

Stress Induces the Retention of Introns in *ECA3* Transcripts in Wheat

Roohi Aslam¹, Emily C. Farthing², Nasar Virk¹, Muhammad Faraz Bhatti¹, Lorraine E. Williams²

Abstract— Alternative splicing involves combining exons in variety of ways to create protein diversity. Sometimes alternative splicing may lead to intron retention in transcripts, which can introduce a premature stop codon, rendering the protein non-functional. It is believed that environmental stresses may cause errors in the molecular machinery leading to the production of transcripts with retained introns. *ECA3* is a P_{2A}-type ATPase believed to have important roles in calcium and manganese transport. In the present study, *ECA3* transcripts in wheat were found which retained an intron near the 3' end followed by exon-exon junctions. To investigate the relationship between stress and the event of intron retention in wheat *ECA3* transcripts, plants were grown under calcium stress. The plants grown under calcium stress produced higher levels of the intron-retaining *ECA3* transcripts compared to control plants, suggesting a possible relationship between stress factor and intron retention events in *ECA3*.

Keywords— intron retention *ECA3*, *Wheat*.

I. INTRODUCTION

The raw RNA transcribed from DNA contains introns as well exons. In order to make it functional i.e. to be able to code for protein, it is necessary to remove any of the intervening introns. This removal of introns and ligation of exons is facilitated by a unique step known as RNA splicing. Interestingly, eukaryotic cells are able to code for more than one type of protein from a single gene. It happens because of alternate splicing (also known as differential splicing), which is a form of splicing that involves combining exons in different ways to generate different proteins (Tilgner et al., 2012). In addition to protein diversity, another important aspect of alternate splicing is its effect on mRNA stability, hence turnover of genes. It happens as some of the spliced forms contain a premature termination codon (PTC), therefore they are not translated and are subjected to NMD (non-sense mediated decay) pathways (Lewis et al., 2003; Maquat et al., 2004) which targets them for degradation (Maquat, 2004). Intron retention is the most common type of alternate splicing in plants while this is the rarest mode of splicing in mammals. In this case a sequence may be spliced out as an intron or may be simply retained. Recent evidences from different studies broadly helped in increasing the knowledge of AS in Arabidopsis hence, the event occurs far more than was previously expected (Alexandrov et al., 2006; Haas et al., 2003; Lida et al., 2004; Nagasaki et al., 2005; Ner-Gaon et al., 2004; Zhu et al., 2003). Also in rice, 33-48% of genes are

found to be alternatively spliced (Chen et al., 2007; Lu et al., 2010; Xiao et al., 2005; Zhang et al., 2010). It is believed that environmental stresses, both biotic and abiotic, have a role in events of AS, mostly in the form of intron retention in plants (Dinesh-Kumar and Baker, 2000; Hori et al. 2005; Iida et al. 2004; Palusa et al. 2006; Reddy 2001; Zhang and Gassman, 2003). For example, heat stress was found to inhibit the splicing of maize poly ubiquitin and *HSP 70*, *HSP 81* in Arabidopsis and *Waxy* gene in rice. Similarly, the exposure of maize seedlings to cadmium stress increases the level of non-spliced intron containing transcripts (Christensen et al., 1992; Hopf et al., 1992; Jordan et al., 2002; Marrs and Welbot, 1997). Alternate splicing events in Arabidopsis genes are largely affected by cold and other stresses (Lida et al., 2004). Such events allow plants to quickly regulate splicing and gene expression of many unrelated genes. During stress, several mechanisms may operate in plants to regulate the splicing events. If a particular transcript is required in lesser amounts during a particular stress, the AS events might result in transcripts with PTC, which are ultimately subjected to NMD pathways. On the other hand, multiple isoforms of a protein might be generated using the same AS events during stress to enable the plant to respond better to the environment. The continued investigations in this direction will lead to better understanding of how plants respond to stress and a changing environment.

Plants possess different types of membrane proteins which facilitate transport of different ions across the membrane. Broadly they are divided into three main types i.e., channel proteins, carriers, and ATP powered pumps. Among ATP powered pumps, the P₂ type ATPases are specialized in Ca²⁺ transport and signaling and are further classified into P_{2A} and P_{2B} type ATPases. The P_{2A} type ATPases from plants (such as OsECA2, CrECA4, PpECA4, SmECA4 and AtECA3) form a distinct set of Ca²⁺ ATPases, generally called ECAs and are closely related to the animal Sarco-endoplasmic reticulum Ca²⁺ pump SERCA1. So far, four such pumps have been identified in Arabidopsis (*ECA1*, *ECA2*, *ECA3*, *ECA4*) while three are present in rice (*ECA1*, *ECA2*, *ECA3*) (Baxter et al., 2003; Kamrul et al., 2013). *AtECA3* (also known as *AtACA6*) is an important calcium transporting P_{2A} type Calcium ATPases. In addition to calcium translocation, studies have shown that this pump is very efficient in manganese transport across the membrane. Arabidopsis *eca3* mutants are detrimentally affected when grown under manganese deficiency (Mills et al., 2008).

The present study reports on the identification of *ECA3* genes in wheat. The gene was identified and amplified to obtain sequence information. Furthermore, *ECA3* transcripts in

¹National University of Science and Technology Islamabad, Pakistan

²University of Southampton, UK.

wheat were found to retain an intron near the 3' end of the gene, a case of alternate splicing. Here the relationship between intron retention in wheat *ECA3* and calcium stress was investigated.

II. MATERIAL AND METHODS

Bioinformatics approach

The *B. distachyon ECA3* gene sequence was retrieved from Aramemnon (<http://aramemnon.uni-koeln.de/>) database. This *ECA3* sequence was used for BLAST searches in INRA, URGI (<https://wheat-urgi.versailles.inra.fr/Seq-Repository/BLAST>) wheat portal to retrieve wheat contigs possibly carrying wheat *ECA3* sequences. The contig (IWGSC_chr4BS_ab_k71_contigs) was aligned (http://www.ebi.ac.uk/Tools/st/emboss_transeq/) with *B. distachyon ECA3* sequence to predict exon-intron boundaries and were edited manually to predict wheat *ECA3* CDS. The resulting CDS correspond to wheat chromosome 4BS and homeologue 4B.

A. Primer designing

The predicted CDS was used to design primers from different regions of the wheat *ECA3* gene. An additional CACC sequence was introduced in forward primers to facilitate cloning in pENTR/D-TOPO vector (Invitrogen). For qRT-PCR, primers were designed within intronic region. All the primers used in this study are outlined in Table 1.

