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 *Ler*. 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* M₃ plant shown in Figure 1H, as also observed in all *ago1-52 mas5* lines shown in Figure 1. *mas5-1* and *mas5-3* F₂ plants, from second backcrosses to *Ler*, 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, or 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) M₃ lines twice to *Ler*. After these backcrosses, *mas5-3* plants exhibited moderately larger rosettes compared to *Ler* and *mas5-1* but were otherwise very similar to *Ler* 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 *Ler* values (Figure 2A–H). The *ago1-52 mas5-1* plants were similar to *Ler*, 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 *Ler*, *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 21st 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 *Ler*, *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 M₄ 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 *angustal-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 *Ler* 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 15th 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

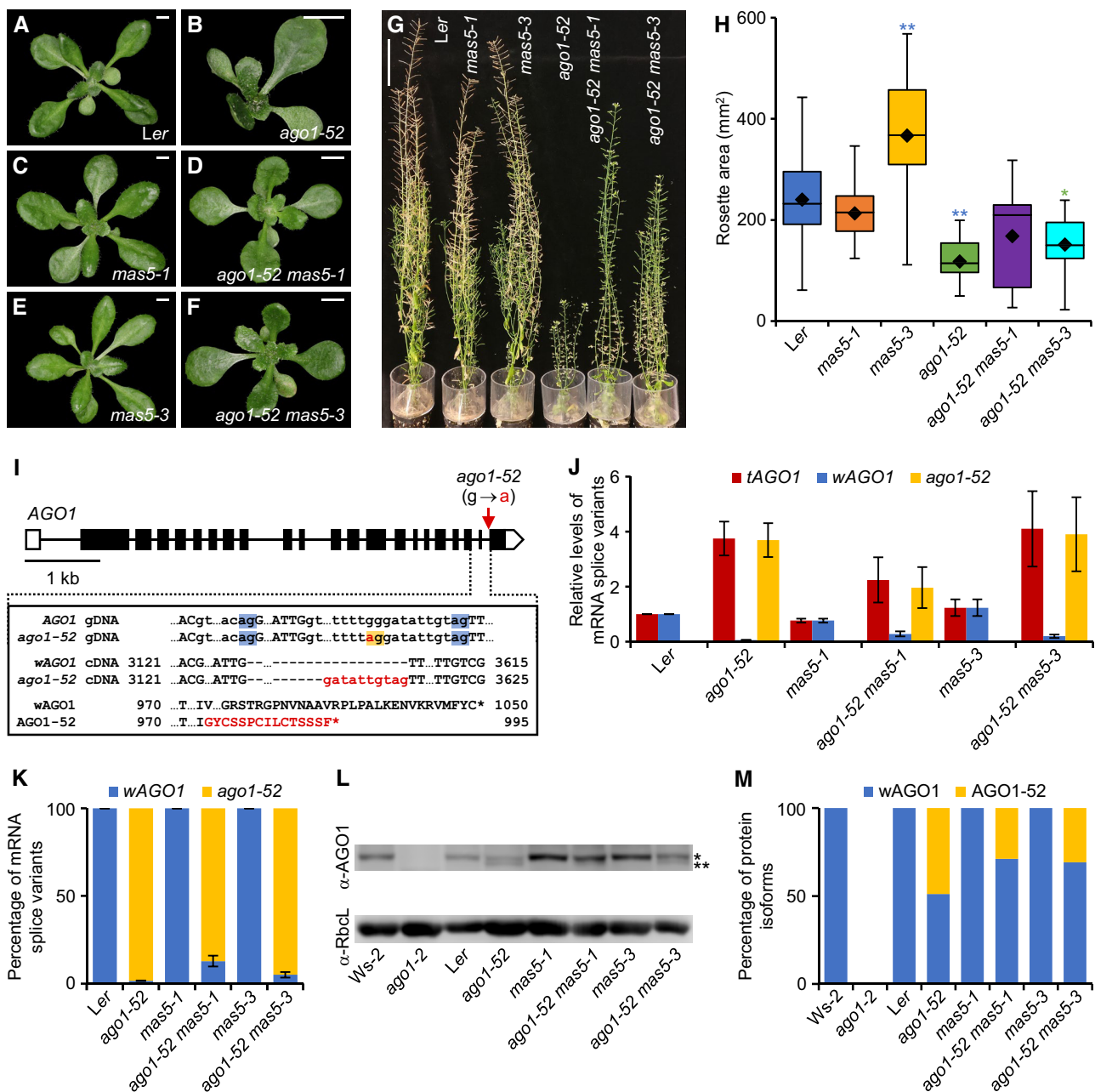


Figure 2. Suppression of the morphological and molecular phenotypes of *ago1-52* by *mas5-1* and *mas5-3*. (A–F) Rosettes of (A) *Ler*, (B) *ago1-52*, (C) *mas5-1*, (D) *ago1-52 mas5-1*, (E) *mas5-3* and (F) *ago1-52 mas5-3* plants. (G) From left to right, adult plants of *Ler*, *mas5-1*, *mas5-3*, *ago1-52*, *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 $Q3 + 1.5 \times IQR$ or no less than $Q1 - 1.5 \times IQR$, 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.

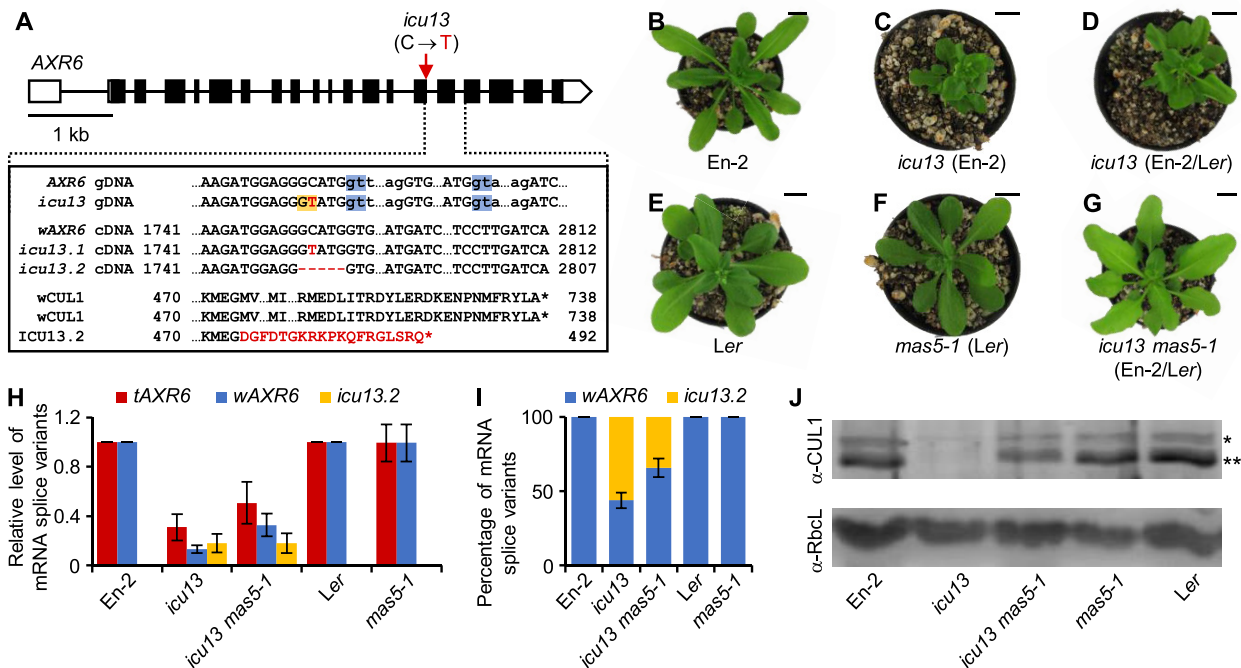


Figure 3. Suppression of the morphological and molecular phenotypes of *icu13* by *mas5-1*. (A) *AXR6* gene structure, mRNA splice variants, and CUL1 isoforms from the translation of *icu13* transcripts, represented as described in the legend of Figure 2. A red arrow indicates the position of the *icu13* mutation. (B–G) Rosettes of (B) En-2, (C) *icu13* (in the En-2 genetic background), (D) *icu13* (in the En-2/Ler hybrid genetic background), (E) Ler, (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*, Ler, 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 15th 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 F₂ 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 15th of its 20 exons (Figure 3A) and produces a PTC at the 16th 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 F₂ 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 F₂ 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 *Ler* 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 *Ler* 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 *Ler* 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 21st 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).

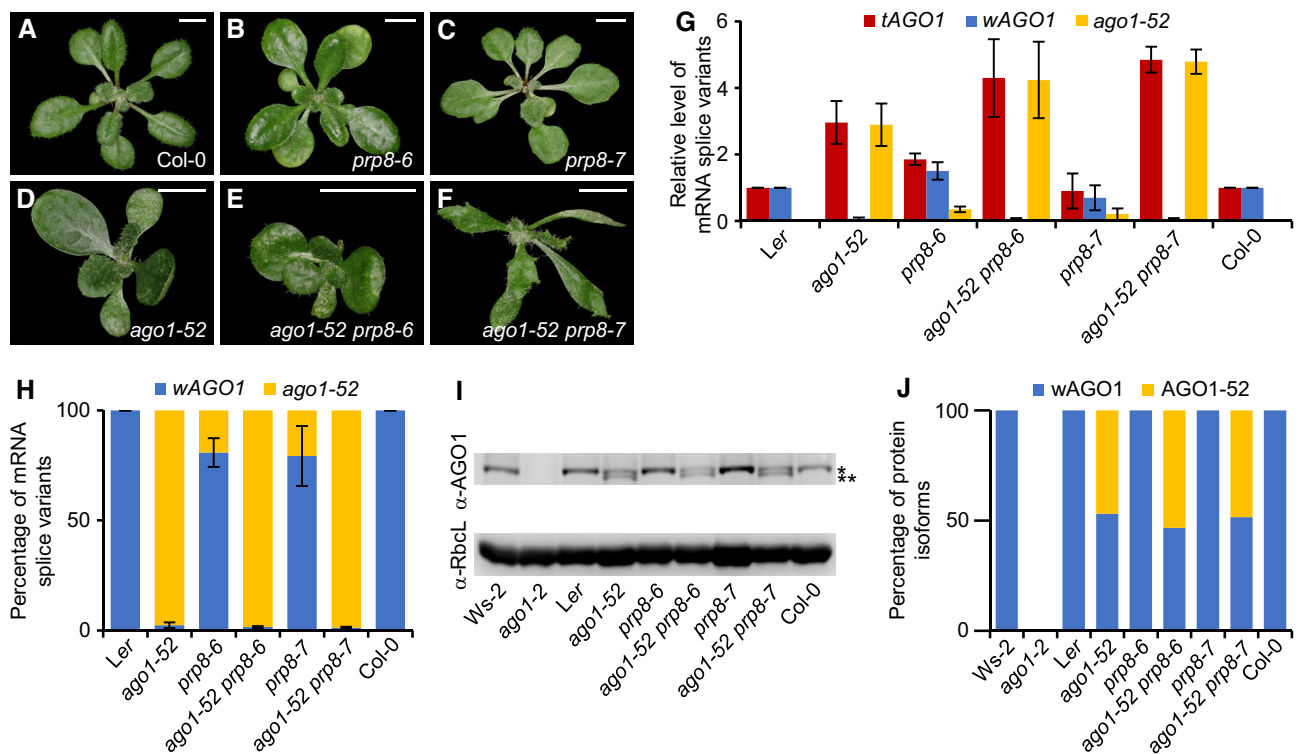


Figure 4. Genetic interactions between *prp8* hypomorphic alleles and *ago1-52*. (A–F) Rosettes of (A) Col-0, (B) *prp8-6*, (C) *prp8-7*, (D) *ago1-52*, (E) *ago1-52 prp8-6*, and (F) *ago1-52 prp8-7*. Photographs were taken 21 das. Scale bars: 4 mm. (G) RT-qPCR analysis of the expression of total (*tAGO1*), 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 *Ler*, *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

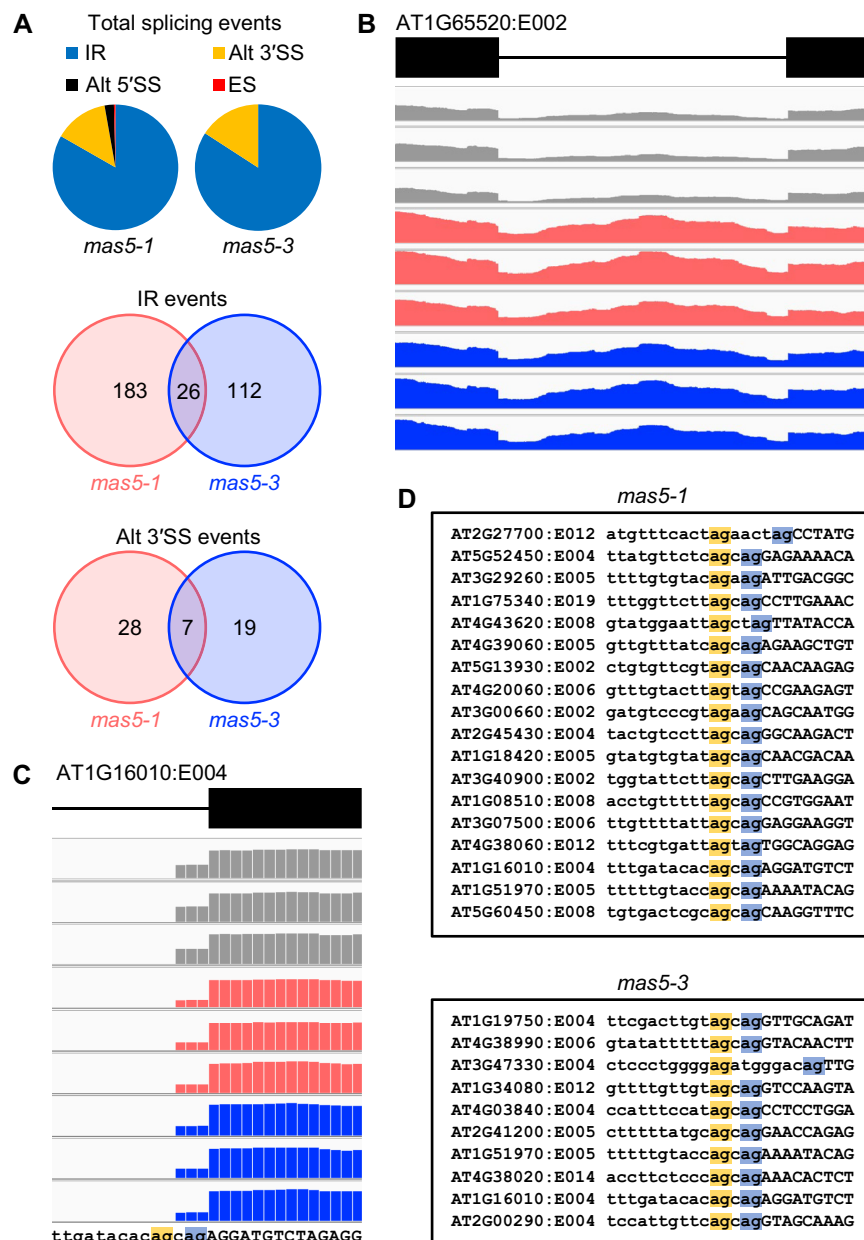


Figure 5. Genome-wide analysis of pre-mRNA splicing in the *mas5* mutants. (A) Percentage of differential splicing events identified in *mas5-1* and *mas5-3*, and Venn diagrams showing the IR and Alt 3'SS events. IR: intron retention; ES: exon skipping; Alt 3'SS/5'SS: alternative 3'/5' splicing site. (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 *Ler* (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* (*SARI*), 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-

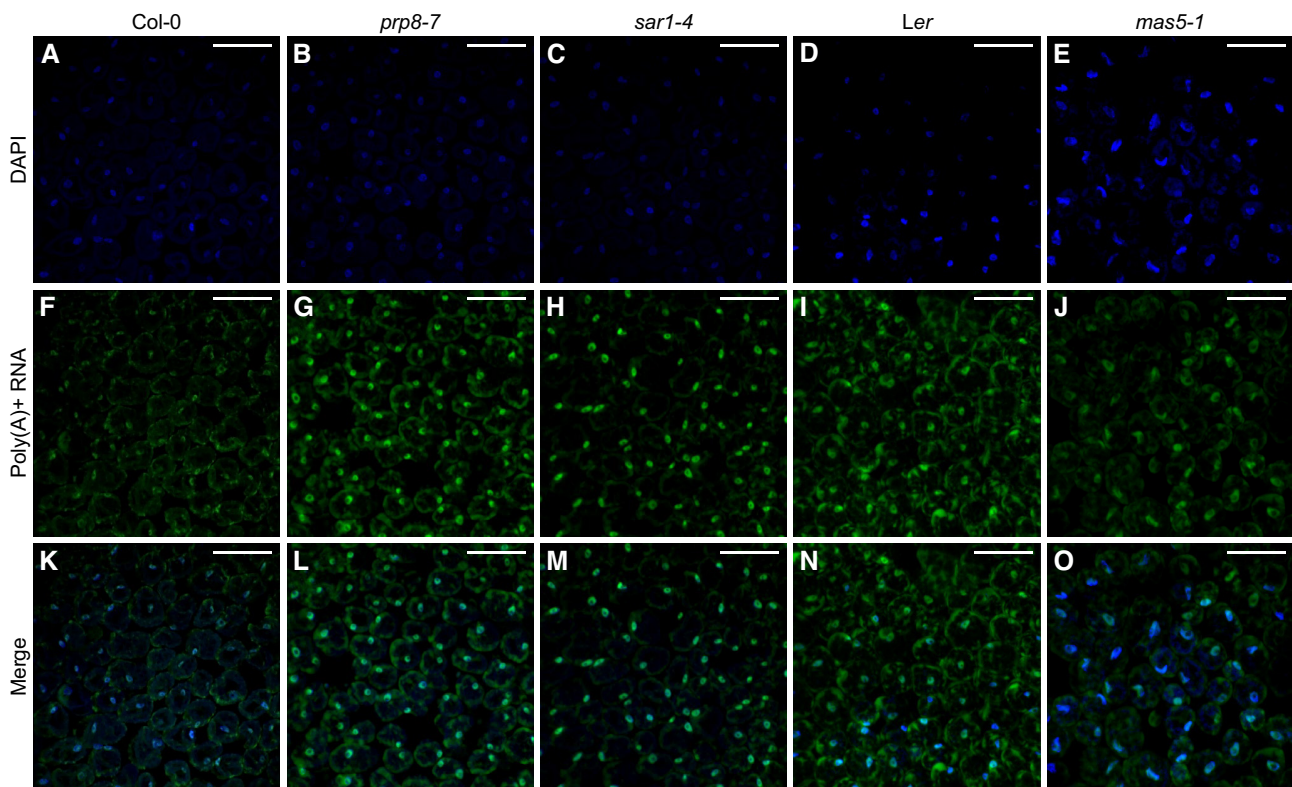


Figure 6. Detection of poly(A)+ RNAs in *prp8-7* and *mas5-1* leaf cells. (A–O) Poly(A)+ RNA-FISH assays in palisade mesophyll cells of (A, F, K) Col-0, (B, G, L) *prp8-7*, (C, H, M) *sar1-4*, (D, I, N) *Ler* and (E, J, O) *mas5-1* leaves. Fluorescent signals correspond to (A–E) nuclear 4',6-diamidino-2-phenylindole (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 struc-

tures 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 *icul3 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 *icul3*, 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 *icul3* 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 *icul3*, 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

- Grainger, R.J. and Beggs, J.D. (2005) Prp8 protein: At the heart of the spliceosome. *RNA*, **11**, 533–557.

- Hoskins, A.A. and Moore, M.J. (2012) The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.*, **37**, 179–188.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Weigel, D. and Glazebrook, J. (2006) EMS mutagenesis of Arabidopsis seed. *Cold Spring Harb. Protoc.*, **2006**, pdb.prot4621.
- 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.
- Kulichová, K., Kumar, V., Steinbachová, L., Klodová, B., Timofejeva, L., Juříček, M., Honyš, D. and Hafidh, S.S. (2020) PRP8A and PRP8B spliceosome subunits act coordinately to control pollen tube attraction in Arabidopsis thaliana. *Development*, **147**, dev186742.
- 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.
- 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.
- Berná, G., Robles, P. and Micol, J.L. (1999) A mutational analysis of leaf morphogenesis in Arabidopsis thaliana. *Genetics*, **152**, 729–742.
- 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.
- Ponce, M.R., Robles, P. and Micol, J.L. (1999) High-throughput genetic mapping in Arabidopsis thaliana. *Mol. Gen. Genet.*, **261**, 408–415.
- 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.
- 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.
- Wingett, S.W. and Andrews, S. (2018) FastQ screen: a tool for multi-genome mapping and quality control. *Fl1000Research*, **7**, 1338.
- 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.
- 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.
- 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-Mañúes, 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., Lüthmann, 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., Lüthmann, 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., Risseuw, 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-Bonmati, 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.

**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.

Supplementary Figures, Tables and References

Ler	(<i>PRP8</i>)	2320	AACCTCAAGCCTGTGAAAACCTTTGACCACCAAAGAGCGAAAGAAGTCACGTTTTGGG
P5 11.1	(<i>mas5-4</i>)	2320	AACCTCAAGCCTGTGAAAACCTTTGACCA T CAAAGAGCGAAAGAAGTCACGTTTTGGG
Ler	(<i>PRP8</i>)	6358	acatgaactcagTTTCCCGACATGGGAGGGACTTTTCTGGGAGAAGGCGTCTGGTTT
P7 22.1	(<i>mas5-6</i>)	6358	acatgaactcagTTTCCCGACATGGGAG A GACTTTTCTGGGAGAAGGCGTCTGGTTT
P7 23.1	(<i>mas5-6</i>)	6358	acatgaactcagTTTCCCGACATGGGAG A GACTTTTCTGGGAGAAGGCGTCTGGTTT
Ler	(<i>PRP8</i>)	6428	AAATATAAGAAGTTGACTAATGCTCAGAGGTCTGGTCTGAACCAGATTCCAAATAGA
P7 24.1	(<i>mas5-3</i>)	6428	AAATATAAGAAGTTGACTAATGCTCAG A AGTCTGGTCTGAACCAGATTCCAAATAGA
P7 26.1	(<i>mas5-3</i>)	6428	AAATATAAGAAGTTGACTAATGCTCAG A AGTCTGGTCTGAACCAGATTCCAAATAGA
Ler	(<i>PRP8</i>)	7266	gTCCAATCCGGCTCTATATGTGTTGAGAGAGAGGATAAGGAAAGGTTTGCAGCTATA
P8 4.1	(<i>mas5-1</i>)	7266	gTCCAATCCGGCTCTATATGTGTTGAG A AGAGGATAAGGAAAGGTTTGCAGCTATA
P8 14.1	(<i>mas5-1</i>)	7266	gTCCAATCCGGCTCTATATGTGTTGAG A AGAGGATAAGGAAAGGTTTGCAGCTATA
P8 25.1	(<i>mas5-1</i>)	7266	gTCCAATCCGGCTCTATATGTGTTGAG A AGAGGATAAGGAAAGGTTTGCAGCTATA
P7 13.1	(<i>mas5-2</i>)	7266	gTCCAATCCGGCTCTATATGTGTTGAG A AGAGGATAAGGAAAGGTTTGCAGCTATA
P7 49.1	(<i>mas5-2</i>)	7266	gTCCAATCCGGCTCTATATGTGTTGAG A AGAGGATAAGGAAAGGTTTGCAGCTATA
Ler	(<i>PRP8</i>)	7769	GATTATTGTCACACGGAAAGGAATGTTGGATCCCCTTGAGGTTCACTTGCTTGATTT
P4 22.2	(<i>mas5-5</i>)	7769	GATTATTGTCACACGGAAAGGAATGTTG A ATCCCCTTGAGGTTCACTTGCTTGATTT

Supplementary Figure S1. Mutations identified in *PRP8* in the *mas5* suppressor lines studied in this work. Numbers indicate nucleotide positions, numbered from the first (5') nucleotide of the 5'-UTR of *PRP8*. Intronic and exonic sequences are shown in lowercase and uppercase, respectively. Mutated nucleotides are shown in red.

S. cerevisiae 1 ----MSGLPPPPPGFEEDSDLALPPPPPPPPGYEIEELDNPMVPSVSNEDTFLPPPPPPP
A. thaliana 1 MWNNNDGMPLAPPGT---GGSMMPPPPAAHPSYT-----ALPPPSNP-
H. sapiens 1 -----MAGVFPYRG---PGNPVP-----
C. elegans 1 -----MANYGG-----
consensus

S. cerevisiae 57 SNFEINAEEIVDFTLPPPPPPPG--LDELETKAEKKVELHGKRRKLDIGKDTFVTRKSRKR
A. thaliana 40 -----TPPVEPTPEEAEAKLEEKARKWMQLNSKR-----
H. sapiens 16 -----GPLAPLPDYMSEEKIQEKARKWQQLQAKR-----
C. elegans 7 -----HPQTEPHAI PDSILEEKSRKWKQLQGKR-----
consensus * * * * *

S. cerevisiae 115 AKKMTKAKRNSLYTPKAEMPPEHLRKIINTHSDMASKMYNTDKKAF LGALKYLPHAILK
A. thaliana 69 ---YGDKRKF GFVETQKEDMPPEHVRKIIRDHGDMSKKFRHDKRVYLGALKFVPHAVFK
H. sapiens 45 ---YAEKRF GFVDAQKEDMPPEHVRKIIRDHGDMTNRKFRHDKRVYLGALKYMPHAVLK
C. elegans 35 ---YSEKKRFGMSDTQKEEMPPEHVRKVIRDHGDMSRKYRHDKRVYLGALKYMPHAVLK
consensus * * * .*****.* * * * . . ** .*****.***. *

S. cerevisiae 175 LLENMPHPWEQAKEVKVLYHTSGAITFVNETPRVIEPVYTAQWSATWIAMRREKRDRRHF
A. thaliana 126 LLENMPMPWEQVRDVKVLYHITGAITFVNEIPWVVEPIYMAQWGTMWIMMRREKRDRRHF
H. sapiens 102 LLENMPMPWEQIRDVPVLYHITGAISFVNEIPWVIEPVYISQWGSWIMMRREKRDRRHF
C. elegans 92 LLENMPMPWEQIRDVKVLYHITGAITFVNDIPRVIEPVYMAQWGTMWIMMRREKRDRRHF
consensus ***** ** * . * * * * . *

S. cerevisiae 235 KRMRFPPFDDEEPLSYEQHIENIEPLDPIINLPLDSQDDEYVKDWLYDSRPLEEDSKKVN
A. thaliana 186 KRMRFPPFDDEEPLDYADNLLDVPLEPIQLELDEEEDSAVHTWFYDHKPLVK-TKLLN
H. sapiens 162 KRMRFPPFDDEEPLDYADNILDVEPLEAIQLELDPEEDAPVLDWFYDHPQRDRSRKYVN
C. elegans 152 KRMRFPPFDDEEPLDYADNILDVEPLEPIQMELDPEEDGAVAEWYFYDHKPLAT-TRFVN
consensus ***** . * * * * * . . *

S. cerevisiae 295 GTSYKWSFDLPEMSNLYRLSTPLRDEVTDKNYYLFDKKSFFNGKALNNAIPGGPKFEP
A. thaliana 245 GPSYRRWNLSLPIMATLHRLAGQLLSDLIDRNYFYLFDMPSFFTAKALNMCIPGGPKFEP
H. sapiens 222 GSTYQRWQFTLPMSTLYRLANQLLTDLVDDNYFYLFDLKAFFTSKALNMAIPGGPKFEP
C. elegans 211 GPTYRKWAFSIPQMSTLYRLANQLLTDLVDDNYFYLFDMKSFFTAKALNVAIPGGPKFEP
consensus * . *

prp8-9 (P347S)

S. cerevisiae 355 LYPRE--EEEDYNEFN SIDRVIIRVPIRSEYKVAFPHLYNSRPRS--VRIPWYNNEVSCI
A. thaliana 305 LYRDMKGD EDWNEFN DINKLIIRSPIRTEYRIAFPHLYNNRPRK--VKLCVYHSEPMIMY
H. sapiens 282 LVRDINLQEDWNEFN DINKLIIRQPIRTEYKIAFPYLYNNLPHH--VHLTWYHTENVVF
C. elegans 271 LVKDLH-TDEDWNEFN DINKVIIRAPIRTEYRIAFPFMYNNLISLSPVQVSWYHTPEVVF
consensus * . *

S. cerevisiae 411 IQNDEEYDTPALFFDPSLNPIPHFIDNNSLNVSNTKENGDFTLPEDFAPLLAE EEEELIL
A. thaliana 363 IK-TEDPDLPAFYDPLIHPIISNTNKEKRERKVYD--DEDDFALPEGVEPLL RD-TQLYT
H. sapiens 340 IK-TEDPDLPAFYDPLINPIISHRHSVKSQEPLPD--DDEEFELPEFVEPEFLKO-TPLYT
C. elegans 330 IK-TEDPDLPAFYDPLINPIVLSNLKATEENLPEGEEDEWELPEDVRPIFED-VPLYT
consensus * * * * * . *

S. cerevisiae 471 PNTK DAMS LYSHPFPFNRTK GKMVRAQDVALAKKWFLOHPDEEY PVKVKVSYQKLLKNYV
A. thaliana 419 DTTAAGISLLFAPRPFNMRSGRTRRAEDIPLVSEWFKEHCPPAYPVKVRVSYQKLLKCYV
H. sapiens 396 DNTANGIALLWAPRPFNLRSGRTRRALDIPLVKNWYREHC PAGQPVKVRVSYQKLLKYYV
C. elegans 388 DNTANGIALLWAPRPFNLRSGRTRRAVDVPLVKS WYREHC PAGMPVKVRVSYQKLLKVEV
consensus * . *

