Translational Landscape of Photomorphogenic Arabidopsis

Ming-Jung Liu, a,1 Szu-Hsien Wu,a,b,c Jing-Fen Wu,a Wen-Dar Lin,a Yi-Chen Wu,a Tsung-Ying Tsai,a Huang-Lung Tsai,a and Shu-Hsing Wu,a,b,c,2

a Institute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan
b Molecular and Biological Agricultural Sciences Program, Taiwan International Graduate Program, Academia Sinica, Taipei 11529, Taiwan
c Graduate Institute of Biotechnology and Department of Life Sciences, National Chung-Hsing University, Taichung 402, Taiwan

ORCID ID: 0000-0002-7179-3138 (Sh.-H.-W.).

Translational control plays a vital role in regulating gene expression. To decipher the molecular basis of translational regulation in photomorphogenic Arabidopsis thaliana, we adopted a ribosome profiling method to map the genome-wide positions of translating ribosomes in Arabidopsis etiolated seedlings in the dark and after light exposure. We found that, in Arabidopsis, a translating ribosome protects an ~30-nucleotide region and moves in three-nucleotide periodicity, characteristics also observed in Saccharomyces cerevisiae and mammals. Light enhances the translation of genes involved in the organization and function of chloroplasts. Upstream open reading frames initiated by ATG but not CTG mediated translational repression of the downstream main open reading frame. Also, we observed widespread translational repression of microRNA target genes in both light- and dark-grown Arabidopsis seedlings. This genome-wide characterization of transcripts undergoing translation at the nucleotide-resolution level reveals that a combination of multiple translational mechanisms orchestrates and fine-tunes the translation of diverse transcripts in plants with environmental responsiveness.

INTRODUCTION

Gene expression involves producing RNA and/or protein products from genetic codes. However, whole-genome proteomic profiling remains a challenge compared with transcriptomic profiling, which has been promoted by technical advances such as microarray and next-generation sequencing.

The association of transcripts with polysomes implies their active translation status. Microarray profiling of mRNAs associated with polysomes has provided fundamental information for identifying transcripts undergoing translation (Melamed and Arava, 2007; Mustroph et al., 2009). With this methodology, differential translations of mRNAs populations were observed in the model plant Arabidopsis thaliana responding to various external stimuli, including dehydration, elevated temperature, high salinity, oxygen deprivation, Suc starvation, and heavy metal, light, and gibberellin treatment (Kawaguchi et al., 2004; Branco-Price et al., 2005, 2008; Nicolai et al., 2006; Matsuura et al., 2010; Sormani et al., 2011; Liu et al., 2012; Ribeiro et al., 2012). However, this approach simply measures the existence and abundance of polysome-associated transcripts; it does not offer the resolution to reveal which part of the transcript is being translated, which is especially important given the multiple in-frame start codons or separate translation units that are commonly present in eukaryotic transcripts. Indeed, several studies have demonstrated that translational regulation could occur at various levels. These include alternative translation initiation with one transcript producing two proteins targeted to different organelles, microRNA (miRNA)-mediated repression of translation by pausing ribosomes or accelerating ribosome drop-off, and lower reinitiation efficiency of main open reading frames (mORFs) by the translation of upstream open reading frames (uORFs) in the 5’ untranslated region (UTR) (reviewed in Morris and Geballe, 2000; Mackenzie, 2005; Fabian et al., 2010; Huntezinger and Izaurralde, 2011). To fully understand translational regulation, we need to measure the dynamic positional information and abundance of translating ribosomes on a transcript. Information about the translation landscape also offers knowledge on how an organism could use translational regulation to cope with the internal developmental program and respond to external environmental stimuli.

Ribosome profiling, combining RNase protection assays with deep sequencing, has allowed for comprehensive and precise mapping of the positions of translating ribosomes on transcripts (Ingolia et al., 2009). Studies of yeast (Saccharomyces cerevisiae) and mammals have revealed sequence features associated with translational regulation under various conditions (Ingolia et al., 2009; Guo et al., 2010; Ingolia et al., 2011; Brar et al., 2012). However, a genome-wide map of active translation in plants has not been reported.

In Arabidopsis, the profiling of mRNAs associated with polysomes has revealed transcripts with translation regulated by environmental light signals (Piques et al., 2009; Juntawong and Bailey-Serres, 2012; Liu et al., 2012). In this study, we used ribosome profiling to build a high-resolution translational map of...
transcripts in Arabidopsis seedlings undergoing photomorphogenic development. We assessed how external light signals coordinate with internal cis-sequence features to regulate the translation of expressed genes in Arabidopsis.

