Recent studies revealed differential usage of Alternative Polyadenylation (APA) in Arabidopsis thaliana wild-type and the oxt6 mutant in which the Oxidative Stress Tolerance 6 gene (OXT6) encoding cleavage and polyadenylation specificity factor 30 (CPSF30) is interrupted. In this thesis I confirmed the differential patterns of APA for the gene LHC4.1, encoding light harvesting complex II subunit, in wild-type versus the oxt6 mutant under either normal conditions or light/temperature stresses by using reverse transcription followed by real time PCR. It suggests the regulatory roles of AtCPSF30 and environmental stresses in the expression level and APA selection of other genes. In addition, leaves of oxt6 plants subjected to dark treatment etiolated faster than that of wild-type under dark treatment. Under dark treatment, LHC4 protein abundance and its associated complexes were significantly lower in the oxt6 plants than in wild-type as determined by using blue native polyacrylamide gel electrophoresis and western blots. These results suggest that oxt6 mutant affects photosystem productivity and may explain why oxt6 plants have reduced stature.
ALTERNATIVE POLYADENYLATION REGULATES THE EXPRESSION OF THE LIGHT HARVESTING GENE LHCBI.1 IN ARABIDOPSIS MUTANT OXT6

A Thesis

Submitted to the
Faculty of Miami University
in partial fulfillment of
the requirements for the degree of
Master of Science
Department of Botany

by
Jie Chen

Miami University
Oxford, Ohio
2011

Advisor________________________
(Dr. Qingshun Quinn Li)

Reader________________________
(Dr. Chun Liang)

Reader________________________
(Dr. Richard Moore)
Table of Contents

Table of Contents ............................................................................................................ ii
Abbreviation .................................................................................................................. vi
List of Tables ................................................................................................................... iv
List of Figures .................................................................................................................. v
Acknowledgment ............................................................................................................ vii
Introduction .................................................................................................................. 1
  Pre-mRNA 3’-end processing ...................................................................................... 1
  Alternative polyadenylation .......................................................................................... 2
  AtCPSF30 and its role in polyadenylation ................................................................. 4
  The OXT6 gene and its mutant in Arabidopsis ............................................................ 5
  Differential usage of poly(A) sites in the oxt6 mutant ................................................. 6
  Light harvesting complex (LHC) and LHC4 protein .................................................... 6
Materials and Methods .................................................................................................. 11
  Plant growth and harvesting ....................................................................................... 11
  Confirmation of polyadenylation sites ........................................................................ 12
  Primer design for quantitative real time PCR ............................................................ 15
  Quantitative real time PCR analysis of APA usage switch ......................................... 15
  Thylakoid membrane preparation ............................................................................. 16
  Determination of chlorophyll content ....................................................................... 17
  Non-denaturing blue native PAGE ........................................................................... 17
  Western blotting ......................................................................................................... 18
Results ....................................................................................................................... 18
  Confirmation of APA sites in seven genes in WT ....................................................... 18
  Reverse primer design for qPCR ............................................................................... 20
  Differential usage of APA in LHC4.1 in WT versus the oxt6 mutant ......................... 20
  Impact on LHC4.1 mRNA level by abiotic stresses ................................................... 23
  Impact on APA usage of LHC4.1 gene under abiotic stresses .................................... 23
  Oxt6 plants turn yellow earlier than WT under dark treatment ................................. 26
  Blue Native PAGE and immunobloting of LHC4 protein .......................................... 26
Discussion .................................................................................................................. 32
  Reverse primer design for determination of specific LHC4.1 transcript isoforms abundance ................................................................. 33
Mutation of OXT6 gene alters APA selection of LHCBl.1 gene.................................33
Smaller stature of the oxt6 mutant might be due to its impact on photosynthesis productivity.........................................................................................................................34
References....................................................................................................................36
List of Tables

Table 1 List of primers................................................................................................... 13
Table 2 Description of main biological functions of genes with APA sites confirmed .... 19
List of Figures

Figure 1 Schematic representation of Alternative Polyadenylation (APA) ......................... 3
Figure 2 Schematic representation of the PSII supercomplex ........................................... 8
Figure 3 Poly(A) sites confirmed by 3’ RACE ................................................................... 16
Figure 4 Primer design for determination of different transcript isoforms abundance by qPCR .......................................................................................................................... 21
Figure 5 Abundance of different LHCB4.1 transcript isoforms using different APA under normal growth conditions in oxt6 .................................................................................. 22
Figure 6 Expression levels of LHCB4.1 mRNA using any APA under dark, high light, and cold stresses ........................................................................................................... 24
Figure 7 APA usage of LHCB4.1 under dark stress ................................................................ 25
Figure 8 APA usage of LHCB4.1 under high light stress .................................................... 27
Figure 9 APA usage of LHCB4.1 under cold stress .............................................................. 28
Figure 10 Days for overall leaves to turn half yellow under dark treatment ....................... 29
Figure 11 5.5-week-old plants before and after 8-day dark treatment ............................... 30
Figure 12 Blue native PAGE and western blot of LHCB4 under dark treatment ............. 32
List of Abbreviations

APA: Alternative polyadenylation
CPSF30: Cleavage Polyadenylation Specificity Factor 30
CstF: Cleavage Stimulation Factor
LHC: Light Harvesting Complex

LHCB4.1: Light Harvesting Chlorophyll a/b Binding 4.1 Gene
LHCB4 (CP29): Light Harvesting Chlorophyll a/b Binding 4 Protein
OXT6: Oxidative Stress Tolerance 6 Gene
Poly(A): Polyadenylation
Acknowledgment

I would like to give many thanks to my supervisor, Dr. Qingshun Quinn Li, for his continuous patience, encouragement, and effort in my research and two years of graduate study at Miami. I would also like to thank my committee members Dr. Chun Liang and Dr. Richard Moore for their valuable time and suggestions on my research and study. I am especially grateful to the two postdoctoral scholars at our lab, Dr. Man Liu and Dr. Denghui Xing for their great help for my problems either in research ideas, technical issues, or my personal life. I would also like to give credits to my lab mates Mr. Jun Zheng and Ms. Xiaohui Wu for their helpful discussions and friendship. I also appreciate a lot of help from Ms. Xiaoyun Deng in the Center of Bioinformatics and Functional Genomics for technical supports. I would also like to thank two undergraduates in our lab, Mr. Stephen Kiss and Mr. Chorus Anthony Jennings for their help in growing plants and many pleasant talks.

I thank Dr. John Z. Kiss, Barbara Wilson and Vickie Sandlin for their kind help over the last two years which makes my life much easier. I have special thanks to Dr. Carole Dabney-Smith in the Department of Chemistry at Miami University and her students Kristen Fite and Amanda Storm for their assistance in performing Blue Native PAGE. I thank all my graduate student fellows, especially Mr. Kokulapalan Wimalanathan for his great help in my study, research, and my graduate life over the last two years. I also thank Hans E. Waldenmaier for his kind help in my English improvement and adaptation to the new environment in the first semester.

