**Running Title:** Reporter gene-specific hyperexpression

**Author:** Hisashi Koiwa[1,2,3], Akihito Fukudome[2]

**Address:**

[1]Vegetable and Fruit Improvement Center, Dept. Horticultural Sciences, Texas A&M University

[2]Molecular and Environmental Plant Science Program, Texas A&M University

[3]**Corresponding author e-mail**: koiwa@tamu.edu

**Title:** Coding sequence of firefly luciferase reporter gene affects specific hyper expression in *cpl1* mutant.

**Material distribution footnote**

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Hisashi Koiwa (koiwa@tamu.edu).

**Abstract**

Monitoring plant gene expression using promoters of interest fused to reporter-genes became a standard practice in plant molecular biology research. Firefly luciferase (LUC) enables non-destructive monitoring gene expression and widely used for not only quantitative analysis but also screening of mutants altered for gene expression levels. Surprisingly, such screenings frequently identified alleles of *CPL1* (Carboxyl-terminal Phosphatase-Like 1) regardless of promoters or pathways studied. In addition, expression of corresponding endogenous genes often show minimal difference between wild type and *cpl1.* Here we show that coding sequence of LUC reporter gene itself is responsible for high luciferase activity in *cpl1* mutant using a classical *RD29a-LUC* reporter gene as a model system. Deletion of *LUC* 3'-UTR of insect origin included in classical *LUC* reporter cassette did not change hyperactivation of *LUC* in *cpl1.* However, a reporter gene based on a codon-modified *LUC* derivative (*LUC2*) produced similar expression levels both in wild type in *cpl1*. These results indicate that coding region of classical *LUC* is responsible for *cpl1*-specific overexpression of reporter gene, which is uncoupled with endogenous counterpart.

Use of a reporter genes to monitor gene expression has become a standard practice in characterization of genes of interest. Various reporter genes with different benefits include bacterial β-glucuronidase (GUS), fluorescent proteins from jellyfish/coral organisms (XFPs), and luciferase from bacteria (LUX), firefly (LUC) and sea pansy (*Renilla* LUC). Of these, LUC from firefly has been widely used to conduct non-invasive monitoring of plant gene expression. Due to the short half-life of *LUC* mRNA (45 min) and protein (15.5 min with luciferin, 155 min without), LUC provides a low-background and highly sensitive way to monitor the plant gene expression in real time and study the inducible gene expression in response to environmental stimuli (Leeuwen et al., 2000). Several groups including ours took advantage of LUC system to identify genetic mutations in *Arabidopsis thaliana,* where *LUC* reporter lines were subjected to mutagenesis and genetic mutations were identified based on the alteration of LUC expression profile (Ishitani et al., 1997). Typically, inducible promoters were fused to LUC, and plants that over- or under-express LUC upon stimulation were identified as potential mutants for regulators of gene expression.

Over the past years it became obvious that the *LUC* reporter-based forward genetic approaches repeatedly identify mutations in *CPL1/FRY2* as well as *HOS5/RCF3* genes regardless of the biological processes studied (Koiwa et al., 2002; Xiong et al., 2002; Matsuda et al., 2009; Manavella et al., 2012; Jiang et al., 2013; Guan et al., 2014). These include salt/osmotic stress, low-temperature-stress, jasmonate signaling/wounding, miRNA expression. Interestingly, in these studies, *cpl1/fry2* and *hos5/rcf3* mutations typically yields 10-100 fold enhancement in LUC mRNA and/or activities, however, the expression of corresponding endogenous transcripts were enhanced only ~2 fold. Because endogenous stress-responsive transcripts are often induced >10 fold, the impact caused by the mutations are relatively subtle. The typical explanation for this inconsistency was the likely presence of additional negative regulatory elements in the endogenous promoter, which was not included in the promoter fragments used in the reporter constructs.

