Running head: SMU1 & SMU2 in pre-mRNA splicing

Brian A. Larkins
Department of Plant Sciences
University of Arizona
Tucson, AZ 85721

Email: Larkins@Ag.Arizona.edu
Phone: 520-621-9958
FAX: 520-621-3692

Journal research area: System Biology, Molecular Biology, and Gene Regulation
Title

Plant SMU-1 and SMU-2 homologues regulate pre-mRNA splicing and multiple aspects of development

Taijoon Chung², Dongfang Wang, Cheol-Soo Kim³, Ramin Yadegari, and Brian A. Larkins*

Department of Plant Sciences, University of Arizona, Tucson, AZ 85721

Key words: alternative RNA splicing, spliceosome, Zm SMU2, microarray
This research was supported by grants to BAL from the Department of Energy (DE-96ER20242), National Science Foundation (DBI-0077676) and the USDA (CSREES 2004-00918); and grants to R.Y. from the Department of Energy (DE-FG02-03ER15438) and the National Science Foundation (IOB-0520008). Partial salary support for CSK came from the Agricultural Plant Stress Research Center (KOSEF, R112001092020080).

Present address: 425-G Henry Mall, Department of Genetics, University of Wisconsin, Madison, WI 53706

Present address: Department of Plant Biotechnology and Agricultural Plant Stress Research Center, Chonnam National University, Kwangju 500-757, Korea

* The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Brian A. Larkins (Larkins@Ag.Arizona.edu)
ABSTRACT

In eukaryotes, alternative splicing of pre-mRNAs contributes significantly to the proper expression of the genome. However, the functions of many auxiliary spliceosomal proteins are still unknown. Here we functionally characterized plant homologues of nematode suppressors of mec-8 and unc-52 (smu). We compared transcript profiles of maize (Zea mays) smu2 endosperm with those of wild-type, and identified pre-mRNA splicing events that depend on the maize SMU2 protein. Consistent with a conserved role of plant SMU-2 homologues, Arabidopsis thaliana (At) smu2 mutants also show altered splicing of similar target pre-mRNAs. The At smu2 mutants occasionally show developmental phenotypes including abnormal cotyledon numbers and higher seed weights. We identified At SMU1 as one of the SMU2-interacting proteins, and At smu1 mutations cause similar developmental phenotypes with higher penetrance than At smu2. The At SMU2 and SMU1 proteins are localized to the nucleus and highly prevalent in actively dividing tissues. Taken together, our data indicated that the plant SMU-1 and SMU-2 homologues appear to be involved in splicing of specific pre-mRNAs that affect multiple aspects of development.
Pre-mRNA splicing is an essential process where the spliceosome removes introns from pre-mRNAs, and it can lead to more than one splice variant (SV), resulting in the production of different proteins. Alternative splicing has been found to be an important regulatory process in global gene expression, since the splicing pattern of a pre-mRNA can vary in different cell types and at different developmental stages, as well as in response to environmental cues (Black, 2003). A model of combinatorial control for pre-mRNA splicing was proposed to explain the apparent complexity of splicing, as the protein complex controlling splice site selection consists of positive- and negative-acting factors, e.g. SR proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs), respectively, that have synergistic and antagonistic interactions (Smith and Valcarcel, 2000; Black, 2003). The balance between splicing factors determines the pattern of pre-mRNA splicing, but tissue-specific regulators appear to be crucial for determining alternative splicing patterns between tissues.

The spliceosome consists of a group of small nuclear RNAs (snRNAs) and proteins that are recruited to pre-mRNAs to remove introns (Jurica and Moore, 2003). Recent proteomic analyses of human spliceosomal complexes (Neubauer et al., 1998; Hartmuth et al., 2002; Jurica et al., 2002; Makarov et al., 2002; Rappsilber et al., 2002; Zhou et al., 2002; Bessonov et al., 2008) revealed nearly 300 putative spliceosomal proteins. Only a fraction of these proteins constitute the catalytic core components for pre-mRNA splicing, while the others are believed to be auxiliary factors. Some of these auxiliary proteins appear to be sequence-specific splicing factors (i.e. recruited to certain sets of pre-mRNAs) or their associated proteins, and provide communication links between splicing and other processes in pre-mRNA and mRNA metabolism, such as transcription, capping, and 3’-end formation (Jurica and Moore, 2003). However, the biochemical functions of many of such auxiliary factors still remain unknown.

Mechanisms of pre-mRNA splicing in plants are assumed to be similar to those in animals, as the consensus sequences around 5’ and 3’ splice sites and branch sites are similar in plants and animals (Brown and Simpson, 1998) and components of animal spliceosomes are conserved in plant genomes (Reddy, 2001). Likewise, precise control of splice site selection during alternative splicing appears to be common and essential in all higher eukaryotes, including plants. The percentage of Arabidopsis genes with alternative splicing is currently estimated to be about 22% (Wang and Brendel, 2006), a value much smaller than is found for
humans. But the Arabidopsis genome contains at least 19 genes for SR proteins (Reddy, 2004), almost twice the number of known human SR proteins (Graveley, 2000). SR proteins and some spliceosomal proteins in plants appear to form a functional complex through their physical interactions (Reddy, 2004).

Despite conserved cis-elements and trans-acting factors, the precise function of plant splicing factors is poorly defined. This is partly because an in vitro splicing assay using plant cell extracts is currently unavailable, and a surprisingly small number of mutations in genes encoding putative plant spliceosomal proteins, such as the sad1 (supersensitive to ABA and drought 1) and stal-1 (stabilized1-1) mutants of Arabidopsis (Xiong et al., 2001; Lee et al., 2006), have been reported. SAD1 and STA1 encode proteins similar to a human Sm-like snRNP and a human U5 snRNP-associated protein, respectively. Both sad1 and stal-1 manifest altered stress responses, as well as morphological defects and reduced growth, indicating that these mutations have pleiotropic effects on plant development. Similarly, developmental defects were obvious in transgenic plants where the production of SR proteins was de-regulated. Over-expression of Arabidopsis SR proteins led to pleiotropic effects on developmental programs and influenced splice site selection among SVs encoding the SR proteins themselves, as well as other splicing factors (Lopato et al., 1999; Kalyna et al., 2003). More recently, it was found that SR proteins in rice, when overexpressed, also affect alternative splicing of their own transcripts and those of other SR protein genes (Isshiki et al., 2006).

We isolated an RNA-splicing maize mutant with a variety of phenotypes, including a reduced level of seed storage proteins and defective development of embryos and flowers (Chung et al., 2007). The mutation results from insertion of a transposable element in the first exon of a gene encoding Zea mays (Zm) SMU2, a homologue of nematode SMU-2 (Spartz et al., 2004) and a human spliceosomal protein (Neubauer et al., 1998; Makarov et al., 2002; Rappsilber et al., 2002; Zhou et al., 2002). Mutations in the nematode smu-2 gene are associated with altered splicing for unc-52 pre-mRNA, supporting a role for SMU-2 in pre-mRNA splicing (Spartz et al., 2004). Nevertheless, it is not clear how SMU-2 homologues regulate pre-mRNA splicing. The human SMU-2 homologue does not appear to be a core component of spliceosome, since it was not uniformly detected in different spliceosomal complexes (Jurica and Moore, 2003; Bessonov et al., 2008). Therefore, it is likely that SMU-2 homologues are auxiliary factors of spliceosomes. Since SMU-2 homologues lack any known RNA binding domain, these
proteins may participate in pre-mRNA splicing by interacting with other spliceosomal proteins. Based on the sequence homology and pleiotropic nature of the maize mutant, we proposed that Zm Smu2 plays a role in pre-mRNA splicing. One important prediction from this hypothesis is that one should observe altered splicing of pre-mRNAs in Zm smu2 mutants, compared to wild-type splicing patterns. Here we report data supporting this hypothesis. Transcript profiling and reverse transcription-PCR (RT-PCR) experiments revealed several potential pre-mRNA targets for Zm SMU2. We further examined the role of other plant SMU-2 homologues using Arabidopsis thaliana (At). We characterized the At SMU2 gene and obtained additional evidence supporting the role of SMU-2 homologues in pre-mRNA splicing. We found interactions of Zm and At SMU2 with proteins that imply a link with pre-mRNA splicing and other mRNA metabolic processes. We identified an interaction of At SMU2 with At SMU1, the homologue of the nematode SMU-1 protein that was previously shown to interact with SMU-2 (Spartz et al., 2004). Our data suggest that SMU-1 and SMU-2 homologues cooperate with other proteins to regulate splicing of specific pre-mRNAs.

RESULTS
Altered pre-mRNA splicing in Zm smu2 mutant endosperms
To investigate the molecular mechanisms that account for the phenotypes of Zm smu2-1 endosperms (Chung et al., 2007), microarray hybridization of RNA from homozygous wild-type and Zm smu2-1 endosperms was performed (Supplemental Protocol S1, Supplemental Table S1 and Supplemental Fig. S1). Data from this experiment can be summarized as three major observations (refer to Supplemental Protocol S1 for full description). First, many genes encoding rRNA processing factors and ribosomal proteins are up-regulated in the mutant (Supplemental Fig. 1F and 2), which appears to result from a feedback mechanism that compensates for defective rRNA processing and inefficient protein synthesis in Zm smu2 mutant endosperm (Chung et al., 2007). Second, genes involved in the biosynthesis of seed storage proteins are mostly down-regulated in the mutant (Supplemental Fig. S1F and S2), which also explains how the level of seed storage proteins is reduced in the mutant endosperm (Chung et al., 2007). Lastly, it was revealed that many probes with differential gene expression values in the mutant were derived from intronic and alternative exonic sequences (Supplemental Fig. S1E). Furthermore, about 8% of the probes with significantly lower gene expression values in the mutant were
estimated to hybridize with specific SVs, while only 4% of randomly chosen probes appeared to be SV-specific (Supplemental Tables S2, S3, S4 and Supplemental Protocol S1). These results strongly suggested altered splicing in Zm *smu2-1* endosperm.

