

PAPER • OPEN ACCESS

## A Novel Serine/Arginine-Rich Like Protein GjSR45a in *Gardenia jasminoides*

To cite this article: Lan Gao 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **242** 042004

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the **collection** - download the first chapter of every title for free.

# A Novel Serine/Arginine-Rich Like Protein GjSR45a in *Gardenia jasminoides*

**Lan Gao**

School of Life Sciences and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou, Guangdong, 510006, China

glgdpu@aliyun.com

**Abstract.** In eukaryotes, the expression of most genes is regulated by alternative pre-mRNA splicing (AS), through which introns are selectively removed to generate multiple transcript variants from a single gene. Serine/arginine-rich (SR) proteins are major modulators of alternative splicing, a key generator of proteomic diversity and flexible means of regulating gene expression likely to be crucial in plant environmental responses. In this work, we aimed to isolate a gene of SR from *Gardenia jasminoides*. A *G. jasminoides* fruit cDNA library was constructed, and the GjSR45a cDNA was isolated from the cDNA library by sequencing method. Sequence analysis revealed an open reading frame encoding a protein of 237 amino acid residues. Alignment of the amino acid sequence with AtSR45a and Human Tra2b revealed that the protein contains RRMs (RNA recognition motifs) motif and two RS (arginine/serine-rich) domains at the N- and the C-terminus, which are modular domains characteristic of RS-rich splicing factor. A three-dimensional dimmer model of GjSR45a was built. Comparison of the 3D structure with HomTra2b revealed highly similarity. The results suggest that GjSR45a is a SR45a like protein mainly responsible for alternative splicing.

## 1. Introduction

Precursor messenger RNA splicing is an essential step in gene expression mediated by the spliceosome, a large protein complex in the cell nucleus that interacts with specific intronic sequences in the pre-mRNA called splice sites for the proper removal of introns and correct joining of exons. These steps include pre-mRNA processing in the form of 5'-end capping, splicing and 3'-end cleavage/polyadenylation, as well as nuclear export of the mature mRNP. Alternative splicing (AS) occurs when splice sites are differentially recognized, allowing for the production of multiple transcripts from a single gene that can potentially result in different protein isoforms. Alternative splicing occurs in a developmental stage-, sex- or tissue-specific manner and in response to the surrounding microenvironment [1, 2]. In humans, about 95% of the multi-exon genes undergo AS [1, 3], of which misregulation is associated with numerous severe diseases [4, 5]. AS also has important biological consequences in plant growth and development, flowering, circadian clock function, and stress responses [6, 7, 8, 9 &10]. In higher plants, half or more of the intron-containing genes are alternatively-spliced [11, 12], with the current estimate in Arabidopsis plants grown under normal conditions being above 60% [13].

The mechanism of AS in plants involves the concerted action of multiple splicing factors, spliceosomal components, and their interaction with cis-acting regulatory sequences on the pre-mRNA such as exonic and intronic splicing enhances and suppressors [6]. Plants and vertebrates differ in prevalence of types of AS events with a high occurrence of intron retention in plants and a high



frequency of exon skipping in animals [6]. Splicing factors primarily comprise a conserved Ser/Arg-rich protein family, heterogeneous nuclear ribonucleoproteins, and lysine homology domain-containing proteins [14].

SR proteins are described as a highly conserved family of RNA-binding proteins (RBPs) with key roles in AS [15]. SR proteins dynamically participate in spliceosome assembly through both protein-protein and protein-RNA interactions. They promote the recruitment of the heterodimeric splicing factor U2AF and the U1 snRNP to the 3' and 5' splice sites, respectively. These proteins contain at least one RNA recognition motifs (RRMs) at the N termini and an arginine/serine-rich (RS) domain at the C terminus. The RRM binds to RNA sequence in a coordinated pattern to determine splicing specificity and commit pre-mRNA substrates to the splicing pathway. The RS domains contain multiple RS dipeptide repeats and mediate specific protein-protein interactions in a number of spliceosomal assembly steps [16]. The 12 human SR proteins have a modular domain structure, with one or two RRM and a C-terminal RS domain comprising multiple Arg-Ser dipeptide repeats [17]. The *Arabidopsis thaliana* genome codes for 18SR proteins that can be grouped into six subfamilies. The SR, RSZ and SC subfamilies include direct orthologs of the mammalian SR splicing factors SRSF1, 7 and 2, respectively, while the SCL, RS2Z and RS subfamilies are plant-specific. The *Arabidopsis* genome encodes two SR-like proteins SR45 and SR45a (a homolog of the *Drosophila* Transformer-2-like protein), which possess two RS domains at the N- and the C-terminus [18]. The N-RS domain is characteristic of plant specific type SR-like proteins. *Arabidopsis* SR45a is sometimes described as *atTra2*. The full-length *Arabidopsis* SR45a protein has been shown to interact with itself and other proteins involved in splicing, notably U1-70K, U2AF35B, and other SR and SR-related proteins [19]. SR45a produces diverse splicing variants and the relative abundance of SR45a splice variants changes under some stress conditions including salinity, high light, and cold.

