Plant Biotechnology 32, 157–163 (2015)
DOI: 10.5511/plantbiotechnology.15.0519a

Original Paper

An upstream open reading frame represses expression of a tomato homologue of Arabidopsis *ANAC096*, a NAC domain transcription factor gene, in a peptide sequence-dependent manner

Abdul Latif Noh[1], Shun Watanabe[2,a], Hiro Takahashi[3], Satoshi Naito[1,2], Hitoshi Onouchi[1,*]

[1] Graduate School of Agriculture, Hokkaido University, Sapporo, Hokkaido 060-8589, Japan; [2] Graduate School of Life Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan; [3] Graduate School of Horticulture, Chiba University, Matsudo, Chiba 271-8510, Japan
* E-mail: onouchi@abs.agr.hokudai.ac.jp Tel: +81-11-706-3887 Fax: +81-11-706-4932

Received March 23, 2015; accepted May 19, 2015 (Edited by T. Mizoguchi)

Abstract Many eukaryotic mRNAs contain one or more upstream open reading frames (uORFs) in their 5′ untranslated regions (5′-UTRs). Some uORFs encode regulatory peptides that repress translation of the main ORF. To comprehensively identify uORFs encoding regulatory peptides, genome-wide searches for uORFs with evolutionarily conserved amino acid sequences, referred to as conserved peptide uORFs (CPuORFs), have been conducted using bioinfomatic approaches. To date, more than 40 homology groups of CPuORFs have been identified in dicotyledonous plants. The *Arabidopsis thaliana ANAC096* gene is one of the CPuORF-containing genes; however, the *ANAC096* CPuORF exerts only little peptide sequence-dependent effect on expression of the main ORF. Here, we investigated the effect of the CPuORF sequence of a tomato *ANAC096* homologue on expression of the main ORF, because it has a more highly conserved amino acid sequence than the *ANAC096* CPuORF. Mutational analyses revealed that the CPuORF of the tomato *ANAC096* homologue represses main ORF expression in a peptide sequence-dependent manner, and determined the critical amino acid residues of the CPuORF peptide responsible for the repression. This study identified a novel peptide sequence-dependent regulatory uORF and demonstrated that the level of uORF peptide-mediated repression can differ among closely related homologues.

Key words: Nascent peptide, ribosome, translational regulation, tomato, upstream open reading frame.

Upstream open reading frames (uORFs) are small open reading frames present in the 5′ untranslated regions (5′-UTRs) of many eukaryotic mRNAs, and often negatively modulate the translational efficiency of the downstream main ORF (Calvo et al. 2009). Usually, the peptides encoded by uORFs are not involved in the modulation of main ORF translation. However, some uORFs have been shown to repress translation of the main ORF in a peptide sequence-dependent manner (Ito and Chiba 2013; Morris and Geballe 2000). In this case, the nascent peptide encoded by a uORF acts inside the ribosome that had synthesized it to cause ribosome stalling, which results in translational repression of the main ORF (Cao and Geballe 1996; Law et al. 2001; Uchiyama-Kadokura et al. 2014; Wang and Sachs 1997; Wang et al. 1999).

To comprehensively identify uORFs encoding regulatory peptides, genome-wide searches for uORFs with conserved amino acid sequences, which are referred to as 'conserved peptide uORFs (CPuORFs)' (Hayden and Jorgensen 2007), have been conducted in various organisms, such as mammals (Crowe et al. 2006), plants (Hayden and Jorgensen 2007; Takahashi et al. 2012; Tran et al. 2008; Vaughn et al. 2012) and insects (Hayden and Bosco 2008), using comparative genomic analyses. In dicotyledonous plants, more than 40 homology groups of CPuORFs have been identified. Recently, we analyzed 16 *Arabidopsis thaliana* CPuORFs for their sequence-dependent effects on expression of the main ORF using a transient expression assay (Ebina et al. 2015). This analysis identified five CPuORFs that have a peptide sequence-dependent inhibitory effect on main ORF expression, and revealed that most of the CPuORFs that showed a sequence-dependent effect have a characteristic feature that the C-terminal amino acid sequence and

Abbreviations: CPuORF, conserved peptide uORF; FLUC, firefly luciferase; fs, frameshift; NAC, NAM, ATAF1,2 and CUC2; ORF, open reading frame; PEG, polyethylene glycol; RLUC, *Renilla reniformis* luciferase; uORF, upstream ORF; UTR, untranslated region; WT, wild-type.

[a] Present address: Wakou Shokuhin Co., Ltd., Otaru, Hokkaido 047-0261, Japan

* To whom correspondence should be addressed

This article can be found at http://www.jspcmb.jp/
Published online June 10, 2015

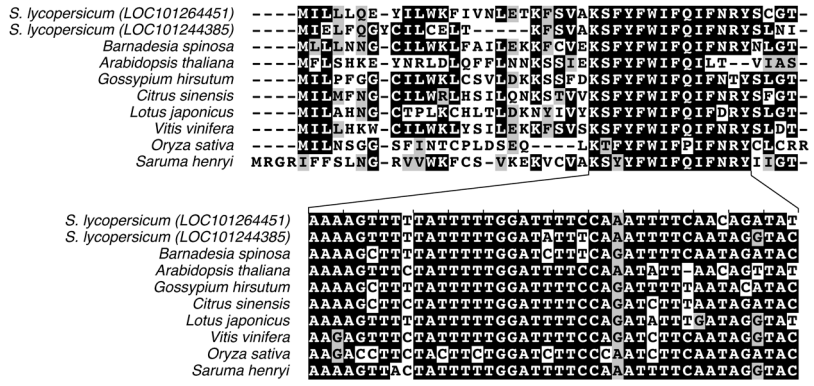


Figure 1. Alignment of the CPuORF sequences of the *ANAC096* homologues. Amino acid sequence alignment of the *ANAC096* homology group CPuORFs (upper) and nucleotide sequence alignment of the highly conserved region of the CPuORFs (lower) are indicated. The amino acid and nucleotide sequences of the CPuORFs were derived from RefSeq, expressed sequence tag or cDNA sequences (accession no. XM_004245832.1, XM_004241924.2, GE533448.1, BT020305.1, ES799867.1, XM_006485373.1, FS335485.1, XM_002280858.3, CT848736.1, and DT587338.1), aligned using ClustalW version 2.1 (http://clustalw.ddbj.nig.ac.jp/), and visualized using Boxshade version 3.21 (http://www.ch.embnet.org/software/BOX_form.html). For *Solanum lycopersicum*, the CPuORF sequences of the *LOC101264451* and *LOC101244385* genes are indicated.