Growth of wheat plants and generating calcium stress

The hydroponic system was used to expose *Triticum aestivum* (spp. Paragon) plants to calcium deficiency/toxicity conditions. Seeds were surface sterilized using 1% bleach solution and were germinated for five days. After germination they were grown for 14 days on standard media according to Lombnaes and Singh (2003). The fourteenth day of growth on standard Lombnaes media is referred to as D0. Plants were then split into three sets and Lombnaes media was modified to induce deficiency and toxicity stress. To induce calcium deficiency no calcium was added to the Lombnaes media while for calcium toxicity 8mM of calcium was added. Normal 2mM calcium concentration was maintained for control plants. The roots of the plants were washed with ddH₂O thrice before transferring to the calcium deficient and calcium toxic media. The plants were grown for a further 21 days. Nine plants (three from each set) were harvested to liquid nitrogen on days 7, 14, and 21. The plants were cultivated in environmentally controlled growth room with temperature set at 21°C/16°C (day/night), humidity maintained at 55-65%. The photoperiod was kept at 16hrs at a quantum flux density (PAR) of 220 $\mu\text{mol m}^{-2} \text{s}^{-1}$.

B. RNA extraction and cDNA synthesis

RNA was extracted from plant roots and shoots. One mL Trizol (Invitrogen, CA, USA) reagent was added to the eppendorf containing about 0.5ml of finely ground wheat tissue and was vortexed vigorously to achieve homogenization. Chloroform 25 % (v/v), was added to the mixture and was incubated at room temperature for five minutes before centrifuging at 12,000 g at 4 °C for 15 min. After centrifugation, the upper phase was carefully transferred to a

new tube followed by the addition of 50 % (v/v) isopropyl alcohol. The mixture was vortexed briefly and then incubated at room temperature for 10 min. After ten minutes the mixture was centrifuged at 12,000 g at 4 °C and RNA pellet was obtained. For washing, the pellet was re suspended in 1 mL 75 % (v/v) ethanol and vortexed briefly and was centrifuged at 7600 g at 4 °C for 5 min. This step was repeated three times. The supernatant was removed and the pellet was completely air dried at room temperature for at least 5-10 minutes. Pellet was re suspended in freshly prepared 30 μL TE buffer (pH 7.0). RNA samples were treated with DNase to prevent any possible genomic contamination. Extracted RNA was used to synthesize first strand complementary DNA (cDNA) using cDNA synthesis kit (Invitrogen), following manufacturer's instructions.

C. Amplification and purification of sequences

For amplification of the target sequences 10mM of template cDNA, along with 10X pfu buffer, 10mM dNTPs, 10 μM forward primer, 10 μM reverse primer, pfu polymerase (1.25 units) and Sterile 18 M Ω H₂O was added in specific PCR tubes. The PCR was carried out at 95°C for 2 minutes followed by 40 cycle of 95°C for 1 minute, 55°C for 30seconds, 72°C (1min/500bases). The final extension was at 72°C for 7 minutes. The amplified band was gel extracted using Qiagen Gel Extraction Kit (Qiagen, CA, USA), according to manufacturer's instructions.

D. Cloning of sequences

Two target sequences for *ECA3* (amplified using primer pair 1 and 2 given in Table 1) were cloned using Gateway technology into the pENTR/D-TOPO vector (Invitrogen, CA, USA). Positive clones were sequenced (Source Bioscience Life Sciences, UK).

E. Real time-PCR

Real-time PCR was performed using the SYBR Green Kit (Invitrogen). For each reaction, 2.5ng of template DNA, 0.3 μM of forward and reverse primers, 1X SYBR-green master mix and sterile 18 Ω H₂O up to 20 μL was used in a 96 well plate format. The reaction was run on an Opticon DNA Engine Continuous Fluorescence Detector (Applied Biosystems 7000 Real-time PCR system). The conditions used were 95°C for 2 minutes before cycling forty times at 95°C for 50 seconds, 60°C for 50second, 70°C for 5 minutes and a final extension time of 71°C for 10 minutes. Actin was used as the reference gene and the expression levels were conducted according to the method of (Pfaffl, 2001).

III. RESULTS

A. To identify the *ECA3* genes from *T. aestivum*, BLAST searches were conducted in the wheat portal (URGI) using *B. distachyon ECA3* predicted sequence. The contigs identified were aligned with *B. distachyon ECA3* sequence for the prediction of exon-intron junctions using EMBOSS pair wise alignment tool. Thirty four exons and thirty three introns were predicted in the sequence. The aligned sequence was edited manually for the prediction of *ECA3* CDS. The predicted CDS was 3009bp in length equating to 1003 amino acids, similar to *B. distachyon*. The predicted wheat *ECA3* sequence showed

95% identity to *B. distachyon ECA3* sequence and 92% to *O. sativa ECA3* sequence Table 2.

B. Amplification and Cloning of *T. aestivum ECA3* gene

The primers were designed using predicted *T. aestivum ECA3* CDS. *T. aestivum* cDNA was used to amplify full length *ECA3* gene using primer pair TaECA3F1 and TaECA3R2 (Table 1). The amplified *ECA3* gene was then used to re amplify the gene in two parts to facilitate cloning as attempts to clone the full-length version failed. The primer TaECA3F1 was used in combination with TaECA3R1 to amplify 1338bp of *ECA3* whereas the primer TaECA3F2 with TaECA3R2 was used to amplify 2781bp. The amplified sequences were cloned using Gateway technology. The clones carrying 1338bp of TaECA3 gene are referred to as pENTRTaECA3-5' (as carried 5' end of TaECA3) while the clones carrying 2781bp was referred to as pENTRECA3-3' (as carried 3' end). The clones were verified using restriction digestion and sequenced.

Intron retention in *ECA3* transcripts in *T. aestivum*

The sequence of pENTRTaECA3-5' was as predicted (Figure 1). For pENTRTaECA3-3' some of the clones appeared similar to the predicted sequence while others were different. The latter all retained an intron near the 3' end of the gene. It was observed that intron number 30 was retained, between exon 31 and 32 (Figure 2). The retained intron disrupted the whole reading frame as it resulted in the introduction of a premature stop codon (Figure 3). The disturbances in the reading frame caused loss of transmembrane (TM) 9 and TM10, hence disrupting the protein structure (Figure 4). This intron retention in *ECA3* transcripts in *T. aestivum* indicated that AS was operating.

Various studies have shown the presence of retained introns in plant transcripts, often happen due to the exposure of the plants to stress factors (Black, 2003; Graveley, 2001; Stamm, 2005). Here calcium stress was imposed to test whether intron retention was increased under stress conditions. The results of qRT-PCR using primers that amplify the intron region are shown in Figure 3.