Last, I would like to extend my appreciation to US National Science Foundation (awarded to my advisor Dr. Quinn Li), the Botany Academic Challenge Grant, Botany Travel Fund, CMSB Travel Fund, and Graduate Funding, all of which support my research financially.
Introduction

Pre-mRNA 3’-end processing

Gene transcription in eukaryotic cells is a very complicated and well-regulated process. The transcription process can be mainly divided into initiation, elongation, and termination, while each nascent mRNA precursor (pre-mRNA) has to go through extensive processing, such as 5’-end capping, splicing, and 3’-end cleavage and polyadenylation to form a mature mRNA. Gene transcription requires three distinct RNA polymerases, RNA polymerase I, II, or III depending on the types of precursor RNA (Cramer et al., 2011). The most important polymerase is RNA polymerase II (RNAPII), which is involved in transcribing protein coding genes as well as many non-coding RNAs like microRNA, small nuclear RNA, and small nucleolar RNA (Richard and Manley, 2009). Pre-mRNA 3’-end processing plays a crucial role in the termination of transcripts by RNAPII (Richard and Manley, 2009). This 3’-end processing includes cleavage at the 3’-end as well as the addition of a polyadenylation tail (Proudfoot, 2004), both of which are highly regulated by a number of polyadenylation factors and sequence elements in the pre-mRNA (cis-elements). It has been demonstrated that mRNA 3’-end formation may be an important hub for eukaryotic gene expression (Danckwardt et al., 2008).

In mammalian cells, it has been found that the termination of transcripts takes place in multiple locations from only a few base pairs to several kb downstream from the 3’-end of the mature transcripts (Proudfoot, 1989). An intact polyadenylation signal is required to terminate protein-coding genes in yeast and humans (Proudfoot and O’sullivan, 2002). Two conserved elements considered as polyadenylation signals in the pre-mRNA molecule have been found in mammalian cells. They are a highly conserved AAUAAA hexamer in the upstream and a poorly conserved GU- or U-rich sequence as the downstream element, and some genes have an element that is upstream of AAUAAA (Gilmartin, 2005). In yeast, polyadenylation signals are less conserved and more complex than those in mammals (Zhao et al., 1999).
In plants, the nucleotide profiles of 3'UTR (untranslated regions) of plants are different from those of mammalian and yeast cells. Plants have three major poly(A) signals, the far upstream element (FUE), the near upstream element (NUE), and the cleavage element (CE) including the cleavage site (Rothine, 1996; Li and Hunt, 1997; Rothine et al., 2001, Loke et al. 2005). Arabidopsis, as a representative of dicot plants, also has very typical polyadenylation signal patterns. All these complex elements contribute to the polyadenylation mechanism (Xing and Li 2010). Polyadenylation signals in rice have been found to share a significant similarity with Arabidopsis (Shen et al., 2008a).

In addition to the conserved poly(A) signals in pre-mRNA, the cleavage and addition of a tract of poly(A) to the 3'-end of a processed pre-mRNA require a complex of multi-subunit protein factors which is largely conserved in eukaryotes (Keller and Minvielle-Sebastia, 1997; Wahle and Ruegsegger, 1999; Zhao et al., 1999). More than 14 proteins and 20 proteins have been identified for the mammalian and yeast complexes, respectively (Mandel et al., 2008). The mammalian machinery contains several sub-complexes, such as Cleavage STimulation Factor (CstF), Cleavage and Polyadenylation Specificity Factor (CPSF), Cleavage Factor I, and Cleavage Factor II (Mandel et al., 2008). Mammalian CPSF contains five subunits, CPSF30, CPSF73, CPSF100, CPSF160 and hFip1, all of which are necessary for efficient cleavage and polyadenylation of pre-mRNAs (Mandel et al., 2008). The gene homologues of Poly(A) polymerase, CPSF, CstF, and Fip1 have been characterized in Arabidopsis (Hunt et al., 2008a).

**Alternative polyadenylation**

Conventional poly(A) sites are located at the end of the 3'UTR of the precursor mRNA sequence (Figure 1). However, a significant number of genes have been found to have poly(A) sites occurring at different places of their genes, including exon, intron, 5'UTR, or alternate sites of 3'UTR (Figure 1). The latter phenomena is called Alternative PolyAdenylation (APA) to differentiate it from polyadenylation at conventional poly(A) sites. APA produces a variety of transcripts because of the different locations of APA sites within a gene. Moreover, if an open reading frame is modified by the selection of
Regular poly(A) sites are located at the end of the 3’UTR of the precursor mRNA sequence. However, a significant number of genes have APA locating at either exon, intron, 5’UTR, or alternate sites of 3’UTR. Thus, APA produces a variety of transcript isoforms and proteins from the same gene.

Figure 1 Schematic representation of Alternative Polyadenylation (APA).
an APA site in a coding region, the resultant mRNA might produce different proteins, or sometimes no protein is produced (Zhang et al., 2005; Lutz, 2008).

Utilization of different APA sites has been shown to be important for gene expression regulation, and potentially functions as a regulatory hub (Danckwardt et al., 2008). Reprogramming of 3'UTRs by APA, which results from regulation of both general polyadenylation activity and cell type-specific factors, can reset post-transcriptional gene regulatory programs in the cell, and is an integral part of induced pluripotent stem cell generation (Ji and Tian, 2009). The APA pattern can also be a good biomarker for cell type and state, and useful for sample classification (Ji and Tian, 2009). APA is also a mechanism by which genes can escape miRNA-mediated repression in cancers (Mayr and Bartel, 2009). It has been demonstrated that APA may play an important role in response of genes to various stresses and developmental cues in *C. elegans* (Mangone et al., 2010). Several recent studies reveal that different selections of APA sites in antisense transcripts can lead to sense gene silencing where flowering time control in plants are involved (Hornyik et al., 2010; Liu et al., 2010).

In animals it is estimated that >50% of genes have APA sites, the majority of which result in 3'UTRs of different lengths while others may generate mRNA with diverse coding capacities (Rosonina and Manley, 2010). The dicot model plant *Arabidopsis thaliana* has typical mRNA polyadenylation signal patterns (Loke et al, 2005). Based on a newly developed deep sequencing approach, it was found that about 70% of *Arabidopsis* genes have more than one poly(A) sites, excluding microheterogeneity (Wu et al., 2011). Alternative mRNAs that differ in their 3'UTRs can exist in different tissues or developmental stages, and studies have shown that these mRNA isoforms can have different stability or translational activity (Edwalds-Gilbert et al., 1997; Lutz, 2008; Shen et al 2011).

**AtCPSF30 and its role in polyadenylation**

The 160-kDa subunit of CPSF, namely CPSF160, is responsible for recognizing the poly(A) signal AAUAAA (Addepalli and Hunt, 2007). AtCPSF73-II (the 73 kDa subunit) and AtCPSF100 (the 100 kDa subunit) play important roles in gametophyte genetic
transmission, and post-transcriptional gene silencing, respectively (Xu et al., 2004; Herr et al., 2006). Another two CPSF subunits, namely CPSF30 (the 30 kDa subunit) and CPSF73 (the 73 kDa subunit), have been suggestive of possessing endonuclease activities subsequent to polyadenylation (Zarudnaya et al., 2002; Mandel et al., 2006).

AtCPSF30 is an RNA-binding protein whose activity is inhibited by calmodulin, suggesting a role of calmodulin in regulating RNA processing (Delaney et al., 2006a). Also, AtCPSF30 possesses a zinc finger motif, CCCH, which is the endonucleolytic module (Addepalli and Hunt, 2007). This endonuclease activity of AtCPSF30 is coordinated with another polyA factor AtFip1 (Addepalli and Hunt, 2008b). It has been also revealed that heavy metal and redox reagents regulate AtCPSF30 activities, which implicates regulatory functions that AtCPSF30 might be involved in (Addepalli and Hunt, 2008a). Further, AtCPSF30 was found to be a hub of an extensive network of protein-protein interactions in plant mRNA polyadenylation process (Hunt et al., 2008). It interacts with many other polyadenylation factors and even with itself (Delaney et al., 2006b; Xu et al., 2006; Hunt et al., 2008). Interestingly, the gene encoding AtCPSF30 is a non-essential gene in Arabidopsis, which is different from its homologues in mammals (Delaney et al., 2006a).

