Histone 2B monoubiquitination complex integrates transcript elongation with RNA processing at circadian clock and flowering regulators

Magdalena Woloszynska[a,b,1,2], Sabine Le Gall[a,b,1], Pia Neyt[a,b], Tommaso M. Boccardi[a,b], Marion Grasser[c], Gernot Längst[d], Stijn Aesaert[a,b], Griet Coussens[a,b], Stijn Dhondt[a,b], Eveline Van De Slijke[a,b], Leonardo Bruno[e], Jorge Fung-Uceda[f], Paloma Mas[f,g], Marc Van Montagu[a,b,3], Dirk Inzé[a,b], Kristiina Himanen[a,b,4,5], Geert De Jaeger[a,b], Klaus D. Grasser[c], and Mieke Van Lijsebettens[a,b,3]

[a]Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium; [b]Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium; [c]Department of Cell Biology and Plant Biochemistry, Biochemistry Centre, University of Regensburg, 93053 Regensburg, Germany; [d]Department for Biochemistry III, Biochemistry Centre, University of Regensburg, 93053 Regensburg, Germany; [e]Dipartimento di Biologia, Ecologia e Scienze della Terra, Università della Calabria, 87036 Arcavacata di Rende, Italy; [f]Center for Research in Agricultural Genomics, Consortium Consejo Superior de Investigaciones Científicas, Instituto Recerca i Tecnologia Agroalimentaries, Universidad Autonoma de Barcelona–Universidad de Barcelona, 08193 Barcelona, Spain; and [g]Consejo Superior de Investigaciones Científicas, 08028 Barcelona, Spain

Contributed by Marc Van Montagu, February 19, 2019 (sent for review April 25, 2018; reviewed by George Coupland, Koen Geuten, and Dorothee Staiger)

HISTONE MONOUBIQUITINATION1 (HUB1) and its paralog HUB2 act in a conserved heterotetrameric complex in the chromatin-mediated transcriptional modulation of developmental programs, such as flowering time, dormancy, and the circadian clock. The KHD1 and SPEN3 proteins were identified as interactors of the HUB1 and HUB2 proteins with in vitro RNA-binding activity. Mutants in *SPEN3* and *KHD1* had reduced rosette and leaf areas. Strikingly, in *spen3* mutants, the flowering time was slightly, but significantly, delayed, as opposed to the early flowering time in the *hub1-4* mutant. The mutant phenotypes in biomass and flowering time suggested a deregulation of their respective regulatory genes *CIRCADIAN CLOCK-ASSOCIATED1* (*CCA1*) and *FLOWERING LOCUS C* (*FLC*) that are known targets of the HUB1-mediated histone H2B monoubiquitination (H2Bub). Indeed, in the *spen3-1* and *hub1-4* mutants, the circadian clock period was shortened as observed by luciferase reporter assays, the levels of the *CCA1α* and *CCA1β* splice forms were altered, and the *CCA1* expression and H2Bub levels were reduced. In the *spen3-1* mutant, the delay in flowering time was correlated with an enhanced *FLC* expression, possibly due to an increased distal versus proximal ratio of its antisense *COOLAIR* transcript. Together with transcriptomic and double-mutant analyses, our data revealed that the HUB1 interaction with SPEN3 links H2Bub during transcript elongation with pre-mRNA processing at *CCA1*. Furthermore, the presence of an intact HUB1 at the *FLC* is required for SPEN3 function in the formation of the *FLC*-derived antisense *COOLAIR* transcripts.

H2Bub | HUB1 interactome | RNA-binding protein | RRM domain | KH domain

In eukaryotic cells, the genomic DNA is organized in nucleosomes that consist of 146-bp-long DNA wrapped around an octamer of the "core" histone dimers of H2A, H2B, H3, and H4 (1), whereas linker DNA and histone H1 connect adjacent nucleosomes. The chromatin structure is highly dynamic, with nucleosomal histone tail modifications, such as methylation, acetylation, and ubiquitination that regulate the DNA availability to the RNA polymerase II (RNAPII) transcription. The major chromatin state for active genes in *Arabidopsis thaliana* is determined by histone H2B monoubiquitination (H2Bub), histone H3 acetylation, and methylation (2). H2Bub is absent from the *Arabidopsis* promoter regions, peaks at the gene bodies, and is required for maximal gene-expression levels, implying that this histone modification is specifically linked with transcript elongation (3–6). In yeast, the E3 ubiquitin ligase Bre1 and the E2-conjugating enzyme Rad6 form a complex that catalyzes the H2B-K123 monoubiquitination (7, 8). In *Arabidopsis*, the relatives HISTONE MONOUBIQUITINATION1 (HUB1) and HUB2 E3 ubiquitin ligases and the UBC1 and UBC2 E2-conjugating enzymes work also together in a complex to monoubiquitinate H2B (9–12). In *hub1*, *hub2*, and *ubc1 ubc2* double mutants, H2Bub was reduced at the *FLOWERING LOCUS C* (*FLC*)/*MADS AFFECTING FLOWERING* (*MAF*) genes, resulting in reduced *FLC*/*MAF* gene-expression levels and early flowering (11–13). Monoubiquitinated H2B associated specifically with the gene body of the *FLC* clade genes and was necessary for enhancement of H3K4me3 and H3K36me2 and transcriptional activation in the *FLC*/*MAF* chromatin. Strikingly, the gene-expression levels of *FLC* and relatives were also reduced in the *Arabidopsis sup32*/*ubp26* deubiquitination

Significance

In eukaryotes, the genomic DNA is organized in chromatin that consists of nucleosomal units formed by histone proteins. The accessibility of genes for RNA polymerase II transcription by the dynamic modification of histone tails during the transcript elongation phase is emerging as an important regulatory mechanism. Here, we show that RNA-binding proteins, together with the conserved HISTONE MONOUBIQUITINATION1–HISTONE MONOUBIQUITINATION2 complex that mediates histone H2B monoubiquitination at the circadian clock, and flowering time-regulatory genes during transcript elongation are required for the processing of their pre-mRNA and antisense RNA, respectively.

Author contributions: M.W., S.L.G., M.V.M., K.H., and M.V.L. designed research; M.W., S.L.G., P.N., T.M.B., M.G., S.A., G.C., E.V.D.S., L.B., and J.F.-U. performed research; M.W., S.L.G., P.N., M.G., G.L., S.D., P.M., D.I., G.D.J., K.D.G., and M.V.L. analyzed data; and M.W., S.L.G., G.L., P.M., K.D.G., and M.V.L. wrote the paper.

Reviewers: G.C., Max Planck Institute for Plant Breeding Research; K.G., Katholieke Universiteit Leuven; and D.S., University of Bielefeld.

The authors declare no conflict of interest.

Data deposition: RNA sequencing data are publicly accessible in the ArrayExpress database, https://www.ebi.ac.uk/arrayexpress/ (accession no. E-MTAB-6780).

[1]M.W. and S.L.G. contributed equally to this work.

[2]Present address: Department of Genetics, Faculty of Biology and Animal Sciences, Wroclaw University of Environmental and Life Sciences, 51-631 Wroclaw, Poland.

[3]To whom correspondence may be addressed. Email: marc.vanmontagu@ugent.be or mieke.vanlijsebettens@psb.ugent.be.

[4]Present address: Department of Agricultural Sciences, University of Helsinki, 00790 Helsinki, Finland.

[5]Present address: Organismal and Evolutionary Biology Research Programme Viikki Plant Science Centre, University of Helsinki, 00014 Helsinki, Finland.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1806541116/-/DCSupplemental.

Published online March 28, 2019.

mutants that were early flowering. The accumulated H2Bub marks at the *FLC* led to a depletion of the activating H3K36me3 marks and an increase in repressive H3K27me3 marks, indicating that both H2B monoubiquitination and deubiquitylation and their steady state are critical in the deposition of other activating histone marks during transcript elongation and in the proper gene activation (14), as also demonstrated in the yeast *Ubp8* deubiquitylation mutant (15).

Proteins interacting with the H2Bub machinery might represent regulators of H2Bub dynamics, transcript elongation efficiency, target gene specificity, or might identify a link to pre-mRNA processing or upstream signaling. In humans, H2B deubiquitylation enzymes are required for efficient cotranscriptional pre-mRNA splicing (16, 17), whereas in yeast, H2B monoubiquitination and deubiquitylation enzymes genetically interact with the Npl3 SR-like protein in splicing (18).

To further elucidate the H2Bub regulation and function in plants, we purified interacting proteins with HUB1 and HUB2 as baits and detected previously uncharacterized RNA-binding motif-containing proteins. SPEN3 and KHD1 bound RNA in an in vitro assay and their respective mutants showed circadian clock and flowering-time phenotypes, two pathways that are also targeted by the HUB1/HUB2 complex activity. SPEN3 functioned in pre-mRNA processing at the *CCA1* regulatory gene in concert with HUB1/HUB2-mediated H2Bub and in antisense *COOLAIR* transcript formation at the *FLC* regulatory gene, independently of H2Bub.

Results and Discussion

KHD1 and SPEN3 Identified as Core Components of the HUB1/HUB2 Complex. To detect HUB1-associated proteins that might represent upstream regulators, cofactors, or components of the core complex, several tandem-affinity purifications (TAPs) were carried out with *Arabidopsis* cell cultures overexpressing N-terminally tagged full-length HUB1 and HUB2 proteins and a modified HUB1, with two point mutations introduced into the RING domain by replacing two cysteines by serines (C826S and C829S), resulting in HUB1pm, similarly to the yBre1 mutations (19). With HUB1 as bait, only HUB2 and the RNA-binding domain-containing SPEN3 and KHD1 were retained in several TAPs (Table 1 and Dataset S1). The identified protein list was filtered for background proteins, according to standard procedures (20), and only the IDs represented by two significant peptides in at least two independent TAP purifications were kept.

Reverse TAP with either HUB2 or SPEN3 as bait purified HUB1, KHD1, and SPEN3, or HUB1, HUB2, and KHD1 proteins, respectively. Hence, the SPEN3 and KHD1 proteins are part of the HUB1–HUB2 complex interactome. Upon introduction of point mutations into the RING domain of HUB1 (HUB1pm), no HUB2, SPEN3, and KHD1 interactions could be detected (Table 1), suggesting that the RING domain could be essential for the heterodimerization and formation of the complex with other interactors. A pairwise yeast two-hybrid system (Y2H) corroborated strong HUB1/HUB1 and HUB1/HUB2, but not HUB2/HUB2 interactions, as already reported (13).

