
Cloning and Functional characterization of seed-specific *LEC1A* promoter from peanut (*Arachis hypogaea* L.)

Guiying Tang¹, Pingli Xu¹, Pengxiang Li^{1,2}, Jieqiong Zhu^{1,2}, Guangxia Chen³,
Lei Shan^{1,2*}, Shubo Wan^{1,2*}

¹ Bio-Tech Research Center, Shandong Academy of Agricultural Sciences / Shandong
Provincial Key Laboratory of Crop Genetic Improvement, Ecology and Physiology,
Jinan, Shandong, China

² College of Life Science, Shandong Normal University, Jinan, Shandong, China

³ Shandong Academy of Grape, Jinan, Shandong, China

*** Corresponding Author**

E-mail: shlei1025@sina.com (LS)

wansb@saas.ac.cn (SW)

Abstract

LEAFY COTYLEDON1 (LEC1) is a HAP3 subunit of CCAAT-binding transcription factor,

which controls several aspects of embryo and postembryo development, including embryo morphogenesis, storage reserve accumulation and skotomorphogenesis. Herein, using the method of chromosomal walking, a 2707bp upstream sequence from the ATG initiation codon site of *AhLEC1A* which is a homolog of Arabidopsis *LEC1* was isolated in peanut. Its transcriptional start site confirmed by 5' RACE was located at 82 nt from 5' upstream of ATG. The bioinformatics analysis revealed that there existed many tissue-specific elements and light responsive motifs in its promoter. To identify the functional region of the *AhLEC1A* promoter, seven plant expression vectors expressing the *GUS* (β -glucuronidase) gene, driven by 5' terminal series deleted fragments of *AhLEC1A* promoter, were constructed and transformed into Arabidopsis. Results of *GUS* histochemical staining showed that the regulatory region containing 82bp of 5' UTR and 2228bp promoter could facilitate *GUS* to express preferentially in the embryos at different development periods of Arabidopsis. Taken together, it was inferred that the expression of *AhLEC1A* during seed development of peanut might be controlled positively by several seed-specific regulatory elements, as well as negatively by some other regulatory elements inhibiting its expression in other organs. Moreover, the *GUS* expression pattern of transgenic seedlings in darkness and in light was relevant to the light-responsive elements scattered in *AhLEC1A* promoter segment, implying that these light-responsive elements harbored in the *AhLEC1A* promoter regulate skotomorphogenesis of peanut seeds, and *AhLEC1A* expression was inhibited after the germinated seedlings were transferred from darkness to light.

Introduction

Seed development is a complex procedure of the flowering plant in life cycle, which can

conceptually be divided into two distinct phases: embryo morphogenesis and seed maturation[1, 2]. Lots of genes highly and specifically expressed in different developmental processes highlight the importance of transcriptional regulations for proper seed formation[3]. The Arabidopsis *LAFL* genes coding for LEAFY COTYLEDON1 (LEC1), ABSCISIC ACID INSENSITIVE3 (ABI3), FUSCA3 (FUS3), and LEC2, respectively form a network involved in regulating seeds development[4-7].

LEC1 is a central regulator controlling embryogenesis and seed maturation in *Arabidopsis thaliana* [8-10]. Loss of *LEC1* function causes a pleiotropic phenotype, including cotyledon with trichomes, distortions in suspensor-cell specification, defects in storage protein and lipid accumulation, embryo desiccation-intolerance in developing seeds, and leaf primordia initiation. [8, 11, 12]. The expression of many genes involved with maturation processes are downregulated in *lec1* mutant seeds [9, 13, 14]. Moreover, the role of *LEC1* was also demonstrated by analyzing its gain-of-function mutant. For example, overexpression of *LEC1* in developing seeds elevates the contents of seed storage macromolecule, as well, upregulates the key genes involved in storage protein and lipid accumulation in a number of plant species[15-19]. Genome-wide analysis of *LEC1* occupancy and interactome indicated that LEC1 regulate many genes involved in embryo development[13, 14].

LEC1 also participates in regulating post-embryonic growth during the developmental transition from germination seeds to seedlings. In higher plants, rapid elongation of hypocotyl in germinating seeds can be induced by darkness. However, *lec1* mutant causes a reduced capability for hypocotyl elongation and apical hook formation[13, 20]. Additionally, the expression of *LEC1* was also detectable in etiolated seedlings [21, 22], and the phenotype of longer hypocotyls and higher expression levels of the genes involved in etiolated growth were observed in LEC1-

overexpression plants[23]. A *LEC1* gain-of-function mutant turnip (*tnp*), displayed the partial de-etiolation at dark-grown condition, including inhibited hypocotyls elongation, and activated SAM (shoots apical meristem)[24].

Despite of such comprehensive knowledge about *LEC1* in the model plant *Arabidopsis*, much less is known about the expression patterns and functions in other higher plants. Here, the 2707bp 5' flanking region of peanut *AhLEC1A* was isolated, and the GUS expression profiles driven by a series of deletion in its 5' flanking region were characterized in transgenic *Arabidopsis*. The results showed that the regulatory region containing 82bp of 5' UTR and 2228bp promoter could specifically regulate *AhLEC1A* expressing in developing seeds. Thus, the *AhLEC1A* promoter could be utilized as a seed-preferential promoter for plant genetic engineering.

Materials and Methods

Plant materials and growth conditions

Peanut (*Arachis hypogaea* L. cv. Luhua 14) seeds, *Arabidopsis thaliana* L. (Ecotypes Col) seeds, *Escherichia coli* strain DH5 α , pCambia3301 plasmid and *Agrobacterium tumefaciens* strain GV3101 used in the present study were maintained at our laboratory. Peanut plants were grown in the experimental field of Shandong Academy of Agricultural Sciences. Developing seeds were collected and kept in -80°C refrigerator for isolation of total RNA.

Cloning of the 5' flanking region of *AhLEC1A*

The peanut genomic DNA was isolated from Luhua 14 leaves using CTAB method[25]. Genome walking was performed to isolate the 5' flanking regulatory region. According to the BD Genome Walker Universal Kit (Clontech, USA) manufacturer's instructions, each of 2.5 μ g genomic

DNA was digested with four restriction enzyme *DraI*, *EcoRV*, *PvuII*, and *StuI* respectively; and then the digested samples were connected with the BD Genome-Walker adaptor resulting in the library containing digestions by *DraI*, *EcoRV*, *PvuII*, and *StuI* (LD, LE, LP, and LS). Based on the sequence of *AhLEC1A* genomic DNA, two nested gene-specific primers (GSP), LEC1AGSP1-2 and LEC1AGSP2-2, were designed. The first round of PCR reaction was done in a 25 μ L reaction system using an AP1 provided by Kit and LEC1A GSP1-2 as 5' terminal and 3' terminus primer, and 1 μ L DNA of each library as template. The nested PCR reaction was also performed using the same volume and conditions with primers AP2 and LEC1AGSP2-2, and 1 μ L of the 10-fold diluted primary PCR products as template. The specific PCR fragments from the second round reaction were isolated and inserted into the vector pEASY-T3. The recombinants harboring the target gene were validated by two-way sequencing using ABI3730 model DNA sequencer. The primer and adaptor sequences of this assay were listed in Table 1.

Table 1. List of primers used in the study.

Serial No.	Name	Sequence(from 5' to 3')	Feature
1	BD GenomeWalker adaptor	GTAATACGACTCACTATAGGGCACGCGTGGTTCGACGGCCC GGGCTGGT	No.1-5 used for the amplification of 5' flanking sequence of <i>AhLEC1A</i>
2	LEC1AGSP1-2	TCAACCCAGAGGTAGTGGTAGGAAGG	
3	LEC1AGSP2-2	TGATAACCGTGAAAGCCTCCTCCAGT	
4	AP1	GTAATACGACTCACTATAGGGC	
5	AP2	ACTATAGGGCACGCGTGGT	
6	GeneRacer RNA Oligo	CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAG UAGAAA	No.6-10 used for the localization of the transcriptional start site of <i>AhLEC1A</i> gene
7	TSS LECGSP1-1	TCTTTTGCCTCGTCGGAGATTTTAGC	
8	TSS <i>AhLEC</i> GSP2-2	TGATAACCGTGAAAGCCTCCTCCAGT	
9	5' GeneRacer Primer	CGACTGGAGCACGAGGACACTGA	
10	5' Nested Primer	GGACACTGACATGGACTGAAGGAGTA	
11	P1	<u>CCATGGGT</u> GTGAAGAAAGATGCAGTG	No.11-18 used for the amplification of different length promoter fragments
12	P2	<u>AAGCTT</u> CATTAGGGTCAAAAGAGTG	
13	P3	<u>AAGCTT</u> CTCTTGCAATAAATGTTGG	
14	P4	<u>AAGCTT</u> CCCGTTAAAAAAATAATAAG	
15	P5	<u>AAGCTT</u> GTAAATTTTGGATAGCTTG	
16	P6	<u>AAGCTT</u> TACATGGCACGCCTCATATC	
17	P7	<u>AAGCTT</u> AGATCGAACTAATTTAAG	
18	P8	<u>AAGCTT</u> AAAAAGTTGAACATTTTATATAG	

