Title

The *Arabidopsis* SR45 splicing factor bridges the splicing machinery and the exon-exon junction complex

Authors

Steven Fanara[1], Marie Schloesser[1], Marine Joris[1], Simona De Franco[2], Marylène Vandevenne[2], Frédéric Kerff[3], Marc Hanikenne[4†] and Patrick Motte[1†]

[1]InBioS-PhytoSystems, Functional Genomics and Plant Molecular Imaging, University of Liège, 4000, Liège, Belgium

[2]InBioS-Center for Protein Engineering, Laboratory of Biological Macromolecules, University of Liège, 4000, Liège, Belgium

[3]InBioS-Center for Protein Engineering, Laboratory of Crystallography, University of Liège, 4000, Liège, Belgium

[4]InBioS-PhytoSystems, Translational Plant Biology, University of Liège, 4000, Liège, Belgium

[†]These two authors contributed equally.

Email address for each author

Steven Fanara: steven.fanara@uliege.be

Marie Schloesser: mschloesser@uliege.be

Marine Joris: mjoris@uliege.be

Simona De Franco: sdefranco4@gmail.com

Marylène Vandevenne: mvandevenne@uliege.be

Frédéric Kerff: fkerff@uliege.be

Marc Hanikenne: marc.hanikenne@uliege.be

Patrick Motte: patrick.motte@uliege.be

Corresponding authors
Steven Fanara (steven.fanara@uliege.be; Tel: +32-4-366-38-11)
Correspondence may also be addressed to Patrick Motte (patrick.motte@uliege.be; Tel: +32-4-366-38-10)
University of Liège, InBioS-PhytoSystems
Quartier de la Vallée, 1
Chemin de la Vallée, 4 - Bât B22, B4000, Liège, Belgium

Date of submission: 07/11/2023

Number of figures: 8

Word count: 6413

Number of supplementary figures: 11

Number of supplementary tables: 15

Running title

Functional characterization of the SR45 splicing factor

Highlight

The contribution of the Arabidopsis SR45 splicing factor individual domains to its nuclear localization, ability to contact *in planta* novel protein partners and specifically bind RNA motifs was examined.

Abstract

The Arabidopsis splicing factor serine/arginine-rich 45 (SR45) contributes to several biological processes. The *sr45-1* loss-of-function mutant exhibits delayed root development, late flowering, unusual numbers of floral organs, shorter siliques with decreased seed sets, narrower leaves and petals, and altered metal distribution. SR45 bears a unique RNA recognition motif (RRM) flanked by one serine/arginine-rich (RS) domain on both sides. Here, we studied the function of each of SR45 domains by examining their involvement in: (i) the spatial distribution of SR45, (ii) the establishment of a protein-protein interaction network including spliceosomal and exon-exon junction complex (EJC) components, and (iii) the RNA binding specificity. We report that the endogenous *SR45* promoter is active during vegetative and reproductive growth, and that the SR45 protein localizes in the nucleus. We demonstrate that the C-terminal arginine/serine-rich domain is a determinant of nuclear localization. We show that the SR45 RNA recognition motif (RRM) domain specifically binds purine-rich RNA motifs via three residues (H101, H141, Y143), and is also involved in protein-protein interactions. We further show that SR45 bridges both mRNA splicing and surveillance machineries as a partner of EJC core components and peripheral factors, which requires phosphoresidues likely phosphorylated by kinases from both CLK and SRPK families. Our findings provide insights into the contribution of each SR45 domain to both spliceosome and EJC assemblies.

Keywords: Serine/arginine-rich splicing factor, SR45, Arabidopsis, splicing, exon-exon junction complex, EJC, protein-RNA and protein-protein interaction, ASAP, PSAP

Introduction

The precise recognition of splice sites at exon-intron boundaries is crucial in order to accurately perform precursor-messenger RNA (pre-mRNA) splicing (Wahl *et al.*, 2009). While constitutive splicing results from the exclusion of each intron from the pre-mRNA, alternative splicing (AS) can lead to intron retention, shortened or lengthened exon or the skipping of an entire exon (Reddy *et al.*, 2013; Wang *et al.*, 2015). This occurs when weak or non-canonical splice sites are recognized and processed by the splicing machinery (Chen and Manley, 2009). It has important consequences for mRNA abundance and protein stability (Graveley, 2001). AS can be linked to an mRNA degrading process called nonsense-mediated mRNA decay (NMD) through the emergence of premature termination codons (PTCs) in alternatively spliced transcripts, which ultimately regulates transcript levels (Palusa and Reddy, 2010; Kalyna *et al.*, 2012). Protein isoforms that arise from alternative splicing usually differ in their primary sequence and domain organization which often alters their localization, stability, function or ability to interact with their protein or RNA partners (Laloum *et al.*, 2018).

Serine/arginine-rich (SR) proteins constitute a highly conserved family in multicellular eukaryotes (Califice *et al.*, 2012), sharing a modular structure characterized by the presence of one or two RNA recognition motif(s) [RRM(s)] in the N-terminus and a C-terminal domain rich in arginine-serine dipeptide repeats (RS) (Manley and Krainer, 2010; Califice *et al.*, 2012). While the RNA binding affinity is mainly dictated by the RRM domain (Maris *et al.*, 2005; Daubner *et al.*, 2013), the RS domain is involved in protein-RNA and protein-protein interactions through reversible phosphorylation (Xiao and Manley, 1997; Shen *et al.*, 2004; Schütze *et al.*, 2016; Schwich *et al.*, 2021). The phosphorylation status of the RS domain is controlled by at least two conserved protein kinase families, serine arginine protein kinases (SRPK) and *cdc2*-like kinases (CLK), as wells as some phosphatases (e.g., Protein Phosphatase 1, PP1) (Aubol *et al.*, 2016; Stankovic *et al.*, 2016; Aubol *et al.*, 2017). The SRPK kinases are located in the cytoplasm and the nucleus and strictly phosphorylate serine-arginine dipeptides (Ma *et al.*, 2009; Koutroumani *et al.*, 2017). They influence the localization of SR proteins by supporting their phosphorylation-dependent transport into the nucleus via an SR-specific transportin protein (TRN-SR, TNPO3) (Kataoka *et al.*, 1999; Lai *et al.*, 2000; Lai *et al.*, 2001; Xu *et al.*, 2011). The CLK members are dual kinases and strictly nuclear (Lee *et al.*, 1996). They phosphorylate not only serine-arginine but also serine-proline dipeptides (Lee *et al.*, 1996; Aubol *et al.*, 2013). Some SR proteins display

a nucleocytoplasmic dynamics and are known to play many additional roles in (pre-)mRNA processing in mammals, including transcription, mRNA export (Huang and Steitz, 2001), stability and surveillance (Zhang and Krainer, 2004), and translation (Sanford *et al.*, 2004). Based on the sole phylogeny of the RRM, the *Arabidopsis thaliana* (Arabidopsis) genome possesses nineteen SR protein-encoding genes divided into seven subgroups, four of which are plant-specific because of a unique multidomain structure (RS, SCL, RS2Z and SR45 subfamilies) (Barta *et al.*, 2010; Califice *et al.*, 2012).

SR45 is an atypical SR protein because it possesses a unique RRM flanked by one RS domain on each side (Barta *et al.*, 2010), a domain organization similar to the one of the human splicing factors Transformer-2β (Tra2β) and RNPS1 (RNA-binding Protein with Ser-rich domain 1) (Cléry *et al.*, 2011), the human ortholog of SR45 (Califice *et al.*, 2012). RNPS1 plays critical roles in splicing (Robinson *et al.*, 1999) and NMD by coupling the deposition of a quality control complex (Exon Junction Complex, EJC) on spliced mRNAs to the UP-FRAMESHIFT (UPF) protein complex, which consists of UPF1, UPF2 and UPF3 and triggers NMD (Lykke-Andersen *et al.*, 2001). Two scaffold proteins, ACINUS and PININ, form with the EJC peripheral factors RNPS1 and SAP18 (Sin3A Associated Protein 18) the ternary complexes ASAP and PSAP, respectively, in order to regulate AS and NMD (Sakashita *et al.*, 2004; Singh *et al.*, 2010; Murachelli *et al.*, 2012; Schlautmann *et al.*, 2022). The EJC itself is a conserved protein complex comprising four highly conserved core proteins named eukaryotic initiation factor 4A3 (eIF4AIII), Mago nashi (MAGOH/Mago), Y14 and Metastatic Lymph Node 51 (MLN51) (Tange *et al.*, 2005). The cytoplasmic disassembly of the EJC results from the action of PYM (Bono *et al.*, 2004). This EJC core serves as a platform for multiple peripheral proteins in order to coordinate several steps of mRNA processing, namely (alternative) splicing, intra-cellular RNA localization (which notably involves ALY/REF), translation efficiency and mRNA stability control by NMD (Le Hir *et al.*, 2001; Nott *et al.*, 2004; Singh *et al.*, 2012).

In Arabidopsis, little is known about the functional characteristics of the EJC components and the associated peripheral factors (Chen *et al.*, 2019). The Arabidopsis homologues of MAGO, Y14, PYM and eIF4AIII have been previously described (Johnson *et al.*, 2004; Park and Muench, 2007; Park *et al.*, 2009). MAGO and Y14 can heterodimerize (Park and Muench, 2007). While the C-terminal portion of MAGO is critical for this interaction, it is dispensable for the association with eIF4AIII (Cilano *et al.*, 2016). PYM interacts with the MAGO-Y14 complex, as

well as with the isolated proteins (Park and Muench, 2007). PYM and Y14 associate only when both proteins are dephosphorylated, but the phosphorylation state does not affect the association with MAGO (Mufarrege *et al.*, 2011). Knockdown of *MAGO* in plants leads to several developmental defects, including vegetative and reproductive development impairment (Park *et al.*, 2009). The DEAD-box RNA helicase eIF4AIII is a nuclear protein whose subcellular localization is altered under hypoxia (Koroleva *et al.*, 2009). It interacts with ALY4 *in vitro* and colocalizes with MAGO and Y14 *in planta* (Koroleva *et al.*, 2009). While RNPS1 was shown to interact with Y14, no link was established between SR45 and Y14 or any other EJC core component in Arabidopsis (Shi and Xu, 2003; Hsu Ia *et al.*, 2005; Wang *et al.*, 2018). However, it was recently shown that ASAP and PSAP complexes containing SR45 exist in Arabidopsis and contribute to the regulation of *FLC* expression (Qüesta *et al.*, 2016; Chen *et al.*, 2019; Bi *et al.*, 2021; Mikulski *et al.*, 2022) and to the reproductive transition (Bi *et al.*, 2021). SR45 was shown to affect the accumulation of ASAP protein components (Chen *et al.*, 2019).

As SR45 presents an atypical modular structure among SR proteins, with a unique RRM flanked by an unstructured RS domain on both sides **(Supplementary Figure S1)**, the function of these domains in the modulation of the SR45 splicing activity, i.e. its localization, ability to interact with proteins and specificity of RNA binding, was examined. We first studied the expression profile of *SR45* during plant development with native promoter-driven reporter constructs to link previously described phenotypes of the loss-of-function mutant to the tissues where SR45 is expressed. By substituting serines from both RS domains to alanines, we established that the SR45 RS domains are required to retain the nuclear localization and play a key role in protein-protein interactions. We further determined the RNA motif recognized by SR45 and demonstrated how three amino acids in the RRM define the RNA binding specificity. Finally, we established how phosphorylatable serines are critical for establishing the SR45 protein-protein interaction network, particularly with kinases, EJC core components and EJC peripheral factors. Our findings provide new insights into the involvement of SR45 in the recruitment of several splicing factors and in EJC assembly.

Materials and methods

Plant material and growth conditions

Arabidopsis plants (Col-0 ecotype) were stably transformed by floral dipping, and independent T3 homozygous lines were analyzed. All observations were performed in at least three independent lines. Seeds were surface-sterilized, then germinated on 1/2 Murashige and Skoog (MS) medium (Duchefa Biochimie) supplemented with sucrose (1% w/v, Duchefa Biochimie) and Select Agar (0.8% w/v, Sigma-Aldrich) after stratification in the dark at 4°C for 48h. For expression profiling, 3-week-old seedlings were grown in hydroponic trays (Araponics) in Hoagland medium for ten weeks (until silique development) as previously reported (Talke *et al.*, 2006; Hanikenne *et al.*, 2008). Hydroponics were grown under a 16-h light (100 µmol m^{-2} sec^{-1})/8-h dark regime in a climate-controlled growth chamber (21°C). The nutrient medium was renewed with fresh solution once a week.

Nicotiana tabacum (cv Petit Havana) transient transformations by *Agrobacterium* infiltration were performed as described (Rausin *et al.*, 2010; Stankovic *et al.*, 2016).

Vector constructions

All binary vector constructions and primers used in this study are listed and detailed in **Supplementary Tables S1-9**. All PCRs were carried out using the Q5 DNA polymerase (NEB) on Arabidopsis cDNA libraries or genomic DNA. All constructs were verified by sequencing. For plant transient and stable transformations, all final plasmids were electroporated into the *Agrobacterium tumefaciens* strain GV3101 (pMP90) and subsequently used for agroinfiltration or floral dipping, respectively. All constructions contain the coding sequence of the *SR45.1* isoform, which we name *SR45* hereafter (Zhang and Mount, 2009).

For yeast two- or three-hybrid assays (Y2H or Y3H), cDNAs coding for the potential interactors were cloned into pGBKT7 and pGADT7 vectors (Clontech) using different restriction sites, with the exception of the *SR45* and mutant variant coding sequences that were cloned into pGBKT7(+) and pGADT7(+) in which *Asc*I and *Pac*I restriction sites were added (see **Supplementary Tables S1 and S2** for primer lists).

The promoter region of *SR45* [1612 bp upstream of the ATG, as described by (Zhang and Mount, 2009)] was amplified by PCR from genomic DNA. The promoter amplicon was ligated at the *Sbf*I/*Kpn*I sites of the pMDC100 vector (Curtis and Grossniklaus, 2003) to create the

pMDC100:pSR45 vector. The *SR45* open reading frame (1242 bp, from the ATG to the last codon before the stop codon) was then cloned at *Asc*I/*Pac*I sites to obtain pMDC100:pSR45:SR45 in which a *nos* terminator (tNOS) was added between *Sac*I restriction sites. Finally, the EGFP coding sequence was cloned at the *Pac*I site to create the pMDC100:pSR45:SR45:EGFP:tNOS vector. To construct the pMDC32:pSR45:EGFP vector, the 2X35S promoter of the pMDC32 vector (Curtis and Grossniklaus, 2003) was replaced by the *SR45* promoter and EGFP was cloned downstream of the promoter at *Kpn*I and *Sac*I restriction sites **(Supplementary Table S3)**.

The *SR45mutRRM* version was obtained by PCR-based site-directed mutagenesis as described (Rausin *et al.*, 2010; Stankovic *et al.*, 2016) **(Supplementary Table S4)**. The other mutated *SR45* versions were assembled by PCR from partially overlapping amplicons, which corresponded to the coding sequence of each domain of SR45. They were amplified using either the native *SR45* gene, the *SR45mutRRM* gene, or the two RS mutated domains (named RA1 and RA2) synthesized by Eurofins Genomics (see **Supplementary Tables S5 and S6**). All assembled mutant versions were then introduced into yeast vectors for Y2H (at *Asc*I/*Bam*HI) or pMDC32 fused to EGFP for protein localization (at *Asc*I/*Pac*I sites), respectively.

The BiFC constructs were obtained as described (Stankovic *et al.*, 2016) (see **Supplementary Tables S7** for primer lists). The TriFC was performed with an additional plasmid, pMDC32 (Curtis and Grossniklaus, 2003), containing no coding sequence or the one of either the RRM domain of ACINUS (765 bp corresponding to the RRM domain and its C-terminal extension, from the 1141st base to the stop codon and preceded by an ATG codon) or PININ.

