1 **Title**

2 **The U1 snRNP subunit LUC7 controls plant development and stress response**
3 **through alternative splicing regulation**

4

5 **Authors**

6 Marcella de Francisco Amorim[1,2,3], Eva-Maria Willing[4], Anchilie G. Francisco-
7 Mangilet[1,2,3,5], Irina Droste-Borel[6], Boris Maček[6], Korbinian Schneeberger[4] and
8 Sascha Laubinger[1,2,3,5*]

9

10 **Affiliations**

11 [1] Centre for Plant Molecular Biology (ZMBP), University of Tuebingen, Tuebingen,
12 Germany
13 [2] Chemical Genomics Centre (CGC) of the Max Planck Society, Dortmund, Germany
14 [3] Max Planck Institute for Developmental Biology, Tuebingen, Germany
15 [4] Max Planck Institute for Plant Breeding Research (MPIPZ), Cologne, Germany
16 [5] Carl von Ossietzky University, Oldenburg, Germany
17 [6] Proteome Centre, University of Tuebingen, Tuebingen, Germany

18

19 Corresponding author: Sascha Laubinger (sascha.laubinger@uol.de)

20

21 Keywords: U1 snRNP, Splicing, Non-sense mediated mRNA decay (NMD),
22 Polyadenylation, RNA transport, Arabidopsis

Abstract

Introns are removed by the spliceosome, a large macromolecular complex composed of five ribonucleoprotein subcomplexes (U snRNP). The U1 snRNP, which binds to 5' splice sites, plays an essential role in early steps of the splicing reaction. Here, we show that Arabidopsis LUC7 proteins, which are encoded by a three-member gene family in Arabidopsis, are important for plant development and stress resistance. We show that LUC7 are U1 snRNP accessory proteins by RNA immunoprecipitation experiments and LUC7 protein complex purifications. Transcriptome analyses revealed that LUC7 proteins are not only important for constitutive splicing, but also affects hundreds of alternative splicing events. Interestingly, LUC7 proteins specifically promote splicing of a subset of terminal introns. Splicing of LUC7-dependent introns is a prerequisite for nuclear export and some splicing events are modulated by stress in a LUC7-dependent manner. Taken together our results highlight the importance of the U1 snRNP component LUC7 in splicing regulation and suggest a previously unrecognized role of a U1 snRNP accessory factor in terminal intron splicing.

41 **Introduction**

42 Eukaryotic genes are often interrupted by non-coding sequences called introns that
43 are removed from pre-mRNAs while the remaining sequence, the exons, are joined
44 together. This process, called splicing, is an essential step before the translation of
45 the mature mRNAs and it offers a wide range of advantages for eukaryotic
46 organisms. For instance, alternative splicing allows the production of more than one
47 isoform from a single gene expanding the genome coding capacity (Kornblihtt et al.,
48 2013; Reddy et al., 2013). Alternative splicing can also regulate gene expression by
49 generating transcripts with premature termination codons (PTC) or/and a long 3'UTR,
50 which may lead to RNA degradation via the nonsense-mediated decay (NMD)
51 pathway (Kalyna et al., 2012; Drechsel et al., 2013; Shaul, 2015). Furthermore,
52 splicing is usually coupled with other RNA processing events, such as 3'end
53 formation and RNA transport to the cytosol (Kaida, 2016; Muller-McNicoll et al.,
54 2016). In plants, alternative splicing contributes to essentially all aspects of
55 development and stress responses (Carvalho et al., 2013; Staiger and Brown, 2013).

56 Intron removal is catalyzed by a large macromolecular complex, the
57 spliceosome, which is formed by five small ribonucleoprotein particles (U snRNP):
58 the U1, U2, U4, U5 and U6 snRNP. Each U snRNP contains a heteroheptameric ring
59 of Sm or Lsm proteins, snRNP-specific proteins and an uridine-rich snRNA.
60 Additional non-core spliceosomal proteins participate during the splicing reaction
61 affecting exon-intron recognition and thus splicing efficiency. The canonical splicing
62 cycle starts with binding of the U1 snRNP to the 5' splice site (5'ss), followed by
63 association of auxiliary proteins such as U2AF to the pre-mRNA, which facilitate the
64 recognition of the 3' splice site (3'ss). The thereby formed complex E recruits the U2
65 snRNP to generate complex A. In the next step, a trimeric complex consisting of
66 U4/U5/U6 snRNPs joins to form complex B. Several rearrangements and ejection of

67 the U1 and U4 snRNP are necessary to generate a catalytically active splicing
68 complex (Wahl et al., 2009; Will and Luhrmann, 2011).

69 The fact that U1 snRNP is recruited to the 5'ss in the initial step of splicing
70 suggests that this complex is necessary for the correct 5' splicing site selection.
71 Indeed, it has been shown that U1-deficient zebrafish mutants accumulate alternative
72 spliced transcripts, suggesting that the U1 snRNP indeed fulfills regulatory roles in
73 splice site selection (Rosel et al., 2011). Although the spliceosome consists of
74 stoichiometrically equal amounts of each subunit, the U1 snRNP is more abundant
75 than all the other spliceosomal subcomplexes (Kaida et al., 2010; Kaida, 2016). One
76 reason for this is that the U1 snRNP executes splicing independent functions. The
77 metazoan U1 snRNP, for instance, binds not only to the 5'ss, but also throughout the
78 nascent transcript blocking a premature cleavage and polyadenylation (Kaida et al.,
79 2010; Berg et al., 2012). Furthermore, the U1 snRNP is important to regulate
80 promoter directionality and transcription in animals (Almada et al., 2013; Guiro and
81 O'Reilly, 2015).

82 U1 snRNP complexes were purified and characterized in yeast and human.
83 The U1 snRNP contains the U1 snRNA, Sm proteins, three U1 core proteins (U1-
84 70K, U1-A and U1-C) and U1-specific accessory proteins, such as LUC7, PRP39
85 and PRP40. All these proteins are conserved in plants suggesting a U1 snRNP
86 composition very similar to the one in yeast and metazoans (Wang and Brendel,
87 2004; Koncz et al., 2012; Reddy et al., 2013). Interaction studies revealed that U1
88 snRNP associates with serine-arginine (SR) proteins, indicating a complex
89 mechanism for splicing site selection that involves also non-snRNP proteins
90 (Golovkin and Reddy, 1998; Cho et al., 2011).

91 The function of the plant U1 snRNP is not well characterized. This might be
92 due to the fact that in *Arabidopsis thaliana* the core U1 snRNP components *U1-70K*

93 or *U1C* are single copy genes and a complete knockout most likely causes severe
94 mutant phenotypes or lethality. On the other hand, proteins such as PRP39, PRP40
95 and LUC7 are encoded by small gene families, which require the generation of
96 multiple mutants for functional studies. Some U1 specific Arabidopsis mutants have
97 been characterized: Mutations in the accessory factor PRP39A cause delayed
98 flowering due to increased expression of the flowering time regulator *FLOWERING*
99 *LOCUS C* (*FLC*), but the mutants do not exhibit severe developmental defects (Wang
100 et al., 2007; Kanno et al., 2017). In a reverse genetic approach, *U1-70K* expression
101 was specifically reduced in flowers by an antisense RNA and the resulting transgenic
102 plants exhibit strong floral defects (Golovkin and Reddy, 2003). Moreover, a mutation
103 in U1A causes an altered salt stress response (Gu et al., 2017). Thus, despite
104 evidences that U1 snRNP is essential for plant development and stress response,
105 the functions of the U1 snRNP in regulating the transcriptome of plants are largely
106 unknown. Other characterized factors, such as GEMIN2 or SRD2 are required for the
107 functionality of all snRNPs, but not specifically for U1 function (Ohtani and Sugiyama,
108 2005; Schlaen et al., 2015).

109 Here, we report on the functional characterization of *Arabidopsis* mutants
110 impaired in U1 snRNP function. For this, we focused in this study on the U1 snRNP
111 components LUC7, which we show to be essential for normal plant development and
112 plant stress resistance. Our whole transcriptome analyses on *luc7* triple mutant show
113 that impairments of LUC7 proteins affect constitutive and alternative splicing.
114 Surprisingly, our results reveal the existence of transcripts, in which terminal introns
115 are preferentially retained in a LUC7-dependent manner. Unspliced LUC7-dependent
116 introns cause a nuclear retention of the pre-mRNAs and the splicing efficiency of
117 LUC7-dependent introns can be modulated by stress. Our results suggests that the
118 plant U1 snRNP components LUC7 regulate alternative splicing of pre-mRNAs and

119 thereby impact their nuclear export, which could be a mechanism to fine-tune gene
120 expression under stress conditions.

121 **Results**

122 **LUC7 proteins, a family of conserved nuclear zinc-finger / arginine-serine (RS)**
123 **proteins, redundantly control plant development**

124 *LETHAL UNLESS CBC 7* (*LUC7*) was first identified in a screen for synthetic lethality
125 in a yeast strain lacking the nuclear cap-binding complex (CBC), which is involved in
126 several RNA processing events (Fortes et al., 1999a; Gonatopoulos-Pournatzis and
127 Cowling, 2014; Sullivan and Howard, 2016). LUC7 proteins carry a C_3H and a C_2H_2-
128 type zinc-fingers, which are located in the conserved LUC7 domain. LUC7 proteins
129 from higher eukaryotes usually contain also an additional C-terminal Arginine/Serine-
130 rich (RS) domain, which is known to mediate protein-protein interactions (Puig et al.,
131 2007; Webby et al., 2009; Heim et al., 2014). *Arabidopsis thaliana* encodes three
132 *LUC7* genes (*AthLUC7A*, *AthLUC7B* and *AthLUC7RL*), which are separated in two
133 clades: *LUC7A/B* and *LUC7RL* (Figure 1A and S1). AthLUC7RL is more similar to its
134 yeast homolog and lacks a conserved stretch of 80 amino acids of unknown function
135 present in AthLUC7A and AthLUC7B (Figure S1). A phylogenetic analysis revealed
136 that algae contain a single *LUC7* gene belonging to the LUC7RL clade reinforcing the
137 idea that LUC7RL proteins are closer to the ancestral LUC7 than LUC7A/B. In the
138 moss Physcomitrella and in the fern Selaginella one can find proteins belonging to
139 both clades, suggesting that the split into *LUC7RL* and *LUC7A/B* occurred early
140 during the evolution of land plants.

141 In order to understand the function of the Arabidopsis U1 snRNP, we analyzed
142 T-DNA insertion lines affecting *LUC7* genes (Figure 1B). Single and double *luc7*
143 mutants were indistinguishable from wild-type plants (WT) (Figure S2). However,
144 *luc7* triple mutant exhibit a wide range of developmental defects, including dwarfism
145 and reduced apical dominance (Figure 1C-E). To test whether the impairment of
146 *LUC7* functions was indeed responsible for the observed phenotypes, we

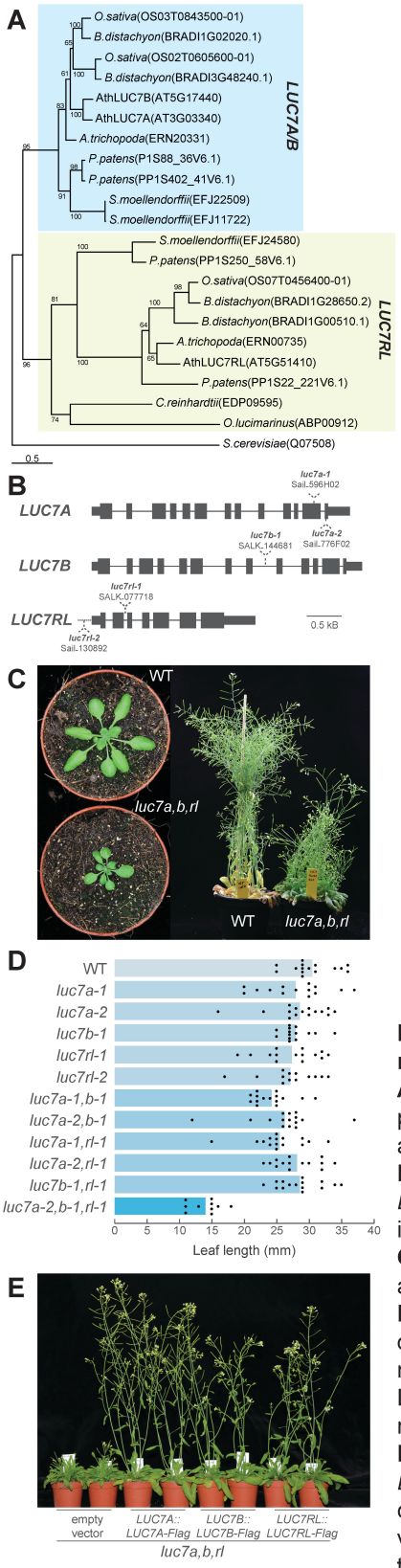

Figure 1: Arabidopsis LUC7 proteins redundantly control plant development
A: Phylogenetic analysis of LUC7 proteins in the plant kingdom using *Saccharomyces cerevisiae* as an external group.
B: Exon/intron structure of *Arabidopsis thaliana* *LUC7A*, *LUC7B* and *LUC7RL*. Dotted lines indicate the positions of T-DNA insertions.
C: WT and *luc7* triple mutant plants 21 days (left) and 45 days (right) after germination.
D: Length of the longest rosette leaf of 21 days-old WT, *luc7* single, double and triple mutant plants growing under long day conditions. Leaves of 10-15 individual plants were measured. Dots indicate individual data points.
E: Complementation of *luc7a,b,rl* mutants by *LUC7A*, *LUC7B* and *LUC7RL* genomic rescue constructs. Transformation of an "empty" binary vector served as a control. Two independent transgenic lines for each construct are shown.