We conducted an RT-PCR analysis of selected gene transcripts to compare pre-mRNA splicing in wild type, Zm *smu2-1* and *smu2-3* endosperms. By an RT-PCR experiment designed to identify SVs, we confirmed at least three pre-mRNAs, SVs of which were differentially accumulated in wild type and Zm *smu2-1* endosperm in the microarray experiment. Gene models corresponding to these SVs were manually constructed, based on EST evidence, similarity to conserved protein sequences, and intronic sequence consensus, as described in the Supplemental Protocol S1 and Supplemental Figure S1A. These three genes encode TRA2 (a SR-like protein similar to *Drosophila* Transformer 2 (Amrein et al., 1988); encoded by GRMZM2G073567 and found on the BAC clone AC187469.3 at location 124,155-127,755), RSp31B (an SR protein (Gupta et al., 2005); encoded by GRMZM2G021272 and found on the BAC clone AC204569.4 at location 146,139-148,813) and OTUX (an OTU-like protease (Makarova et al., 2000); encoded by GRMZM2G047838 and found on the BAC clone AC232266.1 at location 71,091-75,367). RT-PCR analysis of transcripts for these genes revealed one or more SVs that are preferentially increased or decreased in Zm *smu2-1* endosperm compared to wild type (Fig. 1, left panels labeled *Tra2*, *Rsp31b* and *Otux*), and the relative abundance of these SVs was also found to be affected by the Zm *smu2-3* mutation (Fig. 1, right panels). For example, the ratio of *Tra2* SV1 to SV3 is higher in Zm *smu2-1* endosperm than in wild-type or heterozygous siblings (Fig. 1, left panel). Similarly, the ratio of *Rsp31b* SV0 to SV1 is higher in Zm *smu2-1* than in wild-type or heterozygous siblings (Fig. 1, left panel labeled *Rsp31b*). The altered splicing pattern of *Tra2* and *Rsp31b* pre-mRNAs in Zm *smu2-3* is very similar to that in Zm *smu2-1* (Fig. 1). In contrast, the splicing pattern of *Otux* pre-mRNA in Zm *smu2-3* endosperm is opposite to that in Zm *smu2-1*: the ratio of *Otux* SV1 to SV2 is lower in Zm *smu2-3* than wild type or a heterozygote, but the ratio is higher in Zm *smu2-1* than wild-type or heterozygous sibling endosperms (Fig. 1).

Taken together, the microarray and RT-PCR analysis clearly demonstrated that Zm SMU2 affects alternative splicing of specific pre-mRNAs in developing endosperm, supporting the proposed role of Zm SMU2.
Characterization of *smu2* mutants in Arabidopsis

To functionally analyze the Arabidopsis homologue of Zm SMU2, we searched the SIGnAL T-DNA Express database (http://signal.salk.edu/cgi-bin/tdnaexpress) (Alonso et al., 2003) for mutants with a T-DNA insertion in the At *SMU2* gene (At2g26460). Two putative mutants were identified and designated At *smu2-1* and *smu2-2*, respectively. Analysis of the T-DNA-flanking sequences revealed that the At *smu2-1* T-DNA was in the eighth exon of the gene (Fig. 2A). In *smu2-2*, a T-DNA insert was confirmed in the sixth intron of the gene (Fig. 2A). To investigate the effect of these mutations on At *SMU2* gene expression, RT-PCR analysis of RNA from seedlings was performed using different combinations of primers (see Fig. 2A for the location of the primers). The At *smu2-1* (Fig. 2B, lane 2) and *smu2-2* homozygous seedlings (Fig. 2B, lane 3) did not show significant differences in RNA levels corresponding to the 5’ portion of the At *SMU2* coding sequence (see the panel of At *SMU2* (1-2) in Fig. 2B) as compared to wild-type seedlings (Fig. 2B, lane 1). However, when primers flanking the T-DNA insertion sites were used, *SMU2* transcripts were not detected in the mutants (see the panel of At *SMU2* (3-4) in Fig. 2B), indicating that the T-DNA insertions either disrupted proper transcription, pre-mRNA splicing or RNA stability. Thus, At *smu2-1* and *smu2-2* homozygous mutant plants are unlikely to produce a wild-type At SMU2 protein. However, RT-PCR analysis using primers downstream to the T-DNA insertion sites indicated a low level of detectable RNA in the At *smu2-2* mutant plants (see the panel of At *SMU2* (5-6) in Fig. 2B). This suggests that the T-DNA insertion in At *smu2-2* may create a downstream, cryptic transcription start site that can lead to the production of a partial At *SMU2* mRNA.

To determine whether the mutant plants lacked a full-length At SMU2 protein, an immunoblot analysis of extracts from wild-type and At *smu2-1* and *smu2-2* mutant plants was conducted. We previously reported that an At SMU2 protein with a glutathione-S-transferase tag showed strong cross-reaction with an antiserum prepared against Zm SMU2 (Chung et al., 2007). Of the three prominent bands detected in wild-type Arabidopsis seedlings, the largest band corresponding to a 100-kD protein was absent in the mutant extracts (Fig. 2C), indicating that it corresponds to the At SMU2 protein. Furthermore, this result showed that At *smu2-1* and *smu2-2* mutant plants do not produce a full-length At SMU2 protein, suggesting that the two alleles are null for At SMU2 function.
In order to understand the function of the At SMU2 gene during development, we analyzed the phenotypes of the *smu2-1* and *smu2-2* plants. The first generation homozygous-mutant seedlings appeared to grow slower than wild-type seedlings, although their overall architecture was indistinguishable from wild-type plants (data not shown). On the other hand, we observed the appearance of abnormal phenotypes after further selfing of the mutant alleles. For example, the weight of seeds produced in the F4 populations (self-pollinated three times) of the *smu2-1* plants was not significantly different from the wild-type seeds, while that of the F5 populations (self-pollinated 4 times) was considerably greater (Fig. 2D). Homozygous *smu2-2* F4 seeds were significantly heavier than wild-type seeds (Fig. 2D), although the F5 seeds were similar in weight to wild-type seeds (Fig. 2D). In addition, *smu2-1* and *smu2-2* plants produced a low frequency of abnormal seedlings with one, three or more cotyledons in some F3 and F4 populations. For example, 0.4% of *smu2-1* F4 seedlings (N=744) showed abnormal cotyledon number as compared to no abnormal seedlings (N=1443) observed from wild-type Col-0 plants (Table I). However, these phenotypes were not completely penetrant and varied depending on the individual parental plants tested (effectively preventing a genetic complementation of the mutant). Taken together, the genetic evidence indicates that the At SMU2 gene may play a subtle role in seed development under certain conditions.

**Genetic interaction between At smu2 and sta1**

To investigate any role for SMU2 in pre-mRNA splicing in Arabidopsis, we created double-mutant combinations between At *smu2* alleles and a mutation in the *STA1* gene, which encodes a conserved protein with high similarity to the human U5 snRNP-associated 102-kD protein and the yeast proteins, PRP1p and Prp6p (Lee et al., 2006). The *sta1-1* mutant used here is likely a partial loss-of-function allele because it contains a two-codon deletion in the coding region. Homozygous *sta1-1* plants exhibit mild developmental defects, including delayed growth and serrated leaf margins (Lee et al., 2006). We observed that although the size of the mutant plants was typically smaller than wild-type plants, they often showed more shoots per plant and produced a comparable amount of seeds as wild-type plants (Fig. 2E). The weight of the seeds obtained from *sta1-1* mutants was not significantly different from that of wild-type plants (Fig. 2D). The *sta1-1* mutant plants also showed several additional phenotypes, including an abnormal number of cotyledons (N=984, see Table I).
No double mutants were identified among 114 F2 plants analyzed from a cross of homozygous At smu2-1 and sta1-1 plants. Further genotyping of 52 F3 plants that were homozygous for smu2-1 and segregating for sta1-1 and STA1 resulted in identification of four smu2-1 sta1-1 double-mutant plants, three of which produced three cotyledons, while one produced two cotyledons of different sizes. Three of the double-mutant plants died before flowering, but one with three cotyledons developed numerous small leaves and short inflorescences bearing only one or two short siliques (Fig. 2E) with one or two seeds per silique. None of the 15 seeds obtained from this plant germinated. Similar results were obtained using the smu2-2 allele in combination with sta1-1. No double-mutant plants were obtained from F2 plants segregating for smu2-2 and sta1-1, while three double mutants were identified from F3 populations. All three smu2-2 sta1-1 double mutant seedlings ceased to grow at the cotyledon stage or before flowering. Plants that are either homozygous for sta1-1 and heterozygous for smu2-2 or heterozygous for sta1-1 and homozygous for smu2-2 produced partially filled siliques, suggesting that most of the double mutants are aborted in early seed development. Thus, compared with the single-mutant parental plants (Fig. 2E), the smu2 sta1 double-mutant plants exhibited a more severe set of developmental defects. Together with the observation that smu2 single mutants are phenotypically distinct from sta1-1, this genetic interaction indicated that At SMU2 has at least some non-overlapping functions with STA1.

The At smu2 mutations cause pre-mRNA splicing defects

To investigate whether At SMU2 functions in pre-mRNA splicing, we analyzed At smu2 mutant seedlings for genes previously known to produce SVs. Using RT-PCR, we detected three previously described SVs (mRNA1, 2 and 3) for the At1g09140 (AtSRp30) gene (Lopato et al., 1999) both in wild-type and mutant plants (Fig. 3). However, the ratio of mRNA1 to mRNA3 was higher in the two mutants than in wild-type seedlings (Fig. 3). Furthermore, an additional SV candidate was detected specifically in the mutant seedlings (Fig. 3, indicated by an asterisk). Analysis of other genes known to undergo alternative splicing, however, did not indicate any differential accumulation of their SVs in the mutant seedlings (data not shown), suggesting that only a certain number of alternative splicing events were affected by the At smu2-1 mutation.

Interestingly, AtSRp30 encodes a protein showing sequence similarity to Zm RSp31B. Based on this, we speculated that Zm smu2-dependent alternative splicing events might be
evolutionarily conserved, if they are functionally significant for the two species. A search of the Arabidopsis protein sequence database identified At5g04250 (OTUX-like) and At1g07350 (At TRA2) as putative homologues of OTUX and TRA2, respectively. Analysis of the At TRA2 mRNA from wild-type plants indicated the presence of four SVs (Fig. 3), two of which (SV1 and 2) were previously annotated in TAIR (http://www.arabidopsis.org). As shown in Figure 3, the accumulation of SV1 relative to SV2 was elevated in At smu2-1 and smu2-2 seedlings as compared to wild-type seedlings. These data indicate that the At smu2 mutations affect alternative splicing of specific pre-mRNAs. Thus, At SMU2 may function as a pre-mRNA splicing factor similar to the role played by its homologues in nematodes and maize.

Expression of At SMU2 during development

Involvement of At SMU2 in pre-mRNA splicing would suggest that the protein is localized to the nucleus. To test this, we generated transgenic Arabidopsis plants that produced an At SMU2-green fluorescent protein (GFP) fusion product (AtSMU2-GFP) under the control of a modified Cauliflower mosaic virus 35S gene promoter/enhancer sequence. As shown in Figure 4C, the AtSMU2-GFP fusion protein was localized exclusively to the root cell nuclei. In contrast, a non-conjugated GFP protein control was detected in both cytoplasm and the nucleus (Fig. 4A), and a beta-glucuronidase (GUS)-GFP fusion protein control was detected exclusively in the cytoplasm (Fig. 4B). Therefore, the nuclear localization of At SMU2 corresponds to its putative function in pre-mRNA splicing.

Immunoblot analysis of Zm SMU2 in various maize tissues indicated the highest level of expression in rapidly developing tissues or organs, such as the developing endosperm, embryo, shoot apex, and flower (Chung et al., 2007). To identify the highest level of protein accumulation for At SMU2 during development, we extracted proteins from various Arabidopsis organs and performed an immunoblot analysis using anti-Zm SMU2 antibodies. As shown in Figure 5, the highest level of At SMU2 protein was detected in the seven days-after-germination (DAG) seedlings (lane 1), young flowers before anthesis (lane 6), and developing siliques less than 5 mm in length (lane 8). Lower levels of At SMU2 accumulation were detected in roots (lane 2), expanding cauline leaves (lane 4), open flowers (lane 7), and dry seeds (lane 9), while the inflorescence (lane 5) contained only a trace amount of the protein, which was only detectable when the immunoblot was overexposed. No protein was detected in senescing rosette leaves.
Thus, the At SMU2 protein, like Zm SMU2 (Chung et al., 2007), appears to be present throughout the plant, although its abundance varied, with the highest level of accumulation in mitotically active tissues and organs.