the stop codon position are evolutionarily conserved. In this analysis, the CPuORF of the *ANAC096* gene, whose main ORF encodes a NAC (NAM, ATAF1,2 and CUC2) domain-containing transcription factor involved in dehydration and osmotic stress responses (Xu et al. 2013), showed only little sequence-dependent effect on main ORF expression (Ebina et al. 2015). Although the C-terminal amino acid sequence and the stop codon position are evolutionarily conserved among dicot homologues of the *ANAC096* CPuORF, the C-terminal six amino acid residues encoded by the *A. thaliana ANAC096* CPuORF are different from those conserved among other dicot homologues (Figure 1). Therefore, it is possible that the CPuORFs of other dicot *ANAC096* homologues with the conserved C-terminal amino acid sequence may exert a stronger sequence-dependent inhibitory effect than the *A. thaliana ANAC096* CPuORF.

In the present study, we examined the effect of the CPuORF of a tomato *ANAC096* homologue, *LOC101264451*, in which the C-terminal amino acid sequence is highly conserved (Figure 1). Mutational analyses revealed that the CPuORF of the *LOC101264451* gene has a strong sequence-dependent inhibitory effect on main ORF expression, and that its amino acid sequence is responsible for the effect.

Materials and methods

Plant material and growth condition

Tobacco BY-2 cells (*Nicotiana tabacum* L. cv. Bright Yellow 2) (Kato et al. 1972) were cultured in modified Linsmaier and Skoog (LS) medium (Nagata et al. 1992) at 26°C in the dark with orbital shaking at 130 rpm. Cells were transferred to fresh medium every week.

Plasmid construction

To clone the 5′-UTR of the tomato *LOC101264451* gene, total RNA was prepared from tomato (*Solanum lycopersicum* cv. Rutgers) seedlings using a Qiagen Plant RNeasy Mini Kit (Qiagen), and cDNA of the *LOC101264451* 5′-UTR was amplified by reverse transcription PCR using the OneStep RT-PCR Kit (Qiagen) with primers 101264451F (5′-GAC TCT AGA ACT CTA CAC AAC TTC TCT TTC TTC TA-3′) and 101264451R (5′-CAA GTC GAC CCT GCC ATG ATG AAG TAA CAG AAT T-3′), which were designed based on an NCBI reference sequence (RefSeq) of *LOC101264451* mRNA (accession no. XM_004241924.1). The amplified fragment was digested with *Xba*I and *Sal*I, and inserted into the *Xba*I and *Sal*I sites between the cauliflower mosaic virus 35S RNA (35S) promoter and the *Renilla* luciferase (RLUC) coding sequence of pIE0 (Ebina et al. 2015) to generate the *35S::UTR(WT):RLUC* reporter plasmid (Figure 2B). Mutations were introduced into the *LOC101264451* CPuORF of the reporter plasmid using the overlap extension PCR method (Ho et al. 1989) with primers listed in Supplementary Table S1. In all of the constructs, sequence analysis confirmed the integrity of the PCR-amplified regions.

Transient expression assay

To prepare protoplasts, BY-2 cells were collected by centrifugation on the third day after transfer to fresh media, and suspended in modified LS medium containing 1% (w/v) cellulase Onozuka RS (Yakult Pharmaceutical Industry), 0.5% (w/v) pectolyase Y23 (Seishin Pharmaceutical) and 0.4 M mannitol, and incubated at 26°C with gentle shaking until the suspension became turbid with protoplasts. The protoplasts were then washed five times with wash buffer (0.4 M mannitol, 5 mM $CaCl_2$ and 12.5 mM NaOAc, pH 5.8) and suspended in MaMg solution (5 mM morpholinoethanesulfonic acid, 15 mM $MgCl_2$ and 0.4 M mannitol, pH 5.8). Five micrograms each of the *35S::UTR:RLUC* reporter plasmid and the *35S::FLUC* internal control plasmid, which carries a firefly luciferase (FLUC) coding sequence under the control of the 35S promoter (Matsuo et al. 2001), were mixed with 3.0×10^5 protoplasts

