Molecular Plant
Research Article

CellPress
Partner Journal

Ancestral function but divergent epigenetic regulation of *HAIKU2* reveals routes of seed developmental evolution

Di Wu[1], Yiming Wei[1], Xiangyu Zhao[1], Boka Li[2,3], Huankai Zhang[1], Gang Xu[1], Juntong Lv[1], Dajian Zhang[1], Xiansheng Zhang[1,*] and Min Ni[4,*]

[1]National Key Laboratory of Crop Biology, College of Life Sciences, Shandong Agricultural University, Taian 271018, China

[2]State Key Laboratory of Systematic and Evolutionary Botany, CAS Center for Excellence in Molecular Plant Sciences, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

[3]University of Chinese Academy of Sciences, Beijing 100049, China

[4]Department of Plant and Microbial Biology, University of Minnesota at Twin Cities, Saint Paul, MN 55108, USA

***Correspondence: Min Ni (nixxx008@umn.edu), Xiansheng Zhang (zhangxs@sdau.edu.cn)**

https://doi.org/10.1016/j.molp.2022.09.002

ABSTRACT

Evolution is driven by various mechanisms. A directional increase in the embryo to endosperm ratio is an evolutionary trend within the angiosperms. The endosperm constitutes a major portion of the seed volume in *Poales* and some dicots. However, in other dicots such as *Arabidopsis* and soybean, the endosperm proliferates early, followed by embryo growth to replace the endosperm. The *Arabidopsis* leucine-rich repeat receptor protein kinase AtHAIKU2 (AtIKU2) is a key regulator of early endosperm proliferation. In this study, we found that *IKU2s* from *Brachypodium*, rice, and soybean can complement the abnormal seed developmental phenotype of *Atiku2*, while *AtIKU2* also rescues the defective endosperm proliferation in the *Brachypodium BdIKU2* knockout mutant seeds. *AtIKU2* and soybean *GmIKU2* are actively expressed a few days after fertilization. Thereafter, expression of *AtIKU2* is suppressed by the FIS-PRC2 complex-mediated H3K27me3. The soybean *GmIKU2* locus is also enriched with H3K27me3 marks. The histone methyltransferase AtMEA is unique to Brassicaceae, but one GmSWN in soybean plays a similar role in seed development as AtMEA. By contrast, the *BdIKU2* and rice *OsIKU2* loci are continuously expressed and are devoid of H3K27me3 marks. Taken together, these results suggest that *IKU2* genes retain an ancestral function, but the duration of their expression that is controlled by PRC2-mediated epigenetic silencing contributes to silenced or persistent endosperm proliferation in different species. Our study reveals an epigenetic mechanism that drives the development of vastly different seed ontogenies.

Key words: *Brassicales*, endosperm, *Fabales*, IKU2, MEDEA, *Poales*

Wu D., Wei Y., Zhao X., Li B., Zhang H., Xu G., Lv J., Zhang D., Zhang X., and Ni M. (2022). Ancestral function but divergent epigenetic regulation of *HAIKU2* reveals routes of seed developmental evolution. Mol. Plant. **15**, 1575–1589.

INTRODUCTION

Seeds represent an encapsulated embryonic stage unique to angiosperms and gymnosperms. In angiosperms, seeds comprise three major tissues: embryo, endosperm, and maternal tissues, including seed coats or pericarp. A double-fertilization event leads to the formation of a diploid embryo and triploid endosperm (Sun et al., 2010). Meanwhile, the seed coat develops from the sporophytic integument (Figueiredo and Köhler, 2014). Endosperm development involves two main phases: early pattern formation and late functional specialization (Floyd and Friedman, 2000). Early endosperm development involves either cellular or free nuclei division, and endosperm with free nuclei division has been reported only in monocots and several eudicots (Friedman, 1994; Floyd et al., 1999). Evolutionary divergences occur in late stages of endosperm development. The fate of endosperm development may be extensive, moderate, or limited (Floyd and Friedman, 2000). It often varies with the ratio of endosperm to embryo in mature seeds as well as endosperm storage and accessory storage tissue.

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

Arabidopsis endosperm development undergoes an early syncytial stage within a large seed cavity (Olsen, 2001; Berger, 2003). Cellularization then begins and marks the end of rapid endosperm proliferation (Kiyosue et al., 1999; Sørensen et al., 2002). Endosperm nuclear replication also ends after the last round of mitosis of the cellularized endosperm. Finally, the embryo grows into a cotyledon embryo at the expense of the endosperm. At maturation, only one layer of peripheral endosperm is left, along with uncellularized chalazal endosperm (Forbis et al., 2002; Olsen, 2004). Soybean (*Glycine max* [L.] Merr.) has a seed developmental path very similar to that of *Arabidopsis*, except that each stage lasts longer (Goldberg et al., 1989; Le et al., 2007). Grass endosperm development also begins with a coenocyte stage, followed by cellularization. However, grasses have a persistent endosperm with continuing cell division, and almost the entire volume of the seed is filled by the endosperm at maturity (Sabelli and Larkins, 2009; Nowicka et al., 2020). The major discrepancy during the late developmental stage is the continuing endosperm cell division, along with much greater seed growth, durative grain weight, and accumulation of storage compounds (Sreenivasulu et al., 2010; Dante et al., 2014). In *Brachypodium*, developing seeds complete endosperm proliferation and grain milk filling at 18 to 20 days after pollination (DAP) (Opanowicz et al., 2011). At the same time, cells in the central position continuously enlarge and form starchy endosperm. This process also involves endoreduplication and increases in nuclear DNA content (Sabelli and Larkins, 2009). In mature seeds, the outermost aleurone layer cells are still alive.

The *HAIKU* (*IKU*) pathway is critical for endosperm proliferation and initiation of endosperm cellularization (Zhou et al., 2009; Sun et al., 2010). Loss-of-function mutations in the VQ-motif protein HAIKU1 (IKU1), leucine-rich repeat (LRR) receptor kinase HAIKU2 (IKU2), and WRKY transcription factor MINISEED3 (MINI3) lead to precocious cellularization and reductions in endosperm nuclei number and mature seed size (Garcia et al., 2003; Luo et al., 2005; Wang et al., 2010). The expression of *IKU2* is endosperm specific and reaches a peak within a few DAP but is suppressed at 4 to 5 DAP (Kang et al., 2013). Polycomb group (PcG) proteins form transcriptional repression complexes, Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2), to regulate gene expression through histone modification (Orkin and Hochedlinger, 2011; Diez et al., 2014; Laugesen and Helin, 2014). The PRC2 complex is conserved between plants and mammals and suppresses gene expression through H3K27me3 deposited by Enhancer of Zeste (Ez) in mammals or CURLY LEAF/SWINGER/MEDEA (CLF/SWN/MEA) in plants (Grossniklaus et al., 1998; Goodrich et al., 1997; Luo et al., 1999). *Arabidopsis* has three major PRC2 complexes: EMBRYONIC FLOWER (EMF), VERNALIZATION (VRN), and Fertilization-Independent Seed (FIS) (Mozgova et al., 2015). *fis* mutants display autonomous endosperm proliferation in the absence of fertilization and show endosperm over-proliferation and embryo abortion after fertilization (Ohad et al., 1996; Chaudhury et al., 1997; Grossniklaus et al., 1998; Kiyosue et al., 1999). MEA is the histone methyltransferase of FIS-PRC2, and no MEA or FIS2 homologs have been found in 50 taxa, except Brassicaceae (Spillane et al., 2007; Luo et al., 2009). Homologs of CLF, SWN, and FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) are widely found in many species other than Brassicaceae (Miyake et al., 2009; Tonosaki and Kinoshita, 2015).

The size and morphology of seeds affect not only evolutionary fitness but also grain yield of crops. Cereal grains with endosperm as the major component are the foremost source of calories, nutrition, and industrial raw materials for human needs and livestock consumption (Black et al., 2006). Seed size is therefore under continued natural and domestic selection. Developmental evolution is the driving force for various seed morphologies to fit a diverse range of natural environments over millions of years. In this report, we explore the epigenetic regulation of *IKU2* expression in relation to two vastly different endosperm developmental ontogenies: limited proliferation in *Arabidopsis* of *Brassicales* and soybean of *Fabales*, and extensive proliferation in *Brachypodium distachyon* and rice of *Poales*.

RESULTS

Expression of *Brachypodium BdIKU2* driven by the *AtIKU2* promoter rescues *Arabidopsis iku2* mutant seed phenotype

AtIKU2 mutation retards endosperm proliferation and reduces the size of mature seeds in *Arabidopsis* (Garcia et al., 2003; Luo et al., 2005). We selected a number of species with high sequencing quality for phylogenetic analysis (Supplemental Table 1). By searching with the AtIKU2 protein sequence, we found a conserved *IKU2* cluster (Supplemental Figure 1). However, genes from *Amborella trichopoda* with remote similarity to *AtIKU2* are located in various clusters, consistent with the result shown by Man et al. (2020). We tested one *AtIKU2*-like gene and four other genes with remote similarity to *AtIKU2* from *Brachypodium distachyon* (*Bd*): *Bd1G08240*, *Bd1G59290*, *Bd1G58460*, *Bd4G00900*, and *Bd5G15060*. We failed to amplify the *Bd5G15060* genomic sequence. Its genomic sequence is 5539 base pairs (bp), and its CDS is 4812 bp, much larger than that of the other four *BdIKU2* homologs. The genomic sequences of the other four *AtIKU2*-like genes, along with the *AtIKU2* gene driven by an 811-bp *AtIKU2* promoter, were then introduced into *Atiku2*. For *pAtIKU2*::*AtIKU2*/*Atiku2*, we generated 24 independent transgenic plants, and 20 lines (83%) had a wild-type seed mass, ranging from 1.6 to 2.0 mg per 100 seeds. We pursued and presented representative single-insertion lines. *Bd4G00900* rescued the small-seed phenotype of *Atiku2* and was named *BdIKU2* (Figure 1A and 1B, Supplemental Figure 2A and 2B). The other three *BdIKU2* homologs failed to rescue the *Atiku2* seed phenotype (Supplemental Figure 2C). For *pAtIKU2*::*BdIKU2*/*Atiku2*, we generated 36 independent transgenic plants, and 31 lines (86%) had a wild-type seed-size phenotype, ranging from 1.6 to 2.0 mg per 100 seeds. Expression of all transgenes in each line was also examined (Supplemental Figure 2C and 2D), and specificity of the PCR probe for the four *BdIKU2* genes was verified (Supplemental Figure 2E). Interestingly, Bd4G00900 is in one of two clades of grass IKU2 homologs, whereas the other four are in the second clade. It is possible that genes in the second clade have evolved a different function.

Because both *AtIKU2* and *BdIKU2* were driven by the *AtIKU2* promoter, we examined *AtIKU2* or *BdIKU2* expression from 1 to

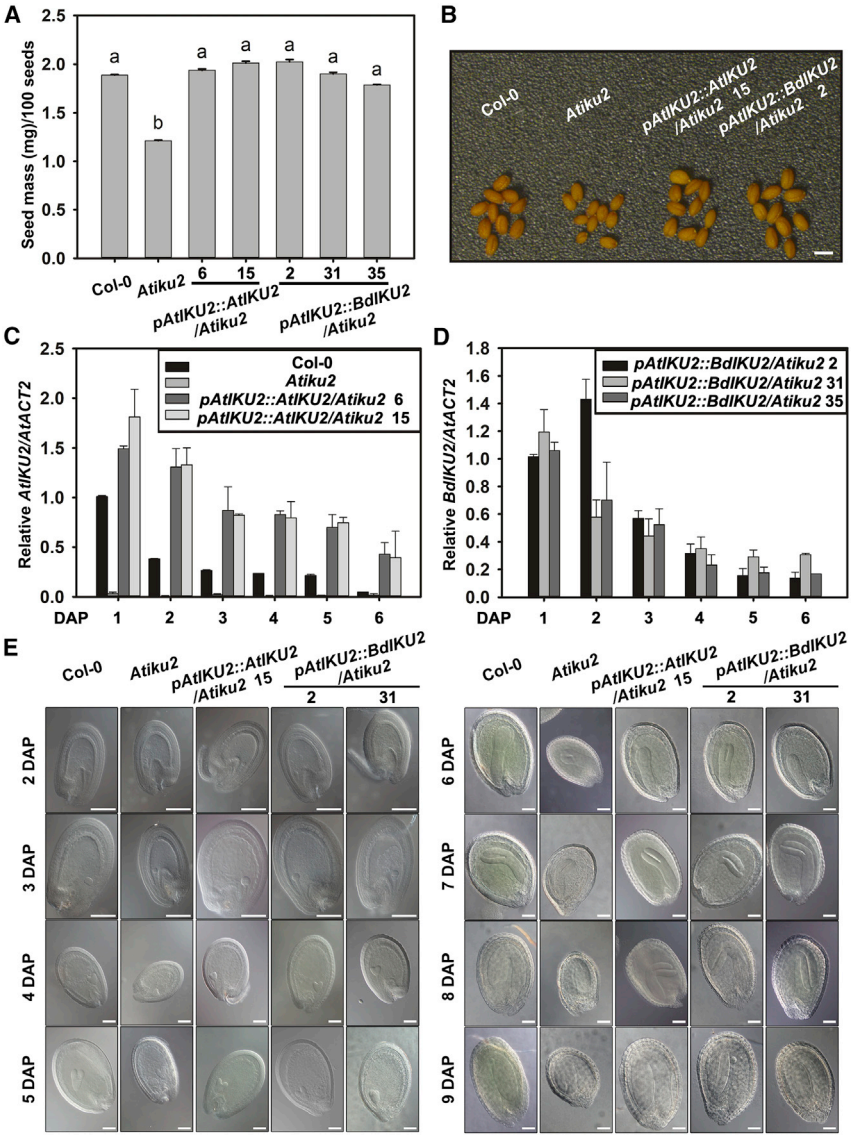

Figure 1. Expression of *BdIKU2* driven by the *AtIKU2* promoter rescues *Atiku2* mutant seed phenotype.