Interestingly, a strong pairwise interaction between HUB1 and SPEN3 was demonstrated at 3 mM and 10 mM 3-AT (*SI Appendix*, Fig. S1 *A* and *B*) and required both N and C termini of SPEN3 (*SI Appendix*, Fig. S1 *B–D*). The N-terminal SPEN3 fragment weakly interacted with HUB1 as shown at 3 mM 3-AT, whereas the C-terminal SPEN3 fragment did not at all interact with HUB1 (*SI Appendix*, Fig. S1 *B–D*). Interactions with KHD1 could not be tested due to self-activation of the fusion protein. In conclusion, two RNA-binding domain proteins, SPEN3 and KHD1, were identified by TAP as integral part of the HUB1/HUB2 core complex and a strong and direct interaction for SPEN3 with HUB1 was confirmed by Y2H.

The fusion constructs GFP–SPEN3 and GFP–KHD1 were transiently expressed upon infiltration of *Nicotiana benthamiana* (tobacco) leaves with agrobacteria and stably expressed in *Arabidopsis* lines obtained by floral dip. The GFP–SPEN3 fluorescence was located exclusively in the nucleus, excluding the nucleolus, of leaf (Fig. 1*A*) and primary root (Fig. 1*B*) epidermis cells, whereas the GFP–KHD1 fluorescence occurred in the nucleus, with the exclusion of the nucleolus, in the leaf (Fig. 1*C*) and root (Fig. 1*D*) epidermis, in the cytoplasm around the nucleus, and near the plasmalemma, similar to the GFP–HUB1 and GFP–HUB2 localization (10).

The SPEN3 and KHD1 Proteins Contain RNA-Binding Domains. The SPEN3 and KHD1 interactors of the HUB1/HUB2 complex contain widely spread protein domains involved in RNA binding in eukaryotes, i.e., an RNA recognition motif (RRM) and the K homology (KH) motif, respectively (21). The *SPEN3* gene (At1g27750) is 4,326 nucleotides long, contains eight exons, and encodes a 117.47-kDa protein. Domain search performed with the InterProScan tool (www.ebi.ac.uk/interpro) revealed the presence of a conserved RRM and a Spen Paralog and Ortholog C-terminal (SPOC) protein-binding domain (*SI Appendix*, Fig. S2*A*), similarly to the mammalian Split Ends (Spen) proteins that function in transcriptional regulation, posttranscriptional processing, and nuclear export of mRNA (22). In humans, Spen proteins interact with histone modification enzymes such as histone deacetylases (23, 24) and histone methyltransferases (25). Some of the mammalian Spen proteins regulate gene expression via the control of splicing activity (26–28). In the *Arabidopsis* genome, only three Spen proteins were retrieved that combined SPOC with one or more RRM motifs: (*i*) the SPEN3 protein studied in this work with one RRM, (*ii*) the SPEN2 (*AT4G12640*) protein with two RRMs with an unknown function [because the knockout and overexpressing lines had no apparent phenotype (29)], and (*iii*) the flowering-time regulator FPA with three RRMs that controls alternative splicing and polyadenylation of antisense transcripts of the floral repressor *FLC* (30). The role in flowering-time regulation has also been suggested in rice (*Oryza sativa*) for the Spen protein encoded by the *OsRRMh* gene with two RRM

Table 1. Proteins identified by MS in TAP eluates of *Arabidopsis* cell cultures with tagged HUB1, HUB1pm (with mutated RING domain), HUB2, and SPEN3 as baits

			No. of TAPs with identified protein			
			HUB1	HUB2	KHD1	SPEN3
Bait	Tag	No.TAPs	At2G44950	At1G55250	At1G51580	At1G27750
HUB1	N-TAP	2	2	2	2	2
HUB1	N-GS	4	4	4	3	2
HUB1pm	N-TAP	3	3	0	0	0
HUB2	N-GS	2	2	2	2	2
SPEN3	C-GS	4	2	2	3	4

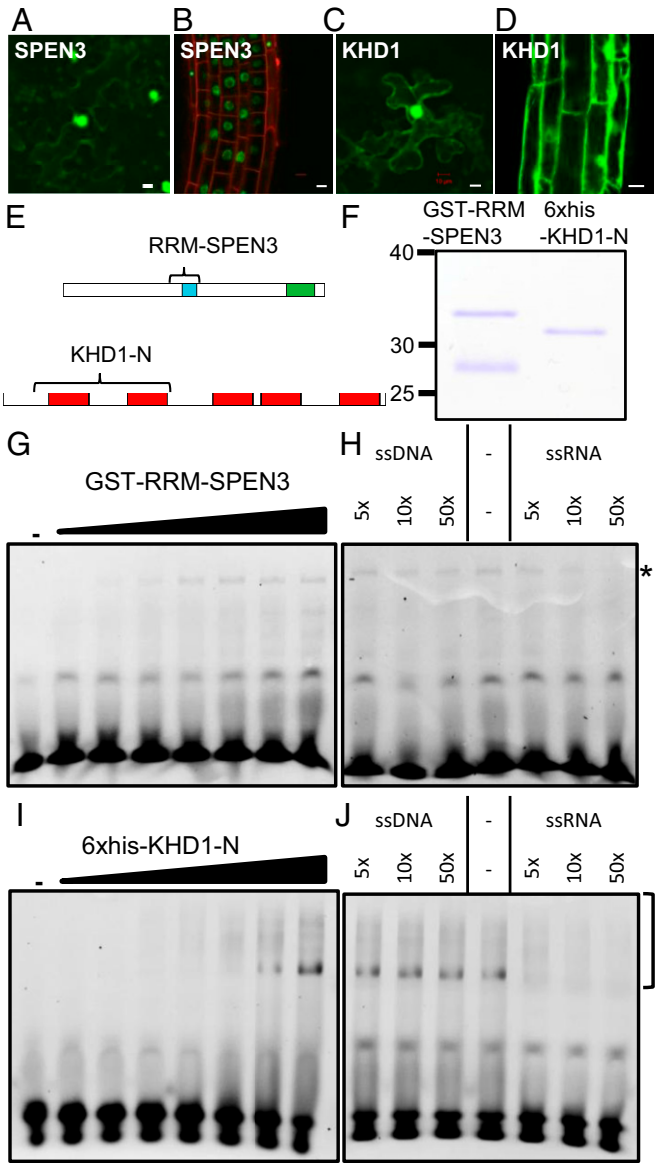

Fig. 1. SPEN3 and KHD1 localization and RNA binding. (*A–D*) Detection of GFP::SPEN3 and GFP::KHD1 upon infiltration of *N. benthamiana* leaves (*A* and *C*) and stable transgenic *A. thaliana* roots (*B* and *D*). (Scale bars, 10 μm.) (*E*) Schemes of the SPEN3 and KHD1 proteins with their functional domains, RRM (blue), SPOC (green), and KH (red). Bracket indicates the part that was expressed in *E. coli* and purified for EMSA. (*F*) SDS/PAGE analysis of purified GST-RRM-SPEN3 and 6×His-KHD1-N proteins, stained with Coomassie brilliant blue. GST-RRM-SPEN3 is partially cleaved in *E. coli* and the band migrating below the fusion protein is free GST, as determined by MS. (*G–J*) Comparison of the binding of GST-RRM-SPEN3 (*G*) or 6×His-KHD1-N (*I*) to ssRNA, and competition assay with unlabeled ssRNA or ssDNA for analysis of GST-RRM-SPEN3 (*H*) or 6×His-KHD1-N (*J*). For titrations (*G* and *I*), the Cy3-labeled 25-nucleotide probe was incubated either in the absence (lane 1) or in the presence of increasing protein concentrations (0.1, 0.2, 0.5, 1, 2, 3, and 5 μM) (lanes 2–8, respectively). Samples were analyzed by PAGE. For competition assays, the Cy3-ssRNA 25-nucleotide probe was incubated with 3 μM of protein in all samples and binding to the labeled probe was competed (as indicated) with increasing concentrations of 25-bp ssDNA or ssRNA (5×, 10×, and 50× excess over Cy3-ssRNA). The bottom bands correspond to the unbound RNA and the asterisk or brackets indicate the protein–RNA complexes.

motifs, as indicated by the flowering-time delay in the knockdown line (31). Spen proteins with one, two, or three RRM domains occur throughout eudicots. The Spen protein cladogram of five selected eudicots created with the MEGAX bioinformatics tool showed clustering of all single-RRM domain Spen proteins, including SPEN3 (*SI Appendix*, Fig. S2*A*).

The *KHD1* gene (At1g51580) is 2,491 nucleotides long, consists of seven exons, and encodes a 67.12-kDa protein. The InterProScan analysis identified five KH domains in the KHD1 protein, ranging from 70 to 77 amino acids (*SI Appendix*, Fig. S2*B*). KH motifs were found in proteins of archeae, bacteria, and eukaryota with five conserved KH domains ranging from 70 to 77 amino acids (*SI Appendix*, Fig. S2*B*). In *Arabidopsis*, 37 unique AGI codes of proteins containing 1–5 KH domains were present. All proteins with three or more KH motifs (including KHD1) had no other domains. Proteins containing only one or two KH motifs frequently combined up to eight different domains. A cladogram was generated of 12 proteins with three, four, or five KH domains (*SI Appendix*, Fig. S2*B*). KHD1 grouped with HUA ENHANCER4 (HEN4; At5g64390) that facilitates the processing of *AGAMOUS* pre-mRNA (32) and REGULATOR OF CBF GENE EXPRESSION 3/SHINY1/HIGH OSMOTIC STRESS GENE EXPRESSION 5 (RCF3/SHI1/HOS5; At5g53060) that is involved in pre-mRNA processing (33). KH proteins with three KH motifs clustered together, among which FLOWERING LOCUS KH DOMAIN (FLK) and PEPPER (PEP), an *FLC* repressor and activator, respectively (34, 35).

The *in planta* mRNA-binding proteome datasets (21, 36–38) that represent RNA interactors in a number of tissues and developmental stages contained KHD1 (21, 37), suggesting a role in poly(A) RNA-related processes. These datasets do not cover all RNA interactors, possibly accounting for the absence of SPEN3 or hinting at pre-mRNA or noncoding RNA rather than poly(A) RNA binding of SPEN3.

