Characterization of Arabidopsis LUC7 proteins reveals functions of the U1 snRNP in alternative splicing and splicing of terminal introns

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Marcella de Francisco Amorim

aus Belém, Brasilien

Tübingen

2017

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	21.09.2017
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Sascha Laubinger
2. Berichterstatter:	Prof. Dr. Klaus Harter

ACKNOWLEDGMENTS

First, I would like to thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil) (N°Processo: 12400-12-2) and DAAD (Deutscher Akademischer Austauschdienst, Germany) for the scholarship to pursue my doctoral studies at the University of Tübingen.

I thank Prof. Dr. Sascha Laubinger for the supervision during my PhD. I am really grateful for the opportunity to work in his research group. I definitely learned a lot during this time.

I would also like to thank Prof. Dr. Klaus Harter and my Technical Advisory Committee, Prof. Dr. Gerd Jürgens and Dr. Andreas Wachter, for all the assistance during these years. I am very thankful for the data and material sharing and for many suggestions and discussions.

I would like to thank Dr. Eva-Maria Willing and Emese Szabó for the bioinformatic analyses and Irina Droste-Borel for the mass spectrometry analyses.

I am grateful to Claudia Martinho and Marília Campos for the great support in the final step of my PhD.

I thank all members of the ZMBP, especially my colleagues from the Plant Physiology department for the nice working atmosphere. I thank also all members of Laubinger's lab for the time and knowledge shared.

Um obrigado especial aos meus pais que sempre me incentivaram a lutar por meus sonhos, apoiando incondicionalmente todas minhas escolhas, mesmo que essas signifiquem, muitas vezes, estar distante fisicamente de casa. Agradeço também a minha irmã pelo apoio em todos os momentos.

I am also especially grateful to Matthias for sharing so many memorable moments during the last years and for giving a great support, when all seems not to work out.

Last, I thank many friends that made my PhD journey in Germany softer, including those that even from far away manage to make me smile and keep me going.

LIST OF MAIN ABBREVIATIONS

3'ss	3' splicing site
5'ss	5' splicing site
7mG	7-methylguanosine
ABA	abscisic acid
ABH1	ABA HYPERSENSITIVE 1
BP	branch point
bp	base pair
CBC	nuclear cap-binding complex
CBP20	CAP-BINDING PROTEIN 20
CBP80	CAP-BINDING PROTEIN 80
CTD	carboxyl terminus domain
EJC	exon junction complex
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
hnRNP	heterogeneous nuclear ribonucleoprotein
ISE	intronic splicing enhancer
ISS	intronic splicing silencer
LUC7	Lethal Unless CBC 7
miRNA	micro RNA
mRNA	messenger RNA
MS	mass spectrometry
uORF	upstream open reading frame
NMD	nonsense mediate decay
nt	nucleotides
pol II	RNA polymerase II
poly(A)	poly adenine
pre-mRNA	precursor messenger RNA
pri-miRNA	primary microRNA
PTC	premature termination codon
PTT	polypyrimidine tract
RIP	RNA immunoprecipitation
RNA-seq	RNA sequencing
SE	SERRATE
snRNA	small nuclear RNA
snRNP	small ribonucleoproteins particle
SR	serine-arginine
TC	termination codon
U	uridine
U2AF	U2 auxiliary factor
uORF	upstream open reading frame
UPF	UP FRAMESHIFT
UTR	untranslated region
WT	wild-type

ABSTRACT

In eukaryotes, RNA polymerase II transcripts often contain intervening sequences called introns that have to be removed from the nascent RNA while the flanking sequences, the exons, are joined together. This process, called splicing, is an essential step before the translation of the mature mRNAs. Splicing is not only important for normal growth and development, but it also allows organisms to respond and adapt fast to changes in the environment. Despite the relevance of splicing and its broad impact in potentially all biological processes, little is known about its mechanism in plants. The removal of introns is catalyzed by the spliceosome, a highly dynamic macromolecular complex formed by five subcomplexes of small nuclear ribonucleoproteins particles (U snRNP). The focus of this study was one of these subcomplexes, the U1 snRNP, which binds to 5' splice sites and plays a fundamental role in the early steps of the splicing reaction.

This study revealed that Arabidopsis LUC7 proteins are U1 snRNP components that act in constitutive and alternative splicing mainly in a redundant manner. In addition, although LUC7A interacts with SERRATE and with the nuclear cap-binding complex -proteins that are known to be involved primarily in splicing of cap-proximal introns-, LUC7 proteins specifically promote splicing of a subset of terminal introns. It was also shown here that splicing of LUC7 dependent terminal intron is a prerequisite for the transcript nuclear export. Moreover, retention of some of these terminal introns is regulated by cold stress in wild type. In agreement with roles under stress conditions, *luc7* triple mutant displays a significant amount of stress-related genes that are up-regulated and in addition, this mutant is salt and ABA hypersensitive. Finally, global analyses revealed that first and last introns are more prone to be regulated under stress conditions uncovering an unknown bias for splicing regulation in Arabidopsis.

ZUSAMMENFASSUNG

In Eukaryonten werden pre-mRNAs oft durch nicht-kodierende Sequenzabschnitte, sogenannte Introns, unterbrochen. Der Prozess des RNA-Speißens dient dazu, Introns zu entfernen und die kodierenden Sequenzen, auch Exons genannt, zu fusionieren. mRNA-Speißen ist ein wesentlicher Schritt vor der Translation von mRNAs in Proteine und kann auf verschiedenste Weisen reguliert werden. Die Entfernung von Introns wird durch das Spleißosom katalysiert. Das Spleißosom setzt sich aus fünf Subkomplexen (U snRNP) zusammen, die von verschiedensten Proteinen und RNAs (snRNA) gebildet werden.

In Pflanzen spielt die Regulation des Speißens eine entscheidende Rolle bei unterschiedlichen Entwicklungsprozessen und Reaktionen auf veränderte Umweltbedingungen. Trotz der Relevanz des Spleißens ist relativ wenig über den Mechanismus des Spleißen und den daran beteiligten Proteinen bekannt. Der Schwerpunkt dieser Studie war die Analyse des U1 snRNP, das an 5' Spleißstellen bindet und dem wichtige Funktionen bei frühen Schritten des Spleißens zukommen. Die Untersuchung ergab, dass Arabidopsis LUC7-Proteine wichtige U1-snRNP-Komponenten sind, die beim konstitutiven und alternativen Spleißen eine vorwiegend redundant Funktion ausüben. Obwohl LUC7 mit Proteinen interagiert, die für die Entfernung von ersten Introns wichtig sind (SERRATE und der nuklearen Cap-bindenden Komplex), weisen LUC7 Proteine eine spezielle Funktion bei der Entfernung von terminalen Introns auf. Das Entfernen der terminalen Introns ist eine Bedingung für den Transport der mRNAs vom Zellkern in das Cytosol. Das Spleißen der LUC7-abhängigen terminalen Introns kann durch abiotischen Stress moduliert werden, was auf eine spezielle Bedeutung von LUC7 bei der pflanzlichen Stressantwort hindeutet. In der Tat sind *luc7* Mutanten weniger stressresistent und weisen eine Missregulation vieler Stress-regulierter Gene auf. Globale Analysen zeigten, dass das Spleißen von ersten und terminalen Introns unter Stressbedingungen vermehrt verändert wird. Zusammenfassend wurde in dieser Arbeit gezeigt, dass eine U1 snRNP Komponente, LUC7, eine wichtige Funktion bei der Entfernung von terminalen Introns spielt und dass das Spleißen von terminalen Introns eine wichtige Rolle bei der pflanzlichen Stressantwort spielen kann.

INTRODUCTION

1. RNA Polymerase II transcripts: from transcription to nuclear export

The majority of eukaryotic genes are transcribed by RNA polymerase II (pol II) and their transcripts are characterized by the addition of a 7-methylguanosine (7mG) cap to the 5' end. This process is known as capping and is the first to occur in the nascent transcript. Apart from protecting the RNA from degradation, the 7mG cap is bound co-transcriptionally by the nuclear cap-binding complex (CBC), which acts as a platform for recruiting proteins involved in further RNA processing events (Calero et al., 2002; Gonatopoulos-Pournatzis & Cowling, 2014; Topisirovic et al., 2011). The CBC consists of two proteins, CBP20 and CBP80/ABH1, that bind synergistically to the 7mG cap and stay usually attached until the transcript reaches the cytoplasm – exceptions are transcripts from the micro RNA (MIR) genes, which are further processed in the nucleus (Achkar et al., 2016; Gonatopoulos-Pournatzis & Cowling, 2014). RNA capping and CBC binding occur when the transcript is only ~25 nt and it indicates a successful pol II initiation (Hallais et al., 2013; Topisirovic et al., 2011).

Following the capping, many nascent pol II transcripts have to overcome a challenge that is the presence of intervening non-coding sequences: the introns. Most of pol II transcripts contain introns that have to be removed, while the remaining sequences, the exons, are joined together to form the mature RNA. This RNA processing is called splicing and usually occurs while the transcript is still being transcribed (co-transcriptionally), but it can also occur after it is released from the chromatin (post-transcriptionally). Splicing events can be divided into two types: constitutive splicing, when the introns are always recognized and removed in the same way; and alternative splicing, which generates, from a single gene, different isoforms based on a differential recognition of exons and introns. Both types of splicing require the same core machinery, namely spliceosome, in addition to regulatory elements that influence the splicing cycle (Bentley, 2014; Fu & Ares, 2014). The spliceosome is a highly dynamic macromolecular complex and seems to be partially recruited by the CBC, at least for the removal of cap-proximal introns (Laubinger et al., 2008; Lewis et al., 1996; Raczynska et al., 2010). Studies in mammals revealed the importance of CBC also for downstream introns. However, currently it is not clear to what extent the CBC is required for splicing throughout the whole genome (Gonatopoulos-Pournatzis & Cowling, 2014; Jiao et al., 2013; Pabis et al., 2013).

In addition to guaranteeing the correct information flow, the splicing reaction also influences the composition of proteins that are bound to the mature RNA. In a splicing-dependent manner, the exon-junction complex (EJC) is typically deposited onto the spliced RNA at ~20-24 nt upstream of the exon-exon junctions (Le Hir et al., 2016). Interestingly, the presence of EJC influences back the splicing by, for instance, enhancing splicing of neighboring introns (Boehm & Gehring, 2016; Hayashi et al., 2014). Furthermore, the EJC seems to facilitate RNA export mainly for shorter transcripts, while longer RNAs are only slightly affected. After the export, the RNA still contains EJC complexes attached, which seem to stimulate translation before they are displaced from the RNA in the first round of translation (Le Hir et al., 2016).

In the final step of transcription, the messenger RNA (mRNA) is cleaved and a poly adenine (poly(A)) tail is added to the 3' end. These processes require several *cis*-elements that are recognized by the cleavage/polyadenylation machinery, a large protein complex formed by four major subcomplexes (Kaida, 2016). The poly(A) signal, which is located at ~15-30 nt upstream to the cleavage site, is a key *cis*-element that dictates where the cleavage occurs and in addition, it acts as signal for transcription termination (Elkon et al., 2013; Proudfoot, 2011). Interestingly, the 3' end formation can be coupled with the removal of terminal introns and these processes influence each other. *In vitro* studies in animals revealed that the recognition of terminal introns by the spliceosome stimulates the 3' end formation; on the other hand, the poly(A) site recognition can stimulate the splicing of the terminal introns (Niwa & Berget, 1991; Niwa et al., 1990). Direct interactions between the splicing components and the cleavage/polyadenylation factors are important for this coupling (Bentley, 2014).

The addition of a poly(A) tail to most of the pol II transcripts is a prerequisite for their export to the cytoplasm and it also protects the RNA from degradation through their 3' end. Many poly(A) binding proteins bind to this tail and assist in the nuclear export, as well as in translation (Hunt et al., 2008). Interestingly, its length varies among species and also during the RNA life cycle. The poly(A) tail length influences RNA stability and translation efficiency by dictating the amount of poly(A) binding proteins that are attached to the RNA (Eckmann et al., 2011; Hunt et al., 2008; Subtelny et al., 2014). Once completely processed the mature RNA is exported from the nucleus to the cytoplasm via the nuclear pore complex. The RNA export can be coupled with splicing. For instance,

splicing factors from the serine/arginine (SR) family seem to act as adaptor proteins for the RNA export (Muller-McNicoll et al., 2016).

Capping, splicing and the 3' end formation, all these RNA processing events occur, at least partially, co-transcriptionally and in a highly coordinated manner. Even for introns that are only removed after transcription termination, the splicing commitment occurs still during the transcription, since the splicing machinery is recruited while the RNA is being synthesized (Bentley, 2014). This co-transcriptionality allows a mechanistic coupling between these RNA processing events and transcription, which is mediated partially by the unstructured carboxyl terminus domain (CTD) from the RNA pol II. The CTD is formed by heptapeptide repeats that are highly modified during the transcription, controlling in this way the recruitment of specific set of proteins at the appropriate time. For instance, factors involved in capping, splicing and 3' end formation are recruited by the CTD and this recruitment is important for proper RNA processing (Harlen & Churchman, 2017; Harlen et al., 2016). In plants, the transcription machinery also associates with factors from the splicing and 3' end formation machinery (Antosz et al., 2017). Interestingly, the co-transcriptionality allows also a crosstalk between the chromatin and RNA processing events. In animals, chromatin with its modified histones is known to affect splicing directly by the binding of adaptor proteins to specific histone modifications, which recruits then specific splicing factors (Luco et al., 2011). The chromatin status can also impact splicing indirectly by affecting the rate of transcription, which in turn influences the time for splicing sites recognition. For instance, slower transcription allows more time for splicing sites to be recognized before a competing splicing site appears, promoting then exon inclusion (Braunschweig et al., 2013). In plants, a very recent study indicates also a direct impact of histone modifications on alternative splicing. In this study, it was shown that histone H3 lysine 36 tri-methylation (H3k36me3) seems to be important for the differential alternative splicing events that occur under changes in temperature (Pajoro et al., 2017). Thus, chromatin modifications and splicing are also connected in plants.

2. The pre-mRNA splicing machinery: constitutive and alternative splicing

Intronic sequences are removed from the precursor messenger RNA (pre-mRNA) by the spliceosome, a macromolecular complex formed by 5 subcomplexes of small nuclear ribonucleoprotein particles (snRNP) and many additional non-snRNP proteins. In

most eukaryotes, including plants, there are two types of spliceosome: (i) U2-dependent, also known as major spliceosome, which catalyzes the removal of U2-type introns; (ii) and U12-dependent, the minor spliceosome, responsible for the less abundant U12-type introns (<0.5% of introns in a genome) (Reddy et al., 2013; Turunen et al., 2013; Will & Luhrmann, 2011). The focus of this study was on the major spliceosome, which is formed by the subcomplexes U1, U2, U4, U5 and U6 snRNP. Each of these subcomplexes contains a specific uridine rich small nuclear RNA (U snRNA), a heptameric ring of Sm proteins (except for U6 snRNP harboring a ring with Sm/Lsm), specific core proteins and associated ones (Will & Luhrmann, 2011).

The spliceosome assembles anew at each intron and requires conserved sequences in the pre-mRNA. Two essential intronic sequences define the exon-intron boundaries: the 5' splicing site (5'ss) and the 3' splicing site (3'ss) (Figure 1). In addition, an adenine (A) close to the 3'ss forms the so-called branch point (BP), which is important for the first nucleophilic attack in the splicing reaction. In higher eukaryotes, there is also another conserved sequence, the polypyrimidine tract (PTT), located between the branch point and the 3'ss. Apart from these intronic sequences, *cis*-acting pre-mRNA elements can also influence the spliceosome assembly and composition. These elements include exonic and intronic splicing enhancers (ESE and ISE) and exonic and intronic splicing silencers (ESS and ISS) (Figure 1). These sequences are typically short and bound by *trans*-acting factors, which are RNA-binding proteins that either promote or repress the recruitment of the spliceosomal complexes (Fu & Ares, 2014; Kornblihtt et al., 2013). Among *trans*-acting factors, the SR family and the heterogeneous nuclear ribonucleoproteins (hnRNP) represent the two major classes (Staiger & Brown, 2013). SR proteins are often viewed as positive regulators promoting exon inclusion by binding to exons and thereby facilitating the U1 snRNP and U2 auxiliary factors (U2AF35 and U2AF65) recruitment in the initial step of splicing (Fu & Ares, 2014). Members of the hnRNP family, on the other hand, are usually seen as negative regulators abolishing the binding of the spliceosome to the 5'ss, 3'ss or BP. Apart from this negative impact, hnRNP may also enhance splicing by facilitating spliceosome recruitment or by avoiding that a splicing repressor binds to the RNA (Wachter et al., 2012). In reality, the regulatory effect of SR and hnRNP proteins seems to depend on their binding position in the transcripts and in the surrounding context. These *cis*-elements and *trans*-acting factors are involved

in both constitutive and alternative splicing and work in coordination with the core splicing machinery to define the functional splice sites (Fu & Ares, 2014; Howard & Sanford, 2015; Wachter et al., 2012).

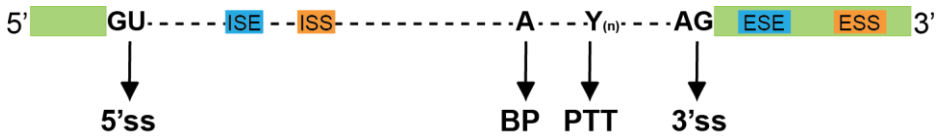


Figure 1: Intronic and exonic sequences involved in RNA splicing. Introns are defined by the 5' splicing site (5'ss) and the 3' splicing site (3'ss) and contains also the polypyrimidine tract (PTT, Y = Pyrimidine, C/T) and the branch point (BP) sequences. Additional *cis*-acting elements also participate in the splicing reaction by promoting or repressing the recruitment of the spliceosomal and they include: exonic and intronic splicing silencers (ESS and ISS) and exonic and intronic splicing enhancers (ESE and ISE). Exons are indicated in green and introns are represented by lines with the consensus sequences.

While in constitutive splicing all these elements cooperate to create always the same outcome, alternative splicing relies basically on the competition among more than one 5'ss and/or 3'ss. Each splicing site has a different strength that is dictated by the deviation from the consensus sequence, which is associated with efficient recognition (Fu & Ares, 2014; Kornblihtt et al., 2013). The splicing sites strengths and also their positions in the transcript, together with the presence of *cis*-elements and *trans*-acting factors, contribute to the splicing decision. This alternative selection of splicing sites generates different alternative splicing events, among which the most common are: intron retention, alternative 5'ss, alternative 3'ss, exon-skipping and mutually exclusive exons (Figure 2). Alternative isoforms might also be created by the combination of many of these events (Reddy et al., 2013).

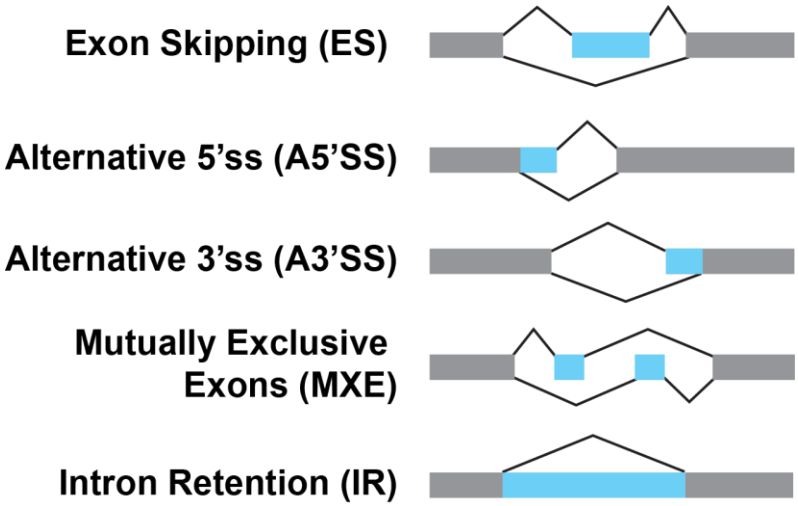


Figure 2: Types of alternative splicing. Constitutive spliced exons are represented in gray, while the alternative splicing exons are in blue.

The spliceosome composition and its assembly on RNAs have been extensively studied in yeast and animals through *in vitro* complex assembly assays and purifications of many spliceosome complexes followed by mass spectrometry (MS) analyses (Koncz et al., 2012; Will & Luhrmann, 2011). On the other hand, the plant spliceosome composition is only inferred based on sequence homology with their yeast and animals counterparts (Koncz et al., 2012; Reddy et al., 2013; Wang & Brendel, 2004). Due to the fact that the components are conserved, it is possible to expect that the basic principle of the splicing mechanism is similar. However, a different mechanism of intron recognition most likely exists, since introns from animals could not be processed in plants, while plant introns could be normally spliced out in an *in vitro* system in animals (Barta et al., 1986; Brown et al., 1986; Hartmuth & Barta, 1986). Part of the difference may rely on the contrasting size of introns, which in plants is around 160 bp, while the average of mammalian introns is about 5 Kb (Meyer et al., 2015; Reddy et al., 2013).

In the canonical view of splicing, the initial step is the binding of the U1 snRNP to the 5'ss followed by the binding of the U2 auxiliary factors U2AF35 and U2AF65 to the 3'ss and polypyrimidine tract, respectively. At the same time, the SF1 factor also binds to the branch point and forms the so-called spliceosomal complex E. These interactions play an essential role in the initial 5'ss and 3'ss recognition. In the next step, the U2 snRNP associates with the branch point, displacing the SF1 factor and thereby forming the complex A. The pre-assembled U4/U5/U6 trimeric complex is then recruited and after

many rearrangements, U1 and U4 snRNP are ejected generating the activated spliceosome. Subsequent steps involve two-transesterification reactions that lead to intron removal and exons ligation. Finally, the spliceosome components dissociate, releasing the mature processed RNA (Wahl et al., 2009; Will & Luhrmann, 2011).

3. U1 snRNP subcomplex composition and functions

The U1 snRNP is the first subcomplex to bind the RNA and it has an important role in the 5'ss selection (Will & Luhrmann, 2011). In yeast and animals, this subcomplex is formed by the U1 snRNA, the heptameric ring of Sm, three core proteins - U1-70k, U1-A and U1-C - and associated proteins. In Arabidopsis, its composition is inferred based on sequence homology with human and yeast components. Apart from the Sm and the core proteins, around 10 U1-accessories proteins seem to be part of this subcomplex, such as LUC7A, LUC7B, LUC7RL, PRP39A, PRP39B, PRP40A and PRP40B (Koncz et al., 2012; Reddy et al., 2013; Wang & Brendel, 2004). In plants, so far only two putative U1 accessories were characterized, but their associations with the U1 snRNP were not shown. Two *prp39a* mutants were analyzed and revealed a late flowering phenotype due to increased expression of the flowering time regulator *FLOWERING LOCUS C* (*FLC) (Wang* et al*., 2007)*. In addition, *rbm25* mutants did not display any development defect under normal conditions, but this gene seems to be required for response to the phytohormone abscisic acid (ABA) (Cheng et al., 2017; Zhan et al., 2015). Moreover, the reverse genetic approach to knock-down the U1 core protein U1-70K in flowers by antisense RNA revealed strong floral defects, indicating a crucial role in flower development (Golovkin & Reddy, 2003). These results indicate the importance of the U1 snRNP subcomplex in essential processes of plant development and during ABA-dependent stress response. Despite this relevance, no further studies are available on U1 snRNP components in plants.

Interestingly, different U1 snRNP subcomplexes exist in animals and thus, most likely also in plants. For instance, analysis of native U1 snRNP components in Hela cells suggests the presence of at least 4 major subcomplexes (Hernandez et al., 2009). In addition, there are many different U1 snRNA that may contribute to the variability of the U1 snRNP subcomplex (Guiro & O'Reilly, 2015; Wang & Brendel, 2004). Two other factors that might also contribute to U1 subcomplex diversification are: alternative splicing of U1

components and their phosphorylation status. For instance, human U1-70k is known to have two isoforms that associate with the U1 snRNA. One of these isoforms contains a serine residue that can be phosphorylated, enhancing the interaction with U1-C, which might then interfere with the 5'ss selection (Guiro & O'Reilly, 2015).

In animal cells, although each U snRNP subcomplex is present in equal amount in the spliceosome, U1 snRNP is more abundant in the cell than the other spliceosomal subcomplexes (Kaida et al., 2010). Therefore, it has long been suggested that U1 snRNP could have different functions apart from splicing. Some older studies pointed out towards a role of U1 snRNP in inhibiting the 3'end formation; however, those were gene-specific studies (Proudfoot, 2011). More recently, genome-wide analyses in animals showed that U1 snRNP protects the nascent RNA from a premature cleavage and polyadenylation and in this way, guarantees the correct transcripts length (Berg et al., 2012; Kaida et al., 2010). Furthermore, this mechanism seems to be used to globally regulate gene expression. During neuron activation, the high increase in transcription rate generates shortened RNAs due to the scarce availability of the U1 snRNP. This effect was suppressed by overexpression of the U1 snRNA (Berg et al., 2012). In addition, the same U1 snRNP property is also responsible for setting the promoter directionality in animals. Most of the active genes in animals are transcribed in both directions, but the upstream antisense transcripts encounter early termination. The reason for this is the low abundance of U1 binding sites and high abundance of poly(A) sites in these upstream antisense transcripts, while transcripts in the sense direction have the opposite tendency (Almada et al., 2013).

Apart from the role in splicing and regulation of the 3' end formation, there are evidences for the involvement of U1 snRNP in transcription initiation. In animals, the U1 snRNA associates with the general transcription initiation factor TFIIH and also with the initiation factor TAF-15 (Jobert et al., 2009; Kwek et al., 2002). Hence, U1 snRNA seems to stimulate transcription initiation, but the exact mechanism of the U1 snRNP effect on transcription remains to be elucidated (Guiro & O'Reilly, 2015).

Finally, *in vitro* studies revealed that the absence of U1 snRNP in the splicing reaction could be compensated by high amount of SR proteins, suggesting that U1-independent splicing could exist (Crispino et al., 1994; Tarn & Steitz, 1994). Later on, the first report on a naturally occurring U1-independent splicing event was described in

humans (Fukumura et al., 2009). Whether U1-independent splicing occurs also in plants is still not known.

4. Alternative splicing coupled to Non-sense Mediate Decay

Alternative splicing allows the expansion of the coding genome and partially explains the increase in complexity of higher eukaryotes (Kornblihtt et al., 2013; Reddy et al., 2013; Yu et al., 2016). In addition, it is also used as a way to regulate gene expression (Braunschweig et al., 2014; Reddy et al., 2013). One way that alternative splicing regulates the levels of transcripts is through the Non-sense Mediate Decay (NMD). NMD is a eukaryotic quality control mechanism that promotes RNA degradation in a translation dependent manner. Fundamentally, it relies on the recognition and elimination of transcripts containing termination codon (TC) in a suboptimal context: (i) typically TC located more than 50-55 nt upstream of the last exon-exon junction; or (ii) TC at more than 300-350 nt to the 3' end of the transcript, generating then a long 3'UTR. Introns at ≥50-55 nt downstream to a TC lead, after their splicing, to the deposition of an EJC that cannot be removed during the translation and therefore may disturb the transcript translation termination. Transcripts with a long 3'UTR in turn seem to usually promote a less efficient translation termination due to a large distance between the TC and the poly(A) tail, which then impairs protein interactions required for a proper termination (Lykke-Andersen & Jensen, 2015; Shaul, 2015).

Similar to mammals, plants transcripts that are targeted to NMD contain one of these features: long 3'UTR (≥300-350 nt), introns located ≥50-55 nt downstream to the TC, an in-frame premature termination codon (PTC) or an upstream open reading frame (uORF) (Figure 3A-D) (Drechsel et al., 2013; Kalyna et al., 2012; Shaul, 2015). In fact, the presence of a PTC generates transcripts harboring a long 3'UTR and/or transcripts with introns downstream (Figure 3C). In all these cases, the termination codon is located in a non-optimal context leading to ribosome stalling or abnormal termination, which then triggers NMD. Although these are hallmarks of NMD targets, there are many transcripts harboring such features that are NMD insensitive (Kalyna et al., 2012; Leviatan et al., 2013).

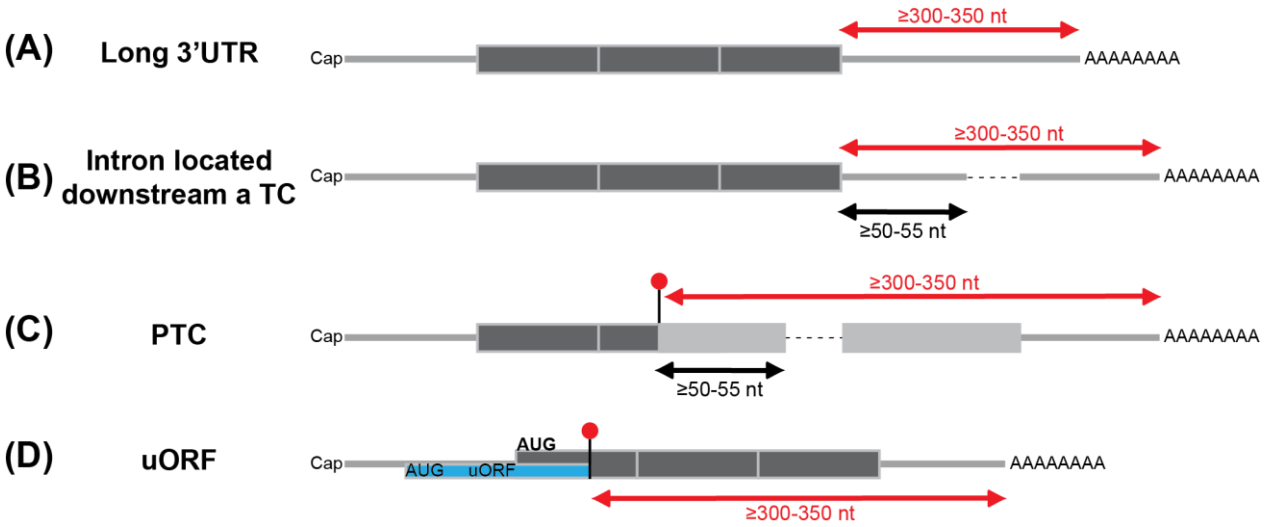


Figure 3: Features of transcripts target to nonsense-mediate decay. (A) Long 3'UTR (≥300-350 nt); (B) Introns located ≥50-55 nt downstream to a termination codon (TC); (C) Premature termination codon (PTC) harboring a long 3'URT or/and introns located ≥50-55 nt downstream to the PTC; (D) an upstream open reading frame (uORF) that, for instance, overlaps with the start codon of the main open reading frame. PTCs are shown in red signals; red arrows indicate the long 3'UTR; and black arrows indicate the distance of the next downstream intron from the TC/PTC.

The three UP FRAMESHIFT (UPF) proteins - UPF1, UFP2 and UFP3 - are core proteins from the NMD machinery and their impairments lead to NMD inactivation and thus, accumulation of NMD-sensitive transcripts. Conserved in plants, these core proteins are responsible for recruiting factors involved in the general mRNA decay, guiding the NMD targets to endonucleolytic cleavage, decapping and/or deadenylation (Lykke-Andersen & Jensen, 2015; Reddy et al., 2013).

Alternative splicing can generate transcripts containing TC in suboptimal context by adding, for instance, an in-frame PTC. Alternative splicing coupled to NMD (AS-NMD) helps to maintain the balance of specific gene expression network by producing isoforms with different RNA stability (Lykke-Andersen & Jensen, 2015). Many genes, including splicing factors, have their expression regulated in this way (Reddy et al., 2013; Song et al., 2009; Sureshkumar et al., 2016). In animals and plants, for instance, splicing factors from the SR family undergo extensive alternative splicing and many isoforms are known to be NMD targets (Morrison et al., 1997; Palusa & Reddy, 2010). Interestingly, many splicing factors seems to auto-regulate their transcripts levels through a negative feedback loop (Reddy et al., 2013). In plants, for instance, it was shown that when PTB1,

PTB2 and PTB3 proteins are present in high amounts, they bind to their own transcripts and lead to the production of an alternative splicing isoform that is NMD sensitive, which in turn reduce the production of PTB proteins. Moreover, PTB1 and PTB2, the two closest related, are also known to cross-regulate their expression. This mechanism of auto- and cross-regulation of PTBs expression is conserved among plants and animals (Stauffer et al., 2010; Wachter et al., 2012).

5. Splicing, development and environmental/stress signals in plants

Splicing is necessary to convey the correct information and it is therefore an essential process during the whole organism's life. Moreover, splicing provides transcriptome and proteome flexibility via alternative splicing, assisting the organisms to adapt and respond fast to changes in the environment (Staiger & Brown, 2013).

Developmental and environmental/stress signals potentially impact alternative splicing in many ways (Figure 4). For instance, changes in the cell content may affect the RNAs structures, which might interfere with the binding of proteins that require a specific RNA secondary structure. A different RNA structure may also impact splicing outcome by masking/unmasking splicing sites or *cis*-regulatory sequences, affecting then their recognition (Reddy et al., 2013). Moreover, developmental/environmental cues could affect the transcription rate by changing, for instance, the chromatin status. Transcription rate impacts in turn splicing by affecting the time available for competition among the splicing sites emerging in the nascent transcript. Chromatin modifications may also interfere with the recruitment of different splicing factors and thus affect the splicing outcome (Luco et al., 2011). Furthermore, changes in the expression of splicing factors, or in their own alternative splicing, impact the splicing pattern of many other transcripts. Finally, the activity/function of splicing factors can be regulated through phosphorylation, which can affect their subcellular localization and their interacting partners (Reddy et al., 2013).

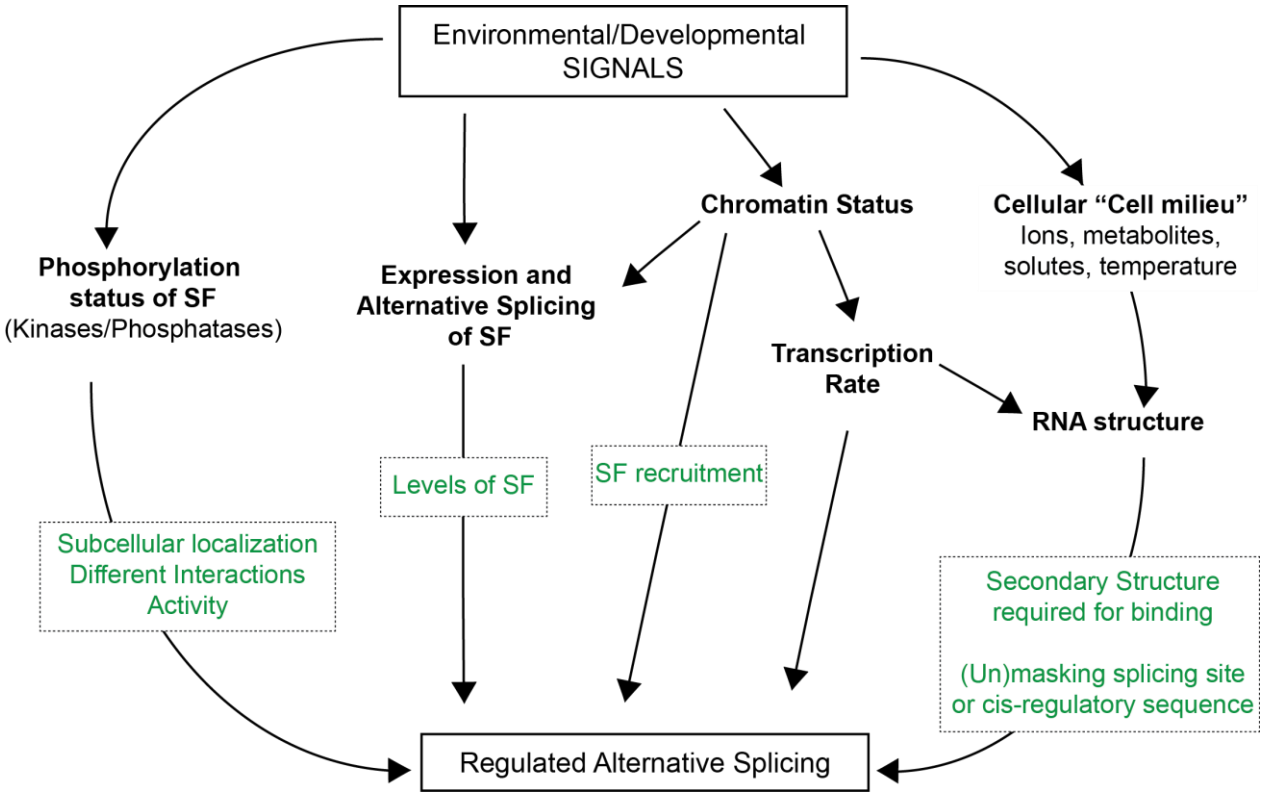


Figure 4: Developmental and environmental/stress signals potentially regulate alternative splicing through different mechanisms. SF, splicing factor. Based on Reddy et al. (2013).

The relevance of splicing in plant development is underscored by a great number of splicing factors' mutants that display a wide range of developmental defects (Ali et al., 2007; Moll et al., 2008; Staiger & Brown, 2013; Volz et al., 2012). Moreover, splicing factors are differentially regulated among tissues and also during the development, indicating that distinct alternative splicing isoforms contribute for the tissue differentiation and plant development (Lopato et al., 1996; Palusa et al., 2007). For instance, alternative splicing regulates flowering time under ambient temperature through a differential splicing of the flowering repressor FLOWERING LOCUS M (FLM), which triggers degradation of the new isoform via NMD (Lee et al., 2013; Pose et al., 2013; Sureshkumar et al., 2016). Interestingly, it was shown recently that minor switch in temperature is sufficient to promote changes in the splicing pattern of many genes apart from the FLM. Among the differentially spliced genes, splicing-related genes are enriched, suggesting that the environmental signal impacts the spliceosome composition via changes in the splicing of

splicing factors, which in turn impact downstream targets, assisting the plant to adapt to the new environmental condition (Verhage et al., 2017).

Drastic changes in the environment are perceived as stresses by the plants and are known to induce global changes in the alternative splicing pattern (Ding et al., 2014; Leviatan et al., 2013). In plants, the majority of the SR proteins undergo alternative splicing under environmental stresses, such as heat and cold (Duque, 2014; Palusa et al., 2007). Furthermore, the stress-related phytohormone ABA is known to affect the expression of many SR genes (Cruz et al., 2014). ABA signaling acts through the kinase SnRNK2 (2.2, 2.3 and 2.6), which is known also to phosphorylate many splicing factors (Umezawa et al., 2013; Wang et al., 2013). Interestingly, it was shown that the localization of splicing factors changes upon stress conditions and upon differential phosphorylation status (Ali et al., 2003). The fact that many splicing factors were found in screening for stress tolerance/sensitivity further indicates that splicing plays a role in stress response (Staiger & Brown, 2013).

All in all, alternative splicing seems to be employed in many developmental and environmental/stress signaling pathways. However, details into the mechanisms used to regulate these varieties of processes is just at the beginning of being understood.

6. The dual role of nuclear cap-binding complex and SERRATE in plants

The nuclear cap-binding complex (CBC) with its two subunits (CBP20 and ABH1) is involved in many RNA processing events (Gonatopoulos-Pournatzis & Cowling, 2014). In some cases, CBC interacts with the zinc-finger protein SERRATE (SE) and forms the SE/CBC complex, which acts as a platform for the recruitment of different proteins that dictate the RNA fate. In plants, it has been described that SE/CBC complex plays a role in at least two processes: the processing of primary microRNA (pri-miRNA) into mature microRNA (miRNA) and splicing of pre-mRNA, mainly of cap-proximal introns (Laubinger et al., 2008; Raczynska et al., 2010; Raczynska et al., 2013).

MiRNAs are a class of endogenous small RNAs with 20-24 nt (predominantly 21 nt in plants) that regulates gene expression post-transcriptionally. Based on sequence complementarity, miRNAs recognize and repress their target transcripts by cleavage or translation repression (Rogers & Chen, 2013). In plants, the miRNA biogenesis occurs inside the nucleus and involves processing of the pri-miRNA by a complex that is recruited

by SE/CBC through SE direct interaction with other proteins such as DICER and HYL1 (Achkar et al., 2016).

The majority of miRNAs in plants originates from microRNA genes that form independent transcript units and are transcribed by pol II. Some of these transcripts contain introns that have to be removed by the splicing machinery. In addition, there is also a portion of miRNAs that is encoded within introns of other genes and, in this case, the splicing of the host transcript can influence the miRNA processing. In both cases, there is a crosstalk between the miRNA biogenesis machinery and the spliceosome, where the U1 snRNP seems to play an important role (Bielewicz et al., 2013; Knop et al., 2016; Schwab et al., 2013; Stepien et al., 2017).

GOALS OF THIS WORK

The focus of this study was the U1 snRNP subcomplex, which is involved in the earliest step of splicing, playing a fundamental role in the 5'ss recognition. In animals, U1 snRNP fulfills also splicing-independent functions turning this subcomplex even more attractive to be studied in plants. Arabidopsis U1 snRNP contain the U1 snRNA, 3 core proteins and apparently at least 10 putative accessories proteins, such as the ones belonging to the LUC7 (Lethal Unless CBC 7) family. In yeast, LUC7 affect 5'ss selection and also seems to mediate the U1 snRNP subcomplex interaction with the nuclear cap binding complex (CBC). In plants, CBC interacts with the zinc finger protein SERRATE (SE) and this complex is required for splicing of cap-proximal introns. It has long been suggested that SE/CBC might act in concert with the U1 snRNP for the removal of first introns. This hypothesis was reinforced by previous data from a yeast two-hybrid screening performed in Laubinger´s lab, where AthLUC7RL was found as putative SE interacting partner. Taking all into account, LUC7 proteins were strong candidates to bridge U1 snRNP and SE/CBC and could potentially act together regulating the removal of first introns. Intending to shed light on U1 snRNP subcomplex in plants and on the SE/CBC-U1 snRNP putative interaction, the functional characterization of LUC7 family was carried out.

MATERIAL AND METHODS

1. Plasmid constructions

For the expression of C-terminal FLAG- and YFP-tagged LUC7 proteins expressed from their endogenous regulatory elements, 2100 bp, 4120 bp and 2106 bp upstream of the ATG start codon of *LUC7A*, *LUC7B* and *LUC7RL*, respectively, to the last coding nucleotide were PCR-amplified using Phusion High-Fidelity DNA Polymerase (Thermo Scientific) and inserts were subcloned in pCR8/GW/TOPO (Invitrogen). Primers used are listed (Supplement S1). Entry clones containing the genomic sequences were recombined with pGWB10 and pGWB540 using Gateway LR Clonase II (Invitrogen) to generate binary plasmids containing LUC7A::LUC7A-FLAG, LUC7B::LUC7B-FLAG, LUC7RL::LUC7RL-FLAG, LUC7A::LUC7A-eYFP, LUC7B::LUC7B-eYFP and LUC7RL::LUC7RL-eYFP. For the co-localization studies, entry vector containing the coding sequence of U1-70k was recombined with pGWB654 for the expression of 35S::U1-70k-mRFP (Nakagawa et al., 2007). The entry vector U1-70k was available in the lab.