S. cerevisiae 1069 LLRLIVDPNIADYITAKNNVVINFKDMSHVNKYGLIRGLKFFASFIQYQYGLVIDLLLLGQ
A. thaliana 1019 LLRLVLDHNIADYVSAKNNVVL SYKDMSTNSYGLIRGLQFASFVVOFYGLLLDLLLLGL
H. sapiens 996 LLRLIVDHNIDYMTAKNNVVINYKDMNHTNSYGIIRGLQFASFIQYQYGLVMDLLVLGL
C. elegans 988 LLRLIVDHNIDYMTSKNNVVLINFKDMNHTNSYGIIRGLQFASFIQYQYGLVLDLLVLGL
consensus ***** * ***** . ***** . ***** * * . ***** ***** . ***** . *****

S. cerevisiae 1129 ERATDLAGPANNPNNEFMQFKSKEVEKAHPRIRLYTRYLDRIYMLFHFEEDGEGEELTDEYLA
A. thaliana 1079 TRASEIAGPPOMPNEFMTEFWDTKVETRHPRIRLYSRYIDKVHIMFKFTHEEARDLIQRILT
H. sapiens 1056 HRASEMAGPPOMPNDFLSFQDIATEAAHPRIRFCRYIDRIHIFFRFETADEARDLIQRILT
C. elegans 1048 RRASEIAGPPQCPNEFLQFQDVATEIGHPIRLYCRYIDRVWIMFRFSADEARDLIQRILT
consensus **...*** . ***. * * ***** . ***... . ** . * . * **

prp8-8 (P1141S)

S. cerevisiae 1189 ENPDPNFENSI GYNNRKCWPKDSRMRLIROQDVNLGRAVFWETQSRVPTSLTISKWENAFV
A. thaliana 1139 EHPDPNNENMVGYNNKCCWPRDARMRLMKHDVNLGRSVFWDMMKRLPRSITTTLEWENGFV
H. sapiens 1116 EHPDPNNENIVGYNNKCCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSVTTVQWENSEV
C. elegans 1108 EHPDPNNENIVGYNNKCCWPRDARMRLMKHDVNLGRAVFWDIKNRLPRSITTTVEWENSEV
consensus * ***** ** . ***** . ***** * ***** . ***** ***** . ** * * . * * * *

S. cerevisiae 1249 SVYSKNNPNLLFSMCGFEVRILPRORM--EEVVSNDGVDLVDERTKORTAKAYLVKVE
A. thaliana 1199 SVYSKDNPNLLFSMCGFEVRILPKIRMTQEAFAFNTKDGWVNLQNEQTKERTAVAFQFLRVDD
H. sapiens 1176 SVYSKDNPNLLFSMCGFEVRILPKCRTSYEEFT-HKDGVWNLQNEVTKERTAQCFQFLRVDD
C. elegans 1168 SVYSKDNPNMLFDMSCGFECRILPKCRTANEEFV-HRDGVWNLQNEVTKERTAQCFQFLRVDE
consensus ***** ** *

S. cerevisiae 1307 EEIKKFDSEIRIGILMASGSTTFTKVAKWNTSLISLFTYFREAVATEPPLLDILVKGETR
A. thaliana 1259 EHMKVFENRVROILMSSGSTTFTKIVNKWNTALIGLMTYFREAVTHTQELLDLILVKCENK
H. sapiens 1235 ESMQRFHNRVROILMASGSTTFTKIVNKWNTALIGLMTYFREAVVNTQELLDLILVKCENK
C. elegans 1227 ESIKSFHNRVROILMSSGSTTFTKIVNKWNTALIGLMTYFREAVVNTQELLDLILVKCENK
consensus * . *

prp8-10 (G1347D)

S. cerevisiae 1367 IQNRVKGILNSKMPTRFPVAVFYTPKEIGGLGMLISASHILIPASDLSWSKQTDG-GITHF
A. thaliana 1319 IQTRIKIGLNSKMPSRFPPVIFYTPKEIGGLGMLSMGHILIPQSDLRYSKQTDV-GVTHF
H. sapiens 1295 IQTRIKIGLNSKMPSRFPPVIFYTPKEIGGLGMLSMGHVLIPOSDLRWSKQTDV-GITHF
C. elegans 1287 IQTRIKIGLNSKMPSRFPPVIFYTPKEIGGLGMLSMGHVLIPOSDLRWMOQTEAGGVTHF
consensus ** * * . ***** . ***** . ***** . ***** . ***** * * * * * * * *

S. cerevisiae 1426 RGMTHEDELIPIFRYIWE EFLDSRVWAEYA KRQEA QNRRL EELE SWD
A. thaliana 1378 RSGMSHEEDQLIPNLYRYIQPWESEFIDSRVWAEYALKRQEAAQNRRLTLEDLEDSDW
H. sapiens 1354 RSGMSHEEDQLIPNLYRYIQPWESEFIDSRVWAEYALKRQEAAQNRRLTLEDLEDSDW
C. elegans 1347 RSGMSHDEDQLIPNLYRYIQPWEAEFVDSRVWAEYALKRQEANAQNRRLTLEDLDDSDW
consensus * * * * . *

S. cerevisiae 1486 RGI PRISTLTFQDRHTLAYDRGHRIRREFKQYSLERNSPFWWTNSHHDGKLNWLNAYRTD
A. thaliana 1438 RGI PRINTLTFQDRHTLAYDKGWRVRTDFKQYQVVKQNPFWWTHQRHDGKLNWLNAYRTD
H. sapiens 1414 RGI PRINTLTFQDRHTLAYDKGWRVRTDFKQYQVVKQNPFWWTHQRHDGKLNWLNAYRTD
C. elegans 1407 RGI PRINTLTFQDRHTLAYDKGWRVRTDFKAYQILKQNPFWWTHQRHDGKLNWLNAYRTD
consensus ***** ***** . ***** . ***** . ***** . ***** . ***** *****

mas5-6 (G1524R)

mas5-3 (R1547T)

S. cerevisiae 1546 VIQALGGIETILEHTLFFKGTGFNSWEGFLWFKASGFEDSMQFKKLTAAQRTGLSQAIPNRR
A. thaliana 1498 VIQALGGVEGILEHTLFFKGTTFPTWEGLFWFKASGFEE SMKYKKLTAQRSGLNQAIPNRR
H. sapiens 1474 MIQALGGVEGILEHTLFFKGTTFPTWEGLFWFKASGFEE SMKWKKLTAQRSGLNQAIPNRR
C. elegans 1467 MIQALGGVEGILEHTLFRGTTFPTWEGLFWERASGFEE SMKFKKLTAAQRSGLNQAIPNRR
consensus . ***** *

S. cerevisiae 1606 FTLWWSPTINRANVYVGFVQLDLTGIFLHGKIPTIKISLIQIFRAHLWQKIHESIVEDI
A. thaliana 1558 FTLWWSPTINRANVYVGFVQLDLTGIFMHGKIPTLKISLIQIFRAHLWQKIHESVVM DL
H. sapiens 1534 FTLWWSPTINRANVYVGFVQLDLTGIFMHGKIPTLKISLIQIFRAHLWQKIHESIVMDL
C. elegans 1527 FTLWWSPTINRANVYVGFVQLDLTGIFMHGKIPTLKISLIQIFRAHLWQKIHESVVM DL
consensus 1621 ***** .***** ** .

S. cerevisiae 1666 COILDGELDVLQIETVTKETVHFRKSYKMNSSAADITMESVHEWEVSKPSLIHETNDSFK
A. thaliana 1618 COVLDQELDALEIETVOKETIHPRKSYKMNSSCADVLLFAAHKWPMSKPSLVAESKDMFD
H. sapiens 1594 COVFDQELDALEIETVOKETIHPRKSYKMNSSCADILLFASYKWNVSRPSLLADSKDVMD
C. elegans 1587 COVFDQELDALEIETVOKETIHPRKSYKMNSSCADVLLFAQYKWNVSRPSLMADSKDVMD
consensus **. * ** * * . * ** * .***** ** . * .* .****. . . *

S. cerevisiae 1726 GLITNKMWFDVQLRYGDYSDHISRYSRAKFLDYTTDNVSMYPSPTGVMIGIDLAYNMYD
A. thaliana 1678 QKASNKYWIDVQLRWGDYSDHDIERYTRAKFMDYTTDNMSIYPSPTGVMIGIDLAYNLHS
H. sapiens 1654 STTTQKYWIDIQLRWGDYSDHDIERYARAKFLDYTTDNMSIYPSPTGVLIAIDLAYNLHS
C. elegans 1647 NTTTQKYWLDVQLRWGDYSDHDVRYARAKFLDYTTDNMSIYPSPTGVLIAIDLAYNLYS
consensus . . * * * .*** .***** . ** **** .***** .* .***** .* .***** .

S. cerevisiae 1786 AYGNWFNGLKPLIIONSMRTIMKANPALYVLRERIRKGLQIYQSSVQEPFLNNSNYAELEN
A. thaliana 1738 AFGNWFPGSKPLIAQAMNKMKSNPALYVLRERIRKGLQLYSSEPTPEYLLSSONYGELES
H. sapiens 1714 AYGNWFPGSKPLIQAAMAKIMKANPALYVLRERIRKGLQLYSSEPTPEYLLSSONYGELES
C. elegans 1707 AYGNWFPGMKPLIROAMAKI I KANPAFYVLRERIRKGLQLYSSEPTPEYLLSSONYGELES
consensus *. **** * ***. . * *. * ** * * .***** .* * * .** * * * * .*

mas5-1, mas5-2 (E1769K) prp8-11 (R1770K)

S. cerevisiae 1846 NDIKLFVDDTINVYRVTIHKTFEGNVAIKAIKNGCIFLNPKTGHLFLKIIHTSVWAGQKRL
A. thaliana 1798 NQIIWFVDDTINVYRVTIHKTFEGNLTTPKPIKNGAIFIFNPRTGQFLKVIHTSVWAGQKRL
H. sapiens 1774 NQIIWFVDDTINVYRVTIHKTFEGNLTTPKPIKNGAIFIFNPRTGQFLKIIHTSVWAGQKRL
C. elegans 1767 NQIIWFVDDTINVYRVTIHKTFEGNLTTPKPIKNGAIFIFNPRTGQFLKIIHTSVWAGQKRL
consensus * * * * .***** .***** . ** * * * * * * .***** .***** .***** .***** .

prp8-7 (G1820E)

S. cerevisiae 1906 SOLAKWKTAEVVSALVRSLSLPEEQPKQIIIVTRKAMLDPLEVHMLDFPNIAIRPTELRPRLPF
A. thaliana 1858 GOLAKWKTAEVAALVRSLSLPEEQPKQIIIVTRKMGMLDPLEVHLLDFPNIVIKGSELQQLPF
H. sapiens 1834 GOLAKWKTAEVAALIRSLPEEQPKQIIIVTRKMGMLDPLEVHLLDFPNIVIKGSELQQLPF
C. elegans 1827 SOLAKWKTAEVAALIRSLPEEQPRQIIIVTRKAMLDPLEVHLLDFPNIVIKGSELMPLPF
consensus ***** ** .***** ***** .***** .***** .***** * . * * ** * * *

prp8-6 (G1891E) mas5-5 (D1894N)

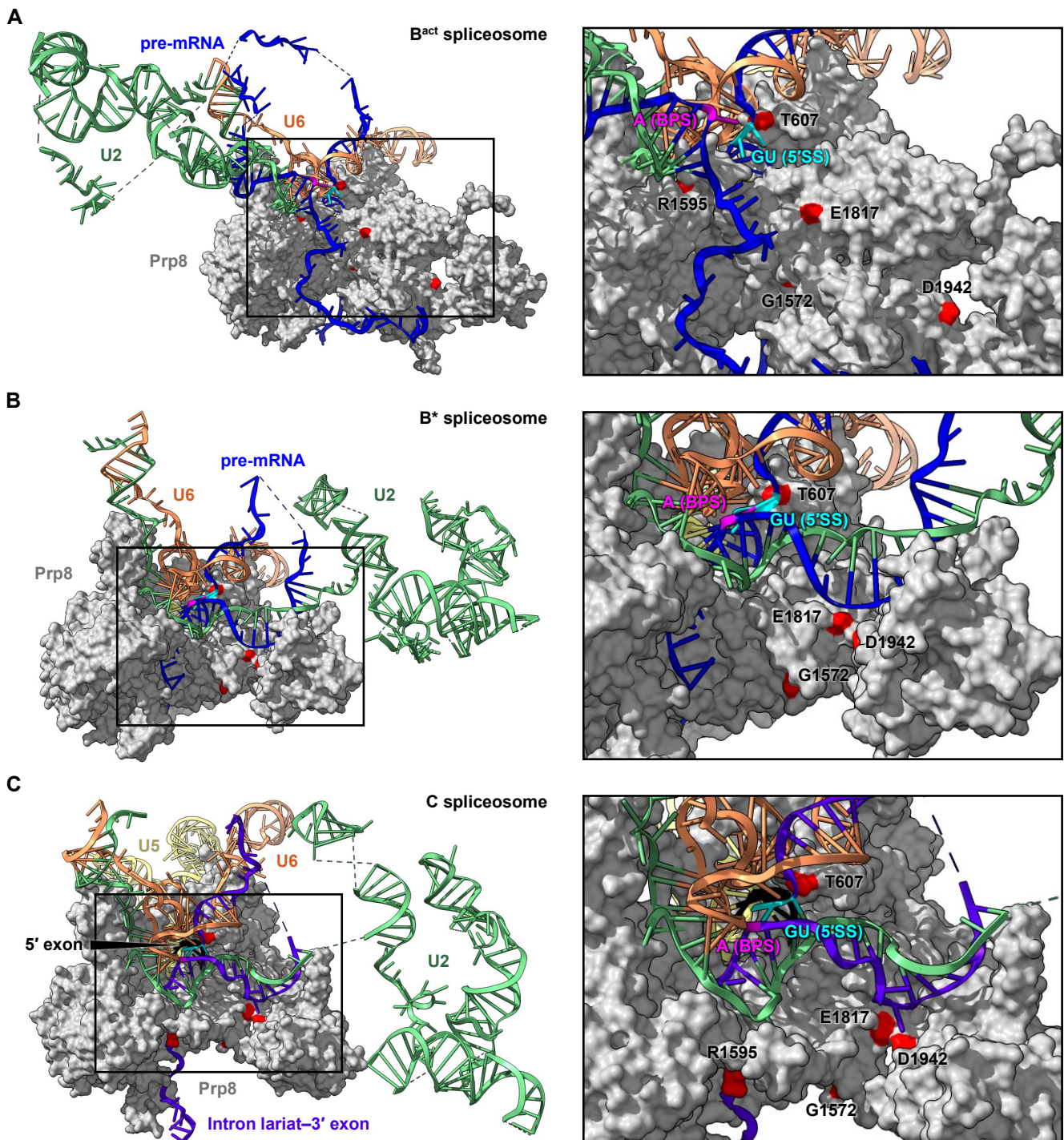
S. cerevisiae 1966 SAAMSIDKLSDVVMKATEPQMVLFNLYDDWLDRISSYAFSRLILILRALKVNNEE SAKMI
A. thaliana 1918 QACLKIEKFGDLILKATEPQMVLFNLYDDWLKSISSYAFSRLILILRALHVNNEKAKML
H. sapiens 1894 QACLKVEKFGDLILKATEPQMVLFNLYDDWLKTISSYAFSRLILILRALHVNNDRAKVI
C. elegans 1887 QAIMKVEKFGDLILKATEPQMVLFNLYDDWLKTISSYAFSRLVLMRGMHINPDKTKVI
consensus * . . . * * . . ***** .***** ***** .***** * . * * .

S. cerevisiae 2026 LLSPTITIKSYHLWPSFTDEQWITIE SQMRDLILTEYGRKYNVNISALTQTEIKDIIILG
A. thaliana 1978 LKPDKSVVTEPHHIWPSLTDDQWMKVEVALRDLILSDYAKKNVNNTSALTQSEIRDIIILG
H. sapiens 1954 LKPKTITITEPHHIWPTLTDEEWIKVEVQLKDLILADYGKKNVN NVASLTQSEIRDIIILG
C. elegans 1947 LKPKTITITEPHHIWPTLSDDDWIKVELALKDMILADYGKKNVN NVASLTQSEVRDIIILG
consensus * * . * * . * . * . * * . * * . * * . * * . * * . * * . * * . * * .

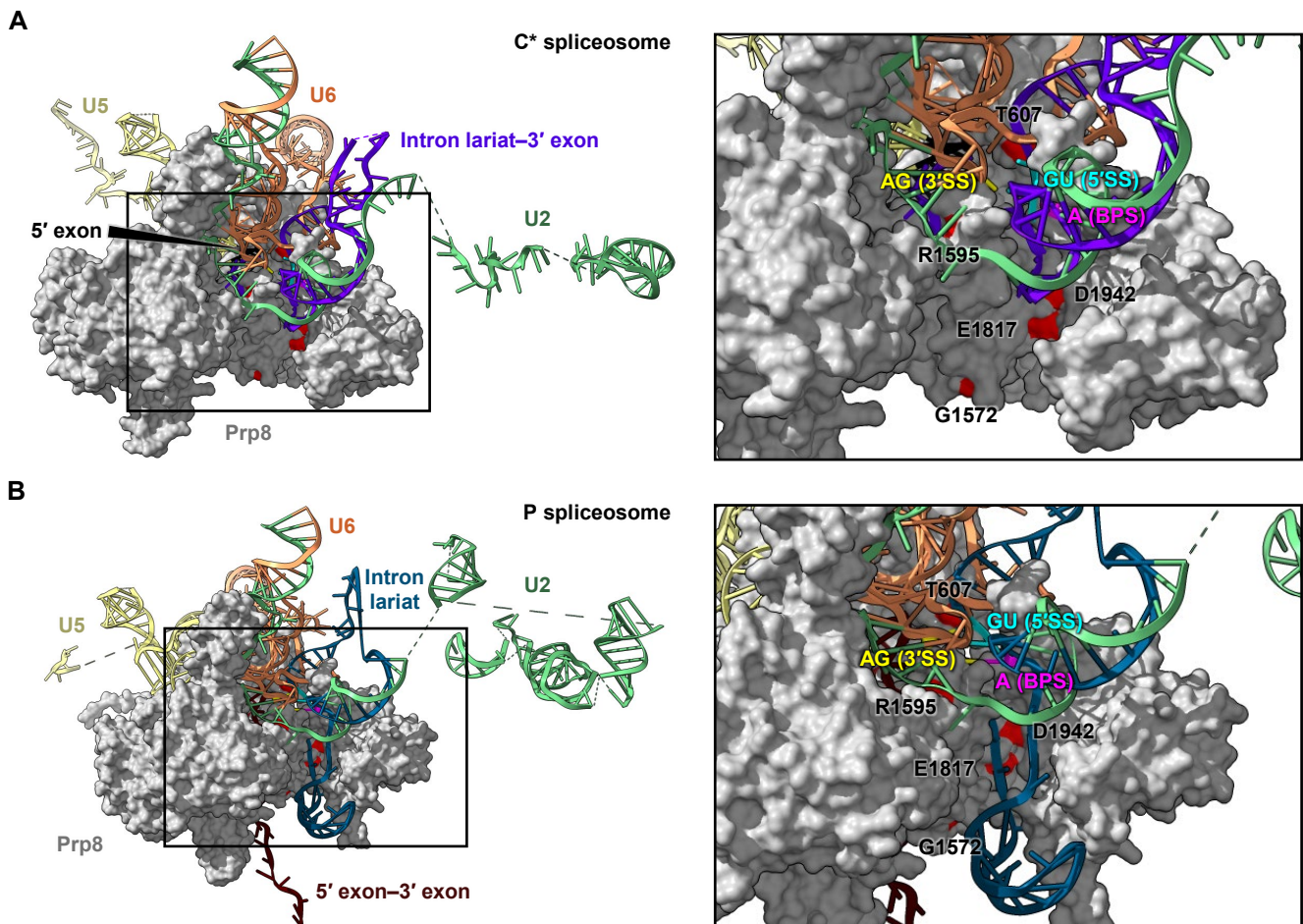
S. cerevisiae 2086 QNIKAPSVKRQKMAE LEAARSEKQNDEEAGASTVMKTKTINAQGEIIVVASADYESQTS
A. thaliana 2038 AEITPPSQQRQIAEIEKQAKEA-----SQLTAVTRTRTNVHGDELIVTTTSPYEQSA
H. sapiens 2014 MEISAPSQQRQIAEIEKQTKEQ-----SQLTATQTRTVNKHGDEIITSTISNYETQT
C. elegans 2007 MEISAPSQQRQIADIEKQTKEQ-----SQVTATTTTRTVNKHGDEIITATISNYETAS
consensus * * * * * * . . * * * * * . * * * * * . * * . * * .

<i>S. cerevisiae</i>	2146	FSSKNEWRKSAIANTLLYLRLKNIYVSADDEVEEQNVYVLPKNLLKKFIEISDVKIQVAA
<i>A. thaliana</i>	2091	FGSKTDWRVRAISATNLVLRVNHIIYVNSDDIKETGYTYIMPKNILKKFICVADLRTQIAG
<i>H. sapiens</i>	2067	FSSKTEWRVRAISAANLHLRTNHIYVSSDDIKETGYTYILPKNVLKKFICISDLRAOIAG
<i>C. elegans</i>	2060	FASRTEWRVRAISSSTNLHLRTQHIYVNSDDVKTDTGYTYILPKNILKKFITISDLRTQIAG
consensus		* * . ** ** * ** *** ** . * . ** . ** . . . * . . .
<hr/>		
<i>S. cerevisiae</i>	2206	FIIYGMSAKDHPKVEIKTVVLPQLGHVGSVQIISNIPDIIGDLPDTEGLELGLWIHTQTEE
<i>A. thaliana</i>	2151	YLYGISPPDNEQVKEIRCVVMVPQWGNHQLVHLPSS--LPEHDFLNDLEPLGWLHTQPNL
<i>H. sapiens</i>	2127	YLYGVSPPDNEQVKEIRCIIVMPQWGTHTQTVHLPQ--LPQHEYLKEMEPLGWIHTQPNL
<i>C. elegans</i>	2120	FMYGVSPPDNEQVKEIRCIIVLPQTGSHQOVNLPQ--LPDHELLRDFEPLGWMHTQPNL
consensus		. . ** . * * * ** . . * ** * * . . . * * * * * * *
<hr/>		
<i>S. cerevisiae</i>	2266	LKFMAASEVATHSKLFADK-----RDCIDISIFSTPGSVLSAYNLTDEGYQWGEENKDI
<i>A. thaliana</i>	2209	LPQLSPQDVTSHSRILENNKQWDGEKCIILTCSFTPGSCSLTTSYKLTQTGYEWGRLNKDN
<i>H. sapiens</i>	2185	SPQLSPQDVTTHAKIMADNPSWDGEKTIITCSFTPGSCTLTAYKLTTPSGYEWGRQNTDK
<i>C. elegans</i>	2178	LPQLSPQDVTTHAKILLTDNISWDGEKTVMITCSFTPGSVSLTAYKLTTPSGYEWGKANTDK
consensus		. . * . * * * * * * . * . * * * * * * *
<hr/>		
<i>S. cerevisiae</i>	2322	MNVLSEGFEPFSTHAQLLSDRITGNFIIPSGNVWNYTFMGTAFNQEGDYNFKYGIPLLE
<i>A. thaliana</i>	2269	GS-NPHGYLPTHYERVOQLLSDRFLGFFYMPVPSGPNYSFTGVKHTLSMKYSVKLGSPKE
<i>H. sapiens</i>	2245	GN-NPKGYLPSHYERVOQLLSDRFLGFFMVPAQSSWNYNFMGVRHDPNMKYELQLANPKE
<i>C. elegans</i>	2238	GN-NPKGYMPTHYERVOQLLSDRFLGYFMVPSNGVWNYNEQGRWSPAMKEDVCLSNPKE
consensus		* . * . . * . * * * * * * . . . * * * * * * . . . * * *
<hr/>		
<i>S. cerevisiae</i>	2382	FYNEMHRPVHFLQFSELAGDEEL-EAEQIDVFS
<i>A. thaliana</i>	2328	FYHEEHRPTHFLEFSNMEED-ITEGDREDTFT
<i>H. sapiens</i>	2304	FYHEVHRPSSHFLNFALLQEGE-VYSADREDLYA
<i>C. elegans</i>	2297	YYHEDHRPVHFNFKAFDDPLGTGSADREDAFA
consensus		. * * * * * * * . . . * . .

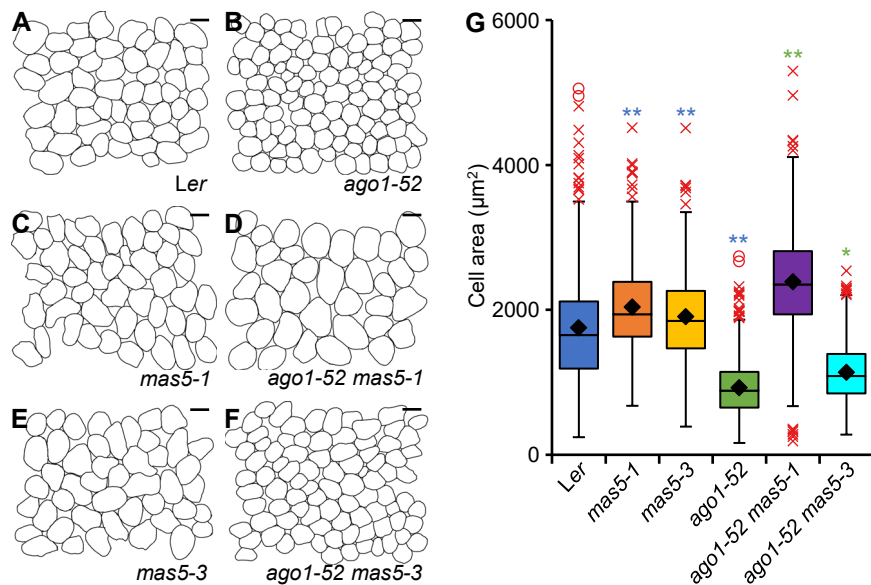
Supplementary Figure S2. Sequence conservation among PRP8 orthologs. Multiple amino acid sequence alignment of full-length PRP8 orthologs from *Saccharomyces cerevisiae* (UniProtKB accession number P33334), *Arabidopsis thaliana* (Q9SSD2), *Homo sapiens* (Q6P2Q9), and *Caenorhabditis elegans* (P34369). Identical and similar residues are shaded in black or gray, respectively. Asterisks and dots in the consensus line indicate identical and conserved residues, respectively. Numbers indicate residue positions. Predicted amino acid changes caused by the mutations studied in this work are highlighted by the following colors: red if they suppress binding of the spliceosome to mutated SSs; green if they suppress novel SSs; light blue if they suppress the effects of mutations in other genes encoding components of the splicing machinery; and purple if they are loss-of-function mutations. Amino acid substitutions in *Arabidopsis* are indicated in brackets. The domains of PRP8 are underlined with the same colors used in Figure 1J. This multiple sequence alignment was obtained using ClustalW2 (74) and shaded with Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).



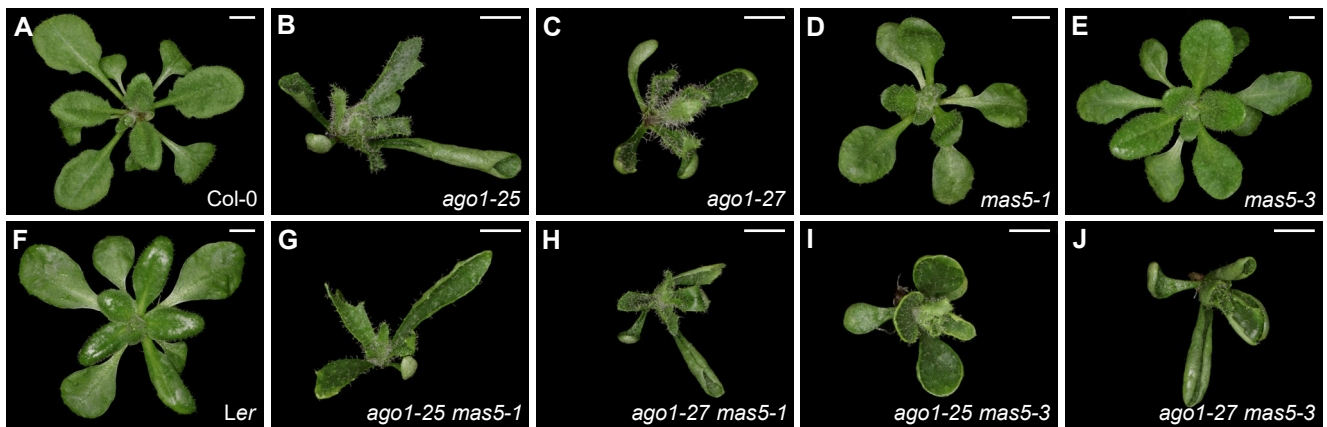
Supplementary Figure S3. Localization in yeast Prp8 of the homologous residues affected by the Arabidopsis *mas5* mutations in the cryo-EM structures of the splicing active site in B^{act}, B^{*} and C spliceosomes. 3D structures of the core of (A) B^{act}, (B) B^{*} and (C) C yeast spliceosomal complexes, determined by cryo-EM at an average resolution of 2.5, 3.7 and 2.8 Å, respectively. The models include U2, U5, and U6 snRNAs (U5 is hidden by Prp8 in B^{act} and B^{*} complexes), a single-intron pre-mRNA (in B^{act} and B^{*} complexes), or the 5' exon and intron lariat-3' exon (in C complex). The conserved GU of the 5'SS (GUAUGU) and A of the BPS (UACUAAC) are highlighted in cyan and magenta, respectively. The surface of yeast Prp8 is also shown, with the conserved residues affected by the Arabidopsis *mas5-1* and *mas5-2* (E1817), *mas5-3* (R1595; not presented in B^{*} complex), *mas5-4* (T607), *mas5-5* (D1942), and *mas5-6* (G1572) mutations marked in red. Structures of yeast spliceosomes were obtained from the Protein Data Bank (<https://www.rcsb.org/>; B^{act} complex PDB: 7DCO, Bai *et al.*, 2021; B^{*} complex PDB: 6J6Q, Wan *et al.*, 2019; C complex PDB: 7B9V, Wilkinson *et al.*, 2021) and visualized using the ChimeraX 1.2.5 software (<https://www.rbvi.ucsf.edu/chimerax/>).