Precise identification of transcription start site in *AhLEC1A*

The transcription start site of *AhLEC1A* gene was identified by 5' RACE (rapid amplification of cDNA ends) using a 5' RACE kit (Invitrogen GeneRacer™ Kit) following the instructions provided by the manufacturer. Total RNA was extracted from the developing seeds of peanut Luhua 14 using the improved CTAB method[26]. The ds-cDNA was synthesized using the full-length mRNA with RNA Oligo as template. The ds-cDNA was cloned into vector pCR4-TOPO to establish the full-length cDNA library. According to the cDNA sequence of *AhLEC1A*, two 3' terminus gene-specific primers TSS LEC1AGSP1-1 and TSS LEC1AGSP2-2 were designed, for use in the nested PCR reaction. The 5' terminus general primer for two rounds of PCR were GeneRacer™ Primer and 5' Nested Primer. 1 μL of the full-length cDNA library as got previously and a 50-fold dilution of the primary PCR product was used respectively as the template of the two rounds of PCR. The nested PCR products were collected and sequenced by ABI3730 model DNA sequencer. The primer sequences used in the assay were listed in Table 1.

In-silico analysis of the *AhLEC1A* promoter for cis-regulatory elements

The cis-elements of the 5' flanking region of *AhLEC1A* gene were analyzed using PLACE (<http://www.dna.affrc.go.jp/PLACE>) and Plant Cis-Acting Regulatory Elements (Plant CARE) (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>).

Plasmid construction and Arabidopsis transformation

A series of 5' -truncated promoter sequences were obtained by PCR using a single reverse primer localized in 5' UTR of *AhLEC1A*, and different forward primers situated in the different sites of the *AhLEC1A* promoter (The primer sequences were listed in Table 1). To construct the vector,

the appropriate restriction sites were introduced into the PCR-amplified promoter (*HindIII* at the 5' end; *NcoI* at the 3' end). The PCR-amplified promoter was then inserted into *HindIII*/*NcoI*-digested pCAMBIA3301, replacing the cauliflower mosaic virus (CaMV) 35S promoter, producing seven deletion constructs containing various fragments (-2228 ~ +82, Q7; -1254 ~ +82, Q6; -935 ~ +82, Q5; -721 ~ +82, Q4; -617 ~ +82, Q3; -354 ~ +82, Q2; -105 ~ +82, Q1).

The constructs including Q1-Q7 and the control pCAMBIA3301 was introduced into *Agrobacterium tumefaciens* strain GV3101 using a freeze-thaw method. Transgenic Arabidopsis plants were generated by the floral dip method. Seeds of the T₀-T₂ generations were germinated on 1/2MS agar medium containing 10µg/L Basta. The identified transgenic plants were transferred to soil under 120 µmol·m⁻²·s⁻¹ light in a growth room at a temperature between 22°C and 25°C. The all Arabidopsis plants grew under a 16h light/8h dark photoperiod, and 65% relative humidity.

Histochemical GUS staining

The GUS assay was performed as described by Jefferson [27]. For each *AhLEC1A* promoter-GUS construct, five Arabidopsis transformants were selected for histochemical staining. The roots and leaves at the 4-leaf stage, stems at the bolting stage, flowers, immature embryos of 6-10 days after pollination and 3-5day etiolated and de-etiolated seedlings in transgenic T₂ lines were incubated in GUS assay buffer with 50mM sodium phosphate(7.0), 0.5mM K₃Fe(CN)₆, 0.5mM K₄Fe(CN)₆·3H₂O, 0.5% Triton X-100, and 1mM X-Gluc at 37°C overnight and then cleared with 70% ethanol. The samples were examined by stereomicroscopy.

Results

Isolation of the promoter of *AhLEC1A* and localization of TSS

The 2739bp DNA fragment was amplified by two rounds of PCR using the method of genome walking. Its sequence analysis found that this fragment includes 2707bp of 5' flanking region upstream from ATG and 32bp of coding sequence (Fig 1). In order to determine the transcription start site (TSS) of *AhLEC1A* gene, the nested 5' RACE was performed to amplify the 5'-end of its cDNA. The 140bp of cDNA fragment, including the 58bp of coding region started from ATG and 82bp 5' UTR, was isolated (Fig 2). Compared with the gDNA sequence of *AhLEC1A*, the sequence of 82bp 5' UTR was identical to the 5' upstream sequence of its gDNA, suggesting that the "A" located at the 82th nucleotid (nt) upstream from ATG is the TSS of *AhLEC1A* gene.

Fig1. PCR amplification of 5' flanking regulation regions of peanut *AhLEC1A* gene by chromosome walking. LD, LE, LP and LS represents the second amplification with different primary product as template respectively. The arrow indicates the targeted band for further cloning and sequencing.

Fig2. Localization of transcription start site of peanut *AhLEC1A* gene using 5' RACE. 1-The products of the first round PCR; 2-The products of the second round PCR.

Analysis of cis-regulatory elements in *AhLEC1A* promoter sequence

In silico analysis of 2707bp 5' flanking region revealed that a number of putative cis-elements were present in the 2625bp of promoter region and 82bp of 5' UTR (Fig 3). In detail, the basic promoter elements, TATA box (TATATAT) and CAAT box (CAAAT)), respectively placed at -36 ~ -30 nt and -143 ~ -148 nt. Many crucial elements required for embryo- or endosperm-specific expression and seed storage compounds accumulation scattered over the promoter, including two

SKn-1 motifs (GTCAT) at -84 ~ -80 nt and -475 ~ -479 nt, two CANBANAPA elements (CNAACAC) at -442 ~ -448 nt and -1597 ~ -1603 nt, and three binding sites for AGL15 (CWWWWWWWWG) at -1245 ~ -1236 nt, -2079 ~ -2070 nt and -2272 ~ -2263 nt. In addition, four DPBFCOREDCDC3 elements (ACACNNG), previously considered to involve in embryo-specific expression and also to respond to ABA were found at -111 ~ -117 nt, -445 ~ -451 nt, -1321 ~ -1327 nt and -1330 ~ -1324 nt. We have also detected many other regulatory elements for the accumulation of seed storage compounds and embryogenesis. For instance, eight EBOX BNNAPA (CANNTG), two 2S SEED PROT BANAPA (CAAACAC) and one SEF3 MOTIF GM (AACCCA). Besides, there were some elements associated with regulating in vegetative organ development on the promoter, such as mesophyll-specific element CACTFTPPCA1 (YACT), root-specific element ROOTMOTIFTAPOX1 (ATATT) and so on. There also exist some elements involved in light responsiveness including more than a dozen I BOX (GATAA) at -366 ~ -362 nt, -372 ~ -368 nt, -405 ~ -401 nt, -357 ~ -353 nt, -511 ~ -507 nt, -810 ~ -806 nt, -905 ~ -901 nt, -913 ~ -909 nt, -983 ~ -979 nt, -1793 ~ -1789 nt, -1916 ~ -1912 nt, and -1946 ~ -1942 nt, three -10PEHVPSBD (TATTCT) at -437 ~ -432 nt, -2584 ~ -2579 nt, and -2606 ~ -2601 nt, and one TCT-motif (TCTTAC) at -830 ~ -835 nt, etc., and some other regulatory elements controlling the chloroplast genes expression, like one GT1 MOTIF PSRBCS (KWGTGRWAAWRW) at -137 ~ -126 nt, and two etiolation-induced expression elements ACGT ATERD1 (ACGT) at -122 ~ -119 nt, and -1334 ~ -1337 nt. We also identified thirteen negative regulatory elements in the promoter, among which five WBOXATNPR1 elements located in -1255 ~ -2228 nt region (-1613 ~ -1616, -1649 ~ -1652, -1953 ~ -1956, -2144 ~ -2147, and -2156 ~ -2159nt), and four WRKY71OS elements densely distributed between the region of -2228 to -2625 nt.

Fig3. The sequence of 5' flanking regulation region of peanut *AhLEC1A* gene and some major elements harbored in this region. The letter “A” in box represents its transcription start site (TSS), and the putative regulatory elements are highlighted by underling or shading.