(A) Seed mass of various genotypes. Letters indicate significant differences from *Atiku2* determined by ANOVA ($p < 0.001$).

(B) Mature seed images of Col-0, *Atiku2*, one *pAtIKU2*::*AtIKU2*/*Atiku2* transgenic line, and one *pAtIKU2*::*BdIKU2*/*Atiku2* transgenic line. Scale bars, 0.4 mm.

(C) Expression of *AtIKU2* in Col-0, *Atiku2*, and two *pAtIKU2*::*AtIKU2* transgenic lines in *Atiku2* from 1 to 6 DAP. Expression of *AtIKU2* in each sample was normalized to that of *AtACT2* and then to that of Col-0 at 1 DAP. Expression of *AtIKU2* in Col-0 at 1 DAP has an artificial value of 1. Expression of *AtIKU2* in Col-0 and the two transgenic lines at 1 and 2 DAP was significantly different from that from 3 to 6 DAP by Student's t test ($p < 0.001$). Expression of *AtIKU2* in Col-0 from 3 to 5 DAP was significantly different from that at 6 DAP by Student's t test ($p < 0.05$).

(D) Expression of *BdIKU2* in Col-0 and the two transgenic lines and three *pAtIKU2*::*BdIKU2* transgenic lines in *Atiku2* from 1 to 6 DAP. Expression of *BdIKU2* in each sample was normalized to that of *AtACT2* and then to that of *pAtIKU2*::*BdIKU2*/*atiku2* line 2 at 1 DAP. Expression of *AtIKU2* in *pAtIKU2*::*BdIKU2*/*atiku2* line 2 at 1 DAP has an artificial value of 1. Expression of *BdIKU2* in the three transgenic lines from 1 to 3 DAP was significantly different from that from 4 to 6 DAP by Student's t test ($p < 0.05$).

(E) DIC images of seed development in Col-0, *Atiku2*, one *pAtIKU2*::*AtIKU2Atiku2* transgenic line, and two *pAtIKU2*::*BdIKU2*/*Atiku2* transgenic lines from 1 to 9 DAP. Scale bars, 100 μm. Data in **(A)**, **(C),** and **(D)** were calculated from three biological replicates. All error bars represent SEM, n = 3.

6 DAP. Expression of *AtIKU2* in Col-0 reached a peak at 1 DAP and gradually declined from 2 to 6 DAP (Figure 1C). Expression of the *AtIKU2* or *BdIKU2* transgene driven by the *AtIKU2* promoter followed a similar pattern (Figure 1C and 1D). Compared with Col-0, *Atiku2* seeds have a small seed cavity due to reduced endosperm proliferation from 3 DAP at the globular stage until 9 DAP, as shown in differential interference contrast (DIC) images (Figure 1E). The *Atiku2* embryo developed normally before 4 DAP, but the size of the embryo was reduced at 5 DAP at the early torpedo stage. Seeds of *pAtIKU2*::*AtIKU2*/*Atiku2* or *pAtIKU2*::*BdIKU2*/*Atiku2* rescue lines were indistinguishable from those of Col-0.

To provide histological and cellular resolution of the *Atiku2* phenotype and phenotypic rescue, we prepared thin sections from the various genotypes (Supplemental Figure 3 and Figure 2A). Cellularization in Col-0 started from the micropylar endosperm (MCE) and several peripheral endosperm (PEN) nuclei at 4 DAP or the triangular to early-heart stage and spread to nearly half of the central cavity at 5 DAP. By 6 DAP, endosperm cellularization was completed. *Atiku2* exhibited precocious cellularization as early as 3 DAP or the mid-globular stage from the MCE and PEN. Cellularization then quickly expanded to nearly half of the central cavity at 4 DAP. By 5 DAP, cellularization was completed in *Atiku2*, but *Atiku2* had fewer endosperm cells than Col-0. Seeds of both *pAtIKU2*::*AtIKU2* and *pAtIKU2*::*BdIKU2* rescue lines showed a normal trend of cellularization, number, and size of endosperm cells like that of Col-0.

Expression of *BdIKU2* by its own promoter not only failed to rescue the *Atiku2* small-seed phenotype but also caused the production of both non-defective and defective developing seeds (Figure 2A and 2B). The expression patterns of *BdIKU2* in several representative lines were random and expression levels remained high from 1 to 6 DAP, as *BdIKU2* misexpression lines and *BdIKU2* over-expression lines (Supplemental Figure 4A). It is likely that the *BdIKU2* promoter bears various *cis* elements and mis-accommodates various *Arabidopsis* transcription factors, driving a different spatial or temporal expression pattern. We also performed *in situ* hybridization on the defective developing seeds. Compared with

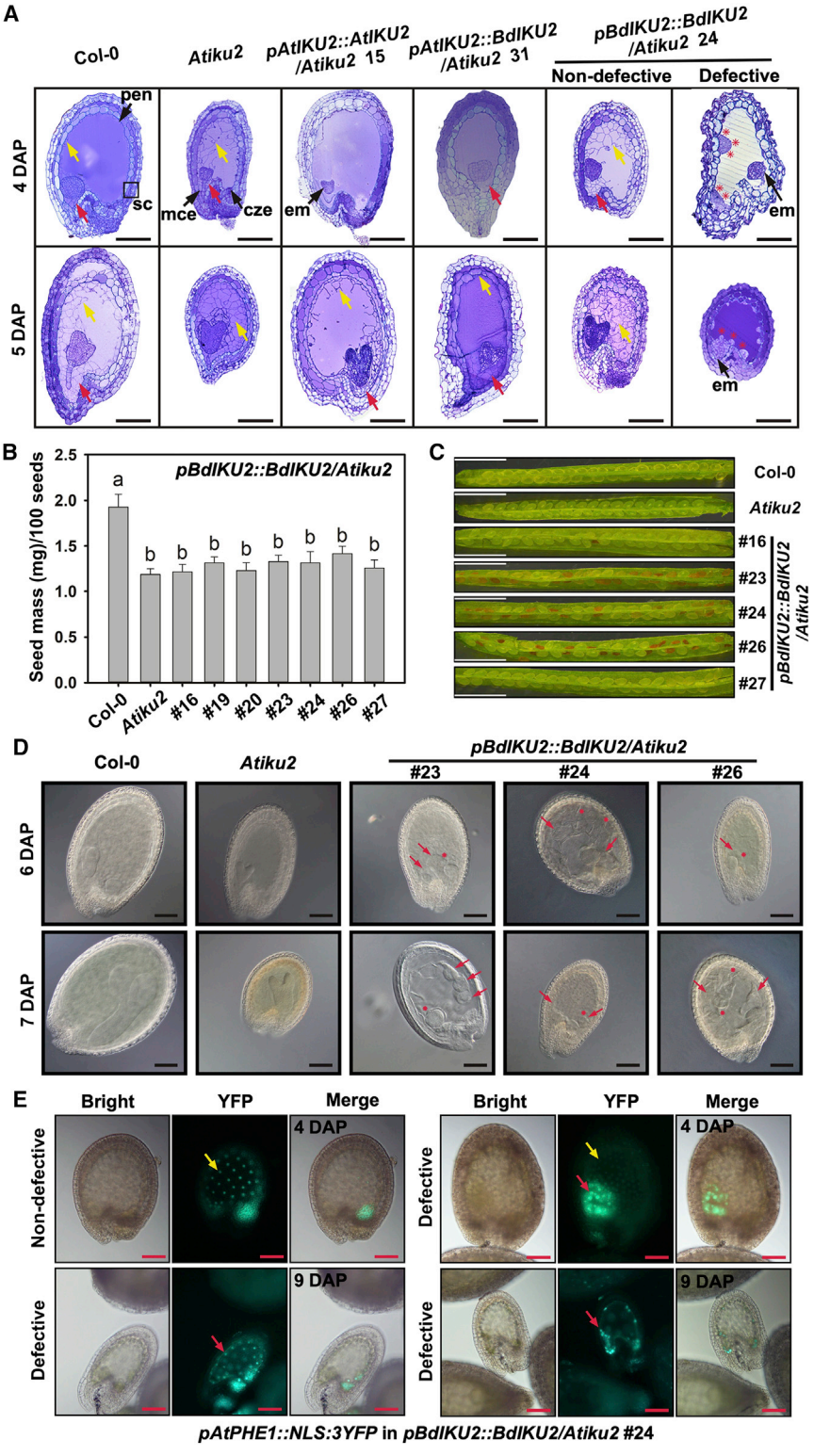

Figure 2. Expression of *BdIKU2* driven by the *Brachypodium BdKU2* promoter in the *Atiku2* mutant causes defective seed development.

(A) Tissue sections of Col-0, *Atiku2*, and various transgenic seeds in *Atiku2* at 4 and 5 DAP. Red arrows indicate cellularized micropylar endosperm (mce). Yellow arrows indicate cellularized peripheral endosperm (pen). Red stars indicate defective structure. cze, chalazal endosperm; em, embryo; sc, seed coat. Scale bars, 100 μm. The complete set of tissue sections at various DAP is shown in Supplemental Figure 3.

(B) Seed mass of Col-0, *Atiku2,* and seven *pBdIKU2*::*BdIKU2*/*Atiku2* transgenic lines. Letters indicate significant differences from *Atiku2* determined by ANOVA ($p < 0.001$). Data were calculated from three biological replicates.

(C) Silique image showing defective seeds from several *pBdIKU2*::*BdIKU2* transgenic lines in *Atiku2*. Scale bars, 2 mm.

(D) Nodule structures in three *pBdIKU2*::*BdIKU2*/*Atiku2* transgenic lines at 6 and 7 DAP. Red arrows indicate defective structures. Red stars indicate free nuclei. Scale bars, 100 μm. The complete set of DIC images at various DAP is shown in Supplemental Figure 5.

(E) Expression of the *pPHE1*::*NLS*::*3YFP* reporter in non-defective and defective seeds. Yellow arrows indicate non-defective endosperm nuclei, and red arrows indicate nodule structure with a cluster of nuclei. Scale bars, 50 μm.

pAtIKU2::*BdIKU2* transgenic plants, signals of *BdIKU2* transcripts in defective developing seeds from *pBdIKU2*::*BdIKU2* transgenic plants were detected from 3 to 7 DAP in the seed coat, embryo, syncytial endosperm nuclei, and cellularized endosperm (Supplemental Figure 4B). Non-defective seeds retained the early cellularization phenotype of *Atiku2*, whereas defective seeds had either an uncellularized or partially cellularized endosperm (Figure 2A).

