This is the **accepted version** of the article:

BIOLOGICAL SCIENCES – Plant biology

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], Dirk Inzé[a,b], Kristiina Himanen[a,b,3], Geert De Jaeger[a,b], Klaus D. Grasser[c], and Mieke Van Lijsebettens[a,b,4]

[a]Department of Plant Biotechnology and Bioinformatics, Gent University, 9052 Ghent, Belgium; [b]Center for Plant Systems Biology, VIB, 9052 Gent, 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 (CRAG), Consortium Consejo Superior de Investigaciones Científicas-Instituto Recerca i Tecnologia Agroalimentaries-Universidad Autonoma de Barcelona-Universidad de Barcelona (CSIC-IRTA-UAB-UB), 08193 Barcelona, Spain; [g]Consejo Superior de Investigaciones Científicas (CSIC), 08028 Barcelona, Spain.

Running title: RNA-binding protein in pre-mRNA and antisense RNA processing

Footnotes

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.

[1]These authors 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]Present address: Department of Agricultural Sciences, University of Helsinki, 00790 Helsinki, and Department of Biosciences, Viikki Plant Science Centre (ViPS), 00014 University of Helsinki, Finland.

[4]To whom correspondence should be addressed. E-mail marc.vanmontagu@ugent.be; mieke.vanlijsebettens@psb.ugent.be

0000-0001-5576-951X (M.W.), 0000-0001-8849-3812 (S.L.G.), 0000-0003-4494-6677 (P.N.), 0000-0002-7618-1974 (M.G.), 0000-0002-8232-1179 (G.L.), 0000-0001-8787-7347 (S.A.), 0000-0003-4402-2191 (S.D.), 0000-0003-2247-2976 (E.V.D.S.), 0000-0001-8260-4729 (L.B.), 0000-0003-3639-8885 (J.F.-U.), 0000-0002-3780-8041 (P.M.), 0000-0003-4711-5131 (M.V.M.), 0000-0002-3217-8407 (D.I.), 0000-0003-4975-9467 (K.H.), 0000-0001-6558-5669 (G.D.J.), 0000-0002-7080-5520 (K.D.G.), 0000-0002-7632-1463 (M.V.L.)

Abstract

HISTONE MONOUBIQUITINATION1 (HUB1) and its ortholog 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 *CCA1*α and *CCA1*β 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.

Key words: RNA-binding protein, KH domain, RRM domain, H2Bub, HUB1 interactome

Significance statement

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 HUB1-HUB2 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 respectively antisense RNA.

\body

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 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 *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 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 human, 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, Supplemental Data S1). The identified protein list was filtered for background proteins, according to standard procedures (20) and only the IDs were kept represented by two significant peptides in at least two independent TAP purifications.

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*, *C,* and *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*, *C,* and *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, likewise 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. the R̲NA r̲ecognition m̲otif (RRM) and the K̲ homology (KH) motif, respectively (21) The *SPEN3* gene (At1g27750) is 4326 nucleotides long, contains eight exons, and encodes a 117.47-kDa protein. Domain search performed with the InterProScan tool (http://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 S̲plit E̲nds (Spen) proteins that function in transcriptional regulation, posttranscriptional processing, and nuclear export of mRNA (22). In human, direct interaction of the Spen proteins with histone modification enzymes has been demonstrated, namely histone deacetylases (23, 24) and histone methyltransferase (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 motif(s), i.e. the SPEN3 protein studied in this work, with one RRM, the SPEN2 (*AT4G12640*) protein with two RRMs with a unknown function, because the knockout and overexpressing lines had no apparent phenotype (29), and 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*

8

gene with two RRM motifs, as indicated by the flowering time delay in the knockdown line (31). Spen proteins with one-, two- or three RRM domains occur throughout eudicot species. 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 2491 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 eucaryota.with five conserved K homology (KH) domains ranging from 70 to 77 amino acids (SI Appendix, Fig. S2*B*). In *Arabidopsis,* 37 unique AGI codes of proteins containing 1 to 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), a *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 the reason 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 Electrophoretic Mobility Shift Assays (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*) as compared to 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 when compared with ssDNA. 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-day-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 tissue. 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' untranslated region (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 days 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 when compared to 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 analysis of variance (ANOVA). Strikingly, in *spen3-1*, the flowering time was significantly delayed by 2 days (Fig. 2*D*), which is an opposite phenotype as compared to *hub1-4* that flowers 4 days earlier as compared to 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; www.wiwam.com) (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 days 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-day-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*; 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 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*) which correlated 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 data were

analyzed for differentially expressed genes (DEGs), down-regulated with $\log_2$ fold change (FC) $\leq$-0.5 and up-regulated with $\log_2$ fold change (FC) $\geq$ 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 together with HUB1 in the transcriptional regulation of a large number of genes.

In contrast, only approximately 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-seq (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 upregulated 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 downregulated 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 downregulated 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 downregulated 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 downregulated 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 downregulated 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 downregulated genes related to the circadian clock and flowering time (*PRR3*, *PRR5*, *ELF3*, *FKF1,* and *SRR1*). In summary, analysis of downregulated genes showed that HUB1, KHD1, and SPEN3 are involved in common pathways, but might only occasionally regulate the same target genes in a complex.

As the number of genes commonly upregulated 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 upregulated 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 with H2Bub antibodies followed by a quantitative polymerase chain reaction (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 knock-down 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 downregulation 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 from the second day under LL, whereas in the *spen1-1* mutant plants, these phenotypes were visible from the third day. The observed period shortening was not very severe (approximately 1 h in *hub1-4* and *spen1-1*), 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*) (43, 44). 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 a 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 contrary 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 (45). 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 (46). 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 (47). 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 (48). 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, 49). In the *hub1-4* mutant, the downregulation of the *FLC* gene expression correlated with reduced H2Bub levels (Fig. 5*B-C*) and early flowering time (Fig. 2 *C-E*). In *khd1-1*, *FLC* was downregulated and in *spen3-1*, *FLC* was upregulated (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 (50, 51). Several regulatory pathways regulate the *FLC* expression at ambient temperature, *i.e.*, 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 downregulates 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, 52, 53). Splicing and 3' processing at the proximal site is 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 Fig. 5*E*; 53, 54).

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 (55, 56). 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 days of 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 days of cold. Hence, as a consequence, no *COOLAIR* transcripts were induced under the cold conditions used in contrast to the vernalization conditions reported previously (57). However, the *FLC* expression was significantly reduced in all mutants upon the 7 days of 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 (58) and for the attenuation of the late flowering time phenotype in *prp8* or autonomous pathway mutants (53, 54), 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 was 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 (59). 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 (60). 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 suggests 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 (48). 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 different 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, while 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 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 Materials and Methods*.

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

Confocal Microscopy and Multiprobe *in situ* Hybridization. Primary roots of 5-day-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 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 (SI Appendix, SI Materials and Methods). Leaf series were prepared from *in vitro*-grown plants (IGIS platform) by aligning all the rosette leaves on 1% (w/v) agar plates ($n = 10$) at day 19. Leaves were photographed and scanned to measure the leaf area by ImageJ 1.41 ((http://rsbweb.nih.gov/ij/). Root growth and root meristem analyses are described in SI Appendix (Materials and Methods).

RNA methods. RNA was isolated with the RNeasy Plant Kit (Qiagen) with on-column DNase digestion. The manufacturer's protocol was modified by including two additional

washes of RNeasy spin columns with the RPE buffer. cDNA was synthesized with the iScript cDNA Synthesis Kit (Bio-Rad). Real-time PCR and transcriptome analysis are described in SI Appendix (Materials and Methods).

ChIP-qPCR. Chromatin immunoprecipitation-qPCR was done as described (4) with 2-week-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 days at 21˚C. For vernalization, seedlings were grown till 10 DAG under control conditions (long day, 21˚C), then transferred for 7 days to cold (short day [8 h light, 16 h darkness] at 7˚C), and finally allowed to recover for 7 days 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 on 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 days 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, *CCA1α* and *CCA1β* (43) (Fig. 3*D*; SI Appendix, Table S3).

References

1. Luger K, et al. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389(6648), 251-260.

2. Roudier F, et al. (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. *EMBO J* 30(10), 1928-1938.

3. Bourbousse C, et al. (2012) Histone H2B monoubiquitination facilitates the rapid modulation of gene expression during Arabidopsis photomorphogenesis. *PLoS Genet* 8(7), 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(2), 249-260.

5. Feng J, Shen W-H (2014) Dynamic regulation and function of histone monoubiquitination in plants. *Front Plant Sci* 5(3), 83.

6. Van Lijsebettens M, Grasser KD (2014) Transcript elongation factors: shaping transcriptomes after transcript initiation. *Trends Plant Sci* 19(11), 717-726.

7. Robzyk K, Recht J, Osley MA (2000) Rad6-dependent ubiquitination of histone H2B in

yeast. *Science* 287(5452), 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(1), 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(2), 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(2), 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(3), 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(2), 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(10), 2586-2602.

14. Schmitz RJ, Tamada Y, Doyle MR, Xiaoyu Z, 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(2), 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(21), 2648-2663.

16. Zhang Z, et al. (2013) USP49 deubiquitinates histone H2B and regulates cotranscriptional pre-mRNA splicing. *Genes Dev* 27(14), 1581-1595.

17. Long, L., Thelen, J.P., Furgason, M., Haj-Yahya, M., Brik, A., Cheng, D., Peng, J., and Yao, T. (2014). The U4/U6 recycling factor SART3 has histone chaperone activity and associates with USP15 to regulate H2B deubiquitination. *J. Biol. Chem*. 289 (13), 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(11), 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 (31) 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(1), 397.

21. Köster T, Marondedze C, Meyer K, Staiger D (2017) RNA-binding proteins revisited – the emerging Arabidopsis mRNA interactome. *Trends Plant Sci* 22(6), 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(15), 1909-1920.

23. Shi Y, et al. (2001) Sharp, an inducible cofactor that integrates nuclear receptor repression and activation. *Genes Dev* 15(9), 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(39), 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(8), 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(11), 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 (44), 36935-36945.

28. Yan D, Perrimon N (2015) *spenito* is required for sex determination in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 112(37), 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(3), 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(2), 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(9), 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(1), 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(10), e1003875.

34. Mockler TC, et al. (2004) Regulation of flowering time in *Arabidopsis* by K homology domain proteins. *Proc Natl Acad Sci USA* 101(34), 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(2), 251-262.

36. Zhang Z, et al. (2016). UV crosslinked mRNA-binding proteins captured from leaf mesophyll protoplasts. *Plant Methods* 12(11), 42.

37. Reichel M, et al. (2016) In planta determination of the mRNA-binding proteome of Arabidopsis etiolated seedlings. *Plant Cell* 28(10), 2435-2452.

38. Marondedze, C., Thomas, L., Serrano, N.L., Lilley, K.S., and Gehring, C. (2016). The RNA-binding protein repertoire of *Arabidopsis thaliana*. Sci. Rep. 6, 29766.

39. Baaske P, et al. (2010) Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew Chem Int Ed* 49(12), 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 (6), 83-91.

41. Koornneef M, Hanhart CJ, van der Veen JH (1991) A genetic and physiologial analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol. Gen. Genet*. 229(1), 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(2), 749-764.

43. 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(6), 2427-2442.

44. Cui Z, Xu Q, Wang X (2014) Regulation of the circadian clock through pre-mRNA splicing in *Arabidopsis*. *J Exp Bot* 65(8), 1973-1980.

45. Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T (2011) Epigenetics in alternative pre-mRNA splicing. *Cell* 144(1), 16-26.

46. Fong N, et al. (2014) Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev* 28(23), 2663-2676.

47. Hérissant L, et al. (2014) H2B ubiquitylation modulates spliceosome assembly and function in budding yeast. *Biol Cell* 106(4), 126-138.

48. 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(4), 854-870.

49. 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(1), 38-44.

50. Michaels SD, Amasino RM (1999) *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11(5), 949-956.

51. Whittaker C, Dean C (2017) The *FLC* locus: a platform for discoveries in epigenetics and adaptation. *Annu Rev Cell Dev Biol* 33, 555-575.

52 Swiezewski S, Liu F, Magusin A, Dean C (2009) Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature* 462(7274), 799-802.

53. Liu F, Marquardt S, Lister C, Swiezewski S, Dean C (2010) Targeted 3' processing of antisense transcripts triggers *Arabidopsis FLC* chromatin silencing. *Science* 327(5961), 94-97.

54 Marquardt S, et al. (2014) Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. *Mol Cell* 54(1), 156-165.

55. Li P, Tao Z, Dean C (2015) Phenotypic evolution through variation in splicing of the noncoding RNA *COOLAIR*. *Genes Dev* 29(7), 696-701.

56. Hawkes EJ, et al. (2016) *COOLAIR* antisense RNAs form evolutionarily conserved elaborate secondary structures. *Cell Rep* 16(12), 3087-3096.

57 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(45), 16160-16165.

58. Johanson U, et al. (2000) Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* 290(5490), 344-347.

59 Ietswaart R, Wu Z, Dean C (2012) Flowering time control: another window to the connection between antisense RNA and chromatin. *Trends Plant Sci* 28(9), 445-453.