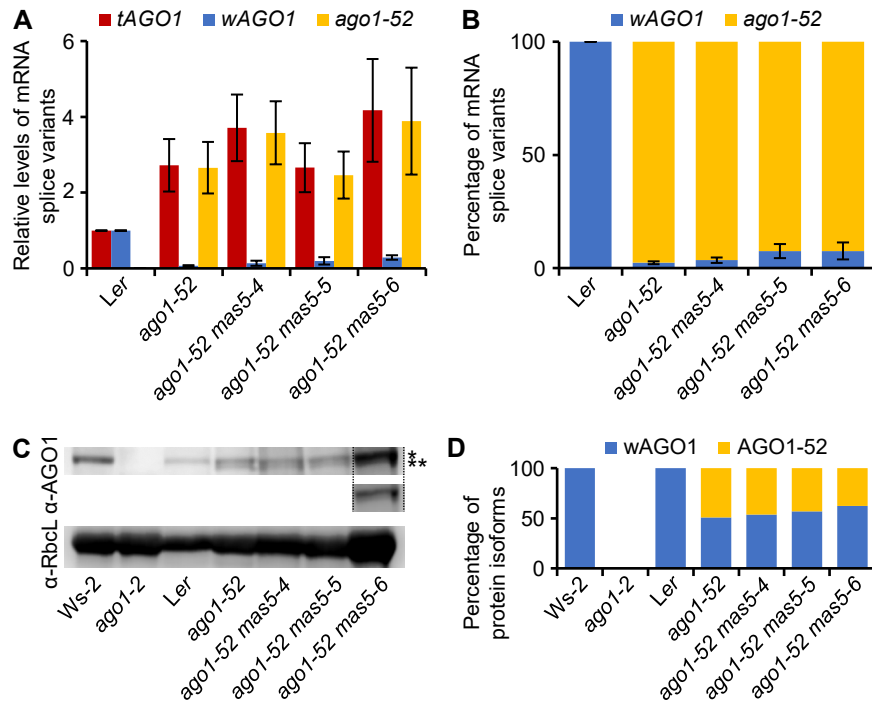
Supplementary Figure S4. Localization in yeast Prp8 of the homologous residues affected by the Arabidopsis *mas5* mutations in the cryo-EM structures of the splicing active site in C* and P spliceosomes. 3D structures of the core of (A) C* and (B) P yeast spliceosomal complexes, determined by cryo-EM at an average resolution of 4.0 and 3.3 Å, respectively. The models include three snRNAs (U2, U5, and U6), and the 5' exon and intron lariat-3' exon (in C* complex), or the spliced exons (5' exon-3' exon) and intron lariat (in P complex). The conserved GU of the 5'SS (GUAUGU), A of the BPS (UACUAAC), and AG of the 3'SS (UAG) are highlighted in cyan, magenta and yellow, respectively. The surface of yeast Prp8 is also shown, with the conserved residues affected by the Arabidopsis *mas5-1* and *mas5-2* (E1817), *mas5-3* (R1595), *mas5-4* (T607), *mas5-5* (D1942), and *mas5-6* (G1572) mutations highlighted in red. Structures of yeast spliceosomes were obtained from the Protein Data Bank (<https://www.rcsb.org/>; C* complex PDB: 5WSG, Yan *et al.*, 2017; P complex PDB: 6BK8, Liu *et al.*, 2017) and visualized using the ChimeraX 1.2.5 software (<https://www.rbvi.ucsf.edu/chimerax/>).



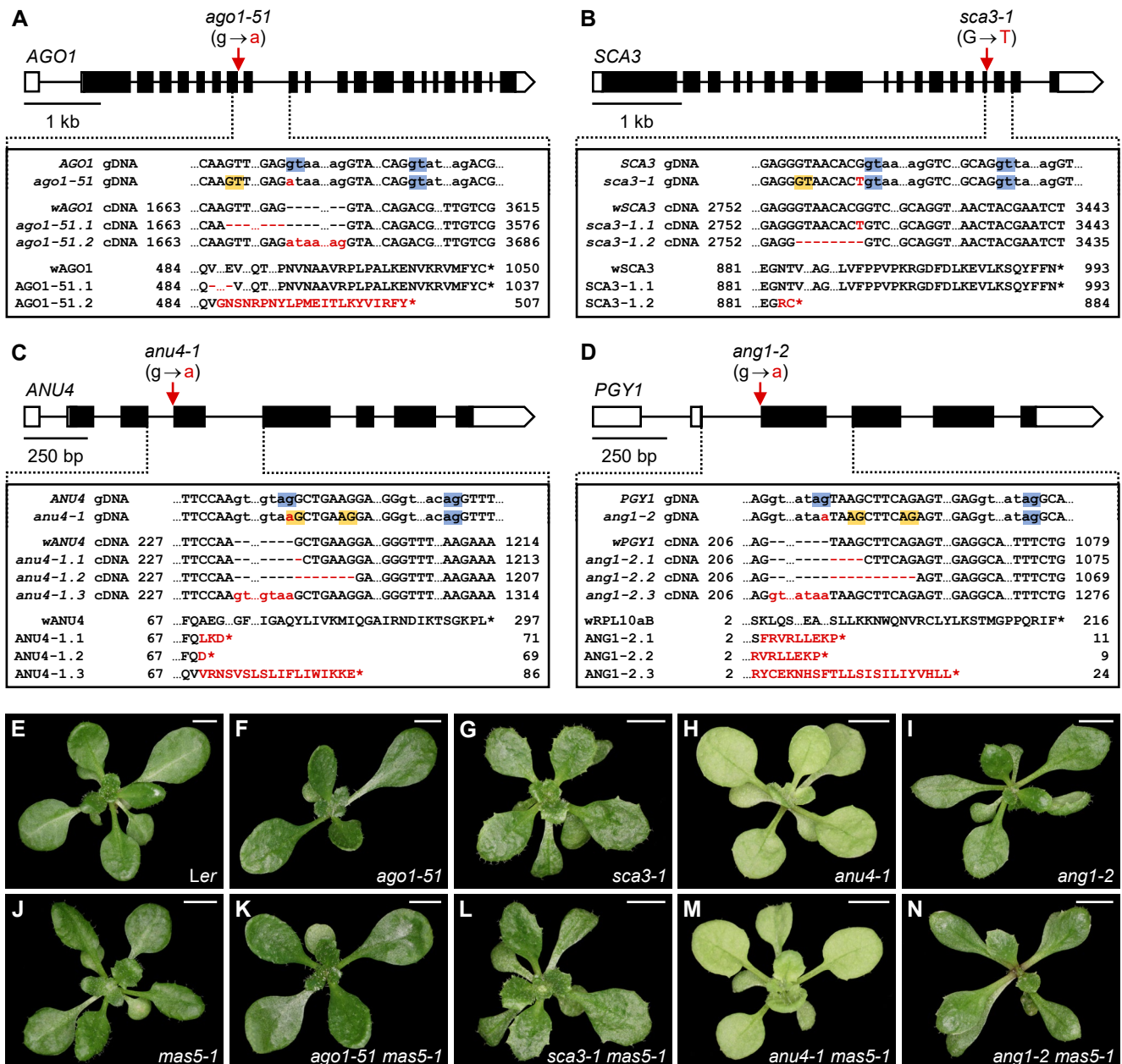
Supplementary Figure S5. Morphological phenotypes of leaf palisade mesophyll cells in *ago1-52*, *mas5-1*, *mas5-3*, and their double mutant combinations. (A–F) Diagrams of palisade mesophyll cells of (A) *Ler*, (B) *ago1-52*, (C) *mas5-1*, (D) *ago1-52 mas5-1*, (E) *mas5-3*, and (F) *ago1-52 mas5-3* plants. Scale bars: 40 µm. (G) Boxplot showing the distribution of cell areas in palisade mesophyll cells from first-node leaves. 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 $Q3 + 1.5 \times IQR$ or no less than $Q1 - 1.5 \times IQR$, where the interquartile range (IQR) is $Q3 - Q1$. ◆: Mean. —: Median. ×: Outliers. ○: Extreme minimum ($<Q1 - 3 \times IQR$) or maximum ($>Q3 + 3 \times IQR$) outliers. Asterisks indicate significant differences from the corresponding parental plants (indicated by color) in a Student's *t*-test (* $P < 0.01$ and ** $P < 0.001$). At least 290 cells per genotype were measured from plants collected 21 das.



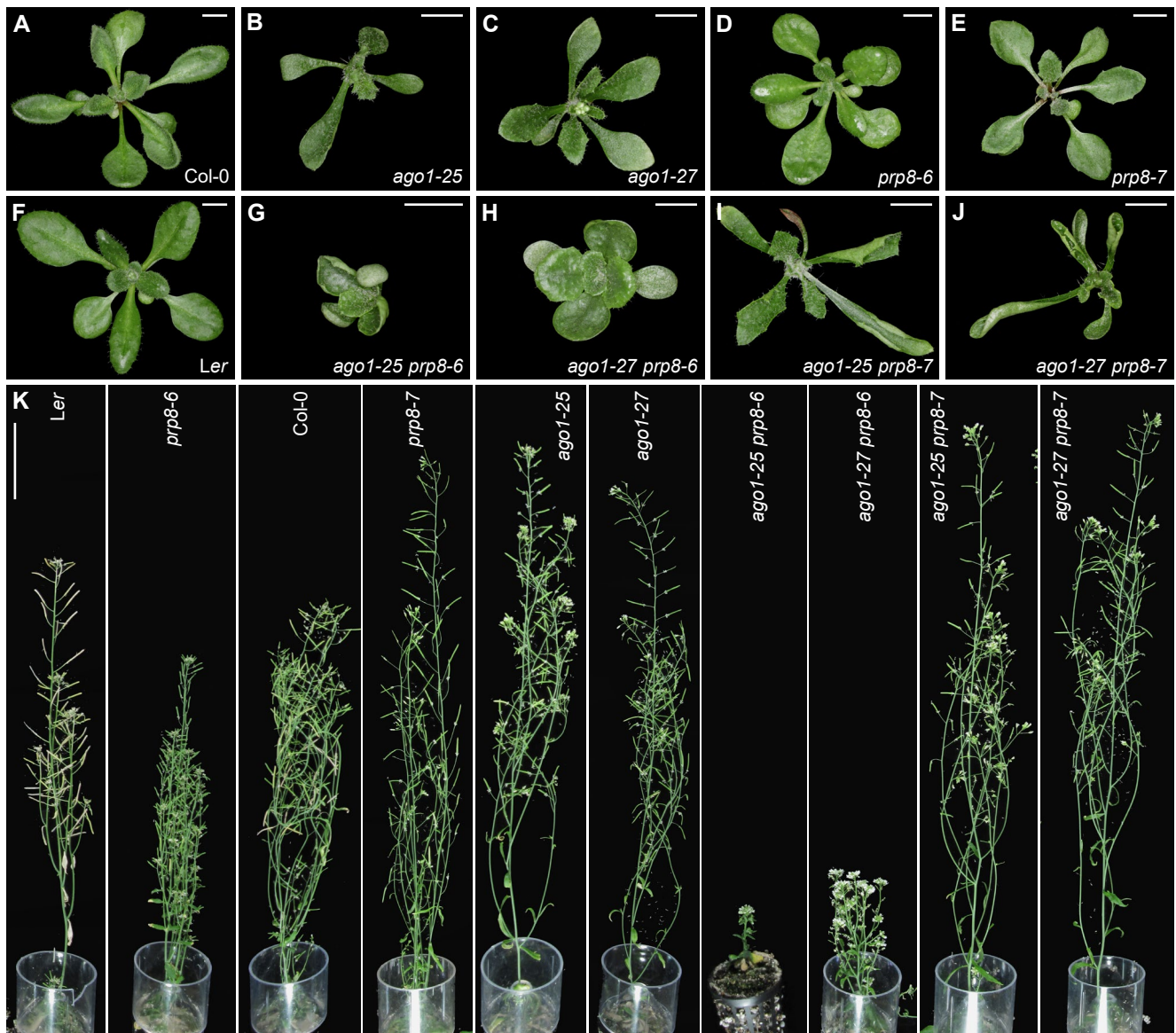
Supplementary Figure S6. Genetic interactions of *mas5-1* and *mas5-3* with *ago1-25* and *ago1-27*. Rosettes of (A) Col-0, (B) *ago1-25*, (C) *ago1-27*, (D) *mas5-1*, (E) *mas5-3*, (F) Ler, (G) *ago1-25 mas5-1*, (H) *ago1-27 mas5-1*, (I) *ago1-25 mas5-3*, and (J) *ago1-27 mas5-3* plants. Photographs were taken 21 das. Scale bars: 4 mm.



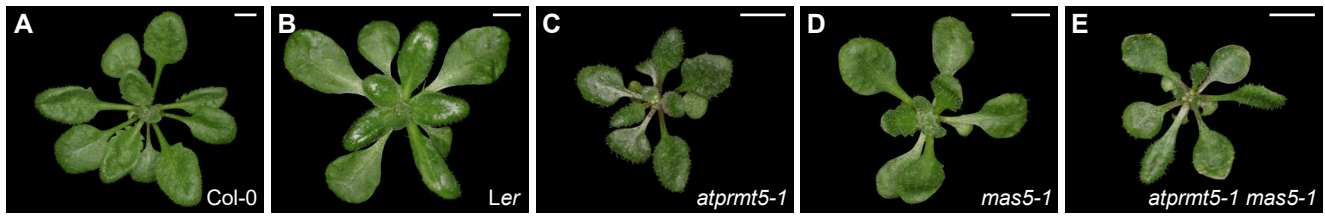
Supplementary Figure S7. Molecular phenotypes of the double mutant combinations of *ago1-52* with *mas5-4*, *mas5-5*, and *mas5-6*. (A) RT-qPCR analysis of the relative expression of total (*tAGO1*), wild-type (*wAGO1*) and mutant (*ago1-52*) AGO1 mRNA splice variants. (B) Percentage of *wAGO1* and *ago1-52* splice variants. Error bars indicate standard deviations. (C) Detection of AGO1 protein isoforms by immunoblot using a primary antibody against AGO1 (α -AGO1). Asterisks indicate the wild-type AGO1 (*) and mutant AGO1-52 (**) proteins. Two views of the bands from the *ago1-52 mas5-6* AGO1 protein sample are shown, corresponding to different exposure time, which allowed *wAGO1* and AGO1-52 proteins to be distinguished. Detection of the RuBisCO large subunit with α -RbcL was used as a loading control. (D) Relative quantification of *wAGO1* and AGO1-52 proteins from (C), using the Image Studio Analysis software (LI-COR). Total RNA and proteins were extracted from plants collected 15 das.



Supplementary Figure S8. Morphological and molecular phenotypes of *ago1-51*, *sca3-1*, *anu4-1*, *ang1-2*, and their double mutant combinations with *mas5-1*. (A–D) Schematic representation of the gene structures of (A) *AGO1*, (B) *SCA3*, (C) *ANU4*, and (D) *PGY1*. The positions of the (A) *ago1-51*, (B) *sca3-1*, (C) *anu4-1*, and (D) *ang1-2* mutations are indicated, and their effects on pre-mRNA splicing and mRNA translation into the (A) *AGO1*, (B) *SCA3*, (C) *ANU4*, and (G) *PGY1* proteins are shown. Gene structures, point mutations, and other molecular changes are represented as described in the legends of Figures 1 and 2. (E–N) Rosettes of (E) *Ler*, (F) *ago1-51*, (G) *sca3-1*, (H) *anu4-1*, (I) *ang1-2*, (J) *mas5-1*, (K) *ago1-51 mas5-1*, (L) *sca3-1 mas5-1*, (M) *anu4-1 mas5-1*, and (N) *ang1-2 mas5-1* plants. Photographs were taken 21 das. Scale bars: 4 mm.



Supplementary Figure S9. Genetic interactions of *prp8-6* and *prp8-7* with *ago1-25* and *ago1-27*. (A–J) Rosettes of (A) Col-0, (B) *ago1-25*, (C) *ago1-27*, (D) *prp8-6*, (E) *prp8-7*, (F) Ler, (G) *ago1-25 prp8-6*, (H) *ago1-27 prp8-6*, (I) *ago1-25 prp8-7*, and (J) *ago1-27 prp8-7*. (K) From left to right, adult plants of Ler, *prp8-6*, Col-0, *prp8-7*, *ago1-25*, *ago1-27*, *ago1-25 prp8-6*, *ago1-27 prp8-6*, *ago1-25 prp8-7*, and *ago1-27 prp8-7*. Photographs were taken (A–J) 21 and (K) 49 das. Scale bars: (A–J) 2 mm, and (K) 5 cm.



Supplementary Figure S10. Genetic interaction between *mas5-1* and *atprmt5-1*. Rosettes of (A) Col-0, (B) Ler, (C) *atprmt5-1*, (D) *mas5-1*, and (E) *atprmt5-1 mas5-1* plants. Photographs were taken 21 das. Scale bars: 4 mm.

A *mas5-1*

AT5G18210:E003	ttcgatctgtagcagCAAGACAAA
AT4G37570:E010	taatataaaaagagGTGCCCAA
AT1G32080:E012	ttcttgaattagtagAGAGAGGCG
AT1G19750:E004	ttcgtctttagcagGTTGCAGAT
AT4G02860:E013	tcattattttagcagTTACGGATT
AT5G19220:E019	tttgatgttcagcagAGGATTAG
AT2G45690:E004	tttctggtccagcagCAGCTTCAA
AT1G30220:E003	ctgcatcattagcagGGTAAGCAG
AT4G22310:E003	tgattggttagtgacaaactcagTTGGCTA
AT2G24310:E010	tatggcacatagcagATGAAGAGG
AT1G66810:E010	aaaccaatgaagcagAGGAAAAG
AT5G10340:E013	atattgatctagtcaggagATTT
AT1G61970:E002	tgtcgacttcagcagAAGTATTTG

B *mas5-3*

AT3G17710:E014	aatTTTTgttagcagTTAGCTGTGATG
AT1G65430:E029	ctTTTTgttagcagGGAAGCTGGAAC
AT4G37870:E010	taatataaaaagagGTGCCCAAGACA
AT4G02860:E013	tcattatTTtagcagTTACGGATTGTT
AT5G15870:E011	cactttgtaaagcagATAATGAGCCAA
AT1G34160:E005	ttgTTTTgttagcagAGCTATTACAAT
AT2G33140:E004	ctTTTTgggagatggaacagTTGGTT
AT4G09360:E009	tgttgTgaacagcagAACACGAGCAGC
AT4G22570:E002	attcaattggagataatgtagCAGCAG
AT5G49640:E003	tgatattgaaagcagGAGGTGTTGAG
AT1G66810:E010	aaaccaatgaagcagAGGAAAAGAAG
AT4G15570:E002	aaggtccattagatacgcagGTGCAA
AT5G50910:E009	ttgaaactgaagcagCCCTCTCGTTC
AT1G30220:E003	ctgcatcattagcagGGTAAGCAGGGC
AT4G28430:E007	ctttgcatctagcagCTTGAAGATTGT
AT4G40130:E002	tccattgTcagcagGTAGCAAAGTGT

Supplementary Figure S11. Alternative 3'SS events in *mas5* mutants. (A–B) DNA sequences corresponding to the statistically significant decreased Alt 3'SS events identified in (A) *mas5-1* and (B) *mas5-3* with an absolute Delta PSI >5% and <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 exón number of each gene.

Supplementary Table S1. Primer sets used for the fine mapping of *mas5-1*

Marker name	Locus	Oligonucleotide sequence (5'→3')		PCR product size (bp)	
		Forward primer	Reverse primer	Ler	Col-0
CER461530	AT1G79050-AT1G79060	GTTGGAGGTTGGTAATCATTAAT	GAAATGAATGCTGTTATCTTAATG	228	246
CER461138	AT1G79520	CAGGGCCTATGCACAAGATC	TCTCTAGAAGTGTCTTGCCC	267	250
CER469862	AT1G79600	GCCCCAAAACCTCGGCGAAA	ACGGCACGCCAAAAGTGTGG	210	220
CER469807	AT1G79830-AT1G79840	CAGCGAGAGGCATTGACCGA	CATGTGTCCAGGTGGCATTAT	183	202
CER449040	AT1G80100	GTGCTCGTAGGGTTCGTAAC	TACCTCCAGTCCTCTCAAGC	188	210
CER470312	AT1G80550	GCAATCCGAACTCGATTGAGT	ATTTTCACCTCTGGATGGCTC	233	258

Supplementary Table S2. Other primers used in this work

Purpose	Oligonucleotide name(s)	Oligonucleotide sequence (5'→3')		
		Forward primer (F)	Reverse primer (R)	
Genotyping of	<i>mas5-1, mas5-2, prp8-7</i>	PRP8_F1/R1	TGGGCTTGATCTGGCATACAA	CTTCAGGAACAACCTGCCAG
	<i>mas5-3, mas5-6</i>	PRP8_F2/R2	AAAATAGGCGTCTTACACTGGAA	ACATATCCTTCGACTCTGCAAC
	<i>mas5-4</i>	PRP8_F3/R3	TGAAAGGAAGGTTTATGATGATGA	CTGGCATAGCATCAAGGACGT
	<i>mas5-5</i>	PRP8_F4/R4	TGTCTATAGCAAGGATAATCCTAA	ACGAGGAATTCCCCTATCCCA
	<i>prp8-6</i>	PRP8_F5/R5	TGGGCTTGATCTGGCATACAA	CTGGAACGGAAGCTGCAGCT
	<i>ago1-27, ago1-52</i>	AGO1_F2/R2	TTACCACGTTCTTTGGGATGAG	GCAGTAGAACATGACACGCTTC
	<i>ago1-25</i>	AGO1_F1 ^a /R1 ^a	GGCTAATATGAGTCTTCTCTGC	CCATCCCTGTGCAGAATAACC
	<i>ago1-51</i>	AGO1_F3 ^a /R3	ATCGACAGCCTTCATAGAGGC	CAACAATCTTGCATACCTGTGCAGCAAC
	<i>sca3-1</i>	AT2G24120_F9 ^b /R10 ^b	TGCAGAAGTGAAAGACATCTG	AACAGCAGTCATCATCATGTG
	<i>anu4-1</i>	AT1G02280_F ^c /R ^c	GATATGAACTCAATGACAGTTCTT	AAGACATCTATTGTTCTGTTTACA
	<i>ang1-2</i>	PGY1_F1/R1	GCCCATGAAACGAAATCAAT	GCGTCTGCCGTTTTTCTTTA
	<i>icu13</i>	AXR6.3_F ^d /R ^d	TCAAGTGCAGAACTACTTGCAACA	TAGGACAGTCTCCATATGGACT
	<i>atprmt5-1</i>	AT4G31120_F1/R1	TCATCCATTTGGCAGGTAA	TGGCCTTTTGAGATGAAAGAG
	<i>sar1-4</i>	AT1G33410_F1/R1	CAGCCCTTGAGCAAGTAGATG	TTACCTGAATATGTCCACCCG
T-DNA insertion verification	Salk_LBb1.3 ^e	GCGTGGACCGCTTGCTGCAACT		
RT-qPCR	AGO1_F6g/AGO1_R6 ^a	GGCATGATAAAGGAGTTGCTCAT	CTGACTCCATCCCTGTAGAAGA	
	AGO1_F4g/AGO1_R7 ^a	TTTACTGCAGATGGACTTCAATC	TAATATGCAGGGGGAACAATT	
	AXR6-1F ^d /1R ^d	TTTCGCTGAGTTCTACAGGAAGA	TGCTTGAGCTTTGTCAGGATACT	
	AXR6-2R ^d		GTGTCAAATCCGTCACCATGC	
	qACT2F/R	GCACCCTGTTCTTCTTACCG	ATCCAGCACAATACCGTTGTA	

Sequences taken from ^a(S1), ^b(S2), ^c(S3), ^d(S4), and ^e<http://signal.salk.edu/tdnaprimers.2.html>.

Supplementary Table S3. Summary of mRNA sequencing data information

Library	Raw reads ^a	Raw data ^b	Aligned reads ^c	Effective reads ^d
<i>Ler</i> 1	49911821 + 49911821	15.0	48494215 (97.2%)	21339388 (44.0%)
<i>Ler</i> 2	51483696 + 51483696	15.4	50128590 (97.4%)	24204849 (48.3%)
<i>Ler</i> 3	55674517 + 55674517	16.7	53811627 (96.6%)	20418267 (38.0%)
<i>mas5-1</i> 1	49504170 + 49504170	14.9	47059088 (95.1%)	22169683 (47.1%)
<i>mas5-1</i> 2	55899607 + 55899607	16.8	53866068 (96.4%)	26952987 (50.0%)
<i>mas5-1</i> 3	50209666 + 50209666	15.1	48778617 (97.1%)	22908231 (47.0%)
<i>mas5-3</i> 1	49297201 + 49297201	14.8	47964189 (97.3%)	22426299 (46.8%)
<i>mas5-3</i> 2	56360185 + 56360185	16.9	54449193 (96.6%)	25618505 (47.0%)
<i>mas5-3</i> 3	49596269 + 49596269	14.9	48100665 (97.0%)	24206176 (50.3%)

^aTotal number of reads of raw data; as it is the result of a paired-end sequencing, it equals the number of read 1 and read 2. ^bTotal number of raw reads multiplied by the sequence length (150 bp), expressed in Gbp. ^cTotal number and percentage of reads mapped on the *Ler* reference genome (NCBI accession GCA_001651475.1). ^dTotal number and percentage of reads after removing optical duplicates and secondary alignments.

Supplementary Table S4. Amino acid substitutions caused in Arabidopsis PRP8 by the *prp8* and *mas5* mutations, and their homologous residues in yeast Prp8 and human PRP8

Arabidopsis	<i>Saccharomyces cerevisiae</i>	Human	Domain of the PRP8
Mutation (predicted effect)	Residue (mutations)*	Residue	protein affected
<i>prp8-9</i> (P347S) ^a	P395	P324	-
<i>mas5-4</i> (T555I)	T607	T532	-
<i>prp8-8</i> (P1141S) ^a	P1191L/S/T (<i>prp8-cat</i>) ^b	P1118	Reverse transcriptase-like
<i>prp8-10</i> (G1347D) ^c	G1395	G1323	Linker
<i>mas5-6</i> (G1524R)	G1572	G1500	Linker
<i>mas5-3</i> (R1547K)	R1595	R1523	Linker
<i>mas5-1</i> (E1769K)	E1817G (<i>D-135</i>) ^d	E1745	Endonuclease-like
<i>prp8-11</i> (R1770K) ^c	R1818	R1746	Endonuclease-like
<i>prp8-7</i> (G1820E) ^e	G1868	G1796	RNase H-like
<i>prp8-6</i> (G1891E) ^f	A1939	G1867	RNase H-like
<i>mas5-5</i> (D1894N)	D1942	D1870	RNase H-like

*Mutations described in *Saccharomyces cerevisiae* that are equivalent to Arabidopsis *prp8-8* (*prp8-cat*; the three mutations shown share the same name) and *mas5-1* (*D-135*). References: ^a(S5), ^b(S6), ^c(S7), ^d(S8), ^e(S9), and ^f(S10).

SUPPLEMENTARY REFERENCES

- S1. 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.
- S2. 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.
- S3. 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.
- S4. 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.
- S5. 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. USA*, **113**, 5447-5452.
- S6. Kuhn, A.N. and Brow, D.A. (2000) Suppressors of a cold-sensitive mutation in yeast U4 RNA define five domains in the splicing factor Prp8 that influence spliceosome activation. *Genetics*, **155**, 1667-1682.
- S7. Kanno, T., Lin, W.D., Fu, J.L., Matzke, A.J.M. and Matzke, M. (2017) A genetic screen implicates a CWC16/Yju2/CCDC130 protein and SMU1 in alternative splicing in *Arabidopsis thaliana*. *RNA*, **23**, 1068-1079.
- S8. 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.
- S9. 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.
- S10. 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.

1 **Cross-kingdom conservation of Arabidopsis RPS24**
2 **function in 18S rRNA maturation**

3
4
5 **Adrián Cabezas-Fuster, Rosa Micol-Ponce, Raquel Sarmiento-**
6 **Mañús, and María Rosa Ponce**

7
8
9 Instituto de Bioingeniería, Universidad Miguel Hernández, Campus de Elche, 03202
10 Elche, Alicante, Spain.

11
12 Running title: RPS24A and RPS24B function in 18S rRNA maturation

13
14 Corresponding author: María Rosa Ponce (telephone: +34 96 665 8503; fax: +34 96
15 665 8511; e-mail: mrponce@umh.es)

16 ABSTRACT

17 All 81 ribosomal proteins (RPs) that form the *Arabidopsis* (*Arabidopsis thaliana*) 80S
18 ribosome are encoded by several paralogous genes. For example, the nearly identical
19 RPS24A and RPS24B proteins are encoded by *RPS24A* and *RPS24B*, respectively.
20 Here we explored the functions of RPS24A and RPS24B in *Arabidopsis*. Their encoding
21 genes exhibit combined haploinsufficiency, as at least two wild-type copies of either
22 *RPS24A* or *RPS24B* are required for plant viability and at least three are required for
23 normal plant development. Loss-of-function of either gene caused a pointed-leaf
24 phenotype, a typical phenotype of null or hypomorphic recessive alleles of genes
25 encoding ribosome biogenesis factors (RBFs) or RPs. We also found that RPS24A and
26 RPS24B act as RBFs during early stages of 18S ribosomal RNA (rRNA) maturation, as
27 loss of RPS24A or RPS24B function reduced the 18S/25S rRNA ratio. An RPS24B-GFP
28 fusion protein predominantly localized to the nucleolus, as expected. The *rps24b-2*
29 mutation strengthened the phenotypes of the RBF mutants *mRNA transporter4-2* and
30 *small organ4-3*, which are defective in 5.8S rRNA maturation. This synergistic interaction
31 might be an effect of increased 45S rDNA transcription, which we also observed in the
32 *rps24* mutants. Therefore, the *Arabidopsis* RPS24 proteins act as RBFs during 18S
33 rRNA maturation, like their human and yeast putative orthologs. Only two plant RPs were
34 previously shown to act not only as structural components of the ribosome but also as
35 RBFs. We provide evidence that RPS24 proteins also regulate 45S rDNA transcription,
36 which has not been described for their yeast or human orthologs.