FIGURE 1

reintroduced a wild-type copy of *LUC7A*, *LUC7B* or *LUC7RL* in the *luc7* triple mutant. Each of the *LUC7* genes was sufficient to restore the growth phenotype of the *luc7*

150 triple mutant (Figure 1E). These results reveal that the phenotype observed in this
151 mutant is attributable to the impairment of *LUC7* function and it suggests that *LUC7*
152 genes act redundantly to control *Arabidopsis* growth and development.
153

154 **LUC7 functions in the ABA pathway and is important for cold and salt stress**
155 **responses**

156 Splicing is essential for plant stress resistance and mutants impaired in splicing often
157 react differently to stress and the stress hormone abscisic acid (ABA) (Filichkin et al.,
158 2015; Zhan et al., 2015). In addition, global impairment of the splicing machinery
159 elicits ABA signaling (AlShareef et al., 2017; Ling et al., 2017). To test whether LUC7
160 is important for plant stress resistance and ABA-mediated stress signaling, we
161 analyzed growth parameters of WT, the *luc7* triple mutant and a *luc7* rescue line in
162 presence of exogenous ABA or salt. A cotyledon greening assay showed that *luc7*
163 triple mutants reacted hypersensitively to exogenous ABA (Figure 2A, B), suggesting
164 that *LUC7* plays an important role in the ABA pathway. Furthermore, salt in the
165 growth medium impaired root growth much more strongly in *luc7* triple mutant than in
166 WT or in a *luc7* rescue line (Figure 2C, D). Similarly, cold temperatures strongly
167 compromised the growth of *luc7* triple mutants when compared to WT (Figure 2E).
168 These results imply that functional *LUC7* proteins are required for plant stress
169 resistance and ABA responses.

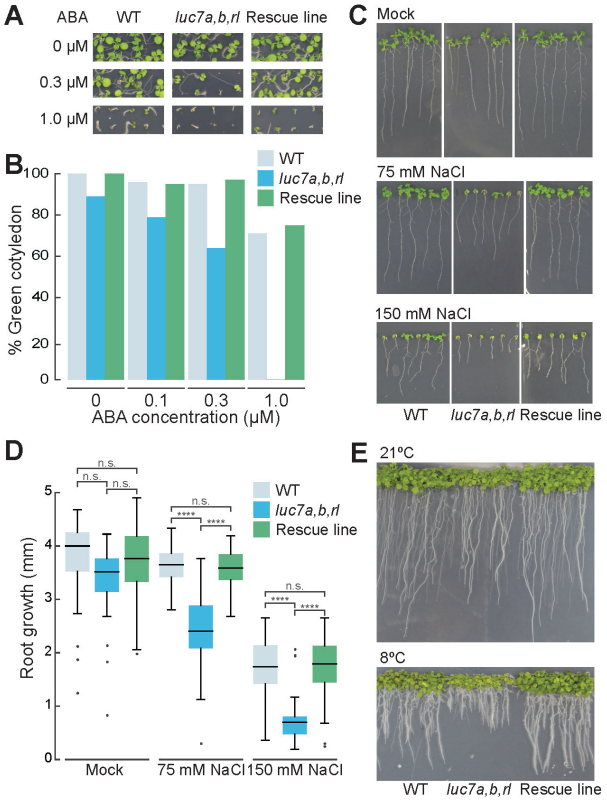

FIGURE 2

Figure 2: Arabidopsis LUC7 is involved in ABA signaling and salt stress responses
A,B: WT, *luc7* triple mutant and a *luc7* rescue line (*luc7a,b,rl; pLUC7A:LUC7A-YFP*) were grown on half-strength MS plates containing 1% sucrose and indicated amount of ABA. Seedling phenotypes (A) and quantification of seedlings with green cotyledons (B) are shown. Green cotyledons were scored ten days after germination. One of two biological replicates is shown.
C,D: WT, *luc7* triple mutant and a *luc7* rescue line (*luc7a,b,rl; pLUC7A:LUC7A-YFP*) were germinated on half-strength MS vertical plates and seedling were transferred on half-strength MS plates containing the indicated amount of NaCl. Plates were always placed vertically and the root growth was scored over 7 days. Phenotypes (C) and root length quantification (D) are shown.
E: Gross phenotype of WT, *luc7* triple mutant and a *luc7* rescue line grown at 22ºC and 8ºC.

LUC7 is a U1 snRNP component in plants

The composition of the U1 snRNP subcomplex is known in yeast and metazoans but not in plants (Will and Luhrmann, 2001; Koncz et al., 2012). Therefore, we asked whether LUC7 is also an U1 component in Arabidopsis. Due to the fact that our genetic analyses of *luc7* mutants suggested that LUC7 proteins act largely redundant, we focused our further analyses mainly on a single LUC7 protein, LUC7A.

178 A protein that is part of the U1 complex is tightly associated with U1 specific
179 components such as the U1 snRNA. To test whether LUC7 is found in a complex
180 with the U1 snRNA, we performed RNA immunoprecipitation (RIP) experiments using
181 a *luc7* triple mutant carrying *pLUC7A:LUC7A-YFP* rescue construct (Figure S3).
182 Immunoprecipitation of LUC7A-YFP enriched the U1 snRNA more than 40-fold, but
183 did not enrich two unrelated, but abundant RNAs, U3 snoRNA and *ACTIN* mRNA
184 (Figure 3A). Small amounts of U2 snRNA was also found associated with LUC7A,
185 which is in agreement with the fact that U1 and U2 snRNP directly interact to form
186 spliceosomal complex A (Figure 3A). However, the amount of recovered U2 snRNA
187 is more than four-fold lower than that of the U1 snRNA (Figure 3A). These results
188 strongly suggest that Arabidopsis LUC7 proteins are bona fide U1 snRNP
189 components.

190 Next we analyzed the subcellular localization of LUC7A and its co-localization
191 with a core U1 snRNP subunit. LUC7A localized to the nucleus, but not to the
192 nucleolus in *Arabidopsis* plants containing the *pLUC7A:LUC7A-YFP* rescue construct
193 (Figure 3B). In addition, LUC7A partially co-localized with U1-70K in the nucleoplasm
194 when transiently expressed in *Nicotiana benthamiana* (Figure 3C). Similar results
195 were obtained for LUC7RL, the Arabidopsis LUC7 most distant in sequence to
196 LUC7A (Figure 3C). In plants, co-localizations studies in protoplasts have shown that
197 also the core U1 components only partially colocalize (Lorkovic and Barta, 2008).
198 These partial colocalizations suggest that plant U1 snRNP proteins may fulfill
199 additional functions as it has been observed in other eukaryotes (Workman et al.,
200 2014).

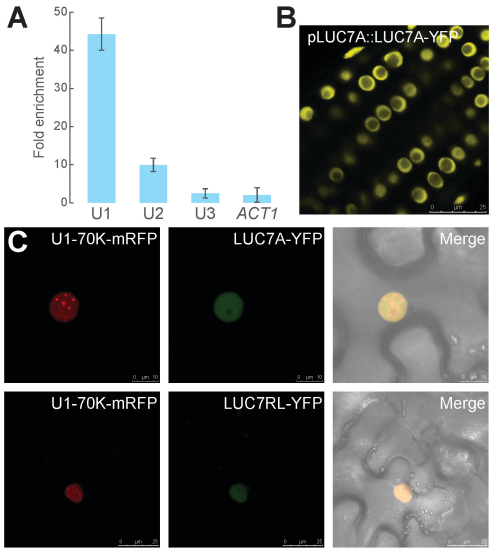

FIGURE 3

Figure 3: Arabidopsis LUC7 is an U1 snRNP component
A: RNA immunoprecipitation using a *pLUC7A:LUC7A-YFP, luc7a,b,rl* complemented line. Proteins were immunoprecipitated using GFP-specific affinity matrix and RNAs were extracted from the input and the immunoprecipitated fraction. U1, U2, U3 snRNAs and *ACTIN* RNA were quantified using qRT-PCR. Enrichment of the respective RNA in *LUC7A:LUC7A-YFP luc7a,b,rl* transgenic line was calculated over WT (negative control). Error bars denote the range of two biological replicates.
B: Subcellular localization of LUC7A in *pLUC7A:LUC7A-YFP luc7a,b,rl* in Arabidopsis transgenic plants. Roots of 9 day-old seedlings were analyzed using confocal microscopy. Scale bar indicates 25 µm.
C: U1-70K-mRFP and LUC7A-YFP or LUC7RL-YFP proteins were transiently expressed in *N. benthamiana*. The subcellular localization of mRFP and YFP fusion proteins was analyzed using confocal microscopy. Scale bars indicate 10 µm and 25 µm for upper and lower panel, respectively.

To further test whether LUC7A associates in planta with known U1 snRNP components, we purified LUC7A-containing complexes. For this, we used *pLUC7A:LUC7A-YFP* complemented lines and as controls wild-type plants and transgenic lines expressing free GFP (*p35S:GFP*). Immunopurifications (IPs) were carried out three to four time independently. We observed that WT often produced more background in mass spectrometry (MS) analyses than the 35S:GFP line and we therefore decided to use WT as a more stringent control (Table S1). Among all identified proteins we considered those putative LUC7 interactors that were found in at least two independent experiment and were at least three time more abundant in *pLUC7A:LUC7A-YFP* IPs than in WT IPs. The mass spectrometry (MS) analysis

212 revealed that LUC7 is indeed found in a complex with core U1 snRNP proteins U1A
213 and U1-70K (Table 1, Table S1). Moreover, we detected peptides corresponding to
214 the spliceosomal complex E components U2AF35 and U2AF65, further suggesting
215 that LUC7 proteins are involved in very early steps of the splicing cycle (Table 1,
216 Table S1). Additional proteins known to be involved in splicing and general RNA
217 metabolism including several serine-arginine (SR) proteins (SR30, SCL30A, SCL33),
218 SR45, SERRATE (SE) and the CBC component ABH1/CBP80 were found in LUC7A-
219 containing complexes (Table 1, Table S1). To test the validity of the LUC7 complex
220 purification experiment, we confirmed the interaction between LUC7 with SE and
221 ABH1/CBP80 by in planta co-immunoprecipitation experiment (Figure S4).
222 Interestingly, we also identified regulatory proteins in LUC7A-containing complexes,
223 among them several kinases and proteins involved in 3'end processing (Table 1,
224 Table S1).

Table 1: List of selected potential LUC7A interacting proteins identified in immunoprecipitation experiments followed by MS analysis.