Several homozygous transgenic lines produced severely defective seeds (Figure 2C). Defective seeds possessed multiple defects shown in DIC images: a mass of uncellularized nuclei surrounding the embryo and inner integument from 4 to 9 DAP, a disorganized cellularization process in some areas, and retarded growth of the embryo after the globular stage (Figure 2D and Supplemental Figure 5). The *BdIKU2* promoter is not targeted by PRC2, and the endosperm was over-proliferated in some of these lines (Figure 2D and Supplemental Figure 4). The defective seeds resembled the weak phenotype of the *fis* mutants, with a many-nodule structure (Kiyosue et al., 1999). Those nodules may represent new endosperm tissues arising from the already cellularized and arrested endosperm after 5 DAP. We introduced an *NLS*:*3YFP* reporter driven by an endosperm-specific *PHE1* promoter to the *pBdIKU2*::*BdIKU2*/*Atiku2* seeds (Zhang et al., 2018). The nodule structures are likely to be endosperm tissues, as the reporter gene is driven by the *PHE1* promoter, which confers endosperm expression (Figure 2E). As positive controls, the reporter signals in non-defective seeds are detected only in regular endosperm

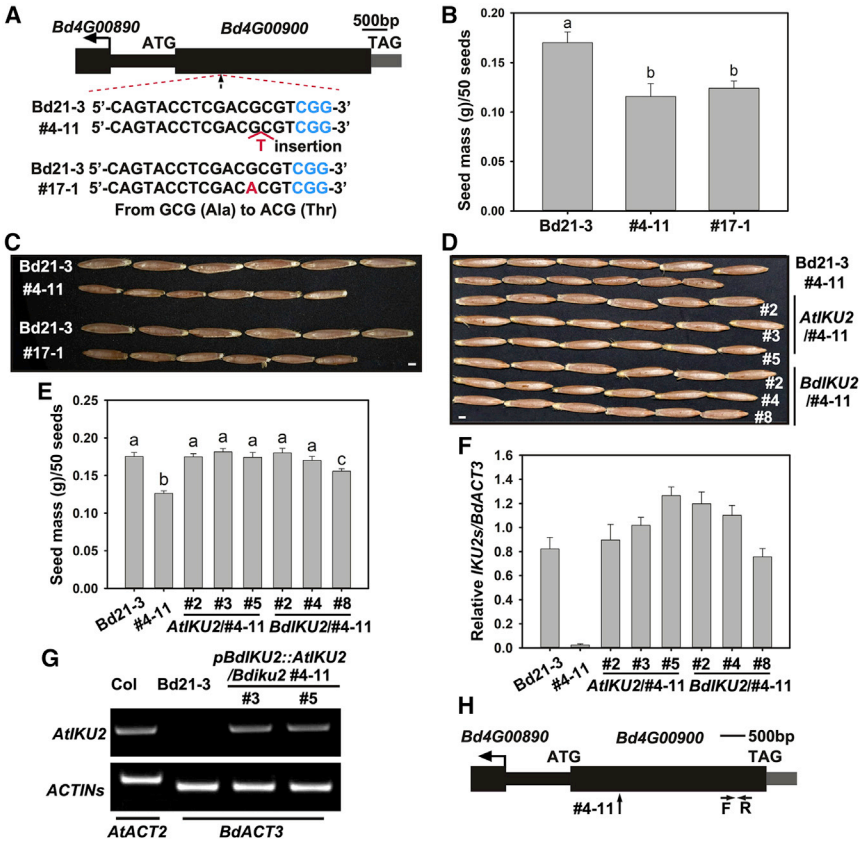

Figure 3. *BdIKU2* is critical for *Brachypodium* endosperm development.

(A) CRISPR-Cas design and lesions of *BdIKU2* mutants.

(B) Wild-type and mutant seed mass. Letters indicate significant differences from Bd21-3 determined by ANOVA ($p < 0.001$).

(C) Mature seed images of Bd21-3 and two *Bdiku2* mutants. Scale bars, 1 mm.

(D) Mature seed images of Bd21-3, *Bdiku2* 4-11 mutant, and *Bdiku2* mutant rescue lines. Scale bars, 1 mm.

(E) Seed mass of Bd21-3, *Bdiku2* 4-11, three *pBdIKU2*::*AtIKU2* transgenic lines in *Bdiku2 4-11,* and three *pBdIKU2*::*BdIKU2* transgenic lines in *Bdiku2* 4-11. Letters indicate significant differences from Bd21-3 determined by ANOVA ($p < 0.001$).

(F) Expression of *BdIKU2* in Bd21-3, *Bdiku2* 4-11, three *pBdIKU2*::*AtIKU2* transgenic lines in *Bdiku2 4-11,* and three *pBdIKU2*::*BdIKU2* transgenic lines in *Bdiku2* 4-11 at 5 DAP. Expression of *BdIKU2* in each sample was normalized to that of *BdACT3*. Expression of *AtIKU2* in transgenic lines 2 and 3 and of *BdIKU2* in transgenic line 8 was not significantly different from expression of *BdIKU2* in Bd21-3 by Student's t test ($p > 0.05$). Expression of *AtIKU2* in transgenic line 5 and of *BdIKU2* in transgenic lines 2 and 4 was significantly different from expression of *BdIKU2* in Bd21-3 by Student's t test ($p < 0.05$).

(G) Expression of *AtIKU2* in Col-0, Bd21-3, and *pBdIKU2*::*AtIKU2* transgenic lines in *Bdiku2* 4-11 for verification of probe specificity.

(H) *BdIKU2* locus, *Bdiku2* 4-11, and qPCR primers used for verification of *BdIKU2* expression in *Bdiku2* 4-11. Data in **(B)**, **(E),** and **(F)** were calculated from three biological replicates. All error bars represent SEM, n = 3.

nuclei but not in the integument. Compared with non-defective seeds, defective seeds at 4 DAP had more endosperm nuclei and abnormal nodules, and they also showed strong fluorescent signals. In the defective seeds at 9 DAP, the embryo did not develop well, and the central cavity still contained many endosperm nuclei.

We also introduced *pBdIKU2*::*BdIKU2* into the Col-0 background. Three transgenic lines produced small seeds (Supplemental Figure 6A). The small-seed phenotype was likely to be due to co-suppression of the transgene as well as the endogenous *AtIKU2* gene (Supplemental Figure 6C and 6D). The other three lines showed defective endosperm and seed development with aborted seeds (Supplemental Figure 6A and 6B). *BdIKU2* in those lines was highly expressed and mis-expressed (Supplemental Figure 6D and 6E). We also observed a cluster of similar nodule structures when we introduced the *BdIKU2* gene into the *Atiku2* mutant (Supplemental Figures 2D and 6F).

BdIKU2 is critical for *Brachypodium* endosperm development

To learn the biological function of BdIKU2 in endosperm development, we performed targeted mutagenesis of *BdIKU2* using the CRISPR-Cas9 system (Feng et al., 2013). Bd21-3 was the wild type used in this study. Two independent mutants were identified (Figure 3A). Line 4-11 had a nucleotide T insertion between nucleotides 820 (G) and 821 (C), causing a frameshift. Line 17-1 had a base change from G to A at nucleotide 820, causing a change of amino acid 274 from alanine (GCG) to threonine (ACG). The seed mass of 4-11 and 17-1 was significantly reduced compared with that of the wild type (Figure 3B). Length, width, and thickness of the caryopses are shown side-by-side for Bd21-3, 4-11, and 17-1 (Figure 3C).

The developing caryopses of Bd21-3 and the *Bdiku2* mutant 4-11 were photographed under a microscope to examine morphological changes in the germinal, transverse, and longitudinal dimensions (Supplemental Figure 7). *Bdiku2* had a smaller caryopsis at 3 DAP and thereafter, with a clear reduction in caryopsis width, length, and thickness. Thin sections of Bd21-3 and *Bdiku2* were examined to detect cytological changes during endosperm development. In transverse and longitudinal sections, Bd21-3 showed no signs of cellularization at 2 DAP, whereas *Bdiku2* had begun to cellularize around a few nuclei in the embryo-surrounding region (Supplemental Figures 8 and 9). Cellularization progressed rapidly toward the central vacuoles from 3 to 4 DAP in the *Bdiku2* mutant compared with the Bd21-3 wild type (Supplemental Figure 9). By 5 DAP, one-third to half of the central cavity was still not cellularized in Bd21-3, but the central cavity was fully cellularized in *Bdiku2*.

We next explored whether *BdIKU2* or *AtIKU2* driven by a 1200-bp *BdIKU2* promoter rescued the *Bdiku2* seed phenotype. Both *pBdIKU2*::*BdIKU2* and *pBdIKU2*::*AtIKU2* constructs complemented the *Bdiku2* seed phenotype in terms of caryopsis width, length, and thickness (Figure 3D and 3E). qRT-PCR analysis also verified transgene expression with a pair of primers near the 3′

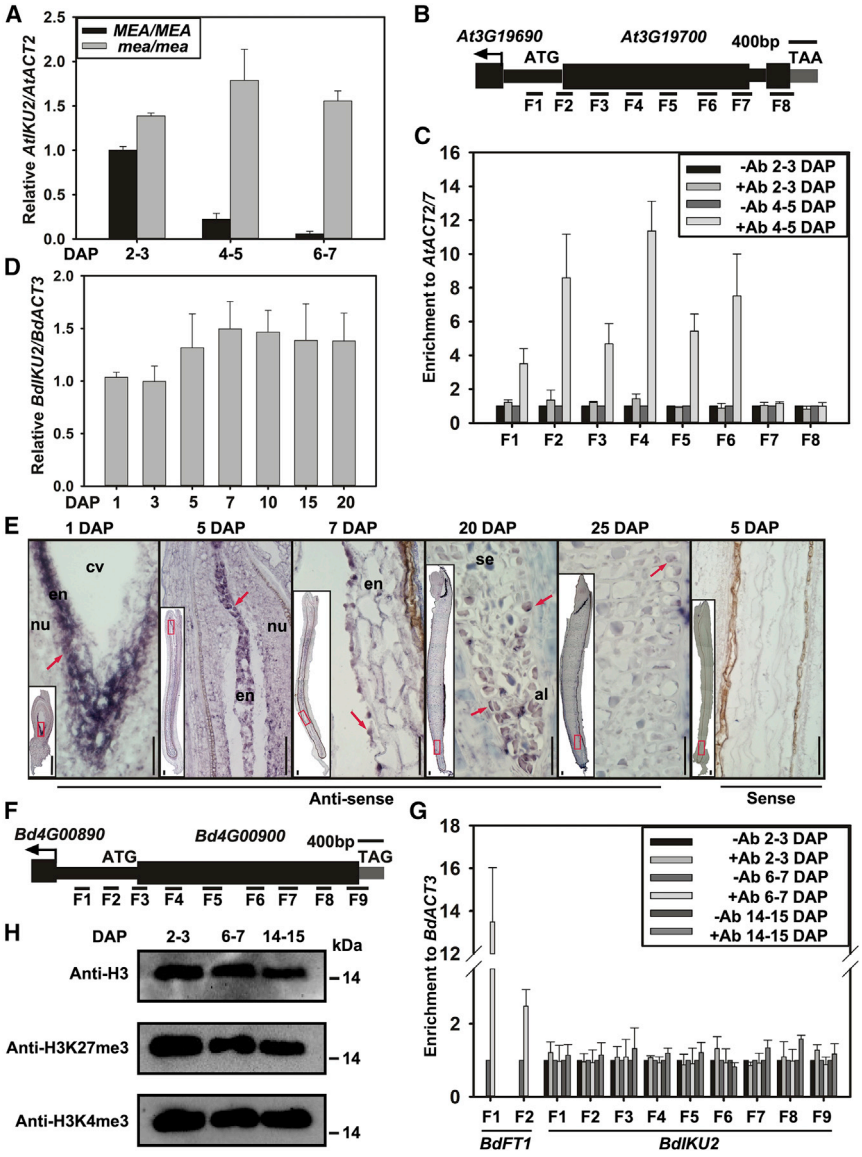

Figure 4. *AtIKU2* expression is suppressed by FIS-PRC2, but *BdIKU2* expression is continuous.

(A) Expression of *IKU2* in L*er* and *Atmea* at various DAP. *Atmea* mutant seeds were selected by their morphology after 3 DAP. Expression of *AtIKU2* in each sample was normalized to that of *AtACT2* and *MEA*/*MEA* at 2 to 3 DAP. Expression of *AtIKU2* in *MEA*/*MEA* at 2 to 3 DAP has an artificial value of 1. Expression of *AtIKU2* in L*er* at 1 to 2 DAP was significantly different from that at 4 to 5 and 6 to 7 DAP by Student's t test ($p < 0.001$). Expression of *AtIKU2* in *Atmea* at 1 to 2 DAP was not significantly different from that at 4 to 5 and 6 to 7 DAP by Student's t test ($p > 0.05$).

(B) Eight ChIP fragments were designed for the *AtIKU2* locus.

(C) ChIP analysis of histone H3K27 trimethylation of the *AtIKU2* locus at 2 to 3 and 4 to 5 DAP. Enrichment of DNA fragments was quantified by qPCR and normalized to that of *AtACT2/7* and –antibodies at 2 to 3 DAP. Enrichment of –antibodies at 2 to 3 DAP for each fragment has an artificial value of 1. Enrichment of +antibody at 4 to 5 DAP was significantly different from that of –or +antibody at 2 to 3 DAP and –antibody at 4 to 5 DAP by Student's t test ($p < 0.05$) for fragments 1 to 6.

(D) Expression of *BdIKU2* in isolated *Brachypodium* endosperm stripped off from nucellar tissue and embryo from 1 to 20 DAP. Expression of *BdIKU2* in each sample was normalized to that of *BdACT3*. Expression of *BdIKU2* in Bd21-3 at 1 DAP was not significantly different from that from 3 to 20 DAP by Student's t test ($p > 0.05$).

(E) *In situ* hybridization of *BdIKU2* with longitudinal sections at various DAP. Red arrows indicate endosperm cells with hybridization signals. Scale bars, 50 μm. al, aleurone endosperm; cv, central vacuole; en, endosperm; nu, nucellar tissue; se, starchy endosperm.