*CPL1/FRY2* encodes a protein with phosphatase activity that resembles eukaryotic TFII-F-interacting carboxyl terminal phosphatase (FCP1), and can dephosphorylate RNA polymerase II in vitro (Koiwa et al., 2002; Xiong et al., 2002; Koiwa et al., 2004). As oppose to the expected function of CPL1 in regulation of RNA polymerase II transcription, however, in vivo data are indicative that CPL1 functions in post-transcriptional metabolism of mRNA (Manavella et al., 2012; Jiang et al., 2013). The first reported evidence that CPL1 affects state of LUC reporter mRNA was that a *cpl1* mutant showed higher capping efficiency and altered polyadenylation site of *LUC* mRNA. Another study demonstrated that CPL1/FRY2 regulate RNA decay pathway of

abnormal transcripts, such as transcripts that retains one or more introns (Cui et al., 2016). These findings prompted me to re-investigate whether stress-inducible promoter/signaling or reporter gene structure are determinant of the enhanced expression of stress-inducible reporter gene in *cpl1* mutant.

In addition to *RD29a-LUC* used in previous studies, we prepared *RD29a-LUC⊿3'* and *RD29a-LUC2*. *RD29a-LUC⊿3'* was prepared by removing 3'-UTR sequence of original firefly *LUC* mRNA from *RD29a-LUC*. The 3'-UTR of *LUC* mRNA contains two cryptic polyadenylation sites, usage of which were affected in *cpl1* mutants (Jiang et al., 2013). On the other hand, *RD29a-LUC2* used codon-modified LUC (LUC2, Promega) coding sequence without firefly *LUC* 3'-UTR. These reporter cassettes were transformed into *Arabidopsis thaliana* Col-0 *cpl1-6* mutant line, and then wild-type lines were prepared by backcrossing T1 plants showing 3:1 segregation ratio to the wild-type Col-0 plants. Homozygous T3 (*cpl1-6*) and F3 (wild type) lines were used to directly compare the reporter gene expression levels between wild type and the *cpl1* mutant.

As shown in Figure 1, the *cpl1-6* mutant host strongly enhanced *RD29a-LUC* as reported previously. Removal of *LUC* 3'-UTR did not alter hyper-induction of *LUC* in *cpl1-6,* suggesting *LUC* hyperexpression in *cpl1-6* was not due to cryptic polyadenylation sites in *LUC* 3'-UTR. Notably, overall expression of *RD29a-LUC⊿3'* was enhanced compared to *RD29a-LUC.* Expression of codon-modified *RD29a-LUC2* substantially altered reporter gene expression profile. Compared to *RD29a-LUC, RD29a-LUC2* activity was similarly increased in both wild type and *cpl1-6*. The induction pattern was close to that of endogenous *RD29a* mRNA level measured by RT-qPCR (Figure 1C), albeit *RD29a-LUC2* response was slower than response of *RD29a* mRNA. This result suggested that hyperexpression of *RD29a-LUC* in *cpl1* mutant are neither due to the enhancement of *RD29a* promoter activity, upstream osmotic-stress signaling events nor alternative polyadenylation in 3'-UTR. Instead, the data showed that *LUC* coding sequence strongly influences *cpl1*-dependent hyperexpression. This explains why mutations in *CPL1/FRY2* as well as in *HOS5/RCF3* encoding a binding partner of CPL1 are repeatedly identified in *LUC* reporter gene-based forward genetic screening, and why there are discrepancies between *LUC* expression level and endogenous counterpart in *cpl1/fry2*. Considering the fact that LUC mRNA has a short half-life in the plant cell (Gallie et al., 1991; Leeuwen et al., 2000), the likely mechanism is *LUC* mRNA but not endogenous counterpart being a target of CPL1-dependent RNA decay pathway (Cui et al., 2016). This observation suggests that we need to be cautious about impact of *cpl1* on synthetic phenotype produced by

artificial transgenes, because *cpl1* may produce phenotype by specifically affecting the level of transgene expression.

