Fall November 11, 2013

Arabidopsis thaliana AHL family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain

Jianfei Zhao, *University of Pennsylvania*
David Favero
Hao Peng
Michael Neff

Arabidopsis thaliana AHL family modulates hypocotyl growth redundantly by interacting with each other via the PPC/DUF296 domain

Jianfei Zhao[a,b], David S. Favero[a,b], Hao Peng[b], and Michael M. Neff[a,b,1]

[a]Molecular Plant Sciences Graduate Program and [b]Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164

Edited* by Diter von Wettstein, Washington State University, Pullman, WA, and approved October 17, 2013 (received for review November 7, 2012)

The *Arabidopsis thaliana* genome encodes 29 *AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED* (*AHL*) genes, which evolved into two phylogenic clades. The AHL proteins contain one or two AT-hook motif(s) and one plant and prokaryote conserved (PPC)/domain of unknown function #296 (DUF296) domain. Seedlings lacking both *SOB3/AHL29* and *ESC/AHL27* confer a subtle long-hypocotyl phenotype compared with the WT or either single-null mutant. In contrast, the missense allele *sob3-6* confers a dramatic long-hypocotyl phenotype in the light. In this study, we examined the dominant-negative feature of *sob3-6* and found that it encodes a protein with a disrupted AT-hook motif that abolishes binding to AT-rich DNA. A loss-of-function approach demonstrated different, yet redundant, contributions of additional *AHL* genes in suppressing hypocotyl elongation in the light. We showed that AHL proteins interact with each other and themselves via the PPC/DUF296 domain. AHLs also share interactions with other nuclear proteins, such as transcription factors, suggesting that these interactions also contribute to the functional redundancy within this gene family. The coordinated action of AHLs requires an AT-hook motif capable of binding AT-rich DNA, as well as a PPC/DUF296 domain containing a conserved Gly-Arg-Phe-Glu-Ile-Leu region. Alteration of this region abolished SOB3/AHL29's physical interaction with transcription factors and resulted in a dominant-negative allele *in planta* that was phenotypically similar to *sob3-6*. We propose a molecular model where AHLs interact with each other and themselves, as well as other nuclear proteins, to form complexes which modulate plant growth and development.

enhanceosome | seedling establishment

Since the colonization of land plants, plant genomes have rapidly expanded and diversified from their common ancestor through the process of evolution. During the long period of natural selection that shaped the morphology and adaptivity of plants to their environment, some genes that regulate important growth and developmental processes of ancient plant species have been amplified and gradually expanded into multimember gene families (1–4). One such family, the *AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED* (*AHL*) genes, exists in all plant species that have been sequenced thus far, from the moss *Physcomitrella patens* (5) to both modern dicot and monocot plants, such as *Arabidopsis thaliana* (6), *Oryza sativa* (7, 8), *Sorghum bicolor* (9), and *Populus trichocarpa* (10). This high conservation through evolution suggests that the *AHL* gene family is important for plant growth and development.

The AHL proteins contain two conserved structural units, the AT-hook motif and the plant and prokaryote conserved (PPC) domain, the latter being also annotated as the domain of unknown function #296 (DUF296). The AT-hook motif enables binding to AT-rich DNA and has been identified in various gene families both in prokaryotes and eukaryotes including the high mobility group A (HMGA) proteins in mammals (11). The AT-hook motif contains a conserved palindromic core sequence, Arg-Gly-Arg, which binds the minor groove of AT-rich B-form DNA (12). On binding with DNA, the Arg-Gly-Arg core sequence adopts a concave conformation with close proximity to the backbone of the DNA, and the side chains of both arginine residues firmly insert into the minor groove (12).

The PPC/DUF296 domain is ∼120 amino acids in length and exists as a single protein in Bacteria and Archaea (13). Crystal structures of several bacterial and archaeal PPC/DUF296 proteins have been determined (14, 15). All of the above prokaryotic PPC/DUF296 proteins share the same tertiary structure with five β-strands forming an antiparallel β-sheet, which partially surrounds a single α-helix. The solved crystal structures suggest that the prokaryotic PPC/DUF296 proteins form a trimer (14). In land plants, the PPC/DUF296 domain is found in AHL proteins, where it is located at the carboxyl end relative to the AT-hook motif(s) (13). In *Arabidopsis* AHL1, a hydrophobic region at the C-terminal end of its PPC/DUF296 domain is essential for its nuclear localization (13). However, there are currently no known biological functions for this domain in regulating plant growth and development.

Based on gene overexpression and biochemical studies, members of the *AHL* family have been proposed to regulate diverse aspects of growth and development in plants. Overexpression of either of two *Arabidopsis AHL* genes, *SUPPRESSOR OF PHYTOCHROME B-4 #3* (*SOB3/AHL29*) or *ESCAROLA* (*ESC/AHL27*) represses hypocotyl elongation for seedlings grown in the light, but not in darkness (16). As adults, these overexpression plants develop larger organs, including expanded leaves and enlarged flowers and fruits, together with delayed flowering and senescence (16). The *AHL* genes modulate several aspects of plant growth and

Significance

Members of the *AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED* (*AHL*) family are involved in various plant biological processes. Our findings reveal a molecular model whereby the AHLs interact with each other via the plant and prokaryote conserved (PPC)/domain of unknown function #296 (DUF296) domain to form homo-/hetero-complexes, possibly trimers. The AHL complex also interacts with other nuclear proteins to form a macromolecular complex that modulates plant growth and development. The coordinated action of AHLs requires an AT-hook motif capable of binding AT-rich DNA, as well as a PPC/DUF296 domain containing a conserved six-amino-acid region. Our proposed model provides a better understanding of the roles of *AHL* genes in regulating plant growth and development, which may in turn lead to better seedling establishment and increased yield.

