

# An E3 ligase complex regulates SET-domain polycomb group protein activity in *Arabidopsis thaliana*

Cheol Woong Jeong<sup>a</sup>, Hyungmin Roh<sup>a</sup>, Tuong Vi Dang<sup>a</sup>, Yang Do Choi<sup>b</sup>, Robert L. Fischer<sup>c,1</sup>, Jong Seob Lee<sup>a,1</sup>, and Yeonhee Choi<sup>a,1</sup>

<sup>a</sup>School of Biological Sciences, Seoul National University, Seoul 151-747, Korea; <sup>b</sup>Department of Agricultural Biotechnology and Center for Agricultural Biomaterials, Seoul National University, Seoul 151-921, Korea; and <sup>c</sup>Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720

Contributed by Robert L. Fischer, March 15, 2011 (sent for review February 23, 2011)

**Transcriptional repression via methylation of histone H3 lysine 27 (H3K27) by the polycomb repressive complex 2 (PRC2) is conserved in higher eukaryotes. The *Arabidopsis* PRC2 controls homeotic gene expression, flowering time, and gene imprinting. Although downstream target genes and the regulatory mechanism of PRC2 are well understood, much less is known about the significance of posttranslational regulation of PRC2 protein activity. Here, we show the posttranslational regulation of CURLY LEAF (CLF) SET-domain polycomb group (PcG) protein by the F-box protein, UPWARD CURLY LEAF1 (UCL1). Overexpression of UCL1 generates mutant phenotypes similar to those observed in plants with a loss-of-function mutation in the CLF gene. Leaf curling and early flowering phenotypes of UCL1 overexpression mutants, like *clf* mutants, are rescued by mutations in the *AGAMOUS* and *FLOWERING LOCUS T* genes, which is consistent with UCL1 and CLF functioning in the same genetic pathway. Overexpression of UCL1 reduces the level of CLF protein and alters expression and H3K27 methylation of CLF-target genes in transgenic plants, suggesting that UCL1 negatively regulates CLF. Interaction of UCL1 with CLF was detected in plant nuclei and in the yeast two-hybrid system. The UCL1 F-box binds *in vivo* to components of the E3 ligase complex, which ubiquitylate proteins that are subsequently degraded via the ubiquitin-26S proteasome pathway. Taken together, these results demonstrate the posttranslational regulation of the CLF SET-domain PcG activity by the UCL1 F-box protein in the E3 ligase complex.**

histone methylation | proteasome | gene silencing | epigenetics | protein stability

**P**olycomb group (PcG) proteins epigenetically silence gene expression and play an important role in controlling eukaryote cell proliferation, stem cell identity, cancer, genomic imprinting, and X chromosome inactivation (1, 2). The term “polycomb” initially referred to mutations in *Drosophila* that caused improper specification of body segment identity. Three different PcG complexes, polycomb repressive complex 1 (PRC1), PRC2, and PcG-like PRC2, are present in animals and work together in a stepwise manner to repress expression of their target genes (3).

The PRC2 complex is highly conserved in *Drosophila*, mammals, and flowering plants (1, 4). The *Drosophila* PRC2 complex is composed of three PcG proteins, Extra Sex Combs (Esc), Suppressor of Zeste 12 [Su(z)12], and Enhancer of Zeste [E(z)], as well as additional core proteins, such as p55 (2). Esc and p55 are WD-40 proteins, Su(z)12 is a C2H2 zinc finger, and E(z) is a SET-domain protein that methylates histone H3 lysine 27 (H3K27), which silences target gene expression. In *Arabidopsis*, there is an Esc homolog, FERTILIZATION INDEPENDENT ENDOSPERM (FIE); p55 homologs, ARABIDOPSIS MULTICOPY SUPPRESSOR OF IRA1 to 5 (MS1 to MS5); zinc finger-containing Su(Z) homologs, EMBRYONIC FLOWER 2 (EMF2), VERNALIZATION2 (VRN2) and FERTILIZATION INDEPENDENT SEED2 (FIS2); and SET-domain E(z) homologs, CURLY LEAF (CLF), SWINGER (SWN), and MEDEA (MEA) (4).

There are three distinct *Arabidopsis* PRC2 complexes, which are named after their respective zinc finger-containing Su(Z) homolog constituents: EMF2, VRN2, and FIS2. The EMF2 PcG complex, composed of FIE, MS1, CLF/SWN, and EMF2, prevents early flowering and regulates vegetative growth (4). During the vegetative state of development, the EMF2 PcG complex interacts with CLF to repress the expression of a key floral promoting gene, *FLOWERING LOCUS T* (*FT*), as well as the floral organ identity genes *AGAMOUS* (*AG*) and *AGAMOUS-LIKE 19* (*AGL19*) (5–7). Mutations in the CLF SET-domain PcG gene cause pleiotropic phenotypes, including hyponastic (curly) leaves, homeotic transformation of floral organs, and early flowering (8). The FIS2 complex is composed of FIE, MS1, MEA, and FIS2, which are required for proper seed development. Mutations in the MEA SET-domain PcG gene cause seed abortion, the formation of nonviable embryos, and disruption of gene imprinting in the endosperm (4, 9).

The ubiquitin-26S proteasome system is conserved in eukaryotes (10). Three ubiquitin-ligase activities (E1, E2, and E3) sequentially transfer ubiquitin to target proteins that are degraded in the 26S proteasome complex. The system degrades abnormal proteins that are produced by biosynthetic errors or have folded into nonnative conformations. It also controls the level of specific regulatory proteins that control cell cycle progression, signal transduction, gene transcription, programmed cell death, and development in animals as well as hormone synthesis and signaling, plant-pathogen interactions, self-incompatibility, circadian rhythms, morphogenesis, and histone modifications in plants (10–12). A well-characterized E3 ligase, SCF, is composed of multiple subunits: a CULLIN1 (CUL1) scaffold protein; a catalytic RING domain protein, RING-BOX1 (RBX1); an adaptor protein called S-PHASE KINASE-ASSOCIATED PROTEIN1 (SKP1) in animals and ARABIDOPSIS SKP1 HOMOLOG1/2 (ASK1/2) in *Arabidopsis*; and an F-box protein, which confers substrate specificity in the transfer of ubiquitin from the E2 ubiquitin-conjugating enzyme to the target protein.