60. Rodríguez-Cazorla E, et al. (2015) K-homology nuclear ribonucleoproteins regulate floral organ identity and determinacy in Arabidopsis. *PLoS Genet* 10(2), e1004983.
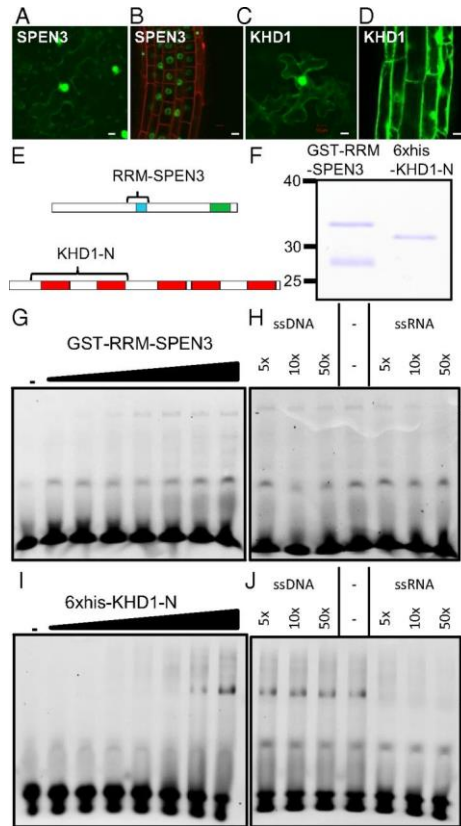
Fig. 1. SPEN3 and KHD1 localization and RNA binding.
(*A-D*) Detection of GFP::SPEN3 and GFP::KHD1 upon infiltration of *N. benthamiana* leaves (*A* and *C*) and stable transgenic *A. thaliana* roots (*B* and *D*). Scale bars = 10 μm. (*E*) Schemes of the SPEN3 and KHD1 proteins with their functional 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 mass spectrometry. (*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 μM, 0.2 μM, 0.5 μM, 1 μM, 2 μM, 3 μM, 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.
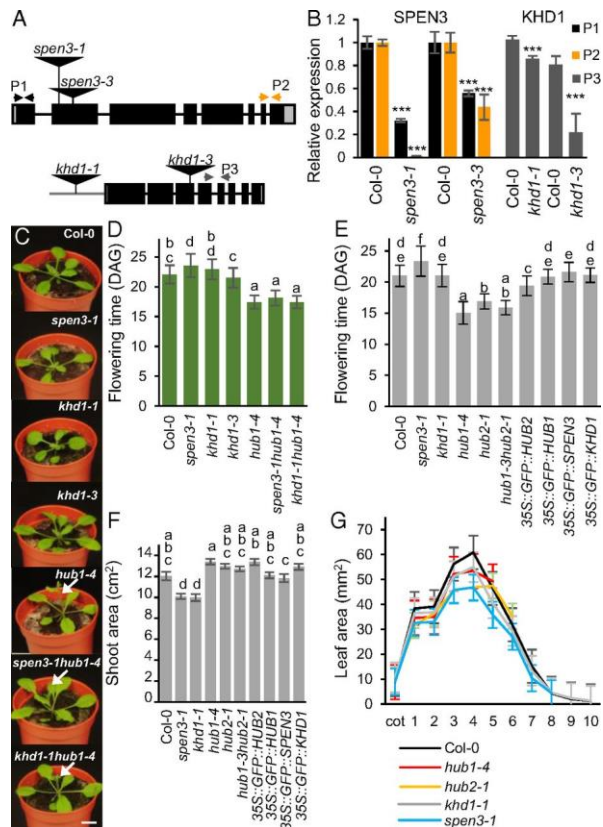
Fig. 2. Phenotypes of single and double mutants and of overexpression lines.
(*A*) Schematic view of the *SPEN3* and *KHD1* T-DNA insertion lines, *spen3-1* (SALK_025388), *spen3-3* (GABI_626H01), *khd1-1* (SALK_046957), and *khd1-3* (SAIL_1285_H03C1). (*B*) Relative expression of *KHD1* and *SPEN3* 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 (**$P < 0.01$, ***$P < 0.001$). (*C*) Representative seedlings at 19 DAG grown in jiffy containers. Arrows mark early emerging inflorescences. (*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 standard errors. Ordinary one-way ANOVA with 95% confidence shows a significant difference between the genotypes, represented by letters (panels *D-F*). (*G*) Individual leaf areas of seedlings at 21 DAS of mutant lines, *in vitro* grown on the IGIS platform.
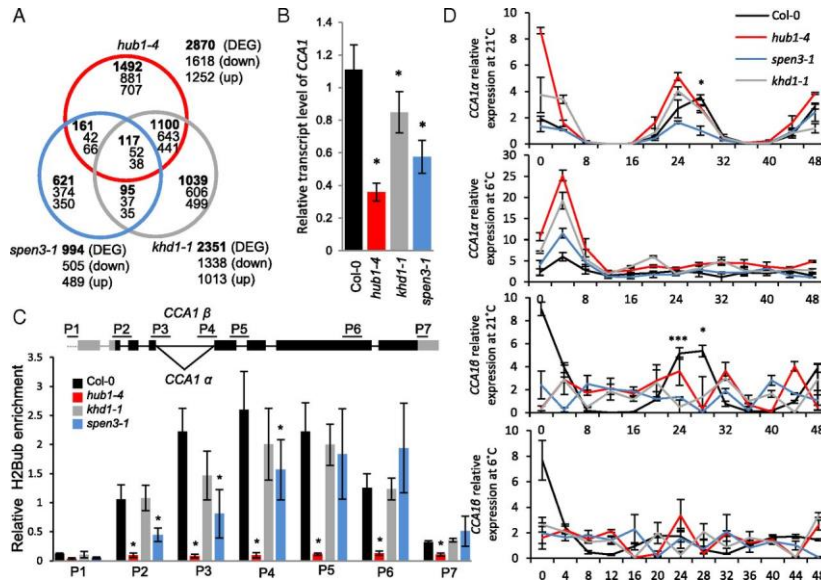
Fig. 3. Transcriptomes, *CCA1* transcripts, and H2Bub in mutants.
(*A*) Venn diagram of the transcriptomes of the *hub1-4*, *spen3-1,* and *khd1-1* expression profiles
compared to Col-0. (*B*) Relative expression of *CCA1* by qPCR in Col-0 and mutants in six biological
replicates. (*C*). Scheme of the *CCA1* gene structure with the splice forms *CCA1α* and *CCA1β* at intron 4
and position of the primers used in ChIP-qPCR (top). Relative enrichment of H2Bub at the *CCA1* gene
established with antibodies against H2Bub and four biological replicates used for ChIP assay (bottom).
Results were normalized versus input samples. Error bars represent standard errors. Asterisks indicate
statistically significant differences from *spen-1* to Col-0 with the Student's *t*-test (*$P < 0.05$, **$P < 0.01$,
***$P < 0.001$) (*B)* and to Col-0 (*$P < 0.05$) (*C*). (*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.
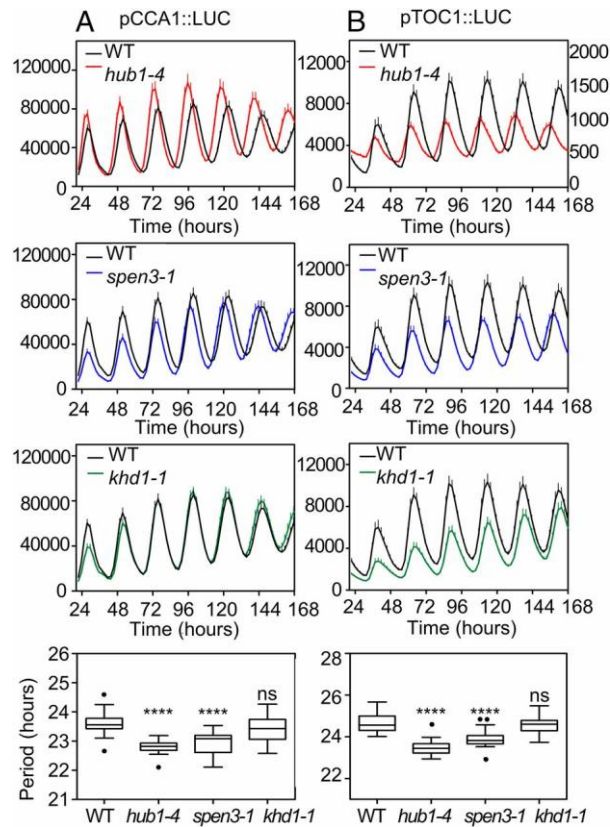
Fig. 4. Circadian waveforms and circadian period in wild type and mutants by means of LUC reporter lines.

Luminescence in *hub1-4*, *spen3-1,* and *khd1-1* mutants containing *pCCA1::LUC* (A) or *pTOC1::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/seedling/5 seconds; lower scale in right-handed y-axis visualizes *pTOC1::LUC* amplitude and period in *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, $P > 0.05$, $*P \le 0.05$, $**P \le 0.01$, $*** P \le 0.001$, $****P \le 0.0001$.
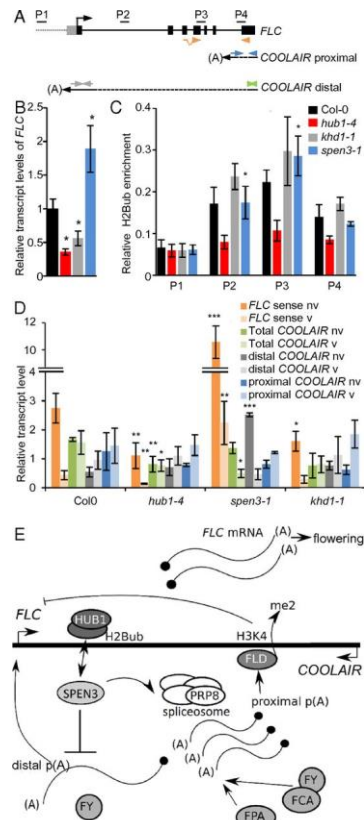
Fig. 5. *FLC* and *COOLAIR* transcripts and H2Bub in mutants.
(*A*) Schematic representation of the *FLC* gene structure and antisense *COOLAIR* with primers (P2-P4) used in ChIP-qPCR and in qPCR for spliced *FLC* (orange), total *COOLAIR* (green), distal *COOLAIR* (grey), and proximal *COOLAIR* (blue). (*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 59). (*B-D*) Error bars represent standard errors. Asterisks indicate statistically significant differences to Col-0 with the Student's *t*-test (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

Supplementary Information for

Histone 2B monoubiquitination complex integrates transcript elongation with splicing at circadian clock and flowering time regulators

Magdalena Woloszynska, Sabine Le Gall, Pia Neyt, Tommaso M. Boccardi, Marion Grasser, Gernot Längst, Stijn Aesaert, Griet Coussens, Stijn Dhondt, Eveline Van De Slijke, Leonardo Bruno, Jorge Fung-Uceda, Paloma Mas, Marc Van Montagu, Dirk Inzé, Kristiina Himanen, Geert De Jaeger, Klaus D. Grasser, and Mieke Van Lijsebettens

Corresponding author: Mieke Van Lijsebettens
Email: mieke.vanlijsebettens@psb.ugent.be

This PDF file includes:

Supplementary text
Figs. S1 to S6
Tables S1 to S3
SI References

Other supplementary materials for this manuscript include the following:

Dataset S1

MATERIALS AND METHODS

Plant material and growth conditions. *hub1-4* (SALK_122512), *spen3-1* (SALK_025388) (1), *spen3-3* (GABI_626H01) (2), *khd1-1* (SALK_046957) (1), *hub1-3* (GABI_276D08) (2), *hub1-4* (SALK_122512) (1), *hub2-1* (GABI_634H04) (2), *khd1-3* (SAIL_1285_H03C1) (3) in Col-0 background were obtained from the Nottingham *Arabidopsis* Stock Centre. T-DNA insertions were confirmed by PCR. A second T-DNA in *spen3-3* (GABI_626H01) was located at the 3' UTR of the *At1g77920* and the 5' end of the *At1g77930* genes, but did not affect their respective gene expression levels (Fig. S4). The *hub1-3 hub2-1* mutant had been described previously (4), whereas the *spen3-1 hub1-4* and *khd1-1 hub1-4* double mutants were constructed in this work and the genotypes were verified by PCR. The clock reporter lines expressing *pCCA1::LUC* (5) and *pTOC1::LUC* (6) were crossed into the *hub1-4*, *spen3-1,* and *khd1-1* mutants and homozygous lines analyzed by *in vivo* luminescence assays. The *p35S::GFP::SPEN3* and *p35S::GFP::KHD1* constructs were obtained by Gateway recombination and were transformed into *Agrobacterium tumefaciens* cells that were used for tobacco (*Nicotiana benthamiana*) leaf infiltration and stable transformation into *Arabidopsis thaliana* (L.) Heynh., accession Columbia-0 (Col-0) by floral dip.

Seedlings were grown in soil (jiffy containers) under growth chamber conditions, namely 16 h day/8 h night with white light and 21°C for flowering time experiments. Seeds for *in vitro* time-lapse analysis on the IGIS platform (7) were sterilized in 3% (v/v) bleach for 15 min and sown on medium containing half-strength Murashige and Skoog (MS) medium (Duchefa), solidified with 0.9 g L^{-1} plant tissue culture agar (Lab M) on round Petri dishes, stratified for 2 days, then incubated in a growth chamber under long-day conditions (16 h light, 8 h darkness) at 21°C. The average light intensity supplied by cool-white fluorescent tubes (Spectralux Plus 36W/840; Radium) was approximately 60 µmol m^{-2} s^{-1} for *in vitro* and in soil-grown experiments. For bioluminescence assays, plants were stratified for 3 days at 4°C on plates with MS agar medium and grown for 7 days under 12 h light, 12 h dark cycles with 60 µmol m^{-2} s^{-1} white light at a constant 22°C temperature. Seedlings were transferred to 96-well plates containing MS agar and 3 mM luciferine (Promega). Luminescence rhythms were monitored under constant white light conditions (60 µmol m^{-2} s^{-1}) with a luminometer LB-960 (Berthold Technologies) and analyzed with the software MikroWin 2000, version 4.34 (Mikrotek Laborsysteme).

Growth parameters and flowering time were also measured by means of an automated weighing, imaging, and watering phenotyping platform, acronym WIWAM XY (www.wiwam.com) according to established protocols (8, 9). WIWAM was placed in an *Arabidopsis* growth room at 21°C, 55% relative humidity, 16 h day/8 h night, and 100 mol m^{-2} s^{-1} light intensity. Seeds were stratified 2 days before sowing in pots with 80-90 g of soil that were randomized by the WIWAM platform. During the entire experiment, the soil water content was set at a constant value of 2.19 g H$_2$O/g dry soil. Images were acquired for each pot on a daily basis and analyzed. The data were validated with the in-house Interface for Plant Phenotype Analysis (PIPPA) (10).

Circadian clock period calculation. The circadian periods were calculated with the Fast Fourier Transform–Non-Linear Least-squares (FFT-NLLS) suite of the BioDare online data repository (https://biodare2.ed.ac.uk/documents/period-methods) (11). This suite is commonly accepted and widely used in the circadian community and the method is clearly described on the web page.