2. T-DNA insertion mutants, Arabidopsis stable lines and complementation assays

All mutants were in the Columbia-0 (Col-0) background. *luc7a-1* (SAIL_596_H02) and *luc7a-2* (SAIL_776_F02), *luc7b-1* (SALK_144681), *luc7rl-1* (SALK_077718) and *luc7rl-2* (SALK_130892C) were isolated by PCR-based genotyping (Supplement S1). *luc7* double and triple mutants were generated by crossing individual mutants. All other mutants used in this study (*abh1-285*, *cbp20-1*, *se*-1, *lba-1,* and *upf3-1*) were described elsewhere (Hori & Watanabe, 2005; Laubinger et al., 2008; Papp et al., 2004; Prigge & Wagner, 2001; Yoine et al., 2006). The Arabidopsis line overexpressing free GFP was generated using the vector pBinarGFP and was kindly provided by Dr. Andreas Wachter (Wachter et al., 2007). For the complementation analyses, the *luc7* triple mutant (*luc7a-2 b-1 rl-1*) were stably transformed with *LUC7A::LUC7A-FLAG*, *LUC7B::LUC7B-FLAG*, *LUC7RL::LUC7RL-FLAG* and *LUC7A::LUC7A-eYFP*. Arabidopsis transformations were performed as described in Clough and Bent (1998).

3. Plant material and growth conditions

Plants on soil were grown in long days conditions (16h light/ 8h dark) at 20°C/18° day/night, 40% humidity. The size of *luc7* mutants was assessed by measuring the longest rosette leaf from the middle point of the plant to the tip of the leaf in plants growing on soil for 21 days. For all experiments performed with seedlings, seeds were surface-sterilized, plated on 1/2 MS medium with 0.8% phytoagar, stratified for 2-4 days and grown for 7 days in continuous light at 22°C. The cold treatment was performed by transferring plates with 7 day-old seedlings to ice-water for 60 min, which was done in triplicates. For the root assay, seedlings growing for 4 days on vertical plates were transferred to mock plates or plates containing 75 mM or 150 mM of NaCl and grown for more 11 days vertically. Root growth rate per day was assessed by measuring in ImageJ the root length in the days 2 and 9 after transfer. To test mutants' sensitivity to ABA (+) (Sigma - A4906), 1/2 MS solid plates were also supplemented with 1% sucrose and seedlings were grown for 10 days.

4. Tobacco transient expression and confocal microscopy

Agrobacterium containing *LUC7A::LUC7A-eYFP*, *LUC7B::LUC7B-eYFP*, *LUC7RL::LUC7RL-eYFP*, *35S::U1-70k-mRFP* and a silencing suppressor p19 were grown overnight at 28°C (Voinnet et al., 2003). These pre-cultures were used in a 1:6 dilution to prepare cultures, which were grown for 4 hours at 28°C. Cultures were centrifuged at 3220 x rcf for 20 min, 4°C and resuspended in half volume of Infiltration Medium (10mM MgCl2, 10mM MES-KOH pH 5.6, 100 µM Acetosyringone). The OD 600nm was measured and adjusted to 0.6-0.8. After adjusting OD, samples were mixed with P19 (1:1) or for the colocalization LUC7A or LUC7RL:U1-70k:P19 were combined in different ratios. *Nicotiana benthamiana* plants were infiltrated and subcellular co-localization checked after 3 days. The subcellular localization studies were performed using confocal microscope (Leica TCS SP2 or SP8).

5. Phylogenetic analysis

AthLUC7A (AT3G03340) protein sequence was analyzed in Interpro (https://www.ebi.ac.uk/interpro/) to retrieve the Interpro ID for the conserved Luc7-related domain (IPR004882). The sequence for *Saccharomyces cerevisiae* (strain ATCC 204508_S288c) was obtained in Interpro. Plants sequences were extracted using

BioMart, selecting for the protein domain IPR004882 on Ensembl Plants (http://plants.ensembl.org/). The following genomes were searched: *Amborella trichopoda* (AMTR1.0 (2014-01-AGD)); Arabidopsis thaliana (TAIR10 (2010-09-TAIR10)); *Brachypodium distachyon* (v1.0); *Chlamydomonas reinhardtii* (v3.1 (2007-11-ENA)); *Physcomitrella patens* (ASM242v1 (2011-03-Phypa1.6)); *Selaginella moellendorffii* (v1.0 (2011-05-ENA)); *Oryza sativa Japonica* (IRGSP-1.0); and *Ostreococcus lucimarinus* (ASM9206v1). The phylogenetic analysis was performed with Seaview (Version 4.6.1), where sequences were aligned using Muscle and Maximum likehood (PhYML) was employed with 1000 bootstraps (Gouy et al., 2010).

6. RNA extractions, RT-PCR and qRT-PCR

RNA extractions were performed with Direct-zol™ RNA MiniPrep Kit (Zymo Research). Samples were treated with DNAseI and cDNA syntheses were carried out with RevertAid First Strand cDNA Synthesis (Thermo Scientific), using oligo dT primers unless different stated. For the quantification of mature miRNAs, cDNA synthesis was performed with a gene specific stem loop-primers (Supplement S1). Standard PCRs for the splicing analysis were performed with DreamTaq DNA Polymerase (Thermo Scientific). Quantitative RT-PCR (qRT-PCR) was performed using the Maxima SYBR Green (Thermo Scientific) in the Bio-Rad CFX 384. The relative expressions were calculated with the 2^(-ΔΔCT) method using PP2A or ACTIN as a control. Primers used for PCR and qRT-PCR are listed (Supplement S1). Statistical test (t-test) on the RT-qPCR data, when performed, was done in the values 2^(-ΔCT). For the experiment to asses intron retention events under cold stress, t-test was performed in the ratio $[2^{(-\Delta CT)}_{\text{Unspliced}} / 2^{(-\Delta CT)}_{\text{Total RNA}}]$.

7. Subcellular fractionation

Two grams of seedlings were ground in N_2 liquid and resuspended in 4 ml of Honda buffer (0.44 M sucrose, 1.25% Ficoll 400, 2.5% Dextran T40, 20 mM HEPES KOH pH 7.4, 10 mM $MgCl_2$, 0.5% Triton X-100, 5 mM DTT, 1 mM PMSF, protease inhibitor cocktail [ROCHE] supplemented with 8U/ml of Ribolock®). The homogenate was filtered through 2 layers of Miracloth, which was washed with 1 ml Honda buffer. From the filtrate, 300 µl was removed as total fraction and kept on ice. The filtrates were centrifuged at 1,500 *g* for 10 min, 4°C for pelleting the nuclei and supernatants were transferred to a new tube. The

supernatants were centrifuged at 13 000 x g, 4 °C, 15 min and 300 µl were kept on ice as cytoplasmic fraction. The nuclei pellet were resuspended in 1 ml of Honda buffer and centrifugation at 1,800 *g* for 5 min. This washing step was repeated four to five times and at the end, the pellets were resuspended in 300 µl of Honda buffer. To all the fractions (total, cytoplasmic and nuclei), 900 µl of TRI Reagent (Sigma) was added. Samples were vortexed and incubated at room temperature for 5 min. After the incubation, 180 µl of chloroform was added to the samples, which were vortexed and incubated at room temperature for 10 min. After centrifugation at 10 000 rpm for 20 min, 4°C, the aqueous phase were transferred to a new tube and RNA extracted with Direct-zol™ RNA MiniPrep Kit (Zymo Research). The organic phase was also saved and protein extractions were performed following the manufacturer instructions (TRI Reagent). The extracted RNAs were measured and the same amount was used for the total and cytoplasmic fractions, while the nuclei fractions were adjusted together. RNAs were treated with DNAseI and cDNA syntheses with random primers were carried out with RevertAid First Strand cDNA Synthesis (Thermo Scientific) and (-) RT reactions were performed for the nuclei fraction with ½V of the treated RNA. Proteins extracted were assessed by western blot and the following antibodies were used: H3 (~ 17KDa / ab 1791, Abcam) and 60S ribosomal (~ 23,7-29KDa / L13-1, AS09478, Agrisera). Two replicates were performed. Standard PCRs performed to assess intron retention events as described above. Primers spanning only the retained last intron were used (Supplement S1).

8. RNA immunoprecipitation

RNA immunoprecipitation (RIP) was performed similarly to previously described protocols (Rowley et al., 2013; Xing et al., 2015). For this, 3 g of seedlings from WT Col-0 and LUC7A::LUC7A-YFP, *luc7 a-2, b-1, rl-1*) were washed four times with milli-Q water. The cross-linking was performed with 1% formaldehyde by applying vacuum at 85 KPa once for 2 min and then reapplying for 13 min. The quenching was done emerging the seedlings in a 0.125M glycine solution and vacuum at 85 KPa was again applied for 5 min. Seedlings washed 3-4 times with milli-Q water were frozen in liquid N_2. Samples were then ground in N_2 liquid and transferred to a falcon tube containing 25 ml of Honda buffer (supplemented with 8U/ml Ribolock®). After resuspending the plant material in the buffer, the extracts were filtrated through two layers of Miracloth. Additional 10 ml of Honda buffer

was used to wash Miracloth pipetting up and down. Filtrates were centrifuged at 1500 x g, 15 min, 4°C and pellets resuspended in 1ml of Honda buffer were transferred to a 2 ml tube followed by centrifugation at 1900 x g, 5 min, 4°C. After washing twice with 1 ml HONDA buffer, pellets were resuspended in 950 µl of freshly prepared Nuclei Lysis Buffer (50 mM Tris-HCl, pH8.0, 10 mM EDTA, 1% SDS, 1 mM PMSF, and 1x protease inhibitor cocktail, 160U/ml Ribolock®) and sonicated using a Covaris E220 (Duty Cycle: 20%; Peak intensity: 140; Cycles per Burst: 200; Cycle time: 3'). The sonicated samples were centrifuged at 16000 x g for 10 min, 4°C and supernatants were transferred to a new tube. The nuclear extracts (NE) were aliquoted (200µl), flash freezed in liquid N_2 and stored at –80°C. Before the immunoprecipitations, 30 µl of GFP-Trap®_A (Chromotek) were washed three times with washing buffer (150 mM NaCl, 20 mM Tris HCl, pH 8.0, 2 mM EDTA, 1% Triton X-100 and 0.1% SDS) and blocked with t-RNA (washing buffer supplemented with 0,1µg/ml tRNA and 40U/ml Ribolock®) for at least 2 hours rotating at 4°C. After blocking, beads were washed once with washing buffer (supplemented with 1mM PMSF and 40U/ml Ribolock®). For the immunoprecipitations, 200 µl of NE were diluted with 1800 µl of ChIP dilution buffer (1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8, 167 mM NaCl, and 350 U/mL Ribolock®) and INPUT (20 µl) was removed. The remaining NE were transferred to a 2 ml tube containing blocked GFP-trap and incubated overnight rotating at 4°C. Inputs were also left overnight at 4°C. The beads were washed four times with 1 ml of washing buffer (supplemented with 1mM PMSF and 40U/ml Ribolock®). The first two washings were done only by inverting the tubes, while in the other two, the beads were incubated for 5 min at 4°C under rotation. For each washing, beads were centrifuged at 2 000 rcf, 2 min. Before the last washing, beads were transferred to a new 1.5 ml tube. For the elution of protein-RNA complexes from the beads, 120 µl of RIP elution buffer (100 mM Tris-HCl, pH8.0, 10 mM EDTA, 1%SDS and 800 U/ml Ribolock®) were added, beads were briefly vortexed and incubated on a rotating wheel for 10 min at room temperature. After centrifugation at 1500 rcf, the eluted was transfered to a new tube. Elution was repeated by adding 120 µl of RIP elution buffer to the beads and incubating at 65°C for 10 min, 800rpm. Both eluates were combined (240 µl). At the same time, 220µl of RIP elution buffer was added to 20 µl of INPUT. To the beads and INPUT, 2.4 µl of 20mg/ml Proteinase K (Roche) was added and samples were incubated at 65°C for 1 hour. RNAs were then extracted using RNeasy Plant Mini Kit

(QIAGEN) following the manufacturer's instructions. The RNA were treated with DNAseI (Thermo Scientific) and samples were split in half for the (-)RT reaction. cDNA syntheses were prepared with SuperScript™ III Reverse Transcriptase (Invitrogen) using hexamers. qRT-PCRs were performed with QuantiNova™ SYBRR Green PCR (QIAGEN).

9. Preparation of mRNA-seq libraries and Illumina sequencing

Total RNAs from 7-day-old seedlings of WT, *se-1* and *luc7 a-2,b-1,rl-1* mutants were extracted with Direct-zol™ RNA MiniPrep Kit (Zymo Research). Poly(A) RNAs were enriched from 4 µg of total RNAs using NEBNext Oligo $d(T)_{25}$ Magnetic Beads (New England Biolabs) and the libraries were then prepared using ScriptSeq™ Plant Leaf kit (Epicentre), following the manufacturer's instruction. Single-end sequencing was performed on Illumina HiSeq2000.

10. RNA-seq libraries: Mapping, differential expression analysis and splicing analysis

(Analysis performed by Dr. Eva-Maria Willing)

RNA-seq reads for each replicate were aligned against the *Arabidopsis thaliana* reference sequence (TAIR10) using tophat (v2.0.10, -p2, -a 10, -g 10, -N 10, --read-edit-dist 10, --library-type fr-secondstrand, --segment-length 31, -G TAIR10.gff). Next, cufflinks (version 2.2.1) was used to extract FPKM counts for each expressed transcript generating a new annotation file (transcripts.gtf), where the coordinates of each expressed transcript can be found. Cuffcompare (version 2.2.1) was then used to generate a non-redundant annotation file containing all reference transcripts in addition to new transcripts expressed in at least one of the nine samples (cuffcmp.combined.gtf). The differential expression analysis was performed with cuffdiff (version 2.2.1) between wt/*luc7* triple using the annotation file generated by cuffcompare (FDR<2 and FC>0,05). For the splicing analysis, the same alignment files generated by tophat and annotation files generated by cuffcompare (cuffcmp.combined.gtf) were used as input for MATS (version 3.0.8) in order to test for differentially spliced transcripts (Shen et al., 2014).

11. Global analysis of intron regulation under stress conditions

(Analysis performed by Emese Xochitl Szabo)

For the analyses of intron retention under stress conditions, published data sets were analyzed (accession numbers SRP035234 and SRP049993) (Ding et al., 2014; Schlaen et al., 2015). Reads were aligned to the *Arabidopsis thaliana* Ensembl3 33 genome and to the annotation GTF file (ftp://ftp.ensemblgenomes.org/pub/release33 /plants/fasta/arabidopsis_thaliana/dna/Arabidopsis_thaliana.TAIR10.dna.toplevel.fa.gz; ftp://ftp.ensemblgenomes.org/pub/release-33/plants/gtf/arabidopsis_thaliana/ Arabidopsis_thaliana.TAIR10.33.gtf.gz) using TopHat2 applying following parameters: tophat2 -p 10 -i 10 -I 1000 -G Arabidopsis_thaliana.TAIR10.33.gtf Arabidopsis_thaliana.TAIR10.dna.toplevel.fa. After alignment, mock-treated samples were used to generate an expressed background for the respective dataset using featureCounts from the Rsubread package (Kersey et al., 2016; Kim et al., 2013; Liao et al., 2013). Read numbers and gene lengths per genes (featureCounts -T 6 -R -p -F GTF -J -G Arabidopsis_thaliana.TAIR10.dna.toplevel.fa -aArabidopsis_thaliana.TAIR10.33. gtf) were collected and TPM values were calculated using an in-house script. Log_2 transformed values of expressed genes were visualized with ggplot2 version 2.1.0, and based on the density plot, threshold of expressed genes was defined as $TPM_{expressed} > 0.6$.

Intron retention events were identified using rMATS with the following parameters: python RNASeq-MATS.py -b1 untreated.bam -b2 treated.bam -gtf Arabidopsis_thaliana.TAIR10.33.gtf -o output_dir -t paired -len 101 (Shen et al., 2014). After filtering the outputs (p value < 0.05, FDR < 0.05), we categorized introns based on their position and annotation. In case of a few ambiguous hits, they were manually re-categorized. For the categorization of introns in first, middle and last introns, we used the GTF annotation file (Ensembl 33) and selected genes with 3 or more introns and TPM > 0.6. The intron distribution of all expressed genes served as a background reference, to which the distribution of retained introns under stress conditions was compared. To test for significance of changes in intron distribution, Fisher's exact test was employed since we assumed a normal distribution.

12. Gene Ontology (GO) Analysis

GO analysis were performed for up and down regulated genes in *luc7 a-2, b-1,rl-1* and *se-1* in AGriGO (http://bioinfo.cau.edu.cn/agriGO/analysis.php) using Singular Enrichment Analysis (SEA) and TAIR10 as background. For the statistical test, Fisher was employed together with the multi-test adjustment method Yekutieli (FDR under dependency) and significance level of 0.05 (Du et al., 2010). GO enrichment for the LUC7 dependent last intron, Bar Utoronto (http://bar.utoronto.ca/ntools/cgi-bin/ntools_classification_superviewer.cgi) was also employed.

13. Immunoprecipitations for mass spectrometry analysis

LUC7A immunoprecipitation was performed using a complemented line LUC7A::LUC7A-eYFP (line 20.3.1) and as controls WT plants and WT plants expressing 35S::GFP were employed. Four independent biological replicates were performed. Seedlings (4g) were ground in N_2 liquid and 1V of extraction buffer (50 mM Tris-Cl pH 7.5, 100 mM NaCl, 0.5% Triton X-100, 5% Glycerol, 1mM PMSF, 100 µM MG132, 3x Complete Protease Inhibitor Cocktail EDTA-free and Plant specific protease Inhibitor, Sigma P9599) was added. After defrosting, samples were incubated on ice for 30 min and centrifuged at 3220 rcf for 30 min at 4°C. Remaining debris were removed by filtering the extracts through two layers of Miracloth. For each immunoprecipitation, 20 µl of GFP-Trap®_MA (gtma-20, Chromotek) in a 5 ml tube was washed twice with 1 ml of washing buffer (50 mM Tris-Cl pH 7.5, 100 mM NaCl, 0.2% Triton X-100) and once with 0.5 ml of IP buffer. The same amount of extract (~5 ml) was then transferred to the washed GFP-trap and samples were incubated for 3 hours in a rotating wheel at 4°C. After this incubation time, tubes were centrifuged at 800 rcf for 1-2 min and supernatant discarded. The beads were resuspended in 1 ml of washing buffer (above) and transferred to a 1.5 ml tube, where beads were washed 4 to 5 times with washing buffer using magnetic separation. At the end, beads were resuspended in 30-40 µl of 2x Laemmili Buffer containing 200 mM DTT and incubated at 80°C for 10min. Short gel purifications (SDS-PAGE) were performed and gels slices were digested with Tripsin. LC-MS/MS analyses were performed in two mass spectrometers. Samples from R10 to R14 were analyzed on a Proxeon Easy-nLC coupled to Orbitrap Elite (method: 90 min, Top10, HCD). Samples from R15 to R17 were analysed on a Proxeon Easy-nLC coupled to OrbitrapXL (method: 90 min, Top10, CID). Samples

from R18 to R20 analysis on a Proxeon Easy-nLC coupled to OrbitrapXL (method: 130 min, Top10, CID). All the replicates were processed together on MaxQuant software (Version 1.5.2.8. with integrated Andromeda Peptide search engine) with a setting of 1% FDR and the spectra were searched against *Arabidopsis thaliana* Uniprot database (UP000006548_3702_complete_20151023. fasta). Putative interacting proteins were selected based on the number of peptide present in the IPs sample compared with the respective controls. Raw data were deposited publically (N° PXD006127). LC-MS/MS analyses were performed by Irina Droste-Borel (Proteome Center, Tübingen).

14. Coimmunoprecipitations

LUC7A immunoprecipitation was performed using a complemented line LUC7A::LUC7A-eYFP (line 20.3.1) and WT plants expressing 35S::GFP were employed. Seedlings (~4g) were ground in N_2 liquid and 1-1.5V of extraction buffer (see below) was added. After defrosting, samples were incubated on ice for 20-30 min and centrifuged at 3220 rcf for 30 min at 4°C. Remaining debris were removed by filtering the extracts through two layers of Miracloth. For each immunoprecipitation, 20 µl of GFP-Trap®_A or GFP-Trap®_MA (Chromotek) in a 5 ml tube was washed twice with 1 ml of washing buffer (50 mM Tris-Cl pH 7.5, 100 mM NaCl, 0,2% Triton X-100) and once with 0,5 ml of IP buffer. After removing 100µl for the INPUT, the same amount of extract (~5 ml) was then transferred to the washed GFP-trap and samples were incubated for 3 hours in a rotating wheel at 4°C. After this incubation time, tubes were centrifuged at 2000 rcf for 1-2 min and supernatant discarded. The beads were resuspended in 1 ml of washing buffer (see above) and transferred to a 1,5 ml tube, where beads were washed 4 to 5 times with washing buffer using magnetic separation. At the end, beads were resuspended in 100 µl of 2x Laemmli with 200mM Laemmili and to the INPUT, 25 µl of 5x Laemmli with 500mM was added. Samples were denaturated at 80°C for 10min.

Two extraction buffers were used with slightly different composition: EB1 (50 mM Tris-Cl pH 7.5, 100 mM NaCl, 0,5% Triton X-100, 1mM PMSF, 1x Complete Protease Inhibitor Cocktail EDTA-free and Plant specific protease Inhibitor, Sigma P9599); and EB2 (50 mM Tris-Cl pH 7.5, 100 mM NaCl, 1% Triton X-100, 5% Glycerol, 1mM PMSF, 100 µM MG132, Complete Protease Inhibitor Cocktail EDTA-free, Plant specific protease

Inhibitor, Sigma P9599 and Phospho Inhibitor cocktail 2 and 3 (Sigma –p5726 and p0044). The interaction between LUC7A and CBP20 was not clearly detected in the EB2.

15. Western Blot Analysis

Proteins samples were resolved in SDS-PAGE (10% self-casted SDS-PAGE or 4–20% Mini-PROTEAN® TGX™ precast gel, Biorad) using 20-25mA/gel. All gels were transferred on PVDF membranes (GE Healthcare) and blocking was performed with 1 x Roti®-Block (Roth). Membranes were incubated with appropriate antibodies in 1XPBS with 1-3% low fat milk. Immunodetections were performed using the following first antibodies CBP20 (AS09 530, Agrisera), ABH80 (AS09 531, Agrisera), SE (Agrisera) and GFP (11 814 460 001, Roche) in the concentrations ranging from 1:1000 to 1:2500 (SE). Appropriate second antibodies were used at 1:10 000 to 1:30 000. Detection was performed using Amersham ECL Kit (GE Healthcare Life Sciences) and a cooled charge coupled devices (CCD) camera to capture chemiluminescence.

10%SDS-PAGE (for 2 gels)

Separating gel 10%		Stacking gel 5%	
30% Acrylamid	5 ml	30% Acrylamid	0.83 ml
1 M TrisHCl, pH 8.8	5.63 ml	1 M TrisHCl, pH 6.8	0.63 ml
10%SDS	150 µl	10%SDS	50 µl
H2O	4.1 ml	H2O	3.46 ml
10% APS	120 µl	10% APS	25 µl
TEMED	12 µl	TEMED	5 µl

Solutions for western-blot:

- 10x SDS-Running buffer: to 1 L, 144g Glycin, 30g Tris base and 10g SDS
- 10x Transfer buffer: 1.5 M Glycin and 200mM Tris-base
  When diluting 1 x for use, add 20% (v/v) ethanol
- 10x PBS: 1.37M NaCl, 81mM Na2HPO4, 27mM KCl, 14.7mM KH_2PO_4
- 5x Laemmli: 250 mM Tris-HCl pH 6.8; 50% (v/v) glycerol, 10% (w/v) SDS; 1 mg/ml bromophenol blue. Add freshly 500 mM DTT.

16. Collaborations

- Dr. Eva-Maria Willing (Max Planck Institute for Plant Breeding Research, Cologne) performed bioinformatic analyses of the RNA sequencing libraries (WT, *luc7 a-2 b-1 rl-1* and *se-1*);
- Emese Xochitl Szabo (Centre for Plant Molecular Biology, Tübingen) performed bioinformatic analyses involving available RNA-seq data sets from Arabidopsis treated with cold and salt stress (Ding et al., 2014; Schlaen et al., 2015);
- Irina Droste-Borel (Proteome Center, Tübingen) performed the LC-MS/MS analyses.

RESULTS

1. Functional characterization of Arabidopsis LUC7 proteins reveals involvement of U1 snRNP in alternative splicing and uncovers a specialized function in removal of terminal introns

The results presented in this section are part of a manuscript submitted and also available in the preprint server *bioRxiv*:

de Francisco Amorim, M., Willing, E.-M., Szabo, E. X., Droste-Borel, I., Macek, B., Schneeberger, K., & Laubinger, S. (2017). Arabidopsis U1 snRNP subunit LUC7 functions in alternative splicing and preferential removal of terminal introns. *bioRxiv*. doi: 10.1101/150805

1.1 LUC7, a family of conserved zinc-finger proteins, redundantly control plant development

Lethal Unless CBC 7 (LUC7) was first identified in a screen for synthetic lethality in a yeast strain lacking the nuclear cap-binding complex (CBC), which is involved in many RNA processing events including splicing (Fortes et al., 1999b; Gonatopoulos-Pournatzis & Cowling, 2014). Further studies revealed that yeast LUC7 (yLUC7) belongs to the U1 snRNP subcomplex and seems to mediate the U1snRNP-CBC interaction (Fortes et al., 1999a). LUC7 proteins form a small family characterized by a conserved LUC7-related domain (IPR004882), which contains two zinc fingers (C_3H and C_2H_2-type) (https://www.ebi.ac.uk/interpro/). It was shown for yLUC7 that the first zinc finger (Znf1) directly binds to the pre-mRNA in the exon upstream to the 5'ss, a region that is in close proximity with the CBC (Puig et al., 2007). LUC7 proteins may have also an additional C-terminal Arginine/Serine-rich (RS) domain, which is known to be target of post-translational modifications and to mediate protein-protein interactions (Heim et al., 2014; Puig et al., 2007; Webby et al., 2009). Yeast has only one LUC7 protein that lacks this RS rich domain and the deletion of this unique LUC7 is lethal (Fortes et al., 1999a).

Arabidopsis thaliana has three *LUC7* genes (*AthLUC7A*, *AthLUC7B* and *AthLUC7RL*), which seem to be broadly expressed in the plant (http://jsp.weigelworld.org/expviz/expviz.jsp) (Schmid et al., 2005). Blastp in NCBI using yeast LUC7 revealed that in Arabidopsis, the LUC7RL has the highest similarity (Supplement S2). Further alignment showed that, similar to yLUC7, the AthLUC7RL lacks a conserved stretch of 80 amino acids that is located between the two zinc fingers in AtLUC7A and

AtLUC7B (Supplement S3). A phylogenetic analysis using *Saccharomyces cerevisiae* as outgroup indicates that the LUC7 family can be divided in two clades: LUC7A/B and LUC7RL *(*Figure 5). In addition, the analysis revealed that both unicellular algae used in the phylogeny -*Chlamydomonas reinhardtii* and *Ostreococcus lucimarinus*- contain a single *LUC7* gene belonging to LUC7RL clade. On the other hand, from the moss *Physcomitrella patens* on, one can find proteins belonging to both clades, suggesting that the split into LUC7A/B and LUC7RL occurred early during the evolution of land plants.

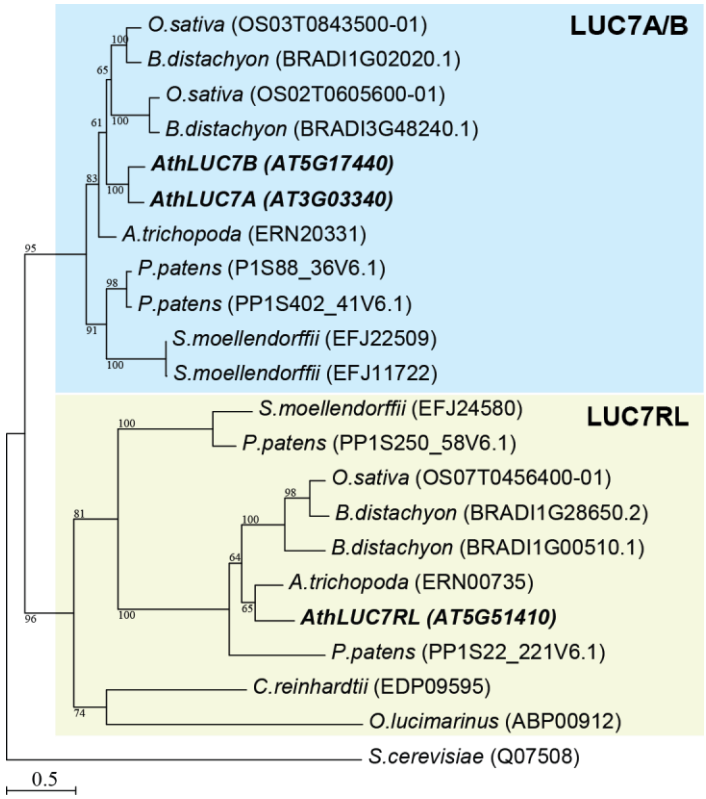


Figure 5: Phylogenetic analysis of LUC7 in plants using *Saccharomyces cerevisiae* as an external group. The analysis was performed in Seaview using Muscle for sequence alignment. Maximum likelihood (PhYML) was employed with 1000 bootstraps.

In order to assess the function of this family in plants, Arabidopsis T-DNA insertion lines in all three *LUC7* genes were analyzed (Figure 6A). Single and double *luc7* mutants growing under long day conditions were indistinguishable from wild-type (WT) plants (Figure 6B). However, *luc7* triple mutant exhibit a wide range of developmental defects, including reduced apical dominance and smaller rosette leaf size (Figure 6B-D). Intending to prove that impairments of *LUC7* functions were indeed responsible for the observed

phenotypes, a WT copy of *AtLUC7A*, *AthLUC7B* and *AthLUC7RL* were separately introduced in the triple mutant. Each of the *LUC7* genes was sufficient to restore the phenotypes of *luc7* triple mutant (Figure 6E). These results reveal that the phenotypes observed in the triple mutant is due to impairments of LUC7 functions and it suggests that *LUC7* genes act redundantly to control Arabidopsis development.

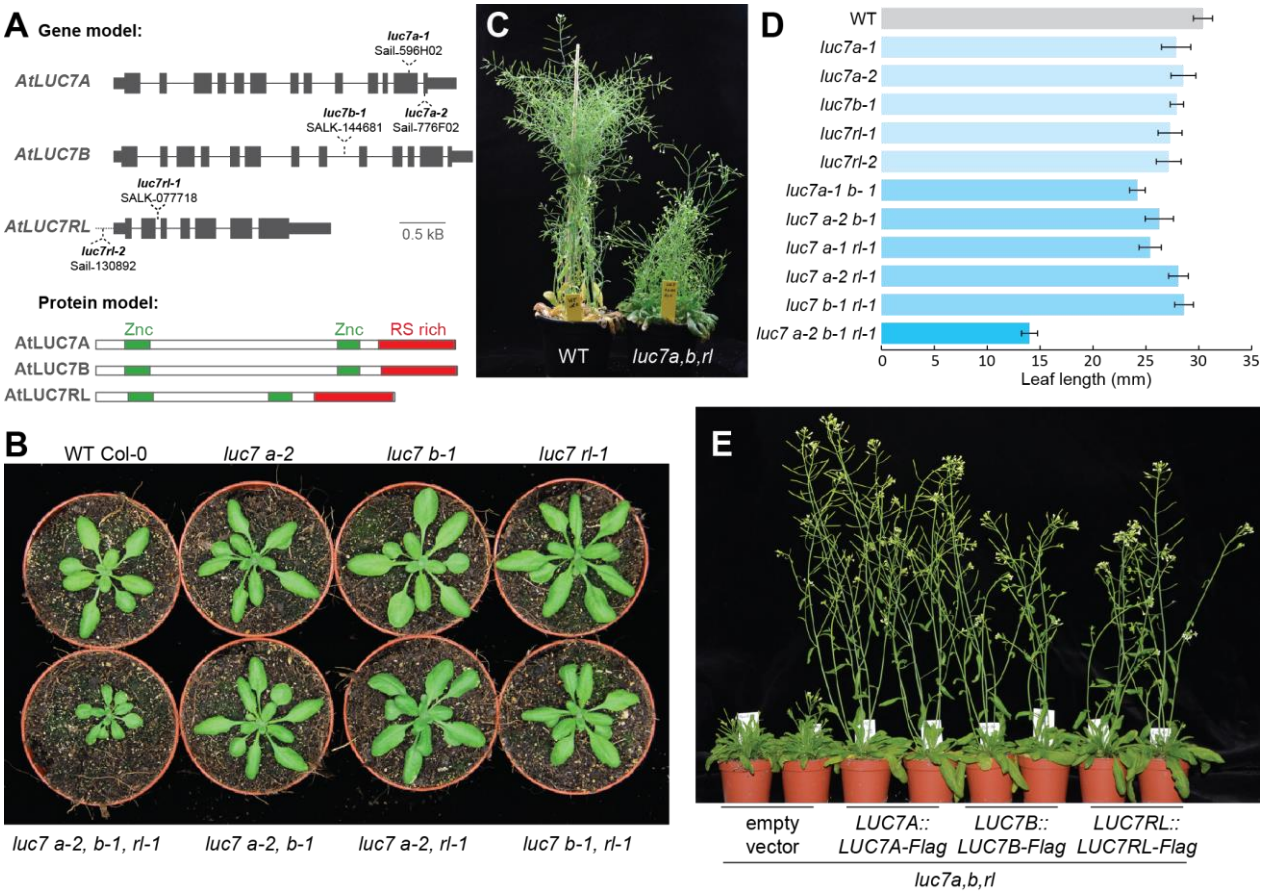


Figure 6: Arabidopsis LUC7 proteins redundantly control plant development. (A) Exon/intron structure of *Arabidopsis thaliana LUC7A*, *LUC7B* and *LUC7RL* indicating the positions of T-DNA insertions (upper); AthLUC7 proteins model displaying the two zinc-fingers (green) and the Arginine-Serine rich C-terminal domains (red) (lower). (B) WT and *luc7* mutants after 21 days growing under long day conditions. (C) WT and *luc7* triple mutant growing for 45 days in the greenhouse. (D) Length of the longest rosette leaf of 21 days-old WT, *luc7* single, double and triple mutants plants growing under long day conditions. Leaves of 10-15 individual plants were measured. Error bars denote ± standard error of the mean (SEM). (E) Complementation of *luc7a-2,b-1,rl-1* mutants by *LUC7A*, *LUC7B* and *LUC7RL* genomic constructs. Transformation with the "empty" binary vector was performed as a control. Two independent transgenic T1 lines for each construct are shown.

1.2 LUC7 proteins as U1 snRNP components and interacting partners in plants

In human and yeast the protein composition of the U1 snRNP is known based on many *in vitro* complex assembly and purifications of spliceosome complexes followed by mass spectrometry analyses (Koncz et al., 2012; Will & Luhrmann, 2011). On the other hand, in plants the composition is inferred from sequence homology with human and yeast counterparts (Koncz et al., 2012; Reddy et al., 2013; Wang & Brendel, 2004). For this reason, it was necessary to prove that Arabidopsis LUC7 proteins are also U1 subunits. If LUC7 is part of the U1 subcomplex it should be associated with U1 specific components: the U1 snRNA and the three core U1 proteins (U1-A, U1-70k and U1-C).

In order to test whether LUC7 is found in a complex with the U1 specific small RNA, RNA immunoprecipitation (RIP) experiments were performed using the *luc7 triple* mutant carrying a functional *LUC7A::LUC7A-eYFP* rescue construct and, as a control, WT plants (Figure 7A). LUC7A-eYFP co-immunoprecipitates the U1 snRNA, but not two unrelated, but abundant RNAs: U3 small nucleolar RNA (U3 snoRNA) and *ACTIN* mRNA (Figure 7B). Moreover, small amounts of U2 snRNA were also found in association with LUC7A-eYFP, which is in agreement with the fact that U1 and U2 snRNP interact in the initial step of splicing. These results strongly suggest that Arabidopsis LUC7s are indeed U1 snRNP components.

The next step was to check the subcellular localization of Arabidopsis LUC7 proteins and their co-localization with a core U1 snRNP subunit. LUC7A localized to the nucleus in the nucleoplasm, but not to the nucleolus in Arabidopsis complemented lines expressing *LUC7A::LUC7A-eYFP* (Figure 7C). Transient expressions in *Nicotiana benthamiana* revealed the same localization for all Arabidopsis LUC7 (Figure 7D). Due to LUC7 proteins redundancy and to the high similarity between LUC7A and LUC7B, co-localizations were performed only with LUC7A and LUC7RL. The results indicate that both, LUC7A and LUC7RL, partially co-localize with U1-70K in the nucleoplasm (Figure 7E). While U1-70K form also some distinct speckles, Arabidopsis LUC7 proteins have only a homogenous distribution in the nucleoplasm. It is interesting to note that the co-expression with U1-70K did not change LUC7A and LUC7RL subcellular localization. To sum up, the fact that these proteins partially colocalize not only indicates that they might be part of the same complex, but in addition suggests that U1-70K complexes without LUC7 also exist.

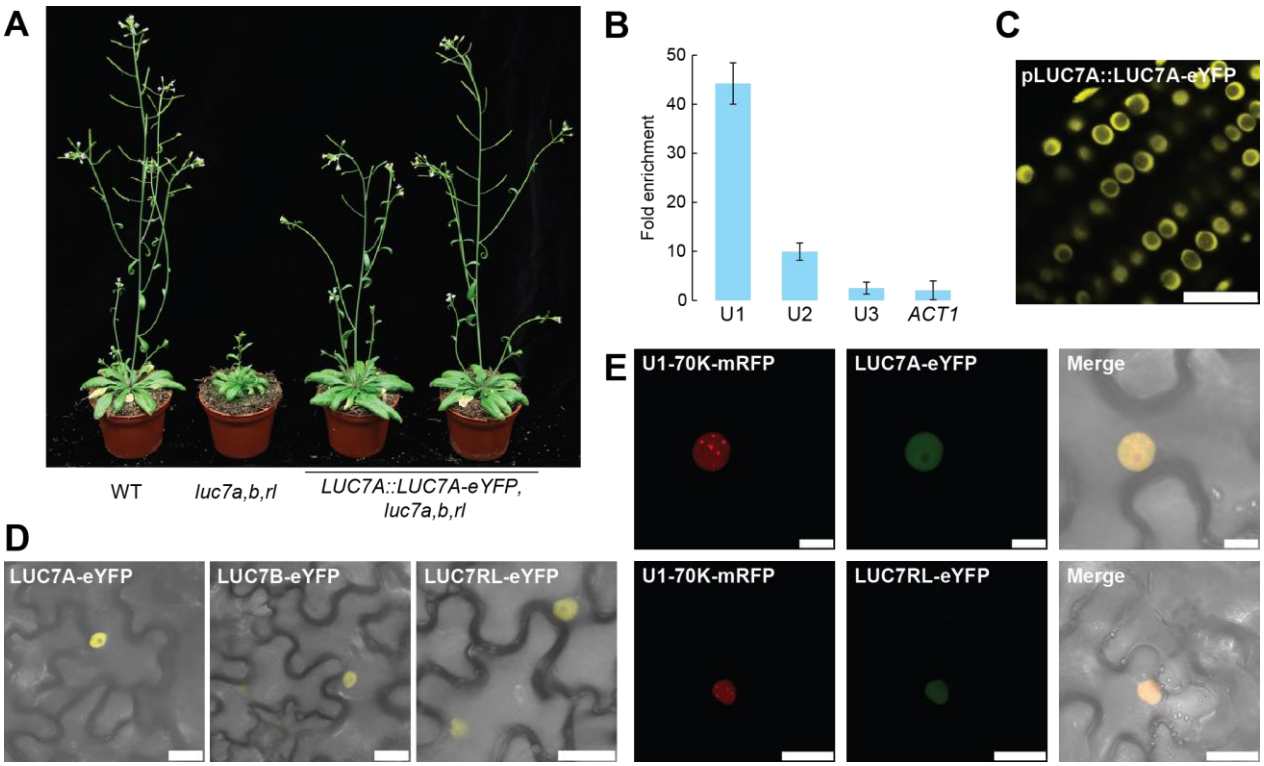


Figure 7: LUC7s are U1 snRNP components in plants. (A) WT, *luc7* triple mutant and two independent *LUC7A::LUC7A-eYFP* complemented line (T4 plants) growing in long day condition for 37 days. (B) RNA immunoprecipitation using a *LUC7A::LUC7A-eYFP, luc7a,b,rl* complemented line. Proteins were immunoprecipitated using GFP affinity matrix and RNAs were extracted from the input and the immunoprecipitated. U1, U2, U3 snRNAs and *ACTIN* RNA were quantified by qRT-PCR. Enrichment of the respective RNA for *LUC7A::LUC7A-eYGFP luc7a,b,rl* transgenic lines was calculated over WT (negative control). Error bars denote the range of two biological replicates. (C) Roots of 9 day-old Arabidopsis transgenic seedlings expressing LUC7A in *LUC7A::LUC7A-eYFP luc7a,b,rl*. (D-E) Transient expressions in leaves of *Nicotiana benthamiana*. (D) Subcellular localization of LUC7A, LUC7B and LUC7RL using the following constructions: *LUC7A::LUC7A-eYFP, LUC7B::LUC7B-eYFP* and *LUC7RL::LUC7RL-eYFP*. (E) Subcellular colocalization of U1-70K-mRFP and LUC7A-eYFP or LUC7RL-eYFP. All scale bars correspond to 25 µm, except for those from the colocalization of LUC7A with U1-70k (E-upper panel), which denote 10µm. For all localization studies confocal microscope (Leica SP2 or SP8) were used.

To further verify whether LUC7A associates *in planta* with proteins known to be U1 core components and also to shed light on putative roles of LUC7 in plants, LUC7A complexes were purified and interacting partners were detected by mass spectrometry (MS). For this, *LUC7A::LUC7A-eYFP* complemented line was used together with two

controls: WT and WT plants expressing free GFP. The analysis revealed that LUC7 is indeed found in complexes with two core U1 snRNP proteins: U1-A and U1-70K (Table 1 and Supplement 4). Additionally, proteins from the Sm heptameric ring (SmB and SmD1) that are part of the U1 snRNP, but not exclusively from this subcomplex, were also detected. These interactions reinforce the previous RIP and colocalization data establishing AthLUC7s as *bona fide* U1 proteins.