Histone methylation by PcG protein complexes maintains gene silencing (2). Alterations in PcG complex composition and association to chromatin may be beneficial as organisms undergo developmental transitions and respond to their environment, however (4, 13). There is increasing evidence for regulation of PRC2 activity by posttranslational modifications of PcG proteins by E3 ubiquitin ligases (14, 15). Here, we show that CLF SET-domain activity is regulated at the posttranslational level by UPWARD CURLY LEAF1 (UCL1), the F-box component of an E3 ligase that binds CLF and targets it for degradation via the ubiquitin-26S proteasome pathway.

Author contributions: C.W.J., Y.D.C., R.L.F., J.S.L., and Y.C. designed research; C.W.J., H.R., and T.V.D. performed research; C.W.J., R.L.F., J.S.L., and Y.C. analyzed data; and C.W.J., R.L.F., J.S.L., and Y.C. wrote the paper.

The authors declare no conflict of interest.

<sup>1</sup>To whom correspondence should be addressed. E-mail: rfischer@berkeley.edu, jongselee@plaza.snu.ac.kr, or yhc@snu.ac.kr.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1104232108/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1104232108/-DCSupplemental).

## Results

### Plants That Overexpress *UCL1*, an F-Box Protein Gene, Have Curly Leaves.

We mutagenized *Arabidopsis* plants using a T-DNA activation vector that activates flanking gene expression (16) and identified a mutation, *upward curly leaf1-dominant* (*ucl1-D*), whose semidominant curly leaf phenotype (Fig. 1*A* and Fig. S1*A*) was also observed in plants with a recessive loss-of-function mutation in the *CLF* gene (8). Plasmid rescue methods (16) revealed that the T-DNA had inserted in the intergenic region between *At1g65740* and *At1g65750* (Fig. 1*B*). Real-time quantitative RT-PCR (qRT-PCR) showed that expression of *At1g65740*, *At1g65760*, and *At1g65770* was elevated in *ucl1-D* homozygous plants (Fig. S1*B*). Among these three genes, we observed the curly leaf phenotype only when overexpressing *At1g65740* in transgenic plants using the *Cauliflower Mosaic Virus* (*CaMV*) promoter (Figs. S1*C* and *D* and S2*B*, *D*, *F*, and *G*). We designated *At1g65740*, which encodes an F-box protein, as *UCL1*.

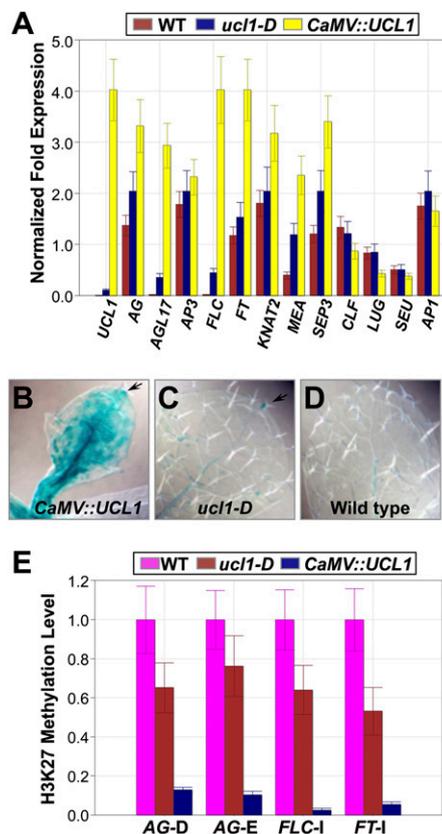
***UCL1* Overexpression and *clf* Loss of Function Produce Similar Phenotypes.** We compared *clf* and *CaMV::UCL1* phenotypes throughout the *Arabidopsis* life cycle. Homozygous *clf* mutant plants display curly leaves, floral homeotic changes attributable to the ectopic expression of *AG*, reduced internode and inflorescence height (8), and early flowering under reduced day-length conditions attributable to the derepression of *FT* (7). Likewise, homozygous *CaMV::UCL1* plants display curly rosette leaves (compare Fig. S2*C* and *D*); reduced internode and inflorescence height (Fig. S3*A*); defects in flower morphology (compare Figs. S3*B* and *C* and *D–I*); and homeotic changes of floral organs, such as petals with stamen-like features (Fig. S3*F* and *I*) and ovule-bearing sepals with stigmatic papilla (Fig. S3*G–J*). *CaMV::UCL1* plants flowered early, especially when grown under short-day conditions (Fig. S3*J*). Thus, *clf* and *CaMV::UCL1* plants display similar pleiotropic phenotypes.

We generated *CaMV::UCL1*; *ft-1/ft-1* plants and *CaMV::UCL1*; *ag-1/ag-1* plants and found that the early flowering and curly leaf phenotypes in *CaMV::UCL1* plants were suppressed by the *ft-1* and *ag-1* mutations, respectively (Fig. S3*K* and *L*). This suggests that *CaMV::UCL1*, like *clf* mutants, causes early flowering and leaf curling by derepression of *FT* and *AG*, respectively. These results suggest that *CaMV::UCL1* and *clf* cause pleiotropic phenotypes by the same genetic pathways.