For SELEX experiments, sequences encoding the RRM of SR45, and the RRM1 or the RRM2 of SR34 were cloned at *Bam*HI/*Eco*RI sites of the pGEX6P-1 as described (De Franco *et al.*, 2019) **(Supplementary Table S8)**. Sequences encoding mutated versions of RRM1 and RRM2 of SR34 were amplified from vectors previously described (Stankovic *et al.*, 2016).

<u>Yeast two- and three-hybrid systems</u>

The vectors and yeast (*Saccharomyces cerevisiae*) strains provided in the Matchmaker Gold Yeast Two-Hybrid System (Clontech) were used to perform directed interaction analyses (dY2H). Prior to (directed) Y2H experiments, toxicity and autoactivation of the wild-type *SR45* and mutant variants were tested simultaneously on a medium -Trp/X-a-gal [for

pGBKT7/pGBKT7(+)] and -Leu/X-a-gal [for pGADT7/pGADT7(+)] using 100 ng of plasmid. Yeast growth was not affected by any of the constructs, none of which led to autoactivation of the reporter gene (*MEL1*) **(Supplementary Figure S2)**. Dimeric interactions were then tested on selective medium -Trp/-Leu/-His/X-a-gal/Aureobasidin A.

For Y2H screening, the Y2HGold strain was transformed with the bait vector (pGBKT7-SR45) and then mated with strain Y187 containing an Arabidopsis cDNA library (Mate and Plate Library-Universal Arabidopsis, Clontech) (Stankovic *et al.*, 2016) or a custom cDNA library made from entire mature leaf material (Make Your Own "Mate & Plate" Library System, Clontech). Trimeric interactions (Y3H) were tested on selective medium -Met/-Trp/-Leu/-His/X-a-gal/Aureobasidin A: yeast expressing SAP18 fused to the binding domain of GAL4 (using SR45 as a bridge) were mated with yeast expressing either ACINUS or PININ fused to the activation domain of GAL4.

Confocal Microscopy

Live cell imaging was performed on Leica TCS SP2 and SP5 inverted confocal laser microscopes (Leica Microsystems). A water-immersion objective was used to collect images at 512x512 pixel resolution. EGFP was visualized using the 488-nm line of the argon laser, and the emission light was dispersed and recorded at 500 to 560 nm (Rausin *et al.*, 2010; Stankovic *et al.*, 2016; Joris *et al.*, 2017). For BiFC and TriFC experiments, YFP was excited at 514 nm using the argon laser and the emission light was recorded at 525 to 600 nm.

Analysis of the mRNA levels

100 mg of plant tissues were used to extract total RNA using NucleoSpin RNA Plant kit (Macherey Nagel) as per manufacturer's instruction. cDNAs were synthesized from 1 µg of total RNAs using the RevertAid H Minus First Strand cDNA Synthesis Kit (Fisher Scientific) with oligo(dT) primers. Quantitative PCR reactions were performed in 384-well plates using a QuantStudio5 system (Applied Biosystems) and Takyon Low Rox SYBR MasterMix dTTP Blue (Eurogentec). Samples were obtained from three independent biological experiments. A total of three technical replicates were performed for each combination of cDNA and primer pair **(Supplementary Table S9)**. Gene expression was normalized relative to At1g58050 as described (Pfaffl, 2001). *At1g58050* expression was the most stable gene among all tested references (*EF1α*

and *UBQ10*) (Czechowski *et al.*, 2005; Spielmann *et al.*, 2020).

Production of recombinant proteins

The recombinant protein production was performed using the expression vector pGEX6P-1 in *E. coli* BL21(DE3) cells as described (De Franco *et al.*, 2019). Protein productions (native and mutated RRM of SR45, or RRM1 or RRM2 of SR34) were carried out overnight at 18°C upon induction with 0.5 mM IPTG. Lysis of the cells was carried out using a sonicator in 50 mM Tris-HCl pH 8, 1 M NaCl, 2 mM $MgCl_2$, 1 mM dithiothreitol (DTT) and 1 mM PMSF. GST fusion proteins were purified on Glutathione-Sepharose® 4B beads (GE Healthcare) and eluted using a solution of 50 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM dithiothreitol (DTT) and 35 mM reduced glutathione.

SELEX experiments

The Systematic Evolution of Ligands by EXponential enrichment (SELEX) experiment was performed using a randomized 25-nt DNA template library as previously reported (De Franco *et al.*, 2019). 500 ng-1 µg of the purified template DNA library was used for the *in vitro* transcription reaction (T7 RiboMAX™ Express Large Scale RNA Production System, Promega). Unincorporated nucleotides were removed using mini Quick Spin RNA Columns (Roche). RNAs were purified with phenol/chloroform and ethanol-precipitated. The RNA pellet was resuspended in RNAse-free water, and the final RNA concentration was quantified by measuring the absorbance at 260/280 nm. Binding reactions were carried out in 200 µL of SELEX buffer (20 mM MOPS pH 7.0, 50 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 0.1% Triton-X-100, and 0.1 mM PMSF) and included 80 pmol of GST-fused protein immobilized on GSH beads (GE Healthcare), 1–8 µg of heparin sulfate, and 0.8–4 nmol of RNAs. Samples were gently agitated at 4 °C for 60 min. Unbound RNA was removed by five consecutive bead washings in 500 µL of SELEX buffer. HCl-25 mM glycine (pH 2.2) was used to dissociate RNA-protein complexes (25 °C, 5 min). Dissociated RNAs were purified using mini Quick Spin RNA Columns (Roche), then ethanol-precipitated, and finally reverse-transcribed (SuperScript® III Reverse Transcriptase, Invitrogen). The resulting cDNAs were PCR amplified by 10 cycles for SELEX rounds 1 to 3, or 15 cycles for last SELEX round. Amplification products were then purified (PCR clean up kit, Thermo Fisher Scientific**)**. The amplified cDNAs were used as a

template to synthesize a new RNA library for the following SELEX rounds. In every SELEX round, selection stringency was intensified by increasing the RNA/protein ratios and the heparin concentration. The final PCR products (round 4) were subcloned into the pJET vector (CloneJET PCR Cloning Kit, Thermo Fisher Scientific) and sequenced. The obtained sequences were analyzed using the MEME tool (version 5.4.1) (Bailey *et al.*, 2009), then logos were redesigned using the WebLogo software (Crooks *et al.*, 2004).

Structural modelling of multiprotein complexes involving SR45

Modelling of the 3D structure of different multiprotein complexes was performed with the AlphaFold software version 2.3.0 (Jumper *et al.*, 2021). The multimer option was selected with a database version from January 1st, 2023, and a total of 25 models generated per prediction. All models were visually checked and the one with the best iptm+ptm (interface predicted template modelling + predicted template modelling) value was energy minimized within AlphaFold and selected for further analysis. Figures were generated with PyMOL v2.5.2 (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.).

Statistical analysis

All data evaluation and statistics were performed using GraphPad Prism 7 (GraphPad Software v7.00). The replication level of each experiment is described in the Figure legends.

Accession numbers and information

All gene sequences are available through *The Arabidopsis Information Resource* (TAIR, http://www.arabidopsis.org/), with the accession numbers: Arabidopsis *ACINUS* (AT4G39680), *AFC1* (AT3G53570), *AFC2* (AT4G24740), *AFC3* (AT4G32660), *ALY4* (AT5G37720), *Cyp59* (AT1G53720), *Cyp65* (AT5G67530), *Cyp71* (AT3G44600), *CypRS64* (AT3G63400), *CypRS92* (AT4G32420), *eIF4A3* (AT3G19760), *GRP7* (AT2G21660), *GRP8* (AT4G39260), *MAGO* (AT1G02140), *MOS12* (AT2G26430), *MOS14* (AT5G62600), *PININ* (AT1G15200), *PRP38* (AT2G40650), *RS2Z232* (AT3G53500), *RS2Z33* (AT2G37340), *RS31* (AT3G61860), *RS31a* (AT2G46610), *RS40* (AT4G25500), *RS41* (AT5G52040), *RSZ21* (AT1G23860), *RSZ22* (AT4G31580), *RSZ22a* (AT2G24590), *RZ-1B* (AT1G60650), *RZ-1C* (AT5G04280), *SAP18* (AT2G45640), *SC35* (AT5G64200), *SCL28* (AT5G18810), *SCL30* (AT3G55460), *SCL30a*

(AT3G13570), *SCL33* (AT1G55310), *SR30* (AT1G09140), *SR34* (AT1G02840), *SR34a* (AT3G49430), *SR34b* (AT4G02430), *SR45* (AT1G16610), *SRPK1* (AT4G35500), *SRPK2* (AT2G17530), *SRPK3a* (AT5G22840), *SRPK3b/SRPK5* (AT3G44850), *SRPK4* (AT3G53030), *Y14* (AT1G51510).

Results

SR45 is expressed at all developmental stages

During plant development, *SR45* was expressed to different extents in all vegetative and floral organs examined as described by quantitative RT-PCR profiling **(Figure 1A)**. *SR45* was significantly more expressed in inflorescences and young seedlings. The expression level was lower and similar in whole roots, stems and (cauline) leaves. We also determined the distribution of the protein using a GFP translational fusion (*pSR45:SR45-EGFP*) **(Figure 1B-R)** and the spatial expression pattern of *SR45* using a promoter:reporter construct (*pSR45:EGFP*) **(Supplementary Figure S3)**. Translational fusion lines confirmed (i) the spatial transcriptional activity of the *SR45* promoter and (ii) the nuclear localization of SR45 (Ali *et al.*, 2003). The seed coat (including hilum) presented GFP fluorescence **(Figure 1B)**. Embryos freshly dissected from seeds displayed fluorescence in cotyledons **(Figure 1C)**. Right after germination, seedlings showed an intense fluorescence in all cell types within the radicle **(Figure 1D-E)**, the hypocotyl and the cotyledons **(Figure 1E)**. During vegetative growth, fluorescence was detected in roots, root hairs **(Figure 1E-G)**, and leaves (including trichomes and stomata) **(Figure 1I-J)** and was particularly visible within the vascular tissues of roots **(Figure 1F)**. During flowering, the receptacle at the flower bases, together with the abscission zone and the gynophore, emitted fluorescence signals **(Figure 1K)**. In flowers, GFP fluorescence resolved to styles and stigmas, stamens (anthers and filaments but not pollen grains), petals (including veins), and sepals **(Figure 1L-R)**. The observations in *pSR45:EGFP* lines were in agreement with those of the *pSR45:SR45-EGFP* translational fusion lines. However, fluorescence of GFP alone, which is not targeted to any specific subcellular compartment, was very weak, hence being sometimes hardly detected above the background **(Supplementary Figure S3)**. Indeed, seeds, embryos and radicles presented a low fluorescence signal in *pSR45:SR45-EGFP* lines that was fainter in *pSR45:EGFP* lines.

Structural determinants modulating SR45 localization

To understand the structural determinants modulating the SR45 cellular localization **(Figure 1)**, we engineered seven mutant variants **(Supplementary Figure S1)**. First, the canonical RRM of SR45, which is defined by a β_1-α_1-β_2-β_3-α_2-β_4 topology, was mutagenized at conserved residues mediating RNA recognition in both RNP1 and RNP2 motifs (Cléry *et al.*,

2008). Both H101 (RNP2) and H141 (RNP1), together with Y143, were substituted to alanine to create the SR45mutRRM variant. Second, all Ser residues were substituted to Ala within the RS1 (nt 1 to 291, 39 S-to-A substitutions) and/or RS2 (nt 529 to 1245, 49 S-to-A substitutions), generating three variants (mutRS1, mutRS2, mutRS1+RS2), respectively. Finally, RRM and RS1/RS2 mutations were combined, producing three more mutant variants (mutRS1+RRM, mutRRM+RS2 and mutRS1+RRM+RS2).

After transient expression in tobacco leaf cells, no distinguishable changes appeared in the protein localization of the SR45mutRS1 or SR45mutRRM variants **(Figure 2)** compared to the native protein. In contrast, SR45mutRS2 partially accumulated in the cytoplasm in addition to a nuclear localization. These observations are in agreement with previous work examining the localization of RS1- or RS2-truncated SR45 mutants (Ali and Reddy, 2006; Ali *et al.*, 2008). SR45mutRS1+RS2 showed nucleolar retention, whereas combining RS2 and RRM mutations (SR45mutRRM+RS2) noticeably increased the cytoplasmic retention compared to the native protein and the SR45mutRS2 variant. With all substitutions combined (SR45mutRS1+RRM+RS2), no fluorescence was detected within the nucleus, and only cytoplasmic signal was observed. Altogether, the integrity of the RS2 domain appears necessary but not sufficient for the nucleoplasmic localization of SR45.

SR45 interactome

To identify interactors of SR45, two Arabidopsis cDNA libraries (a commercially available and a custom-made) were first screened using yeast two-hybrid (Y2H) and four previously unidentified interactors of SR45 were uncovered: CypRS64 (Cyclophilin-containing RS domain of 64 kDa), PRP38 (Pre-mRNA Processing factor 38), RS2Z32 (arginine/serine-rich zinc knuckle-containing protein 32), and MOS12 (*MODIFIER OF SNC1-1*, 12). These interactions were then further confirmed by directed yeast two-hybrid assays (dY2H) **(Figure 3**, **Supplementary Figure S4 and Supplemental Table S2)**.

dY2H assays were performed to examine the SR45 interactions with kinases potentially involved in RS domain phosphorylation, the 3 CLK kinases (AFC1-3) (Bender and Fink, 1994) and the 5 SRPK kinases encoded in the Arabidopsis genome (Lorković *et al.*, 2004; de la Fuente van Bentem *et al.*, 2006; Stankovic *et al.*, 2016). We confirmed that SR45 interacts with AFC2 (Golovkin and Reddy, 1999), and confirmed its recently reported interactions with SRPK3a,

SRPK3b [also named SRPK5 (Wang *et al.*, 2023)] and SRPK4. Among all SR45 mutant variants, only RS1 mutations abolished the SRPK4- and AFC2-SR45 interactions, while RS1 or RS2 mutations both suppressed interactions with SRPK3a and SRPK3b/SRPK5 **(Figure 3)**. Those results suggested that RS domains are involved in interactions with kinases.

SR45 interacts with several other SR proteins, namely SCL33, RSZ21, SR30, SR34, SR34a, as well as other splicing factors such as the hnRNP-like proteins RZ-1B and RZ-1C (Golovkin and Reddy, 1999; Zhang *et al.*, 2014; Stankovic *et al.*, 2016; Wu *et al.*, 2016). We showed here that SR45 was able to interact with all SR proteins but itself (SR45.1 or SR45.2 splicing isoforms (Zhang and Mount, 2009)) **(Supplementary Figure S6)**. The integrity of both RS domains was crucial for SR45 to contact SC35. In the cases of SR34, RS31a and SCL30, the interactions were only weakened by mutations applied to either RS1 or RS2 but were completely abolished for RS1/RS2 variants **(Figure 3)**. Instead, the RS1 variants showed suppressed interactions with RS2Z32, RSZ21, RZ-1B, RZ-1C and MOS12. In contrast to interactions of SR45 with SR proteins, the interactions with RZ-1B, RZ-1C or PRP38 were considerably weakened upon substitutions in the RRM domain **(Figure 3)**. The SR45-PRP38 interaction was, however, not affected by other substitutions **(Figure 3)**. SR45 contacted several nuclear cyclophilins including Cyp59, CypRS64 and CypRS92, but not Cyp65 and Cyp71, which are all involved in splicing regulation (Lorković *et al.*, 2004; Gullerova *et al.*, 2006; Barbosa Dos Santos and Park, 2019). However, none of the mutant variants was able to maintain interactions with any of the cyclophilins tested in yeast. No interaction was observed between SR45 and hnRNP-like proteins GRP7 and GRP8 **(Figure 3)**.