Identified protein	Aver. # of Peptides		Seq.Coverage (%)		MW (kDa)
	IP LUC7	IP WT	IP LUC7	IP WT	
U1 snRNP components					
LUC7A	45.8	0	71.9	0	47.4
U1-70K	5	0	13.3	0	50.4
SmB-a ; SmB-b	2.5	0	9.8	0	27.0
U1A	2	0	7.7	0	28.1
Splicing-related					
U2AF35A;U2AF35B	3.5	1	11.9	4.3	34.6
SCL30A;SCL33	2.8	0	12.3	0	20.2
SR45	1.8	0	5.2	0	44.6
RSZ21;RSZ22;RSZ22A	1.8	0	8.0	0	22.5
SR30	1	0	3.9	0	29.1
RS2Z33	1	0	2.9	0	32.9
SCL30	0.8	0	2.8	0	29.6
U2AF65A	0.8	0	1.7	0	60.7
Kinases					
SRPK4	8.5	0	16.2	0	59.4
SRPK3	4.3	0	8.9	0	61.2
Proteins linked to 3'end processing/polyA binding					
SPT6L (AT1g63210, AT1g65440)	8.5	1	4.8	0.9	185.0
Polyadenylate-binding protein RBP47C	1.5	0.3	4.0	0.7	48.6
La-related protein 6B (LARP6)	2.3	0.7	5.5	1.9	60.6

225

226

227

228 ***LUC7* effects on the Arabidopsis coding and non-coding transcriptome**

229 In order to identify misregulated and misspliced genes in *luc7* mutants, we performed
230 an RNA-sequencing (RNA-seq) analysis with three biological replicates. We decided
231 to use seven days old WT and *luc7* triple mutant seedlings. At this age, *luc7* triple
232 mutant and WT seedlings are morphologically similar and therefore, changes in
233 transcript levels and splicing patterns most likely reflect changes caused by
234 impairments of LUC7 proteins and are not due to different contribution of tissues
235 caused by, for instance, a delay in development or/and different morphology (Figure
236 S5). We sequenced between 22.1 and 27.6 million reads per library.

237 An analysis of differentially expressed genes revealed that 840 genes are up-
238 and 703 are downregulated in *luc7* triple mutant when compared to WT (Table S2,
239 Table S3). The majority of genes that change expression were protein-coding genes
240 (Figure 4A). Nevertheless, non-coding RNAs (ncRNAs) were significantly enriched
241 among the *LUC7* affected genes ($p < 0.05$, hypergeometric test), although the overall
242 number of ncRNA affected in *luc7* triple mutant is relatively small (Figure 4A, B).
243 Previous studies implied that the U1 snRNP regulates microRNA (miRNA)
244 biogenesis (Bielewicz et al., 2013; Schwab et al., 2013; Knop et al., 2016; Stepien et
245 al., 2017). However, the expression of *MIRNA* genes was not affected in *luc7* triple
246 mutants (Figure 4A). In addition, quantification of mature miRNA levels revealed that
247 all tested miRNAs did not change abundance in *luc7* triple mutants (Figure 4C).
248 These results show that LUC7 proteins affect the expression of protein-coding genes
249 and a subset of ncRNAs, but are not involved in the miRNA pathway.

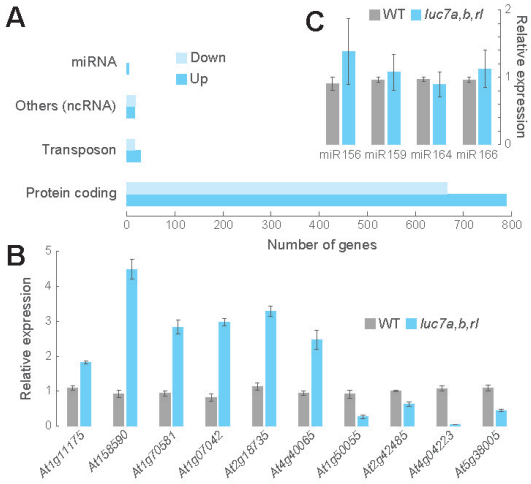

FIGURE 4

Figure 4: Mutations in *LUC7* result in misexpression of protein-coding and non-coding genes, but not in *MIRNA* genes
A: Differentially expressed genes in *luc7a,b,rl* mutant compared to WT.
B,C: qRT-PCR analysis of selected ncRNA (B) and miRNAs (C) in WT and *luc7a,b,rl*. Error bars denote the SEM (n=3).

Arabidopsis *LUC7* functions are important for constitutive and alternative splicing

Because LUC7 proteins are U1 snRNP components, we ask whether misspliced transcripts accumulate in the *luc7* triple mutant. In total, we identified 640 differential splicing events in *luc7* triple mutants compared to WT (Table S4). Only 17 of these alternative splicing events occurred in mRNAs whose expression also differed between *luc7* mutants and WT (Table S5, Table S6). Hence, the splicing differences found were mainly not due to changes in transcript abundance. We detected a large number of intron retention events in the *luc7* triple mutant (Figure 5A). RT-PCR experiments with oligonucleotides flanking selected intron retentions events confirmed the RNA-seq data (Figure 5B). These results suggest that lack and/or impairment of the U1 snRNP component LUC7 disturbs intron recognition and thus splicing. We also identified a large number of exons skipping events in the *luc7* triple mutant. Exon skipping is also a major outcome of impairing U1 snRNP function or binding in metazoans (Lorkovic and Barta, 2008; Rosel et al., 2011). These defects

266 are most likely caused by interactions of the U2 snRNP with U1 snRNPs associated
267 to alternative 5'splice sites (Morcos, 2007). Furthermore, we detected several cases
268 of alternative 5' and 3' splice site selection in *luc7* triple mutants (Figure 5A-F).

269 Some splicing defects observed in *luc7* triple mutants generated transcript
270 variants that did not exist in WT (e.g. *At2g32700*, Figure 5C). In these cases, LUC7
271 proteins affect the splicing of an intron which is constitutively removed in WT plants.
272 On the contrary, in other cases the *luc7* triple mutant lacked specific mRNA isoforms,
273 which exist in wild-type plants (e.g. *At1g10980*, *At4g32060*), or the ratio of two
274 different isoforms was altered in *luc7* triple mutant when compared to WT (e.g.
275 *At3g17310*, *At5g16715*, *At5g48150*, *At2g11000*) (Figure 5D-F). In these cases, LUC7
276 proteins affect a splicing event which is subjected to alternative splicing in WT plants.
277 These results show that LUC7 proteins are involved in both constitutive and
278 alternative splicing in Arabidopsis.

279 Next, we checked whether splicing changes observed in *luc7* triple mutant are
280 actually due to the loss of only a specific *LUC7* gene or whether *LUC7* genes act
281 redundantly. To test this, we analyzed the splicing pattern of some mRNAs in *luc7*
282 single, double and triple mutants. Some splicing defects were detectable even in *luc7*
283 single mutants (Figure S6), but the degree of missplicing increased in *luc7* double
284 and triple mutants suggesting that LUC7 proteins act additively on these introns (e.g.
285 *At5g16715*). Some splicing defects occurred only in *luc7* triple mutants, implying that
286 LUC7 proteins act redundantly to ensure splicing of these introns (e.g. *At1g60995*).
287 Other splicing defects might more likely be due to the lack of LUC7A/B or LUC7RL.
288 For instance, intron removal of *At2g42010* more strongly relied on *LUC7RL*, while
289 removal of an intron in *At5g41220* preferentially depends on LUC7A/LUC7B (Figure
290 S6). These findings suggest that Arabidopsis *LUC7* genes function redundantly,
291 additively or specifically to ensure proper splicing.

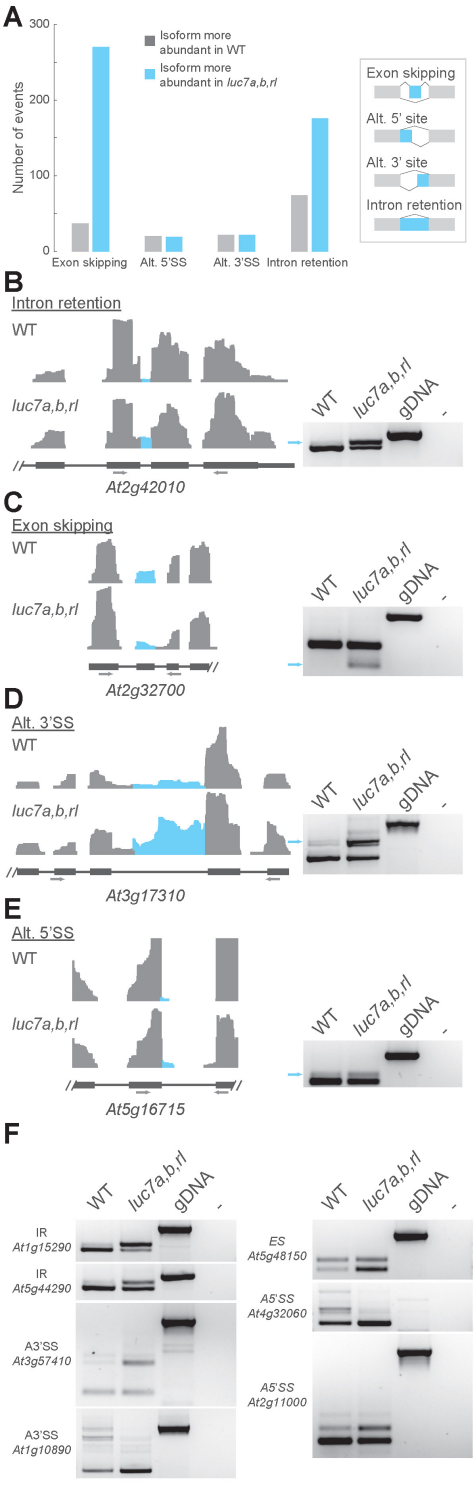

Figure 5: Global analysis of splicing defects in *luc7* triple mutant.
A: Classification of splicing events changes in *luc7* triple mutant compared to WT.
B-F: Coverage plots and RT-PCR validation experiments for selected splicing events in WT and *luc7* triple mutant. Genomic DNA (gDNA) or water (-) served as a control. Primer positions are indicated with gray arrows. IR, intron retention; ES, exon skipping; Alt.3'SS, alternative 3'splicing site; Alt.5'SS, alternative 5'splicing site.

FIGURE 5

292

293

294 **LUC7 proteins are preferentially involved in the removal of terminal introns**

295 In yeast, LUC7 connects the CBC with the U1 snRNP and this interaction is

296 important for the correct 5' splicing site selection (Fortes et al., 1999b). In plants, the

297 CBC associates with SE to ensure splicing of cap-proximal first introns (Laubinger et
298 al., 2008; Raczynska et al., 2010; Raczynska et al., 2013). In addition, we show here
299 that LUC7A form complexes with SE and ABH1/CBP80, one of the CBC competent
300 (Table S1, Figure S4). To investigate the relationship between LUC7 and the
301 CBC/SE in plants, we analyzed the splicing patterns of LUC7 dependent introns in
302 *cbc* mutants (*cbp20* and *cbp80*) and *se-1* by RT-PCR. All tested introns retained in
303 *luc7* triple mutant were correctly spliced in *cbc* and *se* mutants (Figure 6A).
304 Conversely, first introns that were partially retained in *cbp20, cbp80* and *se-1*
305 mutants were completely removed in the *luc7* triple mutant (Figure 6B). These
306 observations suggest that the functions of LUC7 and CBC/SE in splicing of the
307 selected introns do not overlap.

308 Next, we asked whether LUC7 has a preference for promoting splicing of cap-
309 proximal first introns as it has the CBC/SE complex. We classified retained introns in
310 *luc7* triple mutant according to their position within the gene (first, middle or last
311 introns). Only genes with at least 3 introns were considered for this analysis. We
312 found a significant increase in retained last introns, but not first introns, in *luc7* triple
313 mutants (Figure 6C). Although the total number of retained introns was higher among
314 middle introns, the relative amount of retained middle introns in *luc7* triple mutant
315 was significantly reduced (Figure 6C). Retention of terminal introns in *luc7* triple
316 mutants was confirmed by RT-PCR analysis (Figure 6D). In summary, our data
317 revealed that (i) CBC/SE acts independently of LUC7 in splicing of cap-proximal
318 introns and that (ii) LUC7 proteins play an important role for the removal of certain
319 terminal introns.

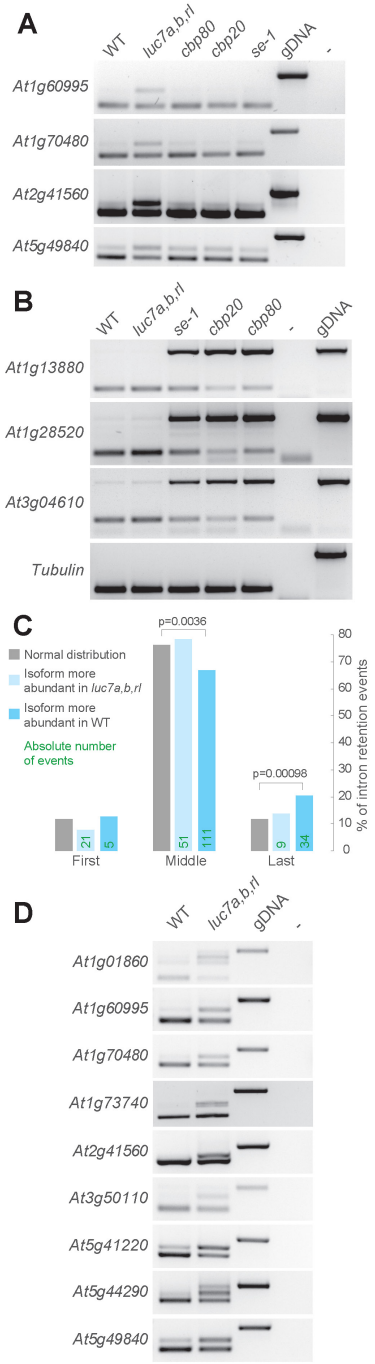

FIGURE 6

Figure 6: LUC7 proteins have a pronounced effect on terminal intron splicing
A: RT-PCR analysis of LUC7-dependent introns in WT, *luc7* triple mutant, *cbp80*, *cbp20* and *se-1* mutants.
B: RT-PCR analysis of CBC/SE-dependent introns in WT, *luc7* triple mutant, *cbp80*, *cbp20* and *se-1* mutants.
C: Classification of intron retention according to the intron position (first, middle, or last). Only genes with 3 or more introns were considered for this analysis.
D: RT-PCR analysis of genes carrying retained terminal introns in WT and *luc7* triple mutants.

