Conserved Peptide Upstream Open Reading Frames Act Via Ribosome Stalling to Regulate Translation in Response to Environmental Signals.

Barry Causier[1]*, Tayah Hopes[1‡], Mary McKay[1], Zachary Paling[1], and Brendan Davies[1]

[1] Centre for Plant Sciences, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, U.K.

[‡] Current address: School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, U.K.

* To whom correspondence should be addressed. Tel: +441133432838; Email: b.e.causier@leeds.ac.uk

FUNDING: The research was funded by the U.K. Biotechnology and Biological Sciences Research Council (BB/L006170/1). The project was an industrial partnership with Bayer CropScience that subsequently transitioned to BASF, whom we also thank for additional funding.

ABSTRACT

The regulation of protein synthesis plays a key role in growth and development in all organisms. Upstream open reading frames (uORFs) are commonly found in eukaryotic mRNA transcripts and typically inhibit translation of downstream ORFs, in part by stalling ribosomes. Conserved peptide uORFs (CPuORFs) are a rare subset

of uORFs, some of which conditionally regulate translation. Here we identify three Arabidopsis CPuORFs that specifically regulate translation of any downstream ORF, in response to the agriculturally significant environmental signals, heat shock and water limitation. Mechanistically, we provide evidence that CPuORF translation causes ribosome stalling, in a peptide sequence-dependent manner, attenuating translation of downstream ORFs. We propose a model in which plant CPuORFs are not simply on/off switches for translation, but rather act conditionally, along a continuum, to fine-tune translation dynamically.

KEYWORDS: Arabidopsis, abiotic stress, 5'-UTR, 5'-leader, uORF, CPuORF, translation, regulation, ribosome stalling

ACKNOWLEDGEMENTS: We thank all members of the Davies lab and Dr Younousse Saidi (BASF) for useful discussions throughout this work. We particularly wish to thank Dr Sam Rayson, Lewis Adams, Andrea Subroto and Katie Poulter for technical assistance and the University of Leeds Plant Growth Facility for help growing Arabidopsis. We are also grateful to Dr James Lloyd and Dr Antoine Larrieu for vital feedback during the preparation of the manuscript.

1 INTRODUCTION

Post-transcriptional regulation is a critical means of controlling protein levels. It provides a mechanism to achieve rapid responses to both internal and external stimuli, without the requirement to initiate or repress transcription (Inolia, 2016; Zhang, Wang, & Lu, 2019). In sessile organisms such as plants, the ability to respond immediately to an ever-changing environment is key to normal growth and development. The untranslated regions of transcripts (UTRs) have been implicated in the control of translation. In particular, up to 50% of transcripts from animals, fungi and plants contain potentially translatable upstream open reading frames (uORFs) within their 5'-UTRs (also known as 5'-leader sequence) (Zhang, Wu, Yue, & Zhao, 2020) (Figure 1a). Typically, uORFs attenuate translation of the downstream major open reading frame (mORF), which encodes the main protein product of the transcript (Barbosa, Peixeiro, & Romão, 2013; Johnstone, Bazzini, & Giraldez, 2016; Kurihara, 2020; von Arnim, Jia, & Vaughn, 2014). uORF-mediated translation inhibition can either occur passively, as ribosomes dissociate after uORF translation, or by an active mechanism, in which uORF translation causes ribosome stalling (Kurihara, 2020; von Arnim et al., 2014). Ribosomal arrest may sequester translating ribosomes, blocking their access to downstream ORFs, or may be interpreted as abnormal translation termination, triggering transcript destruction through pathways such as nonsense-mediated mRNA decay (NMD) (Lloyd, 2018; Yamashita, 2013). In either case, translation of the downstream mORF is inhibited.

Numerous studies have demonstrated that the ribosome exit tunnel plays an important role in regulating translation (Bhushan et al., 2010; Liutkute, Samatova, & Rodnina, 2020). Ribosome stalling during uORF translation is a common regulatory

mechanism that operates in prokaryotes and eukaryotes, to control translation of downstream ORFs. Structural analyses of ribosomes translating uORF regulatory peptides reveals that direct interactions between the nascent peptide and the ribosome exit tunnel cause ribosome stalling (Bhushan et al., 2010: Seidelt et al., 2009). uORF-mediated translational regulation has been shown to be conditional in some cases, where specific small molecules determine whether translation is turned on or off (Seip & Innis, 2016). It is emerging that translating uORF peptides, which cause ribosome stalling, can act as receptors for small effector molecules, either within the ribosome exit tunnel or via an extra-ribosomal domain (Ito & Chiba, 2013; von Arnim et al., 2014). Such effectors are diverse and include antibiotics and metabolites in bacteria, and polyamines, amino acids and *S*-adenosylmethionine in eukaryotes (Ito & Chiba, 2013).

uORF-mediated ribosome stalling is dependent upon the sequence of the nascent peptide (Seip & Innis, 2016). Despite this, most uORFs prevalent in eukaryotic transcripts do not appear to be under selective pressure to conserve their encoded amino acid sequence. However, there is a rare subset of uORFs where peptide sequence is conserved over large evolutionary distances, suggesting functional significance. In plants genome-wide comparisons between Arabidopsis (*Arabidopsis thaliana*), rice (*Oryza sativa*) and other angiosperms has led to the identification of predicted transcripts containing one or more conserved peptide uORFs (CPuORFs), including at least 123 transcripts in Arabidopsis (Hayashi et al., 2017; Hayden & Jorgensen, 2007; Jorgensen & Dorantes-Acosta, 2012; Takahashi et al., 2020; Takahashi, Takahashi, Naito, & Onouchi, 2012; Tran, Schultz, & Baumann, 2008; van der Horst, Filipovska, Hanson, & Smeekens, 2020; van der Horst, Snel, Hanson,

& Smeekens, 2019; Vaughn, Ellingson, Mignone, & von Arnim, 2012) (for a database of Arabidopsis CPuORFs, with details on the function and evolution of these elements, see Table S1). Comparison of CPuORF peptide sequences from diverse angiosperm species places them into over 150 homology groups (HGs) (Takahashi et al., 2020). Interestingly, a proportion of CPuORFs discovered in angiosperms have also been identified in earlier diverging plants, including bryophytes and green algae, suggesting that these short sequences have been maintained for an extraordinary length of time (Hayden & Jorgensen, 2007; Takahashi et al., 2020) (Table S1). In addition, based on patterns of sequence conservation, CPuORFs can be divided into two broad classes: class I is characterised by a highly conserved C-terminal region, while in class II the entire CPuORF peptide sequence or the N-terminal and middle regions are conserved (Takahashi et al., 2012) (Table S1). Several plant CPuORFs have been shown to inhibit translation of downstream ORFs (Alatorre-Cobos, Chiappetta, Bruno, & Bitonti, 2012; Ebina et al., 2015; Hanfrey et al., 2005; Laing et al., 2015; Rahmani et al., 2009; Ribone, Capella, Arce, & Chan, 2017; Takahashi et al., 2020; Zhu, Thalor, Takahashi, Berberich, Kusano, 2012), which in some cases is caused by ribosome stalling on the CPuORF (Bazin et al., 2017; Hayashi et al., 2017; Ribone et al., 2017; Uchiyama-Kadokura et al., 2014; Yamashita et al., 2017). Further, we previously identified CPuORFs as a trigger of NMD in plants (Lloyd & Davies, 2013; Rayson et al., 2012).

The biological functions of most plant CPuORFs are not known, although they are generally associated with mORFs encoding regulatory proteins (Hayden & Jorgensen, 2007; Jorgensen & Dorantes-Acosta, 2012). For the handful of plant CPuORFs that have been functionally characterised, they act as conditional

114 regulators of translation, responding to a range of different signals, including sucrose,
115 ascorbate, phosphocholine, polyamines, galactinol, heat shock, pathogen attack and
116 photosynthetic signals (Alatorre-Cobos et al., 2012; Bazin et al., 2017; Guerrero-
117 González, Ortega-Amaro, Juárez-Montiel, & Jiménez-Bremont, 2016; Hanfrey et al.,
118 2005; Imai et al., 2006; Laing et al., 2015; Rahmani et al., 2009; Ribone et al., 2017;
119 Tabuchi, Okada, Azuma, Nanmori, & Yasuda, 2006; Uchiyama-Kadokura et al.,
120 2014; Wiese, Elzinga, Wobbes, & Smeekens, 2004; Xu et al., 2017; Zhu, Li, &
121 Kusano, 2018; Zhu et al., 2012) (summarised in Table S1).

122 Plant CPuORFs are emerging as important post-transcriptional regulators, acting as
123 sensors that control translation in response to changing intracellular or extracellular
124 conditions. Indeed, the potential of CPuORFs as tools for crop improvement are now
125 being realised (van der Horst et al., 2020; Xing et al., 2020; Xu et al., 2017; Zhang et
126 al., 2018). Signal-dependent induction of translation has potential utility in agriculture,
127 synthetic biology and research, so here we ask whether plant CPuORFs can be
128 used to regulate translation in response to stress conditions that are known to
129 significantly impact global crop production. We discovered three CPuORFs that
130 specifically regulate translation in response to heat stress or water limitation.
131 Mechanistically, we show that most Arabidopsis CPuORFs are translated and block
132 translation of downstream ORFs by sequestering ribosomes. We conclude that
133 CPuORFs encode self-contained regulatory units that include sequences required to
134 conditionally modulate translation. In addition, as proof-of-principle, we demonstrate
135 the utility of a previously described chemically responsive CPuORF in regulating
136 developmental outcomes. Together our data demonstrates the general applicability

of CPuORFs as versatile tools for inducible gene expression with applications both in the laboratory and in the field.