The results of qRT-PCR show that in wheat shoots the event of intron retention is increased when the plants are exposed to deficiency and toxicity stress as compared to the plants grown under control condition. Plants harvested on day7 (i.e., seventh day after exposure to deficiency/toxicity treatment) showed much higher levels of transcripts retaining the intron in plants treated with calcium deficiency/toxicity compared to the level in plants in control conditions. This was also observed at 14 and 21 days but to a lower extent. In the case of wheat roots the plants under calcium deficiency and toxicity contained more transcripts with retained introns on day 7, 14 and 21 as compared to control samples. In this case there was a higher level in plants under toxicity than under deficiency. Also in the roots the number of transcripts with retained introns carried on increasing with the passage of time whereas the opposite affect was observed in shoots.

IV. DISCUSSION

The availability of wheat genome sequence information allows the identification and annotation of various wheat genes

(Brenchley et al., 2012). Here we aimed to determine whether wheat has an *ECA3* gene and to clone it for future functional analysis. *ECA3* was present as indicated by successful amplification by RT-PCR, and two partial length fragments could be cloned. Sequencing of the clones indicated that some transcripts of *ECA3* retained an intron in the 3' end of the gene resulting in an early stop codon. This intron retention was a clear example of alternate splicing.

There are several suggested possibilities responsible for the increased events of intron retention in plants. One possible reason might be the fact that plant introns are generally larger in length and lack a well-defined intron branch point compared to human introns (Tolstrup et al., 1997) which enhances intron retention in the transcripts. Moreover, plants also lack a strict intron border and are more UA (uridine and adenosine) rich (Brown et al., 2002; Lorkovic et al., 2000). The presence of introns in the untranslated regions (UTRs) might be another reason for intron retention in plants. For instance, at least 9% of the Arabidopsis genes have introns in their UTRs. Plants also have a more complex family of spliceosomal-associated serine arginine (SR) factors (Lorkovic and Barta, 2002) as compared to animals, which may also be responsible for increased intron retention in plants. All these factors might enhance the probability of IR, though this needs to be further investigated (Haas et al., 2002; Zhu et al., 2003; Barbazuk et al., 2008; Zhang et al., 2015).

Interestingly, sometimes in plants retained introns have been observed as part of CDS and UTR and, as a result, AS transcript may lead to a new open reading frame. The transcript would thus yield two different proteins which will differ in their activity (Golovkin and Reddy, 1996; Paterno et al., 2002). Hence intron retentions in CDS and UTR might highlight their significant roles in enhancing translational efficiency (Gebauer et al., 1998). Another important aspect of intron retention is the introduction of in-frame stop codons followed by spliced sections containing exon-exon junctions. In the present study, the retention of introns near the 3' end of *ECA3* in some of the transcripts indicate AS through IR. Interestingly, the retained intron resulted in disruption of the reading frame by introducing a premature stop codon (PMC). The presence of PMCs affect the protein structure, due to absence of important protein motifs, which may render the protein functionless (Figure 4). This finding may indicate the fact that a transcript with a retained intron becomes functionless (due to inability to code for full length *ECA3* protein), hence not required by the plants. It is possible that the NMD pathway is activated to destroy/remove these transcripts, similarly to animals suggesting these pathways are present in *T. aestivum*. An evidence for the existence of NMD pathways in plants came earlier from the study based on AS events in Arabidopsis *EARLY FLOWERING 3 (ELF3)* and *ZEITLUPE (ZTL)*. It was found that RNA splice variants of *IMING OF CAB EXPRESSION 1 (TOC1)* and *ELF3* in Arabidopsis were degraded through non-sense mediated decay pathway (Young-Ju Kwon et al., 2014).

A. Why plants produce transcripts with retained introns?

Studies have shown that the degree of the events of AS in plants are correlated with developmental stages as well as

environmental signals, including biotic and abiotic stresses (Black, 2003; Graveley, 2001; Stamm, 2005). Exposure to different stresses affects both the efficiency and patterns of splicing in plants (Luehrsen *et al.*, 1994; Simpson and Filipowicz, 1996). Increased exposure to stress (both biotic and abiotic) may increase the chance of intron retention in the transcripts (Luehrsen *et al.*, 1994; Simpson and Filipowicz, 1996). The relationship between environmental factors and intron retention has been investigated earlier in *Arabidopsis* (Filchikin *et al.*, 2015; Young-Ju Kwon *et al.*, 2014), wheat (Mastrangelo *et al.*, 2005), and maize (Marrs and Walbot, 1997). For instance, it was observed that the event of intron retention is much enhanced in durum wheat *E-COR* gene during stress as compared to control conditions (Mastrangelo *et al.*, 2005). Similarly, it was observed in maize that the *Bronze2* (*Bz2*) locus encodes for glutathione-S-transferase which is induced by heavy metal stresses such as cadmium. The maize seedlings exposed to cadmium stress are found to have a specific 20 fold increase in the *Bz2* message. Interestingly, such seedlings are found to have 50 fold increase in the presence of unspliced intron-containing transcripts (Marrs and Walbot, 1997). These findings indicate the correspondence between plant stress and IR events in the transcripts. The actual mechanism of how stress influences splicing in plants is still not clear and needs to be investigated in detail. However, if a particular stimuli influence the intron retention, then it is conceivable that the presence or absence of introns either helps in stabilizing the transcript or may serve to modify its biological function.

In this study wheat plants were grown under calcium stress (deficiency and toxicity) to investigate the possible relationship between metal stress and intron retention in *ECA3* transcripts. The results indicated that the event of intron retention is much enhanced in the plants grown under stress as

compared to the control conditions. This correlated with the previous findings that environmental factors influence molecular machinery leading to increased events of intron retention. The data from qRT-PCR in roots showed that transcripts with retained introns are more prominent when plants are grown under calcium toxic conditions as compared to the plants grown under calcium deficiency. The elevated calcium conditions may have a more severe effect on the molecular machinery as compared to deficiency conditions leading to more intron retention under toxicity. However, under both toxicity and deficiency conditions the number of transcripts with retained intron is greater as compared to the control conditions. The event of IR was observed more in roots as compared to the shoot. One possible reason may be that stress (both Ca deficiency and toxicity) had greater effect on roots as compared to the shoots. This may cause more errors in splicing in roots as compared to shoots, hence increasing the number of IR transcripts in roots. On the other hand calcium translocation to the shoots is restricted (due to root damage under calcium stress), although not completely. This restriction in the supply could be responsible for fewer events of *ECA3* IR in shoots.