Author contributions: J.Z., D.S.F., H.P., and M.M.N. designed research; J.Z., D.S.F., and H.P. performed research; J.Z., D.S.F., H.P., and M.M.N. analyzed data; and J.Z., D.S.F., H.P., and M.M.N. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]To whom correspondence should be addressed. E-mail: mmneff@wsu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1219277110/-/DCSupplemental.

development, including hypocotyl elongation in the light (*AHL22*, *ESC*/*AHL27,* and *SOB3*/*AHL29*) (16, 17), flower development (16–20), and root growth (21). Overexpression of other *AHL* gene members, such as *HERCULES* (*HRC*/*AHL25*), also enhances adult leaf and stem growth (22). Several *AHL* genes are involved in the homeostasis of phytohormones or the mediation of their responses. For example, overexpression of either *AT-HOOK PROTEIN OF GA FEEDBACK REGULATION 1* (*AGF1*/*AHL25*) or *AGF2*/*AHL15* suggests that these genes play a role in maintaining the negative feedback of *GA 3-oxidase* in gibberellin signaling (23). Based on yeast-one-hybrid and transactivation analysis, *Catharanthus roseus AHL*s have been shown to regulate the jasmonic acid response of AP2 transcription factors (24). Transcript accumulation of *AHL21* was dramatically suppressed in cytokinin-treated *Arabidopsis* roots (25). Overexpression studies also implicate *AHL* genes (*ESC*/*AHL27* and *AHL20*) in the regulation of plant innate immune responses (19, 26). Gene overexpression and related biochemical analyses, however, can be misleading due to neomorphic or hypermorphic phenotypes; in other words, those reactions caused by gene misexpression or biologically irrelevant protein levels. Therefore, it is important to complement gene overexpression and biochemical studies with loss-of-function genetic analysis to unveil the biological functions of the *AHL* genes in plants.

Our previous loss-of-function study showed that two *Arabidopsis AHL*s, *SOB3*/*AHL29* and *ESC*/*AHL27,* repress hypocotyl growth redundantly in light-grown seedlings (16). Single loss-of-function mutants for either *SOB3*/*AHL29* (*sob3-4*) or *ESC*/*AHL27* (*esc-8*) have WT phenotypes. In contrast, the *sob3-4 esc-8* double mutant exhibits slightly increased hypocotyl growth under continuous white, red, far-red, and blue light, providing the first *AHL* loss-of-function analysis and demonstrating that at least *SOB3*/*AHL29* and *ESC*/*AHL27* act as negative regulators of hypocotyl elongation. In support of these observations, an *AHL18-AHL22-ESC*/*AHL27-SOB3*/*AHL29* RNAi line also exhibits a long-hypocotyl phenotype in white, blue, red, and far-red light (17). These studies suggest functional redundancy among multiple *AHL* genes in *Arabidopsis*. In contrast, genetic analyses of single AHL loss-of-function mutants in monocots demonstrated that they regulate inflorescence and floral organ growth, as well as palea formation (27, 28). Although we are beginning to understand the roles of *AHL* genes in plant growth and development, especially with regard to hypocotyl elongation, the molecular mechanism underlying their functional redundancy in *Arabidopsis* and other plants is not known.

In contrast to the subtle phenotypes shown by *sob3-4 esc-8* (16) and the quadruple RNAi line (17), the missense activation-tagged allele *sob3-6*, which was identified in our previous study, confers a much more severe long-hypocotyl phenotype in the light (16). Analysis of this dominant-negative *sob3-6* allele, which is caused by a single amino acid change in the conserved core of the AT-hook, has led to the hypothesis that this DNA-binding domain, as well as protein-protein interactions, is necessary for the biological function of this gene family.

In this study, we show that the *sob3-6* allele encodes a full-length protein that no longer binds AT-rich DNA. Loss-of-function analysis with triple- and quadruple-null mutants demonstrates that multiple *AHL* genes, but not all, suppress hypocotyl growth in the light. We also show that multiple WT AHLs, as well as the SOB3-6 mutant protein, can associate with each other in the nucleus and that the PPC/DUF296 domain mediates this interaction. Overexpression of this domain also leads to a dominant-negative long hypocotyl phenotype, which is similar to that of *sob3-6*. In addition, a yeast two-hybrid (Y2H) library screen identified non-AHL, nuclear-localized, DNA-binding proteins that physically interact with AHLs. We demonstrate that a conserved six-amino-acid region in the PPC/DUF296 domain of SOB3/AHL29 is necessary for transcriptional activation in yeast and interaction with at least one transcription factor. Removal of this region also confers a dominant-negative long hypocotyl phenotype when expressed in plants. Based on these results, we propose a molecular model whereby AHLs interact with themselves and each other, as well as other nuclear proteins, to form complexes similar to the human enhanceosome, which modulate plant growth and development.