Interestingly, the MS analysis showed that LUC7A interacts with SE and ABH-1, one of the CBC components, indicating that the LUC7-SE/CBC complex most likely exists. These interactions were further validated by coimmunoprecipitations (CoIP) followed by western-blot (Figure 8A-B). CoIP to test whether LUC7s could also interact with CBP20 -the small CBC subunit- was also performed and further support the existence of a LUC7-CBC/SE complex (Figure 8C). Additional proteins known to be involved in splicing such as SR45 and the serine-arginine (SR) proteins -SR30, SCL30A, SCL33- were also detected in the immunoprecipitated (Table 1 and Supplement 4). Moreover, LUC7s may associate with the transcriptional machinery since not only a transcription factor (GT-2) was highly enriched in LUC7 immunoprecipitated, but also two proteins involved in transcription elongation: Spt6/GTB1 and ELP1. Furthermore, it was identified three kinases that might influence U1 snRNP activity. Last but not least, it was found in three independent experiments one peptide corresponding to U2AF65A, a protein that belongs to the spliceosomal complex E. Peptides from U2AF35A, another component from this initial splicing complex, were also identified in all four replicates; however a slightly contamination was present in the WT control (Table 1 and Supplement 4). The presence of these two interacting proteins suggests again that LUC7 proteins are involved in the very early steps of the splicing cycle. All in all, the data indicate that: (i) LUC7s are U1 snRNP proteins highly regulated trough phosphorylation; (ii) LUC7s are recruited to the RNA still during transcription; and (iii) LUC7-SE/CBC complex most likely exist *in planta*.

Table 1: List of LUC7 interacting partners. *LUC7A::LUC7A-eYFP* complemented line was used to immunoprecipitate the LUC7A complex, which was analyzed by mass spectrometry. WT and WT expressing free GFP were used as controls. Four replicates/experiments were performed, which are color coded (IP1 to IP4). For the first replicate the WT control was absent. The number of peptides found on each IP is displayed and grouped based on the sample type: GFP, LUC7-YFP or WT. For the sequence coverage of the peptides see Supplement S4. N° proteins: number of proteins present within the group, which corresponds to the number of entries in the column 'Protein IDs' (not displayed here). Peptides: total number of peptide sequences associated with the protein group. Mol. Weight: Molecular weight (KDa).

	Protein Names	Gene Names	N° of proteins	Peptides	N° of Peptide in GFP				N° of Peptide in LUC7-YFP				N° of peptide in WT			Mol. weight [kDa]
					IP1	IP2	IP3	IP4	IP1	IP2	IP3	IP4	IP2	IP3	IP4	
	At3g03340	**LUC7A/UNE6**	**1**	**57**	**1**	**9**	**0**	**1**	**52**	**51**	**37**	**43**	**0**	**0**	**0**	**47,4**
	At5g17440	LUC7B	1	28	0	4	0	0	26	28	21	22	0	0	0	47,2
U1 core proteins	At3g50670	U1-70k	2	9	0	0	0	0	6	9	2	3	0	0	0	50,4
	At2g47580	U1A	1	4	0	0	0	0	2	3	2	1	0	0	0	28,1
Splicing proteins	At5g44500; At4g20440	SmB-a ; SmB-b	2	4	0	0	0	0	3	4	1	2	0	1	0	27,0
(including SRs)	At3g07590; At4g02840	SmD1-a; SmD1-b	3	3	0	0	0	0	0	1	2	0	0	0	0	12,8
	At2g16940	Splicing factor CC1-like	3	5	0	0	0	0	2	3	0	2	0	0	0	63,1
	At5g28740	SYF1	1	4	0	0	0	0	1	4	0	1	0	0	0	107,1
	At1g16610	SR45	3	3	0	0	0	0	2	3	1	1	0	0	0	44,6
	At1g27650;At5g42820	U2AF35A;U2AF35B	3	5	0	0	0	0	3	5	3	3	0	2	1	34,6
	At4g36690	U2AF65A	4	2	0	0	0	0	1	1	0	1	0	0	0	60,7
SR family	At3g13570; At1g55310	SCL30A; SCL33	5	6	0	0	0	0	3	5	2	1	0	1	0	30,2
	At1g23860; At4g31580; At2g24590	RSZ21RSZ22;RSZ22A	5	4	0	0	0	0	2	4	0	1	0	0	0	22,5
	At3g55460	SCL30	1	3	0	0	0	0	1	2	0	0	0	0	0	29,6
	At1g09140	SR30	2	3	0	0	0	0	1	3	0	0	0	0	0	29,1
	At2g37340	RS2Z33	3	3	0	0	0	0	1	3	0	0	0	0	0	32,9
RBP and proteins	At2g27100	SERRATE	1	17	0	0	0	0	10	14	3	4	0	3	0	81,1
involved in others	At2g13540	ABH1	1	3	0	0	0	0	1	3	0	0	0	0	0	96,5
RNA metabolism	At4g25550	mRNA cleav. Factor-25 kDa	2	3	0	0	0	0	1	3	0	3	1	1	1	22,8
	At2g43970	LARP6B	3	5	0	0	0	0	1	3	2	3	0	1	1	60,6
	At1g24800; At1g24881; At1g25055; At1g25211; At1g25150	F-box/kelch-repeat proteins	5	3	0	0	0	0	2	2	1	1	0	0	0	49,6
	At4g39680	SAP domain	1	6	0	0	0	0	0	6	0	1	0	0	0	69,4
	At3g08780	-	2	3	0	0	0	0	0	0	1	2	0	0	0	34,3
	At1g79090	PAT1	1	3	0	0	0	0	1	0	0	2	0	0	0	88,8
	At2g45640	SAP18	2	5	0	0	0	0	2	5	1	1	0	1	0	17,2
Kinases	At3g44850	serine/threonine kinase	1	17	0	0	0	0	13	13	2	5	0	0	0	60,7
	At3g53030	SRPK4	1	16	0	0	0	0	10	9	8	7	0	0	0	59,4
	At5g22840	SRPK3	1	11	0	0	0	0	4	8	2	3	0	0	0	61,2
Transcription	At1g63210; At1g65440	GTB1/Spt6	4	15	0	0	0	0	0	0	5	14	0	0	3	166,4
	At1g76890	GT-2	1	14	0	1	0	0	9	10	3	4	0	1	1	65,8
	At5g13680	ELP1	1	3	0	0	0	0	0	0	1	2	0	0	0	146,6

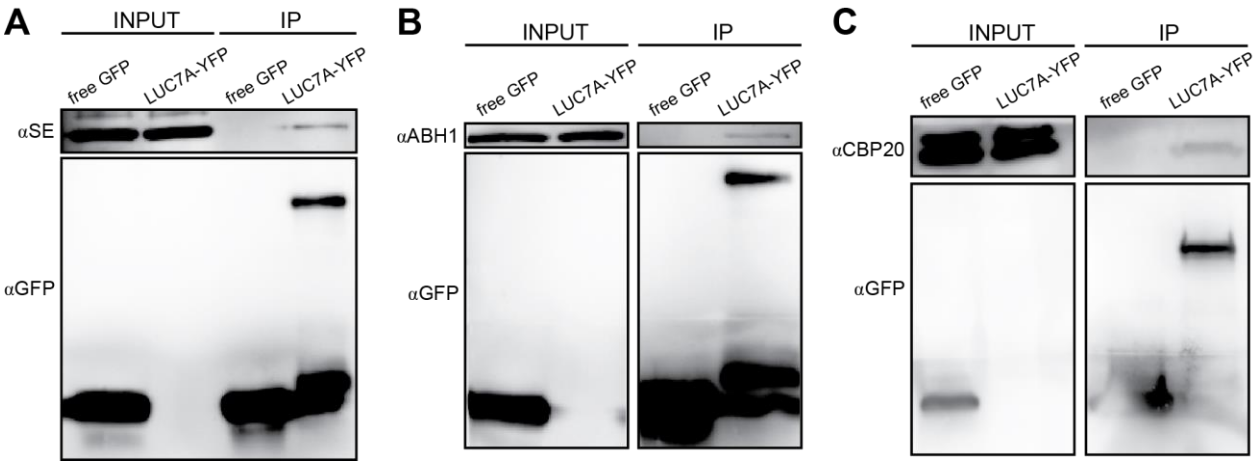


Figure 8: LUC7A interacts with SE/CBC. Coimmunoprecipitations using *LUC7A::LUC7A-eYFP, luc7a,b,rl* complemented line and WT line expressing free GFP were performed followed by western-blot analysis to detect SE (A), ABH1(B), CBP20 (C) and GFP (A-C, lower part). (A-B) 10%SDS-PAGE. (C) 4–20% Mini-PROTEAN® TGX™ precast gel (Biorad). Size of the proteins: SE and ABH ~100KDa; CBP20 ~30 KDa; LUCA-eYFP ~76KDa and GFP ~30KDa.

1.3 LUC7 effect on Arabidopsis coding and non-coding transcriptome

In plants, it has been suggested that U1 snRNP may impact miRNA biogenesis (Bielewicz et al., 2013; Knop et al., 2016; Schwab et al., 2013; Stepien et al., 2017). The majority of plants miRNAs derive from independent pol II transcription units and the pri-miRNA transcripts may contain introns. On the other hand, there are also some cases of miRNAs located inside introns of other genes. In both cases, the 5'ss of the transcript seems to control the efficiency of miRNA production and it has been suggested that the U1 snRNP might play an inhibitory or stimulatory role depending where the miRNA is located. In addition, the LUC7 interacting partners SE, ABH1 and CBP20 are well-known key players in the miRNA biogenesis (Table 1, Figure 8) (Achkar et al., 2016; Rogers & Chen, 2013). Thus, intending to assess whether LUC7 could be involved in this process, levels of 4 mature miRNAs that are affected in *se-1*, *abh1-285* and *cbp20-1* mutants were assessed in the *luc7* triple mutant by RT-qPCR. All tested miRNAs did not change in abundance in *luc7* triple mutant suggesting that LUC7 might not have a role in miRNA biogenesis (Figure 9A). In agreement with this, *luc7* triple does not display serrated leaves as expected for factors involved in miRNA biogenesis (Figure 9B).

In order to identify globally misregulated and misspliced genes in *luc7* mutants, poly(A) enriched RNA-sequencing (RNA-seq) libraries were prepared using seven day old seedlings from WT and *luc7* triple mutant. At this age, the triple mutant and WT seedlings are morphologically similar and therefore, changes in transcript levels and splicing patterns most likely reflect changes caused by LUC7s impairments and are not due to different contribution of tissues cause by, for instance, a delay in development or/and different morphology (Figure 9C). Three biological replicates for each genotype were sequenced.

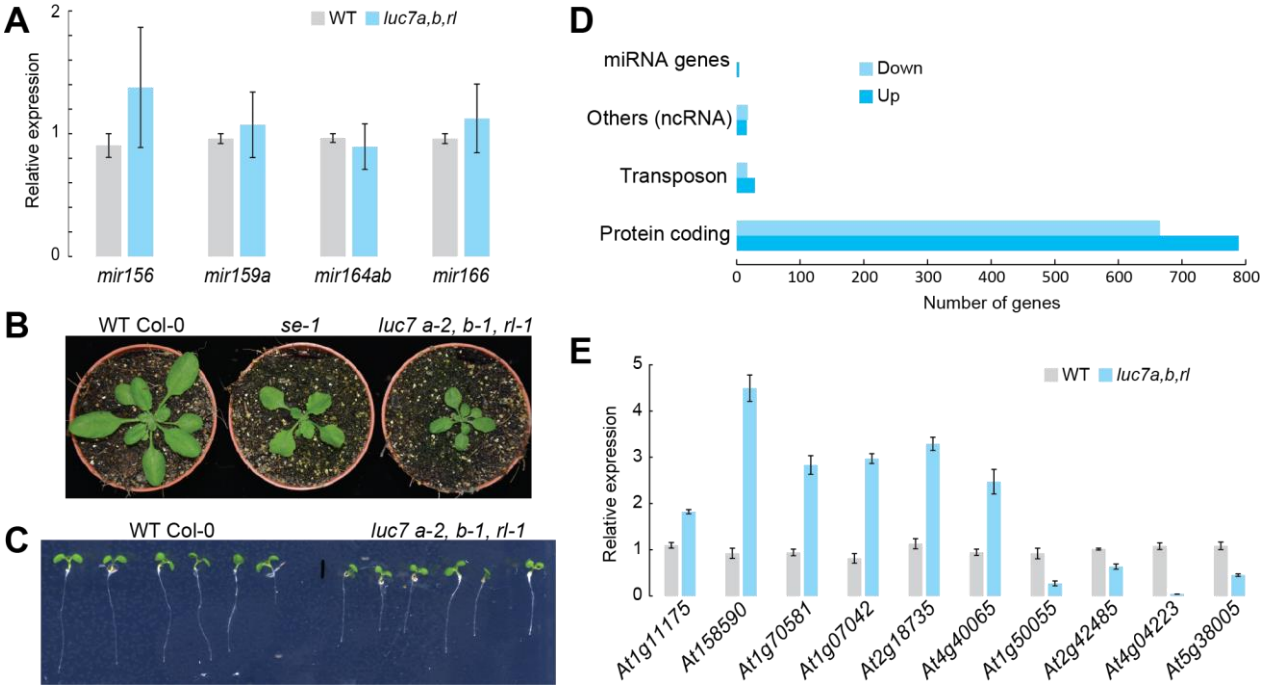


Figure 9: LUC7 effect on Arabidopsis coding and non-coding transcriptome. (A) qRT-PCR of selected mature miRNA. ACTIN was used to normalize. (B). WT, *se-1* and *luc7 a-2,b-1,rl-1* mutants growing under long day conditions for 21 days. (C) Seven days-old WT and *luc7 a-2,b-1,rl-1* seedlings growing on vertical plates. (D) Number of differentially expressed genes in *luc7 a-2,b-1,rl-1* mutant compared to WT. (E) qRT-PCR analysis of selected ncRNA. PP2A was used to normalize. Error bars denote the ±SEM (n=3).

Analysis of differentially expressed genes revealed 840 genes up- and 703 down-regulated in *luc7* mutants when compared to WT. The great majority of these genes (94-95%) encodes for proteins (Figure 9D). Nevertheless, non-coding RNAs (ncRNAs) were significantly enriched, although the overall number of affected ncRNAs is relatively small

($p < 0.05$, hypergeometric test; Figure 9D). The expressions of some ncRNAs were confirmed by RT-qPCR (Figure 9E). Additionally, this global analysis shows that the levels of *MIRNA* genes were not affected in *luc7* triple mutants (Figure 9D). These results reveal that *LUC7* affects the expression of protein-coding genes and a subset of ncRNAs, but is not involved in the miRNA pathway.

1.4 Arabidopsis LUC7 function is important for constitutive and alternative splicing

Since LUC7 is an U1 snRNP component, one would expect that misspliced transcripts accumulate in the *luc7* triple mutant. Therefore a splicing analysis was carried out with the RNA-seq libraries. In total, it was identified 645 differential splicing events in *luc7* triple mutant compared to WT (Figure 10A). A large number of intron retention events were detected and RT-PCRs for some selected intron retentions events confirmed the RNA-seq data (Figure 10A-B). These results suggest that impairments of the U1 snRNP components LUC7s disturb introns recognition. Interestingly, it was also identified a large number of exons that are included in the *luc7* triple mutant when compared to WT, as well as cases of alternative 5' and 3' splice site selection (Figure 10A-F). Some of these affected splicing events generate transcript variants that do not exist in WT (e.g. *At2g32700, At3g57410* - Figure 10C,F). On the contrary, in other cases *luc7* triple mutant were depleted in specific mRNA isoforms, which exist in WT plants (e.g. *At1g10980*, *At4g32060*), or the ratio of two different isoforms was altered in *luc7* mutants (e.g. *At3g17310*, *At5g16715*, *At5g48150*, *At2g11000*) (Figure 10D-F). These results show that the LUC7 proteins are involved in constitutive and alternative splicing in Arabidopsis.

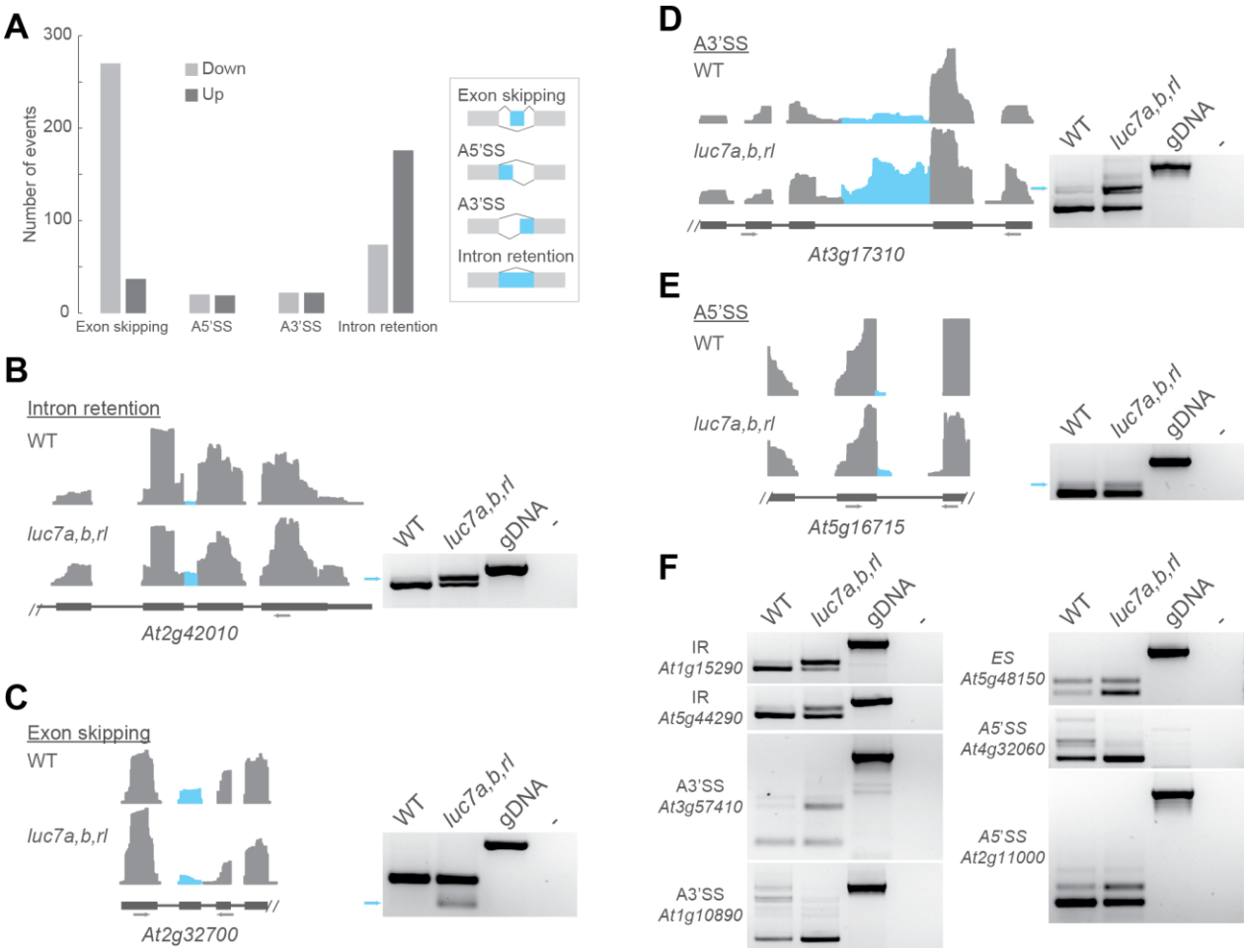


Figure 10: LUC7s impairments affect constitutive and alternative splicing. (A) Classification of differential splicing events altered in *luc7 a-2, b-1, rl-1* compared to WT. (B-F) Coverage plots and validation by RT-PCR for selected splicing events in WT and *luc7* triple mutant. Genomic DNA (gDNA) or water (-) served as a control. Primers positions are indicated with gray arrows. IR, intron retention; ES, exon skipping; A3'SS, alternative 3'splicing site; A5'SS, alternative 5'splicing site.

Next, the aim was to check if LUC7s act redundantly in the molecular level. In other words, investigate whether the splicing changes observed in *luc7* triple mutant are actually due to the loss or impairment of only a specific *LUC7* gene or whether *LUC7* genes have a complete functional overlap. To test this, splicing analysis of some transcripts were assessed in *luc7* single and double mutants. The results show that some splicing defects are detectable even in *luc7* single mutants (Figure 11), but the degree of missplicing increases in double and triple mutants suggesting that LUC7 proteins act additively on these introns (e.g. *At3g57410*). Additionally, some splicing defects are pronounced only in *luc7* triple mutants, implying that LUC7 proteins act redundantly to ensure splicing of

these introns (e.g. *At1g60995*). Interestingly, some defects might more likely be due to the lack of one of the LUC7s. For instance, intron removal of *At2g42010* more strongly relied on *LUC7RL*, while removal of an intron in *At5g41220* seems to preferentially depend on LUC7A/LUC7B (Figure 11). These findings suggest that Arabidopsis *LUC7* genes may function redundantly, additively or specifically to ensure proper splicing of pre-mRNAs.

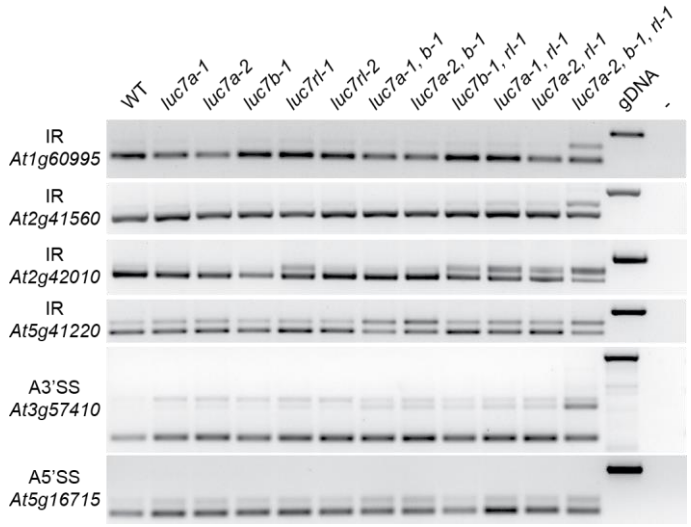


Figure 11: LUC7 proteins act redundantly, additively and specifically on splicing. Selected splicing events affected in *luc7* triple mutant were assessed in the singles and double mutants by RT-PCR. IR, intron retention; A3'SS, alternative 3'splicing site; A5'SS, alternative 5'splicing site.

1.5 LUC7 proteins are preferentially involved in the removal of terminal introns

LUC7A interacts with SE and CBC and in plants, SE/CBC complex plays mainly a role in the splicing of cap-proximal introns (Table 1, Figure 8) (Laubinger et al., 2008; Raczynska et al., 2010; Raczynska et al., 2013). These interactions raise the question whether LUC7s and SE/CBC could act together in the removal of first introns. To test this hypothesis, the splicing patterns of LUC7- and SE/CBC- dependent introns were analyzed in *luc7* triple, *se-1* and *cbc* mutants (*cbp20-1* and *abh1-285*) by RT-PCR. Introns that are retained in the *luc7* triple mutant were correctly spliced in *se-1* and *cbc* mutants (Figure 12A). Conversely, cap-proximal introns that are retained in *cbp20-1*, *abh1-285* and *se-1* mutants were removed in the *luc7* triple mutant (Figure 12B). These results showed that the functions of LUC7s and CBC/SE in splicing of the selected introns do not overlap.

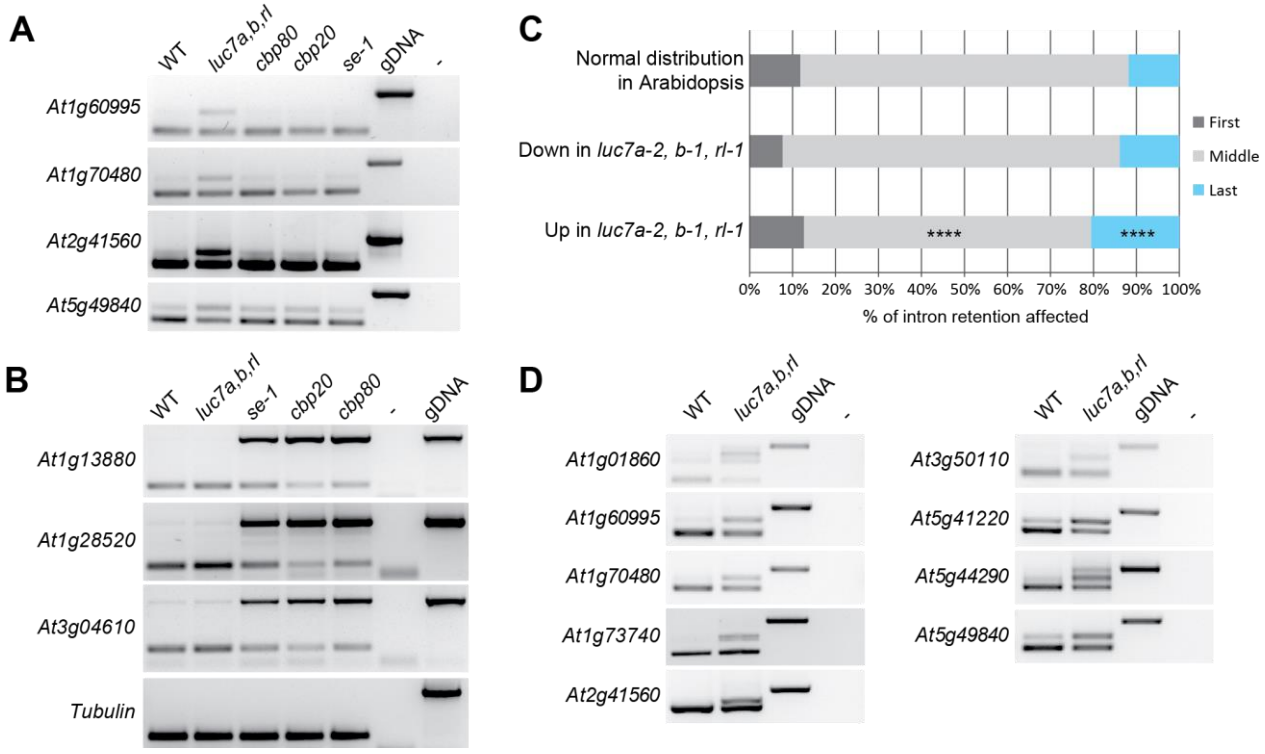


Figure 12: LUC7 proteins preferentially target terminal introns. (A) RT-PCR of LUC7-dependent introns in WT, *luc7* triple mutant, *cbp80*, *cbp20* and *se-1* mutants. (B) RT-PCR of CBC/SE-dependent introns in WT, *luc7* triple mutant, *cbp80*, *cbp20* and *se-1* mutants. (C) Intron retention events detected when luc7 is compared to WT were categorized according to the intron position (first, middle, or last). Only genes with at least 3 introns were considered for this analysis. Terminal introns are significantly enriched among introns retained in *luc7* triple mutant (Fischer test; two side p< 0.0016). (Supplement S5). (D) Validations by RT-PCR of LUC7 dependent terminal introns.

Intending to assess globally if these proteins share a function in splicing of cap-proximal introns, the next step was to check whether LUC7 has also a preference for first introns splicing. To answer this question, all retained introns in the *luc7* triple mutant were classified according to their position within the gene: first, middle or last introns. Only genes with at least 3 introns were considered for this analysis. The same was performed for introns that tend to be more removed in *luc7* triple mutant when this is compared to WT (down in *luc7* triple). As a background for comparing the frequencies, expressed genes (genes with introns>2) were used and their introns were also categorized. Surprisingly, *luc7* triple mutant has a significant increase in the frequency of retained last introns, but not for first introns (Figure 12C, Supplement S5). Some of these events were then confirmed by RT-PCR analyses (Figure 12D). Interestingly, the RT-PCRs, which

were performed with oligo dT primers, also revealed that some LUC7-dependent terminal introns (*At5g42220* and *At5g49840*) are retained in the mature RNA in WT, although in a less degree (Figure 12D). This indicates that at least for some terminal introns the splicing occurs after the 3'end formation. Taking altogether, these observations reveal that: (i) most likely CBC/SE acts on cap-proximal introns splicing independently of LUC7; (ii) LUC7 proteins exhibit a preference for the removal of terminal introns; and (iii) splicing of some LUC7-dependent terminal introns occurs after cleavage and polyadenylation.

1.6 mRNAs harboring unspliced LUC7 dependent terminal introns remain in the nucleus and escape the Nonsense-Mediated Decay (NMD)

Intron retention can generate transcripts harboring an in-frame PTC, which potentially could trigger RNA degradation via the NMD pathway. For eliciting NMD, transcripts harboring a PTC typically may have one of the following features: (i) a long 3'UTR (≥300 - 350 nt); or (ii) an exon-exon junction downstream to the PTC (>50-55 nt) (Drechsel et al., 2013; Kalyna et al., 2012; Shaul, 2015). In the case of terminal introns, only a long 3'UTR could lead to NMD. All analyzed LUC7 dependent last introns generate such NMD feature when the intron is retained, raising the question whether these transcripts are NMD substrates. To test this hypothesis, their splicing patterns were analyzed in two NMD mutants (*lba-1* and *upf3-1*), where the unspliced isoforms should accumulate if they were NMD targets. No difference between WT and NMD mutants was detected (Figure 13A). Thus, one can conclude that retained LUC7 dependent terminal introns do not trigger the RNA degradation via NMD.

Due to the fact that this decay pathway occurs in the cytoplasm, RNAs could escape NMD by not being exported from the nucleus to the cytoplasm (Gohring et al., 2014). For this reason, the subcellular localization (nuclear x cytoplasm) of spliced and unspliced mRNAs from LUC7 dependent terminal introns were assessed. To do this, total, nuclear and cytosolic fractions were isolated from WT and *luc7* triple mutant seedlings and RT-PCR analyses were performed. Spliced mRNA isoforms were mainly found in the cytosol, whereas mRNAs containing the unspliced terminal introns were found in the in nuclear fractions (Figure 13B). These results indicate that retention of these terminal introns correlates with trapping the mRNAs in the nucleus, suggesting that splicing of LUC7-dependent terminal introns is essential for the mRNA transport to the cytosol.

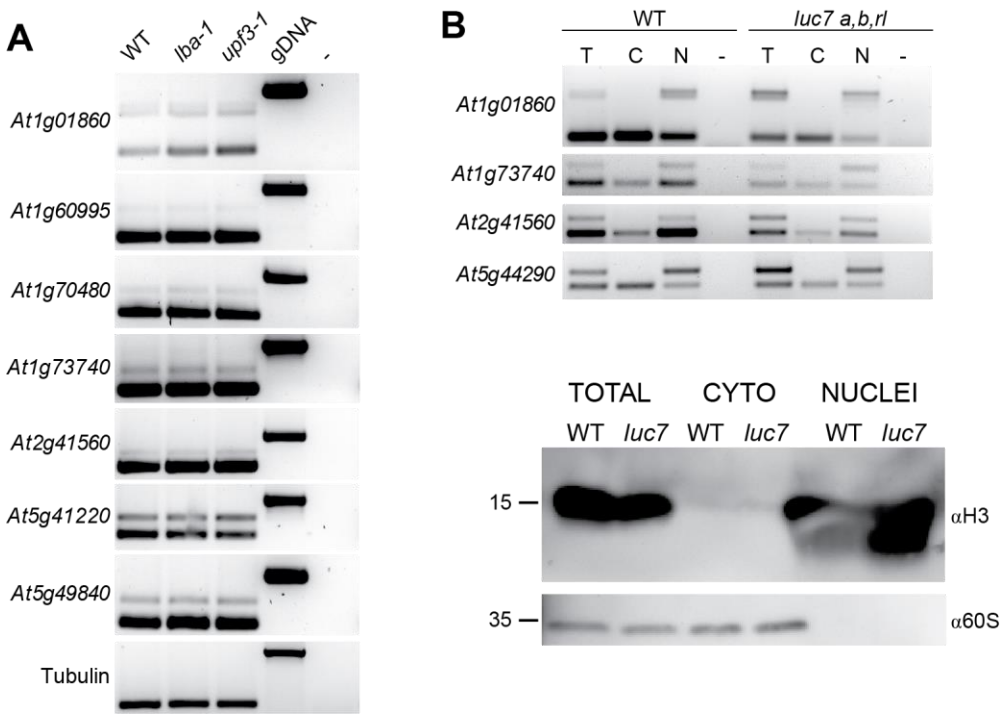


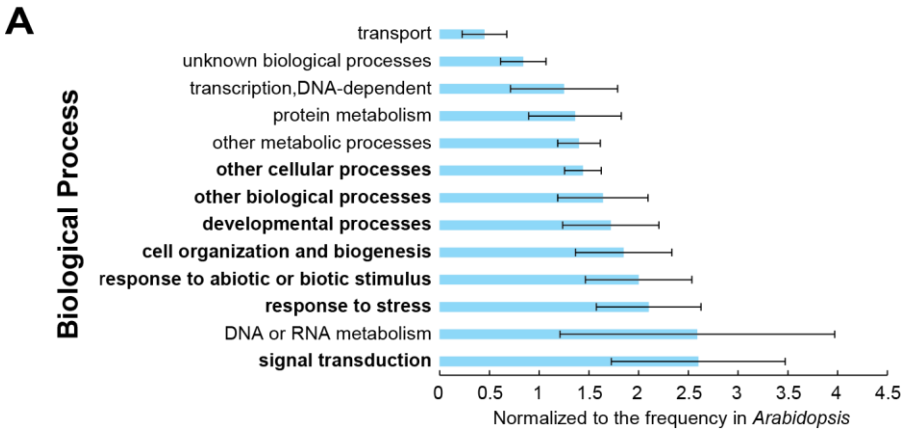
Figure 13: mRNAs carrying retained LUC7-dependent terminal introns are NMD-insensitive and remain nuclear. (A) RT-PCR of LUC7-dependent terminal introns in WT and NMD-related mutants (*lba1* and *upf3-1*). (B) Subcellular fractionation was performed in WT and *luc7* triple mutant, mRNAs were isolated from total [TOTAL], cytosolic [CYTO] and nuclear [NUCLEI] fractions and the splicing pattern of LUC7-dependent terminal introns were analyzed by RT-PCR with primers flanking only the affected intron (upper panel). Proteins were also extracted and western blot was performed to check the fractionation (lower panel). Antibodies against histone 3 (H3) and the ribosomal subunit 60S were used.

1.7 Splicing of LUC7-dependent terminal introns can be modulated by cold stress

Intending to assess if genes with LUC7 dependent terminal introns share a function, GO analyses were performed. In AgriGO no enrichment for terms in the category biological process was found. However, the Classification Super Viewer from Bar Utoronto indicates enrichment for some terms including response to stress (Figure 14). This enrichment indicates a putative role for LUC7 proteins under stress conditions and raises the question whether the retention of these terminal introns could be stress regulated. To test this hypothesis, splicing of LUC7 dependent terminal introns were assessed in WT and in the *luc7* triple mutant under stress condition. One would expect that if this

hypothesis were true then LUC7 dependent terminal introns would be retained under stress conditions in WT; additionally, if this retention were dependent on LUC7, there would be no further increase in intron retention in the *luc7* triple mutant. Among possible stresses to test, cold was chosen since a recent study suggests that U1 snRNP functionality is impaired under cold (Schlaen et al., 2015). Initially, RT-PCRs were used to assess these introns retentions events. Some events of retention were observed in WT under stress (eg. *At1g73740* and *At2g41560*), but it was difficult to estimate the effect especially for the *luc7* triple mutant (Figure 15A). For this reason and also in order to check if the retention events were significant, RT-qPCRs were performed. Due to the fact that most of the tested primers located in the intron-exon boundaries were not specific to the unspliced isoforms, intronic primers were used to assess the amount of unspliced transcripts. Interestingly, 3 out of 4 tested genes display intron retention in WT under cold, while no significant increase was detected in the *luc7* triple mutant (Figure 15B). These results suggest that cold stress can modulate the splicing efficiency of LUC7 dependent terminal introns and this seems to be LUC7 dependent.

To sum up, the results so far reveal that splicing of LUC7 dependent terminal introns are regulated under stress condition and their retentions arrest the RNA in the nucleus.



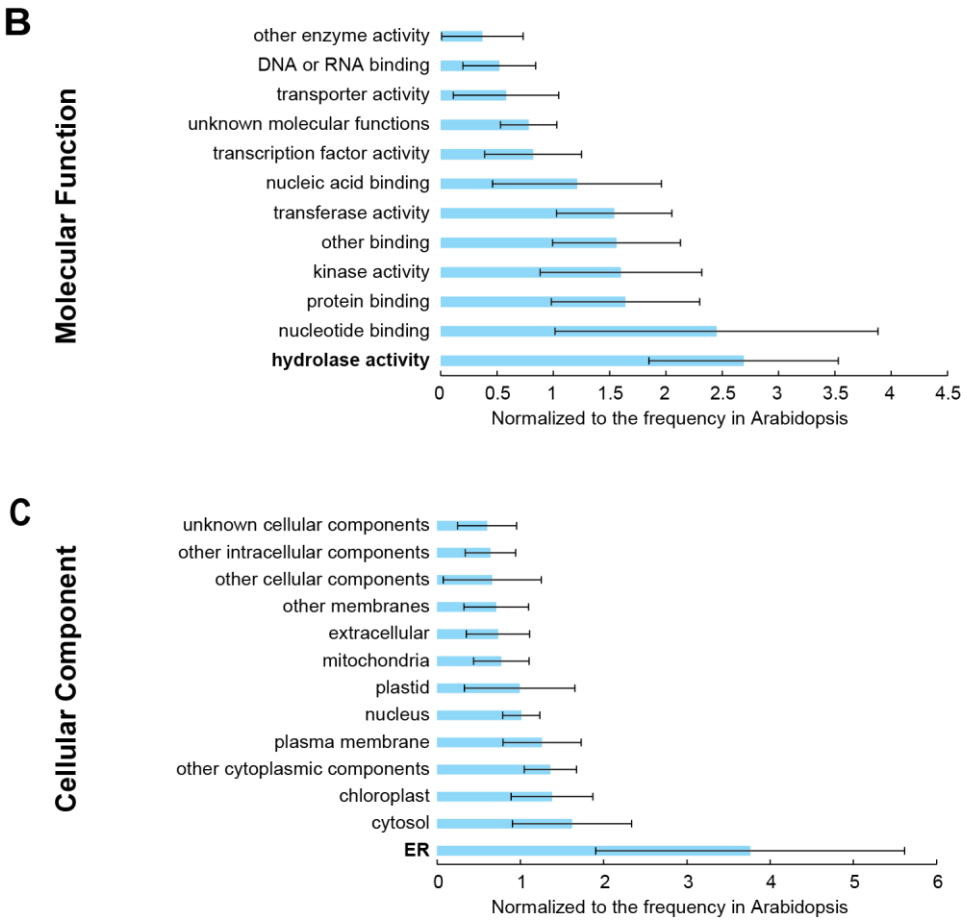


Figure 14: LUC7 dependent terminal introns are enriched for stress related genes. GO categories divided in (A) Biological Process, (B) Molecular Function and (C) Cellular Component. GO analysis was performed in the Classification Super Viewer from Bar Utoronto (http://bar.utoronto.ca/ntools/cgi-bin/ntools_classification_superviewer.cgi). Terms significantly enriched are marked in bold.

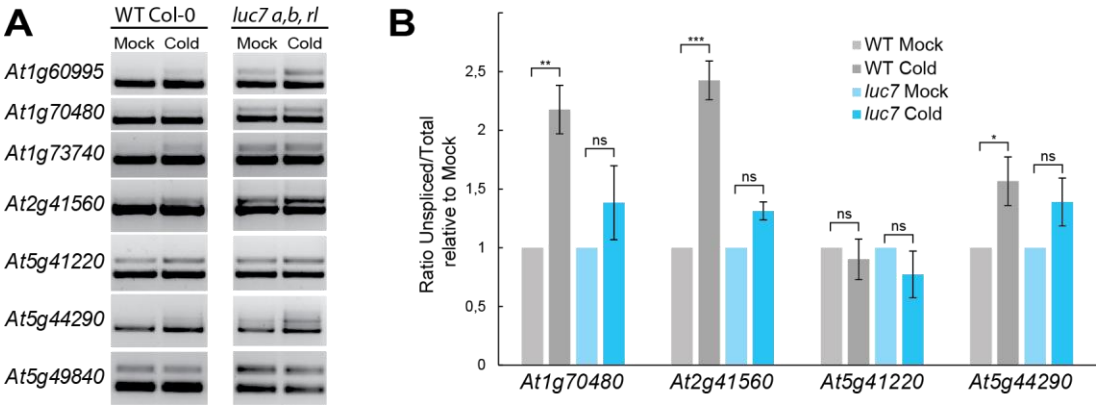


Figure 15: Splicing of LUC7 dependent terminal introns can be modulated by stress. Seven days old WT and *luc7* triple mutant seedlings were treated with cold for 60 min and the splicing pattern were analyzed by (A) RT-PCRs. The splicing ratio (unspliced/total) of four LUC7 dependent terminal introns were also analyzed by RT-qPCR

(B). PP2A was used to normalize. T-test was performed before calculating the relative to respective mock (ns: not significant and p-value: *< 0.05, ** <0.01, ***<0.001).

1.8 Cold and salt stress preferentially affects splicing of first and terminal introns in Arabidopsis

The fact that LUC7 dependent terminal introns can be modulated by stress raises the question if this is a general feature in the Arabidopsis transcriptome for last introns. In order to assess this, global analyses of intron retention under stress condition were performed using available RNA-seq data sets from Arabidopsis treated with cold and salt stress (Ding et al., 2014; Schlaen et al., 2015). For these analyses, the treated samples were compared with the respective control and the differential intron retention events were retrieved. After filtering for transcripts containing at least three introns, the introns were categorized based on their position: first, middle and last. The observed frequency of events on each category was compared to the frequency of introns from all expressed genes. If there were no preference in a particular intron category for being differentially regulated, one would expect to have the same distribution as the distribution of introns from all expressed genes (genes with introns>2). The results show that under salt and cold stresses not only last introns, but also first introns, were in general significantly enriched among the intron retention events. In addition, first and last introns were more prone to be removed in other group of genes, while the middle introns were then underrepresented in all cases (Figure 16, Supplement S6). All in all, these global analyses reveal that first and last introns tend to be more affected under cold and salt stresses. Based on the fact that this tendency was observed for two different stresses, one can expect this to be a widely spread mechanism that might contribute to plant stress responses.