In animals, several shuttling SR proteins (SRSF1, SRSF3 and SRSF7) act as adaptors for mRNA nuclear export through the TAP/NXF1 (Nuclear RNA export factor 1) machinery and compete with ALY/REF for binding in the NXF1 complex (Huang *et al.*, 2003). A gene encoding a homolog of the human TAP/NXF1 is missing in plant genomes (Merkle, 2011), but numerous ALY/REF orthologs were discovered in Arabidopsis (Pfaff *et al.*, 2018; Ehrnsberger *et al.*, 2019). Our dY2H analysis highlighted an interaction between SR45 and ALY4, which was tremendously weakened by RS1 and/or RRM mutations **(Figure 3)**. SR45 presents a nucleocytoplasmic dynamic whose nuclear export is at least partly controlled by the XPO1-dependent export pathway (Stankovic *et al.*, 2016), whereas the nuclear import of plant SR proteins is likely driven

by the activity of MOS14 (*MODIFIER OF SNC1-1*, 14), a transportin-SR (TRN-SR) (Xu *et al.*, 2011). Here, SR45 and MOS14 interacted, and the SR45 RS2 domain was required **(Figure 3)**.

The human RNPS1 contributes, together with SAP18, ACINUS and PININ, to the formation of two EJC peripheral complexes, ASAP and PSAP (Sakashita *et al.*, 2004; Singh *et al.*, 2010; Murachelli *et al.*, 2012). We next performed dY2H and directed yeast three-hybrid (dY3H) assays to investigate interactions between SR45 with core (Y14, MAGO and eIF4AIII) and peripheral (SAP18 and either ACINUS or PININ, which are part of the ASAP and PSAP complexes, respectively) components of the EJC **(Figure 4 and Supplementary Figure S5)**. We demonstrated that, unlike eIF4AIII, MAGO and Y14 were able to interact with SR45, which required an intact RS1 domain **(Figure 4A)**. We further confirmed that SR45 was able to interact with both ACINUS and PININ (Qüesta *et al.*, 2016; Chen *et al.*, 2019; Bi *et al.*, 2021) but not SAP18 (Wang *et al.*, 2023). SAP18 was not able to interact with either ACINUS or PININ **(Figure 4A)**, however, when the three components were co-expressed in a Y3H system, an interaction was established between SR45 and SAP18 through either ACINUS or PININ **(Figure 4B)**. The SR45-ACINUS interaction was greatly diminished upon mutations in the RS1 domain **(Figure 4A)**. Based on animal models where two conserved residues, a threonine and a tyrosine, are critical for the assembly of the ASAP complex (Sakashita *et al.*, 2004; Singh *et al.*, 2010; Murachelli *et al.*, 2012), we identified two amino acids (Thr607 and Tyr615) within the C-terminal extension of the RRM domain of ACINUS that were expected to be critical for the interaction with SR45 **(Supplementary Figure S7D)**. Accordingly, mutations of these two residues (T607A/Y615A) abolished the interaction of ACINUS with SR45. On the other hand, the SR45-PININ interaction was suppressed when mutations were present in the RRM of SR45 **(Figure 4A)**.

Altogether, our observations confirmed the importance of serine residues in the establishment of interactions with kinases and splicing factors but also depicted a role for the canonical RRM (through both RNP1 and RNP2 motifs) in the formation of a protein complex involved in mRNA export and mRNA surveillance.

Structural modelling of multiprotein complexes involving SR45

Structural modelling was conducted to further explore the details of the interactions taking place in two multiprotein complexes involving SR45: ACINUS-SR45-SAP18 (ASAP) and PININ-SR45-SAP18 (PSAP) **(Supplementary Data S1 and S2)**.

For ASAP, clear interactions were observed between structured regions of the three partners, which all contained extended unstructured regions with predicted local distance difference test (pLDDT) values below 50 (ACINUS residues 1-3, 66-444 and 542-597; SR45 residues 1-95 and 177-414; SAP18 residues 1-28). For SR45, this confirmed the unstructured nature of the RS1 (1-97) and RS2 (177-414) domains, surrounding a structured RRM domain (98-176) **(Supplementary Figure S1)**. Outside of these unstructured regions, AlphaFold's estimation of the accuracy was high with an average pLDDT of 89.5 **(Supplementary Figure S7A-B)**. The ACINUS N-terminal SAP domain (residues 4-66) was well folded but did not interact with any other part of the complex, and will not be further discussed. The interactions between the RRM domain of SR45, the ubiquitin-like domain of SAP18 and the last 35 residues of ACINUS were very similar to that observed in the ASAP complex available in the Protein Data Bank (PDB code 4A8X) (**Supplementary Figure S7B-C**) with a Root-Mean-Square Deviation (RMSD) of 1.0Å calculated over 201 Cα of the 218 residues spanning the three chains of the complex (Murachelli *et al.*, 2012). The sequence identity between residues present in this structure (human RNPS1, ACINUS peptide from *Drosophila melanogaster* and *Mus musculus* SAP18) and equivalent residues in the Arabidopsis model was 49%. The model of the core ASAP structure confirmed that (i) ACINUS interacted with SR45 through two residues (T607 and Y615) located in the vicinity of its RRM, and that (ii) the three substituted residues in the RRM of SR45 were not involved in this interaction, corroborating Y2H assays **(Figure 4A)**. In addition to this known part of the ASAP complex in animals, the AlphaFold model revealed a potential additional interaction between the RRM domain of ACINUS (residues 445 to 541) and the ubiquitin-like domain of SAP18. The interface, which covers an area of 588Å^2, involved residues from the L1, L3 and L5 loops **(Supplementary Figure 7D)**, the first two being the most conserved regions on the surface of the domain according to Fernandes and colleagues (Fernandes *et al.*, 2018). With only two salt bridges and some hydrophobic interactions, the interaction between the two partners was not strong according to the PISA server (Krissinel and Henrick, 2007) and could potentially be strengthened upon interactions with other partners including RNA.

In the PSAP complex, which is characterized by an average pLDDT value calculated for structured region (pLDDT > 50) of 86.5, the same interactions are present between the ubiquitin-like domain of SAP18 and the RRM domain of SR45, and as predicted by Murachelli and colleagues (Murachelli *et al.*, 2012), the ^{310}KAEPRIYYAPVKPL^{323} peptide of PININ mostly substituted the C-terminal domain of ACINUS at the junction between SAP18 and SR45 in the ASAP complex **(Supplementary Figure 8A-C)**. In addition to this peptide, the structured part of this model of PININ was made of four α helices (α1-α4) with pLDDT value over 50. Three of these helices were approximately parallel: α1 (residues 8-32), α3 (residues 188-258) and α4 (residues 327-393). The α2 helix (residues 160-186) directly preceded α3 and was perpendicular to the three other helices, but the confidence of the model was relatively weak in this region. The α3 and α4 helices, which were separated by a loop containing the ACINUS-like peptide and a 50 amino acids long unstructured stretch, also interacted with the ubiquitin-like domain of SAP18 **(Supplementary Figure 8D)** and contributed to the interaction with three H-bonds, two salt bridges and few hydrophobic interactions, covering a 764Å^2 surface area. The four α helices might provide an extended platform for the recruitment of multiple additional partners.

The RS1 and RS2 domains of SR45 connect splicing factors and EJC peripheral factors in planta

Observed in yeast and in the ASAP structural model described above, the interactions underlying the involvement of SR45 into the splicing reaction (ACINUS, PININ and SR34) as well as the spliceosome catalytic activation (PRP38) (Schütze *et al.*, 2016), which are also established by RNPS1 in human (Schütze *et al.*, 2016; Schlautmann *et al.*, 2022), were further characterized in living plant cells using Yellow Fluorescent Protein (YFP)-based bimolecular fluorescence complementation (BiFC). BiFC experiments were performed in tobacco leaf cells where the native or mutant variants of SR45 fused to the C-terminal fragment of YFP (^{C}YFP) and interactors fused to the N-terminal fragment of YFP (^{N}YFP) were transiently co-expressed. YFP fluorescence was observed in the nucleus and in structures similar to nuclear speckles, suggesting that SR45 interactions with ACINUS, SR34 and PRP38 occurred in the nucleus and in splicing factors-enriched granules *in planta* **(Figure 5A)**. The study of mutant variants confirmed results obtained in yeast cells **(Figure 3)**, as SR45mutRS1+RS2 was unable to interact with ACINUS and SR34, but remained able to establish an interaction with PRP38 within the nucleolus.

Furthermore, mutations in the RS1 or RRM domains did not affect any of the SR45 associations mentioned above, similarly to what was observed in yeast cells and in the modelling performed for ACINUS **(Supplementary Figure S7)**. Trimolecular fluorescence complementation (TriFC) also confirmed dY3H assays as SAP18 was unable to interact with either SR45, ACINUS or PININ **(Figure 5B, upper panels)** but was able to form trimeric ASAP and PSAP complexes **(Figure 5B, lower panels)**.

Subcellular localization of SR45-associated partners

To better decipher the dynamic phosphorylation of SR45 across cellular compartments and where SR45 might precisely contact its (peripheral) EJC interactors in plant cells, EGFP was fused to the N-terminus of candidate proteins, and translational fusions were transiently expressed in tobacco leaf cells. SRPK1 and SRPK2, two non-interactors of SR45, resided exclusively in the cytoplasm **(Figure 6)**. Unlike SRPK4 which resided predominantly in the cytoplasm (alongside a weak detectable signal within the nucleus), SRPK3a and SRPK3b/SRPK5 localized in the cytoplasm and the nucleus **(Figure 6)**. The CLKs were the only kinases studied that accumulated in nuclear speckles-like structures **(Figure 6)**. ACINUS, SAP18 and PININ all localized in the nucleus, however, SAP18 also displayed a cytoplasmic localization **(Figure 6)**. Similarly, MAGO and Y14 were also detected in the nucleus and in the cytoplasm **(Figure 6)**. PRP38 localized in the nucleoplasm, speckles-like structures and the nucleolus **(Figure 6)**.

These observations suggest that the interactions described by dY2H/dY3H and BiFC/TriFC may actually take place in plant cells, as the intracellular localization of SR45 and its kinase and EJC interactors at least partially overlapped.

SR45 is binding to purine-rich RNA motifs

Then, we investigated the RNA-binding specificity of a recombinant RRM domain of SR45 fused to a glutathione-S-transferase using Systematic Evolution of Ligands by EXponential enrichment (SELEX). Among the twelve clones sequenced after SELEX, two statistically significant consensus sequences recognized by the SR45 RMM were identified, with a specificity for a purine-rich central motif (5'-[A/G]GAAG-3') flanked by pyrimidine-rich regions (5' UC[C/U] and 3' UU[C/U][C/U]C) (*E*-value: 8.6 x 10^{-4}) **(Figure 7A)** or for a purine-rich motif ([A/G]GA[C/A]A) (*E*-value: 1.2 x 10^{-10}) **(Figure 7B)**. Because of the higher variability of bases

at most positions in the 2nd motif, we analyzed the 4-mer and 5-mer frequencies among the sequences used to identify it. The analysis of 4-mer frequency showed a higher occurrence of the GACA and GAAG motifs. The analysis of 5-mer frequency demonstrated higher prevalence of AAGUU, AGACA, GAAGU, GACAU and AGAAG motifs **(Supplementary Table S10)**. This is strikingly similar to G/A-rich motifs bound by RNPS1 (Mabin *et al.*, 2018).

To ensure the specificity of the binding of the identified RNA motifs to the SR45 RRM, a mutated version (H101A, H141A and Y143A) was produced and studied in the SELEX set-up. Four different sets of twelve sequences obtained using the mutated SR45 RRM were analyzed **(Supplementary Table S11)** using the MEME tool. No enriched motif was discovered **(Figure 7C and Supplementary Figure S9)**, and the best motif recovered (lowest *E*-value, #4) did not resemble significant motifs found for the native canonical RRM **(Figures 7A-B)**, indicating that amino acids (H101, H141 and Y143) mutated within the RRM are necessary and sufficient to ensure the accurate binding of SR45 to its purine-enriched target motif.

Discussion

Pre-mRNA splicing is a crucial step in the regulation of gene expression. Splice-site selection is carried out by the binding of many *trans*-acting protein factors including the SR splicing factors (Chen and Manley, 2009). In mammalian cells, it has been shown that SR proteins regulate many aspects of gene expression (Huang and Steitz, 2001). In contrast, their plant counterparts are mainly known to play important roles in constitutive splicing and alternative splicing, in biological processes such as ABA signalling, salt response, metal homeostasis, or response to high-light (Xing *et al.*, 2015; Yan *et al.*, 2017; Zhang *et al.*, 2017; Fanara *et al.*, 2022; Albaqami, 2023) although they are also suspected to be involved in other processes regulating gene expression, e.g. RNA-directed DNA methylation and regulation of transcription (Ausin *et al.*, 2012; Qüesta *et al.*, 2016; Yan *et al.*, 2017; Bi *et al.*, 2021; Mikulski *et al.*, 2022).

The RS domains of SR45 control its subcellular localization and protein-protein interaction network

The SR45 splicing factor displays a modular structure, with two unstructured RS and one structured RRM domains, which is atypical for a plant SR protein (Barta *et al.*, 2010; Califice *et al.*, 2012). In this report, we examined the involvement of each of these domains, as well as some of their constituent amino acid residues, in molecular functions known to be regulated by the human ortholog of SR45 (RNPS1) (Califice *et al.*, 2012), whose association to EJC peripheral factors determines its roles in pre-mRNA splicing, mRNA export and quality control (Le Hir *et al.*, 2016) **(Figure 8)**. To this avail, all serine residues were substituted into alanines in the RS domains and three point mutations were introduced in conserved residues of the RRM domain. One might question the strategy of modifying multiple amino acids and its impact on the global structure of SR45, hence on the functionality of the engineered mutant variants. We studied here the impact of these substitutions on (i) the protein localization, demonstrating that all mutant variants displayed a complete or partial nuclear localization similarly to the native protein (with the exception of the SR45mutRS1+RRM+RS2 that was not further studied); (ii) the protein-protein interactions, for which we demonstrated that the mutations had different impacts on the interactions with different partners in yeast cells and *in planta*, dismissing the possibility of a complete non-functionality of the mutant variants, and (iii) the protein-RNA interactions,

demonstrating that the mutated RRM was still able to engage in interactions with RNA molecules despite an alteration of specificity, as it is well documented (Yang *et al.*, 2011; Pabis *et al.*, 2019; Ni *et al.*, 2023).