V. CONCLUSION

The present study has shown the event of AS via IR in the P_{2A} -type ATPase wheat gene, *ECA3*. This gene has possible roles in calcium and manganese transport. The plants grown under calcium stress produce *ECA3* transcripts with retained introns in larger numbers as compared to plants grown under normal conditions. Interestingly, plants grown under toxicity retained more introns in *ECA3* as compared to plants grown under deficiency in roots. This suggests that both toxic and deficiency conditions effect molecular machinery leading to production of *ECA3* transcripts with retained introns.

TABLE I
PRIMER SEQUENCES USED TO AMPLIFY *ECA3* FROM WHEAT AND PRIMERS USED IN QRT-PCR

Primer pairs	Primers used for gene amplification and qRT-PCR	Sequence
1	TaECA3F1	CACCATGGAGGACGCCTA
	TaECA3R1	AACGCGCAGAGCAACTCA
2	TaECA3F2	CACCTTGGCTCGATTGAATGG
	TaECA3R2	TTAATTATCTCGGGCTTCTTTAGGA
3	TaECA3qF (qRT-PCR)	TCTCTACTTGTCATTCACCCATGG
	TaECA3qr (qRT-PCR)	TCTCTACTTGTCATTCACCCATGG

TABLE II
ECA3 GENE SEQUENCE INFORMATION IN *B. DISTACHYON*, *O. SATIVA* AND *T. AESTIVUM*.

Gene (ID)	Sequence information					
	Genomic (bp)	Chromosome	Exons	Introns	CDS	Protein (aa)
BdECA3 Bradi1g09810.1 (Aramemnon)	37038	I	34	33	3009	1003
TaECA3 Traes_4BS_73F7EF1B8 (Ensembl)	37038	4BS	34	33	3009	1003
OsECA3 LOC_Os03g52090.1 (Aramemnon)	20963	III	37	36	3654	1217

```

Predicted : MEDAVAKSAEVLQAFGVDRTRGLSDSQVEQHAIHDEKHWLPCSESTPFKWLVLQFDLLVRLIAAAVVSFLLARLNGETGLTAFLEPSVIFMILAAANAAGVITETNAEKALPEE : 118
A       : MEDAVAKSVSEVLQAFGVDRTRGLSDSQVEQHAIHDEKHWLPCSESTPFKWLVLQFDLLVRLIAAAVVSFLLARLNGETGLTAFLEPSVIFMILAAANAAGVITETNAEKALPEE : 118

Predicted : RAYQADVATVLRNGCFSEILPATELVPGDIVEVGVGCEVPAIDMRNVMENLSHQLRVDQAILGEGSCSVARELSDTSAMNAVYQDRTNILFSGTVVAGRARAVVIGVGSWTAMGSIRDAM : 236
A       : RAYQADVATVLRNGCFSEILPATELVPGDIVEVGVGCEVPAIDMRNVMENLSHQLRVDQAILGEGSCSVARELSDTSAMNAVYQDRTNILFSGTVVAGRARAVVIGVGSWTAMGSIRDAM : 236

Predicted : LRTEDEATPLRKLDEFGTFLLAKVIBGICILVWVWVNIHGFDPDPSHGGFLRGAHYFRVAVALAVAAIPGGLPAVVTTCLALGTRMARLNAIVRSLPSVETLGGTTVICSDRTGTLTT : 354
A       : LRTEDEATPLRKLDEFGTFLLAKVIBGICILVWVWVNIHGFDPDPSHGGFLRGAHYFRVAVALAVAAIPGGLPAVVTTCLALGTRMARLNAIVRSLPSVETLGGTTVICSDRTGTLTT : 354

Predicted : NMMSVSKVCCVRSVHQRPITDEYSISGTTFAPOGFIYDASDQLLEFPQSPCLLHIA MCSALCNESTLQYNPDRFSTEKIGESTEVALRV : 444
A       : NMMSVSKVCCVRSVHQRPITDEYSISGTTFAPOGFIYDASDQLLEFPQSPCLLHIA MCSALCNESTLQYNPDRFSTEKIGESTEVALRV : 444
    
```

Fig 1: The cloned 5' portion of TaECA3 from *Triticum aestivum* (Paragon) is similar to predicted sequence. A= pENTRTaECA3' amino acid sequence translated from amplified coding sequence.



Fig 2: The intron number 30 between exon number 31 and 32 was retained as a result of AS and introduced a premature stop codon (highlighted violet).

```

Predicted : ALLELRAYQADVATVLRNGCFSEILPATELVKQIIVEVGVGCEVPAIDMRNVMENLSHQLRVDQAILGEGSCSVARELSDTSAMNAVYQDRTNILFSGTVVAGRARAVVIGVGSWTAM : 230
A       : ALLELRAYQADVATVLRNGCFSEILPATELVKQIIVEVGVGCEVPAIDMRNVMENLSHQLRVDQAILGEGSCSVARELSDTSAMNAVYQDRTNILFSGTVVAGRARAVVIGVGSWTAM : 230

Predicted : MGSIRDAMLRTEDEATPLRKLDEFGTFLLAKVIBGICILVWVWVNIHGFDPDPSHGGFLRGAHYFRVAVALAVAAIPGGLPAVVTTCLALGTRMARLNAIVRSLPSVETLGGTTVICSDRTGTLTT : 345
A       : MGSIRDAMLRTEDEATPLRKLDEFGTFLLAKVIBGICILVWVWVNIHGFDPDPSHGGFLRGAHYFRVAVALAVAAIPGGLPAVVTTCLALGTRMARLNAIVRSLPSVETLGGTTVICSDRTGTLTT : 345

Predicted : ILCDFGQTLINNMMSVSKVCCVRSVHQRPITDEYSISGTTFAPOGFIYDASDQLLEFPQSPCLLHIA MCSALCNESTLQYNPDRFSTEKIGESTEVALRVLVEVGLGFGDQNF : 460
A       : ILCDFGQTLINNMMSVSKVCCVRSVHQRPITDEYSISGTTFAPOGFIYDASDQLLEFPQSPCLLHIA MCSALCNESTLQYNPDRFSTEKIGESTEVALRVLVEVGLGFGDQNF : 460

Predicted : SALNMLTMEHASTCNHYNEQFRKISVLPDRDRNMSVLCSEHQEINFQSAPEFVMAKCTHILCNCQGSVFLTMDIRNELEAFQSPAGDTRLCALALAKRNFEGQQSL : 575
A       : SALNMLTMEHASTCNHYNEQFRKISVLPDRDRNMSVLCSEHQEINFQSAPEFVMAKCTHILCNCQGSVFLTMDIRNELEAFQSPAGDTRLCALALAKRNFEGQQSL : 575

Predicted : SYDCEANLTFIQLVQMLLQFVDEVHSAIHSCHSAGIRVIVVTCGRSTAESLCRQIGAFERLDDFTGYSTTASEFEGFLPFLERANALRNVLSRVEDSHRKMVLEALQSHNEV : 690
A       : SYDCEANLTFIQLVQMLLQFVDEVHSAIHSCHSAGIRVIVVTCGRSTAESLCRQIGAFERLDDFTGYSTTASEFEGFLPFLERANALRNVLSRVEDSHRKMVLEALQSHNEV : 690

Predicted : RMTGGGVNDAFALKKADIGIANRSGTAVARASDMVLADDFATVAAVAEGRAIYNNTQFIPYMIERNIGEVVCLFPAAYLGMFQTLVPPVCLLWMLVTCGLDVAIUSWEPD : 805
A       : RMTGGGVNDAFALKKADIGIANRSGTAVARASDMVLADDFATVAAVAEGRAIYNNTQFIPYMIERNIGEVVCLFPAAYLGMFQTLVPPVCLLWMLVTCGLDVAIUSWEPD : 805

Predicted : GNIMAVKPRNVNEAVVSGNLFYRYLVIGATVGLATLQGFVWVYVSEDDGDFLPSSELVNFDCSTRTQTSYPCSIFFEDRNPSTVUMTIVAVVRYHALNLSERQSLLVIRFNS : 920
A       : GNIMAVKPRNVNEAVVSGNLFYRYLVIGATVGLATLQGFVWVYVSEDDGDFLPSSELVNFDCSTRTQTSYPCSIFFEDRNPSTVUMTIVAVVRYHALNLSERQSLLVIRFNS : 919

Predicted : RLVGSSIIIMLISAVAVIETPLSALFSVSPPLTLANETTYVLSPTVILIEVLIFFSRPQQLFPLRLNFRMLPNEARDM : 1002
A       : LSLDFPPHIAAFWTHNEMCYTSGNSWFSYLDWYTWISERSGNSFLTQWLSCLYLVWCITICIFEMIFGTSLSAVSFTHGVTYGLLQGSQCFYQSCINWPOQLF : 1025

Predicted : ----- : -
A       : SQCLRHWLSGKLFPTYPSSIFSNRQVNFQDEELKGVGFLYCYGCAKCFLEKPEIIX : 1079
    
```