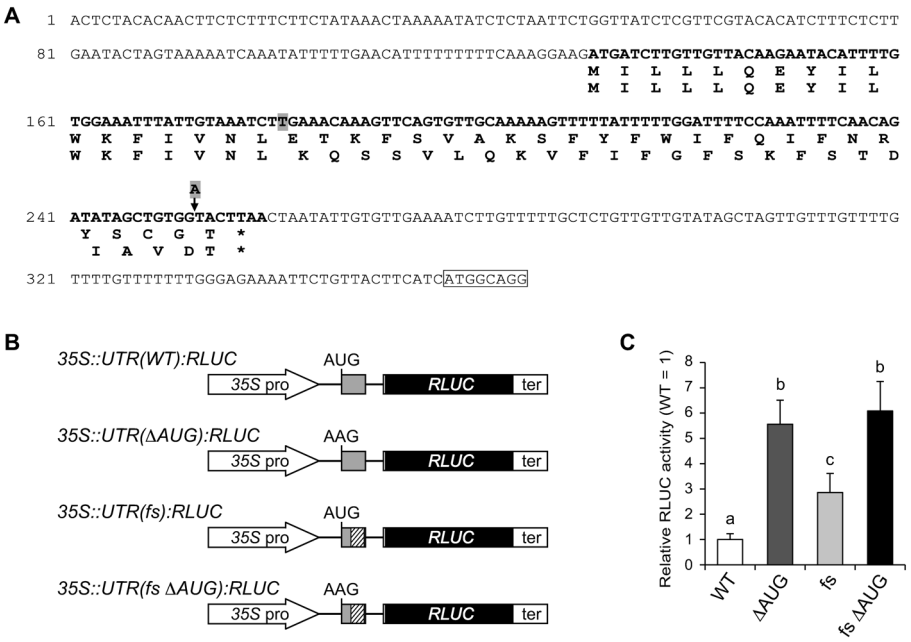


Figure 2. Effects of the presence and the peptide sequence of the *LOC101264451* CPuORF on expression of the main ORF. (A) Nucleotide sequence of the *LOC101264451* 5′-UTR used in this study. The nucleotide sequence of the CPuORF is shown in bold, and the deduced amino acid sequences of the WT and fs mutant CPuORFs are indicated. The nucleotides that were deleted and inserted in the fs mutant are shaded. The 5′-part of the main ORF that was contained in the reporter constructs is boxed. (B) Schematic representation of the reporter constructs. The shaded box represents the *LOC101264451* CPuORF. The hatched box shows the frame-shifted region in the fs mutant. The polyadenylation signal of the *Agrobacterium tumefaciens NOS* gene is designated as 'ter.' (C) Transient expression assay. The *35S::UTR:RLUC* reporter plasmid carrying the WT CPuORF or a mutant CPuORF was co-transfected with the *35S::FLUC* internal control plasmid (Ebina et al. 2015) into tobacco BY-2 protoplasts by polyethylene glycol (PEG) treatment. After 24 h of incubation, the transfected cells were harvested and disrupted for luciferase assay. RLUC activity was normalized to FLUC activity, and the relative activity to that of the WT construct was calculated. Means±SD of at least three biological replicates are shown. The graph is representative of two separate experiments using independently prepared protoplasts. Columns with different letters are significantly different at $p<0.05$ by *t*-test.

in 100 μl of MaMg solution and 110 μl of polyethylene glycol (PEG) solution (40% PEG4000, 0.1 mM $CaCl_2$, 0.2 M mannitol). This mixture was incubated for 15 min at room temperature, and diluted by adding 800 μl of wash buffer. The protoplasts were centrifuged and resuspended in 1 ml of the modified LS medium containing 0.4 M mannitol. After 24 h of incubation at 26°C in the dark, cells were harvested and disrupted in 200 μl of extraction buffer [100 mM $(NaH_2/Na_2H)PO_4$, 5 mM dithiothreitol, pH 7] by sonication on ice with a Branson Sonifier 250. A Dual-LUC Reporter Assay Kit (Promega) was used to measure the RLUC and FLUC activities.

Results

The LOC101264451 *CPuORF has an inhibitory effect on main ORF expression*

Among tomato NAC domain-containing transcription factor genes, *LOC101264451* and *LOC101244385* are the closest homologues of *ANAC096* (Supplementary Figure S1). Although both genes have a CPuORF belonging to the same homology group as the *ANAC096* CPuORF, the *LOC101264451* CPuORF has a more highly conserved amino acid sequence than that of *LOC101244385* (Figure 1). Therefore, we used the *LOC101264451* CPuORF for the functional analysis. For this analysis, the *LOC101264451* 5′-UTR was fused to the *RLUC* coding sequence and placed under the control of the 35S promoter to generate the *35S::UTR(WT):RLUC* construct (Figure 2B). In this construct, the first eight nucleotides of the *LOC101264451* main ORF was in-frame fused to the *RLUC* coding sequence (Figures 2A, B). To test the effect of the *LOC101264451* CPuORF on expression of the main ORF, we eliminated the start codon of the CPuORF from the *35S::UTR(WT):RLUC* construct by changing the start codon to an AAG codon. This mutant construct was designated *35S::UTR(ΔAUG):RLUC* (Figure 2B). The *35S::UTR(WT):RLUC* and *35S::UTR(ΔAUG):RLUC* constructs were separately introduced into protoplasts prepared from tobacco BY-2 suspension cultured cells (Kato et al. 1972). After 24 h of incubation, cells were harvested and disrupted for analysis of RLUC activity. As shown in Figure 2C, the removal of the CPuORF start codon increased the RLUC activity by 5.6-fold compared with the wild-type (WT) construct. This result indicates that the *LOC101264451* CPuORF has an inhibitory effect on expression of the main ORF.

The effect of the LOC101264451 *CPuORF is peptide sequence dependent*

The amino acid sequence of the region comprising the 25th to 38th codons of the *LOC101264451* CPuORF is highly conserved, whereas synonymous codon changes are found for all the codons in this region except for the Trp codon, for which there is no synonymous alternative, in any of the species used to create the alignments shown in Figure 1. This implies that the amino acid sequence of this region is important for the inhibitory effect. To examine the importance of the amino acid sequence of the highly conserved region, we altered the amino acid sequence of this region by introducing −1 and +1 frameshift (fs) mutations upstream and downstream of the highly conserved region, respectively (Figure 2A). These fs mutations were introduced into the *35S::UTR(WT):RLUC* and *35S::UTR(ΔAUG):RLUC* constructs (Figure 2B), and the effect of the fs mutations on the expression of the *RLUC* reporter gene were examined by the transient expression assay. When the fs mutations were introduced into the WT CPuORF, the reporter activity increased by 2.9-fold (Figure 2C). This result shows that the effect of the *LOC101264451* CPuORF on main ORF expression is dependent on its sequence. It is likely that the amino acid sequence rather than the nucleotide sequence of the CPuORF is responsible for the sequence-dependent effect, because 21 amino acid residues were altered in the fs mutant whereas only two nucleotide changes were introduced outside of the highly conserved region. By contrast, in the absence of the start codon (*ΔAUG*), the fs mutations caused no further significant increase in the RLUC activity (Figure 2C), which indicates that the effect of the fs mutations depends on translation of the CPuORF. These results suggest that the peptide translated from the *LOC101264451* CPuORF represses expression of the main ORF.