RESULTS

Monitoring Translation with Single-Nucleotide Resolution in Photomorphogenic Arabidopsis

We used ribosome profiling to monitor dynamic translational changes in Arabidopsis undergoing photomorphogenic development, when a dark-grown seedling first senses light signals (Figure 1A). Arabidopsis 4-d-old etiolated seedlings were exposed to light, and the aerial parts were harvested at 0 min (dark) and after 4-h light (L4h). RNA gel blot analysis revealed that ribosome-protected fragments (RPFs; mRNA_{rp}) in Arabidopsis were ~30 nucleotides (see Supplemental Figure 1 online), similar to that in yeast and mammals (Ingolia et al., 2009; Guo et al., 2010).

We obtained ~5 and ~40 million unique mapped reads for mRNA_{rp} and the randomly fragmented mRNA fragments (steady state mRNA [mRNA_{ss}]), respectively, for replicate 1 (see Supplemental Figures 2A and 2B online). Two independent biological replicates showed high data reproducibility for the 15,384 expressed genes at both mRNA_{ss} and mRNA_{rp} levels ($R^2 = 0.8$ to 0.88; Figure 1B). Despite a significant increase in mRNA_{rp} reads in replicate 2 (~15 and ~41 million unique mapped reads for dark and L4h samples, respectively), similar numbers of genes with comparable read densities were identified (see Supplemental Figure 3A online). This indicated that both replicates exhausted the majority of translating mRNAs in photomorphogenic Arabidopsis seedlings. For most transcripts, mRNA_{rp} densities (in reads per kilobase per million [RPKM]) positively correlated with the mRNA_{ss} density (see Supplemental Figure 3B online) in both replicates.

The reads for mRNA_{ss} were evenly distributed on UTRs and coding sequences (CDSs), but reads for mRNA_{rp} were predominantly mapped to the CDS, especially surrounding the annotated initiation and stop codons (results for both biological replicates in Figure 2 and Supplemental Figure 4 online). This finding is consistent with the differential association of ribosomes with CDS (Ingolia et al., 2009, 2011; Guo et al., 2010) and supports the idea that we have obtained successful ribosome profiling results.

A translating ribosome moves three nucleotides for each translational cycle. Position analyses of aligned reads on the three nucleotides of each codon revealed a three-nucleotide periodicity for only mRNA_{rp} (see Supplemental Figure 5A online) as observed in yeast and humans (Ingolia et al., 2009; Guo et al., 2010). In contrast with the higher ribosome density for the first 30 to 40 codons in yeast (Ingolia et al., 2009), ribosome density was lower in the first 30 codons with both dark and L4h treatment in Arabidopsis (see Supplemental Figure 5B online).

In yeast, more rare codons are found immediately after the start codon. This sequence feature leads to a lower moving speed of the translating ribosomes and a higher ribosome density in the first 30 to 50 codons, known as a ramp phenomenon (Tuller et al., 2010). Although a similar ramp feature was also observed for Arabidopsis transcripts (see Supplemental Figure 5C online), this region showed lower ribosome density (see Supplemental Figure 5B online). Similarly, the ramp feature was not associated with high ribosome density in mouse embryonic stem cells (Ingolia et al., 2011). Alternatively, the addition of the translational inhibitor cycloheximide at the time of polysome extraction may leave a time window allowing for continuing elongation of translating ribosomes.

These results suggest that the translating ribosomes in Arabidopsis, yeast, and mammals possess both conserved and distinct action characteristics when moving along transcripts.
Light Enhances Gene Expression by Adjusting Ribosome Density

Ribosome occupancy refers to the proportion of the transcript associated with polysomes, whereas ribosome density indicates the number of ribosomes on transcripts (Arava et al., 2003; Lackner and Bähler, 2008; Piques et al., 2009). Light-regulated translation could be achieved by adjusting both the ribosome occupancy and ribosome density (Piques et al., 2009; Junta-wong and Bailey-Serres, 2012; Liu et al., 2012). These previous studies mostly revealed mRNA species with significant changes in ribosome occupancy but lacked the resolution to determine the actual distribution of translating ribosomes on a given transcript. For example, steady state transcript levels and ribosome occupancy of AT5G47110.1 and AT4G02510.1 did not significantly differ between the dark and L4h samples in our previous study (Liu et al., 2012). However, these two transcripts were preferentially translated under light by increasing the ribosome density (mRNA_{RP}) (Figure 3A; see Supplemental Figure 6A online). Two transcripts, AT3G25800.1 and AT3G17770.1, not regulated at the translational level are shown for comparison (Figure 3B; see Supplemental Figure 6B online).