(F) Nine ChIP fragments were designed for the *BdIKU2* locus.

(G) ChIP analysis of histone H3K27 trimethylation of the *BdIKU2* locus at various DAP. *BdFTL1* ChIP analysis was performed with leaf tissue. Enrichment of DNA fragments was quantified by qPCR and normalized to that of *BdACT3* and –antibodies at 2 to 3 DAP. Enrichment of –antibodies at 2 to 3 DAP for each fragment has an artificial value of 1. No *BdIKU2* fragments differed significantly from each other in enrichment at 2 to 3, 6 to 7, or 14 to 15 DAP by Student's t test ($p > 0.05$). Enrichment of two *BdFT1* fragments +antibody was significantly different from that of –antibody at 6 to 7 DAP by Student's t test ($p < 0.01$).

(H) Immunoblots of immunoprecipitated histone H3 with anti-H3, anti-H3K27me3, and anti-H3K4me3 antibodies. All data in **(A)**, **(C)**, **(D),** and **(G)** were calculated from three biological replicates. All error bars represent SEM, n = 3.

end (Figure 3F and 3H). The specificity of the *AtIKU2* probe for the *pBdIKU2*::*AtIKU2* transgene is shown in Figure 3G. Although #4-11 had only one nucleotide T insertion in the middle of the gene, we hardly detected *BdIKU2* mRNA in the #4-11 mutant with the same pair of primers. In some CRISPR-Cas9-mediated knockout lines, target gene expression was not readily detected (Huang et al., 2016; Yuva-Aydemir et al., 2019).

FIS-PRC2 suppresses *AtIKU2* expression and this regulation is absent in *Brachypodium*

AtIKU2 was highly expressed at 1 to 3 DAP and then barely expressed at 6 to 7 DAP in wild-type L*er* (Figure 4A). MEA is the H3K27me3 methyltransferase of FIS-PRC2. The *Atmea* mutant was maintained as heterozygous *AtMEA*/*mea*; after selfing, we collected *Atmea*/*mea* mutant seeds based on their defective morphology at various DAP as shown by Kiyosue et al. (1999). In *Atmea*, *AtIKU2* was still strongly expressed at 4 to 5 and 6 to 7 DAP (Figure 4A). *AtIKU2* expression is likely to be repressed by FIS-PRC2 after endosperm proliferation ceases at 4 to 5 DAP. We isolated hexaploid endosperm nuclei at 2 to 3 and 4 to 5 DAP using ploidy-based fluorescence-activated cell sorting (FACS) (Zheng and Gehring, 2019). Triploid 3C and tetraploid 4C nuclei had too large an overlap to be easily separated, whereas hexaploid 6C nuclei offered better resolution and were easily separated from tetraploid 4C nuclei (Supplemental

Figure 10A). The region from the −811 bp promoter region to the +3300 bp 3′ UTR was covered by eight pairs of primers (Figure 4B). H3K27me3 enrichment was not observed for the *AtIKU2* locus at 2 to 3 DAP across the whole region. At 4 to 5 DAP, H3K27me3 marks were densely deposited from fragment 2 to fragment 6, across the transcription start site (TSS) to the exons (Figure 4C). In both leaf and endosperm, trimethylation marks were also detected at the *AtIKU2* locus and the *AtCKX2* locus using either CUT&RUN or chromatin immunoprecipitation sequencing (ChIP-seq) data (Table S2a to S2d in Zheng and Gehring, 2019). AtCKX2 is another FIS-PRC2 target (Li et al., 2013). It had a much less significant role in the deposition of trimethylation marks at the *AtIKU2* locus on the basis of the INTACT-ChIP approach (Table S2e in Zheng and Gehring, 2017), consistent with previous results of Moreno-Romero et al. (2016).

Cereal endosperm exhibits sustained proliferation. Is the expression of *BdIKU2* synchronous with endosperm proliferation? *BdIKU2* indeed maintained a relatively stable expression in isolated *Brachypodium* endosperm stripped of the nucellar tissue and embryo from 1 to 20 DAP (Figure 4D). *BdIKU2* was actively expressed at 1 DAP at the coenocyte stage, mostly in proliferating endosperm between the central vacuole and the nucellar tissue (Figure 4E). *BdIKU2* mRNA was also detected in the cell files at endosperm closure at 5 and 7 DAP during the cellularization process. *BdIKU2* was continuously expressed in the dividing central starchy endosperm and aleurone layer cells at 20 and 25 DAP, with enlarged views (Supplemental Figure 11A and 11B). For example, *BdIKU2* hybridization signals were detected at 20 and 25 DAP in dividing endosperm cells and cells beside the protein body. No signals were detected in dead central starchy endosperm cells. We observed continuous *BdIKU2* expression until 25 DAP. However, developing *Brachypodium* seeds complete endosperm proliferation and grain milk filling at 18 to 20 DAP (Opanowicz et al., 2011). Because growth conditions such as temperature and photoperiod were different, our plants may have completed endosperm proliferation and grain milk filling at a date later than 20 DAP.

Using a similar strategy, we isolated *Brachypodium* hexaploid endosperm nuclei at three different stages: 2 to 3, 6 to 7, and 14 to 15 DAP (Supplemental Figure 10B). No H3K27me3 marks were detected around the *BdIKU2* locus for the whole region at various DAP (Figure 4F and 4G). The *BdIKU2* locus was also not enriched with H3K27me3 marks in 24C endosperm nuclei at the late stage of endoreduplication (Supplemental Figure 10C). In proteins precipitated by anti-H3 antibodies from total protein extract, both H3K27me3 and H3K4me3 modifications were detected from 2 to 3 DAP until 14 to 15 DAP (Figure 4H). *Bd Flowering Locus T1* (*BdFT1*) was used as a control (Wu et al., 2013), and H3K27me3 marks were enriched at this locus (Figure 4G and Supplemental Figure 10D). BdFT1 has a function similar to that of *Arabidopsis* FT (Adrian et al., 2010; Zhang et al., 2020a, 2020b). Various cereal *CLF* and *SWN* genes are expressed from vegetative to reproductive tissues (Zhong et al., 2018; Cheng et al., 2020).

Previous studies predicted 170 Polycomb response elements (PREs) associated with various genes within 600 bp around the Polycomb ChIP-seq peaks, including CTCC, CGG, G-box, AC rich, GA repeat, and Telo-box (Table S2 in Xiao et al., 2017). Xiao et al. also identified 233 PRE-binding transcription factors, such as TCP proteins and E2F proteins (Table S3 in Xiao et al., 2017). We analyzed the promoters of *Arabidopsis*, soybean, rice, and *Brachypodium IKU2s* for enriched and unique *cis* elements. We identified eight *cis* elements that were present only in the *AtIKU2* and *GmIKU2* promoters but not in the *BdIKU2* and *OsIKU2* promoters (Supplemental Table 2). Among these eight *cis* elements, SITEIIATCYTC is likely to be bound by a TCP-domain transcription factor, and E2FCONSENSUS is likely to be bound by an E2F protein (Vandepoele et al., 2005; Welchen and Gonzalez, 2005).

IKU2 function is also conserved in rice and soybean

Rice and soybean have seed developmental paths similar to those of *Brachypodium* and *Arabidopsis*, respectively. Six *IKU2*-like genes from rice and soybean were introduced into *Atiku2* under the control of the *AtIKU2* promoter. One rice homolog (*LOC_Os12g43640*) and one soybean homolog (*Gm6G090800*) rescued the small-seed phenotype of *Atiku2* (Figure 5A and 5F, Supplemental Figure 12A and 12B). Of 36 independent *pAtIKU2*::*OsIKU2*/*Atiku2* transgenic plants, 30 lines (83%) had a seed mass ranging from 1.6 to 2.0 mg per 100 seeds. Of the 36 independent *pAtIKU2*::*GmIKU2*/*Atiku2* transgenic plants, 32 lines (89%) had a seed mass ranging from 1.6 to 2.0 mg per 100 seeds. Seed size increase was positively correlated with transgene expression in at least five independent transgenic plants for each *OsIKU2* or *GmIKU2* construct (Supplemental Figure 12C and 12D). The specificity of the PCR probes is shown in Supplemental Figure 12E and 12F. Seeds from *OsIKU2* and *GmIKU2* transgenic plants were indistinguishable from those of Col-0 in terms of seed cavity, endosperm proliferation, and embryo size from 1 to 9 DAP (Supplemental Figure 13). Their expression patterns were similar to that of endogenous *AtIKU2* (Figure 5B and 5G).

Like *BdIKU2*, endogenous *OsIKU2* showed persistent expression up to 20 DAP (Figure 5C). *OsIKU2* transcripts were detected during the cellularization stage at 3, 4, and 5 DAP (Figure 5D). Strong signals were also visible in the dividing starchy endosperm cells at 7 DAP and in the aleurone layer at 10 and 12 DAP. No H3K27me3 marks were enriched at the *OsIKU2* locus (Figure 5E). Os Heading Date 3a (OsHd3a) is a homolog of AtFT (Tamaki et al., 2007), and the *OsHd3a* locus was enriched with a high level of H3K27me3 marks (Figure 5E and Supplemental Figure 10E).

The IKU2-MEA module exists in soybean

Endosperm reaches its highest volume at the mid-heart stage in *Pisum sativum* and *Glycine max* (Marinos, 1970; Chamberlin et al., 1993). *Glycine max* (L.) Merr has a life cycle of 120 days (Goldberg et al., 1989; Le et al., 2007). Soybean cv. Dongnong 50 that we used as a wild type has a life cycle of about 110 days. In paraffin sections, the embryo derived from double fertilization went through the octant stage at 1 to 2 DAP, the globular stage at 3 to 4 DAP, the heart stage at 5 to 7 DAP, the cotyledon stage at 8 to 13 DAP, and a long maturation stage (Figure 6A). Endosperm proliferated through the free nuclei stage from 1 to 4 DAP, the cellular stage from 5 to 7 DAP, and the absorption stage

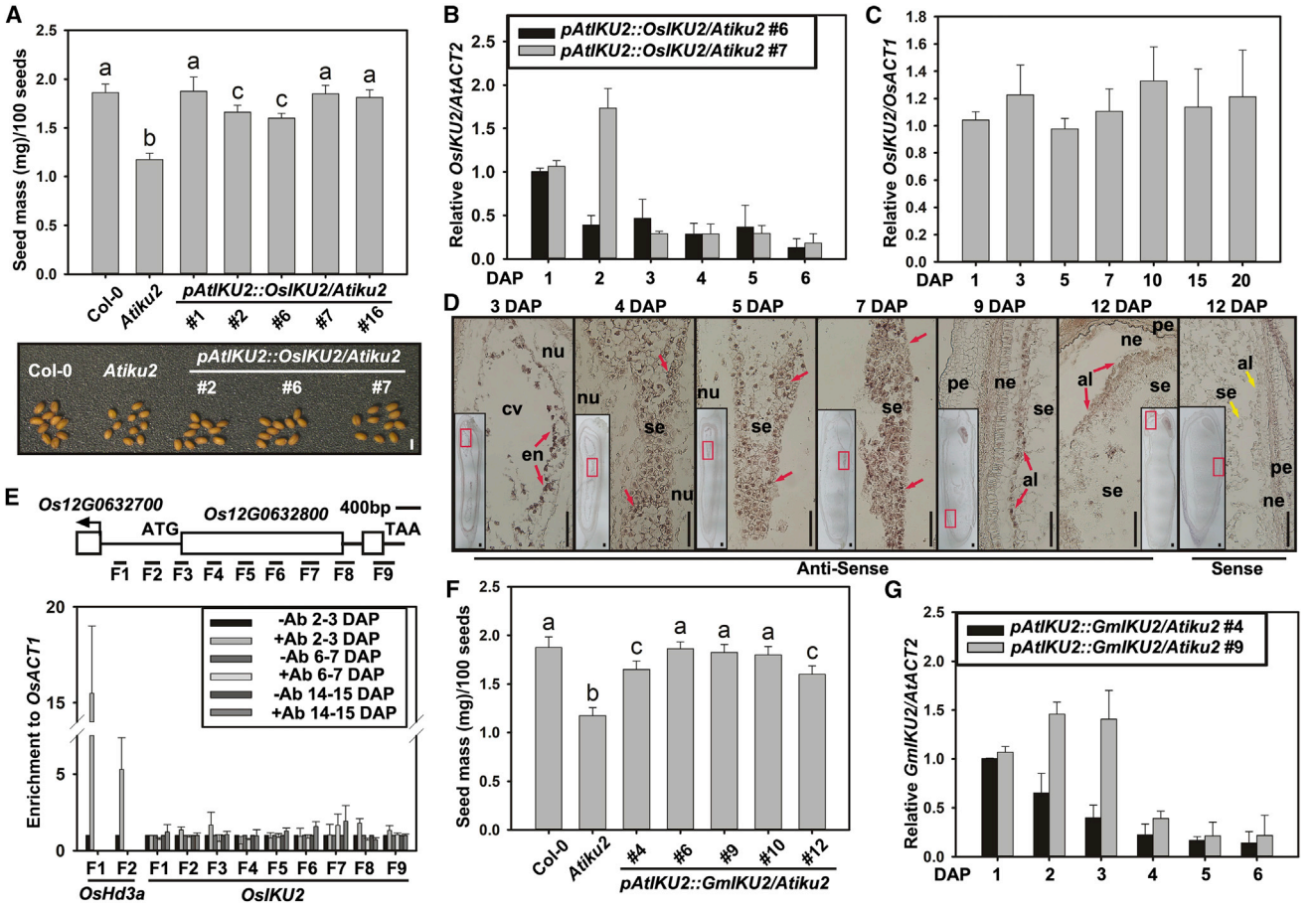

Figure 5. *AtIKU2* homologs are found in rice and soybean.