With this experimental set-up, we show that a number of interactions between SR45 and spliceosomal or EJC components entirely rely on the presence of serine residues in the RS domains of SR45. Indeed, the substitution of only few residues in these domains can either tremendously diminish or abolish the connection between SR45 and its partners, for instance the interactions with SR34 and other SR proteins. Some of those serine residues also control the subcellular localization of SR45. S-to-A substitutions within the RS2 domain alone induce cytoplasmic accumulation of SR45, which might result from its inability to contact the plant transportin-SR (MOS14) that is thought to be involved in the nuclear import of plant SR proteins (Xu *et al.*, 2011) as its human ortholog TNPO3 (Maertens *et al.*, 2014; George *et al.*, 2019; Costa *et al.*, 2021). In contrast, S-to-A substitutions in both RS domains strictly restrict SR45 to the nucleolus where it still can interact with PRP38. The subcellular localizations of RS mutant variants reported here refine the observations made by Ali and colleagues (2006 and 2008), revealing that the accumulation of SR45 within the nucleus or in both the nucleus and cytoplasm upon, respectively, disruption of the RS1 (SR45mutRS1 in this study) or RS2 domains (SR45mutRS2 in this study) are dependent of serine residues (Ali and Reddy, 2006; Ali *et al.*, 2008). The nuclear import of SR proteins further depends on the phosphorylation of serine residues present in the RS domain (Lai *et al.*, 2000; Lai *et al.*, 2001). This process involves several kinases in animals, mainly SRPK1 (ortholog of the Arabidopsis SRPK4) and CLK1 (ortholog of the Arabidopsis AFC2), which act together to regulate the phosphorylation status of the human SRSF1 SR protein (Aubol *et al.*, 2016). We show here that in Arabidopsis, the recognition of SR45 by two kinases of the SRPK family (SRPK3a and SRPK3b/SRPK5) depends on the integrity of the RS2 domain. These kinases might therefore be predominately involved (over other kinases) in the nuclear import of SR45 in conjunction with the transportin-SR MOS14. While S-to-A substitutions within the RS2 does not affect the association between SR45 and SRPK4 or AFC2, all mutant variants harbouring RS1 substitutions were not able to contact any of the kinases identified in this study. It is thus tempting to speculate that the RS1 domain of SR45, which is essential for AFC2 and SRPK4 binding, might serve as a platform for coordinating activity of these two kinases as seen in SRSF1 (Aubol *et al.*, 2017; Wegener and

Müller-McNicoll, 2019). A recent study identified specific serine residues within both RS domains whose phosphorylation levels were downregulated in a *srpk3 srpk4 srpk3b/srpk5* triple mutant, and tested interactions between these three kinases and SR45 mutant variants harbouring phosphoresidue substitutions to alanine (Wang *et al.*, 2023). Here we focused on substituting all serines within a domain to better decipher the importance of the global phosphorylation status to (i) the subcellular localization of SR45 and (ii) to the protein-protein interaction network established by this splicing factor. Both approaches uncovered similar results, as SRPK3, SRPK4 and SRPK3b/SRPK5 interactions with SR45 involve native serine residues within both RS domains, i.e. S-to-A substitutions of all serines (this study) or of only phosphorylated serines (Wang *et al.*, 2023) lead to the disruption of these interactions. The triple mutation *srpkii* notably reduces the phosphorylation level of SR45 at the T264 residue (Wang *et al.*, 2023). This specific residue was in addition shown to be involved in the ABA-controlled phosphodegradation of SR45 (Albuquerque-Martins *et al.*, 2023), suggesting a connection between the SRPK proteins, the stability of their SR45 substrate and ABA signalling.

The SR45 RRM is essential to contact protein partners

In the RRM of SR45, a histidine residue is present at the position of the putative aromatic residue contacting RNA in both RNP2 (2nd position) and RNP1 (3rd position) motifs (Maris *et al.*, 2005; Cléry *et al.*, 2008; Daubner *et al.*, 2013) **(Supplementary Figures S1 and S7D)**. Because histidine residues are known to bind RNA molecules (Morozova *et al.*, 2006; Goroncy *et al.*, 2010), both H101 (RNP2) and H141 (RNP1) together with Y143 were substituted by alanine to engineer the SR45mutRRM variant. The data presented here suggest that the H101, H141 and Y143 residues in the SR45 RRM are not structural determinants modulating SR45 localization but they are partially involved in SR45 association to RZ-1B, RZ-1C, ALY4 and PRP38. According to the structural models obtained in this study **(Supplementary Figures S7 and S8, and Supplementary Data S1 and S2)**, however, these residues are not involved in interactions with the two other partners in the ASAP and PSAP complexes. While the suppression of contact between the mutated RRM domain of SR45 and PININ is not supported by the model obtained with AlphaFold, the structural models are in agreement with a previous report establishing the involvement of the RRM of RNPS1 in the interaction with ACINUS and PININ (Murachelli *et al.*, 2012). Although RNPS1 forms the ASAP complex with ACINUS and SAP18 through its

RRM (Murachelli *et al.*, 2012), SR45 contacts ACINUS using preferentially the RS1 domain. This interaction between unstructured regions, which is not detected in the model, is completely abolished when S-to-A substitutions are applied to both RS domains and is in agreement with recent work (Wang *et al.*, 2023). The model of the core ASAP structure confirms that ACINUS contacts SR45 through two residues (T607 and Y615) located in the C-terminal extension of its RRM, supporting our Y2H data **(Figure 4A)**. Accordingly, the substitutions T607A and Y615A abolished the interaction between both partners in yeast. Our observations indicate that in addition to the core ASAP structure involving the SR45 RRM, the SAP18 ubiquitin-like domain and the ACINUS C-terminal peptide known in animals (i) the ASAP complex involves the RRM of ACINUS according to the model and at least the RS1 domain of SR45 according to Y2H observations, and (ii) the ASAP complex involves SAP18 as a ternary partner of a pre-existing SR45-ACINUS platform as shown by Y3H and TriFC assays.

RNA binding specificity of the SR45 RRM

So far, the characterization of RNA-binding specificity of plant SR proteins mainly involved *in silico* analyses performed on RNA-Seq data generated from loss-of-function mutants. Such an approach suggested that SCL30, one of the interactors of SR45, was able to bind to a specific RNA sequence, 5'-AGAAGA-3' (Yan *et al.*, 2017). Four overrepresented RNA motifs were also previously identified by the identification of direct RNA targets of SR45 (SR45-associated RNAs, SARs) in seedlings (Xing *et al.*, 2015): the identified M1 and M4 motifs are G/A-rich and were predicted to be predominately positioned within exons (Xing *et al.*, 2015); the M2 and M3 motifs were U/C-rich and were preferentially located in intronic regions (Xing *et al.*, 2015). While the "Motif 1" identified here through SELEX cycles combined features of all the motifs previously identified for SR45 (Xing *et al.*, 2015), the "Motif 2" contained invariant G/A nucleotides and tend to be enriched in AGA, GGA and UCC trinucleotides **(Supplementary Figure S10)**. The human RNPS1 was found to be associated to G/A-rich 6-mers RNA regions (Mabin *et al.*, 2018), which is reminiscent of the invariant sequence 5'-[A/G]GAAG-3' found here in "Motif 1" associated to SR45. The RNA-binding specificity of the splicing factor and protein partner of SR45, RZ-1C, was also determined by SELEX and involved G/A-rich motifs (Wu *et al.*, 2016). While the U/C-rich motif associated to SR45 is similar to the binding motifs of human NXF1 and SRSF3 in the 3'UTR (3' untranslated region) of their target RNAs (Müller-

McNicoll *et al.*, 2016), the ”Motif 1” is also strikingly similar to the binding motif of human NXF1 and of SRSF2, UPF3B and MLN51 (Hauer *et al.*, 2016; Wegener and Müller-McNicoll, 2019). Since plant genomes lack the *TAP/NXF1* gene (Serpeloni *et al.*, 2011), it is tempting to question the role of SR45 as an adaptor for the mRNA export machinery as previously reported for human SR proteins (Müller-McNicoll *et al.*, 2016). It is noteworthy that the SR45-associated ACINUS is known to bind both G/A-rich and U/C-rich motifs in humans (Rodor *et al.*, 2016). Different RNA motifs were uncovered for SR34 in this study **(Supplementary Figure S11 and Supplementary Tables S12-15)**, validating our SELEX approach capable of identify motifs specific to the studied RRM domain. Previous reports showed that purine-rich motifs, and specifically GAAG repeats, are typically related to positive splicing regulators (i.e. SR proteins) and are considered as exonic splicing enhancers (ESEs) in animals and plants (Fairbrother *et al.*, 2002; Pertea *et al.*, 2007; Thomas *et al.*, 2012; Wu *et al.*, 2014). Overall, this suggests that SR45 acts, to certain extent, as a positive regulator of alternative splicing (Xing *et al.*, 2015; Zhang *et al.*, 2017).

Dual roles of the SR45 RRM

Our results demonstrated that the RRM of SR45 is involved in the assembly of the ASAP and PSAP complexes, and is able to efficiently bind RNA sequences. This dual role implies that its RNA binding ability and specificity might compete with the RRM contribution to protein complex formation, as previously described in the case of SRSF1 in animals (Aubol *et al.*, 2016; Wegener and Müller-McNicoll, 2019; Aubol *et al.*, 2021). The EJC components, Y14 and MAGO, mediate protein-protein interactions through the β-sheet surface of the RRM, which restricts the RNA binding ability of Y14 (Lau *et al.*, 2003; Shi and Xu, 2003). Although SR45 associates with EJC components and EJC peripheral factors, whether it is or not intrinsically involved in EJC-mediated molecular regulation and in underlying developmental processes still remains undetermined. It is worth noting that *acinus* and *pinin* loss-of-function mutants display phenotypes similar to a *sr45-1* mutant (i.e. dwarfism, delayed flowering, and ABA hypersensitivity) (Bi *et al.*, 2021).

Conclusion

Here, we characterized the contribution of the RS and RRM domains, and their

constitutive residues, to the function of the Arabidopsis SR45 protein considering the many aspects defining a splicing factor: (i) the subcellular localization, (ii) the ability to interact with protein partners, including in multiprotein complexes, as confirmed by structural modelling, and (iii) the specificity of the binding to RNA. Our results demonstrate a dual role for the RRM of SR45 in both protein-protein and protein-RNA interactions, suggesting a possible competition between both functions. Structural analysis will be useful to substantiate these findings, notably in the presence of both a partner requiring the SR45 RRM to interact with and an RNA specifically bound by this domain. Furthermore, it would be of interest to determine whether the enrichment of the RNA motifs identified here also occurs within the pre-mRNA sequences of direct RNA targets of SR45 and whether it is correlated to molecular changes in the expression and in the splicing profiles of these targets. The latter would require the assessment of genome-wide alternative splicing events occurring in the loss-of-function mutant *sr45-1* at developmental stages displaying a higher abundance of SR45, such as in seedlings or in inflorescences as shown here by our expression profiling. The mutant variants engineered in this study could serve in complementation assays of the *sr45-1* mutant with both SR45 isoforms using native promoter-driven translational fusions. Moreover, it was shown that acetylation, methylation, SUMOylation, ubiquitylation, or even prolyl *cis-trans* isomerization can regulate the activity of animal SR proteins and other splicing factors (Slišković *et al.*, 2022; Kretova *et al.*, 2023). Therefore, dissecting more precisely each of the SR45 isoform domains and focusing on all types of post-translational modifications would demonstrate the importance of specific residues associated to the molecular function of SR45 and to its physiological responses in the plant.

Supplementary data

Supplementary Figure S1. Scaled schemes depicting the native *SR45* gene and the engineered mutant variants.

Supplementary Figure S2. Toxicity and autoactivation assays for yeast vectors harboring the native *SR45* or mutant variant coding sequences.

Supplementary Figure S3. Reporter lines depicting the activity of the native *SR45* promoter (*pSR45:EGFP*) at different developmental stages.

Supplementary Figure S4. Interpretive summary about the interactions of SR45 with kinases, splicing-associated factors, and nuclear import and mRNA export factors, as observed by directed

yeast-two-hybrid assays.

Supplementary Figure S5. Interpretive summary about the interactions of SR45 with EJC core components and peripheral factors.

Supplementary Figure S6. SR45 interactions with all the nineteen Arabidopsis SR proteins.

Supplementary Figure S7. Model of the Arabidopsis ASAP assembly.

Supplementary Figure S8. Model of the Arabidopsis PSAP assembly.

Supplementary Figure S9. The mutated RRM of SR45 does not bind to any RNA consensus.

Supplementary Figure S10. Trinucleotide frequency among the sequences used to identify SR45-associated RNA consensuses.

Supplementary Figure S11. RNA motifs binding to RRMs of SR34.

Supplementary Tables S1-9. List of primers used in this study, subdivided into experimental categories.

Supplementary Tables S10-15. List of sequences submitted to MEME to find significant consensus(es) for the native RRMs studied or to find non-significant motifs associated to mutated RRMs.

Supplementary Data S1. Model of the Arabidopsis ASAP assembly (as pdb file).

Supplementary Data S2. Model of the Arabidopsis PSAP assembly (as pdb file).

Acknowledgments

We thank Dr. Julien Spielmann and Prof. Moreno Galleni for helpful discussions.

Author contributions

PM conceived and directed the research. SF conducted most experiments, with contributions of MS, MJ, SDF and MV for SELEX experiments. FK performed AlphaFold computational modelling. PM, FK and SF analyzed the data. PM and MH supervised experiments. SF made the figures, with the help of FK and SDF for computational modelling. SF, FK, MH and PM wrote the manuscript, with comments of all authors.

Conflict of interest

No conflict of interest is declared.

Funding

Funding was provided by the "Fonds de la Recherche Scientifique–FNRS" (FRFC-1.E049.15, PDR-T.0206.13, CDR-J.0009.17, PDR-T0120.18, CDR-J.0082.21, PDR-T.0121.22, PDR-T.0104.22). MH and FK were Senior Research Associate of the F.R.S.-FNRS. SF was a doctoral fellow (F.R.I.A.).

Data availability

All data supporting the findings of the study are available within the paper and within its supplementary materials published online, or upon request.

References

Albaqami, M. (2023). The Splicing Factor SR45 Negatively Regulates Anthocyanin Accumulation under High-Light Stress in *Arabidopsis thaliana*. *Life* **13**.

Albuquerque-Martins, R., Szakonyi, D., Rowe, J., Jones, A.M., and Duque, P. (2023). ABA signaling prevents phosphodegradation of the SR45 splicing factor to alleviate inhibition of early seedling development in *Arabidopsis*. *Plant Communications* **4**: 100495.

Ali, G.S., Golovkin, M., and Reddy, A.S. (2003). Nuclear localization and *in vivo* dynamics of a plant-specific serine/arginine-rich protein. *The Plant Journal* **36**: 883-893.

Ali, G.S., Prasad, K.V., Hanumappa, M., and Reddy, A.S. (2008). Analyses of in vivo interaction and mobility of two spliceosomal proteins using FRAP and BiFC. *PLoS One* **3**: e1953.

Ali, G.S., and Reddy, A.S. (2006). ATP, phosphorylation and transcription regulate the mobility of plant splicing factors. *Journal of Cell Science* **119**: 3527-3538.

Aubol, B.E., Hailey, K.L., Fattet, L., Jennings, P.A., and Adams, J.A. (2017). Redirecting SR Protein Nuclear Trafficking through an Allosteric Platform. *Journal of Molecular Biology* **429**: 2178-2191.

Aubol, B.E., Plocinik, R.M., Hagopian, J.C., Ma, C.T., McGlone, M.L., Bandyopadhyay, R., Fu, X.D., and Adams, J.A. (2013). Partitioning RS domain phosphorylation in an SR protein through the CLK and SRPK protein kinases. *Journal of Molecular Biology* **425**: 2894-2909.

Aubol, B.E., Wozniak, J.M., Fattet, L., Gonzalez, D.J., and Adams, J.A. (2021). CLK1 reorganizes the splicing factor U1-70K for early spliceosomal protein assembly. *Proceedings of the National Academy of Sciences* **118**.

Aubol, B.E., Wu, G., Keshwani, M.M., Movassat, M., Fattet, L., Hertel, K.J., Fu, X.D., and Adams, J.A. (2016). Release of SR Proteins from CLK1 by SRPK1: A Symbiotic Kinase System for Phosphorylation Control of Pre-mRNA Splicing. *Molecular Cell* **63**: 218-228.

Ausin, I., Greenberg, M.V., Li, C.F., and Jacobsen, S.E. (2012). The splicing factor SR45 affects the RNA-directed DNA methylation pathway in Arabidopsis. *Epigenetics* **7**: 29-33.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* **37**: W202-208.

Barbosa Dos Santos, I., and Park, S.W. (2019). Versatility of Cyclophilins in Plant Growth and Survival: A Case Study in *Arabidopsis*. *Biomolecules* **9**.

Barta, A., Kalyna, M., and Reddy, A.S. (2010). Implementing a rational and consistent nomenclature for serine/arginine-rich protein splicing factors (SR proteins) in plants. *The Plant Cell* **22**: 2926-2929.

Bender, J., and Fink, G.R. (1994). AFC1, a LAMMER kinase from *Arabidopsis thaliana*, activates STE12-dependent processes in yeast. *Proceedings of the National Academy of Sciences* **91**: 12105-12109.

Bi, Y., Deng, Z., Ni, W., Shrestha, R., Savage, D., Hartwig, T., Patil, S., Hong, S.H., Zhang, Z., Oses-Prieto, J.A., Li, K.H., Quail, P.H., Burlingame, A.L., Xu, S.L., and Wang, Z.Y. (2021). Arabidopsis ACINUS is O-glycosylated and regulates transcription and alternative splicing of regulators of reproductive transitions. *Nature Communications* **12**: 945.