Ribosome profiling of deetiolating Arabidopsis allowed us to identify genes significantly upregulated (z-score >2) or downregulated (z-score <−2) in translation in both biological replicates (Figure 3C; see Supplemental Figure 6C and Supplemental Data Set 1 online). Many of these genes were not identified in our previous study because of their comparable ribosome occupancy between dark and L4h samples (marked in red in Supplemental Data Set 1 online). Thus, ribosome profiling could effectively identify new target transcripts under translational regulation, especially those with significant changes in ribosome density. To address the functional role of these new target genes, we performed the gene ontology analyses in the category of biological...
process. The newly identified genes upregulated translationally by light were mostly involved in processes related to the organization and function of chloroplasts (P value < 1E-05; Figure 3D). Our previous study showed that genes dedicated to photosynthesis were preferentially translated via increasing their ribosome occupancy (Liu et al., 2012). These results highlight that different levels of translational regulation together promote the efficient biogenesis of proteins for functional chloroplasts and photosynthesis.

In general, genes upregulated translationally had greater mRNAss density than those downregulated translationally. However, the high mRNAss density did not guarantee a high ribosome density (mRNArp) (Figure 3E; see Supplemental Figure 6D online).

We observed a slight light-induced increase of mRNAss density for the upregulated genes, but the increase could not account for the marked increase in mRNArp density (Figure 3E; see Supplemental Figure 6D online). By contrast, genes downregulated translationally showed significantly lower mRNAss density than those upregulated by light or all expressed genes. For downregulated genes, dark and L4h samples showed comparable mRNAss density, but light specifically reduced their ribosome density (Figure 3E; see Supplemental Figure 6D online). These data suggested that, independent of transcript abundance, a light-dependent and selective mechanism exists for increasing or decreasing the ribosome density of specific transcripts (Figure 3E; see Supplemental Figure 6D online).
Global Identification of Expressed uORFs

Previous studies had predicted the presence of uORFs in 5’ UTRs of 20 to 60% of plant transcripts (Pesole et al., 2000; Hayden and Jorgensen, 2007; Kim et al., 2007). However, whether these annotated uORFs can engage in translation or function to regulate gene expression has not been globally studied in Arabidopsis. We used our ribosome profiling data to differentiate between expressed and unexpressed uORFs with ATG as an initiation codon (see Supplemental Figure 7 and Supplemental Data Set 2 online). In all, 1996 expressed genes contained 3177 expressed uORFs on their 5’ UTRs, but 3615 expressed genes contained 6927 unexpressed uORFs. mRNA<sub>ρρ</sub> density was significantly higher for the expressed uORFs than the flanking 5’ UTR sequences, but mRNA<sub>SS</sub> reads did not show a distribution preference (see Supplemental Figure 8A online). mRNA<sub>ρρ</sub> reads also showed a three-nucleotide periodicity associated with uORFs (see Supplemental Figure 8B online). These shared features between uORFs and the annotated open reading frames (ORFs) (Figure 2A; see Supplemental Figure 8A online) suggested that the expressed uORFs we identified were indeed under active translation in deetiolating Arabidopsis.

We next sought unique features associated with the translating uORFs. As compared with unexpressed uORFs, expressed uORFs were longer, more distal to the 5’ termini, and closer to the downstream annotated initiation codon ATG (Figure 4A; see Supplemental Figure 9A online). The shorter distance between the expressed uORFs and the downstream initiation codon was not due to the uORFs residing on transcripts with shorter 5’ UTRs. Rather, uORFs on transcripts with longer 5’ UTRs were more prone to be translated (Figure 4A; see Supplemental Figure 9A online).

Previous studies have shown a strong bias toward nucleotide A or G at −3 or +4 relative to translation start site, termed the Kozak sequence (Kozak, 1986, 1987a). Sequence context analysis showed that, compared with unexpressed uORFs, expressed uORFs had a significantly higher proportion of nucleotide G both in positions −3 and +4 (top panels in Figure 4B and Supplemental Figure 9B online). The expressed uORFs did not have a strong Kozak sequence context as seen for downstream mORFs (Figure 4B; see Supplemental Figure 9B online), which may reflect the regulatory nature of uORFs. Expressed uORFs and mORFs showed some differences in codon usage (Figure 4C; see Supplemental Figure 9C online; Pearson correlation r = 0.3 to 0.33), but this was largely due to the sequence

![Figure 4. Sequence Characteristics of Expressed uORFs in Arabidopsis Transcripts.](image-url)

(A) Box plots of expressed and unexpressed uORFs by mRNA features, including uORF length, distance to 5’ terminus, distance to annotated ATG, and 5’ UTR length of transcripts with uORF(s). P values were determined by Kolmogorov-Smirnov test. Data for 10th to 90th percentiles for each category are as in Figure 2. nt, nucleotides.