2 MATERIAL AND METHODS

2.1 Plant growth

Standard plant growth for Arabidopsis wild-type or transgenic lines was on peat-free soil or on 0.5xMS plates (with or without appropriate treatments as described below) at 21°C, under long days (16h light/8h dark).

2.2 Plasmid construction

All oligonucleotides used in this study are described in Table S2. To clone Arabidopsis or rice CPuORF-containing 5'-UTRs, total RNA was prepared from leaves using the Qiagen Plant RNeasy Mini Kit, according to the manufacturer's instructions (Qiagen). cDNA was prepared from total RNA (1µg) using the SuperScript II reverse transcriptase (Life Technologies). All plant constructs were transformed into Agrobacterium (*Agrobacterium tumefaciens*) strain GV3101, and Arabidopsis plants were subsequently transformed using the floral dip method (Clough & Bent, 1998). Transgenic seed was selected using the GFP seed-coat marker provided on the pALLIGATOR3 vector (Gateway modified pFP101; Bensmihen et al., 2004).

For the *in planta* constructs, 5'-UTRs were PCR amplified using Phusion DNA polymerase (Thermo Scientific) and appropriate primers containing attB1 (forward) or attB5r (reverse) Gateway cloning sites. Subsequently, 5'-UTR sequences were

159 Gateway cloned into the pDONR P1-P5r vector (Invitrogen). The LUC reporter gene
160 was PCR amplified from the 'Luciferase SP6 Control DNA' vector (Promega) using
161 forward (attB5-LUCIFERASE) and reverse (nLUC-B2-R) primers containing attB5
162 and attB2 Gateway sites, respectively. The LUC fragment was subsequently
163 Gateway cloned into the pDONR221 P5-P2 vector (Invitrogen). 5'-UTR and LUC
164 sequences were assembled using MultiSite Gateway technology into plasmid
165 pALLIGATOR3 to generate 35S:CPuORF-LUC constructs. To generate constructs in
166 which the translation initiation codon of the CPuORF was mutated, new forward
167 primers were designed containing the desired mutation (with attB1 site; Table S2).
168 The new forward primer was used in combination with the appropriate reverse primer
169 for PCR amplification using the relevant wild-type pDONR221 P1-P5r construct as
170 template. The mutated PCR product was cloned into pDONR221 P1-P5r and
171 assembled with the LUC gene (in pDONR221 P5-P2) into pALLIGATOR3.

172 *In vitro* LUC reporters were made in one of two ways. First, the CPuORF alone was
173 amplified from the relevant pALLIGATOR3 construct, using a CPuORF-specific
174 forward primer that also contained a Kozak sequence (AACAGACCACC<u>AUG</u>,
175 translation initiation sequence underlined) with a *Hind*III restriction site for cloning,
176 together with a CPuORF-specific reverse primer containing a *Not*I site for cloning
177 (Table S2). PCR products were *Hind*III-*Not*I digested and ligated into the 'Luciferase
178 SP6 Control DNA' vector, digested with the same restriction enzymes, to generate
179 SP6:CPuORF-LUC constructs where the CPuORF was in a non-native 5'-UTR
180 context. As above, versions of these constructs were made in which the start codon
181 of the CPuORF was mutated, through the use of modified forward primers (Table
182 S2). Purified plasmids were used as templates in transcription/translation reactions

as described below. Second, to test the CPuORF in its native 5'-UTR context *in vitro*, the 5'-UTR containing the CPuORF and the LUC gene was amplified from the relevant plant pALLIGATOR3 construct, using a forward CPuORF-specific primer with a tail containing the Kozak sequence and the SP6 promoter sequence, together with a reverse primer (LUC-TNT-R) specific to the 3'-end of the LUC gene (Table S2). Purified PCR products were used directly in transcription/translation reactions as described below.

Mutagenesis of the C-terminal domain of CPuORF47 was carried out using the Q5 site-directed mutagenesis kit, following the manufacturer's instructions (NEB). PCR reactions were performed using the 47SDM primer pairs listed in Supplementary Table S2, with SP6:CPuORF47-LUC (*in vitro*) or pDONR221[P1-P5r]-CPuORF47 (*in planta*) as templates. Following sequence validation, SP6:CPuORF47$^{sdm1\text{-}sdm3}$-LUC constructs were used directly in LUC assays, while pDONR221[P1-P5r]-CPuORF47$^{sdm1\text{-}sdm3}$ constructs were assembled together with the LUC reporter gene into pALLIGATOR3 using Multisite Gateway.

To generate the 35S:SAC51_CPuORF-SEP3 construct, the *SAC51* CPuORF-containing 5'-UTR was amplified using Gateway modified oligos (SAC51-B1-F and SAC51-B5r-R) and was cloned into pDONR221 P1-P5r. The *SEP3* cds was also amplified using Gateway oligos (SEP3-F-B5 and SEP3-R) and cloned into pDONR221 P5-P2. Multisite Gateway was used to assemble the CPuORF and *SEP3* sequences downstream of the 35S promoter in pALLIGATOR3. The *SEP3* cds was also amplified (using SEP3-F-B1 and SEP3-F primers), to allow cloning directly downstream of the 35S promoter, without the CPuORF.

2.3 Luciferase reporter assays

For *in vitro* assays, purified PCR products (250ng) or plasmid constructs (500ng) were transcribed and translated using the TNT SP6 High-Yield Wheat Germ Master Mix. Reactions (25µl) were assembled and incubated for 2 hours at 25°C, according to the manufacturer's instructions (Promega). TNT reactions were performed in triplicate. To measure LUCIFERASE activity, 10µl of each TNT reaction was mixed with an equal volume of LUC assay buffer (0.5mM luciferin, 0.01% (w/v) Triton X-100) in an opaque 96-well plate, and bioluminescence detected using the LB985 NightShade Plant Imaging System (Berthold), with the following settings: exposure 0.1-1s, 4x4 binning, gain high, slow read out. Photon counts per second (cps) was measured for each sample and averaged between replicates.

For *in planta* assays, fully expanded, healthy leaves were selected from the rosettes of transgenic plants, incubated in LUC assay buffer for 5 min at room temperature and placed in the dark for 5 min. Bioluminescence was measured as above, except that exposure was 60s. Cps/mm^2 of leaf was calculated and averaged between replicates. To normalise for possible differences in transgene expression (due to positional effects or CPuORF-mediated transcript decay) we analysed multiple independent T_1 plants for each construct (for details see Table S3).

2.4 Stress and chemical treatments

2.4.1 For LUC reporter assays

To test for response to water limitation in a controlled manner, leaves from transgenic plants were incubated in 0.5xMS liquid medium containing 300mM mannitol for 24h at room temperature (RT). As control, leaves from the same plants were incubated in 0.5xMS for 24h at RT.

To test for response to increased temperatures, leaves from transgenic plants were incubated at 37°C for 6h on damp filter paper. As control, leaves for the same plants were incubated at RT for 6h on damp filter paper.

To test for responses to thermospermine, leaves from transgenic plants were incubated in 0.5xMS liquid medium containing 0.5mM thermospermine for 24h at RT. As control, leaves from the same plants were incubated in 0.5xMS for 24h at RT.

2.4.2 For *SEP3* developmental reporter

Transgenic seed were selected and sown onto 0.5xMS plates with or without 0.5mM thermospermine supplement. Following stratification at 4°C for 2 days, plates were incubated at 21°C under long days. Phenotypes were characterised during the seedling stage and photographed using the Keyence VHX500 digital microscope.

2.5 Bioinformatics

Ribo-seq and PARE data was taken from the GEO database under the accession numbers GSM1226369 (Liu et al., 2013) and GSM280226 (German et al., 2008), respectively. For ribo-seq, the RF reads generated for light-exposed seedlings (Liu et al., 2013) were visualised for CPuORF-containing transcripts annotated in TAIR10 (Lamesch et al., 2012; arabidopsis.org), using the GWIPS-viz genome browser in RiboGalaxy (ribogalaxy.ucc.ie). PARE reads for each CPuORF transcript were retrieved and analysed using the Arabidopsis Next-Gen Sequence DBs (mpss.danforthcenter.org/dbs/index.php?SITE=at_pare).

3 RESULTS

3.1 CPuORFs act as autonomous regulators of mORF translation

To understand the conditional regulation of gene expression by plant CPuORFs we selected Arabidopsis CPuORFs belonging to two previously uncharacterised homology groups (HGs): (1) HG7a and its sole representative CPuORF19 (a class II CPuORF from At1g36730), and (2) HG17, which has four members in Arabidopsis, CPuORF45, 46, 47 and 48, although here we mainly focus on CPuORF46 (At3g53400) and CPuORF47 (At5g03190). For an evolutionary perspective, we also selected a representative member of HG17 from rice (OsCPuORF38 from LOC_Os02g52300).