37

38 KEYWORDS

39 RPS24; ribosome biogenesis; *Arabidopsis*; pre-rRNA processing; 45S rDNA expression,
40 18S rRNA maturation.

41 INTRODUCTION

42 The 80S cytoplasmic ribosome (hereafter, the ribosome) of *Arabidopsis* (*Arabidopsis*
43 *thaliana*) is composed of four ribosomal RNAs (rRNAs) and 81 ribosomal proteins (RPs;
44 Wilson and Cate, 2012), consisting of 33 RPs in its small (40S) subunit and 48 in its large
45 (60S) subunit (Barakat et al., 2001). In plants, rRNAs are encoded by 5S and 45S rDNA
46 genes, which are transcribed in the nucleoplasm and the nucleolus by RNA polymerase
47 III or RNA polymerase I, respectively. All eukaryotic genomes include hundreds of rDNA
48 genes arranged in tandem in several loci. Each 45S rDNA gene is composed (from its 5'
49 to 3' end) of a promoter and a transcriptional unit that contains the 5' external transcribed
50 spacer (5'-ETS), the 18S, 5.8S and 25S rRNA sequences (in plants and yeast; 28S in
51 animals), and a 3'-ETS. These three regions encoding rRNAs are separated by two
52 internal transcribed spacers: ITS1 is located between the sequences of the 18S and 5.8S
53 rRNAs, while ITS2 is located between the sequences of the 5.8S and 28S/25S rRNAs
54 (Supplemental Figure 1).

55 The processing of 45S (47S in animals and 35S in yeast [*Saccharomyces*
56 *cerevisiae*]) pre-rRNA is a multistep procedure involving more than 100 ribosome
57 biogenesis factors (RBFs), which carry out endo- and exo-nucleolytic cleavages and
58 chemical modifications to generate 28S/25S, 18S and 5.8S mature rRNAs. This intricate
59 rRNA production consists of two partially redundant pathways, which are named
60 according to the region of the early 35S(P) pre-rRNA in which the first endonucleolytic
61 cleavage occurs: the 5'-ETS-first and ITS1-first pathways. There is an additional pathway
62 in plants, the ITS2-first pathway (Supplemental Figure 1; Palm et al., 2019). The
63 functions of many RBFs are partially or fully conserved across animals, fungi and plants
64 (reviewed in Sáez-Vásquez and Delseny, 2019).

65 Yeast mRNA transport4 (Mtr4) is an ATP-dependent RNA helicase that acts as
66 a cofactor of the nucleolar exosome, which has 3'→5' exonuclease activity. Loss-of-
67 function mutations of the yeast *Mtr4* gene impair 5.8S and 25S rRNA maturation through
68 the ITS1-first pathway and leads to an imbalance in the ratio of 40S/60S ribosomal
69 subunits (Thoms et al., 2015). *Arabidopsis* MTR4 is required for 18S and 5.8S rRNA
70 maturation and to eliminate 5'-ETS processing by-products produced in the ITS1-first
71 pathway (Lange et al., 2011). Yeast Mtr4 interacts with Nucleolar protein 53 (Nop53),
72 which has orthologs in humans (Glioma Tumor-Suppressor Candidate Region 2;
73 GLTSCR2) and *Arabidopsis* (SMALL ORGAN 4; SMO4), all of which are involved in
74 equivalent steps of 5.8S rRNA maturation (Tafforeau et al., 2013; Micol-Ponce et al.,
75 2020). Ribosomal RNA processing protein 7 (Rrp7) is another yeast RBF involved in 18S
76 rRNA processing that functions in the ITS1-first pathway, as do its human (RRP7A) and

77 Arabidopsis (RRP7) orthologs (Micol-Ponce et al., 2018; Farooq et al., 2020;
78 Supplemental Figure 1).

79 Most RPs that form part of the 40S ribosomal subunit also function as RBFs in
80 the nuclear and cytoplasmic steps of 18S rRNA maturation in yeast, including Ribosomal
81 protein S24 (Rps24; Ferreira-Cerca et al., 2005). Human RPS24 and yeast Rps24 play
82 analogous roles in 5'-ETS processing. Loss-of-function mutations of the yeast *Rps24*
83 and human *RPS24* orthologs lead to the accumulation of 23S and 30S pre-rRNAs,
84 respectively (which are equivalent to Arabidopsis P-A₃ pre-rRNA) and a reduction in the
85 amounts of human 21S and 18S-E, and yeast 21S and 20S pre-rRNAs, which are
86 equivalent to Arabidopsis 18S-A₃ and 20S, respectively. All of these pre-rRNAs are
87 precursors of the 18S rRNA produced by the ITS1-first pathway (Supplemental Figure
88 1). Since this pathway is the major contributor to 18S rRNA production, its level is
89 reduced by human and yeast *rps24* loss-of-function mutations (Ferreira-Cerca et al.,
90 2005; Choismel et al., 2008).

91 Arabidopsis *MORPHOLOGY OF ARGONAUTE1-52 SUPPRESSED 2 (MAS2)* is
92 the ortholog of the human gene encoding NF- κ -B-activating protein (NKAP), which
93 functions in transcriptional repression and splicing in animals (Pajerowski et al., 2009;
94 Burgute et al., 2014; Sánchez-García et al., 2015). We previously identified physical
95 interactors of MAS2 in a yeast two-hybrid assay (Sánchez-García et al., 2015), including
96 RRP7 (Micol-Ponce et al., 2018), SMO4 (Micol-Ponce et al., 2020) and RPS24B. Here,
97 we studied the functions of the Arabidopsis co-orthologs of yeast *Rps24* and human
98 *RPS24*: AT3G04920 and AT5G28060, which encode RPS24A and RPS24B,
99 respectively. We established that both RPS24A and RPS24B act as RBFs in the ITS1-
100 first pathway, like their human and yeast orthologs. Our results also suggest that
101 RPS24A and RPS24B participate in the transcriptional repression of 45S rDNA, a role
102 that has not been proposed for any of their orthologs.

103 **RESULTS**104 ***Arabidopsis RPS24A* and *RPS24B* show combined haploinsufficiency**

105 The *Arabidopsis RPS24A* (AT3G04920) and *RPS24B* (AT5G28060) genes are very
 106 similar in structure, with 5 exons, and their protein products sharing 96% sequence
 107 identity (Supplemental Figure 2). We initially focused our study on *RPS24B*, since we
 108 previously showed that it interacted with *MAS2* in a Y2H assay (Sánchez-García et al.,
 109 2015). Two putatively null alleles of *RPS24B*, *rps24b-1* and *rps24b-2*, were previously
 110 identified as mutants with impaired leaf development; *rps24b-1* harbors a 72-bp deletion
 111 in the coding region of its fifth exon (Horiguchi et al., 2006; Horiguchi et al., 2011), and
 112 *rps24b-2* carries a T-DNA insertion in its third intron (Figure 1A; Horiguchi et al., 2011;
 113 Wang et al., 2018). Both mutants exhibited an almost identical mild pointed-leaf
 114 phenotype (Figure 1F; Horiguchi et al., 2011; Wang et al., 2018).

115 The *apiculata6* (*api6*) mutant was previously isolated in a large-scale screen for
 116 leaf morphological mutants from an EMS mutagenized population in the *Ler* background
 117 (Berná et al., 1999). Using iterative linkage analysis (Ponce et al., 1999; Ponce et al.,
 118 2006), we mapped the *api6* mutation to a genomic region containing 49 genes flanked
 119 by the molecular markers *cer449133* (between AT5G27905 and AT5G27910) and
 120 *cer451402* (in AT5G28200; Supplemental Table 1). Sanger sequencing of AT5G28060
 121 (*RPS24B*), the most plausible candidate gene, revealed an EMS-type G→A transition in
 122 the last nucleotide of the first intron (Figure 1A). Since this nucleotide forms part of a 3'
 123 splicing site (3'SS) of the first intron, the *api6* mutation is predicted to cause the absence
 124 of the second exon of the gene—which is 42-nt in length—from the mature mRNA of
 125 *api6* (Figure 1B and C). The protein produced by the *api6* mutant is predicted to lack 14
 126 residues present in wild-type *RPS24B* (amino acids 25-38; Supplemental Figure 2). An
 127 *api6* × *rps24b-2* cross confirmed that these mutants are allelic (Figure 1E, F, H-J). Like
 128 many mutants in the *Ler* genetic background, the leaf phenotype of *api6* was stronger
 129 than that of *rps24b-1* or *rps24b-2*, which are in the *Col-0* background (Pérez-Pérez et
 130 al., 2009).

131 We also studied SALK_126799, a line that carries a T-DNA insertion in the third
 132 intron of *RPS24A*, which we named *rps24a-1* (Figure 1D and G). Homozygous *rps24a-1*
 133 plants were viable and exhibited a pointed-leaf phenotype, which was milder than that
 134 of *rps24b-2* and *api6* (Figure 1E-I). *RPS24A* and *RPS24B* exhibited similar protein
 135 abundance patterns, as determined in the *Arabidopsis Thaliana* Expression Atlas
 136 (Athena) database
 137 (http://athena.proteomics.wzw.tum.de:5002/master_arabidopsisshiny/; Mergner et al.,
 138 2020), with similar levels throughout all organs and developmental stages, with a few
 139 exceptions (Supplemental Dataset 1 [highlighted in red]). The almost constitutive

Cabezas-Fuster et al., Figure 1

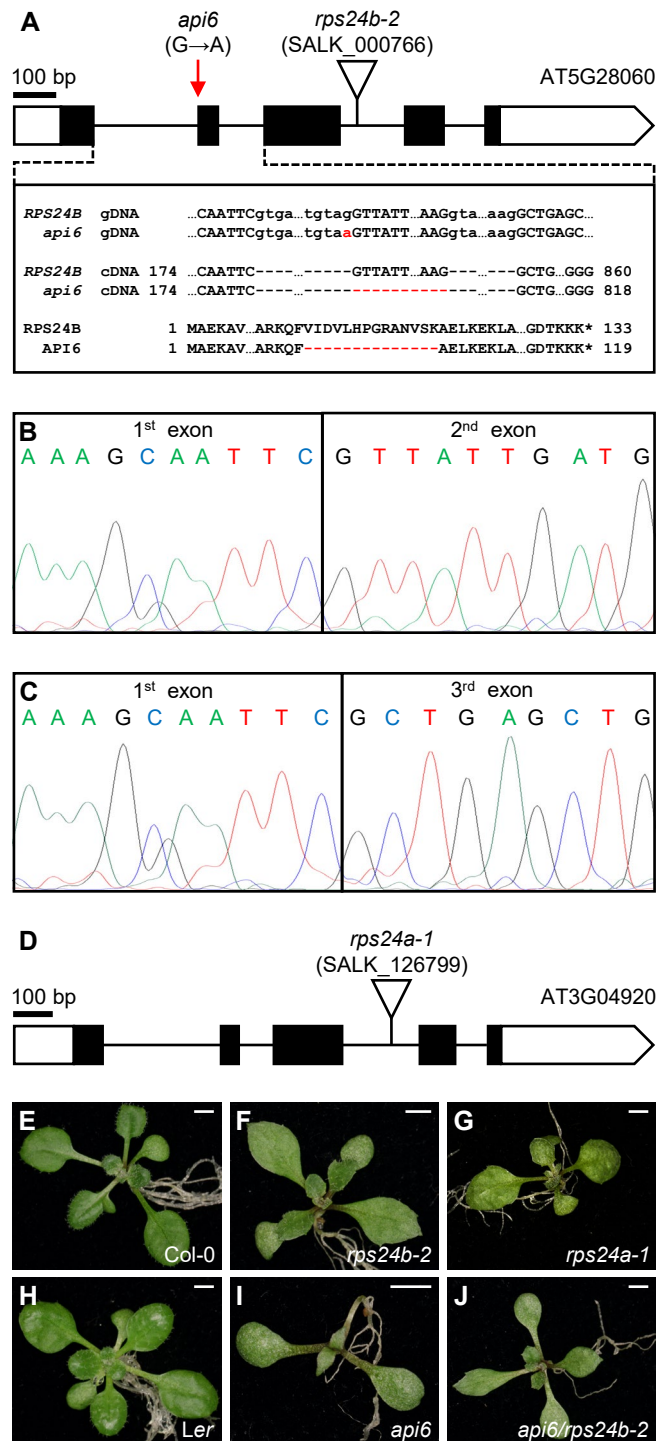


Figure 1. Alleles of the *RPS24A* and *RPS24B* genes studied in this work. (A) Schematic representation of the *RPS24B* locus, with the positions and nature of its mutations indicated. Exons and introns are depicted as boxes and lines, respectively, and white boxes represent untranslated regions (UTRs). A red arrow and a triangle mark the point mutation of *api6* and the T-DNA insertion of *rps24b-2*, respectively. The molecular changes caused by *api6* and *rps24b-2* are shown in red in the corresponding cDNAs and predicted proteins. (B and C) Electropherograms showing the cDNA sequences of *RPS24B* in Ler (B) and *api6* (C). Total RNA was extracted from seedlings collected 15 days after stratification (das). (D) Structure of the *RPS24A* gene, represented as described in A. (E-J) Rosettes of Col-0 (E), *rps24b-2* (F), *rps24a-1* (G), Ler (H), *api6* (I) and *api6/rps24b-2* F₁ seedlings (J). Scale bars, 2 mm. Photographs were taken 14 das.

140 expression of *RPS24A* and *RPS24B* (at a lower level compared to other genes encoding
 141 RPs) was also described by Savada and Bonham-Smith (2014) using microarray data
 142 from the GENEVESTIGATOR platform (Hruz et al., 2008).

143 Since the mutant phenotypes of *rps24a-1* and *rps24b-2* pointed to the functional
 144 redundancy of *RPS24A* and *RPS24B*, we performed *rps24a-1* × *rps24b-2* reciprocal
 145 crosses. All *RPS24A/rps24a-1;RPS24B/rps24b-2* F₁ plants exhibited a mutant
 146 phenotype indistinguishable from that of their parents, suggesting non-allelic non-
 147 complementation (Figure 2A-D). Selfing of such diheterozygotes generated F₂ progeny
 148 with only two phenotypic classes: 221 wild-type and 161 mutant (indistinguishable from
 149 their parents) plants. We genotyped 59 F₂ phenotypically mutant plants, finding 10
 150 *rps24a-1/rps24a-1;RPS24B/RPS24B*, 17 *RPS24A/RPS24A;rps24b-2/rps24b-2* and 32
 151 *RPS24A/rps24a-1;RPS24B/rps24b-2* plants. We identified no *rps24a-1/rps24a-1;*
 152 *RPS24B/rps24b-2* or *RPS24A/rps24a-1;rps24b-2/rps24b-2* sesquimutants or double
 153 mutants among the F₂ progeny (Figure 2A-D); the absence of these genotypes suggests
 154 their lethality. In agreement with the proposed early lethality of the genotypes with fewer
 155 than two wild-type doses of either *RPS24A* or *RPS24B* (expected to be 31.25% of
 156 seeds), the siliques of several diheterozygous F₁ plants exhibited a rate of 27% aborted
 157 or unfertilized ovules compared to Col-0 (Figure 2E-H).

158 To determine whether *rps24a-1*, *rps24b-2* and *api6* were hypomorphic or null
 159 alleles and to test the existence of any dosage compensation mechanism for the
 160 expression of *RPS24A* or *RPS24B*, we analyzed *RPS24B* mRNA levels by RT-PCR. We
 161 used primers upstream and downstream of the T-DNA insertion of *rps24b-2*
 162 (Supplemental Figure 3 and Supplemental Table 2). *RPS24B* mRNA levels were very
 163 similar in *api6* and *Ler*, but *RPS24B* mRNA levels were undetectable in *rps24b-2* and
 164 were not higher in *rps24a-1* than in the wild type (Supplemental Figure 3). We did not
 165 determine whether *rps24a-1* was null, because it was impossible to design RT-PCR
 166 primers specific for this allele that would amplify the region flanking its T-DNA insertion,
 167 due to its sequence similarity with *rps24b-2*. However, the phenotype of the single
 168 mutants and the lethality of the reciprocal sesquimutants strongly suggest that *rps24a-1*
 169 and *rps24b-2* are null alleles.

170

171 **RPS24B predominantly localizes to the nucleolus**

172 Since ribosomal subunits are synthesized and assembled in the nucleolus and the
 173 nucleoplasm, but their final maturation occurs in the cytoplasm, not few RPs have been
 174 detected in all three subcellular compartments (Pendle et al., 2005; Palm et al., 2016;
 175 Montacié et al., 2017; Ayash et al., 2021). As expected from their known roles as
 176 structural components of the ribosome, Arabidopsis *RPS24A* and *RPS24B* have been

Cabezas-Fuster et al., Figure 2

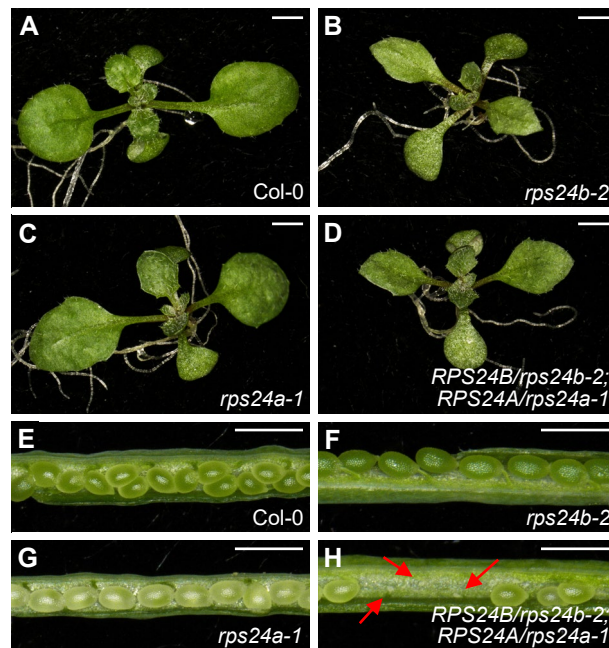


Figure 2. Genetic evidence for the combined haploinsufficiency of *RPS24A* and *RPS24B*. (A-D) Rosettes of Col-0 (A), *rps24b-2* (B), *rps24a-1* (C) and (D) *RPS24B/rps24b-2;RPS24A/rps24a-1* plants. (E-H) Siliques of (E) Col-0, (F) *rps24b-2*, (G) *rps24a-1* and (H) *RPS24B/rps24b-2;RPS24A/rps24a-1* plants. Scale bars, 2 mm (A-D), or 1 mm (E-H). Photographs were taken (A-D) 14 or 39 (E-H) das.

177 found in the cytoplasm in several proteomic studies (Chang et al., 2005; Giavalisco et
 178 al., 2005; Carroll et al., 2008; Hummel et al., 2015), and their abundances appear to be
 179 similar, at least in ribosomes purified from cell cultures (Salih et al., 2020). RPS24A and
 180 RPS24B have also been localized to the nucleolus and nucleoplasm (Pendle et al., 2005;
 181 Palm et al., 2016; Montacié et al., 2017; Ayash et al., 2021).

182 To visualize the subcellular localization of RPS24B, we generated the
 183 *35S_{pro}:RPS24B:GFP* construct and used it to transform Col-0 plants. We primarily
 184 detected green fluorescent protein (GFP) fluorescence in the nucleolus, but also in the
 185 cytoplasm (Figure 3), as observed for other duplicated Arabidopsis RPs of the small and
 186 large subunits: RPS3aA/B, RPS8A/B, RPL7aA/B, RPL15A/B and RPL23aA/B (Savada
 187 and Bonham-Smith, 2014). Three protein-localization predictor tools provided in silico
 188 support to our experimental findings with the *35S_{pro}:RPS24B:GFP* transgenic plants.
 189 LOCALIZER (<https://localizer.csiro.au/>; Sperschneider et al., 2017) identified a putative
 190 nuclear localization signal (NLS) in the C-terminal regions of RPS24A and RPS24B,
 191 which matches the nucleolar localization signal (NoLS) found using Nucleolar
 192 localization sequence Detector (NoD) software (Supplemental Figure 2;
 193 <http://www.compbio.dundee.ac.uk/www-nod/index.jsp>; Scott et al., 2011). MULocDeep
 194 (<https://www.mu-loc.org/>; Jiang et al., 2021), which is able to predict the localization of
 195 any protein in 44 suborganellar compartments, clearly predicted a nucleolar localization
 196 for both RPS24 proteins, in addition to a lower-likelihood prediction for a cytoplasmic and
 197 mitochondrial localization (Supplemental Figures 4 and 5). The predicted nucleolar
 198 localization of RPS24A and experimental confirmation for RPS24B suggest that both
 199 proteins function in early steps of ribosome biogenesis. However, although the
 200 subcellular localization of the RPS24B-GFP fusion protein was the same in transgenic
 201 plants in the Col-0 or *rps24b-2* background, the *35S_{pro}:RPS24B:GFP* transgene did not
 202 rescue the morphological defects caused by the *rps24b-2* mutation.

203

204 **45S pre-rRNA processing is delayed in the *rps24* mutants**

205 As previously mentioned, human and yeast RPS24 act as RBFs in 18S rRNA maturation.
 206 In both species, depletion of RPS24 results in inhibited 5'-ETS processing, the
 207 accumulation of pre-rRNAs that include 5'-ETS (human 30S and yeast 23S pre-rRNAs)
 208 and reduced levels of 21S and 18S-E pre-rRNAs in humans and their corresponding 21S
 209 and 20S pre-rRNAs in yeast (Supplemental Figure 1), all of which are 18S rRNA
 210 precursors (Ferreira-Cerca et al., 2005; Choesmel et al., 2008; Robledo et al., 2008).

211 Taken together, the current identification of RPS24B as a predominantly
 212 nucleolar factor and the previous finding of RPS24A and RPS24B in the Arabidopsis
 213 nucleolar proteome via large-scale analysis (Pendle et al., 2005; Palm et al., 2016) point

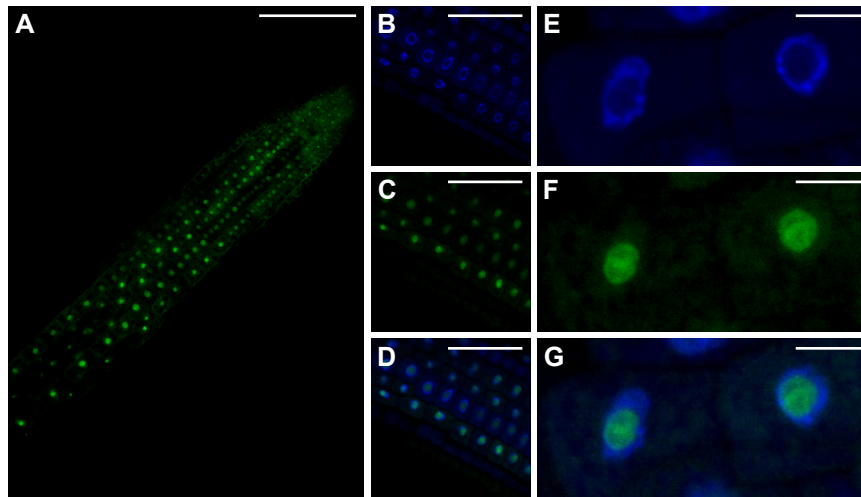
Cabezas-Fuster *et al.*, Figure 3

Figure 3. Subcellular localization of RPS24B. (A-G) Confocal laser-scanning micrographs of root cells from transgenic $35S_{pro}:RPS24B:GFP$ seedlings in the Col-0 background, collected 5 das. Fluorescent signals correspond to GFP (A, C, F), 4',6-diamidino-2-phenylindole (DAPI) (B, E), and their overlay (D, G). Scale bars, 100 μm (A), 50 μm (B-D), or 10 μm (E-G).

214 to a conserved role for these proteins as RBFs. To test this hypothesis, we analyzed 45S
 215 pre-rRNA processing in the *rps24* mutants. We carried out gel blot analysis of total RNA
 216 extracted from *rps24a-1*, *rps24b-2* and *api6*, which was hybridized with the S2, S7 and
 217 S9 probes. These probes are complementary to segments of the 5'-ETS, ITS1, and ITS2
 218 of 45S pre-rRNA, respectively, allowing several 25S, 18S and 5.8S rRNA precursors to
 219 be detected (Figure 4A). We used the *smo4-3* and *mtr4-2* mutants as controls (Lange et
 220 al., 2011; Micol-Ponce et al., 2020). Loss of SMO4 function causes nucleolar hypertrophy
 221 and the accumulation of P-A₃, a precursor of 18S rRNA, and perturbs 5.8S rRNA
 222 maturation (Micol-Ponce et al., 2020). MTR4 is an exosome cofactor involved in the
 223 processing of 5.8S pre-rRNA (like SMO4) but also in the degradation of P-P', a by-
 224 product derived from 5'-ETS processing (Lange et al., 2011).

225 Using the S7 probe, we detected the accumulation of P-A₃ and 35S_{A123B}/35S(P)
 226 pre-rRNAs in *rps24a-1*, *rps24b-2* and (especially) *api6* plants compared to their
 227 respective wild types. We also detected the depletion of 33S(P')/32S pre-rRNA and to a
 228 lesser extent P'-A₃/18S-A₃/18S-A₂ species, which result from the processing of
 229 35S_{A123B}/35S(P) and P-A₃ precursors, respectively (Figure 4B). P-A₃ pre-rRNA is
 230 generated by the cleavage of the A₃ site in 35S(P) pre-rRNA, the first 18S rRNA
 231 precursor in the ITS1-first pathway [Figure 4A and Supplemental Figure 1; reviewed in
 232 Sáez-Vásquez and Delseny (2019)]. We confirmed the accumulation of 35S_{A123B}/35S(P)
 233 and P-A₃ pre-rRNAs using the S2 probe, but we only observed the accumulation of the
 234 P-P' species in *mtr4-2* (Figure 4C), as previously described (Lange et al., 2011). Using
 235 the S9 probe, we did not find any alteration in 5.8S rRNA maturation in any of the three
 236 mutants studied (Figure 4D).

237 Reduced levels of the 21S and 18S-E precursors of 18S pre-rRNA (Supplemental
 238 Figure 1) associated with the lack of RPS24 activity in human cells result in a 40%
 239 reduction in mature 18S rRNA levels (Choismel et al., 2008; Robledo et al., 2008).
 240 Similar effects were found in the absence of RPS24 function in yeast: 35S, 32S and 23S
 241 pre-rRNAs accumulated, 20S pre-rRNA could not be detected, and the levels of mature
 242 18S rRNA were reduced by 95% compared to the wild type (Ferreira-Cerca et al., 2005).
 243 The impaired 18S pre-rRNA processing due to the loss of RPS24A and RPS24B function
 244 in Arabidopsis suggests that the levels of mature 18S rRNA might be reduced in these
 245 mutants. Therefore, we analyzed the rRNA profiles of the mutants using a bioanalyzer.
 246 The 18S/25S rRNA ratio was reduced by 4% in the *rps24a-1* mutant, by 27% in *rps24b-2*
 247 and by 32% in *api6* compared to their corresponding wild types (Supplemental Figure
 248 6). These results provide further evidence for a conserved role for RPS24A and RPS24B
 249 in the early steps of 18S rRNA maturation, particularly in the cleavage of sites internal to

Cabezas-Fuster et al., Figure 4

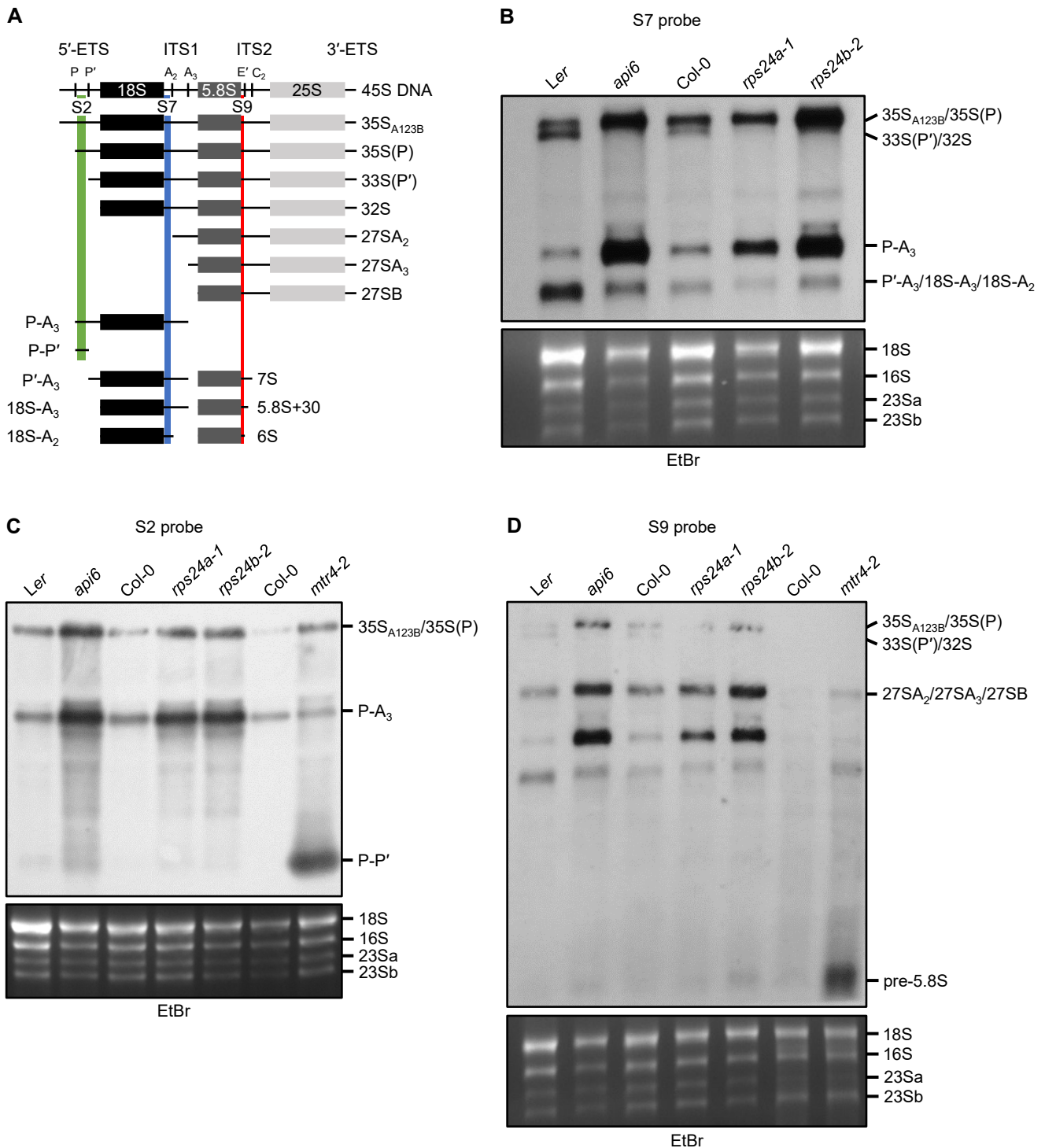


Figure 4. rRNA maturation in the *rps24a* and *rps24b* mutants. (A) Diagram of 45S pre-rRNA processing intermediates detected in RNA gel blots. The regions hybridizing to each probe are highlighted in green (S2 probe), yellow (S7 probe) or blue (S9 probe). Vertical bars in 45S rRNA indicate the endonucleolytic cleavage sites in pre-rRNAs that are relevant to this analysis. ETS and ITS indicate external and internal transcribed spacers, respectively. (B-D) Visualization of the processing of the 5.8S, 18S and 25S rRNA precursors by gel blot analysis. Total RNA from seedlings collected 15 das was separated on a formaldehyde-agarose gel, transferred to a nylon membrane and hybridized with the S7 (B), S2 (C), or S9 probe (D). EtBr, ethidium bromide-stained gels, visualized before blotting, serving as loading controls. Similar results were obtained in at least two independent experiments.