Identification of critical residues of the LOC101264451 *uORF peptide*

Next, we performed Ala-scanning mutagenesis to identify critical amino acid residues of the *LOC101264451* CPuORF-encoded peptide responsible for the repressive function. For this analysis, we selected 11 residues in the highly conserved region and 9 residues outside of the highly conserved region, and individually changed them to Ala. In each region, relatively highly conserved residues were selected although there were some exceptions (Figures 1, 3). The effect of each Ala substitution on expression of the downstream *RLUC* reporter gene was examined by the transient expression assay. To compare the functional importance of each amino acid residue and its evolutionary conservation level, conservation scores were calculated using the CPuORF amino acid sequences of the *ANAC096* homologues used for the alignment presented in Figure 1 except for *LOC101244385* (Figure 3). Among the Ala substitutions introduced into the highly conserved region, most of the Ala substitutions in the region comprising the 25th to 32nd residues significantly increased the RLUC activity compared with the WT (Figure 3). Among the Ala substitutions introduced outside of the highly conserved region, only T42A increased the RLUC activity. These results suggest that at least the region comprising the 25th to 32nd residues and the last residue of the *LOC101264451* CPuORF peptide are critical for its repressive activity.

Synonymous codon changes in the critical region

To further confirm the peptide sequence-dependence of the effect of the *LOC101264451* CPuORF, we introduced synonymous mutations into the critical region comprising the 25th to 32nd codons (Syn25–32) (Figure 4A), and tested its effect on expression of the downstream *RLUC* reporter gene by the transient

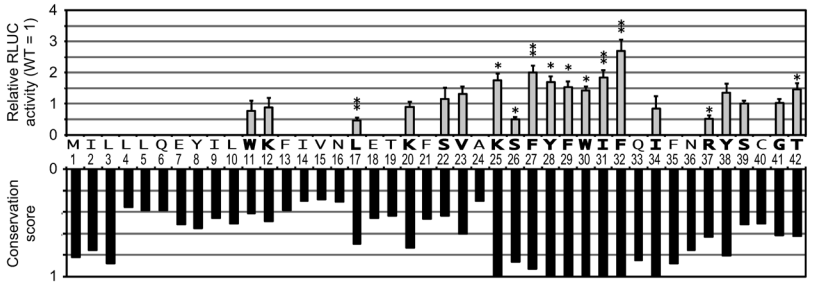


Figure 3. Ala scanning of the *LOC101264451* CPuORF. The *35S::UTR:RLUC* reporter plasmid harboring the WT CPuORF or its mutant form with an Ala substitution was co-transfected with the *35S::FLUC* internal control plasmid into BY-2 protoplasts by PEG treatment, and the reporter activities were analyzed as in Figure 2. Means±SD of at least three biological replicates are shown. The upper graph is representative of two separate transient assay experiments using independently prepared protoplasts. Single and double asterisks indicate significant differences from the WT at $p<0.05$ and $p<0.01$ by *t*-test, respectively. The amino acid sequence of the *LOC101264451* CPuORF is indicated below the graph, and the amino acid residues analyzed in the Ala-scanning mutagenesis are shown in bold. The numbers below the amino acid residues indicate the positions of the residues in the CPuORF peptide. Conservation score of each amino acid residue was calculated using the Scorecons server (https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/valdar/scorecons_server.pl) and shown in the lower graph.

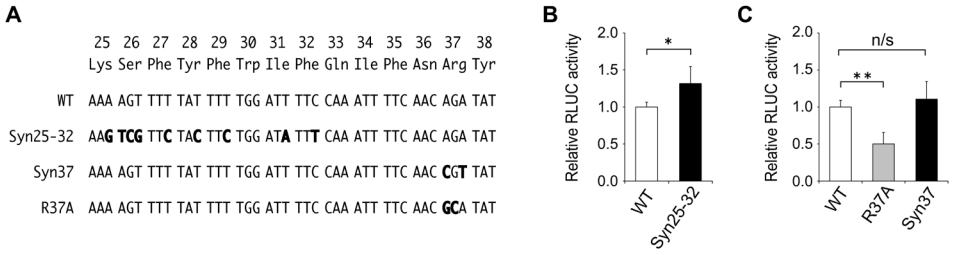


Figure 4. Effects of synonymous mutations in the *LOC101264451* CPuORF. (A) Nucleotide sequences of the synonymous mutations introduced into the *LOC101264451* CPuORF. The amino acid sequence and the WT and mutant (Syn25–32, Syn37 and R37A) nucleotide sequences of the highly conserved region of the CPuORF are indicated. The mutated nucleotides are shown in bold. (B, C) Transient expression assay. The *35S::UTR:RLUC* reporter plasmid harboring the WT or a mutant CPuORF was co-transfected with the *35S::FLUC* internal control plasmid into BY-2 protoplasts by PEG treatment, and the reporter activities were analyzed as in Figure 2. Means±SD of at least three biological replicates are shown. Each graph is representative of three separate experiments using independently prepared protoplasts. Single and double asterisks indicate significant differences between two constructs at $p<0.05$ and $p<0.01$ by *t*-test, respectively, whereas 'n/s' indicate a non-significant difference.

expression assay. As shown in Figure 4B, the Syn25–32 mutation slightly increased the RLUC activity. However, despite the alterations of nine nucleotides in the critical region, the effect of the Syn25–32 mutation was smaller than those of the single amino acid changes in this region that up-regulated the RLUC activity in the Ala-scanning analysis (Figures 3, 4B). Furthermore, the Syn25–32 mutant showed a much weaker effect than the fs mutant, in which no nucleotide change was introduced into the critical region but the amino acid sequence of this region was altered (Figures 2, 4B). Additionally, we introduced a synonymous mutation in R37 (Syn37) (Figure 4A) and examined its effect. In contrast to the effect of R37A, the Syn37 mutation showed no significant effect (Figure 4C). These results support the conclusion that the amino acid sequence of the *LOC101264451* CPuORF is involved in the repression of main ORF expression.