320

321

322 **mRNAs harboring unspliced LUC7-dependent introns remain in the nucleus**
323 **and escape NMD**

324 We were further interested in determining the characteristics and possible functions
325 of LUC7-dependent introns. When introns are retained, the resulting mRNA can
326 contain a premature stop codon and a long 3'UTR, which are hallmarks of NMD
327 targets (Kalyna et al., 2012; Drechsel et al., 2013; Shaul, 2015). To check whether
328 mRNAs containing a retained LUC7-dependent introns are NMD substrates, we
329 analyzed their splicing patterns in two mutants impaired in NMD, *lba-1* and *upf3-1*. If
330 unspliced isoforms were indeed NMD targets, we would expect their abundance to
331 be increased in NMD mutants. Interestingly, we did not observe any change between
332 WT and *upf* mutants (Figure 7A). Thus, we conclude that the tested LUC7-dependent
333 introns do not trigger degradation via the NMD pathway.

334 NMD occurs in the cytoplasm and RNAs can escape NMD by not being
335 transported from the nucleus to the cytosol (Gohring et al., 2014). We therefore
336 checked in which cellular compartment mRNAs with spliced and unspliced LUC7-
337 dependent introns accumulate. To do this, we isolated total, nuclear and cytosolic
338 fractions from wild-type and *luc7* triple mutant plants and performed RT-PCR
339 analyses (Figure 7B). Purity of cytosolic and nuclear fractions was controlled by
340 immunoblot analysis using antibodies against histone H3 (specific for nuclear
341 fractions) and a 60S ribosomal protein (L13-1, specific for cytosolic fractions) (Figure
342 7C). Spliced mRNA isoforms accumulated in the cytosol, whereas mRNAs containing
343 the unspliced LUC7-dependent introns were found in nuclear fractions (Figure 7B).
344 These results indicate that retention of LUC7-dependent introns correlates with
345 trapping mRNAs in the nucleus and suggest that splicing of LUC7-dependent introns
346 is essential for mRNA transport to the cytosol.

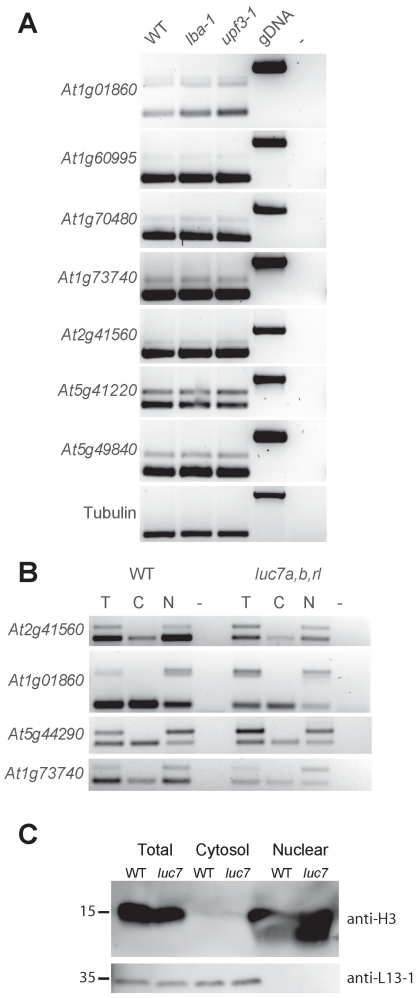

FIGURE 7

Figure 7: mRNAs containing retained LUC7-dependent introns are NMD-insensitive and remain nuclear.
A: RT-PCR analysis of LUC7-dependent introns in WT and NMD mutants (*lba1* and *upf3-1*).
B: Splicing patterns of mRNAs isolated from total (T), cytosolic (C) and nuclear (N) fractions.
C: Immunoblot analysis of proteins isolated from total, cytosolic and nuclear fractions. Blots were probed with antibodies against histone H3 and a ribosomal protein, L13-1.

Splicing of LUC7-dependent introns can be modulated by stress

Our results revealed that a subset of alternatively spliced introns requires LUC7 proteins for efficient splicing and that splicing of these introns is a prerequisite for nuclear export. This mechanism could serve as a nuclear quality control step to prevent that unspliced mRNAs are exported prematurely. Interestingly, a GO analysis of genes containing LUC7-dependent introns indicated an enrichment for stress related genes (Figure S7). This prompted us to speculate that nuclear retention of

356 mRNAs could be exploited as a regulatory mechanism to fine-tune gene expression
357 under stress conditions.

358 To test this hypothesis, we decided to check the splicing of some LUC7-
359 dependent introns in WT under stress condition. We chose cold stress because *luc7*
360 mutants are cold-sensitive (Figure 2) and in addition, it was suggested that U1
361 snRNP functionality is impaired under cold condition (Schlaen et al., 2015). To
362 quantify the amount of unspliced isoforms in cold condition, we designed qPCR-
363 primers specific to unspliced isoforms and the total mRNA pool and calculate the
364 relative amount of mRNA carrying unspliced LUC7-dependent introns compared to
365 the total mRNA pool. mRNAs of *At1g70480*, *At2g41560* and *At5g44290* significantly
366 accumulated unspliced isoforms in responses to cold treatment demonstrating that
367 cold stress modulates the splicing efficiency of these LUC7-dependent introns
368 (Figure 8A). Interestingly, the amount of unspliced mRNA in *luc7* triple mutants does
369 not differ significantly between mock and stress conditions (Figure 8A). This
370 observation suggests that LUC7 is directly involved in the regulation of intron splicing
371 under stress conditions and that LUC7 might be a target for stress response
372 pathways.

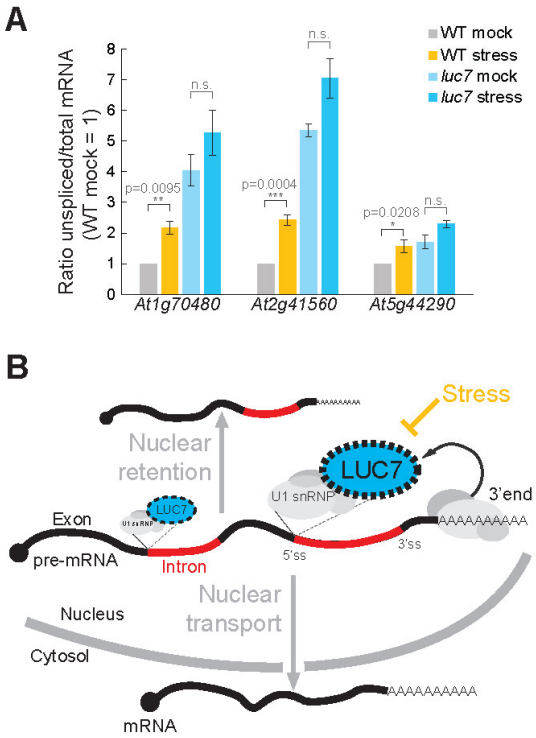

FIGURE 8

Figure 8: Splicing of LUC7 dependent introns can be modulated by stress.
A: Seven days old WT and *luc7* triple mutant seedlings were exposed to cold for 60 min. Splicing ratios (unspliced/total RNA) of four genes featuring a LUC7-dependent introns was analyzed by qPCR. A T-test was performed for statistical analysis.
B: Model for the proposed function of LUC7 in Arabidopsis.

375 **Discussion**

376 **Functions of the Arabidopsis U1 snRNP component LUC7**

377 For this study, we generated an Arabidopsis triple mutant deficient in the U1 snRNP
378 components LUC7 and dissected the genome-wide effects of LUC7s impairments on
379 the Arabidopsis transcriptome. Our results show that LUC7 proteins are *bona-fide* U1
380 components acting mainly redundantly. The reduction of U1 function in the *luc7* triple
381 mutant affects constitutive splicing. A large number of introns are retained in *luc7*
382 triple mutant, suggesting that without a proper recognition of the 5'ss, splicing of the
383 affected introns is impaired. Our results also show that exon-skipping events are
384 impaired in *luc7* triple mutant, revealing that a functional plant U1 snRNP is essential
385 for exon definition. In addition, we show that *luc7* triple mutant affect alternative
386 splicing also by influencing events of alternative 5' and 3' splice site. This implies that
387 the U1 snRNP does not only affect 5' splice site usage, it might also indirectly
388 regulate usage of 3' splice sites via its interaction with U2AFs and the U2 snRNP
389 (Hoffman and Grabowski, 1992; Shao et al., 2012). The functions of LUC7 proteins
390 on the Arabidopsis transcriptome are likely to be underestimated, because
391 misspliced mRNAs in *luc7* mutants might contain hallmarks of NMD and are
392 therefore rapidly turned over and escape detection. Analysis of *luc7* mutants
393 combined with mutations in NMD factors would help to uncover the full set of splicing
394 events affected by LUC7. Furthermore, we found that in our RNA-seq experiments
395 that while the chosen *luc7rl* allele is a RNA-null allele, the *luc7a* and *luc7b* alleles still
396 produced mRNAs that might be translated into truncated proteins. Hence, we can not
397 exclude that a true *luc7* null mutant might exhibit even more severe mutant
398 phenotypes and splicing defects. One has also to consider that U1 snRNP
399 independent splicing has been described in animals, indicating that not all introns

400 require the U1 complex for efficient intron removal (Fukumura et al., 2009). The
401 degree of U1-independent splicing in plants remains to be elucidated.
402 Duplications among genes encoding for U1 snRNP proteins, such as the
403 LUC7 genes, open up a possibility for sub- and neofunctionalization of U1 accessory
404 proteins. Furthermore, the Arabidopsis genome encodes 14 potential U1 snRNAs,
405 which slightly differ in sequence (Wang and Brendel, 2004). Therefore, the plant U1
406 snRNP presumably does not exist as a single complex, but might exist as different
407 sub-complexes exhibiting distinct specificities and functions. In metazoans, the
408 existence of at least four different U1 snRNP subcomplex has been suggested
409 (Hernandez et al., 2009; Guiro and O'Reilly, 2015). Specific combinations of plant U1
410 protein family members and U1 snRNAs could generate an even higher number of
411 such U1 subcomplexes, which could be responsible for specific splicing events. Our
412 results show that LUC7 can act redundantly, but can also fulfill specific functions,
413 suggesting that LUC7 complexes specifically act on certain pre-mRNAs. In this
414 regard, it is important to note that an additional short stretch of amino acids
415 separates the two zinc-finger domains in LUCA and LUC7B (Figure S1). Changing
416 the space in between RNA binding domains affects substrate specificities and could
417 explain different specificities among LUC7 proteins (Chen and Varani, 2013).
418

419 **LUC7 function in terminal intron splicing**

420 Interestingly, *luc7* triple mutant showed a significant higher retention rate of terminal
421 introns compared to first or middle introns. This was surprising because LUC7 was
422 initially found to act in concert with the CBC, a complex involved in the removal of
423 cap-proximal first introns, but not of last introns (Lewis et al., 1996). We found LUC7
424 in a complex with the CBC and the CBC-associated protein SE also in Arabidopsis.