We found the expression of codon-modified *LUC2* reporter gene retains stress-regulation in wild type and is much less affected by *cpl1-6* mutation, therefore, *LUC2* provides a promising alternative for reporter-based forward-genetic screening. It should be noted, however, that *RD29a-LUC2* produced higher background, and there was a delay in timing of induction of *RD29a-LUC2* activity above background level after the onset of cold treatment relative to *RD29a-LUC*, which would require careful optimization of the assay system.

## Materials and methods

### Preparation and analysis of reporter lines

*RD29a-LUC⊿3'* was prepared by inserting a PCR fragment containing 3'-end of *LUC* coding sequence (primer pair 1358 TAGTTGACCGCTTGAAGTCT/
1359 GGAAAAGAGCTCTTACAATTTGGACTTTCCGC) into PacI-SacI sites of original *RD29a-LUC* plasmid (Ishitani et al., 1997). *RD29a-LUC2* was prepared by inserting a PCR fragment encoding *LUC2* (primer pair: 1471
CTATTTACAATTACAGTCGACATGGAAGATGCCAAAAACATT/
1472 cagatcccccgggggtaccgagctcTTACACGGCGATCTTGCCGC) into SalI-SacI site of *RD29a-LUC* plasmid. Resulting plasmids were introduced into *Agrobacterium tumefaciens* ABI, and used for transformation of *cpl1-6* (Jeong et al., 2013) by flower transformation. Kanamycin resistant plants were selected as described (Nagashima and Koiwa, 2017). Transformants whose luciferase expression phenotype were segregating with 3:1 ratio were backcrossed to Col-0. Homozygous T3 (*cpl1-6* background) and F3 (Col-0 background) lines were established for each transgene, which were used for the expression analysis. Luciferase imaging analysis and RT-qPCR analysis was performed as described (Jeong et al., 2013).

## Acknowledgement

This study was partially supported by USDA-NIFA (2015-67013-22816).

## References cited

Cui, P., Chen, T., Qin, T., Ding, F., Wang, Z., Chen, H., and Xiong, L. (2016). The RNA Polymerase II C-Terminal Domain Phosphatase-Like Protein FIERY2/CPL1 Interacts with eIF4AIII and Is Essential for Nonsense-Mediated mRNA Decay in Arabidopsis. Plant Cell 28, 770-785.

Gallie, D.R., Feder, J.N., Schimke, R.T., and Walbot, V. (1991). Post-transcriptional regulation in higher eukaryotes: the role of the reporter gene in controlling expression. Mol Gen Genet 228, 258-264.

Guan, Q., Yue, X., Zeng, H., and Zhu, J. (2014). The protein phosphatase RCF2 and its interacting partner NAC019 are critical for heat stress-responsive gene regulation and thermotolerance in Arabidopsis. Plant Cell 26, 438-453.

Ishitani, M., Xiong, L., Stevenson, B., and Zhu, J.-K. (1997). Genetic analysis of osmotic and cold stress signal transduction in Arabidopsis: interactions and convergence of abscisic acid-dependent and abscisic acid-independent pathways. Plant Cell 9, 1935-1949.

Jeong, I.S., Fukudome, A., Aksoy, E., Bang, W.Y., Kim, S., Guan, Q., Bahk, J.D., May, K.A., Russel, W.K., Zhu, J., and Koiwa, H. (2013). Regulation of abiotic stress signalling by Arabidopsis C-terminal domain phosphatase-like 1 requires interaction with a K-homology domain-containing protein. PLOS One 8, e80509.

Jiang, J., Wang, B., Shen, Y., Wang, H., Feng, Q., and Shi, H. (2013). The arabidopsis RNA binding protein with K homology motifs, SHINY1, interacts with the C-terminal domain phosphatase-like 1 (CPL1) to repress stress-inducible gene expression. PLoS Genet 9, e1003625.