Essentially, the period estimation is based on curve fitting. FFT NLLS starts with a model with a single cosine and determines the parameters ($\tau1$, $\phi1$, $\alpha1$, and c) by means of a non-linear least squares fitting algorithm. This procedure is repeated with models with additional cosine components (increased N), until addition of a supplemental cosine term does not improve significantly the resulting fit. Once the best model and its parameters have been found, the period is taken to be the period of the cosine component lying within a user-defined range of likely circadian periods (typically 15-35 h).

Bioinformatic analysis. With the PLAZA 2.5 bioinformatic tool, common down- or upregulated genes were classified into significantly overrepresented ($P < 0.05$) gene ontology (GO) classes of the Biological Process type (12).

Multiprobe *in situ* hybridization. Short and specific (GSTs) fragments of the *HUB1* (At2g44950), *SPEN3*(At1g27750), and *KHD1* (At1g51580) genes were cloned in the pGEM-T-Easy vector (Promega). Labeled RNA probes were synthesized by means of *in vitro* transcription in the presence of digoxigenin-11-UTP (*SPEN3* probe), biotin-16-UTP (*KHD1* probe), fluoroscein-12-UTP (*HUB1* probe) and processed as reported (13). Four-day-old *Arabidopsis* Col-0 seedlings were fixed, dehydrated, and handled as described (13). *In situ* hybridization was done with a mix of riboprobes and hybridized by *in situ* whole-mount methodology (13) with minor modifications, namely the hybridization step was carried out overnight at 55°C and the mixture of primary and secondary antibodies was diluted 1:500. Samples were imaged with a Leica inverted TCS SP8 confocal scanning laser microscope. The Alexa fluor dyes were detected simultaneously by combining the settings indicated in the sequential scanning facility of the microscope.

Root growth analysis. Root growth was measured on seedlings grown vertically on half-strength MS medium supplemented with 1% (w/v) sucrose, 0.8% (w/v) agarose, pH 5.7, at 21°C under 24-h light conditions (75 µmol m^{-2} s^{-1}). The root meristem size was determined 5 days after germination (DAG) as the number of cells in the cortex cell file from the quiescent center to the first elongated cell (14). The samples were mounted with clearing solution (80 g chloral hydrate, 30 ml glycerol, and 10 ml dH$_2$O) and observed immediately. Root length was marked at 10 DAG and measured with the ImageJ software (http://rsbweb.nih.gov/ij/). Means between samples were compared by a two-tailed Student's *t*-test and variances with an ANOVA.

Tandem Affinity Purification. TAP (15) or GS (16) tags were fused N-terminally to full length cDNAs of HUB1, HUB1pm, and HUB2, and C-terminally to SPEN3. In the HUB1pm, two cysteines of the RING domain (positions 826 and 829) were replaced by serines. The TAP-tagged HUB1 proteins were enzymatically active and complemented partially the *hub1-1* mutation (4). Tagged transgenes were expressed under the control of the constitutive cauliflower tobacco mosaic virus 35S promoter and transformed in *Arabidopsis* cell suspension cultures (17). Protocols of proteolysis and peptide isolation, acquisition of mass spectra by a 4800 Proteomics Analyzer (Applied Biosystems), and mass spectrometry-based protein homology identification based on the TAIR genomic database, were as described (18). Experimental background proteins were subtracted based on approximately 40 TAP experiments on wild-type cultures and cultures expressing the TAP-tagged mock proteins GUS, RFP, and GFP (18).

Production of recombinant proteins. The RRM domain-containing region of SPEN3 and the two KH domain-containing region of KHD1 were amplified by PCR with *HiFi* DNA polymerase (KAPA Biosystems) and the iProof high-fidelidy PCR kit (Bio-Rad), respectively, with an *Arabidopsis* cDNA library as template and primers providing the required restriction enzyme cleavage sites (Table S3). The amplified PCR fragment of SPEN3 was digested with *Bam*HI/*Sal*I and cloned into the *Bam*HI/*Sal*I-digested *E. coli* expression plasmid pGEX-5X-1 (GE Healthcare), providing an N-terminal GST with the pGEX-5X-1-RRM-SPEN3 plasmid for the RRM domain of SPEN3 as a result. The obtained PCR fragment of KHD1 was digested with *Bam*HI/*Sal*I and cloned into the *Bam*HI/*Sal*I-digested *E. coli* expression plasmid pQE9 (Qiagen), providing an N-terminal 6×His-tag, resulting in the pQE9-KHD1-N plasmid for the N-end part of KHD1 that contains two KH domains. Plasmid constructions were checked by DNA sequencing. For protein production, the pGEX-5X-1-RRM-SPEN3 expression vector was transformed into *E. coli* BL21+pRARE cells. After induction by 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG; Sigma-Aldrich), the GST-tagged RRM-SPEN3 was purified by glutathione-sepharose affinity chromatography as previously described (19). *E. coli* M15 cells were transformed with the pQE9-KHD1-N expression vector. After induction by 1 mM IPTG, the 6×His-tagged KHD1-N was purified by metal-chelate chromatography with Ni-NTA agarose (Qiagen) from *E. coli* lysates essentially as described previously (20). By means of PD10 columns (Pharmacia), the purified proteins were collected in buffer (10 mM phosphate buffer, pH 7.0, 1 mM EDTA, 1 mM dithiothreitol [DTT], and 0.5 mM phenylmethanesulfonyl fluoride [PMSF; Sigma-Aldrich]) and the recombinant proteins were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and mass spectrometry.

Fluorescent Electrophoretic Mobility Shift Assay (EMSA) binding. RNA binding of the recombinant proteins was examined by EMSA as described (20) with fluorescently labeled ssRNA oligonucleotides (Table S3) (21) used before to study general RNA interactions of various *Arabidopsis* proteins (20, 21). Different protein concentrations were incubated for 15 min with the Cy3-labeled ssRNA (25 nM) probe in binding buffer (10 mM Hepes, pH 7.9, 3% [w/v] Ficoll, 10 mM $MgCl_2$, 5 mM KCl, 200 mM NaCl, 1 mM EDTA, 0.5 mM DTT, 1 mM spermidine, 0.1 mg/ml bovine serum albumin). Binding reactions were analyzed in 1× Tris/borate/EDTA (TBE) polyacrylamide gels. The RNA was visualized by imaging with a Typhoon 8600 instrument (GE Healthcare). Competition assays were done with constant protein concentrations (3 µM) and labeled ssRNA probes and increasing concentrations of unlabeled ssRNA or ssDNA.

Fluorescent MicroScale Thermophoresis (MST) binding assay. MST binding experiments were carried out essentially as previously described (20) with 200 nM 25-nucleotide-long Cy3-labeled ssRNA or ssDNA oligonucleotides. MST measurements were done in protein buffers with a protein concentration range at 40% MST power, 50% LED power in standard capillaries at 25°C on a Monolith NT.115 device (NanoTemper Technologies). The data were analyzed with the MO.Affinity Analysis software (V2.3, NanoTemper Technologies) and binding reactions were determined by examining Temperature Related Intensity Changes (TRIC effect). To calculate the fraction bound, the ΔF_{norm} value of each point was divided by the amplitude of the fitted curve, resulting in values from 0 to 1 (0 = unbound, 1 = bound), and processed with the KaleidaGraph 4.5 (Synergy Software).

RNA methods. RNA was isolated with the RNeasy Plant Kit (Qiagen) with on-column DNase digestion. The manufacturer's protocol was modified by two additional washes of RNeasy spin columns with RPE buffer. Complementary DNA (cDNA) was synthesized with the SuperScript III First-strand Synthesis Kit (Life-Invitrogen, CAT. 18080051).

Real-time PCR was run in technical triplicates with the LightCycler 480 SYBR Green I Master (Roche Life Science) and the Janus robot (PerkinElmer) for pipetting. The LightCycler 480 Real-Time PCR System was used for amplification (95°C for 10 min, 45 cycles of 95°C/10 s, 60°C/15 s, 72°C/30 s followed by a melting curve analysis). The qPCR results were analyzed with the qBase Plus software (Biogazelle). The PP2A (At1g13320) and UBC (At5g25760) genes were used as references for gene expression normalization. The primer sequences used are presented in Table S3. For the transcriptome, RNA was extracted from shoot apices of 10-day-old seedlings. The A260/A280 and the A260/A230 ratios were measured with a NanoDrop spectrophotometer (Thermo Fisher Scientific) to evaluate the quantity and purity of the samples. Additionally, the high RNA quality was verified by means of the Agilent Bioanalyzer system. After the library preparation by TruSeq, RNA was sequenced on the Illumina HiSeq. Normalization statistics, and bioinformatics were carried out on the raw data to allow pairwise differential gene expression analyses (Nucleomics Core Facility, VIB, Leuven, Belgium). The Gene Ontology categories of the differentially expressed genes were identified with the PLAZA 2.5 software (12).

ChIP-qPCR. The isolated chromatin was sonicated in a Vibra-cell sonicator (Sonics & Materials) with four 15-s pulses at a 20% amplitude and immunoprecipitated with 5 µg of H2Bub antibodies (Medimabs, MM-029). Protein A Agarose (Millipore) was used to collect immunoprecipitated chromatin. After reverse cross-linking and proteinase K digestion, DNA was purified with the MinElute PCR Purification Kit (Qiagen) and eluted with the elution buffer supplemented with RNaseA (10 µg/ml). Samples were analyzed by real-time qPCR with primers in the promoter and coding regions of the *FLC* (22) and *CCA1* (Table S3) genes. The amount of immunoprecipitated DNA was calculated relative to the input.

Yeast Two-Hybrid Analysis. Constructs used for Y2H were obtained by cloning cDNAs of the HUB1, HUB2, SPEN3, SPEN3 N-terminus (761 amino acids, including RRM domain), and SPEN3 C-terminus (488 amino acids, including SPOC domain) by the Gateway Technology (Life Technologies). Constructs were introduced by an LR recombination into the pDEST™22 and pDEST™32 destination vectors, resulting in fusions to the GAL4 activation domain (AD) and GAL4-binding domain (BD), respectively (ProQuest™ Two-HybridSystem, Life Technologies). All plasmids were transformed into yeast strains with the opposite mating types MaV203 MATa and MATα. Transformed yeast strains were selected for the presence of pDEST22 or the pDEST32 vector.

Transformed yeast strains were selected for the presence of pDEST22 or the pDEST32 vector and the abundance of the fusion proteins was assessed by Western-blot. Self-activation of both bait and prey constructs were tested by yeast transformation for non-PREY-specific activation of reporter gene expression by BAIT-constructs or BAIT-GAL-4-DNA-independent activation exerted by PREY constructs in a colony-lift filter assay with X-Gal as substrate. Fusion proteins that showed self-activation, *i.e.*, HUB2 and KHD1, were omitted from the pairwise screens. Diploid transformants were tested for positive interactions by growing the mating strains in SD-leucine-tryptophan-histidine medium with increasing concentrations (0 mM, 3 mM, and 10 mM) of 3-amino-1,2,4-triazole (3-AT) to assess the interaction strengths. Constructs of known interacting

proteins, DmDP and DmE2F, were used as the positive control and the negative control consisted of yeast strains containing an empty AD vector mated with the BD fusion of the protein of interest. For each interaction, three independent biological repeats were done and for the assays the TECAN Genesis Automation and TECAN TEMO-96 pipetting robot were used (TECAN, Munich, Germany).
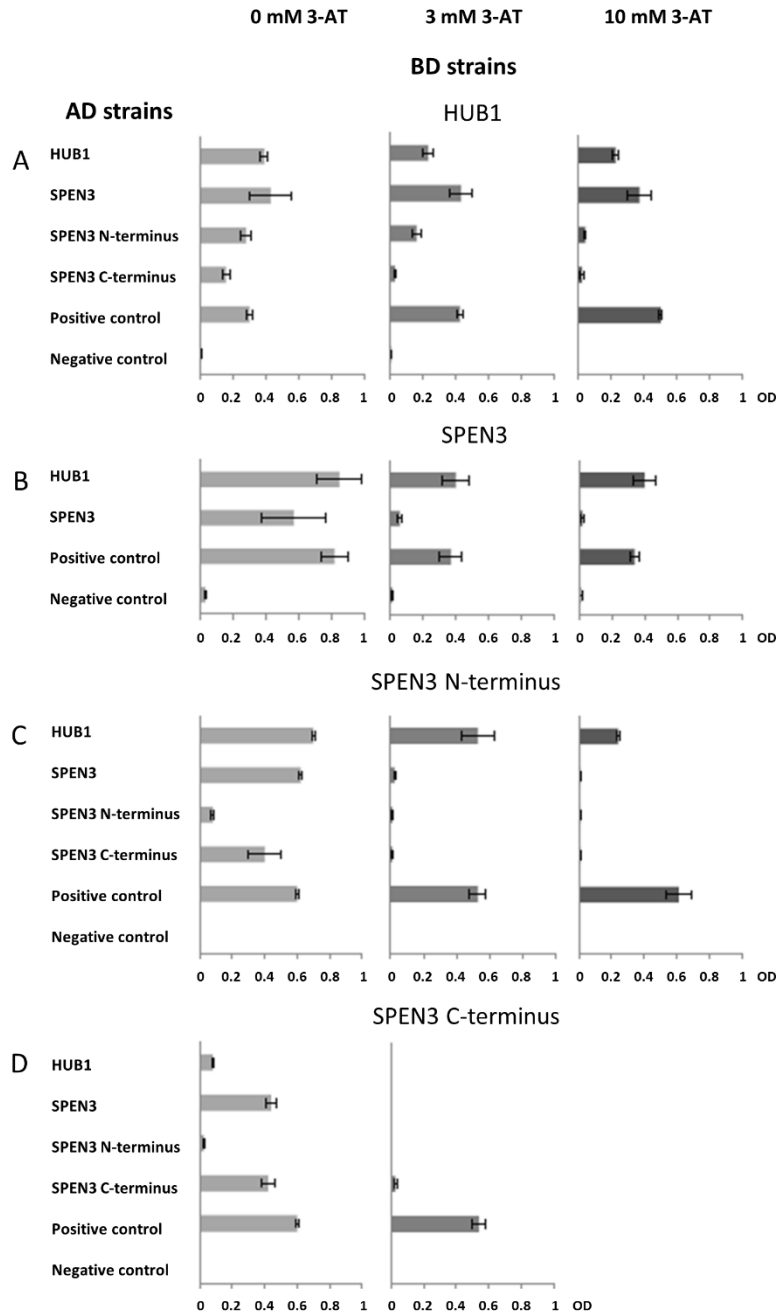
Fig. S1. Yeast two-hybrid interactions between HUB1, SPEN3, SPEN3 N-terminus, and SPEN3 C-terminus.