Discussion

In our recent study, we analyzed 16 of 41 uncharacterized CPuORF homology groups for their sequence-dependent effects on main ORF expression, using *A. thaliana* CPuORFs, and identified five peptide sequence-dependent regulatory uORFs (Ebina et al. 2015). Based on this result and the fact that two uORFs were previously reported to encode regulatory peptides in *A. thaliana* (Hanfrey et al. 2005; Rahmani et al. 2009), we estimated that there would be at least 15 homology groups of peptide sequence-dependent regulatory uORFs in dicotyledonous plants (Ebina et al. 2015). However, in some of the *A. thaliana* CPuORFs that showed no or little sequence-dependent effect, the amino acid sequences are not well conserved compared with those of other plant homologues. In the present study, we analyzed the effect of a tomato homologue of one of such CPuORFs on main ORF expression, and mutational analyses revealed that the tomato *LOC101264451* CPuORF strongly represses main ORF expression in a peptide sequence-dependent manner. This result is in contrast to the previous observation that the CPuORF of *ANAC096*, an *A. thaliana* homologue of *LOC101264451*, showed only little sequence-dependent effect (Ebina et al. 2015). Therefore, the present study demonstrates that the level of uORF peptide-mediated repression can differ among closely related homologues, and that, even if a member of a certain CPuORF homology group show no or little sequence-dependent effect on main ORF expression, a sequence-dependent regulatory uORF can be identified by analyzing another member with higher amino acid sequence conservation. These findings suggest that dicot genomes contain a higher number of peptide sequence-dependent regulatory uORFs than our previous estimation.

Ala-scanning analysis shown in Figure 3 revealed that the region comprising the 25th to 32nd residues and the last residue are critical for the repressive activity of the *LOC101264451* CPuORF peptide. In the previously characterized nascent peptide-mediated regulations, the nascent peptide interacts with the components of the ribosomal exit tunnel at the narrowest constricted region to cause ribosome stalling, and thereby controls expression of the downstream cistron (Cruz-Vera et al. 2005; Ito and Chiba 2013; Nakatogawa and Ito 2002; Wu et al. 2012). Cryo-electron microscopy analysis revealed that, when the nascent peptides encoded by the *Neurospora crassa arg-2* and cytomegalovirus *gpUL4* regulatory uORFs cause ribosome stalling, the region comprising the ninth to 19th residues from the C-terminus and the region around the 11th residue from the C-terminus, respectively, are in contact with exit tunnel components located at the constricted region (Bhushan et al. 2010). If the *LOC101264451* CPuORF peptide causes ribosome stalling by acting as a regulatory nascent peptide, it is likely that ribosome stalling occurs after the last amino acid-coding codon, T42, is translated, because T42A affected the *RLUC* expression (Figure 3). If this is the case, some amino acid residues in the critical region comprising the 25th to 32nd residues may interact with the exit tunnel components at the constricted

region, considering the distance of 10 to 17 residues between the critical region and the C-terminus. In the *A. thaliana ANAC096* CPuORF, the distance between the critical region and the stop codon is different from those conserved among other dicot homologues (Figure 1). Therefore, in *A. thaliana*, such interactions between the exit tunnel components and the nascent peptide might be impaired due to the non-conserved stop codon position and/or the low conservation of the C-terminal amino acid sequence of the *ANAC096* CPuORF. In contrast to the other Ala substitutions introduced into the *LOC101264451* CPuORF that affected the *RLUC* expression, L17A, S26A and R37A decreased the RLUC activity compared with the WT (Figure 3). Possibly, these mutations might change the spatial position of the amino acid residues interacting with the constricted region of the exit tunnel, and thereby make the interaction stronger, resulting in an enhancement of ribosome stalling.

Although the present study revealed that the *LOC101264451* CPuORF represses expression of the main ORF in a peptide sequence-dependent manner, the physiological role of the CPuORF-mediated repression of the *LOC101264451* gene expression remains to be elucidated. Among the previously characterized sequence-dependent regulatory uORFs, the uORF of the human *CHOP* gene has been shown to be involved in the stress-responsive induction of main ORF translation (Palam et al. 2011). The *CHOP* uORF represses translation of the main ORF in a sequence-dependent manner (Jousse et al. 2001), but it is frequently bypassed by scanning ribosomes under endoplasmic reticulum stress condition, resulting in an enhancement of main ORF translation (Palam et al. 2011). Expression of the *ANAC096* gene is induced at the mRNA level in response to dehydration and osmotic stress in *A. thaliana* (Xu et al. 2013). If the protein encoded by the *LOC101264451* main ORF is a functional orthologue of ANAC096, expression of *LOC101264451* could be induced at the translational level in response to these stresses as well as at the mRNA level, and the *LOC101264451* CPuORF could be involved in the stress-responsive induction of main ORF translation by a similar mechanism to the *CHOP* uORF system.

In conclusion, this study identified a novel peptide sequence-dependent regulatory uORF in plants. Further studies on the uORF-medeated regulation of the *LOC101264451* gene will provide a better understanding of the regulatory mechanism and the physiological role of uORF peptide-mediated gene expression control.