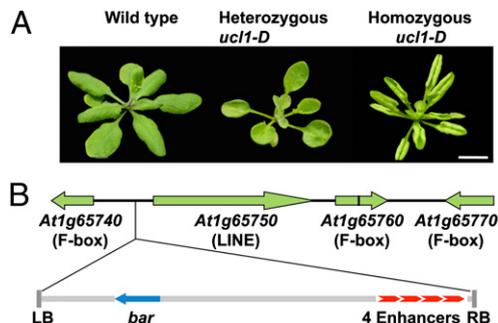
***UCL1* Overexpression Activates Expression of CLF-Target Genes and Alters Their H3K27 Methylation Profiles.** We analyzed CLF-target gene expression by isolating RNA from *CaMV::UCL1* and WT seedlings, which were independently hybridized to Affymetrix microarrays (Table S1). A list of genes showing significant changes in expression includes *AG* and *AGL* genes as well as flowering time regulators *FT* and *FLOWERING LOCUS C* (*FLC*). We verified by real-time qRT-PCR analysis that well-

known CLF-target genes *AG*, *APETALATA3* (*AP3*), *AGL17*, *FT*, *FLC*, *KNOTTED-LIKE FROM ARABIDOPSIS THALIANA 2* (*KNAT2*), *MEA*, and *SEPALLATA3* (*SEP3*) were overexpressed in rosette leaves of *ucl1-D* and *CaMV::UCL1* plants (Fig. 2*A*). By contrast, expression of non-CLF-target genes *SEUSS* (*SEU*), *LEUNIG* (*LEU*), *APETALATA1* (*API*), and *CLF* was not significantly changed by *UCL1* overexpression (Fig. 2*A*). Likewise, overexpression of the *At1g65770* gene, which is adjacent to *UCL1*, did not activate CLF-target genes (Fig. S4*A*). These results suggest that CLF-target genes are specifically up-regulated in plants that overexpress *UCL1*.

We introduced the *pAG-I::GUS* gene, consisting of the *AG* promoter with the *AG* first and second introns fused to the  $\beta$ -glucuronidase (*GUS*) reporter gene (17), into the *CaMV::UCL1* or *ucl1-D* genetic background. In *pAG-I::GUS*; *CaMV::UCL1* plants, we detected ectopic *GUS* expression in young rosette leaves (Fig. 2*B* and Fig. S5*A*), floral stems, and petals (Fig. S5*D*) compared with the control *pAG-I::GUS* plants (Fig. 2*D* and Fig. S5*C* and *F*). In *pAG-I::GUS*; *ucl1-D* plants, we detected a slight increase in ectopic *GUS* expression in young leaves (Fig. 2*C* and Fig. S5*B*) and in the floral stem (Fig. S5*E*, arrow). These results are consistent with the real-time qRT-PCR data shown in Fig. 2*A*. By contrast, we could not detect altered *GUS* expression in control *pAG-I::GUS*; *CaMV::At1g65770* plants (Fig. S4*B–D*).



**Fig. 2.** Overexpression of *UCL1* activates CLF-target genes and alters histone methylation patterns. (A) Real-time qRT-PCR analysis in WT, *ucl1-D*, and *CaMV::UCL1* plants. Values are plotted relative to the *TUBULIN- $\beta$ -CHAIN2* (*TUB2*) reference gene and represent the average of duplicate measurements  $\pm$  SD. *AG-I::GUS* activity in *CaMV::UCL1* (B, arrow), in *ucl1-D* (C, arrow), and in WT (D) rosette leaf. (E) Levels of H3K27me3 at *AG*, *FLC*, and *FT* chromatin in *ucl1-D* and *CaMV::UCL1* seedlings relative to WT. DNA concentration after precipitation with H3K27me3 antibody was quantified by real-time qRT-PCR and normalized to an internal reference (*TUB2*). Values in E are plotted relative to WT for each DNA, which was set at 1.0, and represent the average of triplicate measurements  $\pm$  SD. *TUB*, tubulin  $\beta$ -chain 2.



**Fig. 1.** Insertion of an activation T-DNA next to an F-box gene. (A) WT, heterozygous, and homozygous *ucl1-D* mutant plants. (B) Genomic region flanking the T-DNA insertion site in *ucl1-D* plants. *bar*, Basta-resistance gene; 4 Enhancers, *CaMV* enhancer tetrad; LB, T-DNA left border; RB, T-DNA right border.

These results show that *UCL1* overexpression, similar to loss-of-function *clf* mutations (18), up-regulates *AG* transcription.

The H3K27me3 silencing mark is reduced at CLF-target genes in *clf* mutant plants (7, 19). We found that H3K27me3 levels at the *AG* intron D and E regions (Fig. S5G) were significantly reduced in *UCL1* overexpression plants compared with WT (Fig. 2E and Fig. S5H). Reduced H3K27me3 levels were also detected at specific intron sites (Fig. S5G) in the *FLC* and *FT* loci (Fig. 2E). Thus, *clf* mutations (7, 19) and *UCL1* overexpression reduce H3K27me3 at CLF-target genes.

#### UCL1 Binds CLF in the Yeast Two-Hybrid System and in Plant Nuclei.

We used the yeast two-hybrid system to determine if the UCL1 F-box protein binds to CLF. We failed to detect the interaction when we used full-length UCL1 or CLF (Fig. 3A and B and Discussion). CLF amino acids 1–300 interacted with UCL1 amino acids 201–371 and activated reporter  $\beta$ -galactosidase (*lacZ*) gene expression (Fig. 3A and B). We also checked the binding of UCL1 to MEA, a SET-domain PcG protein that is closely related to CLF. MEA amino acids 1–280, corresponding to CLF amino acids 1–300, did not interact with UCL1 amino acids 201–371 (Fig. 3A and C). Thus, domains within the CLF and UCL1 proteins specifically interact in yeast.