**The OXT6 gene and its mutant in Arabidopsis**

AtCPSF30 is encoded by the gene OXT6, or Oxidative stress tolerant 6 (Zhang et al 2008). The mutation of OXT6 (gene locus ID At1g30460) is due to a T-DNA insertion of the first exon. The oxt6 mutant exhibits tolerance to oxidative stress. When plants are under many biotic and abiotic stresses, such as drought, dark, high light, salinity, and high temperature stresses, the concentration of reactive oxygen species (ROS) increases, which results in damage of most cellular macromolecules and ultimate cell death (Lerner, 1999; Dat et al., 2000). Interestingly, abundance of OXT6 transcripts changes in wild-type under oxidative stresses, which indicates the OXT6 gene might be involved in oxidative stresses (Liu, 2010).

The bulk of polyadenylation sites are affected in the oxt6 mutant (A. G. Hunt and P. Thomas, personal communications). In particular, a group of genes were found to use
different polyadenylation sites or APA between wild-type and *oxt6* indicating that CPSF30 might have an impact in the selection of polyadenylation sites and possibly links these APAs to the regulation of the oxidative stress response (Zhang et al., 2008). In addition, the calmodulin-binding domain (CAM) in CPSF30 connects CPSF30 with calcium signaling, suggesting that *OXT6* gene might be involved in responding to various environmental stresses through calmodulin/calcium sensing mechanisms (Zhang et al., 2008).

**Differential usage of poly(A) sites in the *oxt6* mutant**

Deep sequencing analysis of the *Arabidopsis* genome has been conducted to tally the poly(A) sites and APA profiles of the leaves of the *oxt6* mutant and its parent wild type (WT) supports the presence of a significant number of “switch” APA sites found among them (A. G. Hunt and P. Thomas, personal communications). These “switch” APA sites are defined as locating in genes which use their APA sites differentially in WT and *oxt6* mutant. From these results, 26 genes have been found to use APA significantly differentially in WT and *oxt6*. Therefore, it is reasonable to predict that *oxt6* alters gene expression level and selection of APA sites of genes. One of the genes affected encodes LHCb4.1 (*Light Harvesting Chlorophyll a/b Binding 4*), a component of the light harvesting complex of photosystem II. However, all these data was produced from large scale of sequencing and need to be experimentally validated.

**Light harvesting complex (LHC) and LHCb4 protein**

The light reaction of photosynthesis in plants, algae, and cyanobacteria is mainly controlled by large protein complexes locating in the chloroplast thylakoid membranes with different amounts of antenna proteins (Dekker and Boekema, 2005). A light harvesting complex (LHC) is a complex of functioning light harvesting (or antenna) subunit proteins locating around the reaction center. LHC interacts with pigments such as chlorophylls, xanthophylls, and carotenoids to transfer absorbed energy to chlorophyll a molecules at the reaction center.

Photosystem I (PSI) contains a monomeric core complex and four different LHCI (Light Harvesting Complex I) proteins while photosystem II (PSII) consists of two parts. The
inner part of PSII is a core complex including the reaction center and chlorophyll-a binding light–harvesting complexes CP43 and CP47. The outer part of PSII consists of 6 antenna proteins which are encoded by \textit{LHCB} genes. At least 30 genes homologous to \textit{LHC} superfamily genes in \textit{Arabidopsis} have been found (Jansson, 1999). \textit{LHCB 1,2,3} encode three major antenna proteins, all of which associate in different combinations into trimers of the main complex LHCII (Light Harvesting Complex II). This major complex LHCII binds almost 60% of all chlorophyll in PSII (Peter and Thornber, 1991). More importantly, LHCII plays important roles in organizing thylakoid membrane, the structure of the photosynthetic apparatus, regulating energy flow between PSI and PSII, and controlling dissipation of excess excitation energy under light stress (Barros and Kuhlbrandt, 2009). \textit{LHCB4,5,6} encode the other three antenna proteins which are considered as minor antenna proteins. They are LHCB4 (CP29), LHCB5 (CP26), and LHCB6 (CP24), respectively (Figure 2). These minor antenna proteins form into monomeric complexes (van Oort et al., 2010). They bind to the rest 40% of all chlorophyll (Peter and Thornber, 1991). Particularly, LHCB4 and LHCB5 are the two conserved minor monomeric LHCII proteins among green algae and higher plants, while the other minor monomeric LHCII, LHCB6, is only present in higher plants (Iwai et al., 2008).

It has been demonstrated that the absence of any of these minor antenna complexes, especially LHCB4 protein, influences the packing of the supercomplexes in the chloroplast membrane (Yakushevska et al., 2003; de Bianchi et al., 2008a). No microcrystalline arrays of membranes and no PSII could be observed in the absence of LHCB4 even on very mild solubilization (Yakushevska et al., 2003). The connection between reaction centers and a large fraction of light-harvesting complexes is very weak in the absence of LHCB4 (Yakushevska et al., 2003). A LHCB4 RNAi mutant showed reduction in the PSII antenna size in \textit{Chlamydomonas reinhardtii} (Tokutsu et al., 2009). In contrast, the lack of LHCB5 does not lead to substantial change in PSII organization in \textit{Chlamydomonas reinhardtii} (van Oort et al., 2010). All of these suggest LHCB4 serves as a crucial component for the stability of PSII in green plants.
Figure 2 Schematic representation of the PSII supercomplex

The monomeric minor antenna complexes CP24 (LHCB6), CP26 (LHCB5), CP29(LHCB4) are located in between the dimeric core and the S type and M type of trimeric major antenna complex LHCII, at positions S and M. The same color represents the same proteins. (van Oort et al., 2010).
More importantly, other than maintaining the stability of PSII structure, LHCB4, LHCB5, and LHCB6 proteins play important roles in several other mechanisms, including photoprotection, chlorophyll a triplet quenching (Barzda et al., 1998; Mozzo et al., 2008), and non-photochemical quenching (Ahn et al., 2008). In the antisense plants lacking LHCB4, PSII quantum yield as well as the efficiency of trapping of excitation by the reaction center greatly decreased (Andersson et al., 2001).

These minor antenna proteins including LHCB4 also play important roles in reactive oxygen species (ROS) scavenging (Dall'Osto et al., 2007). The up-regulated amount of ROS is a primary effect of plant exposure to abiotic stresses such as high light, low light, cold, drought, high salinity, and heavy metal stresses (Taylor et al., 2009). Particularly, in chloroplasts, ROS is produced mainly in the reaction centers of the two photosystems embedded in the thylakoid membrane (Asada, 2006). Such increasing amount of ROS induces oxidative stresses in both organelle and the whole cell functions (Moller et al., 2007). Actually, abiotic stresses would also have a significant impact on gene expression regulation either from transcriptional, post-transcriptional, translational, or post-translational levels when the abiotic stresses are sensed (Taylor et al., 2009). A number of photosynthesis-related genes were identified among the differentially regulated genes, including gene LHCB4.1, together with other LHCB genes and those genes encoding ribulose, 1,5-bisphosphate carboxylase small subunit and the components of the two photosystems (Seki et al., 2002a).