SPEN3 and KHD1 Bind RNA. To test the in vitro RNA-binding capabilities of the SPEN3 protein, its RRM-containing region was selected because the production of the full-length protein in *Escherichia coli* proved unsuccessful. A GST-RRM-SPEN3 fusion was expressed and purified by affinity chromatography (Fig. 1 *E* and *F*). The RNA binding of the purified GST-RRM-SPEN3 was examined by EMSAs. Incubating increasing concentrations of the recombinant protein with the fluorescently labeled single-stranded (ss)RNA, followed by EMSA analysis, revealed a dose-dependent interaction of GST-RRM-SPEN3 with the RNA probe, starting at a protein concentration of 0.2 μM (Fig. 1*G*). To test the preference of the protein for RNA, competition experiments with increasing amounts of unlabeled ssRNA or ssDNA were carried out (Fig. 1*H*). The protein–RNA complex formed at a fixed concentration of GST-RRM-SPEN3 (3 μM) and a constant amount of labeled RNA probe is competed by the addition of a 50-fold excess of unlabeled ssRNA. In contrast, protein–RNA complex formation was hardly affected by the addition of a 50-fold excess of unlabeled ssDNA, demonstrating that the GST-RRM-SPEN3 displays a preference for ssRNA over ssDNA.

To test the RNA-binding properties of KHD1, we expressed the region comprising the two N-terminal KH domains, as a 6×His-tagged fusion protein, designated 6×His-KHD1-N, in *E. coli* (Fig. 1 *E* and *F*), because the production of the full-length protein was also unsuccessful. The recombinant protein was purified by metal-chelate affinity chromatography (Fig. 1*F*). In EMSAs, a dose-dependent interaction with ssRNA was observed, starting at a protein concentration of 0.5 μM (Fig. 1*I*). In a competition assay (Fig. 1*J*), the complex formed by a fixed amount of 6×His-KHD1-N (3 μM) and the labeled ssRNA probe was efficiently competed by the addition of excess amounts of unlabeled ssRNA, whereas the addition of ssDNA did not affect the detected protein–RNA complex. Therefore, 6×His-KHD1-N preferentially interacts with the ssRNA probe.

The protein–nucleic acid interactions were additionally analyzed by means of microscale thermophoresis (MST). MST is an in-solution method to detect molecular interactions, due to the

changes in thermophoretic mobility or fluorescence of interacting molecules, in a microscopic temperature gradient (39). Increasing concentrations of GST-RRM-SPEN3 resulted in enhanced binding to ssRNA, visualized as bound fraction (*SI Appendix*, Fig. S3*A*), whereas no binding occurred to ssDNA. Due to the weak RNA-binding activity, no complete binding of the SPEN3 protein to the nucleic acids was observed at the highest protein concentration used. Because of the variance in the three biological replicates and the lack of binding saturation, no exact K_D affinity values could be derived, but the preferential binding of SPEN3 to ssRNA was evident from the sloping curve for ssRNA (*SI Appendix*, Fig. S3*A*) compared with the absence of slope for ssDNA, hence confirming the EMSA data. The qualitative analysis of 6×His-KHD1-N binding to ssRNA and ssDNA by MST revealed a complex pattern of changes in the thermophoretic mobility of the interacting molecules (for details, see *SI Appendix*, Fig. S3), but the experiment supports the EMSA results, showing an enhanced binding to ssRNA relative to ssDNA (*SI Appendix*, Fig. S3*B*).

Thus, EMSA and MST experiments showed that SPEN3 and KHD1 interact preferentially with ssRNA. RNA-binding proteins frequently contain more than one RNA-binding region (i.e., additional KH domains in the case of KHD1), possibly increasing the RNA-binding affinity (40). Hence, the relatively low affinity for ssRNA observed in our assays relates most probably to the use for technical reasons of truncated SPEN3 and KHD1 proteins (Fig. 1*E*).

Growth and Flowering Time in *spen3*, *khd1*, and *hub1/2* Mutants. The expressions of *HUB1*, *KHD1*, and *SPEN3* in the shoot apex and the root apical meristem were analyzed with whole-mount, multiprobe in situ hybridizations of 4-d-old seedlings grown in vitro with sequence-specific probes for *SPEN3*, *KHD1*, and *HUB1* genes. A red fluorochrome label was used for *SPEN3*, a green one for *KHD1*, and a blue one for *HUB1*. Interestingly, *HUB1*, *SPEN3*, and *KHD1* coexpressed strongly in the shoot apical meristem, visible as the white-pink complementary color of green, red, and blue (*SI Appendix*, Fig. S4*A*). The *SPEN3* and *HUB1* genes were coexpressed in cotelydons and *KHD1* and *HUB1* in leaf primordia and vascular tissues. In primary roots, the three genes did not coexpress, but *SPEN3* and *KHD1* coexpressed in the cortex, stele, and root apical meristem and *SPEN3* with *HUB1* in the epidermal cell layers (*SI Appendix*, Fig. S4*B*). The whole-mount in situ expression patterns of *HUB1* correlated with previously described phenotypes in *hub1* leaf, root, and flowering time (9, 13). Hence, flowering time and leaf and root growth were measured in *spen3*, *khd1*, and *hub1* single and double mutants and in overexpression lines. The SALK_025388 and GABI_626H01 lines with T-DNA insertions into exon 2 of At1g27750 (*SPEN3*) with severely reduced *SPEN3* transcript levels were designated *spen3-1* and *spen3-3*, respectively (Fig. 2 *A* and *B*). The SALK_046957 line with T-DNA insertion into the promoter, next to the 5′UTR region of At1g51580 (*KHD1*) with slightly but significantly reduced *KHD1* transcript levels, was designated *khd1-1* (Fig. 2 *A* and *B*); the SAIL_1285_H03C1 line, with T-DNA insertion into exon 3 of At1g51580 with severely reduced *KHD1* transcript levels, was designated *khd1-3* (Fig. 2 *A* and *B*).

Seedlings, 19 d after germination (DAG), grown in soil, showed slightly reduced rosette growth in *spen3-1*, *khd1-1*, and *khd1-3*, and early flowering in *hub1-4* and the double mutants, *spen3-1 hub1-4* and *khd1-1 hub1*-4 compared with the wild type [Columbia-0 accession (Col-0)] (Fig. 2*C*); reduced rosette growth was also observed in the *spen3-3* allele (*SI Appendix*, Fig. S4*C*). Flowering time was determined in four independent biological repeats in a randomization experimental set up in soil under standardized growth room conditions. The genotypes were grouped by statistical ANOVA. Strikingly, in *spen3-1*, the flowering time was significantly delayed by 2 d (Fig. 2*D*), which is an opposite phenotype compared with *hub1-4* that flowers 4 d earlier than the wild type (Fig. 2*D*), as shown previously (13); the *khd1-1* and *khd1-3* mutants had a wild-type flowering time. Also in the *spen3-3* allele, a small but significant delay in flowering time was observed in two independent biological repeat experiments under similar growth room conditions (*SI Appendix*, Fig. S4*D*). A second T-DNA insertion in the *spen3-3* mutant did not affect the gene-expression levels of its adjacent genes, At1g77920 and At1g77930 (*SI Appendix*, Fig. S5), thus, the delay in flowering time in *spen3-3* was caused by the T-DNA insertion in exon 2 of the *SPEN3* gene. Interestingly, the *spen3-1 hub1-4* and *khd1-1 hub1-4* double mutants were flowering as early as the *hub1-4* parental (Fig. 2*D*), suggesting that *HUB1* is epistatic to *SPEN3* and *KHD1* in flowering-time regulation. The small but significant delay in flowering time in *spen3-1*, the normal flowering time in *khd1-1*, and the earlier flowering in *hub1-4* were also shown in an automated weighing, imaging, and watering phenotyping platform (WIWAM XY; https://www.wiwam.be) (Fig. 2*E*). The overexpression lines of *HUB1*, *SPEN3*, and *KHD1* included in this experiment had normal flowering times (Fig. 2*E*). The number of leaves at bolting, which is tightly correlated with flowering time (41, 42), was lower in the early-flowering genotypes and slightly higher in *spen3-1* and *spen3-3* than that in the wild type, as expected (*SI Appendix*, Fig. S4 *E*–*G*). In conclusion, both HUB1/HUB2 and SPEN3 regulate flowering time, but in an opposite manner, suggesting that SPEN3 might also function in the transcriptional control of the flowering-time regulator *FLC*, which is targeted by HUB1/HUB2 for H2Bub activity (13).

Seedling growth of the mutant and overexpression lines was monitored in soil with the automated WIWAM platform. At 23 d after sowing (DAS), the projected rosette area (all green area from photographed rosette) was measured (Fig. 2*F*) and stockiness (leaf shape indicator) and compactness (projected rosette area/area of convex hull) were calculated (*SI Appendix*, Fig. S4 *H* and *I*). The projected rosette area was reduced in *spen3-1* and *khd1-1* by 16% and 17%, respectively, whereas it was similar to that of the wild type in *hub1-4* and *hub2-1* (Fig. 2*F*). In *hub1-4* and *hub2-1*, stockiness was reduced (*SI Appendix*, Fig. S4*H*) and in *spen3-1* compactness was reduced (*SI Appendix*, Fig. S4*I*). In 21-d-old in vitro-grown plants, the individual leaf area was reduced in the *hub1-4*, *hub2-1*, *khd1-1*, and *spen3-1* mutants (Fig. 2*G*), corresponding with the data obtained in soil for *spen3-1* and *khd1-1*, but differing for those for *hub1-4* and *hub2-1* (Fig. 2*F* and *SI Appendix*, Fig. S4 *H* and *I*), possibly hinting at a sensitivity of the *hub* mutants to the more stressful in vitro conditions. The leaf number decreased in the in vitro-grown *hub1-4* and *hub2-1* seedlings, but was normal in *spen3-1* and *khd1-1* (Fig. 2*G*), correlating with their flowering time measured in soil (Fig. 2 *D* and *E*). Leaf series of 26 DAG seedlings of the randomized jiffy-grown experiment at bolting (for flowering time and leaf number, see Fig. 2*D* and *SI Appendix*, Fig. S4*E*) showed that the *spen3-1 hub1-4* and *khd1-1 hub1-4* double mutants had a smaller leaf size and a narrower shape than those of the *hub1-4* parental under these conditions (*SI Appendix*, Fig. S4*J*), suggesting that *HUB1* is also epistatic to *SPEN3* and *KHD1* for these parameters. The primary root length was reduced in all genotypes, except in *hub1-4* (*SI Appendix*, Fig. S6*A*), correlating with a reduced primary root meristem size (*SI Appendix*, Fig. S6 *B* and *C*), hence, indicating that mutation and overexpression of *SPEN3* and *KHD1* affect root growth by cell proliferation.