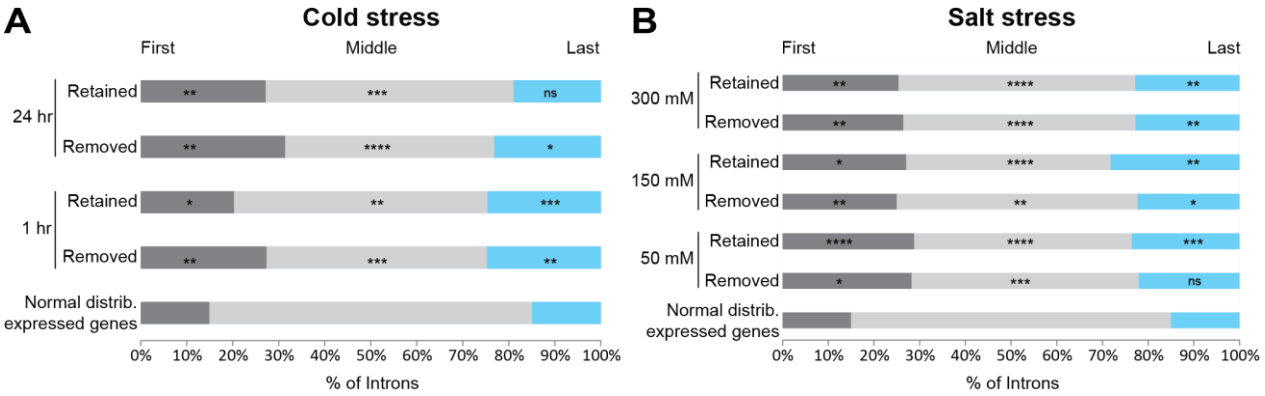


Figure 16: First and last introns are more prone to be regulated under stress. Global analyses on the effect of cold (A) and salt (B) stresses on introns removal/retention were performed. Introns were categorized according to their position in the transcript (genes with introns>2) and a distribution of the affected introns was performed based on their category for each stress. Fischer test was performed (ns: not significant and p-value: *< 0.05, ** <0.01, ***<0.001, ****<0.0001).

2. LUC7, SERRATE and the nuclear cap-binding complex (CBC)

SE/CBC role in the miRNA biogenesis is not shared with LUC7. In addition, SE/CBC and LUC7 proteins seem also to act independently in the removal of cap-proximal introns and terminal introns, respectively. In agreement with the idea that they have independent functions, *se-1*, *abh1-285* and *cbp20-1* mutants have different gross phenotype than *luc7* triple mutant (Figure 17A). However, these proteins indeed interact and supporting these interactions, they display a very similar gene expression pattern (Figure 8 and 17B). Since LUC7, SE and CBC are involved also in other alternative splicing events apart from intron retention, one can expect that they act together in these other events (Raczynska et al., 2010; Raczynska et al., 2013).

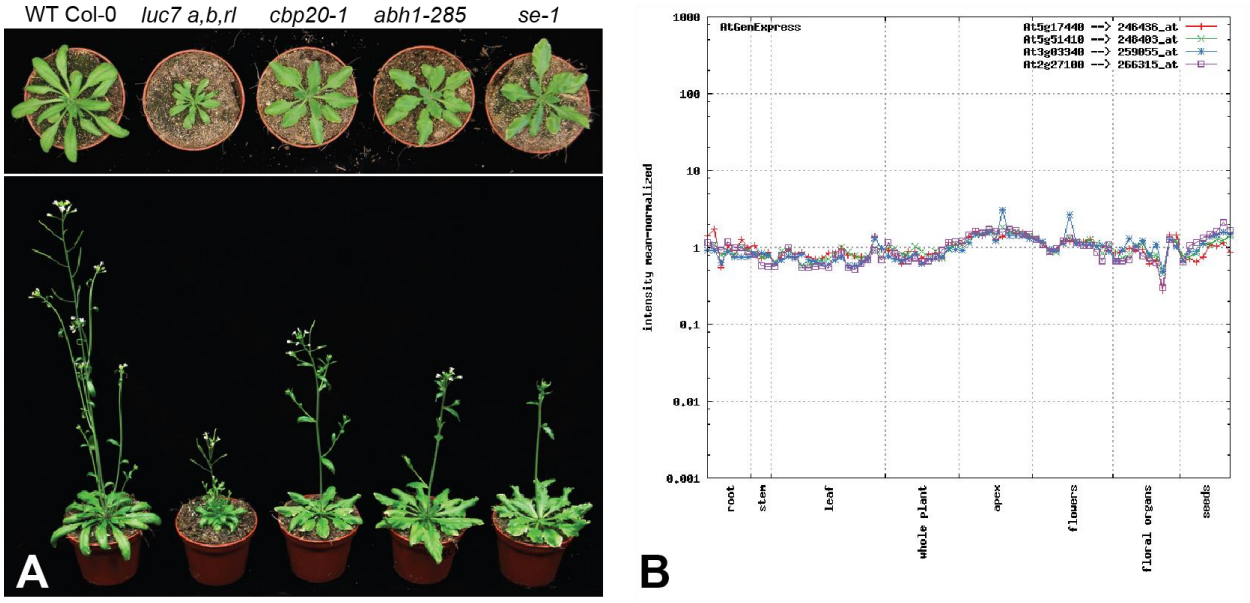


Figure 17: LUC7 and SE/CBC mutants and genes expressions. (A) Gross phenotype of *luc7* triple, *cbc* and *se-1* mutants growing in long day conditions for 37 days. (B) Gene expression pattern of *LUC7A* (*At3g03340*), *LUC7B* (*At5g17440*), *LUC7RL (At5g51410*), *SE* (*AT2G27100*) *ABH1/CBP80* (*At2g13540*) and *CBP20* (*At5g44200*). Mean normalized Affymetrix microarray data from the AtGenExpress across different tissues/organ are shown (Data source: AtGenExpress Visualization Tool - http://jsp.weigelworld.org/expviz/expviz.jsp) (Schmid et al., 2005).

2.1 LUC7 and SE affects the transcriptome in the similar way

Intending to assess globally splicing events that might require LUC7-SE and in addition, to shed light on the role of the LUC7-SE complex, poly(A) enriched RNA sequencing libraries were also prepared from *se-1* mutant. Analysis of the differential splicing events reveals 225 events affected in *se-1* when this is compared to WT (Figure 18A). Clearly the distribution of splicing events affected in *se-1* is different from the *luc7* triple distribution (Figure 10A and 18A). A global comparison checking the overlap between the *luc7* triple and *se-1* splicing events remains to be done. However, due to the fact that intron retention is indeed the most enriched event in *se-1* and that there is a deviation for cap-proximal introns among these events of intron retention, one can expect that globally the splicing overlap between *luc7* triple mutant and *se-1* might not be significant. This supports a different LUC7 and SE function in splicing.

Next, the SE impact on gene expression was evaluated and compared with LUC7 misregulated genes. A differential gene expression analyses indicate that *se-1* has 1174

up regulated and 950 down-regulated genes when compared to WT. Interestingly, there is a significant overlap in up- and downregulated genes in *se-1* and *luc7a-2,b-1,rl-1*, supporting their physical interaction data (Figure 18B-C). GO analyses on the genes up-regulated in *luc7* triple, *se-1* or in both mutants reveal a significant amount of stress related genes while down-regulated genes are enriched for general metabolic process (Supplement S7-S12). Among the up-regulated stress genes is RD29A, a well-known stress response markers involved in ABA, cold and drought (Ishitani et al., 1997; Msanne et al., 2011). This enrichment suggests that LUC7 and SE might act as negative regulators in stress response.

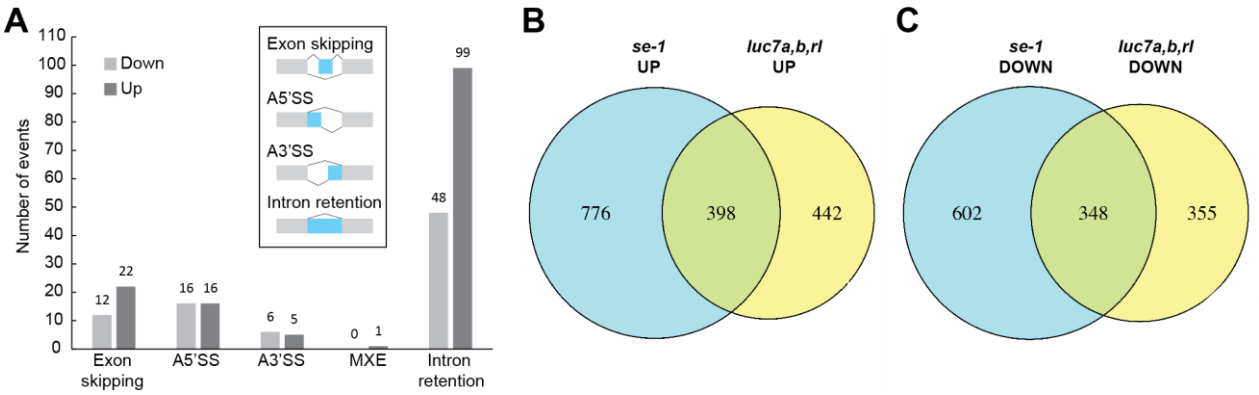


Figure 18: SE impairment effect on transcriptome. (A) Distribution of splicing events affected in *se-1*. Overlap between up- (B) and downregulated (C) genes in *se-1* and *luc7* triple.

2.2 *luc7* triple mutant display hypersensitivity to NaCl and ABA as reported for SE and CBC

If *LUC7* genes have a role in response to abiotic stress, one would expect that plants impaired in LUC7 function display an altered sensitivity to stress. Due to the fact that CBP20 and ABH1 seem to regulate salt stress response, *luc7* triple mutant growth was analyzed under salt stress conditions (Kong et al., 2014). For this, WT, *luc7* triple mutant and *LUC7A::LUC7A-eYFP* complement line seedlings growing on vertical plates for 4 days were transferred to plates containing 75 mM, 150 mM NaCl or plates without salt (Mock). Their root growths were assessed for up to 11 days after transferring. Clearly salt impact the root growth rate in the *luc7* triple mutant and this effect is rescued in the complemented line indicating that LUC7s impairments are indeed responsible for this

phenotype (Figure 19). Interestingly, this hypersensitivity to salt has also been observed for SE (Speth and Laubinger, unpublished data).

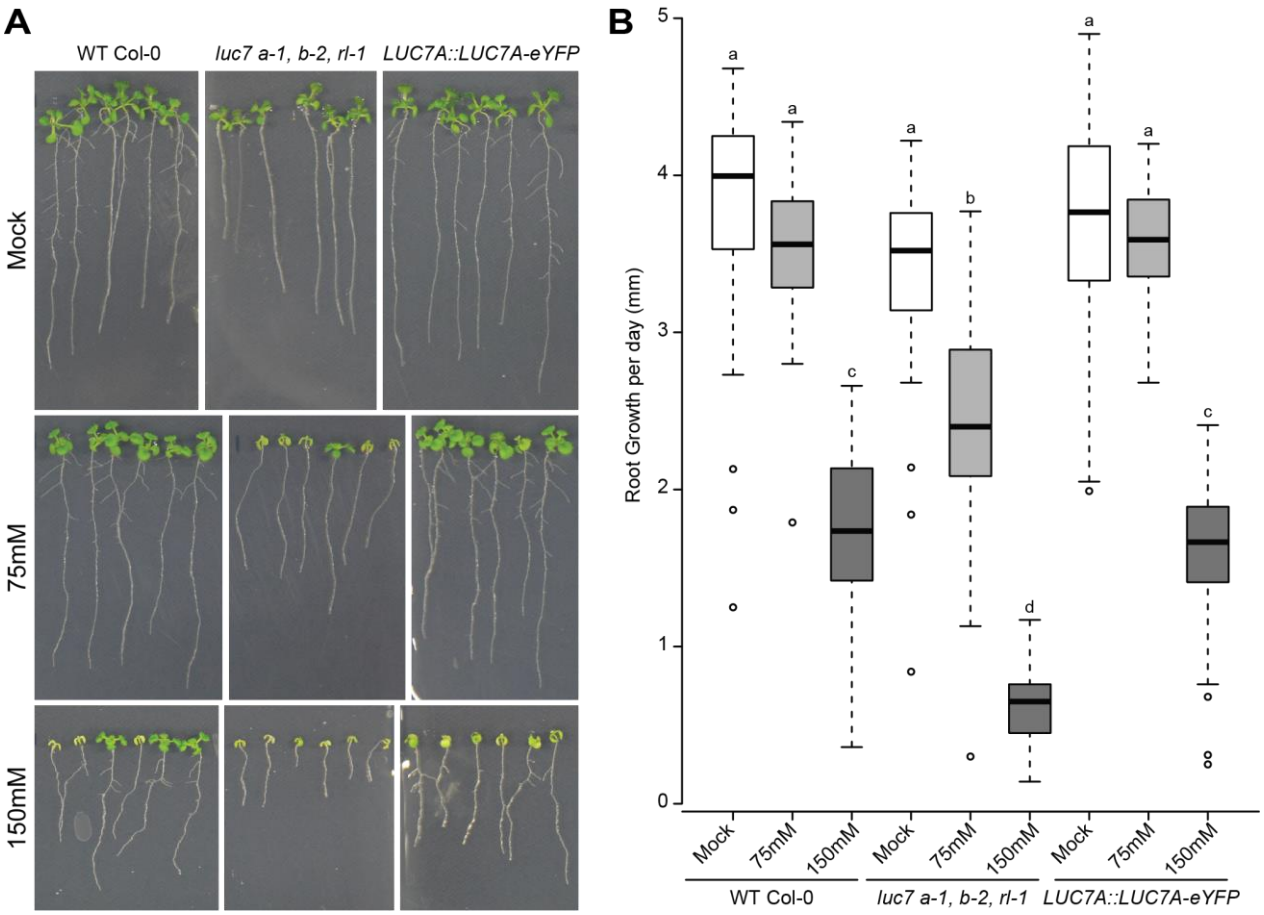


Figure 19: LUC7 impairment leads to salt hypersensitivity. (A) WT, *luc7* triple mutant and *LUC7A::LUC7A-eYFP* complemented line (#20.3.1) seedlings growing under continuous light for further 11 days after the transferring. (B) Root growth rate per day was calculated considering the root length from day 2 and 9 after transferring. ANOVA was performed followed by Tukey *post-hoc* test for multiple comparison. Letters on top of each boxplot indicate similar samples.

Since many of LUC7 interacting partners -SE, ABH1, CBP20 and SR45- display hypersensitivity to the phytohormone ABA, the impact of exogenous ABA on *luc7* triple mutant growth was analyzed sr45 (Bezerra et al., 2004; Carvalho et al., 2010; Hugouvieux et al., 2001; Papp et al., 2004) . For this, seeds from WT, *luc7* triple mutant and two independent complement lines from *LUC7A::LUC7A-eYFP* (#9.10 And #20.5) were sowed out on plates with and without ABA. Initial test with 4 different concentrations of

ABA (100 nM, 300 nM, 1 µM and 3 µM) was performed. After establishing in which of these concentration a differential effect could be easily seen, the experiment was repeated with 300 nM but now *abh1-285* and *sr45-1* were also included. The greening of the seedlings was assessed after 10 days. The *luc7* triple mutant displays a similar sensibility than *abh-1* and *sr45-1*, which was rescued in the two independent complemented lines (Figure 20). All in all, *luc7* triple mutant displays hypersensitivity to salt and ABA similar to many LUC7 interacting partners, indicating that they most likely share functions together.

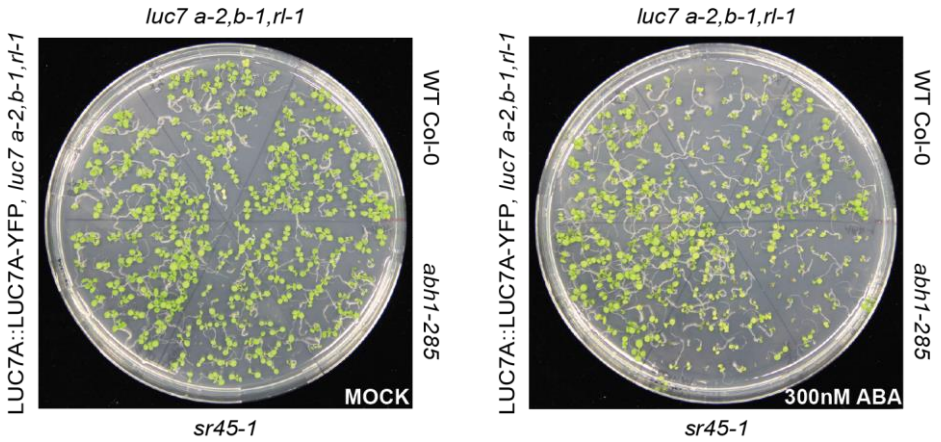


Figure 20: *luc7* triple mutant displays similar sensitivity to exogenous ABA than *abh1-285* and *sr45-1.* WT, *luc7* triple mutant, two *LUC7A::LUC7A-eYFP* complemented lines, *abh1-285* and *sr45-1* were sow out on 1/2MS plates supplemented with 1% sucrose and in the presence or absence of ABA. Seedlings were grown for 10 days under continuous light.

DISCUSSION

1. Function of Arabidopsis U1 snRNP subcomplex

In this study, the function of Arabidopsis U1 snRNP subcomplex was assessed through the characterization of Lethal Unless CBC7 (LUC7) family. For this, Arabidopsis mutants specifically impaired in LUC7 functions were generated and the genome-wide effects were analyzed. After showing that LUC7 belongs to the U1 snRNP subcomplex in plants, complementation of the *luc7* triple mutant revealed that all three Arabidopsis LUC7 act redundantly. Thus, the genome-wide analysis was carried out in the triple mutant.

The impairment of U1 snRNP function in the *luc7* triple mutant affects constitutive splicing since a large number of introns are retained in this mutant (Figure 10A). This data reveals introns that have a strong requirement for LUC7 and suggests that their splicing depends on a proper recognition of the 5'ss, which is probably assisted by LUC7. This difference in LUC7 requirement could be due to, for instance, different strength in the 5'ss. Yeast strains defective in LUC7 are more affected in splicing of introns with non-consensus 5'ss or branch point sequence (Fortes et al., 1999a). It is plausible to envision that the same happens in the *luc7* mutants in plants; however, further analysis to assess the 5'ss strength of the retained introns is required to confirm this hypothesis. Interestingly, events of alternative 5'ss and 3'ss were also detected in *luc7* triple mutant, further suggesting that LUC7 proteins are important for 5'ss selection and in addition affect 3'ss selection, respectively (Figure 10). The effect on the 3'ss could be mediated indirectly through LUC7 interactions with the U2 auxiliary proteins U2AF35 and U2AF65, both detected in LUC7A MS analysis (Table 1). In the very early steps of splicing, when the spliceosomal complex E is formed, the interactions between U1 snRNP and other proteins including U2AF35 and U2AF65 have a fundamental role in the 5'ss and 3'ss recognition (Hoffman & Grabowski, 1992; Wahl et al., 2009). This indicates that an impairment of U1 snRNP proteins could indirectly impact 3'ss usage due to impaired interactions with U2AF proteins. Moreover, it was observed that some of these events are also present in WT, but have a different ratio in *luc7* triple mutant (Figure 10F, e.g. *At1g10890*). This indicates that LUC7 proteins play a role also in alternative splicing. Furthermore, *luc7* triple mutant displays many changes in exon-skipping events, showing that the LUC7s are involved in regulating several aspects of alternative splicing. Strikingly, among these events, there are a great number of exons (270) that are not anymore skipped in the triple mutant (see

exon-skipping events down in *luc7* - Figure 10A). Hence, LUC7 proteins seem to promote many exon-skipping events. Interestingly, the general splicing factor U2AF65, which as mention above interact with LUC7, can also promote alternative exon exclusion in humans (Cho et al., 2015). One hypothesis is that LUC7s could act in concert with U2AF65 in the selection of alternative exons. Moreover, since the MS data revealed several SR proteins as LUC7 interacting partners, one could also expect LUC7 proteins to exert their role in alternative splicing through these interactions. SR proteins belong to a huge family of regulatory proteins and are often associated with alternative splicing events. However, since SR proteins tend to be seen as positive regulators promoting exon inclusion via its binding to exons and U1 snRNP recruitment, their interaction with LUC7 would better explain the *luc7* triple mutant events where exons are excluded (see exon-skipping events up in *luc7*- Figure 10A) (Fu & Ares, 2014). In this case, the impairment of LUC7 proteins would abolish the recruitment of SR proteins, which would then lead to exon skipping. Although SR proteins are often associated with alternative splicing events, they play also a role in constitutive splicing (Howard & Sanford, 2015; Reddy & Shad Ali, 2011). In fact, the separation between splicing factors involved only in constitutive versus alternative splicing is challenging, since proteins considered core components seems also to be involved in alternative splicing (Saltzman et al., 2011). For instance, U1-C, a core U1 snRNP component, also affect alternative splicing (Rosel et al., 2011). Thus, LUC7s roles in alternative splicing are reasonable and reflect an intricate network among many factors.

In the future, one important step to gain insight into LUC7 mode of action is the identification of LUC7s direct targets. At the moment, it is not known which of the differential splicing events in *luc7* triple are due to LUC7 impairments and, which are secondary effects caused by, for instance, missplicing of splicing factors. The identification of LUC7 direct targets could be done by using RIP-seq, where all RNA bound by LUC7 would be detected. In addition, analyses of differential splicing events in an inducible artificial microRNA (amiRNA) lines against LUC7A/B in *rl1* background would also show events that are directly dependent on the LUC7 downregulation. Knowing the direct targets of LUC7s would help in a search for a consensus motif that LUC7 proteins might bind. Based on what is known from yeast, it is possible to expect that LUC7s are binding directly to the RNA in an exonic region upstream to the 5’ss (Puig et al., 2007).

Furthermore, the identification of LUC7 direct targets followed by a GO analysis may help in assessing whether LUC7 has a direct role in a specific biological process via splicing.

Analyses of some splicing events in single and double *luc7* mutants confirm that molecularly LUC7 proteins act mainly redundantly (Figure 11). However, in some introns they act additively or in other cases, introns seem to display a preference for LUC7A/B or LUC7RL. What could make some introns more dependent on a specific LUC7 remains unknown. In this regard, it is important to note that LUC7A and LUC7B differ from LUC7RL in the presence of an additional stretch of amino acids separating the two zinc-finger motifs. Different lengths separating two RNA binding domains may affect substrate specificity and thus, it might explain the difference in substrate binding among the LUC7 proteins (Chen & Varani, 2013).

Interestingly, it has been suggested that U1 snRNP subcomplex may exist in different compositions (Guiro & O'Reilly, 2015; Hernandez et al., 2009). Duplications among genes encoding for U1 proteins, such as the LUC7 family, may contribute for the diversity of complexes. In addition, in humans many U1 snRNA variants exist and they seem to be packaged into different ribonucleoprotein complexes (Guiro & O'Reilly, 2015). Arabidopsis genome has not only duplicated U1 proteins, but also 14 potential U1 snRNAs, which slightly differ in sequence (Koncz et al., 2012; Reddy et al., 2013; Wang & Brendel, 2004). Specific combinations of protein family members and U1 snRNAs could potentially generate many distinct sub-complexes also in plants. Moreover, in metazoan, it was shown that U1-70K has two splicing isoforms that can be found in association with U1 snRNA and one of this isoform can be specifically phosphorylated (Serine 226). This difference in phosphorylation interferes with the U1-70k binding strength to U1-C, which could impact the 5’ss selection (Guiro & O'Reilly, 2015). Thus, splicing of splicing factors and changes in the strength of interaction between U1 components via phosphorylation could potentially regulate the complex function also in plants. In Arabidopsis at least LUC7A and LUC7RL are known to be phosphorylated (Durek et al., 2010; Heazlewood et al., 2008; Roitinger et al., 2015). Therefore one could speculate that changes in LUC7A and LUC7RL phosphorylation status could change their contribution to the U1 snRNP. On one hand, phosphorylation can interfere with the strength of interactions; on the other hand, it can promote different interaction network, modifying thereby complex function/activity (Guiro & O'Reilly, 2015; Reddy et al., 2013).

The effect of LUC7 proteins on Arabidopsis transcriptome is underestimated. The *luc7* triple mutant is a knockout for LUC7RL and produces most likely a nonfunctional LUC7B protein since in the *luc7 b-1* mutant the second zinc finger is absent, which has been shown to be essential in yeast (Supplement S13) (Agarwal et al., 2016). However, in the *luc7* triple mutant, LUC7A is almost complete with both zinc-fingers intact that might be still able to bind RNAs. Considering the fact that *LUC7* deletion is lethal in yeast and that Arabidopsis *LUC7A* was found in a screening for female gametophyte defects displaying problems in fertilization, LUC7A in *luc7a-2* most likely can partially accomplish its function (Fortes et al., 1999a; Pagnussat et al., 2005). Thus, *luc7* triple is a hypomorphic mutant and the detected effects on the whole transcriptome represent the most sensitive LUC7 events. Nevertheless, one would not expect to have all introns misregulated in an U1 mutant since U1 snRNP independent splicing has been reported in animals (Fukumura et al., 2009). Whether U1-independent splicing exists in plants and its impact remains to be assessed.

Splicing and 3’ end formation can be tightly associated and influence each other (Bentley, 2014). Recognition of the terminal intron can enhance cleavage and polyadenylation; conversely, poly(A) signal recognition can stimulate terminal intron splicing (Bentley, 2014; Niwa & Berget, 1991; Niwa et al., 1990; Proudfoot, 2011). This stimulatory effect of terminal intron splicing on the cleavage/polyadenylation process seems not to be mediated by U1 snRNP binding to the 5’ss of the terminal intron, but by splicing factors binding to the 3’ss and BP: the U2AF proteins and the U2 snRNP, respectively. This is because *in vitro* studies revealed that mutations in the 3’ss or BP from the terminal intron impair 3’end formation, while mutation in its 5’ss has no effect on the efficiency on cleavage/polyadenylation (Cooke et al., 1999). Nevertheless, U1 snRNP seems still to modulate the coupling of the last intron splicing and the 3’end formation (Bentley, 2014). Remarkably, *luc7* triple mutant has a significant higher retention rate for terminal introns comparing with first or middle introns, indicating the existence of a special mechanism for splicing of last introns by LUC7 (Figure 12C,D). This mechanism could involve the 3’ end formation machinery. Supporting this idea is the fact that the pre-mRNA cleavage factor 25KDa (AT4G25550) was detected as interacting partner of LUC7A (Table 1). Therefore, the interaction between LUC7 and the 3’ end formation complexes may contribute to a LUC7 specific function in the splicing of some terminal introns. In

humans, the same factor was detected as U1-70K interacting partner, but how they act together remains to be elucidated (Awasthi & Alwine, 2003). The current data reveal that removal of LUC7-dependent terminal introns may occur in some cases after their 3'end formation since some of these introns are still retained in transcripts containing poly(A) tail in WT (Figure 12D, e.g. *At5g41220*). This suggests that splicing of these introns are independent of cleavage/polyadenylation. Although this might be true for a set of genes, one cannot rule out that LUC7/U1snRNP has an effect in the 3'end formation for those removed co-transcriptionally.

In mammals, splicing-independent functions have been described for the U1 snRNP, explaining its higher abundance in a cell, when compared to others spliceosome subcomplexes (Guiro & O'Reilly, 2015; Kaida et al., 2010). Interestingly, U1 snRNP has a role in protecting the nascent RNA from a premature cleavage and polyadenylation and, in this way, determines also the RNA length (Berg et al., 2012; Kaida et al., 2010). To carry out this function U1 snRNP seems to bind throughout the whole nascent transcript inhibiting its 3'end formation. Moreover, due to the same U1 snRNP inhibition function, this complex is also responsible for establishing the promoter directionally in animals (Almada et al., 2013). In plants, nothing is known about this U1 splicing-independent function. The current data generated for the Arabidopsis LUC7 family characterization do not allow any inference about this putative U1 function. However, further analysis of the *luc7* triple mutant to detect where the 3' end formation occurs at genome-wide level, using direct RNA sequencing, could reveal whether these U1 components impact the position of the 3´end formation (Sherstnev et al., 2012). It is possible that LUC7 proteins are not involved in this process at all, since, as mention before, many different U1 snRNP complexes might also exist in plants. Therefore, the most straightforward strategy to assess this putative U1 snRNP function would be the downregulation of the U1 snRNA using an inducible system. It is important to keep in mind that recently a genome-wide study revealed that divergent transcripts are lacking in plants (Hetzel et al., 2016). In animals, the promoters are often bidirectional and the restriction of transcription to the sense direction is accomplished by the high amount of U1 snRNP binding sites in the transcripts that protect them from a premature 3' end formation. On the other hand, the upstream antisense transcripts encounter early termination due to depletion of U1 binding sites (Almada et al., 2013). Considering the lack of divergent transcription in plants, it is

possible that this U1 splicing independent function in protecting RNA from early termination is a novelty in some metazoan and might not even exist in plants. Future works to assess whether plant U1 snRNP has also this U1 putative splicing-independent function would definitely provide more insights in the function of U1 snRNP in plants.

2. Regulation of gene expression through intron retention: LUC7, terminal introns and stress in plants

In plants, intron retention is the most abundant alternative splicing event and transcripts with retained introns can be found in association with polyribosomes, indicating that they may potentially contribute to proteome diversity (Ner-Gaon et al., 2004; Yu et al., 2016). Apart from increasing the proteome, intron retention can be exploited as a mechanism to control gene expression. Even in animals, where it is the less represented event of alternative splicing, intron retention emerged as a way of regulating gene expression, for instance, during development (Boutz et al., 2015; Braunschweig et al., 2014; Naro et al., 2017; Pimentel et al., 2016). There are at least two ways how intron retention could modulate gene expression. In the first way, transcripts harboring a retained intron may lead to an in-frame PTC, which might trigger their degradation via NMD in the cytoplasm (Wong et al., 2013). Alternatively, intron retention can inhibit the nuclear export of unspliced transcripts, which can be stored and spliced later or, be directly degraded (Boothby et al., 2013; Naro et al., 2017; Yap et al., 2012). In both cases, the consequence is a reduction of the translatable RNA in the cytoplasm. Any retained intron can potentially control gene expression through one of this mechanism.

Interestingly, this study revealed that LUC7 regulates in part the removal of some terminal introns (Figure 12C,D), which is special in what concerns degradation via NMD (Ner-Gaon et al., 2004). To target a transcript to NMD, a retained intron that causes an in-frame PTC requires typically the deposition of an exon-junction-complex (EJC) downstream of the retained intron or the generation of a long 3'UTR (≥300 - 350 nt) (Lykke-Andersen & Jensen, 2015; Shaul, 2015). Since EJC is deposited upstream to an exon-exon junction, there will be no EJC downstream to terminal introns and thus, their retention would lead to degradation via NMD only if they have a long 3'UTR. This suggests that removal of terminal intron should be strictly controlled and in case retention occurs, another mechanism might exist to avoid the production of defective proteins. In this

regard, nuclear retention could serve as quality control mechanism to avoid the unspliced transcript to be translated. The current data suggest that LUC7 dependent terminal introns are controlled in this way since their unspliced isoforms are trapped in the nucleus (Figure 13B). One can hypothesize that these LUC7 dependent terminal introns might contain binding sites for *trans*-acting factors that inhibit the RNA export and therefore the intron splicing is required prior to the RNA export. In this regard, in animals the polypyrimidine tract-binding protein 1 (PTB1), a regulatory hnRNP protein, was found to control gene expression of nonneuronal cells by binding and inhibiting the 3' terminal intron splicing of some transcripts, which then leads to their nuclear retention and further degradation in the nucleus (Yap et al., 2012). A similar mechanism can be envisioned for LUC7 dependent terminal introns in case the splicing is impaired.

Environmental signals, including those generated by stresses, affect alternative splicing of many transcripts and this seems to be an independent layer of gene expression regulation apart from transcriptional regulation (Ding et al., 2014; Feng et al., 2015; Verhage et al., 2017). Remarkably, cold stress can promote retention of LUC7 dependent terminal introns in wild-type plants and this response is lost in *luc7* triple mutant, indicating that the observed effect requires functional LUC7 proteins (Figure 15). Due to the fact that retention of these LUC7 dependent introns causes nuclear trapping, it is possible to hypothesize that this is a broad mechanism to adjust gene expression under stress conditions. Interestingly, in a recent study, nuclear retention of unspliced RNAs was suggested as a survival strategy for plants under hypoxia (Niedojadlo et al., 2016). The retention of LUC7 dependent terminal introns under stress raised the question whether stress preferentially promotes retention of terminal introns. Remarkably, global analyses of intron retention under salt and cold stress conditions revealed that not only last, but also first introns are more prone to be regulated (Figure 16). A reasonable explanation for this preference is the fact that first and last introns are in close proximity to the 5' cap and the poly(A) tail, respectively, and these positions offer more possibilities for splicing regulation via crosstalk between the spliceosome and the machineries involved in capping and 3'end formation. Whether the same tendency is found in other species under stress conditions remains to be assessed, but one can expect that this is a conserved effect of stress.

In addition to this stress regulation of LUC7 dependent terminal introns, two other facts further indicate that LUC7 proteins have a role in stress responses: (i) the genes up-regulated in *luc7* triple mutant are enriched for functions related to stress (Supplement S7); and (ii) *luc7* triple mutant is hypersensitive to salt and ABA (Figure 19 and 20). Further experiments are required in order to assess in details LUC7 function in stress responses. For instance, a genome-wide analysis of WT and *luc7* triple mutant under stress conditions would enable the identification of splicing events that change upon exposure to stress in a LUC7 dependent manner. Another question that emerge is: how LUC7 proteins are regulated in response to adverse environmental conditions. Since LUC7 proteins are phosphorylated and in this study, three kinases were identified as LUC7A interacting partners (Table 1), it is tempting to propose that stress might regulate LUC7/U1 snRNP functions via phosphorylation (Durek et al., 2010; Heazlewood et al., 2008; Roitinger et al., 2015). This idea is supported by the observation that stress signaling pathways triggered by the phytohormone ABA results in phosphorylation of several splicing factors (Umezawa et al., 2013; Wang et al., 2013). It would be interesting to check whether LUC7 phosphorylation status changes under stress conditions and whether this causes a change in their subcellular localization.

Recently, it was shown in plants that inhibition of the whole splicing machinery via chemicals leads to activation of ABA signaling (AlShareef et al., 2017; Ling et al., 2017). This indicates a role for splicing in stress conditions, at least in abiotic stresses that trigger ABA signaling. This idea is supported by the fact that many splicing factors seem to be involved in controlling abiotic stress response and some are even directly involved via splicing of stress related genes (Feng et al., 2015; Staiger & Brown, 2013; Zhan et al., 2015). Moreover, genome-wide analyses of alternative splicing events revealed stress response genes overrepresented among those undergoing alternative splicing (Eichner et al., 2011; Ner-Gaon et al., 2004). Taking into account the results presented here, the novel role of LUC7 in splicing under stress conditions, reflects and reinforces a broad function of the splicing machinery in stress responses in plants.

3. LUC7, SE and CBC, still a common role in splicing?

The interactions between LUC7 and SE/CBC suggest a shared function, but the data indicate that it is not in the miRNA biogenesis pathways (Table 1, Figure 8 and 9A,D). To test whether LUC7 proteins act together with SE/CBC in splicing of cap-proximal introns, a global analysis of splicing events in *luc7* triple mutant was performed. Remarkably, *luc7* triple mutant displays a preference for retention of last introns, which is the opposite found in *se-1* and *cbc* mutants. However, LUC7, SE and CBC are also involved in other alternative splicing events (exon skipping, A3'SS and A5'SS), raising then the question whether LUC7 and SE/CBC functions would overlap in these other alternative splicing events (Raczynska et al., 2010; Raczynska et al., 2013). Thus, RNA-seq data from *se-1* were generated and used for a whole transcriptome analysis of SE dependent splicing events. This analysis reveals a low number (225) of alternative splicing events especially if intron retention events are not considered (78) (Figure 18A). A possible reason for finding a small number of splicing events affected in *se-1* is that this is a hypomorphic allele, where a 7 bp deletion is found in SE gene that causes a frameshift affecting the last 27 aa (Prigge & Wagner, 2001). In addition, very recently it was shown that *se-1* can still interact with U1 snRNP and therefore part of SE function in splicing is not affected. Moreover, this further suggests that the splicing events in *se-1* might not significantly overlap with the ones in *luc7* triple mutant since the SE-U1 snRNP common functions seem not to be impaired in *se-1*. Another SE allele, *se-2*, was shown to have the U1 snRNP recruitment impaired (Grigg et al., 2005; Knop et al., 2016). Further analysis with this stronger SE allele would help in answering whether LUC7 and SE have an overlap in splicing function.

Interestingly, evidences suggest that SE role in removal of first introns is not related to U1 snRNP components since *se-1* mutant display cap-proximal defect although it still can interact with U1 components (Knop et al., 2016; Laubinger et al., 2008; Raczynska et al., 2013). One interesting point is that the *cbp20-1* and *abh1-285* are both knock out mutants and have the same preference for cap-proximal introns like *se-1* (Laubinger et al., 2008; Raczynska et al., 2010). This is in agreement with the hypothesis that SE/CBC act in cap-proximal introns independently of U1 proteins. In this scenario, U1 snRNP recruitment might be done through SE, which is binding directly to the RNA through its zinc finger and the C-terminal tail, and even the absence of CBC would not impair SE-

U1sNRNP recruitment (Machida et al., 2011). What could then be involved in cap-proximal splicing if not U1 snRNP splicing factors? One reasonable hypothesis is that the removal of first introns involving SE/CBC are pretty much dependent on a factor in the chromatin. In this regard, CBC associates with two histone methyltransferases complexes: COMPASS-like and EFS/SDG8 (EARLY FLOWING IN SHORT DAYS/SET DOMAIN 8), which have been recently shown to promote efficient splicing of cap-proximal introns. COMPASS-like promotes the deposition of H3K4me3 in the proximity of the transcription start site and in the early transcribed 5'regions, while EFS deposits H3K36m3 mainly in the gene body (Li et al., 2016). Both complexes change then the chromatin context by adding markers of active genes. The exact mechanism by which these methyltransferases complexes act in concert with CBC remains to be elucidated. However, it is possible to imagine that the changes in histone methylation promoted by these complexes may allow/facilitate the recruitment of other components necessary for cap-proximal introns splicing. One cannot rule out that others splicing factors are being recruited to assist cap-proximal intron removal.

All in all, the differential expression analyses from *se-1* and *luc7* triple mutants support the existence of a shared function between LUC7 and SE. Furthermore, although in the most likely scenario SE and LUC7 share a role in alternative splicing, one cannot rule out the existence of other functions apart from splicing. The fact that LUC7 seems to associate with transcription factors (Table 1, *e.g.* GT-2) opens the possibility for a role in transcription regulation, a function that could be shared with SE. Interestingly, in animals SE homolog (ARS2) associates with the chromatin in an RNA-independent manner and activates the transcription of the gene Sox2, acting then like a transcription factor (Andreu-Agullo et al., 2011). In plants, although SE seems to interact with many transcription factors in yeast-two-hybrid, the mechanism and the significance of this interaction remain to be uncovered (Geyer and Laubinger, unpublished data). Future experiments, such as a ChIP-seq (currently in progress in the lab by Claudia Martinho and Corinna Speth), to assess whether LUC7 and SE can associate with the chromatin and whether there is an overlap in this association may help in uncovering a putative LUC7-SE role in transcription.

CONCLUSION

This study provides insights into the function of the U1 snRNP subcomplex in plants through the characterization of the LUC7 family. Previous studies in yeast and animals revealed LUC7 proteins as zinc finger proteins that belong to the U1 snRNP subcomplex (Fortes et al., 1999a; Puig et al., 2007). It was also shown formerly that yeast LUC7 is able to bind directly the RNA in an exonic region upstream to the 5'ss and close to the 7mG cap (Puig et al., 2007). In plants, apart from the identification of LUC7A in a screening for female gametophyte defects, nothing was known about the function of the LUC7 family before this study (Pagnussat et al., 2005).

The work presented here revealed that Arabidopsis LUC7A is found in complex with core U1 proteins (U1A and U1-70K) and also with the U1 snRNA, allowing to conclude that AthLUC7 proteins are indeed U1 snRNP components. Further analyses showed that LUC7 proteins act in constitutive and alternative splicing mainly in a redundant manner. LUC7 proteins emerge as the first splicing factors to show a preference for removing introns based on their position in the transcript. A position preference has only been described for SE and CBC proteins (ABH1 and CBP20), which are required primarily for cap-proximal intron splicing (Laubinger et al., 2008; Raczynska et al., 2010; Raczynska et al., 2013). Although LUC7 proteins interact with SE/CBC, LUC7 proteins have a preference for splicing of terminal introns.

LUC7 dependent terminal intron splicing is a pre-requisite for the transcript export, explaining the NMD insensitivity of the unspliced isoform. Moreover, some of these introns are retained under cold stress in WT. Therefore, U1 snRNP/LUC7 might be involved in a broader mechanism to fine-tune gene expression under stress conditions. In agreement with LUC7s roles under stress condition, *luc7* triple mutant displays a significant amount of stress related genes that are up-regulated and in addition, this mutant is salt and ABA hypersensitive. At least three kinases interact with LUC7A and a switch in LUC7 phosphorylation status would explain a fast change in their activities under stress conditions. Global analyses using publically available data sets revealed that not only last introns, but also first are more prone to be regulated under stress uncovering an unknown preference for splicing regulation in Arabidopsis.

REFERENCES

Achkar, N. P., Cambiagno, D. A., & Manavella, P. A. (2016). miRNA Biogenesis: A Dynamic Pathway. *Trends Plant Sci, 21*(12), 1034-1044. doi: 10.1016/j.tplants.2016.09.003

Agarwal, R., Schwer, B., & Shuman, S. (2016). Structure-function analysis and genetic interactions of the Luc7 subunit of the Saccharomyces cerevisiae U1 snRNP. *RNA, 22*(9), 1302-1310. doi: 10.1261/rna.056911.116

Ali, G. S., Golovkin, M., & Reddy, A. S. (2003). Nuclear localization and in vivo dynamics of a plant-specific serine/arginine-rich protein. *Plant J, 36*(6), 883-893.

Ali, G. S., Palusa, S. G., Golovkin, M., Prasad, J., Manley, J. L., & Reddy, A. S. (2007). Regulation of plant developmental processes by a novel splicing factor. *PLoS One, 2*(5), e471. doi: 10.1371/journal.pone.0000471

Almada, A. E., Wu, X., Kriz, A. J., Burge, C. B., & Sharp, P. A. (2013). Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature, 499*(7458), 360-363. doi: 10.1038/nature12349

AlShareef, S., Ling, Y., Butt, H., Mariappan, K. G., Benhamed, M., & Mahfouz, M. M. (2017). Herboxidiene triggers splicing repression and abiotic stress responses in plants. *Bmc Genomics, 18*(1), 260. doi: 10.1186/s12864-017-3656-z

Andreu-Agullo, C., Maurin, T., Thompson, C. B., & Lai, E. C. (2011). Ars2 maintains neural stem-cell identity through direct transcriptional activation of Sox2. *Nature, 481*(7380), 195-198. doi: 10.1038/nature10712

Antosz, W., Pfab, A., Ehrnsberger, H. F., Holzinger, P., Kollen, K., Mortensen, S. A., Bruckmann, A., Schubert, T., Langst, G., Griesenbeck, J., Schubert, V., Grasser, M., & Grasser, K. D. (2017). The Composition of the Arabidopsis RNA Polymerase II Transcript Elongation Complex Reveals the Interplay between Elongation and mRNA Processing Factors. *Plant Cell, 29*(4), 854-870. doi: 10.1105/tpc.16.00735

Awasthi, S., & Alwine, J. C. (2003). Association of polyadenylation cleavage factor I with U1 snRNP. *RNA, 9*(11), 1400-1409.