Results

AHL Gene Family in *A. thaliana*. The *A. thaliana* genome encodes 29 *AHL* genes (Fig. 1*A*). A Bayesian phylogenetic analysis showed that these 29 *Arabidopsis AHL* genes evolved into two phylogenetic clades (clades A and B; Fig. 1*B*). Clade A *AHL*s are intronless with only one AT-hook motif and a single PPC/DUF296 domain. Clade B consists of intron-containing *AHL*s with either one or two AT-hook motif(s) and a single PPC/DUF296 domain (Fig. 1*B*). The type 1 and 2 AT-hook motifs are distinguished by the sequences that flank the conserved Arg-Gly-Arg core, especially those at the carboxyl end. The consensus sequence at the carboxyl end of the conserved core in type 1 AT-hook motifs in AHL proteins is Gly-Ser-Lys-Asn-Lys, whereas Arg-Lys-Tyr-X is found at the carboxyl end of the core in type 2 AT-hook motifs (Fig. 1 *C* and *D*). Within each clade, the *Arabidopsis AHL*s tend to coexpress with other members (Fig. S1) (29, 30).

The PPC/DUF296 domain of SOB3/AHL29 is predicted to share the same order of secondary structural elements, β_1-α-β_2-β_3-β_4-β_5-β_6-β_7-β_8-β_9, as observed in their prokaryotic counterparts (Fig. S2). We predicted the tertiary structure of the PPC/DUF296 domain from SOB3/AHL29 via homology modeling (Fig. S3). It is also predicted to adopt a similar tertiary structure as observed in the crystal structures of prokaryotic PPC/DUF296 proteins. The PPC/DUF296 proteins form a trimer, where the β-sheets mediate interactions among the three PPC monomers (14). This structure suggests that the PPC/DUF296 domains in *Arabidopsis* AHL proteins could associate with each other as their counterparts do in prokaryotes.

***sob3-6* Is a Dominant-Negative Allele Disrupting AHL Protein Interactions with DNA.** To obtain loss-of-function alleles, we treated activation-tagged *SOB3-D* seeds with ethyl methanesulfonate (EMS) and screened for intragenic suppressors of the short-hypocotyl phenotype conferred by overexpression of *SOB3*/*AHL29* (16). In addition to the *sob3-4* null allele, we identified two additional missense mutant alleles, *sob3-5* and *sob3-6*, which produced longer hypocotyls in white light than the WT (Fig. 2*A*) (16). The *sob3-5* allele has a Gly^{80} to Gln mutation, which is adjacent to the carboxyl end of the core sequence of the type 1 AT-hook motif (Fig. 1*C*). The *sob3-6* mutation is at the second arginine (Arg^{77} to His) in the conserved AT-hook core motif (Fig. 1*C*). Mutations in this Arg-Gly-Arg core motif have been shown to abolish DNA binding by the AT-hook motif in non-AHL proteins (31, 32).

Due to the nature of the mutation, the SOB3-6 protein is likely to have minimal or no AT-rich DNA binding capability. To test this hypothesis, we performed an EMSA with SOB3 and SOB3-6 proteins using a probe sequence taken from the AT-rich *PRA2* promoter from pea (*Pisum sativum*) (17, 19). SOB3 readily binds with the biotin-labeled oligonucleotide *PRA2* probe (Fig. 2*B*). Incubation with an anti-SOB3 antibody resulted in a supershift of the *PRA2* probe, indicating the presence of SOB3 in the protein-DNA complex. In addition, binding of the *PRA2* probe by SOB3 could be competed away by adding a cold (unlabeled) competitive probe. In contrast, no DNA-binding could be detected when using SOB3-6 protein instead of SOB3 (Fig. 2*B* and Fig. S4*A*).

sob3-6 was identified as an intragenic suppressor of the short-hypocotyl phenotype conferred by the gain-of-function *SOB3-D* mutant, leading to a light-grown hypocotyl that is even longer than the WT (16). The observation that the *sob3-6* allele confers