Since the paradigm is that translation of a uORF/CPuORF is required for translational repression, we compared the activity of the wild-type CPuORFs with that of altered versions in which the translation initiation codon (uAUG) of the CPuORF was mutated to prevent its translation (Figure 1b). To establish that CPuORFs were acting at the level of translation, we first examined reporter activity in a cell-free expression system. For this we cloned the relevant CPuORF coding region between the *SP6* promoter and the *LUCIFERASE* (*LUC*) reporter gene. As shown in Figure 1c (see also Table S3), all reporters in which the uAUG was mutated had increased LUC activity relative to WT versions, indicating that translation of the CPuORF is necessary in order to attenuate translation of the associated mORF. Importantly, these CPuORFs were tested in a completely heterologous system (non-native promoter, 5'-UTR and mORF) indicating that the CPuORF regulates translation autonomously. We also tested CPuORF19 and CPuORF46 (±uAUG) in their native 5'-UTR context. As above, reporters with a mutated uAUG had significantly higher LUC activity relative to WT constructs, and to a similar level observed with the non-native 5'-UTR (Table S3), demonstrating that

the information required for CPuORF-induced attenuation of mORF translation is confined to the CPuORF. The fact that the presence of a translatable CPuORF alone, independent of 5'-UTR sequence context, is sufficient to attenuate downstream mORF translation, also rules out the influence of surrounding RNA structure.

To verify that both classes of CPuORFs also function as translational regulators *in planta*, we tested CPuORF19 (HG7a, class II) and CPuORF47 (HG17, class I). The 5'-UTR containing the relevant CPuORF coding region (±uAUG) was cloned between the constitutive *35S* promoter and the *LUC* reporter gene (see Figure 1b) and the constructs transformed into wild-type Arabidopsis plants. Relative to the WT constructs, those with a mutated uAUG had significantly higher levels of LUC activity (Figure 1d and Table S3), closely mirroring the *in vitro* data.

Taken together, the data shows that translation of HG7a and HG17 CPuORFs inhibits translation of the mORF and that these CPuORFs are self-contained units that can attenuate translation of downstream ORFs.

3.2 HG17 CPuORFs repress translation in a peptide sequence-dependent manner and cause ribosome stalling

HG17 are class I CPuORFs, which are highly conserved at the C-terminus (Arabidopsis HG17 CPuORFs share >57% identity across the C-terminal 14aa of the peptide), but more divergent at the N-terminus of the encoded peptide (the Arabidopsis HG17 peptides share no overall identity outside the C-termini) (Figure 2a). To examine the functional relevance of the conserved C-terminus of these peptides we generated a series of mutations within this region of CPuORF47 (Figure

2b) and measured LUC activity both *in vitro* and *in planta* relative to the WT constructs. We found that mutations anywhere within the C-terminus of CPuORF47 significantly reduced the effectiveness of the CPuORF to attenuate LUC activity *in vitro*, and to a similar level as observed for the CPuORF with mutated uAUG (Figure 2b and Table S3). A similar trend was also seen *in planta* (Figure 2b and Table S3). Together the data suggests that the C-terminus of the CPuORF47 peptide includes sequences that attenuate translation of a downstream mORF, potentially by arresting translation at the CPuORF.

To determine whether HG17, class I CPuORFs encode sequences that cause ribosome stalling, we examined the positional distribution of translating ribosomes using ribosome profiling (ribo-seq) and parallel analysis of RNA ends (PARE) datasets (German et al., 2008; Liu et al., 2013). Ribo-seq and PARE allow the precise identification of transcript sequences bound by translating ribosomes (ribosome footprints; RFs). While no ribo-seq data is available for CPuORF45, visual inspection of RFs associated with CPuORF46, CPuORF47 and CPuORF48 transcripts revealed an accumulation of RF reads within the coding regions of the CPuORFs (Figure 3a-c), showing that HG17 CPuORFs are translated under control conditions (seedlings grown under long day light conditions; Liu et al., 2013). Further, the distribution of ribo-seq reads along the CPuORF sequence showed a peak at the end of the CPuORF, and in the case of CPuORF46 and CPuORF48, greater ribosome occupancy in the CPuORF relative to the mORF (Figure 3). Together, this indicates that translation of these CPuORF sequences stalls ribosomes, preventing them from accessing and translating the mORF. By analysing PARE data, Hou et al. (2016) found that some Arabidopsis and rice CPuORFs accumulate PARE reads at

nucleotide positions -16 and/or -46 (where position 0 is the first nucleotide of the CPuORF stop codon), consistent with ribosome stalling at the termination codon and stacking of ribosomes at 30 nucleotide intervals upstream (the approximate size of an Arabidopsis ribosome footprint). Using the same PARE dataset (German et al., 2008) we observed enrichment of PARE reads at positions -16 and -46 for CPuORF46 and CPuORF47 (Figure 3a,b), and at -16 for CPuORF48 (Figure 3c), supporting the idea that the highly conserved C-terminal ends of these class I CPuORFs include sequences that stall ribosomes.

Examination of ribo-seq and PARE data for the class II CPuORF19 also shows an abundance of reads on the CPuORF, indicating that this CPuORF is translated under control conditions, which may cause ribosome stalling. However, the PARE data for this class II CPuORF shows a different pattern of read peaks to the class I HG17 CPuORFs, with no obvious enrichment of reads at positions -16 or -46 (Figure 3d).

3.3 Class I and class II CPuORFs show different patterns of ribosome occupancy

The different patterns of PARE reads between the HG17 CPuORFs (class I) and CPuORF19 (class II) prompted us to look more closely at ribosome stalling by these two classes of CPuORFs. For this, we expanded our analysis of the available ribo-seq and PARE datasets to include all CPuORF transcripts annotated in TAIR10 (arabidopsis.org). Visual inspection of the ribo-seq reads mapped to 46 different CPuORF-containing transcripts showed ribosome occupancy within almost all of the

CPuORF sequences (Figure S1), supporting the idea that CPuORFs are translated under control conditions (Hsu et al., 2016; Hu, Merchante, Stepanova, Alonso, & Heber, 2016). We also noted that for many CPuORF-containing transcripts, there was higher density of ribosome occupancy on the CPuORF coding region relative to the mORF (Figure S1), consistent with CPuORF-mediated attenuation of mORF translation. By examining the PARE analysis published by Hou et al. (2016), we identified different patterns of ribosome occupancy for class I and class II CPuORFs. Class I CPuORFs, which are mainly conserved at the C-terminus, predominantly accumulate PARE reads at positions -16 and -46 (Figure S2). Indeed, we found that approximately 50% of class I CPuORFs had an enrichment of PARE reads at position -16, and ~35% accumulated reads at position -46 (Figure 3e). In contrast, class II CPuORFs, which are conserved along the entire length of the peptide, or at the N-terminus and/or middle regions, showed a more even distribution of PARE reads across the 3'-end of the CPuORF coding sequence, with only ~26% of these CPuORFs accumulating reads at position -16, and <1% accumulating reads at position -46 (Figure 3e and Figure S2). Together, these analyses suggest that ribosome stalling is a common mechanism by which CPuORFs modulate translation of associated mORFs, and that arrest in the vicinity of a stop codon is a major mechanism for class I CPuORF function, but not necessarily for those belonging to class II.

3.4 CPuORFs control mORF activity conditionally

It is known that for some uORFs, including some plant CPuORFs, their inhibitory function can be modulated by intrinsic or extrinsic factors (Dever et al., 2020; van der

Horst et al., 2020). Arabidopsis transcripts containing HG7a or HG17 CPuORFs are up-regulated in response to abiotic stress related to water limitation (Matsui et al., 2008; Higashi, Okazaki, Myouga, Shinozaki, & Saito, 2015; Shaar-Moshe, Hübner, & Peleg, 2015; Sham et al., 2015; Rasheed, Bashir, Matsui, Tanaka, & Seki, 2016). Here we examined whether these CPuORFs play any additional role in a translational response to these stress conditions.

Water limitation is one of the more economically significant environmental stresses faced by plants in the field. To test whether HG17 CPuORFs mediate responses to drought-like conditions, we compared our *in planta* reporter lines treated with mannitol (often used as a highly controllable drought mimic (Dubois & Inzé, 2020)) with mock-treated samples. In response to mannitol, *35S:CPuORF47-LUC* samples showed a dramatic increase in LUC activity relative to controls (~10-fold; $p<0.01$) (Figure 4a and Table S3). We also found that for *35S:CPuORF47-LUC* plants grown in soil and subjected to drought treatment, LUC activity increased approximately 2.4-fold relative to well-watered controls (Table S3), suggesting that CPuORF47 is a *bona fide* autonomous, post-transcriptional drought-responsive element. To confirm that CPuORF translation was important for this conditional response, we also tested reporter lines in which the translation initiation codon of CPuORF47 had been mutated (see Figure S3a) and found that unlike the WT constructs there was no response to mannitol when CPuORF translation was inhibited (Figure S3b, and Table S3). Similarly, *35S:CPuORF46-LUC* plants showed a response to mannitol (~1.7-fold; Figure 4a), confirming that HG17 CPuORFs modulate translation in response to water limitation. Interestingly, however, the rice *35S:OsCPuORF38-LUC* plants showed no mannitol response (Figure 4a), suggesting that this divergent

HG17 CPuORF lacks the ability to respond to drought, at least when expressed in Arabidopsis.