*Gardenia jasminoides* originates in Asia and has been in cultivation for at least a thousand years. The fruit of *G. jasminoides* is used in Asian countries as a natural colorant, and as a traditional herbal medicine. Crocin, crocetin and geniposide are the main secondary metabolites in the fruit, and they all exhibit a wide range of pharmacological activities [20]. In this paper, we identified and analyzed a light-harvesting chlorophyll a/b-binding protein (GjSR45a) in *G. jasminoides*.

## 2. Materials and Methods

### 2.1. Plant and Growth Conditions

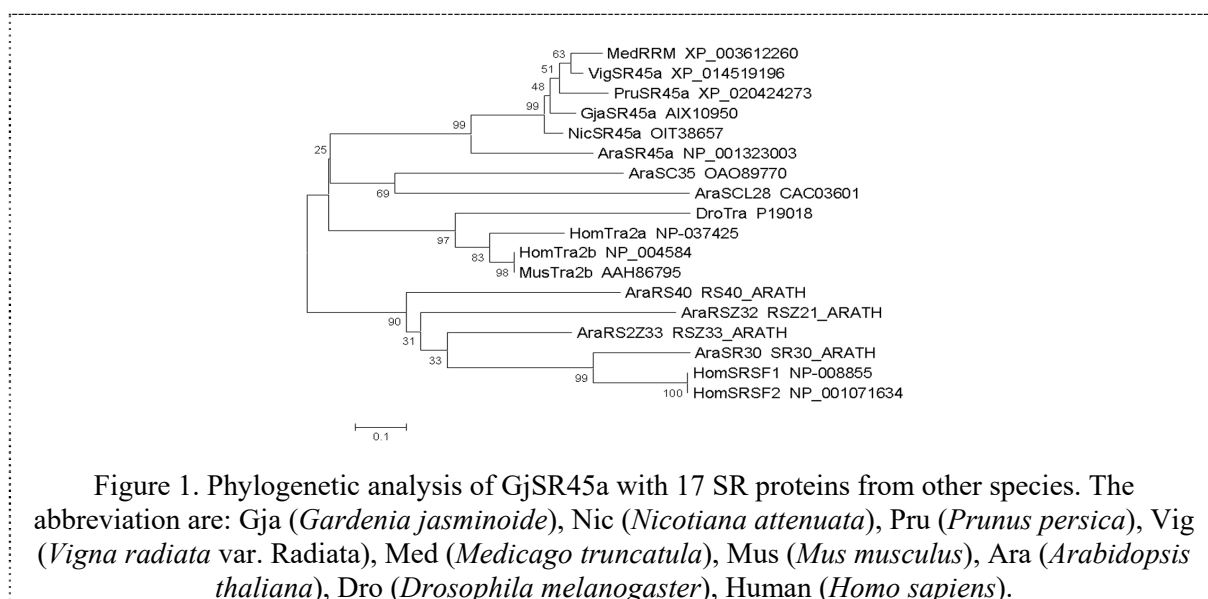
*Gardenia jasminoides* plants cultivated at Guangdong Pharmaceutical University were used as materials. Fruits were collected at development stage II, closed with yellowish green exocarp and orange mesocarp. The samples were stored at -80°C until required.