425 However, the functional significance of the of LUC7-CBC/SE interaction remains to
426 be established.

427 Often, the removal of terminal introns is coupled to polyadenylation (Cooke et
428 al., 1999; Cooke and Alwine, 2002; Rigo and Martinson, 2008). Interestingly, we
429 detected components involved in RNA 3'end processing or polyA-binding as part of
430 LUC7A complexes, suggesting that such interactions may contribute to the specific
431 functions of LUC7 in terminal intron splicing. We found LUC7 in complexes with
432 RBP47C, a polyA-binding protein of unknown function (Lorkovic et al., 2000), and
433 LARP6, which is targeted to 3'ends of mRNA through interaction with polyA-binding
434 protein 2 (PAB2) (Merret et al., 2013). Interestingly, we also found an SPT6-like
435 transcription factor associated with LUC7 complexes. SPT6 binds the pol II C-
436 terminal domain (CTD) phosphorylated at serine 2 (Ser2P), which accumulates at
437 3'end of genes (Kaplan et al., 2000; Sun et al., 2010). None of these LUC7 complex
438 components has been studied functionally and it will be a major effort for future
439 studies to determine the function of these proteins in terminal intron splicing.

440

441 **Possible functions of regulated intron retention for plant stress responses**

442 We found that splicing of LUC7-dependent introns is required for transport of mRNAs
443 from the nucleus to the cytosol. The fact that we cannot detect unspliced transcript in
444 the cytosol suggests a nuclear retention mechanism for such mRNAs. One possibility
445 is that LUC7-dependent introns might contain binding sites for specific trans-
446 regulatory factors that upon binding inhibit export. Polypyrimidine tract-binding
447 protein 1 (PTB1) is a candidate for such a trans-regulatory protein, because binding
448 of PTB1 to introns represses nuclear export of certain RNAs (Yap et al., 2012; Roy et
449 al., 2013).

450 Nuclear retention of unspliced mRNAs might be a much more general
451 mechanism to escape NMD and to regulate gene expression (Marquez et al., 2015;
452 Wong et al., 2016). In plants, some specific transcript isoforms have been detected
453 only in the nucleus, but not in the cytosol (Gohring et al., 2014). Also in metazoans,
454 intron retention might have a more general role in regulating gene expression (Yap et
455 al., 2012; Braunschweig et al., 2014; Pimentel et al., 2016; Naro et al., 2017). The
456 so-called detained introns are evolutionary conserved, NMD insensitive and retained
457 in the nucleus (Boutz et al., 2015). The functional importance of intron retention was
458 also suggested in the fern *Marsilea vestita*, in which many mRNAs contain introns
459 that are only spliced shortly before gametophyte development (Boothby et al., 2013).
460 We found that splicing of LUC7 dependent introns can be modulated by cold
461 stress. Because retention of these introns causes nuclear trapping, it is prompting to
462 speculate that environmental cues affect splicing and nuclear retention of mRNAs.
463 Such a mechanism would regulate the amount of translatable mRNAs in the cytosol
464 in a cost-efficient and rapid manner (Figure 8B). Since the stress-dependent
465 regulation of splicing of LUC7-dependent introns is lost in *luc7* mutants, one can
466 expect that LUC7 function might be regulated under stress conditions. Interestingly,
467 the RS domains of LUC7 proteins are phosphorylated and we identified three kinases
468 as potential LUC7A interactors (Heazlewood et al., 2008; Durek et al., 2010). In
469 addition, stress signaling triggered by the phytohormone ABA causes differential
470 phosphorylation of several splicing factors (Umezawa et al., 2013; Wang et al.,
471 2013). Whether stress-induced changes in phosphorylation play a role in regulating
472 LUC7 proteins and whether the LUC7-interacting kinases here identified are involved
473 in this process remains to be elucidated.

474 **Material and Methods**

475

476 **Plant material and growth conditions**

477 All mutants were in the Columbia-0 (Col-0) background. *luc7a-1* (SAIL_596_H02)
478 and *luc7a-2* (SAIL_776_F02), *luc7b-1* (SALK_144681), *luc7rl-1* (SALK_077718) and
479 *luc7rl-2* (SALK_130892C) were isolated by PCR-based genotyping (Table S7). *luc7*
480 double and triple mutants were generate by crossing individual mutants. All other
481 mutants used in this study (*abh1-285*, *cbp20-1*, *se*-1, *lba-1* and *upf3-1*) were
482 described elsewhere (Prigge and Wagner, 2001; Papp et al., 2004; Hori and
483 Watanabe, 2005; Yoine et al., 2006; Laubinger et al., 2008). The line expressing
484 GFP was generated using the vector pBinarGFP and was kindly provided by Dr.
485 Andreas Wachter (Wachter et al., 2007). For complementation analyses,
486 *pLUC7A:LUC7A-FLAG*, *pLUC7B:LUC7B-FLAG*, *pLUC7RL:LUC7RL-FLAG* and
487 *pLUC7A:LUC7A-YFP* constructs were introduced in *luc7* triple mutant by
488 Agrobacterium-mediated transformation (Clough and Bent, 1998). All plants were
489 grown on soil in long days conditions (16-h light/8-h dark) at 20°C/18°C day/night.
490 The size of *luc7* mutants was assessed by measuring the longest rosette leaf after 21
491 days. For all molecular studies, seeds were surface-sterilized, plated on 1/2 MS
492 medium with 0.8% phytoagar and grown for 7 days in continuous light at 22°C. For
493 the cold treatment, plates with *Arabidopsis* seedlings were transferred to ice-water for
494 60 min. For the root growth assay, 4 days old seedlings growing on vertical plates
495 were transferred to mock plates or plates containing indicated amount of NaCl and
496 grown for more 11 days vertically. Root growth rate per day was assessed by
497 measuring in ImageJ the root length in the days 2 and 9 after transfer. For ABA
498 sensitivity assays, seedlings were grown for 10 days on 1/2 MS plates supplemented
499 with 0.8 % phytoagar, 1 % sucrose and indicated amounts of ABA (+) (Sigma -

500 A4906). For cold stress experiments, seeds were grown at 20°C for 5 days and then
501 transferred to 8°C for two weeks.
502
503 **Plasmid constructions and transient expression analyses**
504 For the expression of C-terminal FLAG- and YFP-tagged LUC7 proteins expressed
505 from their endogenous regulatory elements, 2100 bp, 4120 bp and 2106 bp upstream
506 of the ATG start codon of *LUC7A*, *LUC7B* and *LUC7RL*, respectively, to the last
507 coding nucleotide were PCR-amplified and subcloned in pCR8/GW/TOPO®
508 (Invitrogen). Oligonucleotides are listed in Table S7. Entry clones were recombined
509 with pGWB10 and pGWB540 using Gateway LR clonase II (Invitrogen) to generate
510 binary plasmids containing *pLUC7A:LUC7A-FLAG*, *pLUC7B:LUC7B-FLAG*,
511 *pLUC7RL:LUC7RL-FLAG* and *pLUC7A:LUC7A-YFP*. For the co-localization studies,
512 entry vector containing the coding sequence of U1-70k was recombined with
513 pGWB654 for the expression of *p35S:U1-70k-mRFP* (Nakagawa et al., 2007).
514 Agrobacterium-mediated transient transformation of *Nicotiana benthamiana* plants
515 was conducted as following. Overnight Agrobacterium culture were diluted in 1:6 and
516 grown for 4 hours at 28°C. After centrifugation, pellets were resuspended in
517 infiltration medium (10 mM $MgCl_2$, 10 mM MES-KOH pH 5.6, 100 μM
518 Acetosyringone). The OD 600nm was adjusted to 0.6-0.8 and samples were mixed
519 when required. *N. benthamiana* were infiltrated and subcellular localization was
520 checked after 3 days. Subcellular localization of fluorescent proteins was analyzed by
521 confocal microscopy using a Leica TCS SP8.
522
523 **Phylogenetic analysis**
524 AthLUC7A (AT3G03340) protein sequence was analyzed in Interpro
525 (https://www.ebi.ac.uk/interpro/) to retrieve the Interpro ID for the conserved Luc7-

526 related domain (IPR004882). The sequence for *Saccharomyces cerevisiae* (strain
527 ATCC 204508_S288c) was obtained in Interpro. Plants sequences were extracted
528 using BioMart selecting for the protein domain IPR004882 on Ensembl Plants
529 (http://plants.ensembl.org/). The following genomes were included in our analyses:
530 *Amborella trichopoda* (AMTR1.0 (2014-01-AGD)); *Arabidopsis thaliana* (TAIR10
531 (2010-09-TAIR10)); *Brachypodium distachyon* (v1.0); *Chlamydomonas reinhardtii*
532 (v3.1 (2007-11-ENA)); *Physcomitrella patens* (ASM242v1 (2011-03-Phypa1.6));
533 *Selaginella moellendorffii* (v1.0 (2011-05-ENA)); *Oryza sativa Japonica* (IRGSP-1.0);
534 and *Ostreococcus lucimarinus* genes (ASM9206v1). The phylogenetic analysis was
535 performed in Seaview (Version 4.6.1) using Muscle for sequence alignment.
536 Maximum likehood (PhYML) was employed with 1000 bootstraps (Gouy et al., 2010).
537

538 **RNA extractions, RT-PCR and qRT-PCR**

539 RNAs extractions were performed with Direct-zol™ RNA MiniPrep Kit (Zymo
540 Research). Total RNAs were treated with DNAse I and cDNA synthesis carried out
541 with RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific) using usually
542 oligo dT primers or a mixture of hexamer and miRNA-specific stem-loop primers
543 (Table S7). Standard PCRs for the splicing analysis were performed with DreamTaq
544 DNA Polymerase (Thermo Scientific). Quantitative RT-PCR (qRT-PCR) was
545 performed using the Maxima SYBR Green (Thermo Scientific) in a Bio-Rad CFX 384.
546 For all qPCR-primers, primer efficiencies were determined by a serial dilution of
547 cDNA template. The relative expressions were calculated using the 2ˆ(-ΔΔCT)
548 method with *PP2A* or *ACTIN* as control. For the qRT-PCR to detect splicing ratio
549 changes under cold condition, the ratio $2^{(-\Delta CT_{unspliced})}/2^{(-\Delta CT_{total\ RNA})}$ was
550 calculated separately for each replicate and t-test was performed before calculating
551 the relative to WT mock. Oligonucleotides are listed in Table S7. For RNA-

552 sequencing analysis, polyA RNAs were enriched from 4 µg of total RNAs using
553 NEBNext Oligo d(T)$_{25}$ Magnetic Beads (New England Biolabs). The libraries were
554 prepared using ScriptSeq™ Plant Leaf kit (Epicentre) following the manufacturer's
555 instruction. Single end sequencing was performed on an Illumina HiSeq2000.
556 Sequencing data were deposited at Gene Expression Omnibus under accession
557 number GSE98779.
558
559
560 **RNA-seq libraries: Mapping, differential expression analysis and splicing**
561 **analysis**
562 RNA-seq reads for each replicate were aligned against the *Arabidopsis thaliana*
563 reference sequence (TAIR10) using tophat (v2.0.10, -p2, -a 10, -g 10, -N 10, --read-
564 edit-dist 10, --library-type fr-secondstrand, --segment-length 31, -G TAIR10.gff). Next,
565 cufflinks (version 2.2.1) was used to extract FPKM counts for each expressed
566 transcript generating a new annotation file (transcripts.gtf), where the coordinates of
567 each expressed transcript can be found. Cuffcompare (version 2.2.1) was then used
568 to generate a non-redundant annotation file containing all reference transcripts in
569 addition to new transcripts expressed in at least one of the nine samples
570 (cuffcmp.combined.gtf). The differential expression analysis was performed with
571 cuffdiff (version 2.2.1) between wt/*luc7* triple using the annotation file generated by
572 cuffcompare (false discovery rate (FDR) < 0.05 and fold change (FC) > 2). For the
573 splicing analysis, the same alignment files generated by tophat and annotation files
574 generated by cuffcompare (cuffcmp.combined.gtf) were used as input for MATS
575 (version 3.0.8) in order to test for differentially spliced transcripts (Shen et al., 2014).
576
577 **Subcellular fractionation**