Another important function about LHCB4 is its role in state transition (Tokutsu et al., 2009; Iwai et al., 2010). To achieve optimal photosynthetic electron transport efficiency and avoid damage during light stress conditions or under changing light conditions, plants and green algae control the balance of excitation energy between PSI and PSII, or as termed “state transition” (Allen et al., 1981). This energy balance is maintained by redistributing LHCII, which results from LHCII phosphorylation. LHCII migrates from PSII when LHCII is phosphorylated under State 2 (where PSII is more excited than PSI). In the unicellular green alga *Chlamydomonas reinhardtii*, the number of phosphorylated residues in LHCB4 protein increases from 2 to 4 under a transfer of the cells from State 1 (where PSI is preferentially excited) to State 2 (Turkina et al., 2006). Because these
minor LHCII complexes locate in the outer layer of the PSII, they border the major LHCII trimers and even PSII core dimer. The whole PSII-LHCII supercomplex is thus forced to migrate when minor LHCII are undocked (Kargul et al., 2005; Takahashi et al., 2006). Together with this, the newly result indicates that LHCB4 is crucial when mobile LHCII dissociates with PSII and re-associates with PSI under State 2 (Iwai et al., 2010). De Bianchi and coworkers found an Arabidopsis mutant lack of LHCB5 protein did not find abnormal state transition in the absence of LHCB5 (de Bianchi et al., 2008b). However, it is still unclear whether LHCB4 protein is as dispensable for state transitions in Arabidopsis as in Chlamydomonas reinhardtii (Tokutsu et al., 2009).

In summary, LHCB4 protein, as minor antenna complex in Photosystem II, associates with ROS production, which can be promoted by various stresses. On the other hand, the oxt6 mutant lacking of AtCPSF30 has increased tolerance to oxidative stress. It is reasonable to predict certain regulatory roles of AtCPSF30 and/or APA selection by studying gene and protein expression level of LHCB4 under light/temperature stresses.

I hypothesize that the mutation of OXT6 may alter APA selection of the LHCB4.1 gene and depress its protein productivity. This, in turn, may reduce the photosynthetic productivity of oxt6 plants and reduce their stature To test these hypotheses, I characterized the APA profiles of LHCB4.1 in WT versus oxt6 mutants. I used qPCR to confirm whether the presence of OXT6 would change the expression level and/or APA selection of the gene LHCB4.1 under normal growth conditions. Because LHCB4.1 gene associates with different environmental stresses, I examined mRNA level and APA usage of LHCB4.1 gene in WT and oxt6 under dark, high light, and cold stresses by qPCR. I then tested physiological change of different age-old of WT and oxt6 plants under dark stress. Finally, I used blue native PAGE and western blot to examine LHCB4 protein abundance under dark treatment.
Materials and Methods

**Plant growth and harvesting**

Seeds of *Arabidopsis thaliana* ecotype Columbia (Col-0) wild type and the *oxt6* mutant were used as described by Zhang et al (2008). Plants were grown on Metro-Mix 360 (Sun Gro Inc.) in growth chambers with a long-day photoperiod (16 hours light /8 hours dark), an irradiance of 120 µE and a temperature cycle of 23°C/20°C (day/night). Plant leaves were harvested during daytime 11 am -12 pm to avoid variance caused by circadian clock regulation.

To study the impact on APA usage in high light, 3 week-old WT and *oxt6* plants were placed under 2000 µE of high light and 4°C for 2 days while maintaining the same light cycle and humidity. To study the impact on APA usage by dark treatment, 3 week-old WT and *oxt6* plants were placed under dark treatment in the same growth chamber with the same temperature and humidity. To study the impact on APA usage by cold stress, 3 week-old WT and *oxt6* plants were placed under 4°C for 2 days while maintaining all other environmental conditions.

To study the physiological change trend under dark treatment, 3 week- to 8 week-old WT and *oxt6* plants were placed under dark treatment in the same growth chamber with the same temperature and humidity until all the leaves turn yellow. These plants were observed and recorded everyday during 5-6 pm under the green light and a weak white fluorescence light (10watts, 520 lumens) for the purpose of clear observation. Plants were subjected to the white fluorescence light for as short time as possible. The observation of color change is qualitative but not quantitative.

To study the abundance of LHCB4 protein, 4-week-old plants under the normal growth conditions and dark treatment (5 days dark after 4 weeks of normal conditions) were used for protein analysis. Plants used comprise WT, *oxt6*, and its complementation lines C30G, C30GM, C30YG, and C30YGM. Two of these lines, C30G and C30YG, were complemented to the *oxt6* mutant by introducing the short and long transcript of OXT6, respectively. The other two lines, C30GM and C30YGM, were complemented to the
oxt6 mutant by introducing the short and long transcript of OXT6 with mutated Calmodulin-binding site, respectively.

**Confirmation of polyadenylation sites**

To confirm APA sites individually, 3’-RACE was performed to a group of 7 selected genes. These 7 genes were picked up according to 2 criteria. One criterion was the gene had sufficient information of its biological functions; the other criterion was that the poly(A) signals were supported by relatively more poly(A) tags, indicating a strong poly(A) signal.

Total RNA was extracted from leaves of 3-week-old Arabidopsis plants by using TURBO DNA-free Kit from Ambion (http://www.ambion.com/) and treated with DNase. cDNA was synthesized from 1 µg of total RNA by using SuperScript® III Reverse Transcriptase Kit from Invitrogen (http://www.invitrogen.com/).

Rapid Amplification of cDNA 3'Ends (3'-RACE) was used to authenticate the Poly(A) sites (Xing et al., 2008). Such confirmations were needed because these poly(A) sites were collected by large scale sequencing approach. For each gene, 3'RACE was performed as described previously (Xing et al., 2008). The common reverse primer oligo-d(T) linked to an adaptor sequence (5’-TTCTAGAATTCCAGCATTCGCTTTTTTTTTTTTTTTTTTN-3’) was used together with specific primers designed as forward primers for each gene. A nested forward primer was used together with oligo-d(T) adaptor primer when an unspecific amplification occurred in 3’RACE. Primers used are listed in Table 1.

Thirty five PCR cycles were employed to achieve optimum amplicon yield, PCR products were then size fractionated on 2% (m/V) agarose gel with Ethidium Bromide (EB) staining. PCR products were then purified by QIAEX II Gel Extraction Kit (Qiagen, http://www.qiagen.com/). To perform sequencing, the PCR products were ligated into the vector by Invitrogen TOPO TA Cloning kit (pCR2.1-TOPO) and transformed into E. Coli DH5α strain. After overnight incubation, plasmids from single colonies were extracted by using QuickClean II Plasmid Miniprep Kit from GenScript. Sequencing was
### Table 1 List of primers