Transcriptomes of *spen3*, *khd1*, and *hub1* Mutants. RNA deep-sequencing was done on *hub1-4*, *khd1-1*, *spen3-1*, and Col-0 total RNA prepared from shoot apices. The dataset (ArrayExpress ID E-MTAB-6780) (43) was analyzed for differentially expressed genes (DEGs), down-regulated with $\log_2$ fold-change ≤ -0.5 and up-regulated with $\log_2$ fold-change ≥ 0.5 ($P < 0.05$) (Fig. 3*A*). More than 40% and almost 50% of the DEGs in *hub1-4* and *khd1-1*, respectively, were common, suggesting that KHD1 might act

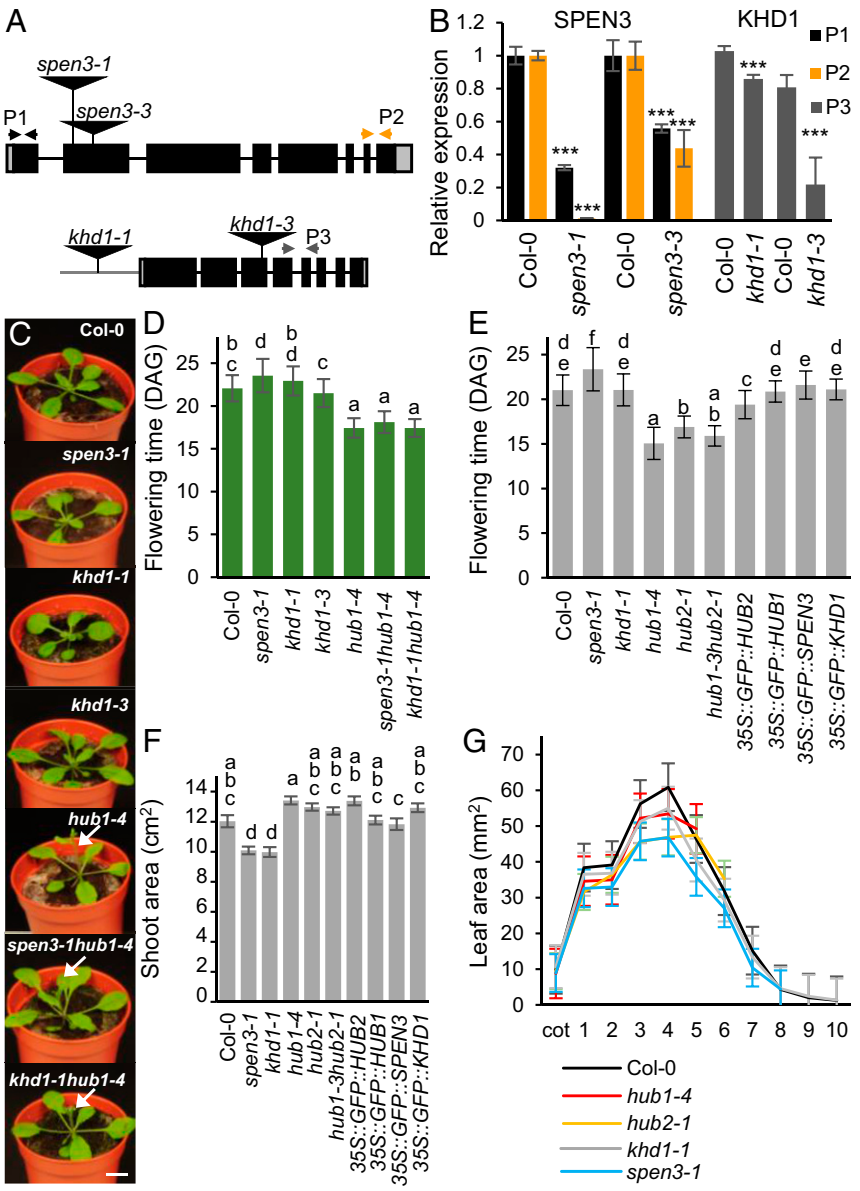

Fig. 2. Phenotypes of single and double mutants and of overexpression lines. (*A*) Schematic view of the *SPEN3* and *KHD1* T-DNA insertion lines, *spen3-1* (SALK_025388), *spen3-3* (GABI_626H01), *khd1-1* (SALK_046957), and *khd1-3* (SAIL_1285_H03C1). (*B*) Relative expression of *KHD1* and *SPEN3* in Col-0 and mutant lines estimated with the qPCR assay in six biological replicates. Asterisks indicate statistically significant differences by Student's *t* test (***$P < 0.001$). (*C*) Representative seedlings at 19 DAG grown in jiffy containers. Arrows mark early-emerging inflorescences. (Scale bar, 1 cm.) (*D*) Flowering time of lines germinated in jiffy pots ($n = 51$). (*E* and *F*) Flowering time and projected rosette area of seedlings at 23 DAS ($n = 24$) respectively, grown on the WIWAM phenotyping platform. Error bars represent SEs. Ordinary one-way ANOVA with 95% confidence shows a significant difference between the genotypes, represented by letters (*D*–*F*). (*G*) Individual leaf areas of seedlings at 21 DAS of mutant lines, in vitro grown on the IGIS platform.

together with HUB1 in the transcriptional regulation of a large number of genes.

In contrast, only ~5.5% and 16% of the DEGs in *hub1-4* and *spen3-1* were common, indicating that SPEN3 and HUB1 coregulate limited but specific classes of genes, possibly related to functions of the tissue used for RNA sequencing (such as shoot apices), and that SPEN3 controls the expression of a number of other genes via a mechanism that is unrelated to HUB1-mediated H2Bub. Very few DEGs occurred in all three mutants, hinting at a limited combined activity of KHD1, SPEN3, and HUB1 in transcriptional regulation (*SI Appendix*, Table S1). Substantial portions of the DEGs were unique to each mutant (i.e., 44% in *khd1-1*, 52% in *hub1-4*, and 63% in *spen3-1*), implying additional specific roles for HUB1, KHD1, and SPEN1.

Next, gene ontology (GO) classes were identified among down- or up-regulated genes common for all three or two mutants, or unique for only one mutant, to define the molecular functions and biological processes affected (*SI Appendix*, Table S2). Genes down-regulated in all three mutants and in the two mutants *hub1-4* and *spen3-1* fall into the same GO classes coding for cell cycle proteins, histone kinases, and ribosomal proteins. The detailed analysis of cell cycle-related genes commonly down-regulated in *hub1-4* and *spen3-1* revealed several genes crucial for cell division—*KNOLLE*, *HINKEL*, *AURORA2*, *AURORA3*, *CYCB1;4*, *CDKB1;2*, *RSW7*, *BUBR1*, and *POK2*—confirming specialized function for the HUB1–SPEN3 interaction. Genes down-regulated in *hub1-4* and *khd1-1* encode proteins related to cell cycle, DNA replication, response to stimuli, or involved in secondary metabolism. The *hub1-4*–specific down-regulated genes grouped mainly into the GO classes related to defense and stress response or cell wall organization or biosynthesis, but included also the GO flower development, containing the flowering repressors *FLC*, *FLM*, *SMZ*, and *BOP2*. Many *khd1-1* unique down-regulated genes clustered into the classes organ morphogenesis, shoot and leaf development, reproductive structure development (growth regulator *GRF9*, phase transition regulator *SPL15*, and *RBR1* and *VIM3* repressors of flowering activator *FWA*), or were involved in cell cycle or nucleic acid metabolism. The *spen3-1* uniquely down-regulated genes related to the circadian clock and flowering time (*PRR3*, *PRR5*, *ELF3*, *FKF1*, and *SRR1*). In summary, analysis of down-regulated genes showed that HUB1, KHD1, and SPEN3 are involved in common pathways, but might only occasionally regulate the same target genes in a complex.