### 2.2. CDNA Library Construction, ESTs Sequencing and Cloning of GjSR45a

Total RNA was extracted from *Gardenia* fruit (stage II), using a modified CTAB (hexadecyl trimethyl ammonium bromide) based extraction protocol [21]. From total RNA, the cDNA library construction and amplification were performed following the manual of the Creator™ SMART™ cDNA Library construction Kit (Clontech, USA). The SMART cDNAs were ligated into SfiI-digested pDNR-LIB vector and transformed into *Escherichia coli* strain DH5α. Colonies were randomly picked, inoculate each colony to a separated PCR reaction solutions. The colony was lysed by heating the mixed solutions at 95°C in a PTC-200 Thermocycler (MJ Research, USA) for 5 min. After then, went to PCR amplification procedure with M13 primers provided by the Creator™ SMART™ cDNA Library construction Kit. The amplified PCR products (ESTs, expressed sequence tags) were analyzed by 1.2% agarose gel electrophoresis. When the amplified PCR products were longer than 1000 bp, incubated the isolated colonies and sequenced the ESTs. There are 40 ESTs were sequenced. After sequencing and analysis, the colony containing the predicted pDNR-LIB-GjSR45a was isolated.

### 3. Results and discussion

We identified novel SR homologues in *Gardenia jasminoides* (named GjSR45a) by exploiting the fruit cDNA library of *G. jasminoides*. The full-length GjSR45a cDNA (Genbank accession No. KM371242) was obtained. The cDNA contains a predicted 714bp ORF that encodes 237 amino acids. The predicted protein sequence of GjSR45a was compared to Genbank database, multiple sequence alignment was performed using the program Clustal Omega, the best homology was found to SR45a-like protein of *Nicotiana attenuata*. The two proteins share 78.1% identical amino acids. And the homology to some SR proteins from other species is: SR45a like protein from *Prunus persica* 74.7%, Tra2b from *Homo sapiens* 33.9%, SRSF2 from *Homo sapiens* 16.3%, Ara45a from *Arabidopsis thaliana* 30.99%. AraRS40 from *Arabidopsis thaliana* 16.9% (Figure1, 2).