The yeast strains expressing the HUB1, SPEN3, SPEN3 N-terminus, and SPEN3 C-terminus proteins fused to the activation domain (AD strains, ordinate) or binding domain (BD strains, panels *A-D*) did not show self-activation and were mated pairwise to test for direct interactions between proteins that allowed yeast growth on selective medium and quantification as the optical density (OD_{600}) of the culture. Different concentrations (0 mM, 3 mM, and 10 mM) of 3-amino-1,2,4-triazole (3-AT) were applied to the medium to detect the high-affinity binding between two interactors allowing yeast to survive increased 3-AT concentrations. For each interaction, the average of three independent biological repeats are shown. As a positive control, known interactors (DmDP and DmE2F) were used and, as negative control, the empty AD vector strain was mated with the BD fusion of the protein of interest.

7
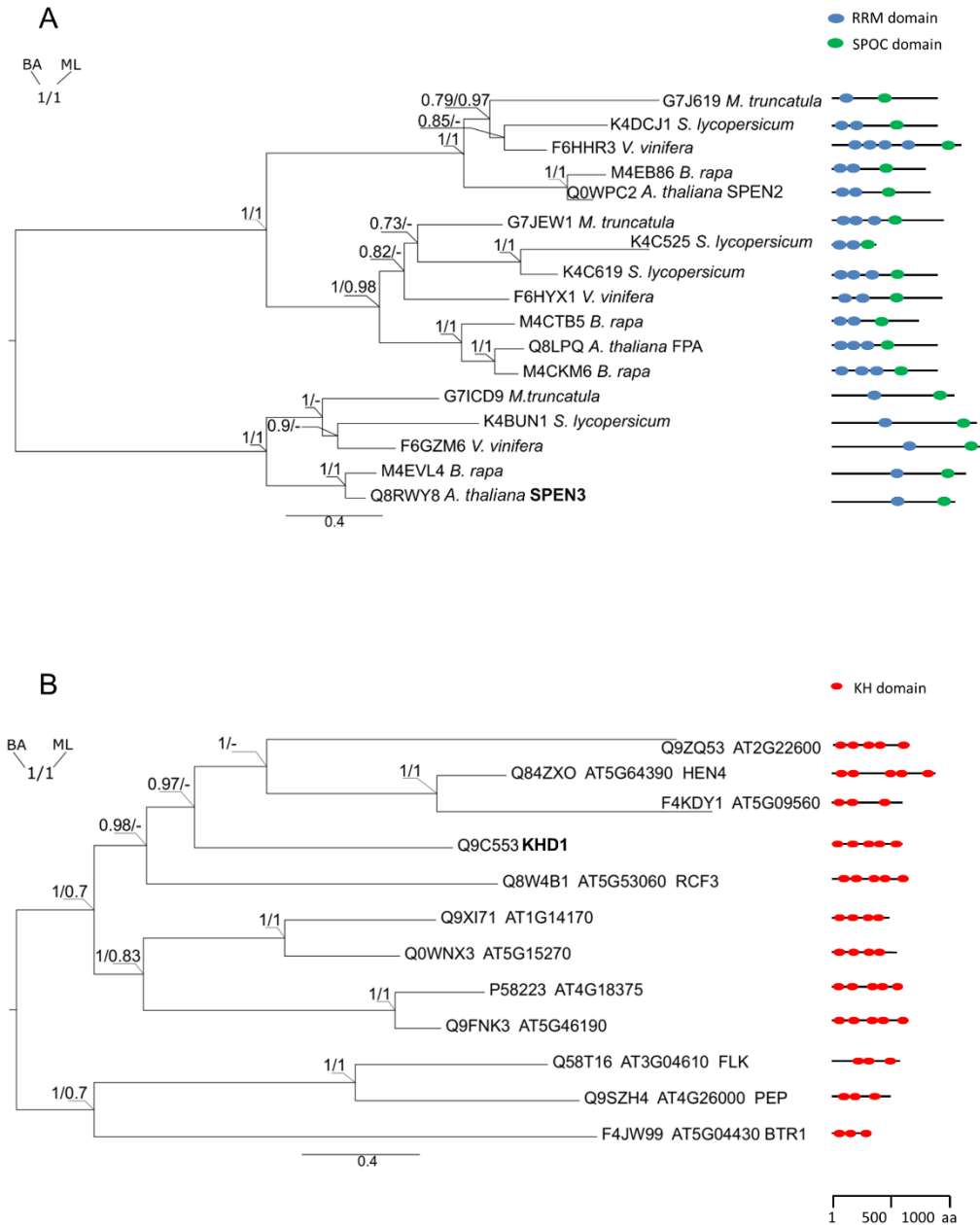
Fig. S2. Evolutionary relationships of SPEN3 (*A*) and KHD1 (*B*) proteins over taxa and within *Arabidopsis thaliana*. The evolutionary history was inferred from two independent maximum likelihood (ML) and Bayesian approaches (BA) with PhyML (23, 24) and MrBayes 3.2.6 (25), respectively. Both trees were estimated with a best-fit model obtained by MEGA X (26) as JTT + G for SPEN3 (*A*) and LG + G for KHD1 (*B*). The analysis involved 17 (*A*) and 12 (*B*) amino acid sequences retrieved from the UniProt database. In MrBayes, two independent runs were applied and trees were sampled every 200th generation for 5 000 000 generations (with a 25% burn-in). Overall, deviation of split frequencies for all trees used for consensus was much below 0.01. The reliability of the nodes is indicated by their posterior probability values (BA) and bootstrap values (1000 replicates) (ML) presented along the nodes. Node support values below 70% were not shown. The scale bar represents the genetic distance.
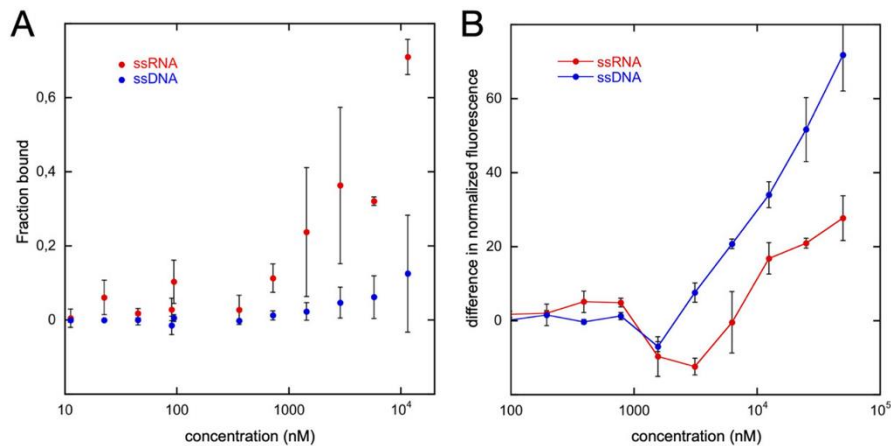
Fig. S3. MST analysis of SPEN3 and KHD1 interactions with nucleic acids.
(*A*) Increasing concentrations of the GST-RRM-SPEN3 protein incubated with 200 nM of 25-nucleotide-long ssRNA or ssDNA of the same sequence. Protein-nucleic acid interactions were quantified by MST. The approximate bound fraction of nucleic acids per tested protein concentration is plotted. (*B*) Increasing concentrations of binding of 6×His-KHD1-N protein incubated with ssRNA and ssDNA as in (*A*). Protein-nucleic acid interactions were measured by MST. In case of 6×His-KHD1-N binding to ssRNA and ssDNA, the MST analysis revealed a complex pattern of changes in the thermophoretic mobility of the interacting molecules. Instead of following a typical binding curve with increasing protein concentrations, a significant change in the mobility direction at a protein concentration of 1 μM. At lower concentrations, the protein-nucleic acid complexes moved toward the heated region in the capillary, whereas at higher concentrations (above 4 μM) they left the heated region. This behavior resulted in a "local minimum" of the plotted data that can be interpreted as two dependent binding events, creating two different protein-nucleic acid complexes with different thermophoretic behaviors established each one after the other with increasing protein concentration. The overlay of the thermophoretic properties can result in such plots. As no K_D or Hill fitting could be applied, the curves can only be described in a qualitative manner. The plot is displayed as changes in normalized fluorescence exhibiting qualitative differences in the presence of RNA or DNA. The "local minimum" was much more pronounced in the presence of RNA and the curve reached a plateau at high protein concentrations. In contrast, in the presence of DNA, the curve did not reach a plateau, hinting at partial DNA binding. Hence, this qualitative analysis supports the EMSA experiments, suggesting improved RNA binding. Measurements were done with three biological and three technical replicates. Error bars indicate standard deviation of the biological replicates.
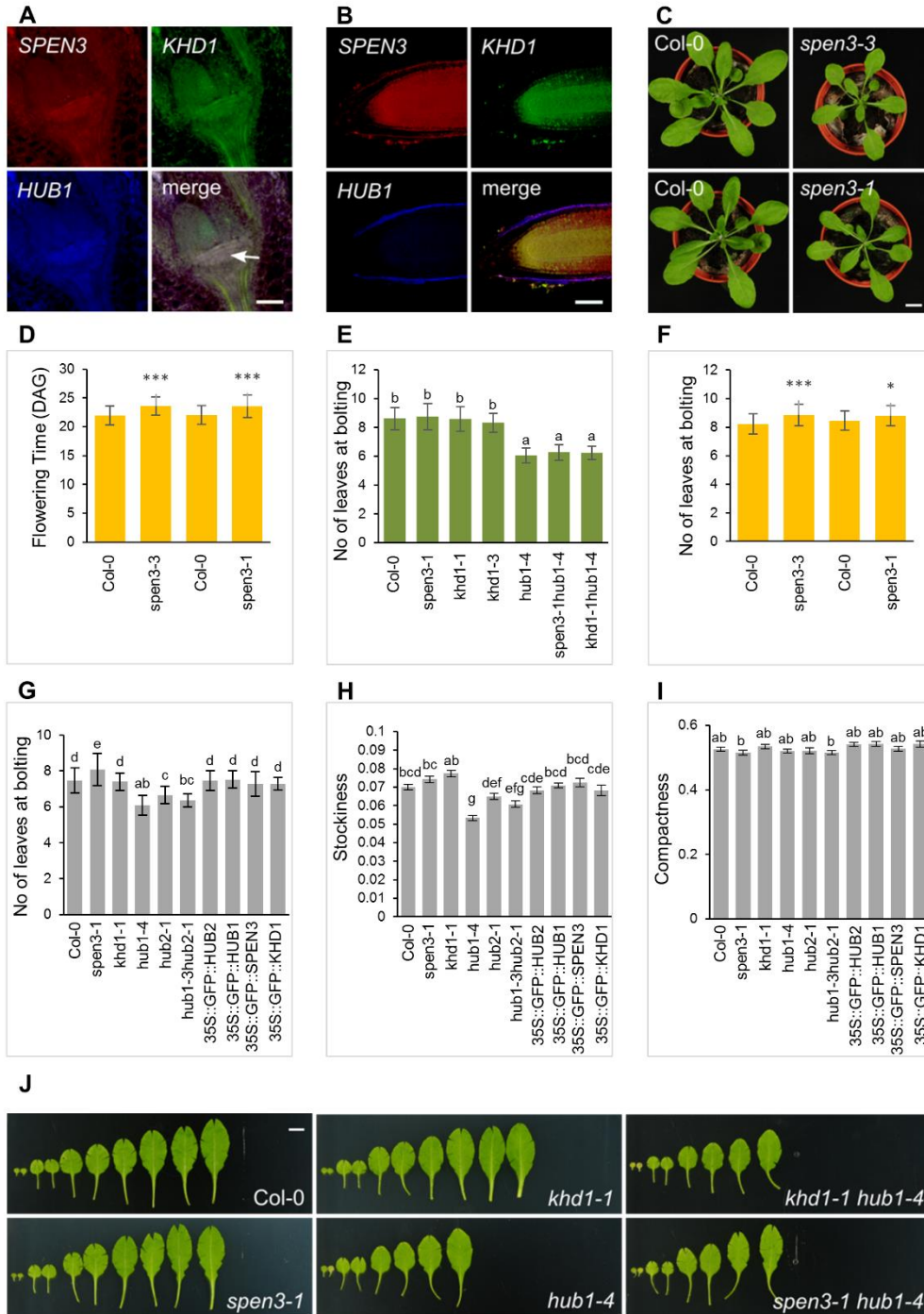
Fig. S4. *in situ* hybridization and leaf phenotypes of mutants and overexpression lines.
(*A* and *B*) Whole-mount, multiprobe *in situ* hybridization of the shoot apex (*A*) and the primary root meristem
(*B*). Arrow indicates white-pink merge of *HUB1* (blue samples treated with HUB1 FITC riboprobe and rabbit
anti-FITC followed by AF647 chicken antirabbit), *SPEN3* (red samples treated with *SPEN3* Dig-riboprobe
and sheep anti-Dig, followed by AF555 Donkey anti-sheep), and *KHD1* (green samples treated with *KHD1*
Bio-riboprobe and mouse anti-Bio followed by AF488-Donkey anti-mouse IgG) expression pattern in the
shoot apex. (*C*) Rosette phenotype of the *spen3-3* and *spen3-1* alleles and their Col-0 controls at 25 DAG.