Functional analysis of the regulatory regions of the *AhLEC1A* promoter

To validate the role of the crucial regulatory region in *AhLEC1A* promoter, a series of GUS expression vectors (Q1~ Q7) (Fig 4), driven by different length of the promoters with truncated 5' terminal were established, and the GUS expression patterns in stable transgenic plants of *Arabidopsis* was investigated. In the histochemical assay, GUS expression was visualized specifically in the developing embryos of transgenic plants containing Q7 construct (including 2228bp promoter region and 82bp 5' UTR). Otherwise, the GUS staining was observed in all detected organs of transgenic plants carrying Q3, Q4, Q5, and Q6 construct (Fig 5). These four promoter segments are respectively 617bp, 721bp, 935bp and 1254bp in size with 5' terminal deletion of 1611bp ~ 974bp. It was suggested that there exist some key motifs in the promoter region between -2228bp and -1255bp, which related to inhibit the expression in the other organs except for the developing seeds. Moreover, the further deletional promoter fragment Q2 with 354bp drove the GUS to express only in embryos and rosette leaves. The shortest fragment Q1 containing 105bp promoter region and 82bp 5' UTR couldn't drive the GUS to express in any detected organs of transgenic *Arabidopsis* (Fig 5), implying that it might be caused by the deletion of the necessary component for gene expression.

Fig4. The vector diagram expressing GUS in plants driven by different length *AhLEC1A* promoters with 5' terminal deletion. Q1~Q7 indicates its promoters with different length. The rectangles in light and dark gray respectively show the promoter region upstream of the TSS, and the 5' UTR region of gene.

Fig5. GUS Histochemical staining of different organs in transgenic *Arabidopsis*. Q1~Q7

respectively shows the GUS expression patterns in flower, stem and cauline leaf, and rosette leaf and root harboring different GUS expression structures, and the CK-N and CK-P respectively show the GUS expression profiles in COL, and in positive control harboring 35S::GUS constructs.

To explore the role of *AhLEC1A* on seedling establishment, the transgenic lines with Q7, Q5, Q3 and Q2 constructs were chosen for further analysis. Transgenic Arabidopsis seeds were kept in the dark till their germinating. The results of GUS staining indicated that the unexpanded cotyledon and apex hook of the seedlings harboring Q7 or Q5 construct showed dark blue color, and the hypocotyls were light blue. However, after the etiolated transgenic seedlings had been moved to the light for 2 days, the plants with Q7 construct hardly got dyed, and only the expanded cotyledons with Q5 construct were stained blue. By contrast, the whole seedlings with Q3 or Q2 construct were dyed dark blue under both growth conditions (Fig 6). The results suggested that there existed some negatively regulatory elements at the region of -2228bp ~ -618bp in *AhLEC1A* promoter to control the expression of *AhLEC1A* in hypocotyls and radicles at the stage of seedling formation, and some of them mentioned above might associate with light response.

Fig6. GUS staining of the seedlings germinated for 3~5d in transgenic Arabidopsis. Q2, Q3, Q5 and Q7 indicate respectively the GUS expression patterns in transgenic Arabidopsis lines harboring different GUS expression structures. CK-N and CK-P indicate the GUS expression profiles in COL and in positive control harboring 35S::GUS construct, respectively. A-F indicate the seedlings germinated for 3d in dark, and G-L represent the seedlings after transferring from dark to light for 2 days.

Discussion

Identifying and characterizing the 5' flanking region of gene is helpful for revealing its temporal and spatial expression pattern, and facilitating its utilization in plant genetic engineering[28]. In the present study, we have cloned and analyzed the 5' flanking regulatory sequence of *AhLEC1A*. Several cis-elements in *AhLEC1A* promoter, such as SKn-1, CANBANAPA (CA)_n, AGL15, DOF

core, SEF3 motif and the like, which previously were demonstrated to be required for seed development and storage accumulation, were identified. Skn-1 motif and (CA)_n element were reported to play a vital role in determining the seed-specific expression; the deletion of Skn-1 motif or (CA)_n element in *glutelin* and *napin* promoter decreased their transcription in seeds[29, 30]. The element of DOF core was considered to confer the endosperm-specific expression in *Zea mays* [31, 32]. SEF3 motif is the binding site of Soybean Embryo Factor 3, which regulate the transcription of the β -conglycinin (a storage protein) gene and participate in seed development[33, 34]. Our results of GUS staining assay revealed *GUS* gene, driven by the full-length *AhLEC1A* promoter (Q7), specially expressed in the embryos of the transgenic Arabidopsis, which is well in agreement with our results of gene expression analyzed by RT-PCR method[35]. These data showed that *AhLEC1A* functioned in a seed-specific manner. Otherwise, the transgenic lines with 1611bp ~ 974bp deletion constructs from 5' terminal of Q7 promoter showed the constitutive expression at higher GUS levels in roots, rosettes, stems, flowers, and seeds. Meanwhile, in silico analysis of *AhLEC1A* promoter displayed several tissue-specific elements like mesophyll-specific element CACTFTPPCA1, root-specific element ROOTMOTIFTAPOX1 and pollen-specific element POLLEN1LELAT52 distributed on its upstream regulatory region, as well as many negatively regulatory elements including four WRKY71OS and five WBOXATNPR1 dispersed intensively in the fragment of -1225 ~ -2228bp which was deleted in Q3 ~ Q6. These results demonstrate that *AhLEC1A* expression in seed-specific pattern might be attributed to be negatively regulated its transcription in vegetative organs by some cis-element existed in the distal region of its promoter, and simultaneously to be controlled its expression in seeds positively by some seed-specific elements in the proximal region of its promoter. This regulatory model was also found in *AtLEC1* promoter of Arabidopsis[24], *D540*

promoter of rice[36], and *C-hordein* promoter of barley[37].

Beyond embryogenesis and embryo development, LEC1 also regulate skotomorphogenesis of seedlings at the post-germination stage. The unexpanded cotyledons and apical hook of seedlings with Q7 construct germinated in dark dyed obviously in blue, while the whole seedlings were scarcely stained after transferring to light for 2 days. It was suggested that light might repress the expression of *AhLEC1A* by recruiting some proteins to bind the particular elements in its promoter. Our results found that total 23 light-responsive elements I BOX CORE/GATA BOX (GATA) scattered on Q7 segment of *AhLEC1A* promoter, 9, 15 and 21 out of them were respectively deleted in Q5, Q3 and Q2 promoter, resulting in that in GUS assay, the staining patterns of Q2 and Q3 transgenic plants in darkness were similar to those in light, and the hypocotyls of Q5 transgenic plants were dyed in blue when growing in darkness while there were no dyeing after transferring them to light. The core sequence of I BOX, and the GATA BOX with similar function had been shown to be essential for light-regulated transcriptional activation[38-40]. Furthermore, it has been demonstrated that I BOX as a negative cis-element can inhibit the expression of *GalUR* in strawberry, and the inhibited role is strictly depended on light[41]. Yamagata et al. also found that I BOX, as a negative regulatory element, was necessary for down-regulating the expression of *cucumis* gene by binding fruit nuclear protein in Musc melons (*Cucumis melo* L.)[42]. Previous study found that *AtLEC1* promoter exists several I BOX CORE elements, and deleting some of them localized on 5' upstream segment from -436 nt in mutant *tnp* restrains the hypocotyl elongation of etiolated seedlings in darkness[24]. These data suggested that some of I BOX elements function as a negative regulator in response to illumination. In our GUS histochemical assay, the degree and range of dyeing changed with the number of I BOX, demonstrating that some of them might be involved in negative regulating the

expression of *AhLEC1A* gene during the procedure of seedling growth from dark condition to light condition.

The *cis*- elements comparison in the promoters of *AhLEC1A* and *AhLEC1B* showed that lots of similar elements are dispersed in the both promoters, but their amounts and positions were much different (Table 2). *AhLEC1A* promoter contained a number of distinct seed-development related components such as 2S SEED PROT BANAPA, SP8BFIBSP8BIB, CANBNNAPA. However, *AhLEC1B* promoter contained numerous specific elements involved in abiotic stress or hormones responding, including GCCCORE, ASF1MOTIFCAWV, and several regulatory elements known to modulate gene expression at higher transcription level in different plant species[43]. The similarities and differences between two *AhLEC1* promoters implied that their functions might be partially same and redundant, and to some extent *AhLEC1A* and *AhLEC1B* might play different roles during the particular growth and development period of peanuts, respectively. The point of view was consistent with the study of predecessors who thought *LEC1* genes originated from a common ancestor and neofunctionalization and /or subfunctionalization processes were responsible for the emergence of a different role for *LEC1* genes in seeds plants[44]. Moreover, during the evolution of cultivated peanuts, A and B subgenomes were subjected to asymmetric homoeologous exchanges and homoeolog expression bias. Yin et al. considered that A subgenome were significantly affected by domestication, while natural selection preferred to B subgenome[45]. It was speculated that during genome evolution, to satisfy the demands for seed growth and development, the orthologous genes *AhLEC1A* and *AhLEC1B* suffered from the different selection pressure at different life stages to produce their functional divergence.

Table 2. Comparison of regulatory elements in *AhLEC1A* promoter and *AhLEC1B* promoter.