**Zm SMU2 and At SMU2 interact with At SMU1**

To gain additional insight into the molecular function of plant SMU-2 homologues, we used the yeast two-hybrid system (Y2H) to identify proteins that interacted with Zm SMU2. Transformants containing a GAL4 binding domain fused to a full-length Zm SMU2 protein exhibited self-activation (Fig. 6A). Removal of the N-terminus (residues 1-374, construct IV) or C-terminus (residues 360-565, construct V) substantially reduced the self-activation (Fig. 6A), allowing us to create a fusion-protein bait to screen two maize and one Arabidopsis library. In a screen of a maize endosperm library using construct IV as a bait, six colonies were obtained all of which encoded a full-length histone H4 cDNA. To expand the repertoire of Zm SMU2-interacting proteins, we used construct V as a bait to screen a cDNA library prepared from young maize ears. This analysis resulted in the identification of six Zm SMU2-interacting proteins represented by 38 clones (Table II, clone IDs starting with Zm) including an SMU-1-like protein containing WD-40 repeats (Spartz et al., 2004). Three additional putative Zm SMU2-interacting proteins were identified from a screen of an Arabidopsis library of 3-DAG seedlings using construct V as a bait (Table II, clone IDs starting with At).

We primarily focused on the interaction of At SMU2 and Zm SMU2 with the SMU-1-like protein because previous findings indicated that SMU-1 and SMU-2 homologues are likely partners in formation of a functional spliceosome complex. The *C. elegans* SMU-1, which also contains WD-40-repeat motifs, interacts with SMU-2 (Spartz et al., 2004), and the human SMU1 and SMU2 proteins were found within human spliceosomal complexes (Jurica and Moore, 2003). Therefore, we used *in vitro* pull-down assays to confirm the Y2H interactions indicated above. As shown in Figure 6B, the GST-At SMU2 and GST-Zm SMU2 recombinant proteins when used as baits were capable of pulling down His-tagged At SMU2 and histone H4 proteins *in vitro*. No interaction was detected when GST alone was used to pull down the target proteins (Fig. 6B, lane 2). In addition, a His-tagged At MEC8 used as a control also did not show interaction with the GST-At SMU2 and GST-Zm SMU2 proteins. The At MEC8 protein is an Arabidopsis counterpart of the nematode MEC-8 that contains two RNA recognition motifs and has been
shown to regulate alternative splicing of the unc-52 transcripts (Lundquist et al., 1996). Taken together, these data indicate that the At and Zm SMU2 proteins interact with SMU1, a highly conserved component of the eukaryotic spliceosome.

Expression of At SMU1 during development

To examine the expression of At SMU1 during development, we generated a GFP protein fusion construct under the control of the endogenous promoter (AtSMU1-GFPP-Kan) and introduced it into Arabidopsis plants. The At SMU1-GFP fusion protein was detected mainly in actively dividing cells, such as the newly emerged leaf and the root tip of seedlings (Fig. 7A-C). A close examination of these structures using confocal microscopy revealed that the fusion protein was likely nuclear-localized (Fig. 7D-E). This was further confirmed in the root hair cells using co-localization of DAPI and GFP signals (Fig. 7F-H). GFP expression was detected during embryogenesis including the heart stage of embryo development (Fig. 7L-N). The AtSMU1-GFP fusion product was also detected in the nuclei of all cell types within the female gametophyte, including egg cell, central cell, synergid cells (Fig. 7J), and the antipodal cells (data not shown). Within the mature pollen, GFP signal was localized to both vegetative and generative nuclei (Fig. 7K). These results indicate that the sub-cellular localization and distribution of At SMU1 protein are similar to those of At SMU2.

Molecular and genetic characterization of At SMU1

Identification of Zm and At SMU1 proteins as interacting partners for Zm and At SMU2 prompted us to investigate the phenotype of smu1 loss-of-function alleles. Like At SMU2, SMU1 appeared to be a single-copy gene in the Arabidopsis genome. We identified three putative T-DNA insertional mutant alleles in the At SMU1 gene (At1g73720) and designated them smu1-1 (SALK_123852), smu1-2 (SALK_051163) and smu1-3 (SAIL_95_E04). Sequencing of the genomic DNA flanking these insertions revealed that the T-DNAs in the smu1-1 and smu1-2 alleles were located within exon 11 and intron 15, respectively (Fig. 8A). By contrast, the smu1-3 allele contained a T-DNA insertion in the promoter region. Analysis of the individual segregants from selfing populations of heterozygous smu1-1/+, smu1-2/+ or smu1-3/+ plants indicated that At SMU1 may be required for plant viability. Homozygous smu1-1 seedlings could not be maintained under normal growth conditions as they were extremely dwarfed (1-5 mm in
diameter), developed many small leaves, and did not produce any flowers (data not shown). This suggested that in comparison to smu1-2 allele (see below), smu1-1 is likely a complete loss-of-function allele. On the other hand, homozygous smu1-2 plants were generally more viable and reached the flowering stage despite exhibiting a syndrome of developmental abnormalities including (but were not limited to) abnormal cotyledon number (Table I, Fig. 8B), increased seed weight (Fig. 2D), abnormal flower positions (node spacing) (Fig. 8B), and seed germination defects (Table I). However, some of these phenotypes were associated with a reduced penetrance that varied significantly from generation to generation (Table I). The reduced severity of the phenotypes observed in smu1-2 plants was consistent with the intronic T-DNA insertion site in this allele, indicating that smu1-2 is a weaker allele than smu1-1. Finally, in contrast to the smu1-1 and smu1-2 alleles, smu1-3 did not exhibit any phenotypic changes, indicating that the upstream T-DNA insertion in this allele did not affect At SMU1 gene expression.

To see if the smu1-2 phenotypes are caused by T-DNA insertion in the corresponding gene, we performed complementation of the mutant with the wild-type At SMU1 gene. In addition to construct At SMU1-GFP<sup>Kan</sup>, we also generated construct At SMU1-GFP<sup>Bar</sup> and At SMU1<sup>Bar</sup>, and transformed smu1-2 heterozygous plants. For construct At SMU1-GFP<sup>Bar</sup>, we obtained 41 T1 transformants, and 5 of them were homozygous for the smu1-2 mutant allele. The GFP expression pattern of At SMU1-GFP<sup>Bar</sup> was similar to that of At SMU1-GFP<sup>Kan</sup>. For construct At SMU1<sup>Bar</sup>, we obtained 21 T1 transformants, and 7 of them were homozygous for smu1-2 mutant allele. The T1 transformants homozygous for smu1-2 germinated and grew normally, developed two cotyledons, and showed normal node spacing in the inflorescence. The T2 populations segregating for the transgenes displayed close to normal germination rate and cotyledon morphology (Table I). This phenotypic rescue by the transgenes clearly demonstrates that the atsmu1-2 phenotypes are caused by the T-DNA insertion in the AtSMU1 gene.

To confirm that the T-DNA insertion in smu1-2 affected expression of At SMU1, we used RT-PCR to amplify At SMU1 RNA in wild-type and mutant plants. As shown in Figure 8C, the RT-PCR product obtained with At SMU1-specific primers indicated that the transcript in smu1-2 seedlings lacked the 3’ terminal region of the coding sequence. The RT-PCR product with another pair of At SMU1-specific primers (see panel of At SMU1 (3-4) in Fig. 8C) showed that the mutant seedlings accumulated reduced levels of At SMU1 transcript compared with wild type. The smu1-2 mutation did not appear to affect At SMU2 mRNA levels (panel of At SMU2 in
Nevertheless, At SMU2 protein was not detected in *smu1-2* mutant extract (see Fig. 2C, lane 4), whereas its level in the *sta1-1* mutant was comparable to that of wild-type (lanes 1 and 5). These data indicate that the expression of At *SMU1* is required for proper accumulation of the At SMU2 protein in the cell.

Protein-interaction studies (see above) indicated that At SMU1 may function in pre-mRNA splicing in a complex with At SMU2. Therefore, we predicted that the *smu1-2* mutant would manifest altered splicing of At TRA2 and SRp30 pre-mRNAs, as was observed in At *smu2-1* and *smu2-2* mutant plants. As shown in Figure 8C, the relative abundance of At TRA2 and SRp30 SVs in *smu1-2* resembled those seen in *smu2-1* and *smu2-2* mutant plants (see Fig. 3). These data indicate that At *SMU1* and *SMU2* are involved in the alternative splicing of common pre-mRNAs.

**DISCUSSION**

We hypothesized that Zm SMU2 is a pre-mRNA splicing factor (Chung et al., 2007), and data presented in this paper provide additional evidence supporting this hypothesis. We found an evolutionarily conserved protein interaction between SMU-1 and SMU-2 homologues (Fig. 6) and nuclear localization of At SMU2-GFP (Fig. 4) and At SMU1-GFP (Fig. 7). Intriguingly, Zm *smu2*, At *smu2* and At *smu1* mutants showed an altered pattern of pre-mRNA splicing (Figs. 1, 3 and 8). It is known that over-expression of splicing regulators in plants and animals, including SR proteins, often leads to altered splicing of their own RNAs and that of other splicing regulators (Bell et al., 1991; Jumaa and Nielsen, 1997; Lopato et al., 1999; Kalyna et al., 2003; Isshiki et al., 2006). In this regard, the altered splicing patterns of *Rsp31b* and *Tra2* in Zm *smu2-1* and *smu2-3* endosperm (Fig. 1) are particularly interesting, because these genes encode an SR and an SR-related protein, respectively. The altered splicing that occurs when splicing regulators are over-expressed may be due to an auto-regulatory mechanism of the splicing regulatory network. This mechanism, though not understood, is being actively investigated (Wollerton et al., 2004; Kumar and Lopez, 2005). Altered splicing of *Tra2* and *Rsp31b* pre-mRNAs in Zm *smu2-1* could also be a secondary consequence of the Zm *smu2-1* and *smu2-3* mutations, i.e., *Tra2* and *Rsp31b* might not normally be primary targets of the Zm SMU2 protein. Consistent with this, the microarray data indicated that expression of genes encoding splicing regulators was down- or up-regulated in Zm *smu2-1* endosperm (Supplemental Table S1). Among these putative splicing
regulators are an SRp75-like protein, a KH-domain-containing protein, a TIA-1-like oligouridylate binding protein and a hnRNP A2/B1-like protein (Supplemental Table S1).