(*D*) Flowering time of *spen3-3* and *spen3-1* and their Col-0 controls grown in jiffy pots ($n = 50$). (*E*) Number of rosette leaves per seedling at bolting in jiffy pots ($n = 51$). (*F*) Number of rosette leaves per seedling at bolting in jiffy pots (($n = 50$). (*G*) Number of rosette leaves per seedling at bolting grown in the WIWAM-automated platform ($n = 24$). (*H*) Stockiness and (*I*) compactness of the rosettes in WIWAM experiment calculated at 23 DAS. (*J*) Leaf series of 26 DAG seedlings grown in jiffy container experiment. Control, Col-0; single mutants, *spen3-1*, *khd1-1* and *hub1-4*; double mutants, *spen3-1 hub1-4* and *khd1-1 hub1-4*. Error bars represent standard deviations. Ordinary one-way ANOVA with 95% confidence shows a significant difference between the genotypes, represented by the letters (panels *E,G,H,I*). Asterisks (*D* and *F*) indicate statistically significant differences by Student's *t*-test (*$P < 0.05$, **$P < 0.01$).
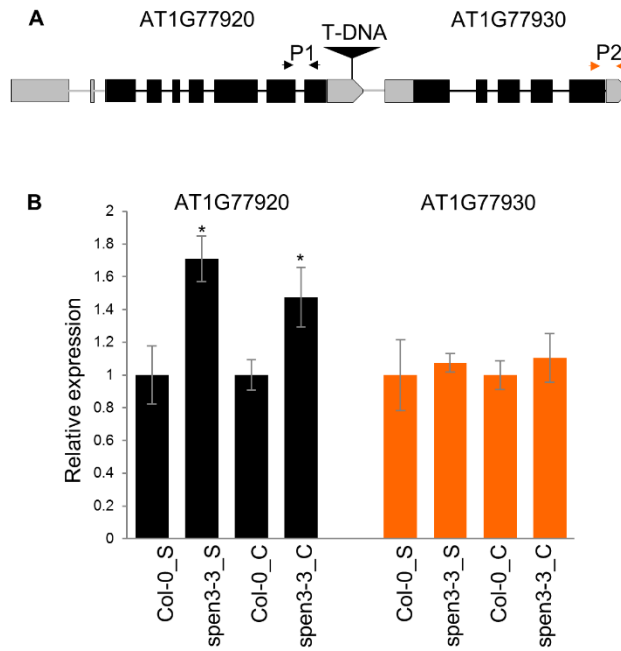
Fig. S5. Expression levels of genes flanking the second T-DNA insertion in *spen3-3*.
(*A*) Scheme of a second T-DNA insertion at the 3' end of *At1g77920* and the 5' end of *At1g77930* in the *spen3-3* allele. (*B*) qPCR analysis of *At1g77920* and *At1g77930* gene expression levels in seedling (S) and cauline (C) leaf tissue of *spen3-3* and Col-0 control (primers, see Table S3), five biological replicates.
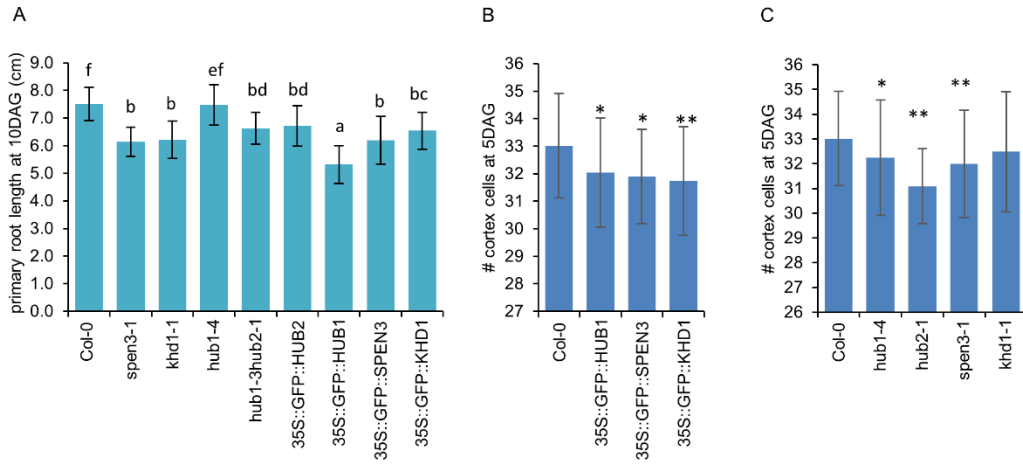
Fig. S6. Primary root phenotypes of mutants and overexpression lines.
(*A*) Primary root length at 10 DAG ($n \geq 15$). (*B* and *C*) Root meristem size measured by the number of cortex cells of 5 DAG seedlings of overexpression lines ($n > 15$) (*B*) and mutant lines ($n \geq 10$) (*C*). Error bars represent standard deviations. Ordinary one-way ANOVA with 95% confidence shows a significant difference between the genotypes, represented by the letters (panel *A*). Asterisks (panels *B* and *C*) indicate statistically significant differences by Student's *t*-test (*$P < 0.05$, **$P < 0.01$).

Table S1. AGI codes of the 117 DEGs common to the *hub1-4*, *khd1-1*, and *spen3-1* transcriptomes (log$_2$FC≥0.5 or log$_2$FC≤-0.5, $P \le 0.05$)

Gene ID	Gene name	Function/process	*hub1-4* vs COL-0: log$_2$FC	*khd1-1* vs COL-0: log$_2$FC	*spen3-1* vs COL-0: log$_2$FC
AT3G30720	QQS	Starch biosynthetic process	-5.6626	-5.5745	-6.5275
AT4G08093	NA	Unknown	-3.7844	-5.5504	-4.0931
AT2G01422	NA	Unknown	-3.3395	-3.7738	-2.5119
AT4G04223	NA	Unknown	-3.1981	-2.3271	-3.6096
AT4G15320	ATCSLB06	Cellulose biosynthetic process	-1.9482	-2.2482	-2.0571
AT1G48740	F11I4_9	Oxidation reduction process	-1.7234	-1.2481	-1.8338
AT1G51055	NA	Unknown	-1.4514	-1.3610	-1.3864
AT1G55320	AAE18	Auxin metabolic process	-1.2355	-0.8119	-0.6587
AT4G08991	NA	Unknown	-1.2329	-1.9179	-1.3986
AT5G25970	T1N24.19	Transferase activity	-1.1978	-0.8172	-1.1685
AT5G07640	NA	Zinc ion binding	-1.1420	-0.9536	-0.9802
AT2G39460	ATRPL23A	RNA binding	-1.0806	-1.0022	-0.6850
AT3G05727	NA	Unknown	-0.9690	-0.6627	-1.0264
AT2G29570	PCNA2	DNA methylation	-0.9243	-0.7950	-0.5766
AT3G55660	ATROPGEF6	Unknown	-0.8981	-0.6359	-0.5877
AT5G22440	RPL10AC	RNA methylation	-0.8793	-0.8620	-0.7202
AT5G37010	NA	DNA replication	-0.8736	-0.7098	-0.4696
AT1G52770	F14G24.4	Response to light stimulus	-0.8589	-1.3647	-0.8099
AT3G16490	IQD26	Calmodulin binding	-0.8390	-0.8087	-0.8740
AT2G25880	AtAUR2	Histone kinase	-0.8332	-0.8656	-0.5894
AT2G28620	NA	DNA replication	-0.8305	-0.7375	-0.5404
AT2G01020	NA	Peptide biosynthetic process	-0.7982	-0.9755	-0.8181
AT2G33400	F4P9.17	Unknown	-0.7936	-0.5135	-0.6040
AT1G18370	HIK	Microtubule movement	-0.7906	-0.6073	-0.6042
AT5G01600	ATFER1	Iron ion binding	-0.7852	-1.0765	-0.6556
AT4G02800	T5J8.12	Microtubule cytoskeleton	-0.7802	-0.6539	-0.5214
AT4G03100	F4C21.2	Microtubule cytoskeleton	-0.7608	-0.8148	-0.5114
AT1G02780	emb2386	RNA methylation	-0.7545	-0.6800	-0.9208
AT3G23890	TOPII	DNA topoisomerase	-0.7470	-0.5620	-0.5724
AT4G35810	NA	Oxidation reduction process	-0.7201	-0.7203	-0.9608
AT3G01710	NA	Unknown	-0.7187	-0.5457	-0.5202
AT4G22505	NA	Lipid transport	-0.6991	-0.5741	-0.8263
AT5G38940	NA	Response to salt stress	-0.6947	-0.7681	-1.0002
AT2G38620	CDKB1	Regulation of cell cycle	-0.6756	-0.5173	-0.7126
AT5G15200	RPS9B	RNA methylation	-0.6751	-0.5741	-0.6052
AT3G58650	F14P22.240	DNA replication	-0.6746	-0.6901	-0.6336
AT5G60150	NA	Petal formation	-0.6702	-0.7296	-0.5222
AT1G05440	NA	DNA methylation	-0.6700	-0.7066	-0.5283
AT5G44560	VPS2.2	Protein binding	-0.6659	-0.5251	-0.5011
AT2G26760	CYCB1	Regulation of cell cycle	-0.6633	-0.7223	-0.6465
AT1G23790	F5O8.34	Cell proliferation	-0.6442	-0.7267	-0.5978
AT3G26050	NA	Unknown	-0.6299	-0.6320	-0.5259
AT5G67270	ATEB1C	Microtubule binding	-0.6257	-0.6774	-0.6385
AT3G19050	POK2	Microtubule movement	-0.6187	-0.5180	-0.5445
AT5G26742	emb1138	Embryo development	-0.5808	-0.9022	-0.6706
AT2G36885	NA	Unknown	-0.5808	-0.8775	-0.8374
AT4G24670	TAR2	Cotyledon development	-0.5388	-0.6692	-0.6080
AT2G45490	AtAUR3	Histone kinase	-0.5322	-0.5612	-0.7462
AT2G33560	BUBR1	Cell proliferation	-0.5258	-0.6428	-0.5698
AT4G37490	CYC1	Regulation of cell cycle	-0.5244	-0.7092	-0.6084
AT5G35935	NA	Transposon	3.8585	4.0484	3.9318
AT1G19510	ATRL5	Regulation of transcription	2.6802	1.7416	1.6497
AT3G10420	SPD1	Nucleoside-triphosphatase activity	2.4560	2.4592	2.6330
AT3G05660	AtRLP33	Kinase activity	2.4204	1.4367	1.4173
AT4G08040	ACS11	Biosynthetic process	1.8670	0.9910	1.2030
AT2G34010	T14G11.13	Negative regulation of transcription	1.7997	1.0896	0.8731
AT1G51820	NA	Proline transport	1.6036	1.9470	1.5766
AT2G26560	PLP2	Lipase activity	1.5401	1.4479	1.6085
AT1G43910	F9C16_7	Response to ABA stimulus	1.4203	1.3555	1.1442
AT3G16030	CES101	Immune response	1.3967	1.1568	1.0114
AT4G36280	CRH1	ATP binding	1.2300	0.7077	0.5734
AT5G15510	NA	Cell proliferation	1.1310	0.7409	1.2072

14

AT5G17860	CAX7	Transmembrane transport	1.1310	0.7409	1.2072
AT1G62510	NA	Lipid transport	1.1214	0.9785	0.8740
AT5G38970	BR6OX1	Oxidation reduction process	1.1193	1.2404	1.2257
AT1G52880	NAM	Regulation of transcription	1.1080	0.7682	0.9576
AT4G11900	NA	Protein phosphorylation	1.0676	0.7431	0.7588
AT3G53250	T4D2.180	Response to auxin stimulus	1.0341	0.9364	0.7971
AT3G62860	F26K9_290	Catalytic activity	1.0282	1.0083	0.7175
AT1G72430	T10D10.10	Response to auxin stimulus	0.9981	1.0623	0.8578
AT1G21270	WAK2	Oligosaccharide metabolic process	0.9727	0.7971	0.6735
AT5G64780	NA	Unknown	0.9604	0.6438	0.5608
AT2G39980	NA	Response to karrikin	0.9578	0.6103	0.5734
AT3G61430	PIP1A	Water channel activity	0.9367	0.7735	0.5416
AT4G18010	IP5PII	Inositol tri-phosphate metabolic process	0.8834	1.2286	0.6841
AT3G56000	ATCSLA14	Transferase activity	0.8726	0.9119	0.7815
AT3G48720	T8P19.230	Cutin biosynthetic process	0.8433	0.8613	0.6017
AT5G08150	SOB5	Cytokinin metabolism process	0.8382	1.1877	0.6384
AT1G23130	T26J12.10	Defense response	0.8248	0.9054	0.6017
AT3G12090	TET6	Transition metal ion transport	0.8139	0.6686	0.6863
AT5G24030	SLAH3	Nitrate transport	0.7619	0.9383	0.6074
AT2G31730	BHLH-BETA	Response to ethylene and GA stimulus	0.7106	0.8444	0.6729
AT5G07000	ATST2B	Sulfotransferase activity	0.6955	0.9848	0.6295
AT3G07340	BHLH62	Regulation of transcription	0.6624	0.6468	0.5896
AT4G23190	CRK11	Kinase activity	0.6350	0.7084	0.8002
AT5G23750	MRO11.3	Cell wall biogenesis	0.5664	0.6600	0.7512
AT4G32790	F4D11.10	Catalytic activity	0.5664	0.5992	0.5092
AT2G41330	NA	Cell redox homeostasis	0.5613	0.6276	0.7785
AT2G23690	NA	Protein myristoylation	0.5438	0.7003	0.8397
AT1G02820	F22D16.18	Embryo development	-1.7547	-1.2409	1.2362
AT1G62540	FMO GS-OX2	Oxidation reduction process	-1.4021	-0.9912	0.9452
AT4G26790	NA	Lipid metabolic process	-1.1570	-1.1056	0.8255
AT3G21670	NRT1.3	Oligopeptide transport	-1.1033	-1.0078	0.9084
AT1G16730	UP6	Fatty acid beta oxidation	-1.0651	-1.2906	1.0012
AT4G37310	CYP81H1	Oxidation reduction process	-1.0019	-0.7621	0.5620
AT2G46680	ATHB-7	Regulation of transcription	-0.9295	-0.6854	0.6970
AT1G73390	T9L24.40	Protein myristoylation	-0.9262	-0.6509	0.6113
AT2G17300	NA	Unknown	-0.7890	-0.5806	0.5410
AT5G03760	ATCSLA09	Calcium ion transport	-0.7349	-0.7330	0.7095
AT5G61290	NA	Oxidation reduction process	-0.6413	-0.7854	0.7625
AT1G64770	NDF2	Carbohydrate metabolic process	-0.5301	-0.5406	0.5175
AT4G37770	ACS8	Biosynthetic process	1.7545	2.0780	-1.8580
AT4G40065	NA	Unknown	1.4381	1.1595	-1.3468
AT2G23170	GH3.3	Response to auxin stimulus	1.4334	1.2295	-0.7410
AT2G18010	SAUR10	Response to auxin stimulus	1.3172	1.5038	-1.1684
AT3G42800	T21C14_20	Unknown	1.1613	1.4009	-0.8314
AT1G04610	YUC3	Oxidation reduction process	0.8388	1.2292	-1.9934
AT1G30420	ATMRP12	Transmembrane transport	0.5738	0.5111	-1.0747
AT1G51830	T14L22.4	Nitrate transport	-1.5491	1.2138	-1.0585
AT4G20320	F1C12.230	Pyrimidine nucleotide biosynthetic process	-1.4217	-0.8890	-0.5056
AT4G30170	PER45	Oxidation reduction process	-1.2226	1.4707	-1.1787
AT3G32925	NA	Transposon	2.4493	-2.0186	1.4498
AT5G59670	NA	Protein phosphorylation	1.2640	-1.1608	1.5756
AT1G73000	PYL3	Unknown	1.2088	-1.0563	1.6080
AT1G11070	NA	Unknown	0.9290	-0.6423	0.8314
AT1G78000	SULTR1	Nitrate transport	-1.5906	0.6725	0.8204
AT3G54160	F24B22.120	Unknown	0.9710	-1.2585	-1.0686

*P values for all log$_2$FC ≤ 0.05.