<i>cis</i> -element	<i>AhLEC1A</i> promoter	<i>AhLEC1B</i> promoter	Motif ^a	Putative function
Skn-1 motif	+	+	GTCAT	Cis-acting regulatory element required for endosperm expression[46]
CARGCW8GAT	+	+	CWWWWWWWW WG	Motif with a longer A/T-rich core providing binding site for AGL15 which accumulates during embryo development[47]
CACTFT PPCA1	+	+	YACT	Cis-acting regulatory element required for mesophyll-specific expression[48]
-10PEHVPSBD	+	+	TATTCT	Cis-acting regulatory element involved in light responsiveness[49]
ROOT MOTIF TAPOX1	+	+	ATATT	Motif found in the promoter of rolD, which expresses strongly in roots[50]
POLLEN1 LELAT52	+	+	AGAAA	A regulatory element responsible for pollen specific activation of gene[51]
WRKY71OS	+	+	TGAC	Binding site of rice WRKY71, a transcriptional repressor of the gibberellin signaling pathway[52]
W BOX ATNPR1	+	+	TTGAC	A cluster of WRKY binding sites act as negative A regulatory element for the inducible expression of genes[53]
CPB CSPOR	+	+	TATTAG	Cis-Acting regulatory element involved in Cytokinin responsiveness[54]
DOF CORE ZM	+	+	AAAG	Core site required for binding of Dof proteins which may be associated with the plant-specific pathway for carbon metabolism[31]
E BOX BNNAPA	+	+	CANNTG	The cis-elements in the promoter regions of most genes encoding the storage protein[55]
ERE LEE4	+	+	AWTTCAAA	The ethylene responsive element mediate ethylene-induced activity of transcription[56]
SEF3 MOTIF GM	+	+	AACCCA	Binding with SEF3, one of soybean embryo factor (SEF)[33]
I BOX	+	+	GATAA	Conserved sequence upstream of light-regulated genes[38]
ARFAT	+	-	TGTCTC	Cis-Acting regulatory element involved in auxin responsiveness[57]
CAN BANAPA	+	-	CNAACAC	Core of (CA) _n element in storage protein genes[29]
2S SEED PROT BANAPA	+	-	CAAACAC	Cis-Regulatory element conserved in many storage-protein gene promoters[55]
ATC-motif	+	-	AGCTATCCA	part of a conserved DNA module involved in light responsiveness[58]
ERE	+	-	ATTTCAAA	ethylene-responsive element[59]
GARE AT	+	-	TAACAAR	gibberellin-responsive element[60]
SP8BFIBSP8BIB	+	-	TACTATT	“SP8b” found in the 5' upstream region of three differnt genes coding for sporamin and beta-amylase[61]
TCT motif	+	-	TCTTAC	part of a light responsive element[62]
ASF1MOTIFCAMV	-	+	TGACG	Motif involved in transcriptional activation of genes by auxin or salicylic acid, may be relevant to light regulation[63]
GCCCORE	-	+	GCCGCC	Core of GCC-box found in pathogen-responsive, ethylene-responsive and jasmonate-responsive gene[64]
TGACGTVMAMY	-	+	TGACGT	Motif required for high level expression in cotyledons of the germinated seeds[65]
5'UTR Py-rich stretch	-	+	TTTCTTCTCT	Cis-acting element conferring high transcription levels[66]
CGTCA-motif	-	+	CGTCA	Cis-acting regulatory element involved in the MeJA-responsiveness[67]
GAG-motif	-	+	AGAGAGT	Part of a light responsive element[68]
GARE-motif	-	+	AAACAGA	Gibberellin-responsive element[69]
LTRE1HVBLT49	-	+	CCGAAA	“LTRE-1” (low-temperature- responsive element) in barley (H.v.) blt4.9 gene promoter[62]

Note: “+” means the element existing in the promoter, “-” means the element not existing in the promoter.^aW=A/T; Y=T/C; N=G/C/A/T.

In summary, we identified and characterized the promoter of *AhLEC1A*. It was found that

during the process of seed development and maturation, its expression in embryo were regulated by the positive cis-elements in seed-specific mode and the negative elements restricting its expression in other organs. Moreover, *AhLEC1A* was also involved in skotomorphogenesis of peanut seeds, its expression level in the hypocotyls germinated in darkness was inhibited by some light-responsive elements. The results will be helpful for understanding the function of *AhLEC1A* in peanuts.

Reference

1. Goldberg RB, de Paiva G, Yadegari R. Plant embryogenesis: zygote to seed. *Science*. 1994; 266(5185): 605-614. doi:10.1126/science.266.5185.605 PMID: 17793455
2. Gutierrez L, Van Wuytswinkel O, Castelain M, Bellini C. Combined networks regulating seed maturation. *Trends Plant Sci*. 2007; 12(7): 294-300. doi:10.1016/j.tplants.2007.06.003 PMID: 17588801
3. Harada JJ, Pelletier J. Genome-wide analyses of gene activity during seed development. *Seed ence Research*. 2012; 22(S1): S15-S22. doi:10.1017/s0960258511000304.
4. Braybrook SA, Harada JJ. LECs go crazy in embryo development. *Trends Plant Sci*. 2008; 13(12): 624-630. doi:10.1016/j.tplants.2008.09.008 PMID: 19010711
5. Santos-Mendoza M, Dubreucq B, Baud S, Parcy F, Caboche M, Lepiniec L. Deciphering gene regulatory networks that control seed development and maturation in *Arabidopsis*. *Plant J*. 2008; 54(4): 608-620. doi:10.1111/j.1365-313X.2008.03461.x PMID: 18476867
6. Verdier J, Thompson RD. Transcriptional regulation of storage protein synthesis during dicotyledon seed filling. *Plant Cell Physiol*. 2008; 49(9): 1263-1271. doi:10.1093/pcp/pcn116 PMID: 18701524
7. Roscoe TT, Guillemot J, Bessoule JJ, Berger F, Devic M. Complementation of Seed Maturation Phenotypes by Ectopic Expression of ABSCISIC ACID INSENSITIVE3, FUSCA3 and LEAFY COTYLEDON2 in *Arabidopsis*. *Plant Cell Physiol*. 2015; 56(6): 1215-1228. doi:10.1093/pcp/pcv049 PMID: 25840088
8. West M, Yee KM, Danao J, Zimmerman JL, Fischer RL, Goldberg RB, et al. LEAFY COTYLEDON1 Is an Essential Regulator of Late Embryogenesis and Cotyledon Identity in *Arabidopsis*. *Plant Cell*. 1994; 6(12): 1731-1745. doi:10.1105/tpc.6.12.1731 PMID: 16055812244233
9. Lotan T, Ohto M, Yee KM, West MA, Lo R, Kwong RW, et al. *Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell*. 1998; 93(7): 1195-1205. doi:10.1016/s0092-8674(00)81463-4 PMID: 9657152
10. Harada JJ. Role of *Arabidopsis* LEAFY COTYLEDON genes in seed development. *Journal of Plant Physiology*. 2001; 158(4): 405-409. doi:10.1078/0176-1617-00351.
11. Meinke DW. A Homoeotic Mutant of *Arabidopsis thaliana* with Leafy Cotyledons. *Science*. 1992; 258(5088): 1647-1650. doi:10.1126/science.258.5088.1647 PMID: 17742538