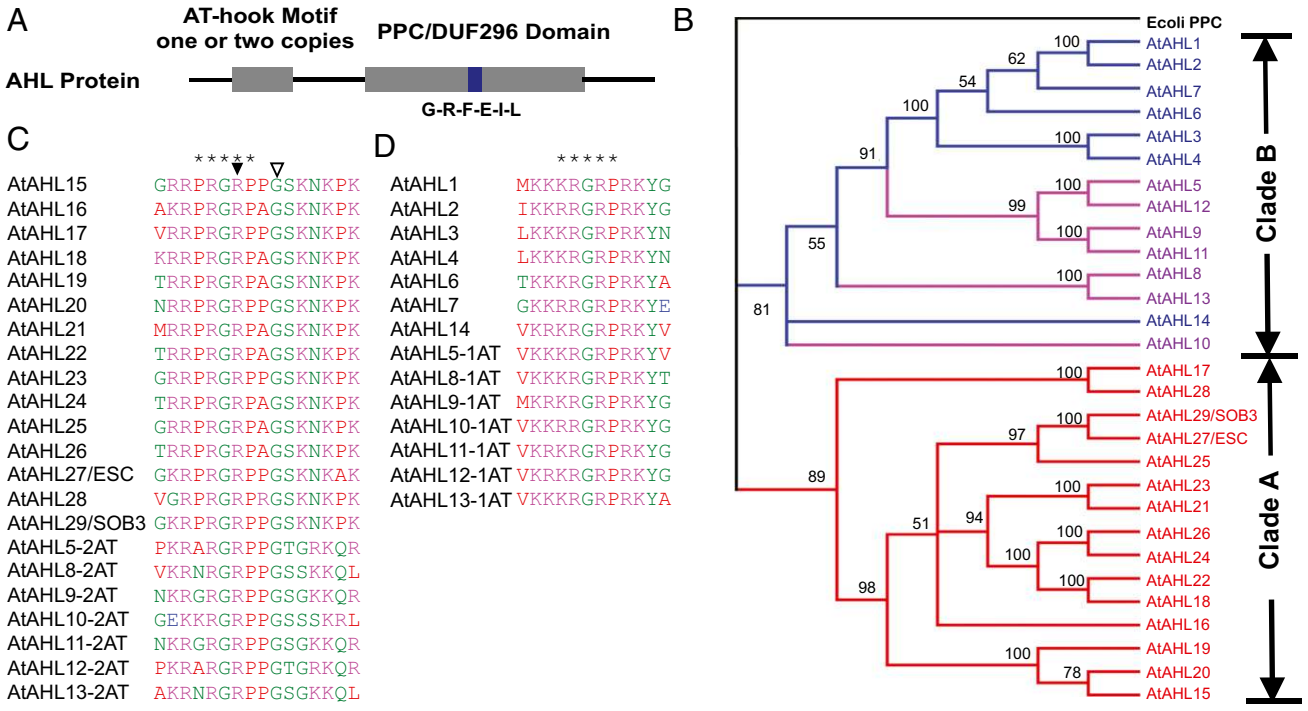

Fig. 1. The *AHL* gene family in *A. thaliana*. (*A*) Topologies of the AHL proteins. A conserved six-amino-acid region in the PPC/DUF296 domain is highlighted by the blue box. (*B*) Phylogeny of the *AHL* gene family using Bayesian analysis. Numbers near branches indicate Bayesian posterior probabilities for given clades. *AHL* genes containing only one type 1 AT-hook motif are shown in red. *AHL* genes containing one type 2 AT-hook motif are shown in blue. *AHL* genes containing two AT-hook motifs (one each of type 1 and 2) are shown in purple. Two forms of the AT-hook motif, types 1 (*C*) and 2 (*D*), were identified in the *Arabidopsis* AHL proteins. The stars indicate the sequences of the AT-hook motifs. The arrowhead and open arrowhead represent the *sob3-6* mutation (R-to-H) and *sob3-5* mutation (G-to-Q) in SOB3/AHL29, respectively.

the opposite phenotype of *SOB3-D* suggests that the Arg^{77} to His missense mutation disrupts AT-hook motif function, rendering *sob3-6* a negative allele. The *sob3-6* negative allele was initially identified as exhibiting a long-hypocotyl phenotype in white light–grown heterozygous *sob3-6*/*SOB3-D* plants (16). Because one copy of *sob3-6* was sufficient to suppress the *SOB3-D* gain-of-function phenotype, we hypothesized that this missense mutation functions as a dominant-negative allele. We tested this hypothesis by overexpressing the *sob3-6* cDNA under the constitutive *Cauliflower Mosaic Virus* (*CaMV*) *35S* promoter in WT *Arabidopsis*. Multiple independent T1 (hemizygous) and T2 transgenic lines recapitulated the long-hypocotyl phenotype of the original *sob3-6* allele when grown in white light, with some lines conferring an even longer hypocotyl phenotype than the original *sob3-6* mutant line (Fig. 2*A*).

We also overexpressed another *AHL*, *ESC*/*AHL27*, driven by the *CaMV 35S* promoter in *Arabidopsis*. The *ESC*/*AHL27* overexpression lines exhibited similar suppressed hypocotyl growth as seen in *SOB3-D* seedlings (16) (Fig. 2*C*). We further created the same *sob3-6*-like mutation (Arg^{91} to His) in the AT-hook motif of the *ESC*/*AHL27* gene (designated as *esc-11*). Multiple independent T1 and T2 *esc-11* overexpression lines (also driven by the *CaMV35S* promoter) exhibited long hypocotyls in white light, recapitulating the *sob3-6* phenotype (Fig. 2*C*) and demonstrating that multiple *AHLs* with a *sob3-6*–like mutation function as dominant-negative alleles.