Table S2. PLAZA enrichment of Biological Process Gene Ontology (GO) categories identified within genes down- or upregulated in the indicated mutants. For the single mutants, only the mutant-specific down- or upregulated genes are presented. The GO ratio is the ratio between the number of analyzed annotated genes that belong to the given GO category and the number of all analyzed annotated genes.

GO term	Log2-Enrichment	*P* value	GO ratio (%)	Description
Down-regulated in *hub1-4, khd1-1,* and *spen3-1*				
GO:0051726	4.63	5.67E-04	12	Regulation of cell cycle
GO:0016572	8.32	4,.0E-03	5	Histone phosphorylation
GO:0007049	3.40	3.00E-02	12	Cell cycle
Down-regulated in *hub1-4* and *spen3-1*				
GO:0051726	3.71	2.00E-02	6	Regulation of cell cycle
GO:0016572	7.40	3.00E-02	2	Histone phosphorylation
Down-regulated in *hub1-4* and *khd1-1*				
GO:0008152	0.54	4.52E-17	61	Metabolic process
GO:0019748	2.13	4.61E-16	9	Secondary metabolic process
GO:0006260	3.09	9.67E-13	4	Cellular DNA replication
GO:0006259	2.22	1.73E-12	7	Cellular DNA metabolic process
GO:0044238	0.53	6.76E-12	51	Primary metabolic process
GO:0044237	0.53	3.52E-11	49	Cellular metabolic process
GO:0044281	1.24	4.94E-11	16	Small molecule metabolic process
GO:0006519	1.76	9.67E-11	9	Cellular amino acid and derivative metabolic process
GO:0016144	3.62	1.52E-10	3	*S*-glycoside biosynthetic process
GO:0019758	3.62	1.52E-10	3	Glycosinolate biosynthetic process
GO:0019761	3.62	1.52E-10	3	Glucosinolate biosynthetic process
GO:0009812	3.23	1.64E-10	3	Flavonoid metabolic process
GO:0009987	0.42	2.07E-10	60	Cellular process
GO:0009813	3.30	2.84E-10	3	Flavonoid biosynthetic process
GO:0009058	0.70	4.05E-10	33	Biosynthetic process
GO:0034637	2.36	6.90E-10	5	Cellular carbohydrate biosynthetic process
GO:0044249	0.70	8.93E-10	33	Cellular biosynthetic process
GO:0044283	1.59	3.03E-09	9	Small molecule biosynthetic process
GO:0009698	2.50	3.42E-09	4	Phenylpropanoid metabolic process
GO:0006807	0.79	4.56E-09	26	Nitrogen compound metabolic process
GO:0016137	2.62	4.79E-09	4	Glycoside metabolic process
GO:0016138	3.07	5.27E-09	3	Glycoside biosynthetic process
GO:0016143	2.92	8.79E-09	3	S-glycoside metabolic process
GO:0019757	2.92	8.79E-09	3	Glycosinolate metabolic process
GO:0019760	2.92	8.79E-09	3	Glucosinolate metabolic process
GO:0009699	2.63	1.12E-08	4	Phenylpropanoid biosynthetic process
GO:0042398	2.26	5.03E-08	5	Cellular amino acid derivative biosynthetic process
GO:0019438	2.21	5.06E-08	5	Aromatic compound biosynthetic process
GO:0006725	1.87	5.39E-08	6	Cellular aromatic compound metabolic process
GO:0006790	2.25	5.93E-08	5	Sulfur metabolic process
GO:0016051	2.07	9.32E-08	5	Carbohydrate biosynthetic process
GO:0007049	2.14	1.47E-07	5	Cell cycle
GO:0006575	2.00	1.50E-07	5	Cellular amino acid derivative metabolic process
GO:0044272	2.73	2.87E-07	3	Sulfur compound biosynthetic process
GO:0005975	1.21	2.53E-06	10	Carbohydrate metabolic process
GO:0050896	0.61	2.58E-06	29	Response to stimulus
GO:0044262	1.49	6.89E-06	7	Cellular carbohydrate metabolic process
GO:0022402	2.31	1.22E-05	3	Cell cycle process
GO:0042440	2.48	1.22E-05	3	Pigment metabolic process
GO:0043455	3.52	1.31E-05	2	Regulation of secondary metabolic process
GO:0010439	5.36	2.42E-05	1	Regulation of glucosinolate biosynthetic process
GO:0000103	4.34	2.64E-05	1	Sulfate assimilation
GO:0006791	4.18	6.57E-05	1	Sulfur utilization
GO:0006261	3.08	6.78E-05	2	DNA-dependent DNA replication
GO:0034285	2.85	1.07E-04	2	Response to disaccharide stimulus
GO:0006270	4.51	1.23E-04	1	DNA replication initiation
GO:0006950	0.71	1.68E-04	19	Response to stress
GO:0046148	2.45	3.03E-04	3	Pigment biosynthetic process
GO:0006281	2.07	3.43E-04	3	DNA repair

GO:0034984	2.07	3.67E-04	3	Cellular response to DNA damage stimulus
GO:0009744	2.82	4.69E-04	2	Response to sucrose stimulus
GO:0043255	4.78	4.79E-04	1	Regulation of carbohydrate biosynthetic process
GO:0044271	1.57	4.94E-04	5	Cellular nitrogen compound biosynthetic process
GO:0006139	0.66	5.11E-04	19	Cellular nucleobase, nucleoside, nucleotide, and nucleic acid metabolic process
GO:0006974	2.01	6.27E-04	3	Response to DNA damage stimulus
GO:0009309	2.01	6.27E-04	3	Amine biosynthetic process
GO:0010675	4.62	9.38E-04	1	Regulation of cellular carbohydrate metabolic process
GO:0009753	1.90	9.81E-04	3	Response to jasmonic acid stimulus
GO:0006996	1.20	2.20E-03	7	Organelle organization
GO:0008652	2.02	2.50E-03	3	Cellular amino acid biosynthetic process
GO:0033205	3.48	2.70E-03	1	Cytokinesis during cell cycle
GO:0006268	4.36	2.80E-03	1	DNA unwinding during replication
GO:0033554	1.38	4.00E-03	5	Cellular response to stress
GO:0016043	0.93	4.20E-03	10	Cellular component organization
GO:0042762	4.25	4.50E-03	1	Regulation of sulfur metabolic process
GO:0008283	2.64	4.80E-03	2	Cell proliferation
GO:0042180	1.09	1.00E-02	7	Cellular ketone metabolic process
GO:0051716	1.02	1.00E-02	8	Cellular response to stimulus
GO:0046283	3.30	1.00E-02	1	Anthocyanin metabolic process
GO:0006109	4.14	1.00E-02	1	Regulation of carbohydrate metabolic process
GO:0044106	1.40	1.00E-02	4	Cellular amine metabolic process
GO:0034641	1.16	1.00E-02	6	Cellular nitrogen compound metabolic process
GO:0042221	0.58	1.00E-02	18	Response to chemical stimulus
GO:0051726	2.20	1.00E-02	2	Regulation of cell cycle
GO:0009308	1.14	2.00E-02	6	Amine metabolic process
GO:0006520	1.42	2.00E-02	4	Cellular amino acid metabolic process
GO:0019752	1.03	2.00E-02	7	Carboxylic acid metabolic process
GO:0043436	1.03	2.00E-02	7	Oxoacid metabolic process
GO:0032392	3.86	2.00E-02	1	DNA geometric change
GO:0032508	3.86	2.00E-02	1	DNA duplex unwinding
GO:0006082	1.03	2.00E-02	7	Organic acid metabolic process
GO:0009628	0.69	3.00E-02	13	Response to abiotic stimulus
GO:0000910	2.72	3.00E-02	1	Cytokinesis
GO:0007018	2.49	3.00E-02	2	Microtubule-based movement
GO:0007017	2.06	3.00E-02	2	Microtubule-based process
GO:0009411	1.76	4.00E-02	3	Response to UV
GO:0019419	5.21	4.00E-02	1	Sulfate reduction
GO:0051276	1.60	5.00E-02	3	Chromosome organization

Down-regulated specifically in *hub1-4*

GO:0050896	0.96	2.40E-30	37	Response to stimulus
GO:0006950	1.20	5.99E-29	26	Response to stress
GO:0042221	1.13	3.33E-25	26	Response to chemical stimulus
GO:0019748	2.12	2.47E-23	9	Secondary metabolic process
GO:0006970	1.85	3.67E-22	11	Response to osmotic stress
GO:0009651	1.86	7.93E-21	10	Response to salt stress
GO:0010035	1.82	3.71E-20	10	Response to inorganic substance
GO:0016137	2.88	9.30E-20	5	Glycoside metabolic process
GO:0010038	1.86	4.49E-18	9	Response to metal ion
GO:0016143	3.13	1.15E-17	4	*S*-glycoside metabolic process
GO:0019757	3.13	1.15E-17	4	Glycosinolate metabolic process
GO:0019760	3.13	1.15E-17	4	Glucosinolate metabolic process
GO:0006952	1.59	1.36E-17	11	Defense response
GO:0005975	1.48	1.64E-17	13	Carbohydrate metabolic process
GO:0009628	1.17	6.67E-17	18	Response to abiotic stimulus
GO:0009607	1.49	1.24E-16	12	Response to biotic stimulus
GO:0051707	1.49	2.36E-16	12	Response to other organism
GO:0006790	2.45	3.19E-16	5	Sulfur metabolic process
GO:0009611	2.36	9.08E-15	5	Response to wounding
GO:0051704	1.30	4.60E-14	13	Multi-organism process
GO:0044283	1.59	8.36E-14	9	Small molecule biosynthetic process
GO:0008152	0.41	5.47E-13	56	Metabolic process
GO:0044262	1.67	1.10E-12	8	Cellular carbohydrate metabolic process
GO:0009605	1.73	7.82E-12	7	Response to external stimulus
GO:0009753	2.29	1.06E-11	4	Response to jasmonic acid stimulus
GO:0009266	1.58	2.51E-11	8	Response to temperature stimulus
GO:0010043	2.70	3.23E-11	3	Response to zinc ion
GO:0006030	2.31	9.64E-11	4	Chitin metabolic process
GO:0006022	2.29	1.39E-10	4	Aminoglycan metabolic process