(B) Sequence logo results show the probability of base composition for sequence contexts of uATGs and annotated ATGs. Sequence logos for the expressed or unexpressed uORFs and their downstream mORFs are shown separately. * G is overpresented at the −3 position (P value = 2.8E-04) ** G is overpresented at the +4 position (P value = 7.1E-15). P values were determined by one-tailed Fisher’s exact test.

(C) Comparison of codon usage for annotated ORFs and 3067 expressed uORFs (red square), 1996 mORFs with expressed uORFs (gray square), or 5’ UTRs of 1996 mORFs with expressed uORFs (black square). Usage frequency of the 61 codons (excluding the three stop codons) was ranked from low to high according to usage frequency in annotated ORFs.
composition of the 5’ UTRs rather than to a selective presence of specific codons in uORFs (Figure 4C; see Supplemental Figure 9C online; Pearson correlation r = 0.96 to 0.97 for frequencies of codon usage between 5’ UTR and uORFs).

Translation of ATG-Initiated uORFs Represses the Translation of Main ORFs in a Light-Dependent Manner

In a few Arabidopsis genes studied, uORFs could trigger a translational repression of their downstream mORFs in response to various environmental stimuli (Hanfrey et al., 2002; Wiese et al., 2004; Nishimura et al., 2005; Imai et al., 2006; Alatorre-Cobos et al., 2012; Rosado et al., 2012). We wondered whether this function was widespread in Arabidopsis. Compared with transcripts with unexpressed uORFs, those with increasing numbers of expressed uORFs in their 5’ UTRs showed significantly lower translation efficiencies in both dark and L4h samples (Figure 5A; see Supplemental Figure 10A online). Thus, the association of ribosomes with uORFs could impose translation inhibition on their downstream mORFs. The inhibitory effects of uORFs on the translation of mORFs were more evident for L4h than dark samples (translation efficiency [TE] at L4h/dark; D-value = 1.9E-01, P value = 9.3E-08 for genes with at least three expressed uORFs; right panel in Figure 5A), perhaps to attenuate the light-enhanced translation in L4h observed previously (Liu et al., 2012).

Recent studies in both yeast and mammals suggested that the near-cognate codons, especially CTG, could also be translational initiation codons for uORFs (Ingolia et al., 2011; Brar et al., 2012; Fritsch et al., 2012; Lee et al., 2012). The identification of expressed uORFs starting with CTG in Arabidopsis indicated that this near-cognate initiation codon is also used by Arabidopsis uORFs (see Supplemental Figure 7 and Supplemental Data Set 2 online). Interestingly, TE was similar for mORFs on transcripts with CTG-mediated expressed uORFs and unexpressed uORFs in both dark and L4h samples (Figure 5B; see Supplemental Figure 10B online). Thus, only uORFs initiated with ATG but not CTG could negatively regulate the translation of the downstream mORFs in deetiolating Arabidopsis. Similarly, only ATG-mediated uORFs were found to act competitively on the translation of downstream mORFs in yeast (Brar et al., 2012).

Characterization of Downstream Translation Start Sites

In addition to identifying the expressed uORFs, the features of the RPFs predominantly mapped to translating regions also provides the possibility of revealing the alternative translational start sites downstream of annotated start sites. A higher read coverage at downstream ATGs other than the annotated ATG initiation codon may indicate an alternative translation start site. We identified 35 downstream ATGs in 31 genes (see Supplemental Figure 11 and Supplemental Data Set 3 online). Among them, FARNESYL-DIPHOSPHATE SYNTHASE1 (FPS1) and GLUTATHIONE S-TRANSFERASE PHI8 (GSTF8) were previously described to encode dual-targeting (cytosol and mitochondria) polypeptides (Cunillera

Figure 5. ATG-Initiated uORFs Repress the Translation of the Downstream mORF.

(A) Cumulative curves of the TE (in arbitrary units) of mORFs in transcripts without expressed uORFs or with 1, 2, or ≥3 expressed uORFs. Shown are fold changes in TE for dark and L4h samples for transcripts in four categories and D- and P values from the Kolmogorov-Smirnov test.