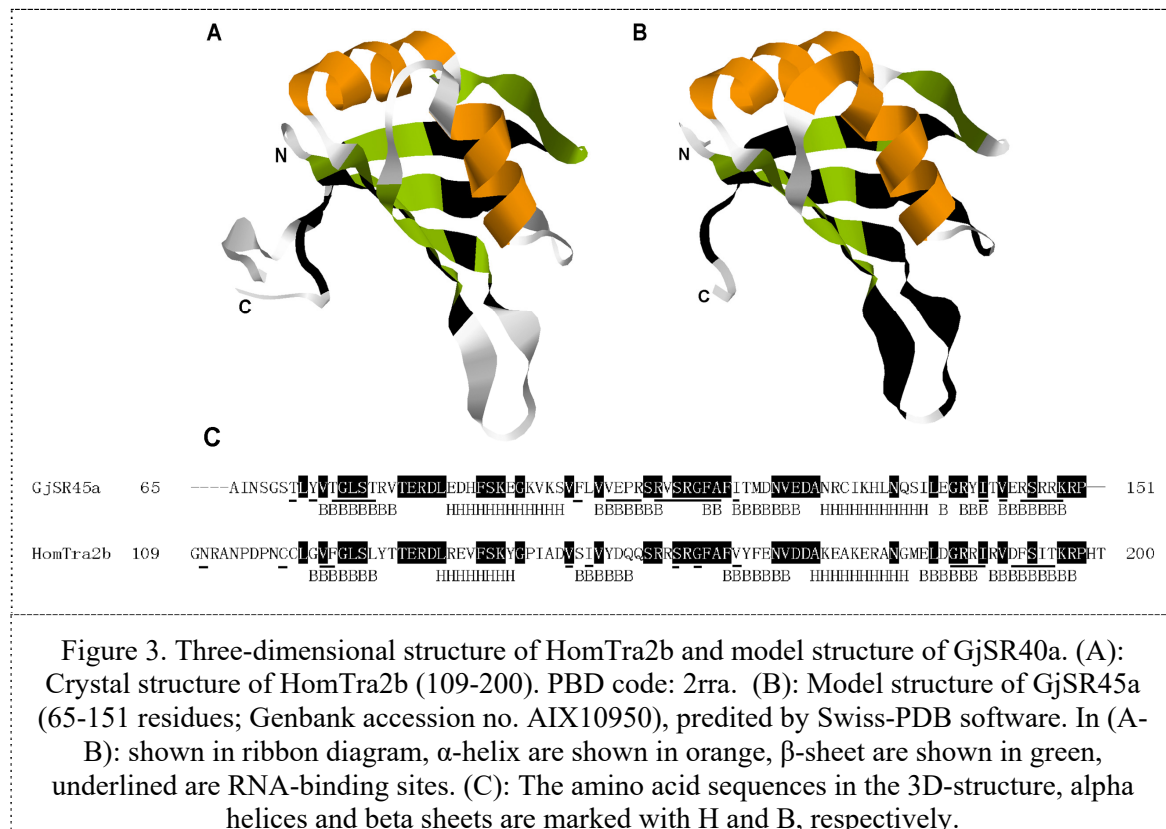


The GjSR45a secondary structure was predicted by Swiss-model software (Fig 3). GjSR45a have a common RRM domain between amino acid residues 69 and 148, and there were two RS domains at N- and C- termini. The RRM domain of GjSR45a is composed of mainly four  $\beta$ -strands and two  $\alpha$ -helices. Phylogenetic analysis of GjSR45a (using MEGA4) with representative SR proteins from database indicated that GjSR45a is clustered into the leaf type subgroup which consists of *Medicago truncatula* RRM protein, SR45a protein from *Vigna radiata* var. Radiata OsFd, *Prunus persica*, *Nicotiana attenuate* and Arabisopsis. GjSR45a is distanced to SRSF1 and SRSF2 from human.



Figure 2. Sequence alignment of GjSR45a with SR45a -like protein from other organism. Amino acids are numbered at the right of the sequence. The abbreviation and Genbank accession number are: Gja (*Gardenia jasminoides*, AIX10950), Nic (*Nicotiana attenuata*, OIT38657), Pru (*Prunus persica*, XP\_020424273), Vig (*Vigna radiata* var. *Radiata*, XP\_014519196). Underline indicates RRM. Identical residues are shaded.

The 3D (three dimension) model structure of GjSR45a was predicted using SWISS-PDB software, the NMR structure of Arabisopsis SR45a (residues from 109-200) was used as template (Figure. 3A), the GjSR45a including residues 65-151 was predicated. The two amino acid sequences have 40.9% amino acid identity (Figure. 3C). The structure was successfully built as a monomer (Figure. 3B), the fold of GjSR45a consists of two  $\alpha$ -helices flanked by four  $\beta$ -sheets. Indicate a high level of structural similarity with Arabisopsis SR45a.



Splicing is one of the most important cellular processes in maintaining the integrity of the transcriptome in eukaryotic cells. In eukaryotes, the expression of most genes is regulated by alternative pre-mRNA splicing. The importance of splicing is underlined by the increasing number of diseases associated with miss-splicing. Alternative splicing of pre-mRNAs generates protein diversity from a limited number of genes.

Serine/Arginine Splicing Factor1 (SRSF1) is the archetype member of the SR protein family of splicing regulators. The multiple functions of SRSF1 are a consequence of its RNA-binding potential, nuclear-cytoplasmic shuttling, and interactions with diverse proteins, as dictated by its structure. The modular domains of SRSF1 consist of two RRM-s-a canonical RRM at the N-terminus, followed by a pseudo-RRM-and a C-terminal RS domain that is shorter than that of most other SR proteins.