578 Two grams of seedlings were ground in N_2 liquid and resuspended in 4 ml of Honda
579 buffer (0.44 M sucrose, 1.25% Ficoll 400, 2.5% Dextran T40, 20 mM HEPES KOH
580 pH 7.4, 10 mM $MgCl_2$, 0.5% Triton X-100, 5 mM DTT, 1 mM PMSF, protease
581 inhibitor cocktail (Roche) supplemented with 40U/ml of Ribolock®). The homogenate
582 was filtered through 2 layer of Miracloth, which was washed with 1ml of Honda
583 Buffer. From the filtrate, 300 µl was removed as "total" fraction and kept on ice.
584 Filtrates were centrifuged at 1,500 *g* for 10 min, 4°C for pelleting nuclei and
585 supernatants were transferred to a new tube. Supernatants were centrifuged at 13
586 000 x g, 4 °C, 15 min and 300 µl were kept on ice as cytoplasmic fraction. Nuclei
587 pellets were washed five times in 1 ml of Honda buffer (supplemented with 8U/ml of
588 Ribolock®, centrifugation at 1,800 *g* for 5 min. The final pellet was resuspended in
589 300 µl of Honda buffer. To all the fractions (total, cytoplasmic and nuclei), 900 µl of
590 TRI Reagent (Sigma) was added. After homogenization, 180 µl of chloroform was
591 added and samples were incubated at room temperature for 10 min. After
592 centrifugation at 10 0000 rpm for 20 min, 4°C, the aqueous phase were transferred to
593 a new tube and RNA extracted with Direct-zol™ RNA MiniPrep Kit (Zymo Research).
594 The organic phase was collected and proteins were isolated according to
595 manufacturer's instructions. cDNA synthesis with random primes was performed as
596 above. Proteins extracted were analyzed by standard western blot techniques using
597 the following antibodies: H3 (~ 17 KDa / ab 1791, Abcam) and 60S ribosomal (~ 23,7-
598 29 KDa / L13, Agrisera).
599

600 **RNA immunoprecipitation**

601 RNA immunoprecipitation (RIP) using WT and a *pLUC7A:LUC7A-eYFP* rescue line
602 was performed as described elsewhere with minor modifications (Rowley et al., 2013;
603 Xing et al., 2015). Isolated nuclei were sonicated in nuclear lysis buffer in a Covaris

604 E220 (Duty Cycle: 20%; Peak intensity: 140; Cycles per Burst: 200; Cycle time: 3').
605 RNAs were extracted using RNeasy Plant Mini Kit (QIAGEN) following the
606 manufacturer's instructions. The RNA were treated with DNAseI (Thermo Scientific)
607 and samples were split in half for the (-)RT reaction. cDNA synthesis were perform
608 with SuperScript™ III Reverse Transcriptase (Invitrogen). qRT-PCRs were performed
609 with QuantiNova™ SYBRR Green PCR (QIAGEN).
610
611 **GO Analysis**
612 GO analysis was performed in Bar Utoronto (http://bar.utoronto.ca/ntools/cgi-
613 bin/ntools_classification_superviewer.cgi).
614
615 **Protein complex purification and mass spectrometry (MS) Analysis**
616 LUC7A immunoprecipitation was performed using a complemented line
617 *pLUC7A:LUC7A-eYFP* (line 20.3.1) and a transgenic *p35S:GFP* and WT as controls.
618 Four independent biological replicates were performed. Seedlings (4 g) were ground
619 in N_2 liquid and respuspended in 1 volumes of extraction buffer (50 mM Tris-Cl pH
620 7.5, 100 mM NaCl, 0.5 % Triton X-100, 5 % Glycerol, 1 mM PMSF, 100 µM MG132,
621 Complete Protease Inhibitor Cocktail EDTA-free [Roche] and Plant specific protease
622 Inhibitor, Sigma P9599). After thawing, samples were incubated on ice for 30 min,
623 centrifuged at 3220 rcf for 30 min at 4°C and filtrated with two layers of Miracloth. For
624 each immunoprecipitation, 20 µl of GFP-trap (Chromotek) was washed twice with 1
625 ml of washing buffer (50 mM Tris-Cl pH 7.5, 100 mM NaCl, 0.2 % Triton X-100) and
626 once with 0.5 ml of IP buffer. For each replicate, the same amount of plant extracts
627 (~5 ml) were incubated with GFP-trap and incubated on a rotating wheel at 4°C for 3
628 hours. Samples were centrifuged at 800-2000 rcf for 1-2 min and the supernatant
629 discarded. GFP-beads were resuspended in 1 ml of washing buffer, transferred into a

630 new tube and washed 4 to 5 times. Then, beads were ressuspended in ~40 µl of 2x
631 Laemmli Buffer and incubated at 80°C for 10 min. Short gel purifications (SDS-
632 PAGE) were performed and gels slices were digested with Trypsin. LC-MS/MS
633 analyses were performed in two mass spectrometer. Samples from R10 to R14 were
634 analyzed on a Proxeon Easy-nLC coupled to Orbitrap Elite (method: 90min, Top10,
635 HCD). Samples from R15 to R17 were analysed on a Proxeon Easy-nLC coupled to
636 OrbitrapXL (method: 90min, Top10, CID). Samples from R18 to R20 analysis on a
637 Proxeon Easy-nLC coupled to OrbitrapXL (method: 130min, Top10, CID). All the
638 replicates were processed together on MaxQuant software (Version 1.5.2.8. with
639 integrated Andromeda Peptide search engine) with a setting of 1% FDR and the
640 spectra were searched against an *Arabidopsis thaliana* Uniprot database
641 (UP000006548_3702_complete_20151023.fasta). All peptides identified are listed in
642 Supplementary Table S1 and raw data were deposited publically (accession
643 PXD006127). For co-immunoprecipitation experiments shown in Figure S4,
644 experiments were conducted as described above and IPed protein fractions were
645 analyzed using SDS-PAGE followed by detection with GFP- (Roche), SE- (Agrisera)
646 and CBP80- (Agrisera) specific antibodies.
647

648 **Acknowledgments**

649 This work was supported by the Deutsche Forschungsgemeinschaft DFG (to S.L.,
650 LA2633-4/1), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
651 (CAPES - Brazil) for doctoral fellowship (to M.F.A.), the Max Planck Society (to K.S.),
652 the Max Planck Society Chemical Genomics Centre (CGC) through its supporting
653 companies AstraZeneca, Bayer CropScience, Bayer Healthcare, Boehringer-
654 Ingelheim and Merck (to S.L). We are grateful to Andreas Wachter (ZMBP, University
655 of Tuebingen, Germany) and members of the lab for critical reading of the

656 manuscript, Christa Lanz for her invaluable help with Illumina sequencing, and
657 Johanna Schröter and her team for excellent care of our plants, Anja Hoffmann for
658 excellent technical assistance, and Andreas Wachter (ZMBP, University of
659 Tuebingen, Germany), Tsuyoshi Nakagawa (Department of Molecular and Functional
660 Genomics, Center for Integrated Research in Science, Shimane University, Matsue,
661 Japan) and the Notthingham Arabidopsis Stock Centre for providing seeds and DNA
662 constructs.

663

664 **Author contributions**

665 M.d.F.A. and S.L. designed research. M.d.F.A., A.G.F.-M., I.D-B. and S.L. performed
666 experiments. E.-M. W., M.d.F.A., A.G.F.-M., I.D-B., B.M., K.S. and S.L. analyzed the
667 data. M.d.F.A. and S.L. wrote the manuscript with contributions from all authors.

668

669 **Conflict of interest**

670 The authors declare no conflict of interest.

671

672 **Figure legends**

673 **Figure 1: *Arabidopsis* LUC7 proteins redundantly control plant development**

674 **A:** Phylogenetic analysis of LUC7 proteins in the plant kingdom using
675 *Saccharomyces cerevisiae* as an external group.

676 **B:** Exon/intron structure of *Arabidopsis thaliana LUC7A*, *LUC7B* and *LUC7RL*.
677 Dotted lines indicate the positions of T-DNA insertions.

678 **C:** WT and *luc7* triple mutant plants 21 days (left) and 45 days (right) after
679 germination.

680 **D:** Length of the longest rosette leaf of 21 days-old WT, *luc7* single, double and triple
681 mutant plants growing under long day conditions. Leaves of 10-15 individual plants
682 were measured. Dots indicate individual data points.
683 **E:** Complementation of *luc7a,b,rl* mutants by *LUC7A*, *LUC7B* and *LUC7RL* genomic
684 rescue constructs. Transformation of an "empty" binary vector served as a control.
685 Two independent transgenic lines for each construct are shown.
686
687 **Figure 2: *Arabidopsis* LUC7 is involved in ABA signaling and salt stress**
688 **responses**
689 **A,B:** WT, *luc7* triple mutant and a *luc7* rescue line (*luc7a,b,rl; pLUC7A:LUC7A-YFP*)
690 were grown on half-strength MS plates containing 1% sucrose and indicated amount
691 of ABA. Seedling phenotypes (A) and quantification of seedlings with green
692 cotyledons (B) are shown. Green cotyledons were scored ten days after germination.
693 One of two biological replicates is shown.
694 **C,D:** WT, *luc7* triple mutant and a *luc7* rescue line (*luc7a,b,rl; pLUC7A:LUC7A-YFP*)
695 were germinated on half-strength MS vertical plates and seedling were transferred on
696 half-strength MS plates containing the indicated amount of NaCl. Plates were always
697 placed vertically and the root growth was scored over 7 days. Phenotypes (C) and
698 root length quantification (D) are shown.
699 **E:** Gross phenotype of WT, *luc7* triple mutant and a *luc7* rescue line grown at 22°C
700 and 8°C.
701
702 **Figure 3: *Arabidopsis* LUC7 is an U1 snRNP component**
703 **A:** RNA immunoprecipitation using a *pLUC7A:LUC7A-YFP, luc7a,b,rl* complemented
704 line. Proteins were immunoprecipitated using GFP-specific affinity matrix and RNAs
705 were extracted from the input and the immunoprecipitated fraction. U1, U2, U3

706 snRNAs and *ACTIN* RNA were quantified using qRT-PCR. Enrichment of the
707 respective RNA in *LUC7A:LUC7A-YFP luc7a,b,rl* transgenic line was calculated over
708 WT (negative control). Error bars denote the range of two biological replicates.
709 **B:** Subcellular localization of LUC7A in *pLUC7A:LUC7A-YFP luc7a,b,rl* in
710 Arabidopsis transgenic plants. Roots of 9 day-old seedlings were analyzed using
711 confocal microscopy. Scale bar indicates 25 µm.
712 **C:** U1-70K-mRFP and LUC7A-YFP or LUC7RL-YFP proteins were transiently
713 expressed in *N. benthamiana*. The subcellular localization of mRFP and YFP fusion
714 proteins was analyzed using confocal microscopy. Scale bars indicate 10 µm and 25
715 µm for upper and lower panel, respectively.
716
717 **Figure 4: Mutations in *LUC7* result in misexpression of protein-coding and non-**
718 **coding genes, but not in *MIRNA* genes**
719 **A:** Differentially expressed genes in *luc7a,b,rl* mutant compared to WT.
720 **B,C:** qRT-PCR analysis of selected ncRNA **(B)** and miRNAs **(C)** in WT and
721 *luc7a,b,rl*. Error bars denote the SEM (n=3).
722
723 **Figure 5: Global analysis of splicing defects in *luc7* triple mutant.**
724 **A:** Classification of splicing events changes in *luc7* triple mutant compared to WT.
725 **B-F:** Coverage plots and RT-PCR validation experiments for selected splicing events
726 in WT and *luc7* triple mutant. Genomic DNA (gDNA) or water (-) served as a control.
727 Primer positions are indicated with gray arrows. IR, intron retention; ES, exon
728 skipping; Alt.3'SS, alternative 3'splicing site; Alt.5'SS, alternative 5'splicing site.
729
730 **Figure 6: LUC7 proteins have a pronounced effect on terminal intron splicing**

731 **A:** RT-PCR analysis of LUC7-dependent introns in WT, *luc7* triple mutant, *cbp80*,
732 *cbp20* and *se-1* mutants.
733 **B:** RT-PCR analysis of CBC/SE-dependent introns in WT, *luc7* triple mutant, *cbp80*,
734 *cbp20* and *se-1* mutants.
735 **C:** Classification of intron retention according to the intron position (first, middle, or
736 last). Only genes with 3 or more introns were considered for this analysis.
737 **D:** RT-PCR analysis of genes carrying retained terminal introns in WT and *luc7* triple
738 mutants.
739
740 **Figure 7: mRNAs containing retained LUC7-dependent introns are NMD-**
741 **insensitive and remain nuclear.**
742 **A:** RT-PCR analysis of LUC7-dependent introns in WT and NMD mutants (*lba1* and
743 *upf3-1*).
744 **B:** Splicing patterns of mRNAs isolated from total (T), cytosolic (C) and nuclear (N)
745 fractions.
746 **C:** Immunoblot analysis of proteins isolated from total, cytosolic and nuclear
747 fractions. Blots were probed with antibodies against histone H3 and a ribosomal
748 protein, L13-1.
749
750 **Figure 8: Splicing of LUC7 dependent introns can be modulated by stress.**
751 **A:** Seven days old WT and *luc7* triple mutant seedlings were exposed to cold for 60
752 min. Splicing ratios (unspliced/total RNA) of four genes featuring a LUC7-dependent
753 intron was analyzed by qPCR. A T-test was performed for statistical analysis.
754 **B:** Model for the proposed function of LUC7 in Arabidopsis.
755