***AHL* Genes Contribute Differently to the Suppression of Hypocotyl Elongation in White Light.** The *sob3-6* and *esc-11* overexpression lines both exhibited longer hypocotyls than the *sob3-4 esc-8* double-null mutant. This observation suggests that other *AHL*s with similar functions are also affected by these dominant-negative alleles. To test this hypothesis and further examine the function of additional *AHL*s with regard to suppression of hypocotyl elongation in the light, we generated four *AHL* triple-null mutant combinations (Fig. 2*D* and Fig. S4 *B*–*E*). Three triple-null mutants (*ahl6 sob3-4 esc-8*, *ahl15 sob3-4 esc-8,* and *ahl22 sob3-4 esc-8*) exhibited longer hypocotyls than the *sob3-4 esc-8* double-null, although they varied with regard to the severity of the phenotype ($P < 4.5E-4$). A fourth triple-null mutant, *sob3-4 esc-8 ahl5*, had a similar light-grown hypocotyl length as the *sob3-4 esc-8* double-null (Fig. 2*D*). We also made the *sob3-4 esc-8 ahl6 ahl22* quadruple-null mutant. The *sob3-4 esc-8 ahl6 ahl22* quadruple null conferred an even longer hypocotyl phenotype than both triple-null lines when grown in white light, although this phenotype was still shorter than that of *sob3-6* ($P < 3.2E-7$; Fig. 2*D*). Together these data demonstrate that these *AHL* genes contribute differently to white light–mediated suppression of hypocotyl elongation.

AHL Proteins Form Homo- and Hetero-Complexes. The dominant-negative long-hypocotyl phenotype of *sob3-6* and *esc-11* overexpressing seedlings suggests a model whereby these mutant proteins physically interact with other AHL family members and/or other shared interacting partners, rendering the complex functionally inactive due to abolished DNA binding by the mutated AT-hook motif. To test this hypothesis, we first performed a targeted GAL4-based Y2H assay (Fig. 3*A*). Using the lowest concentration of 3-amino-1, 2, 4-triazol (3-AT) that prevented transcriptional autoactivation by the AHL baits, we demonstrated that SOB3/AHL29 and ESC/AHL27 interact with themselves as well as with each other (Fig. 3*A* and Fig. S5*A*).

We further examined the AHL-AHL interactions *in planta* using a bimolecular fluorescence complementation (BiFC) assay via transient expression in onion epidermal cells (33, 34). In this experiment, two halves of the enhanced YFP (EYFP) (nEYFP, N terminus half of EYFP; cEYFP, C terminus half of EYFP)

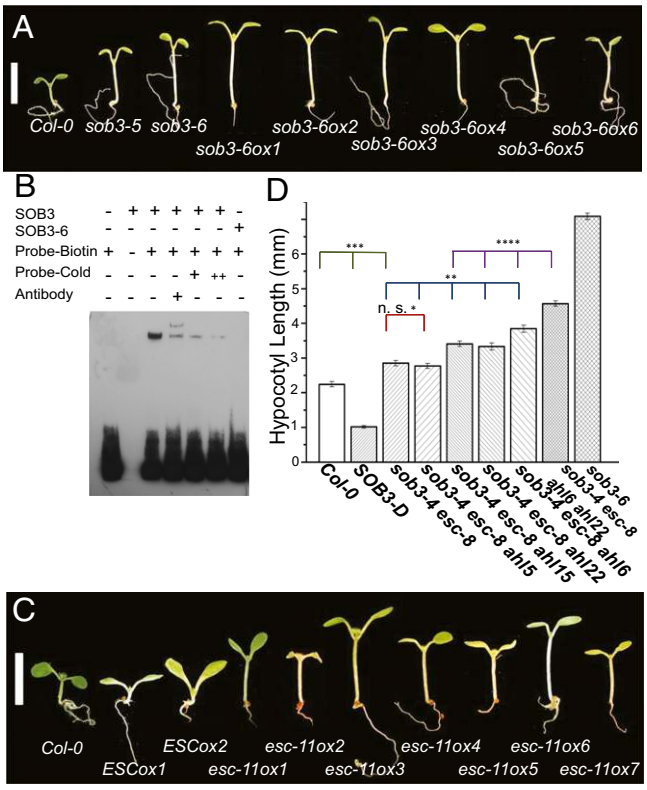

Fig. 2. *sob3-6* is a dominant-negative allele involved in regulating hypocotyl elongation in light growth. (*A*) Hypocotyl growth of WT *Col-0*, *sob3-5*, *sob3-6*, and multiple independent transgenic lines overexpressing the *sob3-6* allele, growing in 30 μmol·s^{-1}·m^{-2} white light. (*B*) SOB3 bound to the biotin-labeled pea *PRA2* probe in EMSA. A super shift was detected with the addition of the anti-SOB3 antibody. With increasing concentrations of cold probe, the binding of *PRA2* probe by SOB3 could be competed away. However, SOB3-6 did not bind to the *PRA2* probe in EMSA. (*C*) Hypocotyl growth of *Col-0*, *ESC/AHL27* overexpression lines (*ESCox1* and *ESCox2*) and multiple *esc-11* overexpression lines, growing in 30 μmol·s^{-1}·m^{-2} white light. (Scale bar, 5 mm.) (*D*) Six-day-old *Arabidopsis* seedlings were grown in 25 μmol·m^{-2}·s^{-1} white light at 25 °C. The error bar denotes the SEM. In a Student *t* test (unpaired two-tailed *t* test with unequal variance). *not significant with $P = 0.39$. **$P < 4.5E-4$. ***$P < 7.5E-7$. ****$P < 3.2E-7$. *Col-0*, $n = 24$; *SOB3-D*, $n = 29$; *sob3-4 esc-8*, $n = 34$; *sob3-4 esc-8 ahl5*, $n = 50$; *sob3-4 esc-8 ahl15*, $n = 47$; *sob3-4 esc-8 ahl22*, $n = 32$; *sob3-4 esc-8 ahl6*, $n = 31$; quadruple null, $n = 38$; *sob3-6*, $n = 32$.