GO:0044036	2.18	4.61E-10	4	Cell wall macromolecule metabolic process
GO:0009409	1.67	5.73E-10	6	Response to cold
GO:0044281	1.04	5.85E-10	14	Small molecule metabolic process
GO:0046686	1.68	7.44E-10	6	Response to cadmium ion
GO:0006037	2.33	8.66E-10	4	Cell wall chitin metabolic process
GO:0071555	2.33	8.66E-10	4	Cell wall organization
GO:0010383	2.25	1.73E-09	4	Cell wall polysaccharide metabolic process
GO:0016138	2.72	2.24E-08	3	Glycoside biosynthetic process
GO:0055114	1.19	2.30E-08	10	Oxidation reduction
GO:0016144	3.13	2.68E-08	2	S-glycoside biosynthetic process
GO:0019758	3.13	2.68E-08	2	Glycosinolate biosynthetic process
GO:0019761	3.13	2.68E-08	2	Glucosinolate biosynthetic process
GO:0005976	1.77	2.96E-08	5	Polysaccharide metabolic process
GO:0010033	0.91	3.74E-08	14	Response to organic substance
GO:0044272	2.52	3.87E-08	3	Sulfur compound biosynthetic process
GO:0006629	1.31	4.34E-08	8	Lipid metabolic process
GO:0071554	1.87	8.57E-08	4	Cell wall organization or biogenesis
GO:0009308	1.38	1.12E-07	7	Amine metabolic process
GO:0016053	1.62	2.00E-07	5	Organic acid biosynthetic process
GO:0046394	1.62	2.00E-07	5	Carboxylic acid biosynthetic process
GO:0016145	3.21	2.16E-07	2	*S*-glycoside catabolic process
GO:0019759	3.21	2.16E-07	2	Glycosinolate catabolic process
GO:0019762	3.21	2.16E-07	2	Glucosinolate catabolic process
GO:0016139	2.94	2.24E-07	2	Glycoside catabolic process
GO:0009058	0.53	2.84E-07	30	Biosynthetic process
GO:0044238	0.37	3.22E-07	46	Primary metabolic process
GO:0009414	1.67	6.66E-07	5	Response to water deprivation
GO:0044273	3.08	7.88E-07	2	Sulfur compound catabolic process
GO:0080028	3.39	8.63E-07	2	Nitrile biosynthetic process
GO:0042398	1.93	8.77E-07	4	Cellular amino acid derivative biosynthetic process
GO:0019438	1.86	1.21E-06	4	Aromatic compound biosynthetic process
GO:0009415	1.63	1.29E-06	5	Response to water
GO:0050898	3.32	1.61E-06	2	Nitrile metabolic process
GO:0080027	3.32	1.61E-06	2	Response to herbivore
GO:0044249	0.51	1.89E-06	29	Cellular biosynthetic process
GO:0006575	1.68	2.49E-06	4	Cellular amino acid derivative metabolic process
GO:0050832	1.93	2.90E-06	3	Defense response to fungus
GO:0016051	1.72	4.12E-06	4	Carbohydrate biosynthetic process
GO:0006519	1.30	4.79E-06	6	Cellular amino acid and derivative metabolic process
GO:0009699	2.14	6.73E-06	3	Phenylpropanoid biosynthetic process
GO:0034637	1.82	8.96E-06	4	Cellular carbohydrate biosynthetic process
GO:0019752	1.15	1.21E-05	7	Carboxylic acid metabolic process
GO:0043436	1.15	1.21E-05	7	Oxoacid metabolic process
GO:0006082	1.14	1.31E-05	7	Organic acid metabolic process
GO:0016052	1.82	1.51E-05	3	Carbohydrate catabolic process
GO:0009698	1.94	1.90E-05	3	Phenylpropanoid metabolic process
GO:0042180	1.13	1.96E-05	7	Cellular ketone metabolic process
GO:0006631	1.83	2.36E-05	3	Fatty acid metabolic process
GO:0044271	1.49	2.56E-05	5	Cellular nitrogen compound biosynthetic process
GO:0006725	1.45	3.46E-05	5	Cellular aromatic compound metabolic process
GO:0033037	3.32	3.87E-05	1	Polysaccharide localization
GO:0051716	1.04	3.88E-05	8	Cellular response to stimulus
GO:0009719	0.91	3.96E-05	10	Response to endogenous stimulus
GO:0042742	1.60	4.78E-05	4	Defense response to bacterium
GO:0009987	0.26	5.97E-05	54	Cellular process
GO:0006979	1.50	8.26E-05	4	Response to oxidative stress
GO:0032787	1.40	1.19E-04	5	Monocarboxylic acid metabolic process
GO:0052543	3.39	1.20E-04	1	Callose deposition in cell wall
GO:0052386	3.34	1.64E-04	1	Cell wall thickening
GO:0009620	1.42	1.89E-04	4	Response to fungus
GO:0044275	1.81	2.11E-04	3	Cellular carbohydrate catabolic process
GO:0007155	2.74	2.13E-04	2	Cell adhesion
GO:0022610	2.74	2.13E-04	2	Biological adhesion
GO:0052545	3.30	2.21E-04	1	Callose localization
GO:0009664	2.46	2.45E-04	2	Plant-type cell wall organization
GO:0006633	1.96	4.90E-04	2	Fatty acid biosynthetic process
GO:0009269	3.17	5.11E-04	1	Response to desiccation
GO:0034641	1.10	9.25E-04	6	Cellular nitrogen compound metabolic process
GO:0042430	2.48	1.50E-03	2	Indole and derivative metabolic process
GO:0042434	2.48	1.50E-03	2	Indole derivative metabolic process

GO:0071669	1.96	1.90E-03	2	Plant-type cell wall organization or biogenesis
GO:0006807	0.47	2.20E-03	21	Nitrogen compound metabolic process
GO:0044255	1.19	2.50E-03	5	Cellular lipid metabolic process
GO:0042435	2.52	3.20E-03	1	Indole derivative biosynthetic process
GO:0009908	1.31	3.20E-03	4	Flower development
GO:0052482	3.23	1.00E-02	1	Cell wall thickening during defense response
GO:0052544	3.23	1.00E-02	1	Callose deposition in cell wall during defense response
GO:0042343	3.54	1.00E-02	1	Indole glucosinolate metabolic process
GO:0010143	3.99	1.00E-02	1	Cutin biosynthetic process
GO:0007275	0.54	1.00E-02	15	Multicellular organismal development
GO:0005985	2.88	1.00E-02	1	Sucrose metabolic process
GO:0052542	3.13	1.00E-02	1	Callose deposition during defense response
GO:0031668	1.73	2.00E-02	2	Cellular response to extracellular stimulus
GO:0071496	1.73	2.00E-02	2	Cellular response to external stimulus
GO:0009759	3.86	2.00E-02	1	Indole glucosinolate biosynthetic process
GO:0080119	3.86	2.00E-02	1	ER body organization
GO:0009684	3.39	2.00E-02	1	Indoleacetic acid biosynthetic process
GO:0044237	0.25	2.00E-02	40	Cellular metabolic process
GO:0033554	1.12	2.00E-02	4	Cellular response to stress
GO:0032501	0.51	2.00E-02	15	Multicellular organismal process
GO:0009408	1.68	2.00E-02	2	Response to heat
GO:0009683	3.32	2.00E-02	1	Indoleacetic acid metabolic process
GO:0032502	0.50	3.00E-02	15	Developmental process
GO:0031669	1.80	3.00E-02	2	Cellular response to nutrient levels
GO:0009751	1.60	3.00E-02	2	Response to salicylic acid stimulus
GO:0042545	1.77	4.00E-02	2	Cell wall modification
GO:0009991	1.61	4.00E-02	2	Response to extracellular stimulus
GO:0006869	1.73	5.00E-02	2	Lipid transport

Down-regulated specifically in *khd1-1*

GO:0006259	2.56	2.35E-17	9	Cellular DNA metabolic process
GO:0006139	1.15	6.27E-16	27	Cellular nucleobase, nucleoside, nucleotide, and nucleic acid metabolic process
GO:0006807	0.90	1.11E-10	28	Nitrogen compound metabolic process
GO:0006260	3.06	3.86E-10	4	Cellular DNA replication
GO:0022402	2.46	5.21E-06	4	Cell cycle process
GO:0007049	2.01	6.90E-05	4	Cell cycle
GO:0048367	1.89	6.92E-05	5	Shoot development
GO:0022621	1.87	9.28E-05	5	Shoot system development
GO:0044260	0.49	9.93E-05	36	Cellular macromolecule metabolic process
GO:0006261	3.18	1.37E-04	2	DNA-dependent DNA replication
GO:0048827	2.05	1.90E-04	4	Phyllome development
GO:0043170	0.44	3.85E-04	39	Macromolecule metabolic process
GO:0006281	2.13	7.89E-04	3	DNA repair
GO:0034984	2.13	8.39E-04	3	Cellular response to DNA damage stimulus
GO:0009987	0.30	1.20E-03	55	Cellular process
GO:0016070	0.93	1.30E-03	12	Cellular RNA metabolic process
GO:0006974	2.07	1.40E-03	3	Response to DNA damage stimulus
GO:0022403	2.48	1.50E-03	3	Cell cycle phase
GO:0031323	0.84	1.70E-03	14	Regulation of cellular metabolic process
GO:0044237	0.37	1.90E-03	44	Cellular metabolic process
GO:0051171	0.85	2.80E-03	13	Regulation of nitrogen compound metabolic process
GO:0019219	0.84	3.80E-03	13	Regulation of cellular nucleobase, nucleoside, nucleotide, and nucleic acid metabolic process
GO:0009887	2.03	4.20E-03	3	Organ morphogenesis
GO:0044238	0.34	4.30E-03	45	Primary metabolic process
GO:0019222	0.75	1.00E-02	15	Regulation of metabolic process
GO:0031326	0.80	1.00E-02	13	Regulation of cellular biosynthetic process
GO:0009889	0.80	1.00E-02	13	Regulation of biosynthetic process
GO:0006350	0.79	1.00E-02	13	Cellular transcription
GO:0061018	0.79	1.00E-02	13	Transcription
GO:0008152	0.28	1.00E-02	51	Metabolic process
GO:0010467	0.56	1.00E-02	21	Gene expression
GO:0080090	0.77	1.00E-02	13	Regulation of primary metabolic process
GO:0048608	0.95	1.00E-02	9	Reproductive structure development
GO:0007126	2.83	1.00E-02	2	Meiosis
GO:0051327	2.83	1.00E-02	2	M phase of meiotic cell cycle
GO:0000279	2.43	2.00E-02	2	M phase
GO:0045449	0.79	2.00E-02	12	Regulation of cellular transcription
GO:0061019	0.79	2.00E-02	12	Regulation of transcription
GO:0010556	0.78	2.00E-02	12	Regulation of macromolecule biosynthetic process

GO:0048366	1.83	2.00E-02	3	Leaf development
GO:0022414	0.85	2.00E-02	11	Reproductive process
GO:0010468	0.74	2.00E-02	13	Regulation of gene expression
GO:0060255	0.73	2.00E-02	14	Regulation of macromolecule metabolic process
GO:0009451	2.20	3.00E-02	2	RNA modification
GO:0032501	0.64	3.00E-02	16	Multicellular organismal process
GO:0000003	0.83	3.00E-02	11	Reproduction
GO:0007275	0.64	3.00E-02	16	Multicellular organismal development
GO:0007017	2.16	3.00E-02	2	Microtubule-based process
GO:0042254	1.84	3.00E-02	3	Ribosome biogenesis
GO:0009888	1.49	4.00E-02	4	Tissue development
GO:0003006	0.86	5.00E-02	9	Reproductive developmental process

Down-regulated specifically in *spen3-1*

GO:0023046	1.18	3.00E-02	10	Signaling process
GO:0023060	1.18	3.00E-02	10	Signal transmission
GO:0007623	2.80	5.00E-02	3	Circadian rhythm
GO:0048511	2.80	5.00E-02	3	Rhythmic process

Up-regulated in *hub1-4* and *khd1-1*

GO:0050896	1.15	1.09E-25	42	Response to stimulus
GO:0009725	2.00	1.78E-24	20	Response to hormone stimulus
GO:0010033	1.67	7.16E-24	25	Response to organic substance
GO:0009719	1.91	8.64E-24	20	Response to endogenous stimulus
GO:0009733	2.76	2.17E-23	12	Response to auxin stimulus
GO:0042221	1.27	9.13E-18	28	Response to chemical stimulus
GO:0009741	3.53	2.69E-09	4	Response to brassinosteroid stimulus
GO:0023052	1.26	9.32E-08	15	Signaling
GO:0032870	2.25	1.09E-07	6	Cellular response to hormone stimulus
GO:0007165	1.43	1.25E-07	12	Signal transduction
GO:0009755	2.24	3.00E-07	6	Hormone-mediated signaling pathway
GO:0051704	1.32	4.12E-07	13	Multi-organism process
GO:0009607	1.42	4.20E-07	11	Response to biotic stimulus
GO:0023046	1.38	4.87E-07	12	Signaling process
GO:0023060	1.38	4.87E-07	12	Signal transmission
GO:0051707	1.43	5.50E-07	11	Response to other organism
GO:0071495	2.10	1.00E-06	6	Cellular response to endogenous stimulus
GO:0007242	1.61	1.34E-06	9	Intracellular signaling cascade
GO:0051716	1.43	1.97E-06	10	Cellular response to stimulus
GO:0006952	1.44	2.24E-06	10	Defense response
GO:0065008	1.51	3.19E-06	9	Regulation of biological quality
GO:0009606	3.47	3.45E-06	3	Tropism
GO:0006468	1.39	5.67E-06	10	Protein amino acid phosphorylation
GO:0065007	0.72	8.30E-06	27	Biological regulation
GO:0006796	1.25	1.70E-05	11	Phosphate metabolic process
GO:0006793	1.24	1.78E-05	11	Phosphorus metabolic process
GO:0071310	1.83	2.12E-05	6	Cellular response to organic substance
GO:0055085	1.59	4.24E-05	7	Transmembrane transport
GO:0006464	1.00	4.57E-05	15	Protein modification process
GO:0043687	1.12	4.94E-05	13	Post-translational protein modification
GO:0016310	1.24	7.49E-05	10	Phosphorylation
GO:0006950	0.80	1.07E-04	20	Response to stress
GO:0009630	3.39	1.44E-04	2	Gravitropism
GO:0070887	1.66	2.05E-04	6	Cellular response to chemical stimulus
GO:0008219	1.90	4.16E-04	5	Cell death
GO:0016265	1.90	4.16E-04	5	Death
GO:0016049	1.83	4.56E-04	5	Cell growth
GO:0009639	2.17	4.97E-04	4	Response to red or far red light
GO:0009629	3.18	5.31E-04	2	Response to gravity
GO:0043412	0.89	6.22E-04	15	Macromolecule modification
GO:0012501	1.97	7.42E-04	4	Programmed cell death
GO:0008361	1.76	9.31E-04	5	Regulation of cell size
GO:0032535	1.74	1.20E-03	5	Regulation of cellular component size
GO:0090066	1.74	1.20E-03	5	Regulation of anatomical structure size
GO:0051179	0.90	1.40E-03	14	Localization
GO:0040007	1.61	1.70E-03	5	Growth
GO:0048589	1.81	1.80E-03	5	Developmental growth
GO:0006810	0.90	1.90E-03	14	Transport
GO:0009638	4.44	1.90E-03	1	Phototropism
GO:0051234	0.90	2.10E-03	14	Establishment of localization