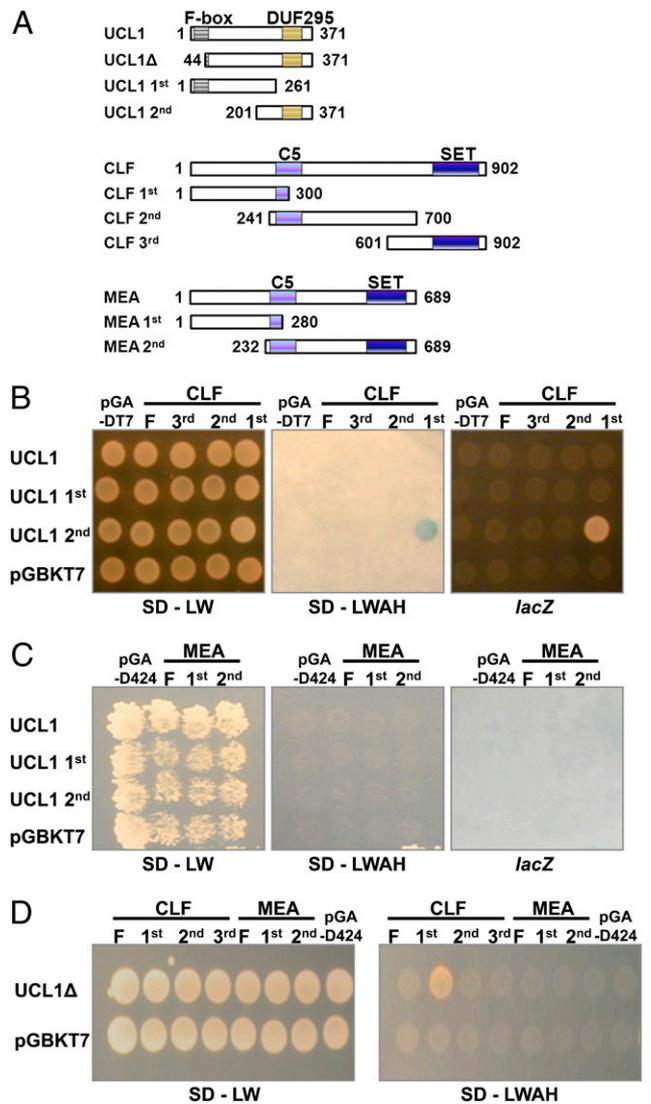
To determine UCL1 localization, we generated *CaMV::UCL1::GFP* transgenic plants. *GFP* activity was detected in root cell nuclei (Fig. 4A), indicating that UCL1 is a nuclear protein like CLF (20). We investigated the *in vivo* interaction between full-length UCL1 and CLF using bimolecular fluorescence complementation (BiFC) (21). We detected a strong reconstituted YFP signal in the nucleus when *CaMV::nEYFP::CLF* and *CaMV::cEYFP::UCL1* were cotransformed into *Arabidopsis* protoplasts (Fig. 4B, Lower). These results demonstrate that UCL1 interacts with CLF in plant nuclei.

**UCL1 Binds Proteins in the E3 Ligase Complex.** We checked the interaction of F-box UCL1 with other polypeptides in an SCF E3 ubiquitin ligase complex with the yeast two-hybrid system. UCL1 interacted with ASK1, an *Arabidopsis* SKP1 homolog, and also with ASK2, albeit with weaker binding activity (Fig. 4C). UCL1 $\Delta$  with most of its F-box deleted (amino acids 44–371) (Fig. 3A) failed to interact with ASK1 (Fig. 4C), indicating that the F-box motif in UCL1 is essential for ASK1 binding activity. As expected, UCL1 did not directly interact with CUL1 (Fig. 4C), which needs ASK1 to mediate UCL1 F-box protein binding to the CUL1 scaffold protein to form an SCF complex (22, 23).

We performed coimmunoprecipitation experiments to determine if UCL1 is part of the SCF complex *in vivo*. Using *Arabidopsis* protoplasts, UCL1 and CUL1 with hemagglutinin (UCL1-HA and CUL1-HA) were coimmunoprecipitated when GFP-tagged ASK1 (ASK1-GFP) was coexpressed and reacted with anti-GFP antibodies (Fig. 4D). These results show that UCL1 is a component of the SCF complex in plant nuclei.

**UCL1 Overexpression Reduces CLF Protein Level.** F-box proteins bind target proteins that are ubiquitinated and subsequently degraded in the proteasome (10, 24). In protein blot experiments with GFP antibody, we detected GFP:CLF fusion protein in control homozygous *CaMV::GFP::CLF* seedlings as well as in F1 seedlings generated by crossing *CaMV::GFP::CLF* with WT plants (Fig. 4E). By contrast, the level of GFP:CLF fusion protein was significantly reduced in F1 seedlings generated by crossing *CaMV::GFP::CLF* with *CaMV::UCL1* plants (Fig. 4E). To confirm that *GFP::CLF* RNA levels are not altered in F1 seedlings, we measured RNA expression in these plants by qRT-PCR. *CLF* and *GFP* RNA is expressed at a similar level in F1 seedlings generated by crossing *CaMV::GFP::CLF* with WT plants and with *CaMV::UCL1* plants (Fig. S6A and B).

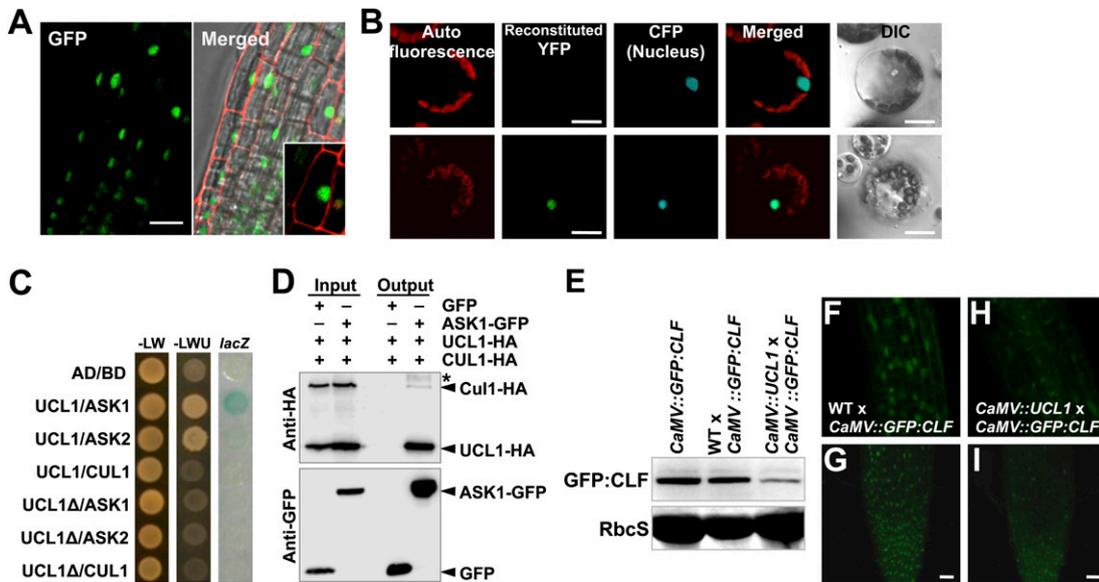
We checked *UCL1* and *AG* RNA expression as a positive control, as well as *SWN* and *MSI1* expression as a negative control (Fig. S6C–F).