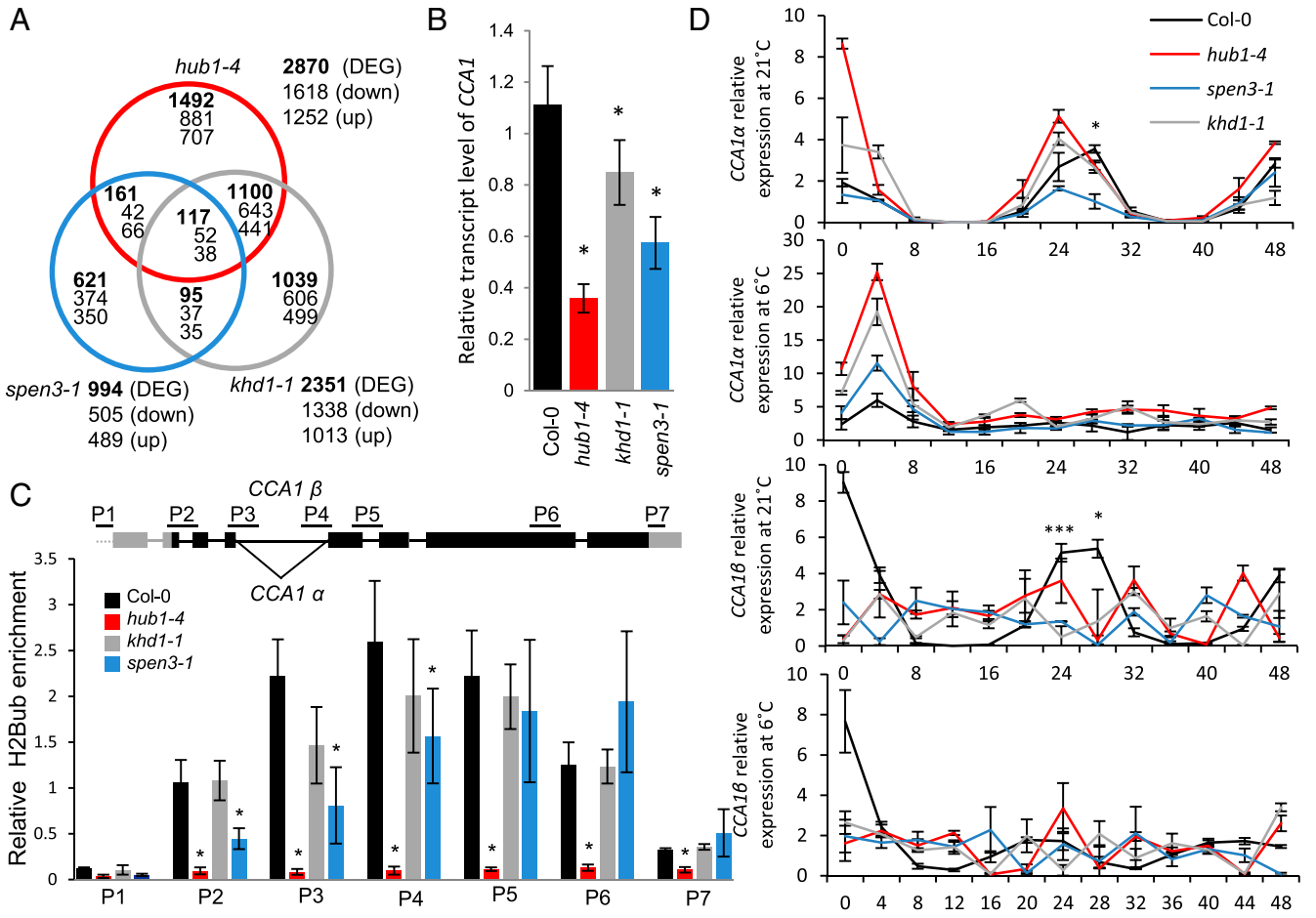

Fig. 3. Transcriptomes, *CCA1* transcripts, and H2Bub in mutants. (*A*) Venn diagram of the transcriptomes of the *hub1-4*, *spen3-1*, and *khd1-1* expression profiles compared with Col-0. (*B*) Relative expression of *CCA1* by qPCR in Col-0 and mutants in six biological replicates. (*C*) Scheme of the *CCA1* gene structure with the splice forms *CCA1α* and *CCA1β* at intron 4 and position of the primers used in ChIP-qPCR (*Upper*). Relative enrichment of H2Bub at the *CCA1* gene established with antibodies against H2Bub and four biological replicates used for ChIP assay (*Lower*). Results were normalized versus input samples. Error bars represent SDs (*B* and *C*) or SEs (*D*). Asterisks indicate statistically significant differences to Col-0 with the Student's *t* test (**P* < 0.05, *B* and *C*) or from *spen3-1* to Col-0 (**P* < 0.05, ****P* < 0.001, *D*). (*D*) Relative expression in a 48-h time course experiment of the *CCA1* splice forms, *CCA1α* and *CCA1β*, by qPCR at 21 °C and 6 °C in three biological replicates.

Because the number of genes commonly up-regulated in all three mutants was too low for a reliable GO analysis (i.e., 38 genes), we compared the *hub1-4* mutant individually with *khd1-1* or *spen3-1* (*SI Appendix*, Table S2). Both comparisons identified genes involved mainly in programmed cell death, regulatory processes, and response to different stimuli. Among the overlapping genes between the *hub1-4* and *khd1-1* transcriptomes, additional categories were detected, such as tropism, growth, cell wall, transmembrane transport, and response to hormones. Genes up-regulated only in individual mutants clustered predominantly to the response to stimulus or signaling classes, but GO categories specific for only one mutant and highly enriched were also identified: circadian clock-related genes in *hub1-4* ($\log_2$ enrichment 2.53) and growth-related genes in *khd1-1* ($\log_2$ enrichment 1.77).

SPEN3 and HUB1 Regulate *CCA1* Gene Expression Through H2Bub and Pre-mRNA Processing. The H2Bub at the clock regulator *CCA1*, a known HUB1 target (4), was determined in the *khd1-1* and *spen3-1* mutants to investigate whether it could explain the reduced *CCA1* expression in mutant seedlings (Fig. 3*B*) and whether the HUB1/HUB2-mediated histone H2B monoubiquitylase activity could be affected by KHD1 and SPEN3 during transcript elongation. Chromatin immunoprecipitation (ChIP) with H2Bub antibodies followed by a qPCR with primers annealing to the promoter and coding regions of the *CCA1* gene (Fig. 3*C*) revealed that H2Bub was absent from the promoter and peaked centrally in the gene body in the wild type (Fig. 3*C*), but was very low over the whole gene in the *hub1-4* mutant (Fig. 3*C*), characteristically for HUB1/HUB2 target genes (4). In the *spen3-1* mutant, the reduced gene expression correlated with a significant decrease in H2Bub at the 5′ and central part of the *CCA1* gene, suggesting that SPEN3 affects the HUB1-mediated H2Bub activity at the *CCA1* locus (Fig. 3*C*). In the knockdown allele, *khd1-1*, the H2Bub at the *CCA1* gene is reduced, but not significantly (Fig. 3*C*); hence, it is unclear whether KHD1 functions in HUB1-mediated H2Bub at the *CCA1* locus.

The down-regulation of the *CCA1* gene expression in *spen3-1* and *khd1-1*, as well as in *hub1-4*, prompted us to investigate circadian rhythms by means of reporter lines expressing the *LUCIFERASE* (*LUC*) fused to the promoters of *CCA1* and *TIMING OF CHLOROPHYLL A/B* (*CAB*) *EXPRESSION 1* (*TOC1*) (*pCCA1::LUC* and *pTOC1::LUC*) that were introgressed into the *spen3-1*, *khd1-1*, and *hub1-4* mutants. Bioluminescence analysis showed that for both reporters and in *hub1-4* and *spen3-1*, the amplitude of the rhythms was not altered, but the circadian period was significantly shorter than that of wild-type plants (Fig. 4). In *hub1-4* mutant plants, the advanced phase and short period were already detected

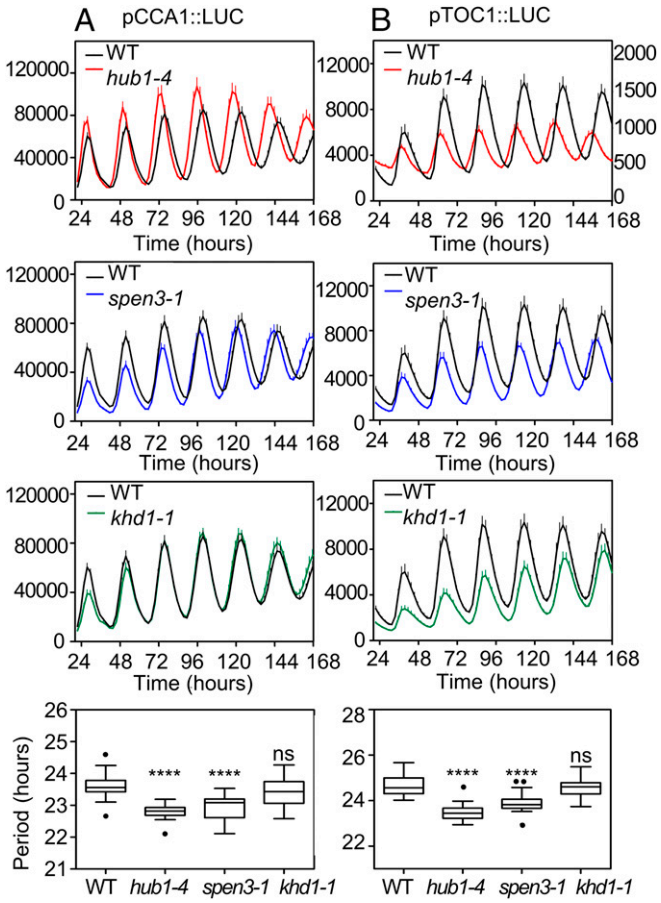

Fig. 4. Circadian waveforms and circadian period in wild type and mutants by means of LUC reporter lines. Luminescence in *hub1-4*, *spen3-1*, and *khd1-1* mutants containing p*CCA1::LUC* (A) or p*TOC1::LUC* (B) reporter genes was recorded under constant white-light conditions, following synchronization under 12-h light/12-h dark cycles; *y* axes represent counts per seedling per 5 s; lower scale in right-handed *y* axis visualizes the p*TOC1::LUC* amplitude and period in the *hub1-4* graph. Data are means + SEM of at least 12 individual seedlings. Statistical significance was calculated with an ordinary one-way ANOVA with 95% confidence value: ns, not significant, $****P \leq 0.0001$. The dots represent outliers.

from the second day under constant light, whereas in the *spen3-1* mutant plants these phenotypes were visible from the third day. The observed period shortening was not very severe (~1 h in *hub1-4* and *spen3-1*), but it was reproducible in the biological triplicates (with 12 independent seedlings per genotype and replicate) and statistically significant. The results were fully consistent among the triplicates in terms of period, phase, and amplitude, confirming the reliability of the conclusions. In *khd1-1*, the circadian period was not affected, but the amplitude of p*TOC1::LUC* was significantly reduced (Fig. 4). The short circadian period in *hub1-4* and *spen3-1* suggests that the loss of the HUB1 and SPEN3 activities makes the clock run faster than in wild-type plants.