GO:0050789	0.63	2.40E-03	23	Regulation of biological process
GO:0009605	1.49	2.60E-03	6	Response to external stimulus
GO:0080086	5.03	3.20E-03	1	Stamen filament development
GO:0009692	3.76	3.80E-03	1	Ethylene metabolic process
GO:0009693	3.76	3.80E-03	1	Ethylene biosynthetic process
GO:0043449	3.71	4.70E-03	1	Cellular alkene metabolic process
GO:0043450	3.71	4.70E-03	1	Alkene biosynthetic process
GO:0050832	1.99	1.00E-02	4	Defense response to fungus
GO:0050794	0.64	1.00E-02	20	Regulation of cellular process
GO:0042742	1.71	1.00E-02	4	Defense response to bacterium
GO:0009742	3.16	1.00E-02	2	Brassinosteroid mediated signaling pathway
GO:0043401	3.16	1.00E-02	2	Steroid hormone mediated signaling pathway
GO:0048545	3.16	1.00E-02	2	Response to steroid hormone stimulus
GO:0071367	3.16	1.00E-02	2	Cellular response to brassinosteroid stimulus
GO:0071383	3.16	1.00E-02	2	Cellular response to steroid hormone stimulus
GO:0044042	2.26	2.00E-02	3	Glucan metabolic process
GO:0010359	4.44	2.00E-02	1	Regulation of anion channel activity
GO:0022898	4.44	2.00E-02	1	Regulation of transmembrane transporter activity
GO:0032409	4.44	2.00E-02	1	Regulation of transporter activity
GO:0032412	4.44	2.00E-02	1	Regulation of ion transmembrane transporter activity
GO:0034762	4.44	2.00E-02	1	Regulation of transmembrane transport
GO:0034765	4.44	2.00E-02	1	Regulation of ion transmembrane transport
GO:0044070	4.44	2.00E-02	1	Regulation of anion transport
GO:0009416	1.21	2.00E-02	7	Response to light stimulus
GO:0000902	1.66	2.00E-02	4	Cell morphogenesis
GO:0009314	1.18	3.00E-02	7	Response to radiation
GO:0005976	1.55	3.00E-02	4	Polysaccharide metabolic process
GO:0010200	1.72	4.00E-02	4	Response to chitin

Up-regulated in *hub1-4* and *spen3-1*

GO:0050896	1.39	2.17E-08	50	Response to stimulus
GO:0042221	1.49	4.43E-05	33	Response to chemical stimulus
GO:0006952	2.27	6.45E-05	18	Defense response
GO:0010033	1.64	8.04E-04	24	Response to organic substance
GO:0051704	1.90	8.14E-04	19	Multi-organism process
GO:0051707	2.03	1.20E-03	17	Response to other organism
GO:0009607	2.00	1.60E-03	17	Response to biotic stimulus
GO:0065008	2.18	2.00E-03	15	Regulation of biological quality
GO:0012501	3.05	2.30E-03	9	Programmed cell death
GO:0006950	1.31	3.70E-03	28	Response to stress
GO:0009626	4.13	4.50E-03	6	Plant-type hypersensitive response
GO:0042592	2.92	4.60E-03	9	Homeostatic process
GO:0034050	4.11	4.70E-03	6	Host programmed cell death induced by symbiont
GO:0009719	1.76	1.00E-02	18	Response to endogenous stimulus
GO:0008219	2.83	1.00E-02	9	Cell death
GO:0016265	2.83	1.00E-02	9	Death
GO:0009987	0.52	4.00E-02	65	Cellular process

Up-regulated specifically in *hub1-4*

GO:0050896	0.82	3.06E-11	34	Response to stimulus
GO:0006950	0.69	2.10E-03	19	Response to stress
GO:0065007	0.56	3.30E-03	25	Biological regulation
GO:0007623	2.53	1.00E-02	2	Circadian rhythm
GO:0048511	2.53	1.00E-02	2	Rhythmic process
GO:0009987	0.27	1.00E-02	55	Cellular process
GO:0023052	0.86	1.00E-02	11	Signaling
GO:0016310	1.00	1.00E-02	9	Phosphorylation
GO:0000160	2.43	2.00E-02	2	Two-component signal transduction system (phosphorelay)
GO:0006796	0.94	2.00E-02	9	Phosphate metabolic process
GO:0006793	0.94	2.00E-02	9	Phosphorus metabolic process
GO:0006468	1.03	2.00E-02	8	Protein amino acid phosphorylation
GO:0044237	0.32	3.00E-02	42	Cellular metabolic process

Up-regulated specifically in *khd1-1*

GO:0050896	1.1	3.53E-19	41	Response to stimulus
GO:0042221	1.23	1.19E-13	28	Response to chemical stimulus
GO:0006950	1.19	4.49E-12	26	Response to stress
GO:0010033	1.21	1.28E-07	18	Response to organic substance
GO:0006979	2.11	4.37E-06	6	Response to oxidative stress
GO:0006952	1.41	7.54E-05	10	Defense response

GO:0065007	0.71	1.15E-04	27	Biological regulation
GO:0009741	3.17	1.55E-04	3	Response to brassinosteroid stimulus
GO:0040007	1.77	4.93E-04	6	Growth
GO:0009607	1.27	5.13E-04	10	Response to biotic stimulus
GO:0009628	0.97	6.55E-04	15	Response to abiotic stimulus
GO:0051707	1.27	7.53E-04	10	Response to other organism
GO:0051704	1.14	1.10E-03	11	Multi-organism process
GO:0009651	1.45	1.20E-03	8	Response to salt stress
GO:0006970	1.41	1.30E-03	8	Response to osmotic stress
GO:0065008	1.36	1.70E-03	8	Regulation of biological quality
GO:0009719	1.09	2.80E-03	11	Response to endogenous stimulus
GO:0016049	1.83	3.00E-03	5	Cell growth
GO:0050832	2.11	4.30E-03	4	Defense response to fungus
GO:0009725	1.11	4.30E-03	11	Response to hormone stimulus
GO:0008361	1.77	1.00E-02	5	Regulation of cell size
GO:0006869	2.41	1.00E-02	3	Lipid transport
GO:0032535	1.75	1.00E-02	5	Regulation of cellular component size
GO:0090066	1.75	1.00E-02	5	Regulation of anatomical structure size
GO:0019748	1.52	1.00E-02	6	Secondary metabolic process
GO:0050789	0.63	1.00E-02	23	Regulation of biological process
GO:0023052	0.96	1.00E-02	12	Signaling
GO:0010876	2.25	2.00E-02	3	Lipid localization
GO:0032502	0.72	2.00E-02	18	Developmental process
GO:0050794	0.65	2.00E-02	21	Regulation of cellular process
GO:0055114	1.09	3.00E-02	9	Oxidation reduction
GO:0006810	0.85	3.00E-02	13	Transport
GO:0051234	0.84	4.00E-02	13	Establishment of localization

Up-regulated specifically in *spen3-1*

GO:0050896	0.79	3.05E-05	33	Response to stimulus
GO:0042221	1.01	3.68E-05	24	Response to chemical stimulus
GO:0009607	1.42	8.29E-04	11	Response to biotic stimulus
GO:0051707	1.40	0.0015	11	Response to other organism
GO:0008152	0.40	0.0037	56	Metabolic process
GO:0009628	1.01	0.01	16	Response to abiotic stimulus
GO:0051704	1.19	0.01	12	Multi-organism process
GO:0055114	1.29	0.01	10	Oxidation reduction
GO:0042542	3.43	0.01	2	Response to hydrogen peroxide
GO:0015893	3.05	0.02	3	Drug transport
GO:0042493	3.04	0.02	3	Response to drug
GO:0006979	1.89	0.02	6	Response to oxidative stress
GO:0000302	2.96	0.02	3	Response to reactive oxygen species
GO:0010035	1.43	0.03	8	Response to inorganic substance

Table S3. Primer sequences

Gene	AGI code	Primer set	Forward Primer Sequence	Reverse Primer Sequence
Real-time qPCR				
SPEN3	At1g27750		CCCTGCATCAAGTCCCATGT	ACCGATCAAGCATTCCGAGG
KHD1	At1g51580		CCCCATTTGGACCGAGACAA	CCAGGACCATGACAATGCCT
CCA1	At2g46830		CCATGGAAGCCAAAGAAAGT	GGAAGCTTGAGTTTCCAACC
PP2A	At1g13320		TAACGTGGCCAAAATGATGC	GTTCTCCACAACCGCTTGGT
UBC	At5g25760		CTGCGACTCAGGGAATCTTCTAA	TTGTGCCATTGAATTGAACCC
FLC	At5g10140		CCTCTCCGTGACTAGAGCCAAG	AGGTGACATCTCCATCTCAGCTTC
FLC2			TTTGTCCAGCAGGTGACATC	AGCCAAGAAGACCGAACTCA
totalCOOLAIR			GCCGTAGGCTTCTTCACTGT	TGTATGTGTTCTTCACTTCTGTCAA
proxCOOLAIR			CACACCACCAAATAACAACCA	TTTTTTTTTTTTTTTACTGCTTCCA
distCOOLAIR			GGGGTAAACGAGAGTGATGC	TTTTTTTTTTTTTTTGCGGTACAC
CCA1α			GATCTGGTTATTAAGACTCGGAAGCCATATAC	GCCTCTTTCTCTACCTTGGAGA
CCA1β			GAATGTTCCTTGTGATAAGCCATAGAGG	AGGATCGTTCCACTTCCCGTCTT
	At1g77920		CGAGTCCACGCATTATCCCA	CCTCGAGCCGATTGTCTTGT
	At1g77930		TCGAGCTTGATACCGAAGCAG	AGAGACAGAGAGGGAGCAAGT
ChIP-qPCR				
CCA1	At2g46830	P1	GAACAAGTTGATGTTAAGATGGAC	GGAGAAATCTCAGCCACTATAATTATC
		P2	GAAGTTGTGTAGAGGAGCTTAGTG	CTTCCTCAGTCCACCTTTCACGTTGC
		P3	ATCCTCGAAAGACGGGAAGT	GTCGATCTTCATTGGCCATC
		P4	AAGGCTCGATCTTCACTGGA	CCATCCTCTTGCCTTTCTGA
		P5	CTCAAGCTTCCACATGAGACTC	GTTACAGGAAGACTATGGACAAG
FLC	At5g10140	P1	GTTCGGGAGATTAACACAAATAATAAAGG	GAAAACAAGCTGATACAAGCATTTCAC
		P2	GCTGGACCTAACTAGGGGTGAAC	CCTCTTTGGTACGGATCTATAATGAATC
		P3	CCTCTCCGTGACTAGAGCCAAG	CTTCAACATGAGTTCGGTCTGC
		P4	CCTTGGATAGAAGACAAAAAGAGAAAGTG	AGGTGACATCTCCATCTCAGCTTC
Cloning				
KHD1N			ACAGGATCCAAACGTCCGGCGACGACA	ACAGTCGACTCAACTCGTCCCATGTTGGA
RRMSPEN3			CACGGATCCACTCTACGGATCGTAGGAA	CACGTCGACTCAAATCCTTTCACTGGATCAAA
EMSA				
RNA			AAAACAAAAUAGCACCGUAAAGCAC	
DNA			AAAACAAAATACCAGCGTAAAGCAC	

Supplemental dataset S1 (separate file)

Protein Identification details obtained with the 4800 MALDI TOF/TOFTM Proteomics analyzer (AB SCIEX) and the GPS explorer v3.6 (AB SCIEX) software package combined with the search engine Mascot version 2.1 (Matrix Science) and database TAIR8.

Supporting references

1. Alonso JM, et al. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301(5633), 653-657 [Err. *Science* 301(5641), 1849].
2. Rosso MG, et al. (2003) An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol Biol* 53(1-2), 247-259.
3. Sessions A, et al. (2002) A high-throughput Arabidopsis reverse genetics system. *Plant Cell* 14(12), 2985-2994.
4. 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(2), 417-432.
5. Salomé PA, McClung R (2005) *PSEUDO-RESPONSE REGULATOR7* and *9* are partially redundant genes essential for the temperature responsiveness of the Arabidopsis circadian clock. *Plant Cell* 17(3), 791-803.
6. Portolés S, Más P (2007) Altered oscillator function affects clock resonance and is responsible for the reduced day-length sensitivity of CKB4 overexpressing plants. *Plant J* 51(6), 966-977.
7. Dhondt S, et al. (2014) High-resolution time-resolved imaging of *in vitro* Arabidopsis rosette growth. *Plant J* 80(1), 172-184.
8. Skirycz A, et al. (2011) Survival and growth of *Arabidopsis* plants given limited water are not equal. *Nat. Biotechnol* 29(3), 212-214.
9. Clauw P, et al. (2015) Leaf responses to mild drought stress in natural variants of Arabidopsis. *Plant Physiol* 167(3), 800-816. [Corrigendum Plant Physiol. 168(3), 1180].
10. Coppens F, Wuyts N, Inzé D, Dhondt S (2017) Unlocking the potential of plant phenotyping data through integration and data-driven approaches. *Curr Opin Syst Biol* 4, 58-63.
11. Moore A, Zielinski T, Millar AJ (2014) Online period estimation and determination of rhythmicity in circadian data, using the BioDare data infrastructure. *Methods Mol Biol* 1158, 13-44.
12. Van Bel M, et al. (2012) Dissecting plant genomes with the PLAZA comparative genomics platform. *Plant Physiol* 158(2), 590-600.
13. Bruno L, et al. (2011) Multi-probe *in situ* hybridization to whole mount *Arabidopsis* seedlings. *Int J Dev Biol* 55 (2), 197-203.
14. Casamitjana-Martínez E, et al. (2003) Root-specific *CLE19* overexpression and the *sol1/2* suppressors implicate a CLV-like pathway in the control of *Arabidopsis* root meristem maintenance. *Curr Biol* 13(16), 1435-1441.
15. Rigaut G, et al. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* 17(10), 1030-1032.
16. Bürckstümmer T, et al. (2006) An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. *Nat Methods* 3(12), 1013-1019.
17. Van Leene J, et al. (2007) A tandem affinity purification-based technology platform to study the cell cycle interactome in *Arabidopsis thaliana*. *Mol Cell Proteomics* 6(7), 1226-1238.
18. Van Leene J, et al. (2010) Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*. *Mol Syst Biol* 6(1), 397.
19. Krohn NM, Yanagisawa S, Grasser KD (2002) Specificity of the stimulatory interaction between chromosomal HMGB proteins and the transcription factor Dof2 and its negative regulation by protein kinase CK2-mediated phosphorylation. *J Biol Chem* 277(36), 32438-32444.
20. Kammel C, et al. (2013) *Arabidopsis* DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. *PLoS ONE* 8(3), e60644.
21. Shen J, Zhang L, Zhao R (2007) Biochemical characterization of the ATPase and helicase activity of UAP56, an essential pre-mRNA splicing and mRNA export factor. *J Biol Chem* 282(31), 22544-22550.

22. 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(10), 2586-2602.
23. Guindon S, et al. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59(3), 307–321.
24. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52(5), 696-704.
25. Ronquist F, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61(3), 539-542.
26. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* 35(6), 1547-1549.