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence (5' - 3')</th>
<th>Gene Locus ID</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>oligo(dT)V</td>
<td>TTCTAGAATTCAGCATTGCTTC</td>
<td>At4g31800</td>
<td>Reverse transcription</td>
</tr>
<tr>
<td>oligo(dT) adapter</td>
<td>TTCTAGAATTCAGCATTGCTTC</td>
<td>At4g31800</td>
<td>3'RACE confirmation</td>
</tr>
<tr>
<td>WRKY18-F1</td>
<td>TGCTGCTCATAGACTAGCGTTGG</td>
<td>At1g17730</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>WRKY18-F1.2</td>
<td>TGCTGCTCATATATCTGGAGGT</td>
<td>At1g17730</td>
<td>A nested forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>ESCRT-F1</td>
<td>TCGGAAAGAAACACGCTTGAGA</td>
<td>At3g24050</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>ESCRT-F1.2</td>
<td>CCAGAAGAACACGCTTGAGA</td>
<td>At3g24050</td>
<td>A nested forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>GATA-F1</td>
<td>TCAACTTTTCACGCGGAGTAA</td>
<td>At5g46110</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>GATA-F1.2</td>
<td>AACCCTAAACGCCGTTGTCTGA</td>
<td>At5g46110</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>GATA-F1.3</td>
<td>CCGACGACCTTGTAACTTCCC</td>
<td>At5g46110</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>GATA-F2</td>
<td>AGACGACGACGAACACGCAAA</td>
<td>At5g46110</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>CTPT-F1</td>
<td>TGCAGCTGGCTACCAATACCTT</td>
<td>At5g46110</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>CTPT-F2</td>
<td>GCGGTCGACAGTTCATAAAACA</td>
<td>At5g46110</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>T18C6.1-F1</td>
<td>AATGAGTCCCGCCTGAAGAAA</td>
<td>At2g07698</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>T18C6.1-F2</td>
<td>AGTGACGAGCCTATGCAAACA</td>
<td>At2g07698</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>Primer ID</td>
<td>Sequence</td>
<td>Gene ID</td>
<td>Description</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>ATMED14</td>
<td>AGTCATGGAGGTCGAGCTTGGT</td>
<td>At3g04740</td>
<td>A forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>FP</td>
<td>GATATGGATCGAGGTGTTAG</td>
<td>At5g01530</td>
<td>A common forward primer to confirm APA of its gene</td>
</tr>
<tr>
<td>RP11</td>
<td>TTTTTTTTTTTTTTTTNNGCCCAAAAGTC</td>
<td>At5g01530</td>
<td>A half-specific primer to confirm pA1 of its gene</td>
</tr>
<tr>
<td>RP01</td>
<td>AAAAAAAAAAAAAAAAAAGCCCCCCAGTC</td>
<td>At5g01530</td>
<td>A negative control of the half-specific primer to confirm pA1 of its gene</td>
</tr>
<tr>
<td>RP12</td>
<td>TTTTTTTTTTTTTTTTTTAGTAAATAAGGAG</td>
<td>At5g01530</td>
<td>A half-specific primer to confirm pA2 of its gene</td>
</tr>
<tr>
<td>RP02</td>
<td>AAAAAAAAAAAAAAAAAATAGTAAATAAAGGAG</td>
<td>At5g01530</td>
<td>A negative control of half-specific primer to confirm pA2 of its gene</td>
</tr>
<tr>
<td>RP022</td>
<td>ATAGTAAATAAAGGAG</td>
<td>At5g01530</td>
<td>A negative control of half-specific primer to confirm pA2 of its gene</td>
</tr>
<tr>
<td>RP13</td>
<td>TTTTTTTTTTTTTTTTTTTTTTTTTNTATGCGAC</td>
<td>At5g01530</td>
<td>A half-specific primer to confirm pA3 of its gene</td>
</tr>
<tr>
<td>RP03</td>
<td>AAAAAAAAAAAAAAAAAAAAAAAAAANNTATGCGAC</td>
<td>At5g01530</td>
<td>A negative control of half-specific primer to confirm pA3 of its gene</td>
</tr>
<tr>
<td>RP_pA1+2+3</td>
<td>GATGATGGTGGTGTGGAGTG</td>
<td>At5g01530</td>
<td>A reverse primer to amplify transcripts regardless of APA selection</td>
</tr>
<tr>
<td>RP_pA3</td>
<td>GAAAAATGTACGACGAGTAGGTATAG</td>
<td>At5g01530</td>
<td>A reverse primer to amplify transcripts using either pA2 or pA3</td>
</tr>
<tr>
<td>RP_pA2+3</td>
<td>GCCCCAATTTAATTGTGTCATGGGCC</td>
<td>At5g01530</td>
<td>A reverse primer to amplify transcripts using pA3</td>
</tr>
<tr>
<td>Tub4-F</td>
<td>GAGGGAGCCATTGACAACATCTT</td>
<td>---------</td>
<td>A primer for the reference gene in real time PCR</td>
</tr>
<tr>
<td>Tub4-R</td>
<td>GCGLACAGTTCACAGCTATGTTCA</td>
<td>---------</td>
<td>A primer for the reference gene in real time PCR</td>
</tr>
</tbody>
</table>
later performed by BigDye® Terminator v3.1 Cycle Sequencing Kit from Applied Biosystems, and then aligned to the genome using BLAST.

**Primer design for quantitative real time PCR**

An oligo-dT adaptor primer was used in 3’-RACE to confirm major poly(A) sites for individual genes (Figure 3); however, it cannot be used in qPCR to tell the specific abundance of different transcript isoforms, because it can anneal to the DNA template no matter where poly(A) sites locate. To solve this problem, sequence-specific primers (i.e. 5’- GAAAATGTACGAGTAGGTATAG-3’) locating upstream of the target poly(A) sites were used as reverse primers. In this case, PCR using a sequence-specific reverse primer was able to amplify certain transcript isoforms using poly(A) sites downstream of the primer annealing positions. Primers are listed in Table 1.

**Quantitative real time PCR analysis of APA usage switch**

qPCR was performed in gene *LHCB4.1* (At5g01530) with three biological replicates, each of which has three technical replicates for measurement by using Bio-Rad IQ cycler apparatus with iQ™ SYBR® Green Supermix from Bio-Rad Inc. These data were normalized by the threshold cycle number of the reference gene *Tublin6* under their corresponding conditions. The abundance of *LHCB4.1* transcript isoforms using RP_pA1+2+3, RP_pA2+3, and RP_pA3 denoted x, y, and z, respectively. Therefore, the percentage of pA1 usage was calculated by this equation:

\[
\text{Percentage of pA1 usage} = \frac{x - y}{x} \times 100\%
\]

The pA2 usage percentage was calculated by this equation:

\[
\text{Percentage of pA2 usage} = \frac{y - z}{x} \times 100\%
\]

The pA3 usage percentage was calculated by this equation:

\[
\text{Percentage of pA3 usage} = \frac{z}{x} \times 100\%
\]
Vertical line marks the unique poly(A) site from large scale of sequencing (Wu et al., 2011). The number above the vertical line denotes the supported number of tags for that poly(A) site. Red triangles mark the poly(A) sites confirmed by 3' RACE. The number below the triangle denotes the supported number of sequences for each poly(A) site. Light blue boxes and deep blue boxes represent the untranslated regions (UTR) and coding sequences (CDS), respectively. The lines between two CDS are the introns, while the lines after 3'-UTR are the extended region.

Figure 3 Poly(A) sites confirmed by 3' RACE
**Thylakoid membrane preparation**

Thylakoid membranes were isolated as described with minor modifications (Suorsa et al., 2006). Plants were dark-adapted for 30 min before protein extraction. Mature leaves were harvested at 9-10 am and were excised in ice-cold grinding buffer (50 mM HEPES/KOH, pH 7.5, containing 0.33 M sorbitol, 1 mM MgCl2, 1 mM MnCl2, 2 mM EDTA, 5 mM Na-ascorbate, and 1% BSA). The resulting lysate was filtered through one layer of miracloth and centrifuged at 2500g for 4 min at 4°C The pellet was resuspended and centrifuged twice in 1X Import Buffer (50 mM HEPES/KOH, pH 8, containing 33 mM sorbitol). The final pellet was resuspended in a small volume of the same 1X Import buffer to a concentration of 0.1 mg chl/ml.