In contrast to SVs for Tra2 and Rsp31b genes, Zm smu2-1 and smu2-3 endosperms showed opposite effects on the splicing pattern in Otux pre-mRNA (Fig. 1). Importantly, the Zm SMU2 protein level is higher in smu2-1 mutant endosperm than in its wild-type sibling, while it is lower in smu2-3 than the corresponding wild type (Chung et al., 2007). Therefore, if the action of Zm SMU2 as a splicing factor follows a model of combinatorial control for pre-mRNA splicing (Smith and Valcarcel, 2000; Black, 2003), Otux pre-mRNA may represent a true target of the Zm SMU2 splicing factor, since its splicing depends on the concentration of Zm SMU2 protein.

Our genetic data also indicate that At SMU1 and SMU2 function in a similar pathway to control pre-mRNA splicing in Arabidopsis. Similar to Zm smu2 mutants, At smu2-1, smu2-2 and smu1-2 showed alterations in At TRA2 and SRp30 splicing patterns (Fig. 3 and Fig. 8C). Consistent with the proposed role of the plant SMU-2 homologues in pre-mRNA splicing, mutants of Zm smu2, At smu2 and At smu1 are all pleiotropic. The Zm smu2-1 mutant embryos often produce twin shoots and roots (Chung et al., 2007). Similarly, the At smu1-2 mutation is associated with an abnormal number of cotyledons and irregular inflorescence architecture (Fig. 8), indicating an abnormal meristematic function. Multiple developmental phenotypes are also associated with At smu2 mutations, albeit low in penetrance. In addition, the At smu1 and smu2 mutants showed similar phenotypes including increased seed weight (Fig. 2D) and abnormal cotyledon number (Table I). Increased seed weight was not associated with a mutation in the STA1 gene which presumably encodes a core spliceosomal protein, although sta1-1 mutant seedlings have similar defects in germination and cotyledon development. Furthermore, At smu2 sta1 double mutants showed very severe developmental defects (Fig. 2E), suggesting that the targets of STA1 may not overlap with those of the At SMU2 gene in the main.

It remains to be seen what pre-mRNA targets for plant SMU-2 homologues are responsible for the developmental phenotypes of Zm and At smu mutants, although our microarray analysis of Zm smu2 endosperm suggested extensive changes in the expression of genes controlling the biosynthesis and perception of plant hormones (see Supplemental Table S1). Notably, an abnormal number of cotyledons and twin embryos were also caused by ectopic expression of atRSZ33, which encodes an Arabidopsis SR protein (Kalyna et al., 2003). Spartz et
al. (2004) proposed that SMU-2 may act similarly to SR-related proteins, which also lack RNA binding motifs. In fact, we observed several features shared by Zm SMU2 and SR proteins. First, mutations in Zm smu2 are associated with altered splicing patterns in multiple pre-mRNAs, including those encoding SR proteins (Fig. 1). This effect has been observed in plant and animal cells over-expressing SR proteins (Mizzen et al., 1996; Jumaa and Nielsen, 1997; Lopato et al., 1999; Kalyna et al., 2003; Kumar and Lopez, 2005). In addition, expression of Arabidopsis SR protein genes appeared to be tissue-specific and particularly high in pollen and in actively dividing cells (Lopato et al., 1999; Lopato et al., 2002; Kalyna et al., 2003). This pattern of expression is very similar to that observed in At SMU2 and SMU1 gene expression (Figs. 5 and 7).

Our study on protein interactions with Zm SMU2 and At SMU2 suggests that plant SMU-2 homologues work in a protein complex to regulate not only pre-mRNA splicing but also other aspects of mRNA metabolism. No Zm SMU2-interacting proteins, except Zm SMU1, have human homologues that were previously identified as spliceosomal proteins or are known to function in pre-mRNA splicing (Table II). However, this does not mean that the observed protein interactions conflict with the proposed function of Zm SMU2 in pre-mRNA splicing. Rather, these interactions suggest that Zm SMU2 could also serve as a link to related processes in mRNA metabolism, such as transcription initiation and chromatin modification. Splicing factors, especially SR proteins, were shown to interact with the preinitiation complex, transcription elongation factors, 5’ and 3’ processing factors and the mRNA export complex (Maniatis and Reed, 2002). In this regard, it is not surprising that Zm SMU2-interacting proteins include proteins that are found in preinitiation complexes. For instance, a yeast two-hybrid screen for Zm SMU2-interacting proteins identified a partial HAF01 clone (Bertrand et al., 2005) showing a high degree of sequence identity to human TAFII250, the largest polypeptide in TFIID complex. Notably, animal homologues of TAFII250 can acetylate histone H4, another Zm SMU2-interacting protein (Mizzen et al., 1996), although it is still unclear whether histone H4 is a genuine in vivo substrate for the histone acetyltransferase activity of TAFII250. One intriguing possibility is that SMU2 may mediate coupling of pre-mRNA splicing with chromatin modification. For example, SMU2 could recruit a chromatin modification complex to an alternative exon. Alternatively, SMU2 could help recognize a specific histone code during pre-mRNA splicing and affect alternative splicing of transcripts from genes with a particular
In addition to interacting with a histone or a histone modification factor, Zm SMU2 might also interact with sequence-specific transcription factors in vivo. One of the Zm SMU2-interacting proteins identified in our Y2H screen encodes a protein containing a zinc-coordinating motif and a conserved basic region with high sequence similarity to the Arabidopsis sequence-specific transcription factors At VOZ1 and VOZ2 (Mitsuda et al., 2004). Taken together, protein interactions identified from the yeast two-hybrid screen suggest a role for Zm SMU2 in connecting pre-mRNA splicing and transcription initiation.

More recently, a biochemical role for animal and plant SMU-1 homologues has been proposed. The human SMU-1 homologue was identified from immuno-purified CUL4B-DDB1 ubiquitin ligase complexes (Higa et al., 2006), along with other proteins with a similar type of WD40 motif, or DWD (DDB1 binding WD40) motif. These DWD proteins were found to be specific substrate receptors for the corresponding ubiquitin ligase complex. In addition, a subset of Arabidopsis and rice proteins containing a DWD motif was shown to interact with DDB1 in vitro and in vivo (Lee et al., 2008), although an interaction for At SMU1 was not tested in this study. Considering the highly conserved interactions of animal and plant DWD proteins with DDB1, plant SMU-1 homologues are most likely involved in the recognition of specific target proteins for ubiquitination. If the proposed CUL4-DDB1-SMU1 ubiquitin ligase complex is involved in proteasome-dependent proteolysis, what can be the target? Probably it is not At SMU2, since At SMU2 is unstable in smu1-2 rather than stabilized (see Fig. 2C). Alternatively, the CUL4-DDB1-SMU1 complex may be involved in a non-proteolytic function. No matter what biochemical activity SMU-1 homologues have, their target proteins for ubiquitination are most likely to include spliceosomal proteins.

Comparison of the C. elegans smu-1 and smu-2 mutants, the At smu1 and smu2 mutants, and the Zm smu2 mutant indicates that SMU orthologues vary in terms of their dispensability. In the nematode, smu-1 and smu-2 null mutants are viable and do not show a highly-penetrant phenotype (Spike et al., 2001; Spartz et al., 2004), indicating both loci are dispensable. Genetic interaction between smu-1 and smu-2 was also found in C. elegans, where these two loci were originally identified from a genetic screen for recessive, loss-of-function suppressors of mec-8 unc-52 synthetic lethality (Lundquist and Herman, 1994). In the nematode, smu-1 and smu-2...
double mutants were phenotypically indistinguishable from the single mutants (Spartz et al., 2004). In Arabidopsis, *smu2* null mutants are viable, while a strong *smu1-1* allele is lethal, and a weak *smu1-2* allele has several developmental defects. In contrast, the maize *smu2* mutants, which are not null, show severe developmental defects and are sterile (Chung et al., 2007), indicating that Zm *smu2* is an essential gene. Three models could explain the variation among the three species. First, although At SMU2 and SMU-2 are encoded by single-copy genes, Arabidopsis and *C. elegans* genomes may encode genes that function in a partially redundant manner to the SMU-2 homologues. Second, although the amino acid sequences of SMU-1 and SMU-2 homologues are highly conserved, these proteins may have diverged to gain additional functions in various eukaryotic organisms. Third, the SMU-1 and SMU-2 homologues may have a similar biochemical role, but their target pre-mRNAs (and proteins if SMU-1 homologues are involved in ubiquitination of target proteins) may differ among various species.

In the last model, Zm SMU2-dependent pre-mRNA targets would be critical for viability, while At SMU2- and SMU-2-dependent targets may not be as critical for cellular functions to cause early developmental defects. Alternatively, Zm SMU2 may be required for processing of a larger set of pre-mRNA targets than At SMU2 and SMU-2, and this could result from differences in the respective genomes. For example, the average size of Arabidopsis, rice and maize introns is estimated to be 167 bp, 413 bp, and 607 bp, respectively (Haberer et al., 2005). This variation in average intron size is probably due to the frequency of transposon insertions in each genome (Haberer et al., 2005). The higher frequency of transposon-related sequences in maize introns could be responsible for the indispensability of Zm SMU2. Consistent with this, transcript profiling of Zm *smu2-1* endosperm revealed several transposon-like sequences that were differentially expressed in wild-type and *smu2-1* mutant endosperm (Supplemental Table S1). In this regard, it would be interesting to know if Zm *Smu2* is the endosperm splicing regulator that was proposed to be responsible for differential splicing of *waxy* pre-mRNA for the *wxG* allele, which contains a retrotransposon in an intron (Marillonnet and Wessler, 1997). There have been several reports in maize describing similar alternative splicing events that were triggered by insertion of different transposon types, and at least some of the alternative splicing events were found to be tissue-specific (Ortiz and Strommer, 1990; Marillonnet and Wessler, 1997; Lal and Hannah, 1999).
MATERIALS AND METHODS

Arabidopsis mutants

The T-DNA insertional Arabidopsis (Arabidopsis thaliana) mutants, smu2-1 (stock ID SALK_039202), smu2-2 (stock ID WiscDsLox320H09), smu1-1 (stock ID SALK_123852), smu1-2 (stock ID SALK_051163), and smu1-3 (stock ID SAIL_95_E04) were obtained from the Arabidopsis Biological Resource Center (Columbus, OH). The EMS mutant, sta1-1 (Lee et al., 2006), was obtained from Dr. Jian-Kang Zhu (University of California, Riverside, CA). smu2-1 and smu2-2 were crossed to the Col-0 ecotype twice and smu1-2 once before any phenotypic analyses were carried out. Homozygous mutant plants were identified using PCR analysis of genomic DNA isolated from individual plants (Edwards et al., 1991) with a combination of the following primers: for the At SMU2-1 allele, 2646Ex7_F_In7 (5’-ccggtgccccgaggtatgtatactt-3’) and AtMTO_R2 (5’-cttggctccaaactgaagaccag-3’); for At smu2-1, 2646Ex7_F_In7 and T-DNA_LB_b1 (5’-gcgtggaccgcttgctgcaact-3’); for At SMU2-2, AtMTO_F3 (5’-ccagtggacttgcttgctgaact-3’) and 2646ln7_R_Exp1 (5’-aagtatacatactgcaagcgg-3’); for At smu2-2, AtMTO_F3 and pDS-Lox_LB (5’-aacgtccgacagtctttaggctg-3’); for At SMU1-2, either AtSMU-1_F3 (5’-agagtttctcatcggtatagggaga-3’) and AtSMU-1_R1 (5’-tcaggttctcatctgtatgc-3’) or SMU-4f (5’-taatttccgacggtctcc -3’) and atsmu1-2r (5’-gacatggcaaagctgctgatac-3’); for At smu1-2, either AtSMU-1_F3 and T-DNA_LB_b1 or SMU-4f and LBag (5’-ttggcagtgctggcagctgc-3’); for STA1-1, primers AtPRP6_R2 (5’-ccactctctcagctgcgag-3’) and AtPRP6_F2 (5’-agaggggtcagagttg-3’); and for sta1-1, AtPRP6_R2 and AtPRP6_F2hc (5’-agaggggtcagagaaatat-3’).