(B) Cumulative curves of TE (in arbitrary units [A.U.]) for mORFs in transcripts without expressed uORFs (unATG/unCTG), with expressed ATG (exATG) or expressed CTG (exCTG) initiated uORFs in their 5’ UTRs.
et al., 1997; Thatcher et al., 2007). In deetiolating *Arabidopsis*, FPS1 and GSTF8 were preferentially translated from the downstream ATGs for cytosolic forms (Figure 6A; see Supplemental Figure 12A and Supplemental Data Set 4 online). Our data further supported that a potential dual-targeting gene, *GLUTATHIONE PEROXIDASE6* (GPX6) (Rodriguez Milla et al., 2003), encoding the cytosolic isoform in deetiolating Arabidopsis (top panels in Figure 6B and Supplemental Figure 12B online; see Supplemental Data Set 4 online), and a transcription factor gene, *CONSTANS-LIKE4* (COL4) (bottom panels in Figure 6B and Supplemental Figure 12B online; see Supplemental Data Set 4 online), may produce a protein isoform with N-terminal truncation. Whether the choice of downstream ATGs was determined by alternative transcription sites or alternative transcript processing (Cunillera et al., 1997; Thatcher et al., 2007) remains to be clarified.

**MiRNA Target Genes Have Lower Translational Efficiency**

In plants, miRNAs can negatively regulate gene expression through miRNA-mediated mRNA cleavage and translation inhibition (Aukerman and Sakai, 2003; Chen, 2004; Gandikota et al., 2007; Brodersen et al., 2008; Lanet et al., 2009; Beauclair et al., 2010; Chen, 2010; Yang et al., 2012). However, the impact of miRNAs on translational repression has not been evaluated globally in *Arabidopsis*. To analyze this, we retrieved 1155 expressed genes as targets of 228 expressed miRNA from L4h data (transcripts per million >0; see Supplemental Figure 13 and 13B). For each gene, the annotated gene model is shown underneath the read distribution plot, with the UTR, CDS, annotated ATG, alternative ATG initiation sites, internal downstream ATGs, and stop codon marked with a black line, gray box, gray triangle, red triangle, filled black triangle, and open triangle, respectively. nt, nucleotides; RPM, reads per million.

**Figure 6.** Representative Proteins with Alternative Translation Start Sites in Photomorphogenic *Arabidopsis*.

Read densities for mRNA_{SS} and mRNA_{RP} were plotted for *FARNESYL-DIPHOSPHATE SYNTHASE1* (FPS1) and *GLUTATHIONE S-TRANSFERASE PHI8* (GSTF8) (A) and *GLUTATHIONE PEROXIDASE6* (GPX6) and *CONSTANS-LIKE4* (COL4) (B).
Supplemental Data Set 5 online). We found that miRNA targets had significantly lower TE (Figure 7A; see Supplemental Figure 14A online). A similar conclusion was drawn for 971 expressed genes targeted by 203 expressed miRNAs in the dark (see Supplemental Figure 15A online). The mRNA\textsubscript{rs} density remained comparable for miRNA targets and nontargets (Figure 7B; see Supplemental Figures 14B and 15B online). However, the mRNA\textsubscript{par} density was uniformly lower for miRNA targets than for nontargets spanning the CDS (Figure 7C; see Supplemental Figures 14C and 15C online), so miRNAs may exert their functions in inhibiting translation initiation or elongation, as was seen in Drosophila melanogaster and zebra fish (Dario rario; Bazzini et al., 2012; Djuranovic et al., 2012).

Unlike the attenuating impact of uORFs on light-mediated translation enhancement (Figure 5A), light-regulated global changes in TE were comparable for miRNA targets and nontargets (see Supplemental Figure 15D online). Therefore, L4h treatment may affect the translation of only a subset of or specific miRNA target genes.

**DISCUSSION**

This study has expanded the current understanding of translation control in Arabidopsis. We identify additional light-responsive genes regulated at the translational level, provide molecular evidence for protein isoforms being translated, and show the differential influence of uORFs and miRNAs in the regulation of translation in photomorphogenic Arabidopsis.

**Translating Ribosomes in Plants, Mammals, and Yeasts**

Analogous to yeast and mammals (Ingolia et al., 2009; Guo et al., 2010; Ingolia et al., 2011), in Arabidopsis, we found that translating ribosomes were predominantly located in coding regions, protected ~30 nucleotides and moved in three-nucleotide periodicity (Figures 1 and 2; see Supplemental Figures 1 and 5A online), which suggests that these are conserved features across kingdoms.