12. Meinke DW, Franzmann LH, Nickle TC, Yeung EC. Leafy Cotyledon Mutants of Arabidopsis. *Plant Cell*. 1994; 6(8): 1049-1064. doi:10.1105/tpc.6.8.1049 PMID: 16050012244265
13. Junker A, Mönke G, Rutten T, Keilwagen J, Seifert M, Thi TM, et al. Elongation-related functions of LEAFY COTYLEDON1 during the development of Arabidopsis thaliana. *Plant J*. 2012; 71(3): 427-442. doi:10.1111/j.1365-313X.2012.04999.x PMID: 22429691
14. Pelletier JM, Kwong RW, Park S, Le BH, Baden R, Cagliari A, et al. LEC1 sequentially regulates the transcription of genes involved in diverse developmental processes during seed development. *Proc Natl Acad Sci U S A*. 2017; 114(32): E6710-e6719. doi:10.1073/pnas.1707957114 PMID: 555904728739919
15. Kagaya Y, Toyoshima R, Okuda R, Usui H, Yamamoto A, Hattori T. LEAFY COTYLEDON1 controls seed storage protein genes through its regulation of FUSCA3 and ABSCISIC ACID INSENSITIVE3. *Plant Cell Physiol*. 2005; 46(3): 399-406. doi:10.1093/pcp/pci048 PMID: 15695450
16. Mu JY, Tan HL, Zheng Q, Fu FY, Liang Y, J Z. LEAFY COTYLEDON1 is a key regulator of fatty acid biosynthesis in Arabidopsis. *Plant Physiol*. 2008; 148(2): 1042-1054. doi:10.1104/pp.108.126342 PMID: 255682718689444
17. Tan HL, Yang XH, Zhang F, Zheng X, Qu CM, Mu JY, et al. Enhanced seed oil production in canola by conditional expression of Brassica napus LEAFY COTYLEDON1 and LEC1-LIKE in developing seeds. *Plant Physiol*. 2011; 156(3): 1577-1588. doi:10.1104/pp.111.175000 PMID: 313596521562329
18. Elahi N, Duncan RW, Stasolla C. Modification of oil and glucosinolate content in canola seeds with altered expression of Brassica napus LEAFY COTYLEDON1. *Plant Physiol Biochem*. 2016; 100(52-63). doi:10.1016/j.plaphy.2015.12.022 PMID: 26773545
19. Tang GY, Xu PL, Ma WH, Wang F, Liu ZJ, Wan SB, et al. Seed-Specific Expression of AtLEC1 Increased Oil Content and Altered Fatty Acid Composition in Seeds of Peanut (Arachis hypogaea L.). *Front Plant Sci*. 2018; 9(260). doi:10.3389/fpls.2018.00260 PMID: 584566829559985
20. Brocard-Gifford IM, Lynch TJ, Finkelstein RR. Regulatory networks in seeds integrating developmental, abscisic acid, sugar, and light signaling. *Plant Physiol*. 2003; 131(1): 78-92. doi:10.1104/pp.011916 PMID: 16678912529517
21. Warpeha KM, Upadhyay S, Yeh J, Adamiak J, Hawkins SI, Lapik YR, et al. The GCR1, GPA1, PRN1, NF-Y signal chain mediates both blue light and abscisic acid responses in Arabidopsis. *Plant Physiol*. 2007; 143(4): 1590-1600. doi:10.1104/pp.106.089904 PMID: 185183517322342
22. Siefers N, Dang KK, Kumimoto RW, Bynum WEt, Tayrose G, Holt BF, 3rd. Tissue-specific expression patterns of Arabidopsis NF-Y transcription factors suggest potential for extensive combinatorial complexity. *Plant Physiol*. 2009; 149(2): 625-641. doi:10.1104/pp.108.130591 PMID: 263383319019982
23. Huang MK, Hu YL, Liu X, Li YG, Hou XL. Arabidopsis LEAFY COTYLEDON1 Mediates Postembryonic Development via Interacting with PHYTOCHROME-INTERACTING FACTOR4. *Plant Cell*. 2015; 27(11): 3099-3111. doi:10.1105/tpc.15.00750 PMID: 468230726566918
24. Casson SA, Lindsey K. The turnip mutant of Arabidopsis reveals that LEAFY COTYLEDON1 expression mediates the effects of auxin and sugars to promote embryonic cell identity. *Plant Physiol*. 2006; 142(2): 526-541. doi:10.1104/pp.106.080895 PMID: 158604016935993

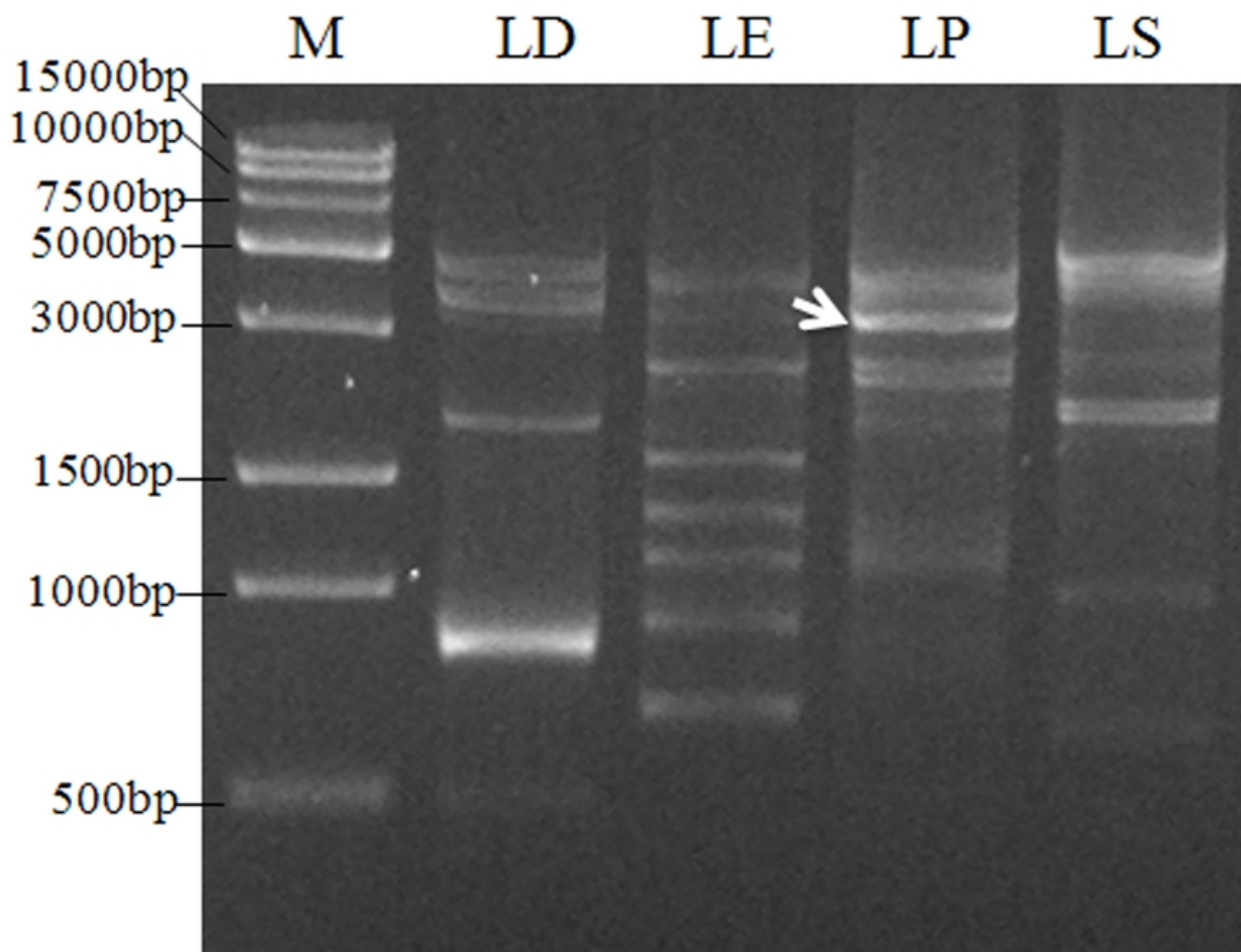
- 389 25. Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA. *Nucleic*
390 *Acids Res.* 1980; 8(19): 4321-4325. doi:10.1093/nar/8.19.4321 PMID: 3242417433111
- 391 26. Chen G, Shan L, Zhou LX, Tang GY, Bi YP. The Comparison of Different Methods for Isolating
392 Total RNA from Peanuts. *Chinese Agricultural ence Bulletin.* 2011; 27(1): 214-218.
393 doi:10.1007/s11606-010-1494-7 (In Chinese)
- 394 27. Jefferson RA, Kavanagh TA, Bevan MW. GUS fusions: beta-glucuronidase as a sensitive and
395 versatile gene fusion marker in higher plants. *Embo j.* 1987; 6(13): 3901-3907.
396 doi:10.1089/dna.1987.6.583 PMID: 5538673327686
- 397 28. Ayoubi TA, Van De Ven WJ. Regulation of gene expression by alternative promoters. *Faseb j.*
398 1996; 10(4): 453-460. doi:10.1089/dna.1996.15.263 PMID: 8647344
- 399 29. Bustos MM, Begum D, Kalkan FA, Battraw MJ, Hall TC. Positive and negative cis-acting DNA
400 domains are required for spatial and temporal regulation of gene expression by a seed storage
401 protein promoter. *Embo j.* 1991; 10(6): 1469-1479. doi:10.1089/dna.1991.10.397 PMID:
402 4528102026144
- 403 30. Washida H, Wu CY, Suzuki A, Yamanouchi U, Akihama T, Harada K, et al. Identification of cis-
404 regulatory elements required for endosperm expression of the rice storage protein glutelin gene
405 GluB-1. *Plant Mol Biol.* 1999; 40(1): 1-12. doi:10.1023/a:1026459229671 PMID: 10394940
- 406 31. Yanagisawa S. Dof1 and Dof2 transcription factors are associated with expression of multiple
407 genes involved in carbon metabolism in maize. *Plant J.* 2000; 21(3): 281-288.
408 doi:10.1046/j.1365-313x.2000.00685.x PMID: 10758479
- 409 32. Yanagisawa S, Schmidt RJ. Diversity and similarity among recognition sequences of Dof
410 transcription factors. *Plant J.* 1999; 17(2): 209-214. doi:10.1046/j.1365-313x.1999.00363.x
411 PMID: 10074718
- 412 33. Lessard PA, Allen RD, Bernier F, Crispino JD, Fujiwara T, Beachy RN. Multiple nuclear factors
413 interact with upstream sequences of differentially regulated beta-conglycinin genes. *Plant Mol*
414 *Biol.* 1991; 16(3): 397-413. doi:10.1007/bf00023991 PMID: 1893110
- 415 34. Allen RD, Bernier F, Lessard PA, Beachy RN. Nuclear factors interact with a soybean beta-
416 conglycinin enhancer. *Plant Cell.* 1989; 1(6): 623-631. doi:10.1105/tpc.1.6.623 PMID:
417 1597972535514
- 418 35. Li AQ, Xia H, Wang XJ, Li CS, Zhao CZ, Bi YP. Cloning and expression analysis of peanut
419 (*Arachis hypogaea* L.) LEC1. *Acta Bot Boreal-OccidentSin.* 2009; 29(9)(1730-1735.
420 doi:10.1007/978-1-4020-9623-5_5 (In Chinese)
- 421 36. Cai M, Wei J, Li XH, Xu CG, Wang SP. A rice promoter containing both novel positive and
422 negative cis-elements for regulation of green tissue-specific gene expression in transgenic
423 plants. *Plant Biotechnol J.* 2007; 5(5): 664-674. doi:10.1111/j.1467-7652.2007.00271.x PMID:
424 17596180
- 425 37. Müller M, Knudsen S. The nitrogen response of a barley C-hordein promoter is controlled by
426 positive and negative regulation of the GCN4 and endosperm box. *Plant J.* 1993; 4(2): 343-355.
427 doi:10.1046/j.1365-313x.1993.04020343.x PMID: 8220485
- 428 38. Borello U, Ceccarelli E, Giuliano G. Constitutive, light-responsive and circadian clock-
429 responsive factors compete for the different l box elements in plant light-regulated promoters.
430 *Plant J.* 1993; 4(4): 611-619. doi:10.1046/j.1365-313x.1993.04040611.x PMID: 8252065
- 431 39. Reyes JC, Muro-Pastor MI, Florencio FJ. The GATA family of transcription factors in
432 *Arabidopsis* and rice. *Plant Physiol.* 2004; 134(4): 1718-1732. doi:10.1104/pp.103.037788

