Coupling of H3K27me3 recognition with transcriptional repression through the BAH-PHD-CPL2 complex in Arabidopsis

Yi-Zhe Zhang1,2,6, Jianlong Yuan1,2,6, Lingrui Zhang3,6, Chunxiang Chen1, Yuhua Wang1, Guiping Zhang1, Li Peng1, Si-Si Xie1,2, Jing Jiang4, Jian-Kang Zhu1,3✉, Jiamu Du5✉ & Cheng-Guo Duan1,4✉

Histone 3 Lys 27 trimethylation (H3K27me3)-mediated epigenetic silencing plays a critical role in multiple biological processes. However, the H3K27me3 recognition and transcriptional repression mechanisms are only partially understood. Here, we report a mechanism for H3K27me3 recognition and transcriptional repression. Our structural and biochemical data showed that the BAH domain protein AIPP3 and the PHD proteins AIPP2 and PAIPP2 cooperate to read H3K27me3 and unmodified H3K4 histone marks, respectively, in Arabidopsis. The BAH-PHD bivalent histone reader complex silences a substantial subset of H3K27me3-enriched loci, including a number of development and stress response-related genes such as the RNA silencing effector gene ARGONAUTE 5 (AGO5). We found that the BAH-PHD module associates with CPL2, a plant-specific Pol II carboxyl terminal domain (CTD) phosphatase, to form the BAH-PHD-CPL2 complex (BPC) for transcriptional repression. The BPC complex represses transcription through CPL2-mediated CTD dephosphorylation, thereby causing inhibition of Pol II release from the transcriptional start site. Our work reveals a mechanism coupling H3K27me3 recognition with transcriptional repression through the alteration of Pol II phosphorylation states, thereby contributing to our understanding of the mechanism of H3K27me3-dependent silencing.
In eukaryotic cells, the N-terminal histone tails undergo numerous posttranslational modifications (PTMs). Histone modification acts as a mark to specify the chromatin status, as well as potential functional indications. The deposition, recognition, and removal of specific histone PTMs are dynamically regulated by different proteins or protein complexes called “writer”, “reader”, and “eraser” modules, respectively. The reader module can specifically recognize certain histone mark in both sequence- and modification-specific manners, and subsequently transmits the signal to downstream effectors. As a repressive epigenetic mark localized in euchromatin, deposition of trimethylation on histone H3 lysine 27 (H3K27me3) has been observed in many important functional genes. The polycomb repressive complexes (PRCs), consisting of a different polycomb group (PcG) of proteins, have been shown to be involved in the deposition and downstream action of H3K27me3 mark. Different PcG proteins associate to form two functionally distinct complexes, PRC1 and PRC2. The PRC1 complex has E3 ligase activity which has been shown to catalyze the monoubiquitination of histone H2A at lysine (H2Aub1), and the PRC2 complex catalyzes H3K27me2 and H3K27me3. Three major models have been proposed to explain the mechanisms of PRC complex-mediated transcription repression. For these bivalent promoters marked by both H3K27me3 and H3K4me3 marks, PcG complexes are believed to hold the poised Pol II at the transcription start site (TSS), resulting in the inhibition of Pol II release. Alternatively, PcG complexes can alter the chromatin environment by inducing chromatin condensation, thereby blocking the accessibility of chromatin remodeling complexes that is required for transcription activation. The histone PTMs might directly prevent Pol II processivity during transcription elongation. For example, studies in Drosophila have indicated that H3K27me3 could limit Pol II recruitment to gene promoters. H2Aub1 has been implicated in restraining Pol II elongation. However, the detailed mechanisms through which H3K27me3 reading is connected to transcriptional repression are not fully understood.

Histone mark recognition in plants is generally similar to that in animals, but sometimes possesses plant-specific mark–reader pairs. The PHD and BAH domains are two types of histone-binding domains in eukaryotes. The PHD finger has been reported to recognize methylated/unmethylated H3K4 marks and lysine acetylation marks, and the BAH domain can bind distinct histone marks, including H3K9me2 (ref. 2), H4K20me2 (ref. 21), unmodified H3K4 (ref. 22), nucleosome core particle, and the more recently identified H3K27me3 (ref. 23–25). In plants and animals, a large number of development and environmental response-related processes are subjected to H3K27me3-dependent regulation. Among them, flowering control has been a paradigmatic model for PRC complexes-mediated transcriptional repression in plants. The H3K27me3 dynamics in the flowering repressor gene FLOWERING LOCUS C (FLC) and the flowering gene FLOWERING LOCUS T (FT) play essential roles in the flowering time control, and the H3K27me3 regulators influence flowering time in different ways. Here, we demonstrated that the BAH domain-containing protein anti-silencing 1 (AS1)-IMMUNOPRECIPITATED PROTEIN 3 (AIPP3) and two PHD domain-containing proteins AIPP2/PARALOG OF AIPP2 (PAIPP2) could form a BAH−PHD module to read H3K27me3 and unmethylated H3K4, respectively, and coordinate in implementing transcriptional repression of hundreds of genes, particularly those development and stress-responsive genes in Arabidopsis, such as the flowering gene FT and the RNA silencing effector gene AGOS. Moreover, our structural and biochemical studies further revealed the molecular basis for the specific recognition of these histone marks. We also revealed that the BAH−PHD module represses the release of Pol II from TSS regions by cooperating with CPL2, a known plant-specific Pol II carboxyl terminal domain (CTD)-Ser5 phosphatase. Collectively, our findings reveal a coupling of the H3K27me3 recognition and downstream transcriptional repression through the BPC complex. This pathway may represent a mechanism of H3K27me3-mediated gene silencing.

### Results

**BAH protein AIPP3 associates with PHD proteins and CPL2 to form a protein complex in Arabidopsis.** Through a mass spectrometry (MS) analysis of the chromatin regulator ASI1, we previously demonstrated that BAH domain-containing protein AIPP3, and PHD protein AIPP2 and CPL2 are associated with ASI1 (refs. 41,42). AIPP2 is known to interact with AIPP3 and CPL2 (ref. 41). This association was further confirmed by immunoprecipitation assays coupled to a mass spectrometry analysis (IP−MS) of AIPP3, AIPP2, and CPL2 in which AIPP3, AIPP2, and CPL2 could be mutually co-purified with one another except for ASI1 (Fig. 1a and Supplementary Data 1). Interestingly, another PHD protein encoded by AT3G16680 (Supplementary Fig. 1), the closest paralog of AIPP2 (hereafter referred to as PAIPP2) in Arabidopsis, was co-purified with AIPP3 and CPL2, and AIPP3 and CPL2 were also present in the IP−MS of PAIPP2 (Fig. 1a). The yeast two-hybrid (Y2H) and split luciferase assays indicated that PAIPP2 could also interact with AIPP3 and CPL2, but not with AIPP2 (Fig. 1b and Supplementary Fig. 2). Regarding the domain requirements for protein interactions, the Y2H results indicated that the BAH domain-containing N-terminus of AIPP3 and the RBM motif-containing C-terminus region of CPL2 are required for their interactions with AIPP2 and PAIPP2, respectively (Supplementary Fig. 3). AIPP2 and PAIPP2 were divided into three parts: the N-terminus (N), PHD that is followed by a frequently associated polybasic region (PHD−PBR), and the C-terminus (C) (Supplementary Fig. 4). The PHD−PBR part is dispensable for AIPP2/PAIPP2–AIPP3 interactions. Intriguingly, the PHD−PBR interaction with AIPP3 could be strengthened, and inhibited by the N and C termini of AIPP2/PAIPP2, respectively (Supplementary Fig. 4), indicating the presence of an intramolecular regulation mode within AIPP2 and PAIPP2. The C termini of AIPP2 and PAIPP2 were fully responsible for the interaction with CPL2. Thus, we reasoned that the PHD proteins AIPP2 and PAIPP2 interact with AIPP3 and CPL2 independently to associate in vivo (Fig. 1c). Moreover, the gel filtration assay that was performed using epitope-tagged transgenic lines indicated that these four proteins co-eluted in the same fractions (Fig. 1d). Thus, these data support the notion that the BAH protein AIPP3 associates with two PHD proteins and CPL2 to form a protein complex in vivo (which are hereafter referred to as the BAH−PHD−CPL2 complex or the BPC complex).