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science [Grant-in-Aid for Scientific Research (C) (25440122) to H.O.]; the Ministry of Education, Culture, Sports, Science and Technology of Japan [Grant-in-Aid for Scientific Research on Innovative Areas (21112501) to H.O. and (22119006) to S.N.]. We are grateful to Dr. Shunsuke Imanishi of NARO Institute of Vegetable and Tea Science, National Agriculture and Food Research Organization for providing seeds of *S. lycopersicum* cv. Rutgers. We thank Ms. Hitomi Sekihara and Ms. Eriko Tanaka for skillful technical assistance, and Ms. Maki Mori for general assistance. We used the DNA sequencing facility of the Graduate School of Agriculture, Hokkaido University.

References

Bhushan S, Meyer H, Starosta AL, Becker T, Mielke T, Berninghausen O, Sattler M, Wilson DN, Beckmann R (2010) Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide. *Mol Cell* 40: 138–146

Calvo SE, Pagliarini DJ, Mootha VK (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci USA* 106: 7507–7512

Cao J, Geballe AP (1996) Coding sequence-dependent ribosomal arrest at termination of translation. *Mol Cell Biol* 16: 603–608

Crowe ML, Wang XQ, Rothnagel JA (2006) Evidence for conservation and selection of upstream open reading frames suggests probable encoding of bioactive peptides. *BMC Genomics* 7: 16

Cruz-Vera LR, Rajagopal S, Squires C, Yanofsky C (2005) Features of ribosome-peptidyl-tRNA interactions essential for tryptophan induction of tna operon expression. *Mol Cell* 19: 333–343

Ebina I, Takemoto-Tsutsumi M, Watanabe S, Koyama H, Endo Y, Kimata K, Igarashi T, Murakami K, Kudo R, Ohsumi A, et al. (2015) Identification of novel *Arabidopsis thaliana* upstream open reading frames that control expression of the main coding sequences in a peptide sequence-dependent manner. *Nucleic Acids Res* 43: 1562–1576

Hanfrey C, Elliott KA, Franceschetti M, Mayer MJ, Illingworth C, Michael AJ (2005) A dual upstream open reading frame-based autoregulatory circuit controlling polyamine-responsive translation. *J Biol Chem* 280: 39229–39237

Hayden CA, Bosco G (2008) Comparative genomic analysis of novel conserved peptide upstream open reading frames in *Drosophila melanogaster* and other dipteran species. *BMC Genomics* 9: 61

Hayden CA, Jorgensen RA (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 Biol* 5: 32

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77: 51–59

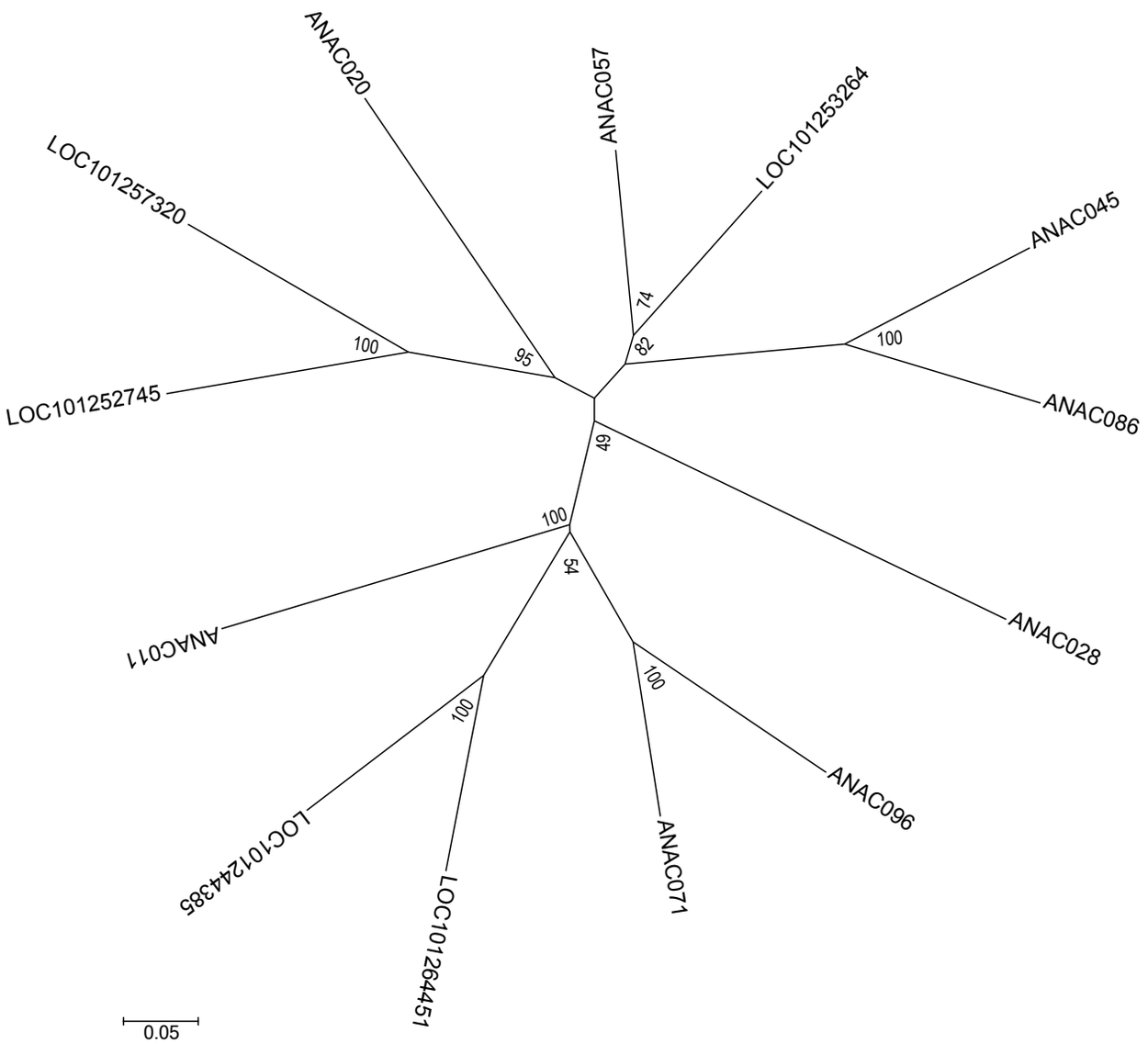
Ito K, Chiba S (2013) Arrest peptides: *cis*-acting modulators of translation. *Annu Rev Biochem* 82: 171–202