250 the 5'-ETS, as already observed for their yeast and human putative orthologs (Ferreira-
251 Cerca et al., 2005; Choemmel et al., 2008).

252

253 **The *rps24b-2* mutation enhances the morphological and molecular defects of**
254 **mutants in genes encoding RBFs that function in 18S and 5.8S rRNA maturation**

255 To genetically confirm the involvement of RPS24 in ribosome biogenesis, we obtained
256 double mutant combinations between *rps24b-2* and mutant alleles of *SMO4* and *MTR4*,
257 which participate in 5.8S rRNA maturation (Lange et al., 2011; Micol-Ponce et al., 2020)
258 and *RRP7* and *PARALLEL1* (*PARL1*; also named *NUCLEOLIN1*), which act in 18S rRNA
259 maturation (Pontvianne et al., 2007; Micol-Ponce et al., 2018).

260 All of the double mutants showed synergistic phenotypes (Figure 5). Leaves of
261 *rps24b-2 smo4-2* plants exhibited increased serration and took on a twisted appearance
262 into a spiral (Figure 5F). *rps24b-2 mtr4-2* plants exhibited delayed growth, lanceolate
263 leaves and strong drops in rosette size and plant height (Figure 5G). *rps24b-2 parl1-2*
264 plants strongly accumulated anthocyanins, exhibiting dark, serrated, pointed leaves and
265 small, compact rosettes (Figure 5H). The *rps24b-2 mtr4-2* and *rps24b-2 parl1-2* double
266 mutants completed their life cycles but had poor fertility (Figure 5I, J, M-P). We sowed
267 108 F₂ seeds from an *rps24b-2* × *rrp7-1* cross and isolated two double mutants, which
268 exhibited very strong mutant phenotypes (Supplemental Figure 7). Only one plant
269 completed its life cycle and produced 20 F₃ seeds, only one of which germinated, but the
270 seedling died a few days later, suggesting early postembryonic lethality of the *rps24b-2*
271 *rrp7-1* genotype.

272 We analyzed 45S pre-rRNA maturation in the viable double mutants, finding
273 similar levels of P-A₃ or P'-A₃/18S-A₃/18S-A₂ pre-rRNAs in *rps24b-2 smo4-3*, *rps24b-*
274 *2 mtr4-2* and *rps24b-2 parl1-2* compared to the single mutants using the S2 and S7
275 probes (Figure 6A and B). The *smo4-3* and *mtr4-2* single mutants accumulate 7S and
276 5.8S+10 pre-rRNAs, respectively (Lange et al., 2011; Micol-Ponce et al., 2020).
277 Unexpectedly, using the S9 probe, we detected higher levels of 7S pre-rRNA in *rps24b-*
278 *2 smo4-3* and of 5.8S+10 pre-rRNA in *rps24b-2 mtr4-2* compared to the *smo4-3* and
279 *mtr4-2* single mutants, even though *rps24b-2* did not accumulate either of these pre-
280 rRNA species (Figure 6C and D).

281 To assess the subcellular localization of 5.8S pre-rRNA in the *rps24b-2 smo4-3*
282 and *rps24b-2 mtr4-2* double mutants, we performed an RNA-FISH assay. We used the
283 23-nt long S9 probe, which hybridizes to the ITS2 region adjacent to the 5.8S rRNA
284 coding sequence, and allows detection of all 5.8S pre-rRNA, but not mature 5.8S rRNA
285 (Figure 4A). In agreement with the results of RNA gel blot analysis (Figure 6C and D),
286 we observed very low fluorescence in the nucleoli of wild-type Col-0 and Ler, higher

Cabezas-Fuster et al., Figure 5

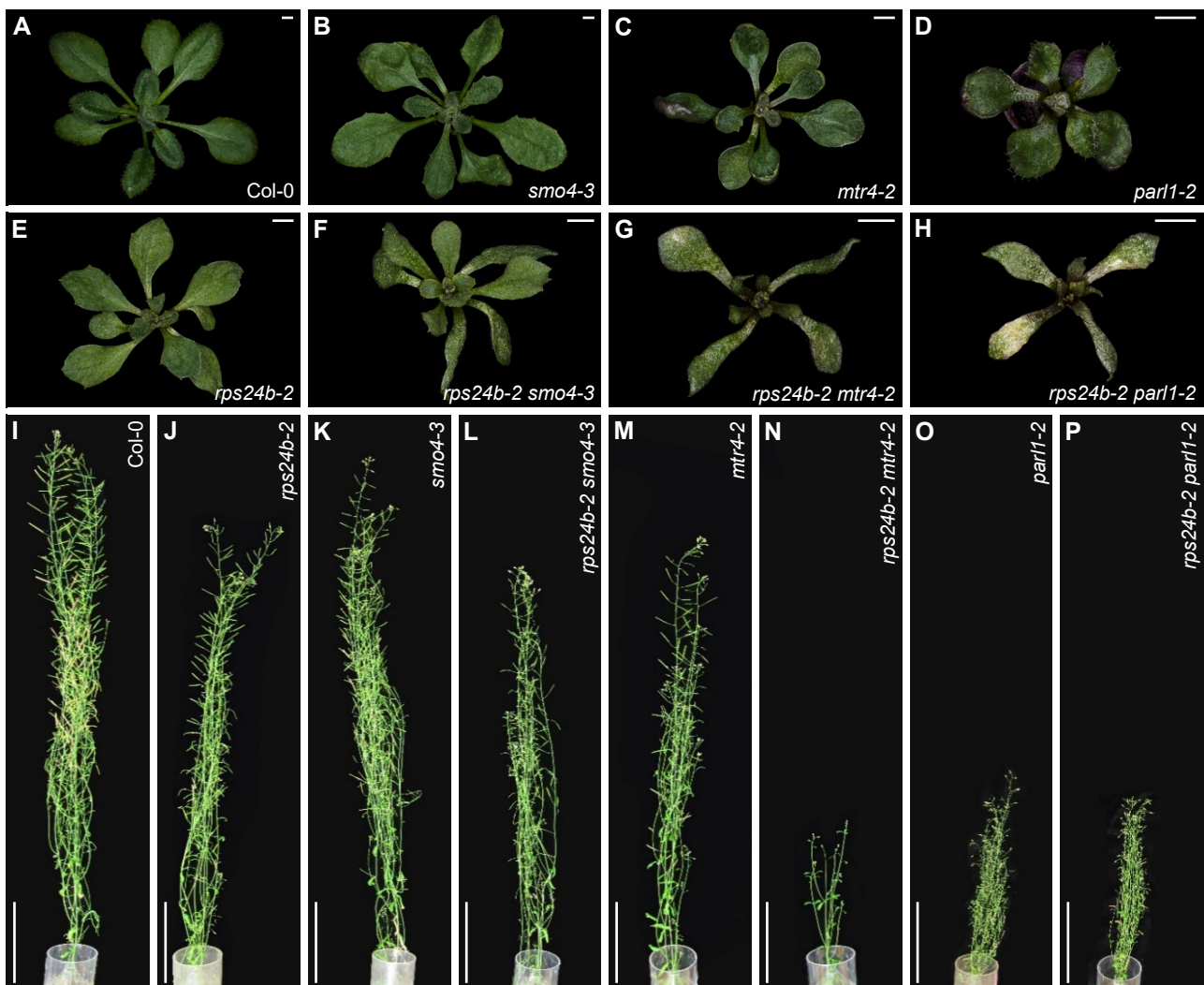


Figure 5. Genetic interactions between *rps24b-2* and mutant alleles of genes encoding RBFs. Rosettes of Col-0 (A), *smo4-3* (B), *mtr4-2* (C), *par11-2* (D), *rps24b-2* (E), *rps24b-2 smo4-3* (F), *rps24b-2 mtr4-2* (G) and *rps24b-2 par11-2* plants (H). From left to right, adult plants of Col-0 (I), *rps24b-2* (J), *smo4-3* (K), *rps24b-2 smo4-3* (L), *mtr4-2* (M), *rps24b-2 mtr4-2* (N), *par11-2* (O) and *rps24b-2 par11-2* (P). Photographs were taken 21 das (A-H) or 57 das (I-P). Scale bars, 2mm (A-H), or 5 cm (I-P).

287 fluorescence in the *api6*, *rps24a-1*, *rps24b-2*, *smo4-3* and *mtr4-2* single mutants, and
 288 much higher fluorescence in the *rps24b-2 smo4-3* and *rps24b-2 mtr4-2* double mutants
 289 (Supplemental Figure 8). The enhanced fluorescence that we observed in the nucleoli of
 290 *rps24* mutants compared to Col-0 and Ler is likely due to the accumulation of the
 291 35S_{A123B}/35S(P) species that we detected in RNA gel blots using the S7 and S9 probes
 292 (Figure 4B-D).

293 In the RNA-FISH assays, we also observed enlarged nucleoli in *rps24b-2* and
 294 (especially) *smo4-3 rps24b-2* and *mtr4-2 rps24b-2* double mutant plants compared to
 295 Col-0 (Supplemental Figure 8). Perturbations in the size and organization of the
 296 nucleolus are hallmarks of nucleolar stress, which occurs when ribosome biogenesis is
 297 defective, as previously observed in the nucleoli of *smo4* plants (Micol-Ponce et al.,
 298 2020). Therefore, we measured the areas of the nucleolus and nucleoplasm using the
 299 fluorescence of the S9 probe and of nuclei stained with DAPI. Both the nucleus and
 300 nucleolus of *api6* and *rps24b-2* (but not *rps24a-1*) were larger than those of the
 301 respective wild types (Supplemental Figure 8A2 and A3). Moreover, the nucleus and
 302 nucleolus were greatly enlarged in both the *smo4-3 rps24b-2* and *mtr4-2 rps24b-2*
 303 double mutants (Supplemental Figure 8D-F, J-L, P-R, V-A3).

304 To determine whether the organization of the nucleolus was perturbed in the
 305 *rps24* mutants, we used an antibody against the nucleolar marker fibrillarin to examine
 306 the roots of Col-0, *rps24a-1* and *rps24b-2* seedlings. We did not observe differences in
 307 nucleolar organization between the wild type and the mutants (Supplemental Figure 9),
 308 suggesting that the defective 45S pre-rRNA processing in the *rps24* mutants causes the
 309 enlargement but not disorganization of the nucleolus.

310 Taken together, these results suggest that RPS24B participates (directly or
 311 indirectly) in the maturation of 5.8S rRNA. Since the *rps24a-1* and *rps24b-2* single
 312 mutants do not accumulate any 5.8S pre-rRNA, RPS24B might play a lesser role in 5.8S
 313 maturation compared to SMO4 and MTR4; therefore, the absence of RPS24B is only
 314 perceived in the absence of SMO4 and MTR4. Another interpretation is that RPS24
 315 participates in the repression of 45S rDNA transcription. Under this second scenario, the
 316 accumulation of 5.8S pre-rRNAs would be more evident in the *rps24b* background than
 317 in the *smo4-3* and *mtr4-2* single mutant backgrounds due to the presence of more 5.8S
 318 pre-rRNAs to process.

319

320 **45S pre-rRNAs accumulate in the *rps24* mutants**

321 Kinetic analysis of Arabidopsis 45S pre-rRNA processing has revealed three rate-limiting
 322 steps: the processing of 35S, 32S, and P-A₃ pre-rRNAs; 32S and P-A₃ pre-rRNAs are
 323 the first two precursors in the two alternative pathways, the 5'-ETS first pathway and the

Cabezas-Fuster et al., Figure 6

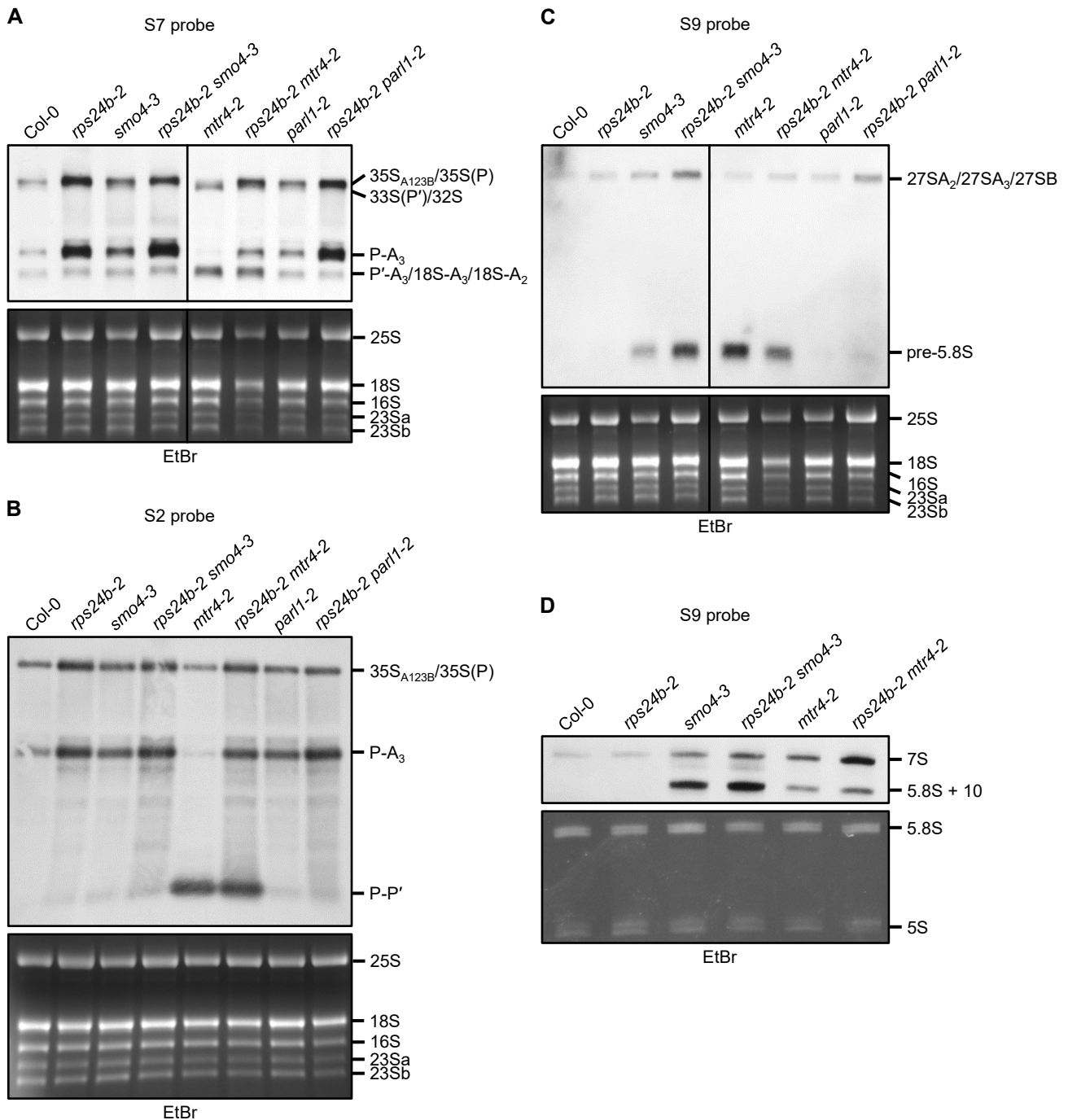


Figure 6. rRNA maturation in double mutants between *rps24b-2* and alleles of genes encoding RBFs. (A-D) Visualization of the processing of the 5.8S, 18S and 25S rRNA precursors using gel blot analysis. Total RNA from seedlings collected 15 das was separated on (A-C) formaldehyde-agarose or (D) polyacrylamide-urea gels, transferred to a nylon membrane and hybridized with the S7 (A), S2 (B), or S9 probe (C and D). EtBr, ethidium bromide-stained gels, visualized before blotting, serving as loading controls. Similar results were obtained in at least two independent experiments.

324 ITS1-first pathway, respectively (Supplemental Figure 1). The level of 35S pre-rRNA is
 325 the major indicator of the transcription rate of 45S rDNA (Shanmugam et al., 2021). Our
 326 results show that a reduction in RPS24 activity lower the rate of 45S pre-rRNA
 327 processing, as shown by the accumulation of pre-rRNAs produced by two of the rate-
 328 limiting steps, 35S_{A123B}/35S(P) and P-A₃. Our results also suggest that 45S rDNA
 329 transcription is upregulated in the *rps24* mutants.

330 Therefore, we performed RT-PCR to investigate the presence of the four different
 331 45S rDNA variants (*VARs*), which differ in their 3'-ETS (Figure 8A), using the p3+p4
 332 primer pair (Figure 7A and B; Supplemental Table 2). We did not detect any differences
 333 in the intensities of the PCR bands using genomic DNA as a template (Figure 7C).
 334 However, the intensities of the PCR bands corresponding to *VAR4*, *VAR2* and *VAR3*
 335 were higher in *rps24a-1*, and those corresponding to all four *VARs* were higher in *rps24b-*
 336 *2*, compared to the wild type (Figure 7D). Since changes in the relative abundance of
 337 45S rRNA variants are usually interpreted as changes in 45S rDNA transcription (Kojima
 338 et al., 2007; Pontvianne et al., 2010; Durut et al., 2014), these results suggest that
 339 RPS24B represses 45S rDNA expression.

340 To confirm these results, we carried out RT-qPCR analysis using the 45S pre-
 341 rRNA-45S-F and 45S pre-rRNA-45S-R primers, which hybridize to the 5'-ETS
 342 (Supplemental Table 2; Zhu et al., 2016). We detected higher levels of 45S pre-rRNA in
 343 the *rps24* mutants than in wild-type plants (Figure 7E).

344

345 **RPS24B-GFP co-precipitates with nucleolar factors involved in ribosome** 346 **biogenesis and pre-mRNA splicing**

347 To identify nucleolar interactors of RPS24B, we carried out co-immunoprecipitation (Co-
 348 IP) assays with an anti-GFP antibody in Col-0 plants harboring the *35S_{pro}:RPS24B:GFP*
 349 transgene. The raw list of 2,635 proteins that co-purified with RPS24B-GFP was filtered
 350 by considering only those classified as high-confidence proteins (q-value <0.01) with at
 351 least two peptide spectra matches (PSMs) in each of the three biological replicates
 352 (Supplemental Dataset 2). Several additional filters were used, which discarded proteins
 353 that co-immunoprecipitated with all four other proteins examined (none related to
 354 ribosome biogenesis), as we considered these to be spurious interactions. We analyzed
 355 the predicted suborganellar localizations of the interactors using MuLOCDeep (Jiang et
 356 al., 2021), discarding those that were not predicted to be nucleolar proteins, which
 357 reduced the list of putative interactors to 25 proteins (Supplemental Dataset 2). We
 358 refined the annotations of these proteins using the TAIR or Aramemnon database and
 359 those identified in two independent nucleolar proteomic analyses (Palm et al., 2016;
 360 Montacié et al., 2017). In addition, we added the pathways that were enriched in genes

Cabezas-Fuster et al., Figure 7

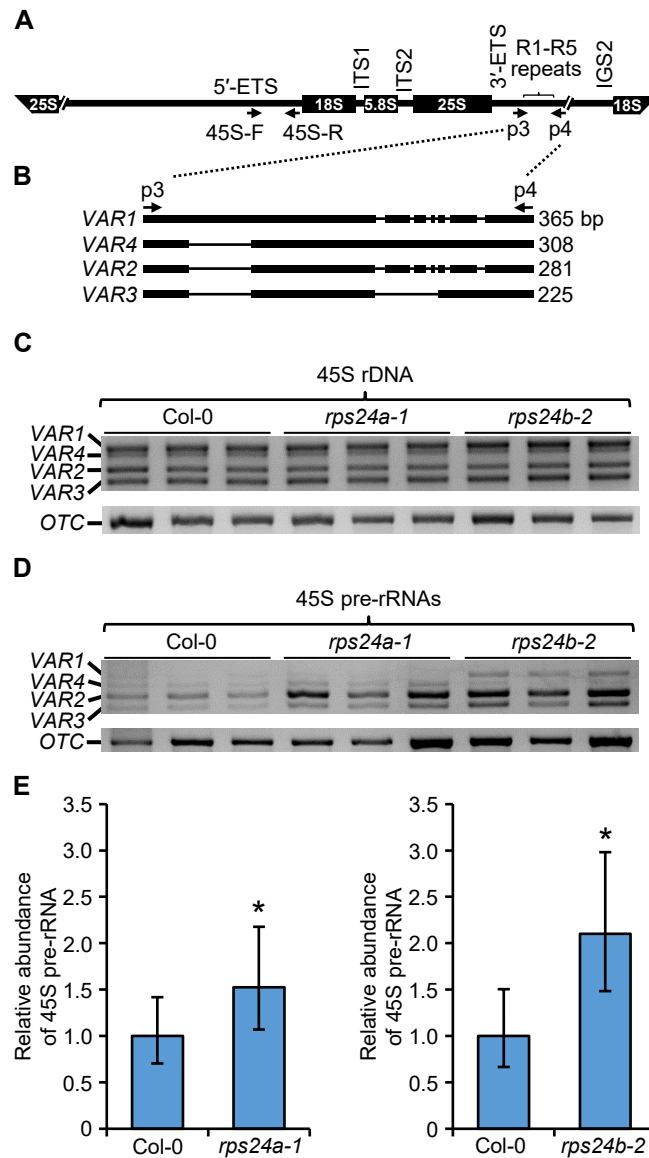


Figure 7. 45S rDNA expression in the *rps24a-1* and *rps24b-2* mutants. (A and B) Schematic representation of 45S pre-rRNA (A) and its 3'-ETS polymorphic region (B), modified from Micol-Ponce *et al.* (2018). (C and D) PCR analysis of the relative abundance of 45S rDNA variants using genomic DNA (C) and cDNA templates (D) from three biological replicates obtained from Col-0, *rps24a-1* and *rps24b-2* plants. The analysis was performed twice with similar results. (E) RT-qPCR analysis of 45S pre-rRNA abundance. Eight biological replicates were performed per genotype (Col-0, *rps24a-1*, and *rps24b-2*), with three technical replicates in each experiment, resulting in 24 amplifications per genotype. The analyses were carried out using ExpressionSuite Software. Asterisks indicate a significant difference with Col-0 in a Mann-Whitney U test ($n = 8$) ($*P < 0.05$).

361 that were co-expressed with the genes encoding the 25 putative interactors of RPS24B
362 obtained from the ATTEDII database (Obayashi et al., 2022).

363 We found ten proteins encoded by genes whose co-expression was enriched in
364 ribosome components, which included six RPs, six whose co-expression was enriched
365 in genes involved in ribosome biogenesis, three in spliceosome components, and two in
366 RNA transport, but that were components of the spliceosome: PRE-MRNA
367 PROCESSING FACTOR 8 (PRP8), the main component of the spliceosome, and one of
368 its major interactors during pre-mRNA splicing, the putative RNA helicase BAD
369 RESPONSE TO REFRIGERATION2 (BRR2; Nguyen et al., 2013).

370 Among the putative interactors identified by co-immunoprecipitation that were
371 encoded by genes that are co-expressed with genes involved in ribosome biogenesis,
372 we found AT3G05060 (*NOP58-2*) and AT5G27120 (*NOP58-1*); these genes encode the
373 co-orthologs of *NOP58* in yeast and humans, which are putative components of the C/D
374 small nucleolar ribonucleoprotein (snoRNP) complex. AT1G31970 and AT1G77030,
375 encoding DEA(D/H)-box RNA helicases, were two other co-expressed genes that are
376 putative orthologs of yeast *DEAD-box protein 3 (DBP3)* and *DBP10*, respectively, both
377 of which are involved in 60S subunit biogenesis. DBP3 functions in the endonucleolytic
378 cleavage of site A₃ in ITS1, upstream of 5.8S rRNA (Weaver et al., 1997), and DBP10
379 functions in the cleavage of sites C₁ and C₂ in the ITS2 (Burger et al., 2000). AT1G31970
380 encodes STRESS RESPONSE SUPPRESSOR 1 (STRS1), a nucleolar- and
381 chromocenter-localized protein that undergo stress-mediated relocalization and is
382 involved in epigenetic gene silencing at heterochromatic loci; its mutants show reduced
383 DNA methylation at these loci (Khan et al., 2014). AT1G77030 encodes the putative
384 DEAD-box ATP-dependent RNA HELICASE 29 (RH29), which has not been studied in
385 Arabidopsis. AT3G56510, AT5G04600 and AT5G57120 have not been studied either.
386 AT3G56510 is annotated at the Aramemnon database (Schwacke et al., 2003) as the
387 putative ortholog of yeast Eighteen S rRNA Factor 2 (*Esf2*), a nucleolar component of
388 the small subunit processome (SSU) involved in early steps of 18S rRNA maturation.
389 Yeast *Esf2* associates with 5'-ETS and is required for SSU assembly and compaction.
390 The *esf2* mutants show defective cleavage of sites A0 (within the 5'-ETS) through A2
391 (within the ITS1), therefore exhibiting reduced 18S rRNA production (Hoang et al., 2005).
392 AT5G04600 encodes the putative ortholog of human NUCLEOLAR PROTEIN
393 INTERACTING WITH THE FHA DOMAIN OF MKI67 (NIFK), an RBF involved in the
394 endonucleolytic ITS2 cleavage that is required for 28S and 5.8S rRNA maturation (Pan
395 et al., 2015). Finally, AT5G57120 encodes the putative ortholog of yeast serine rich
396 protein 40 (SRP40) and human nucleolar and coiled-body phosphoprotein 1 (NOLC1),

397 also named 140 kDa nucleolar phosphoprotein (Nopp140), which might function as
398 chaperones of snoRNP complexes (Yang et al., 2000).

399 **DISCUSSION**400 **The Arabidopsis *RPS24A* and *RPS24B* paralogs exhibit combined**
401 **haploinsufficiency**

402 The ribosome has traditionally been considered to be a highly conserved housekeeping
403 machinery, as it is present in all cells of all organisms. In prokaryotic and eukaryotic
404 ribosomes, the number of RPs is almost identical, i.e., 55 and 79, respectively. However,
405 proteomic analyses of animals, plants and fungi have revealed the existence of different
406 ribosomes with heterogeneous compositions and stoichiometry, suggesting that
407 specialized ribosomes might translate different sets of mRNAs. Such heterogeneity in
408 composition results from the existence of different paralogous genes produced by gene
409 duplications. These genes, with similar architectures and variable numbers among
410 species, encode quasi-identical RPs [reviewed in Martinez-Seidel et al. (2020; Petibon
411 et al. (2021)].

412 Plants have undergone more frequent gene and genomic duplications during
413 evolution compared to metazoans, leading to families of RPs with two to seven paralogs
414 in Arabidopsis (Carroll et al., 2008). In Arabidopsis, the RPs of the cytoplasmic ribosome
415 appear to be encoded by 234 functional genes, whereas the ribosomes of humans and
416 yeast include proteins encoded by 85 and 137 genes, respectively (Barakat et al., 2001;
417 Chang et al., 2005; Carroll et al., 2008). In the quasi-identical RP pairs studied to date,
418 one of the two paralogous genes is sometimes expressed at a higher level and
419 contributes more to ribosome structure than the other gene. In other cases, their levels
420 of expression and contributions to the ribosome are equivalent [reviewed in Petibon et
421 al. (2021)]. Loss of function of either gene of a pair produces a similar mutant phenotype,
422 indicating that both paralogs are required for normal development (Savada and Bonham-
423 Smith, 2014; Salih et al., 2020; Xiong et al., 2021).