Barta, A., Sommergruber, K., Thompson, D., Hartmuth, K., Matzke, M. A., & Matzke, A. J. (1986). The expression of a nopaline synthase - human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. *Plant Mol Biol, 6*(5), 347-357. doi: 10.1007/BF00034942

Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space. *Nat Rev Genet, 15*(3), 163-175. doi: 10.1038/nrg3662

Berg, M. G., Singh, L. N., Younis, I., Liu, Q., Pinto, A. M., Kaida, D., Zhang, Z., Cho, S., Sherrill-Mix, S., Wan, L., & Dreyfuss, G. (2012). U1 snRNP determines mRNA length and regulates isoform expression. *Cell, 150*(1), 53-64. doi: 10.1016/j.cell.2012.05.029

Bezerra, I. C., Michaels, S. D., Schomburg, F. M., & Amasino, R. M. (2004). Lesions in the mRNA cap-binding gene ABA HYPERSENSITIVE 1 suppress FRIGIDA-mediated delayed flowering in Arabidopsis. *Plant J, 40*(1), 112-119. doi: 10.1111/j.1365-313X.2004.02194.x

Bielewicz, D., Kalak, M., Kalyna, M., Windels, D., Barta, A., Vazquez, F., Szweykowska-Kulinska, Z., & Jarmolowski, A. (2013). Introns of plant pri-miRNAs enhance miRNA biogenesis. *EMBO Rep, 14*(7), 622-628. doi: 10.1038/embor.2013.62

Boehm, V., & Gehring, N. H. (2016). Exon Junction Complexes: Supervising the Gene Expression Assembly Line. *Trends Genet, 32*(11), 724-735. doi: 10.1016/j.tig.2016.09.003

Boothby, T. C., Zipper, R. S., van der Weele, C. M., & Wolniak, S. M. (2013). Removal of retained introns regulates translation in the rapidly developing gametophyte of Marsilea vestita. *Dev Cell, 24*(5), 517-529. doi: 10.1016/j.devcel.2013.01.015

Boutz, P. L., Bhutkar, A., & Sharp, P. A. (2015). Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes Dev, 29*(1), 63-80. doi: 10.1101/gad.247361.114

Braunschweig, U., Barbosa-Morais, N. L., Pan, Q., Nachman, E. N., Alipanahi, B., Gonatopoulos-Pournatzis, T., Frey, B., Irimia, M., & Blencowe, B. J. (2014). Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res, 24*(11), 1774-1786. doi: 10.1101/gr.177790.114

Braunschweig, U., Gueroussov, S., Plocik, A. M., Graveley, B. R., & Blencowe, B. J. (2013). Dynamic integration of splicing within gene regulatory pathways. *Cell, 152*(6), 1252-1269. doi: 10.1016/j.cell.2013.02.034

Brown, J. W., Feix, G., & Frendewey, D. (1986). Accurate in vitro splicing of two pre-mRNA plant introns in a HeLa cell nuclear extract. *EMBO J, 5*(11), 2749-2758.

Calero, G., Wilson, K. F., Ly, T., Rios-Steiner, J. L., Clardy, J. C., & Cerione, R. A. (2002). Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. *Nat Struct Biol, 9*(12), 912-917. doi: 10.1038/nsb874

Carvalho, R. F., Carvalho, S. D., & Duque, P. (2010). The plant-specific SR45 protein negatively regulates glucose and ABA signaling during early seedling development in Arabidopsis. *Plant Physiol, 154*(2), 772-783. doi: 10.1104/pp.110.155523

Chen, Y., & Varani, G. (2013). Engineering RNA-binding proteins for biology. *FEBS J, 280*(16), 3734-3754. doi: 10.1111/febs.12375

Cheng, C., Wang, Z., Yuan, B., & Li, X. (2017). RBM25 Mediates Abiotic Responses in Plants. *Front Plant Sci, 8*, 292. doi: 10.3389/fpls.2017.00292

Cho, S., Moon, H., Loh, T. J., Jang, H. N., Liu, Y., Zhou, J., Ohn, T., Zheng, X., & Shen, H. (2015). Splicing inhibition of U2AF65 leads to alternative exon skipping. *Proc Natl Acad Sci U S A, 112*(32), 9926-9931. doi: 10.1073/pnas.1500639112

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J, 16*(6), 735-743.

Cooke, C., Hans, H., & Alwine, J. C. (1999). Utilization of splicing elements and polyadenylation signal elements in the coupling of polyadenylation and last-intron removal. *Mol Cell Biol, 19*(7), 4971-4979.

Crispino, J. D., Blencowe, B. J., & Sharp, P. A. (1994). Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science, 265*(5180), 1866-1869.

Cruz, T. M., Carvalho, R. F., Richardson, D. N., & Duque, P. (2014). Abscisic acid (ABA) regulation of Arabidopsis SR protein gene expression. *Int J Mol Sci, 15*(10), 17541-17564. doi: 10.3390/ijms151017541

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., & Scheible, W. R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol, 139*(1), 5-17. doi: 10.1104/pp.105.063743

Ding, F., Cui, P., Wang, Z., Zhang, S., Ali, S., & Xiong, L. (2014). Genome-wide analysis of alternative splicing of pre-mRNA under salt stress in Arabidopsis. *Bmc Genomics, 15*, 431. doi: 10.1186/1471-2164-15-431

Drechsel, G., Kahles, A., Kesarwani, A. K., Stauffer, E., Behr, J., Drewe, P., Ratsch, G., & Wachter, A. (2013). Nonsense-mediated decay of alternative precursor mRNA splicing variants is a major determinant of the Arabidopsis steady state transcriptome. *Plant Cell, 25*(10), 3726-3742. doi: 10.1105/tpc.113.115485

Du, Z., Zhou, X., Ling, Y., Zhang, Z., & Su, Z. (2010). agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res, 38*(Web Server issue), W64-70. doi: 10.1093/nar/gkq310

Duque, P. (2014). A role for SR proteins in plant stress responses. *Plant Signal Behav, 6*(1), 49-54. doi: 10.4161/psb.6.1.14063

Durek, P., Schmidt, R., Heazlewood, J. L., Jones, A., MacLean, D., Nagel, A., Kersten, B., & Schulze, W. X. (2010). PhosPhAt: the Arabidopsis thaliana phosphorylation site database. An update. *Nucleic Acids Res, 38*(Database issue), D828-834. doi: 10.1093/nar/gkp810

Eckmann, C. R., Rammelt, C., & Wahle, E. (2011). Control of poly(A) tail length. *Wiley Interdiscip Rev RNA, 2*(3), 348-361. doi: 10.1002/wrna.56

Eichner, J., Zeller, G., Laubinger, S., & Ratsch, G. (2011). Support vector machines-based identification of alternative splicing in Arabidopsis thaliana from whole-genome tiling arrays. *BMC Bioinformatics, 12*, 55. doi: 10.1186/1471-2105-12-55

Elkon, R., Ugalde, A. P., & Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet, 14*(7), 496-506. doi: 10.1038/nrg3482

Feng, J., Li, J., Gao, Z., Lu, Y., Yu, J., Zheng, Q., Yan, S., Zhang, W., He, H., Ma, L., & Zhu, Z. (2015). SKIP Confers Osmotic Tolerance during Salt Stress by Controlling Alternative Gene Splicing in Arabidopsis. *Mol Plant*. doi: 10.1016/j.molp.2015.01.011

Fortes, P., Bilbao-Cortes, D., Fornerod, M., Rigaut, G., Raymond, W., Seraphin, B., & Mattaj, I. W. (1999a). Luc7p, a novel yeast U1 snRNP protein with a role in 5' splice site recognition. *Genes Dev, 13*(18), 2425-2438.

Fortes, P., Kufel, J., Fornerod, M., Polycarpou-Schwarz, M., Lafontaine, D., Tollervey, D., & Mattaj, I. W. (1999b). Genetic and physical interactions involving the yeast nuclear cap-binding complex. *Mol Cell Biol, 19*(10), 6543-6553.

Fu, X. D., & Ares, M. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. *Nature Reviews Genetics, 15*(10), 689-701. doi: 10.1038/nrg3778

Fukumura, K., Taniguchi, I., Sakamoto, H., Ohno, M., & Inoue, K. (2009). U1-independent pre-mRNA splicing contributes to the regulation of alternative splicing. *Nucleic Acids Res, 37*(6), 1907-1914. doi: 10.1093/nar/gkp050

Gohring, J., Jacak, J., & Barta, A. (2014). Imaging of endogenous messenger RNA splice variants in living cells reveals nuclear retention of transcripts inaccessible to nonsense-mediated decay in Arabidopsis. *Plant Cell, 26*(2), 754-764. doi: 10.1105/tpc.113.118075

Golovkin, M., & Reddy, A. S. (2003). Expression of U1 small nuclear ribonucleoprotein 70K antisense transcript using APETALA3 promoter suppresses the development of sepals and petals. *Plant Physiol, 132*(4), 1884-1891.

Gonatopoulos-Pournatzis, T., & Cowling, V. H. (2014). Cap-binding complex (CBC). *Biochem J, 457*(2), 231-242. doi: 10.1042/BJ20131214

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol, 27*(2), 221-224. doi: 10.1093/molbev/msp259

Grigg, S. P., Canales, C., Hay, A., & Tsiantis, M. (2005). SERRATE coordinates shoot meristem function and leaf axial patterning in Arabidopsis. *Nature, 437*(7061), 1022-1026. doi: 10.1038/nature04052

Guiro, J., & O'Reilly, D. (2015). Insights into the U1 small nuclear ribonucleoprotein complex superfamily. *Wiley Interdisciplinary Reviews-Rna, 6*(1), 79-92. doi: Doi 10.1002/Wrna.1257

Hallais, M., Pontvianne, F., Andersen, P. R., Clerici, M., Lener, D., Benbahouche Nel, H., Gostan, T., Vandermoere, F., Robert, M. C., Cusack, S., Verheggen, C., Jensen, T. H., & Bertrand, E. (2013). CBC-ARS2 stimulates 3'-end maturation of multiple RNA families and favors cap-proximal processing. *Nat Struct Mol Biol, 20*(12), 1358-1366. doi: 10.1038/nsmb.2720

Harlen, K. M., & Churchman, L. S. (2017). The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat Rev Mol Cell Biol, 18*(4), 263-273. doi: 10.1038/nrm.2017.10

Harlen, K. M., Trotta, K. L., Smith, E. E., Mosaheb, M. M., Fuchs, S. M., & Churchman, L. S. (2016). Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Rep, 15*(10), 2147-2158. doi: 10.1016/j.celrep.2016.05.010

Hartmuth, K., & Barta, A. (1986). In vitro processing of a plant pre-mRNA in a HeLa cell nuclear extract. *Nucleic Acids Res, 14*(19), 7513-7528.

Hayashi, R., Handler, D., Ish-Horowicz, D., & Brennecke, J. (2014). The exon junction complex is required for definition and excision of neighboring introns in Drosophila. *Genes Dev, 28*(16), 1772-1785. doi: 10.1101/gad.245738.114

Heazlewood, J. L., Durek, P., Hummel, J., Selbig, J., Weckwerth, W., Walther, D., & Schulze, W. X. (2008). PhosPhAt: a database of phosphorylation sites in Arabidopsis thaliana and a plant-specific phosphorylation site predictor. *Nucleic Acids Res, 36*(Database issue), D1015-1021. doi: 10.1093/nar/gkm812

Heim, A., Grimm, C., Muller, U., Haussler, S., Mackeen, M. M., Merl, J., Hauck, S. M., Kessler, B. M., Schofield, C. J., Wolf, A., & Bottger, A. (2014). Jumonji domain containing protein 6 (Jmjd6) modulates splicing and specifically interacts with arginine-serine-rich (RS) domains of SR- and SR-like proteins. *Nucleic Acids Res, 42*(12), 7833-7850. doi: 10.1093/nar/gku488

Hernandez, H., Makarova, O. V., Makarov, E. M., Morgner, N., Muto, Y., Krummel, D. P., & Robinson, C. V. (2009). Isoforms of U1-70k control subunit dynamics in the human spliceosomal U1 snRNP. *PLoS One, 4*(9), e7202. doi: 10.1371/journal.pone.0007202

Hetzel, J., Duttke, S. H., Benner, C., & Chory, J. (2016). Nascent RNA sequencing reveals distinct features in plant transcription. *Proc Natl Acad Sci U S A, 113*(43), 12316-12321. doi: 10.1073/pnas.1603217113

Hoffman, B. E., & Grabowski, P. J. (1992). U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes Dev, 6*(12B), 2554-2568.

Hori, K., & Watanabe, Y. (2005). UPF3 suppresses aberrant spliced mRNA in Arabidopsis. *Plant J, 43*(4), 530-540. doi: 10.1111/j.1365-313X.2005.02473.x

Howard, J. M., & Sanford, J. R. (2015). The RNAissance family: SR proteins as multifaceted regulators of gene expression. *Wiley Interdiscip Rev RNA, 6*(1), 93-110. doi: 10.1002/wrna.1260

Hugouvieux, V., Kwak, J. M., & Schroeder, J. I. (2001). An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in Arabidopsis. *Cell, 106*(4), 477-487.

Hunt, A. G., Xu, R., Addepalli, B., Rao, S., Forbes, K. P., Meeks, L. R., Xing, D., Mo, M., Zhao, H., Bandyopadhyay, A., Dampanaboina, L., Marion, A., Von Lanken, C., & Li, Q. Q. (2008). Arabidopsis mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling. *Bmc Genomics, 9*, 220. doi: 10.1186/1471-2164-9-220

Ishitani, M., Xiong, L., Stevenson, B., & Zhu, J. K. (1997). Genetic analysis of osmotic and cold stress signal transduction in Arabidopsis: interactions and convergence of abscisic acid-dependent and abscisic acid-independent pathways. *Plant Cell, 9*(11), 1935-1949. doi: 10.1105/tpc.9.11.1935

Jiao, X., Chang, J. H., Kilic, T., Tong, L., & Kiledjian, M. (2013). A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol Cell, 50*(1), 104-115. doi: 10.1016/j.molcel.2013.02.017

Jobert, L., Pinzon, N., Van Herreweghe, E., Jady, B. E., Guialis, A., Kiss, T., & Tora, L. (2009). Human U1 snRNA forms a new chromatin-associated snRNP with TAF15. *EMBO Rep, 10*(5), 494-500. doi: 10.1038/embor.2009.24

Kaida, D. (2016). The reciprocal regulation between splicing and 3'-end processing. *Wiley Interdiscip Rev RNA, 7*(4), 499-511. doi: 10.1002/wrna.1348

Kaida, D., Berg, M. G., Younis, I., Kasim, M., Singh, L. N., Wan, L., & Dreyfuss, G. (2010). U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature, 468*(7324), 664-668. doi: 10.1038/nature09479

Kalyna, M., Simpson, C. G., Syed, N. H., Lewandowska, D., Marquez, Y., Kusenda, B., Marshall, J., Fuller, J., Cardle, L., McNicol, J., Dinh, H. Q., Barta, A., & Brown, J. W. (2012). Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. *Nucleic Acids Res, 40*(6), 2454-2469. doi: 10.1093/nar/gkr932

Kersey, P. J., Allen, J. E., Armean, I., Boddu, S., Bolt, B. J., Carvalho-Silva, D., Christensen, M., Davis, P., Falin, L. J., Grabmueller, C., Humphrey, J., Kerhornou, A., Khobova, J., Aranganathan, N. K., Langridge, N., Lowy, E., McDowall, M. D., Maheswari, U., Nuhn, M., Ong, C. K., Overduin, B., Paulini, M., Pedro, H., Perry, E., Spudich, G., Tapanari, E., Walts, B., Williams, G., Tello-Ruiz, M., Stein, J., Wei, S., Ware, D., Bolser, D. M., Howe, K. L., Kulesha, E., Lawson, D., Maslen, G., & Staines, D. M. (2016). Ensembl Genomes 2016: more genomes, more complexity. *Nucleic Acids Res, 44*(D1), D574-580. doi: 10.1093/nar/gkv1209

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol, 14*(4), R36. doi: 10.1186/gb-2013-14-4-r36

Knop, K., Stepien, A., Barciszewska-Pacak, M., Taube, M., Bielewicz, D., Michalak, M., Borst, J. W., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2016). Active 5' splice sites regulate the biogenesis efficiency of Arabidopsis microRNAs derived from intron-containing genes. *Nucleic Acids Res*. doi: 10.1093/nar/gkw895

Koncz, C., Dejong, F., Villacorta, N., Szakonyi, D., & Koncz, Z. (2012). The spliceosome-activating complex: molecular mechanisms underlying the function of a pleiotropic regulator. *Front Plant Sci, 3*, 9. doi: 10.3389/fpls.2012.00009

Kong, X., Ma, L., Yang, L., Chen, Q., Xiang, N., Yang, Y., & Hu, X. (2014). Quantitative proteomics analysis reveals that the nuclear cap-binding complex proteins arabidopsis CBP20 and CBP80 modulate the salt stress response. *J Proteome Res, 13*(5), 2495-2510. doi: 10.1021/pr4012624

Kornblihtt, A. R., Schor, I. E., Allo, M., Dujardin, G., Petrillo, E., & Munoz, M. J. (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat Rev Mol Cell Biol, 14*(3), 153-165. doi: 10.1038/nrm3525

Kwek, K. Y., Murphy, S., Furger, A., Thomas, B., O'Gorman, W., Kimura, H., Proudfoot, N. J., & Akoulitchev, A. (2002). U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nat Struct Biol, 9*(11), 800-805. doi: 10.1038/nsb862

Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J. U., Ratsch, G., & Weigel, D. (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. *Proc Natl Acad Sci U S A, 105*(25), 8795-8800. doi: 10.1073/pnas.0802493105

Le Hir, H., Sauliere, J., & Wang, Z. (2016). The exon junction complex as a node of post-transcriptional networks. *Nat Rev Mol Cell Biol, 17*(1), 41-54. doi: 10.1038/nrm.2015.7

Lee, J. H., Ryu, H. S., Chung, K. S., Pose, D., Kim, S., Schmid, M., & Ahn, J. H. (2013). Regulation of temperature-responsive flowering by MADS-box transcription factor repressors. *Science, 342*(6158), 628-632. doi: 10.1126/science.1241097

Leviatan, N., Alkan, N., Leshkowitz, D., & Fluhr, R. (2013). Genome-wide survey of cold stress regulated alternative splicing in Arabidopsis thaliana with tiling microarray. *PLoS One, 8*(6), e66511. doi: 10.1371/journal.pone.0066511

Lewis, J. D., Izaurralde, E., Jarmolowski, A., McGuigan, C., & Mattaj, I. W. (1996). A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes & Development, 10*(13), 1683-1698. doi: 10.1101/gad.10.13.1683

Li, Z., Jiang, D., Fu, X., Luo, X., Liu, R., & He, Y. (2016). Coupling of histone methylation and RNA processing by the nuclear mRNA cap-binding complex. *Nat Plants, 2*, 16015. doi: 10.1038/nplants.2016.15

Liao, Y., Smyth, G. K., & Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res, 41*(10), e108. doi: 10.1093/nar/gkt214

Ling, Y., Alshareef, S., Butt, H., Lozano-Juste, J., Li, L., Galal, A. A., Moustafa, A., Momin, A. A., Tashkandi, M., Richardson, D. N., Fujii, H., Arold, S., Rodriguez, P. L., Duque, P., & Mahfouz, M. M. (2017). Pre-mRNA splicing repression triggers abiotic stress signaling in plants. *Plant J, 89*(2), 291-309. doi: 10.1111/tpj.13383

Lopato, S., Waigmann, E., & Barta, A. (1996). Characterization of a novel arginine/serine-rich splicing factor in Arabidopsis. *Plant Cell, 8*(12), 2255-2264. doi: 10.1105/tpc.8.12.2255

Luco, R. F., Allo, M., Schor, I. E., Kornblihtt, A. R., & Misteli, T. (2011). Epigenetics in alternative pre-mRNA splicing. *Cell, 144*(1), 16-26. doi: 10.1016/j.cell.2010.11.056

Lykke-Andersen, S., & Jensen, T. H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol, 16*(11), 665-677. doi: 10.1038/nrm4063

Machida, S., Chen, H. Y., & Adam Yuan, Y. (2011). Molecular insights into miRNA processing by Arabidopsis thaliana SERRATE. *Nucleic Acids Res, 39*(17), 7828-7836. doi: 10.1093/nar/gkr428

Meyer, K., Koester, T., & Staiger, D. (2015). Pre-mRNA Splicing in Plants: In Vivo Functions of RNA-Binding Proteins Implicated in the Splicing Process. *Biomolecules, 5*(3), 1717-1740. doi: 10.3390/biom5031717

Moll, C., von Lyncker, L., Zimmermann, S., Kagi, C., Baumann, N., Twell, D., Grossniklaus, U., & Gross-Hardt, R. (2008). CLO/GFA1 and ATO are novel regulators of gametic cell fate in plants. *Plant J, 56*(6), 913-921. doi: 10.1111/j.1365-313X.2008.03650.x

Morrison, M., Harris, K. S., & Roth, M. B. (1997). smg mutants affect the expression of alternatively spliced SR protein mRNAs in Caenorhabditis elegans. *Proc Natl Acad Sci U S A, 94*(18), 9782-9785.

Msanne, J., Lin, J., Stone, J. M., & Awada, T. (2011). Characterization of abiotic stress-responsive Arabidopsis thaliana RD29A and RD29B genes and evaluation of transgenes. *Planta, 234*(1), 97-107. doi: 10.1007/s00425-011-1387-y

Muller-McNicoll, M., Botti, V., de Jesus Domingues, A. M., Brandl, H., Schwich, O. D., Steiner, M. C., Curk, T., Poser, I., Zarnack, K., & Neugebauer, K. M. (2016). SR proteins are NXF1 adaptors that link alternative RNA processing to mRNA export. *Genes Dev, 30*(5), 553-566. doi: 10.1101/gad.276477.115

Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y., Watanabe, Y., Nakamura, K., Kimura, T., & Ishiguro, S. (2007). Improved Gateway binary vectors: high-performance vectors for creation of fusion constructs in transgenic analysis of plants. *Biosci Biotechnol Biochem, 71*(8), 2095-2100. doi: 10.1271/bbb.70216

Naro, C., Jolly, A., Di Persio, S., Bielli, P., Setterblad, N., Alberdi, A. J., Vicini, E., Geremia, R., De la Grange, P., & Sette, C. (2017). An Orchestrated Intron Retention Program in Meiosis Controls Timely Usage of Transcripts during Germ Cell Differentiation. *Dev Cell, 41*(1), 82-93 e84. doi: 10.1016/j.devcel.2017.03.003

Ner-Gaon, H., Halachmi, R., Savaldi-Goldstein, S., Rubin, E., Ophir, R., & Fluhr, R. (2004). Intron retention is a major phenomenon in alternative splicing in Arabidopsis. *Plant J, 39*(6), 877-885. doi: 10.1111/j.1365-313X.2004.02172.x

Niedojadlo, J., Delenko, K., & Niedojadlo, K. (2016). Regulation of poly(A) RNA retention in the nucleus as a survival strategy of plants during hypoxia. *RNA Biol, 13*(5), 531-543. doi: 10.1080/15476286.2016.1166331

Niwa, M., & Berget, S. M. (1991). Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes Dev, 5*(11), 2086-2095.

Niwa, M., Rose, S. D., & Berget, S. M. (1990). In vitro polyadenylation is stimulated by the presence of an upstream intron. *Genes Dev, 4*(9), 1552-1559.

Pabis, M., Neufeld, N., Steiner, M. C., Bojic, T., Shav-Tal, Y., & Neugebauer, K. M. (2013). The nuclear cap-binding complex interacts with the U4/U6.U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA, 19*(8), 1054-1063. doi: 10.1261/rna.037069.112

Pagnussat, G. C., Yu, H. J., Ngo, Q. A., Rajani, S., Mayalagu, S., Johnson, C. S., Capron, A., Xie, L. F., Ye, D., & Sundaresan, V. (2005). Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. *Development, 132*(3), 603-614. doi: 10.1242/dev.01595

Pajoro, A., Severing, E., Angenent, G. C., & Immink, R. G. H. (2017). Histone H3 lysine 36 methylation affects temperature-induced alternative splicing and flowering in plants. *Genome Biol, 18*(1), 102. doi: 10.1186/s13059-017-1235-x

Palusa, S. G., Ali, G. S., & Reddy, A. S. (2007). Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses. *Plant J, 49*(6), 1091-1107. doi: 10.1111/j.1365-313X.2006.03020.x

Palusa, S. G., & Reddy, A. S. (2010). Extensive coupling of alternative splicing of pre-mRNAs of serine/arginine (SR) genes with nonsense-mediated decay. *New Phytol, 185*(1), 83-89. doi: 10.1111/j.1469-8137.2009.03065.x

Papp, I., Mur, L. A., Dalmadi, A., Dulai, S., & Koncz, C. (2004). A mutation in the Cap Binding Protein 20 gene confers drought tolerance to Arabidopsis. *Plant Mol Biol, 55*(5), 679-686. doi: 10.1007/s11103-004-1680-2

Pimentel, H., Parra, M., Gee, S. L., Mohandas, N., Pachter, L., & Conboy, J. G. (2016). A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis. *Nucleic Acids Res, 44*(2), 838-851. doi: 10.1093/nar/gkv1168

Pose, D., Verhage, L., Ott, F., Yant, L., Mathieu, J., Angenent, G. C., Immink, R. G., & Schmid, M. (2013). Temperature-dependent regulation of flowering by antagonistic FLM variants. *Nature, 503*(7476), 414-417. doi: 10.1038/nature12633

Prigge, M. J., & Wagner, D. R. (2001). The arabidopsis serrate gene encodes a zinc-finger protein required for normal shoot development. *Plant Cell, 13*(6), 1263-1279.

Proudfoot, N. J. (2011). Ending the message: poly(A) signals then and now. *Genes Dev, 25*(17), 1770-1782. doi: 10.1101/gad.17268411

Puig, O., Bragado-Nilsson, E., Koski, T., & Seraphin, B. (2007). The U1 snRNP-associated factor Luc7p affects 5' splice site selection in yeast and human. *Nucleic Acids Res, 35*(17), 5874-5885. doi: 10.1093/nar/gkm505

Raczynska, K. D., Simpson, C. G., Ciesiolka, A., Szewc, L., Lewandowska, D., McNicol, J., Szweykowska-Kulinska, Z., Brown, J. W., & Jarmolowski, A. (2010). Involvement of the nuclear cap-binding protein complex in alternative splicing in Arabidopsis thaliana. *Nucleic Acids Res, 38*(1), 265-278. doi: 10.1093/nar/gkp869

Raczynska, K. D., Stepien, A., Kierzkowski, D., Kalak, M., Bajczyk, M., McNicol, J., Simpson, C. G., Szweykowska-Kulinska, Z., Brown, J. W., & Jarmolowski, A. (2013). The SERRATE protein is involved in alternative splicing in Arabidopsis thaliana. *Nucleic Acids Res*. doi: 10.1093/nar/gkt894

Reddy, A. S., Marquez, Y., Kalyna, M., & Barta, A. (2013). Complexity of the alternative splicing landscape in plants. *Plant Cell, 25*(10), 3657-3683. doi: 10.1105/tpc.113.117523

Reddy, A. S., & Shad Ali, G. (2011). Plant serine/arginine-rich proteins: roles in precursor messenger RNA splicing, plant development, and stress responses. *Wiley Interdiscip Rev RNA, 2*(6), 875-889. doi: 10.1002/wrna.98

Rogers, K., & Chen, X. (2013). Biogenesis, Turnover, and Mode of Action of Plant MicroRNAs. *Plant Cell, 25*(7), 2383-2399. doi: 10.1105/tpc.113.113159

Roitinger, E., Hofer, M., Kocher, T., Pichler, P., Novatchkova, M., Yang, J., Schlogelhofer, P., & Mechtler, K. (2015). Quantitative phosphoproteomics of the ataxia telangiectasia-mutated (ATM) and ataxia telangiectasia-mutated and rad3-related (ATR) dependent DNA damage response in Arabidopsis thaliana. *Mol Cell Proteomics, 14*(3), 556-571. doi: 10.1074/mcp.M114.040352

Rosel, T. D., Hung, L. H., Medenbach, J., Donde, K., Starke, S., Benes, V., Ratsch, G., & Bindereif, A. (2011). RNA-Seq analysis in mutant zebrafish reveals role of U1C protein in alternative splicing regulation. *EMBO J, 30*(10), 1965-1976. doi: 10.1038/emboj.2011.106

Rowley, M. J., Bohmdorfer, G., & Wierzbicki, A. T. (2013). Analysis of long non-coding RNAs produced by a specialized RNA polymerase in Arabidopsis thaliana. *Methods, 63*(2), 160-169. doi: 10.1016/j.ymeth.2013.05.006

Saltzman, A. L., Pan, Q., & Blencowe, B. J. (2011). Regulation of alternative splicing by the core spliceosomal machinery. *Genes Dev, 25*(4), 373-384. doi: 10.1101/gad.2004811

Schlaen, R. G., Mancini, E., Sanchez, S. E., Perez-Santangelo, S., Rugnone, M. L., Simpson, C. G., Brown, J. W., Zhang, X., Chernomoretz, A., & Yanovsky, M. J. (2015). The spliceosome assembly factor GEMIN2 attenuates the effects of temperature on alternative splicing and circadian rhythms. *Proc Natl Acad Sci U S A, 112*(30), 9382-9387. doi: 10.1073/pnas.1504541112

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., & Lohmann, J. U. (2005). A gene expression map of Arabidopsis thaliana development. *Nat Genet, 37*(5), 501-506. doi: 10.1038/ng1543

Schwab, R., Speth, C., Laubinger, S., & Voinnet, O. (2013). Enhanced microRNA accumulation through stemloop-adjacent introns. *EMBO Rep, 14*(7), 615-621. doi: 10.1038/embor.2013.58

Shaul, O. (2015). Unique Aspects of Plant Nonsense-Mediated mRNA Decay. *Trends Plant Sci, 20*(11), 767-779. doi: 10.1016/j.tplants.2015.08.011

Shen, S., Park, J. W., Lu, Z. X., Lin, L., Henry, M. D., Wu, Y. N., Zhou, Q., & Xing, Y. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A, 111*(51), E5593-5601. doi: 10.1073/pnas.1419161111

Sherstnev, A., Duc, C., Cole, C., Zacharaki, V., Hornyik, C., Ozsolak, F., Milos, P. M., Barton, G. J., & Simpson, G. G. (2012). Direct sequencing of Arabidopsis thaliana RNA reveals patterns of cleavage and polyadenylation. *Nat Struct Mol Biol, 19*(8), 845-852. doi: 10.1038/nsmb.2345

Song, H. R., Song, J. D., Cho, J. N., Amasino, R. M., Noh, B., & Noh, Y. S. (2009). The RNA binding protein ELF9 directly reduces SUPPRESSOR OF OVEREXPRESSION OF CO1 transcript levels in arabidopsis, possibly via nonsense-mediated mRNA decay. *Plant Cell, 21*(4), 1195-1211. doi: 10.1105/tpc.108.064774

Staiger, D., & Brown, J. W. (2013). Alternative splicing at the intersection of biological timing, development, and stress responses. *Plant Cell, 25*(10), 3640-3656. doi: 10.1105/tpc.113.113803

Stauffer, E., Westermann, A., Wagner, G., & Wachter, A. (2010). Polypyrimidine tract-binding protein homologues from Arabidopsis underlie regulatory circuits based on alternative splicing and downstream control. *Plant J, 64*(2), 243-255. doi: 10.1111/j.1365-313X.2010.04321.x

Stepien, A., Knop, K., Dolata, J., Taube, M., Bajczyk, M., Barciszewska-Pacak, M., Pacak, A., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2017). Posttranscriptional coordination of splicing and miRNA biogenesis in plants. *Wiley Interdiscip Rev RNA, 8*(3). doi: 10.1002/wrna.1403

Subtelny, A. O., Eichhorn, S. W., Chen, G. R., Sive, H., & Bartel, D. P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. *Nature, 508*(7494), 66-71. doi: 10.1038/nature13007

Sureshkumar, S., Dent, C., Seleznev, A., Tasset, C., & Balasubramanian, S. (2016). Nonsense-mediated mRNA decay modulates FLM-dependent thermosensory flowering response in Arabidopsis. *Nat Plants, 2*(5), 16055. doi: 10.1038/nplants.2016.55

Tarn, W. Y., & Steitz, J. A. (1994). Sr Proteins Can Compensate for the Loss of U1 Snrnp Functions in-Vitro. *Genes & Development, 8*(22), 2704-2717. doi: DOI 10.1101/gad.8.22.2704

Topisirovic, I., Svitkin, Y. V., Sonenberg, N., & Shatkin, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA, 2*(2), 277-298. doi: 10.1002/wrna.52

Turunen, J. J., Niemela, E. H., Verma, B., & Frilander, M. J. (2013). The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA, 4*(1), 61-76. doi: 10.1002/wrna.1141

Umezawa, T., Sugiyama, N., Takahashi, F., Anderson, J. C., Ishihama, Y., Peck, S. C., & Shinozaki, K. (2013). Genetics and phosphoproteomics reveal a protein phosphorylation network in the abscisic acid signaling pathway in Arabidopsis thaliana. *Sci Signal, 6*(270), rs8. doi: 10.1126/scisignal.2003509

Verhage, L., Severing, E. I., Bucher, J., Lammers, M., Busscher-Lange, J., Bonnema, G., Rodenburg, N., Proveniers, M. C., Angenent, G. C., & Immink, R. G. (2017). Splicing-related genes are alternatively spliced upon changes in ambient temperatures in plants. *PLoS One, 12*(3), e0172950. doi: 10.1371/journal.pone.0172950

Voinnet, O., Rivas, S., Mestre, P., & Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J, 33*(5), 949-956.

Volz, R., von Lyncker, L., Baumann, N., Dresselhaus, T., Sprunck, S., & Gross-Hardt, R. (2012). LACHESIS-dependent egg-cell signaling regulates the development of female gametophytic cells. *Development, 139*(3), 498-502. doi: 10.1242/dev.075234

Wachter, A., Ruhl, C., & Stauffer, E. (2012). The Role of Polypyrimidine Tract-Binding Proteins and Other hnRNP Proteins in Plant Splicing Regulation. *Front Plant Sci, 3*, 81. doi: 10.3389/fpls.2012.00081

Wachter, A., Tunc-Ozdemir, M., Grove, B. C., Green, P. J., Shintani, D. K., & Breaker, R. R. (2007). Riboswitch control of gene expression in plants by splicing and alternative 3' end processing of mRNAs. *Plant Cell, 19*(11), 3437-3450. doi: 10.1105/tpc.107.053645

Wahl, M. C., Will, C. L., & Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell, 136*(4), 701-718. doi: 10.1016/j.cell.2009.02.009

Wang, B. B., & Brendel, V. (2004). The ASRG database: identification and survey of Arabidopsis thaliana genes involved in pre-mRNA splicing. *Genome Biol, 5*(12), R102. doi: 10.1186/gb-2004-5-12-r102

Wang, C., Tian, Q., Hou, Z., Mucha, M., Aukerman, M., & Olsen, O. A. (2007). The Arabidopsis thaliana AT PRP39-1 gene, encoding a tetratricopeptide repeat protein with similarity to the yeast pre-mRNA processing protein PRP39, affects flowering time. *Plant Cell Rep, 26*(8), 1357-1366. doi: 10.1007/s00299-007-0336-5

Wang, P., Xue, L., Batelli, G., Lee, S., Hou, Y. J., Van Oosten, M. J., Zhang, H., Tao, W. A., & Zhu, J. K. (2013). Quantitative phosphoproteomics identifies SnRK2 protein kinase substrates and reveals the effectors of abscisic acid action. *Proc Natl Acad Sci U S A, 110*(27), 11205-11210. doi: 10.1073/pnas.1308974110

Webby, C. J., Wolf, A., Gromak, N., Dreger, M., Kramer, H., Kessler, B., Nielsen, M. L., Schmitz, C., Butler, D. S., Yates, J. R., 3rd, Delahunty, C. M., Hahn, P., Lengeling, A., Mann, M., Proudfoot, N. J., Schofield, C. J., & Bottger, A. (2009). Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science, 325*(5936), 90-93. doi: 10.1126/science.1175865

Will, C. L., & Luhrmann, R. (2011). Spliceosome structure and function. *Cold Spring Harb Perspect Biol, 3*(7). doi: 10.1101/cshperspect.a003707

Wong, J. J., Ritchie, W., Ebner, O. A., Selbach, M., Wong, J. W., Huang, Y., Gao, D., Pinello, N., Gonzalez, M., Baidya, K., Thoeng, A., Khoo, T. L., Bailey, C. G., Holst, J., & Rasko, J. E. (2013). Orchestrated intron retention regulates normal granulocyte differentiation. *Cell, 154*(3), 583-595. doi: 10.1016/j.cell.2013.06.052

Xing, D., Wang, Y., Hamilton, M., Ben-Hur, A., & Reddy, A. S. (2015). Transcriptome-Wide Identification of RNA Targets of Arabidopsis SERINE/ARGININE-RICH45 Uncovers the Unexpected Roles of This RNA Binding Protein in RNA Processing. *Plant Cell*. doi: 10.1105/tpc.15.00641

Yap, K., Lim, Z. Q., Khandelia, P., Friedman, B., & Makeyev, E. V. (2012). Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. *Genes Dev, 26*(11), 1209-1223. doi: 10.1101/gad.188037.112

Yoine, M., Ohto, M. A., Onai, K., Mita, S., & Nakamura, K. (2006). The lba1 mutation of UPF1 RNA helicase involved in nonsense-mediated mRNA decay causes pleiotropic phenotypic changes and altered sugar signalling in Arabidopsis. *Plant J, 47*(1), 49-62. doi: 10.1111/j.1365-313X.2006.02771.x

Yu, H., Tian, C., Yu, Y., & Jiao, Y. (2016). Transcriptome Survey of the Contribution of Alternative Splicing to Proteome Diversity in Arabidopsis thaliana. *Mol Plant, 9*(5), 749-752. doi: 10.1016/j.molp.2015.12.018

Zhan, X., Qian, B., Cao, F., Wu, W., Yang, L., Guan, Q., Gu, X., Wang, P., Okusolubo, T. A., Dunn, S. L., Zhu, J. K., & Zhu, J. (2015). An Arabidopsis PWI and RRM motif-containing protein is critical for pre-mRNA splicing and ABA responses. *Nat Commun, 6*, 8139. doi: 10.1038/ncomms9139

SUPPLEMENTAL MATERIAL

S1: Primer list

S2: Protein sequence of yeast LUC7 was used for a blastp in NCBI.

S3: Alignment of yeast LUC7 with all Arabidopsis LUC7.

S4: Sequence coverage of the peptides found in the MS analysis.

S5: Fischer test for intron retention events in WT and *luc7* triple mutant.

S6: Fisher test for cold and salt samples.

S7: GO analysis performed in AgriGO for genes upregulated in *luc7a-2 b-1 rl-1*.

S8: GO analysis performed in AgriGO for genes downregulated in *luc7a-2 b-1 rl-1*.

S9: GO analysis performed in AgriGO for genes upregulated in *se-1*.

S10: GO analysis performed in AgriGO for genes downregulated in *se-1*.

S11: GO analysis performed in AgriGO for genes upregulated in *luc7 a-2, b-1, rl-1* and *se-1*.

S12: GO analysis performed in AgriGO for genes downregulated in *luc7 a-2, b-1, rl-1* and *se-1*.

S13: RNA sequencing coverage plot for *AthLUC7A*, *AthLUC7B* and *AthLUC7RL* and relative expression of *AthLUC7A* and *AthLUC7B* in *luc7* triple.

S1. Primer list.