To further examine the specificity of CPuORF responses to water limitation, we performed a set of additional experiments. First, we checked to see if mannitol had a direct impact on LUC activity. LUC activity of *35S:LUC* plants treated with mannitol showed no significant difference to controls (Figure 4a). Next, we examined the effect of mannitol on the function of the well-characterised *SUPPRESSOR OF ACAULIS 51* (*SAC51*; At5g64340) CPuORFs that respond to the polyamine thermospermine (Imai et al., 2006). As expected, *35S:SAC51_CPuORF-LUC* plants showed an increase in LUC activity after thermospermine treatment, while the same treatment was inhibitory to LUC activity in both *35S:LUC* and *35S:CPuORF47-LUC* plants (Figure 4b). The LUC activity of *35S:SAC51_CPuORF-LUC* plants was not significantly altered by mannitol treatment (Figure 4a), suggesting that the responses are specific to CPuORF46 and CPuORF47.

Heat stress is another important environmental stress that can significantly impact plant productivity (Fahad et al., 2017). Furthermore, heat stress can lead to drought-like stress in plants (Lamaoui, Jemo, Datla, & Bekkaoui, 2018). Here we examined whether HG17 CPuORFs also respond to heat shock (37ºC for 6h). Although we detected no response to heat in *35S:CPuORF47-LUC* plants, *35S:CPuORF46-LUC* plants showed a statistically significant increase in LUC reporter activity (~8.85-fold; $p<0.01$) relative to controls (Figure 4c and Table S3). In addition, *35S:CPuORF48-LUC* plants also showed a moderate increase in LUC activity in response to heat (Table S3), suggesting that HG17 CPuORFs can modulate translation in response to

increased temperature. As with water limitation, 35S:*OsCPuORF38-LUC* plants did not show a response to heat shock in our reporter assays (Figure 4c).

The C-terminus of the OsCPuORF38 peptide shares significant amino acid identity with that of Arabidopsis CPuORF46 and CPuORF47 (~86%; Figure S4). In contrast, the OsCPuORF38 N-terminus shares little overall homology with the same regions of CPuORF46 (~27%) or CPuORF47 (~18%) (Figure S4), suggesting that the different conditional responses may be mediated by the divergent N-terminal region of these CPuORFs.

As with mannitol, we also measured LUC activity in 35S:LUC and 35S:SAC51_CPuORF-LUC plants in response to heat shock. In neither case was LUC activity significantly increased following incubation for 6h at 37ºC (Figure 4c), confirming that the response of CPuORF46 to heat shock is specific.

Next, we examined the conditional responses of CPuORF19. We found LUC activity to be significantly elevated in *35S:CPuORF19-LUC* plants treated with mannitol, compared to controls (~19-fold; Figure 4a), while heat shock treatment of *35S:CPuORF19-LUC* plants resulted in reduced LUC signal (Figure 4c). Thus, CPuORF19 is a second element identified here that specifically regulates translation in response to water limitation.

In summary, we have identified three previously uncharacterised stress responsive CPuORFs with potential as tools for the control of inducible gene expression. In particular, these CPuORFs respond to specific signals, which is linked to their peptide sequence. The SAC51 CPuORFs (HG15) modulate translation in response to thermospermine (as published previously; Imai et al., 2006) but do not respond to

abiotic stress treatments. In contrast, the newly characterised CPuORF19 (HG7a), CPuORF46 and CPuORF47 (both HG17) fail to respond to thermospermine, but specifically respond to water limitation or heat shock. CPuORF47 and CPuORF19 strongly enhance translation in response to mannitol treatment, whereas CPuORF46 strongly enhances translation in response to elevated temperature.

3.5 CPuORFs have the potential to regulate agronomically important traits

The *SAC51* CPuORFs represent a set of regulatory elements that respond to specific effector molecules. The use of chemicals to precisely regulate agronomically important traits, such as flowering time, in response to the application of particular chemicals is an emerging strategy in crop engineering (e.g. Ionescu, Møller, & Sánchez-Pérez, 2016 – J Exp Bot 68:369). Understanding the mechanism(s) through which CPuORFs conditionally regulate translation of an associated mORF may ultimately allow the design of novel CPuORFs that respond to chemical signals of choice. For this to be achieved, the *in vivo* CPuORF-induced attenuation of translation needs to maintain expression of the mORF below the threshold required for biological activity and the signal-induced increase in translation must be sufficient to exceed that threshold. As a first step to show that CPuORFs can fulfil these criteria, we aimed to determine whether the thermospermine-responsive *SAC51* CPuORFs (Imai et al., 2006; Figure 4b) were capable of altering the activity of the floral regulator *SEPALLATA3* (*SEP3*). Constitutive mis-expression of *SEP3* results in seedlings that flower extremely early, and that produce terminal flowers and curled leaves (Honma & Goto, 2001; Castillejo, Romera-Branchat, & Pelaz, 2005; Figure 5a). To test whether the *SAC51* CPuORFs could regulate the developmental

outcomes of *SEP3* mis-expression, we placed the *SAC51* 5'-leader between the *35S* promoter and the *SEP3* coding sequence and examined seedling phenotypes in the T_2 generation. First, we asked whether the *SAC51* CPuORFs were capable of attenuating *SEP3* translation and maintaining translation levels below the threshold required to induce flowering. As shown in Figure 5, compared to *35S:SEP3* plants that had severe flowering time defects, all of the *35S:SAC51_CPuORF-SEP3* plants were aphenotypic under control conditions, indicating that the *SAC51* CPuORFs stringently attenuate the effects of constitutive *SEP3* expression. Next, we determined whether thermospermine treatment would release the repressive activity of the *SAC51* CPuORFs on *SEP3* translation. Although all plants treated with thermospermine showed a stress response (retarded growth and increased anthocyanin production; Figure 5a), critically ~20% of the *35S:SAC51_CPuORF-SEP3* seedlings grown in the presence of 0.5mM thermospermine exhibited the strong *SEP3* over-expression phenotype (Figure 5), which was not seen for thermospermine treated wild-type plants. The data indicates that at least for the *SAC51* CPuORFs, it is feasible to regulate a specific developmental outcome in a user-defined manner. Future work to understand how translational stalling can be conditionally relieved, will reveal the true potential of CPuORFs to engineer conditional plant responses.

4 DISCUSSION

4.1 Mechanisms of CPuORF function

Ribosome stalling during uORF/CPuORF translation is an important mechanism of post-transcriptional regulation that has been demonstrated for some Arabidopsis and rice CPuORFs (Hayashi et al., 2017; Hou et al., 2016; Juntawong, Girke, Bazin, & Bailey-Serres, 2014; Ribone et al., 2017; Uchiyama-Kadokura et al., 2014; Yamashita et al., 2017). Using ribosome profiling (Liu et al., 2013) and PARE (German et al., 2008; Hou et al., 2016) datasets, we determined that most CPuORFs are translated under control conditions, resulting in stalled ribosomes (Figure S1 and Figure S2). Furthermore, the distribution of PARE reads for class I CPuORFs reveals that ribosomes predominantly accumulate near the stop codon (Figure S2). Accordingly, we show that, at least for HG17 CPuORFs, translation inhibition is dependent on their conserved C-termini (Figure 2), which are likely to include ribosome stalling sequences (Figure 3). How translation of these CPuORFs cause conditional ribosome stalling is not understood, but data from other eukaryotes implicates interactions between the uORF peptide and effector molecules within the ribosome exit tunnel (Bhushan et al., 2010). Interestingly, point mutations in genes encoding the Arabidopsis ribosome exit tunnel proteins RPL4A, and RPL10A, alleviate *SAC51* CPuORF-mediated translation attenuation (Imai, Komura, Kawano, Kuwashiro, & Takahashi, 2008; Kakehi et al., 2015). Human and fungal RPL4 and RPL10 ribosomal subunits are also associated with translation stalling caused by the unrelated CMV and AAP uORF peptides, respectively (Bhushan et al., 2010). Together these data suggest that plant CPuORFs act as self-contained regulatory units, containing the features required for the conditional stalling of translation, potentially through interactions with common ribosomal components.

4.2 Conditional CPuORFs are not simple on/off switches

The current paradigm is that uORFs/CPuORFs are conditional switches that turn translation on or off in response to particular effectors. However, we propose that CPuORFs are more akin to a rheostat or dimmer switch, functioning along a continuum to fine-tune translation (Figure 6). Like rheostats, CPuORFs are discrete components that adjust an output, in this case mORF translation, in response to signal levels. Under default conditions, CPuORFs block the passage of ribosomes (Figure S1 and Figure 6), constitutively attenuating translation of downstream ORFs. In response to specific signals, the degree of translation inhibition can be modulated through a mechanism that we call Conditional uORF-dependent Translation Stalling (CUTS). Examples of uORFs acting through CUTS have been reported for prokaryotes and eukaryotes (Ito & Chiba, 2013), suggesting that it is an ancient mechanism of translation regulation. CUTS has two modes of action (Figure 6). In repressive CUTS (rCUTS), the constitutive translational attenuation is enhanced by the presence of a small effector molecule, potentially interacting with the nascent uORF peptide within the ribosome exit tunnel. The majority of Arabidopsis CPuORFs examined so far regulate translation of the mORF via rCUTS, each responding to a different signal, such as sucrose, ascorbate, phosphocholine, polyamines, galactinol and photosynthetic signals (Alatorre-Cobos et al., 2012; Guerrero-González et al., 2016; Hanfrey et al., 2005; Laing et al., 2015; Rahmani et al., 2009; Ribone et al., 2017; Tabuchi et al., 2006; Uchiyama-Kadokura et al., 2014; Wiese et al., 2004; Zhu et al., 2018) (see Table S1). In contrast, in activating CUTS (aCUTS), the effector interacts with the stalling peptide to release translation inhibition (Figure 6). There are fewer reported examples of aCUTS CPuORFs, but our data (Figure 4), together

with that of others, show that these respond to metabolites such as polyamines (Imai et al., 2006), and to environmental and biotic stresses (Bazin et al., 2017; Xu et al., 2017; Zhu et al., 2018; Zhu et al., 2012). The CUTS model is supported by the fact that different signals may have opposing effects on the regulatory activity of the same CPuORF. For example, the TBF1 cassette that includes CPuORF49 (Xu et al., 2017), acts as a potent inhibitor of translation in response to galactinol (rCUTS) (Zhu et al., 2018), while heat shock or pathogen attack alleviate TBF1 cassette-mediated translation repression (aCUTS) (Xu et al., 2017; Zhu et al., 2012).