**The BPC complex represses flowering by inhibiting FT expression.** By generating native promoter-driven β-glucuronidase (GUS) reporter transgenes, we observed that the BPC complex genes were ubiquitously expressed in Arabidopsis (Supplementary Fig. 5). To explore the biological function of the BPC complex, the null mutants of PAIPP2 were generated using CRISPR/Cas9-mediated gene editing (Supplementary Fig. 6), and the morphological and flowering phenotypes of the generated paipp2-1 mutant, as well as the reported aipp3-1, aipp2-1, and cpl2-2 mutants were investigated. The aipp3-1, cpl2-2, and paipp2-1-aipp2-1 displayed multiple developmental defects, such as a dwarfed size, small leaves, and poor fertility (Fig. 2a). Instead, the aipp2-1 and paipp2-1 mutants only showed mild...
displayed similar early flowering compared with the colorectal mutant flc-3 single mutants. Unfortunately, we failed to obtain the aipp3-1/aipp2-1/paipp2-1/cpl2-2 quadruple mutants due to a severe developmental defect. Compared to Col-0, the aipp3-1/aipp2-1/paipp2-1/cpl2-2 mutants showed earlier flowering compared with the flc-3 single mutant (Fig. 2g). By contrast, the early flowering phenotypes of aipp3-1, aipp2-2, and aipp2-1/paipp2-1 mutants were completely rescued by ft-10 (Fig. 2h), indicating that the BPC complex represses flowering primarily by repressing the expression of FT.

The AIPP3-BAH domain specifically recognizes the H3K27me3 mark. The chromatin-based mechanisms play vital roles in the flowering time control. The BAH domain is commonly identified as an epigenetic reader module of a particular histone mark. To decipher the molecular function of the AIPP3-BAH domain (Fig. 3a), we first performed a histone peptide pull-down assay using purified AIPP3-BAH protein. Among all the tested histone peptides, AIPP3-BAH could only be pulled down by H3K27me1, H3K27me2, and H3K27me3 peptides in a sequentially increasing manner (Fig. 3b). The H3K27me3-binding activity was further confirmed by isothermal titration calorimetry (ITC) binding analysis (Fig. 3c). Among all four tested histone methylation marks, H3K4me3, H3K9me2, H3K27me3, and H3K36me3, the AIPP3-BAH domain showed a significant preference for the H3K27me3 mark. Moreover, consistent with the histone pull-down result, AIPP3-BAH has a preference for the higher methylation level in H3K27 (Fig. 3c). Thus, the evidence above fully demonstrated that AIPP3-BAH is a H3K27me3-reader module.

Structure of the AIPP3-BAH domain in complex with an H3K27me3 peptide. To gain molecular insight into the interaction between the AIPP3-BAH domain and H3K27me3, we successfully determined the crystal structure of the AIPP3-BAH domain in complex with an H3K27me3 peptide at a resolution of 2.4 Å (Fig. 3d and Supplementary Table 1). Overall, the AIPP3-BAH domain has a classic β-barrel structure that is similar to other reported BAH domain structures. The H3K27me3 peptide has a good electron density map and a β-strand-like extended conformation (Supplementary Fig. 7a). The peptide is captured by a negatively charged cavity that is formed on the surface of the

---

**Fig. 1** The BAH protein AIPP3 associates with PHD proteins and CPL2 to form a protein complex. a Mass spectrometry analysis of epitope-tagged AIPP3, AIPP2, PAIPP2, and CPL2 transgenic plants. b Y2H results showing the reciprocal interactions within the tested proteins. c A diagram showing the interaction network within AIPP3, AIPP2, PAIPP2, and CPL2. d Immunoblotting results of gel filtration assays. One of three independent experiments was shown.

---

**Table:**

| Peptides       | AIPP2–4Myc | AIPP3–3Flag | CPL2–4Myc | PAIPP2–3Flag |
|----------------|------------|-------------|-----------|--------------|
| AIPP2          | 78         | 9           | 9         | /            |
| AIPP3          | 34         | 128         | 1         | 40           |
| CPL2           | 22         | 4           | 138       | 85           |
| PAIPP2         | /          | 22          | 40        | 169          |

---

Developmental defects. Moreover, the aipp3-1 and cpl2-2 mutants showed obvious earlier flowering during both long day (LD) and short day (SD) photoperiods in comparison with Col-0 (Fig. 2b, c). Although only mild early flowering was observed in the aipp2-1 and paipp2-1 single mutants, aipp2-1/paipp2-1 showed similar early flowering compared with the aipp3-1 and cpl2-2 mutants (Fig. 2b, c), suggesting a redundancy in these two PHD proteins in relation to the flowering time control. To dissect their genetic relationship, we attempted to generate double, triple, and quadruple mutants. Unfortunately, we failed to obtain the aipp3/aipp2/paipp2/cpl2 quadruple mutant due to a severe developmental defect. Compared to Col-0, the aipp3-1/aipp2-1/paipp2-1, aipp3-1/cpl2-2, and aipp2-1/paipp2-1/cpl2-2 mutants flowered earlier, and the time to flower was similar to the single mutants under the LD condition (Fig. 2d), suggesting that the BPC complex acts in the same genetic pathway in flowering time control.

In Arabidopsis, FLC, which is a MADS-box transcription factor that integrates multiple flowering signals, acts as a key floral repressor. FLC directly represses the expression of the flowering gene FT and SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1) by binding to the promoter of SOC1 and the first intron of FT. The quantitative RT-PCR (RT-qPCR) indicated that the FLC RNA levels were reduced in the aipp3-1, aipp2-1, cpl2-2, paipp2-1, and aipp2-1/paipp2-1 mutants compared to Col-0 (Fig. 2e). Instead, the FT RNA levels were significantly increased in the aipp3-1, cpl2-2, and aipp2-1/paipp2-1 mutants, but not in the aipp2-1 and paipp2-1 single mutants, further supporting the functional redundancy of AIPP2 and PAIPP2. By contrast, the SOC1 RNA level was not significantly changed in all the tested mutants. It is known that the accumulation of FT protein has a circadian rhythm that peaks before dusk during LD photoperiod. We noticed that the loss of the BPC complex did not change the circadian rhythm of FT mRNA, but led to a constitutive increase (Fig. 2f). To determine the genetic relationship between FLC, FT, and the BPC complex in relation to the flowering time control, flc-3, and ft-10 mutants (in Col-0 background) were crossed with the tested mutants. Surprisingly, the aipp3-1/flc-3, cpl2-2/flc-3, and aipp2-1/paipp2-1/flc-3 mutants displayed similar early flowering compared with the aipp3-1, cpl2-2, and aipp2-1/paipp2-1 mutants, but they flowered earlier than the flc-3 single mutant (Fig. 2g). By contrast, the early flowering phenotypes of aipp3-1, cpl2-2, and aipp2-1/paipp2-1 mutants were completely rescued by ft-10 (Fig. 2h), indicating that the BPC complex represses flowering primarily by repressing the expression of FT.
AIPP3-BAH domain with extensive hydrophobic and hydrophilic interactions (Fig. 3e). In detail, the three aromatic residues of the AIPP3-BAH domain, Tyr149, Trp170, and Tyr172, form an aromatic cage to accommodate the trimethyl-lysine of the H3K27me3 peptide (Fig. 3f), and it resembles other typical methyl-lysine-reading histone readers. At the N-terminus, the H3A25 forms two main chain–main chain hydrogen bonds with the AIPP3 Val140 and Glu142 (Fig. 3f). In the middle, the AIPP3 residues Trp170, Tyr172, His198, and Asp200 form an extensive hydrogen-binding network with H3S28, which further fixes the imidazole ring of His198 in a special rotamer state that is parallel with the prolyl ring of H3P30. This parallel alignment of the two rings enables the CH–π, and stacking interactions between AIPP3 His198 and H3P30, which resemble the recognition of...
H3K27me3 by the EBS and SHL BAH domains with conserved key residues (Supplementary Fig. 7b). In the C-terminus, H3G33 interacts with AIPPP3 Val240 through a main chain–main chain hydrogen bond. To validate these structural observations, we performed a mutagenesis analysis on the key residues. The mutations D200A and H198A, which are essential for H3P30 recognition, in addition to Y149A, W170A, and Y172A, which represent important aromatic cage residues, showed reduced or no detectable binding to the H3K27me3 peptide (Fig. 3g, h).