The use of the 15th nucleotide of the RPFs in the mapping process revealed a clear offset between the UTR and the annotated initiation codon ATG (Figure 2A; see Supplemental Figure 4 online). This finding suggested the existence of a 14-nucleotide distance between the 5'-termini of protected fragments and the P-sites in Arabidopsis compared with the 12- or 13-nucleotide offset in yeast and mammals (Guo et al., 2010; Ingolia et al., 2011, 2012; Lee et al., 2012). The sizes of sequence fragments (27 to 30 nucleotides) in Arabidopsis are similar to those in yeast and mammals (Guo et al., 2010; Ingolia et al., 2011, 2012; Lee et al., 2012), which is unlikely to cause this slight offset. Alternatively, the differential composition of ribosomal proteins or rRNAs (18S/25S in Arabidopsis, 18S/26S in yeast, and 18S/28S in mammals) may contribute to different ribosome conformations, thus protecting mRNAs with slightly different sizes. Also, the magnesium concentration and ion strength in buffers used for Arabidopsis ribosome profiling could be optimized to further improve the nuclease digestion process, as was previously suggested (Ingolia et al., 2012). Nevertheless, this slight offset should not have affected the expression analyses used in this study.

**Inhibitory Roles of Expressed uORFs on Translation of Downstream mORFs**

Our genome-wide analyses provide systemic identification of the expressed uORFs in Arabidopsis. We also revealed inhibitory roles of ATG-initiated uORFs in the translation of downstream mORFs (Figure 5A; see Supplemental Figure 10A online). Because the expressed uORF(s) and the mORF are separate ORFs on a single transcript (see Supplemental Figure 7 online), a leaky scanning or reinitiation mechanism (Jackson et al., 2010) is needed for successful translation of the downstream mORF. A recent study in humans showed that the reinitiation process was less efficient than was leaking scanning (Lee et al., 2012), which is consistent with our results showing that transcripts with expressed uORFs have lower translational efficiency for mORFs. Also, this translational repression role of uORFs may be associated with their unique features (Figure 4; see Supplemental

![Figure 7](image-url)
Global Translational Repression of miRNA Target Transcripts

Our results showed that miRNA target transcripts have significantly lower translational efficiency, likely because of the even decrease in ribosome density on the CDS (Figure 7; see Supplemental Figures 14 and 15 online). A recent report also demonstrated a predominant translation repression mediated by miRNAs in multiple plant species (Li et al., 2013a). The translation inhibition likely occurs on the endoplasmic reticulum in Arabidopsis (Li et al., 2013b).

The association of miRNAs or an RNA-induced silencing complex and target transcripts undergoing translation may result in a “traffic jam” of ribosomes that would appear as a stalling of ribosomes, possibly at the 5′ end of miRNA target sites. We did not observe such ribosome-stalling sites upstream of putative miRNA target sites, which is consistent with the lack of ribosome pausing in heterochronous miRNA target transcripts in Caenorhabditis elegans (Stadler et al., 2012).

miRNAs were found to act first on mRNA translation, then decapping or decay of mRNAs (Bazzini et al., 2012; Djuranovic et al., 2012). A recent study concluded that miRNA targeting at the CDS leads to translational inhibition, likely because targeting at the 3′ UTR preferentially results in mRNA degradation (Haussler et al., 2013). In our study, despite the clear decrease in mRNA<sub>100</sub> level for miRNA targets (Figure 7C; see Supplemental Figures 14C and 15C online), mRNA<sub>100</sub> levels did not differ between miRNA targets and nontargets (Figure 7B; see Supplemental Figures 14B and 15B online), which is consistent with the predominant presence of miRNA target sites in CDS regions (77%). In plants, miRNA-mediated cleavage of target transcripts account for many of the miRNA-triggered gene-silencing events reported previously. By adopting a stringent expression threshold in defining expressed miRNA target genes in our analyses (see Supplemental Figure 13 online), we may have eliminated genes with low expression resulting from miRNA-mediated decay. Nevertheless, our study clearly highlights a global effect of miRNA-mediated translational repression via reduced ribosome occupancy for most miRNA targets.

METHODS

Preparation of Ribosome-Protected and Randomly Fragmented mRNAs

Arabidopsis thaliana ecotype Columbia-0 plants were grown and treated (dark or L4h) as described previously (Liu et al., 2012). By pouring liquid N<sub>2</sub> directly into Petri dishes, aerial parts of seedlings were harvested, frozen, and stored at −80°C. The Arabidopsis polySome complexes were extracted with polySome extraction buffer (200 mM Tris-HCl, pH 8, 50 mM KCl, 25 mM MgCl<sub>2</sub>, 100 μg/mL cycloheximide, 2% polyoxyethylene 10 tridecyl ether, 1% deoxycholic acid, and 1 mM DTT). The digestion and purification of RPFs were as described (Ingolia et al., 2011), except that 14 mL of polySome extract was digested with RNase I (10 units/μg RNA; New England Biolabs) at room temperature for 1 h. The reaction was stopped by adding Superase In (4 units/μg RNA; Ambion) and immediately loaded on a 1 M Suc cushion containing 0.1 units/μL Superase In for pelleting the ribosomes. The SDS/phenol method was used to purify the RPFs.