From Phylogenetic analysis, GjSR45a is similar to Tra2 than SRSF from human. A single tra2 protein is found in fruit flies, where Tra2 is one of the classical splicing regulators controlling sexual differentiation as well as being essential for spermatogenesis (Baker, 1989). The mammalian transformer-2b belongs to the SR-like protein family and has an RRM and two RS domains. One RS domain is located at the N terminus and the other at the C terminus, separated by an RRM. The pattern is same with GjSR45a. The Tra2 gene has duplicated in vertebrates, resulting in two mammalian Tra2 proteins with 63% amino acid identity. These proteins are called Tra2a and Tra2b. Tra2b binds to exons to regulate their alternative splicing inclusion. Recently the details of exactly how Tra2b protein

binds to both AGAA and CAA target RNA sequences have been revealed at atomic resolution, and involve protein-RNA interactions with both the RRM and flanking regions [22]. Tra2b is essential in the embryonic brain [23]. Besides in the central nervous system, the abnormal splicing events elicited by dysfunction of Tra2B have also been observed in cancer [24,25], stroke [26] and vascular smooth muscle diversification [27]. The expression of Tra2b mRNA was upregulated almost 2-fold during meiosis. In contrast, the expression levels of the classical SR proteins remained similar at both RNA and protein levels between the pre-meiotic and meiotic testis.

SR45a protein contains two RS (arginine/serine-rich) domains flanking a central RNA recognition motif (RRM). Its most closely related counterparts in animals are members of the transformer 2 protein family. Due to this similarity, Arabidopsis SR45a is sometimes described as atTra2.

Members of the plant-specific SC35-like (SCL) Arabidopsis SR protein subfamily are distinctively responsive to exogenous ABA, while the expression of seven SR and SR-related genes is affected by alterations in key components of the ABA pathway. SR34, SR34b, SCL30a, SCL28, SCL33, RS40, SR45 and SR45a are promising candidates for involvement in ABA-mediated stress responses [28]. Expression of SR45a is markedly induced by high light stress [29]. While DNA microarray studies [30] and RNA-seq analysis of abiotic stress samples [31] have shown that SR45a mRNA abundance is also altered upon heat and water deprivation stress.

We demonstrated in this work that, GjSR45a belongs to a plant type splicing factor and it might be a key regulator of alternative splicing, and more functions in plant need to be explored.

## References

- [1] Wang, E.T., Sandberg, R., Luo, S. et al. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, 456:470-476.
- [2] Baralle, F.E., Giudice, J. (2017) Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol.*, 18:437-451.
- [3] Ramanouskaya, T.V., Grinev, V.V. (2017) The determinants of alternative RNA splicing in human cells. *Mol Genet Genomics.*, 292:1175-1195.
- [4] Kornblihtt, A.R., Schor, I.E., Allo, M. et al. (2013) Alternative splicing: A pivotal step between eukaryotic transcription and translation. *Nat. Rev. Mol. Cell Biol.*, 14:153-165.
- [5] Climente-González, H., Porta-Pardo, E., Godzik, A. et al. (2017) The Functional Impact of Alternative Splicing in Cancer. *Cell Rep.*, 20:2215-2226.
- [6] Reddy, A.S.N., Marquez, Y., Kalyna, M. et al. (2013) Complexity of the alternative splicing landscape in plants. *Plant Cell.*, 25: 3657-3683.
- [7] Staiger, D., Brown, J.W.S. (2013). Alternative splicing at the intersection of biological timing, development, and stress responses. *Plant Cell.*, 25:3640-3656.
- [8] Laloum, T., Martín, G., Duque, P. (2018) Alternative Splicing Control of Abiotic Stress Responses. *Trends Plant Sci.*, 23:140-150.
- [9] Shakhmantsir, I., Nayak, S., Grant, G.R. et al. (2018) Spliceosome factors target timeless (tim) mRNA to control clock protein accumulation and circadian behavior in *Drosophila*. *Elife*, 7. pii: e39821.
- [10] Wang, G., Sun, P., Gong, Z. et al. (2018) Srk1 kinase, a SR protein-specific kinase, is important for sexual reproduction, plant infection and pre-mRNA processing in *Fusarium graminearum*. *Environ Microbiol.*, 20:3261-3277.
- [11] Lu, T., Lu, G., Fan, D. et al. (2010) Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. *Genome Res.*, 20:1238-1249.
- [12] Li, S., Yamada, M., Han, X. et al. (2016) High-Resolution Expression Map of the Arabidopsis Root Reveals Alternative Splicing and lincRNA Regulation. *Dev Cell.*, 39:508-522.
- [13] Marquez, Y., Brown, J.W., Simpson, C. et al. (2012) Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res.*, 22:1184-1195.