4.3 Future prospects

Agriculture faces challenges from climate change, population expansion and land availability, which together significantly impact crop productivity (Jaganathan, Ramasamy, Sellamuthu, Jayabalan, & Venkataraman, 2018; Mall, Gupta, & Sonkar, 2017). Consequently, new tools are required to facilitate the sustainable and secure supply of plant products. Recently, the potential of conditional CPuORFs for crop improvement has been explored (Xing et al., 2020; Xu et al., 2017; Zhang et al., 2018). However, only a handful of CPuORFs have been shown to respond to agriculturally relevant conditions (Table S1). For example, Bazin et al. (2017) identified CPuORFs that respond to depleted phosphorous, while the TBF1 cassette has been shown to regulate translation in response to pathogen attack and heat shock (Xu et al., 2017; Zhu et al., 2012). Here we identify three uncharacterised Arabidopsis CPuORFs, that respond to specific agronomically significant environmental stresses: water limitation (CPuORF19 and CPuORF47) and increased temperatures (CPuORF46) (Figure 4). They function independently of transcription,

so are an additional layer of regulation that responds rapidly to change. We show that all the information required for ribosome stalling and conditional responses is contained wholly within the CPuORF coding region. We also show that they can function upstream of any mORF, indicating that they have the potential to be used to control the expression of any gene. These, and other, CPuORFs could be used to engineer conditional enhancement or suppression of the product of an endogenous or exogenous transgene, by inserting a CPuORF sequence into its 5'-leader. We have investigated a limited number of conditions, but expect that, as more conditions are tested, new conditional CPuORFs will be identified. Although a significant number of plant CPuORFs have been identified (e.g. Takahashi et al., 2020; see Table S1), finding these elements is not straightforward. *In silico* searches for plant CPuORFs have tended to focus on those containing a canonical AUG initiation codon (van der Horst et al., 2019). However, a small number of non-AUG plant CPuORFs have been identified (Laing et al., 2015; van der Horst et al., 2019), suggesting that more CPuORFs remain to be discovered. Since CPuORFs provide a layer of gene regulation that is independent of transcription, it may be possible to increase the stringency of conditional responses by using different CPuORFs in series, and/or in combination with conditional promoters, increasing the versatility of the elements. Others have begun to explore the possibility of using CRISPR/Cas9 genome editing to modify plant uORF sequences *in situ*, thereby manipulating translation (Si, Zhang, Wang, Chen, & Gao, 2020; Zhang et al., 2018). We show that CPuORFs are valid targets to engineer induced developmental responses (Figure 5). It may ultimately be possible to design synthetic uORFs that selectively regulate translation in response to applied chemicals or abiotic/biotic challenges.

ACKNOWLEDGEMENTS

We thank all members of the Davies lab and Dr Younousse Saidi (BASF) for useful discussions throughout this work. We particularly wish to thank Dr Sam Rayson, Lewis Adams, Andrea Subroto and Katie Poulter for technical assistance and the University of Leeds Plant Growth Facility for assistance growing Arabidopsis. We are also grateful to Dr James Lloyd and Dr Antoine Larrieu for vital feedback during the preparation of the manuscript. B.D. and B.C. thank the U.K. Biotechnology and Biological Sciences Research Council for funding this research (BB/L006170/1). The project was an industrial partnership with Bayer CropScience that subsequently transitioned to BASF, whom we also thank for additional funding.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

Alatorre-Cobos, F., Cruz-Ramírez, A., Hayden, C. A., Pérez-Torres, C. A., Chauvin, A. L., Ibarra-Laclette, E., Alva-Cortés, E., Jorgensen, R. A., & Herrera-Estrella, L. (2012). Translational regulation of Arabidopsis XIPOTL1 is modulated by phosphocholine levels via the phylogenetically conserved upstream open reading frame 30. *Journal of experimental botany*, *63*(14), 5203–5221. https://doi.org/10.1093/jxb/ers180

Barbosa, C., Peixeiro, I., & Romão, L. (2013). Gene expression regulation by upstream open reading frames and human disease. *PLoS genetics*, *9*(8), e1003529. https://doi.org/10.1371/journal.pgen.1003529

Bazin, J., Baerenfaller, K., Gosai, S. J., Gregory, B. D., Crespi, M., & Bailey-Serres, J. (2017). Global analysis of ribosome-associated noncoding RNAs unveils new modes of translational regulation. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(46), E10018–E10027. https://doi.org/10.1073/pnas.1708433114

Bensmihen S, To A, Lambert G, Kroj T, Giraudat J, Parcy F. Analysis of an activated ABI5 allele using a new selection method for transgenic Arabidopsis seeds. FEBS Lett. 2004 Mar 12;561(1-3):127-31. doi: 10.1016/S0014-5793(04)00148-6. PMID: 15013763.

Bhushan, S., Meyer, H., Starosta, A. L., Becker, T., Mielke, T., Berninghausen, O., Sattler, M., Wilson, D. N., & Beckmann, R. (2010). Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide. *Molecular cell*, *40*(1), 138–146. https://doi.org/10.1016/j.molcel.2010.09.009

624 Castillejo, C., Romera-Branchat, M., & Pelaz, S. (2005). A new role of the
625 Arabidopsis SEPALLATA3 gene revealed by its constitutive expression. *The Plant*
626 *journal*, *43*(4), 586–596. https://doi.org/10.1111/j.1365-313X.2005.02476.x

627

628 Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-
629 mediated transformation of Arabidopsis thaliana. *The Plant journal*, *16*(6), 735–743.
630 https://doi.org/10.1046/j.1365-313x.1998.00343.x

631

632 Dejonghe, W., & Russinova, E. (2017). Plant Chemical Genetics: From Phenotype-
633 Based Screens to Synthetic Biology. *Plant physiology*, *174*(1), 5–20.
634 https://doi.org/10.1104/pp.16.01805

635

636 Dever, T. E., Ivanov, I. P., & Sachs, M. S. (2020). Conserved Upstream Open
637 Reading Frame Nascent Peptides That Control Translation. *Annual review of*
638 *genetics*, *54*, 237–264. https://doi.org/10.1146/annurev-genet-112618-043822

639

640 Dubois, M., & Inzé, D. (2020). Plant growth under suboptimal water conditions: early
641 responses and methods to study them. *Journal of experimental botany*, *71*(5), 1706–
642 1722. https://doi.org/10.1093/jxb/eraa037

643

644 Ebina, I., Takemoto-Tsutsumi, M., Watanabe, S., Koyama, H., Endo, Y., Kimata, K.,
645 Igarashi, T., Murakami, K., Kudo, R., Ohsumi, A., Noh, A. L., Takahashi, H., Naito, S.,
646 & Onouchi, H. (2015). Identification of novel Arabidopsis thaliana upstream open
647 reading frames that control expression of the main coding sequences in a peptide
648 sequence-dependent manner. *Nucleic acids research*, *43*(3), 1562–1576.
649 https://doi.org/10.1093/nar/gkv018

650

651 Fahad, S., Bajwa, A. A., Nazir, U., Anjum, S. A., Farooq, A., Zohaib, A., Sadia, S.,
652 Nasim, W., Adkins, S., Saud, S., Ihsan, M. Z., Alharby, H., Wu, C., Wang, D., &
653 Huang, J. (2017). Crop Production under Drought and Heat Stress: Plant Responses
654 and Management Options. *Frontiers in plant science*, *8*, 1147.
655 https://doi.org/10.3389/fpls.2017.01147

656

657 German, M. A., Pillay, M., Jeong, D. H., Hetawal, A., Luo, S., Janardhanan, P.,
658 Kannan, V., Rymarquis, L. A., Nobuta, K., German, R., De Paoli, E., Lu, C., Schroth,
659 G., Meyers, B. C., & Green, P. J. (2008). Global identification of microRNA-target
660 RNA pairs by parallel analysis of RNA ends. *Nature biotechnology*, *26*(8), 941–946.
661 https://doi.org/10.1038/nbt1417

662

663 Guerrero-González, M. L., Ortega-Amaro, M. A., Juárez-Montiel, M., & Jiménez-
664 Bremont, J. F. (2016). Arabidopsis Polyamine oxidase-2 uORF is required for
665 downstream translational regulation. *Plant physiology and biochemistry : PPB*, *108*,
666 381–390. https://doi.org/10.1016/j.plaphy.2016.08.006

Hanfrey, C., Elliott, K. A., Franceschetti, M., Mayer, M. J., Illingworth, C., & Michael, A. J. (2005). A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. *The Journal of biological chemistry*, *280*(47), 39229–39237. https://doi.org/10.1074/jbc.M509340200