Average seed mass per 100 seeds from Col-0, *Atiku2*, and five *pAtIKU2*::*OsIKU2* transgenic lines **(A)** or five *pAtIKU2*::*GmIKU2* transgenic lines **(F)** in *Atiku2*. Letters indicate significant differences determined by ANOVA ($p < 0.001$). The panels below are images of rescued seeds in **(A)**. Scale bars, 0.4 mm. Expression of *OsIKU2* **(B)** and *GmIKU2* **(G)** in two transgenic lines from 1 to 6 DAP. Expression of *OsIKU2* or *GmIKU2* in each sample was normalized to that of *AtACT2* and *pAtIKU2*::*OsIKU2* line #6 or *pAtIKU2*::*GmIKU2* line #4 at 1 DAP. Expression of *OsIKU2* in *pAtIKU2*::*OsIKU2* line 6 or *GmIKU2* in *pAtIKU2*::*GmIKU2* line 4 has an artificial value of 1. Expression of *OsIKU2* in transgenic line 6 at 1 DAP was significantly different from that from 2 to 6 DAP by Student's t test ($p < 0.05$). Expression of *OsIKU2* in transgenic line 7 at 1 and 2 DAP was significantly different from that from 3 to 6 DAP by Student's t test ($p < 0.001$). Expression of *GmIKU2* in the two transgenic lines from 1 to 3 DAP was significantly different from that from 4 to 6 DAP by Student's t test ($p < 0.001$).

(C) *OsIKU2* expression in the developing rice caryopsis without chaff at various DAP. Expression of *OsIKU2* in each sample was normalized to that of *OsACT1* at 1 DAP. Expression of *OsIKU2* at 1 DAP was not significantly different from that from 3 to 20 DAP by Student's t test ($p > 0.05$).

(D) RNA *in situ* hybridization of *OsIKU2* in developing endosperm of Zhonghua 11 rice at various DAP. Red arrows indicate endosperm cells with hybridization signals. Yellow arrows indicate starch endosperm and aleurone endosperm cells without signals. Scale bars, 50 μm. al, aleurone endosperm; cv, central vacuole; en, endosperm; ne, nucellar epidermis; nu, nucellar tissue; pe, pericarp; se, starchy endosperm.

(E) ChIP analysis of histone H3K27me3 of the *OsIKU2* locus in spikelets without chaff at various DAP. Nine ChIP fragments were designed for the *OsIKU2* locus. Enrichment of DNA fragments was quantified by qPCR and normalized to that of *OsACT1* and –antibodies at 2 to 3 DAP. Enrichment of –antibodies at 2 to 3 DAP for each fragment has an artificial value of 1. *OsIKU2* fragments did not differ significantly from each other in enrichment at 2 to 3, 6 to 7, or 14 to 15 DAP by Student's t test ($p > 0.05$). *OsHd3a* ChIP analysis was performed with leaf tissue. Enrichment of two *OsHd3a* fragments +antibody was significantly different from that of –antibody at 6 to 7 DAP by Student's t test ($p < 0.01$). Data in **(A)**, **(B)**, **(C)**, **(E)**, **(F),** and **(G)** were calculated from three biological replicates. Error bars represent SEM, n = 3.

thereafter. *GmIKU2* expression was active from 1 to 6 DAP and then gradually declined from 7 to 18 DAP (Figure 6B). We isolated hexaploid endosperm nuclei at 2 to 4 DAP and 7 to 9 DAP (Supplemental Figure 14). At the *GmIKU2* locus, H3K27me3 marks were barely detected at 2 to 4 DAP but were highly enriched at 7 to 9 DAP at the TSS and the first exon (Figure 6C).

MEA exists only in the Brassicaceae family (Spillane et al., 2007; Luo et al., 2009), but there are many SWN or CLF homologs in a number of taxa (Supplemental Figure 15). The SET proteins from *Arabidopsis* and soybean clustered on one branch, whereas SETs from *Poales* were clustered on the other branch. Because of genome duplications, 75% of the genes in paleopolyploid soybean are present in multiple copies (Schmutz et al., 2010). Using the three *Arabidopsis* SET proteins, we identified eight soybean SET proteins. Two SET proteins aligned closely with AtCLF, and the other six aligned closely with AtSWN and AtMEDEA (Supplemental Figure 16).

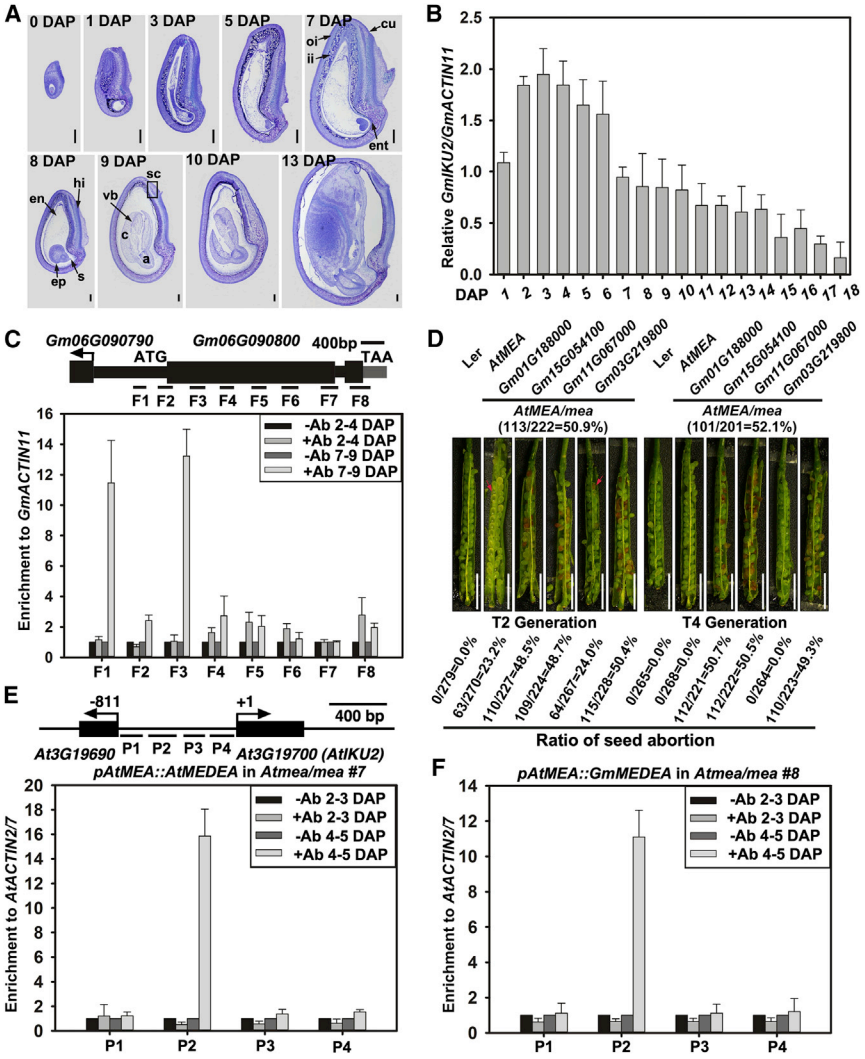

Figure 6. IKU2-MEA module exists in soybean.

(A) Soybean seed development at various DAP in paraffin sections. a, axis; c, cotyledon; cu, cuticle; en, endosperm; ent, endothelium; ep, embryo proper; hi, hilum; ii, inner integument; oi, outer integument; s, suspensor; sc, seed coat; vb, vascular bundle. Scale bars, 0.4 mm.

(B) *GmIKU2* expression in soybean seeds without pods at 1 to 18 DAP. Expression of *GmIKU2* in each sample was normalized to that of *GmACT11*. Expression of *GmIKU2* at 2 to 6 DAP was significantly different from that at 1 DAP and from 7 to 18 DAP by Student's t test ($p < 0.05$).

(C) ChIP analysis of histone H3K27me3 of the *GmIKU2* locus at various DAP. Eight ChIP fragments were designed for the *GmIKU2* locus. Enrichment of DNA fragments was quantified by qPCR and normalized to that of *GmACT11* and –antibodies at 2 to 4 DAP. Enrichment of –antibodies at 2 to 4 DAP for each fragment has an artificial value of 1. Enrichment of +antibody at 7 to 9 DAP was significantly different from that of – or +antibody at 2 to 4 DAP and –antibody at 7 to 9 DAP ($p < 0.001$) for fragments 1 and 3 and ($p < 0.05$) for fragments 2 and 5 by Student's t test.

(D) Rescue of the *Atmea* phenotype by *AtMEDEA* and *Gm11G067000*. All genes were driven by the *AtMEDEA* promoter, and seed phenotypes of the T2 and T4 generations were examined. Red arrows indicate defective seeds. Scale bars, 2 mm. Association of AtMEDEA:GFP **(E)** and GmMEDEA:GFP **(F)** with the *AtIKU2* promoter at 2 to 3 and 4 to 5 DAP. Four ChIP fragments were designed for the *AtIKU2* promoter. Enrichment of DNA fragments was quantified by qPCR and normalized to that of *AtACT2/7* and –antibodies at 2 to 3 DAP. Enrichment of –antibodies at 2 to 3 DAP for each fragment has an artificial value of 1. Enrichment of +antibody at 4 to 5 DAP was significantly different from that of – or +antibody at 2 to 3 DAP and –antibody at 4 to 5 DAP ($p < 0.001$) for fragment P2 by Student's t test. Data in **(B)**, **(C)**, **(E),** and **(F)** were calculated from three biological replicates. Error bars represent SEM, n = 3.

AtMEDEA is expressed mainly in the endosperm (Kiyosue et al., 1999; Spillane et al., 2007). We examined the expression of two CLF-like and six MEA/SWN-like soybean SET genes in leaves, pods, and seeds. The *AtCLF* homologs *Gm3G0188000* and *Gm15G054100* were expressed in all three organs (Supplemental Figure 17A). Of the *AtSWN*-like genes, *Gm10G012100*, *Gm11G067000*, and *Gm3G219800* were highly expressed in seeds compared with leaves. Three core subunits of the human PRC2 complex interact with each other to enhance their stabilization and activity (Kouznetsova et al., 2019). Only one of the three SET homologs, Gm11G067000, interacted with AtFIE, whereas both Gm11G067000 and Gm3G219800 interacted with AtFIS2 (Supplemental Figure 17B).

Mutations in *AtMEDEA* lead to autonomous endosperm development and seed lethality (Kiyosue et al., 1999). We introduced two *AtCLF*-like genes (*Gm1G188000* and *Gm15G054100*) and two *AtSWN*-like genes (*Gm3G219800* and *Gm11G06700*) under the control of a 650-bp *AtMEDEA* promoter to *Atmedea-3* (Figure 6D). If one of the four soybean SET proteins functions like AtMEA, it should rescue the endosperm phenotype and seed lethality rate of *Atmedea-3*. The transgenic plants of *pAtMEA*::*AtMEA* and *pAtMEA*::*GmMEA* in *Atmea/mea* were genotyped (Supplemental Figure 18). The defective rate of *AtMEA/mea* is about 50% because *AtMEA* is a maternally imprinted gene. In the T2 generation of *pAtMEA*::*AtMEA* in *AtMEA/mea* transgenic plants, both the transgene and the *AtMEA/mea* allele were heterozygous. The defective rate was 25% owing to their homozygous *Atmea* progeny. In the T4 generation, *Atmea* should be completely complemented. Reduced seed lethality rate of *Atmedea-3* was observed for T2 transgenic plants that carried either *AtMEA* or *Gm11G067000*, one of the four soybean SET proteins tested. In the T4 generation, *AtMEA* and *Gm11G067000* (*GmMEA*) fully complemented the *Atmedea-3* lethal phenotype at the expected expression level (Figure 6D, Supplemental Figure 19A and 19B). Therefore, we named Gm11G067000 GmMEA.