were translationally fused to the 5′ end of *SOB3/AHL29* and *ESC/AHL27* (Fig. 3*B* and Fig. S5 *B–F*). A separate plasmid expressing a monomeric red fluorescent protein (mRFP) was also cotransformed with the nEYFP and cEYFP fusions as an indicator of successful transformation. The mRFP protein was expressed in the cytoplasm and in the nucleus. The reconstructed yellow fluorescence signal could be detected in the transformed onion epidermal cells and colocalized with the red fluorescent signal only in the nucleus (Fig. 3*B*). No yellow fluorescence was observed with the negative controls using empty BiFC vectors (Fig. S5 *G–L*).

We also used the BiFC assay to demonstrate interactions among SOB3/AHL29, ESC/AHL27, and another clade A AHL, HRC/AHL25 (Fig. 3*C*). In addition, we used the Y2H and BiFC assays to test and demonstrate that two clade B AHLs, AHL5 and AHL12, interact with SOB3/AHL29 and ESC/AHL27, as well as with themselves (Fig. S5 *M–P*). AHL5 and AHL12 each contain two AT-hook motifs (one each of type 1 and 2) and one PPC/DUF296 domain. Thus, the tested AHL proteins in both clades can associate with each other and with themselves, regardless of the number and type of the AT-hook motif(s).

To further characterize the AHL protein complex *in planta*, we performed a BiFC-FRET assay in onion epidermal cells (Fig. 3*D*). Three copies of SOB3/AHL29 were fused separately to the two halves of the EYFP, as well as to intact cyan fluorescent protein (CFP). The resulting three plasmids were simultaneously transformed into onion epidermal cells (Fig. 3*D*). Yellow fluorescence

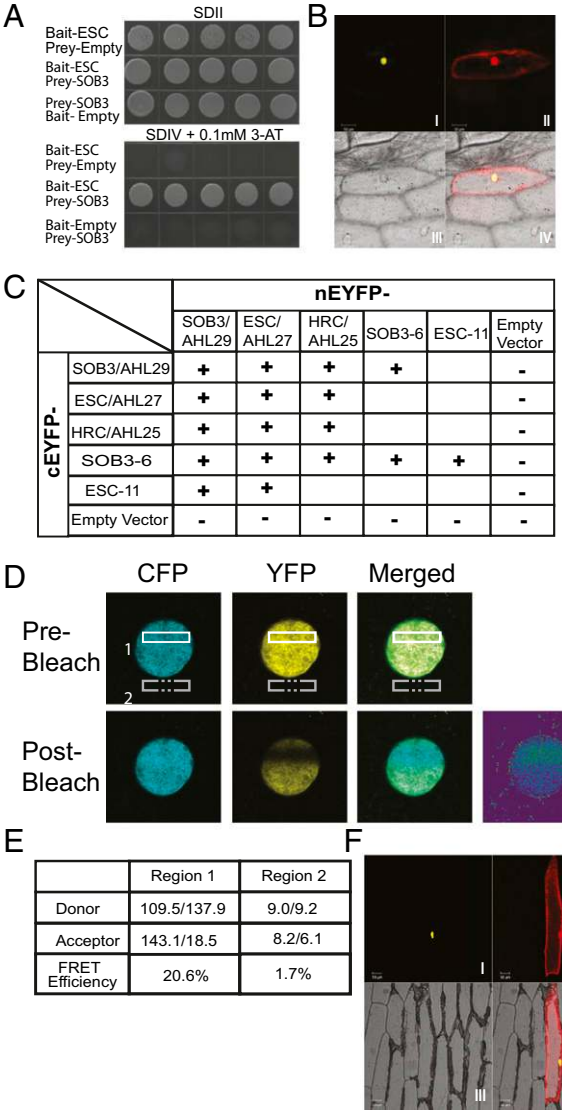

Fig. 3. The AHL proteins interact with each other and themselves. SOB3/AHL29 interacts with ESC/AHL27 in Y2H (*A*) and BiFC (*B*) assays. Yeast transformed with the indicated prey and bait plasmids were plated on SDII, as well as on SDIV media (supplemented with 3-AT of indicated concentration), and checked for growth after 5 d in 30 °C. Five independent biological replicates are shown. Onion epidermal cells, transformed with plasmids expressing nEYFP-SOB3 and cEYFP-ESC, were checked for reconstructed yellow fluorescence signal (*I*) and red fluorescence signal (*II*), observed in bright field (*III*) and merged (*IV*). The protein-protein interactions tested by BiFC and/or Y2H were summarized as in *C*. Interaction among three copies of SOB3/AHL29 proteins was examined by BiFC-FRET assay (*D*). Onion epidermal cells were transformed with plasmids expressing nEYFP-SOB3, cEYFP-SOB3, and CFP-SOB3. (*E*) Fluorescent intensities (before bleach/after bleach) were examined individually in region 1 and control region 2 with laser bleaching at the acceptor's (YFP) excitation wavelength. (*F*) SOB3-6 and ESC-11 interact with each other in the BiFC assay.