**Fig. 3.** UCL1 specifically binds to CLF but not to MEA in the yeast two-hybrid assay. (A) Schematic diagram of UCL1, CLF, and MEA. The amino acid numbers of proteins used in the yeast two-hybrid analysis are shown. (B) Interaction between UCL1 and CLF in the yeast two-hybrid assay. (C) Interaction between UCL1 and MEA in the yeast two-hybrid assay. (D) Interaction test between UCL1 $\Delta$  and CLF, as well as between UCL1 $\Delta$  and MEA, in the yeast two-hybrid assay. F, clones encoding full-length proteins; *lacZ*,  $\beta$ -galactosidase enzyme activity; SD-LW, synthetic dextrose minimal medium without leucine and tryptophan; SD-LWAH, synthetic dextrose minimal medium without leucine, tryptophan, adenine, and histidine.

We also measured *in vivo* GFP fluorescence in root tissues. Control F1 plants generated by crossing *CaMV::GFP::CLF* with WT displayed strong GFP fluorescence in the root elongation zone (Fig. 4F) and root tip (Fig. 4G and Fig. S4H). By contrast, we detected a significant reduction of GFP fluorescence in the root elongation zone (Fig. 4H) and root tip (Fig. 4I and Fig. S4I) in F1 plants generated by crossing *CaMV::GFP::CLF* with *CaMV::UCL1*. In control experiments, ectopic expression of *CaMV::At1g65770* did not diminish GFP fluorescence (Fig. S4J). Taken together, these results indicate that overexpression of *UCL1* specifically decreases the level of GFP:CLF protein *in planta*.

**UCL1 Function During Endosperm and Stamen Development.** qRT-PCR experiments revealed significant *UCL1* expression in flowers, particularly in siliques with developing seeds and in male reproductive structures, stamens (Fig. S7A). We measured *UCL1*



**Fig. 4.** UCL1 forms an SCF complex, binds CLF, and reduces CLF protein level. (A) Subcellular localization of UCL1:GFP fusion protein in root cells. Subcellular localization of UCL1:GFP fusion protein in *CaMV::UCL1:GFP* transgenic plants in the root tip of light-grown seedlings is shown by confocal laser scanning microscopy using a GFP filter set. A merged image using DIC, PI, and GFP filter sets is shown. (Inset) Close-up view using GFP and PI filter sets. (B) BiFC assay for protein-protein interaction of UCL1 and CLF in *Arabidopsis* cells. (Upper) No YFP signal in the nucleus when *CaMV::nEYFP:CLF* and *CaMV::cEYFP* (without UCL1) were cotransfected as a negative control. (Lower) Reconstituted YFP signal in the nucleus when *CaMV::nEYFP:CLF* and *CaMV::cEYFP:UCL1* were cotransfected. The known transcription factor *FES*, fused with *CFP*, was used as a nucleus marker after cotransformation. DIC, differential interference contrast microscopy. (C) Yeast two-hybrid assay for interaction of UCL1 with ASK1 and ASK2. UCL1 $\Delta$  has a deletion (amino acids 1–43) and does not contain most of the N-terminal F-box motif. –LW, synthetic dextrose minimal (SD) medium without leucine and tryptophan; –LWU, synthetic dextrose minimal (SD) medium without leucine, tryptophan, and uracil; *lacZ*,  $\beta$ -galactosidase enzyme activity. (D) Coimmunoprecipitation of UCL1, CUL1, and ASK1. Proteins from protoplasts transformed with constructs expressing GFP, UCL1-HA, CUL1-HA, and ASK1-GFP were precipitated (Output) or not precipitated (Input) with anti-GFP antibody, run on SDS/PAGE gels, blotted, and reacted with Anti-HA or Anti-GFP antibodies. \*Neddylated CUL1. (E) Protein blot with anti-GFP antibody in *CaMV::GFP:CLF* 15-d-old seedlings or from F1 seedlings generated by crossing *CaMV::GFP:CLF* with the indicated lines. For a loading control, ribulose-bisphosphate carboxylase (RbcS) protein was stained with Ponceau. GFP fluorescence in the root elongation (F and H) and differentiation zone (G and I) from F1 plants generated by crossing *CaMV::GFP:CLF* with WT (F and G) or with *CaMV::UCL1* (H and I). (Scale bars: A, 20  $\mu$ m; B, 10  $\mu$ m; G–I, 20  $\mu$ m.)

promoter activity by analyzing transgenic plants with *UCL1::GUS* (transcriptional fusion) or *UCL1::UCL:GUS* (translational fusion) genes. Cytoplasmic *GUS* activity in *UCL1::GUS* plants was detected in young stamens (Fig. S7 B and C) and in early endosperm development (Fig. S7 D–F). In *UCL1::UCL:GUS* plants, *GUS* activity was detected soon after fertilization in the primary endosperm nucleus (Fig. 5A) and in proliferating endosperm nuclei (Fig. 5 B–D). We did not detect *GUS* activity in the zygote or embryo (Fig. 5 A–D).

To understand the function of *UCL1* in WT plant development, we examined the phenotype of *ucl1* loss-of-function mutations caused by insertion of a T-DNA in the *UCL1* gene. No developmental abnormalities were detected in homozygous *ucl1-1* seed (Fig. S2E). It is possible that the extensive genetic redundancy within the F-box gene family (12) might have masked the effects of the *ucl1-1* mutation (Discussion).