Hayashi, N., Sasaki, S., Takahashi, H., Yamashita, Y., Naito, S., & Onouchi, H. (2017). Identification of Arabidopsis thaliana upstream open reading frames encoding peptide sequences that cause ribosomal arrest. *Nucleic acids research*, *45*(15), 8844–8858. https://doi.org/10.1093/nar/gkx528

Hayden, C. A., & Jorgensen, R. A. (2007). Identification of novel conserved peptide uORF homology groups in Arabidopsis and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. *BMC biology*, *5*, 32. https://doi.org/10.1186/1741-7007-5-32

Higashi, Y., Okazaki, Y., Myouga, F., Shinozaki, K., & Saito, K. (2015). Landscape of the lipidome and transcriptome under heat stress in Arabidopsis thaliana. *Scientific reports*, *5*, 10533. https://doi.org/10.1038/srep10533

687 Honma, T., & Goto, K. (2001). Complexes of MADS-box proteins are sufficient to
688 convert leaves into floral organs. *Nature*, *409*(6819), 525–529.
689 https://doi.org/10.1038/35054083

690

691 Hou, C. Y., Lee, W. C., Chou, H. C., Chen, A. P., Chou, S. J., & Chen, H. M. (2016).
692 Global Analysis of Truncated RNA Ends Reveals New Insights into Ribosome
693 Stalling in Plants. *The Plant cell*, *28*(10), 2398–2416.
694 https://doi.org/10.1105/tpc.16.00295

695

696 Hsu, P. Y., Calviello, L., Wu, H. L., Li, F. W., Rothfels, C. J., Ohler, U., & Benfey, P.
697 N. (2016). Super-resolution ribosome profiling reveals unannotated translation
698 events in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the*
699 *United States of America*, *113*(45), E7126–E7135.
700 https://doi.org/10.1073/pnas.1614788113

701

702 Hu, Q., Merchante, C., Stepanova, A. N., Alonso, J. M., & Heber, S. (2016).
703 Genome-Wide Search for Translated Upstream Open Reading Frames in
704 Arabidopsis Thaliana. *IEEE transactions on nanobioscience*, *15*(2), 148–157.
705 https://doi.org/10.1109/TNB.2016.2516950

706

707 Imai, A., Hanzawa, Y., Komura, M., Yamamoto, K. T., Komeda, Y., & Takahashi, T.
708 (2006). The dwarf phenotype of the Arabidopsis acl5 mutant is suppressed by a

mutation in an upstream ORF of a bHLH gene. *Development (Cambridge, England)*, *133*(18), 3575–3585. https://doi.org/10.1242/dev.02535

Imai, A., Komura, M., Kawano, E., Kuwashiro, Y., & Takahashi, T. (2008). A semi-dominant mutation in the ribosomal protein L10 gene suppresses the dwarf phenotype of the acl5 mutant in Arabidopsis thaliana. *The Plant journal : for cell and molecular biology*, *56*(6), 881–890. https://doi.org/10.1111/j.1365-313X.2008.03647.x

Ingolia N. T. (2016). Ribosome Footprint Profiling of Translation throughout the Genome. *Cell*, *165*(1), 22–33. https://doi.org/10.1016/j.cell.2016.02.066

Ito, K., & Chiba, S. (2013). Arrest peptides: cis-acting modulators of translation. *Annual review of biochemistry*, *82*, 171–202. https://doi.org/10.1146/annurev-biochem-080211-105026

Jaganathan, D., Ramasamy, K., Sellamuthu, G., Jayabalan, S., & Venkataraman, G. (2018). CRISPR for Crop Improvement: An Update Review. *Frontiers in plant science*, *9*, 985. https://doi.org/10.3389/fpls.2018.00985

Johnstone, T. G., Bazzini, A. A., & Giraldez, A. J. (2016). Upstream ORFs are prevalent translational repressors in vertebrates. *The EMBO journal*, *35*(7), 706–723. https://doi.org/10.15252/embj.201592759

Jorgensen, R. A., & Dorantes-Acosta, A. E. (2012). Conserved Peptide Upstream Open Reading Frames are Associated with Regulatory Genes in Angiosperms. *Frontiers in plant science*, *3*, 191. https://doi.org/10.3389/fpls.2012.00191

Juntawong, P., Girke, T., Bazin, J., & Bailey-Serres, J. (2014). Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(1), E203–E212. https://doi.org/10.1073/pnas.1317811111

Kakehi, J., Kawano, E., Yoshimoto, K., Cai, Q., Imai, A., & Takahashi, T. (2015). Mutations in ribosomal proteins, RPL4 and RACK1, suppress the phenotype of a thermospermine-deficient mutant of Arabidopsis thaliana. *PloS one*, *10*(1), e0117309. https://doi.org/10.1371/journal.pone.0117309

Kurihara Y. (2020). uORF Shuffling Fine-Tunes Gene Expression at a Deep Level of the Process. *Plants (Basel, Switzerland)*, *9*(5), 608. https://doi.org/10.3390/plants9050608

Laing, W. A., Martínez-Sánchez, M., Wright, M. A., Bulley, S. M., Brewster, D., Dare, A. P., Rassam, M., Wang, D., Storey, R., Macknight, R. C., & Hellens, R. P. (2015). An upstream open reading frame is essential for feedback regulation of ascorbate

753 biosynthesis in Arabidopsis. *The Plant cell*, *27*(3), 772–786.
754 https://doi.org/10.1105/tpc.114.133777

755

756 Lamaoui, M., Jemo, M., Datla, R., & Bekkaoui, F. (2018). Heat and Drought Stresses
757 in Crops and Approaches for Their Mitigation. *Frontiers in chemistry*, *6*, 26.
758 https://doi.org/10.3389/fchem.2018.00026

759

760 Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller,
761 R., Dreher, K., Alexander, D. L., Garcia-Hernandez, M., Karthikeyan, A. S., Lee, C.
762 H., Nelson, W. D., Ploetz, L., Singh, S., Wensel, A., & Huala, E. (2012). The
763 Arabidopsis Information Resource (TAIR): improved gene annotation and new tools.
764 *Nucleic acids research*, *40*(Database issue), D1202–D1210.
765 https://doi.org/10.1093/nar/gkr1090

766

767 Liu, M. J., Wu, S. H., Wu, J. F., Lin, W. D., Wu, Y. C., Tsai, T. Y., Tsai, H. L., & Wu,
768 S. H. (2013). Translational landscape of photomorphogenic Arabidopsis. *The Plant*
769 *cell*, *25*(10), 3699–3710. https://doi.org/10.1105/tpc.113.114769

770

771 Liutkute, M., Samatova, E., & Rodnina, M. V. (2020). Cotranslational Folding of
772 Proteins on the Ribosome. *Biomolecules*, *10*(1), 97.
773 https://doi.org/10.3390/biom10010097

774

775 Lloyd, J. P., & Davies, B. (2013). SMG1 is an ancient nonsense-mediated mRNA
776 decay effector. *The Plant journal : for cell and molecular biology*, *76*(5), 800–810.
777 https://doi.org/10.1111/tpj.12329

778

779 Lloyd J. (2018). The evolution and diversity of the nonsense-mediated mRNA decay
780 pathway. *F1000Research*, *7*, 1299. https://doi.org/10.12688/f1000research.15872.2

781

782 Mall, R. K., Gupta, A., & Sonkar, G. (2017) Effect of climate change on agricultural
783 crops. In Current Developments in Biotechnology and Bioengineering: Crop
784 Modification, Nutrition, and Food Production (eds S. K. Dubey, A. Pandey, & R. S.
785 Sangwan), pp. 23-46. Elsevier B.V.

786

787 Matsui, A., Ishida, J., Morosawa, T., Mochizuki, Y., Kaminuma, E., Endo, T. A.,
788 Okamoto, M., Nambara, E., Nakajima, M., Kawashima, M., Satou, M., Kim, J. M.,
789 Kobayashi, N., Toyoda, T., Shinozaki, K., & Seki, M. (2008). Arabidopsis
790 transcriptome analysis under drought, cold, high-salinity and ABA treatment
791 conditions using a tiling array. *Plant & cell physiology*, *49*(8), 1135–1149.
792 https://doi.org/10.1093/pcp/pcn101

793

794 Merchante, C., Brumos, J., Yun, J., Hu, Q., Spencer, K. R., Enríquez, P., Binder, B.
795 M., Heber, S., Stepanova, A. N., & Alonso, J. M. (2015). Gene-specific translation

regulation mediated by the hormone-signaling molecule EIN2. *Cell*, *163*(3), 684–697. https://doi.org/10.1016/j.cell.2015.09.036

Park, S. Y., Peterson, F. C., Mosquna, A., Yao, J., Volkman, B. F., & Cutler, S. R. (2015). Agrochemical control of plant water use using engineered abscisic acid receptors. *Nature*, *520*(7548), 545–548. https://doi.org/10.1038/nature14123

Rahmani, F., Hummel, M., Schuurmans, J., Wiese-Klinkenberg, A., Smeekens, S., & Hanson, J. (2009). Sucrose control of translation mediated by an upstream open reading frame-encoded peptide. *Plant physiology*, *150*(3), 1356–1367. https://doi.org/10.1104/pp.109.136036