could be observed in the nucleus due to reconstitution of EYFP by the physical interaction between the two SOB3/AHL29 proteins that were each fused either to nEYFP or cEYFP. Cyan fluorescent signal could also be detected in the nucleus due to the nuclear localization of SOB3/AHL29. A laser, set at the acceptor's excitation wavelength, was then used to bleach a selected region in the nucleus. Following this treatment, the intensity of yellow fluorescence diminished, whereas the cyan fluorescence increased, suggesting that, before bleaching, FRET was occurring between the donor (CFP) and the acceptor (reconstructed EYFP) (Fig. 3*E*). The high efficiency of FRET indicated that the three SOB3/AHL29 proteins form at least a trimer complex *in planta*.

***sob3-6* and *esc-11* Mutations Do Not Abolish AHL Physical Interactions or Nuclear Localization.** We further used the Y2H and BiFC assays to test whether SOB3-6 and ESC-11 proteins interact with their WT forms in yeast and *in planta*. In both assays, the mutant proteins interacted with WT SOB3 and ESC proteins and with themselves (Fig. 3*F* and Fig. S5*Q*). In addition, these interactions still occurred in the nucleus (Fig. 3*F*), demonstrating that the dominant-negative effect of these alleles is not acting through the disruption of nuclear localization and further that disruption of the AT-hook motif did not abolish the ability of AHLs to form homo- or hetero-complexes.

AHLs Interact via the PPC/DUF296 Domain. We fused the PPC/DUF296 domain from SOB3/AHL29 to the carboxyl end of CFP to examine its subcellular localization via transient expression in onion epidermal cells. The PPC/DUF296 domain of SOB3/AHL29 localized to the nucleus of onion epidermal cells (Fig. 4 *A* and *B*). This finding is consistent with the report that the C-terminal hydrophobic region of the PPC/DUF296 domain is essential for the nuclear localization of AHL1 (13). The SOB3/AHL29 PPC/DUF296 domain also interacted with full-length SOB3/AHL29 (Fig. S5 *R* and *S*) and ESC/AHL27 (Fig. S5*T*) proteins in Y2H assays. A role for the PPC/DUF296 domain in mediating protein-protein interactions *in planta* was further supported by BiFC analysis (Fig. 4*C*). The SOB3/AHL29 PPC/DUF296 domain also interacted with SOB3-6 and ESC-11 mutant proteins in both assays (Fig. 4*D* and Fig. S5 *U* and *V*).

To test the hypothesis that the dominant-negative nature of *sob3-6* is caused by the association of its PPC/DUF296 domain with other AHL proteins, we overexpressed this domain under the control of the constitutive *CaMV 35S* promoter in WT *Arabidopsis* and examined the hypocotyl phenotype in white light (Fig. 4*E*). Multiple independent T1 seedlings exhibited a long-hypocotyl phenotype in white light. Many lines recapitulated the phenotypes of *sob3-6*/*esc-11* overexpression, as well as the original *sob3-6* line (Fig. 2 *A* and *C*), demonstrating that overexpression of the SOB3 PPC/DUF296 domain alone functions in a dominant-negative manner (Fig. 4*E*). Together these results support the hypothesis that the AHLs suppress hypocotyl elongation in *Arabidopsis* seedlings by associating with each other via the PPC/DUF296 domain and that this suppression requires functional AT-hook motifs that recognize AT-rich chromosomal regions.

AHLs Interact with Other Nuclear-Localized Non-AHL Proteins. To further test the hypothesis that AHL proteins interact with each other and possibly have shared interacting partners, we performed a blind Y2H library screen using multiple AHL proteins as the bait. We identified multiple AHLs as interacting partners (Table 1 and Table S1) supporting the AHL-AHL physical interactions identified in our study. Interestingly, using just the SOB3/AHL29 PPC/DUF296 domain alone as bait in the library screen, we identified SOB3/AHL29, as well as AHL3, as interaction partners. In addition, we also identified several transcription factors that interact with the AHLs (Table 1, Table S1, and Fig. S6 *A*–*H*), suggesting that the redundancy among AHLs is not only attributable to their physical interactions with each other but also due to other common interacting partners. We also found that core histones (histones H2B, H3, and H4) in