Bono, F., Ebert, J., Unterholzner, L., Güttler, T., Izaurralde, E., and Conti, E. (2004). Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex. *EMBO Reports* **5**: 304-310.

Califice, S., Baurain, D., Hanikenne, M., and Motte, P. (2012). A single ancient origin for prototypical serine/arginine-rich splicing factors. *Plant physiology* **158**: 546-560.

Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature Reviews Molecular Cell Biology* **10**: 741-754.

Chen, S.L., Rooney, T.J., Hu, A.R., Beard, H.S., Garrett, W.M., Mangalath, L.M., Powers, J.J., Cooper, B., and Zhang, X.N. (2019). Quantitative Proteomics Reveals a Role for SERINE/ARGININE-Rich 45 in Regulating RNA Metabolism and Modulating Transcriptional Suppression via the ASAP Complex in *Arabidopsis thaliana*. *Frontiers in Plant Science* **10**: 1116.

Cilano, K., Mazanek, Z., Khan, M., Metcalfe, S., and Zhang, X.N. (2016). A New Mutation, *hap1-2*, Reveals a C Terminal Domain Function in AtMago Protein and Its Biological Effects in Male Gametophyte Development in *Arabidopsis thaliana*. *PLoS One* **11**: e0148200.

Cléry, A., Blatter, M., and Allain, F.H. (2008). RNA recognition motifs: boring? Not quite. *Current Opinion in Structural Biology* **18**: 290-298.

Cléry, A., Jayne, S., Benderska, N., Dominguez, C., Stamm, S., and Allain, F.H. (2011). Molecular basis of purine-rich RNA recognition by the human SR-like protein Tra2-β1. *Nature structural & molecular biology* **18**: 443-450.

Cléry, A., Krepl, M., Nguyen, C.K.X., Moursy, A., Jorjani, H., Katsantoni, M., Okoniewski, M., Mittal, N., Zavolan, M., Sponer, J., and Allain, F.H. (2021). Structure of SRSF1 RRM1 bound to

RNA reveals an unexpected bimodal mode of interaction and explains its involvement in SMN1 exon7 splicing. *Nature Communications* **12**: 428.

Cléry, A., Sinha, R., Anczuków, O., Corrionero, A., Moursy, A., Daubner, G.M., Valcárcel, J., Krainer, A.R., and Allain, F.H. (2013). Isolated pseudo-RNA-recognition motifs of SR proteins can regulate splicing using a noncanonical mode of RNA recognition. *Proceedings of the National Academy of Sciences* **110**: E2802-2811.

Costa, R., Rodia, M.T., Zini, N., Pegoraro, V., Marozzo, R., Capanni, C., Angelini, C., Lattanzi, G., Santi, S., and Cenacchi, G. (2021). Morphological study of TNPO3 and SRSF1 interaction during myogenesis by combining confocal, structured illumination and electron microscopy analysis. *Molecular and Cellular Biochemistry* **476**: 1797-1811.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. *Genome Research* **14**: 1188-1190.

Curtis, M.D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant physiology* **133**: 462-469.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant physiology* **139**: 5-17.

Daubner, G.M., Cléry, A., and Allain, F.H. (2013). RRM-RNA recognition: NMR or crystallography...and new findings. *Current Opinion in Structural Biology* **23**: 100-108.

De Franco, S., Vandenameele, J., Brans, A., Verlaine, O., Bendak, K., Damblon, C., Matagne, A., Segal, D.J., Galleni, M., Mackay, J.P., and Vandevenne, M. (2019). Exploring the suitability of RanBP2-type Zinc Fingers for RNA-binding protein design. *Scientific Reports* **9**: 2484.

de la Fuente van Bentem, S., Anrather, D., Roitinger, E., Djamei, A., Hufnagl, T., Barta, A., Csaszar, E., Dohnal, I., Lecourieux, D., and Hirt, H. (2006). Phosphoproteomics reveals extensive *in vivo* phosphorylation of Arabidopsis proteins involved in RNA metabolism. *Nucleic Acids Research* **34**: 3267-3278.

Ehrnsberger, H.F., Grasser, M., and Grasser, K.D. (2019). Nucleocytosolic mRNA transport in plants : export factors and their influence on growth and development. *Journal of Experimental Botany* **70**: 3757-3763.

Fairbrother, W.G., Yeh, R.F., Sharp, P.A., and Burge, C.B. (2002). Predictive identification of exonic splicing enhancers in human genes. *Science* **297**: 1007-1013.

Fanara, S., Schloesser, M., Hanikenne, M., and Motte, P. (2022). Altered metal distribution in the *sr45-1* Arabidopsis mutant causes developmental defects. *The Plant Journal* **110**: 1332-1352.

Fernandes, H., Czapinska, H., Grudziaz, K., Bujnicki, J.M., and Nowacka, M. (2018). Crystal structure of human Acinus RNA recognition motif domain. *PeerJ* **6**: e5163.

George, A., Aubol, B.E., Fattet, L., and Adams, J.A. (2019). Disordered protein interactions for an ordered cellular transition: Cdc2-like kinase 1 is transported to the nucleus via its Ser-Arg protein substrate. *Journal of Biological Chemistry* **294**: 9631-9641.

Golovkin, M., and Reddy, A.S. (1999). An SC35-like protein and a novel serine/arginine-rich protein interact with *Arabidopsis* U1-70K protein. *Journal of Biological Chemistry* **274**: 36428-36438.

Goroncy, A.K., Koshiba, S., Tochio, N., Tomizawa, T., Inoue, M., Watanabe, S., Harada, T., Tanaka, A., Ohara, O., Kigawa, T., and Yokoyama, S. (2010). The NMR solution structures of the five constituent cold-shock domains (CSD) of the human UNR (upstream of N-ras) protein. *Journal of Structural and Functional Genomics* **11**: 181-188.

Graveley, B.R. (2001). Alternative splicing: increasing diversity in the proteomic world. *Trends in Genetics* **17**: 100-107.

Gullerova, M., Barta, A., and Lorković, Z.J. (2006). AtCyp59 is a multidomain cyclophilin from *Arabidopsis thaliana* that interacts with SR proteins and the C-terminal domain of the RNA polymerase II. *RNA* **12**: 631-643.

Hanikenne, M., Talke, I.N., Haydon, M.J., Lanz, C., Nolte, A., Motte, P., Kroymann, J., Weigel, D., and Krämer, U. (2008). Evolution of metal hyperaccumulation required cis-regulatory changes and triplication of *HMA4*. *Nature* **453**: 391-395.

Hauer, C., Sieber, J., Schwarzl, T., Hollerer, I., Curk, T., Alleaume, A.M., Hentze, M.W., and Kulozik, A.E. (2016). Exon Junction Complexes Show a Distributional Bias toward Alternatively Spliced mRNAs and against mRNAs Coding for Ribosomal Proteins. *Cell Reports* **16**: 1588-1603.

Hsu Ia, W., Hsu, M., Li, C., Chuang, T.W., Lin, R.I., and Tarn, W.Y. (2005). Phosphorylation of Y14 modulates its interaction with proteins involved in mRNA metabolism and influences its methylation. *Journal of Biological Chemistry* **280**: 34507-34512.

Huang, Y., Gattoni, R., Stévenin, J., and Steitz, J.A. (2003). SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Molecular Cell* **11**: 837-843.

Huang, Y., and Steitz, J.A. (2001). Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Molecular Cell* **7**: 899-905.

Johnson, M.A., von Besser, K., Zhou, Q., Smith, E., Aux, G., Patton, D., Levin, J.Z., and Preuss, D. (2004). Arabidopsis *hapless* mutations define essential gametophytic functions. *Genetics* **168**: 971-982.

Joris, M., Schloesser, M., Baurain, D., Hanikenne, M., Muller, M., and Motte, P. (2017). Number of inadvertent RNA targets for morpholino knockdown in *Danio rerio* is largely underestimated: evidence from the study of Ser/Arg-rich splicing factors. *Nucleic Acids Research* **45**: 9547-9557.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**: 583-589.

Kalyna, M., Simpson, C.G., Syed, N.H., Lewandowska, D., Marquez, Y., Kusenda, B., Marshall, J., Fuller, J., Cardle, L., McNicol, J., Dinh, H.Q., Barta, A., and Brown, J.W. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in *Arabidopsis*. *Nucleic Acids Research* **40**: 2454-2469.

Kataoka, N., Bachorik, J.L., and Dreyfuss, G. (1999). Transportin-SR, a nuclear import receptor for SR proteins. *Journal of Cell Biology* **145**: 1145-1152.

Koroleva, O.A., Calder, G., Pendle, A.F., Kim, S.H., Lewandowska, D., Simpson, C.G., Jones, I.M., Brown, J.W., and Shaw, P.J. (2009). Dynamic behavior of *Arabidopsis* eIF4A-III, putative core protein of exon junction complex: fast relocation to nucleolus and splicing speckles under hypoxia. *The Plant Cell* **21**: 1592-1606.

Koutroumani, M., Papadopoulos, G.E., Vlassi, M., Nikolakaki, E., and Giannakouros, T. (2017). Evidence for disulfide bonds in SR Protein Kinase 1 (SRPK1) that are required for activity and nuclear localization. *PLoS One* **12**: e0171328.

Kretova, M., Selicky, T., Cipakova, I., and Cipak, L. (2023). Regulation of Pre-mRNA Splicing: Indispensable Role of Post-Translational Modifications of Splicing Factors. *Life* **13**.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **372**: 774-797.

Lai, M.C., Lin, R.I., Huang, S.Y., Tsai, C.W., and Tarn, W.Y. (2000). A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. *Journal of Biological Chemistry* **275**: 7950-7957.

Lai, M.C., Lin, R.I., and Tarn, W.Y. (2001). Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. *Proceedings of the National Academy of Sciences* **98**: 10154-10159.

Laloum, T., Martín, G., and Duque, P. (2018). Alternative Splicing Control of Abiotic Stress Responses. *Trends in Plant Science* **23**: 140-150.

Lau, C.K., Diem, M.D., Dreyfuss, G., and Van Duyne, G.D. (2003). Structure of the Y14-Magoh core of the exon junction complex. *Current Biology* **13**: 933-941.

Le Hir, H., Gatfield, D., Izaurralde, E., and Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *The EMBO Journal* **20**: 4987-4997.

Le Hir, H., Saulière, J., and Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. *Nature Reviews Molecular Cell Biology* **17**: 41-54.

Lee, K., Du, C., Horn, M., and Rabinow, L. (1996). Activity and autophosphorylation of LAMMER protein kinases. *Journal of Biological Chemistry* **271**: 27299-27303.

Lorković, Z.J., Lopato, S., Pexa, M., Lehner, R., and Barta, A. (2004). Interactions of *Arabidopsis* RS domain containing cyclophilins with SR proteins and U1 and U11 small nuclear ribonucleoprotein-specific proteins suggest their involvement in pre-mRNA Splicing. *Journal of Biological Chemistry* **279**: 33890-33898.

Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. (2001). Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. *Science* **293**: 1836-1839.

Ma, C.T., Hagopian, J.C., Ghosh, G., Fu, X.D., and Adams, J.A. (2009). Regiospecific phosphorylation control of the SR protein ASF/SF2 by SRPK1. *Journal of Molecular Biology* **390**: 618-634.

Mabin, J.W., Woodward, L.A., Patton, R.D., Yi, Z., Jia, M., Wysocki, V.H., Bundschuh, R., and Singh, G. (2018). The Exon Junction Complex Undergoes a Compositional Switch that Alters mRNP Structure and Nonsense-Mediated mRNA Decay Activity. *Cell Reports* **25**: 2431-2446 e2437.

Maertens, G.N., Cook, N.J., Wang, W., Hare, S., Gupta, S.S., Öztop, I., Lee, K., Pye, V.E.,

Cosnefroy, O., Snijders, A.P., KewalRamani, V.N., Fassati, A., Engelman, A., and Cherepanov, P. (2014). Structural basis for nuclear import of splicing factors by human Transportin 3. *Proc Natl Acad Sci U S A* **111**: 2728-2733.

Manley, J.L., and Krainer, A.R. (2010). A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes & development* **24**: 1073-1074.

Maris, C., Dominguez, C., and Allain, F.H. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *The FEBS Journal* **272**: 2118-2131.

Merkle, T. (2011). Nucleo-cytoplasmic transport of proteins and RNA in plants. *Plant Cell Reports* **30**: 153-176.

Mikulski, P., Wolff, P., Lu, T., Nielsen, M., Echevarria, E.F., Zhu, D., Questa, J.I., Saalbach, G., Martins, C., and Dean, C. (2022). VAL1 acts as an assembly platform co-ordinating co-transcriptional repression and chromatin regulation at Arabidopsis *FLC*. *Nature Communications* **13**: 5542.

Morozova, N., Allers, J., Myers, J., and Shamoo, Y. (2006). Protein-RNA interactions: exploring binding patterns with a three-dimensional superposition analysis of high resolution structures. *Bioinformatics* **22**: 2746-2752.

Mufarrege, E.F., Gonzalez, D.H., and Curi, G.C. (2011). Functional interconnections of *Arabidopsis* exon junction complex proteins and genes at multiple steps of gene expression. *Journal of Experimental Botany* **62**: 5025-5036.

Müller-McNicoll, M., Botti, V., de Jesus Domingues, A.M., Brandl, H., Schwich, O.D., Steiner, M.C., Curk, T., Poser, I., Zarnack, K., and Neugebauer, K.M. (2016). SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. *Genes & development* **30**: 553-566.

Murachelli, A.G., Ebert, J., Basquin, C., Le Hir, H., and Conti, E. (2012). The structure of the ASAP core complex reveals the existence of a Pinin-containing PSAP complex. *Nature Structural & Molecular Biology* **19**: 378-386.

Ni, X., Joerger, A.C., Chaikuad, A., and Knapp, S. (2023). Structural insights into a regulatory mechanism of FIR RRM1-FUSE interaction. *Open Biology* **13**: 230031.

Nott, A., Le Hir, H., and Moore, M.J. (2004). Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes & development* **18**: 210-222.

Pabis, M., Popowicz, G.M., Stehle, R., Fernández-Ramos, D., Asami, S., Warner, L., García-Mauriño, S.M., Schlundt, A., Martínez-Chantar, M.L., Díaz-Moreno, I., and Sattler, M. (2019). HuR biological function involves RRM3-mediated dimerization and RNA binding by all three RRMs. *Nucleic Acids Research* **47**: 1011-1029.

Palusa, S.G., and Reddy, A.S. (2010). Extensive coupling of alternative splicing of pre-mRNAs of serine/arginine (SR) genes with nonsense-mediated decay. *New Phytologist* **185**: 83-89.

Park, N.I., and Muench, D.G. (2007). Biochemical and cellular characterization of the plant ortholog of PYM, a protein that interacts with the exon junction complex core proteins Mago and Y14. *Planta* **225**: 625-639.

Park, N.I., Yeung, E.C., and Muench, D.G. (2009). Mago Nashi is involved in meristem organization, pollen formation, and seed development in Arabidopsis. *Plant Science* **176**: 461-469.

Pertea, M., Mount, S.M., and Salzberg, S.L. (2007). A computational survey of candidate exonic splicing enhancer motifs in the model plant *Arabidopsis thaliana*. *BMC Bioinformatics* **8**: 159.

Pfaff, C., Ehrnsberger, H.F., Flores-Tornero, M., Sørensen, B.B., Schubert, T., Längst, G., Griesenbeck, J., Sprunck, S., Grasser, M., and Grasser, K.D. (2018). ALY RNA-Binding Proteins Are Required for Nucleocytosolic mRNA Transport and Modulate Plant Growth and Development. *Plant physiology* **177**: 226-240.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29**: e45.

Qüesta, J.I., Song, J., Geraldo, N., An, H., and Dean, C. (2016). *Arabidopsis* transcriptional repressor VAL1 triggers Polycomb silencing at *FLC* during vernalization. *Science* **353**: 485-488.