PHD fingers of AIPPP2 and PAIPPP2 recognize the unmodified H3K4 mark. In addition to AIPPP3-BAH, the PHD fingers of AIPPP2 and PAIPPP2 may also be involved in the recognition of histone marks. AIPPP2 and PAIPPP2 share a conserved PHD finger in their sequences (Supplementary Fig. 8a), and their sequences of PHD fingers possess the typical signatures of unmodified H3 recognition PHD fingers. We first detected their histone substrate binding properties by ITC method. Although the AIPPP2-PHD finger does not behave well in vitro and tends to precipitate, we successfully detected the binding between the PAIPPP2-PHD finger and the differentially methylated H3K4 peptides (Fig. 3i).

The PAIPPP2-PHD finger prefers to bind to unmethylated H3K4 and the binding affinities were clearly decreased when the methylation level of the H3K4 peptide was increased. Considering the high sequence similarity between the two PHD fingers (Supplementary Fig. 8a), we believe that the AIPPP2-PHD finger may possess the same binding preference on the unmodified H3K4. Moreover, both AIPPP2 and PAIPPP2 PHD fingers share ~35% sequence identity with the PHD finger of Glycine max ATXR5 (PDB code: 5VAB; Supplementary Fig. 8a), which recognizes the unmodified H3K4 (ref. 52). We modeled the AIPPP2 and PAIPPP2-PHD fingers using the ATXR5 PHD finger, as a template to analyze the interactions with unmodified H3K4 (Fig. 3j and Supplementary Fig. 8b). In the modeled structure, we noticed that almost all the peptide-binding residues are conserved (Supplementary Fig. 8a). For instance, the Pro322 and Gly324 of PAIPPP2, which correspond to the ATXR5 Pro60 and Gly62, respectively, are involved in the hydrogen-binding interaction with H3A1 (Fig. 3k). Similarly, the Ile300, Cys301, and Asp306 of PAIPPP2, which are equivalent to ATXR5 Leu39, Cys40, and Asp44, respectively, contribute to the recognition of H3R2 (Fig. 3k). The Val285, Gly292, and Leu298 of PAIPPP2, which correspond to the Val23, Gly31, and Leu37 of ATXR5, respectively, participate in the recognition of the unmodified H3K4 (Fig. 3k). To validate the modeling results, we performed mutagenesis experiment, too. As most of the peptide recognition is achieved by the hydrogen-bonding interactions of the main chain of the PHD finger, we only mutated Asp306 of PAIPPP2, which is involved in the recognition of H3R2 by its side chain. As shown in Fig. 3l, the D306K mutation of PAIPPP2-PHD finger almost totally disrupts the peptide binding, further supporting our modeling data. The AIPPP2-PHD finger possesses similar unmodified H3K4 recognition residues and interactions (Supplementary Fig. 8c).

The BPC complex represses the expression of the genes marked by H3K27me3 and low-methylated H3K4. H3K27me3 is usually considered a repressive mark that functions in transcriptional repression by recruiting recognition proteins or protein complexes. To dissect the molecular function of the BPC complex, mRNA-seq was performed using their null mutants, as well as CLF and LHP1 mutants, which are one of the known H3K27me3 methyltransferases and a reader protein in Arabidopsis, respectively. Using a twofold change cutoff, we noticed that the numbers of upregulated differentially expressed genes (up-DEGs) were far greater than the numbers of down-DEGs in the aipp3-1, cpl2-2, and aipp2-1/paipp2-1 mutants (Fig. 4a). Supporting the functional redundancy of AIPPP2 and PAIPPP2, very few DEGs were identified in aipp2-1 and paipp2-1 single mutants. It is noteworthy that most of the up-DEG genes regulated by the BPC complex display very low-expression levels in the wild type, which is a similar pattern to that of the clf and lhp1 mutants (Fig. 4b). Under a strict criterion, 155 genes were commonly upregulated in the bphd–cpl2 mutants (Fig. 4c, d and Supplementary Data 2), and a substantial subset are stress-responsive genes (Supplementary Fig. 9), such as SEC31A (AT1G18830), which participates in endoplasmic reticulum stress responses55–57, and AGO5 (AT2G27880), which is involved in antiviral RNA silencing58–60 and gametophyte development61,62.

In fact, for most of the up-DEGs, higher expression was also observed in the aipp2-1/paipp2-1 up-DEGs, higher expression was also observed in the aipp3-1 and cpl2-2 mutants (Fig. 4d), indicating that the BPC complex targets a common subset of genes for repression.

We next explored the chromatin feature of the target genes by plotting the distributions of the histone marks on the up-DEGs, using published histone modifications chromatin immunoprecipitation (ChIP)-seq data. As expected, the H3K27me3 mark was heavily enriched on the bodies of the common target genes (Fig. 4e). Consistent with the ITC result in which the higher methylation of H3K4 inhibits the binding of PAIPPP2-PHD (Fig. 3i), the levels of active H3K4me3 mark were lower in the regions around the TSS. H3K36me3 deposition was also low on the whole gene bodies. The depositions of H3K9me2 and H3K27me1 were markedly lower on the target genes, suggesting that the BPC complex primarily targets euchromatic genes. The distribution patterns of H3K27me3 and H3K4me3 marks at BPC complex target genes were confirmed by ChIP-qPCR assays at representative target genes AGO5, SUC5, and FT (Fig. 4f, g). We next investigated the impacts of BPC complex dysfunction on global histone mark levels. Interestingly, compared to the reduction of global H3K27me2/3 levels in clf-81 mutant, the absence of the BPC complex did not result in obvious changes in...
global levels of H3K27me1/2/3 and H3K4me1/2/3 (Supplementary Fig. 10a). Consistent with this pattern, H3K27me3 deposition was not obviously changed between Col-0 and bpc mutants at selected target genes AGO5, SUC5, and FT (Fig. 4g), whereas H3K4me3 levels are slightly increased (Fig. 4h). To further verify this observation, more target genes were selected. As shown in Supplementary Fig. 10b–d, nuclear run-on assay showed that AT2G43570, AT3G59480, and AT3G53650 genes were transcriptionally upregulated compared to wild type, indicating that these three genes are subjected to BPC complex-mediated transcriptional repression. Interestingly, ChIP-qPCR results indicated that H3K27me3 levels were not obviously changed in bpc mutants at AT2G43570 and AT3G59480, consistent with our observation at AGO5, SUC5, and FT (Fig. 4g). While H3K27me3 deposition showed significant reduction at AT3G53650, implying that BPC dysfunction has substantial impacts on H3K27me3 deposition at this gene. Combined with these data, we speculate that, for most target genes, the BPC complex may serve as a surveillance system to prevent reactivation of H3K27me3-marked genes, which are already silenced by PRC2, but in some specific
genes, BPC is required for H3K27me3 deposition through unknown mechanism.