Subsequently, the alternative splicing of *CCA1* was analyzed in the *hub1-4*, *spen3-1*, and *khd1-1* mutants by qPCR measurements of the relative transcript levels of the *CCA1α* and *CCA1β* splice forms over a time course of 48 h under continuous light and under control (21 °C) and cold (6 °C) conditions (Fig. 3*D*) (44, 45). In *CCA1α*, the fourth intron was spliced, whereas in *CCA1β* the fourth intron was retained (Fig. 3*C*). The Col-0 wild type showed rhythmic patterns for both splice forms at the control temperature with a peak at the circadian time point CT28 (Fig. 3*D*). In the cold, Col-0 loses the CT28 peak and the relative levels of *CCA1α* and *CCA1β* are reduced. In the *hub1-4*, *khd1-1*, and *spen3-1* mutants, under control conditions, the *CCA1α* splice variant peak shifts toward the time point CT24 (Fig. 3*D*), which correlates with the shorter circadian period observed with the LUC reporter lines for *hub1-4* and *spen3-1* (Fig. 4). Interestingly, under control conditions, *spen3-1* has reduced *CCA1α* transcript levels, which might reveal a positive role for SPEN3 in the processing of the *CCA1* pre-mRNA at the fourth intron. In contrast, in *hub1-4*, the *CCA1α* levels had increased, suggesting a regulatory effect of H2Bub on transcript processing. No peak for the *CCA1β* splice variant was observed in any of the mutants, in contrast to the Col-0 control with a clear peak, indicating that the intron retention mechanism necessary for the formation of the β transcript is affected in the mutants. Under cold conditions, Col-0 and the three mutants do not differ, because the rhythmic patterns of the *CCA1α* and *CCA1β* transcript levels were absent and all peaks were lost. The alternative splicing is regulated by RNA-binding proteins, chromatin structure, histone modifications, and the RNA polymerase II elongation rate (46). Slow transcript elongation expands and fast transcript elongation compresses the "window of opportunity" for recognition of upstream splice sites, thereby decreasing or increasing intron retention (47). Moreover, alternatively spliced introns are removed more slowly than constitutive introns and, therefore, their splicing requires a longer transcript elongation time.

In summary, at control temperature, the general *CCA1* transcript level is reduced in *hub1-4*, but the level of the *CCA1α* splice form is increased and that of the *CCA1β* splice form is reduced, indicating that the transcript elongation rate might be slowed down by the reduced H2Bub, enhancing splicing of intron 4 and shifting the *CCA1α*/*CCA1β* balance toward *CCA1α*. In *spen3-1*, the H2Bub at *CCA1* is reduced in the first part of the coding region until intron 4, where the intron retention/splicing occurs, suggesting that SPEN3 plays a role in the establishment of a H2Bub maximum at the splice site that might function as a signal for splice site selection. Consequently, in *spen3-1*, a decrease in H2Bub at the intron 4 splice site might slow down the transcript elongation rate and retard the splicing, which might explain the reduced total *CCA1*, *CCA1α*, and *CCA1β* transcript levels without shift in the *CCA1α*/*CCA1β* balance. In yeast, H2Bub facilitates the early spliceosome assembly at certain genes (48). Our data suggest that SPEN3 might provide an important link between the splicing machinery and the HUB1-mediated H2Bub and that SPEN3 might be an adaptor protein between the histone mark H2Bub and the splicing factors. Moreover, the RRM domain of SPEN3 might bind the nascent ssRNA and its SPOC domain might associate with proteins of the spliceosome or the transcript elongation complex. Indeed, the RNA polymerase II transcript elongation complex interacts with the mRNA-splicing factors (49). We hypothesize that the interaction between the HUB1-mediated H2Bub and SPEN3 regulates the coupling of transcript elongation and the pre-mRNA processing at *CCA1*, identifying an until now unknown transcription regulation mechanism in plants.

SPEN3 Function in *FLC*-Derived Antisense *COOLAIR* Transcript Formation. The flowering-time repressor gene *FLC* (Fig. 5*A*) is also a known target of the HUB1-mediated H2bub that promotes FLC transcription (13, 50). In the *hub1-4* mutant, the down-regulation of the *FLC* gene expression correlated with reduced H2Bub levels (Fig. 5 *B* and *C*) and early flowering time (Fig. 2 *C*–*E*). In *khd1-1*, *FLC* was down-regulated and in *spen3-1*, *FLC* was up-regulated (Fig. 5*B*); however, the H2Bub level at *FLC* was normal in both mutants (Fig. 5*C*). Thus, SPEN3 and KHD1 affect *FLC* gene expression through a mechanism that is not related to HUB1-mediated activity in H2Bub.

The FLC transcription factor prevents flowering by repression of the floral pathway integrator genes that activate the transition

to flowering and the *FLC* expression level quantitatively correlates with the time to flower (51, 52). Several regulatory pathways regulate the *FLC* expression at ambient temperatures, namely the FRIGIDA (FRI) pathway that triggers the *FLC* expression through recruitment of activating chromatin-modifying complexes, promotion of its 5′ cap cotranscriptional processing, and the autonomous pathway that down-regulates the *FLC* expression through an antisense RNA-mediated chromatin-silencing mechanism. Indeed, noncoding *FLC* antisense transcripts, designated *COOLAIR*, might fully encompass the *FLC* gene or be shorter depending on the use of distal or proximal splice and polyadenylation sites at the *FLC* (Fig. 5*A*) (30, 53, 54). Splicing and 3′ processing at the proximal site are promoted by the core spliceosome component and pre-mRNA–processing splicing factor 8 (PRP8), respectively, and by components of the autonomous pathway, including the RNA-binding proteins, FCA and FPA and the 3′ processing factors that trigger histone demethylation by FLOWERING LOCUS D (FLD) at the *FLC* gene body, which suppresses the *FLC* expression (see model in Fig. 5*E*) (54, 55).

In the *spen3* alleles, the *FLC* expression is enhanced and the flowering time is delayed, which are phenotypes similar to those in the autonomous pathway mutants or the *prp8* spliceosome-defective mutant affecting the *COOLAIR* proximal site selection/processing. Hence, we investigated the *FLC* antisense *COOLAIR* transcripts by using total RNA of *spen3-1*, *hub1-4*, and *khd1-1* mutant seedlings by means of qPCR in four biological replicates (Fig. 5*D*). In ambient temperature, the *FLC* expression was reduced in *hub1-4* and *khd1-1* and increased in *spen3-1*, confirming the results shown in Fig. 5*B*. In the *hub1-4* and *khd1-1* mutants, the proximal and distal *COOLAIR* transcript levels and their ratios were comparable to those in the Col-0 control. However, in *spen3-1*, significantly increased distal and reduced proximal transcript levels and altered distal/proximal ratios were observed (Fig. 5*D*). Thus, SPEN3 might inhibit the *COOLAIR* distal or promote the proximal transcript formation. Distal transcript secondary structure correlated with the *FLC* expression and flowering time in natural accessions (56, 57). Hence, it would be interesting to investigate whether SPEN3 affects the secondary structure or stability of distal transcripts in addition to a postulated role in distal/proximal site selection.

In our qPCR experiment (Fig. 5*D*), the 7-d cold treatment only slightly increased the proximal and distal *COOLAIR* transcripts in all genotypes, including the Col-0 control, and might have been too short for mimicking vernalization, usually 30 d of cold. Hence, as a consequence, no *COOLAIR* transcripts were induced under the cold conditions used in contrast to the vernalization conditions reported previously (58). However, the *FLC* expression was significantly reduced in all mutants upon the 7-d cold treatment (Fig. 5*D*), indicating that the applied cold conditions did induce the *FLC* repression. The Col-0 background contains a mutated *fri* allele that is responsible for a weak *FLC* expression (59) and for the attenuation of the late flowering-time phenotype in *prp8* or autonomous pathway mutants (54, 55), a possible reason for the mild but significant delay in flowering time in the *spen3-1* allele.

In conclusion, our data link the increase in *FLC* transcript in *spen3-1* to a function in distal or proximal antisense *COOLAIR* formation without effect on the H2Bub level at the *FLC*, whereas the decrease in the *FLC* transcript in *hub1-4* is only associated with a reduced H2Bub at the *FLC*. The reduced *FLC* expression in *khd1-1* does not result in a flowering-time phenotype and does not correspond to altered *COOLAIR* splice variant ratios or H2Bub levels.

The delayed flowering time in *spen3-1* and earlier flowering in the *hub1* mutants were correlated with increased and decreased *FLC* expression levels, respectively. In the *spen3-1* mutant, an increased distal antisense *COOLAIR* transcript and a distal/proximal *COOLAIR* ratio indicate that SPEN3 plays a role in *COOLAIR* polyadenylation or splicing to control the sense *FLC* transcript level and that it acts antagonistically to HUB1 in the *FLC* regulation. A number of flowering-time regulators with RNA-binding capacity, such as FPA and FCA (RRM domain), and FLK and PEP (KH domain) have been identified and are part of a regulatory loop in which FPA and FCA independently regulate 3′ end formation of antisense *COOLAIR* RNA at the *FLC* locus, triggering the *FLD* demethylation of H3K4me2, with a repressed chromatin state as a consequence (60). FLK and PEP have an antagonistic effect on the *FLC* expression, but their action mechanisms are unknown (35). They interact with the KH proteins HEN4 and HUA1 to form a complex that assists in transcript elongation and facilitates correct splicing (61). The SPOC domain in SPEN3 is important for its copurification with HUB1 and KHD1; hence, SPEN3 might function together with these proteins in the RNA-mediated control of *FLC* and might represent an antagonistic regulatory loop for the HUB1-mediated H2Bub transcriptional regulation. Both activities might cross-talk with environmental and developmental cues. The early flowering phenotype of the *spen3-1 hub1-4* double mutant, together with the molecular data, suggest that at *FLC* the presence of the intact HUB1 complex is required for SPEN3 to function in antisense *COOLAIR* transcript formation. Thus, SPEN3 represents an, until now, unidentified player requiring HUB1 to regulate *FLC* expression and flowering time.

Conclusion

We demonstrated that the HUB1 and HUB2 subunits of the H2B monoubiquitination complex and the RNA-binding proteins, SPEN3 and KHD1, strongly interact. Moreover, SPEN3 is involved in the processing of the pre-mRNA of *CCA1* and in the formation of antisense *COOLAIR* at *FLC*, key regulators of the circadian clock and flowering time, respectively. Consequently, these pathways were affected in the *spen3* knockout mutants. Our data on *CCA1* indicate that during transcript elongation the HUB1/HUB2 complex is a platform for integration of H2Bub deposition activity with pre-mRNA processing and that these processes influence each other reciprocally. Indeed, in the *spen3-1* mutant, a reduced H2Bub level at the *CCA1* intron 4 correlated with a decrease in *CCA1*α and *CCA1*β splice forms, a lowered *CCA1* expression, and a shortened circadian clock period. These results are in line with the absence of SPEN3 in the mRNA-binding proteome datasets (21) and with the composition of the RNAPII transcript elongation complex that contains transcript elongation factors (such as PAF and FACT), histone modification enzymes (such as HUB1, Elongator, and histone methyltransferases), but also spliceosome and polyadenylation factors involved in pre-mRNA processing (49). In yeast, Npl3, an arginine-serine–like protein with an RRM motif and a domain enriched in arginine-serine dipeptides, is involved in pre-mRNA processing of predominantly ribosomal proteins in close cooperation with Bre1 (the yeast HUB1 ortholog) and its activity in H2B monoubiquitination during transcription (18). We showed that HUB1 is an anchor point for the SPEN3 RNA-binding protein that functions in transcript formation at specific genes, thus providing insight into the mechanism, players, and target genes in plants. Hence, the integration of transcript elongation and transcript processing occurs through a conserved mechanism in yeast and plants, in which HUB1/Bre1 acts as anchor point for RNA-binding proteins, such as SPEN3 with a function in transcript formation. Importantly, SPEN3 is not a homolog of the yeast Npl3 and, moreover, both proteins differ in their protein-interacting domain, which might explain the difference in target proteins between plants and yeast. The particular function of SPEN3 related to the formation of the noncoding antisense *COOLAIR* transcript might be facilitated by the presence of HUB1 at the *FLC* gene during transcript elongation, whereas the SPOC domain of SPEN3 might promote interaction with RNA-processing factors.