**Determination of chlorophyll content**

Chlorophyll of each isolated thylakoid membrane was extracted with 80% aqueous acetone, and the amount of chlorophyll was determined as described (Mackinney, 1941). Absorption under two wavelengths of 645 nm and 663 nm were measured by using spectrophotometer. The chlorophyll content (c) was determined by the following equation:

\[
c = \frac{20.2 \times A_{645} + 8.02 \times A_{663}}{V_1 \times 1000} \times V_2 \times V_3
\]

V1=Chlorophyll Volume (ml)

V2=Acetone Volume (ml)

V3=Volume of Resuspended Isolated Thylakoid Membrane (ml)

**Non-denaturing blue native PAGE**

Thylakoid preparations were solubilized with 1% digitonin (100 ul 1X Import Buffer (50 mM HEPES/KOH pH8, 330 mM sorbitol), 40 ul 10% digitonin, 60 ul H2O). After solubilization, they were agitated for 1 hour at 4 °C and then spin at 100,000 x g for 30 min at 4 °C. Electrophoresis was performed using 5%-13.5% gradient blue native polyacrylamide gel (PAGE) as described in Rokka et al. (2005). Gels were loaded on an
equal chlorophyll content (2 µg) basis measured as described above. Samples of an
equal amount of chlorophyll were loaded in each lane.

**Western blotting**

After Blue Native PAGE, the proteins were incubated in 0.1% SDS, transfer buffer
(25mM Tris, 192mM Glycine, 20% Methanol) for 10 min at RT. Proteins were then
transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore) at a constant
voltage of 50V for 45 min. For immunodetection of proteins on a PVDF membrane,
specific rabbit polyclonal antibody against LHCB4.1 from Agrisera (dilution: 1:5000) and
conjugated anti-rabbit IgG was used as a secondary antibody which is suited for
chemiluminescence detection (Xu et al., 2006b).

**Results**

**Confirmation of APA sites in seven genes in WT**

Based on genome-wide analysis, more than 70% of *Arabidopsis* genes have more than
one poly(A) site (Wu et al., 2011). Most of the poly(A) sites tallied by the deep
sequencing analysis for the 7 targeted genes were supported by the RACE results
Figure 3 vertical bars and Table 2.) However, because of ubiquitous microheterogeneity
in plants (Shen et al., 2008b; Tian et al., 2005), poly(A) sites within adjacent 24 nts were
clustered into a group, which was defined as a poly(A) cluster (Wu et al., 2011). The
unique center poly(A) site of such clusters were used as its representative of its poly(A)
cluster. Therefore, these Sanger sequencing confirmed sites are in the vicinity of the
poly(A) sites. Because individual poly(A) sites were confirmed from individual Sanger
sequencing, the number of sequences that supports each poly(A) by 3’RACE (Figure 3
numbers below red triangles) is lower than the number of poly(A) tags that supports
each poly(A) cluster (Figure 3 numbers above the vertical lines). Each of these genes
with confirmed APA has interesting and unique biological functions (TAIR 9, Table 2).
Table 2 Description of main biological functions of genes with APA sites confirmed

The resource used is TAIR 9 website (http://www.arabidopsis.org/index.jsp).

<table>
<thead>
<tr>
<th>Gene Locus ID</th>
<th>Main Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Other name)</td>
<td></td>
</tr>
<tr>
<td>AT4G31800</td>
<td>Involved in defense response to Pseudomonas syringae</td>
</tr>
<tr>
<td>(WRK18)</td>
<td>Potential miRNA-781 target site between adjacent poly(A) sites.</td>
</tr>
<tr>
<td>AT1G17730</td>
<td>Encodes ESCRT-rated proteins which are auxin carriers.</td>
</tr>
<tr>
<td>(CHMP1B)</td>
<td>Potential miRNA-417 and miRNA-400 target sites between adjacent poly(A) sites.</td>
</tr>
<tr>
<td>AT3G24050</td>
<td>Encodes a member of the GATA factor family of zinc finger transcription factors.</td>
</tr>
<tr>
<td>(GATA1)</td>
<td>Potential miRNA-835 target site between adjacent poly(A) sites.</td>
</tr>
<tr>
<td>AT5G46110</td>
<td>Encodes Chloroplast Triose Phosphate Translocator.</td>
</tr>
<tr>
<td>(APE2)</td>
<td>Mutant has altered acclimation responses;</td>
</tr>
<tr>
<td></td>
<td>Potential miRNA -418 target site between adjacent poly(A) sites.</td>
</tr>
<tr>
<td>AT5G01530</td>
<td>Encodes chlorophyll A-B binding protein (CP29) in photosystem II</td>
</tr>
<tr>
<td>(LHCB4.1)</td>
<td>Involved in photosynthesis; Response to blue, red, and far red light.</td>
</tr>
<tr>
<td>AT2G07698</td>
<td>Encodes ATPase, F1 complex, Alpha subunit protein</td>
</tr>
<tr>
<td>(ATPase)</td>
<td>Involved in hydrogen ion transporting ATP synthase activity, rotational mechanism, poly(U) RNA binding, and zinc ion binding</td>
</tr>
<tr>
<td>AT3G04740</td>
<td>Encodes a protein with similarities to subunits of Mediator Complex, required for RNA POLII recruitment at target promoters in response to specific activators.</td>
</tr>
<tr>
<td>(Mediator Component 14)</td>
<td>Its mutant shows reduction in cell numbers in aerial organs.</td>
</tr>
</tbody>
</table>
Reverse primer design for qPCR

After confirmation of the three unique poly(A) sites of the LHCB4.1 (At5g01530) gene (Figure 3), relative abundance of the three corresponding LHCB4.1 transcript isoforms were further tested by qPCR after reverse transcription. Since amplification using an oligo-dT reverse primer cannot separate different transcript isoforms of this gene, a set of reverse primers with a common forward primer (FP) which was located in the coding region were used.

The first reverse primer RP_pA1+2+3 was located in the coding region to amplify all transcript isoforms using any APA. Amplicons abundance by using RP_pA1+2+3 (see materials and methods, denoted as x) was considered as LHCB4.1 total mRNA level or transcript isoforms using any of pA1, pA2, or pA3 sites (Figure 4). The second reverse primer RP_pA2+3 was located between pA1 and pA2. Amplicons abundance by using RP_pA2+3 (see materials and methods, denoted as y) was considered as level of LHCB4.1 transcript isoforms using either pA2 or pA3. The third reverse primer RP_pA3 was located between pA2 and pA3. Amplicons abundance by using RP_pA3 (see materials and methods, denoted as z) was considered as level of transcript isoforms using only pA3. Therefore, when we subtracted z from y, we have pA2 abundance (level of LHCB4.1 transcript isoforms using only pA2). Similarly, when we subtracted y from x, we have pA1 abundance (level of LHCB4.1 transcript isoforms using only pA1; see materials and methods). All these amplicons have length of 400-500bp.

Differential usage of APA in LHCB4.1 in WT versus the oxt6 mutant

The large extent of alternative polyadenylation in Arabidopsis is revealed by genome-wide analysis, in which usage of APA sites has been shown to vary in a tissue-specific manner (Wu et al., 2011). Also, the difference in APA selection in the oxt6 mutant and oxt6 complemented has been confirmed in particular genes, such as At5g36910, At3g09390, At5g38410, and At5g36910 (Zhang et al., 2008). Here, I tested the APA usage of LHCB4.1 gene in oxt6 mutant (Figure 5). Quantitative PCR results showed no significant difference in total LHCB4.1 mRNA level in between oxt6 and WT (P-value=0.12, Figure 5A), but the APA was used dramatically differently. The predominant
Figure 4 Primer design for determination of different transcript isoforms abundance by qPCR.

Vertical lines denote the three poly(A) sites, pA1, pA2, and pA3, all of which are 40-50nt away from each other. FP was used as the common forward primer; RP_pA1+2+3, RP_pA2+3, and RP_pA3 were used as three reverse primers. The first reverse primer RP_pA1+2+3 was located in the coding region to amplify all transcript isoforms using any APA. The second reverse primer RP_pA2+3 was located between pA1 and pA2 to amplify *LHCB4.1* transcript isoforms using either pA2 or pA3. The third reverse primer RP_pA3 was located between pA2 and pA3 to amplify transcript isoforms using only pA3.
Figure 5 Abundance of different \textit{LHCB4.1} transcript isoforms using different APA under normal growth conditions in \textit{oxt6}.