We next examined whether GmMEA is targeted to the *AtIKU2* locus using ChIP-PCR analysis. Both AtMEA and GmMEA proteins associated with the *AtIKU2* promoter in the P2 region at 4 to 5 DAP (Figure 6E and 6F). We also tested whether *Arabidopsis*

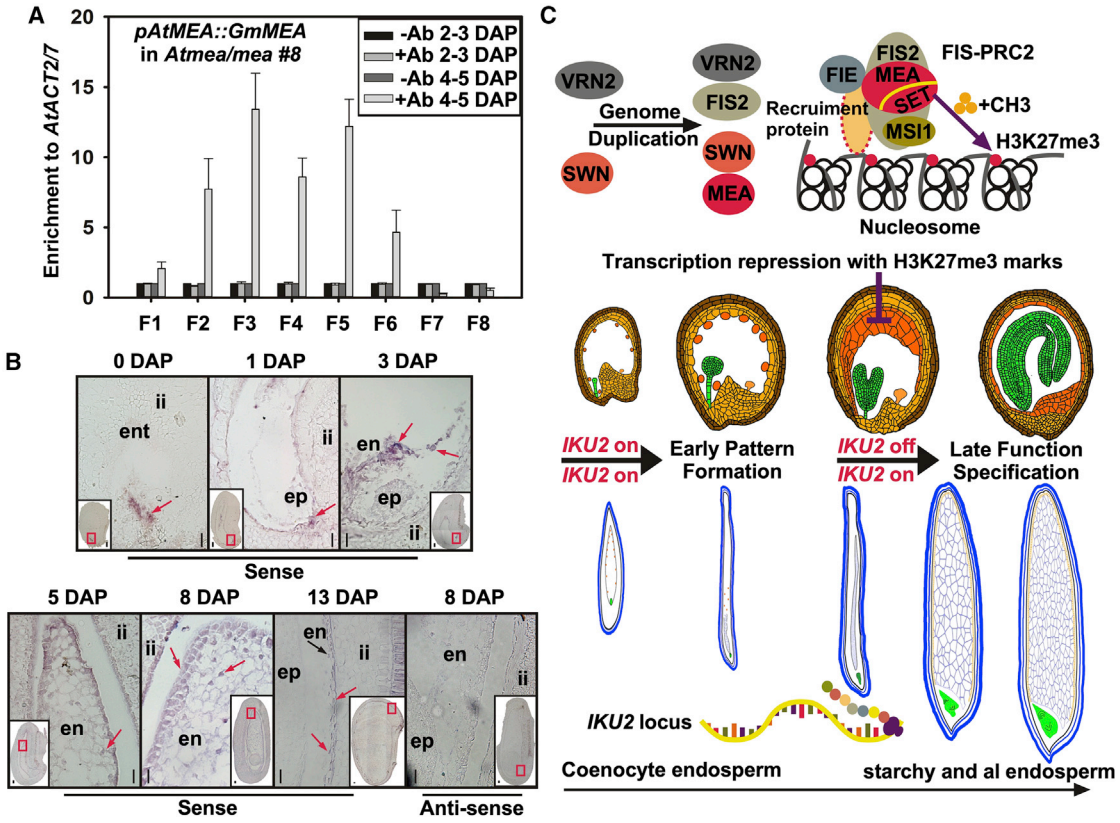

Figure 7. GmMEA restores trimethylation marks at the *IKU2* locus in *Atmea/mea*.

(A) ChIP analysis of histone H3K27me3 at the *AtIKU2* locus in *pAtMEA*::*GmMEA* transgenic line in the *Atmea/mea* background. Enrichment of DNA fragments was quantified by qPCR and normalized to that of *AtACT2/7* and –antibodies at 2 to 3 DAP. Enrichment of –antibodies at 2 to 3 DAP for each fragment has an artificial value of 1. Enrichment of +antibody at 4 to 5 DAP was significantly different from that of – or +antibody at 2 to 3 DAP and –antibody at 4 to 5 DAP ($p < 0.05$) for fragments 1 to 6 by Student's t test. Data were calculated from three biological replicates. All error bars indicate SEM, n = 3.

(B) *In situ* hybridization analysis for *GmMEA* expression in developing soybean seeds at various DAP. Red arrows indicate signals. Scale bars, 100 μm. en, endosperm; ent, endothelium; ep, embryo proper; ii, inner integument.

(C) A working model for two different seed developmental ontogenies associated with trimethylation marks on key endosperm proliferation genes such as *IKU2*. FIS2 and MEDEA may have evolved from VRN2 and SWN through a genome duplication event. al, aleurone.

and soybean MEA proteins restore H3K27 marks in *Atmea* complemented lines. No significant level of H3K27me3 marks was detected at the *AtIKU2* locus in *Atmea/mea* at either 2 to 3 or 4 to 5 DAP (Supplemental Figure 19C). Both *AtMEA* and *GmMEA* restored H3K27me3 marks at the *AtIKU2* locus (Figure 7A and Supplemental Figure 19D). In a series of *in situ* hybridization sections, *GmMEDEA* was expressed only before fertilization (0 DAP), during endosperm proliferation from 1 to 3 DAP, and at the cellularization stage from 5 to 8 DAP to 13 DAP (Figure 7B).

DISCUSSION

Embryo and endosperm in ancient cereals do not develop within a seed, and the ongoing endosperm does not invade the space that the embryo occupies. A directional increase in the ratio of embryo to endosperm is an evolutionary trend within the angiosperms (Stebbins, 1974; Takhtajan, 1991). Early divergent taxa have a small, underdeveloped, primitive embryo, whereas some advanced taxa have a larger and fully developed embryo. Seeds of some angiosperms such as cereals are filled almost entirely by endosperm. Seeds in other angiosperms such as *Arabidopsis* are completely dominated by the embryo and lack endosperm. Therefore, endosperm diversification becomes dominant for seed ontogeny. Independent evolution of seed morphogenesis and specific modification occurred within early flowering plants and after the divergences of the two other angiosperm lineages, monocots and eudicots (Sanderson and Donoghue, 1994).

In the Brassicaceae lineage, *Arabidopsis* endosperm and embryo within a central cavity approach an almost fully mature volume around the heart stage. Endosperm becomes the main reservoir for storing nutrients and providing energy for the growth of the embryo. The *Arabidopsis* IKU pathway regulates endosperm development; the *Atiku2* mutant exhibits precocious endosperm cellularization and produces only half the number of nuclei found in the wild type (Garcia et al., 2003; Luo et al., 2005). *IKU2s* from *Brachypodium*, rice, and soybean functionally complement the *Atiku2* seed phenotype (Figures 1 and 5). IKU2 has retained its original and ancestral function over millions of years of evolution based on a conserved IKU2 function during endosperm proliferation.

Cereal *IKU2* transcription is sustained for persistent endosperm development. Mutations in *BdIKU2* caused precocious cellularization, reduced the number of endosperm nuclei without altering cell size, and produced small seeds (Figure 3). Because IKU2 functions are conserved among the four species, the expression duration of *IKU2s* contributes to persistent or silenced endosperm proliferation. After the heart stage, the *Arabidopsis* embryo undergoes rapid growth and requires space. The ongoing endosperm structure must give way and perish so that embryos persist in the species and gene pool as the next plant generation. Prolonged expression of *BdIKU2* at 4 to 9 DAP in *Arabidopsis* seeds from *pBdIKU2*::*BdIKU2*/*Atiku2* transgenic plants produced undesirable free endosperm nuclei and continuous endosperm proliferation (Figure 2).

AtIKU2 expression is suppressed by the FIS-PRC2 complex, and H3K27me3 marks are enriched over the *AtIKU2* locus at 4 to 5 DAP (Figure 7C). AtMEA catalyzes the trimethylation of histone 3 lysine 27. Onset of cellularization and recruitment of FIS-PRC2 suppresses *AtIKU2* transcription and terminates endosperm proliferation. In either *mea iku2* or *fis iku2* double mutants, removal of *IKU2* from the *mea* mutant does not restore cellularization (Garcia et al., 2003; Luo et al., 2005). FIS-PRC2 may target an array of genes that includes *IKU2*. Because IKU2 is one of the FIS-PRC2 targets, loss of *IKU2* may be one of the causes of the endosperm phenotype of the *mea* mutant, as *IKU2* expression is increased in *mea*/*mea* (Figure 4A). In addition, *IKU2* expression driven by the *BdIKU2* promoter did not follow the normal *IKU2* expression pattern and caused endosperm over-proliferation that partially phenocopied *mea* (Figure 2D and 2E).

MEA homologs are unique to the Brassicaceae lineage (Spillane et al., 2007; Luo et al., 2009). By contrast, CLF and SWN homologs are found in a large variety of species and retain their ancient function. Some dicots appear to have experienced one or two more genome-wide duplication events, which enabled the evolution of many crucial genes (Soltis et al., 2008). SWN is also likely to be specific to Brassicaceae in terms of recent gene duplication events, as six soybean genes are similar to both AtSWN and AtMEA (Supplemental Figure 15). Whether the six soybean genes share the functions of AtSWN and AtMEA requires further tests. It is possible that SWN has retained an ancestral function, whereas AtMEA may have acquired a new function. In species with a similar developmental pattern, such as soybean, GmSWN plays an essential role like AtMEA (Figure 6). By contrast, *Brachypodium* and rice *IKU2* loci maintain active transcription for coenocytic endosperm proliferation in the early phase and starchy and aleurone endosperm proliferation in the late phase (Figure 7C). Although CLF and SWN are highly expressed from the vegetative stage to the reproductive stage (Tonosaki and Kinoshita, 2015), the *Brachypodium* and rice *IKU2* loci are devoid of H3K27me3 marks. Conserved IKU2 function but divergent MEA evolution in *Poales*, *Fabales*, and *Brassicales* may also account for evolutionary diversification among angiosperm taxa.

METHODS

Plant materials and growing conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) and var. Landsberg *erecta* (*Ler*) were used as the *Arabidopsis* wild types. The mutant lines *Atiku2-4* (SALK_073260) and *Atmedea-3* were described previously (Kiyosue et al., 1999; Zhou et al., 2009). All *Arabidopsis* plants were grown under a photoperiod of 16 h light and 8 h dark at 22°C and 80% relative humidity. The Bd21-3 accession of *Brachypodium distachyon* was used as the wild type and grown under conditions of 18 h light and 6 h dark at 21°C and 70% relative humidity. Soybean cv. Dongnong 50 was grown under conditions of 16 h light and 8 h dark and 60% relative humidity. *Japonica* rice and Zhonghua 11 were grown in a growth room with 16 h light and 8 h dark at 23°C and 80% relative humidity.

Phylogenetic analysis and multiple protein sequence alignment

Protein sequences and the corresponding coding sequences of the *IKU2*- and *E(z)*-like genes were retrieved from published genomes of representative species through BLAST searches using *Arabidopsis* sequences as queries (Supplemental Table 1). For each gene family, multiple sequence alignment was performed using the MAFFT program in PhyloSuite (Zhang et al., 2020a; 2020b) with manual adjustment. The codon-based nucleotide matrices corresponding to the protein alignments were generated by PAL2NAL (Suyama et al., 2006). The phylogenetic trees were constructed from matrices containing only alignable nucleotide sequences in IQ-TREE using the maximum-likelihood method (Nguyen et al., 2015). The best-fit substitution models detected by ModelFinder (Kalyaanamoorthy et al., 2017) were TVMe + I + G4 and GTR + F + I + G4 (based on the Bayesian Information Criterion) for *IKU2*- and *E(z)*-like genes, respectively. Branch support was estimated by performing 10 000 replicates of ultrafast bootstrapping (Hoang et al., 2018). The alignment matrix in fasta format and the tree topology in Newick format can be found in Supplemental Figures 1 and 15.

Multiple protein sequence alignment was performed using the ClustalW program, and a phylogenetic tree was constructed.

Plasmid construction and generation of transgenic plants

All primers used to make plasmid constructs are listed in Supplemental Table 3. The genes used to complement the *Atiku2* phenotype were cloned into a modified pCAMBIA1302 vector, which was reconstructed with Kpn1, Xba1, Apa1, and Swa1 sites. First, an *AtIKU2* promoter region from −806 to −1 was PCR-amplified from Col-0 genomic DNA and ligated to Kpn1 and Xba1 sites of the modified pCAMBIA1302 vector. Various *IKU2* genes were then PCR-amplified from wild-type genomic DNA of different species and ligated to the Apa1 and Swa1 sites of the vector with the *AtIKU2* promoter. All vectors used to complement *Atmedea* and *Bdiku2 #4-11* phenotypes were constructed the same way as with various *IKU2* genes, except that the *AtMEDEA* promoter was 650 bp from −650 to −1, and the *BdIKU2* promoter was 1200 bp from −1200 to −1. All other genomic gene sequences driven by their native promoters were PCR-amplified from various genomic DNA preparations and cloned into the pCR8/GW/TOPO TA Cloning entry vector. The TA fragments were then cloned into the pEARLYGATE303 vector or other destination vectors using the gateway system. All stable *Arabidopsis* transgenic plants were created by the floral dip method with *Agrobacterium tumefaciens* GV3101.