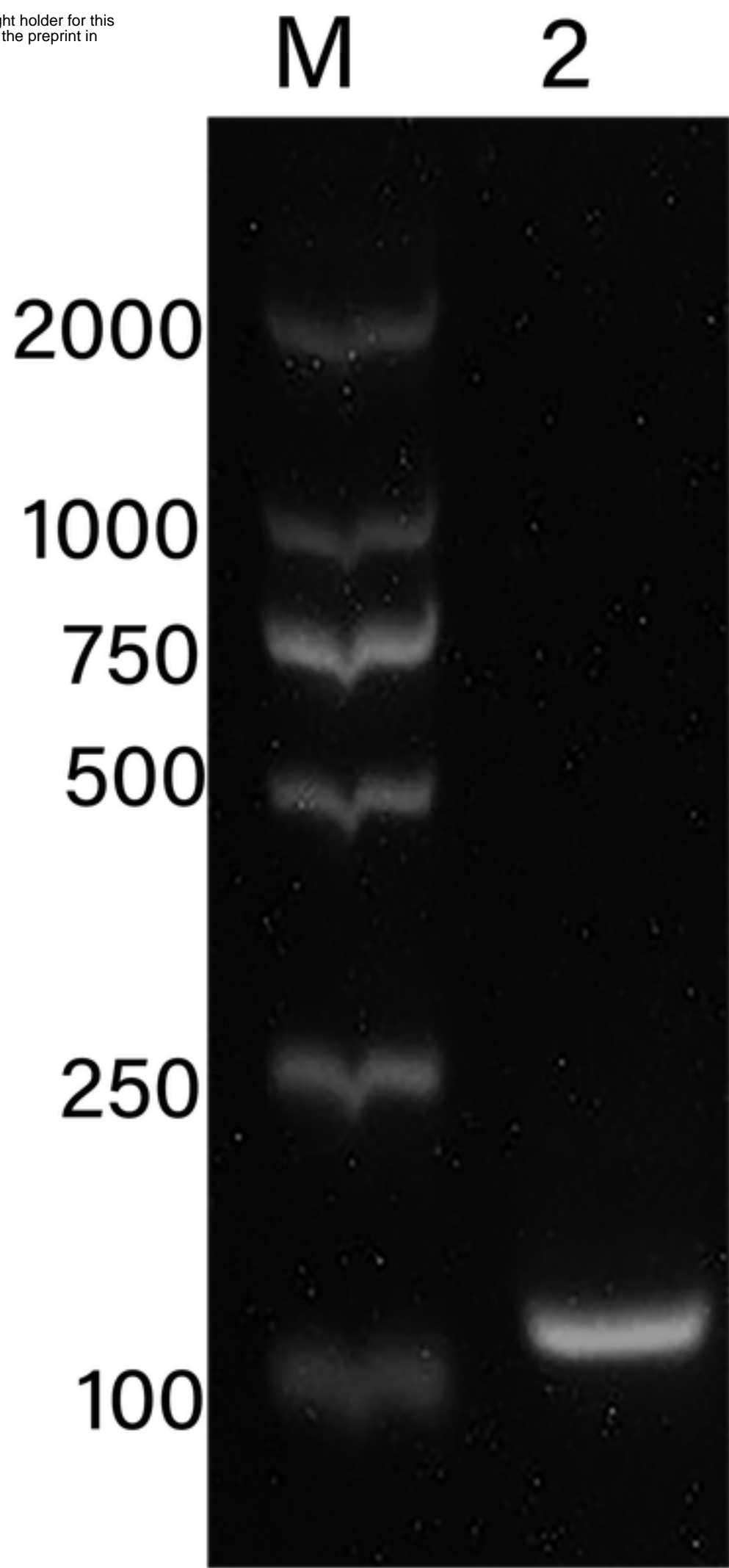
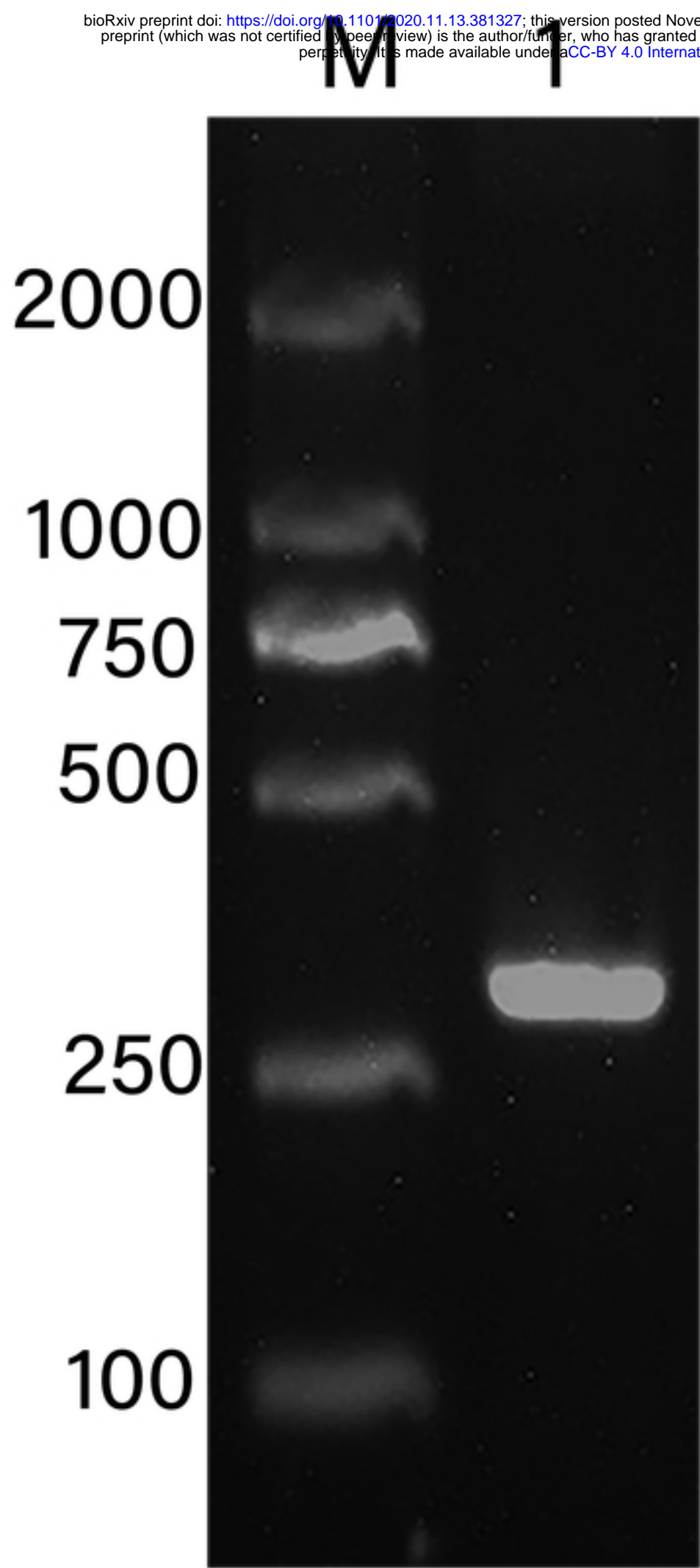
PMID: 41984515084732