The first pair of columns shows total \textit{LHCB4.1} mRNA level regardless of APA selection. The other three pairs of columns indicate abundance of \textit{LHCB4.1} transcript isoforms using pA1, pA2, and pA3, respectively. The dominant APA site-pA1 (85% of usage, \textit{p}-value<0.01) in WT is switched to pA2 (\textit{p}-value<0.01) in \textit{oxt6} under the same growth condition. Also, pA3 is used more in \textit{oxt6} compared with that in WT.
APA site in WT was pA1 (85%, $P$-value<0.01, Figure 5B), while the predominant APA was switched to pA2 (87%, $P$-value<0.01) in oxt6. pA3 was used more frequently in oxt6 (11%) than that in WT (2%). This indicates differential usage of APA in \textit{LHCB4.1} in oxt6, compared to wild-type, giving another individual gene example showing the effect of the deficit in AtCPSF30 on APA usage.

\textbf{Impact on \textit{LHCB4.1} mRNA level by abiotic stresses}

To have a comprehensive view of the effect of stresses on APA usage of \textit{LHCB4.1} gene in oxt6, total \textit{LHCB4.1} transcript abundance in WT and oxt6 was first measured. The result of qPCR showed that expression levels of \textit{LHCB4.1} mRNA under different growth stresses, including dark, high light, and cold stresses (Figure 6). Under normal growth conditions no significant difference was detected in \textit{LHCB4.1} mRNA levels in between WT and oxt6. Under all stresses tested above (dark, high light, and cold), \textit{LHCB4.1} mRNA levels decreased in WT and to a greater extent in oxt6. Thus, extreme conditions of either light or temperature significantly reduce \textit{LHCB4.1} gene expression level, and more interestingly, the absence of AtCPSF30 due to mutation of \textit{OXT6}, has an extra effect on repressing \textit{LHCB4.1} transcription.

\textbf{Impact on APA usage of \textit{LHCB4.1} gene under abiotic stresses}

Based on qPCR using primers described above, we can see the change in APA usage between WT and oxt6 when changes in light and/or temperature stimuli occur. Under 2-day of dark stress after 3-week of normal growth conditions, APA usage profile of \textit{LHCB4.1} changed in either WT or oxt6 (Figure 7). In WT, pA1 usage decreased while pA2 usage increased significantly under dark stresses. However, the change in APA usage due to dark stress is different in oxt6 compared with WT. In oxt6, pA1 usage increased while pA2 usage decreased significantly under dark stresses. In addition, under the same dark treatment, APA usage profiles were altered in oxt6. pA1 and pA2 were used equivalently in WT; whereas pA1 was used predominantly in oxt6. The results indicate that APA usage profiles of \textit{LHCB4.1} are obviously different in the presence or absence of light. More interestingly, oxt6 mutant affects the APA selection under dark stress.
Figure 6 Expression levels of \textit{LHC}B4.1 mRNA using any APA under dark, high light, and cold stresses.

Under normal growth conditions, no significant difference was detected in \textit{LHC}B4.1 mRNA level in between WT and \textit{oxt}6. Under all stresses above, \textit{LHC}B4.1 mRNA levels decreased significantly in both WT and \textit{oxt}6, and they decreased even more significantly in \textit{oxt}6 mutant. \textit{LHC}B4.1 mRNA level is much lower in \textit{oxt}6 than WT under dark stress ($P$-value <0.05), high light ($P$-value <0.01), and cold stress ($P$-value <0.01).
Figure 7 APA usage of \textit{LHCB4.1} under dark stress.

A) Bar chart APA usage profiles. B) Pie chart showed APA usage profiles change significantly in both WT and \textit{oxt6} under the same dark treatment.
Under 2-day of high light stress (2000 µE of light irradiance, 4°C) after 3-week of normal growth conditions, APA usage of \( \text{LHCB4.1} \) in WT did not change much; in contrast, APA usage of \( \text{LHCB4.1} \) in \( \text{oxt6} \) changed significantly, which basically restored the APA usage profile in WT under normal growth conditions (Figure 8). Under 2-day of cold stress (4°C) after 3-week of normal growth conditions, APA profiles of \( \text{LHCB4.1} \) did not change obviously in either WT or \( \text{oxt6} \) other than more usage of pA3 (Figure 9). This, compared with the change due to high light stress, indicates that the change in high light stress is mainly due to high light irradiance but not low temperature. It is possible that high light irradiance triggers some factors which restore the functions of \( \text{AtCPSF30} \), so that the APA usage of \( \text{LHCB4.1} \) in \( \text{oxt6} \) under high light stress is close to that in WT under normal growth conditions. Taken together, both \( \text{oxt6} \) mutant and light/temperature stresses affect \( \text{LHCB4.1} \) gene transcription level and its APA usage selection.

**Oxt6 plants turn yellow earlier than WT under dark treatment**

In order to further investigate the potential role of APA and \( \text{oxt6} \) under various stresses, dark stress was studied in more depth from the perspective of physiology. Two to eight-week-old WT and \( \text{oxt6} \) mutant plants were treated in the dark. We found that the leaves of \( \text{oxt6} \) plants turned the same extent of yellow 2-3 days earlier. Similarly, older plants turned yellow earlier than the younger plants. This kind of leaf etiolation was observed qualitatively but not quantitatively. Most plants died after leaf etiolation. An example of 5.5 week-old plants of WT and \( \text{oxt6} \) were shown (Figures 10 and 11). This physiology difference between WT and \( \text{oxt6} \) indicates that \( \text{oxt6} \) may not have efficient photosynthesis system as WT, so that the ability of \( \text{oxt6} \) to resist dark is lower than WT.

**Blue Native PAGE and immunobloting of LHCB4 protein**

To investigate the abundance of LHCB4 protein in \( \text{oxt6} \) mutant, WT, \( \text{oxt6} \), and its complementation lines \( \text{c30G, c30GM, c30YG, and c30YGM} \) plants under the normal growth conditions (4 weeks) and dark treatment (5 days of dark after 4 weeks of normal conditions) were used by conducting blue native PAGE and western blot against LHCB4.
Figure 8 APA usage of *LHCB4.1* under high light stress.

A) Bar chart of APA usage profiles. B) Pie chart of APA usage profiles showed APA usage of *LHCB4.1* in WT did not change much; in contrast, APA usage of *LHCB4.1* in oxt6 changed significantly, which basically restored the APA usage profile in WT under normal growth conditions.
Figure 9 APA usage of *LHCB4.1* under cold stress.

A) Bar chart of APA usage profiles. B) Pie chart showed APA usage profiles were not much changed under the same cold stresses other than higher usage of pA3 in both WT and *oxt6*.
Figure 10 Days for overall leaves to turn half yellow under dark treatment.

The trend lines showed old plants leaves took a shorter time to turn yellow in the absence of light compared to young plants. For the same age of plants, oxt6 plants leaves turn the same extent of yellow 2-3 days earlier than WT.
Figure 11 5.5-week-old plants before and after 8-day dark treatment.
(A) 5.5-week-old WT and oxt6 plants before dark treatment. WT plants have bigger stature and longer bolting than the oxt6 plants of the same age. (B) Both of WT and oxt6 turned yellow after 8 days of dark treatment, in which oxt6 plants were overall yellower than WT. In fact, WT took another 3 days to turn as yellow as the current oxt6 plants had.
Thylakoid were extracted from leaves of each line and then solubilized in a very weak detergent (1% digitonin). All wells were loaded on an equivalent chlorophyll content basis (2 µg). Because only a very weak detergent was used in blue native PAGE, proteins were not denatured. Therefore, chloroplast protein complexes were separated only based on their sizes and structures. Since LHCB4 protein was incorporated into different complexes, it showed multiple bands in western blot after blue native PAGE.