- [14] Ali, G.S., Reddy, A.S.N. (2008) Spatiotemporal organization of pre-mRNA splicing proteins in plants. In Nuclear pre-mRNA processing in plants. A.N. Reddy and M. Golovkin eds (Berlin, Heidelberg, Springer), PP. 103-118.
- [15] Wu, Z., Zhu, D., Lin, X. et al. (2016) RNA Binding Proteins RZ-1B and RZ-1C Play Critical Roles in Regulating Pre-mRNA Splicing and Gene Expression during Development in Arabidopsis. *Plant Cell.*, 28:55-73.
- [16] Long, J.C., Caceres, J.F. (2009) The SR protein family of splicing factors: Master regulators of gene expression. *Biochem J.*, 417:15-27.
- [17] Manley, J.L., Krainer, A.R. (2010) A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes Dev.*, 24:1073-1074.
- [18] Barta, A., Kalyna, M., Reddy, A.S. (2010) Implementing a rational and consistent nomenclature for serine/arginine-rich protein splicing factors (SR proteins) in plants. *Plant Cell.*, 22:2926-2929.
- [19] Tanabe, N., Yoshimura, K., Kimura, A. et al. (2007) Differential expression of alternatively spliced mRNAs of Arabidopsis SR protein homologs, atSR30 and atSR45a, in response to environmental stress. *Plant Cell Physiol.*, 48:1036-1049.
- [20] Hong, Y.J., Yang, K.S. (2013) Anti-inflammatory activities of crocetin derivatives from processed *Gardenia jasminoides*. *Arch Pharm Res.*, 36:933-940.
- [21] Bekesiova, I, Nap, J.P, Mlynarova, L. (1999) Isolation of high quality DNA and RNA from leaves of the carnivorous plant *Drosera rotundifolia*. *Plant Mol. Biol. Rep.*, 17: 269-277.
- [22] Tsuda, K., Someya, T., Kuwasako, K. et al. (2011) Structural basis for the dual RNA recognition modes of human Tra2  $\beta$  RRM. *Nucleic Acids Res.*, 39:1538-1553.
- [23] Grellscheid, S., Dalgliesh, C., Storbeck, M. et al. (2011) Identification of evolutionarily conserved exons as regulated targets for the splicing activator tra2 $\beta$  in development. *PLoS Genet.*, 7:e1002390.
- [24] Satake, Y., Kuwano, Y., Nishikawa, T. et al. (2017) Nucleolin facilitates nuclear retention of an ultraconserved region containing TRA2 $\beta$ 4 and accelerates colon cancer cell growth. *Oncotarget.*, 9:26817-26833.
- [25] Munkley, J., Livermore, K., Rajan, P. et al. (2017) RNA splicing and splicing regulator changes in prostate cancer pathology. RNA splicing and splicing regulator changes in prostate cancer pathology. *Hum Genet.*, 136:1143-1154.
- [26] Daoud, R., Mies, G., Smialowska, A. et al. (2002) Ischemia induces a translocation of the splicing factor tra2-1 and changes alternative splicing patterns in the brain. *J. Neurosci.*, 22:5889-5899.
- [27] Shukla, S., Fisher, S.A. (2008) Tra2 as a novel mediator of vascular smooth muscle diversification. *Circ. Res.*, 103:485-492.
- [28] Cruz T.M., Carvalho R.F., Richardson D.N. et al. (2014) Absciscic acid (ABA) regulation of Arabidopsis SR protein gene expression. *Int J Mol Sci.*, 15: 17541-17564.
- [29] Yoshimura, K., Mori, T., Yokoyama, K. et al. (2011) Identification of alternative splicing events regulated by an Arabidopsis serine/arginine-like protein, atSR45a, in response to high-light stress using a tiling array. *Plant Cell Physiol.*, 52:1786-1805.
- [30] Larkindale, J., Vierling, E. (2008) Core genome responses involved in acclimation to high temperature. *Plant Physiol.*, 146:748-761.
- [31] Gullledge, A.A., Roberts, A.D., Vora, H. et al. (2012) Mining Arabidopsis thaliana RNA-seq data with Integrated Genome Browser reveals stress-induced alternative splicing of the putative splicing regulator SR45a. *Am. J. Bot.*, 99:219-231.