40. Terzaghi WB, Cashmore AR. Light-Regulated Transcription. Annual Review of Plant Physiology and Plant Molecular Biology. 1995; 46(1): 445-474.
doi:10.1146/annurev.pp.46.060195.002305
41. Agius F, Amaya I, Botella MA, Valpuesta V. Functional analysis of homologous and heterologous promoters in strawberry fruits using transient expression. J Exp Bot. 2005; 56(409): 37-46. doi:10.1093/jxb/eri004 PMID: 15533885
42. Yamagata H, Yonesu K, Hirata A, Aizono Y. TGTCACA motif is a novel cis-regulatory enhancer element involved in fruit-specific expression of the cucumisin gene. J Biol Chem. 2002; 277(13): 11582-11590. doi:10.1074/jbc.M109946200 PMID: 11782472
43. Tang G, Xu P, Liu W, Liu Z, Shan L. Cloning and Characterization of 5' Flanking Regulatory Sequences of AhLEC1B Gene from Arachis Hypogaea L. PLoS One. 2015; 10(10): e0139213. doi:10.1371/journal.pone.0139213 PMID: 459127726426444
44. Cagliari A, Turchetto-Zolet AC, Korbes AP, Maraschin Fdos S, Margis R, Margis-Pinheiro M. New insights on the evolution of Leafy cotyledon1 (LEC1) type genes in vascular plants. Genomics. 2014; 103(5-6): 380-387. doi:10.1016/j.ygeno.2014.03.005 PMID: 24704532
45. Yin D, Ji C, Song Q, Zhang W, Zhang X, Zhao K, et al. Comparison of Arachis monticola with Diploid and Cultivated Tetraploid Genomes Reveals Asymmetric Subgenome Evolution and Improvement of Peanut. Adv Sci (Weinh). 2020; 7(4): 1901672. doi:10.1002/advs.201901672 PMID: 702964732099754
46. Takaiwa F, Oono K, Wing D, Kato A. Sequence of three members and expression of a new major subfamily of glutelin genes from rice. Plant Mol Biol. 1991; 17(4): 875-885. doi:10.1007/bf00037068 PMID: 1680490
47. Tang W, Perry SE. Binding site selection for the plant MADS domain protein AGL15: an in vitro and in vivo study. J Biol Chem. 2003; 278(30): 28154-28159. doi:10.1074/jbc.M212976200 PMID: 12743119
48. Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, et al. cis-Regulatory elements for mesophyll-specific gene expression in the C4 plant Flaveria trinervia, the promoter of the C4 phosphoenolpyruvate carboxylase gene. Plant Cell. 2004; 16(5): 1077-1090. doi:10.1105/tpc.019729 PMID: 42320115100398
49. Thum KE, Kim M, Morishige DT, Eibl C, Koop HU, Mullet JE. Analysis of barley chloroplast psbD light-responsive promoter elements in transplastomic tobacco. Plant Mol Biol. 2001; 47(3): 353-366. doi:10.1023/a:1011616400264 PMID: 11587507
50. Elmayan T, Tepfer M. Evaluation in tobacco of the organ specificity and strength of the rolD promoter, domain A of the 35S promoter and the 35S2 promoter. Transgenic Res. 1995; 4(6): 388-396. doi:10.1007/bf01973757 PMID: 7581519
51. Filichkin SA, Leonard JM, Monteros A, Liu PP, Nonogaki H. A novel endo-beta-mannanase gene in tomato LeMAN5 is associated with anther and pollen development. Plant Physiol. 2004; 134(3): 1080-1087. doi:10.1104/pp.103.035998 PMID: 38993214976239
52. Zhang ZL, Xie Z, Zou X, Casaretto J, Ho TH, Shen QJ. A rice WRKY gene encodes a transcriptional repressor of the gibberellin signaling pathway in aleurone cells. Plant Physiol. 2004; 134(4): 1500-1513. doi:10.1104/pp.103.034967 PMID: 41982615047897
53. Chen C, Chen Z. Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced Arabidopsis transcription factor. Plant Physiol. 2002; 129(2):

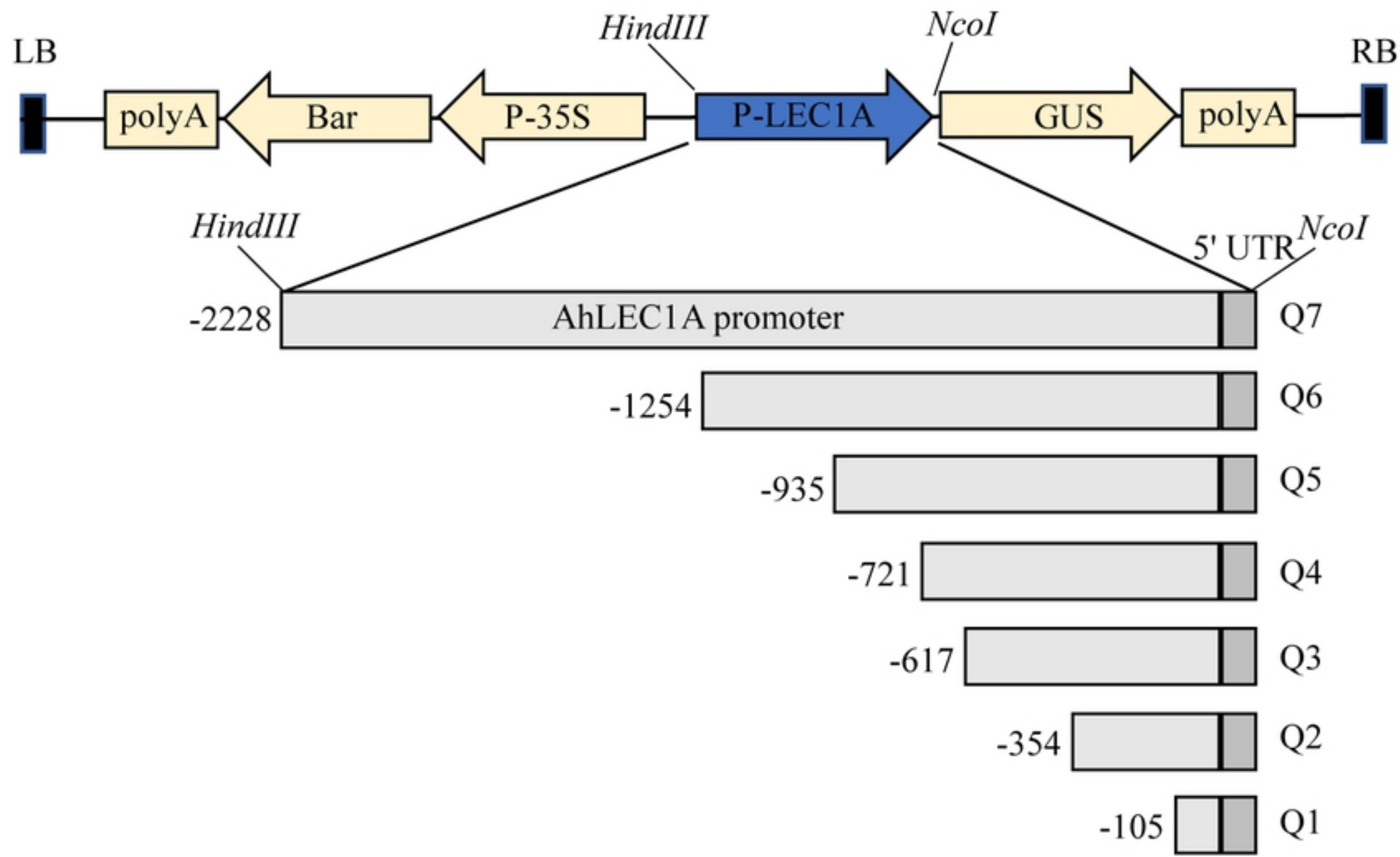
-
- 706-716. doi:10.1104/pp.001057 PMID: 16169512068113
54. Fusada N, Masuda T, Kuroda H, Shimada H, Ohta H, Takamiya K. Identification of a novel cis-element exhibiting cytokinin-dependent protein binding in vitro in the 5'-region of NADPH-protochlorophyllide oxidoreductase gene in cucumber. *Plant Mol Biol*. 2005; 59(4): 631-645. doi:10.1007/s11103-005-0579-x PMID: 16244912
55. Stålberg K, Ellerstöm M, Ezcurra I, Ablov S, Rask L. Disruption of an overlapping E-box/ABRE motif abolished high transcription of the napA storage-protein promoter in transgenic *Brassica napus* seeds. *Planta*. 1996; 199(4): 515-519. doi:10.1007/bf00195181 PMID: 8818291
56. Tapia G, Verdugo I, Yañez M, Ahumada I, Theoduloz C, Cordero C, et al. Involvement of ethylene in stress-induced expression of the TLC1.1 retrotransposon from *Lycopersicon chilense* Dun. *Plant Physiol*. 2005; 138(4): 2075-2086. doi:10.1104/pp.105.059766 PMID: 118339616040666
57. Hagen G, Guilfoyle T. Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Mol Biol*. 2002; 49(3-4): 373-385. doi:10.1023/A:1015207114117 PMID: 12036261
58. Welsch R, Medina J, Giuliano G, Beyer P, Von Lintig J. Structural and functional characterization of the phytoene synthase promoter from *Arabidopsis thaliana*. *Planta*. 2003; 216(3): 523-534. doi:10.1007/s00425-002-0885-3 PMID: 12520345
59. Rawat R, Xu ZF, Yao KM, Chye ML. Identification of cis-elements for ethylene and circadian regulation of the *Solanum melongena* gene encoding cysteine proteinase. *Plant Mol Biol*. 2005; 57(5): 629-643. doi:10.1007/s11103-005-0954-7 PMID: 15988560
60. Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S. Gibberellin biosynthesis and response during *Arabidopsis* seed germination. *Plant Cell*. 2003; 15(7): 1591-1604. doi:10.1105/tpc.011650 PMID: 16540312837949
61. Ishiguro S, Nakamura K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. *Mol Gen Genet*. 1994; 244(6): 563-571. doi:10.1007/bf00282746 PMID: 7969025
62. Luo QL, Li YG, Gu HQ, Zhao L, Gu XP, Li WB. The promoter of soybean photoreceptor GmPLP1 gene enhances gene expression under plant growth regulator and light stresses. *Plant Cell, Tissue and Organ Culture (PCTOC)*. 2013; 114(1): 109-119. doi:10.1007/s11240-013-0310-6
63. Redman J, Whitcraft J, Johnson C, Arias J. Abiotic and biotic stress differentially stimulate as-1 element activity in *Arabidopsis*. *Plant Cell Reports*. 2002; 21(2): 180-185. doi:10.1007/s00299-002-0472-x.
64. Chakravarthy S, Tuori RP, D'Ascenzo MD, Fobert PR, Despres C, Martin GB. The tomato transcription factor Pti4 regulates defense-related gene expression via GCC box and non-GCC box cis elements. *Plant Cell*. 2003; 15(12): 3033-3050. doi:10.1105/tpc.017574 PMID: 28285414630974
65. Yamauchi D. A TGACGT motif in the 5'-upstream region of alpha-amylase gene from *Vigna mungo* is a cis-element for expression in cotyledons of germinated seeds. *Plant Cell Physiol*. 2001; 42(6): 635-641. doi:10.1093/pcp/pce079 PMID: 11427683
66. Ye J, Li WF, Ai G, Li CX, Liu GZ, Chen WF, et al. Genome-wide association analysis identifies