Jousse C, Bruhat A, Carraro V, Urano F, Ferrara M, Ron D, Fafournoux P (2001) Inhibition of CHOP translation by a peptide encoded by an open reading frame localized in the chop 5′UTR. *Nucleic Acids Res* 29: 4341–4351

Kato K, Matsumoto T, Koiwai A, Mizusaki S, Nishida K, Noguchi M, Tamaki E (1972) Liquid suspension culture of tobacco cells. In: Terui G (ed) *Ferment Technology Today*. Society of

Fermentation Technology, Osaka, pp 689–695
Law GL, Raney A, Heusner C, Morris DR (2001) Polyamine regulation of ribosome pausing at the upstream open reading frame of *S*-adenosylmethionine decarboxylase. *J Biol Chem* 276: 38036–38043
Matsuo N, Minami M, Maeda T, Hiratsuka K (2001) Dual luciferase assay for monitoring gene expression in higher plants. *Plant Biotechnol* 18: 71–75
Morris DR, Geballe AP (2000) Upstream open reading frames as regulators of mRNA translation. *Mol Cell Biol* 20: 8635–8642
Nagata T, Nemoto Y, Hasezawa S (1992) Tobacco BY-2 cell line as the "Hela" cell in the cell biology of higher plants. *Int Rev Cytol* 132: 1–30
Nakatogawa H, Ito K (2002) The ribosomal exit tunnel functions as a discriminating gate. *Cell* 108: 629–636
Palam LR, Baird TD, Wek RC (2011) Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. *J Biol Chem* 286: 10939–10949
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 Physiol* 150: 1356–1367
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* 28: 2231–2241
Tran MK, Schultz CJ, Baumann U (2008) Conserved upstream open reading frames in higher plants. *BMC Genomics* 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 Physiol* 55: 1556–1567
Vaughn JN, Ellingson SR, Mignone F, Arnim A (2012) Known and novel post-transcriptional regulatory sequences are conserved across plant families. *RNA* 18: 368–384
Wang Z, Gaba A, Sachs MS (1999) A highly conserved mechanism of regulated ribosome stalling mediated by fungal arginine attenuator peptides that appears independent of the charging status of arginyl-tRNAs. *J Biol Chem* 274: 37565–37574
Wang Z, Sachs MS (1997) Ribosome stalling is responsible for arginine-specific translational attenuation in *Neurospora crassa*. *Mol Cell Biol* 17: 4904–4913
Wu C, Wei J, Lin PJ, Tu L, Deutsch C, Johnson AE, Sachs MS (2012) Arginine changes the conformation of the arginine attenuator peptide relative to the ribosome tunnel. *J Mol Biol* 416: 518–533
Xu ZY, Kim SY, Hyeon DY, Kim DH, Dong T, Park Y, Jin JB, Joo SH, Kim SK, Hong JC, et al. (2013) The *Arabidopsis* NAC transcription factor ANAC096 cooperates with bZIP-type transcription factors in dehydration and osmotic stress responses. *Plant Cell* 25: 4708–4724

Supplementary Figure S1



Supplementary Figure S1. Phylogenetic tree of tomato and *A. thaliana ANAC011* family NAC domain-containing proteins. To generate the phylogenetic tree, the amino acid sequences of *A. thaliana ANAC011* family proteins (accession no. NP_199471.1, NP_193532.1, NP_174529.2, NP_188400.1, NP_176766.1, NP_186970.2, NP_175835.2 and NP_197228.1) and their tomato homologues (accession no. XP_004245880.1, XP_004241972.1, XP_010316553.1, XP_004235381 and XP_010315723.1) were analyzed using ClustalW version 2.1. The phylogenetic tree was constructed and bootstrap values were calculated using the MEGA software version 6.06 (http://www.megasoftware.net/index.php). The bootstrap values are indicated at the nodes as a percentage of 1,000 replications.

Noh et al.

Supplementary Table S1. List of primers used to introduce the mutations into the *LOC101264451* CPuORF.