The BPC complex reads the H3K27me3 at a genome-wide level. We next investigated the interplay between the BPC complex and the chromatin of the target genes by performing a ChIP assay in epitope-tagged transgenic plants. The ChIP-qPCR results indicated that AIPP3, AIPP2, and PAIPP3 were enriched in the selected target genes, particularly in the regions close to TSSs (Fig. 5a), but they were low in the adjacent high H3K4me3/low H3K27me3 genes. To confirm this point, an AIPP3 ChIP-seq was performed. As shown in Fig. 5b, AIPP3 specifically binds to the selected target genes. At the genome-wide level, AIPP3-binding peaks were enriched in the DEGs that were upregulated in the selected target genes, particularly in the regions close to TSSs (Fig. 5c, d). Moreover, a pattern of high peaks were enriched in the DEGs that were upregulated in the selected target genes. At the genome-wide level, AIPP3-binding H3K27me3-marked regions1. As shown in Fig. 5f, 455 loci, or methylation levels (right panel). NDB no detectable binding.

The overall structure of the AIPP3-BAH domain H3K27me3 peptide complex shown in

Fig. 3 Structural analysis of the AIPP3-BAH domain and PAIPP2-PHD finger in complex with H3K27me3 and unmodified H3K4 peptide, respectively.

a The domain architecture of AIPP3 (lower panel) and the BAH domain construct used for structural and biochemical studies (upper panel). b The immunoblotting results showing the pull-down results for AIPP3-BAH using different histone peptides. The 0.5% inputs serve as positive controls. c The ITC binding curves showing the AIPP3-BAH domain-binding preference for different histone methylation marks (left panel) and different H3K27 methylation levels (right panel). NDB no detectable binding. d The overall structure of the AIPP3-BAH domain H3K27me3 peptide complex shown in ribbon with the AIPP3-BAH domain, and the H3K27me3 peptide colored in green and yellow, respectively. e An electrostatic surface view of AIPP3-BAH domain with the H3K27me3 peptide in the space-filling model showing that the peptide fits inside a negatively charged surface cleft of the AIPP3-BAH domain. f The detailed interaction between AIPP3-BAH domain and the H3K27me3 peptide with the interacting residues highlighted in sticks and the hydrogen bonds highlighted in dashed yellow lines. g, h The ITC binding curves show that the mutations of essential residues for the H3P30 recognition (g) and aromatic cage residues (h) of the AIPP3-BAH domain significantly decrease the binding toward the H3K27me3 peptide. i The ITC binding curves showing the specific preference of the PAIPP2-PHD finger for the unmodified H3K4. j The overall modeled structure of the PAIPP2-PHD finger in complex with the unmodified H3 peptide with the PHD finger and peptide shown in ribbon and space-filling models. The modeled PAIPP2-PHD finger and the modeling template ATXR5 PDH finger are colored in cyan and silver, respectively, and they were superimposed together. k The detailed interaction between the PAIPP2-PHD finger (in cyan) and the unmodified H3K4 peptide with the interacting residues are highlighted in stick model and hydrogen bonds highlighted in dashed yellow lines. The corresponding residues from modeling template ATXR5 (in silver, PDB code: 5VAB) were overlain and highlighted, showing that almost all the interacting residues are conserved. l The ITC binding curves showing that the D306K mutation of PAIPP2-PHD finger, which potentially disrupts H3R2 recognition, totally abolishes the unmodified H3K4 binding by PAIPP2.

BACH–PHD–CPL2 couples the recognition of H3K27me3 and the repression of Pol II release. It has been well documented that differential phosphorylation on the Ser2 (Ser2P) and Ser5 (Ser5P) forms of Pol II in the target genes. Recently, Zhu et al. revealed that the BAH- and PHD-mediated recognition of H3K27me3 and unmodified H3K4 directly represses transcription through the CPL2-mediated phosphorylation of Pol II. To confirm this hypothesis, we first determined the states of different forms of Pol II in the target genes. Recently, Zhu et al. revealed that the BAH- and PHD-mediated recognition of H3K27me3 and unmodified H3K4 directly represses transcription through the CPL2-mediated phosphorylation of Pol II. To confirm this hypothesis, we first determined the states of different forms of Pol II in the target genes. Recently, Zhu et al. revealed that the BAH- and PHD-mediated recognition of H3K27me3 and unmodified H3K4 directly represses transcription through the CPL2-mediated phosphorylation of Pol II. To confirm this hypothesis, we first determined the states of different forms of Pol II in the target genes. Recently, Zhu et al. revealed that the BAH- and PHD-mediated recognition of H3K27me3 and unmodified H3K4 directly represses transcription through the CPL2-mediated phosphorylation of Pol II.
Fig. 4 BAH–PHD–CPL2 complex represses the expression of H3K27me3-enriched genes. a The numbers of up- and down-DEGs identified in the selected mutants. A twofold cutoff was used for the DEG criterion. b The BAH–PHD–CPL2 complex primarily represses low-expression genes. The gene expression levels of up-DEGs were classified into four groups according to their FPKM values in Col-0 mRNA-seq. All the annotated genes serve as controls. c A Venn diagram showing the overlap of up-DEGs between the selected mutants. d A heatmap showing the expression analysis in selected mutants. The genes in the upper and lower panels represent the common 155 up-DEGs and up-DEGs of aipp2–1/paipp2–1, respectively. e The distribution of different histone marks in the respective up-DEGs of selected mutants. TTS transcription termination site. f Snapshots showing the expression of the selected target genes. One representative replicate of mRNA-seq and histone ChIP-seq are shown. The adjacent genes with high H3K4me3/low H3K27me3 were also shown as parallel controls. g, h ChiP-qPCR showing the H3K27me3 (g) and H3K4me3 (h) density at the selected target genes. The ChiP signals were normalized to histone H3. The data are means ± S.D. of three technical repeats. One representative result of three biological replicates is shown. The lowercase letters represent the ChiP-qPCR examined regions as shown in f.
unphosphorylated Pol II were detected at approximately +200 bp and sharply decreased to a very low level, and Ser5P and Ser2P-Pol II peaked in the promoter-proximal regions, but decreased to a mild level when entering the gene body regions (Fig. 7a). These results strongly support the idea that transcription initiation and subsequent elongation of target genes were repressed in the wild type by the BPC complex. We then checked what happens to the Pol II occupancy when the complex is absent. Compared to the wild type, higher accumulations of Ser5P-Pol II were observed at the selected target genes in aipp3-1, cpl2-2, and aipp2-1/paipp2-1 mutants (Fig. 7b), and the Pol II signals maintained high levels toward the 3′ end of the selected genes. Similar to Ser5P-Pol II, the occupancy of total (unphosphorylated) and Ser2P-Pol II also displayed higher levels at selected target genes AGO5, SUC5, and FT in the bpc mutants in comparison with Col-0 (Supplementary Fig. 12), demonstrating that Pol II occupancy when the complex is absent. Compared to the wild type, higher accumulations of Ser5P-Pol II were observed at the selected target genes in aipp3-1, cpl2-2, and aipp2-1/paipp2-1 mutants (Fig. 7b), and the Pol II signals maintained high levels toward the 3′ end of the selected genes. Similar to Ser5P-Pol II, the occupancy of total (unphosphorylated) and Ser2P-Pol II also displayed higher levels at selected target genes AGO5, SUC5, and FT in the bpc mutants in comparison with Col-0 (Supplementary Fig. 12). In contrast, the occupancies of all types of Pol II at AGO5 and SUC5 downstream nontarget genes were not significantly changed (Supplementary Fig. 12), demonstrating that bpc mutations led to reactivation of Pol II initiation specifically at BPC complex target genes and the initiated Pol II successfully switched to an elongating state in the bpc mutants. These evidence, combined with the known knowledge that CPL2 is a Ser5P-Pol II phosphatase, prompts us to hypothesize that the BPC complex represses gene expression by connecting the BAH–PHD module-mediated recognition of H3K27me3/unmodified H3K4 and the CPL2-mediated dephosphorylation of Pol II CTD-Ser5-P. To support the above hypothesis, flavopiridol (FLA) treatment assay was performed which can reduce the phosphorylation of Pol II CTD by inhibiting the activity of CDK kinases, and the nascent RNA levels of the selected target genes were measured by nuclear run-on assay. The results indicated that the nascent RNA levels of AGO5 and SUC5, but not that of the downstream nontarget genes, were dramatically increased in the bpc mutants compared to Col-0 in mock condition. This evidence strongly supports our conclusion of transcriptional repression.
implemented by the BPC complex. In contrast, the nascent RNA levels of both AGOS, SUC5, and the downstream nontarget genes were greatly reduced by FLA treatment, and no obvious changes were observed between different genotypes (Fig. 7c), indicating that inhibition of Pol II CTD phosphorylation dramatically repressed the transcription reactivation caused by the malfunction of the BPC complex. Considering the likelihood of general effects of FLA treatment on transcriptional at a global level, to repressed the transcription reactivation caused by the malfunction of Pol II by ChIP-qPCR assay was performed in CPL2-GFP/Col-0, CPL2-GFP/ipp3-1, and CPL2-GFP/ipp2-1/paipp2-1 plants. The result indicated that CPL2 has significant binding at selected target genes in wild-type background, whereas this binding was completely abolished in the ipp3-1 and ipp2-1/paipp2-1 mutants (Fig. 7d). This result provides a link between chromatin marks and CPL2-mediated dephosphorylation of Pol II, in which the key amino acids for the recognition of H3K27me3/H3K4me0 marks by BAH complex (Fig. 8). In this model, the BAH complex is essential for proper CPL2-mediated repression of flowering and gene expression. A The flowering phenotypes of the ipp3-1 mutant for its complementary lines when transformed with the wild-type AIPP3 genomic DNA, mutated AIPP3 in which the key amino acids for H3K27me3-binding were mutated. For the mutated transgen, two randomly selected transgenic lines were used for the analysis. The plants were grown under LD. B The column showing the numbers of rosette leaves (LD) of different AIPP3 transgenic plants upon flowering under the LD condition. Black horizontal lines represent the mean, and the error bars represent ±S.D. from the number of plants counted for each genotype. C Western blotting result showing the accumulation levels of AIPP3 proteins in different transgenic lines. Ponceaux staining serves as protein loading controls. Relative mRNA levels of the selected target genes of the BAH–PHD–CPL2 complex in the wild-type, ipp3-1, and different AIPP3 transgenic plants. Their levels are presented relative to Col-0. The data are the means ± S.D. of three biological repeats. Unpaired one-tailed t test was performed and *p value < 0.01. ChIP-qPCR showing AIPP3 occupancy at the selected target genes in AIPP3 and W170A transgenes. The data are the means ± S.D. of three biological repeats. Unpaired one-tailed t test was performed and *p value < 0.01. ns no significance.
Fig. 7 BPC complex directly connects H3K27me3 recognition with transcriptional repression. a Snapshots showing the occupancies of different Pol II forms on selected target genes using a published Pol II NET-seq database. The red and blue lines indicate the Pol II signals in the plus and minus strands, respectively. The dashed red boxes indicate the occupancies of Pol II on the selected target genes. The black arrows indicate the transcriptional direction. b The ChIP-qPCR validation of the occupancy of Ser5P-Pol II on selected genes in different mutants. The lowercase letters indicate the examined regions (the same below). The occupancy was normalized to the ACT7. The data are the means ± S.D. of three biological repeats. Unpaired one-tailed t test was performed and *p value < 0.01. ns no significance. c Nuclear run-on analysis showing the relative Pol II transcription rate at the selected target genes in Col-0 and bpc mutants with or without FLA treatment (mock). AT2G27900 and AT1G71900 genes serve as control genes. The relative transcription rate was normalized to 18 S rRNA. The data are the means ± S.D. of three biological repeats. d CPL2 ChIP-qPCR results showing the relative occupancy of CPL2 at selected target genes in the presence and absence of AIP3 and AIP2/PAIP2. The occupancy was first normalized to AtSN1 and then normalized to Col-0. The data are the means ± S.D. of three biological repeats. Unpaired one-tailed t test was performed and *p value < 0.01.
Discussion