521 a natural variation in basic helix-loop-helix transcription factor regulating ascorbate biosynthesis
522 via D-mannose/L-galactose pathway in tomato. PLoS Genet. 2019; 15(5): e1008149.
523 doi:10.1371/journal.pgen.1008149 PMID: PMC652724431067226
524 67. Wang Y, Liu GJ, Yan XF, Wei ZG, Xu ZR. MeJA-inducible expression of the heterologous JAZ2
525 promoter from Arabidopsis in Populus trichocarpa protoplasts. J PLANT DIS PROTECT. 2011.
526 68. Jin B, Sheng ZL, Ishfaq M, Chen JQ, Yang HL. Cloning and functional analysis of the promoter
527 of a stress-inducible gene (Zmap) in maize. PLoS One. 2019; 14(2): e0211941.
528 doi:10.1371/journal.pone.0211941 PMID: 636832130735543
529 69. Sutoh K, Yamauchi D. Two cis-acting elements necessary and sufficient for gibberellin-
530 upregulated proteinase expression in rice seeds. Plant J. 2003; 34(5): 635-645.
531 doi:10.1046/j.1365-3113x.2003.01753.x PMID: 12787245
532
533





bioRxiv preprint doi: <https://doi.org/10.1101/2020.11.13.381327>; this version posted November 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.





CK-N

Q1

Q2

Q3

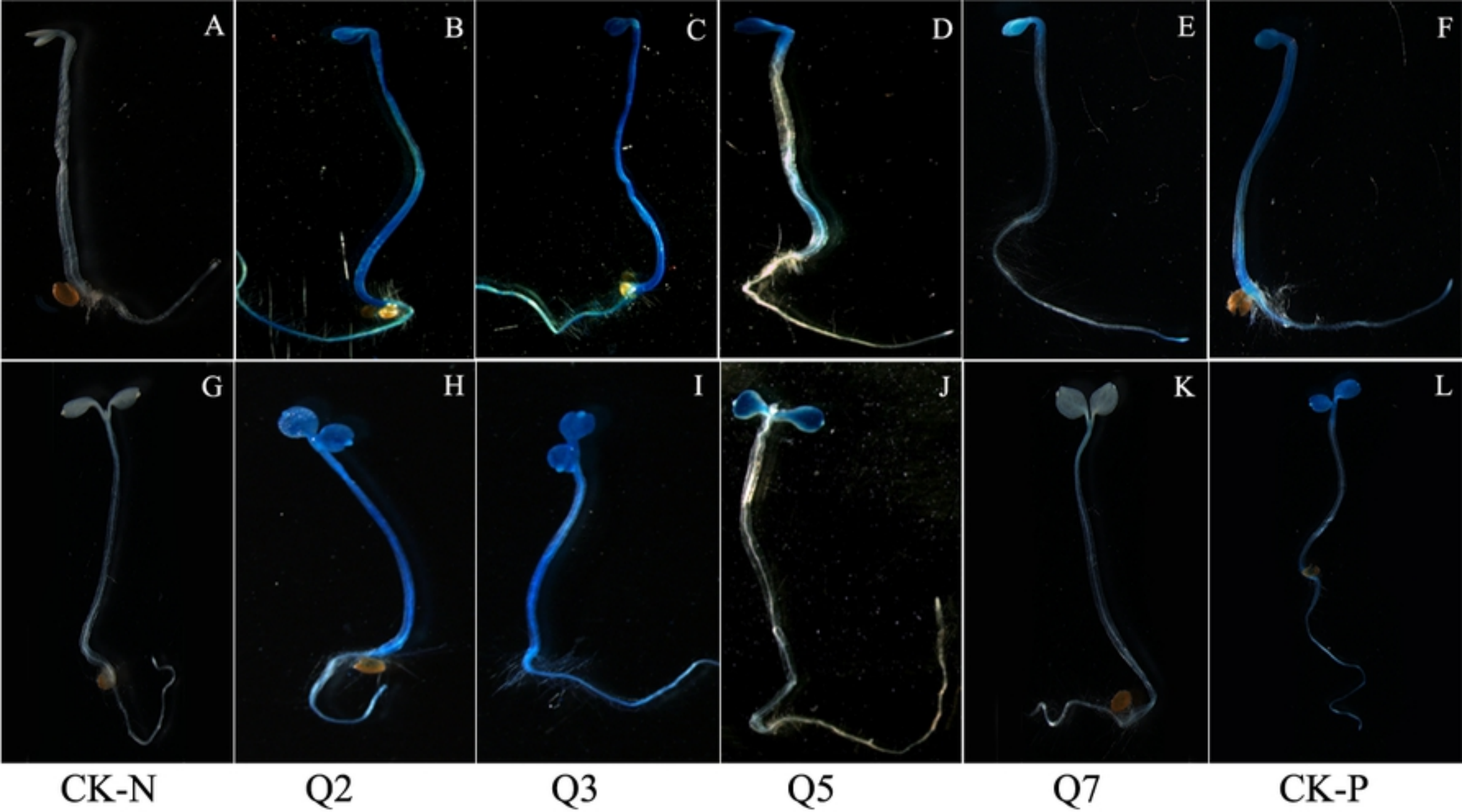
Q4

Q5

Q6

Q7

CK-P



Figure