	Sequence	**Info/ Reference**
Genotyping		
SALK-LB	GTTCACGTAGTGGGCCATCG	SALK-lines
SAIL-LB	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC	SAIL-lines
LUC7a-1F	GAGAAGGCCAATATACGCAAG	N825430, N834734
LUC7a-1R	GAGTAACATACATGTAGACAG	N825430, N834735
LUC7b-F1	GCTAGACAAGAACCTGTAGTG	N802553, N644681
LUC7b-R1	CTCCGCATATGTCACATAGACG	N802553, N644681
LUC7rl-F1	GAGAGAAGAGGGTTCAAGGAG	N577718
LUC7rl-R1	CGCAAAACTGTGCAAGCTCAG	N577718
LUC7rl-F2	TCGTTATGGCCATTAATGGTG	N664006
LUC7rl-R2	GTCCAACAGAGCTCGCTGAG	N664006
Cloning		
Luc7a-PromFw	GAAACCATGCACAGATGATTG	genomic
Luc7a-nostopRv	GTAGCGGTTGTGACGCCTGCA	genomic
Luc7b_seq1F	TCTTCCGAGGCGACATATTATGC	genomic
Luc7b-nostopRv	GTAGCGGTCATGGCGTCTGCAAG	genomic
Luc7rl-PromFw	CGAATCCTTGTTCTTCATGCG	genomic
Luc7rl-nostopRv	GTATTGGCGAATGGGGCTTC	genomic
Sequencing		
Luc7a_seq1F	CATCGAATGTAGTCTGCACAC	
Luc7a_seq2F	AACCAGCGAGAGACCAAGTC	
Luc7a_seq3F	TGCAGCTTCGGAAAGAGTATC	
Luc7a_seq4F	TCATGTAGGCGGTGTTGCTG	
Luc7a_seq5F	AGTACACCGCTGTTGATGTG	
Luc7b_seq2F	TGTCTTAGTGGCTACAGTTC	
Luc7b_seq3F	AATACGAAAGAACGGCTGAG	
Luc7b_seq4F	GTGGAAATGATGAGGGTCTG	
Luc7b_seq5F	TTGTTCAGGATTCCATTCTG	
Luc7b_seq6F	CACTGGCAAATTGGTTGAGTG	
Luc7b_seq7F	GGTGTTGACAACTATGATAGG	
Luc7rl_seq1F	TCAAGGTTCTTGCTGCTACTG	
Luc7rl_seq2R	TGCTCTGTTCCAATGACTTAG	
GW1	GTTGCAACAAATTGATGAGCAATGC	TOPO vector
GW2	GTTGCAACAAATTGATGAGCAATTA	TOPO vector

RIP		
U1 snRNA_Fw	TACCTGGACGGGGTCAAC	
U1 snRNA_Rv	CCCTCTGCCACAAATAATGAC	
U2 snRNA_Fw	TCGGCCCACACGATATTAAC	
U2 snRNA_Rv	GCAGTAGTGCAACGCATGG	
U3 snRNA_Fw	GGCTCGTACCTCTGTTTCC	
U3 snRNA_Rv	GCCGTCAATCACGCTCTA	
Actin2 -AT3G18780	CTTGCACCAAGCAGCATGAA	(Czechowski et al., 2005)
Actin2 -AT3G18780	CCGATCCAGACACTGTACTTCCTT	(Czechowski et al., 2005)
miRNA -qPCR/ cDNA synthesis		
mir156-RTprimer	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGA TACGACGTGCTC	stem loop oligo for cDNA
mir166-RTprimer	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGA TACGACGGGGAA	stem loop oligo for cDNA
miRNA159a-RTprimer	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGA TACGACTAGAGC	stem loop oligo for cDNA
miRNA164ab-RTprimer	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGA TACGACTGCACG	stem loop oligo for cDNA
RT-universal-Reverse-primer	GTGCAGGGTCCGAGGT	
mir156- Fwprimer	GCGGCGGTGACAGAAGAGAGT	
mir166-Fwprimer	TCGCTTCGGACCAGGCTTCA	
mir159- Fwprimer	GCGGCGTTTGGATTGAAGGGA	
mir164- Fwprimer	AGGACATGGAGAAGCAGGGCA	
Actin2 -AT3G18780	CTTGCACCAAGCAGCATGAA	(Czechowski et al., 2005)
Actin2 -AT3G18780	CCGATCCAGACACTGTACTTCCTT	(Czechowski et al., 2005)
LUC7A and LUC7B - qPCR		
LUC7a-Fw-qPCR	AGGAACGGAACAGTAAGGAG	
LUC7a-Rv-qPCR	TCTACTCCTGCGATCATAATC	
LUC7b-Fw-qPCR	AGCACTAGAAGAGGCGGAAG	
LUC7b-Rv-qPCR	ACGTCAGCAGCAGTGTACTTG	
PP2A	GGCAGAAGTTCGGATAGCAG	
PP2A	CAATGCAGATCTGACGTGCT	
ncRNA- qPCR		
At1g11175	AGATGTAGAACTTTCATGGAG	
At1g11175	TTCACCGACACGACGCAATC	
At1g58590	TTCAGGCGGTGAGAGAGTTA	
At1g58590	TGTGAAAGCAAAGATTCCATG	
At1g70581	AGCTCCGGTGAGTTCTAAGC	
At1g70581	AGGACACTCAAATAGCGG	
At2g07042	TAGATCCAGAGTTTGACGAG	
At2g07042	ACTCACTGTATGACGTTTGAC	

At2g18735	AGAGTTTCATTGCTCGGTGC	
At2g18735	AAGCTAAGGACACGTGAATC	
At4g40065	AAGGTTATCGACTTCTCGG	
At4g40065	AGATCGACGCATAATTGTC	
At1g50055	AGAGTAAACATGAGCGCCGTC	
At1g50055	TGATCTAAGAACCTATCTC	
At2g42485	AAGAGCGACTCCATGGGACTTG	
At2g42485	AACACTGGAGTTACCTGTTGC	
At4g04223	ATAATGCTTCCACCTTCTC	
At4g04223	TGTGGTGTACAAGCATTC	
At5g38005	TGATAAGATGAGAATGGAG	
At5g38005	AGTAATGCAATCGATCCAAG	
PP2A	GGCAGAAGTTCGGATAGCAG	
PP2A	CAATGCAGATCTGACGTGCT	
Splicing validations (RT-PCR)		
Tubulin-F	GAGCCTTACAACGCTACTCTGTCTGTC	Control
Tubulin-R	ACACCAGACATAGTAGCAGAAATCAAG	Control
Intron Retention (cap-dependent introns)		
At1g13880	CAC CAT TTT CAA CCC TAG CCG CTT TC	
At1g13880	AAA CGT TCG CTT GAA ACC CAT	
At3g04610	CAC CCC CGC CTC CAT TGT TAC CGG A	
At3g04610	TTC AGC TTC AGC CAT GAC TAG	
At1g28520	CAC CCA TTA CTC TTC TTT TTA TTT TCC	
At1g28520	CGT CAT AAG AAA TCT CAC CTC	
Intron Retention (LUC7-dependent)		
At2g42010_F	GATGATGTATGAGACAATCTAC	
At2g42010_R	AGATTCTCTTGTATGGCTATG	
At1g15290_F	ATCAAGAGCCATAGAGCAGCAG	
At1g15290_R	TTG AGT ATG CCT CTT GCC AG	
At5g44290_F	TGGAAGACGTGTAGCAGAG	
At5g44290_R	TTCATCATCGGACCTGAGTG	
At2g41560_F1	AACTGTTAATGTGGTTGCTCTG	
At2g41560_R1	GTCACAGTCATTACCCAAGTG	
At1g60995_F1	AGGACACCAATATCGATGAAG	
At1g60995_R1	TGGGTGTTGCTGAAGGTGTG	
At1g70480_F1	TTGACCCTGACAATAGAAGAG	
At1g70480_R1	TTTACCTCCAGTGCACCTG	
At5g41220_F1	CAACCATCTATAGCTGATCTG	
At5g41220_R1	CGACAATTGGATCCTTGCTAG	
At5g49840_F1	GTGGCTTACGGTCTATACTG	
At5g49840_R1	ACATACTCTGAGCTCTCTTAG	
At1g01860_F1	GCTGCTGTCGTAACATTCAAG	

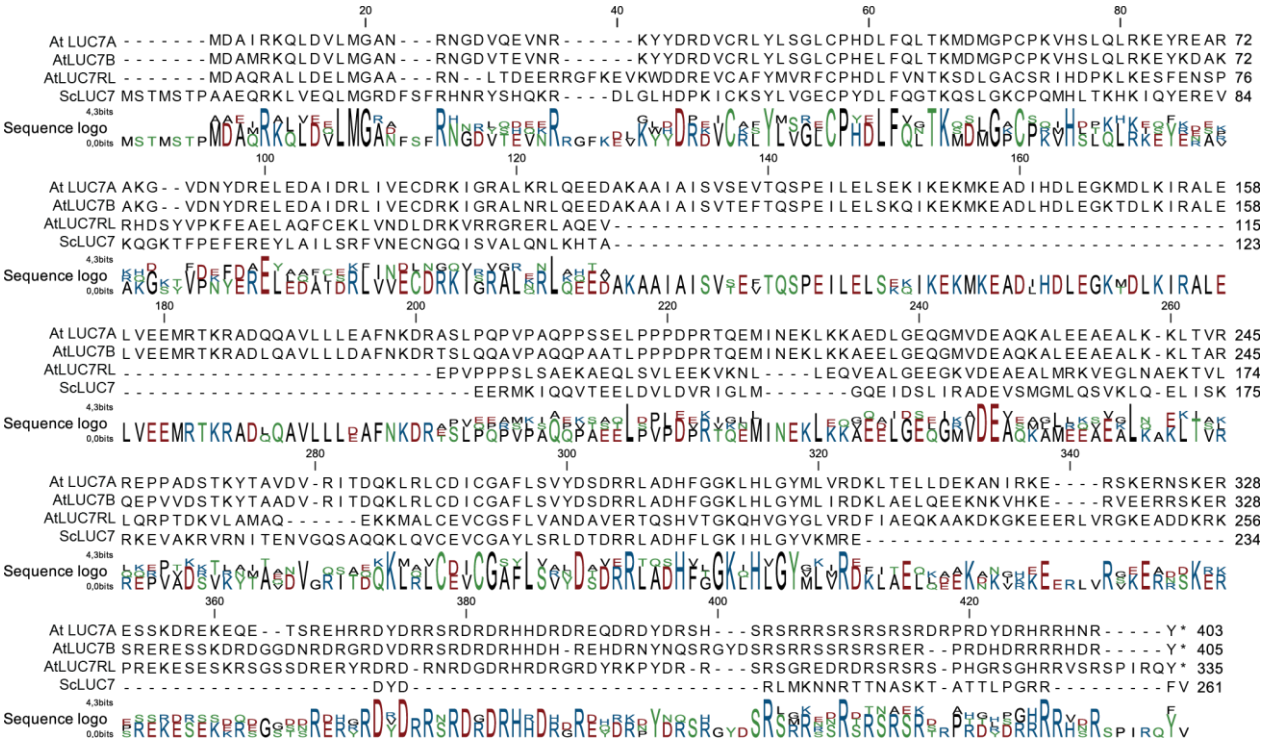
At1g01860_R1	GCGATTACGTTATGCAACTTG	
At5g44290_F1	TCACTTTCAACAGCTGTGGAG	
At5g44290_R1	CCAGTAATGGACCGGACATG	
At3g50110_F1	ACTTTCAAGATAGCGATGGAG	
At3g50110_R1	CTATCATTACGATCTCGACATG	
At1g73740_F1	GGTGCAATGACTTGCTCAGAG	
At1g73740_R1	CTCGCTCATTTCCATCATCAG	
Exon Skipping (LUC7-dependent)		
XE 5 - At2g32700	CTCTCTTCCAGCTGTTTCATG	
XE 5 - At2g32700	AAACATCAAGCATCTTGTCAG	
XE 9 - At5g48150	ATGTGCCTTGTCTCCGACAAC	
XE 9 - At5g48150	CGAGAGTTGTTAACCGGTAAG	
Aternative 3'ss (LUC7-dependent)		
A3SS-1 - AT3G17310	GAAAGACATGTACGATTCACTG	
A3SS-1 - AT3G17310	GAAGTAATGTTAAGTCATTGCTG	
A3SS-2 - AT3G57410	TCAAGGCAGCTAAAGCAGCAG	
A3SS-2 - AT3G57410	GCTGGTTCGCTGGAAGGCTTG	
A3SS-6 - AT1G10890	ACTCATCTCATTCGTATAGCAG	
A3SS-6 - AT1G10890	GTGACAAACACCGCATAACAC	
Alternative 5'ss (LUC7-dependent)		
A5SS-3 - AT4G32060	TCTCCTTCCGAGTTCTTCATG	
A5SS-3 - AT4G32060	CAGCCTCAGCATCTCTTCAG	
A5SS-6 - AT2G11000	TAGGTCAAATCCCAGCATGAG	
A5SS-6 - AT2G11000	CATGGCCTCCATACGAATGAG	
A5SS-8 - AT5G16715	ACCGATGCCACCTCCTAATG	
A5SS-8 - AT5G16715	TGGCAATACCAGCATGATCAG	
Subcellular fractionation (Primers flanking only affected intron)		
AT2G41560	TACCAGTTGATTGTCTTAGG	
AT2G41560_R1	GTCACAGTCATTACCCAAGTG	
AT1G01860	TCCAGCACATATCATCATCTC	
AT1G01860_R1	GCGATTACGTTATGCAACTTG	
At5g44290	ATACGTGAAGGACAATGCAG	
At5g44290_R1	CCAGTAATGGACCGGACATG	
At1g73740	GATGAAGGGGATCAAGTAAG	

At1g73740_R1	CTCGCTCATTTCCATCATCAG	
At1g70480	TGGTCAGGAACAATCTCAG	
At1g70480_R1	TTTACCTCCAGTGCACCTG	
At5g41220	TCAGAAGCAGCGAGAGATG	
At5g41220_R1	CGACAATTGGATCCTTGCTAG	
AT5G49840	AAGGCAGTTCTGGTGGATG	
AT5G49840_R1	ACATACTCTGAGCTCTCTTAG	
Cold stress - qPCR		
AT2G41560_intronic	GTAACAAGAGAATCAAACGGTG	
AT2G41560_ Total	GATCAATAGCCGGGAAATAGAG	
AT2G41560_R1	GTCACAGTCATTACCCAAGTG	
At1g70480_intronic	CTCAAGTTTCTTGAATGGTTCGTG	
At1g70480_Total	CAAGCAAGTGAAAAGTGGATTC	
At1g70480_R1	TTTACCTCCAGTGCACCTG	
At5g41220_unspliced	TCAGAAGCAGCGAGAGATG	
At5g41220_R2_unspliced	AGATCCATGAAATTGGATGAG	
At5g41220_total	AGTACTATTGGTGAGAAGAGTG	
At5g41220_R1_total	CGACAATTGGATCCTTGCTAG	
At5g44290_unspliced	AACATTCAACAGAGCTGACAC	
At5g44290_unspliced2	AGTCTCTTGAGATGACTTAC	
At5g44290_total_1	TCCATAGTGTCAGCTCTGTTG	
At5g44290_total_2	ATACGTGAAGGACAATGCAG	
PP2A	GGCAGAAGTTCGGATAGCAG	
PP2A	CAATGCAGATCTGACGTGCT	

```
Query= Saccharomyces_cerevisiae(strain_ATCC_204508/S288c)_Q07508
Length=261
                                                                      Score     E
Sequences producing significant alignments:                           (Bits)  Value

NP_199954.1  LUC7 N terminus domain-containing protein [Arabid...  85.5    9e-19
AAF01597.1  unknown protein [Arabidopsis thaliana]                  84.0    4e-18
CAC01888.1  putative protein [Arabidopsis thaliana]                67.8    2e-12

ALIGNMENTS
>NP_199954.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 NP_851170.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 NP_001190514.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 NP_001318778.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 NP_001331059.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 AAB68037.1 putative arginine-aspartate-rich RNA binding protein [Arabidopsis thaliana]
 AAB68040.1 putative aspartate-arginine-rich mRNA binding protein [Arabidopsis thaliana]
 BAB09752.1 arginine-aspartate-rich RNA binding protein-like [Arabidopsis thaliana]
 AAL67007.1 putative arginine-aspartate-rich RNA binding protein [Arabidopsis thaliana]
 AAM91771.1 putative arginine-aspartate-rich RNA binding protein [Arabidopsis thaliana]
 AED96076.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 AED96077.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 AED96078.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 OAO89560.1 hypothetical protein AXX17_AT5G50200 [Arabidopsis thaliana]
 ANM69375.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
 ANM69376.1 LUC7 N terminus domain-containing protein [Arabidopsis thaliana]
Length=334

 Score = 85.5 bits (210),  Expect = 9e-19, Method: Compositional matrix adjust.
 Identities = 63/230 (27%), Positives = 114/230 (50%), Gaps = 12/230 (5%)

Query  11   QRKLVEQLMG--RDFSFRHNRYSHQKRDLGLHDPKICKSYLVGECPYDLFQGTKQSLGKC  68
            QR L+++LMG  R+ +    R     +++   D ++C  Y+V  CP+DLF  TK  LG C
Sbjct  4    QRALLDELMGAARNLTDEERRGF---KEVKWDDREVCAFYMVRFCPHDLFVNTKSDLGAC  60

Query  69   PQMHLTKHKIQYEREVKQGKTFPEFEREYLAILSRFVNECNGQISVALQNLKHTAEERMK  128
             ++H  K K  +E   +     P+FE E      + VN+ + ++    + L    E
Sbjct  61   SRIHDPKLKESFENSPRHDSYVPKFEAELAQFCEKLVNDLDRKVRRGRERLAQEVEPVPP  120

Query  129  IQ---QVTEELDVLDVRIGLMGQEIDSLIRADEVSMGMLQSVKLQELISKRKEVAKRVRN  185
                 +  E+L VL+ ++  + +++++L    +V        K++ L +++  + +R
Sbjct  121  PSLSAEKAEQLSVLEEKVKNLLEQVEALGEEGKVDEAEALMRKVEGLNAEKTVLLQR---  177

Query  186  ITENVGQSAQQ-KLQVCEVCGAYLSRLDTDRRLADHFLGKIHLGYVKMRE  234
             T+ V   AQ+ K+ +CEVCG++L   D   R   H  GK H+GY  +R+
Sbjct  178  PTDKVLAMAQEKKMALCEVCGSFLVANDAVERTQSHVTGKQHVGYGLVRD  227


>AAF01597.1 unknown protein [Arabidopsis thaliana]
Length=385

 Score = 84.0 bits (206),  Expect = 4e-18, Method: Compositional matrix adjust.
 Identities = 69/260 (27%), Positives = 115/260 (44%), Gaps = 62/260 (24%)

Query  40   HDPKICKSYLVGECPYDLFQGTKQSLGKCPQMHLTKHKIQYE------------------  81
            +D  +C+ YL G CP+DLFQ TK  +G CP++H  + + +Y
Sbjct  28   YDRDVCRLYLSGLCPHDLFQLTKMDMGPCPKVHSLQLRKEYPLFTIGRALKRLQEEDAKA  87
```

```
Query  82   ------REVKQGKTFPEFEREYLAILSRF-VNECNGQISVALQNLKHTAEERMK------  128
                  EV Q     E   +    +    +++  G++ + ++ L+   E R K
Sbjct  88   AIAISVSEVTQSPEILELSEKIKEKMKEADIHDLEGKMDLKIRALELVEEMRTKRADQQA  147

Query  129  -------------------IQQVTEELDVLDVRIGLMGQEIDSLIRADEV-SMGMLQSVK  168
                              Q  + EL   D R   M  E   L +A+++   GM+   +
Sbjct  148  VLLLEAFNKDRASLPQPVPAQPPSSELPPPDPRTQEMINE--KLKKAEDLGEQGMVDEAQ  205

Query  169  --------LQELISKRKEVAKRVRNITENVGQSAQQKLQVCEVCGAYLSRLDTDRRLADH  220
                   L++L  +R+  A   +    +V +   QKL++C++CGA+LS  D+DRRLADH
Sbjct  206  KALEEAEALKKLTVRREPPADSTKYTAVDV-RITDQKLRLCDICGAFLSVYDSDRRLADH  264

Query  221  FLGKIHLGYVKMREDYDRLM  240
            F GK+HLGY+ +R+    L+
Sbjct  265  FGGKLHLGYMLVRDKLTELL  284


>CAC01888.1 putative protein [Arabidopsis thaliana]
Length=381

 Score = 67.8 bits (164),  Expect = 2e-12, Method: Compositional matrix adjust.
 Identities = 31/78 (40%), Positives = 53/78 (68%), Gaps = 1/78 (1%)

Query  157  DEVSMGMLQSVKLQELISKRKEVAKRVRNITENVGQSAQQKLQVCEVCGAYLSRLDTDRR  216
            DE    + ++  L++L ++++ V    +    +V +   QKL++C++CGA+LS  D+DRR
Sbjct  204  DEAQKALEEAEALKKLTARQEPVVDSTKYTAADV-RITDQKLRLCDICGAFLSVYDSDRR  262

Query  217  LADHFLGKIHLGYVKMRE  234
            LADHF GK+HLGY+ +R+
Sbjct  263  LADHFGGKLHLGYMLIRD  280


 Score = 35.4 bits (80),  Expect = 0.091, Method: Compositional matrix adjust.
 Identities = 21/82 (26%), Positives = 33/82 (40%), Gaps = 25/82 (30%)

Query  40   HDPKICKSYLVGECPYDLFQGTKQSLGKCPQMHLTKHKIQYEREVKQGKTFPEFEREYLA  99
            +D  +C+ YL G CP++LFQ T                          K    ++RE
Sbjct  28   YDRDVCRLYLSGLCPHELFQLT--------------------------AKGVDNYDRELED  62

Query  100  ILSRFVNECNGQISVALQNLKH  121
             + R + EC+ +I  AL  L+
Sbjct  63   AIDRLIVECDRKIGRALNRLQE  84
```

S2: AtLUC7RL is the closest *Arabidopsis thaliana* homolog in yeast. Sequence of yeast LUC7 was used for Blastp search against *Arabidopsis thaliana* database in NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi). NP_1990.54.1 correspond to *LUC7RL* (*At5g51410*).



S3: Alignment of all three *Arabidopsis thaliana* LUC7 proteins with the unique *Saccharomyces cerevisiae* LUC7. The alignment was performed using Muscle and visualized in CLC Main Workbench 7.5.1.

S4: List of LUC7 interacting partners displaying the sequence coverage (Seq. Cov). *LUC7A::LUC7A-eYFP* complemented line was used to immunoprecipitate the LUC7A complex, which was analyzed by mass spectrometry. WT and WT expressing free GFP were used as controls. Four replicates/experiments were performed, which are color coded (IP1 to IP4). For the first replicate the WT control was absent. The sequence coverage for the peptides found in each IP is displayed and are grouped based on the sample type: GFP, LUC7-YFP or WT.

	Protein Names	Gene Names	N° of proteins	Peptides	Seq.Cov. in GFP (%)				Seq. Cov. In LUC7-YFP (%)				Seq. Cov. in WT (%)			Mol. weight [kDa]
					IP1	IP2	IP3	IP4	IP1	IP2	IP3	IP4	IP2	IP3	IP4	
	At3g03340	**LUC7A/UNE6**	**1**	**57**	**5,5**	**21,9**	**0**	**5,5**	**72,4**	**74,1**	**69,2**	**71,9**	**0**	**0**	**0**	**47,4**
	At5g17440	LUC7B	1	28	0	8,2	0	0	34,9	35,4	32,9	32,9	0	0	0	47,2
U1 core proteins	At3g50670	U1-70k	2	9	0	0	0	0	16,6	20,8	6,8	8,9	0	0	0	50,4
	At2g47580	U1A	1	4	0	0	0	0	7,6	10,8	9,2	3,2	0	0	0	28,1
Splicing proteins	At5g44500; At4g20440	SmB-a ; SmB-b	2	4	0	0	0	0	13,4	16,5	3,1	6,3	0	3,1	0	27,0
(including SRs)	At3g07590; At4g02840	SmD1-a; SmD1-b	3	3	0	0	0	0	0	11,2	19	0	0	0	0	12,8
	At2g16940	Splicing factor CC1-like	3	5	0	0	0	0	3,4	7	0	4,8	0	0	0	63,1
	At5g28740	SYF1	1	4	0	0	0	0	1,3	4,8	0	1,3	0	0	0	107,1
	At1g16610	SR45	3	3	0	0	0	0	5,2	8,8	3,2	3,7	0	0	0	44,6
	At1g27650;At5g42820	U2AF35A;U2AF35B	3	5	0	0	0	0	13,5	15,9	9,1	9,1	0	8,4	4,4	34,6
	At4g36690	U2AF65A	4	2	0	0	0	0	2	2	0	2,9	0	0	0	60,7
SR family	At3g13570; At1g55310	SCL30A; SCL33	5	6	0	0	0	0	12,6	19,5	10,7	6,5	0	4,2	0	30,2
	At1g23860; At4g31580; At2g24590	RSZ21RSZ22;RSZ22A	5	4	0	0	0	0	10,5	17	0	4,5	0	0	0	22,5
	At3g55460	SCL30	1	3	0	0	0	0	4,2	6,9	0	0	0	0	0	29,6
	At1g09140	SR30	2	3	0	0	0	0	4,7	10,9	0	0	0	0	0	29,1
	At2g37340	RS2Z33	3	3	0	0	0	0	2,4	9,3	0	0	0	0	0	32,9
RBP and proteins	At2g27100	SERRATE	1	17	0	0	0	0	15,8	20,7	4,7	6,9	0	5,3	0	81,1
involved in others	At2g13540	ABH1	1	3	0	0	0	0	1,1	4,4	0	0	0	0	0	96,5
RNA metabolism	At4g25550	mRNA cleav. Factor-25 kDa	2	3	0	0	0	0	7	18,5	0	18,5	7	7	6,5	22,8
	At2g43970	LARP6B	3	5	0	0	0	0	2,2	7,5	5	7,2	0	2,8	2,8	60,6
	At1g24800; At1g24881; At1g25055; At1g25211; At1g25150	F-box/kelch-repeat proteins	5	3	0	0	0	0	5,5	5,5	5,1	2,8	0	0	0	49,6
	At4g39680	SAP domain	1	6	0	0	0	0	0	11,1	0	2,2	0	0	0	69,4
	At3g08780	-	2	3	0	0	0	0	0	0	4,5	7,1	0	0	0	34,3
	At1g79090	PAT1	1	3	0	0	0	0	1,6	0	0	2,8	0	0	0	88,8
	At2g45640	SAP18	2	5	0	0	0	0	14,5	42,1	7,2	7,2	0	8,6	0	17,2
Kinases	At3g44850	serine/threonine kinase	1	17	0	0	0	0	23,4	24	4,9	11,2	0	0	0	60,7
	At3g53030	SRPK4	1	16	0	0	0	0	15,9	15,7	19,7	13,4	0	0	0	59,4
	At5g22840	SRPK3	1	11	0	0	0	0	8	16,2	4,8	6,5	0	0	0	61,2
Transcription	At1g63210; At1g65440	GTB1/Spt6	4	15	0	0	0	0	0	0	4,3	13,1	0	0	2,7	166,4
	At1g76890	GT-2	1	14	0	1,6	0	0	18,8	21,2	6,6	8,7	0	2,1	2,1	65,8
	At5g13680	ELP1	1	3	0	0	0	0	0	0	1,2	1,8	0	0	0	146,6

S5: Fischer test for intron retention events in WT and *luc7* triple mutant. Only transcripts with at least 3 introns are considered.

IR events specific to WT: Are last intron more often affected?

		Is part of RI		
	+	**-**		
is last	+	9	41.159	41.168
	-	56	306.937	306.993
		65	348.096	348.161

Two – sided p > 0.56
=> normal

		Is part of RI		
	+	**-**		
is middle	+	51	265.774	265.825
	-	14	82.322	82.336
		65	348.096	348.161

Two – sided p > 0.77
=> normal

		Is part of RI		
	+	**-**		
is first	+	5	41.163	41.168
	-	60	306.933	306.993
		65	348.096	348.161

Two – sided p > 0.43
=> normal

IR events specific to *luc7a,b,rl*: Are last intron more often affected?

		Is part of RI		
	+	**-**		
is last	+	34	41.134	41.168
	-	132	306.861	306.993
		166	347.995	348.161

Two – sided p < 0.0016
One sided (greater) p < 0.00098
=> **over-represented**

		Is part of RI		
	+	**-**		
is middle	+	111	265.714	265.825
	-	55	82.281	82.336
		166	347.995	348.161

Two – sided p < 0.0059
One sided (less) p < 0.0036
=> Under represented

		Is part of RI		
	+	**-**		
is first	+	21	41.147	41.168
	-	145	306.848	306.993
		166	347.995	348.161

Two – sided p > 0.71
=> normal

S6: Fisher test for cold and salt samples.

Libraries:

Cold: SRR1655104, SRR1655105, SRR1655106, SRR1655110, SRR1655111, SRR1655112, SRR1655116, SRR1655117, SRR1655118

Salt: SRR1104133, SRR1104134, SRR1104135, SRR1104136

Cold: expressed and significant events of IR in untreated samples (INK1)

WT_1h

		involved in IR "+"	"-"	Total (?)		
					Two-tailed:	p-value = 0.005883
	+	29	25821	25850	One-tailed:	**p-value = 0.003611**
last?	-	88	146931	147019		Overrepresented
	total	117	172752	172869		

		involved in IR "+"	"-"	Total (?)		
					Two-tailed:	p-value = 7.48e-07
	+	56	121113	121169	One-tailed:	**p-value = 4.202e-07**
middle?	-	61	51639	51700		Underrepresented
	total	117	172752	172869		

		involved in IR "+"	"-"	Total (?)		
					Two-tailed:	p-value = 0.0006061
	+	32	25818	25850	One-tailed:	**p-value = 0.0003898**
first?	-	85	146934	147019		Overrepresented
	total	117	172752	172869		

WT_24h

		involved in IR "+"	"-"	Total (?)		
					Two-tailed:	p-value = 0.01515
	+	28	25822	25850	One-tailed:	**p-value = 0.01114**
last?	-	93	146926	147019		Overrepresented
	total	121	172748	172869		

		involved in IR "+"	"-"	Total (?)		
					Two-tailed:	p-value = 2.113e-08
	+	55	121114	121169	One-tailed:	**p-value = 1.445e-08**
middle?	-	66	51634	51700		Underrepresented
	total	121	172748	172869		

		involved in IR "+"	"-"	Total (?)		
					Two-tailed:	p-value = 4.644e-06
	+	38	25812	25850	One-tailed:	**p-value = 3.889e-06**
first?	-	83	146936	147019		Overrepresented
	total	121	172748	172869		

Cold: expressed and significant events of IR in treated samples (INK2)

WT_1h

		involved in IR				
		"+"	"-"	Total (?)	Two-tailed:	p-value = 0.0002691
	+	51	25799	25850	One-tailed:	**p-value = 0.0001771**
last?	-	156	146863	147019		Overrepresented
	total	207	172662	172869		

		involved in IR				
		"+"	"-"	Total (?)	Two-tailed:	p-value = 6.099e-06
	+	114	121055	121169	One-tailed:	**p-value = 3.558e-06**
middle?	-	93	51607	51700		Underrepresented
	total	207	172662	172869		

		involved in IR				
		"+"	"-"	Total (?)	Two-tailed:	p-value = 0.03983
	+	42	25808	25850	One-tailed:	**p-value = 0.02307**
first?	-	165	146854	147019		Overrepresented
	total	207	172662	172869		

WT_24h

		involved in IR				
		"+"	"-"	Total (?)	Two-tailed:	p-value = 0.1172
	+	39	25811	25850		not significant
last?	-	167	146852	147019		
	total	206	172663	172869		

		involved in IR				
		"+"	"-"	Total (?)	Two-tailed:	p-value = 1.24e-06
	+	111	121058	121169	One-tailed:	**p-value = 6.879e-07**
middle?	-	95	51605	51700		underrepresented
	total	206	172663	172869		

		involved in IR				
		"+"	"-"	Total (?)	Two-tailed:	p-value = 5.32e-06
	+	56	25794	25850	One-tailed:	**p-value = 4.267e-06**
first?	-	150	146869	147019		Overrepresented
	total	206	172663	172869		

Salt: expressed and significant events of IR in untreated samples (INK1)

Salt 50 mM

		involved in IR				
Salt 50 mM		"+"	"-"	Total (?)	Two-tailed:	p-value = 0.05534
	+	32	26229	26261		not significant
last?	-	113	148616	148729		
	total	145	174845	174990		

		involved in IR				
Salt 50 mM		"+"	"-"	Total (?)	Two-tailed:	p-value = 3.951e-07
	+	72	122396	122468	One-tailed:	**p-value = 2.387e-07**
middle?	-	73	52449	52522		underrepresented
	total	145	174845	174990		

		involved in IR				
Salt 50 mM		"+"	"-"	Total (?)	Two-tailed:	p-value = 3.702e-05
	+	41	26220	26261	One-tailed:	**p-value = 3.097e-05**
first?	-	104	148625	148729		Overrepresented
	total	145	174845	174990		

Salt 150 mM

		involved in IR				
Salt 150 mM		"+"	"-"	Total (?)	Two-tailed:	p-value = 0.03434
	+	25	26236	26261	One-tailed:	**p-value = 0.02516**
last?	-	87	148642	148729		Overrepresented
	total	112	174878	174990		

		involved in IR				
Salt 150 mM		"+"	"-"	Total (?)	Two-tailed:	p-value = 0.0001186
	+	59	122409	122468	One-tailed:	**p-value = 8.68e-05**
middle?	-	53	52469	52522		Underrepresented
	total	112	174878	174990		

		involved in IR				
Salt 150 mM		"+"	"-"	Total (?)	Two-tailed:	p-value = 0.00514
	+	28	26233	26261	One-tailed:	**p-value = 0.003873**
first?	-	84	148645	148729		Overrepresented
	total	112	174878	174990		

Salt 300 mM

		involved in IR "+"	"-"	Total (?)		
Salt 300 mM					Two-tailed:	p-value = 0.004233
	+	43	26218	26261	One-tailed:	**p-value = 0.00307**
last?	-	146	148583	148729		Overrepresented
	total	189	174801	174990		

		involved in IR "+"	"-"	Total (?)		
Salt 300 mM					Two-tailed:	p-value = 4.194e-08
	+	96	122372	122468	One-tailed:	**p-value = 2.496e-08**
middle?	-	93	52429	52522		Underrepresented
	total	189	174801	174990		

		involved in IR "+"	"-"	Total (?)		
Salt 300 mM					Two-tailed:	p-value = 5.619e-05
	+	50	26211	26261	One-tailed:	**p-value = 3.25e-05**
first?	-	139	148590	148729		Overrepresented
	total	189	174801	174990		

Salt: expressed and significant events of IR in treated samples (INK2)

Salt 50 mM

		involved in IR "+"	"-"	Total (?)		
Salt 50 mM					Two-tailed:	p-value = 0.0005737
	+	54	26207	26261	One-tailed:	**p-value = 0.0004132**
last?	-	175	148554	148729		Overrepresented
	total	229	174761	174990		

		involved in IR "+"	"-"	Total (?)		
Salt 50 mM					Two-tailed:	p-value = 1.666e-12
	+	109	122359	122468	One-tailed:	**p-value = 1.325e-12**
middle?	-	120	52402	52522		Underrepresenetd
	total	229	174761	174990		

		involved in IR "+"	"-"	Total (?)		
Salt 50 mM					Two-tailed:	p-value = 8.286e-08
	+	66	26195	26261	One-tailed:	**p-value = 6.991e-08**
first?	-	163	148566	148729		Overrepresented
	total	229	174761	174990		

Salt 150 mM

		involved in IR "+"	"-"	Total (?)		
Salt 150 mM					Two-tailed:	p-value = 1.214e-05
	+	48	26213	26261	One-tailed:	**p-value = 7.23e-06**
last?	-	122	148607	148729		Overrepresented
	total	170	174820	174990		

		involved in IR "+"	"-"	Total (?)		
Salt 150 mM					Two-tailed:	p-value = 1.08e-11
	+	76	122392	122468	One-tailed:	**p-value = 6.42e-12**
middle?	-	94	52428	52522		Underrepresented
	total	170	174820	174990		

		involved in IR "+"	"-"	Total (?)		
Salt 150 mM					Two-tailed:	p-value = 5.441e-05
	+	46	26215	26261	One-tailed:	**p-value = 3.592e-05**
first?	-	124	148605	148729		Overrepresented
	total	170	174820	174990		

Salt 300 mM

		involved in IR "+"	"-"	Total (?)		
Salt 300 mM					Two-tailed:	p-value = 0.003581
	+	45	26216	26261	One-tailed:	**p-value = 0.002308**
last?	-	152	148577	148729		Overrepresented
	total	197	174793	174990		

		involved in IR "+"	"-"	Total (?)		
Salt 300 mM					Two-tailed:	p-value = 8.611e-08
	+	102	122366	122468	One-tailed:	**p-value = 6.185e-08**
middle?	-	95	52427	52522		Underepresented
	total	197	174793	174990		

		involved in IR "+"	"-"	Total (?)		
Salt 300 mM					Two-tailed:	p-value = 0.0001323
	+	50	26211	26261	One-tailed:	**p-value = 0.0001019**
first?	-	147	148582	148729		Overrepresented
	total	197	174793	174990		

S7: GO analysis for genes upregulated in *luc7a-2 b-1 rl-1*.

			Query		Observ.	Background		Expect.		
GO_acc	type	Term	item	total	Freq.	item	total	Freq.	pvalue	FDR
GO:0042221	P	response to chemical stimulus	213	806	0,264	3978	28397	0,140	6,90E-20	2,40E-16
GO:0009743	P	response to carbohydrate stimulus	78	806	0,097	812	28397	0,029	2,20E-19	3,80E-16
GO:0010200	P	response to chitin	54	806	0,067	421	28397	0,015	1,40E-18	1,60E-15
GO:0010033	P	response to organic substance	161	806	0,200	2754	28397	0,097	5,00E-18	4,40E-15
GO:0050896	P	response to stimulus	287	806	0,356	6292	28397	0,222	8,30E-18	5,80E-15
GO:0006950	P	response to stress	210	806	0,261	4089	28397	0,144	1,30E-17	7,80E-15
GO:0006952	P	defense response	112	806	0,139	1653	28397	0,058	1,20E-16	6,20E-14
GO:0045087	P	innate immune response	69	806	0,086	930	28397	0,033	3,30E-12	1,50E-09
GO:0050832	P	defense response to fungus	38	806	0,047	342	28397	0,012	9,80E-12	3,80E-09
GO:0009751	P	response to salicylic acid stimulus	45	806	0,056	470	28397	0,017	1,30E-11	4,50E-09
GO:0009620	P	response to fungus	46	806	0,057	499	28397	0,018	2,40E-11	7,80E-09
GO:0051707	P	response to other organism	88	806	0,109	1421	28397	0,050	3,00E-11	8,20E-09
GO:0070482	P	response to oxygen levels	21	806	0,026	104	28397	0,004	3,40E-11	8,20E-09
GO:0002376	P	immune system process	69	806	0,086	984	28397	0,035	3,50E-11	8,20E-09
GO:0006955	P	immune response	69	806	0,086	984	28397	0,035	3,50E-11	8,20E-09
GO:0002679	P	respiratory burst during defense response	22	806	0,027	121	28397	0,004	6,90E-11	1,40E-08
GO:0045730	P	respiratory burst	22	806	0,027	121	28397	0,004	6,90E-11	1,40E-08
GO:0009719	P	response to endogenous stimulus	92	806	0,114	1615	28397	0,057	6,40E-10	1,30E-07
GO:0001666	P	response to hypoxia	19	806	0,024	100	28397	0,004	7,10E-10	1,30E-07
GO:0023052	P	signaling	121	806	0,150	2376	28397	0,084	7,40E-10	1,30E-07
GO:0071446	P	cellular response to salicylic acid stimulus	35	806	0,043	351	28397	0,012	9,30E-10	1,60E-07
GO:0070887	P	cellular response to chemical stimulus	83	806	0,103	1417	28397	0,050	1,40E-09	2,30E-07
GO:0009863	P	salicylic acid mediated signaling pathway	34	806	0,042	349	28397	0,012	2,80E-09	4,10E-07
GO:0051704	P	multi-organism process	98	806	0,122	1820	28397	0,064	2,70E-09	4,10E-07
GO:0007165	P	signal transduction	92	806	0,114	1670	28397	0,059	3,10E-09	4,40E-07
GO:0009605	P	response to external stimulus	68	806	0,084	1087	28397	0,038	4,30E-09	5,80E-07
GO:0009414	P	response to water deprivation	37	806	0,046	416	28397	0,015	5,40E-09	7,00E-07
GO:0009415	P	response to water	37	806	0,046	424	28397	0,015	8,60E-09	1,10E-06
GO:0009753	P	response to jasmonic acid stimulus	39	806	0,048	471	28397	0,017	1,30E-08	1,60E-06
GO:0009814	P	defense response, incompatible interaction	42	806	0,052	536	28397	0,019	1,50E-08	1,80E-06
GO:0009611	P	response to wounding	32	806	0,040	340	28397	0,012	1,70E-08	2,00E-06
GO:0009862	P	systemic acquired resistance, salicylic acid mediated signaling pathway	27	806	0,033	251	28397	0,009	1,80E-08	2,00E-06
GO:0051716	P	cellular response to stimulus	115	806	0,143	2355	28397	0,083	1,90E-08	2,10E-06
GO:0010310	P	regulation of hydrogen peroxide metabolic process	23	806	0,029	187	28397	0,007	2,20E-08	2,30E-06
GO:0080010	P	regulation of oxygen and reactive oxygen species metabolic process	23	806	0,029	188	28397	0,007	2,50E-08	2,50E-06
GO:0009750	P	response to fructose stimulus	20	806	0,025	144	28397	0,005	3,00E-08	3,00E-06
GO:0071453	P	cellular response to oxygen levels	10	806	0,012	26	28397	0,001	3,20E-08	3,10E-06
GO:0009607	P	response to biotic stimulus	89	806	0,110	1687	28397	0,059	3,90E-08	3,60E-06
GO:0023046	P	signaling process	92	806	0,114	1768	28397	0,062	4,10E-08	3,60E-06
GO:0023060	P	signal transmission	92	806	0,114	1767	28397	0,062	4,00E-08	3,60E-06
GO:0009646	P	response to absence of light	11	806	0,014	37	28397	0,001	5,90E-08	5,10E-06
GO:0034284	P	response to monosaccharide stimulus	21	806	0,026	170	28397	0,006	8,50E-08	7,00E-06
GO:0009746	P	response to hexose stimulus	21	806	0,026	170	28397	0,006	8,50E-08	7,00E-06
GO:0009725	P	response to hormone stimulus	75	806	0,093	1375	28397	0,048	1,50E-07	1,20E-05
GO:0071456	P	cellular response to hypoxia	9	806	0,011	24	28397	0,001	1,90E-07	1,50E-05
GO:0007242	P	intracellular signaling cascade	69	806	0,086	1252	28397	0,044	3,30E-07	2,50E-05
GO:0009627	P	systemic acquired resistance	34	806	0,042	445	28397	0,016	6,20E-07	4,70E-05

GO:0071495	P	cellular response to endogenous stimulus	50	806	0,062	815	28397	0,029	8,70E-07	6,40E-05
GO:0009723	P	response to ethylene stimulus	29	806	0,036	353	28397	0,012	1,10E-06	7,80E-05
GO:0009626	P	plant-type hypersensitive response	31	806	0,038	401	28397	0,014	1,50E-06	0,00011
GO:0034050	P	host programmed cell death induced by symbiont	31	806	0,038	402	28397	0,014	1,60E-06	0,00011
GO:0071310	P	cellular response to organic substance	66	806	0,082	1234	28397	0,043	1,60E-06	0,00011
GO:0034285	P	response to disaccharide stimulus	21	806	0,026	213	28397	0,008	2,50E-06	0,00016
GO:0010363	P	regulation of plant-type hypersensitive response	29	806	0,036	371	28397	0,013	2,70E-06	0,00018
GO:0045088	P	regulation of innate immune response	31	806	0,038	415	28397	0,015	3,00E-06	0,00019
GO:0002682	P	regulation of immune system process	31	806	0,038	419	28397	0,015	3,60E-06	0,00022
GO:0050776	P	regulation of immune response	31	806	0,038	419	28397	0,015	3,60E-06	0,00022
GO:0080135	P	regulation of cellular response to stress	29	806	0,036	379	28397	0,013	4,00E-06	0,00024
GO:0010941	P	regulation of cell death	30	806	0,037	405	28397	0,014	5,00E-06	0,0003
GO:0012501	P	programmed cell death	32	806	0,040	451	28397	0,016	5,60E-06	0,00033
GO:0009744	P	response to sucrose stimulus	20	806	0,025	210	28397	0,007	6,90E-06	0,0004
GO:0006612	P	protein targeting to membrane	29	806	0,036	392	28397	0,014	7,30E-06	0,00042
GO:0043067	P	regulation of programmed cell death	29	806	0,036	397	28397	0,014	9,10E-06	0,00051
GO:0009617	P	response to bacterium	37	806	0,046	577	28397	0,020	9,30E-06	0,00051
GO:0008219	P	cell death	33	806	0,041	500	28397	0,018	1,60E-05	0,00087
GO:0016265	P	death	33	806	0,041	500	28397	0,018	1,60E-05	0,00087
GO:0048583	P	regulation of response to stimulus	40	806	0,050	667	28397	0,023	1,80E-05	0,00093
GO:0048585	P	negative regulation of response to stimulus	26	806	0,032	349	28397	0,012	1,90E-05	0,00099
GO:0031347	P	regulation of defense response	34	806	0,042	529	28397	0,019	2,00E-05	0,001
GO:0000165	P	MAPKKK cascade	19	806	0,024	209	28397	0,007	2,10E-05	0,0011
GO:0042743	P	hydrogen peroxide metabolic process	25	806	0,031	335	28397	0,012	2,70E-05	0,0013
GO:0031348	P	negative regulation of defense response	22	806	0,027	273	28397	0,010	2,80E-05	0,0013
GO:0002252	P	immune effector process	22	806	0,027	273	28397	0,010	2,80E-05	0,0013
GO:0009612	P	response to mechanical stimulus	10	806	0,012	63	28397	0,002	3,10E-05	0,0015
GO:0080134	P	regulation of response to stress	34	806	0,042	544	28397	0,019	3,50E-05	0,0016
GO:0009867	P	jasmonic acid mediated signaling pathway	22	806	0,027	282	28397	0,010	4,30E-05	0,002
GO:0071395	P	cellular response to jasmonic acid stimulus	22	806	0,027	282	28397	0,010	4,30E-05	0,002
GO:0006800	P	oxygen and reactive oxygen species metabolic process	25	806	0,031	347	28397	0,012	4,60E-05	0,0021
GO:0007243	P	protein kinase cascade	19	806	0,024	223	28397	0,008	4,80E-05	0,0021
GO:0042742	P	defense response to bacterium	27	806	0,033	394	28397	0,014	5,20E-05	0,0023
GO:0009061	P	anaerobic respiration	6	806	0,007	20	28397	0,001	6,20E-05	0,0027
GO:0031668	P	cellular response to extracellular stimulus	26	806	0,032	388	28397	0,014	0,0001	0,0042
GO:0033554	P	cellular response to stress	68	806	0,084	1473	28397	0,052	0,0001	0,0042
GO:0071496	P	cellular response to external stimulus	26	806	0,032	389	28397	0,014	0,0001	0,0043
GO:0042538	P	hyperosmotic salinity response	15	806	0,019	162	28397	0,006	0,00013	0,0052
GO:0009696	P	salicylic acid metabolic process	18	806	0,022	222	28397	0,008	0,00013	0,0055
GO:0071215	P	cellular response to abscisic acid stimulus	20	806	0,025	267	28397	0,009	0,00016	0,0064
GO:0009737	P	response to abscisic acid stimulus	35	806	0,043	621	28397	0,022	0,00018	0,0073
GO:0050794	P	regulation of cellular process	170	806	0,211	4595	28397	0,162	0,00019	0,0074
GO:0009697	P	salicylic acid biosynthetic process	17	806	0,021	209	28397	0,007	0,0002	0,0076
GO:0009991	P	response to extracellular stimulus	26	806	0,032	406	28397	0,014	0,0002	0,0076
GO:0043069	P	negative regulation of programmed cell death	15	806	0,019	170	28397	0,006	0,0002	0,0078
GO:0060548	P	negative regulation of cell death	15	806	0,019	174	28397	0,006	0,00026	0,0098
GO:0007568	P	aging	13	806	0,016	145	28397	0,005	0,00046	0,017
GO:0009738	P	abscisic acid mediated signaling pathway	18	806	0,022	252	28397	0,009	0,00056	0,021
GO:0009682	P	induced systemic resistance	5	806	0,006	22	28397	0,001	0,00077	0,028
GO:0045454	P	cell redox homeostasis	12	806	0,015	136	28397	0,005	0,00086	0,031
GO:0002237	P	response to molecule of bacterial origin	10	806	0,012	102	28397	0,004	0,0011	0,039