Rasheed, S., Bashir, K., Matsui, A., Tanaka, M., & Seki, M. (2016). Transcriptomic Analysis of Soil-Grown Arabidopsis thaliana Roots and Shoots in Response to a Drought Stress. *Frontiers in plant science*, *7*, 180. https://doi.org/10.3389/fpls.2016.00180

Rayson, S., Arciga-Reyes, L., Wootton, L., De Torres Zabala, M., Truman, W., Graham, N., Grant, M., & Davies, B. (2012). A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in Arabidopsis thaliana NMD mutants. *PloS one*, *7*(2), e31917. https://doi.org/10.1371/journal.pone.0031917

Ribone, P. A., Capella, M., Arce, A. L., & Chan, R. L. (2017). A uORF Represses the Transcription Factor AtHB1 in Aerial Tissues to Avoid a Deleterious Phenotype. *Plant physiology*, *175*(3), 1238–1253. https://doi.org/10.1104/pp.17.01060

Seidelt, B., Innis, C. A., Wilson, D. N., Gartmann, M., Armache, J. P., Villa, E., Trabuco, L. G., Becker, T., Mielke, T., Schulten, K., Steitz, T. A., & Beckmann, R. (2009). Structural insight into nascent polypeptide chain-mediated translational stalling. *Science (New York, N.Y.)*, *326*(5958), 1412–1415. https://doi.org/10.1126/science.1177662

Seip, B., & Innis, C. A. (2016). How Widespread is Metabolite Sensing by Ribosome-Arresting Nascent Peptides?. *Journal of molecular biology*, *428*(10 Pt B), 2217–2227. https://doi.org/10.1016/j.jmb.2016.04.019

Shaar-Moshe, L., Hübner, S., & Peleg, Z. (2015). Identification of conserved drought-adaptive genes using a cross-species meta-analysis approach. *BMC plant biology*, *15*, 111. https://doi.org/10.1186/s12870-015-0493-6

Sham, A., Moustafa, K., Al-Ameri, S., Al-Azzawi, A., Iratni, R., & AbuQamar, S. (2015). Identification of Arabidopsis candidate genes in response to biotic and abiotic stresses using comparative microarrays. *PloS one*, *10*(5), e0125666. https://doi.org/10.1371/journal.pone.0125666

Si, X., Zhang, H., Wang, Y., Chen, K., & Gao, C. (2020). Manipulating gene translation in plants by CRISPR-Cas9-mediated genome editing of upstream open reading frames. *Nature protocols*, *15*(2), 338–363. https://doi.org/10.1038/s41596-019-0238-3

Tabuchi, T., Okada, T., Azuma, T., Nanmori, T., & Yasuda, T. (2006). Posttranscriptional regulation by the upstream open reading frame of the phosphoethanolamine N-methyltransferase gene. *Bioscience, biotechnology, and biochemistry*, *70*(9), 2330–2334. https://doi.org/10.1271/bbb.60309

Takahashi, H., Takahashi, A., Naito, S., & Onouchi, H. (2012). BAIUCAS: a novel BLAST-based algorithm for the identification of upstream open reading frames with conserved amino acid sequences and its application to the Arabidopsis thaliana genome. *Bioinformatics (Oxford, England)*, *28*(17), 2231–2241. https://doi.org/10.1093/bioinformatics/bts303

Takahashi, H., Hayashi, N., Hiragori, Y., Sasaki, S., Motomura, T., Yamashita, Y., Naito, S., Takahashi, A., Fuse, K., Satou, K., Endo, T., Kojima, S., & Onouchi, H. (2020). Comprehensive genome-wide identification of angiosperm upstream ORFs with peptide sequences conserved in various taxonomic ranges using a novel

pipeline, ESUCA. *BMC genomics*, *21*(1), 260. https://doi.org/10.1186/s12864-020-6662-5

Tran, M. K., Schultz, C. J., & Baumann, U. (2008). Conserved upstream open reading frames in higher plants. *BMC genomics*, *9*, 361. https://doi.org/10.1186/1471-2164-9-361

Uchiyama-Kadokura, N., Murakami, K., Takemoto, M., Koyanagi, N., Murota, K., Naito, S., & Onouchi, H. (2014). Polyamine-responsive ribosomal arrest at the stop codon of an upstream open reading frame of the AdoMetDC1 gene triggers nonsense-mediated mRNA decay in Arabidopsis thaliana. *Plant & cell physiology*, *55*(9), 1556–1567. https://doi.org/10.1093/pcp/pcu086

van der Horst, S., Snel, B., Hanson, J., & Smeekens, S. (2019). Novel pipeline identifies new upstream ORFs and non-AUG initiating main ORFs with conserved amino acid sequences in the 5' leader of mRNAs in *Arabidopsis thaliana*. *RNA (New York, N.Y.)*, *25*(3), 292–304. https://doi.org/10.1261/rna.067983.118

van der Horst, S., Filipovska, T., Hanson, J., & Smeekens, S. (2020). Metabolite Control of Translation by Conserved Peptide uORFs: The Ribosome as a Metabolite Multisensor. *Plant physiology*, *182*(1), 110–122. https://doi.org/10.1104/pp.19.00940

Vaughn, J. N., Ellingson, S. R., Mignone, F., & Arnim, A. v. (2012). Known and novel post-transcriptional regulatory sequences are conserved across plant families. *RNA (New York, N.Y.)*, *18*(3), 368–384. https://doi.org/10.1261/rna.031179.111

von Arnim, A. G., Jia, Q., & Vaughn, J. N. (2014). Regulation of plant translation by upstream open reading frames. *Plant science : an international journal of experimental plant biology*, *214*, 1–12. https://doi.org/10.1016/j.plantsci.2013.09.006

Wiese, A., Elzinga, N., Wobbes, B., & Smeekens, S. (2004). A conserved upstream open reading frame mediates sucrose-induced repression of translation. *The Plant cell*, *16*(7), 1717–1729. https://doi.org/10.1105/tpc.019349

Xing, S., Chen, K., Zhu, H., Zhang, R., Zhang, H., Li, B., & Gao, C. (2020). Fine-tuning sugar content in strawberry. *Genome biology*, *21*(1), 230. https://doi.org/10.1186/s13059-020-02146-5

Xu, G., Yuan, M., Ai, C., Liu, L., Zhuang, E., Karapetyan, S., Wang, S., & Dong, X. (2017). uORF-mediated translation allows engineered plant disease resistance without fitness costs. *Nature*, *545*(7655), 491–494. https://doi.org/10.1038/nature22372

Yamashita A. (2013). Role of SMG-1-mediated Upf1 phosphorylation in mammalian nonsense-mediated mRNA decay. *Genes to cells : devoted to molecular & cellular mechanisms*, *18*(3), 161–175. https://doi.org/10.1111/gtc.12033

Yamashita, Y., Takamatsu, S., Glasbrenner, M., Becker, T., Naito, S., & Beckmann, R. (2017). Sucrose sensing through nascent peptide-meditated ribosome stalling at the stop codon of Arabidopsis bZIP11 uORF2. *FEBS letters*, *591*(9), 1266–1277. https://doi.org/10.1002/1873-3468.12634

Zhang, H., Si, X., Ji, X., Fan, R., Liu, J., Chen, K., Wang, D., & Gao, C. (2018). Genome editing of upstream open reading frames enables translational control in plants. *Nature biotechnology*, *36*(9), 894–898. https://doi.org/10.1038/nbt.4202

Zhang, H., Wang, Y., & Lu, J. (2019). Function and Evolution of Upstream ORFs in Eukaryotes. *Trends in biochemical sciences*, *44*(9), 782–794. https://doi.org/10.1016/j.tibs.2019.03.002

Zhang, T., Wu, A., Yue, Y., & Zhao, Y. (2020). uORFs: Important *Cis*-Regulatory Elements in Plants. *International journal of molecular sciences*, *21*(17), 6238. https://doi.org/10.3390/ijms21176238

Zhu, X., Thalor, S. K., Takahashi, Y., Berberich, T., & Kusano, T. (2012). An inhibitory effect of the sequence-conserved upstream open-reading frame on the translation of the main open-reading frame of HsfB1 transcripts in Arabidopsis. *Plant, cell & environment*, *35*(11), 2014–2030. https://doi.org/10.1111/j.1365-3040.2012.02533.x

Zhu, X., Li, Y., Fang, W., & Kusano, T. (2018). Galactinol is involved in sequence-conserved upstream open reading frame-mediated repression of Arabidopsis HsfB1 translation. Environmental and Experimental Botany, 156, 120-129. https://doi.org/10.1016/j.envexpbot.2018.09.006

FIGURE LEGENDS

Figure 1

Plant CPuORFs attenuate activity of downstream ORFs. (a) Illustration of typical eukaryote transcripts. Top: textbook transcript with 5'-m^7G cap (red), a 5'-untranslated region (5'-UTR; blue), the major open reading frame (mORF; dark grey) of the transcript, which encodes the main protein product, a 3'-UTR (yellow) and the poly-A tail ($[A]_n$) at the 3'-end of the transcript. Bottom: a transcript with a protein-encoding upstream open reading frame (uORF; light grey) within the 5'-UTR. (b) Summary of constructs used. Arrows represent the promoter (SP6 for *in vitro* assays, 35S *in planta*). Lines represent the 5'-UTR. The black box represents the major open reading frame (mORF), which encodes the luciferase (LUC) reporter. The coloured

box represents the various CPuORFs studied. The dotted box represents CPuORFs in which the start codon (uAUG) has been mutated (red cross) to prevent CPuORF translation, releasing the inhibition of mORF translation. (c) LUC activity measured in wheat germ extracts charged with mutant CPuORF (-uAUG) reporters, relative to the appropriate WT control (CPuORF +uAUG). (d) LUC activity measured in leaves of transgenic Arabidopsis plants transformed with mutant CPuORF (-uAUG) reporters, relative to the appropriate WT control (CPuORF +uAUG). Means ± standard error of the mean are shown. Significant differences between the +uAUG and -uAUG reporter for each CPuORF at $p<0.05$ and $p<0.01$ (Tukey HSD inference) are represented by a single or double asterisk, respectively.