As a result, in blue native PAGE, no obvious chloroplast proteins were detected in any line after dark treatment (Figure 12A). In contrast, three blue bands were detected in lines WT, oxt6, c30G, c30GM under normal growth conditions. Protein abundance of all the three main blue bands was higher in WT than in either oxt6 or the four complementation lines. Interestingly, protein abundance of blue bands was lower in the complementation lines c30YG and c30YGM relative to the other two complementation lines c30G and c30GM. All these blue bands represented certain chloroplast complexes, which needed further experiments to tell the components of each complex. They did not necessarily represent LHCB4 protein; however, the location and abundance of LHCB4 protein can be told by western blot using LHCB4 antibody (Figure 12B). In western blots, no LHCB4 bands were detected in any line under dark treatment; while three LHCB4 bands were detected in either WT, oxt6, or three complementation lines c30G, c30GM, c30YG under normal growth conditions (Figure 12B). The western blot results were consistent with those of the blue native PAGE. The results showed the levels of LHCB4 protein and its associated complex in oxt6 were significantly lower than that in WT, indicating that oxt6 mutant affects the protein abundance of LHCB4 and its incorporated photosynthesis related complexes. This supports the hypothesis that the smaller stature of the oxt6 plant is due to its impact on photosynthesis productivity. More interestingly, the LHCB4 protein abundance and its associated complexes greatly decreased in c30YG and c30YGM lines compared with the other two complementation lines c30G and c30GM.
WT, oxt6, and its complementation lines c30G, c30GM, c30YG, and c30YGM plants under the normal growth conditions (4 weeks) and dark treatment (5 days of dark after 4 weeks of normal conditions) were used for protein analysis. All wells were loaded with an equivalent chlorophyll content of 2ug. (A) As shown in blue native PAGE, no obvious chloroplast proteins were detected in any line after dark treatment. On the contrary, under normal growth conditions, protein abundance of all the three main blue bands was higher in WT than in either oxt6 or the four complementation lines. Interestingly, protein abundance of blue bands was lower in the complementation lines c30YG and c30YGM relative to the other two complementation lines c30G and c30GM. All these blue bands represented certain chloroplast complexes, but did not necessarily represent LHCB4 protein. (B) Western blot against LHCB4 showed the location and abundance of LHCB4 protein. No bands were detected in the dark treatment group (not shown); while three LHCB4 bands were detected in either WT, oxt6, or three complementation lines c30G, c30GM, c30YG under normal growth conditions. Interestingly, the LHCB4 protein abundance and its associated complexes greatly decreased in c30YG and c30YGM lines compared with the other two complementation lines c30G and c30GM. The western blot results were consistent with those of the blue native PAGE.
Discussion

Reverse primer design for determination of specific LHC4.1 transcript isoforms abundance

To identify the authentic APA of LHC4.1 gene, an oligo-dT adaptor primer was used in 3'-RACE, but it cannot tell the abundance of specific transcript isoforms of different lengths, because it can anneal to the DNA template no matter where APA sites locate. Instead, sequence-specific primers locating between adjacent APA sites were used as reverse primers. In this case, when we calculate abundance of particular transcript isoform using specific APA, we had to do a set of subtraction on abundance (See Materials and Methods). It might be challenged that this way of calculation is based on the same efficiency of utilizing different primers in qPCR instrument, which cannot be guaranteed. However, this is currently the best way to determine abundance of specific transcript isoforms, considering poor success in distinguishing different LHC4.1 transcript isoforms by using the way of designing half-specific primers suggested in amplifying FLC gene (Liu et al., 2007).

Mutation of OXT6 gene alters APA selection of LHC4.1 gene

Recent next generation sequencing has revealed a very wide spread of APA phenomenon in Arabidopsis, more than 70% of its genes have multiple poly(A) sites (Wu et al., 2011). In oxt6 mutant which mutates OXT6 gene and thus lacks of a poly(A) factor subunit AtCPSF30 protein (Addepalli and Hunt, 2007), many genes use their APA sites very differently from those in WT (Wu et al., 2011). This has been confirmed in particular genes such as At5g36910, At3g09390, At5g38410, and At5g36910 by 3'RACE and qPCR (Zhang et al., 2008). Here, I tested usage of APA in another gene LHC4.1 in oxt6 under normal growth conditions as well as light/temperature stresses, all results confirmed that APA site choice of LHC4.1 gene was altered significantly in the oxt6 mutant. The three APA sites were confirmed in 3'UTR of the gene LHC4.1. Each of these APA sites only located 40-50 nt away from the other two APAs, but qPCR results showed that its gene LHC4.1 tend to use APA very differently after mutation of OXT6 gene. The predominant APA sites switched from pA1 in WT to pA2 in oxt6 mutant,
indicating a higher probability of including approximately 50 nt longer length of 3‘UTR sequence of this \textit{LHCB4.1} gene in \textit{oxt6} mutant. This length extension raises the possibility of including regulatory features, such as mRNA stability factors and microRNA target sites.

This piece of evidence validates the alternative APA choice is a consequence of the absence or presence of AtCPSF30. In addition, because different mRNA lengths due to different APA selection potentially include or exclude regulatory features, the absence or presence of AtCPSF30 probably also affect activities of many regulatory factors and thus has impact on gene functions to a great extent.

**Smaller stature of the \textit{oxt6} mutant might be due to its impact on photosynthesis productivity**

The \textit{oxt6} mutant is due to T-DNA insertion into the first exon of non-essential \textit{OXT6} gene in \textit{Arabidopsis}, and thus the protein AtCPSF30 encoded by \textit{OXT6} is absent. The \textit{oxt6} mutant presents smaller stature compared with wild-type, but the reason is still unclear. This study reveals that LHCB4 protein abundance is much lower in \textit{oxt6} than either WT or its complementation lines. LHCB4 protein is one of the light harvesting antenna protein in photosystem II, it is critical in excitation state transition between two photosystems for the purpose of maximum photosystem yield (Tokutsu et al., 2009; Iwai et al., 2010). Therefore, when LHCB4 protein abundance is depressed in \textit{oxt6}, it is reasonable to predict the photosynthesis productivity is probably reduced too. On the other hand, leaves of \textit{oxt6} plants underwent etiolation much faster than WT under the dark treatment, indicating the photosystems might be more labile in \textit{oxt6} compared with WT. Taken together, photosynthesis productivity might be reduced in the \textit{oxt6}, therefore, it may be the reason why \textit{oxt6} mutant has a smaller stature.

To address this, photosynthesis efficiency should be measured in WT, \textit{oxt6}, and its complementation lines under the same growth conditions. In addition, to further explore how LHCB4 protein incorporates into different complexes, qualifications of the blue native gel should be performed. Interestingly, the LHCB4 protein abundance and its associated complexes greatly decreased in c30YG and c30YGM lines, compared with
the other two complementation lines c30G and c30GM, based on the western blots results. Since we have already repeated these results, we believe some special characteristics exist in c30YG and c30YGM lines, which leads to future exploration.
References


pattern of around 7,000 Arabidopsis genes under ABA treatments using a full-length cDNA microarray. Funct Integr Genomics 2, 282 - 291.