GO:0032870	P	cellular response to hormone stimulus	33	806	0,041	641	28397	0,023	0,0012	0,043
GO:0009595	P	detection of biotic stimulus	10	806	0,012	104	28397	0,004	0,0013	0,044
GO:0006972	P	hyperosmotic response	17	806	0,021	251	28397	0,009	0,0014	0,047
GO:0000160	P	two-component signal transduction system (phosphorelay)	13	806	0,016	165	28397	0,006	0,0014	0,048
GO:0010150	P	leaf senescence	8	806	0,010	70	28397	0,002	0,0014	0,048
GO:0009055	F	electron carrier activity	39	806	0,048	525	28397	0,018	1,80E-07	0,00013
GO:0015035	F	protein disulfide oxidoreductase activity	13	806	0,016	94	28397	0,003	8,20E-06	0,0029
GO:0015036	F	disulfide oxidoreductase activity	13	806	0,016	102	28397	0,004	1,80E-05	0,0042
GO:0016701	F	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen	8	806	0,010	40	28397	0,001	4,60E-05	0,0082
GO:0030613	F	oxidoreductase activity, acting on phosphorus or arsenic in donors	5	806	0,006	14	28397	0,000	0,00013	0,013
GO:0016491	F	oxidoreductase activity	75	806	0,093	1691	28397	0,060	0,00014	0,013
GO:0008794	F	arsenate reductase (glutaredoxin) activity	5	806	0,006	14	28397	0,000	0,00013	0,013
GO:0030614	F	oxidoreductase activity, acting on phosphorus or arsenic in donors, with disulfide as acceptor	5	806	0,006	14	28397	0,000	0,00013	0,013
GO:0030611	F	arsenate reductase activity	5	806	0,006	15	28397	0,001	0,00017	0,014
GO:0030528	F	transcription regulator activity	74	806	0,092	1740	28397	0,061	0,00052	0,037

S8: GO analysis for genes downregulated in *luc7a-2 b-1 rl-1*

			Query		Observ.	Background		Expect.		
GO_acc	type	Term	item	total	Freq.	item	total	Freq.	pvalue	FDR
GO:0010264	P	myo-inositol hexakisphosphate biosynthetic process	19	680	0,028	65	28397	0,002	6,40E-14	7,90E-11
GO:0046173	P	polyol biosynthetic process	21	680	0,031	81	28397	0,003	2,30E-14	7,90E-11
GO:0032958	P	inositol phosphate biosynthetic process	19	680	0,028	66	28397	0,002	8,00E-14	7,90E-11
GO:0033517	P	myo-inositol hexakisphosphate metabolic process	19	680	0,028	65	28397	0,002	6,40E-14	7,90E-11
GO:0006021	P	inositol biosynthetic process	19	680	0,028	67	28397	0,002	1,00E-13	8,00E-11
GO:0006020	P	inositol metabolic process	20	680	0,029	98	28397	0,003	4,40E-12	2,50E-09
GO:0043647	P	inositol phosphate metabolic process	19	680	0,028	86	28397	0,003	4,50E-12	2,50E-09
GO:0019751	P	polyol metabolic process	22	680	0,032	136	28397	0,005	2,20E-11	1,10E-08
GO:0019748	P	secondary metabolic process	72	680	0,106	1247	28397	0,044	2,50E-11	1,10E-08
GO:0010114	P	response to red light	16	680	0,024	104	28397	0,004	2,20E-08	8,60E-06
GO:0034637	P	cellular carbohydrate biosynthetic process	53	680	0,078	941	28397	0,033	2,70E-08	9,70E-06
GO:0055114	P	oxidation reduction	67	680	0,099	1364	28397	0,048	5,90E-08	2,00E-05
GO:0015979	P	photosynthesis	32	680	0,047	435	28397	0,015	7,60E-08	2,30E-05
GO:0006733	P	oxidoreduction coenzyme metabolic process	24	680	0,035	267	28397	0,009	1,10E-07	3,20E-05
GO:0006066	P	alcohol metabolic process	58	680	0,085	1143	28397	0,040	1,80E-07	3,70E-05
GO:0042440	P	pigment metabolic process	28	680	0,041	361	28397	0,013	1,80E-07	3,70E-05
GO:0046496	P	nicotinamide nucleotide metabolic process	21	680	0,031	212	28397	0,007	1,60E-07	3,70E-05
GO:0005975	P	carbohydrate metabolic process	94	680	0,138	2249	28397	0,079	1,80E-07	3,70E-05
GO:0019362	P	pyridine nucleotide metabolic process	21	680	0,031	214	28397	0,008	1,90E-07	3,70E-05
GO:0006769	P	nicotinamide metabolic process	21	680	0,031	212	28397	0,007	1,60E-07	3,70E-05
GO:0044262	P	cellular carbohydrate metabolic process	79	680	0,116	1778	28397	0,063	2,00E-07	3,80E-05
GO:0006098	P	pentose-phosphate shunt	20	680	0,029	200	28397	0,007	2,80E-07	4,90E-05
GO:0006740	P	NADPH regeneration	20	680	0,029	201	28397	0,007	3,00E-07	5,10E-05
GO:0043603	P	cellular amide metabolic process	21	680	0,031	222	28397	0,008	3,30E-07	5,40E-05
GO:0006739	P	NADP metabolic process	20	680	0,029	204	28397	0,007	3,70E-07	5,80E-05
GO:0009820	P	alkaloid metabolic process	21	680	0,031	225	28397	0,008	4,00E-07	5,90E-05
GO:0051186	P	cofactor metabolic process	43	680	0,063	753	28397	0,027	4,10E-07	5,90E-05
GO:0010218	P	response to far red light	14	680	0,021	99	28397	0,003	4,20E-07	5,90E-05
GO:0009416	P	response to light stimulus	58	680	0,085	1188	28397	0,042	5,80E-07	7,90E-05
GO:0046148	P	pigment biosynthetic process	23	680	0,034	282	28397	0,010	1,00E-06	1,30E-04
GO:0016051	P	carbohydrate biosynthetic process	53	680	0,078	1070	28397	0,038	1,20E-06	1,50E-04
GO:0046165	P	alcohol biosynthetic process	23	680	0,034	293	28397	0,010	1,90E-06	2,30E-04
GO:0015994	P	chlorophyll metabolic process	18	680	0,026	189	28397	0,007	2,00E-06	2,40E-04
GO:0009314	P	response to radiation	58	680	0,085	1263	28397	0,044	3,60E-06	4,20E-04
GO:0044283	P	small molecule biosynthetic process	77	680	0,113	1865	28397	0,066	4,10E-06	4,60E-04
GO:0019684	P	photosynthesis, light reaction	24	680	0,035	333	28397	0,012	4,40E-06	4,80E-04
GO:0009637	P	response to blue light	14	680	0,021	126	28397	0,004	5,60E-06	5,90E-04
GO:0006778	P	porphyrin metabolic process	19	680	0,028	226	28397	0,008	5,80E-06	6,00E-04
GO:0033013	P	tetrapyrrole metabolic process	19	680	0,028	227	28397	0,008	6,20E-06	6,20E-04
GO:0019761	P	glucosinolate biosynthetic process	16	680	0,024	172	28397	0,006	9,80E-06	9,20E-04
GO:0016144	P	S-glycoside biosynthetic process	16	680	0,024	172	28397	0,006	9,80E-06	9,20E-04
GO:0019758	P	glycosinolate biosynthetic process	16	680	0,024	172	28397	0,006	9,80E-06	9,20E-04
GO:0008152	P	metabolic process	383	680	0,563	13633	28397	0,480	1,10E-05	9,80E-04
GO:0009639	P	response to red or far red light	25	680	0,037	406	28397	0,014	3,50E-05	3,10E-03
GO:0016138	P	glycoside biosynthetic process	17	680	0,025	215	28397	0,008	3,70E-05	3,20E-03
GO:0016137	P	glycoside metabolic process	19	680	0,028	268	28397	0,009	5,30E-05	4,50E-03
GO:0034641	P	cellular nitrogen compound metabolic process	57	680	0,084	1373	28397	0,048	6,80E-05	5,60E-03
GO:0016143	P	S-glycoside metabolic process	16	680	0,024	205	28397	0,007	7,10E-05	5,60E-03

GO:0019757	P	glycosinolate metabolic process	16	680	0,024	205	28397	0,007	7,10E-05	5,60E-03
GO:0019760	P	glucosinolate metabolic process	16	680	0,024	205	28397	0,007	7,10E-05	0,0056
GO:0006732	P	coenzyme metabolic process	28	680	0,041	508	28397	0,018	7,60E-05	0,0059
GO:0015995	P	chlorophyll biosynthetic process	12	680	0,018	125	28397	0,004	9,40E-05	0,0072
GO:0009658	P	chloroplast organization	17	680	0,025	237	28397	0,008	1,10E-04	0,0084
GO:0051188	P	cofactor biosynthetic process	23	680	0,034	388	28397	0,014	1,20E-04	0,0088
GO:0006007	P	glucose catabolic process	26	680	0,038	474	28397	0,017	1,50E-04	0,01
GO:0019320	P	hexose catabolic process	26	680	0,038	476	28397	0,017	1,60E-04	0,011
GO:0046365	P	monosaccharide catabolic process	26	680	0,038	480	28397	0,017	1,80E-04	0,012
GO:0006779	P	porphyrin biosynthetic process	13	680	0,019	157	28397	0,006	1,90E-04	0,013
GO:0044275	P	cellular carbohydrate catabolic process	28	680	0,041	545	28397	0,019	2,30E-04	0,015
GO:0033014	P	tetrapyrrole biosynthetic process	13	680	0,019	160	28397	0,006	2,30E-04	0,015
GO:0046164	P	alcohol catabolic process	26	680	0,038	491	28397	0,017	2,50E-04	0,016
GO:0016556	P	mRNA modification	10	680	0,015	101	28397	0,004	2,80E-04	0,018
GO:0046483	P	heterocycle metabolic process	43	680	0,063	1023	28397	0,036	4,20E-04	0,026
GO:0009657	P	plastid organization	22	680	0,032	398	28397	0,014	4,20E-04	0,026
GO:0016052	P	carbohydrate catabolic process	28	680	0,041	573	28397	0,020	4,90E-04	0,03
GO:0044281	P	small molecule metabolic process	128	680	0,188	4065	28397	0,143	8,20E-04	0,049
GO:0003824	F	catalytic activity	282	680	0,415	8787	28397	0,309	6,70E-09	4,30E-06
GO:0016491	F	oxidoreductase activity	76	680	0,112	1691	28397	0,060	2,30E-07	7,30E-05
GO:0046527	F	glucosyltransferase activity	15	680	0,022	142	28397	0,005	4,60E-06	0,00098
GO:0035251	F	UDP-glucosyltransferase activity	12	680	0,018	114	28397	0,004	4,20E-05	0,0067
GO:0016759	F	cellulose synthase activity	7	680	0,010	37	28397	0,001	6,70E-05	0,0072
GO:0016758	F	transferase activity, transferring hexosyl groups	24	680	0,035	396	28397	0,014	6,30E-05	0,0072
GO:0008194	F	UDP-glycosyltransferase activity	17	680	0,025	239	28397	0,008	0,00012	0,011
GO:0004497	F	monooxygenase activity	11	680	0,016	112	28397	0,004	0,00015	0,012
GO:0016765	F	transferase activity, transferring alkyl or aryl (other than methyl) groups	11	680	0,016	123	28397	0,004	0,00032	0,023
GO:0043167	F	ion binding	101	680	0,149	3030	28397	0,107	0,00053	0,026
GO:0003700	F	transcription factor activity	63	680	0,093	1682	28397	0,059	0,00044	0,026
GO:0016760	F	cellulose synthase (UDP-forming) activity	5	680	0,007	24	28397	0,001	0,00051	0,026
GO:0043169	F	cation binding	101	680	0,149	3029	28397	0,107	0,00053	0,026
GO:0030528	F	transcription regulator activity	64	680	0,094	1740	28397	0,061	0,00062	0,028
GO:0005506	F	iron ion binding	25	680	0,037	499	28397	0,018	0,00069	0,03
GO:0016757	F	transferase activity, transferring glycosyl groups	28	680	0,041	596	28397	0,021	0,00088	0,035
GO:0044435	C	plastid part	62	680	0,091	1252	28397	0,044	1,50E-07	2,80E-05
GO:0044434	C	chloroplast part	61	680	0,090	1211	28397	0,043	1,10E-07	2,80E-05
GO:0009570	C	chloroplast stroma	36	680	0,053	603	28397	0,021	1,40E-06	0,00017
GO:0009532	C	plastid stroma	36	680	0,053	637	28397	0,022	4,50E-06	0,00042
GO:0009536	C	plastid	139	680	0,204	4037	28397	0,142	8,00E-06	0,00061
GO:0009507	C	chloroplast	136	680	0,200	3959	28397	0,139	1,20E-05	0,00073
GO:0009579	C	thylakoid	30	680	0,044	518	28397	0,018	1,80E-05	0,00097
GO:0010319	C	stromule	7	680	0,010	37	28397	0,001	6,70E-05	0,003
GO:0005576	C	extracellular region	100	680	0,147	2824	28397	0,099	7,00E-05	0,003
GO:0009526	C	plastid envelope	30	680	0,044	598	28397	0,021	0,00021	0,0065
GO:0055035	C	plastid thylakoid membrane	20	680	0,029	324	28397	0,011	0,0002	0,0065
GO:0009535	C	chloroplast thylakoid membrane	20	680	0,029	322	28397	0,011	0,00018	0,0065
GO:0031976	C	plastid thylakoid	23	680	0,034	425	28397	0,015	0,00041	0,01
GO:0042651	C	thylakoid membrane	20	680	0,029	341	28397	0,012	0,00037	0,01
GO:0009534	C	chloroplast thylakoid	23	680	0,034	425	28397	0,015	0,00041	0,01
GO:0031984	C	organelle subcompartment	23	680	0,034	428	28397	0,015	0,00045	0,011

GO:0009941	C	chloroplast envelope	28	680	0,041	573	28397	0,020	0,00049	0,011
GO:0034357	C	photosynthetic membrane	20	680	0,029	355	28397	0,013	0,0006	0,013
GO:0044436	C	thylakoid part	21	680	0,031	388	28397	0,014	0,00072	0,014

S9: GO analysis for genes upregulated in *se-1*.

			Query		Observ.	Background		Expect.		
GO_acc	type	Term	item	total	Freq.	item	total	Freq.	pvalue	FDR
GO:0050896	P	response to stimulus	388	1147	0,338	6292	28397	0,222	5,80E-19	1,40E-15
GO:0042221	P	response to chemical stimulus	277	1147	0,241	3978	28397	0,140	3,00E-19	1,40E-15
GO:0031668	P	cellular response to extracellular stimulus	58	1147	0,051	388	28397	0,014	8,00E-16	1,10E-12
GO:0071496	P	cellular response to external stimulus	58	1147	0,051	389	28397	0,014	8,90E-16	1,10E-12
GO:0010033	P	response to organic substance	201	1147	0,175	2754	28397	0,097	1,20E-15	1,10E-12
GO:0009991	P	response to extracellular stimulus	59	1147	0,051	406	28397	0,014	1,40E-15	1,10E-12
GO:0009605	P	response to external stimulus	106	1147	0,092	1087	28397	0,038	1,60E-15	1,20E-12
GO:0031669	P	cellular response to nutrient levels	54	1147	0,047	350	28397	0,012	2,50E-15	1,50E-12
GO:0031667	P	response to nutrient levels	55	1147	0,048	367	28397	0,013	4,10E-15	2,30E-12
GO:0009719	P	response to endogenous stimulus	136	1147	0,119	1615	28397	0,057	6,60E-15	3,30E-12
GO:0009267	P	cellular response to starvation	51	1147	0,044	336	28397	0,012	2,70E-14	1,20E-11
GO:0016036	P	cellular response to phosphate starvation	36	1147	0,031	169	28397	0,006	2,90E-14	1,20E-11
GO:0042594	P	response to starvation	51	1147	0,044	344	28397	0,012	6,00E-14	2,30E-11
GO:0035195	P	gene silencing by miRNA	38	1147	0,033	197	28397	0,007	8,70E-14	3,10E-11
GO:0009725	P	response to hormone stimulus	118	1147	0,103	1375	28397	0,048	1,50E-13	5,00E-11
GO:0010468	P	regulation of gene expression	186	1147	0,162	2695	28397	0,095	2,00E-12	6,20E-10
GO:0010629	P	negative regulation of gene expression	67	1147	0,058	611	28397	0,022	3,10E-12	8,90E-10
GO:0010605	P	negative regulation of macromolecule metabolic process	67	1147	0,058	636	28397	0,022	1,50E-11	4,20E-09
GO:0035194	P	posttranscriptional gene silencing by RNA	39	1147	0,034	254	28397	0,009	2,20E-11	5,60E-09
GO:0007154	P	cell communication	74	1147	0,065	758	28397	0,027	3,50E-11	8,60E-09
GO:0051716	P	cellular response to stimulus	163	1147	0,142	2355	28397	0,083	5,00E-11	1,10E-08
GO:0016441	P	posttranscriptional gene silencing	39	1147	0,034	262	28397	0,009	4,90E-11	1,10E-08
GO:0009743	P	response to carbohydrate stimulus	77	1147	0,067	812	28397	0,029	4,90E-11	1,10E-08
GO:0023052	P	signaling	164	1147	0,143	2376	28397	0,084	5,30E-11	1,10E-08
GO:0071495	P	cellular response to endogenous stimulus	77	1147	0,067	815	28397	0,029	5,80E-11	1,10E-08
GO:0060255	P	regulation of macromolecule metabolic process	187	1147	0,163	2829	28397	0,100	5,60E-11	1,10E-08
GO:0009892	P	negative regulation of metabolic process	67	1147	0,058	659	28397	0,023	6,30E-11	1,20E-08
GO:0006950	P	response to stress	247	1147	0,215	4089	28397	0,144	1,30E-10	2,30E-08
GO:0032870	P	cellular response to hormone stimulus	63	1147	0,055	641	28397	0,023	8,20E-10	1,40E-07
GO:0010608	P	posttranscriptional regulation of gene expression	43	1147	0,037	347	28397	0,012	9,20E-10	1,50E-07
GO:0019222	P	regulation of metabolic process	199	1147	0,173	3186	28397	0,112	1,10E-09	1,70E-07
GO:0009755	P	hormone-mediated signaling pathway	60	1147	0,052	600	28397	0,021	1,10E-09	1,80E-07
GO:0009415	P	response to water	47	1147	0,041	424	28397	0,015	4,10E-09	6,10E-07
GO:0009414	P	response to water deprivation	46	1147	0,040	416	28397	0,015	6,40E-09	9,20E-07
GO:0065007	P	biological regulation	336	1147	0,293	6222	28397	0,219	6,60E-09	9,20E-07
GO:0034285	P	response to disaccharide stimulus	31	1147	0,027	213	28397	0,008	7,50E-09	1,00E-06
GO:0048519	P	negative regulation of biological process	95	1147	0,083	1243	28397	0,044	1,10E-08	1,40E-06
GO:0050789	P	regulation of biological process	290	1147	0,253	5235	28397	0,184	1,20E-08	1,50E-06
GO:0009744	P	response to sucrose stimulus	30	1147	0,026	210	28397	0,007	1,90E-08	2,40E-06
GO:0009737	P	response to abscisic acid stimulus	56	1147	0,049	621	28397	0,022	1,10E-07	1,30E-05
GO:0031047	P	gene silencing by RNA	40	1147	0,035	371	28397	0,013	1,10E-07	1,40E-05
GO:0009750	P	response to fructose stimulus	23	1147	0,020	144	28397	0,005	1,40E-07	1,70E-05
GO:0007568	P	aging	23	1147	0,020	145	28397	0,005	1,60E-07	1,80E-05
GO:0010150	P	leaf senescence	16	1147	0,014	70	28397	0,002	1,80E-07	2,00E-05
GO:0042538	P	hyperosmotic salinity response	24	1147	0,021	162	28397	0,006	2,70E-07	2,90E-05
GO:0007242	P	intracellular signaling cascade	90	1147	0,078	1252	28397	0,044	3,40E-07	3,60E-05
GO:0006355	P	regulation of transcription, DNA-dependent	147	1147	0,128	2372	28397	0,084	3,70E-07	3,90E-05
GO:0045449	P	regulation of transcription	147	1147	0,128	2376	28397	0,084	4,10E-07	4,20E-05
GO:0051252	P	regulation of RNA metabolic process	147	1147	0,128	2388	28397	0,084	5,40E-07	5,40E-05
GO:0034284	P	response to monosaccharide stimulus	24	1147	0,021	170	28397	0,006	5,90E-07	5,70E-05

GO:0009746	P	response to hexose stimulus	24	1147	0,021	170	28397	0,006	5,90E-07	5,70E-05
GO:0009611	P	response to wounding	36	1147	0,031	340	28397	0,012	7,10E-07	6,70E-05
GO:0070887	P	cellular response to chemical stimulus	97	1147	0,085	1417	28397	0,050	9,20E-07	8,50E-05
GO:0009723	P	response to ethylene stimulus	36	1147	0,031	353	28397	0,012	1,60E-06	0,00014
GO:0010260	P	organ senescence	17	1147	0,015	96	28397	0,003	1,70E-06	0,00015
GO:0010149	P	senescence	17	1147	0,015	96	28397	0,003	1,70E-06	0,00015
GO:0031326	P	regulation of cellular biosynthetic process	156	1147	0,136	2631	28397	0,093	1,90E-06	0,00016
GO:0009889	P	regulation of biosynthetic process	156	1147	0,136	2634	28397	0,093	2,00E-06	0,00017
GO:0071310	P	cellular response to organic substance	86	1147	0,075	1234	28397	0,043	2,00E-06	0,00017
GO:0010556	P	regulation of macromolecule biosynthetic process	149	1147	0,130	2491	28397	0,088	2,10E-06	0,00017
GO:0019219	P	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	148	1147	0,129	2496	28397	0,088	3,50E-06	0,00028
GO:0000160	P	two-component signal transduction system (phosphorelay)	22	1147	0,019	165	28397	0,006	4,00E-06	0,00031
GO:0016458	P	gene silencing	41	1147	0,036	450	28397	0,016	4,20E-06	0,00033
GO:0006972	P	hyperosmotic response	28	1147	0,024	251	28397	0,009	4,90E-06	0,00037
GO:0010200	P	response to chitin	39	1147	0,034	421	28397	0,015	5,00E-06	0,00038
GO:0009753	P	response to jasmonic acid stimulus	42	1147	0,037	471	28397	0,017	5,30E-06	0,00039
GO:0051171	P	regulation of nitrogen compound metabolic process	148	1147	0,129	2517	28397	0,089	5,40E-06	0,0004
GO:0009692	P	ethylene metabolic process	19	1147	0,017	131	28397	0,005	6,00E-06	0,00043
GO:0009693	P	ethylene biosynthetic process	19	1147	0,017	131	28397	0,005	6,00E-06	0,00043
GO:0009828	P	plant-type cell wall loosening	10	1147	0,009	35	28397	0,001	6,90E-06	0,00049
GO:0043449	P	cellular alkene metabolic process	19	1147	0,017	133	28397	0,005	7,40E-06	0,0005
GO:0043450	P	alkene biosynthetic process	19	1147	0,017	133	28397	0,005	7,40E-06	0,0005
GO:0009646	P	response to absence of light	10	1147	0,009	37	28397	0,001	1,00E-05	0,0007
GO:0009628	P	response to abiotic stimulus	152	1147	0,133	2635	28397	0,093	1,00E-05	0,0007
GO:0006351	P	transcription, DNA-dependent	151	1147	0,132	2618	28397	0,092	1,10E-05	0,00074
GO:0006350	P	transcription	151	1147	0,132	2620	28397	0,092	1,20E-05	0,00076
GO:0032774	P	RNA biosynthetic process	151	1147	0,132	2621	28397	0,092	1,20E-05	0,00077
GO:0009738	P	abscisic acid mediated signaling pathway	27	1147	0,024	252	28397	0,009	1,40E-05	0,00087
GO:0071215	P	cellular response to abscisic acid stimulus	28	1147	0,024	267	28397	0,009	1,40E-05	0,00087
GO:0031323	P	regulation of cellular metabolic process	165	1147	0,144	2928	28397	0,103	1,40E-05	0,00088
GO:0080090	P	regulation of primary metabolic process	157	1147	0,137	2761	28397	0,097	1,50E-05	0,00093
GO:0050794	P	regulation of cellular process	241	1147	0,210	4595	28397	0,162	1,60E-05	0,00094
GO:0009718	P	anthocyanin biosynthetic process	12	1147	0,010	63	28397	0,002	3,00E-05	0,0018
GO:0033554	P	cellular response to stress	93	1147	0,081	1473	28397	0,052	3,10E-05	0,0018
GO:0046283	P	anthocyanin metabolic process	14	1147	0,012	87	28397	0,003	3,60E-05	0,0021
GO:0009733	P	response to auxin stimulus	37	1147	0,032	431	28397	0,015	4,00E-05	0,0023
GO:0040029	P	regulation of gene expression, epigenetic	41	1147	0,036	508	28397	0,018	5,70E-05	0,0032
GO:0071369	P	cellular response to ethylene stimulus	18	1147	0,016	144	28397	0,005	6,20E-05	0,0035
GO:0007165	P	signal transduction	101	1147	0,088	1670	28397	0,059	7,00E-05	0,0039
GO:0006970	P	response to osmotic stress	59	1147	0,051	842	28397	0,030	7,10E-05	0,0039
GO:0009873	P	ethylene mediated signaling pathway	16	1147	0,014	125	28397	0,004	0,00012	0,0066
GO:0019375	P	galactolipid biosynthetic process	14	1147	0,012	99	28397	0,003	0,00012	0,0067
GO:0019374	P	galactolipid metabolic process	14	1147	0,012	100	28397	0,004	0,00014	0,0073
GO:0042398	P	cellular amino acid derivative biosynthetic process	41	1147	0,036	534	28397	0,019	0,00015	0,0081
GO:0007623	P	circadian rhythm	19	1147	0,017	171	28397	0,006	0,00016	0,0083
GO:0048511	P	rhythmic process	19	1147	0,017	171	28397	0,006	0,00016	0,0083
GO:0009838	P	abscission	10	1147	0,009	54	28397	0,002	0,00017	0,0085
GO:0010053	P	root epidermal cell differentiation	30	1147	0,026	350	28397	0,012	0,00021	0,011

GO:0009831	P	plant-type cell wall modification during multidimensional cell growth	7	1147	0,006	26	28397	0,001	0,00023	0,011
GO:0009651	P	response to salt stress	54	1147	0,047	788	28397	0,028	0,00024	0,012
GO:0006575	P	cellular amino acid derivative metabolic process	50	1147	0,044	714	28397	0,025	0,00025	0,012
GO:0009247	P	glycolipid biosynthetic process	14	1147	0,012	107	28397	0,004	0,00026	0,012
GO:0010054	P	trichoblast differentiation	29	1147	0,025	339	28397	0,012	0,00027	0,013
GO:0009739	P	response to gibberellin stimulus	18	1147	0,016	166	28397	0,006	0,00031	0,015
GO:0023060	P	signal transmission	102	1147	0,089	1767	28397	0,062	0,00032	0,015
GO:0023046	P	signaling process	102	1147	0,089	1768	28397	0,062	0,00033	0,015
GO:0042547	P	cell wall modification during multidimensional cell growth	7	1147	0,006	28	28397	0,001	0,00034	0,016
GO:0006979	P	response to oxidative stress	42	1147	0,037	582	28397	0,020	0,00043	0,02
GO:0046688	P	response to copper ion	6	1147	0,005	21	28397	0,001	0,0005	0,022
GO:0006664	P	glycolipid metabolic process	14	1147	0,012	116	28397	0,004	0,00054	0,024
GO:0055072	P	iron ion homeostasis	7	1147	0,006	33	28397	0,001	0,00079	0,035
GO:0030528	F	transcription regulator activity	113	1147	0,099	1740	28397	0,061	1,20E-06	0,0011
GO:0003700	F	transcription factor activity	105	1147	0,092	1682	28397	0,059	1,50E-05	0,0067
GO:0005576	C	extracellular region	158	1147	0,138	2824	28397	0,099	3,20E-05	0,01

S10: GO analysis for genes downregulated in *se-1*.

			Query		Observ.	Background		Expect.		
GO_acc	type	Term	item	total	Freq.	item	total	Freq.	pvalue	FDR
GO:0055114	P	oxidation reduction	144	926	0,156	1364	28397	0,048	3,20E-33	1,50E-29
GO:0046173	P	polyol biosynthetic process	38	926	0,041	81	28397	0,003	7,30E-28	1,70E-24
GO:0015979	P	photosynthesis	72	926	0,078	435	28397	0,015	7,70E-27	1,00E-23
GO:0009657	P	plastid organization	69	926	0,075	398	28397	0,014	8,40E-27	1,00E-23
GO:0010264	P	myo-inositol hexakisphosphate biosynthetic process	34	926	0,037	65	28397	0,002	2,90E-26	2,30E-23
GO:0033517	P	myo-inositol hexakisphosphate metabolic process	34	926	0,037	65	28397	0,002	2,90E-26	2,30E-23
GO:0032958	P	inositol phosphate biosynthetic process	34	926	0,037	66	28397	0,002	4,20E-26	2,80E-23
GO:0006021	P	inositol biosynthetic process	34	926	0,037	67	28397	0,002	6,20E-26	3,60E-23
GO:0019684	P	photosynthesis, light reaction	59	926	0,064	333	28397	0,012	1,70E-23	9,20E-21
GO:0043647	P	inositol phosphate metabolic process	34	926	0,037	86	28397	0,003	3,70E-23	1,80E-20
GO:0019751	P	polyol metabolic process	39	926	0,042	136	28397	0,005	3,50E-22	1,50E-19
GO:0006066	P	alcohol metabolic process	109	926	0,118	1143	28397	0,040	6,60E-22	2,40E-19
GO:0042440	P	pigment metabolic process	59	926	0,064	361	28397	0,013	6,10E-22	2,40E-19
GO:0006020	P	inositol metabolic process	34	926	0,037	98	28397	0,003	1,10E-21	3,80E-19
GO:0046148	P	pigment biosynthetic process	51	926	0,055	282	28397	0,010	8,40E-21	2,60E-18
GO:0051186	P	cofactor metabolic process	79	926	0,085	753	28397	0,027	3,10E-18	9,20E-16
GO:0010114	P	response to red light	30	926	0,032	104	28397	0,004	1,90E-17	5,40E-15
GO:0009637	P	response to blue light	32	926	0,035	126	28397	0,004	3,80E-17	1,00E-14
GO:0010218	P	response to far red light	29	926	0,031	99	28397	0,003	4,70E-17	1,20E-14
GO:0009820	P	alkaloid metabolic process	40	926	0,043	225	28397	0,008	2,30E-16	5,40E-14
GO:0019748	P	secondary metabolic process	102	926	0,110	1247	28397	0,044	2,80E-16	6,30E-14
GO:0006733	P	oxidoreduction coenzyme metabolic process	43	926	0,046	267	28397	0,009	4,00E-16	8,60E-14
GO:0006098	P	pentose-phosphate shunt	37	926	0,040	200	28397	0,007	1,00E-15	1,90E-13
GO:0046496	P	nicotinamide nucleotide metabolic process	38	926	0,041	212	28397	0,007	1,00E-15	1,90E-13
GO:0006769	P	nicotinamide metabolic process	38	926	0,041	212	28397	0,007	1,00E-15	1,90E-13
GO:0006740	P	NADPH regeneration	37	926	0,040	201	28397	0,007	1,20E-15	2,10E-13
GO:0019362	P	pyridine nucleotide metabolic process	38	926	0,041	214	28397	0,008	1,30E-15	2,30E-13
GO:0006739	P	NADP metabolic process	37	926	0,040	204	28397	0,007	1,70E-15	2,90E-13
GO:0043603	P	cellular amide metabolic process	38	926	0,041	222	28397	0,008	3,70E-15	6,10E-13
GO:0008299	P	isoprenoid biosynthetic process	56	926	0,060	477	28397	0,017	4,20E-15	6,60E-13
GO:0044262	P	cellular carbohydrate metabolic process	125	926	0,135	1778	28397	0,063	4,40E-15	6,70E-13
GO:0009658	P	chloroplast organization	39	926	0,042	237	28397	0,008	5,10E-15	7,50E-13
GO:0010207	P	photosystem II assembly	34	926	0,037	177	28397	0,006	5,60E-15	8,00E-13
GO:0015994	P	chlorophyll metabolic process	35	926	0,038	189	28397	0,007	5,80E-15	8,10E-13
GO:0006720	P	isoprenoid metabolic process	57	926	0,062	500	28397	0,018	7,50E-15	1,00E-12
GO:0019288	P	isopentenyl diphosphate biosynthetic process, mevalonate-independent pathway	38	926	0,041	229	28397	0,008	9,00E-15	1,20E-12
GO:0019682	P	glyceraldehyde-3-phosphate metabolic process	38	926	0,041	232	28397	0,008	1,30E-14	1,70E-12
GO:0009240	P	isopentenyl diphosphate biosynthetic process	38	926	0,041	233	28397	0,008	1,50E-14	1,80E-12
GO:0046490	P	isopentenyl diphosphate metabolic process	38	926	0,041	233	28397	0,008	1,50E-14	1,80E-12
GO:0016108	P	tetraterpenoid metabolic process	27	926	0,029	111	28397	0,004	2,70E-14	3,10E-12
GO:0016116	P	carotenoid metabolic process	27	926	0,029	111	28397	0,004	2,70E-14	3,10E-12
GO:0005975	P	carbohydrate metabolic process	144	926	0,156	2249	28397	0,079	3,30E-14	3,80E-12
GO:0006364	P	rRNA processing	44	926	0,048	324	28397	0,011	4,30E-14	4,70E-12
GO:0016072	P	rRNA metabolic process	44	926	0,048	325	28397	0,011	4,70E-14	5,00E-12
GO:0016109	P	tetraterpenoid biosynthetic process	26	926	0,028	106	28397	0,004	6,90E-14	7,10E-12
GO:0016117	P	carotenoid biosynthetic process	26	926	0,028	106	28397	0,004	6,90E-14	7,10E-12

GO:0006091	P	generation of precursor metabolites and energy	68	926	0,073	730	28397	0,026	1,50E-13	1,50E-11
GO:0008152	P	metabolic process	557	926	0,602	13633	28397	0,480	2,00E-13	1,90E-11
GO:0010027	P	thylakoid membrane organization	33	926	0,036	198	28397	0,007	4,70E-13	4,40E-11
GO:0009668	P	plastid membrane organization	33	926	0,036	198	28397	0,007	4,70E-13	4,40E-11
GO:0006778	P	porphyrin metabolic process	35	926	0,038	226	28397	0,008	6,10E-13	5,70E-11
GO:0033013	P	tetrapyrrole metabolic process	35	926	0,038	227	28397	0,008	6,90E-13	6,30E-11
GO:0006081	P	cellular aldehyde metabolic process	42	926	0,045	334	28397	0,012	1,60E-12	1,40E-10
GO:0046165	P	alcohol biosynthetic process	39	926	0,042	293	28397	0,010	2,10E-12	1,80E-10
GO:0044283	P	small molecule biosynthetic process	120	926	0,130	1865	28397	0,066	4,80E-12	4,10E-10
GO:0034470	P	ncRNA processing	44	926	0,048	387	28397	0,014	1,00E-11	8,50E-10
GO:0015995	P	chlorophyll biosynthetic process	25	926	0,027	125	28397	0,004	1,00E-11	8,50E-10
GO:0009416	P	response to light stimulus	87	926	0,094	1188	28397	0,042	1,40E-11	1,10E-09
GO:0034660	P	ncRNA metabolic process	48	926	0,052	481	28397	0,017	7,40E-11	6,00E-09
GO:0042254	P	ribosome biogenesis	45	926	0,049	437	28397	0,015	1,20E-10	9,10E-09
GO:0006732	P	coenzyme metabolic process	49	926	0,053	508	28397	0,018	1,40E-10	1,10E-08
GO:0034637	P	cellular carbohydrate biosynthetic process	72	926	0,078	941	28397	0,033	1,50E-10	1,10E-08
GO:0009773	P	photosynthetic electron transport in photosystem I	16	926	0,017	51	28397	0,002	2,10E-10	1,60E-08
GO:0022613	P	ribonucleoprotein complex biogenesis	45	926	0,049	447	28397	0,016	2,30E-10	1,70E-08
GO:0009314	P	response to radiation	87	926	0,094	1263	28397	0,044	2,50E-10	1,80E-08
GO:0009767	P	photosynthetic electron transport chain	19	926	0,021	80	28397	0,003	2,50E-10	1,80E-08
GO:0034641	P	cellular nitrogen compound metabolic process	92	926	0,099	1373	28397	0,048	2,80E-10	2,00E-08
GO:0005996	P	monosaccharide metabolic process	64	926	0,069	804	28397	0,028	3,70E-10	2,60E-08
GO:0051667	P	establishment of plastid localization	21	926	0,023	104	28397	0,004	3,90E-10	2,60E-08
GO:0051188	P	cofactor biosynthetic process	41	926	0,044	388	28397	0,014	3,90E-10	2,60E-08
GO:0051644	P	plastid localization	21	926	0,023	104	28397	0,004	3,90E-10	2,60E-08
GO:0009902	P	chloroplast relocation	21	926	0,023	104	28397	0,004	3,90E-10	2,60E-08
GO:0051656	P	establishment of organelle localization	21	926	0,023	108	28397	0,004	7,20E-10	4,70E-08
GO:0006779	P	porphyrin biosynthetic process	25	926	0,027	157	28397	0,006	7,30E-10	4,70E-08
GO:0033014	P	tetrapyrrole biosynthetic process	25	926	0,027	160	28397	0,006	1,00E-09	6,50E-08
GO:0044281	P	small molecule metabolic process	202	926	0,218	4065	28397	0,143	1,10E-09	6,70E-08
GO:0009639	P	response to red or far red light	41	926	0,044	406	28397	0,014	1,30E-09	8,10E-08
GO:0016051	P	carbohydrate biosynthetic process	75	926	0,081	1070	28397	0,038	2,30E-09	1,40E-07
GO:0006644	P	phospholipid metabolic process	44	926	0,048	472	28397	0,017	3,20E-09	1,90E-07
GO:0008654	P	phospholipid biosynthetic process	40	926	0,043	405	28397	0,014	3,70E-09	2,20E-07
GO:0006629	P	lipid metabolic process	106	926	0,114	1772	28397	0,062	4,40E-09	2,50E-07
GO:0019637	P	organophosphate metabolic process	44	926	0,048	483	28397	0,017	6,00E-09	3,50E-07
GO:0022900	P	electron transport chain	21	926	0,023	133	28397	0,005	1,90E-08	1,10E-06
GO:0006090	P	pyruvate metabolic process	38	926	0,041	398	28397	0,014	2,10E-08	1,20E-06
GO:0008610	P	lipid biosynthetic process	76	926	0,082	1159	28397	0,041	2,60E-08	1,40E-06
GO:0016114	P	terpenoid biosynthetic process	30	926	0,032	268	28397	0,009	2,70E-08	1,50E-06
GO:0006721	P	terpenoid metabolic process	31	926	0,033	290	28397	0,010	4,20E-08	2,30E-06
GO:0006007	P	glucose catabolic process	41	926	0,044	474	28397	0,017	7,30E-08	3,90E-06
GO:0019320	P	hexose catabolic process	41	926	0,044	476	28397	0,017	8,20E-08	4,30E-06
GO:0046365	P	monosaccharide catabolic process	41	926	0,044	480	28397	0,017	1,00E-07	5,30E-06
GO:0018130	P	heterocycle biosynthetic process	28	926	0,030	258	28397	0,009	1,40E-07	7,30E-06
GO:0009117	P	nucleotide metabolic process	51	926	0,055	685	28397	0,024	1,70E-07	8,60E-06
GO:0046164	P	alcohol catabolic process	41	926	0,044	491	28397	0,017	1,80E-07	8,90E-06
GO:0051640	P	organelle localization	21	926	0,023	154	28397	0,005	1,80E-07	8,90E-06
GO:0006753	P	nucleoside phosphate metabolic process	51	926	0,055	687	28397	0,024	1,80E-07	9,00E-06
GO:0016556	P	mRNA modification	17	926	0,018	101	28397	0,004	1,90E-07	9,30E-06
GO:0044255	P	cellular lipid metabolic process	83	926	0,090	1395	28397	0,049	2,90E-07	1,40E-05