Koiwa, H., Hausmann, S., Bang, W.Y., Ueda, A., Kondo, N., Hiraguri, A., Fukuhara, T., Bahk, J.D., Yun, D.J., Bressan, R.A., Hasegawa, P.M., and Shuman, S. (2004). Arabidopsis C-terminal domain phosphatase-like 1 and 2 are essential Ser-5-specific C-terminal domain phosphatases. Proc. Natl. Acad. Sci. USA 101, 14539-14544.

Koiwa, H., Barb, A.W., Xiong, L., Li, F., McCully, M.G., Lee, B.-h., Sokolchik, I., Zhu, J., Gong, Z., Reddy, M., Sharkhuu, A., Manabe, Y., Yokoi, S., Zhu, J.-K., Bressan, R.A., and Hasegawa, P.M. (2002). C-terminal domain phosphatase-like family members (AtCPLs) differentially regulate *Arabidopsis thaliana* abiotic stress signaling, growth, and development. Proc Natl Acad Sci USA 99, 10893-10898.

Leeuwen, W., Hagendoorn, M., Ruttink , T., Poecke, R., Plas, L., and Krol , A. (2000). The use of the luciferase reporter system forin planta gene expression studies. Plant Mol. Biol. Rep. 18, 143-144.

Manavella, P.A., Hagmann, J., Ott, F., Laubinger, S., Franz, M., Macek, B., and Weigel, D. (2012). Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. Cell 151, 859-870.

Matsuda, O., Sakamoto, H., Nakao, Y., Oda, K., and Iba, K. (2009). CTD phosphatases in the attenuation of wound-induced transcription of jasmonic acid biosynthetic genes in

Arabidopsis. Plant J. 57, 96-108.

Nagashima, Y., and Koiwa, H. (2017). High throughput selection of antibiotic-resistant transgenic Arabidopsis plants. Anal Biochem 525, 44-45.

Xiong, L., Lee, H., Ishitani, M., Tanaka, Y., Stevenson, B., Koiwa, H., Bressan, R.A., Hasegawa, P.M., and Zhu, J.-K. (2002). Repression of stress-responsive genes by *FIERY2*, a novel transcriptional regulator in Arabidopsis. Proc Natl Acad Sci U S A 99, 10899-10904.

**Figure legend**

**Figure 1. Expression of *RD29a*-luciferase reporter genes in Col-0 and *cpl1-6.* A)** Bioluminescence images of *RD29a-LUC* WT, D3', and *LUC2* in 25°C and after 2 or 3-day 0°C treatment. Plants were grown on medium containing 1/4x MS salts, 1% sucrose, and 0.7% agar for 7 days. For cold treatment, plates were incubated at 0°C in dark for indicated duration. Exposure time for CCD camera was 10 min (5 min for short exposure panels). **B**) Bioluminescence image quantification results. Each datapoint was calculated using 50-70 plants from 3-4 independent plates. Error bars indicate standard error of the mean. ND, not determined. **C**) RT-qPCR analysis of *RD29a* transcripts in Col-0 and *cpl1-6.* Total RNA was extracted from plants prepared as described in A. qPCR data were normalized using *GAPDH*. Error bars indicate standard error of the mean from biological triplicates.