Total RNA isolated as described previously (Chang et al., 2008) was used for total mRNA purification with the Illustra mRNA purification kit (GE Healthcare). The mRNA was fragmented with NEBNext RNA fragmentation reaction buffer (NEB) at 94°C for 40 min.

cDNA Library Construction and Deep Sequencing

Ribosome-protected or random mRNA fragments were used to construct libraries as described (Ingolia et al., 2011) with the RT primer sequence 5′-GATCGTCGGACTGTAGAACTCTGAACGTGTAGATC(Sp18)CACTC-3′ (Sp18 is an 18-atom hexa-ethylenglycol spacer). Amplification of cDNA involved use of 0.5 units of Phusion polymerase (NEB) with 30-s denaturation at 98°C, then eight cycles of 10-s denaturation at 98°C, 10-s annealing at 60°C, and 5-s extension at 72°C. Deep sequencing involved use of Illumina HiSeq2000 at Yourgene Bio Science (Taipei). For mRNA<sub>3p</sub> reads, dark and L4h samples were separately indexed and sequenced in the same lane. For mRNA<sub>5p</sub> reads, dark and L4h samples were each sequenced in one lane.

Sequencing Data Mapping and Analyses

Reference transcript data sets including the 5′ UTRs, CDS, and 3′ UTRs for 27,416 annotated protein-coding gene models were retrieved from the Arabidopsis Information Resources database (TAIR10; http://www.Arabidopsis.org). Among them, 8572 protein-coding genes lacked annotated UTRs (see Supplemental Figure 16 online). We have updated gene models of some genes lacking 5′/3′ UTRs using our sequencing data sets outlined in Supplemental Figure 16 online. The updated gene models are referred to as the extended protein-coding gene models in this study. For mapping reads to the transcripts of the extended protein-coding gene models, raw reads were processed and mapped as described in Supplemental Figure 2 online. Only unique mapped reads were aligned and included for calculating the read density of each gene. The 15th nucleotide of aligned reads was used for calculating the read count of each individual transcript. AT2G01021.1 was removed from the reference transcript data set because its sequence perfectly matched the Arabidopsis rRNA gene (GenBank: X52320.1) and had a large read count.
in our mRNARP data. The filtering processes and criteria used for identifying translationally regulated genes, position-based read densities, ATG- or CTG-initiated uORFs, the downstream in-frame ATG initiation codon, and miRNA target genes are outlined in Supplemental Figures 2, 7, 11, and 13 online. The usage frequency for each Arabidopsis codon (Figure 4C and see Supplemental Figure 9 online) was from a previous report (Nakamura et al., 2000). As a control for data shown in Supplemental Figure 5C online, the CDSs for each protein coding gene were randomly shuffled by the shuffle function in perl, and the codon usages were calculated as described (Tuller et al., 2010). The TE of mORFs and uORFs (Figures 3, 5, and 7 and see Supplemental Figures 6, 10, 14, and 15 online) was computed by normalizing the densities of mRNA_{up} for mORFs or uORFs to those of mRNA_{SS} for the whole transcript.

RNA Gel Blot Analysis

Total polysomal RNAs treated with or without RNase I from dark or L4h samples were separated on 15% denaturing polyacrylamide Tris-borate ethylenediaminetetraacetic acid-Urea gel (Invitrogen) and transferred to nylon membrane (Hybond N+; Amersham GE Healthcare) using a transblot apparatus. Total polysomal RNAs treated with or without RNase I from dark or L4h ecotype Columbia-0 plants were grown, treated (dark or 3-h light), and harvested as described above. Total RNA was isolated with use of the mirVana miRNA isolation kit (Invitrogen). The size-fractionated small RNAs were used to generate libraries for sequencing by the Illumina platform. Mapping of the sequencing reads to known miRNAs resulted in 203 (dark) and 228 (3-h light) expressed miRNAs (transcripts per million > 0) used in this study.

Gene Ontology Analyses

The enriched functional groups were revealed with use of the elim method of the TopGO package (Alexa et al., 2006) implemented in the MultiView plugin of GOBU (Lin et al., 2006). The Fisher exact test was used to evaluate the representation differences between genes in the light-upregulated group and those not upregulated by light in the whole genome. Only GO terms with preferential representation in the category of biological process (P value < 1E-05) from the light-upregulated group were selected.