To gain clues about *UCL1* function, we examined the effect of overexpressing its target, *CLF*, in a region of the plant where *UCL1* is expressed, the endosperm. We used the endosperm-specific *MEA* promoter to express *CLF* and found a high level of *CLF* expression in seeds of transgenic *MEA::CLF* plants compared with WT (Fig. 5E). Interestingly, we detected phenotypes associated with loss-of-function *mea* mutations (25, 26), including aborted seeds (Fig. 5F), excessive endosperm cell proliferation (Fig. 5I), proliferation of unfertilized central cells, and seed-like structure formation in the absence of fertilization (Fig. 5K), which were not detected in WT seeds (Fig. 5 G, H, and J). We also detected ectopic expression of the *AGL62* transcription factor that controls endosperm cellularization, which is also observed in *mea* mutant seed (27) (Fig. S7G). Control experiments showed that the *MEA::CLF* transgene did not affect *MEA* gene expression in the seed (Fig. S7H). These data show that ectopic expression of *CLF* in the endosperm is detrimental to

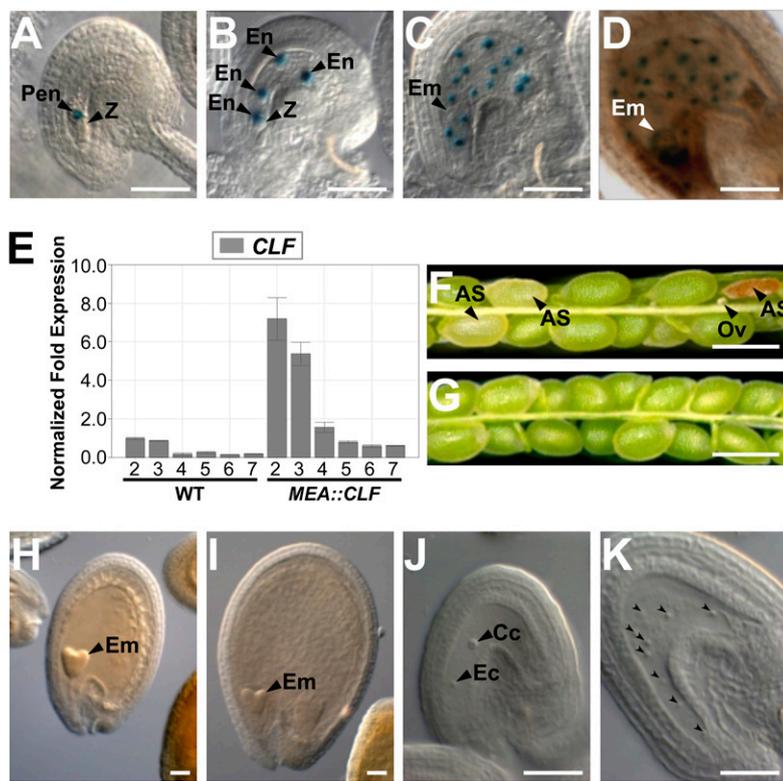
seed development. One possibility is that overexpression of *CLF* in the endosperm suppresses the activity of the closely related MEA PcG protein.

## Discussion

**Posttranslational Regulation of PcG Proteins.** In *Arabidopsis*, the ubiquitin-26S proteasome system's role as a regulatory system is comparable in breadth and depth to transcriptional gene regulation (12). The *Arabidopsis* and rice F-box gene families each have ~700 members (28). By comparison, there are 68 and 74 F-box genes in the human and mouse genomes, respectively (29). Some *Arabidopsis* F-box proteins have clear animal homologs, but most, like UCL1, are plant-specific.

Cullin serves as the backbone of the SCF complex and interacts with RBX1 via its C terminus and with the SKP1 (ASK1) adapter protein via its N terminus (23). SKP1 (ASK1) protein also binds to the F-box motif of F-box proteins to form the complete complex (24). In F-box proteins, the 60-aa F-box motif is at the N terminus and the domain that binds to the target protein is at the C terminus (30, 31). Likewise, the well-conserved N-terminal F-box motif of UCL1 interacts with ASK1 protein (Fig. 4), and its C-terminal domain interacts with its target protein, CLF (Fig. 3).

Full-length UCL1 or CLF protein failed to interact in the yeast two-hybrid system (Fig. 3). Similar results have been reported for the binding of full-length F-box proteins and their targets in yeast, however. For example, the *Arabidopsis* full-length F-box protein, UNUSUAL FLORAL ORGANS (UFO), and its full-length target, the LEAFY (LFY) transcription factor, bind at a low level in the yeast two-hybrid system (32). Deletion of the F-box domain resulted in more than an eightfold enhancement of binding, however, and it was concluded that the protein-protein interaction domain in the UFO F-box became



**Fig. 5.** Ectopic expression of *CLF* in the endosperm causes *mea*-like phenotypes. GUS activity in a mature *UCL1::UCL1::GUS* ovule right after fertilization (A) and in seeds 4 (B), 18 (C), and 36 (D) h after self-pollination. Em, embryo; En, endosperm nucleus; Pen, primary endosperm nucleus; Z, zygote. (E) Real-time qRT-PCR analysis of *CLF* RNA level in WT and *MEA::CLF* transgenic plants at the indicated days after self-pollination (DAP). Values are plotted relative to expression of *CLF* in WT at 2 DAP, which was set at 1.0, and represent the average of duplicate measurements  $\pm$  SD. (F) *MEA::CLF* transgenic plants containing abnormal aborted seeds (arrowhead) and small white unfertilized ovules. AS, aborted seed; Ov, ovule. (G) WT open silique. WT seed (H) or *MEA::CLF* seed (I). Unfertilized WT ovule (J) or *MEA::CLF* ovule (K) at 4 d after removal of anthers is shown. Cc, central cell nucleus; Ec, egg cell nucleus. Arrowheads point to nuclei of proliferating endosperm cells. (Scale bars: A–D and H–K, 50  $\mu$ m; F and G, 500  $\mu$ m.)

accessible when the F-box domain was deleted. It has also been suggested that the unoccupied F-box motif antagonizes the target-binding activity (33). Likewise, full-length UCL1 (amino acids 1–371) did not bind CLF amino acids 1–300, whereas UCL1 with most of its F-box deleted (amino acids 44–371) bound CLF amino acids 1–300 (Fig. 3D). Indeed, it is a common problem with the yeast two-hybrid system that full-length proteins have buried domains not available for interactions with other proteins (34). This likely explains why full-length UCL1 and its target, CLF, fail to interact in the yeast two-hybrid system. Nevertheless, we successfully detected full length-protein interaction using BiFC *in planta* (Fig. 4B).