424 Haploinsufficiency describes the requirement for more than one wild-type allele
425 for normal development or viability for single-copy genes (Meinke, 2013). Combined
426 haploinsufficiency describes a situation in which two paralogous genes behave as a
427 single haploinsufficient unit with four alleles (Hawley and Gilliland, 2006). *RPS24A* and
428 *RPS24B* provide such an example of combined haploinsufficiency, since three wild-type
429 doses of *RPS24* are required for proper plant development and two for viability. It is worth
430 noting that the single-copy human *RPS24* gene is a haploinsufficient locus.
431 Haploinsufficiency is observed in several human genetic diseases, such as Diamond-
432 Blackfan anemia (DBA), a ribosomopathy (Gazda et al., 2006) associated with the loss-
433 of-function of *RPS24*, as well as other genes encoding RPs (Tyagi et al., 2020).
434 Ribosomopathies are diseases caused by mutations in genes encoding components of

435 the ribosome, with tissue-specific effects and an increased risk of cancer [reviewed in
436 Kang et al. (2021)].

437 *RPS24A* and *RPS24B* have similar expression patterns and contributions to the
438 composition of Arabidopsis ribosomes, pointing to their functional equivalence and
439 partial redundancy. However, the mRNA levels of *RPS24A* and *RPS24B* are lower than
440 those of other genes encoding RPs (Savada and Bonham-Smith, 2014). Here we
441 demonstrated that these genes do not show dosage compensation; i.e., *RPS24A* is not
442 more expressed in the absence of *RPS24B*. A proper stoichiometry of ribosome
443 structural components is crucial for ribosome function (Warner, 1999; Slavov et al., 2015;
444 Ni and Buszczak, 2023), and our results suggest that the amount of RPS24A and
445 RPS24B produced in Arabidopsis cells is limiting for ribosome production. Indeed,
446 reduction in the wild-type doses of either protein in *rps24a* and *rps24b* single mutants
447 and diheterozygous plants (Figures 1 and 2) decreases the number of productive
448 ribosomes, resulting in an inadequate rate of protein synthesis. This phenomenon of
449 non-allelic non-complementation is also observed in mutants affected in other
450 Arabidopsis genes encoding RPs, including *RPS6*, *RPL5*, *RPL23* and *RPL36*
451 (Degenhardt and Bonham-Smith, 2008; Fujikura et al., 2009; Creff et al., 2010;
452 Casanova-Sáez et al., 2014).

453

454 **RPS24A and RPS24B function as RBFs during 18S rRNA maturation, like their** 455 **yeast and human orthologs**

456 Gene duplications have allowed some RPs to acquire additional functions that might or
457 might not be related to translation. In Arabidopsis, besides their functions as structural
458 proteins of the ribosome, only RPS2 was shown to function as an RBF during the
459 maturation of 25S and 18S rRNA, and RPS6 in the transcriptional regulation of 45S rDNA
460 (Kim et al., 2014; Hang et al., 2021). PROTEIN ARGININE METHYLTRANSFERASE 3
461 (PRMT3) and RPS2B (one of the six paralogs encoded in the Arabidopsis genome)
462 facilitate the assembly and disassembly of the SSU processome and repress nucleolar
463 stress (Hang et al., 2021). RPS6, encoded by two functionally equivalent paralogs,
464 interacts with HISTONE DEACETYLASE 2 (HDA2) in the nucleus and nucleolus and
465 binds to the promoter of 45S rDNA, controlling its transcription (Kim et al., 2014).
466 Furthermore, RPL10 is related to responses to ultraviolet light stress and viral infection,
467 and RPL24 is involved in microRNA biogenesis (Carvalho et al., 2008; Ferreyra et al.,
468 2010a; Ferreyra et al., 2010b; Li et al., 2017).

469 Our results show that both Arabidopsis RPS24A and RPS24B function as RBFs
470 in the early steps of 45S pre-rRNA processing, specifically at the cleavage of the P' site
471 of the 5'-ETS. In the *rps24b-2*, *api6* and *rps24a-1* mutants, we detected the accumulation

472 of 35S_{A123B}/35S(P) pre-rRNA (the first intermediate in 45S pre-rRNA processing) and a
 473 reduction in the levels of 33S(P')/32S species (the second intermediate in this
 474 processing; Supplemental Figure 1). We also detected defects in the processing of 18S
 475 pre-rRNA: the *rps24b-2*, *api6* and *rps24a-1* mutants accumulated P-A₃ pre-rRNA and
 476 showed reduced levels of P'-A₃/18S-A₃/18S-A₂ pre-rRNAs (Figure 4). These results
 477 suggest that the loss-of-function of *RPS24A* and *RPS24B* globally reduces the
 478 processing of 5'-ETS, as occurs in response to the loss-of-function of yeast and human
 479 *RPS24* (Ferreira-Cerca et al., 2005; Choemsel et al., 2008; Robledo et al., 2008).
 480 Moreover, the *rps24b* mutant showed an imbalance in the 25S/18S ratio, suggesting that
 481 alterations in the processing of pre-rRNAs of the ITS1-first pathway lead to a decrease
 482 in 18S rRNA levels.

483

484 ***RPS24B* genetically interacts with *NOP53* (*SMO4*) and *MTR4***

485 We obtained double mutant combinations of *rps24b-2* with mutant alleles of genes
 486 involved in different steps of 45S pre-rRNA processing. Two such double mutants
 487 exhibited synergistic phenotypes: *rps24b-2 rrp7-1* was embryonic lethal, and *rps24b-2*
 488 *mtr4-2* was dwarf, with poor fertility. The phenotype of *rps24b-2 smo4-3* was also
 489 synergistic, as this mutant was also nearly fully sterile (Figure 5).

490 We previously demonstrated that RRP7 is involved in 18S rRNA maturation and
 491 acts as a repressor of 45S rDNA transcription. The *rrp7* mutant has a reduced 18S/25S
 492 rRNA ratio and accumulated P-A₃ pre-rRNA (Micol-Ponce et al., 2018), similar to the
 493 *rps24* mutant. Moreover, we reasoned that RRP7 might function as a repressor of 45S
 494 rDNA, as *rrp7* showed heterochronic expression of *VAR1* and increased *VAR3*
 495 expression (Micol-Ponce et al., 2018). Here, we determined that *rps24a-1* and *rps24b-2*
 496 contained increased levels of 45S pre-rRNA (Figure 7), which might explain the
 497 synergistic phenotype of the *rps24b-2 rrp7-1* double mutant. However, due to the lethality
 498 of *rps24b-2 rrp7-1*, we were unable to examine the effects of the combination of both
 499 mutations on the maturation of 18S rRNA and/or the transcription of 45S rDNA.

500 *NOP53* (*SMO4* in Arabidopsis) and *MTR4* are RBFs that function in the
 501 maturation of 5.8S rRNA (Lange et al., 2011; Micol-Ponce et al., 2020). *MTR4* is a
 502 cofactor of the exosome, which also functions in the degradation of P-P', the by-product
 503 of 5'-ETS (Lange et al., 2011). Loss-of-function of *SMO4* or *MTR4* leads to the
 504 accumulation of 7S or 5.8S + 70-nt pre-rRNAs, respectively (which are precursors of
 505 5.8S rRNA that are processed in the nucleolus), but not 6S rRNA (the final precursor of
 506 5.8S rRNA maturation that is processed after its export to the cytoplasm; Lange et al.,
 507 2011; Micol-Ponce et al., 2020). In the current study, we did not detect the accumulation
 508 of P-P' or any 5.8S pre-rRNA in the *rps24a* or *rps24b* mutants, suggesting that *RPS24A*

509 and RPS24B do not participate in these steps of 45S pre-rRNA processing. However,
 510 we unexpectedly detected the over-accumulation of specific 5.8S pre-rRNAs in *smo4-3*
 511 and *mtr4-2* and in the double mutants *rps24b-2 smo4-3* and *rps24b-2 mtr4-2* (Figure 6).
 512 Such over-accumulation may explain their synergistic phenotypic defects. Our RNA-
 513 FISH analysis with the S9 probe confirmed these results, revealing strong fluorescent
 514 signals in the nucleoli of the double mutants (Supplemental Figure 8).

515 One possible explanation of the above-mentioned results is that RPS24B (and
 516 possibly RPS24A) functions in 5.8S rRNA maturation, albeit to a lesser extent than
 517 SMO4 and MTR4. An alternative hypothesis is that the over-accumulation of 7S and 5.8S
 518 + 70-nt pre-rRNAs, which we detected in *rps24b-2 mtr4-2* and *rps24b-2 smo4-3* plants,
 519 respectively, is due to an increase in rDNA transcription caused by the *rps24b-2*
 520 mutation. If more pre-rRNAs are available for processing, the steps involving MTR4 and
 521 SMO4 are likely to be more strongly affected in the double mutants. We consider this
 522 second hypothesis to be more plausible, as it also might explain the synergistic
 523 phenotype of the *rps24b-2 rrp7-1* and *rps24b-2 par1-2* double mutants.

524

525 **The predominant nucleolar localization of RPS24B and its putative interactors also** 526 **support its role as an RBF**

527 Our results using the RPS24B-GFP fusion protein indicate that RPS24B is primarily
 528 located in the nucleolus, with some presence in the cytoplasm; these observations are
 529 consistent with previous predictions and proteomic analyses. This distribution reflects
 530 the dual role of RPS24B as both a structural protein of the ribosome and an RBF involved
 531 in early steps of ribosome biogenesis and the transcriptional repression of 45S rDNA.

532 Using different filters that included sub-organellar localization, we identified 25
 533 nucleolar proteins that co-immunoprecipitated with RPS24B-GFP, which we considered
 534 to be its most likely interactors. Among these were RBFs involved in the synthesis of the
 535 60S and 40S subunits, including STRS1 (AT1G31970), AT1G77030 and AT3G56510,
 536 which are putative orthologs of yeast DBP3, DBP10 and ESF2, respectively, putative
 537 ortholog of yeast, as well as NIFK. We speculate that STRS1 might also be involved in
 538 repressing 45S rDNA, which would be difficult in the *rps24* background if RPS24A,
 539 RPS24B, or both proteins facilitate access to the epigenetic silencing machinery at the
 540 promoter of 45S rDNA. In fact, STRS1 colocalized with HDA2 in the nucleolus (Khan et
 541 al., 2014). Like RPS6, another RP that plays extra-ribosomal roles in the maturation of
 542 18S rRNA and the transcriptional regulation of 45S rDNA (Kim et al., 2014), RPS24A
 543 and RPS24B might also act in this capacity.

544 We also identified AT5G57120 as an RPS24B-GFP interactor. AT5G57120 is the
 545 putative ortholog of human NOLC1, which colocalizes with RNA Pol I in the nucleolus at

546 47S rDNA and might activate its transcription (Chen et al., 1999). The interaction
547 between the protein encoded by AT5G57120 and RPS24B might inhibit access of the-
548 encoded protein to the 45S rDNA promoter. RPS24B, and probably RPS24A, might also
549 act as chaperones to facilitate the incorporation of RBFs into their corresponding pre-
550 rRNA substrates, as observed for other RPs. Although we did not find SMO4 or MTR4
551 among the RPS24B-GFP interactors, we cannot exclude the possibility that RPS24B and
552 RPS24A function in the maturation of 5.8S rRNA. However, we consider this possibility
553 unlikely, as previously argued. Interestingly, we also identified two well-known splicing
554 factors among the proteins that co-immunoprecipitated with RPS24B-GFP: PRP8 and
555 the RNA helicase BRR2. PRP8 was previously identified in a U3 snoRNA
556 ribonucleoprotein complex isolated from cauliflower (*Brassica oleracea*) inflorescences
557 (Samaha et al., 2010) and in a proteomic analysis of the Arabidopsis nucleolus (Palm et
558 al., 2016). We cannot exclude the possibility that PRP8 and BRR2 function in ribosome
559 biogenesis, as the expression of *RPS24B-GFP* was unable to rescue the mutant
560 phenotype of *rps24b-2* plants, suggesting that many RNA helicases act in more than one
561 pathway. Indeed, yeast Prp43 and its human ortholog DEAD-box helicase 5 (DDX5) are
562 multifunctional RNA helicases that participate in ribosome biogenesis and pre-mRNA
563 splicing, among other RNA metabolic steps [reviewed in Bohnsack et al. (2022)].
564 However, our assay may have generated false positives, since RPS24B-GFP was
565 unable to rescue the mutant phenotype of *rps24b-2* plants.

566

567 MATERIALS AND METHODS

568 Plant material and growth conditions

569 The wild-type *Arabidopsis* (*Arabidopsis thaliana*) (L) Heynh. Columbia-0 (Col-0) and
570 Landsberg *erecta* (Ler) accessions were provided by the Nottingham Arabidopsis Stock
571 Centre (NASC) and propagated in our laboratory. The *api6* mutant in the Ler background
572 was isolated in the laboratory of J.L. Micol after ethyl methanesulfonate (EMS)
573 mutagenesis (Berná et al., 1999). Seeds of *rps24b-2* (SALK_000766; Wang et al., 2018),
574 *rps24a-1* (SALK_126799), *smo4-3* (SALK_071764; Micol-Ponce et al., 2018), *mtr4-2*
575 (SAIL_50_C11; Lange et al., 2011), *rrp7-1* (SAIL_628_F08; Micol-Ponce et al., 2018)
576 and *parl1-2* (SALK_002764; Petricka and Nelson, 2007) were also provided by NASC.
577 Seed sterilization, sowing, plant culture and crosses were performed as previously
578 described (Berná et al., 1999; Ponce et al., 1999).

579

580 Positional cloning of the *api6* mutation and genotyping of single and double 581 mutants

582 Genomic DNA extraction and mapping of the *api6* mutation were performed as
583 previously described (Ponce et al., 1999; Ponce et al., 2006). The primers used for fine-
584 mapping are listed in Supplemental Table 1. The *api6* point mutation was identified by
585 Sanger sequencing using the primers listed in Supplemental Table 2. The presence of
586 T-DNA insertions in *RPS24A*, *RPS24B*, *SMO4*, *RRP7*, *MTR4* and *PARL1* was verified
587 by PCR using the primers shown in Supplemental Table 2.

588 Most Sanger sequencing reactions and electrophoreses were carried out in our
589 laboratory with ABI PRISM BigDye Terminator Cycle Sequencing kits and an ABI PRISM
590 3130xl Genetic Analyzer (Applied Biosystems). Some Sanger sequencing reactions
591 were carried out by Stab Vida (Portugal).

592

593 Construction of transgenic lines

594 The constructs used in this study were generated by Gateway cloning as described in
595 Sánchez-García et al. (2015) using the pGEM-T Easy221 entry vector and the pMDC83
596 destination vector (Curtis and Grossniklaus, 2003). To generate the *35S_{pro}:RPS24B:GFP*
597 transgene, the full-length coding region (without the stop codon) of *RPS24B* was
598 amplified by PCR using genomic DNA from Col-0 as a template and primers that included
599 the *attB1* and *attB2* sequences (Supplemental Table 2). The integrity of the constructs
600 was verified by Sanger sequencing. Chemically competent *Escherichia coli* DH5α cells
601 were transformed with the Gateway cloning products by the heat shock method, and
602 *Agrobacterium tumefaciens* C58C1 cells were transformed by electroporation with the
603 individual destination vectors carrying each insert of interest.

604

605 Plant morphological and histochemical analyses

606 The rosettes of plants were photographed under a Nikon SMZ1500 stereomicroscope
607 equipped with a Nikon DS-Ri2 digital camera. For large rosettes, several photographs
608 from the same plant were assembled with the Photomerge tool of Adobe Photoshop CS3
609 software (Adobe). The backgrounds of the photographs were homogenized using Adobe
610 Photoshop CS3 software without modifying the rosettes.

611 Fluorescence and confocal laser-scanning microscopy images were taken under
612 a Leica STELLARIS microscope and processed using Leica Application Suite X (LAS X)
613 software. For nuclear staining, the roots of seedlings collected 5 days after stratification
614 (das) were immersed in 0.5 µg/mL 4',6-diamidino-2-phenylindole (DAPI) for 5 min and
615 washed two times with water. DAPI, GFP, tetramethylrhodamine-5-isothiocyanate
616 (TRITC) and cyanine 3 (Cy3) were excited at 405, 488, 503 and 506 nm, respectively,
617 and their emissions collected at 425/727, 515/730, 530/732 and 511/726 nm,
618 respectively.

619

620 RT-PCR and RNA gel-blot analysis

621 Genomic DNA was extracted using a DNeasy Plant mini kit (Qiagen) from three
622 biological replicates per genotype (leaves from a pool of three seedlings collected 15 das
623 from three different Petri plates). Total RNA was extracted using TRIzol Reagent
624 (Invitrogen) for RT-PCR and semiquantitative RT-PCR and an RNeasy Plant Mini Kit
625 (Qiagen) for RT-qPCR RNA gel blots. Three biological replicates per genotype were
626 used for RT-PCR and semiquantitative RT-PCR, and eight were used for RT-qPCR.
627 Each biological replicate included the aerial tissues from a pool of three seedlings
628 collected 15 das from three different Petri plates. In all RT-PCR assays, RNA was treated
629 twice with Turbo DNase (Invitrogen), and reverse transcription was carried out using
630 random hexamer primers and Maxima Reverse Transcriptase (Invitrogen). The
631 *ORNITHINE CARBAMOYLTRANSFERASE (OTC)* and *ACTIN2 (ACT2)* housekeeping
632 genes were used as internal controls: *OTC* for RT-PCR and semiquantitative RT-PCR,
633 and *ACT2* for RT-qPCR assays. Primers for all RT-PCR assays and those used as
634 controls for genomic DNA amplification in the analysis of 45S rDNA variant expression
635 are listed in Supplemental Table 2.

636 For RNA gel-blot analysis, four digoxigenin (DIG) labeled probes were used,
637 whose sequences were obtained from Lange et al. (2011); the S6, S7 and S9 probes
638 were oligonucleotides labeled at both the 3' and 5' ends, synthesized by Sigma-Aldrich.
639 The S2 probe was synthesized by PCR using genomic DNA as a template, DIG-11-dUTP
640 (Roche) and the S2-Fw and S2-Rv primers (Supplemental Table 2). Two µg of total RNA

641 per sample were loaded onto 1.2% (w/v) agarose and 2.12% (v/v) formaldehyde or 6%
642 (w/v) polyacrylamide gels. Electrophoresis, hybridization and detection were performed
643 as described in Micol-Ponce et al. (2018; Micol-Ponce et al. (2020).

644

645 **RNA-FISH and immunolocalization**

646 RNA fluorescence *in situ* hybridization (RNA-FISH) was performed as described by Parry
647 et al. (2006) and modified in Micol-Ponce et al. (2020) using the S9 probe labeled with
648 Cy3 dye (Eurofins Genomics) at a concentration of 0.5 µg/mL. Leaves were mounted
649 onto slides with Vectashield antifade mounting medium (Vector Laboratories) containing
650 0.01 µg/mL DAPI, which was used as a nuclear marker.

651 Fibrillarlin immunolocalization was performed following Pasternak et al. (2015), as
652 modified by Micol-Ponce et al. (2020). In brief, roots of seedlings were collected at 5 das
653 and fixed for 40 min at 37°C in a solution containing 2% (w/v) paraformaldehyde in 1×
654 microtubule-stabilizing buffer (50 mM PIPES, 5 mM MgSO₄, and 5 mM EGTA, pH 6.9)
655 and 0.1% (v/v) Triton X-100. An anti-fibrillarlin [38F3] (Abcam) primary antibody was used
656 at a 1:250 dilution, and a TRITC-conjugated anti-mouse IgG (Sigma Aldrich) secondary
657 antibody was used at a 1:500 dilution. Nuclei were stained for 10 min with 0.2 µg/mL
658 DAPI and washed for 5 min before mounting the samples on slides with water. Nuclei
659 and nucleoli areas were measured as described in Micol-Ponce et al. (2020) using the
660 NIS Elements AR3.1 (Nikon) image-analysis package.

661

662 **Co-immunoprecipitation assay**

663 For co-immunoprecipitation assays, three biological replicates were used, each
664 consisting in 1 g of rosettes of *35S_{pro}:RPS24B:GFP* transgenic plants in the Col-0
665 background collected 15 das. The tissue was manually ground and proteins were
666 extracted as described in Navarro-Quiles et al. (2022). Co-immunoprecipitation was
667 carried out using a µMACS GFP Isolation kit (Milteny Biotec). The effectiveness of the
668 immunoprecipitation of RPS24B-GFP was checked by immunoblotting using anti-GFP-
669 HRP antibody (Milteny Biotec).

670 The co-immunoprecipitated proteins were identified at the Centro Nacional de
671 Biotecnología (CNB) Proteomics facility (Madrid, Spain) by liquid chromatography
672 followed by tandem mass spectrometry (LC-MS/MS) using an Orbitrap Exploris 240
673 mass spectrometer. Tandem mass peptide spectra were searched against Arabidopsis
674 protein sequences in the Araport11 database using the MASCOT search engine (Matrix
675 Science). Peptide sequences identified with a false discovery rate (FDR) < 1% were
676 considered to be valid. To increase the confidence of the results, only the proteins

677 identified with at least two peptides in each of the three biological replicates were
678 considered for analysis.

679

680 **Accession numbers**

681 Sequence data from this article can be found at The Arabidopsis Information Resource
682 (<https://www.arabidopsis.org/>) under the following accession numbers: *RPS24A*
683 (AT3G04920), *RPS24B* (AT5G28060), *RRP7* (AT5G38720), *SMO4* (AT2G40430),
684 *MTR4* (AT1G59760), and *PARL1* (AT1G48920). Although AT3G04920 and AT5G28060,
685 which encode RPS24A and RPS24B, have been renamed as ES24Z and ES24Y,
686 respectively, according to a new nomenclature for RPs (Scarpin et al., 2022), we used
687 their old names in this work, since the *rps24b-2* allele of *RPS24* were previously studied
688 (Horiguchi et al., 2011).

689

690 **FUNDING**

691 This work was supported by grants from the Ministerio de Ciencia e Innovación of Spain
692 (BIO2017-89728-R and PID2020-117125RB-I00 [MCI/AEI/FEDER, UE]) and the
693 Generalitat Valenciana [PROMETEO/2019/117] to M.R.P. R.M.-P. holds a María
694 Zambrano postdoctoral contract funded by the Next Generation EU programs and
695 administered by the Universidad Miguel Hernández.

696

697 **ACKNOWLEDGEMENTS**

698 The authors thank F. Lozano, J. Castelló, D. Navarro and M. Gomariz for their excellent
699 technical assistance, and J.L. Micol for useful discussions, comments on the manuscript,
700 and providing the *api6* mutant, as well as for the use of his facilities.

701

702 **AUTHOR CONTRIBUTIONS**

703 M.R.P. obtained funding and conceived, designed and supervised research; all authors
704 performed research and wrote the paper.

830 **REFERENCES**

- 831 **Ayash M, Abukhalaf M, Thieme D, Proksch C, Heilmann M, Schattat MH,**
832 **Hoehenwarter W** (2021) LC-MS based draft map of the *Arabidopsis thaliana* nuclear
833 proteome and protein import in pattern triggered immunity. *Front Plant Sci* **12**: 744103
- 834 **Barakat A, Szick-Miranda K, Chang IF, Guyot R, Blanc G, Cooke R, Delseny M,**
835 **Bailey-Serres J** (2001) The organization of cytoplasmic ribosomal protein genes in
836 the *Arabidopsis* genome. *Plant Physiol* **127**: 398-415
- 837 **Berná G, Robles P, Micol JL** (1999) A mutational analysis of leaf morphogenesis in
838 *Arabidopsis thaliana*. *Genetics* **152**: 729-742
- 839 **Bohnsack KE, Henras AK, Nielsen H, Bohnsack MT** (2022) Making ends meet: a
840 universal driver of large ribosomal subunit biogenesis. *Trends Biochem Sci* **48**: 213-
841 215
- 842 **Burger F, Daugeron MC, Linder P** (2000) Dbp10p, a putative RNA helicase from
843 *Saccharomyces cerevisiae*, is required for ribosome biogenesis. *Nucleic Acids Res*
844 **28**: 2315-2323
- 845 **Burgute BD, Peche VS, Steckelberg AL, Glockner G, Gassen B, Gehring NH,**
846 **Noegel AA** (2014) NKAP is a novel RS-related protein that interacts with RNA and
847 RNA binding proteins. *Nucleic Acids Res* **42**: 3177-3193
- 848 **Carroll AJ, Heazlewood JL, Ito J, Millar AH** (2008) Analysis of the *Arabidopsis*
849 cytosolic ribosome proteome provides detailed insights into its components and their
850 post-translational modification. *Mol Cell Proteomics* **7**: 347-369
- 851 **Carvalho CM, Santos AA, Pires SR, Rocha CS, Saraiva DI, Machado JP, Mattos EC,**
852 **Fietto LG, Fontes EP** (2008) Regulated nuclear trafficking of rpL10A mediated by
853 NIK1 represents a defense strategy of plant cells against virus. *PLOS Pathog* **4**:
854 e1000247
- 855 **Casanova-Sáez R, Candela H, Micol JL** (2014) Combined haploinsufficiency and
856 purifying selection drive retention of *RPL36a* paralogs in *Arabidopsis*. *Sci Rep* **4**: 4122
- 857 **Creff A, Sormani R, Desnos T** (2010) The two *Arabidopsis* *RPS6* genes, encoding for
858 cytoplasmic ribosomal proteins S6, are functionally equivalent. *Plant Mol Biol* **73**: 533-
859 546
- 860 **Curtis MD, Grossniklaus U** (2003) A Gateway cloning vector set for high-throughput
861 functional analysis of genes in planta. *Plant Physiol* **133**: 462-469
- 862 **Chang IF, Szick-Miranda K, Pan S, Bailey-Serres J** (2005) Proteomic characterization
863 of evolutionarily conserved and variable proteins of *Arabidopsis* cytosolic ribosomes.
864 *Plant Physiol* **137**: 848-862

- 865 **Chen HK, Pai CY, Huang JY, Yeh NH** (1999) Human Nopp140, which interacts with
866 RNA polymerase I: implications for rRNA gene transcription and nucleolar structural
867 organization. *Mol Cell Biol* **19**: 8536-8546
- 868 **Choesmel V, Fribourg S, Aguisa-Toure AH, Pinaud N, Legrand P, Gazda HT,**
869 **Gleizes PE** (2008) Mutation of ribosomal protein RPS24 in Diamond-Blackfan anemia
870 results in a ribosome biogenesis disorder. *Hum Mol Genet* **17**: 1253-1263
- 871 **Degenhardt RF, Bonham-Smith PC** (2008) Arabidopsis ribosomal proteins RPL23aA
872 and RPL23aB are differentially targeted to the nucleolus and are disparately required
873 for normal development. *Plant Physiol* **147**: 128-142
- 874 **Durut N, Abou-Ellail M, Pontvianne F, Das S, Kojima H, Ukai S, de Bures A, Comella**
875 **P, Nidelet S, Rialle S, Merret R, Echeverria M, Bouvet P, Nakamura K, Saez-**
876 **Vasquez J** (2014) A duplicated NUCLEOLIN gene with antagonistic activity is
877 required for chromatin organization of silent 45S rDNA in Arabidopsis. *Plant Cell* **26**:
878 1330-1344
- 879 **Farooq M, Lindbaek L, Krogh N, Doganli C, Keller C, Monnich M, Goncalves AB,**
880 **Sakthivel S, Mang Y, Fatima A, Andersen VS, Hussain MS, Eiberg H, Hansen L,**
881 **Kjaer KW, Gopalakrishnan J, Pedersen LB, Mollgard K, Nielsen H, Baig SM,**
882 **Tommerup N, Christensen ST, Larsen LA** (2020) RRP7A links primary
883 microcephaly to dysfunction of ribosome biogenesis, resorption of primary cilia, and
884 neurogenesis. *Nat Commun* **11**: 5816
- 885 **Ferreira-Cerca S, Poll G, Gleizes PE, Tschochner H, Milkereit P** (2005) Roles of
886 eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and
887 ribosome function. *Mol Cell* **20**: 263-275
- 888 **Ferreyra ML, Biarc J, Burlingame AL, Casati P** (2010a) Arabidopsis L10 ribosomal
889 proteins in UV-B responses. *Plant Signal Behav* **5**: 1222-1225
- 890 **Ferreyra MLF, Pezza A, Biarc J, Burlingame AL, Casati P** (2010b) Plant L10
891 ribosomal proteins have different roles during development and translation under
892 ultraviolet-B stress. *Plant Physiol* **153**: 1878-1894
- 893 **Fujikura U, Horiguchi G, Ponce MR, Micol JL, Tsukaya H** (2009) Coordination of cell
894 proliferation and cell expansion mediated by ribosome-related processes in the leaves
895 of *Arabidopsis thaliana*. *Plant J* **59**: 499-508
- 896 **Gazda HT, Grabowska A, Merida-Long LB, Latawiec E, Schneider HE, Lipton JM,**
897 **Vlachos A, Atsidaftos E, Ball SE, Orfali KA, Niewiadomska E, Da Costa L,**
898 **Tchernia G, Niemeyer C, Meerpohl JJ, Stahl J, Schratt G, Glader B, Backer K,**
899 **Wong C, Nathan DG, Beggs AH, Sieff CA** (2006) Ribosomal protein S24 gene is
900 mutated in Diamond-Blackfan anemia. *Am J Hum Genet* **79**: 1110-1118