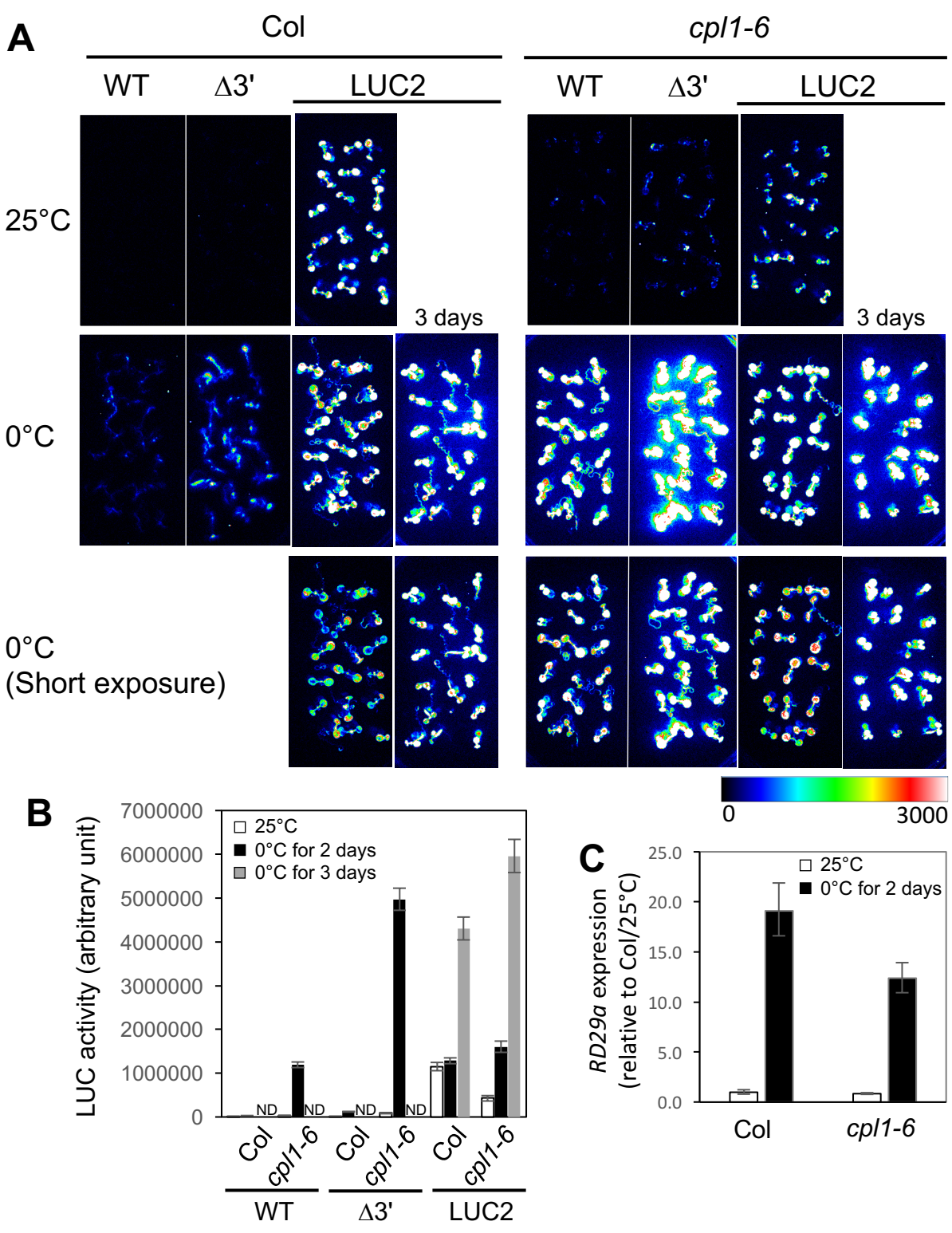


**Figure 1. Expression of *RD29a*-luciferase reporter genes in Col-0 and *cpl1-6.*** **A)** Bioluminescence images of *RD29a-LUC* WT, D3', and *LUC2* in 25°C and after 2 or 3-day 0°C treatment. Plants were grown on medium containing 1/4x MS salts, 1% sucrose, and 0.7% agar for 7 days. For cold treatment, plates were incubated at 0°C in dark for indicated duration. Exposure time for CCD camera was 10 min (5 min for short exposure panels). **B**) Bioluminescence image quantification results. Each datapoint was calculated using 50-70 plants from 3-4 independent plates. Error bars indicate standard error of the mean. ND, not determined. **C**) RT-qPCR analysis of *RD29a* transcripts in Col-0 and *cpl1-6.* Total RNA was extracted from plants prepared as described in A. qPCR data were normalized using GAPDH. Error bars indicate standard error of the mean from biological triplicates.