Characterizations of mutant phenotypes

To measure seed weights, four to six plants of the F3 or F4 generation were self-pollinated. All of the F4 plants analyzed were obtained from one of the F3 plants. At least 200 F4 or F5 seeds harvested from each self-pollinated plant were weighed and counted to calculate average seed weights. Seeds were surface-sterilized, placed on solid media (0.5 X MS salt, 1% (w/v) sucrose and 0.7% (w/v) phytoagar), and stored at 4°C for three days. Percentage of germination failure or arrested growth, and the number of cotyledons was analyzed at 7 DAG.
**Reporter constructs and plant transformation**

For GFP, GUS-GFP and At SMU2-GFP fusion proteins, the cDNA for GFP was amplified by PCR using primers NcoI_gfpF (5’-ctcaccatggtgagcaagggcga-3’) and BamHI_gfpR (5’-taaggatcctacttgacgctgc-3’), and the DNA product was digested with NcoI and BamHI and introduced into pRTL2 (Carrington et al., 1990). pRTL2 contains two tandem repeats of the cauliflower mosaic virus 35S promoter and the tobacco etch virus leader sequence upstream of multiple cloning sites. The recombinant expression cassette was transferred into the HindIII site of the binary vector pBIN19 (Bevan, 1984), and the resulting plasmid, pBIN19-35Sp-GFP, was used to transform Arabidopsis Col-0 *gl1* using the floral dip method (Clough and Bent, 1998).

To create transgenic Arabidopsis plants expressing *GUS-GFP*, a GUS cDNA was amplified with primers NcoI_gusF (5’-tctaccatggtagatctgactag-3’) and gusRgfp (5’-ttgctcaccatgctagcttgtttgcctc-3’) and the cDNA for GFP was amplified with primers gusFgfp (5’-aatcaccatggtgagcaagggcga-3’) and BamHI_gfpR. A second PCR was performed using NcoI_gusF and BamHI_gfpR to create the fusion product, *GUS-GFP*. For a transgenic plant over-expressing At *SMU2-GFP*, the At *SMU2* cDNA was amplified with primers F+60BspHI (5’-ttctcatgaaaccttcaaaatcgcatc-3’) and 2646Rgfp (5’-ttgctcaccatgcttggatctcttag-3’) and the cDNA for GFP was amplified with primers 2646Fgfp (5’-gatccaagcatatggtgagcaagggcga-3’) and BamHI_gfpR. A second PCR was performed using F+60BspHI and BamHI_gfpR to make the fusion product, *AtSMU2-GFP*. The fusion PCR products, *GUS-GFP* and *AtSMU2-GFP*, were cloned into the NcoI and BamHI sites of pRTL2, and the expression cassettes were transferred into the HindIII site of a binary vector, pBIN19. The resulting plasmids, pBIN19-35Sp-GUS-GFP and pBIN19-35Sp-AtSMU2-GFP, were used to transform Arabidopsis *gl1*. Kanamycin-resistant seedlings were rescued, and images of GFP expressed in roots of the T2 plants at 5 DAG were collected on a BioRad MRC 1024 confocal laser scanning microscope and processed with NIH Image software (http://rsb.info.nih.gov/nih-image/) and Adobe Photoshop 6.0 or CS (San Jose, CA).

For the At *SMU1-GFP* translational fusion, a 7.4kb fragment containing 2.7kb of 5’-flanking sequence and the coding region of *SMU1* was amplified from Col-0 genomic DNA by Phusion polymerase (Finnzymes, Helsinki, Finland) with primer SUM1(-2700)Hind-f (5’-gctgaagcttggagcattgaggc-3’) and SUM1-BamStop-r (5’-gaaggatcgggctgccatactct-3’), restriction digested with *HindIII* and *BamHI*, and cloned into the pBN-GFP vector (Wang et al.,...
2006), resulting in construct pBN-SMU1-GFP (SMU1-GFP\textsuperscript{Kan}). The HindIII/EcoRI fragment containing SMU1-GFP and the nopaline synthase polyadenylation sequence (NOS-ter) was subcloned from pBN-SMU1-GFP into pGPTV-BAR vector (Becker et al., 1992), resulting in construct pGPTV-SMU1-GFP (SMU1-GFP\textsuperscript{Bar}).

For complementation of the At smu1-2 mutant, a 7.8kb fragment containing 2.7kb of 5’-flanking sequence, the coding region, and 433bp of 3’ flanking sequence of SMU1 was amplified from Col-0 genomic DNA by Phusion polymerase with primer SUM1(-2700)Hind-f (5’-gctgaagcttgggacattgaggcagc-3’) and SUM1-Bam-r (5’-ctgcggatccgttgtaatggtctctca -3’), restriction digested with HindIII and BamHI, and cloned into a modified pBI101 vector where the NOS-ter and GUS coding region were replaced with a synthetic linker (5’-gaattcgagctcggtacccggggatcc-3’), resulting in construct pBI-SMU1 (SMU1\textsuperscript{Kan}). The HindIII/EcoRI fragment containing SMU1 and NOS-ter was subcloned from pBI-SMU1 into pGPTV-BAR vector (Becker et al., 1992), resulting in construct pGPTV-SMU1 (SMU1\textsuperscript{Bar}).

The standard flower dip method (Clough and Bent, 1998) was used to transform Arabidopsis ecotype Columbia-0 or smu1-2/+ plants with Agrobacterium tumefaciens strain GV3101 pMP90 (Koncz and Schell, 1986) containing appropriate constructs. T1 transformants were selected on MS media (Gibco BRL Life Technologies, Gaithersburg, MD) containing 35 \mu g/mL kanamycin or 50 \mu g/mL gluphosinate ammonium (Crescent Chemical Company, Islandia, NY). Kanamycin or BASTA resistant plants were transferred to soil, and the presence of the transgene was confirmed by PCR.

Confocal images of the At SMU1-GFP plants were obtained as describe above. Bright-field and epifluorescence images were obtained using a Zeiss Axiophot compound epifluorescence microscope (Carl Zeiss, Jena, Germany) equipped with a GFP bandpass filter (exciter 450 to 490, dichroic 495, emitter 500-550; Chroma Technology, Inc., Brattleboro, VT), a fluorescein isothiocyanate (FITC) longpass filter (exciter 450 to 490, dichroic 510, emitter 515; Carl Zeiss), a DAPI longpass filter (exciter 365, dichroic 395, emitter 420; Carl Zeiss), and an Optronics MicroFire CCD camera (Goleta, CA).
Analysis of RNA and protein

Arabidopsis plants were grown under a 16/8 hour light/dark cycle condition at 20°C. RNA samples were prepared from whole Arabidopsis seedlings grown on solid media 14 DAG, and roots, leaves and flowers of Arabidopsis plants grown on soil at 30 DAG.

RNA was isolated using Trizol according to the manufacturer’s instructions (Invitrogen, Carlsbad, CA). Total RNA was used in a reverse transcription reaction with Superscript II (Invitrogen, Carlsbad, CA) and oligo-dT primers to generate first-strand cDNA which was subsequently used to amplify specific cDNAs using standard techniques. Primers for PCR were as follows: At ACT8, AtACT8_F (5’-taaactaaagacagatcgtttca-3’) and AtACT8_R (5’-ttttatcggattgaagaggg-3’); the 5’ coding sequence of At SMU2, AtSMU2_1 (5’-atgaaaccttcaaatcgttca-3’) and AtSMU2_2 (5’-agttcctcaccatc-3’); the middle region of At SMU2 coding sequence, AtSMU2_3 (5’-ccaggttagttgaagctcag-3’) and AtSMU2_4 (5’-ccacaagaattcattc-3’); the 3’ coding sequence of At SMU2, AtSMU2_5 (5’-gagagatagggttcgta-3’) and AtSMU2_6 (5’-gaatcgttagtctcttagggtc-3’); At SRp30, AtSRp30_F1 (5’-atgagagagatggtaaatgga-3’) and AtSRp30_R1 (5’-cgtttgatctcctcagatcagatc-3’); At TRA2, AtTRA2_F2 (5’-attgccttactcaagaagttcag-3’) and AtTRA2_R3 (5’-cgagggagctcagagagact-3’); and At SMU1, either AtSMU1_1 (5’-atgaaaccttcatcggatg-3’) and AtSMU1_2 (5’-ccaggtttaaccatcagatc-3’) or AtSMU1_3 (5’-caggttagttcctcggatg-3’) and AtSMU1_4 (5’-ccaggttgaagttcctgcagtcttc-3’).

For protein extraction, tissues were homogenized in a 2X Laemmli buffer. The extract was centrifuged and the supernatant was analyzed by SDS-PAGE. Western blot analysis was conducted as described in Chung et al. (2007).

RNA and protein analyses using maize (Zea mays) endosperm extracts were performed as described (Chung et al., 2007). Primers for PCR were as follows: rpl15, ZmrpL15_F1 (5’-agccgcatcgcctggttgcctggta-3’) and ZmrpL15_R2 (5’-ctacgttagcgcgctctctc-3’); rpl7a, ZmrpL7a_F2 (5’-ccaggtttaaccatcgcgtc-3’) and ZmrpL7a_R3 (5’-ttgtgatctcctcgggttcgta-3’); eEf-1α, ZmEF1a_F2 (5’-ctcccccctagctctcag-3’) and ZmEF1a_R3 (5’-gcaacacaggagccgacac-3’); O2 SV1 (major transcript), O2_F1 (5’-atgagacggtctcctcagagag-3’) and O2_R3 (5’-tgattcctttttactacc-3’); O2 SV2, O2_F1 and O2_R2a (5’-tctgcgattctctcctcagc-3’); tra2, ZmTRA2_F1 (5’-aggtcctacaagggctcag-3’), ZmTRA2_R3 (5’-gactctccaggtggggtgcctc-3’) and ZmTRA2_R2a (5’-caaccctttgtgagttggcat-3’).
3’); *rsp31b*, ZmRSp31B_F2 (5’-aaccctaggtgccgctcctttttc-3’), ZmRSp31B_R3 (5’-atggaaggatgtgaagacccgtgg-3’) and ZmRSp31B_R5 (5’-catacgtctaccagctgtgctc-3’); and *otux*, ZmOTUX1_F2 (5’-ctatgatggatacgttcccatggc-3’) and ZmOTUX1_R1 (5’-aagagagctgaaccggccatcaat-3’).