- 901 **Giavalisco P, Wilson D, Kreitler T, Lehrach H, Klose J, Gobom J, Fucini P** (2005)
902 High heterogeneity within the ribosomal proteins of the *Arabidopsis thaliana* 80S
903 ribosome. *Plant Mol Biol* **57**: 577-591
- 904 **Hang R, Wang Z, Yang C, Luo L, Mo B, Chen X, Sun J, Liu C, Cao X** (2021) Protein
905 arginine methyltransferase 3 fine-tunes the assembly/disassembly of pre-ribosomes
906 to repress nucleolar stress by interacting with RPS2B in *Arabidopsis*. *Mol Plant* **14**:
907 223-236
- 908 **Hawley RS, Gilliland WD** (2006) Sometimes the result is not the answer: the truths and
909 the lies that come from using the complementation test. *Genetics* **174**: 5-15
- 910 **Hoang T, Peng WT, Vanrobays E, Krogan N, Hiley S, Beyer AL, Osheim YN,**
911 **Greenblatt J, Hughes TR, Lafontaine DL** (2005) Esf2p, a U3-associated factor
912 required for small-subunit processome assembly and compaction. *Molecular and Cell*
913 *Biology* **25**: 5523-5534
- 914 **Horiguchi G, Fujikura U, Ferjani A, Ishikawa N, Tsukaya H** (2006) Large-scale
915 histological analysis of leaf mutants using two simple leaf observation methods:
916 identification of novel genetic pathways governing the size and shape of leaves. *Plant*
917 *J* **48**: 638-644
- 918 **Horiguchi G, Mollá-Morales A, Pérez-Pérez JM, Kojima K, Robles P, Ponce MR,**
919 **Micol JL, Tsukaya H** (2011) Differential contributions of ribosomal protein genes to
920 *Arabidopsis thaliana* leaf development. *Plant J* **65**: 724-736
- 921 **Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P,**
922 **Gruissem W, Zimmermann P** (2008) Genevestigator v3: a reference expression
923 database for the meta-analysis of transcriptomes. *Adv Bioinf* **2008**: 420747
- 924 **Hummel M, Dobrenel T, Cordewener JJ, Davanture M, Meyer C, Smeekens SJ,**
925 **Bailey-Serres J, America TA, Hanson J** (2015) Proteomic LC-MS analysis of
926 *Arabidopsis* cytosolic ribosomes: Identification of ribosomal protein paralogs and re-
927 annotation of the ribosomal protein genes. *J Proteom* **128**: 436-449
- 928 **Jiang Y, Wang D, Yao Y, Eubel H, Kunzler P, Moller IM, Xu D** (2021) MULocDeep: a
929 deep-learning framework for protein subcellular and suborganellar localization
930 prediction with residue-level interpretation. *Comput Struct Biotechnol J* **19**: 4825-4839
- 931 **Kang J, Brajanovski N, Chan KT, Xuan J, Pearson RB, Sanij E** (2021) Ribosomal
932 proteins and human diseases: molecular mechanisms and targeted therapy. *Signal*
933 *Transduct Target Ther* **6**: 323
- 934 **Khan A, Garbelli A, Grossi S, Florentin A, Batelli G, Acuna T, Zolla G, Kaye Y, Paul**
935 **LK, Zhu JK, Maga G, Grafi G, Barak S** (2014) The *Arabidopsis* STRESS
936 RESPONSE SUPPRESSOR DEAD-box RNA helicases are nucleolar- and

- 937 chromocenter-localized proteins that undergo stress-mediated relocalization and are
 938 involved in epigenetic gene silencing. *Plant J* **79**: 28-43
- 939 **Kim YK, Kim S, Shin YJ, Hur YS, Kim WY, Lee MS, Cheon CI, Verma DP** (2014)
 940 Ribosomal protein S6, a target of rapamycin, is involved in the regulation of rRNA
 941 genes by possible epigenetic changes in *Arabidopsis*. *J Biol Chem* **289**: 3901-3912
- 942 **Kojima H, Suzuki T, Kato T, Enomoto K, Sato S, Kato T, Tabata S, Sáez-Vásquez J,**
 943 **Echeverría M, Nakagawa T, Ishiguro S, Nakamura K** (2007) Sugar-inducible
 944 expression of the nucleolin-1 gene of *Arabidopsis thaliana* and its role in ribosome
 945 synthesis, growth and development. *Plant J* **49**: 1053-1063
- 946 **Lange H, Sement FM, Gagliardi D** (2011) MTR4, a putative RNA helicase and exosome
 947 co-factor, is required for proper rRNA biogenesis and development in *Arabidopsis*
 948 *thaliana*. *Plant J* **68**: 51-63
- 949 **Li S, Liu K, Zhang S, Wang X, Rogers K, Ren G, Zhang C, Yu B** (2017) STV1, a
 950 ribosomal protein, binds primary microRNA transcripts to promote their interaction
 951 with the processing complex in *Arabidopsis*. *Proc Natl Acad Sci USA* **114**: 1424-1429
- 952 **Martinez-Seidel F, Beine-Golovchuk O, Hsieh YC, Kopka J** (2020) Systematic review
 953 of plant ribosome heterogeneity and specialization. *Front Plant Sci* **11**: 948
- 954 **Meinke DW** (2013) A survey of dominant mutations in *Arabidopsis thaliana*. *Trends Plant*
 955 *Sci* **18**: 84-91
- 956 **Mergner J, Frejno M, List M, Papacek M, Chen X, Chaudhary A, Samaras P, Richter**
 957 **S, Shikata H, Messerer M, Lang D, Altmann S, Cyprys P, Zolg DP, Mathieson T,**
 958 **Bantscheff M, Hazarika RR, Schmidt T, Dawid C, Dunkel A, Hofmann T, Sprunck**
 959 **S, Falter-Braun P, Johannes F, Mayer KFX, Jurgens G, Wilhelm M, Baumbach J,**
 960 **Grill E, Schneitz K, Schwechheimer C, Kuster B** (2020) Mass-spectrometry-based
 961 draft of the *Arabidopsis* proteome. *Nature* **579**: 409-414
- 962 **Micol-Ponce R, Sarmiento-Mañús R, Fontcuberta-Cervera S, Cabezas-Fuster A, de**
 963 **Bures A, Sáez-Vásquez J, Ponce MR** (2020) SMALL ORGAN4 is a ribosome
 964 biogenesis factor involved in 5.8S ribosomal RNA maturation. *Plant Physiol* **184**:
 965 2022-2039
- 966 **Micol-Ponce R, Sarmiento-Mañús R, Ruiz-Bayón A, Montacié C, Sáez-Vásquez J,**
 967 **Ponce MR** (2018) *Arabidopsis* RIBOSOMAL RNA PROCESSING7 is required for 18S
 968 rRNA maturation. *Plant Cell* **30**: 2855-2872
- 969 **Montacié C, Durut N, Opsomer A, Palm D, Comella P, Picart C, Carpentier MC,**
 970 **Pontvianne F, Carapito C, Schleiff E, Sáez-Vásquez J** (2017) Nucleolar proteome
 971 analysis and proteasomal activity assays reveal a link between nucleolus and 26S
 972 proteasome in *A. thaliana*. *Front Plant Sci* **8**: 1815

- 973 **Navarro-Quiles C, Mateo-Bonmatí E, Candela H, Robles P, Martínez-Laborda A,**
 974 **Fernández Y, Simura J, Ljung K, Rubio V, Ponce MR, Micol JL** (2022) The
 975 Arabidopsis ATP-Binding Cassette E protein ABCE2 is a conserved component of the
 976 translation machinery. *Front Plant Sci* **13**: 1009895
- 977 **Nguyen TH, Li J, Galej WP, Oshikane H, Newman AJ, Nagai K** (2013) Structural basis
 978 of Brr2-Prp8 interactions and implications for U5 snRNP biogenesis and the
 979 spliceosome active site. *Structure* **21**: 910-919
- 980 **Ni C, Buszczak M** (2023) The homeostatic regulation of ribosome biogenesis. *Semin*
 981 *Cell Dev Biol* **136**: 13-26
- 982 **Obayashi T, Hibara H, Kagaya Y, Aoki Y, Kinoshita K** (2022) ATTED-II v11: A plant
 983 gene coexpression database using a sample balancing technique by subagging of
 984 principal components. *Plant Cell Physiol* **63**: 869-881
- 985 **Pajerowski AG, Nguyen C, Aghajanian H, Shapiro MJ, Shapiro VS** (2009) NKAP is
 986 a transcriptional repressor of notch signaling and is required for T cell development.
 987 *Immunity* **30**: 696-707
- 988 **Palm D, Simm S, Darm K, Weis BL, Ruprecht M, Schleiff E, Scharf C** (2016)
 989 Proteome distribution between nucleoplasm and nucleolus and its relation to
 990 ribosome biogenesis in *Arabidopsis thaliana*. *RNA Biol* **13**: 441-454
- 991 **Palm D, Streit D, Shanmugam T, Weis BL, Ruprecht M, Simm S, Schleiff E** (2019)
 992 Plant-specific ribosome biogenesis factors in *Arabidopsis thaliana* with essential
 993 function in rRNA processing. *Nucleic Acids Res* **47**: 1880-1895
- 994 **Pan WA, Tsai HY, Wang SC, Hsiao M, Wu PY, Tsai MD** (2015) The RNA recognition
 995 motif of NIFK is required for rRNA maturation during cell cycle progression. *RNA Biol*
 996 **12**: 255-267
- 997 **Parry G, Ward S, Cernac A, Dharmasiri S, Estelle M** (2006) The Arabidopsis
 998 SUPPRESSOR OF AUXIN RESISTANCE proteins are nucleoporins with an
 999 important role in hormone signaling and development. *Plant Cell* **18**: 1590-1603
- 1000 **Pasternak T, Tietz O, Rapp K, Begheldo M, Nitschke R, Ruperti B, Palme K** (2015)
 1001 Protocol: an improved and universal procedure for whole-mount immunolocalization
 1002 in plants. *Plant Methods* **11**: 50
- 1003 **Pendle AF, Clark GP, Boon R, Lewandowska D, Lam YW, Andersen J, Mann M,**
 1004 **Lamond AI, Brown JW, Shaw PJ** (2005) Proteomic analysis of the Arabidopsis
 1005 nucleolus suggests novel nucleolar functions. *Mol Biol Cell* **16**: 260-269
- 1006 **Pérez-Pérez JM, Candela H, Micol JL** (2009) Understanding synergy in genetic
 1007 interactions. *Trends Genet* **25**: 368-376
- 1008 **Petibon C, Malik Ghulam M, Catala M, Abou Elela S** (2021) Regulation of ribosomal
 1009 protein genes: an ordered anarchy. *WIREs RNA* **12**: e1632

- 1010 **Petricka JJ, Nelson TM** (2007) Arabidopsis nucleolin affects plant development and
1011 patterning. *Plant Physiol* **144**: 173-186
- 1012 **Ponce MR, Robles P, Lozano FM, Brotóns MA, Micol JL** (2006) Low-resolution
1013 mapping of untagged mutations. *Methods Mol Biol* **323**: 105-113
- 1014 **Ponce MR, Robles P, Micol JL** (1999) High-throughput genetic mapping in *Arabidopsis*
1015 *thaliana*. *Mol Gen Genet* **261**: 408-415
- 1016 **Pontvianne F, Abou-Ellail M, Douet J, Comella P, Matia I, Chandrasekhara C,**
1017 **Debures A, Blevins T, Cooke R, Medina FJ, Tourmente S, Pikaard CS, Sáez-**
1018 **Vásquez J** (2010) Nucleolin is required for DNA methylation state and the expression
1019 of *rRNA* gene variants in *Arabidopsis thaliana*. *PLOS Genet* **6**: e1001225
- 1020 **Pontvianne F, Matia I, Douet J, Tourmente S, Medina FJ, Echeverria M, Sáez-**
1021 **Vásquez J** (2007) Characterization of *AtNUC-L1* reveals a central role of nucleolin in
1022 nucleolus organization and silencing of *AtNUC-L2* gene in Arabidopsis. *Mol Biol Cell*
1023 **18**: 369-379
- 1024 **Robledo S, Idol RA, Crimmins DL, Ladenson JH, Mason PJ, Bessler M** (2008) The
1025 role of human ribosomal proteins in the maturation of rRNA and ribosome production.
1026 *RNA* **14**: 1918-1929
- 1027 **Sáez-Vásquez J, Delseny M** (2019) Ribosome biogenesis in plants: from functional 45S
1028 ribosomal DNA organization to ribosome assembly factors. *Plant Cell* **31**: 1945-1967
- 1029 **Salih KJ, Duncan O, Li L, Trosch J, Millar AH** (2020) The composition and turnover of
1030 the *Arabidopsis thaliana* 80S cytosolic ribosome. *Biochem J* **477**: 3019-3032
- 1031 **Samaha H, Delorme V, Pontvianne F, Cooke R, Delalande F, Van Dorsseleer A,**
1032 **Echeverria M, Saez-Vasquez J** (2010) Identification of protein factors and U3
1033 snoRNAs from a Brassica oleracea RNP complex involved in the processing of pre-
1034 rRNA. *Plant J* **61**: 383-398
- 1035 **Sánchez-García AB, Aguilera V, Micol-Ponce R, Jover-Gil S, Ponce MR** (2015)
1036 *Arabidopsis MAS2*, an essential gene that encodes a homolog of animal NF- κ B
1037 activating protein, is involved in 45S ribosomal DNA silencing. *Plant Cell* **27**: 1999-
1038 2015
- 1039 **Savada RP, Bonham-Smith PC** (2014) Differential transcript accumulation and
1040 subcellular localization of Arabidopsis ribosomal proteins. *Plant Sci* **223**: 134-145
- 1041 **Scarpin MR, Busche M, Martinez RE, Harper LC, Reiser L, Szakonyi D, Merchante**
1042 **C, Lan T, Xiong W, Mo B, Tang G, Chen X, Bailey-Serres J, Browning KS,**
1043 **Brunkard JO** (2022) An updated nomenclature for plant ribosomal protein genes.
1044 *Plant Cell* **35**: 640–643
- 1045 **Scott MS, Troshin PV, Barton GJ** (2011) NoD: a nucleolar localization sequence
1046 detector for eukaryotic and viral proteins. *BMC Bioinformatics* **12**: 317

- 1047 **Schwacke R, Schneider A, van der Graaff E, Fischer K, Catoni E, Desimone M,**
 1048 **Frommer WB, Flugge UI, Kunze R** (2003) ARAMEMNON, a novel database for
 1049 Arabidopsis integral membrane proteins. *Plant Physiol* **131**: 16-26
- 1050 **Shanmugam T, Streit D, Schroll F, Kovacevic J, Schleiff E** (2021) Dynamics and
 1051 thermal sensitivity of ribosomal RNA maturation paths in plants. *J Exp Bot* **72**: 7626-
 1052 7644
- 1053 **Slavov N, Semrau S, Airoidi E, Budnik B, van Oudenaarden A** (2015) Differential
 1054 stoichiometry among core ribosomal proteins. *Cell Rep* **13**: 865-873
- 1055 **Sperschneider J, Catanzariti AM, DeBoer K, Petre B, Gardiner DM, Singh KB,**
 1056 **Dodds PN, Taylor JM** (2017) LOCALIZER: subcellular localization prediction of both
 1057 plant and effector proteins in the plant cell. *Sci Rep* **7**: 44598
- 1058 **Tafforeau L, Zorbas C, Langhendries JL, Mullineux ST, Stamatopoulou V, Mullier**
 1059 **R, Wacheul L, Lafontaine DL** (2013) The complexity of human ribosome biogenesis
 1060 revealed by systematic nucleolar screening of pre-rRNA processing factors. *Mol Cell*
 1061 **51**: 539-551
- 1062 **Thoms M, Thomson E, Bassler J, Gnadig M, Griesel S, Hurt E** (2015) The exosome
 1063 is recruited to RNA substrates through specific adaptor proteins. *Cell* **162**: 1029-1038
- 1064 **Tyagi A, Gupta A, Dutta A, Potluri P, Batti B** (2020) A review of Diamond-Blackfan
 1065 anemia: current evidence on involved genes and treatment modalities. *Cureus* **12**:
 1066 e10019
- 1067 **Wang R, Zhao J, Jia M, Xu N, Liang S, Shao J, Qi Y, Liu X, An L, Yu F** (2018) Balance
 1068 between cytosolic and chloroplast translation affects leaf variegation. *Plant Physiol*
 1069 **176**: 804-818
- 1070 **Warner JR** (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem*
 1071 *Sci* **24**: 437-440
- 1072 **Weaver PL, Sun C, Chang TH** (1997) Dbp3p, a putative RNA helicase in
 1073 *Saccharomyces cerevisiae*, is required for efficient pre-rRNA processing
 1074 predominantly at site A₃. *Mol Cell Biol* **17**: 1354-1365
- 1075 **Wilson DN, Cate JHD** (2012) The structure and function of the eukaryotic ribosome.
 1076 *Cold Spring Harb Perspect Biol* **4**: a011536
- 1077 **Xiong W, Zhang J, Lan T, Kong W, Wang X, Liu L, Chen X, Mo B** (2021) High
 1078 resolution RNA-seq profiling of genes encoding ribosomal proteins across different
 1079 organs and developmental stages in *Arabidopsis thaliana*. *Plant Direct* **5**: e00320
- 1080 **Yang Y, Isaac C, Wang C, Dragon F, Pogacic V, Meier UT** (2000) Conserved
 1081 composition of mammalian box H/ACA and box C/D small nucleolar ribonucleoprotein
 1082 particles and their interaction with the common factor Nopp140. *Mol Biol Cell* **11**: 567-
 1083 577

1084 **Zhu P, Wang Y, Qin N, Wang F, Wang J, Deng XW, Zhu D** (2016) Arabidopsis small
1085 nucleolar RNA monitors the efficient pre-rRNA processing during ribosome
1086 biogenesis. Proc Natl Acad Sci U S A **113**: 11967-11972

1087

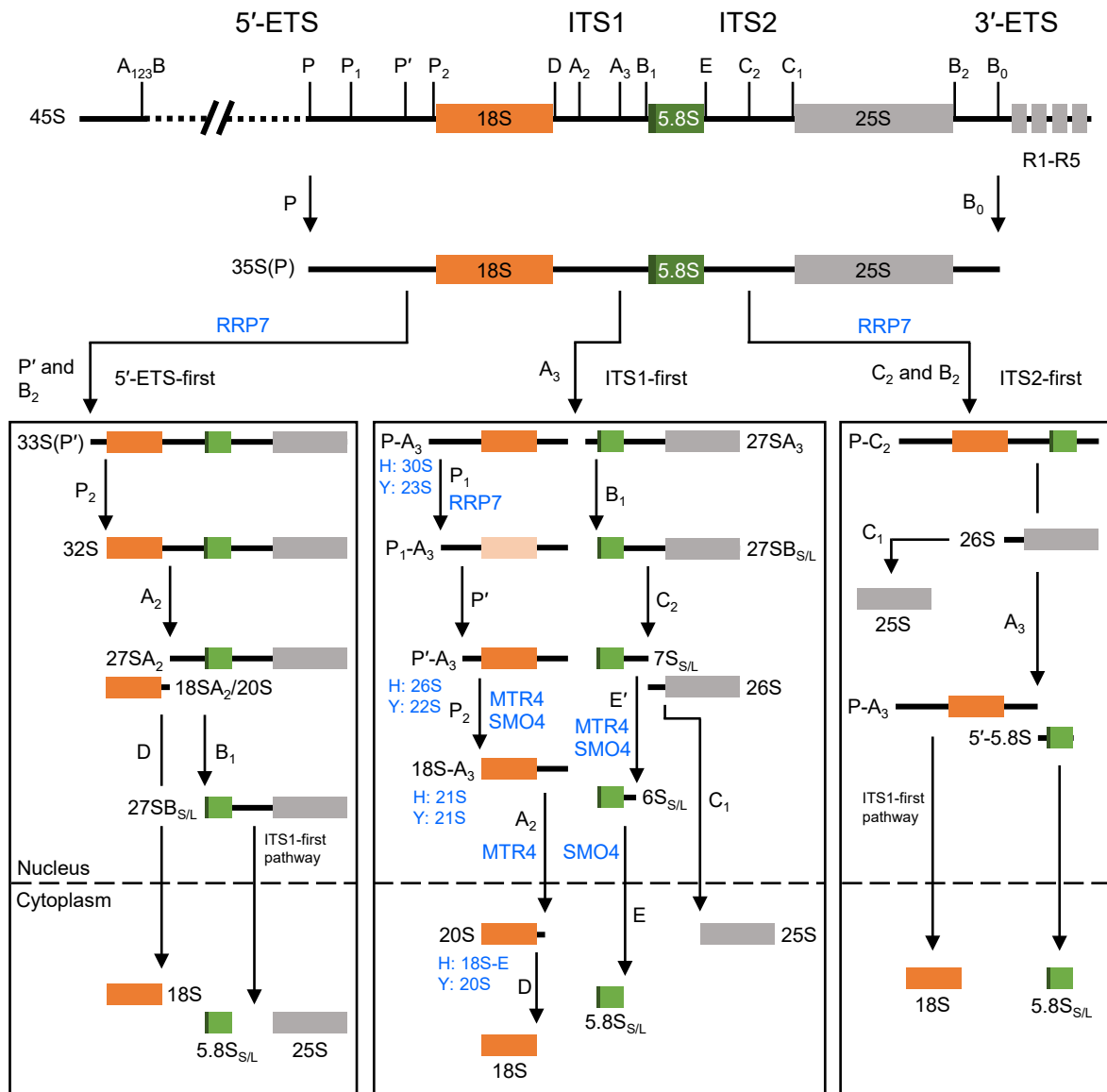
1088

Cross-kingdom conservation of Arabidopsis RPS24 function in 18S rRNA maturation

**Adrián Cabezas-Fuster, Rosa Micol-Ponce,
Raquel Sarmiento-Mañús and María Rosa Ponce**

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

Supplemental Figures, Tables and References



Supplemental Figure 1. Overview of 45S pre-rRNA processing in Arabidopsis. Colored boxes represent the sequences of the three mature rRNAs transcribed from the 45S rDNA genes. Vertical arrows indicate endonucleolytic cleavages; letters indicate the cleavage site in the corresponding pre-rRNA. Only the relevant factors and the human (H) and yeast (Y) 18S pre-rRNAs relevant to this study are represented. Based on information from Sáez-Vasquez and Delseny (2019).

```

RPS24B      1 MAEKAVTIRTRNFMTNRLLARKQFVIDVLHPGRANVSKAELKEKLARMYEVKDPNAIFCF
RPS24A      1 MAEKAVTIRTRKFMNRLLSRKQFVIDVLHPGRANVSKAELKEKLARMYEVKDPNAIFVE
consensus   1 ***** * ***** * ***** * ***** * ***** *

```



```

RPS24B      61 KFRTHFGGGKSSGYGLIYDTVENAKKFEPKYRLIRNGLDTKIEKSRQIKERKNRAKKIR
RPS24A      61 KFRTHFGGGKSSGFGLIYDTVESAKKFEPKYRLIRNGLDTKIEKSRQIKERKNRAKKIR
consensus   61 ***** * ***** * ***** * ***** * ***** *

```

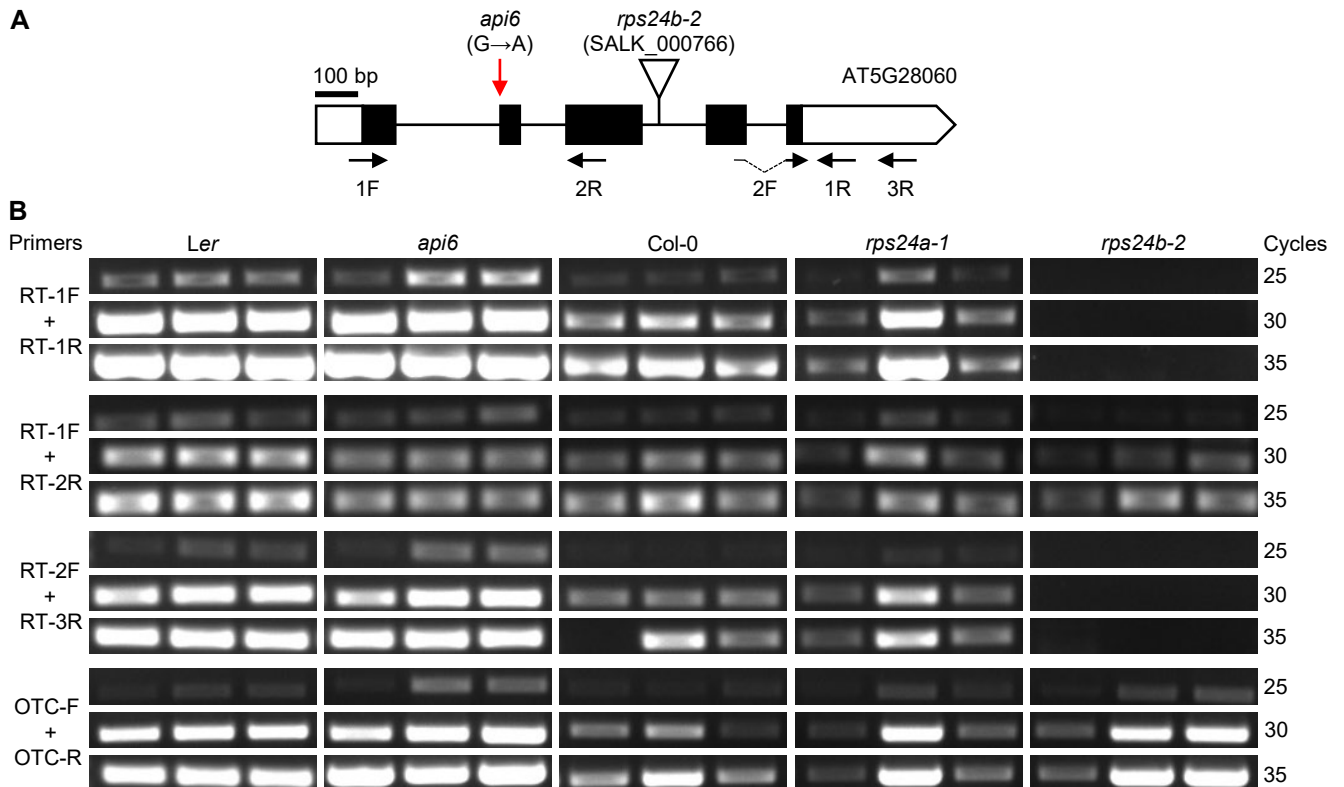


```

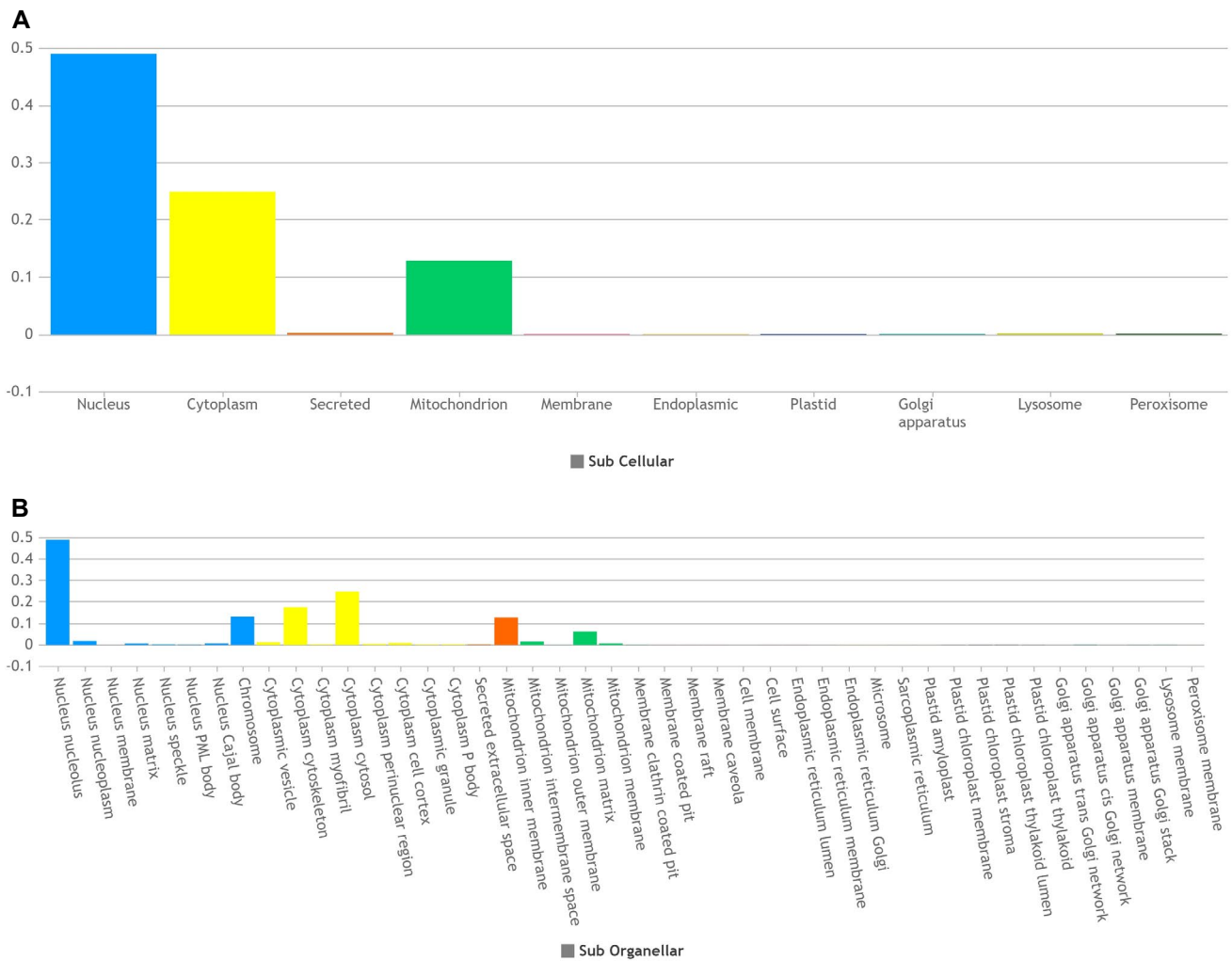
RPS24B      121 GVKKTKAGDTKKK
RPS24A      121 GVKKTKAGDAKKK
consensus   121 ***** *

```

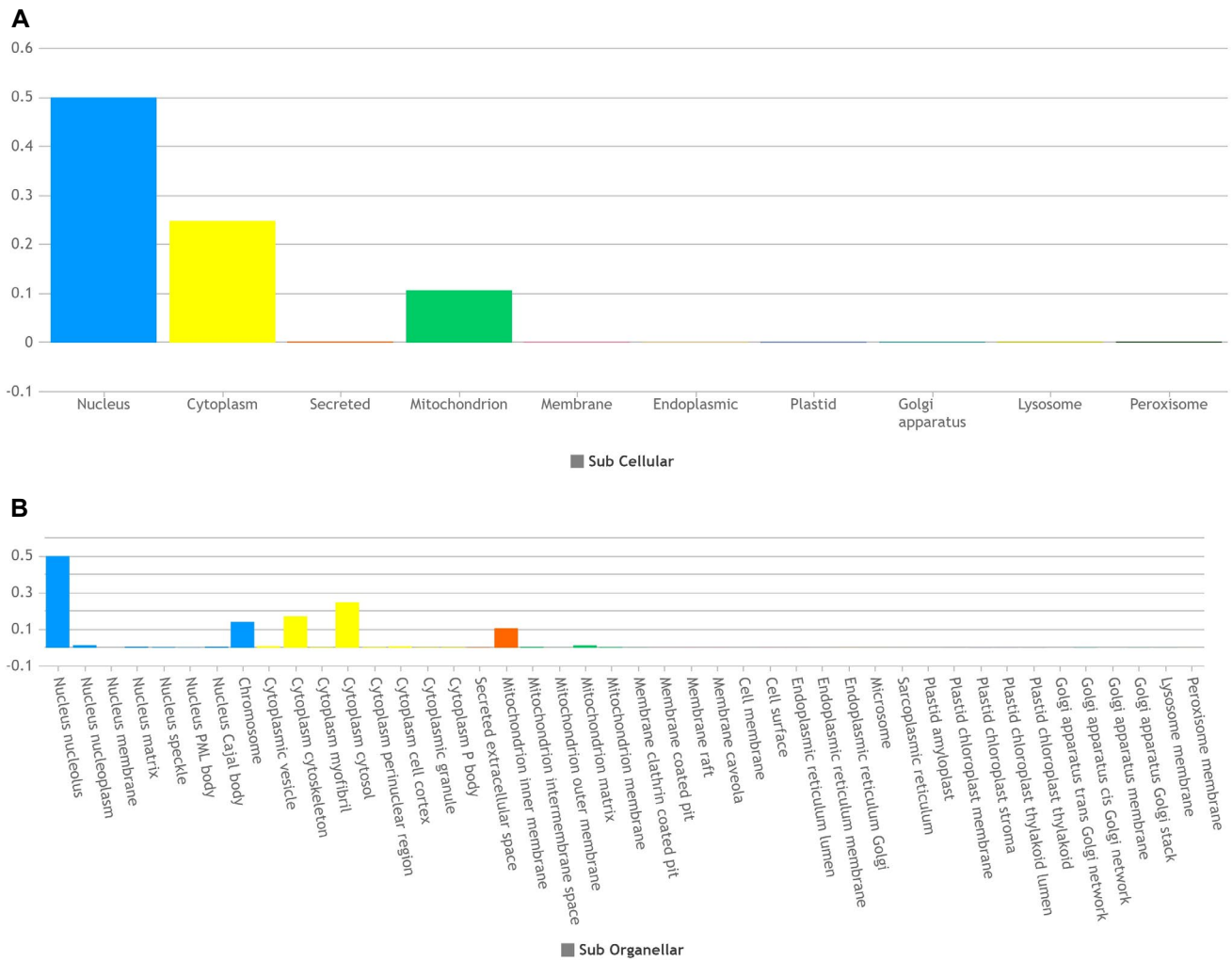
Supplemental Figure 2. Sequence conservation between the RPS24A and RPS24B paralogs. Amino acid sequence alignment of Arabidopsis RPS24B and RPS24A. The 14 amino acids that are absent from the RPS24B variant produced by the expression of the *api6* allele are highlighted in red letters. Identical and similar residues are shaded in black and gray, respectively. Asterisks and dots in the consensus line indicate identical and conserved residues, respectively. Numbers indicate residue positions. The alignment was obtained using ClustalW2 and shaded with Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html). Nuclear and nucleolar localization sequences predicted with LOCALIZER (<https://localizer.csiro.au/>) and NoD software (<http://www.compbio.dundee.ac.uk/www-nod/index.jsp>) are underlined in black and green, respectively.