Recently, we structurally described several plant H3K27me3-reading BAH domain proteins, such as EBS, SHL, and AIPP3 in this research\(^{27,28}\), which has encouraged us to characterize the H3K27me3-reading BAH domains. The recognitions of H3K27me3 by all these BAH domains depend on the aromatic cage to recognize the methyl-lysine, and a histidine and an aspartic acid residue to specifically interact with H3P30 for sequence specificity. Using the aromatic cage, and the specific His and Asp residues as a criterion, we identified a subfamily of BAH domain proteins that can potentially recognize the H3K27me3 mark, which are widely distributed in plants, fungi, and animals (Supplementary Fig. 13). Interestingly, the predicted H3K27me3-binding BAH domain-containing protein human BAHD1 was reported to be an H3K27me3 reader\(^{30}\), which further supported our prediction. Therefore, we believe that the aromatic cage, and conserved His and Asp residues are the key features of H3K27me3-recognition BAH domains.

Different histone marks do not function in a totally independent manner. Instead, they engage in communications and cooperation with each other. The cooperation between different histone marks can be a combinatorial reading to enhance binding or be mutually exclusive to balance the binding between different marks, which occurs at both the single protein level and the multiple protein complex level\(^{38}\). Recently, we reported that the two flowering regulators EBS and SHL could dynamically recognize the antagonist histone marks H3K27me3 and H3K4me0 at the single protein level to regulate floral phase transitions\(^{27,28}\). Here, AIPP3 and AIPP2/PAIPP2 could recognize H3K27me3 and H3K4me0, respectively, and form a bivalent histone reader complex. This observation is consistent with our functional data showing that the BPC complex colocalizes with higher H3K27me3 and lower H3K4me3 marked chromatin regions. The BAH domain of AIPP3 is responsible for the colocalization of the complex with H3K27me3. The binding of unmodified H3K4 by the PHD finger of AIPP2/PAIPP2 may have two roles. First, the binding of unmodified H3K4 may prevent the binding toward methylated H3K4 to make sure the complex is targeted to the gene repressive H3K4me3 depletion region. Second, the binding of H3K4me0 together with the H3K27me3 binding by AIPP3 may combine to enhance the overall binding of the complex toward a certain chromatin region. Therefore, the proper targeting of the BPC complex relies on the crosstalk between the H3K4me0 and H3K27me3 marks.

RNA Pol II-dependent transcription is a stepwise process involving the formation of the preinitiation complex (PIC), initiation, elongation, and polyadenylation/termination stages. Each of the stages is associated with a distinct pattern of CTD phosphorylation\(^{8,65}\). The transcriptional machinery is first recruited to the promoter regions. Once incorporated into the PIC, the mediator stimulates cyclin-dependent kinase to phosphorylate serine 5 of the CTD heptad repeat, and Ser5P helps in the release of Pol II from the PIC complex, thereby allowing Pol II to escape the promoter and the subsequent initiation of transcription. Ser5P is retained during the first several hundred nucleotides, in preparation for productive elongation\(^{69}\).