Fig. 4: The introduction of PMC due to IR in *T.aestivum* ECA3 gene caused disruption of the protein structure resulting in loss of two important TM domains (TM9 and 10). Where A= pENTRTaECA3' with retained intron, Highlighted blue= TM domains, Highlighted red= Key motifs.

```

Predicted : LARLNGETGLTAFLEPSVIFMILAANAAGVITETNAERALEELRAYQADVATVLRNGCFSILPATELVPGDIVVGVGCKVPAIDMRMVEMLSHQLRVDQAILTGESCSVAKELH : 115
A : LARLNGETGLTAFLEPSVIFMILAANAAGVITETNAERALEELRAYQADVATVLRNGCFSILPATELVPGDIVVGVGCKVPAIDMRMVEMLSHQLRVDQAILTGESCSVAKELH : 115
B : LARLNGETGLTAFLEPSVIFMILAANAAGVITETNAERALEELRAYQADVATVLRNGCFSILPATELVPGDIVVGVGCKVPAIDMRMVEMLSHQLRVDQAILTGESCSVAKELH : 115

Predicted : STSAMNAVYQDKTNILFSGTVVWVAGRARAVVIGVGSNTAMGSIKRDAMLTEDEATPLKKLDEFGTFLAKVIAGICILWVVVNIHGFDPSPHGGFLRGAIHFFVAVALAVAAIH : 230
A : STSAMNAVYQDKTNILFSGTVVWVAGRARAVVIGVGSNTAMGSIKRDAMLTEDEATPLKKLDEFGTFLAKVIAGICILWVVVNIHGFDPSPHGGFLRGAIHFFVAVALAVAAIH : 230
B : STSAMNAVYQDKTNILFSGTVVWVAGRARAVVIGVGSNTAMGSIKRDAMLTEDEATPLKKLDEFGTFLAKVIAGICILWVVVNIHGFDPSPHGGFLRGAIHFFVAVALAVAAIH : 230

Predicted : EGLPAVVTTCALGTRMARLNAIVRSLPSVETLGCITVIOCDRTGTLFTNMSVSKVCVVRSVQRPITDEYSISGTTFAPDGFIVDASELQLEFPQSPCLLHIAMCSALCNE : 345
A : EGLPAVVTTCALGTRMARLNAIVRSLPSVETLGCITVIOCDRTGTLFTNMSVSKVCVVRSVQRPITDEYSISGTTFAPDGFIVDASELQLEFPQSPCLLHIAMCSALCNE : 345
B : EGLPAVVTTCALGTRMARLNAIVRSLPSVETLGCITVIOCDRTGTLFTNMSVSKVCVVRSVQRPITDEYSISGTTFAPDGFIVDASELQLEFPQSPCLLHIAMCSALCNE : 345

Predicted : STLQYNPDRKSYEKIGESTEVALRVLVEKVGLPGFDSMPSALNMLTRHERASYCNIHWENQFRKISVLDVSRDRKMSVLCSSKQEIFMFSRGAPESVMARCTHILCNDGSSVF : 460
A : STLQYNPDRKSYEKIGESTEVALRVLVEKVGLPGFDSMPSALNMLTRHERASYCNIHWENQFRKISVLDVSRDRKMSVLCSSKQEIFMFSRGAPESVMARCTHILCNDGSSVF : 460
B : STLQYNPDRKSYEKIGESTEVALRVLVEKVGLPGFDSMPSALNMLTRHERASYCNIHWENQFRKISVLDVSRDRKMSVLCSSKQEIFMFSRGAPESVMARCTHILCNDGSSVF : 460

Predicted : LTMDIRNELEARFQSFAGKDTLRCLALALRRNPEGQQLSYDDEANLTFIQLVGMDDPPRDEVHSAIHSKMSAGIRVIVVTGDNFSTAESLCRQIGAFHELDLDFGYSYTAEEF : 575
A : LTMDIRNELEARFQSFAGKDTLRCLALALRRNPEGQQLSYDDEANLTFIQLVGMDDPPRDEVHSAIHSKMSAGIRVIVVTGDNFSTAESLCRQIGAFHELDLDFGYSYTAEEF : 575
B : LTMDIRNELEARFQSFAGKDTLRCLALALRRNPEGQQLSYDDEANLTFIQLVGMDDPPRDEVHSAIHSKMSAGIRVIVVTGDNFSTAESLCRQIGAFHELDLDFGYSYTAEEF : 575

Predicted : GLPPLERANALRRMVLFSRVEPSHKMLVEALQSHNEVWAMTGDGVNDAPALKKADIGIANGSGTAVAKSASDMVLADDFATIVAARVAEGRAYNNTRQFIRYMISNIGEVVC : 690
A : GLPPLERANALRRMVLFSRVEPSHKMLVEALQSHNEVWAMTGDGVNDAPALKKADIGIANGSGTAVAKSASDMVLADDFATIVAARVAEGRAYNNTRQFIRYMISNIGEVVC : 690
B : GLPPLERANALRRMVLFSRVEPSHKMLVEALQSHNEVWAMTGDGVNDAPALKKADIGIANGSGTAVAKSASDMVLADDFATIVAARVAEGRAYNNTRQFIRYMISNIGEVVC : 690

Predicted : IFVAAVLGMPTLVVQQLLWVNLVTDGLPATAIGFNFPGDGNIMAVKPRVNEAVVSGWLFRRVLYIGAYVGLATIAGFVWVWVYSEDGPRLPYSELVNFDFSCSTRQTSYPCSIFF : 805
A : IFVAAVLGMPTLVVQQLLWVNLVTDGLPATAIGFNFPGDGNIMAVKPRVNEAVVSGWLFRRVLYIGAYVGLATIAGFVWVWVYSEDGPRLPYSELVNFDFSCSTRQTSYPCSIFF : 805
B : IFVAAVLGMPTLVVQQLLWVNLVTDGLPATAIGFNFPGDGNIMAVKPRVNEAVVSGWLFRRVLYIGAYVGLATIAGFVWVWVYSEDGPRLPYSELVNFDFSCSTRQTSYPCSIFF : 805

Predicted : DRBPSVSMTVLVVWVEMFNALNLSNQSLNLIHFNWNLVWVSIILTMLEHVAVLYTEPLSALRVSSELTLAENKVVLLSFPVILIDEVLKFFSRPQCMSPFLRLWRSEMLH : 920
A : DRBPSVSMTVLVVWVEMFNALNLSNQSLNLIHFNWNLVWVSIILTMLEHVAVLYTEPLSALRVSSELTLAENKVVLLSFPVILIDEVLKFFSRPQCMSPFLRLWRSEMLH : 920
B : DRBPSVSMTVLVVWVEMFNALNLSNQSLNLIHFNWNLVWVSIILTMLEHVAVLYTEPLSALRVSSELTLAENKVVLLSFPVILIDEVLKFFSRPQCMSPFLRLWRSEMLH : 920

Predicted : *----- : 926
A : *----- : 926
B : TSLASVSTHGVTYGLLQSFQCFYQSCIRNPOQLFSQCLELWLSGKLFYTPSLFISMRCNFSQEDLWGVFLYGYGDARCLKHPETIX : 1003
    
```