For *BdIKU2* CRISPR-Cas9 knockout, a *pPZP211-35S*::*Cas9*-derived vector was constructed from *pYLCRISPR*/*Cas9* PUbi-H and a single guide RNA expression cassette against the *BdIKU2* target sequence under the control of the *Brachypodium* U6 promoter by an overlapping PCR approach (Feng et al., 2013). The PCR products were ligated into the predigested binary vector with Mlu1 and Sal1. The resulting binary vectors were transformed into the *Agrobacterium tumefaciens* EHA105 strain and introduced to embryogenic callus generated from immature embryos (Alves et al., 2009). The selection markers were neomycin phosphotransferase II for the CRISPR-Cas9 system and hygromycin for *Bdiku2* rescue transgenic plants.

Full-length coding sequences of *AtMEDEA*, *AtCLF*, *AtSWN*, *AtFIE*, and *AtFIS2* were PCR-amplified from Col-0 cDNA. Full-length coding sequences of *Gm02G012100*, *Gm11G067000*, and *Gm03G219800* were PCR-amplified from cv. Dongnong 50 cDNA. All cDNAs were reverse transcribed from mRNA extracted from mixed *Arabidopsis* or soybean seeds at various DAP. All primers used are listed in Supplemental Table 3. The PCR products were cloned into the pCR8/GW/TOPO TA Cloning entry vector. The cDNAs were then recombined into the GAL4-DNA-binding domain vector pDEST-GBKT7 or the GAL4-activation domain vector pDEST-GADT7.

Isolation of *Brachypodium* endosperm tissue, RNA extraction, and qRT-PCR analysis

Brachypodium endosperm tissue from 1 to 3 DAP was harvested from the outlined area shown in Supplemental Figure 9 under a microscope. The nucellus and tiny embryo were carefully stripped off. Endosperm tissue isolated from 1- to 3-DAP seeds had more maternal and embryo contamination. After 5 to 7 DAP, it was much easier to separate the endosperm tissue from other tissues, with less maternal and embryo contamination. Total RNA was isolated using the GeneJET RNA Purification Mini Kit (Thermo Scientific). Two micrograms of total RNA were digested using DNase I (Thermo Scientific), and first-strand cDNA was synthesized using the SuperScript III reverse transcriptase kit (Invitrogen). qRT-PCR was performed with SYBR Premix Ex Taq II (Takara) on an Applied Biosystems QuantStudio 6 Flex Real-Time PCR machine (Thermo Scientific). Data were generated from three biological replicates. The primers used for quantitative RT-PCR are listed in Supplemental Table 3.

DIC microscopy

Arabidopsis plants were hand-pollinated and harvested from 1 to 9 DAP. Developing seeds were fixed in FAA (10% formalin, 5% acetic acid, 45% ethanol, and 0.01% Triton X-100) for 1 h and washed with a graded ethanol series (30%, 50%, 70%, 80%, 90%, 95%, 100%, and 100%). The developing seeds from 1 to 5 DAP were then incubated in Hoyer buffer (3 mL glycerol, 1 mL water, and 8 g chloral hydrate). Herr solution (10 mL phenol, 10 mL 85% lactic acid, 5 mL xylene, 10 mL clove oil, and 10 g chloral hydrate) was used for developing seeds from 6 to 9 DAP. Seeds with cleared embryos were photographed under DIC optics using an Olympus BX51 microscope (Waltham, MA) with an Olympus DP26 CCD camera.

Thin sections for histological analysis

Arabidopsis siliques and *Brachypodium* caryopses were harvested at various DAP. The materials were fixed in a 50% ethanol, 5% formaldehyde, and 10% acetic acid solution for 12 to 48 h at 4°C. The fixed materials were then dehydrated in a series of ethanol dilutions (60%, 75%, 85%, 95%, 100%, and 100%). Each step was repeated twice for 30 min each time, and the final dilution was left in 100% ethanol overnight. Seed materials were stepwise infiltrated in Technovit 7100 (Kulzer) according to the manufacturer's instructions. The embedded seeds were sectioned into 2-μm sections using a Leica RM2255 microtome, and sections were stained with 0.05% toluidine blue and dried on a heating plate at 55°C. Stained sections were photographed using a Cool Cam color CCD camera (Cool Camera) coupled to an Olympus BH-2 microscope.

Paraffin sections and *in situ* hybridization assay

Arabidopsis seeds, *Brachypodium* caryopses, and soybean seeds were fixed in formalin-acetic acid-alcohol (FAA; 5% formaldehyde, 5% acetic acid, and 50% ethanol). After 48 h of fixation at 4°C, a graded ethanol series (70%, 80%, 90%, 95%, and 100%) was used to dehydrate the seeds; each step took 60 min, and the samples were finally left in 100% ethanol overnight at 4°C. After washing with xylenes and pre-incubation with paraffin for 2 days, the seeds were embedded in paraffin and sectioned to 8 mm using a Leica RM2125 Microtome. For studies of soybean seed developmental stages, paraffin sections were dewaxed and photographed using an Olympus BX53 microscope with an Olympus DP26 CCD camera.

For *in situ* hybridization, *Arabidopsis*, *Brachypodium,* and soybean sections were dewaxed and hybridized with sense and antisense digoxigenin-labeled RNA probes at 55°C overnight. The probes were synthesized with a digoxigenin RNA labeling kit. Antibody incubation and color detection were performed according to the manufacturer's instructions (DIG nucleic acid detection kit; Boehringer Mannheim).

Isolation of hexaploid endosperm nuclei

FACS was used to isolate endosperm nuclei (Zheng and Gehring 2019). *Arabidopsis* siliques were harvested after hand pollination and stripped of silique walls. Soybean seeds without pods and *Brachypodium* caryopses were harvested at different developmental stages. Seeds were freshly ground (avoiding flash-freezing if possible), and unbroken nuclei were extracted with the Partec CyStain UV Precise P kit (Sysmex America). Nuclei with different DNA ploidy contents were separated by their DAPI staining intensity into various peaks indicative of different ploidies. Peaks of 3C and 6C for *Arabidopsis* and soybean and peaks of 3C, 6C, 12C, and 24C for *Brachypodium* were sorted using a BD FACSAria II (BD) machine. About 200 000 nuclei were isolated for histone modification assays.

ChIP-PCR

Material for regular *Arabidopsis* ChIP and sorted nuclei for histone modification were fixed with 1% formaldehyde, and cross-linking was stopped with 2 M glycine (Sun et al., 2019). About 200 000 sorted 6C nuclei were used for ChIP analysis. After sonication, the chromatin complexes were incubated with the respective antibodies or anti-H3K27me3 (EMD Millipore) for 6 to 12 h and purified with protein A beads (Abcam). For ChIP analysis of AtMEA:GFP or GmMEA:GFP to the *AtIKU2* locus, we used a 1:1 ratio of monoclonal (HT801-01, TransGen) and polyclonal (A01388-40, GenScript) anti-GFP antibodies. Chromatin complexes incubated with BSA were used as negative controls. After reversal of the cross-linking and treatment with proteinase K and RNase, immunoprecipitated DNA was quantified by qRT-PCR. All primers used are listed in Supplemental Table 3. Fold enrichment for various specific chromatin fragments was normalized to the enrichment level of the *Arabidopsis AtACTIN2/7* amplicon, *Brachypodium BdACTIN3* amplicon, or soybean *GmACTIN11* amplicon. Enrichment of each amplicon was calculated using the following equation: $2^{-(Ct\ Fragment\ ChIP - Ct\ Fragment\ MOCK)} / 2^{-(Ct\ ACTIN\ ChIP - Ct\ ACTIN\ MOCK)}$ and was normalized to the negative control (immunoglobulin G) of each fragment.

Protein isolation and immunoblotting

To detect histone 3 accumulation and H3K27me3 or H3K4me3 marks in *B. distachyon* caryopses, developing caryopses without chaff or pericarp were harvested at 2 to 3, 6 to 7, and 14 to 15 DAP. Total protein was isolated using IP buffer that contained 25 mM Tris-HCl (pH 8.0), 1 mM EDTA, 150 mM NaCl, 10% glycerol, 0.1% Triton X-100, 2 mM β-mercaptoethanol, 1 mM PMSF, and 1× protease inhibitors (Roche). The protein extract was incubated with anti-H3 antibodies (Beyotime, AF0009) for 3 h. The beads were washed three times with IP buffer, boiled in 40 μL 2× SDS buffer, and separated by 15% SDS-PAGE. Western blots were probed using the respective anti-H3, anti-H3K27me3 (Millipore, #07–449), and anti-H3K4me3 (Millipore, #07–473) antibodies at a 1:5000 dilution.

Yeast two-hybrid assays

Both pDEST-GBKT7 and pDEST-GADT7 vectors were co-transformed into yeast strain Y2HGold using a lithium acetate-based transformation protocol. Positive colonies were selected on SD/-Trp/-Leu (-TL) medium. Interaction assays were performed on selective SD/-Trp/-Leu/-His/-Ade medium (-TLHA) at 30°C for 3 days.

AtMEA/*Atmea* genotyping

AtMEA/*Atmea* was genotyped by PCR and FspBI restriction digestion. PCR products digested with this enzyme were 611 bp for the *AtMEA allele* and 508 bp and 103 bp for the *Atmea allele*.

Sequence analysis of *AtIKU2*, *GmIKU2*, *BdIKU2*, and *OsIKU2* promoters

The entire 811-bp upstream region of *AtIKU2* genes and the 1500-bp upstream region of *GmIKU2*, *BdIKU2*, and *OsIKU2* genes were used as query promoter sequences to predict possible *cis*-acting elements with the PLACE motif database (https://www.dna.affrc.go.jp/PLACE/?action=newplace).

ACCESSION NUMBERS

Sequence data from this article can be found in the EMBL/GenBank data libraries under accession numbers NC_003074.8 *At3G19700* (*AtIKU2*), XM_003579052.4 *Bd4G00900* (*BdIKU2*), XM_003526471.4 *Gm06G090800* (*GmIKU2*), XM_015764749.2 *LOC_Os12g43640* (*OsIKU2*), NC_003071.7 *At2G23380* (*AtCLF*), NC_003075.7 *At4G02020* (*AtSWN*), NC_003070.9 *At1G02580* (*AtMEDEA*), and KRH28654.1 *Gm11G067000* (*GmMEDEA*).

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

FUNDING

We thank Professor Ramin Yadegari at the University of Arizona for *Atmedea-3* seeds. We thank Professor Hongzhi Kong at the University of the Chinese Academy of Sciences for *IKU2* and *MEDEA* phylogenetic tree analysis. This work was supported by grants from the National Natural Science Foundation of China 32071921 (X.Y.Z.), National Key Research and Development Program 2021YFF1001203 (D.J.Z.), National Natural Science Foundation of China 31730008 (X.S.Z.), and National Science Foundation IOS-1933291 (M.N.).

AUTHOR CONTRIBUTIONS

S.W., X.Y.Z., X.S.Z., and M.N. conceptualized and designed the research plan. D.W., Y.W., H.Z., G.X., J.L., and D.J.Z. performed the experiments. D.W. and M.N. wrote the manuscript. X.S.Z. reviewed and edited the manuscript.

ACKNOWLEDGMENTS

We declare no conflicts of interest.

Received: May 13, 2022
Revised: July 19, 2022
Accepted: September 2, 2022
Published: September 7, 2022

REFERENCES

Adrian, J., Farrona, S., Reimer, J.J., Albani, M.C., Coupland, G., and Turck, F. (2010). *Cis*-Regulatory elements and chromatin state coordinately control temporal and spatial expression of *FLOWERING LOCUS T* in *Arabidopsis*. Plant Cell **22**:1425–1440.

Alves, S.C., Worland, B., Thole, V., Snape, J.W., Bevan, M.W., and Vain, P. (2009). A protocol for Agrobacterium-mediated transformation of *Brachypodium distachyon* community standard line *Bd21*. Nat. Protoc. **4**:638–649.

Berger, F. (2003). Endosperm, the crossroad of seed development. Curr. Opin. Plant Biol. **6**:42–50.

Black, M., Bewley, J.D., and Halmer, P. (2006). The encyclopaedia of seeds: science, technology and uses. Ann. Bot. **100**:1379.

Chamberlin, M.A., Horner, H.T., and Palmer, R.G. (1993). Nuclear size and DNA content of the embryo and endosperm during their initial stages of development in Glycine max (Fabaceae). Am. J. Bot. **80**:1209–1215.

Chaudhury, A.M., Ming, L., Miller, C., Craig, S., Dennis, E.S., and Peacock, W.J. (1997). Fertilization-independent seed development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U. S. A. **94**:4223–4228.