GO:0055086	P	nucleobase, nucleoside and nucleotide metabolic process	52	926	0,056	726	28397	0,026	3,80E-07	1,80E-05
GO:0050896	P	response to stimulus	269	926	0,290	6292	28397	0,222	9,10E-07	4,30E-05
GO:0044275	P	cellular carbohydrate catabolic process	42	926	0,045	545	28397	0,019	9,00E-07	4,30E-05
GO:0010155	P	regulation of proton transport	14	926	0,015	77	28397	0,003	9,90E-07	4,70E-05
GO:0016052	P	carbohydrate catabolic process	43	926	0,046	573	28397	0,020	1,30E-06	5,90E-05
GO:0046483	P	heterocycle metabolic process	63	926	0,068	1023	28397	0,036	3,00E-06	0,00014
GO:0006006	P	glucose metabolic process	41	926	0,044	554	28397	0,020	3,10E-06	0,00014
GO:0006073	P	cellular glucan metabolic process	32	926	0,035	388	28397	0,014	4,90E-06	0,00022
GO:0019344	P	cysteine biosynthetic process	22	926	0,024	210	28397	0,007	5,10E-06	0,00023
GO:0010016	P	shoot morphogenesis	39	926	0,042	529	28397	0,019	5,80E-06	0,00026
GO:0006534	P	cysteine metabolic process	22	926	0,024	213	28397	0,008	6,20E-06	0,00027
GO:0035304	P	regulation of protein amino acid dephosphorylation	17	926	0,018	135	28397	0,005	6,80E-06	0,00029
GO:0035303	P	regulation of dephosphorylation	17	926	0,018	137	28397	0,005	8,10E-06	0,00035
GO:0009069	P	serine family amino acid metabolic process	25	926	0,027	271	28397	0,010	9,20E-06	0,00039
GO:0043623	P	cellular protein complex assembly	37	926	0,040	504	28397	0,018	1,10E-05	0,00046
GO:0009070	P	serine family amino acid biosynthetic process	22	926	0,024	222	28397	0,008	1,10E-05	0,00047
GO:0044271	P	cellular nitrogen compound biosynthetic process	51	926	0,055	804	28397	0,028	1,20E-05	0,00052
GO:0044042	P	glucan metabolic process	33	926	0,036	432	28397	0,015	1,50E-05	0,00063
GO:0009628	P	response to abiotic stimulus	126	926	0,136	2635	28397	0,093	1,60E-05	0,00065
GO:0006461	P	protein complex assembly	40	926	0,043	577	28397	0,020	1,70E-05	0,00068
GO:0070271	P	protein complex biogenesis	40	926	0,043	577	28397	0,020	1,70E-05	0,00068
GO:0005982	P	starch metabolic process	22	926	0,024	230	28397	0,008	1,90E-05	0,00074
GO:0019318	P	hexose metabolic process	41	926	0,044	602	28397	0,021	2,00E-05	0,00077
GO:0010103	P	stomatal complex morphogenesis	17	926	0,018	149	28397	0,005	2,20E-05	0,00085
GO:0010224	P	response to UV-B	14	926	0,015	104	28397	0,004	2,20E-05	0,00086
GO:0061024	P	membrane organization	38	926	0,041	543	28397	0,019	2,30E-05	0,00086
GO:0016044	P	cellular membrane organization	38	926	0,041	543	28397	0,019	2,30E-05	0,00086
GO:0031399	P	regulation of protein modification process	17	926	0,018	150	28397	0,005	2,40E-05	0,00089
GO:0009250	P	glucan biosynthetic process	26	926	0,028	313	28397	0,011	3,30E-05	0,0012
GO:0019252	P	starch biosynthetic process	19	926	0,021	191	28397	0,007	4,20E-05	0,0016
GO:0006636	P	unsaturated fatty acid biosynthetic process	11	926	0,012	70	28397	0,002	4,70E-05	0,0018
GO:0033559	P	unsaturated fatty acid metabolic process	11	926	0,012	71	28397	0,003	5,30E-05	0,0019
GO:0022621	P	shoot system development	45	926	0,049	723	28397	0,025	6,10E-05	0,0022
GO:0048367	P	shoot development	45	926	0,049	723	28397	0,025	6,10E-05	0,0022
GO:0034622	P	cellular macromolecular complex assembly	39	926	0,042	592	28397	0,021	6,10E-05	0,0022
GO:0042180	P	cellular ketone metabolic process	103	926	0,111	2123	28397	0,075	6,10E-05	0,0022
GO:0032268	P	regulation of cellular protein metabolic process	20	926	0,022	215	28397	0,008	6,30E-05	0,0022
GO:0065003	P	macromolecular complex assembly	42	926	0,045	665	28397	0,023	7,90E-05	0,0028
GO:0051246	P	regulation of protein metabolic process	21	926	0,023	237	28397	0,008	8,00E-05	0,0028
GO:0006725	P	cellular aromatic compound metabolic process	57	926	0,062	1022	28397	0,036	0,00012	0,004
GO:0032787	P	monocarboxylic acid metabolic process	76	926	0,082	1481	28397	0,052	0,00012	0,004
GO:0044272	P	sulfur compound biosynthetic process	33	926	0,036	486	28397	0,017	0,00013	0,0044
GO:0009744	P	response to sucrose stimulus	19	926	0,021	210	28397	0,007	0,00013	0,0045
GO:0006790	P	sulfur metabolic process	42	926	0,045	683	28397	0,024	0,00014	0,0046
GO:0034621	P	cellular macromolecular complex subunit organization	39	926	0,042	617	28397	0,022	0,00014	0,0046
GO:0034285	P	response to disaccharide stimulus	19	926	0,021	213	28397	0,008	0,00016	0,0052
GO:0043436	P	oxoacid metabolic process	100	926	0,108	2103	28397	0,074	0,00016	0,0052
GO:0019752	P	carboxylic acid metabolic process	100	926	0,108	2103	28397	0,074	0,00016	0,0052

GO:0006082	P	organic acid metabolic process	100	926	0,108	2105	28397	0,074	0,00016	0,0053
GO:0043933	P	macromolecular complex subunit organization	42	926	0,045	691	28397	0,024	0,00017	0,0055
GO:0010374	P	stomatal complex development	19	926	0,021	217	28397	0,008	0,0002	0,0063
GO:0043269	P	regulation of ion transport	15	926	0,016	147	28397	0,005	0,0002	0,0064
GO:0042793	P	transcription from plastid promoter	10	926	0,011	72	28397	0,003	0,00026	0,0081
GO:0009411	P	response to UV	20	926	0,022	247	28397	0,009	0,00035	0,011
GO:0044237	P	cellular metabolic process	427	926	0,461	11509	28397	0,405	0,0004	0,012
GO:0071482	P	cellular response to light stimulus	10	926	0,011	77	28397	0,003	0,00042	0,013
GO:0071478	P	cellular response to radiation	10	926	0,011	77	28397	0,003	0,00042	0,013
GO:0006519	P	cellular amino acid and derivative metabolic process	67	926	0,072	1324	28397	0,047	0,00043	0,013
GO:0006952	P	defense response	80	926	0,086	1653	28397	0,058	0,00044	0,013
GO:0006811	P	ion transport	54	926	0,058	1019	28397	0,036	0,00056	0,017
GO:0015992	P	proton transport	14	926	0,015	147	28397	0,005	0,00061	0,018
GO:0006818	P	hydrogen transport	14	926	0,015	147	28397	0,005	0,00061	0,018
GO:0006950	P	response to stress	170	926	0,184	4089	28397	0,144	0,00063	0,019
GO:0044282	P	small molecule catabolic process	50	926	0,054	932	28397	0,033	0,00069	0,02
GO:0006633	P	fatty acid biosynthetic process	22	926	0,024	303	28397	0,011	0,00073	0,021
GO:0009644	P	response to high light intensity	18	926	0,019	224	28397	0,008	0,00073	0,021
GO:0032879	P	regulation of localization	17	926	0,018	207	28397	0,007	0,00082	0,024
GO:0019438	P	aromatic compound biosynthetic process	39	926	0,042	680	28397	0,024	0,00083	0,024
GO:0009699	P	phenylpropanoid biosynthetic process	23	926	0,025	327	28397	0,012	0,00083	0,024
GO:0006396	P	RNA processing	51	926	0,055	967	28397	0,034	0,00087	0,025
GO:0051049	P	regulation of transport	15	926	0,016	172	28397	0,006	0,00092	0,026
GO:0046394	P	carboxylic acid biosynthetic process	57	926	0,062	1116	28397	0,039	0,00092	0,026
GO:0000097	P	sulfur amino acid biosynthetic process	22	926	0,024	309	28397	0,011	0,00092	0,026
GO:0016053	P	organic acid biosynthetic process	57	926	0,062	1116	28397	0,039	0,00092	0,026
GO:0045036	P	protein targeting to chloroplast	9	926	0,010	71	28397	0,003	0,00094	0,026
GO:0071704	P	organic substance metabolic process	5	926	0,005	20	28397	0,001	0,00098	0,027
GO:0015977	P	carbon fixation	5	926	0,005	20	28397	0,001	0,00098	0,027
GO:0022607	P	cellular component assembly	45	926	0,049	832	28397	0,029	0,0011	0,029
GO:0015996	P	chlorophyll catabolic process	8	926	0,009	58	28397	0,002	0,0011	0,029
GO:0046149	P	pigment catabolic process	8	926	0,009	58	28397	0,002	0,0011	0,029
GO:0006470	P	protein amino acid dephosphorylation	17	926	0,018	214	28397	0,008	0,0011	0,03
GO:0055072	P	iron ion homeostasis	6	926	0,006	33	28397	0,001	0,0013	0,035
GO:0009698	P	phenylpropanoid metabolic process	26	926	0,028	405	28397	0,014	0,0014	0,036
GO:0051187	P	cofactor catabolic process	10	926	0,011	92	28397	0,003	0,0015	0,038
GO:0042221	P	response to chemical stimulus	163	926	0,176	3978	28397	0,140	0,0015	0,04
GO:0006787	P	porphyrin catabolic process	8	926	0,009	62	28397	0,002	0,0016	0,041
GO:0033015	P	tetrapyrrole catabolic process	8	926	0,009	62	28397	0,002	0,0016	0,041
GO:0043085	P	positive regulation of catalytic activity	12	926	0,013	129	28397	0,005	0,0018	0,045
GO:0048513	P	organ development	93	926	0,100	2083	28397	0,073	0,0018	0,046
GO:0048731	P	system development	93	926	0,100	2083	28397	0,073	0,0018	0,046
GO:0044093	P	positive regulation of molecular function	12	926	0,013	130	28397	0,005	0,0019	0,047
GO:0009813	P	flavonoid biosynthetic process	17	926	0,018	225	28397	0,008	0,0019	0,047
GO:0016491	F	oxidoreductase activity	144	926	0,156	1691	28397	0,060	2,00E-24	1,60E-21
GO:0047134	F	protein-disulfide reductase activity	35	926	0,038	146	28397	0,005	6,00E-18	1,60E-15
GO:0016668	F	oxidoreductase activity, acting on sulfur group of donors, NAD or NADP as acceptor	36	926	0,039	156	28397	0,005	5,80E-18	1,60E-15
GO:0003824	F	catalytic activity	410	926	0,443	8787	28397	0,309	3,90E-17	7,60E-15
GO:0016651	F	oxidoreductase activity, acting on NADH or NADPH	39	926	0,042	248	28397	0,009	1,90E-14	3,00E-12
GO:0016667	F	oxidoreductase activity, acting on sulfur group of donors	38	926	0,041	283	28397	0,010	3,20E-12	4,10E-10

GO:0016705	F	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	47	926	0,051	462	28397	0,016	6,50E-11	7,30E-09
GO:0046906	F	tetrapyrrole binding	41	926	0,044	407	28397	0,014	1,40E-09	1,40E-07
GO:0005506	F	iron ion binding	45	926	0,049	499	28397	0,018	5,30E-09	4,70E-07
GO:0020037	F	heme binding	37	926	0,040	373	28397	0,013	1,30E-08	1,00E-06
GO:0019825	F	oxygen binding	24	926	0,026	234	28397	0,008	2,70E-06	0,00019
GO:0004497	F	monooxygenase activity	16	926	0,017	112	28397	0,004	3,00E-06	0,00019
GO:0046914	F	transition metal ion binding	115	926	0,124	2313	28397	0,081	7,70E-06	0,00047
GO:0009055	F	electron carrier activity	38	926	0,041	525	28397	0,018	1,10E-05	0,00063
GO:0043169	F	cation binding	139	926	0,150	3029	28397	0,107	3,80E-05	0,0019
GO:0043167	F	ion binding	139	926	0,150	3030	28397	0,107	3,90E-05	0,0019
GO:0016740	F	transferase activity	137	926	0,148	3012	28397	0,106	6,50E-05	0,0028
GO:0046872	F	metal ion binding	131	926	0,141	2853	28397	0,100	6,40E-05	0,0028
GO:0050660	F	FAD binding	17	926	0,018	166	28397	0,006	7,50E-05	0,0031
GO:0008194	F	UDP-glycosyltransferase activity	21	926	0,023	239	28397	0,008	8,90E-05	0,0035
GO:0048037	F	cofactor binding	33	926	0,036	479	28397	0,017	0,0001	0,0038
GO:0050661	F	NADP or NADPH binding	10	926	0,011	66	28397	0,002	0,00014	0,0049
GO:0046527	F	glucosyltransferase activity	15	926	0,016	142	28397	0,005	0,00014	0,0049
GO:0050662	F	coenzyme binding	26	926	0,028	354	28397	0,012	2,10E-04	0,007
GO:0016758	F	transferase activity, transferring hexosyl groups	28	926	0,030	396	28397	0,014	2,20E-04	0,007
GO:0043531	F	ADP binding	16	926	0,017	167	28397	0,006	0,00024	0,0074
GO:0001883	F	purine nucleoside binding	107	926	0,116	2333	28397	0,082	0,00033	0,0094
GO:0030554	F	adenyl nucleotide binding	107	926	0,116	2331	28397	0,082	0,00032	0,0094
GO:0001882	F	nucleoside binding	107	926	0,116	2341	28397	0,082	0,00038	0,01
GO:0016209	F	antioxidant activity	14	926	0,015	142	28397	0,005	0,00045	0,012
GO:0005488	F	binding	415	926	0,448	11247	28397	0,396	0,00085	0,022
GO:0017076	F	purine nucleotide binding	113	926	0,122	2595	28397	0,091	0,0013	0,033
GO:0035251	F	UDP-glucosyltransferase activity	11	926	0,012	114	28397	0,004	0,0021	0,049
GO:0044435	C	plastid part	130	926	0,140	1252	28397	0,044	2,70E-29	6,00E-27
GO:0044434	C	chloroplast part	128	926	0,138	1211	28397	0,043	1,60E-29	6,00E-27
GO:0009579	C	thylakoid	69	926	0,075	518	28397	0,018	5,30E-21	7,90E-19
GO:0009570	C	chloroplast stroma	71	926	0,077	603	28397	0,021	6,60E-19	7,30E-17
GO:0009532	C	plastid stroma	73	926	0,079	637	28397	0,022	8,10E-19	7,30E-17
GO:0009507	C	chloroplast	232	926	0,251	3959	28397	0,139	1,10E-18	8,20E-17
GO:0009536	C	plastid	235	926	0,254	4037	28397	0,142	1,30E-18	8,30E-17
GO:0031976	C	plastid thylakoid	56	926	0,060	425	28397	0,015	4,50E-17	2,30E-15
GO:0009534	C	chloroplast thylakoid	56	926	0,060	425	28397	0,015	4,50E-17	2,30E-15
GO:0031984	C	organelle subcompartment	56	926	0,060	428	28397	0,015	6,00E-17	2,70E-15
GO:0009526	C	plastid envelope	66	926	0,071	598	28397	0,021	2,20E-16	9,10E-15
GO:0009941	C	chloroplast envelope	64	926	0,069	573	28397	0,020	3,80E-16	1,40E-14
GO:0044436	C	thylakoid part	51	926	0,055	388	28397	0,014	1,30E-15	4,40E-14
GO:0034357	C	photosynthetic membrane	47	926	0,051	355	28397	0,013	1,30E-14	4,20E-13
GO:0042651	C	thylakoid membrane	46	926	0,050	341	28397	0,012	1,40E-14	4,20E-13
GO:0009535	C	chloroplast thylakoid membrane	44	926	0,048	322	28397	0,011	3,50E-14	9,90E-13
GO:0055035	C	plastid thylakoid membrane	44	926	0,048	324	28397	0,011	4,30E-14	1,10E-12
GO:0031975	C	envelope	70	926	0,076	929	28397	0,033	4,90E-10	1,20E-08
GO:0031967	C	organelle envelope	70	926	0,076	929	28397	0,033	4,90E-10	1,20E-08
GO:0048046	C	apoplast	37	926	0,040	406	28397	0,014	9,60E-08	2,20E-06
GO:0010319	C	stromule	10	926	0,011	37	28397	0,001	1,70E-06	3,60E-05
GO:0005576	C	extracellular region	137	926	0,148	2824	28397	0,099	3,20E-06	6,50E-05
GO:0044446	C	intracellular organelle part	150	926	0,162	3379	28397	0,119	8,70E-05	0,0017
GO:0044422	C	organelle part	150	926	0,162	3385	28397	0,119	9,40E-05	0,0018

GO:0010598	C	NAD(P)H dehydrogenase complex (plastoquinone)	5	926	0,005	11	28397	0,000	0,0001	0,0018
GO:0031977	C	thylakoid lumen	11	926	0,012	88	28397	0,003	0,00029	0,0051
GO:0031978	C	plastid thylakoid lumen	9	926	0,010	71	28397	0,003	0,00094	0,015
GO:0009543	C	chloroplast thylakoid lumen	9	926	0,010	71	28397	0,003	0,00094	0,015
GO:0009521	C	photosystem	8	926	0,009	68	28397	0,002	0,0027	0,041
GO:0010287	C	plastoglobule	8	926	0,009	68	28397	0,002	0,0027	0,041

S11: GO analysis for genes upregulated in *luc7 a-2*, *b-1*, *rl-1* and *se-1*.

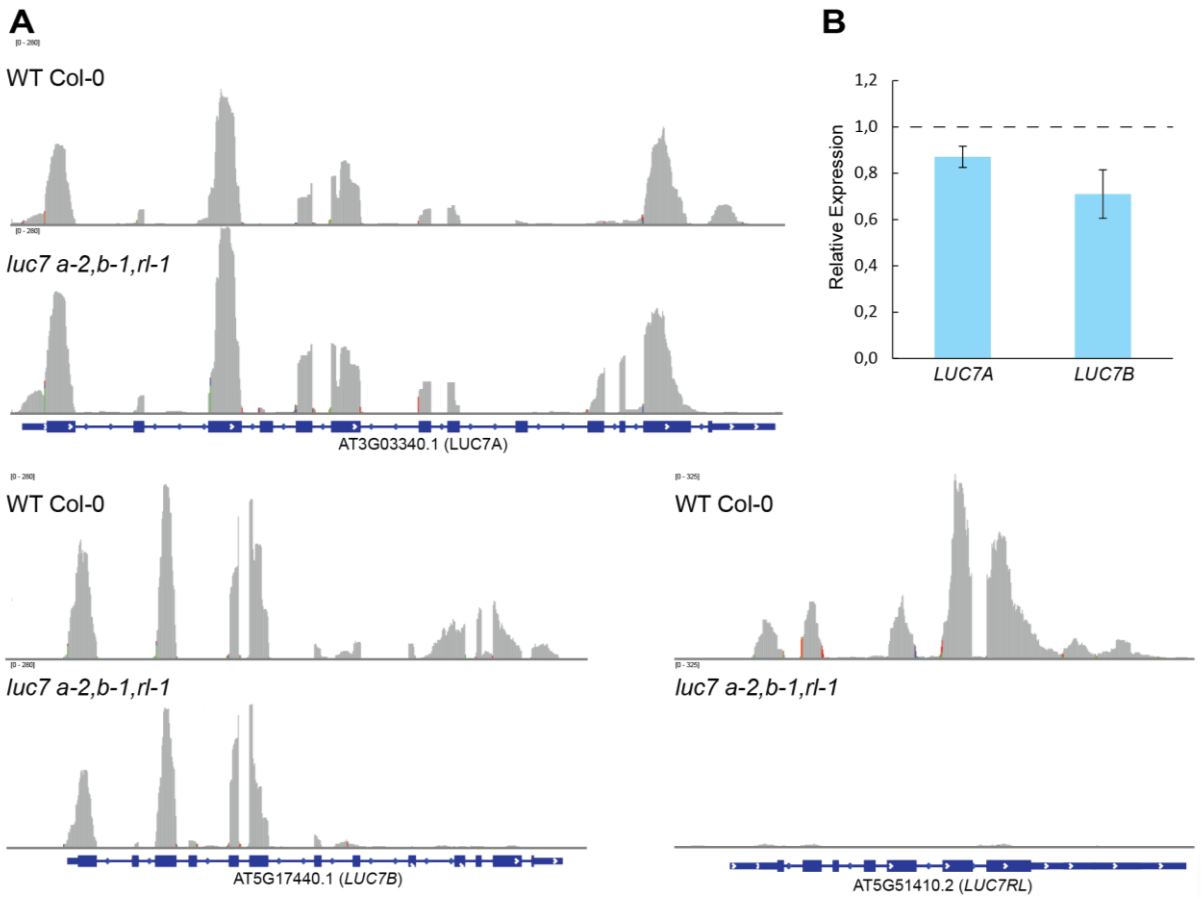
			Query		Observ.	Background		Expect.		
GO_acc	type	Term	item	total	Freq.	item	total	Freq.	pvalue	FDR
GO:0042221	P	response to chemical stimulus	115	390	0,295	3978	28397	0,140	4,10E-15	9,50E-12
GO:0009743	P	response to carbohydrate stimulus	45	390	0,115	812	28397	0,029	9,70E-15	1,10E-11
GO:0050896	P	response to stimulus	148	390	0,379	6292	28397	0,222	2,00E-12	1,50E-09
GO:0010033	P	response to organic substance	84	390	0,215	2754	28397	0,097	4,60E-12	2,60E-09
GO:0009750	P	response to fructose stimulus	18	390	0,046	144	28397	0,005	9,40E-12	4,30E-09
GO:0034284	P	response to monosaccharide stimulus	18	390	0,046	170	28397	0,006	1,10E-10	3,70E-08
GO:0009746	P	response to hexose stimulus	18	390	0,046	170	28397	0,006	1,10E-10	3,70E-08
GO:0006950	P	response to stress	104	390	0,267	4089	28397	0,144	2,80E-10	8,00E-08
GO:0034285	P	response to disaccharide stimulus	19	390	0,049	213	28397	0,008	5,10E-10	1,30E-07
GO:0009744	P	response to sucrose stimulus	18	390	0,046	210	28397	0,007	2,60E-09	6,00E-07
GO:0010200	P	response to chitin	24	390	0,062	421	28397	0,015	1,20E-08	2,60E-06
GO:0009414	P	response to water deprivation	23	390	0,059	416	28397	0,015	4,30E-08	8,20E-06
GO:0009415	P	response to water	23	390	0,059	424	28397	0,015	6,00E-08	1,00E-05
GO:0009646	P	response to absence of light	8	390	0,021	37	28397	0,001	1,50E-07	2,40E-05
GO:0009605	P	response to external stimulus	38	390	0,097	1087	28397	0,038	2,70E-07	4,00E-05
GO:0009719	P	response to endogenous stimulus	47	390	0,121	1615	28397	0,057	1,60E-06	2,30E-04
GO:0051716	P	cellular response to stimulus	60	390	0,154	2355	28397	0,083	3,50E-06	4,70E-04
GO:0009725	P	response to hormone stimulus	41	390	0,105	1375	28397	0,048	4,40E-06	5,60E-04
GO:0031668	P	cellular response to extracellular stimulus	18	390	0,046	388	28397	0,014	1,30E-05	1,50E-03
GO:0071496	P	cellular response to external stimulus	18	390	0,046	389	28397	0,014	1,30E-05	1,50E-03
GO:0023052	P	signaling	58	390	0,149	2376	28397	0,084	1,80E-05	2,00E-03
GO:0009991	P	response to extracellular stimulus	18	390	0,046	406	28397	0,014	2,30E-05	2,40E-03
GO:0009628	P	response to abiotic stimulus	62	390	0,159	2635	28397	0,093	2,60E-05	2,60E-03
GO:0071495	P	cellular response to endogenous stimulus	27	390	0,069	815	28397	0,029	3,70E-05	3,40E-03
GO:0007242	P	intracellular signaling cascade	36	390	0,092	1252	28397	0,044	3,60E-05	3,40E-03
GO:0031669	P	cellular response to nutrient levels	16	390	0,041	350	28397	0,012	4,60E-05	3,90E-03
GO:0009751	P	response to salicylic acid stimulus	19	390	0,049	470	28397	0,017	4,60E-05	3,90E-03
GO:0031667	P	response to nutrient levels	16	390	0,041	367	28397	0,013	7,80E-05	6,20E-03
GO:0007154	P	cell communication	25	390	0,064	758	28397	0,027	7,70E-05	6,20E-03
GO:0010150	P	leaf senescence	7	390	0,018	70	28397	0,002	8,50E-05	6,50E-03
GO:0009620	P	response to fungus	19	390	0,049	499	28397	0,018	9,70E-05	7,20E-03
GO:0009267	P	cellular response to starvation	15	390	0,038	336	28397	0,012	1,00E-04	7,30E-03
GO:0009611	P	response to wounding	15	390	0,038	340	28397	0,012	1,10E-04	8,00E-03
GO:0050832	P	defense response to fungus	15	390	0,038	342	28397	0,012	1,20E-04	8,20E-03
GO:0042594	P	response to starvation	15	390	0,038	344	28397	0,012	1,30E-04	8,50E-03
GO:0055072	P	iron ion homeostasis	5	390	0,013	33	28397	0,001	1,50E-04	9,80E-03
GO:0048511	P	rhythmic process	10	390	0,026	171	28397	0,006	1,90E-04	1,20E-02
GO:0007623	P	circadian rhythm	10	390	0,026	171	28397	0,006	1,90E-04	1,20E-02
GO:0006952	P	defense response	41	390	0,105	1653	28397	0,058	2,50E-04	1,40E-02
GO:0007568	P	aging	9	390	0,023	145	28397	0,005	2,60E-04	1,50E-02
GO:0045087	P	innate immune response	27	390	0,069	930	28397	0,033	3,00E-04	1,70E-02
GO:0032870	P	cellular response to hormone stimulus	21	390	0,054	641	28397	0,023	3,10E-04	1,70E-02
GO:0009755	P	hormone-mediated signaling pathway	20	390	0,051	600	28397	0,021	3,50E-04	1,90E-02
GO:0070887	P	cellular response to chemical stimulus	36	390	0,092	1417	28397	0,050	3,90E-04	2,00E-02
GO:0051707	P	response to other organism	36	390	0,092	1421	28397	0,050	4,10E-04	2,00E-02
GO:0009753	P	response to jasmonic acid stimulus	17	390	0,044	471	28397	0,017	4,10E-04	2,00E-02
GO:0050794	P	regulation of cellular process	89	390	0,228	4595	28397	0,162	4,50E-04	2,20E-02
GO:0009266	P	response to temperature stimulus	27	390	0,069	962	28397	0,034	5,00E-04	2,40E-02
GO:0010260	P	organ senescence	7	390	0,018	96	28397	0,003	5,20E-04	2,40E-02
GO:0010149	P	senescence	7	390	0,018	96	28397	0,003	5,20E-04	2,40E-02
GO:0009737	P	response to abscisic acid stimulus	20	390	0,051	621	28397	0,022	5,30E-04	2,40E-02
GO:0007165	P	signal transduction	40	390	0,103	1670	28397	0,059	5,70E-04	2,50E-02
GO:0010017	P	red or far-red light signaling pathway	5	390	0,013	45	28397	0,002	5,70E-04	2,50E-02

GO:0006979	P	response to oxidative stress	19	390	0,049	582	28397	0,020	6,20E-04	2,60E-02
GO:0009409	P	response to cold	20	390	0,051	629	28397	0,022	6,20E-04	2,60E-02
GO:0071489	P	cellular response to red or far red light	5	390	0,013	47	28397	0,002	6,90E-04	2,70E-02
GO:0002376	P	immune system process	27	390	0,069	984	28397	0,035	7,00E-04	2,70E-02
GO:0006955	P	immune response	27	390	0,069	984	28397	0,035	7,00E-04	2,70E-02
GO:0045454	P	cell redox homeostasis	8	390	0,021	136	28397	0,005	8,00E-04	3,10E-02
GO:0050789	P	regulation of biological process	97	390	0,249	5235	28397	0,184	1,00E-03	4,00E-02
GO:0006355	P	regulation of transcription, DNA-dependent	51	390	0,131	2372	28397	0,084	1,10E-03	4,20E-02
GO:0045449	P	regulation of transcription	51	390	0,131	2376	28397	0,084	1,20E-03	4,30E-02
GO:0071310	P	cellular response to organic substance	31	390	0,079	1234	28397	0,043	1,20E-03	4,30E-02
GO:0051252	P	regulation of RNA metabolic process	51	390	0,131	2388	28397	0,084	1,30E-03	4,70E-02
GO:0031326	P	regulation of cellular biosynthetic process	55	390	0,141	2631	28397	0,093	1,30E-03	4,70E-02
GO:0009889	P	regulation of biosynthetic process	55	390	0,141	2634	28397	0,093	1,40E-03	4,80E-02
GO:0030528	F	transcription regulator activity	45	390	0,115	1740	28397	0,061	4,70E-05	0,015
GO:0015035	F	protein disulfide oxidoreductase activity	8	390	0,021	94	28397	0,003	7,60E-05	0,015
GO:0015036	F	disulfide oxidoreductase activity	8	390	0,021	102	28397	0,004	0,00013	0,016
GO:0003700	F	transcription factor activity	41	390	0,105	1682	28397	0,059	0,00035	0,028
GO:0008270	F	zinc ion binding	38	390	0,097	1521	28397	0,054	0,00036	0,028
GO:0031072	F	heat shock protein binding	8	390	0,021	132	28397	0,005	0,00067	0,036
GO:0016407	F	acetyltransferase activity	6	390	0,015	72	28397	0,003	0,00067	0,036

S12: GO analysis for genes downregulated in *luc7 a-2, b-1, rl-1* and *se-1*.

			Query		Observ.	Background		Expect.		
GO_acc	type	Term	item	total	Freq.	item	total	Freq.	pvalue	FDR
GO:0046173	P	polyol biosynthetic process	20	339	0,059	8,10E+01	2,84E+04	0,003	4,40E-19	1,10E-15
GO:0032958	P	inositol phosphate biosynthetic process	18	339	0,053	6,60E+01	2,84E+04	0,002	5,90E-18	3,80E-15
GO:0010264	P	myo-inositol hexakisphosphate biosynthetic process	18	339	0,053	6,50E+01	2,84E+04	0,002	4,70E-18	3,80E-15
GO:0033517	P	myo-inositol hexakisphosphate metabolic process	18	339	0,053	6,50E+01	2,84E+04	0,002	4,70E-18	3,80E-15
GO:0006021	P	inositol biosynthetic process	18	339	0,053	6,70E+01	2,84E+04	0,002	7,40E-18	3,80E-15
GO:0043647	P	inositol phosphate metabolic process	18	339	0,053	8,60E+01	2,84E+04	0,003	3,30E-16	1,40E-13
GO:0006020	P	inositol metabolic process	18	339	0,053	9,80E+01	2,84E+04	0,003	2,50E-15	9,10E-13
GO:0019751	P	polyol metabolic process	20	339	0,059	1,36E+02	2,84E+04	0,005	3,10E-15	9,90E-13
GO:0009416	P	response to light stimulus	44	339	0,130	1,19E+03	2,84E+04	0,042	6,70E-11	1,90E-08
GO:0046165	P	alcohol biosynthetic process	21	339	0,062	2,93E+02	2,84E+04	0,010	2,10E-10	5,30E-08
GO:0009314	P	response to radiation	44	339	0,130	1,26E+03	2,84E+04	0,044	4,30E-10	9,90E-08
GO:0055114	P	oxidation reduction	46	339	0,136	1,36E+03	2,84E+04	0,048	4,60E-10	9,90E-08
GO:0006066	P	alcohol metabolic process	41	339	0,121	1,14E+03	2,84E+04	0,040	8,00E-10	1,60E-07
GO:0051186	P	cofactor metabolic process	31	339	0,091	7,53E+02	2,84E+04	0,027	5,30E-09	9,70E-07
GO:0015979	P	photosynthesis	23	339	0,068	4,35E+02	2,84E+04	0,015	7,20E-09	1,20E-06
GO:0019748	P	secondary metabolic process	41	339	0,121	1,25E+03	28397	0,044	8,90E-09	1,40E-06
GO:0006733	P	oxidoreduction coenzyme metabolic process	18	339	0,053	2,67E+02	28397	0,009	1,00E-08	1,50E-06
GO:0010114	P	response to red light	12	339	0,035	1,04E+02	28397	0,004	1,40E-08	2,00E-06
GO:0042440	P	pigment metabolic process	20	339	0,059	3,61E+02	28397	0,013	3,40E-08	4,70E-06
GO:0005975	P	carbohydrate metabolic process	57	339	0,168	2,25E+03	28397	0,079	7,90E-08	1,00E-05
GO:0009637	P	response to blue light	12	339	0,035	1,26E+02	28397	0,004	9,60E-08	1,10E-05
GO:0046496	P	nicotinamide nucleotide metabolic process	15	339	0,044	2,12E+02	28397	0,007	9,60E-08	1,10E-05
GO:0006769	P	nicotinamide metabolic process	15	339	0,044	2,12E+02	28397	0,007	9,60E-08	1,10E-05
GO:0019362	P	pyridine nucleotide metabolic process	15	339	0,044	2,14E+02	28397	0,008	1,10E-07	1,20E-05
GO:0044262	P	cellular carbohydrate metabolic process	48	339	0,142	1,78E+03	28397	0,063	1,70E-07	1,70E-05
GO:0043603	P	cellular amide metabolic process	15	339	0,044	2,22E+02	28397	0,008	1,70E-07	1,70E-05
GO:0009820	P	alkaloid metabolic process	15	339	0,044	2,25E+02	28397	0,008	2,00E-07	1,90E-05
GO:0019684	P	photosynthesis, light reaction	18	339	0,053	3,33E+02	28397	0,012	2,40E-07	2,20E-05
GO:0044283	P	small molecule biosynthetic process	49	339	0,145	1,87E+03	28397	0,066	2,70E-07	2,40E-05
GO:0006098	P	pentose-phosphate shunt	14	339	0,041	2,00E+02	28397	0,007	2,90E-07	2,50E-05
GO:0006740	P	NADPH regeneration	14	339	0,041	2,01E+02	28397	0,007	3,10E-07	2,60E-05
GO:0006739	P	NADP metabolic process	14	339	0,041	204	28397	0,007	3,60E-07	3,00E-05
GO:0034637	P	cellular carbohydrate biosynthetic process	31	339	0,091	941	28397	0,033	6,30E-07	4,60E-05
GO:0046148	P	pigment biosynthetic process	16	339	0,047	282	28397	0,010	6,00E-07	4,60E-05
GO:0008152	P	metabolic process	208	339	0,614	13633	28397	0,480	6,30E-07	4,60E-05
GO:0009657	P	plastid organization	19	339	0,056	398	28397	0,014	6,60E-07	4,70E-05
GO:0010218	P	response to far red light	10	339	0,029	99	28397	0,003	6,90E-07	4,90E-05
GO:0015994	P	chlorophyll metabolic process	13	339	0,038	189	28397	0,007	9,30E-07	6,40E-05
GO:0006732	P	coenzyme metabolic process	21	339	0,062	508	28397	0,018	1,60E-06	1,00E-04
GO:0009658	P	chloroplast organization	14	339	0,041	237	28397	0,008	2,00E-06	1,30E-04
GO:0009639	P	response to red or far red light	18	339	0,053	406	28397	0,014	3,50E-06	2,20E-04
GO:0006778	P	porphyrin metabolic process	13	339	0,038	226	28397	0,008	6,00E-06	3,70E-04
GO:0033013	P	tetrapyrrole metabolic process	13	339	0,038	227	28397	0,008	6,30E-06	3,80E-04
GO:0016051	P	carbohydrate biosynthetic process	31	339	0,091	1070	28397	0,038	8,00E-06	4,70E-04
GO:0044281	P	small molecule metabolic process	77	339	0,227	4065	28397	0,143	2,60E-05	1,50E-03
GO:0055086	P	nucleobase, nucleoside and nucleotide metabolic process	22	339	0,065	726	28397	0,026	9,20E-05	5,20E-03
GO:0051188	P	cofactor biosynthetic process	15	339	0,044	388	28397	0,014	1,00E-04	5,60E-03
GO:0009117	P	nucleotide metabolic process	21	339	0,062	685	28397	0,024	1,10E-04	6,10E-03
GO:0006753	P	nucleoside phosphate metabolic process	21	339	0,062	687	28397	0,024	1,20E-04	6,20E-03

GO:0015995	P	chlorophyll biosynthetic process	8	339	0,024	125	28397	0,004	1,90E-04	9,80E-03
GO:0009739	P	response to gibberellin stimulus	9	339	0,027	166	28397	0,006	2,50E-04	1,30E-02
GO:0006091	P	generation of precursor metabolites and energy	21	339	0,062	730	28397	0,026	2,60E-04	1,30E-02
GO:0009628	P	response to abiotic stimulus	52	339	0,153	2635	28397	0,093	2,70E-04	1,30E-02
GO:0006007	P	glucose catabolic process	16	339	0,047	474	28397	0,017	2,70E-04	1,30E-02
GO:0019320	P	hexose catabolic process	16	339	0,047	476	28397	0,017	2,80E-04	1,30E-02
GO:0046365	P	monosaccharide catabolic process	16	339	0,047	480	28397	0,017	3,00E-04	1,40E-02
GO:0046164	P	alcohol catabolic process	16	339	0,047	491	28397	0,017	3,90E-04	1,80E-02
GO:0044275	P	cellular carbohydrate catabolic process	17	339	0,050	545	28397	0,019	4,20E-04	1,90E-02
GO:0009767	P	photosynthetic electron transport chain	6	339	0,018	80	28397	0,003	5,50E-04	2,40E-02
GO:0016052	P	carbohydrate catabolic process	17	339	0,050	573	28397	0,020	7,20E-04	3,10E-02
GO:0046483	P	heterocycle metabolic process	25	339	0,074	1023	28397	0,036	7,50E-04	3,20E-02
GO:0006779	P	porphyrin biosynthetic process	8	339	0,024	157	28397	0,006	8,00E-04	3,30E-02
GO:0033014	P	tetrapyrrole biosynthetic process	8	339	0,024	160	28397	0,006	9,00E-04	3,70E-02
GO:0016491	F	oxidoreductase activity	50	339	0,147	1691	28397	0,060	5,30E-09	2,10E-06
GO:0003824	F	catalytic activity	148	339	0,437	8787	28397	0,309	6,70E-07	0,00013
GO:0004497	F	monooxygenase activity	10	339	0,029	112	28397	0,004	2,00E-06	0,00025
GO:0046527	F	glucosyltransferase activity	10	339	0,029	142	28397	0,005	1,40E-05	0,0013
GO:0035251	F	UDP-glucosyltransferase activity	8	339	0,024	114	28397	0,004	0,0001	0,008
GO:0016705	F	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	16	339	0,047	462	28397	0,016	0,0002	0,013
GO:0008194	F	UDP-glycosyltransferase activity	10	339	0,029	239	28397	0,008	0,0008	0,04
GO:0016758	F	transferase activity, transferring hexosyl groups	13	339	0,038	396	28397	0,014	0,0013	0,04
GO:0043167	F	ion binding	55	339	0,162	3030	28397	0,107	0,0012	0,04
GO:0016765	F	transferase activity, transferring alkyl or aryl (other than methyl) groups	7	339	0,021	123	28397	0,004	0,00091	0,04
GO:0043169	F	cation binding	55	339	0,162	3029	28397	0,107	0,0012	0,04
GO:0050660	F	FAD binding	8	339	0,024	166	28397	0,006	0,0011	0,04
GO:0050661	F	NADP or NADPH binding	5	339	0,015	66	28397	0,002	0,0015	0,045
GO:0044435	C	plastid part	42	339	0,124	1252	28397	0,044	3,30E-09	5,80E-07
GO:0044434	C	chloroplast part	41	339	0,121	1211	28397	0,043	4,00E-09	5,80E-07
GO:0009526	C	plastid envelope	24	339	0,071	598	28397	0,021	4,60E-07	3,40E-05
GO:0009579	C	thylakoid	22	339	0,065	518	28397	0,018	5,70E-07	3,40E-05
GO:0009570	C	chloroplast stroma	24	339	0,071	603	28397	0,021	5,40E-07	3,40E-05
GO:0009536	C	plastid	82	339	0,242	4037	28397	0,142	8,90E-07	4,30E-05
GO:0009532	C	plastid stroma	24	339	0,071	637	28397	0,022	1,40E-06	5,70E-05
GO:0009507	C	chloroplast	80	339	0,236	3959	28397	0,139	1,60E-06	5,70E-05
GO:0009941	C	chloroplast envelope	22	339	0,065	573	28397	0,020	2,80E-06	9,10E-05
GO:0010319	C	stromule	6	339	0,018	37	28397	0,001	1,10E-05	0,00032
GO:0031976	C	plastid thylakoid	16	339	0,047	425	28397	0,015	8,00E-05	0,0017
GO:0031975	C	envelope	26	339	0,077	929	28397	0,033	7,60E-05	0,0017
GO:0009534	C	chloroplast thylakoid	16	339	0,047	425	28397	0,015	8,00E-05	0,0017
GO:0031967	C	organelle envelope	26	339	0,077	929	28397	0,033	7,60E-05	0,0017
GO:0031984	C	organelle subcompartment	16	339	0,047	428	28397	0,015	8,70E-05	0,0017
GO:0009535	C	chloroplast thylakoid membrane	13	339	0,038	322	28397	0,011	0,00019	0,0035
GO:0055035	C	plastid thylakoid membrane	13	339	0,038	324	28397	0,011	0,00021	0,0035
GO:0044436	C	thylakoid part	14	339	0,041	388	28397	0,014	0,00034	0,0052
GO:0042651	C	thylakoid membrane	13	339	0,038	341	28397	0,012	0,00033	0,0052
GO:0034357	C	photosynthetic membrane	13	339	0,038	355	28397	0,013	0,00048	0,007



S13: *luc7 a-2, b-1, rl-1* is a hypomorphic mutant. (A) RNA sequencing coverage plot for *AthLUC7A*, *AthLUC7B* and *AthLUC7RL*. (B) RT-qPCR for *LUC7A* and *LUC7B* with primers positioned before the T-DNA insertions. Two biological replicates were performed and PP2A was used to normalize. Expression level is relative to WT (dashed line).