All RT-PCR and immunoblot analyses were performed with at least three biological replicates. Representative gel images are shown in the figures.

**Detection of Zm SMU2-interacting proteins with the yeast two-hybrid system**

A full-length Zm *Smu2* cDNA bait sequence for the yeast two-hybrid screen was prepared by PCR using primers E1 (5’-ggaattcatgtcatcgaagaagaactataag-3’) and S1 (5’-tcccccggggatcagccacgctgtttcttcagct-3’) for insertion between the EcoRI and SmaI sites of the pBD-GAL4 Cam phagemid vector (Stratagene, La Jolla, CA). Because this construct self-activated in the yeast two-hybrid system, a series of deletion clones encoding portions of the Zm SMU2 protein were made as follows: the ZmSMU2(I) construct (nucleotides 148 to 1698 from the ATG) was created by PCR with primers 38YCONST-I (5’-ggaattcatgtcgtttcatgcagtgg-3’ and S1, and the amplified DNA fragment was cloned into pBD-GAL4 between the EcoRI and SmaI sites; the ZmSMU2(II) construct was created by PCR using primers 38YCONST-II (5’-ggaattcatgaagaagatcgacggctg-3’) and S1, and the amplified DNA was cloned into pBD-GAL4, as with ZmSMU2(I); the ZmSMU2(III) construct was created by PCR with primers 38YCONST-III (5’-ggaattcatgcccaccacccggctcc-3’) and S1, and the amplified DNA was cloned into pBD-GAL4, as with ZmSMU2(I); the ZmSMU2(IV) construct was created by PCR with primers 38YCONST-IV (5’-ggaattcatgggttatccagaagctg-3’) and S1, and the amplified DNA was cloned into pBD-GAL4, as with ZmSMU2(I); the ZmSMU2(V) construct was created by PCR with primers E1 and R4 (5’-ctctggtcatgaaggctgccg-3’) and the amplified DNA was digested with EcoRI and EcoRV and inserted into pBD-GAL4, as with ZmSMU2(I). Each construct was verified using DNA sequencing. The pBD-GAL4-ZmSMU2(IV) construct produced a detectable X-gal product in two and a half hours, and we monitored the yeast colonies for a color reaction that developed within 30 minutes. On the other hand, deletion of the COOH-terminal 207 amino acids of the Zm SMU2 protein dramatically reduced self-activation in the yeast two-hybrid system. Consequently, we used the ZmSMU2(V) bait clone to screen maize and Arabidopsis target libraries. Developing maize endosperm (10-14 days after
pollination) and immature ear libraries (HybriZAP®-2.1 Two-Hybrid Predigested Vector Kit, Stratagene, La Jolla, CA) were obtained from Dr. Bob Schmidt (Dept. of Biology, UC San Diego). The Arabidopsis library was from the Arabidopsis Biological Resource Center (http://www.arabidopsis.org/abrc/libraries.jsp, Stock # CD4-22). Yeast transformation and screening procedures were performed according to manufacturer’s instructions. To identify the interacting proteins encoded by Arabidopsis target DNAs, colony PCR reactions were carried out with primers pACTForward (5’-ctatctatcgaagatgaag-3’) and pACTReverse (5’-acagttgaagtgacctagtgcc-3’); for maize target genes, the pACTForward and PADT7 (5’-taatacgactcactataagg-3’) primers were used. The amplified PCR product was inserted into the pCR4-TOPO vector (Invitrogen, Carlsbad, CA) for DNA sequencing.

Expression of recombinant proteins in E. coli and in vitro pull-down assays

The following PCR primers were used to produce target proteins fused with a 6XHis tag: for histone H4, AtHIS4_F1 (5’-atgctcaggaagaaatagggaa-3’) and AtHIS4_R1 (5’-taaatcctcgaacacctaggg-3’); for At SMU1, AtSMU-1_F1 (5’-atggtgcctggaatctgctc-3’) and AtSMU-1_R1 (5’-tcaggcccttccataacttcatag-3’); and for At MEC8, AtMEC-8_F1 (5’-atggcgtatcaccaaccgtacga-3’) and AtMEC-8_R1 (5’-ttaccktatcacaacccgta-3’) and AtMEC-8_R1 (5’-ttaccktatcacaacctcgctc-3’). The RT-PCR products were cloned into pGEMT-easy (Promega, Madison, WI) and subsequently transferred into the EcoRI site of pRSET (Invitrogen, Carlsbad, CA) and used to transform the BL21(DE3) strain of E. coli. For pull-down experiments where GST fusion proteins were used as a bait protein, cell lysates containing the fused target proteins with a 6XHis tag were prepared in lysis buffer (50 mM Tris, pH 8.5, 150 mM NaCl, 1 mM dithiothreitol, 1 mM EDTA, 0.2 mM phenylmethanesulphonylfluoride) containing 1% (v/v) Triton X-100. Purification of GST and GST-ZmSMU2 was performed as described in Chung et al. (2007). Cell lysates containing 0.1-0.5 μg of target proteins were pre-incubated with 0.5 μg of GST and 50 μL of a 50% (v/v) glutathione-agarose slurry (Sigma, St. Louis, MO) at 4 °C for an hour. The cleared cell lysates were added to 50 μL of a 50% (v/v) glutathione-agarose slurry and 10 μg of GST, GST-AtSMU2, or GST-ZmSMU2 proteins. The mixtures were adjusted to 1 mL with lysis buffer and incubated at 4 °C for two hours. The beads were then washed four times with lysis buffer containing 0.1% (v/v) Triton X-100. Proteins were eluted in 1X Laemmli sample buffer containing 10 mM glutathione, separated by SDS-PAGE, and transferred onto nitrocellulose membrane.
Immunoblot analysis was performed with mouse HisG antibodies (Sigma, St. Louis, MO) at 1:3,000 dilution, and then with goat anti-mouse IgG antibodies conjugated with horseradish peroxidase (Sigma, St. Louis, MO) at 1:80,000 dilution.

**Supplemental Material**

The following materials are available in the online version of this article.

**Supplemental Figure S1.** Microarray data analysis of Zm smu2-1 endosperm.

**Supplemental Figure S2.** Semi-quantitative RT-PCR analysis of RNAs corresponding to co-regulated and consistent SP classes of DHPs identified in the microarray experiment.

**Supplemental Table S1.** List of DHP-Hs and DHP-Ls (Microsoft Excel file)

**Supplemental Table S2.** Correlation of random probes and DHP-L with the number of SVs in MAGI

**Supplemental Table S3.** Identification of hybridization sites for random probes and DHP-L, according to the \textit{ab initio} prediction-based exons and introns in MAGI

**Supplemental Table S4.** Identification of hybridization sites for random probes and DHP-L, according to EST-based gene annotation in MAGI

**Supplemental Protocol S1.** Results and Methods for Zm smu2-1 microarray

**ACKNOWLEDGMENT**

We thank colleagues in the Larkins lab for review and editing of the manuscript.
Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen HM, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Nubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseuw E, Brogden D, Zeke A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science 301: 653-657

Amrein H, Gorman M, Nothiger R (1988) The sex-determining gene Tra-2 of Drosophila encodes a putative RNA-binding protein. Cell 55: 1025-1035

Becker D, Kemper E, Schell J, Masterson R (1992) New plant binary vectors with selectable markers located proximal to the left T-DNA border. Plant Mol Biol 20: 1195-1197

Bell LR, Horabin JI, Schedl P, Cline TW (1991) Positive autoregulation of Sex-lethal by alternative splicing maintains the female-determined state in Drosophila. Cell 65: 229-239

Bertrand C, Benhamed M, Li YF, Ayadi M, Lemonnier G, Renou JP, Delarue M, Zhou DX (2005) Arabidopsis HAF2 gene encoding TATA-binding protein (TBP)-associated factor TAF1, is required to integrate light signals to regulate gene expression and growth. J Biol Chem 280: 1465-1473

Bessonov S, Anokhina M, Will CL, Urlaub H, Luhrmann R (2008) Isolation of an active step I spliceosome and composition of its RNP core. Nature 452: 846-U843

Bevan M (1984) Binary agrobacterium vectors for plant transformation. Nucleic Acids Res 12: 8711-8721

Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. Ann Rev Biochem 72: 291-336

Brown JWS, Simpson CG (1998) Splice site selection in plant pre-mRNA splicing. Ann Rev Plant Physiol Plant Mol Biol 49: 77-95

Carrington JC, Freed DD, Oh CS (1990) Expression of potyviral polyproteins in transgenic plants reveals 3 proteolytic activities required for complete processing. EMBO J 9: 1347-1353

Chung T, Kim CS, Nguyen HN, Meeley RB, Larkins BA (2007) The maize Zmsmu2 gene encodes a putative RNA-splicing factor that affects protein synthesis and RNA processing during endosperm development. Plant Physiol 144: 821-835

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16: 735-743

Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res 19: 1349-1349

Graveley BR (2000) Sorting out the complexity of SR protein functions. RNA 6: 1197-1211

Gupta S, Wang BB, Stryker GA, Zanetti ME, Lal SK (2005) Two novel arginine/serine (SR) proteins in maize are differentially spliced and utilize non-canonical splice sites. Biochim Biophys Acta 1728: 105-114

Haberer G, Young S, Bhati AK, Gundlach H, Raymond C, Fuks G, Butler E, Wing RA, Rounsley S, Birren B, Nusbaum C, Mayer KFX, Messing J (2005) Structure and architecture of the maize genome. Plant Physiol 139: 1612-1624
Hartmuth K, Urlaub H, Vornlocher HP, Will CL, Gentzel M, Wilm M, Luhrmann R (2002) Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. Proc Natl Acad Sci USA 99: 16719-16724

Higa LA, Wu M, Ye T, Kobayashi R, Sun H, Zhang H (2006) CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. Nat Cell Biol 8: 1277-U1248

Isshiki M, Tsumoto A, Shimamoto K (2006) The serine/arginine-rich protein family in rice plays important roles in constitutive and alternative splicing of pre-mRNA. Plant Cell 18: 146-158

Jumaa H, Nielsen PJ (1997) The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. EMBO J 16: 5077-5085

Jurica MS, Licklider LJ, Gygi SP, Grigorieff N, Moore MJ (2002) Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. RNA 8: 426-439

Jurica MS, Moore MJ (2003) Pre-mRNA splicing: Awash in a sea of proteins. Mol Cell 12: 5-14

Kalyna M, Lopato S, Barta A (2003) Ectopic expression of atRSZ33 reveals its function in splicing and causes pleiotropic changes in development. Mol Biol Cell 14: 3565-3577

Koncz C, Schell J (1986) The promoter of T1-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of Agrobacterium binary vector. Mol Genet Genomics 204: 383-396

Kumar S, Lopez AJ (2005) Negative feedback regulation among SR splicing factors encoded by Rbp1 and Rbp1-like in Drosophila. EMBO J 24: 2646-2655

Lal SK, Hannah LC (1999) Maize transposable element Ds is differentially spliced from primary transcripts in endosperm and suspension cells. Biochem Biophys Res Comm 261: 798-801