756 **Table 1: List of selected potential LUC7A interacting proteins identified in**
757 **immunoprecipitation experiments followed by MS analysis.**

758

759 **Supplementary Material**

760 Figure S1-S8, Table S1-S7

761

762 **Literature**

763 **Almada, A.E., Wu, X., Kriz, A.J., Burge, C.B., and Sharp, P.A.** (2013). Promoter
764 directionality is controlled by U1 snRNP and polyadenylation signals. Nature
765 **499,** 360-363.
766 **AlShareef, S., Ling, Y., Butt, H., Mariappan, K.G., Benhamed, M., and Mahfouz,**
767 **M.M.** (2017). Herboxidiene triggers splicing repression and abiotic stress
768 responses in plants. BMC genomics **18,** 260.
769 **Berg, M.G., Singh, L.N., Younis, I., Liu, Q., Pinto, A.M., Kaida, D., Zhang, Z.,**
770 **Cho, S., Sherrill-Mix, S., Wan, L., and Dreyfuss, G.** (2012). U1 snRNP
771 determines mRNA length and regulates isoform expression. Cell **150,** 53-64.
772 **Bielewicz, D., Kalak, M., Kalyna, M., Windels, D., Barta, A., Vazquez, F.,**
773 **Szweykowska-Kulinska, Z., and Jarmolowski, A.** (2013). Introns of plant
774 pri-miRNAs enhance miRNA biogenesis. EMBO reports **14,** 622-628.
775 **Boothby, T.C., Zipper, R.S., van der Weele, C.M., and Wolniak, S.M.** (2013).
776 Removal of retained introns regulates translation in the rapidly developing
777 gametophyte of Marsilea vestita. Developmental cell **24,** 517-529.
778 **Boutz, P.L., Bhutkar, A., and Sharp, P.A.** (2015). Detained introns are a novel,
779 widespread class of post-transcriptionally spliced introns. Genes &
780 development **29,** 63-80.
781 **Braunschweig, U., Barbosa-Morais, N.L., Pan, Q., Nachman, E.N., Alipanahi, B.,**
782 **Gonatopoulos-Pournatzis, T., Frey, B., Irimia, M., and Blencowe, B.J.**
783 (2014). Widespread intron retention in mammals functionally tunes
784 transcriptomes. Genome research **24,** 1774-1786.
785 **Carvalho, R.F., Feijao, C.V., and Duque, P.** (2013). On the physiological
786 significance of alternative splicing events in higher plants. Protoplasma **250,**
787 639-650.
788 **Chen, Y., and Varani, G.** (2013). Engineering RNA-binding proteins for biology. The
789 FEBS journal **280,** 3734-3754.
790 **Cho, S., Hoang, A., Sinha, R., Zhong, X.Y., Fu, X.D., Krainer, A.R., and Ghosh,**
791 **G.** (2011). Interaction between the RNA binding domains of Ser-Arg splicing
792 factor 1 and U1-70K snRNP protein determines early spliceosome assembly.
793 Proc Natl Acad Sci U S A **108,** 8233-8238.
794 **Clough, S.J., and Bent, A.F.** (1998). Floral dip: a simplified method for
795 Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant
796 journal : for cell and molecular biology **16,** 735-743.
797 **Cooke, C., and Alwine, J.C.** (2002). Characterization of specific protein-RNA
798 complexes associated with the coupling of polyadenylation and last-intron
799 removal. Molecular and cellular biology **22,** 4579-4586.

800 **Cooke, C., Hans, H., and Alwine, J.C.** (1999). Utilization of splicing elements and
801 polyadenylation signal elements in the coupling of polyadenylation and last-
802 intron removal. Molecular and cellular biology **19,** 4971-4979.
803 **Drechsel, G., Kahles, A., Kesarwani, A.K., Stauffer, E., Behr, J., Drewe, P.,**
804 **Ratsch, G., and Wachter, A.** (2013). Nonsense-mediated decay of alternative
805 precursor mRNA splicing variants is a major determinant of the Arabidopsis
806 steady state transcriptome. Plant Cell **25,** 3726-3742.
807 **Durek, P., Schmidt, R., Heazlewood, J.L., Jones, A., MacLean, D., Nagel, A.,**
808 **Kersten, B., and Schulze, W.X.** (2010). PhosPhAt: the Arabidopsis thaliana
809 phosphorylation site database. An update. Nucleic acids research **38,** D828-
810 834.
811 **Filichkin, S., Priest, H.D., Megraw, M., and Mockler, T.C.** (2015). Alternative
812 splicing in plants: directing traffic at the crossroads of adaptation and
813 environmental stress. Current opinion in plant biology **24,** 125-135.
814 **Fortes, P., Kufel, J., Fornerod, M., Polycarpou-Schwarz, M., Lafontaine, D.,**
815 **Tollervey, D., and Mattaj, I.W.** (1999a). Genetic and physical interactions
816 involving the yeast nuclear cap-binding complex. Molecular and cellular
817 biology **19,** 6543-6553.
818 **Fortes, P., Bilbao-Cortes, D., Fornerod, M., Rigaut, G., Raymond, W., Seraphin,**
819 **B., and Mattaj, I.W.** (1999b). Luc7p, a novel yeast U1 snRNP protein with a
820 role in 5' splice site recognition. Genes & development **13,** 2425-2438.
821 **Fukumura, K., Taniguchi, I., Sakamoto, H., Ohno, M., and Inoue, K.** (2009). U1-
822 independent pre-mRNA splicing contributes to the regulation of alternative
823 splicing. Nucleic Acids Res **37,** 1907-1914.
824 **Gohring, J., Jacak, J., and Barta, A.** (2014). Imaging of endogenous messenger
825 RNA splice variants in living cells reveals nuclear retention of transcripts
826 inaccessible to nonsense-mediated decay in Arabidopsis. Plant Cell **26,** 754-
827 764.
828 **Golovkin, M., and Reddy, A.S.** (1998). The plant U1 small nuclear ribonucleoprotein
829 particle 70K protein interacts with two novel serine/arginine-rich proteins. Plant
830 Cell **10,** 1637-1648.
831 **Golovkin, M., and Reddy, A.S.** (2003). Expression of U1 small nuclear
832 ribonucleoprotein 70K antisense transcript using APETALA3 promoter
833 suppresses the development of sepals and petals. Plant physiology **132,**
834 1884-1891.
835 **Gonatopoulos-Pournatzis, T., and Cowling, V.H.** (2014). Cap-binding complex
836 (CBC). The Biochemical journal **457,** 231-242.
837 **Gouy, M., Guindon, S., and Gascuel, O.** (2010). SeaView version 4: A multiplatform
838 graphical user interface for sequence alignment and phylogenetic tree
839 building. Molecular biology and evolution **27,** 221-224.
840 **Gu, J., Xia, Z., Luo, Y., Jiang, X., Qian, B., Xie, H., Zhu, J.K., Xiong, L., Zhu, J.,**
841 **and Wang, Z.Y.** (2017). Spliceosomal protein U1A is involved in alternative
842 splicing and salt stress tolerance in Arabidopsis thaliana. Nucleic Acids Res.
843 **Guiro, J., and O'Reilly, D.** (2015). Insights into the U1 small nuclear
844 ribonucleoprotein complex superfamily. Wires Rna **6,** 79-92.
845 **Heazlewood, J.L., Durek, P., Hummel, J., Selbig, J., Weckwerth, W., Walther, D.,**
846 **and Schulze, W.X.** (2008). PhosPhAt: a database of phosphorylation sites in
847 Arabidopsis thaliana and a plant-specific phosphorylation site predictor.
848 Nucleic acids research **36,** D1015-1021.
849 **Heim, A., Grimm, C., Muller, U., Haussler, S., Mackeen, M.M., Merl, J., Hauck,**
850 **S.M., Kessler, B.M., Schofield, C.J., Wolf, A., and Bottger, A.** (2014).

851 Jumonji domain containing protein 6 (Jmjd6) modulates splicing and
852 specifically interacts with arginine-serine-rich (RS) domains of SR- and SR-like
853 proteins. Nucleic Acids Res **42,** 7833-7850.
854 **Hernandez, H., Makarova, O.V., Makarov, E.M., Morgner, N., Muto, Y., Krummel,**
855 **D.P., and Robinson, C.V.** (2009). Isoforms of U1-70k control subunit
856 dynamics in the human spliceosomal U1 snRNP. PloS one **4,** e7202.
857 **Hoffman, B.E., and Grabowski, P.J.** (1992). U1 snRNP targets an essential splicing
858 factor, U2AF65, to the 3' splice site by a network of interactions spanning the
859 exon. Genes & development **6,** 2554-2568.
860 **Hori, K., and Watanabe, Y.** (2005). UPF3 suppresses aberrant spliced mRNA in
861 Arabidopsis. The Plant journal : for cell and molecular biology **43,** 530-540.
862 **Kaida, D.** (2016). The reciprocal regulation between splicing and 3'-end processing.
863 Wiley Interdiscip Rev RNA **7,** 499-511.
864 **Kaida, D., Berg, M.G., Younis, I., Kasim, M., Singh, L.N., Wan, L., and Dreyfuss,**
865 **G.** (2010). U1 snRNP protects pre-mRNAs from premature cleavage and
866 polyadenylation. Nature **468,** 664-668.
867 **Kalyna, M., Simpson, C.G., Syed, N.H., Lewandowska, D., Marquez, Y.,**
868 **Kusenda, B., Marshall, J., Fuller, J., Cardle, L., McNicol, J., Dinh, H.Q.,**
869 **Barta, A., and Brown, J.W.** (2012). Alternative splicing and nonsense-
870 mediated decay modulate expression of important regulatory genes in
871 Arabidopsis. Nucleic Acids Res **40,** 2454-2469.
872 **Kanno, T., Lin, W.D., Fu, J.L., Chang, C.L., Matzke, A.J.M., and Matzke, M.**
873 (2017). A Genetic Screen for Pre-mRNA Splicing Mutants of Arabidopsis
874 thaliana Identifies Putative U1 snRNP Components RBM25 and PRP39a.
875 Genetics **207,** 1347-1359.
876 **Kaplan, C.D., Morris, J.R., Wu, C., and Winston, F.** (2000). Spt5 and spt6 are
877 associated with active transcription and have characteristics of general
878 elongation factors in D. melanogaster. Genes & development **14,** 2623-2634.
879 **Knop, K., Stepien, A., Barciszewska-Pacak, M., Taube, M., Bielewicz, D.,**
880 **Michalak, M., Borst, J.W., Jarmolowski, A., and Szweykowska-Kulinska,**
881 **Z.** (2016). Active 5' splice sites regulate the biogenesis efficiency of
882 Arabidopsis microRNAs derived from intron-containing genes. Nucleic Acids
883 Res.
884 **Koncz, C., Dejong, F., Villacorta, N., Szakonyi, D., and Koncz, Z.** (2012). The
885 spliceosome-activating complex: molecular mechanisms underlying the
886 function of a pleiotropic regulator. Frontiers in plant science **3,** 9.
887 **Kornblihtt, A.R., Schor, I.E., Allo, M., Dujardin, G., Petrillo, E., and Munoz, M.J.**
888 (2013). Alternative splicing: a pivotal step between eukaryotic transcription
889 and translation. Nature reviews. Molecular cell biology **14,** 153-165.
890 **Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J.U., Ratsch,**
891 **G., and Weigel, D.** (2008). Dual roles of the nuclear cap-binding complex and
892 SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis
893 thaliana. Proc Natl Acad Sci U S A **105,** 8795-8800.
894 **Lewis, J.D., Izaurralde, E., Jarmolowski, A., McGuigan, C., and Mattaj, I.W.**
895 (1996). A nuclear cap-binding complex facilitates association of U1 snRNP
896 with the cap-proximal 5' splice site. Genes & Development **10,** 1683-1698.
897 **Ling, Y., Alshareef, S., Butt, H., Lozano-Juste, J., Li, L., Galal, A.A., Moustafa,**
898 **A., Momin, A.A., Tashkandi, M., Richardson, D.N., Fujii, H., Arold, S.,**
899 **Rodriguez, P.L., Duque, P., and Mahfouz, M.M.** (2017). Pre-mRNA splicing
900 repression triggers abiotic stress signaling in plants. The Plant journal : for cell
901 and molecular biology **89,** 291-309.