Within the UCL1 C-terminal domain is a conserved DOMAIN OF UNKNOWN FUNCTION 295 (DUF295) that may be responsible for binding CLF. If so, there is considerable genetic redundancy because there are  $\sim$ 40 genes in *Arabidopsis* that encode proteins with an F-box domain and a DUF295 domain. There are three domains in CLF: a C-terminal SET domain involved in H3K27 methylation (19); a central C5 domain that interacts with the VEFS domain of EMF2, FIS2, and VRN2 PcG proteins (18); and an N-terminal domain of unknown function. It is the CLF N-terminal domain, whose sequence is specific to CLF, that binds to UCL1 in the yeast two-hybrid system (Fig. 3). Thus, the specificity of the interaction in *Arabidopsis* likely resides in the interaction between the redundant UCL1 DUF295 domain and the unique CLF N-terminal domain.

**Function of UCL1 During Arabidopsis Development.** Mutations in the *CLF* and *MEA* genes primarily cause defects during vegetative and reproductive development, respectively, suggesting that they function primarily in distinct developmental stages (4). Consistent with this idea, it was recently reported that *CLF* might be the ancestral PcG gene and that *MEA* may have acquired a unique reproductive role in seed development during evolution (35). It has previously been shown that *CLF* functions in the vegetative stage of development and is likely to be regulated transcriptionally (18). Surprisingly, the expression patterns of *CLF* and *MEA* were shown not to be mutually exclusive. For

example, *CLF* and *MEA* have overlapping expression patterns during plant reproduction (35) and both are expressed in the endosperm, where *UCL1* promoter activity is detected (Fig. 5). Ectopic expression of *CLF* from the *MEA* promoter results in *mea*-like mutant phenotypes (Fig. 5), suggesting that *CLF* competes with *MEA* in the formation of PRC2 complexes. This is consistent with the fact that both *CLF* and *MEA* are able to interact *in vitro* with the same Su(z) homologs, including FIS2, that are expressed in endosperm (18, 36). So, the question arises, how is *CLF* and *MEA* function separated during endosperm development when their respective RNA expression patterns and ability to interact with PcG proteins overlap? One possibility is that UCL1, expressed on fertilization in the primary endosperm cell and in actively dividing endosperm cells, specifically binds and ubiquitinylates CLF, resulting in its degradation, leaving *MEA* free to form PRC2 complexes in the endosperm that are essential for seed viability (4).

Understanding the validity of the above model at the molecular level requires a more thorough knowledge of the levels of *MEA* and *CLF* RNA and protein in the *Arabidopsis* endosperm than is currently available as well as the relative affinity of FIS2 binding to *MEA* and *CLF*. Generating multiple KO mutants for endosperm-specific F-box proteins that target CLF could be used to test this model at the genetic level, however. *In silico* and other molecular analyses (e.g., binding to CLF in the yeast two-hybrid system) could be used to find those F-box proteins. Plants with loss-of-function mutations in the appropriate F-box genes could be generated by genetic crosses to determine whether they show defects in seed development.

In summary, we have identified a previously undescribed post-translational mechanism for the control of PRC2 PcG activity and have demonstrated an interaction between the ubiquitin-26S proteasome pathway and PcG-mediated gene silencing during *Arabidopsis* development.

## Materials and Methods

Full details of methods used in this study are presented in *SI Materials and Methods*.

**Plant Materials and Growing Conditions.** The *ucl1-D* mutant was isolated from an activation-tagging mutant library as described previously (37). Plants were grown under previously described conditions (38).

**Histochemical GUS Staining and Microscopy.** Methods for GUS staining, fixing tissues, and microscopy are as described previously (38).

**Recombinant Plasmid Construction.** Methods for generating CaMV::UCL1, 4x Enh::UCL1, CaMV::UCL1:GFP, UCL1::GUS, UCL1::UCL1:GUS, and MEA::CLF constructs are described in *SI Materials and Methods*.

**Microarray Analysis.** Total RNA was isolated using an RNeasy Mini kit (Qiagen). Probe synthesis, detection, and scanning were performed according to protocols from Affymetrix, Inc.

**Real-Time qRT-PCR.** RNA levels were quantified by real-time qRT-PCR (iQ5; Bio-Rad), and data were analyzed with iCycle iQ system software (Bio-Rad). Primer sequences are listed in [Table S2](#).

**ChIP.** ChIP was performed as described previously (39) with minor modifications. Sonication was performed at 30% for 5 × 10 s, 0.6 s on/0.4 s off in each 1 s, with a Fisher Scientific Sonic Dismembrator 500.

**BiFC Assay.** The plant expression vectors *pSAT4-nEYFP-C1* (E3801) and *pSAT4-cEYFP-C1-B* (E3802) were used to generate constructs, which were introduced in pairs into *Arabidopsis* protoplasts by PEG transfection as described previously (40).

**Yeast Two-Hybrid Assay.** The pGBKT7 bait vector and pGADT7 or pGAD424 prey vector in the Matchmaker Two-Hybrid system (Clontech Laboratories) were used. Assay conditions were as described by the manufacturer.

**Coimmunoprecipitation Assay.** Coimmunoprecipitation was carried out as described previously (41). Protein was immunoprecipitated with polyclonal rabbit anti-GFP antibody (Invitrogen) and detected with monoclonal mouse anti-GFP (Clontech Laboratories) and monoclonal rat anti-HA (Roche) monoclonal antibodies.

**ACKNOWLEDGMENTS.** We thank J. Goodrich for *CaMV::GFP:CLF* and *pAG-l::GUS* reporter lines and I. Lee for the *FES::CFP* nucleus marker. We thank J. Sheen for the *HA*-tagging vector and H. G. Nam for the *GFP*-tagging vector. We thank Y. Jin for technical help and S. Lee for helpful discussions about protoplast analyses. We thank T.-F. Hsieh for critically reading this manuscript. This work was supported in part by Grant CG1121 from the Crop Functional Genomics Center (to J.S.L.) funded by the Korea Ministry of Science and Technology, by the BK21 Program (to J.S.L. and Y.C.) funded by the Korean government, and by National Institutes of Health Grant R01-GM069415 (to R.L.F.).