Lee BH, Kapoor A, Zhu JH, Zhu JK (2006) STABILIZED1, a stress-upregulated nuclear protein, is required for pre-mRNA splicing, mRNA turnover, and stress tolerance in Arabidopsis. Plant Cell 18: 1736-1749

Lee JH, Terzaghi W, Gusmaroli G, Charron JBF, Yoon HJ, Chen HD, He YJ, Xiong Y, Deng XW (2008) Characterization of Arabidopsis and rice DWD proteins and their roles as substrate receptors for CUL4-RING E3 ubiquitin ligases. Plant Cell 20: 152-167

Lopato S, Forstner C, Kalyna M, Hilscher J, Langhammer U, Indrapichate K, Lorkovic ZJ, Barta A (2002) Network of interactions of a novel plant-specific Arg/Ser-rich protein, atRSZ33, with atSC35-like splicing factors. J Biol Chem 277: 39989-39998

Lopato S, Kalyna M, Dorner S, Kobayashi R, Krainer AR, Barta A (1999) atSRp30, one of two SF2/ASF-like proteins from Arabidopsis thaliana, regulates splicing of specific plant genes. Genes Dev 13: 987-1001

Lundquist EA, Herman RK (1994) The Mec-8 Gene of Caenorhabditis-Elegans Affects Muscle and Sensory Neuron Function and Interacts with 3 Other Genes - Unc-52, Smu-1 and Smu-2. Genetics 138: 83-101

Lundquist EA, Herman RK, Rogalski TM, Mullen GP, Moerman DG, Shaw JE (1996) The mec-8 gene of C-elegans encodes a protein with two RNA recognition motifs and regulates alternative splicing of unc-52 transcripts. Development 122: 1601-1610
Makarov EM, Makarova OV, Urlaub H, Gentzel M, Will CL, Wilm M, Luhrmann R (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. Science 298: 2205-2208

Makarova KS, Aravind L, Koonin EV (2000) A novel superfamily of predicted cysteine proteases from eukaryotes, viruses and Chlamydia pneumoniae. Trends Biochem Sci 25: 50-52

Maniatis T, Reed R (2002) An extensive network of coupling among gene expression machines. Nature 416: 499-506

Marillonnet S, Wessler SR (1997) Retrotransposon insertion into the maize waxy gene results in tissue-specific RNA processing. Plant Cell 9: 967-978

Mitsuda N, Hisabori T, Takeyasu K, Sato MH (2004) VOZ: Isolation and characterization of novel vascular plant transcription factors with a one-zinc finger from Arabidopsis thaliana. Plant Cell Physiol 45: 845-854

Mizzen CA, Yang XJ, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, Workman J, Wang L, Berger SL, Kouzarides T, Nakatani Y, Allis CD (1996) The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. Cell 87: 1261-1270

Neubauer G, King A, Rappsilber J, Calvio C, Watson M, Ajuh P, Sleeman J, Lamond A, Mann M (1998) Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. Nat Genet 20: 46-50

Ortiz DF, Strommer JN (1990) The Mul Maize Transposable Element Induces Tissue-Specific Aberrant Splicing and Polyadenylation in 2 Adh1 Mutants. Mol Cell Biol 10: 2090-2095

Pazman C, Mayes CA, Fanto M, Haynes SR, Mlodzik M (2000) Rasputin, the Drosophila homologue of the RasGAP SH3 binding protein, functions in Ras- and Rho-mediated signaling. Development 127: 1715-1725

Rappsilber J, Ryder U, Lamond AI, Mann M (2002) Large-scale proteomic analysis of the human spliceosome. Genome Res 12: 1231-1245

Reddy ASN (2001) Nuclear pre-mRNA splicing in plants. Crit. Rev Plant Sci 20: 523-571

Reddy ASN (2004) Plant serine/arginine-rich proteins and their role in pre-mRNA splicing. Trends Plant Sci 9: 541-547

Smith CWJ, Valcarcel J (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem Sci 25: 381-388

Spartz AK, Herman RK, Shaw JE (2004) SMU-2 and SMU-1, Caenorhabditis elegans homologs of mammalian spliceosome-associated proteins RED and fSAP57, work together to affect splice site choice. Mol Cell Biol 24: 6811-6823

Spike CA, Shaw JE, Herman RK (2001) Analysis of smu-1, a gene that regulates the alternative splicing of unc-52 pre-mRNA in Caenorhabditis elegans. Mol Cell Biol 21: 4985-4995

Wang BB, Brendel V (2006) Genomewide comparative analysis of alternative splicing in plants. Proc Natl Acad Sci USA 103: 7175-7180

Wang D, Tyson MD, Jackson SS, Yadegari R (2006) Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis. Proc Natl Acad Sci USA 103: 13244-13249

Wollerton MC, Gooding C, Wagner EJ, Garcia-Blanco MA, Smith CWJ (2004) Autoregulation of polypyrimidine tract binding protein by alternative splicing leading to nonsense-mediated decay. Mol Cell 13: 91-100
Xiong LM, Gong ZZ, Rock CD, Subramanian S, Guo Y, Xu WY, Galbraith D, Zhu JK (2001) Modulation of abscisic acid signal transduction and biosynthesis by an Sm-like protein in Arabidopsis. Dev Cell 1: 771-781

Zhou ZL, Licklider LJ, Gygi SP, Reed R (2002) Comprehensive proteomic analysis of the human spliceosome. Nature 419: 182-185
FIGURE LEGEND

Figure 1. Altered RNA splicing in Zm smu2-1 and smu2-3 endosperm. Heterozygous (smu2-1/+ or smu2-3+/+) plants were self-pollinated and 16-days-after-pollination kernels dissected into embryo and endosperm tissues. Embryo DNA was used to genotype the Zm smu2 locus, and RNA was extracted from homozygous wild-type, heterozygous, and homozygous mutant endosperms and used for semi-quantitative RT-PCR. +/+, endosperm with a homozygous wild-type embryo; +/-, endosperm with a heterozygous embryo; -/-, endosperm with a homozygous smu2-1 mutant embryo. Solid arrowheads indicate splice variants (SVs) with different band intensities between the wild-type and mutants. Pre-mRNA splicing patterns of SVs are shown at right, with RT-PCR primers indicated by arrows. Ratios of differentially accumulated SVs are shown below the panels.

Figure 2. (A) Structure of the At SMU2 gene. Dark gray boxes and light gray bars between the boxes represent exons and introns, respectively. Solid triangles indicate T-DNA insertion sites in the At smu2-1 and smu2-2 mutants. (B) RT-PCR analysis of At SMU2 transcripts. Four pairs of primers were used for amplifying different regions of the At SMU2 mRNA sequence prepared from wild type 14-DAG Col-0 (lane 1), smu2-1 (lane 2) and smu2-2 (lane 3) seedlings. Numbers in parenthesis indicate the primers used for RT-PCR, and their positions are shown as arrows in (A). At ACT8 transcript encoding an actin was used as an internal control. (C) Immunoblot analysis of At SMU2 using protein extracts from four 7-DAG Col-0 (lane 1), smu2-1 (lane 2), smu2-2 (lane 3), smu1-2 (lane 4) and sta1-1 (lane 5) seedlings as described by Chung et al., 2007. The arrowhead indicates the At SMU2 protein. (D) Phenotypic analyses of various mutants. Homozygous mutant plants (smu2-1, smu2-2, smu1-2 and sta1-1) were identified among segregating populations, and were subsequently self-pollinated two or three times. Phenotypic analysis was conducted as described in Methods. (E) Dry sta1-1 (left), smu2-1 sta1-1 double-mutant (dwarf plant in the middle) and smu2-1 (right) plants.

Figure 3. RT-PCR detection of SVs in At smu2-1 and smu2-2. RT-PCR was performed using total RNA from 14-DAG Col-0 (lanes 1), smu2-1 (lane 2) and smu2-2 (lane 3) seedlings. Arrowheads indicate the positions of known SVs of At1g09140 (At SRp30) and At1g07350 (At
TRA2), while asterisks correspond to putative SVs first observed here. Pre-mRNA splicing patterns are shown on the right side, with dark gray boxes and light gray bars between the boxes representing exons and introns, respectively. Arrows depict primers used in each reaction. Ratios of differentially accumulated SVs are shown below the panels.

**Figure 4.** Subcellular localization of At SMU2 fusion protein. (A) Arabidopsis root cells overexpressing GFP showed a diffused localization pattern in the cytosol and the nuclei. (B) GUS-GFP in Arabidopsis root cells was exclusively localized in the cytosol. (C) Arabidopsis root cells overexpressing At SMU2-GFP exhibited nuclear-specific localization.

**Figure 5.** Immunoblot analysis of At SMU2 showing accumulation of At SMU2 protein in various organs and structures of Arabidopsis. Each lane contained approximately 10 μg of proteins from reproductive and vegetative organs of Col-0 plants. Lane 1, 7-DAG seedlings; lane 2, roots of 30-DAG plants; lane 3, senescing rosette leaves; lane 4, young cauline leaves; lane 5, inflorescence; lane 6, immature flowers; lane 7, open flowers; lane 8, developing siliques (less than 5 mm in length); lane 9, dry seeds.

**Figure 6.** Interactions of Zm and At SMU2 with other proteins. (A) Yeast two-hybrid (Y2H) bait constructs were tested for self-activation and the corresponding beta-GAL assays. A construct containing the full-length Zm SMU2 protein (pBD-GAL4-ZmSMU2, top) exhibited strong self-activation of the beta-GAL reporter gene when introduced into Y190 cells. N-terminal deletions of Zm SMU2 (bait constructs I-IV) also showed self-activation, albeit at lower levels than for the full-length protein. Only a C-terminal deletion (bait construct V) and a negative control (pBD-GAL4, bottom) did not activate the reporter gene. Black bars represent the Zm SMU2 polypeptides. Numbers above the bars indicate amino acid residues corresponding to the full-length Zm SMU2 polypeptide. (B) Confirmation of yeast two-hybrid screen by *in vitro* pull-down assay using GST or GST fusion proteins as a bait. GST (lane 2), GST-AtSMU2 (lane 3) or GST-ZmSMU2 (lane 4) bound to glutathione-agarose beads were used to pull down 6XHis-histone H4, 6XHis-AtSMU1 or 6XHis-AtMEC8. Lane 1, 10% input; lanes 2, 3 and 4, 30% pull-down by GST, GST-AtSMU2 and GST-ZmSMU2, respectively.
**Figure 7.** Expression of SMU1-GFP<sub>kan</sub> during plant development. (A-C) Dark-field, epifluorescence GFP, and merged images of a Kan-resistant seedling. The arrow points to a newly emerged leaf. The arrowhead points to the root tip. (D) Confocal image of the area indicated by the arrow in panel B. (E) Confocal image of the area indicated by the arrowhead in panel B. (F-H) Bright-field, epifluorescence DAPI and GFP images of a root hair. (I-J) Bright-field and epifluorescence GFP images of a flower stage13 ovule. (K) Epifluorescence GFP image of mature pollens. (L-N) Bright-field, epifluorescence GFP, and FITC images of a heart stage embryo. ccn, central cell nucleus; en, egg cell nucleus; gn, generative nuclei; rn, root hair nucleus; sn, synergid nuclei; vn, vegetative nucleus. Bars in A-C = 1mm; Bars in D-N = 50μm.