Here, on the selected target genes AGO5 and SUC5, phosphorylated Pol II is restricted in the TSS region, whereas Ser5P and Ser2P were below detectable levels (Fig. 7a), indicating that Ser5P-dependent transcription initiation is inhibited, possibly by CPL2, in the presence of the BPC complex. Consistent with this notion, the density of the Ser5P-Pol II was significantly increased, and it peaked within the first several hundred nucleotides in the BPC complex mutants (Fig. 7b), indicating that the Ser5P-dependent transcription initiation was derepressed. In the canonical model of H3K27me3-mediated transcription repression, the recognition of H3K27me3 recruits the PcG proteins in the PRC1 complex to impose the monoubiquitination of H2A, which represses transcription through three possible mechanisms, as mentioned in the introduction\(^{8}\). Consistent with this model, a recent report showed that the mutations in H3K27me3 reader proteins LHP1, EBS, and SHL in Arabidopsis led to a reduction in H2Aub1 (ref.\(^{29}\)). We found that the H2Aub1 levels were reduced in the hiji-3 mutant, but were not affected in the aipp3-1/cpl2-2, aipp3-1/aipp2-1/paipp2-1, and aipp2-1/paipp2-1/cpl2-2 mutants (Supplementary Fig. 10a). This observation is consistent with our hypothesis that the BPC complex represses transcription mainly through the inhibition of Ser5P-dependent transcription initiation. While, we cannot rule out the possibility that BPC the complex has direct/indirect interaction with H2Aub1 at specific target genes. This study unveiled a direct connection between H3K27me3 recognition and transcription repression; the BAH–PHD module recognizes H3K27me3/H3K4me0 and recruits CPL2 to dephosphorylate Pol II CTD, resulting in the
failure of Pol II to enter into the initiation and elongation form. Although Pol II CTD phosphatases are conserved in eukaryotes, no evidence shows that histone modifiers manipulate CTD phosphatas, including CPL2 analogies in other systems, to affect transcription. Regarding histone modifiers, several reports have indicated that histone PTM modifiers affect transcription via modulating Pol II CTD phosphorylation state.

JMJ3D, an H3K27me3 demethylase in human, has been shown to directly interact with CTD-Ser2P to affect gene expression. Knocking down JMJ3D or JHDM1D, a H3K2me3/2 demethylase, reduces the enrichment of Pol II CTD-Ser2P at specific genes in human promyelocytic leukemia cells. In addition to histone methyla, other histone PTMs including histone ubiquitylation and phosphorylation, are also associated with Pol II CTD phosphorylation-dependent transcriptional elongation. For example, knocking down histone ubiquitylation modifiers has been shown to affect CTD-Ser2 phosphorylation. The phosphorylation of histone H3 on Ser10 and Ser28 has been reported to be associated with phosphorylated Pol II during transcriptional activation in human and Drosophila.

These studies report a notion that modulation of Pol II CTD phosphorylation represents an important regulatory mechanism adopted by chromatin regula, to regulate gene expression. Our finding that the BPC complex reading histone information and conferring transcriptional repression through CPL2 phosphate-mediated modulation of Pol II CTD phosphorylation state provides a direct evidence to support this notion. Considering that CPL2 is a plant-specific Pol II phosphatase that bears a unique RMB domain, and that this domain is required for its interaction with PHD proteins, the BAH–PHD–CPL2 pathway may represent a newly evolved silencing pathway in plants. Our findings suggest a greater complexity and diversification of H3K27me3-mediated transcriptional regulation. In addition, it is well-known that H3K27me3-mediated silencing mechanisms participate in multiple biological processes in both plants and animals, particularly development and stress-responsive genes. The obvious developmenta, defects and reactivation of many development- and stress-response-related genes (Supplementary Fig. 9) in bpc mutants imply that the BPC complex may play more important roles in plant development and stress responses in addition to flowering time control.

**Methods**

**Plant materials and growth conditions.** All the plant seeds were sown and grown on 1/2 MS medium containing 1% sucrose. The seedlings were grown under a LD (16 h light/8 h dark) or SD (8 h light/16 h dark) photoperiod at 23 °C. The T-DNA insertion mutants aipp2-1, aipp3-1, and cpl2-1 were described in our previous study.

The paipp2-1 mutant was generated by CRISPR/Cas9-mediated mutat,ion in a col-0 background (Supplementary Fig. 4). elf-81 and bph-1 have been previously described. For the epitope-tagged transgenic expression of the AIPP2, AIPP3, PAIPP2, and CPL2 genes, the wild-type and mutated genomic DNA driven by their native promoters were cloned into binary vectors with different tags and then transformed into the corresponding mutants, using the flowering dip method. T3 generation transgenic plants were used for analysis.

**RT-qPCR and RNA-seq analysis.** Total RNA was extracted from 2-week-old seedlings using Trizol reagent (Thermo), and cDNAs were synthesized using HiScript II Reverse Transcriptase (Vazyme). qPCR was performed using a CFX96 Touch Deep Well Real-Time PCR Detection System (Bio-Rad). Three biological replicates were created. The primers used in this study are listed in Supplementary Table 2. For the RNA-seq analysis, total RNAs were extracted from 12-day-old seedlings grown during LDs using a NReasy Plant Mini kit (Qiagen). Following RNA purification, reverse transcription and library construction, the libraries were quantified by TBS380, and a paired-end RNA sequencing library was performed with Illumina NovaSeq 6000 (2 × 150 bp read length). The raw paired-end reads were trimmed and subject to quality control with SeqPrep (https://github.com/)

**jstor/jn/Seq(Prep) and Sickle (https://github.com/najoshi/sickle), using the default parameters.**

**IP and MS analysis.** IP and MS analyses were performed as previously described. In brief, the total proteins were extracted from the inflorescence tissues with IP buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM MgCl2, 10% glycerol, 0.1% NP-40, 0.5 mM DTT, and protease inhibitor cocktail), and then precipitated with anti-Flag (Sigma-Aldrich) or anti-Myc (Millipore) antibodies for 2 h at 4 °C. After five washes, the precipitated protein mixtures were subjected to MS analysis.

**Protein interaction analysis and gel filtration assays.** For the Y2H assays, the full-length and truncated coding sequences of AIPP2, AIPP3, CPL2, and PAIPP2 were cloned into the pGADT7 and pGBKTK7 vectors to generate AD and BD constructs. After the transformation, the yeast cultures were spotted onto SD plates lacking Leu andTry (-1/-) and incubated at 30 °C for 3 d.

For gel filtration assay, the total proteins were extracted from 4 g of seedling tissues expressing AIPP2-FLAG, AIPP2-4MYC, PAIPP2-3FLAG, and CPL2-4MYC via buffer and then loaded on a Superdex 200 10/300 GL column (GE Healthcare). The eluted fractions were collected in 96-well plates and the target proteins were detected by standard western blotting.

**Protein expression and purification.** The sequence containing the AIPP3-BAH domain (residues 112–279) was constructed into a self-modified pMal-p2x vector to fuse a hexahistidine tag plus a maltose-binding protein (MBP) tag to the N-termi,ns of the target protein. The plasmids were transformed into Escherichia coli strain BL21 (DE3) RIL (Stratagene). Expression was induced at 16 °C overnight with 0.2 mM of IPTG. The recombinant proteins were purified with a prepac,aged HisTrap FF column (GE Healthcare). The His-MBP tags were cleaved by TEV prot,ease overnight and removed by flowing through a HisTrap FF column (GE Healthcare) again. The target protein was further purified using a Heparin column (GE Healthcare) and a Superdex 2000 column (GE Healthcare). All the AIPP3-BAH mutations were generated by standard PCR-based mutagenesis protocol. The mutations of AIPP3-BAH and truncated AIPP2/PAIPP2 fragments were purified using the same protocols, as those used for the wild-type AIPP3-BAH. For the GST-AIPP3-BAH proteins used in histone peptide pull-down assays, the wild-type and mutated AIPP3-BAH proteins were purified with glutathione-Sepharose beads (GE Healthcare) and eluted with elution buffer (50 mM Tris-HCl pH 8.0, and 10 mM reduced glutathione). The peptides were purchased from GL Biochem or Epicypher.

**Histone peptide pull-down.** For histone peptide pull-down assay, 1.5 µg of biotinylated histone peptides were incubated with streptavidin beads (NEB) in binding buffer (50 mM Tris-HCl 8.0, 300 mM NaCl, and 0.1% NP-40) for 1 h at 4 °C, and then washed with binding buffer. A 1.5 µg quantity of AIPP3-BAH proteins was incubated with a peptide–bead mixture in 0.5 ml of binding buffer for 3 h at 4 °C, and then washed with binding buffer five times. The protein–bead mixtures were subjected to immunoblotting using anti-GST antibody (Abmart, #12G8, 1:2000 dilution).