Figure 2

CPuORFs function in a peptide sequence-dependent manner. (a) Clustal Omega alignment of HG17 family CPuORF peptide sequences from diverse angiosperms. Aligned sequences are from the following species: *Solanum lycopersicum* (Sl), *Arabidopsis thaliana* (CPuORF45-48), *Oryza sativa* (OsCPuORF38 and OsHG17b), *Zea mays* (Zm), *Brachypodium distachyon* (Bd), *Hordeum vulgare* (Hv), *Setaria italica* (Si), *Brassica napus* (Bn), *Glycine max* (Gm) and *Solanum lycopersicum* (Sl). Arrows indicate the Arabidopsis (green) and rice (red) HG17 CPuORFs investigated in this study. The green bar below the alignment highlights the highly conserved residues that are likely to include sequences that cause translation stalling. (b) Chart comparing LUC activity for CPuORF47 CPuORF reporters mutated within the conserved C-terminus (sdm1-3; mutations shown in alignment beneath the chart) relative to appropriate wild-type constructs. Bars represent mean relative LUC

activity (± s.e.m) tested *in vitro* (black bars) or *in planta* (striped bars). Significant differences between WT and mutant CPuORFs are indicated ($p<0.05$ (*) and $p<0.01$ (**); Tukey HSD inference). 'ns' indicates no statistical difference.

Figure 3

CPuORFs are translated and cause ribosome stalling. (a-d) positional distribution of ribo-seq reads (red peaks; top) and PARE reads (blue peaks; bottom) on CPuORF-containing transcripts. Each transcript is shown diagrammatically with the CPuORF (black box) and the mORF (grey box). Ribo-seq reads (Liu et al., 2013) are shown for the entirety of each transcript. PARE reads (German et al., 2008) are shown only for the CPuORF sequence (inset). The first nucleotide of the CPuORF stop codon is assigned position 0. (e) The proportion of class I (red bars) and class II (green bars) that accumulate PARE reads at positions -16 and -46, relative to the CPuORF stop codon (position 0) (from data analyses published by Hou et al., 2016).

Figure 4

CPuORFs function as conditional regulators of translation. Translational responses following (a) mannitol, (b) thermospermine, or (c) heat treatments relative to appropriate controls are shown. Means ± standard error of the mean are indicated for HG17 CPuORFs (striped bars), the HG7a CPuORF19 (solid grey bars), the HG15 SAC51_CPuORF (solid black bars), and 35S:LUC (without CPuORF; white bars). Significant differences between controls and treatments for each CPuORF

reporter at $p<0.05$ and $p<0.01$ (Tukey HSD inference) are represented by a single or double asterisk, respectively.

Figure 5

The *SAC51* CPuORF is able to control the phenotypic effects of constitutive *SEP3* expression in a thermospermine-dependent manner. (a) Seedling phenotypes of transgenic lines constitutively expressing *SEP3* (central column) or *SAC51-SEP3* (right column) relative to wild-type plants (left column), with (bottom row) or without (top row) thermospermine treatment. White arrows indicate early, terminal flowers. Red arrows indicate curled/rolled leaves. Both phenotypes are typical of constitutive *SEP3* expression. Note that for *35S:SAC51-SEP3* plants, this phenotype is only evident after thermospermine treatment. (b) Proportion of *35S:SAC51-SEP3* transgenic seedlings showing a typical 35S:*SEP3* phenotype with or without thermospermine (tspm) treatment. Numbers of seedlings examined under each treatment are shown at the top of the chart.

Figure 6

Mechanisms of Conditional uORF-dependent Translational Stalling (CUTS).

Centre: By default, translation of a uORF/CPuORF results in few ribosomes translating the mORF. The image shows a uORF-containing transcript with arrested ribosome at the uORF stop codon. The translated uORF peptide interacts with the

ribosome exit tunnel, causing ribosome stalling and suppressed translation of the mORF on the same transcript.

Top: The repressive CUTS (rCUTS) mode of translational regulation. In this case, a small signal molecule interacts with the nascent uORF peptide in the ribosome exit tunnel, stabilising interactions between the peptide and discrete components of the tunnel. As a consequence, ribosome stalling is significantly enhanced and results in potent inhibition of mORF translation. Therefore, in rCUTS no mORF protein is made when the signal is present.

Bottom: The activating CUTS (aCUTS) mode of translational regulation. Here, the signal molecule interacts with an extra-ribosomal domain of nascent uORF/CPuORF peptide, promoting release of the newly synthesised peptide and 60S ribosomal subunit during translation termination. The 40S subunit remains in contact with the transcript and continues scanning to the AUG of the mORF, where a complete ribosome reassembles, and translation is reinitiated. Therefore, in aCUTS synthesis of the mORF protein is significantly increased when the signal is present.

Blue bars above the transcripts in each panel show the relative occupancy of ribosomes between the CPuORF and mORF, which might be observed in ribosome profiling experiments. In the default situation (centre) ribosomes are arrested in the vicinity of the stop codon, resulting in an accumulation of ribosome profiling reads at the 3’-end of the CPuORF. However, a relatively small number of ribosomes escape arrest and are able to reinitiate translation of the mORF, resulting in the synthesis of a limited amount of mORF peptide. In rCUTS (top), all ribosomes are sequestered by the CPuORF (reads accumulate on CPuORF relative to mORF), while in aCUTS

(bottom), the majority of scanning ribosomes accumulate on the mORF relative to CPuORF.

SUPPORTING INFORMATION

Supporting Figure Legends

Figure S1

Positional distribution of ribo-seq reads (Liu et al., 2013) on CPuORF-containing transcripts annotated on TAIR (arabidopsis.org). Transcripts for CPuORF-containing genes are shown, with the appropriate CPuORF number listed to the left. Black boxes represent the mORF (where annotated). Green boxes represent the annotated CPuORF. Ribo-seq reads (ribosome footprint data retrieved and visualised using the GWIPS-viz genome browser in RiboGalaxy (ribogalaxy.ucc.ie)) for each transcript are shown as red peaks. Note that CPuORFs show a high density of ribosome footprints, indicating that these are translatable elements.

Figure S2

The distribution of PARE reads in Arabidopsis class I and class II CPuORFs. Clustered heat map of PARE reads redrawn from that published by Hou et al. (2016). The first nucleotide of the CPuORF stop codon is position 0. Black/grey blocks represent the peak index value (calculated by dividing the number of PARE reads at a particular position by the number of total reads in a 31-nucleotide flanking region; Hou et al., 2016), with darker colours representing a greater accumulation of reads.

CPuORF number, homology group and whether it is a class I or class II CPuORF, is listed to the left. To highlight the distinct patterns of PARE read accumulation seen for class I and class II CPuORFs, class I are at the top of the heat map (in red), with class II at the bottom (in green). Note that for class I CPuORFs, there is an accumulation of PARE reads at positions -16 and -46, suggesting that this class of CPuORFs stall ribosomes at or near the stop codon.

Figure S3

(a) Summary of constructs used. Arrows represent the *35S* promoter. Lines represent the 5'-UTR. The black box represents the major open reading frame (mORF), which encodes the luciferase (LUC) reporter. The coloured box represents CPuORF47. The dotted box represents CPuORF47 in which the start codon (uAUG) has been mutated (red cross) to prevent its translation, releasing the inhibition of mORF translation. (b) LUC activity measured in leaves from Arabidopsis transformed with the CPuORF47-containing reporter constructs. Activity of CPuORFs with a mutated uAUG (-CPuORF; grey bars) is presented relative to WT CPuORFs (+CPuORF; orange bars), following mock (-) or mannitol (+) treatment of leaf samples. Mean fold-changes ± s.e.m. are shown. Significant differences are shown. Bars with different letters are significantly different from one another (Tukey HSD inference; $p<0.05$). Note that translation of CPuORF47 attenuates that of the mORF (LUC), and is also required for the response of CPuORF47 to mannitol. Data used to generate the chart is presented in the table below.

Figure S4

(a) Clustal Omega alignment of Arabidopsis CPuORF46, CPuORF47 and rice OsCPuORF38 peptide sequences. The yellow box highlights a block of residues that differ between CPuORF46 and CPuORF47. Also note the N-terminal extension of CPuORF47. These may explain the difference in conditional responses seen for these two CPuORFs, which needs further testing. (b) Heat maps of percent amino acid identities shared between the N-terminal (left) and C-terminal (right) regions of HG17 CPuORFs from Arabidopsis and rice. % identities are also shown.

The dotted line in (a) through (b) separates the conserved C-terminal and the divergent N-terminal sequences.

Supporting Tables

Table S1. Summary of all known Arabidopsis CPuORFs.

Table S2. Primers used in this study.

Table S3. *In vitro* and *in planta* reporter assay data for various CPuORF-LUC constructs