Mutation	Primer name	Direction	Primer sequence [a)]
ΔAUG	101264451ΔAUGf	Forward	5′-CAAAGGAAGAaGATCTTGTTGTTACAAGAA-3′
	101264451ΔAUGr	Reverse	5′-CTTGTAACAACAAGATCtTCTTCCTTTGAA-3′
fs	101264451FSf1	Forward	5′-GGAAATTTATTGTAAATCTGAAACA-3′
	101264451FSr1	Reverse	5′-TGTTTCAGATTTACAATAAATTTCCAAC-3′
	101264451FSf2	Forward	5′-GCTGTGGATACTTAACTAATA-3′
	101264451FSr2	Reverse	5′-TAGTTAAGTATCCACAGCTA-3′
W11A	101264451W11Af	Forward	5′-GAATACATTTTGgcGAAATTTATTGTAAATCTTGAA-3′
	101264451W11Ar	Reverse	5′-TTCAAGATTTACAATAAATTTCgcCAAAATGTATTC-3′
K12A	101264451K12Af	Forward	5′-GAATACATTTTGTGGgcATTTATTGTAAATCTTGAAACA-3′
	101264451K12Ar	Reverse	5′-TGTTTCAAGATTTACAATAAATgcCCACAAAATGTATTC-3′
L17A	101264451L17Af	Forward	5′-TGGAAATTTATTGTAAATgcTGAAACAAAGTTCAGT-3′
	101264451L17Ar	Reverse	5′-ACTGAACTTTGTTTCAgcATTTACAATAAATTTCCA-3′
K20A	101264451K20Af	Forward	5′-GTAAATCTTGAAACAgcGTTCAGTGTTGCA-3′
	101264451K20Ar	Reverse	5′-TGCAACACTGAACgcTGTTTCAAGATTTAC-3′
S22A	101264451S22Af	Forward	5′-GAAACAAAGTTCgcTGTTGCAAAAAGT-3′
	101264451S22Ar	Reverse	5′-ACTTTTTGCAACAgcGAACTTTGTTTC-3′
V23A	101264451V23Af	Forward	5′-GAAACAAAGTTCAGTGcTGCAAAAAGTTTTTATTTTTGGA-3′
	101264451V23Ar	Reverse	5′-CTTTTTGCAgCACTGAACTTTGTTTCAAGA-3′
K25A	101264451K25Af	Forward	5′-GTTCAGTGTTGCAgcAAGTTTTTATTTTTGGATTTTC-3′
	101264451K25Ar	Reverse	5′-CCAAAAATAAAAACTTgcTGCAACACTGAACT-3′
S26A	101264451S26Af	Forward	5′-GAAACAAAGTTCAGTGTTGCAAAAgcTTTTTATTTTTGGATTTTC-3′
	101264451S26Ar	Reverse	5′-GAAAATCCAAAAATAAAAAgcTTTTGCAACACTGAACTTTGTTTC-3′
F27A	101264451F27Af	Forward	5′-GTGTTGCAAAAAGTgcTTATTTTTGGATTTTCCAAATTTTC-3′
	101264451F27Ar	Reverse	5′-GGAAAATCCAAAAATAAgcACTTTTTGCAACACTGAA-3′
Y28A	101264451Y28Af	Forward	5′-GTGTTGCAAAAAGTTTTgcTTTTTGGATTTTCCAAATTTTC-3′
	101264451Y28Ar	Reverse	5′-GGAAAATCCAAAAAgcAAAACTTTTTGCAACACTGA-3′
F29A	101264451F29Af	Forward	5′-GTGTTGCAAAAAGTTTTTATgcTTGGATTTTCCAAATTTTCAAC-3′
	101264451F29Ar	Reverse	5′-GTTGAAAATTTGGAAAATCCAAgcATAAAAACTTTTTGCAACAC-3′
W30A	101264451W30Af	Forward	5′-GTTGCAAAAAGTTTTTATTTTgcGATTTTCCAAATTTTCAAC-3′
	101264451W30Ar	Reverse	5′-GTTGAAAATTTGGAAAATCgcAAAATAAAAACTTTTTGCAAC-3′
I31A	101264451I31Af	Forward	5′-GTTTTTATTTTTGGgcTTTCCAAATTTTCAACAGATATAGC-3′
	101264451I31Ar	Reverse	5′-GCTATATCTGTTGAAAATTTGGAAAgcCCAAAAATAAAAAC-3′
F32A	101264451F32Af	Forward	5′-GTTTTTATTTTTGGATTgcCCAAATTTTCAACAGATATAGC-3′
	101264451F32Ar	Reverse	5′-GCTATATCTGTTGAAAATTTGGgcAATCCAAAAATAAAAAC-3′
I34A	101264451I34Af	Forward	5′-GTTTTTATTTTTGGATTTTCCAAgcTTTCAACAGATATAGC-3′
	101264451I34Ar	Reverse	5′-GCTATATCTGTTGAAAgcTTGGAAAATCCAAAAATAAAAAC-3′
R37A	101264451R37Af	Forward	5′-CCAAATTTTCAACgcATATAGCTGTGGTACTTAAC-3′
	101264451R37Ar	Reverse	5′-CCACAGCTATATgcGTTGAAAATTTGGAAAATCCAA-3′
Y38A	101264451Y38Af	Forward	5′-CCAAATTTTCAACAGAgcTAGCTGTGGTACTTAAC-3′
	101264451Y38Ar	Reverse	5′-CCACAGCTAgcTCTGTTGAAAATTTGGAAAATCCAA-3′
S39A	101264451S39Af	Forward	5′-CCAAATTTTCAACAGATATgcCTGTGGTACTTAACTAATATTG-3′
	101264451S39Ar	Reverse	5′-CAATATTAGTTAAGTACCACAGgcATATCTGTTGAAAATTTGG-3′
G41A	101264451G41Af	Forward	5′-GATATAGCTGTGcTACTTAACTAATATTGTGTTGA-3′
	101264451G41Ar	Reverse	5′-AACACAATATTAGTTAAGTAgCACAGCTATATCT-3′
T42A	101264451T42Af	Forward	5′-GCTGTGGTgCTTAACTAATATTGTGTTGA-3′
	101264451T42Ar	Reverse	5′-CAACACAATATTAGTTAAGcACCACAGCTA-3′
Syn25-32	101264451Syn1F	Forward	5′-gtcgTTcTAcTTcTGGATaTTtCAAATTTTCAACAGATATAGCTGTGGTA-3′
	101264451Syn1R	Reverse	5′-GaAAtATCCAgAAgTAgAAcgacTTTGCAACACTGAACTTTGTTTCAAGA-3′
Syn37	101264451Syn37F	Forward	5′-CCAAATTTTCAACcGtTATAGCTGTGGTACTTAAC-3′
	101264451Syn37R	Reverse	5′-CCACAGCTATAaCgGTTGAAAATTTGGAAAATCCAA-3′

[a)] The nucleotide that was inserted in the fs mutant is boxed. Shaded letters denote the position of the deletion in the fs mutant, in which a nucleotide between the shaded nucleotides was deleted. Lowercase letters indicate substituted nucleotides.