**Chromatin immunoprecipitation assays and ChIP-seq analysis.** ChIP assays were performed according to a reported procedure. In brief, 3 g of seedlings was harvested and fixed with 1% formaldehyde in the cross-linking buffer (0.4 M sucrose, 10 mM Tris-HCl pH 8, 1 mM PMSF, and 1 mM EDTA). After nuclei isolation with isolation buffer (0.25 M sucrose, 15 mM PIPES pH 6.8, 5 mM MgCl2, 60 mM KCl, 15 mM NaCl, 1 mM CaCl2, 0.9% Triton X-100, and protease inhibitor cocktail) and the following centrifugation, nuclei were resuspended in 500 µl of lysis buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% PMSF, 1% SDS, 0.1% Na deoxycholate, 1% Triton X-100, and protease inhibitor cocktail) and sonicated with a Bioruptor (Diagenode). The nuclear lysate was precipitated with anti-Flag (Sigma-Aldrich, #F1804, 1:100 dilution), anti-Myc (Millipore, #05-724, 1:100 dilution), anti-H3 (Abcam, #ab1791, 1:100 dilution), anti-H3K27me3 (Millipore, #07-449, 1:100 dilution), Ser2-Pol II (Abcam, #ab5095, 1:100 dilution), Ser5-Pol II (Abcam, #ab5131, 1:100 dilution), unphosphorylated Pol II (Abcam, #aback7, 1:100 dilution), and GFP (Abcam, #ab290, 1:100 dilution) antibodies overnight and incubated with Dynabeads (Thermo) for 2 h. The precipitated protein–DNA mixtures were washed and eluted with elution buffer (0.5% SDS and 0.1 M NaHCO3) at room temperature. The DNA was recovered after reverse cross-linking and proteinase K treatment.

For ChIP-seq analysis, clean reads were mapped to the Arabidopsis thaliana genome (TAIR10) by Bowtie2 (version 2.2.8) with default parameters. Enriched peaks were identified by MACS (version 1.4) with default parameters. We defined the region of a target gene as the range from 1 kb upstream of TSS to TTS. The target genes of each peak were annotated by annotatePeak function in ChIPseeker package. The visualization of RNA-seq data was performed using GAGE and DESeq2. The average read coverage over gene body and additional 1 kb upstream and downstream of the TSS and TES was performed by deepTools (version 2.4.1)79.
Crystallization, data collection, and structure determination. Crystalization screening was performed using the sitting drop vapor diffusion method at 4 °C. The sample was concentrated and mixed with peptide at a molar ratio of 1:4. All the crystals emerged in a solution of 0.1 M HEPES, pH 7.0, and 2.4 M ammonium sulfate. The crystals were cryo-protected in the reservoir solution supplemented with 20% glycerol and flash-cold in liquid nitrogen for X-ray diffraction. All the diffraction data were collected at beamline BL19U1 of the National Center for Protein Sciences Shanghai (NCPCS) at the Shanghai Synchrotron Radiation Facility (SSRF). The data were processed with the HKL3000 program.\(^{80}\) The structure was determined by molecular replacement method as implemented in the Phenix program\(^{81}\), using the ZMET2 BAH domain (PDB ID: 4FT2) as the searching model.\(^{82}\) Model building and structure refinement were performed using the Coot and Phenix programs, respectively.\(^{81,82}\) The statistics of data collection and structure refinement are listed in Supplementary Table 1.

**IC** T calorimetry. All the binding experiments were performed on a Microcal PE-AQITC instrument (Malvern) at 25 °C. The proteins were dialyzed against a buffer consisting of 50 mM Tris-HCl, pH 8.0, 5 mM MgCl\(_2\), and 40% glycerol and mixed with titration buffer (10 mM Tris-HCl, pH 8.0, 5 mM MgCl\(_2\), 150 mM KCl, 2 mM DTT, 80 units RNase inhibitor, 0.5 mM BrUTP, 1 mM ATP, 1 mM GTP, 1 mM CTP, 0.5 mM UTP, and 0.2% sarkosyl). After incubation at 30 °C for 15 min, the run-on reaction was stopped by adding Trizol reagent (Thermo). The nuclear RNAs were extracted and treated with DNase. The RNA-free RNAs were then incoated with 2 μg anti-BrU antibody (abcam) at room temperature for 30 min and precipitated by Dynabeads (Thermo).

After two times washing, RNAs were extracted using Trizol reagent (Thermo). cDNAs were synthesized using SuperScript IV Reverse Transcriptase (Thermo) and subjected to qPCR analysis.

**GUS** staining assay. The plant materials were stained in staining solution overnight at 37 °C and then washed with 70% ethanol at room temperature. The plants were cleared after being washed and were observed under a Zeiss Stemi 305 microscope.

**Histone** immunoblotting. The nuclei were extracted from 2-week-old seedlings with isolation buffer (0.25 M sucrose, 15 mM PIPES pH 6.8, 5 mM MgCl\(_2\), 60 mM KCl, 15 mM NaCl, 1 mM CaCl\(_2\), 0.9% Triton X-100, and protease inhibitor cocktail), the nuclear pellet was resuspended with nuclear storage buffer (50 mM Tris-HCl pH 8.0, 0.1 mM EDTA, 5 mM MgCl\(_2\) and 40% glycerol) and mixed with titration buffer (10 mM Tris-HCl, pH 8.0, 2.5 mM MgCl\(_2\), 150 mM KCl, 2 mM DTT, 80 units RNase inhibitor, 0.5 mM BrUTP, 1 mM ATP, 1 mM GTP, 1 mM CTP, 0.5 mM UTP, and 0.2% sarkosyl). After incubation at 30 °C for 15 min, the run-on reaction was stopped by adding Trizol reagent (Thermo).