The KHD1 interactor of HUB1 localizes in the nucleus and cytoplasm, binds to ssRNA; its mutant, *khd1-1*, has decreased leaf and

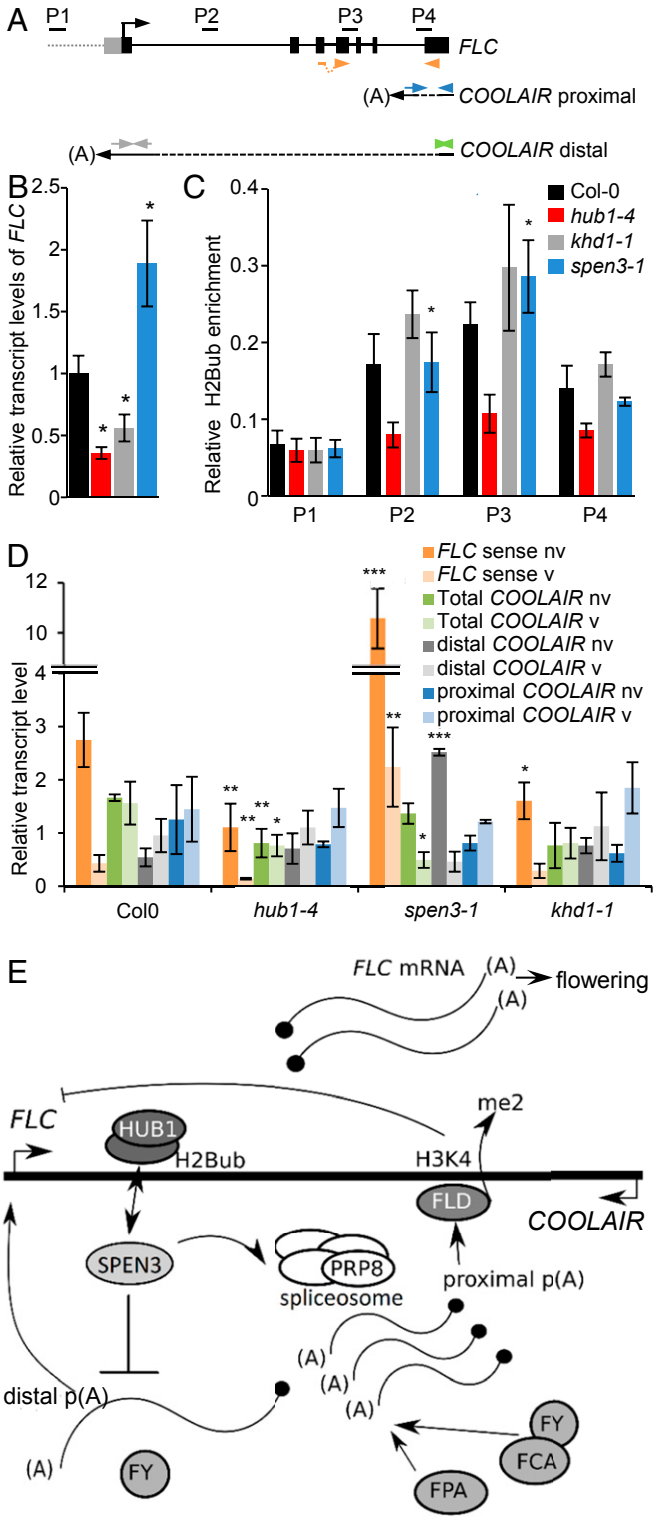

Fig. 5. *FLC* and *COOLAIR* transcripts and H2Bub in mutants. (*A*) Schematic representation of the *FLC* gene structure and antisense *COOLAIR* with primers (P2–P4) used in ChIP-qPCR and in qPCR for spliced *FLC* (orange), total *COOLAIR* (green), distal *COOLAIR* (gray), and proximal *COOLAIR* (blue). (*B*) Relative expression of *FLC* by qPCR in Col-0 and mutants in six biological replicates. (*C*) Relative enrichment of H2Bub at the *FLC* gene established with antibodies against H2Bub and four biological replicates used for ChIP assay. Results were normalized versus input samples. (*D*) qRT-PCR showing the relative expression of the *FLC* and *COOLAIR* forms before (bv) and after (av) vernalization in four biological replicates. (*E*) Hypothetical position and function of SPEN3 in *FLC*/*COOLAIR* regulation under normal conditions (model modified from ref. 60). (*B*–*D*) Error bars represent SEs. Asterisks indicate statistically significant differences to Col-0 with the Student's *t* test (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

root growth, and a transcriptome that overlapped substantially with that in *hub1-4*. However, flowering time and circadian clock period were normal in the *khd1-1* mutant, coinciding with a normal H2Bub at the *FLC* and *CCA1*, a normal splicing at the *CCA1* and *FLC*-derived antisense *COOLAIR*, despite the reduced expression of *FLC* and *CCA1*. The recently studied knockout allele *khd1-3* has a normal flowering time as well, indicating that the complete loss-of-function of KHD1 might not affect the *FLC* biological function and that *FLC* might not be a target of the KHD1 activity. We postulate that KHD1 might function in mRNA stability, export, or translation rather than in transcript formation, which is supported by its presence in an mRNA-binding interactome (21, 37).

Materials and Methods

Plant Material and Growth Conditions. The mutant lines and growth conditions are described in *SI Appendix, Materials and Methods*.

Tandem Affinity Purification and Y2H. Cloning, tandem affinity purification of protein complexes, mass spectrometry, and Y2H are described in *SI Appendix, Materials and Methods*.

RNA-Binding Assays. Cloning into the *E. coli* expression vector, protein production in *E. coli*, EMSA, and MST RNA-binding assays are described in *SI Appendix, Materials and Methods*.

Confocal Microscopy and Multiprobe in Situ Hybridization. Primary roots of 5-d-old transgenic *Arabidopsis* seedlings, transformed with 35S::GFP::SPEN3 and 35S::GFP::KHD1, were grown vertically under continuous light and analyzed by confocal microscopy (Olympus, FV10 ASW) for fusion protein localization. Multiprobe in situ hybridization is described in *SI Appendix, Materials and Methods*.

Flowering-Time Determination and Shoot and Root Growth Analyses. Flowering time was determined in several biological repeats in soil as the number of days between germination and the initiation of the floral stem elongation at 2-mm height. The number of rosette leaves produced by the apical meristem was recorded at that time ($n \geq 28$). Image acquisition and data analyses on growth parameters are described in *SI Appendix, Materials and Methods*. Leaf series were prepared from in vitro-grown plants (IGIS platform) by aligning all of the rosette leaves on 1% (wt/vol) agar plates ($n = 10$) at day 19. Leaves were photographed and scanned to measure the leaf area by ImageJ 1.41 (https://imagej.nih.gov/ij/). Root growth and root meristem analyses are described in *SI Appendix, Materials and Methods*.

RNA Methods. RNA isolation, real-time PCR, and transcriptome analysis are described in *SI Appendix, Materials and Methods*.

ChIP-qPCR. ChIP-qPCR was done as described previously (4) with 2-wk-old seedlings. For details, see *SI Appendix, Materials and Methods*.

Detection and Quantification of Polyadenylated COOLAIR. For nonvernalized samples, seedlings were grown under long-day (16-h light/8-h darkness) conditions for 10 d at 21 °C. For vernalization, seedlings were grown until 10 DAG under control conditions (long day, 21 °C), then transferred for 7 d to cold [short day (8-h light, 16-h darkness) at 7 °C], and finally allowed to recover for 7 d under control conditions (long day, 21 °C). Primer pairs used (Fig. 5*A*) are given in *SI Appendix*, Table S3.

Detection and Quantification of Alternatively Spliced *CCA1* over a 48-h Time Course. Circadian time is defined under free-running conditions (constant light) based on the previous synchronization under light/dark cycles. By convention, the onset of activity of diurnal organisms (lights on under entraining conditions) is defined as circadian time 0 (CT0). Seedlings were grown under long-day conditions for 15 d at 21 °C, then transferred at the circadian time point CT0, under continuous light conditions at 21 °C or 6 °C. Seedling pools were harvested in triplicate every 4 h from time point CT0 until CT48. Primer pairs were used to identify the two splice variants, *CCA1*α and *CCA1*β (44) (Fig. 3*D* and *SI Appendix*, Table S3).

ACKNOWLEDGMENTS. The work was supported by the European Commission Marie Curie Initial Research Training network (FP7-PEOPLE-2013-ITN-607880) (to P.M., K.D.G., and M.V.L.); the Spanish Ministry of Economy and Competitiveness, by the Generalitat de Catalunya and by the Spanish Ministry of Economy and Competitiveness through the "Severo Ochoa Program for Centers of Excellence in R&D" 2016–2019 (to P.M.); and Deutsche Forschungsgemeinschaft Grant SFB960 (to K.D.G.). M.W. was the recipient of a Marie Curie Intra-European fellowship (FP7-PEOPLE-2010-IEF-273068; acronym, LightEr) and S.D. was a postdoctoral fellow of the Research Foundation-Flanders.

1. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389:251–260.
2. Roudier F, et al. (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J* 30:1928–1938.
3. Bourbousse C, et al. (2012) Histone H2B monoubiquitination facilitates the rapid modulation of gene expression during Arabidopsis photomorphogenesis. *PLoS Genet* 8:e1002825.
4. Himanen K, et al. (2012) Histone H2B monoubiquitination is required to reach maximal transcript levels of circadian clock genes in Arabidopsis. *Plant J* 72:249–260.
5. Feng J, Shen W-H (2014) Dynamic regulation and function of histone monoubiquitination in plants. *Front Plant Sci* 5:83.
6. Van Lijsebettens M, Grasser KD (2014) Transcript elongation factors: Shaping transcriptomes after transcript initiation. *Trends Plant Sci* 19:717–726.
7. Robzyk K, Recht J, Osley MA (2000) Rad6-dependent ubiquitination of histone H2B in yeast. *Science* 287:501–504.
8. Wood A, et al. (2003) Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell* 11:267–274.
9. Fleury D, et al. (2007) The *Arabidopsis thaliana* homolog of yeast *BRE1* has a function in cell cycle regulation during early leaf and root growth. *Plant Cell* 19:417–432.
10. Liu Y, Koornneef M, Soppe WJJ (2007) The absence of histone H2B monoubiquitination in the *Arabidopsis hub1* (*rdo4*) mutant reveals a role for chromatin remodeling in seed dormancy. *Plant Cell* 19:433–444.
11. Gu X, Jiang D, Wang Y, Bachmair A, He Y (2009) Repression of the floral transition via histone H2B monoubiquitination. *Plant J* 57:522–533.
12. Xu L, et al. (2009) The E2 ubiquitin-conjugating enzymes, AtUBC1 and AtUBC2, play redundant roles and are involved in activation of *FLC* expression and repression of flowering in *Arabidopsis thaliana*. *Plant J* 57:279–288.
13. Cao Y, Dai Y, Cui S, Ma L (2008) Histone H2B monoubiquitination in the chromatin of *FLOWERING LOCUS C* regulates flowering time in *Arabidopsis*. *Plant Cell* 20:2586–2602.
14. Schmitz RJ, Tamada Y, Doyle MR, Zhang X, Amasino RM (2009) Histone H2B deubiquitination is required for transcriptional activation of *FLOWERING LOCUS C* and for proper control of flowering in Arabidopsis. *Plant Physiol* 149:1196–1204.
15. Henry KW, et al. (2003) Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev* 17:2648–2663.
16. Zhang Z, et al. (2013) USP49 deubiquitinates histone H2B and regulates cotranscriptional pre-mRNA splicing. *Genes Dev* 27:1581–1595.
17. Long L, et al. (2014) The U4/U6 recycling factor SART3 has histone chaperone activity and associates with USP15 to regulate H2B deubiquitination. *J Biol Chem* 289:8916–8930.
18. Moehle EA, Ryan CJ, Krogan NJ, Kress TL, Guthrie C (2012) The yeast SR-like protein Npl3 links chromatin modification to mRNA processing. *PLoS Genet* 8:e1003101.
19. Kim J, Roeder RG (2009) Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast. *J Biol Chem* 284:20582–20592.
20. Van Leene J, et al. (2010) Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*. *Mol Syst Biol* 6:397.
21. Köster T, Marondedze C, Meyer K, Staiger D (2017) RNA-binding proteins revisited—The emerging Arabidopsis mRNA interactome. *Trends Plant Sci* 22:512–526.
22. Ariyoshi M, Schwabe JWR (2003) A conserved structural motif reveals the essential transcriptional repression function of Spen proteins and their role in developmental signaling. *Genes Dev* 17:1909–1920.
23. Shi Y, et al. (2001) Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev* 15:1140–1151.
24. Sawada T, et al. (2008) Fusion of OTT to BSAC results in aberrant up-regulation of transcriptional activity. *J Biol Chem* 283:26820–26828.
25. Lee J-H, Skalnik DG (2012) Rbm15-Mkl1 interacts with the Setd1b histone H3-Lys4 methyltransferase via a SPOC domain that is required for cytokine-independent proliferation. *PLoS One* 7:e42965.
26. Majerciak V, Lu M, Li X, Zheng Z-M (2014) Attenuation of the suppressive activity of cellular splicing factor SRSF3 by Kaposi sarcoma-associated herpesvirus ORF57 protein is required for RNA splicing. *RNA* 20:1747–1758.
27. Hiriart E, et al. (2005) Interaction of the Epstein-Barr virus mRNA export factor EB2 with human Spen proteins SHARP, OTT1, and a novel member of the family, OTT3, links Spen proteins with splicing regulation and mRNA export. *J Biol Chem* 280:36935–36945.
28. Yan D, Perrimon N (2015) *spenito* is required for sex determination in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 112:11606–11611.
29. Solís-Guzmán MG, et al. (2017) Expression analysis of the *Arabidopsis thaliana AtSpen2* gene, and its relationship with other plant genes encoding Spen proteins. *Genet Mol Biol* 40:643–655.
30. Hornyik C, Terzi LC, Simpson GG (2010) The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. *Dev Cell* 18:203–213.
31. Liu D, Cai X (2013) *OsRRMh*, a *Spen*-like gene, plays an important role during the vegetative to reproductive transition in rice. *J Integr Plant Biol* 55:876–887.
32. Cheng Y, Kato N, Wang W, Li J, Chen X (2003) Two RNA binding proteins, HEN4 and HUA1, act in the processing of *AGAMOUS* pre-mRNA in *Arabidopsis thaliana*. *Dev Cell* 4:53–66.
33. Chen T, et al. (2013) A KH-domain RNA-binding protein interacts with FIERY2/CTD phosphatase-like 1 and splicing factors and is important for pre-mRNA splicing in Arabidopsis. *PLoS Genet* 9:e1003875.
34. Mockler TC, et al. (2004) Regulation of flowering time in *Arabidopsis* by K homology domain proteins. *Proc Natl Acad Sci USA* 101:12759–12764.
35. Ripoll JJ, et al. (2009) Antagonistic interactions between *Arabidopsis* K-homology domain genes uncover *PEPPER* as a positive regulator of the central floral repressor *FLOWERING LOCUS C*. *Dev Biol* 333:251–262.
36. Zhang Z, et al. (2016) UV crosslinked mRNA-binding proteins captured from leaf mesophyll protoplasts. *Plant Methods* 12:42.
37. Reichel M, et al. (2016) In planta determination of the mRNA-binding proteome of Arabidopsis etiolated seedlings. *Plant Cell* 28:2435–2452.
38. Marondedze C, Thomas L, Serrano NL, Lilley KS, Gehring C (2016) The RNA-binding protein repertoire of *Arabidopsis thaliana*. *Sci Rep* 6:29766.
39. Baaske P, Wienken CJ, Reineck P, Duhr S, Braun D (2010) Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew Chem Int Ed Engl* 49:2238–2241.
40. Helder S, Blythe AJ, Bond CS, Mackay JP (2016) Determinants of affinity and specificity in RNA-binding proteins. *Curr Opin Struct Biol* 38:83–91.
41. Koornneef M, Hanhart CJ, van der Veen JH (1991) A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol Gen Genet* 229:57–66.
42. Alonso-Blanco C, El-Din El-Assal S, Coupland G, Koornneef M (1998) Analysis of natural allelic variation at flowering time loci in the Landsberg *erecta* and Cape Verde Islands ecotypes of *Arabidopsis thaliana*. *Genetics* 149:749–764.
43. Van Lijsebettens M (2018) Transcriptome analysis of *khd1-1*, *spen3-1* and *hub1-4* Arabidopsis mutants. ArrayExpress. Available at https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-6780/. Deposited November 1, 2014.
44. Seo PJ, et al. (2012) A self-regulatory circuit of CIRCADIAN CLOCK-ASSOCIATED1 underlies the circadian clock regulation of temperature responses in *Arabidopsis*. *Plant Cell* 24:2427–2442.
45. Cui Z, Xu Q, Wang X (2014) Regulation of the circadian clock through pre-mRNA splicing in *Arabidopsis*. *J Exp Bot* 65:1973–1980.
46. Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T (2011) Epigenetics in alternative pre-mRNA splicing. *Cell* 144:16–26.
47. Fong N, et al. (2014) Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev* 28:2663–2676.
48. Hérissant L, et al. (2014) H2B ubiquitylation modulates spliceosome assembly and function in budding yeast. *Biol Cell* 106:126–138.
49. Antosz W, et al. (2017) The composition of the Arabidopsis RNA polymerase II transcript elongation complex reveals the interplay between elongation and mRNA processing factors. *Plant Cell* 29:854–870.
50. Crevillén P, Dean C (2011) Regulation of the floral repressor gene *FLC*: The complexity of transcription in a chromatin context. *Curr Opin Plant Biol* 14:38–44.
51. Michaels SD, Amasino RM (1999) *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11:949–956.
52. Whittaker C, Dean C (2017) The *FLC* locus: A platform for discoveries in epigenetics and adaptation. *Annu Rev Cell Dev Biol* 33:555–575.
53. Swiezewski S, Liu F, Magusin A, Dean C (2009) Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature* 462:799–802.
54. Liu F, Marquardt S, Lister C, Swiezewski S, Dean C (2010) Targeted 3′ processing of antisense transcripts triggers *Arabidopsis FLC* chromatin silencing. *Science* 327:94–97.
55. Marquardt S, et al. (2014) Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. *Mol Cell* 54:156–165.
56. Li P, Tao Z, Dean C (2015) Phenotypic evolution through variation in splicing of the noncoding RNA *COOLAIR*. *Genes Dev* 29:696–701.
57. Hawkes EJ, et al. (2016) *COOLAIR* antisense RNAs form evolutionarily conserved elaborate secondary structures. *Cell Rep* 16:3087–3096.
58. Csorba T, Questa JI, Sun Q, Dean C (2014) Antisense *COOLAIR* mediates the coordinated switching of chromatin states at *FLC* during vernalization. *Proc Natl Acad Sci USA* 111:16160–16165.
59. Johanson U, et al. (2000) Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* 290:344–347.
60. Ietswaart R, Wu Z, Dean C (2012) Flowering time control: Another window to the connection between antisense RNA and chromatin. *Trends Genet* 28:445–453.
61. Rodríguez-Cazorla E, et al. (2015) K-homology nuclear ribonucleoproteins regulate floral organ identity and determinacy in Arabidopsis. *PLoS Genet* 11:e1004983.