- Margueron R, Reinberg D (2011) The Polycomb complex PRC2 and its mark in life. *Nature* 469:343–349.
- Schwartz YB, Pirrotta V (2008) Polycomb complexes and epigenetic states. *Curr Opin Cell Biol* 20:266–273.
- Müller J, Verrijzer P (2009) Biochemical mechanisms of gene regulation by polycomb group protein complexes. *Curr Opin Genet Dev* 19:150–158.
- Hennig L, Derkacheva M (2009) Diversity of Polycomb group complexes in plants: Same rules, different players? *Trends Genet* 25:414–423.
- Yoshida N, et al. (2001) EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in *Arabidopsis*. *Plant Cell* 13:2471–2481.
- Schönrock N, et al. (2006) Polycomb-group proteins repress the floral activator AGL19 in the FLC-independent vernalization pathway. *Genes Dev* 20:1667–1678.
- Jiang DH, Wang YQ, Wang YZ, He YH (2008) Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the *Arabidopsis* Polycomb repressive complex 2 components. *PLoS ONE* 3:e3404.
- Goodrich J, et al. (1997) A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature* 386:44–51.
- Huh JH, Bauer MJ, Hsieh TF, Fischer RL (2008) Cellular programming of plant gene imprinting. *Cell* 132:735–744.
- Hershko A, Ciechanover A (1998) The ubiquitin system. *Annu Rev Biochem* 67:425–479.
- Ho MS, Ou C, Chan YR, Chien CT, Pi H (2008) The utility F-box for protein destruction. *Cell Mol Life Sci* 65:1977–2000.
- Vierstra RD (2009) The ubiquitin-26S proteasome system at the nexus of plant biology. *Nat Rev Mol Cell Biol* 10:385–397.
- Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K (2006) Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* 20:1123–1136.
- Dumbliuskas E, et al. (2011) The *Arabidopsis* CUL4-DDB1 complex interacts with MSI1 and is required to maintain MEDEA parental imprinting. *EMBO J* 30:731–743.
- Pazhouhandeh M, Molinier J, Berr A, Genschik P (2011) MSI4/FVE interacts with CUL4-DDB1 and a PRC2-like complex to control epigenetic regulation of flowering time in *Arabidopsis*. *Proc Natl Acad Sci USA* 108:3430–3435.
- Weigel D, et al. (2000) Activation tagging in *Arabidopsis*. *Plant Physiol* 122:1003–1013.
- Sieburth LE, Meyerowitz EM (1997) Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. *Plant Cell* 9:355–365.
- Chanvivattana Y, et al. (2004) Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. *Development* 131:5263–5276.
- Schubert D, et al. (2006) Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J* 25:4638–4649.
- Baumbusch LO, et al. (2001) The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res* 29:4319–4333.
- Kerppola TK (2006) Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nat Protoc* 1:1278–1286.
- Gray WM, et al. (1999) Identification of an SCF ubiquitin-ligase complex required for auxin response in *Arabidopsis thaliana*. *Genes Dev* 13:1678–1691.
- Zheng N, et al. (2002) Structure of the Cul1-Rbx1-5kp1-F box5kp2 SCF ubiquitin ligase complex. *Nature* 416:703–709.
- Santner A, Estelle M (2010) The ubiquitin-proteasome system regulates plant hormone signaling. *Plant J* 61:1029–1040.
- Kiyosue T, et al. (1999) Control of fertilization-independent endosperm development by the MEDEA polycomb gene in *Arabidopsis*. *Proc Natl Acad Sci USA* 96:4186–4191.
- Grossniklaus U, Vielle-Calzada J-P, Hoepfner MA, Gagliano WB (1998) Maternal control of embryogenesis by MEDEA, a polycomb group gene in *Arabidopsis*. *Science* 280:446–450.
- Kang IH, Steffen JG, Portereiko MF, Lloyd A, Drews GN (2008) The AGL62 MADS domain protein regulates cellularization during endosperm development in *Arabidopsis*. *Plant Cell* 20:635–647.
- Xu G, Ma H, Nei M, Kong H (2009) Evolution of F-box genes in plants: Different modes of sequence divergence and their relationships with functional diversification. *Proc Natl Acad Sci USA* 106:835–840.
- Jin J, et al. (2004) Systematic analysis and nomenclature of mammalian F-box proteins. *Genes Dev* 18:2573–2580.
- Gagne JM, Downes BP, Shiu SH, Durski AM, Vierstra RD (2002) The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis*. *Proc Natl Acad Sci USA* 99:11519–11524.
- Xu L, Shen WH (2008) Polycomb silencing of KNOX genes confines shoot stem cell niches in *Arabidopsis*. *Curr Biol* 18:1966–1971.
- Chae E, Tan QK, Hill TA, Irish VF (2008) An *Arabidopsis* F-box protein acts as a transcriptional co-factor to regulate floral development. *Development* 135:1235–1245.
- Deshais RJ (1999) SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu Rev Cell Dev Biol* 15:435–467.
- Bonetta L (2010) Protein-protein interactions: Interactome under construction. *Nature* 468:851–854.
- Spillane C, et al. (2007) Positive darwinian selection at the imprinted MEDEA locus in plants. *Nature* 448:349–352.
- Wang D, Tyson MD, Jackson SS, Yadegari R (2006) Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in *Arabidopsis*. *Proc Natl Acad Sci USA* 103:13244–13249.
- Ahn JH, et al. (2007) Isolation of 151 mutants that have developmental defects from T-DNA tagging. *Plant Cell Physiol* 48:169–178.
- Yadegari R, et al. (2000) Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms. *Plant Cell* 12:2367–2382.
- Saleh A, Alvarez-Venegas R, Avramova Z (2008) An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in *Arabidopsis* plants. *Nat Protoc* 3:1018–1025.
- Yoo SD, Cho YH, Sheen J (2007) *Arabidopsis* mesophyll protoplasts: A versatile cell system for transient gene expression analysis. *Nat Protoc* 2:1565–1572.
- Kim HJ, et al. (2008) Control of plant germline proliferation by SCF(FBL1) degradation of cell cycle inhibitors. *Nature* 455:1134–1137.