Rausin, G., Tillemans, V., Stankovic, N., Hanikenne, M., and Motte, P. (2010). Dynamic nucleocytoplasmic shuttling of an Arabidopsis SR splicing factor: role of the RNA-binding domains. *Plant physiology* **153**: 273-284.

Reddy, A.S., Marquez, Y., Kalyna, M., and Barta, A. (2013). Complexity of the alternative splicing landscape in plants. *The Plant Cell* **25**: 3657-3683.

Robinson, N.J., Procter, C.M., Connolly, E.L., and Guerinot, M.L. (1999). A ferric-chelate reductase for iron uptake from soils. *Nature* **397**: 694-697.

Rodor, J., Pan, Q., Blencowe, B.J., Eyras, E., and Cáceres, J.F. (2016). The RNA-binding profile of Acinus, a peripheral component of the exon junction complex, reveals its role in splicing

regulation. *RNA* **22**: 1411-1426.

Sakashita, E., Tatsumi, S., Werner, D., Endo, H., and Mayeda, A. (2004). Human RNPS1 and its associated factors: a versatile alternative pre-mRNA splicing regulator in vivo. *Molecular and Cellular Biology* **24**: 1174-1187.

Sanford, J.R., Gray, N.K., Beckmann, K., and Cáceres, J.F. (2004). A novel role for shuttling SR proteins in mRNA translation. *Genes & development* **18**: 755-768.

Schlautmann, L.P., Lackmann, J.W., Altmüller, J., Dieterich, C., Boehm, V., and Gehring, N.H. (2022). Exon junction complex-associated multi-adapter RNPS1 nucleates splicing regulatory complexes to maintain transcriptome surveillance. *Nucleic Acids Research* **50**: 5899-5918.

Schütze, T., Ulrich, A.K., Apelt, L., Will, C.L., Bartlick, N., Seeger, M., Weber, G., Lührmann, R., Stelzl, U., and Wahl, M.C. (2016). Multiple protein-protein interactions converging on the Prp38 protein during activation of the human spliceosome. *RNA* **22**: 265-277.

Schwich, O.D., Blümel, N., Keller, M., Wegener, M., Setty, S.T., Brunstein, M.E., Poser, I., Mozos, I.R.L., Suess, B., Münch, C., McNicoll, F., Zarnack, K., and Müller-McNicoll, M. (2021). SRSF3 and SRSF7 modulate 3'UTR length through suppression or activation of proximal polyadenylation sites and regulation of CFIm levels. *Genome Biology* **22**: 82.

Serpeloni, M., Vidal, N.M., Goldenberg, S., Ávila, A.R., and Hoffmann, F.G. (2011). Comparative genomics of proteins involved in RNA nucleocytoplasmic export. *BMC Ecology and Evolution* **11**: 7.

Shen, H., Kan, J.L., and Green, M.R. (2004). Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. *Molecular Cell* **13**: 367-376.

Shi, H., and Xu, R.M. (2003). Crystal structure of the *Drosophila* Mago nashi-Y14 complex. *Genes & development* **17**: 971-976.

Singh, G., Kucukural, A., Cenik, C., Leszyk, J.D., Shaffer, S.A., Weng, Z., and Moore, M.J. (2012). The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151**: 750-764.

Singh, K.K., Erkelenz, S., Rattay, S., Dehof, A.K., Hildebrandt, A., Schulze-Osthoff, K., Schaal, H., and Schwerk, C. (2010). Human SAP18 mediates assembly of a splicing regulatory multiprotein complex via its ubiquitin-like fold. *RNA* **16**: 2442-2454.

Slišković, I., Eich, H., and Müller-McNicoll, M. (2022). Exploring the multifunctionality of SR

proteins. *Biochemical Society Transactions* **50**: 187-198.

Spielmann, J., Ahmadi, H., Scheepers, M., Weber, M., Nitsche, S., Carnol, M., Bosman, B., Kroymann, J., Motte, P., Clemens, S., and Hanikenne, M. (2020). The two copies of the zinc and cadmium ZIP6 transporter of *Arabidopsis halleri* have distinct effects on cadmium tolerance. *Plant, Cell & Environment* **43**: 2143-2157.

Stankovic, N., Schloesser, M., Joris, M., Sauvage, E., Hanikenne, M., and Motte, P. (2016). Dynamic Distribution and Interaction of the Arabidopsis SRSF1 Subfamily Splicing Factors. *Plant physiology* **170**: 1000-1013.

Talke, I.N., Hanikenne, M., and Krämer, U. (2006). Zinc-dependent global transcriptional control, transcriptional deregulation, and higher gene copy number for genes in metal homeostasis of the hyperaccumulator *Arabidopsis halleri*. *Plant physiology* **142**: 148-167.

Tange, T.O., Shibuya, T., Jurica, M.S., and Moore, M.J. (2005). Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA* **11**: 1869-1883.

Thomas, J., Palusa, S.G., Prasad, K.V., Ali, G.S., Surabhi, G.K., Ben-Hur, A., Abdel-Ghany, S.E., and Reddy, A.S. (2012). Identification of an intronic splicing regulatory element involved in auto-regulation of alternative splicing of *SCL33* pre-mRNA. *The Plant Journal* **72**: 935-946.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**: 701-718.

Wang, T., Wang, X., Wang, H., Yu, C., Xiao, C., Zhao, Y., Han, H., Zhao, S., Shao, Q., Zhu, J., Zhao, Y., Wang, P., and Ma, C. (2023). *Arabidopsis* SRPKII family proteins regulate flowering via phosphorylation of SR proteins and effects on gene expression and alternative splicing. *New Phytologist* **238**: 1889-1907.

Wang, Y., Liu, J., Huang, B.O., Xu, Y.M., Li, J., Huang, L.F., Lin, J., Zhang, J., Min, Q.H., Yang, W.M., and Wang, X.Z. (2015). Mechanism of alternative splicing and its regulation. *Biomedical Reports* **3**: 152-158.

Wang, Z., Ballut, L., Barbosa, I., and Le Hir, H. (2018). Exon Junction Complexes can have distinct functional flavours to regulate specific splicing events. *Scientific Reports* **8**: 9509.

Wegener, M., and Müller-McNicoll, M. (2019). View from an mRNP: The Roles of SR Proteins in Assembly, Maturation and Turnover. *Advances in Experimental Medicine and Biology* **1203**: 83-112.

Wu, H.P., Su, Y.S., Chen, H.C., Chen, Y.R., Wu, C.C., Lin, W.D., and Tu, S.L. (2014). Genome-

wide analysis of light-regulated alternative splicing mediated by photoreceptors in *Physcomitrella patens*. *Genome Biology* **15**: R10.

Wu, Z., Zhu, D., Lin, X., Miao, J., Gu, L., Deng, X., Yang, Q., Sun, K., Zhu, D., Cao, X., Tsuge, T., Dean, C., Aoyama, T., Gu, H., and Qu, L.J. (2016). RNA Binding Proteins RZ-1B and RZ-1C Play Critical Roles in Regulating Pre-mRNA Splicing and Gene Expression during Development in Arabidopsis. *The Plant Cell* **28**: 55-73.

Xiao, S.H., and Manley, J.L. (1997). Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes & development* **11**: 334-344.

Xing, D., Wang, Y., Hamilton, M., Ben-Hur, A., and Reddy, A.S. (2015). Transcriptome-Wide Identification of RNA Targets of Arabidopsis SERINE/ARGININE-RICH45 Uncovers the Unexpected Roles of This RNA Binding Protein in RNA Processing. *The Plant Cell* **27**: 3294-3308.

Xu, S., Zhang, Z., Jing, B., Gannon, P., Ding, J., Xu, F., Li, X., and Zhang, Y. (2011). Transportin-SR is required for proper splicing of *Resistance* genes and plant immunity. *PLoS Genetics* **7**: e1002159.

Yan, Q., Xia, X., Sun, Z., and Fang, Y. (2017). Depletion of *Arabidopsis* SC35 and SC35-like serine/arginine-rich proteins affects the transcription and splicing of a subset of genes. *PLoS Genetics* **13**: e1006663.

Yang, Q., Coseno, M., Gilmartin, G.M., and Doublié, S. (2011). Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. *Structure* **19**: 368-377.

Zhang, X.N., Mo, C., Garrett, W.M., and Cooper, B. (2014). Phosphothreonine 218 is required for the function of SR45.1 in regulating flower petal development in *Arabidopsis*. *Plant signaling & behavior* **9**: e29134.

Zhang, X.N., and Mount, S.M. (2009). Two alternatively spliced isoforms of the Arabidopsis SR45 protein have distinct roles during normal plant development. *Plant physiology* **150**: 1450-1458.

Zhang, X.N., Shi, Y., Powers, J.J., Gowda, N.B., Zhang, C., Ibrahim, H.M.M., Ball, H.B., Chen, S.L., Lu, H., and Mount, S.M. (2017). Transcriptome analyses reveal SR45 to be a neutral splicing regulator and a suppressor of innate immunity in *Arabidopsis thaliana*. *BMC Genomics*

18: 772.

Zhang, Z., and Krainer, A.R. (2004). Involvement of SR proteins in mRNA surveillance. *Molecular Cell* **16**: 597-607.

Figure legends

Figure 1. Expression profile of the *SR45* gene and of the corresponding protein. **(A)** Quantitative RT-PCR analyses of *SR45* gene expression in Arabidopsis vegetative and reproductive organs. Values represent means ± SEM (from three biological replicates, each consisting of pools of organs from three plants) and are relative to *AT1G58050*. Data were analyzed by one-way ANOVA followed by Dunnett multiple comparison post-tests. Statistically significant differences between means are indicated by different letters ($P<0.05$). (**B-R**) Expression profile of the SR45 protein. The fusion protein SR45:GFP is expressed in **(B)** seed coat and hilum region, **(C)** embryo (cotyledons), **(D)** radicle, **(E)** young seedling (cotyledons and root, including root hairs as indicated by the arrow), **(F-H)** two-week-old root vasculature **(F, H)** and epidermis **(G)**, **(I-J)** leaves, including stomata [**(I)**, arrow] and trichome [**(J)**, arrow], **(K-L)** gynoecium (valves, gynophore, abscission zone and receptacle) **(K)**, and stigma and style **(L)**, **(M-N)** androecium (anther and filament), **(O-Q)** petals (epidermis and veins), and **(R)** sepal epidermis. Red signals represent chlorophyll autofluoresence. At least three independent T3 lines were generated and analyzed, depicting similar fluorescence profiles.

Figure 2. Structural determinants modulating SR45 localization. Subcellular fluorescence distribution in transient expression assays in tobacco leaf cells upon N-terminal EGFP-tagging of the native SR45 and SR45 mutant variants. Dotted lines delimit the nucleus. Scale bars = 10 µm. Red signals represent chlorophyll autofluoresence. Whenever cytoplasmic and nuclear signals were detected, insets depict the detailed intranuclear localization. At least three independent transient events were generated and analyzed, depicting similar fluorescence profiles.

Figure 3. SR45 interactions detected by (directed) yeast two-hybrid screenings. From the mated culture, dilutions to an OD600 of 1 were spotted on -Trp/-Leu/-His/X-a-Gal/AurA agar plates. Positive interactions were confirmed by growth and blue staining as seen with the positive control (pGADT7-T x pGBKT7-53). Yellow and purple squares are indicative of a weakened or abolished interaction upon substitutions within the corresponding SR45 domain, respectively. A weakened interaction is characterized by the ability of several single colonies to grow, while an abolished interaction is characterized at most by a white halo of dead cells (similar to the negative control pGADT7-T x pGBKT7-Lam). An interpretive summary of the results is

presented in **Supplementary Figure S4**.

Figure 4. SR45 associates with EJC core components and peripheral factors. Interaction between the native SR45, and SR45 mutant variants, and ACINUS(RRM) and PININ (from plant ASAP and PSAP complexes, respectively), and EJC core components (MAGO and Y14) **(A)**. Ternary complexes formed in the simultaneous presence of SAP18, SR45 and either ACINUS(RRM) or PININ **(B)**. From the mated culture, dilutions to an OD600 of 1 were spotted on -Trp/-Leu/-His/X- a-Gal/AurA agar plates (for Y2H assays) **(A)** or -Met/-Trp/-Leu/-His/X-a-gal/Aureobasidin A (for Y3H assays) **(B)**. Positive interactions were confirmed by growth and blue staining. Yellow and purple squares are indicative of a weakened or abolished interaction upon substitutions within the corresponding SR45 domain, respectively. A weakened interaction is characterized by the ability of several single colonies to grow, while an abolished interaction is characterized at most by a white halo of dead cells. An interpretive summary of the results is presented in **Supplementary Figure S5**.

Figure 5. SR45 interacts with splicing factors and EJC peripheral factors in plant cells. **(A)** Bimolecular Fluorescence Complementation (BiFC) through co-expression of SR45 or SR45 mutant variants fused to the C-terminal half of YFP (SR45, SR45mutRS1, or SR45mutRS1+RS2:^{C}YFP) and ACINUS(RRM), SR34, or PRP38 fused to the N-terminal half of YFP (ACINUS(RRM), SR34, or PRP38:^{N}YFP). Dotted lines delimit the nucleus. **(B)** Trimolecular Fluorescence Complementation (TriFC) through co-expression of SR45:^{C}YFP, ACINUS:^{C}YFP or PININ:^{C}YFP with SAP18:^{N}YFP and a pMDC32 plasmid harboring no coding sequence (Ø, upper panels) or through co-expression of SR45:^{C}YFP with SAP18:^{N}YFP and a pMDC32 harboring the coding sequence of either the RRM of ACINUS (bottom left panel) or PININ (bottom right panel). Pictures represent the maximum intensity projections of 30 z-stack images from a 29.21 µm section (with a z-step size of 1.01 µm). The voltage of the EYFP PMT was set up at 700V in all experiments. For both BiFC **(A)** and TriFC **(B)** observations, representative images of fluorescence reveal the interaction of SR45 variants in the nucleus, nucleolus, and in speckles-like structures (arrowheads) in transient expression assay in tobacco leaf cells. At least three independent transient events were generated and analyzed, depicting similar fluorescence profiles. Scale bars = 10 µm. Red signals represent chlorophyll

autofluoresence.

Figure 6. Subcellular localization of SR45-associated proteins. **(A)** Subcellular fluorescence distribution in transient expression assays in tobacco leaf cells upon N-terminal EGFP-tagging (except for SRPK4 which was C-terminally tagged). Scale bars = 10 µm. Red signals represent chlorophyll autofluoresence. Whenever cytoplasmic and nuclear signals were detected, insets depict the detailed intranuclear localization. At least three independent transient events were generated and analyzed, depicting similar fluorescence profiles.

Figure 7. RNA motifs binding to the SR45 RRM. RNA motif 1 **(A)** and RNA motif 2 **(B)** identified through 4 rounds of SELEX selection with the RRM of SR45. Non-significant RNA motifs **(C)** identified using the mutated version of the RRM of SR45. The statistical significance (*E*-value) is indicated at the bottom of each consensus. Logos were redesigned using WebLogo (Crooks *et al.*, 2004). The fourth non-significant RNA motif (#4) identified using the mutated version of SR45 RRM (which presents the best *E*-value) is displayed here, but the three other RNA motifs (#1-3) are depicted in **Supplementary Figure S9**. All RNA motifs were discovered using the MEME tool (version 5.4.1).

Figure 8. A working model for the molecular function of SR45-binding proteins. Schematic diagram showing the different binding partners of SR45 identified in this study and the functional implications of these interactions. Green and red double-headed arrows indicate, respectively, the presence or the absence of interaction during directed Y2H experiments. Green circles represent positive trimeric complexes observed by Y3H experiments and TriFC *in planta*. The grey divider represents functional interconnection between nuclear complexes. The absence of interaction discussed in this study is not displayed [e.g., between either SR45, ACINUS or PININ and SAP18, or between SR45 and either eIF4IIIA and ACINUS(T607A/Y615A)].