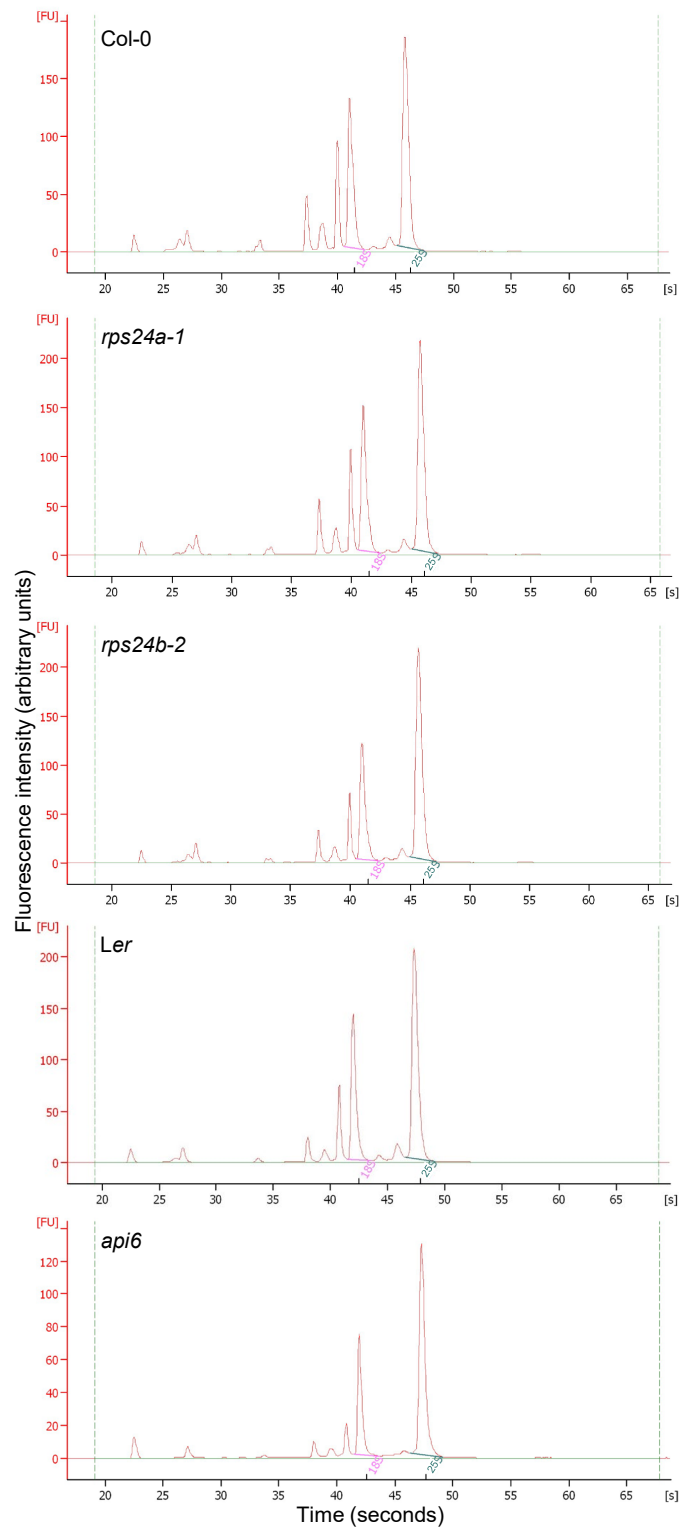
Supplemental Figure 3. *RPS24B* expression analysis. (A) Schematic representation of *RPS24B* and its mutant alleles used in this study. Gene structure is represented as described in the legend of Figure 1. Black arrows represent the primers used for semiquantitative RT-PCR amplifications in (B). (B) Semiquantitative RT-PCR analysis of *RPS24B* in the *rps24a* and *rps24b* mutants used in this study. The bands for each PCR were visualized after 25, 30 and 35 cycles of amplification. Total RNA was extracted from seedlings collected at 15 das. Transcripts from the *OTC* gene were used as an internal control. The primer sequences used are shown in Supplemental Table 2.



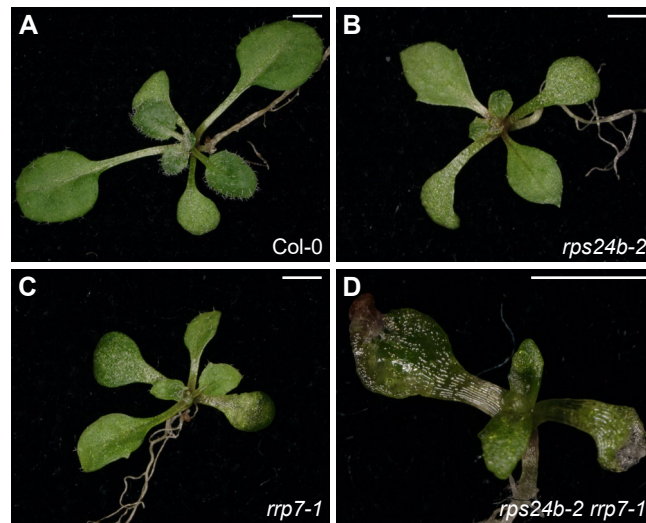
Supplemental Figure 4. Predicted localization of RPS24A. (A-B) Predicted subcellular (A) and sub-organellar (B) localization of RPS24A using MULocDeep software (<https://mu-loc.org/>).



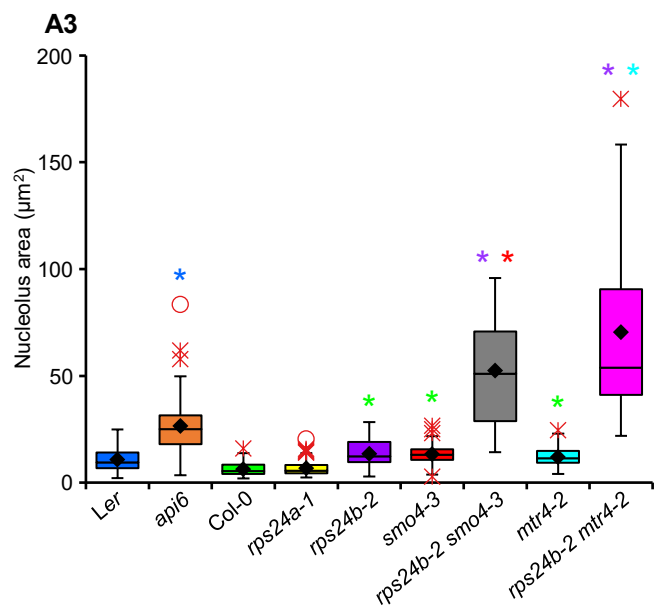
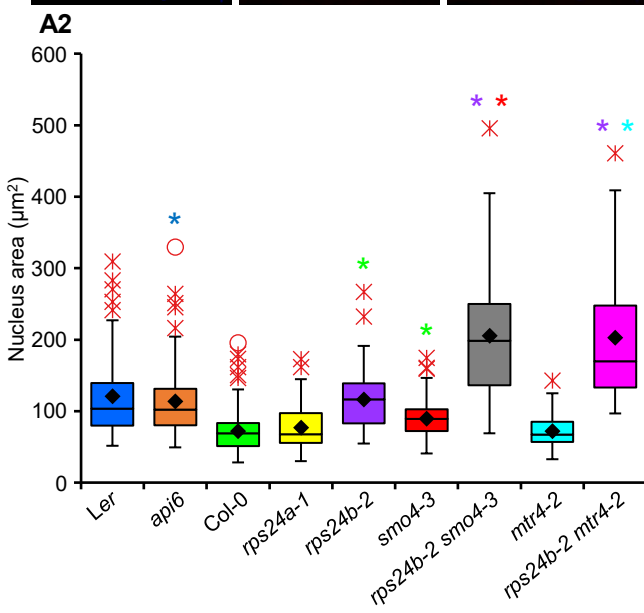
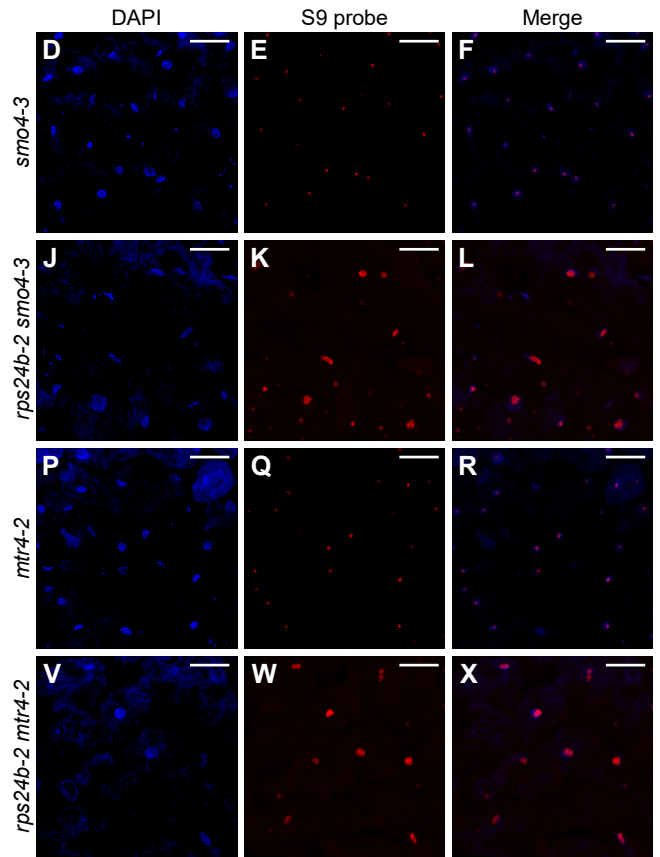
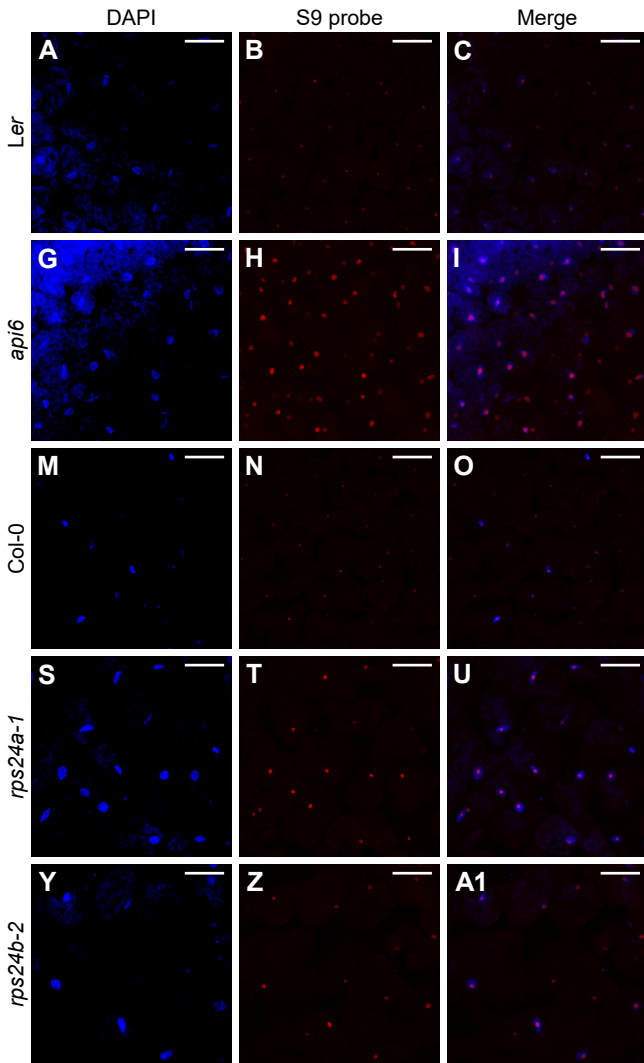
Supplemental Figure 5. Predicted localization of RPS24B. (A-B) Predicted subcellular (A) and sub-organellar (B) localization of RPS24B protein using MULocDeep software (<https://mu-loc.org/>).



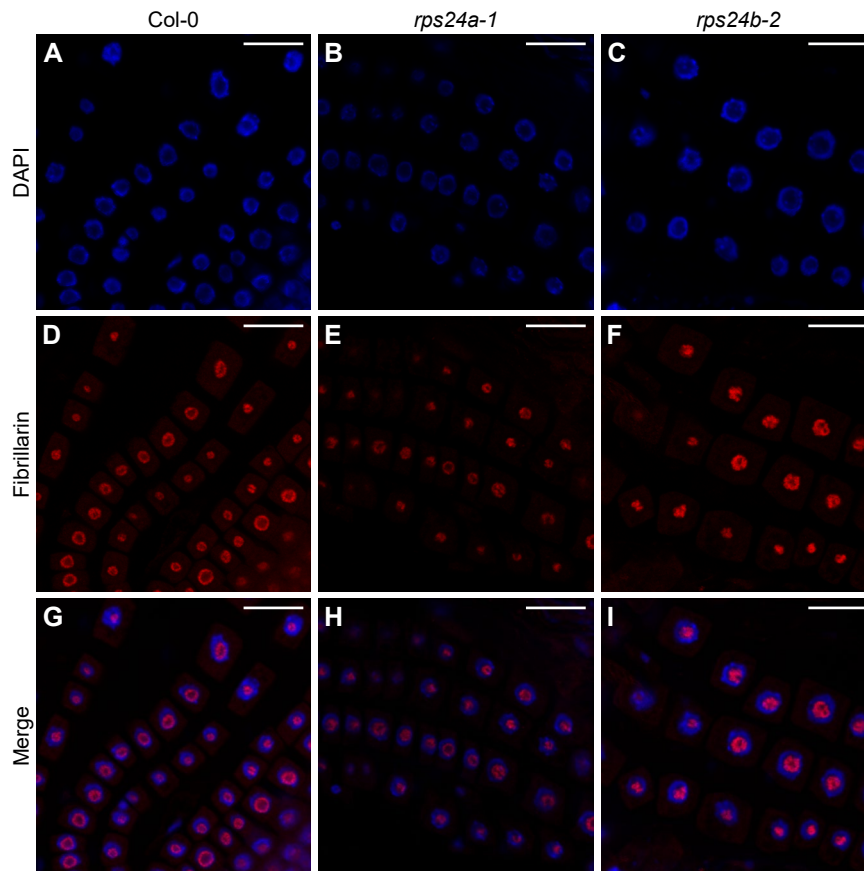
Supplemental Figure 6. Relative amounts of 18S and 25S RNA in the *rps24* mutants. Agilent 2100 Bioanalyzer electropherogram profiles of total RNA from Col-0, *rps24a-1*, *rps24b-2*, Ler and *api6*. Total RNA was extracted from seedlings collected 15 das.



Supplemental Figure 7. Genetic interactions between *rps24b-2* and *rrp7-1*. Rosettes of Col-0 (A), *rps24b-2* (B), *rrp7-1* (C) and *rps24b-2 rrp7-1* plants (D). Photographs were taken from seedlings collected 14 das. Scale bars, 2 mm.



Supplemental Figure 8. Subcellular localization of 5.8S rRNA precursors. (A-A1) RNA-FISH assay in palisade mesophyll cells from first-node leaves of *Ler* (A-C), *smo4-3* (D-F), *api6* (G-I), *rps24b-2 smo4-3* (J-L), *Col-0* (M-O), *mtr4-2* (P-R), *rps24a-1* (S-U), *rps24b-2 mtr4-2* (V-X) and *rps24b-2* plants (Y-A1). Fluorescent signals correspond to DAPI (in blue; A, D, G, J, M, P, S, V, and Y), which was used as a nuclear marker; the S9 probe labeled with Cy3 (in red; B, E, H, K, N, Q, T, W, and Z); and (C, F, I, L, O, R, U, X, and A1) their overlay. Plants were collected 21 das. Scale bars, 49 μ m. (A2 and A3) Nucleus (A2) and nucleolus (A3) areas measured as DAPI and S9 probe fluorescence, respectively. Asterisks indicate significant differences from the wild type and parental lines (indicated by color) in a Student's *t*-test ($*P < 0.0001$).



Supplemental Figure 9. Nucleolus organization in *rps24* root cells. (A-I) Visualization by immunolocalization of the fibrillarlin nuclear marker in root cells of Col-0 (A, D, and G), *rps24a-1* (B, E, and H) and *rps24b-2* plants (C, F, and I). Fluorescent signals correspond to DAPI (A-C), the secondary antibody conjugated with TRITC for fibrillarlin detection (D-E) and their overlay (G-I). Immunolocalization was performed in at least 5 seedlings per genotype, collected 5 das.

Supplemental Table 1. Primer sets used for the fine mapping of *api6*

Marker name	Locus	Oligonucleotide sequences (5'→3')		PCR product size (bp)	
		Forward primer	Reverse primer	Ler	Col-0
cer449133	AT5G27905-AT5G27910	CGATTGTTTGTTCACCTTTCAA ^a	GCGTATGTTTGGGGATAGG ^b	262	287
cer451402	AT5G28200	GTGTATCATGGACCATCGCG ^c	AATTGTTTTGTAGGTCGCTAAC ^d	237	229

The oligonucleotide names are: ^acer449133-F, ^bcer449133-R, ^ccer451402-F and ^dcer451402-R.

Supplemental Table 2. Other primers used in this work

Purpose	Oligonucleotide name(s)	Oligonucleotide sequences (5'→3')		
		Forward primer (F)	Reverse primer (R)	
Genotyping of	<i>rps24b-2/api6</i>	AT5G28060-F1/R1	CTAATCTATTCTCTGGGCATGG	CAGCCTTGGTCTACACTCAC
	<i>rps24a-1</i>	AT3G04920-F1/R1	CCTGGAAGAGCCAATGTTTCA	ATGGGAATGGTGGAAAGAGAC
	<i>smo4-3</i>	NOP53-F1/R1 ^a	GTCTCGAACTTTTTCTTGGG	AGTATTCCTCGCTTCTCGAGG
	<i>mtr4-2</i>	MTR4-F/R ^a	TTTGTCAATACCTCGACGTCC	ATTGTCTGCGTACTGTGGGTC
	<i>rrp7-1</i>	RRP7-F1/R1	CTCATGAAGAACGCCTTGAAC	GTGGAGATCGTGGAGATGAAG
	<i>par11-2</i>	PARL1-F/R ^b	AGTTGCTGTCACCAAGAAG	TGGCCTACCATGGAATTCA
T-DNA insertion verification	Salk_LBb1.3 ^c	GCGTGGACCGCTTGCTGCAACT		
	Salk_Rb1 ^c	CGTGACTIONCCTTAATTCTCCGC		
	Sail_LB1 ^c	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC		
Construction of transgenes	35Spro:RPS24B:GFP-F/R	<i>GGGGACCACCTTTGTACAAGAAAGCTGGGTG</i>	<i>GGGGACCACCTTTGTACAAGAAAGCTGG</i>	
		<i>TTACTTCTTCTTGGTATCACCAGC</i>	<i>GTGCTTCTTCTTGGTATCACCAGC</i>	
Semiquantitative RT-PCR	RPS24B-RT-1F/1R	AAGTAAATCGCAGCCATGGC	TTGTTCTGCATCACTCCTTCTT	
	RPS24B-RT-2F/2R	AGTACAGACTTATCAGGAATGGA	GGCGAGGATGTATGAGGTTAAG	
	RPS24B-RT-3R		CCTCTTGCGTTTTCGGAGATT	
	OTC-REV/3D	GCATGCATGCGATTCTCCGC	TCCTTGCCAAATCATGGCCG	
Quantitative RT-PCR	45S pre-rRNA-45S-F/R ^d	CGGTCCGGTCATTCCTCGTGTGATATC	TATAGGGGGGTGGGTGTTGAGGGA	
45S rDNA variant expression	p3/p4 ^e	GACAGACTTGTCCAAAACGCCACC	CTGGTCCGAGGAATCCTGGACGATT	
	OTC-F/R ^f	TGAAGGGACAAAGGTTGTGTATGTT	CGCAGACAAGTGGAAATGGA	
Probes for RNA gel blots and <i>in situ</i> hybridization	S2-F/R ^g	TAGGCTGTCCCGAAGTATC ^h	TCACTTCGAGTCACCGTCGACA ⁱ	
	S7 ^g	GTCGTTCTGTTTTGGACAGGTATCGA ^h		
	S9 ^g	AGGATGGTGAGGGACGACGATTT ^{h,i}		

Sequences taken from ^aMicol-Ponce et al., 2020, ^bMicol-Ponce et al., 2018, ^c<http://signal.salk.edu/tdnaprimers.2.html>, ^dZhu et al., 2016, ^ePontvianne et al., 2010, ^fCnops et al., 2004, and ^gLange et al., 2011. ^{h,i}Oligonucleotides labeled with ^hDIG (Digoxigenin) and ⁱCy3 (Cyanine 3). The *attB* sequences are shown in italics.

SUPPLEMENTAL REFERENCES

- Cnops G, Jover-Gil S, Peters JL, Neyt P, De Block S, Robles P, Ponce MR, Gerats T, Micol JL, Van Lijsebettens M** (2004) The *rotunda2* mutants identify a role for the *LEUNIG* gene in vegetative leaf morphogenesis. *J Exp Bot* **55**: 1529-1539
- Lange H, Sement FM, Gagliardi D** (2011) MTR4, a putative RNA helicase and exosome co-factor, is required for proper rRNA biogenesis and development in *Arabidopsis thaliana*. *Plant J* **68**: 51-63
- Micol-Ponce R, Sarmiento-Mañús R, Fontcuberta-Cervera S, Cabezas-Fuster A, de Bures A, Sáez-Vásquez J, Ponce MR** (2020) SMALL ORGAN4 is a ribosome biogenesis factor involved in 5.8S ribosomal RNA maturation. *Plant Physiol* **184**: 2022-2039
- Micol-Ponce R, Sarmiento-Mañús R, Ruiz-Bayón A, Montacié C, Sáez-Vásquez J, Ponce MR** (2018) Arabidopsis RIBOSOMAL RNA PROCESSING7 is required for 18S rRNA maturation. *Plant Cell* **30**: 2855-2872
- Pontvianne F, Abou-Elail M, Douet J, Comella P, Matia I, Chandrasekhara C, Debures A, Blevins T, Cooke R, Medina FJ, Tourmente S, Pikaard CS, Sáez-Vásquez J** (2010) Nucleolin is required for DNA methylation state and the expression of *rRNA* gene variants in *Arabidopsis thaliana*. *PLOS Genet* **6**: e1001225
- Zhu P, Wang Y, Qin N, Wang F, Wang J, Deng XW, Zhu D** (2016) Arabidopsis small nucleolar RNA monitors the efficient pre-rRNA processing during ribosome biogenesis. *Proc Natl Acad Sci USA* **113**: 11967-11972

X.- AGRADECIMIENTOS

X.- AGRADECIMIENTOS

La realización de esta Tesis ha sido posible gracias a la financiación del trabajo que se realiza en el laboratorio de María Rosa Ponce por la Generalitat Valenciana (PROMETEO/2019/117) y el Ministerio de Ciencia e Innovación (BIO2014-56889-R, BIO2017-89728-R y PID2020-117125RB).

En primer lugar, quiero dar las gracias a mi directora, María Rosa Ponce, por darme la oportunidad de realizar esta Tesis en su laboratorio y bajo su dirección, por transmitirme su pasión por la ciencia y la investigación y enseñarme que de los errores es de donde más se aprende.

A José Luis Micol, por su interés en mi trabajo con sus preguntas y discusiones, y por aportar siempre un punto de vista crítico y constructivo.

A todos los profesores del área de Genética, por haberme iniciado en un campo que se ha acabado convirtiendo en el centro de mi carrera científica, enseñarme todo lo que se esconde tras un gen y, sobre todo, lo que queda por descubrir de ellos.

A los técnicos de los laboratorios de María Rosa Ponce y José Luis Micol con los que he coincidido: María José, Diana, Juan, María y, en especial, José Manuel Serrano, por todo lo que me ha aportado como persona fuera del laboratorio. Gracias a todos ellos he aprendido que el día a día en un laboratorio es mucho más complejo e importante que lo que nuestros experimentos finales reflejan.

A todos los alumnos que he tenido bajo mi dirección: Yaiza, Eloy, Jesús, Antonio, Noel, Alejandro, Patricia, Lucía y Gabriela, por su voluntad de aprender, su inestimable ayuda y enseñarme a ver y explicar la ciencia desde los ojos de otros.

A todos mis compañeros de laboratorio. A los más antiguos, Tamara, David y Edu, que me arroparon en mis inicios, y en especial a Rosa, por haber dirigido mi TFG cuando no sabía nada, y ayudarme en los primeros pasos de mi Tesis cuando decidí emprender un camino totalmente desconocido. A Raquel, por su confianza, su apoyo y soportar mis constantes preguntas sin dejar nunca de responderme. También a los más nuevos, Emilio, Gabriela, Juanfe y Rossy, con los que he compartido menos tiempo, pero de los que me llevo un gran recuerdo.

A mis contemporáneos, Carla, Riad, Lucía, Samuel, Uri y Sara con los que he compartido 7 años de mi vida, e infinidad de horas de laboratorio, congresos, comidas, cenas, deporte, risas y alguna lágrima. A Alejandro y Sergio, por haber sido unos buenos compañeros durante estos años, y llevarme unos grandes amigos para siempre. A Àngela, la que ha passat de ser la nova del laborator i a ser una part fonamental de la meua vida. Gràcies per haver

estat quan més ho necessitava. En aquest món de ciència, per a mi has sigut un gran descobriment.

A mis chicos de Biotec: Mateo, Adri, Carlos, Dani y Alex, por su apoyo constante, en la vida y en la ciencia, y por las madrugadas intentando resolver esta Tesis sin saber muy bien por qué. Als meus amics de Crevillent, els de sempre, per evadir-me de la ciència i el treball i fer més fàcil aquest recorregut. A Raquel, per tot el camí que vam fer junts, per creure en mi i donar-me suport sempre. Per la Lluna.

A la part més important de mi, la meua família. Els meus pares, Galo i Salud, la meua germana Esther y la meua àvia, Pepi. També a la meua tia Blanca, i als meus dos cosins, Carla i Álvaro. Per donar-me la vida, estimar-me i estar sempre al meu costat. Tot el que soc es per vosaltres. Gràcies.