**Figure 8.** Analysis of At smu1 mutant alleles. (A) Structure of the At SMU1 gene and the T-DNA insertions in the At smu1-1 and smu1-2 mutants (solid reverse triangles). (B) Phenotypes of Columbia wild type, At smu2-1 (F4) and smu1-2 (F3) plants. Seedlings of smu1-2 at 7 DAG often showed an abnormal number of cotyledons (left panels; see Table I) and arrangement of flowers/siliques in the primary inflorescence (right panels). (C) RT-PCR comparison of At SMU2, SMU1, SRp30 and TRA2 transcripts in Col-0 and smu1-2. Two pairs of primers were used to amplify different regions of the At SMU1 cDNA prepared from 14-DAG seedlings of Col-0 (lane 1) and smu1-2 (lane 2) plants. Numbers in parenthesis indicate the primers used for the RT-PCR; the positions of the primers are shown as arrows in (A). Arrowheads indicate the positions of known SVs of SRp30 and TRA2, while asterisks correspond to the potentially novel SVs. Ratios of differentially accumulated SVs are shown below the panels.
Table I. Arabidopsis mutant phenotypes in seedling development

| Genotypes         | Generation | \(^{a}\)Abnormal cotyledon number (%) | \(^{a}\)Germination failure or arrested growth (%) |
|-------------------|------------|---------------------------------------|--------------------------------------------------|
| Col-0             | F4         | 0 ± 0                                 | 0.2 ± 0.3                                         |
|                   | F5         | 0 ± 0                                 | 0.1 ± 0.2                                         |
| smu2-1            | F4         | 0.4 ± 0.5                             | 0.8 ± 1.0                                         |
|                   | F5         | 0 ± 0                                 | 0.1 ± 0.3                                         |
| smu2-2            | F4         | 0.3 ± 0.4                             | 0.6 ± 0.4                                         |
|                   | F5         | 0.2 ± 0.3                             | 0.8 ± 0.7                                         |
| smu1-2            | F4         | 45.8 ± 4.8                            | 13.9 ± 8.6                                        |
|                   | F5         | 10.0 ± 2.0                            | 10.5 ± 8.8                                        |
| \(SMU1^{bar} /_{-}; smu1-2\) | T2         | 0.4 ± 0.5                             | 0 ± 0                                             |
| \(SMU1-GFP^{bar} /_{-}; smu1-2\) | T2         | 2.3 ± 1.4                             | 0.3 ± 0.8                                         |
| \(sta1-1\)        | F4         | 2.8 ± 1.0                             | 0.5 ± 0.7                                         |
|                   | F5         | 0.3 ± 0.3                             | 0.2 ± 0.3                                         |

Values are presented as mean ± s.d.

\(^{a}\) Phenotypes scored as described in Materials and Methods.
Table II. Summary of a yeast two-hybrid screen for proteins interacting with Zm SMU2

| clone ID | TopBlastn   | % id | Predicted protein                                      |
|---------|-------------|------|--------------------------------------------------------|
| Zm6-1   | DV495718.1  | 100  | rasputin-like protein (Pazman et al., 2000)            |
| Zm8-1   | BU051217.1  | 98   | Histone H4                                              |
| Zm17-3  | DN226990.1  | 97   | SMU-1-like protein (Spike et al., 2001)                 |
| Zm31-7  | DN218694.1  | 96   | VOZ-like protein (Mitsuda et al., 2004)                 |
| Zm33-8  | CA404370.1  | 97   | EF-1beta                                               |
| Zm59-20 | DT654009.1  | 96   | Glycerol kinase-like protein                            |
| At1-1   | NM_103009.3 | 97   | TAFII250-like HAF01, partial (Bertrand et al., 2005)   |
| At2-3   | NM_118947.2 | 98   | TPR protein, partial                                    |
| At3-3   | NM_117736.2 | 99   | DUF751-containing protein                               |

*a* GenBank accessions with the highest nucleotide sequence similarity to the Y2H clone shown in the left column.

*b* The percentage of nucleotide sequence identity based on BLASTN analysis.

*c* Based on the BLASTX result for the TopBlastn sequences.
Figure 1. Altered RNA splicing in Zm smu2-1 and smu2-3 endosperm. Heterozygous (smu2-1/+ or smu2-3/+) plants were self-pollinated and 16-days-after-pollination kernels dissected into embryo and endosperm tissues. Embryo DNA was used to genotype the Zm smu2 locus, and RNA was extracted from homozygous wild-type, heterozygous, and homozygous mutant endosperms and used for semi-quantitative RT-PCR. +/-, endosperm with a homozygous wild-type embryo; +/-, endosperm with a heterozygous embryo; -/-, endosperm with a homozygous smu2-1 mutant embryo. Solid arrowheads indicate splice variants (SVs) with no predicted intergenic regions between the wild-type and mutants. Pre-mRNA splicing patterns of SVs are shown at right, with RT-PCR primers indicated by arrows. Ratios of differentially accumulated SVs are shown below the panels.
Figure 2. (A) Structure of the At SMU2 gene. Dark gray boxes and light gray bars between the boxes represent exons and introns, respectively. Solid triangles indicate T-DNA insertion sites in the At smu2-1 and smu2-2 mutants. (B) RT-PCR analysis of At SMU2 transcripts. Four pairs of primers were used for amplifying different regions of the At SMU2 mRNA sequence prepared from wild type 14-DAG Col-0 (lane 1), smu2-1 (lane 2) and smu2-2 (lane 3) seedlings. Numbers in parenthesis indicate the primers used for RT-PCR, and their positions are shown as arrows in (A). At ACT8 transcript encoding an actin was used as an internal control. (C) Immunoblot analysis of At SMU2 using protein extracts from four 7-DAG Col-0 (lane 1), smu2-1 (lane 2), smu2-2 (lane 3), smu1-2 (lane 4) and sta1-1 (lane 5) seedlings as described by Chung et al., 2007. The arrowhead indicates the At SMU2 protein. (D) Phenotypic analyses of various mutants. Homozygous mutant plants (smu2-1, smu2-2, smu1-2, and sta1-1) were identified among segregating populations, and were subsequently self-fertilized three times. Phenotypic analysis was conducted as described in Methods. (E) Dry sta1-1 (left), smu2-1 sta1-1 double-mutant (dwarf plant in the middle) and smu2-1 (right) plants.
Figure 3. RT-PCR detection of SVs in At smu2-1 and smu2-2. RT-PCR was performed using total RNA from 14-DAG Col-0 (lanes 1), smu2-1 (lane 2) and smu2-2 (lane 3) seedlings. Arrowheads indicate the positions of known SVs of At1g09140 (At SRp30) and At1g07350 (At TRA2), while asterisks correspond to putative SVs first observed here. Pre-mRNA splicing patterns are shown on the right side, with dark gray boxes and light gray bars between the boxes representing exons and introns, respectively. Arrows depict primers used in each reaction. Ratios of differentially accumulated SVs are shown below the panels.
Figure 4. Subcellular localization of At SMU2 fusion protein. (A) Arabidopsis root cells overexpressing GFP showed a diffused localization pattern in the cytosol and the nuclei. (B) GUS-GFP in Arabidopsis root cells was exclusively localized in the cytosol. (C) Arabidopsis root cells overexpressing At SMU2-GFP exhibited nuclear-specific localization.
Figure 5. Immunoblot analysis of At SMU2 showing accumulation of At SMU2 protein in various organs and structures of Arabidopsis. Each lane contained approximately 10 μg of proteins from reproductive and vegetative organs of Col-0 plants. Lane 1, 7-DAG seedlings; lane 2, roots of 30-DAG plants; lane 3, senescing rosette leaves; lane 4, young cauline leaves; lane 5, inflorescence; lane 6, immature flowers; lane 7, open flowers; lane 8, developing siliques (less than 5 mm in length); lane 9, dry seeds.
Figure 6. Interactions of Zm and At SMU2 with other proteins. (A) Yeast two-hybrid (Y2H) bait constructs were tested for self-activation and the corresponding beta-GAL assays. A construct containing the full-length Zm SMU2 protein (pBD-GAL4-ZmSMU2, top) exhibited strong self-activation of the beta-GAL reporter gene when introduced into Y190 cells. N-terminal deletions of Zm SMU2 (bait constructs I-IV) also showed self-activation, albeit at lower levels than for the full-length protein. Only a C-terminal deletion (bait construct V) and a negative control (pBD-GAL4, bottom) did not activate the reporter gene. Black bars represent the Zm SMU2 polypeptides. Numbers above the bars indicate amino acid residues corresponding to the full-length Zm SMU2 polypeptide. (B) Confirmation of yeast two-hybrid screen by in vitro pull-down assay using GST or GST fusion proteins as a bait. GST (lane 2), GST-AtSMU2 (lane 3) or GST-ZmSMU2 (lane 4) bound to glutathione-agarose beads were used to pull down 6XHis-Histone H4, 6XHis-AtSMU1 or 6XHis-AtMEC8. Lane 1, 10% input; lanes 2, 3 and 4, 30% pull-down by GST, GST-AtSMU2 and GST-ZmSMU2, respectively.
Figure 7. Expression of SMU1-GFP<sup>kan</sup> during plant development.
(A-C) Dark-field, epifluorescence GFP, and merged images of a Kan-resistant seedling. The arrow points to a newly emerged leaf. The arrowhead points to the root tip. (D) Confocal image of the area indicated by the arrow in panel B. (E) Confocal image of the area indicated by the arrowhead in panel B. (F-H) Bright-field, epifluorescence DAPI and GFP images of a root hair. (I-J) Bright-field and epifluorescence GFP images of a flower stage13 ovule. (K) Epifluorescence GFP image of mature pollens. (L-N) Bright-field, epifluorescence GFP, and FITC images of a heart stage embryo. ccn, central cell nucleus; en, egg cell nucleus; gn, generative nucleus; sn, synergid nuclei; vn, vegetative nucleus. Bars in A-C = 1mm; Bars in D-N = 50μm.
Figure 8. Analysis of At smu1 mutant alleles. (A) Structure of the At SMU1 gene and the T-DNA insertions in the At smu1-1 and smu1-2 mutants (solid reverse triangles). (B) Phenotypes of At smu2-1 (F4) and smu1-2 (F3) plants. Seedlings of smu1-2 at 7 DAG often showed an abnormal number of cotyledons (left panels; see Table 1) and arrangement of flowers/siliques in the primary inflorescence (right panels). (C) RT-PCR comparison of At SMU2, SMU1, SRp30 and TRA2 transcripts in Col-0 and smu1-2. Two pairs of primers were used to amplify different regions of the At SMU1 cDNA prepared from 14-DAG seedlings of Col-0 (lane 1) and smu1-2 (lane 2) plants. Numbers in parenthesis indicate the primers used for the RT-PCR; the positions of the primers are shown as arrows in (A). Arrowheads indicate the positions of known SVs of SRp30 and TRA2, while asterisks correspond to the potentially novel SVs. Ratios of differentially accumulated SVs are shown below the panels.