Fig. 3: A= pENTRTaECA3 3' sequence without intron, B= pENTRTaECA3 3' sequence with intron. The reading frame is disrupted in B and a premature stop codon appears in it due to intron retention. *= premature stop codon, *-stop codon

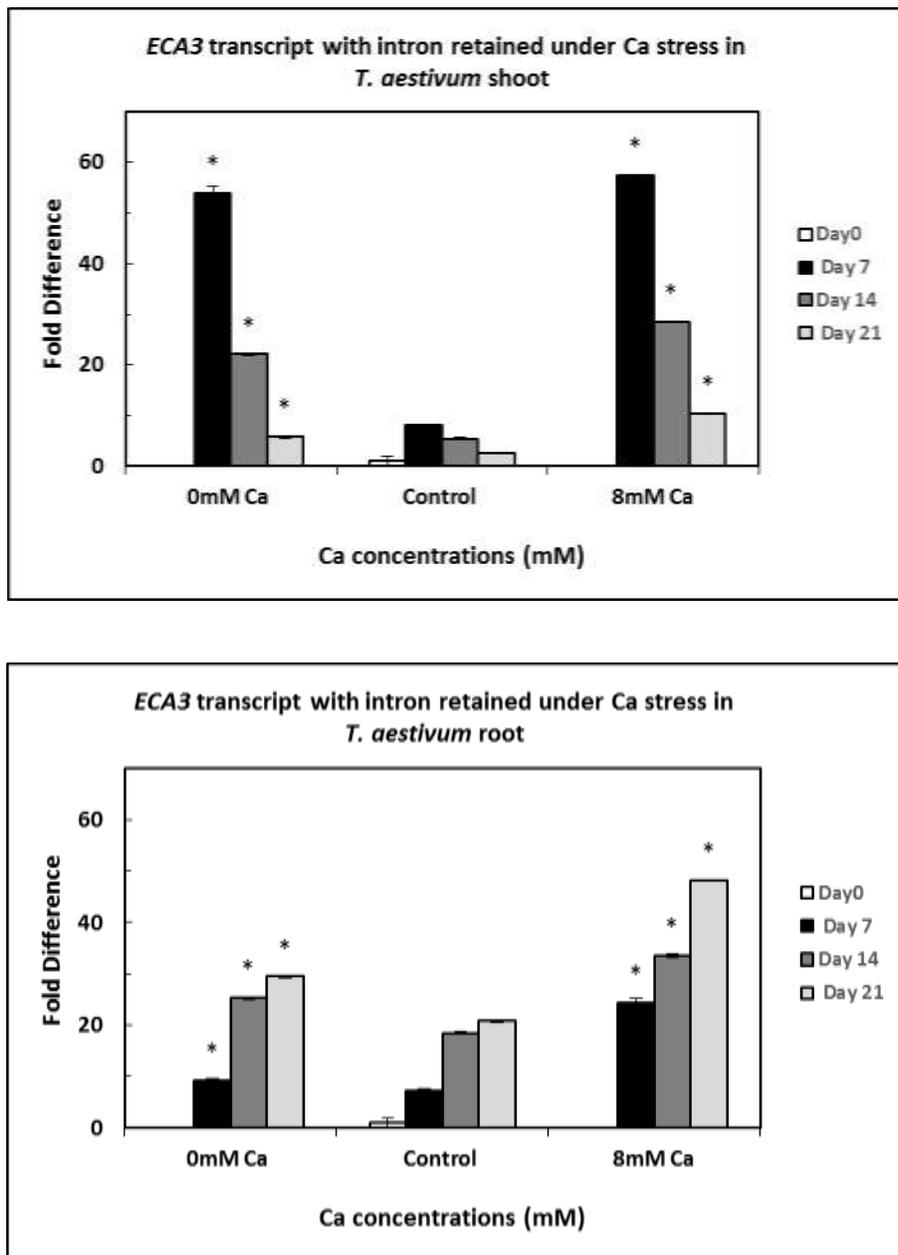


Fig. 5: qRT-PCR data indicating the retention of intron in *ECA3* in *Triticum aestivum* shoots and roots under calcium deficiency/toxicity and control. The fold difference was evaluated relative to baseline D0 control. The significant difference in expression from control values (student's t-test) are indicated by * where, $P < 0.05$