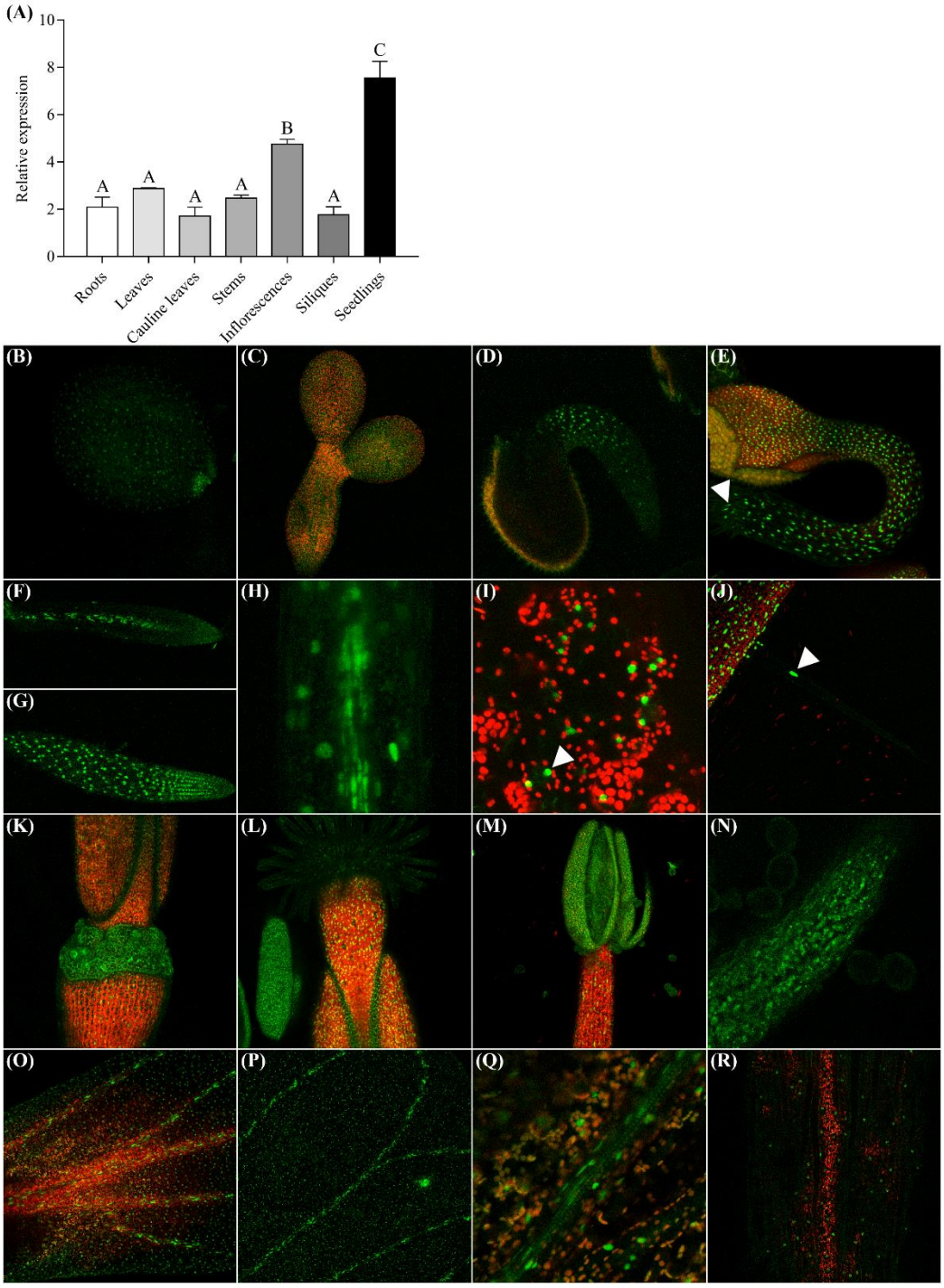

Figure 1. Expression profile of the *SR45* gene and of the corresponding protein. **(A)** Quantitative RT-PCR analyses of *SR45* gene expression in Arabidopsis vegetative and reproductive organs. Values represent means ± SEM (from three biological replicates, each consisting of pools of organs from three plants) and are relative to *AT1G58050*. Data were analyzed by one-way ANOVA followed by Dunnett multiple comparison post-tests. Statistically significant differences between means are indicated by different letters ($P<0.05$). (**B-R**) Expression profile of the SR45 protein. The fusion protein SR45:GFP is expressed in **(B)** seed coat and hilum region, **(C)** embryo (cotyledons), **(D)** radicle, **(E)** young seedling (cotyledons and root, including root hairs as indicated by the arrow), **(F-H)** two-week-old root vasculature **(F, H)** and epidermis **(G)**, **(I-J)** leaves, including stomata [**(I)**, arrow] and trichome [**(J)**, arrow], **(K-L)** gynoecium (valves, gynophore, abscission zone and receptacle) **(K)**, and stigma and style **(L)**, **(M-N)** androecium (anther and filament), **(O-Q)** petals (epidermis and veins), and **(R)** sepal epidermis. Red signals represent chlorophyll autofluoresence. At least three independent T3 lines were generated and analyzed, depicting similar fluorescence profiles.

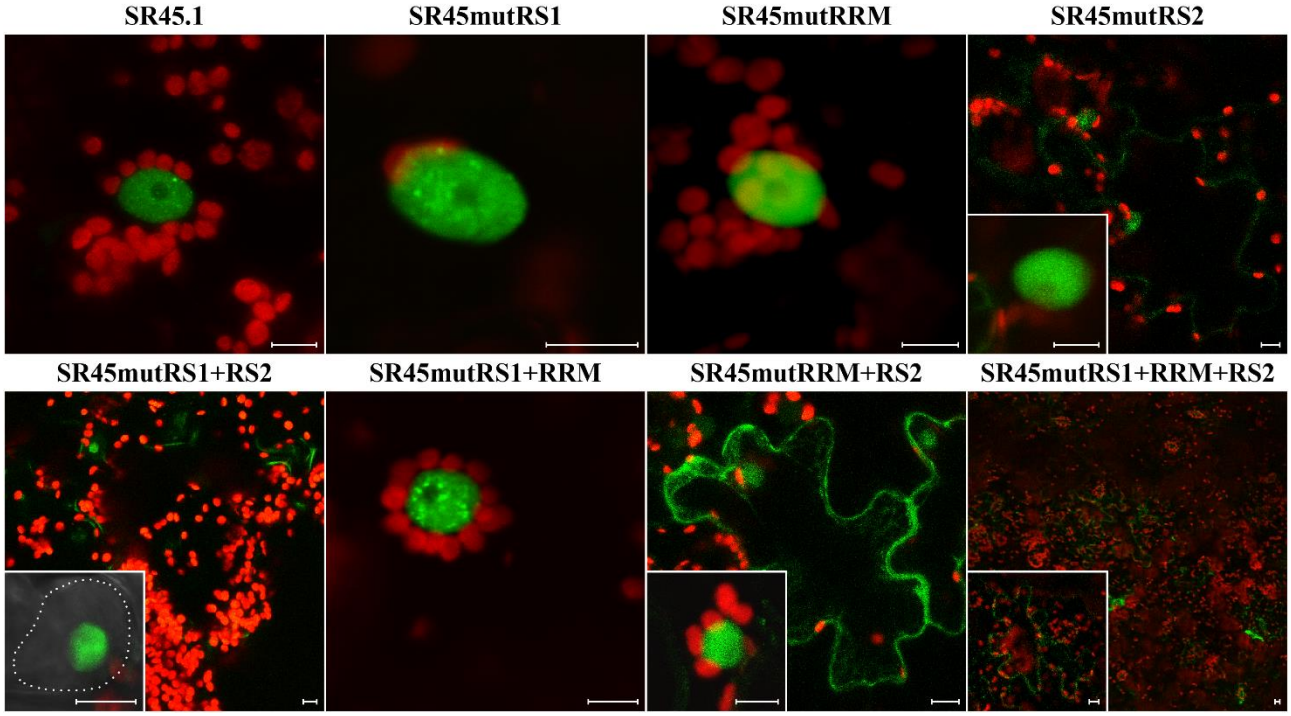

Figure 2. Structural determinants modulating SR45 localization. Subcellular fluorescence distribution in transient expression assays in tobacco leaf cells upon N-terminal EGFP-tagging of the native SR45 and SR45 mutant variants. Dotted lines delimit the nucleus. Scale bars = 10 µm. Red signals represent chlorophyll autofluoresence. Whenever cytoplasmic and nuclear signals were detected, insets depict the detailed intranuclear localization. At least three independent transient events were generated and analyzed, depicting similar fluorescence profiles.

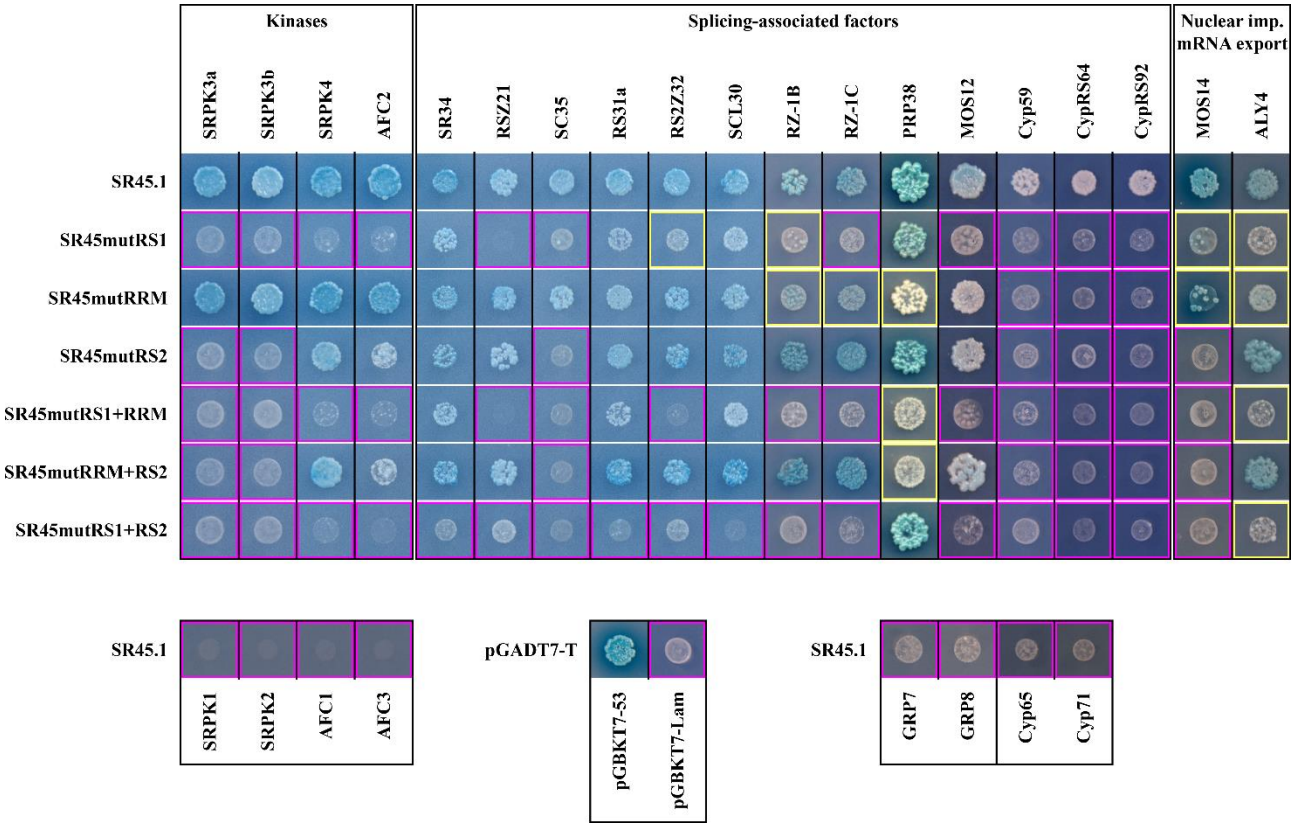

Figure 3. SR45 interactions detected by (directed) yeast two-hybrid screenings. From the mated culture, dilutions to an OD600 of 1 were spotted on -Trp/-Leu/-His/X-a-Gal/AurA agar plates. Positive interactions were confirmed by growth and blue staining as seen with the positive control (pGADT7-T x pGBKT7-53). Yellow and purple squares are indicative of a weakened or abolished interaction upon substitutions within the corresponding SR45 domain, respectively. A weakened interaction is characterized by the ability of several single colonies to grow, while an abolished interaction is characterized at most by a white halo of dead cells (similar to the negative control pGADT7-T x pGBKT7-Lam). An interpretive summary of the results is presented in **Supplementary Figure S4**.

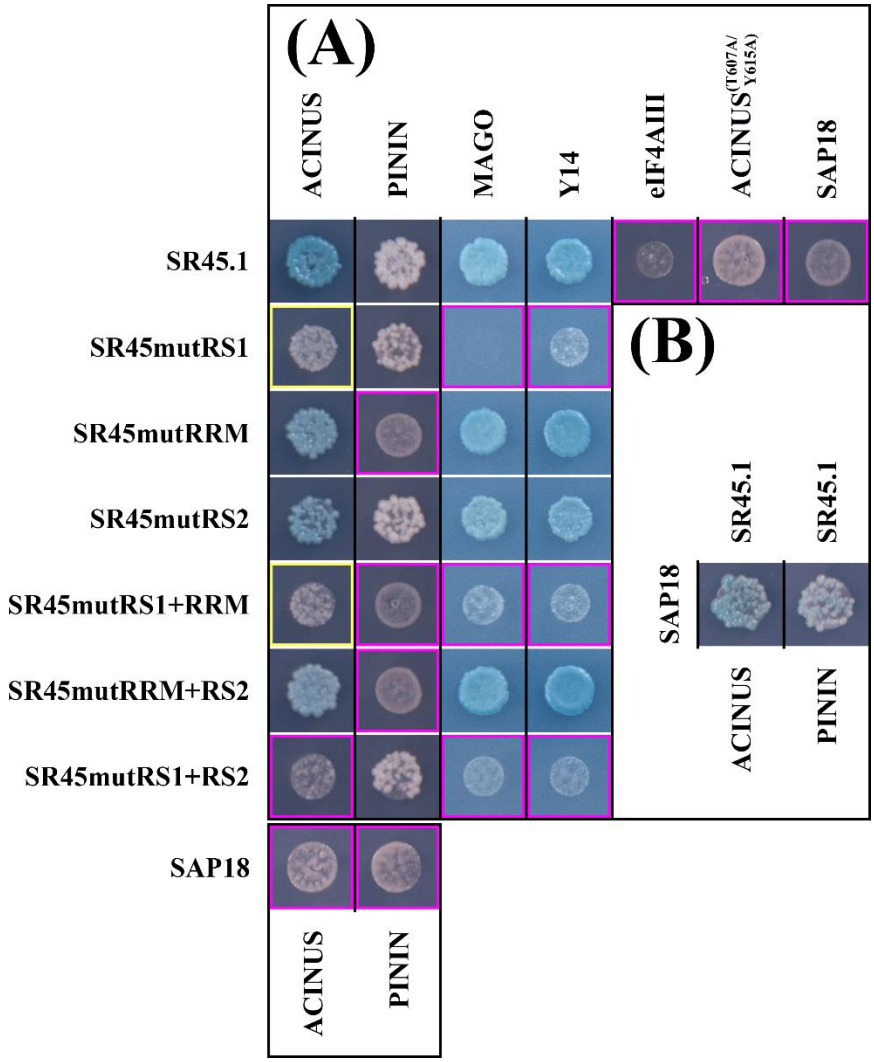

Figure 4. SR45 associates with EJC core components and peripheral factors. Interaction between the native SR45, and SR45 mutant variants, and ACINUS(RRM) and PININ (from plant ASAP and PSAP complexes, respectively), and EJC core components (MAGO and Y14) **(A)**. Ternary complexes formed in the simultaneous presence of SAP18, SR45 and either ACINUS(RRM) or PININ **(B)**. From the mated culture, dilutions to an OD600 of 1 were spotted on -Trp/-Leu/-His/X- a-Gal/AurA agar plates (for Y2H assays) **(A)** or -Met/-Trp/-Leu/-His/X-a-gal/Aureobasidin A (for Y3H assays) **(B)**. Positive interactions were confirmed by growth and blue staining. Yellow and purple squares are indicative of a weakened or abolished interaction upon substitutions within the corresponding SR45 domain, respectively. A weakened interaction is characterized by the ability of several single colonies to grow, while an abolished interaction is characterized at most by a white halo of dead cells. An interpretive summary of the results is presented in **Supplementary Figure S5**.