**References**

1. Kim, S. Y., Lee, J., Eshed-Williams, L., Zilberman, D. & Sung, Z. R. EMF1 and PRC2 cooperate to repress key regulators of Arabidopsis development. PLoS Genet. 8, e1002512 (2012).
2. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. Nature 403, 41–45 (2000).
3. Campos, E. I. & Reinberg, D. Histones: annotating chromatin. Annu. Rev. Genet. 43, 559–599 (2009).
4. Liu, C., Lu, F., Cui, X. & Cao, X. Histone methylation in higher plants. Annu. Rev. Plant Biol. 61, 395–420 (2010).
5. Chang, Y. N. et al. Epigenetic regulation in plant abiotic stress responses. J. Integr. Plant Biol. 62, 563–580 (2019).
6. Kini, D. H. & Sung, S. Polycomb-mediated gene silencing in Arabidopsis thaliana. Molecules Cells 37, 841–850 (2014).
7. Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. Nat. Rev. Mol. Cell Biol. 10, 697–708 (2009).
8. Aranda, S., Mas, G. & Di Croce, L. Regulation of gene transcription by Polycomb proteins. Sci. Adv. 1, e1500737 (2015).
9. Di Croce, L. & Helin, K. Transcriptional regulation by Polycomb group proteins. Nat. Struct. Mol. Biol. 20, 1147–1155 (2013).
10. Blackledge, N. P., Rose, N. R. & Kloee, R. J. Targeting Polycomb systems to regulate gene expression: modifications to a complex story. Nat. Rev. Mol. Cell Biol. 16, 643–649 (2015).
11. Shao, Z. et al. Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell 98, 37–46 (1999).
12. Levine, S. S. et al. The core of the chromatin repressive complex is functionally and compositionally conserved in flies and humans. Mol. Cell. Biol. 22, 6070–6078 (2002).
13. Francis, N. J., Kingston, R. E. & Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. Science 306, 1574–1577 (2004).
14. Cen, Y. V. et al. The polycomb group mutant esc leads to augmented levels of paused Pol II in the Drosophila embryo. Mol. Cell 42, 837–844 (2011).
15. Zhou, W. et al. Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. Mol. Cell 29, 69–80 (2008).
16. Stock, J. K. et al. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat. Cell Biol. 9, 1428–1435 (2007).
17. Liu, R., Li, X., Chen, W. & Du, J. Structure and mechanism of plant histone mark readers. Acta Biochim. Biophys. Sin. 44, 28–39 (2012).
18. Yang, N. & Xu, R. M. Structure and function of the BAH domain in chromatin biology. Crit. Rev. Biochem. Mol. Biol. 48, 211–221 (2013).
19. Li, Y. & Li, H. Many keys to push: diversifying the ‘readership’ of plant homeodomain fingers. Acta Biochim. Biophys. Sin. 44, 28–39 (2012).
20. Du, J. et al. Dual binding of chromodomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell 151, 167–180 (2012).
21. Koo, A. J. et al. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. Nature 484, 115–119 (2012).
22. Li, S. et al. Structural basis for the unique multivalent readout of unmodified H3 tail by Arabidopsis ORC1b BAH-PHD cassette. Structure 24, 488–494 (2016).
23. Armache, K. J., Garlick, J. D., Canzio, D., Narlikar, G. J. & Kingston, R. E. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. Science 334, 977–982 (2011).
24. Yang, D. et al. Nalpha-acetylated Sir3 stabilizes the conformation of a nucleosome-binding loop in the BAH domain. Nat. Struct. Mol. Biol. 20, 1116–1118 (2013).
25. Arnaudo, N. et al. The N-terminal acetylation of Sir3 stabilizes its binding to the nucleosome core particle. Nat. Struct. Mol. Biol. 20, 1119–1121 (2013).
26. Wang, F. et al. Heterochromatin protein Sir3 induces contacts between the amino terminus of histone H4 and nucleosomal DNA. Proc. Natl Acad. Sci. USA 110, 8495–8500 (2013).
27. Yang, Z. et al. EBS is a bivalent histone reader that regulates floral phase transition in Arabidopsis. Nat. Genet. 50, 1247–1253 (2018).
28. Qian, S. et al. Dual recognition of H3K4me3 and H3K27me3 by a plant histone reader SHL. Nat. Commun. 9, 2425 (2018).
29. Li, Z., Fu, X., Wang, Y., Liu, R. & He, Y. Polycomb-mediated gene silencing by the BAH-EMF1 complex in plants. Nat. Genet. 50, 1254–1261 (2018).
30. Zhao, D. et al. The BAH domain of BAHD1 is a histone H3K27me3 reader. Protein Cell 7, 222–226 (2016).
31. Luo X. & He Y. Experiencing winter for spring flowering: a molecular epigenetic process on vernalization. J. Integr. Plant Biol. 62, 104–117 (2019).
32. Noh, B. et al. Divergent roles of a pair of homologous jumonji/zinc-finger-class transcription factor proteins in the regulation of Arabidopsis flowering time. Plant Cell 16, 2601–2613 (2004).
33. Goodrich, J. et al. A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature 386, 41–51 (1997).

34. Yan, W. et al. Dynamic and spatial distribution of Polycomb activity by plant histone demethylases. Nat. Plants 4, 681–689 (2018).

35. Zheng, S. et al. The Arabidopsis H3K27me3 demethylase JUMONJI 113 is a temperature and photoperiod dependent flowering repressor. Nat. Commun. 10, 1303 (2019).

36. Dong, H., Howard, M. & Dean, C. Physical coupling of activation and repression activities to maintain an active transcriptional state at FLC. Proc. Natl Acad. Sci. USA 113, 9369–9374 (2016).

37. Gaudin, V. et al. Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in Arabidopsis. Development 128, 4847–4858 (2001).

38. Gomez-Mena, C. et al. early bolting in short days: an Arabidopsis mutation that causes early flowering and partially suppresses the floral phenotype of leafy. Plant Cell 13, 1011–1024 (2001).

39. Sung, S. et al. Epigenetic maintenance of the vernalized state in Arabidopsis thaliana requires LIKE HETEROCHROMATIN PROTEIN 1. Nat. Genet. 38, 706–710 (2006).

40. Mylona, S. et al. LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. Proc. Natl Acad. Sci. USA 103, 5012–5017 (2006).

41. Duan, C. G. et al. A protein complex regulates RNA processing of intrinsic heterochromatin-containing genes in Arabidopsis. Proc. Natl Acad. Sci. USA 111, 20389–20394 (2014).

42. Wang, X. et al. RNA-binding protein regulates plant DNA methylation by controlling mRNA processing at the intrinsic heterochromatin-containing gene IBB1. Proc. Natl Acad. Sci. USA 110, 15467–15472 (2013).

43. Michaels, S. D. & Amasino, R. M. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11, 949–956 (1999).

44. Sheldon, C. C. et al. The FLC MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. Plant Cell 11, 445–458 (1999).

45. Searle, I. et al. The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. Genetics 204, 898–912 (2016).

46. Suarez-Lopez, P. et al. CONSTANS mediates between the circadian clock and flowering in Arabidopsis. Genes Dev. 21, 1754–1759 (2007).

47. Imaizumi, T. & Kay, S. A. Photoperiodic control of flowering time and plant architecture in Arabidopsis. Development 143, 2661–2666 (2016).

48. Bergamin, E. et al. Molecular basis for the methylation specificity of the euchromatic histone H3 Lys 27 demethylase JMJD3. Proc. Natl Acad. Sci. USA 113, 6961–6966 (2016).

49. Brooks, A. et al. Histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. Genes Dev. 26, 1364–1375 (2012).

50. Ivadi, M. S., Karam, C. S. & Corces, V. G. Phosphorylation of histone H3 at Ser10 facilitates RNA polymerase II release from promoter-proximal pausing in Drosophila. Genes Dev. 21, 2818–2831 (2007).

51. Rossetto, D., Avvakumov, N. & Cote, J. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. Epigenetics 7, 1098–1108 (2012).

52. Duan, C. G. et al. A pair of transposon-derived proteins function in a histone acetyltransferase complex for active DNA demethylation. Cell Res. 27, 326–341 (2017).

53. Salah, A., Alvarez-Venegas, R. & Avramova, Z. An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. Nat. Protoc. 3, 1018–1025 (2008).

54. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012).

55. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44, W160–W165 (2016).

56. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 267, 307–326 (1997).

57. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D. Biol. Crystallogr. 66, 212–221 (2010).

58. Emley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. Sect. D. Biol. Crystallogr. 66, 486–501 (2010).

59. Roberts, T. C. et al. Quantification of nascent transcription by bromouridine immunocapture nuclear run-on RT-qPCR. Nat. Protoc. 10, 1198–1211 (2015).

Acknowledgements

We thank Dr. Yuehua He for the seeds of cfl-81, lhp1-3, and flc-3 mutants. We thank staff at beamline BL19U1 of the National Center for Protein Sciences Shanghai (NCPS) at the Shanghai Synchrotron Radiation Facility (SSRF) for data collection. C.-G.D. was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB20040000) and by the National Natural Science Foundation of China (31570351). J.-K.Z. was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences(XDB20040000). I.-J.S. was supported by the National Natural Science Foundation of China (31770579). S.-X.S. conducted the experiments. Y.Z.Z., J.Y., L.Z., Y.W., G.Z., and S.-X.S. conducted the experiments. Y.Z.Z., J.Y., L.Z., J.D., J.-K.Z., and C.-G.D. analyzed the data.
data. C.C. and P.L. performed bioinformatic analysis. J.D., J.-K.Z., and C.-G.D. wrote the paper.

**Competing interests**
The authors declare no competing interests.

**Additional information**
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-20089-0.

**Correspondence** and requests for materials should be addressed to J.-K.Z., J.D. or C.-G.D.

**Peer review information** Nature Communications thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[Open Access] This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit [http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/).

© The Author(s) 2020