VI. REFERENCES

- [1] Alexandrov NN, Troukhan ME, Brover VV, Tatarinova T, Flavell RB, Feldmann KA. 2006. Features of Arabidopsis genes and genome discovered using full-length cDNAs. *Plant Mol. Biol.* 60:69–85. <https://doi.org/10.1007/s11103-005-2564-9>
- [2] Barbazuk WB, Yan F and McGinnis KM. 2008. Genome-wide analyses of alternative splicing in plants: Opportunities and challenges. *Genome Res.* 2008. 18:1381-1392 <https://doi.org/10.1101/gr.053678.106>
- [3] Baxter I, Tchiew J, Sussman MR, Boutry M, Palmgren MG, Gribskov M, Harper JF, Axelsen KB. 2003. Genomic comparison of P-type ATPase ion pumps in Arabidopsis and rice. *Plant Physiology* 132: 618–628 <https://doi.org/10.1104/pp.103.021923>
- [4] Black, DL. 2003. "Mechanisms of alternative pre-messenger RNA splicing". *Ann. Rev of Biochem.* 72 (1): 291–336. <https://doi.org/10.15242/IICBE.C1216009>
- [5] Brown JW, Simpson CG, Thow G, Clark GP, Jennings SN, Medina-Escobar, N, Haupt S, Chapman SC and Oparka, KJ. 2002. Splicing signals and factors in plant intron removal. *Biochem. Soc. Trans.* 30, 146–149. <https://doi.org/10.1042/bst0300146>
- [6] Chen FC, Wang SS, Chaw, SM, Huang YT and Chuang TJ. 2007. Plant gene and alternatively spliced variant annotator: a plant genome annotation pipeline from rice gene and alternatively spliced variant identification with cross-species expressed sequence tag conservation from seven plant species. *Plant Phys.* 143, 1086–1095. <https://doi.org/10.1104/pp.106.092460>
- [7] Christensen AH, Sharrock RA, Quail PH. 1992. Maize polyubiquitin genes: Structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol. Biol.* 18:675–89. <https://doi.org/10.1007/BF00020010>