902 **Lorkovic, Z.J., and Barta, A.** (2008). Role of Cajal bodies and nucleolus in the
903 maturation of the U1 snRNP in Arabidopsis. PLoS One **3,** e3989.
904 **Lorkovic, Z.J., Wieczorek Kirk, D.A., Klahre, U., Hemmings-Mieszczak, M., and**
905 **Filipowicz, W.** (2000). RBP45 and RBP47, two oligouridylate-specific hnRNP-
906 like proteins interacting with poly(A)+ RNA in nuclei of plant cells. RNA **6,**
907 1610-1624.
908 **Marquez, Y., Hopfler, M., Ayatollahi, Z., Barta, A., and Kalyna, M.** (2015).
909 Unmasking alternative splicing inside protein-coding exons defines exitrons
910 and their role in proteome plasticity. Genome research **25,** 995-1007.
911 **Merret, R., Martino, L., Bousquet-Antonelli, C., Fneich, S., Descombin, J., Billey,**
912 **E., Conte, M.R., and Deragon, J.M.** (2013). The association of a La module
913 with the PABP-interacting motif PAM2 is a recurrent evolutionary process that
914 led to the neofunctionalization of La-related proteins. RNA **19,** 36-50.
915 **Muller-McNicoll, M., Botti, V., de Jesus Domingues, A.M., Brandl, H., Schwich,**
916 **O.D., Steiner, M.C., Curk, T., Poser, I., Zarnack, K., and Neugebauer, K.M.**
917 (2016). SR proteins are NXF1 adaptors that link alternative RNA processing to
918 mRNA export. Genes & development **30,** 553-566.
919 **Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata,**
920 **R., Kawai, T., Tanaka, K., Niwa, Y., Watanabe, Y., Nakamura, K., Kimura,**
921 **T., and Ishiguro, S.** (2007). Improved Gateway binary vectors: high-
922 performance vectors for creation of fusion constructs in transgenic analysis of
923 plants. Bioscience, biotechnology, and biochemistry **71,** 2095-2100.
924 **Naro, C., Jolly, A., Di Persio, S., Bielli, P., Setterblad, N., Alberdi, A.J., Vicini, E.,**
925 **Geremia, R., De la Grange, P., and Sette, C.** (2017). An Orchestrated Intron
926 Retention Program in Meiosis Controls Timely Usage of Transcripts during
927 Germ Cell Differentiation. Developmental cell **41,** 82-93 e84.
928 **Ohtani, M., and Sugiyama, M.** (2005). Involvement of SRD2-mediated activation of
929 snRNA transcription in the control of cell proliferation competence in
930 Arabidopsis. The Plant journal : for cell and molecular biology **43,** 479-490.
931 **Papp, I., Mur, L.A., Dalmadi, A., Dulai, S., and Koncz, C.** (2004). A mutation in the
932 Cap Binding Protein 20 gene confers drought tolerance to Arabidopsis. Plant
933 molecular biology **55,** 679-686.
934 **Pimentel, H., Parra, M., Gee, S.L., Mohandas, N., Pachter, L., and Conboy, J.G.**
935 (2016). A dynamic intron retention program enriched in RNA processing genes
936 regulates gene expression during terminal erythropoiesis. Nucleic Acids Res
937 **44,** 838-851.
938 **Prigge, M.J., and Wagner, D.R.** (2001). The arabidopsis serrate gene encodes a
939 zinc-finger protein required for normal shoot development. Plant Cell **13,** 1263-
940 1279.
941 **Puig, O., Bragado-Nilsson, E., Koski, T., and Seraphin, B.** (2007). The U1
942 snRNP-associated factor Luc7p affects 5' splice site selection in yeast and
943 human. Nucleic Acids Res **35,** 5874-5885.
944 **Raczynska, K.D., Simpson, C.G., Ciesiolka, A., Szewc, L., Lewandowska, D.,**
945 **McNicol, J., Szweykowska-Kulinska, Z., Brown, J.W., and Jarmolowski,**
946 **A.** (2010). Involvement of the nuclear cap-binding protein complex in
947 alternative splicing in Arabidopsis thaliana. Nucleic Acids Res **38,** 265-278.
948 **Raczynska, K.D., Stepien, A., Kierzkowski, D., Kalak, M., Bajczyk, M., McNicol,**
949 **J., Simpson, C.G., Szweykowska-Kulinska, Z., Brown, J.W., and**
950 **Jarmolowski, A.** (2013). The SERRATE protein is involved in alternative
951 splicing in Arabidopsis thaliana. Nucleic Acids Res.

952 **Reddy, A.S., Marquez, Y., Kalyna, M., and Barta, A.** (2013). Complexity of the
953 alternative splicing landscape in plants. Plant Cell **25,** 3657-3683.
954 **Rigo, F., and Martinson, H.G.** (2008). Functional coupling of last-intron splicing and
955 3'-end processing to transcription in vitro: the poly(A) signal couples to splicing
956 before committing to cleavage. Molecular and cellular biology **28,** 849-862.
957 **Rosel, T.D., Hung, L.H., Medenbach, J., Donde, K., Starke, S., Benes, V., Ratsch,**
958 **G., and Bindereif, A.** (2011). RNA-Seq analysis in mutant zebrafish reveals
959 role of U1C protein in alternative splicing regulation. The EMBO journal **30,**
960 1965-1976.
961 **Rowley, M.J., Bohmdorfer, G., and Wierzbicki, A.T.** (2013). Analysis of long non-
962 coding RNAs produced by a specialized RNA polymerase in Arabidopsis
963 thaliana. Methods **63,** 160-169.
964 **Roy, D., Bhanja Chowdhury, J., and Ghosh, S.** (2013). Polypyrimidine tract binding
965 protein (PTB) associates with intronic and exonic domains to squelch nuclear
966 export of unspliced RNA. FEBS letters **587,** 3802-3807.
967 **Schlaen, R.G., Mancini, E., Sanchez, S.E., Perez-Santangelo, S., Rugnone, M.L.,**
968 **Simpson, C.G., Brown, J.W., Zhang, X., Chernomoretz, A., and Yanovsky,**
969 **M.J.** (2015). The spliceosome assembly factor GEMIN2 attenuates the effects
970 of temperature on alternative splicing and circadian rhythms. Proc Natl Acad
971 Sci U S A **112,** 9382-9387.
972 **Schwab, R., Speth, C., Laubinger, S., and Voinnet, O.** (2013). Enhanced
973 microRNA accumulation through stemloop-adjacent introns. EMBO reports **14,**
974 615-621.
975 **Shao, W., Kim, H.S., Cao, Y., Xu, Y.Z., and Query, C.C.** (2012). A U1-U2 snRNP
976 interaction network during intron definition. Molecular and cellular biology **32,**
977 470-478.
978 **Shaul, O.** (2015). Unique Aspects of Plant Nonsense-Mediated mRNA Decay.
979 Trends Plant Sci **20,** 767-779.
980 **Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing,**
981 **Y.** (2014). rMATS: robust and flexible detection of differential alternative
982 splicing from replicate RNA-Seq data. Proc Natl Acad Sci U S A **111,** E5593-
983 5601.
984 **Staiger, D., and Brown, J.W.** (2013). Alternative splicing at the intersection of
985 biological timing, development, and stress responses. Plant Cell **25,** 3640-
986 3656.
987 **Stepien, A., Knop, K., Dolata, J., Taube, M., Bajczyk, M., Barciszewska-Pacak,**
988 **M., Pacak, A., Jarmolowski, A., and Szweykowska-Kulinska, Z.** (2017).
989 Posttranscriptional coordination of splicing and miRNA biogenesis in plants.
990 Wiley Interdiscip Rev RNA **8**.
991 **Sullivan, C., and Howard, P.L.** (2016). The Diverse Requirements of ARS2 in
992 nuclear cap-binding complex-dependent RNA Processing. Rna & Disease.
993 **Sun, M., Lariviere, L., Dengl, S., Mayer, A., and Cramer, P.** (2010). A tandem SH2
994 domain in transcription elongation factor Spt6 binds the phosphorylated RNA
995 polymerase II C-terminal repeat domain (CTD). The Journal of biological
996 chemistry **285,** 41597-41603.
997 **Umezawa, T., Sugiyama, N., Takahashi, F., Anderson, J.C., Ishihama, Y., Peck,**
998 **S.C., and Shinozaki, K.** (2013). Genetics and phosphoproteomics reveal a
999 protein phosphorylation network in the abscisic acid signaling pathway in
1000 Arabidopsis thaliana. Sci Signal **6,** rs8.
1001 **Wachter, A., Tunc-Ozdemir, M., Grove, B.C., Green, P.J., Shintani, D.K., and**
1002 **Breaker, R.R.** (2007). Riboswitch control of gene expression in plants by

1003 splicing and alternative 3' end processing of mRNAs. The Plant cell **19,** 3437-
1004 3450.
1005 **Wahl, M.C., Will, C.L., and Luhrmann, R.** (2009). The spliceosome: design
1006 principles of a dynamic RNP machine. Cell **136,** 701-718.
1007 **Wang, B.B., and Brendel, V.** (2004). The ASRG database: identification and survey
1008 of Arabidopsis thaliana genes involved in pre-mRNA splicing. Genome biology
1009 **5,** R102.
1010 **Wang, C., Tian, Q., Hou, Z., Mucha, M., Aukerman, M., and Olsen, O.A.** (2007).
1011 The Arabidopsis thaliana AT PRP39-1 gene, encoding a tetratricopeptide
1012 repeat protein with similarity to the yeast pre-mRNA processing protein
1013 PRP39, affects flowering time. Plant cell reports **26,** 1357-1366.
1014 **Wang, P., Xue, L., Batelli, G., Lee, S., Hou, Y.J., Van Oosten, M.J., Zhang, H.,**
1015 **Tao, W.A., and Zhu, J.K.** (2013). Quantitative phosphoproteomics identifies
1016 SnRK2 protein kinase substrates and reveals the effectors of abscisic acid
1017 action. Proc Natl Acad Sci U S A **110,** 11205-11210.
1018 **Webby, C.J., Wolf, A., Gromak, N., Dreger, M., Kramer, H., Kessler, B., Nielsen,**
1019 **M.L., Schmitz, C., Butler, D.S., Yates, J.R., 3rd, Delahunty, C.M., Hahn, P.,**
1020 **Lengeling, A., Mann, M., Proudfoot, N.J., Schofield, C.J., and Bottger, A.**
1021 (2009). Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated
1022 with RNA splicing. Science **325,** 90-93.
1023 **Will, C.L., and Luhrmann, R.** (2001). Spliceosomal UsnRNP biogenesis, structure
1024 and function. Current opinion in cell biology **13,** 290-301.
1025 **Will, C.L., and Luhrmann, R.** (2011). Spliceosome structure and function. Cold
1026 Spring Harbor perspectives in biology **3**.
1027 **Wong, J.J., Au, A.Y., Ritchie, W., and Rasko, J.E.** (2016). Intron retention in
1028 mRNA: No longer nonsense: Known and putative roles of intron retention in
1029 normal and disease biology. BioEssays : news and reviews in molecular,
1030 cellular and developmental biology **38,** 41-49.
1031 **Workman, E., Veith, A., and Battle, D.J.** (2014). U1A regulates 3' processing of the
1032 survival motor neuron mRNA. The Journal of biological chemistry **289,** 3703-
1033 3712.
1034 **Xing, D., Wang, Y., Hamilton, M., Ben-Hur, A., and Reddy, A.S.** (2015).
1035 Transcriptome-Wide Identification of RNA Targets of Arabidopsis
1036 SERINE/ARGININE-RICH45 Uncovers the Unexpected Roles of This RNA
1037 Binding Protein in RNA Processing. Plant Cell.
1038 **Yap, K., Lim, Z.Q., Khandelia, P., Friedman, B., and Makeyev, E.V.** (2012).
1039 Coordinated regulation of neuronal mRNA steady-state levels through
1040 developmentally controlled intron retention. Genes Dev **26,** 1209-1223.
1041 **Yoine, M., Ohto, M.A., Onai, K., Mita, S., and Nakamura, K.** (2006). The lba1
1042 mutation of UPF1 RNA helicase involved in nonsense-mediated mRNA decay
1043 causes pleiotropic phenotypic changes and altered sugar signalling in
1044 Arabidopsis. The Plant journal : for cell and molecular biology **47,** 49-62.
1045 **Zhan, X., Qian, B., Cao, F., Wu, W., Yang, L., Guan, Q., Gu, X., Wang, P.,**
1046 **Okusolubo, T.A., Dunn, S.L., Zhu, J.K., and Zhu, J.** (2015). An Arabidopsis
1047 PWI and RRM motif-containing protein is critical for pre-mRNA splicing and
1048 ABA responses. Nature communications **6,** 8139.
1049