Statistical Analyses

The Kalmogorov-Smirnov test downloaded from the Arizona Laserchron Center at the University of Arizona (https://docs.google.com/View?id=dbcr8b2_7c366pxft) was used for calculating D- and P values with the no error option. One-tailed Fisher’s exact test was performed online (http://www.matforsk.no/ola/fisher.htm). Z-scores were calculated by the Standardize formula in Microsoft Excel.

Accession Numbers

Sequencing data from this article can be found in the National Center for Biotechnology Information’s Gene Expression Omnibus under accession number GSE43703. Names and locus numbers for genes described in this study are listed in Figures 3 and 6, Supplemental Figures 6 and 12 online, and Supplemental Data Sets 1 to 5 online.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. RNA Gel Blot Analysis of Ribosome-Protected mRNA Fragments in Arabidopsis.

Supplemental Figure 2. A Flowchart and Summary of the Read Mapping.

Supplemental Figure 3. Two Biological Replicates Provide Comparable Data Resolution.

Supplemental Figure 4. mRNA_{up} Reads Are Enriched in CDS (Rep 2).

Supplemental Figure 5. Positional Analyses of mRNA_{SS} and mRNA_{up} in CDS Regions.

Supplemental Figure 6. Ribosome Profiling Identifies Additional Genes Regulated at the Translational Level in Photomorphogenic Arabidopsis (Rep 2).

Supplemental Figure 7. A Flowchart of the Data Analysis Procedure for Identifying and Categorizing ATG- or CTG-Initiated uORFs.

Supplemental Figure 8. Expressed uORFs Are Active Translating Units.

Supplemental Figure 9. Sequence Characteristics of Expressed Upstream Open Reading Frames in Arabidopsis Transcripts (Rep 2).

Supplemental Figure 10. ATG-Initiated uORFs Repress the Translation of the Downstream mORF (Rep 2).

Supplemental Figure 11. A Flowchart of the Procedure for Identifying Downstream in-Frame Translation Start Codons.

Supplemental Figure 12. Representative Proteins with Alternative Translation Start Sites in Photomorphogenic Arabidopsis (Rep 2).

Supplemental Figure 13. A Flowchart of the Procedure for Identifying MiRNA Target Genes.

Supplemental Figure 14. Widespread Low Translation Efficiency for MiRNA Target Genes in Photomorphogenic Arabidopsis (Rep 2).

Supplemental Figure 15. Widespread Lower Translation Efficiency for MiRNA Target Genes in Etiolated Arabidopsis.

Supplemental Data Set 1. List of Genes Up- (z-score >2) or Down-regulated (z-score < -2) by Light at the Translational Level.

Supplemental Data Set 2. List of Expressed Genes with One or More uORFs.

Supplemental Data Set 3. List of 31 Genes with 35 in-Frame dATGs.

Supplemental Data Set 4. dATGs for Representative Loci in Figure 6 and Supplemental Figure 12.

Supplemental Data Set 5. Lists of the 971 and 1155 Genes Predicted to Be Targets of Expressed MiRNA(s) under Dark and L4h Conditions, Respectively.

ACKNOWLEDGMENTS

This research was supported by a research grant and Investigator Award (to Sh.-H.W.) from Academia Sinica and a fellowship (to Sz.-H.W.) from the Taiwan International Graduate Program.

AUTHOR CONTRIBUTIONS

REFERENCES


**Translational Landscape of Photomorphogenic Arabidopsis**
Ming-Jung Liu, Szu-Hsien Wu, Jing-Fen Wu, Wen-Dar Lin, Yi-Chen Wu, Tsung-Ying Tsai, Huang-Lung Tsai and Shu-Hsing Wu

*Plant Cell* 2013;25;3699-3710; originally published online October 31, 2013;
DOI 10.1105/tpc.113.114769

This information is current as of December 24, 2017

<table>
<thead>
<tr>
<th>Supplemental Data</th>
<th>/content/suppl/2013/10/21/tpc.113.114769.DC1.html</th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>This article cites 68 articles, 28 of which can be accessed free at: /content/25/10/3699.full.html#ref-list-1</td>
</tr>
<tr>
<td>eTOCs</td>
<td>Sign up for eTOCs at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>CiteTrack Alerts</td>
<td>Sign up for CiteTrack Alerts at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>Subscription Information</td>
<td>Subscription Information for <em>The Plant Cell</em> and <em>Plant Physiology</em> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a></td>
</tr>
</tbody>
</table>