Cheng, X., Pan, M., E, Z.G., Zhou, Y., Niu, B., and Chen, C. (2020). The maternally expressed polycomb group gene *OsEMF2a* is essential for endosperm cellularization and imprinting in rice. Plant Commun. **2**:100092.

Dante, R.A., Larkins, B.A., and Sabelli, P.A. (2014). Cell cycle control and seed development. Front. Plant Sci. **5**:493.

Diez, C.M., Roessler, K., and Gaut, B.S. (2014). Epigenetics and plant genome evolution. Curr. Opin. Plant Biol. **18**:1–8.

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y., et al. (2013). Efficient genome editing in plants using a CRISPR/Cas system. Cell Res. **23**:1229–1232.

Figueiredo, D.D., and Köhler, C. (2014). Signalling events regulating seed coat development. Biochem. Soc. Trans. **42**:358–363.

Floyd, S., and Friedman, W. (2000). Evolution of endosperm developmental patterns among basal angiosperms. Int. J. Plant Sci. **161**:57–81.

Floyd, S.K., Lerner, V.T., and Friedman, W.E. (1999). A developmental and evolutionary analysis of embryology in *Platanus* (Platanaceae), a basal eudicot. Am. J. Bot. **86**:1523–1537.

Forbis, T.A., Floyd, S.K., and de Queiroz, A. (2002). The evolution of embryo size in angiosperm and other seed plant. Evolution **56**:2112–2125.

Friedman, W.E. (1994). The evolution of embryogeny in seed plants and the developmental origin and early history of endosperm. Am. J. Bot. **81**:1468–1486.

Garcia, D., Saingery, V., Chambrier, P., Mayer, U., Jürgens, G., and Berger, F. (2003). Arabidopsis haiku mutants reveal new controls of seed size by endosperm. Plant Physiol. **131**:1661–1670.

Goldberg, R.B., Barker, S.J., and Perez-Grau, L. (1989). Regulation of gene expression during plant embryogenesis. Cell **56**:149–160.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G. (1997). A polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. Nature **386**:44–51.

Grossniklaus, U., Vielle-Calzada, J.P., Hoeppner, M.A., and Gagliano, W.B. (1998). Maternal control of embryogenesis by *MEDEA*, a Polycomb group gene in *Arabidopsis*. Science **280**:446–450.

Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. **35**:518–522.

Huang, J., Zhang, A., Ho, T.T., Zhang, Z., Zhou, N., Ding, X., Zhang, X., Xu, M., and Mo, Y.Y. (2016). Linc-RoR promotes c-Myc expression through hnRNP I and AUF1. Nucleic Acids Res. **44**:3059–3069.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods **14**:587–589.

Kang, X., Li, W., Zhou, Y., and Ni, M. (2013). A WRKY transcription factor recruits the SYG1-like protein SHB1 to activate gene expression and seed cavity enlargement. PLoS Genet. **9**:e1003347.

Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J., Wells, D., Katz, A., Margossian, L., Harada, J.J., Goldberg, R.B., et al. (1999). Control of fertilization-independent endosperm development by the *MEDEA* polycomb gene in *Arabidopsis*. Proc. Natl. Acad. Sci. U. S. A. **96**:4186–4191.

Kouznetsova, V.L., Tchekanov, A., Li, X., Yan, X., and Tsigelny, I.F. (2019). Polycomb repressive 2 complex-Molecular mechanisms of function. Protein Sci. **28**:1387–1399.

Laugesen, A., and Helin, K. (2014). Chromatin repressive complexes in stem cells, development, and cancer. Cell Stem Cell **14**:735–751.

Le, B.H., Wagmaister, J.A., Kawashima, T., Bui, A.Q., Harada, J.J., and Goldberg, R.B. (2007). Using genomics to study legume seed development. Plant Physiol. **144**:562–574.

Li, J., Nie, X., Tan, J.L.H., and Berger, F. (2013). Integration of epigenetic and genetic controls of seed size by cytokinin in. Proc. Natl. Acad. Sci. U. S. A. **110**:15479–15484.

Luo, M., Bilodeau, P., Koltunow, A., Dennis, E.S., Peacock, W.J., and Chaudhury, A.M. (1999). Genes controlling fertilization-independent seed development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U. S. A. **96**:296–301.

Luo, M., Dennis, E.S., Berger, F., Peacock, W.J., and Chaudhury, A. (2005). *MINISEED3* (*MINI3*), a *WRKY* family gene, and *HAIKU2* (*IKU2*), a leucine-rich repeat (*LRR*) *KINASE* gene, are regulators of seed size in *Arabidopsis*. Proc. Natl. Acad. Sci. U. S. A. **102**:17531–17536.

Luo, M., Platten, D., Chaudhury, A., Peacock, W.J., and Dennis, E.S. (2009). Expression, imprinting, and evolution of rice homologs of the polycomb group genes. Mol. Plant **2**:711–723.

Man, J., Gallagher, J.P., and Bartlett, M. (2020). Structural evolution drives diversification of the large LRR-RLK gene family. New Phytol. **226**:1492–1505.

Marinos, N.G. (1970). Embryogenesis of the pea (Pisum sativum). Protoplasma **70**:261–279.

Miyake, T., Takebayashi, N., and Wolf, D.E. (2009). Possible diversifying selection in the imprinted gene, *MEDEA*, in *Arabidopsis*. Mol. Biol. Evol. **26**:843–857.

Moreno-Romero, J., Jiang, H., Santos-González, J., and Köhler, C. (2016). Parental epigenetic asymmetry of PRC2-mediated histone modifications in the Arabidopsis endosperm. EMBO J. **35**:1298–1311.

Mozgova, I., Köhler, C., and Hennig, L. (2015). Keeping the gate closed: functions of the polycomb repressive complex PRC2 in development. Plant J. **83**:121–132.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol. Biol. Evol. **32**:268–274.

Nowicka, A., Kovacik, M., Tokarz, B., Vrána, J., Zhang, Y., Weigt, D., Doležel, J., and Pecinka, A. (2020). Dynamics of endoreduplication in developing barley seeds. J. Exp. Bot. **72**:268–282.

Ohad, N., Margossian, L., Hsu, Y.C., Williams, C., Repetti, P., and Fischer, R.L. (1996). A mutation that allows endosperm development without fertilization. Proc. Natl. Acad. Sci. U. S. A. **93**:5319–5324.

Olsen, O.A. (2001). Endosperm development: cellularization and cell fate specification. Annu. Rev. Plant Physiol. Plant Mol. Biol. **52**:233–267.

Olsen, O.-A. (2004). Nuclear endosperm development in cereals and *Arabidopsis thaliana*. Plant Cell Suppl. **16**:S214–S227.

Opanowicz, M., Hands, P., Betts, D., Parker, M.L., Toole, G.A., Mills, E.N.C., Doonan, J.H., and Drea, S. (2011). Endosperm development in Brachypodium distachyon. J. Exp. Bot. **62**:735–748.

Orkin, S.H., and Hochedlinger, K. (2011). Chromatin connections to pluripotency and cellular reprogramming. Cell **145**:835–850.

Sabelli, P.A., and Larkins, B.A. (2009). The development of endosperm in grasses. Plant Physiol. **149**:14–26.

Sanderson, M.J., and Donoghue, M.J. (1994). Shifts in diversification rate with the origin of angiosperms. Science **264**:1590–1593.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., et al. (2010). Genome sequence of the palaeopolyploid soybean. Nature **463**:178–183.

Soltis, D.E., Bell, C.D., Kim, S., and Soltis, P.S. (2008). Origin and early evolution of angiosperm. Ann. N. Y. Acad. Sci. **1133**:3–25.

Sørensen, M.B., Mayer, U., Lukowitz, W., Robert, H., Chambrier, P., Jürgens, G., Somerville, C., Lepiniec, L., and Berger, F. (2002). Cellularisation in the endosperm of *Arabidopsis thaliana* is coupled to mitosis and shares multiple components with cytokinesis. Development **129**:5567–5576.

Spillane, C., Schmid, K.J., Laoueillé-Duprat, S., Pien, S., Escobar-Restrepo, J.M., Baroux, C., Gagliardini, V., Page, D.R., Wolfe, K.H., and Grossniklaus, U. (2007). Positive darwinian selection at the imprinted *MEDEA* locus in plants. Nature **448**:349–352.

Sreenivasulu, N., Borisjuk, L., Junker, B.H., Mock, H.P., Rolletschek, H., Seiffert, U., Weschke, W., and Wobus, U. (2010). Barley grain development: toward an integrative view. Int. Rev. Cell Mol. Biol. **281**:49–89.

Stebbins, G. (1974). Flowering Plants: Evolution above the Species Level (Cambridge: Belknap Press).

Sun, X., Shantharaj, D., Kang, X., and Ni, M. (2010). Transcriptional and hormonal signaling control of Arabidopsis seed development. Curr. Opin. Plant Biol. **13**:611–620.

Sun, Q., Wang, S., Xu, G., Kang, X., Zhang, M., and Ni, M. (2019). SHB1 and CCA1 interaction desensitizes light responses and enhances thermomorphogenesis. Nature Commun. **10**:3110. https://doi.org/10.1038/s41467-019-11071-6.

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. **34**:W609–W612.

Takhtajan, A. (1991). Evolutionary Trends in Flowering Plants (New York: Columbia University Press).

Tamaki, S., Matsuo, S., Wong, H.L., Yokoi, S., and Shimamoto, K. (2007). Hd3a protein is a mobile flowering signal in rice. Science **316**:1033–1036.

Tonosaki, K., and Kinoshita, T. (2015). Possible roles for polycomb repressive complex 2 in cereal endosperm. Front. Plant Sci. **6**:144.

Vandepoele, K., Vlieghe, K., Florquin, K., Hennig, L., Beemster, G.T.S., Gruissem, W., Van de Peer, Y., Inzé, D., and De Veylder, L. (2005). Genome-wide identification of potential plant E2F target genes. Plant Physiol. **139**:316–328.

Wang, A., Garcia, D., Zhang, H., Feng, K., Chaudhury, A., Berger, F., Peacock, W.J., Dennis, E.S., and Luo, M. (2010). The VQ motif protein IKU1 regulates endosperm growth and seed size in Arabidopsis. Plant J. **63**:670–679.

Welchen, E., and Gonzalez, D.H. (2005). Differential expression of the Arabidopsis cytochrome c genes *Cytc-1* and *Cytc-2*. Evidence for the involvement of TCP-domain protein-binding elements in anther- and meristem-specific expression of the *Cytc-1* gene. Plant Physiol. **139**:88–100.

Wu, L., Liu, D., Wu, J., Zhang, R., Qin, Z., Liu, D., Li, A., Fu, D., Zhai, W., and Mao, L. (2013). Regulation of *FLOWERING LOCUS T* by a microRNA in *Brachypodium distachyon*. Plant Cell **25**:4363–4377.

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J.D., Pai, A., Song, C., Zhuang, M., Klasfeld, S., He, C., et al. (2017). Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in *Arabidopsis*. Nat. Genet. **49**:1546–1552.

Yuva-Aydemir, Y., Almeida, S., Krishnan, G., Gendron, T.F., and Gao, F.B. (2019). Transcription elongation factor AFF2/FMR2 regulates expression of expanded GGGGCC repeat-containing C9ORF72 allele in ALS/FTD. Nat. Commun. **10**:5466.

Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W.X., and Wang, G.T. (2020a). PhyloSuite: an integrated and scalable desktop platform

for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol. Ecol. Resour. **20**:348–355.

Zhang, S., Wang, D., Zhang, H., Skaggs, M.I., Lloyd, A., Ran, D., An, L., Schumaker, K.S., Drews, G.N., and Yadegari, R. (2018). FERTILIZATION-INDEPENDENT SEED-polycomb repressive complex 2 plays a dual role in regulating type I MADS-box genes in early endosperm development. Plant Physiol. **177**:285–299.

Zhang, Y.Z., Yuan, J., Zhang, L., Chen, C., Wang, Y., Zhang, G., Peng, L., Xie, S.S., Jiang, J., Zhu, J.K., et al. (2020b). Coupling of H3K27me3 recognition with transcriptional repression through the BAH-PHD-CPL2 complex in *Arabidopsis*. Nat. Commun. **11**:6212.

Zheng, X.Y., and Gehring, M. (2019). Low-input chromatin profiling in Arabidopsis endosperm using CUT&RUN. Plant Reprod. **32**:63–75.

Zhong, J., Peng, Z., Peng, Q., Cai, Q., Peng, W., Chen, M., and Yao, J. (2018). Regulation of plant height in rice by the Polycomb group genes *OsEMF2b, OsFIE2* and *OsCLF*. Plant Sci. **267**:157–167.

Zhou, Y., Zhang, X., Kang, X., Zhao, X., Zhang, X., and Ni, M. (2009). SHORT HYPOCOTYL UNDER BLUE1 associates with *MINISEED3* and *HAIKU2* promoters in vivo to regulate *Arabidopsis* seed development. Plant Cell **21**:106–117.