- [8] Dinesh-Kumar SP, Baker BJ. 2000. Alternatively spliced N resistance gene transcripts: their possible role in tobacco mosaic virus resistance. *Proc. Natl. Acad. Sci. USA* 97:1908–13. <https://doi.org/10.1073/pnas.020367497>
- [9] Filichkin S, Henry DP, Molly M and Todd CM. 2015. Alternative splicing in plants: directing traffic at the crossroads of adaptation and environmental stress *Current Opinion in Plant Biology*, 24:125–135 <https://doi.org/10.1016/j.pbi.2015.02.008>
- [10] Gebauer F, Merendino L, Hentze MW and Valcarcel J. 1998. The Drosophila splicing regulator sex-lethal directly inhibits translation of male-specific-lethal 2 mRNA. *RNA*, 4, 142–150.
- [11] Graveley BR. 2001. Alternative splicing: increasing diversity in the proteomic world. *Trends Genet.* 17, 100–107 [https://doi.org/10.1016/S0168-9525\(00\)02176-4](https://doi.org/10.1016/S0168-9525(00)02176-4)
- [12] Golovkin M and Reddy AS. 1996. Structure and expression of a plant U1 snRNP 70K gene: Alternative splicing of U1 snRNP 70K pre-mRNAs produces two different transcripts. *Plant Cell* <https://doi.org/10.1105/tpc.8.8.1421>
- [13] 8: 1421–1435.
- [14] Haas BJ, Volfovsky, N, Town CD, Troukhan M, Alexandrov N, Feldmann KA, Flavell RB, White, O and Salzberg, SL. 2002. Full-length messenger RNA sequences greatly improve genome annotation. *Genome Biol.* 3, 0029.1–0029.12.
- [15] Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RKJ, et al. 2003. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* 31:5654–66. <https://doi.org/10.1093/nar/gkg770>
- [16] Hopf N, Plesofsky-Vig N, Brambl R. 1992. The heat shock response of pollen and other tissues of maize. *Plant Mol. Biol.* 19:623–30 <https://doi.org/10.1007/BF00026788>
- [17] Hori K, Watanabe Y. 2005. UPF3 suppresses aberrant spliced mRNA in Arabidopsis. *Plant J.* 43:530–40. <https://doi.org/10.1111/j.1365-313X.2005.02473.x>
- [18] Jordan T, Schornack S, Lahaye T. 2002. Alternative splicing of transcripts encoding Toll-like plant resistance proteins—what's the functional relevance to innate immunity? *Trends Plant Sci.* 7:392–98. [https://doi.org/10.1016/S1360-1385\(02\)02311-7](https://doi.org/10.1016/S1360-1385(02)02311-7)
- [19] KAMRUL HKM, YADAV S, AKHTER BANU MS, TRIVEDI DK, TUTEJA N. 2013. GENOME-WIDE ANALYSIS OF PLANT-TYPE II Ca²⁺ ATPASES GENE FAMILY FROM RICE AND ARABIDOPSIS: POTENTIAL ROLE IN ABIOTIC STRESSES. *PLANT PHYSIOL BIOCHEM.*;65:32-47 <https://doi.org/10.1016/j.plaphy.2013.01.002>
- [20] Lewis BP, Green RE and Brenner SE. 2003. Evidence for the wide spread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc. Natl. Acad. Sci. U.S.A.* 7, 189–192. <https://doi.org/10.1073/pnas.0136770100>
- [21] Lida K, Seki M, Sakurai T, Satou M, Akiyama K, et al. 2004. Genome-wide analysis of alternative pre-mRNA splicing in Arabidopsis thaliana based on full-length cDNA sequences. *Nucleic Acids Res.* 32:5096–103. <https://doi.org/10.1093/nar/gkh845>
- [22] Lombnaes P, Sing BR. 2003. Varietal tolerance to zinc deficiency in wheat and barley grown in chelator buffered nutrient solution and its effect on uptake of Cu, Fe and Mn. *J. Plant Nutr. Soil. Sci.* 166, 76-83. <https://doi.org/10.1002/jpln.200390015>
- [23] Lorkovic ZJ, Wiczorek Kirk DA, Lambermon MH and Filipowicz W. 2000. Pre-mRNA splicing in higher plants. *Trends Plant Sci.* 5, 160–167. [https://doi.org/10.1016/S1360-1385\(00\)01595-8](https://doi.org/10.1016/S1360-1385(00)01595-8)
- [24] Lorkovic, ZJ and Barta A. 2002. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. *Nucl. Acids Res.* 30, 623–635. <https://doi.org/10.1093/nar/30.3.623>
- [25] Luehrsen KR, Taha S, Walbot V. 1994. Nuclear pre-mRNA processing in higher plants. *NucleicAcids Res.* 47:149–93.
- [26] Lu T, Lu G, Fan D, Zhu C, Li W, Zhao Q, Feng Q, Zhao Y, Guo Y, Li W, Huang X, Han B. Function annotation of the rice transcriptome at single-nucleotide resolution by RNA-seq. *Genome Res.* 2010 Sep; 20(9):1238-49. <https://doi.org/10.1101/gr.106120.110>
- [27] Mastrangelo AM, Belloni S, Barilli S, Ruperti B, Di Fonzo N, Stanca AM, Cattivelli L. 2005. Low temperature promotes intron retention in two e-cor genes of durum wheat, 221(5):705-15
- [28] Maquat, L E. 2004. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat. Rev. Mol. Cell Biol.* 5, 89–99. <https://doi.org/10.1038/nrm1310>
- [29] Marrs KA, Walbot V. 1997. Expression and RNA splicing of the maize glutathione S-transferase Bronze2 gene is regulated by cadmium and other stresses. *Plant Physiol.* 113:93–102. <https://doi.org/10.1104/pp.113.1.93>
- [30] Mills R.F., Doherty M.L., Lopez-Marque's R.L., Weimar T., Dupre P., Palmgren M.G., Pittman J.K., Williams L.E. (2008) ECA3, a Golgi-localized P_{2A}-type ATPase, plays a crucial role in manganese nutrition in Arabidopsis. *Plant Physiology*, 146, 116–1128. <https://doi.org/10.1104/pp.107.110817>
- [31] Nagasaki H, Arita M, Nishizawa T, Suwa M, Gotoh O. 2005. Species-specific variation of alternative splicing and transcriptional initiation in six eukaryotes. *Gene* 364:53–62 <https://doi.org/10.1016/j.gene.2005.07.027>
- [32] 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.* 39:877–85 <https://doi.org/10.1111/j.1365-313X.2004.02172.x>
- [33] Palusa SG, Ali GS, Reddy ASN. 2006. Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins and its regulation by hormones and stresses. *Plant J.* In press
- [34] Paterno, G.D., Ding, Z., Lew, Y.Y., Nash, G.W., Mercer, F.C. and Gillespie, L.L. 2002. Genomic organization of the human mi-er1 gene and characterization of alternatively spliced isoforms: regulated use of a facultative intron determines subcellular localization. *Gene*, 295, 79–88. [https://doi.org/10.1016/S0378-1119\(02\)00823-5](https://doi.org/10.1016/S0378-1119(02)00823-5)
- [35] Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time PCR. *Nucleic Acids Res* 29 2002–2007 <https://doi.org/10.1093/nar/29.9.e45>
- [36] Rachel Brenchley, Manuel Spannagl, Matthias Pfeifer, Gary L. A. Barker, Rosalinda D'Amore, Alexandra M. Allen, Neil McKenzie, Melissa Kramer, Arnaud Kerhornou, Dan Bolser, Suzanne Kay, Darren Waite, Martin Trick, Ian Bancroft, Yong Gu, Naxin Huo, Ming-Cheng Luo, Sunish Sehgal, Bikram Gill, Sharyar Kianian, Olin Anderson, Paul Kersey, Jan Dvorak, W. Richard McCombie, Anthony Hall et al. 2012. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*:491,705–710 <https://doi.org/10.1038/nature11650>
- [37] Reddy ASN. 2001. Nuclear pre-mRNA splicing in plants. *Crit. Rev. Plant Sci.* 20:523–71 [https://doi.org/10.1016/S0735-2689\(01\)80004-6](https://doi.org/10.1016/S0735-2689(01)80004-6)
- [38] Simpson GG and Filipowicz, W. 1996. Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organization of the spliceosomal machinery. *Plant Mol. Biol.* 32, 1–41. <https://doi.org/10.1007/BF00039375>
- [39] Stamm S, Ben-Ari S, Rafalska I, Tang Y, Zhang Z, Toiber D, Thanaraj TA, and Soreq H. 2005. Function of alternative splicing. *Gene* 344: 1–20. <https://doi.org/10.1007/BF00039375>
- [40] Tilgner H, Knowles DG, Knowles CA, Davis S, Chakraborty S, Djebali J, Curado I M, Snyder TR, Gingeras and Guigo R. 2012. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res.* 2012. 22:1616-1625 <https://doi.org/10.1101/gr.134445.111>
- [41] Tolstrup N, Rouze P and Brunak S. 1997. A branch point consensus from Arabidopsis found by non-circular analysis allows for better prediction of acceptor sites. *Nucl. Acids Res.* 25, 3159–3163. <https://doi.org/10.1093/nar/25.15.3159>
- [42] Xiao YL, Smith SR, Ishmael N, Redman JC, Kumar N, Monaghan EL, Ayele M, Haas BJ, Wu HC, and Town CD. 2005. Analysis of the cDNAs of hypothetical genes on Arabidopsis chromosome 2 reveals numerous transcript variants. *Plant Physiol.* 139: 1323–1337. <https://doi.org/10.1104/pp.105.063479>
- [43] Young-Ju K, Mi-Jeong P, Sang-Gyu K, Ian T B and Chung-Mo P. 2014. Alternative splicing and nonsense-mediated decay of circadian clock genes under environmental stress conditions in Arabidopsis. Zhang XC,

- Gassmann W. 2003. RPS4-mediated disease resistance requires the combined presence of RPS4 transcripts with full-length and truncated open reading frames. *Plant Cell* 15:2333–42.
- [44] Zhang G, Guo G, Hu X, Zhang Y, Li Q, Li R, Zhuang R, Lu Z, He Z, Fang X, Chen L, Tian W, Tao Y, Kristiansen K, Zhang X, Li S, Yang H, Wang J, Wang J. Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome.
- [45] *Genome Res.* 2010 May; 20(5):646-54.
<https://doi.org/10.1101/gr.100677.109>
- [46] ZHANG C, YANG H, AND YANG H. 2015. EVOLUTIONARY CHARACTER OF ALTERNATIVE SPLICING IN PLANTS. *BIOINFORM BIOL INSIGHTS*. 9: 47–52.
- [47] Zhu W, Schlueter SD and Brendel V. 2003. Refined annotation of the Arabidopsis genome by complete expressed sequence tag mapping. *Plant Physiol.* 132, 469–484.
<https://doi.org/10.1104/pp.102.018101>