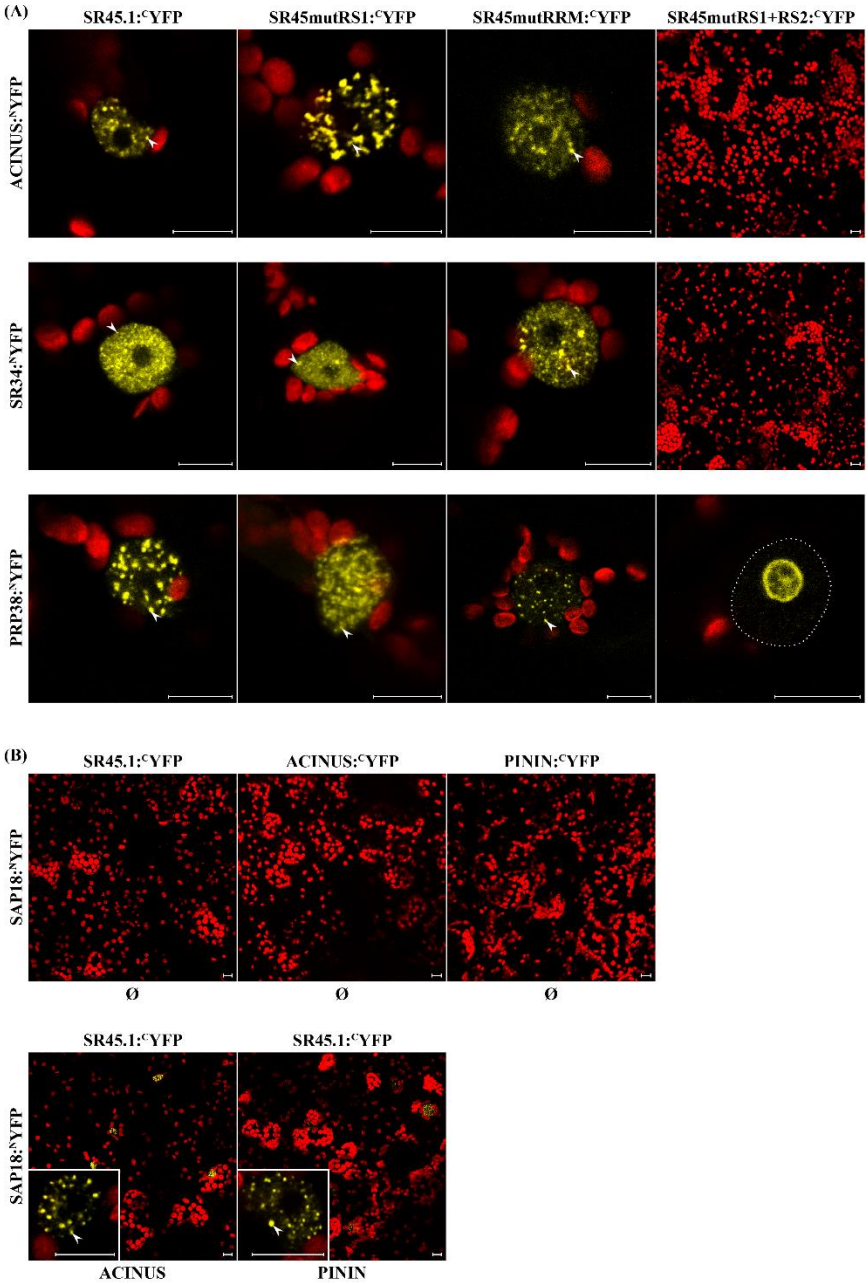

Figure 5. SR45 interacts with splicing factors and EJC peripheral factors in plant cells. **(A)** Bimolecular Fluorescence Complementation (BiFC) through co-expression of SR45 or SR45 mutant variants fused to the C-terminal half of YFP (SR45, SR45mutRS1, or SR45mutRS1+RS2:^{C}YFP) and ACINUS(RRM), SR34, or PRP38 fused to the N-terminal half of YFP (ACINUS(RRM), SR34, or PRP38:^{N}YFP). Dotted lines delimit the nucleus. **(B)** Trimolecular Fluorescence Complementation (TriFC) through co-expression of SR45:^{C}YFP, ACINUS:^{C}YFP or PININ:^{C}YFP with SAP18:^{N}YFP and a pMDC32 plasmid harboring no coding sequence (Ø, upper panels) or through co-expression of SR45:^{C}YFP with SAP18:^{N}YFP and a pMDC32 harboring the coding sequence of either the RRM of ACINUS (bottom left panel) or PININ (bottom right panel). Pictures represent the maximum intensity projections of 30 z-stack images from a 29.21 µm section (with a z-step size of 1.01 µm). The voltage of the EYFP PMT was set up at 700V in all experiments. For both BiFC **(A)** and TriFC **(B)** observations, representative images of fluorescence reveal the interaction of SR45 variants in the nucleus, nucleolus, and in speckles-like structures (arrowheads) in transient expression assay in tobacco leaf cells. At least three independent transient events were generated and analyzed, depicting similar fluorescence profiles. Scale bars = 10 µm. Red signals represent chlorophyll autofluoresence.

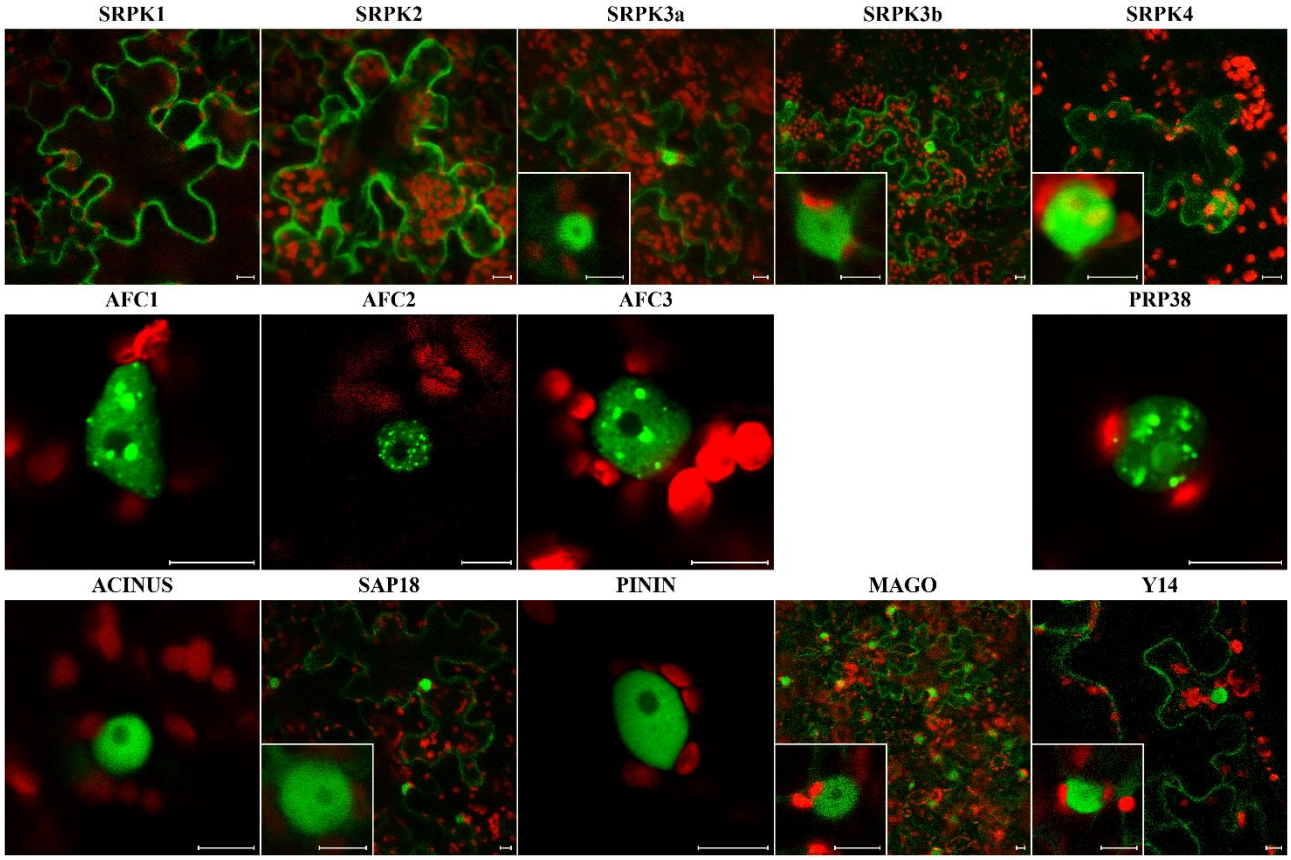

Figure 6. Subcellular localization of SR45-associated proteins. **(A)** Subcellular fluorescence distribution in transient expression assays in tobacco leaf cells upon N-terminal EGFP-tagging (except for SRPK4 which was C-terminally tagged). Scale bars = 10 µm. Red signals represent chlorophyll autofluoresence. Whenever cytoplasmic and nuclear signals were detected, insets depict the detailed intranuclear localization. At least three independent transient events were generated and analyzed, depicting similar fluorescence profiles.

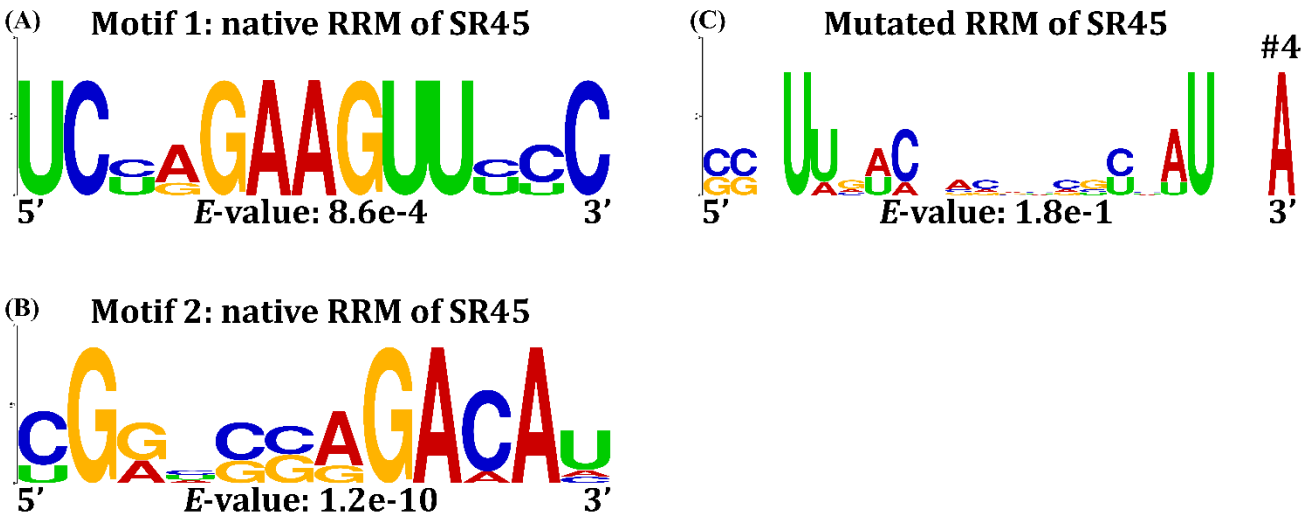

Figure 7. RNA motifs binding to the SR45 RRM. RNA motif 1 **(A)** and RNA motif 2 **(B)** identified through 4 rounds of SELEX selection with the RRM of SR45. Non-significant RNA motifs **(C)** identified using the mutated version of the RRM of SR45. The statistical significance (*E*-value) is indicated at the bottom of each consensus. Logos were redesigned using WebLogo (Crooks *et al.*, 2004). The fourth non-significant RNA motif (#4) identified using the mutated version of SR45 RRM (which presents the best *E*-value) is displayed here, but the three other RNA motifs (#1-3) are depicted in **Supplementary Figure S9**. All RNA motifs were discovered using the MEME tool (version 5.4.1).

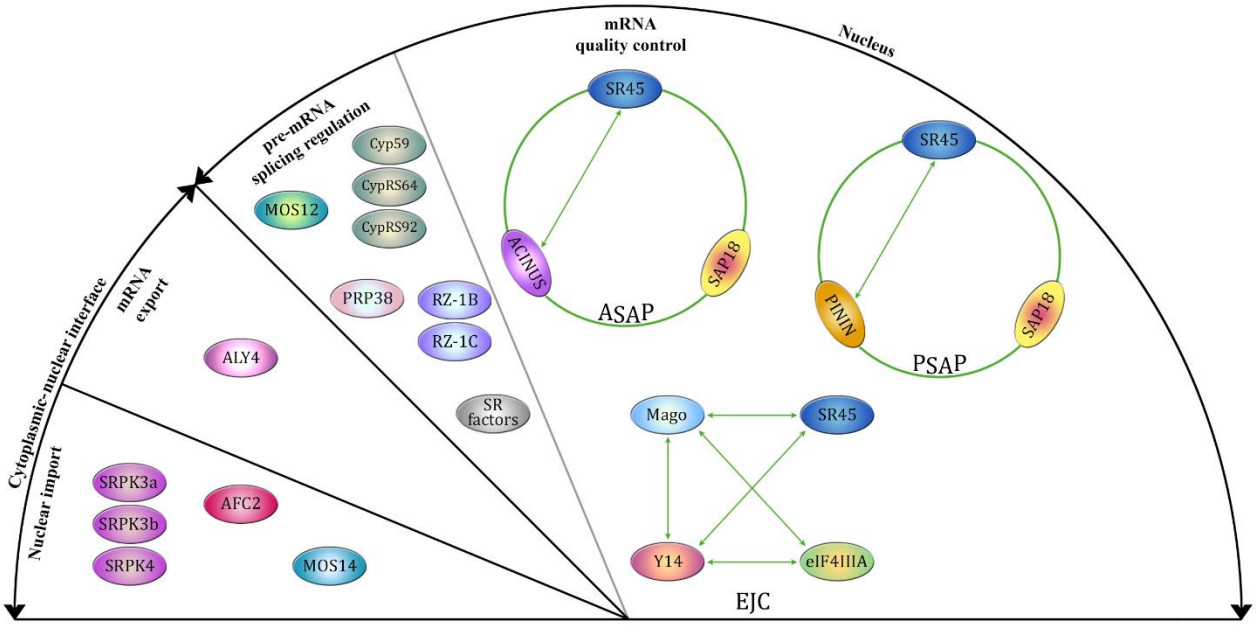

Figure 8. A working model for the molecular function of SR45-binding proteins. Schematic diagram showing the different binding partners of SR45 identified in this study and the functional implications of these interactions. Green and red double-headed arrows indicate, respectively, the presence or the absence of interaction during directed Y2H experiments. Green circles represent positive trimeric complexes observed by Y3H experiments and TriFC *in planta*. The grey divider represents functional interconnection between nuclear complexes. The absence of interaction discussed in this study is not displayed [e.g., between either SR45, ACINUS or PININ and SAP18, or between SR45 and either eIF4IIIA and ACINUS(T607A/Y615A)].