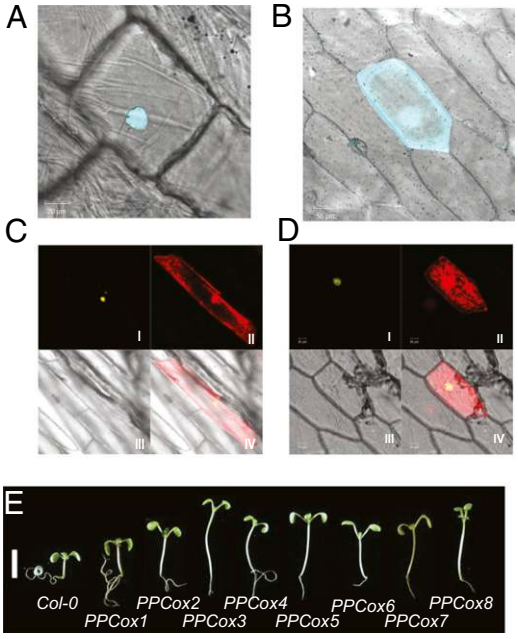

Fig. 4. The PPC/DUF296 domain mediates interactions between AHL proteins. The SOB3 PPC/DUF296 domain tagged with CFP localizes to the nucleus in transformed onion epidermal cells (*A*) in contrast to free CFP (*B*). The SOB3 PPC/DUF296 domain interacts with SOB3 in the BiFC assay (*C*). The PPC/DUF296 domain of SOB3/AHL29 also interacts with SOB3-6, which contains a disrupted AT-hook motif, in the BiFC assay (*D*). (*E*) Compared with WT *Col-0* (*Left*), overexpression of the PPC/DUF296 domain in SOB3 leads to longer hypocotyl growth in the light in multiple independent transformation events.

Table 1. Combined interactors identified in Y2H library screens and/or targeted Y2H assays

Interaction	AGI number
AHLs interacted with AHLs	
AHL1 (as bait)—AHL8	
AHL5 (as bait)—AHL3	
AHL27 (as bait)—AHL23	
AHL27 (as bait)—AHL20	
PPC/DUF296 of SOB3 (as bait)—AHL29/SOB3	
PPC/DUF296 of SOB3 (as bait)—AHL3	
Transcription factors interacted with AHLs	
AP2/EREBP	At5G61890
NAC transcription factor (ATAF2)	At5G08790
Zinc-finger protein	At1G09520, At2G36930
TCP transcription factor 4	At3G15030
TCP transcription factor 13	At3G02150
TCP transcription factor 14	At3G47620
Transcriptional coactivator (KIWI)	At5G09250
Methyltransferase interacted with AHLs	
Methyltransferase	AT5G10830
Histones interacted with AHLs	
Histone H2B	At1G07790
Histone H3.3	At5G10980, At4G40040
Histone H4	At3G53730

nucleosomes interact with multiple AHLs in the Y2H analysis (Fig. S6 *I–K* and Table S1).

Conserved Region in the PPC/DUF296 Domain of SOB3/AHL29 Is Essential for Its Activation of Transcription and Physical Interaction with Non-AHL Interactors. Alignment of amino acid sequences of the PPC/DUF296 domain from the AHLs of several sequenced plant genomes reveals a conserved six-amino-acid region, Gly-Arg-Phe-Glu-Ile-Leu, which is predicted to exist as one β-strand in its secondary structure (Fig. S2). We deleted this six-amino-acid region from the coding sequence of *SOB3/AHL29* and designated it as *sob3-ΔG*. The deletion did not affect the nuclear localization of the mutant SOB3-ΔG protein (Fig. 5*A*). In addition, it did not disrupt interactions with WT SOB3/AHL29 or ESC/AHL27 (Fig. 5*B* and Fig. S6 *L* and *M*). The PPC/DUF296 domain of SOB3/AHL29 also interacted with SOB3-ΔG in both assays (Fig. 5*C* and Fig. S6 *N* and *O*). These results, together with the observed interactions of SOB3-ΔG with both SOB3-6 and the ESC protein with the same deletion (ESC-ΔG) (Fig. S6 *P–R*), suggest that protein stability, nuclear localization, and association among AHL proteins do not depend on this region.

AHL25 from *Arabidopsis* has been shown to activate transcription in a yeast one-hybrid assay (23). We tested SOB3/AHL29 for transcriptional activation in yeast (Fig. 5*D* and Fig. S7*A*). With no 3-AT supplemented in synthetic dropout IV (SDIV) minimal media, yeast expressing SOB3/AHL29 as a Y2H bait exhibited strong autoactivation activity (Fig. 5*D*). Interestingly, the SOB3/AHL29 PPC/DUF296 domain alone tagged with the DNA-binding domain used in the Y2H assay also conferred autoactivation (Fig. S5*S*). This observation suggests that the PPC/DUF296 domain is responsible for transcriptional activation by the AHL proteins in yeast.

In contrast, SOB3-ΔG demonstrated a complete loss of transcriptional activation compared with the WT SOB3 (Fig. 5*D* and Fig. S7*A*). Yeast transformed with WT *SOB3/AHL29* or *SOB3-ΔG* survived on synthetic dropout II (SDII) media. When transferred to SDIV media without 3-AT, only the yeast transformed with WT *SOB3/AHL29* survived, due to activation of transcription by the AHL protein. Neither growth on SDIV media nor β-gal activity was observed in the five biological yeast replicates that were transformed with *SOB3-ΔG* (Fig. 5*D*). We also created another mutant allele (designated as *SOB3-GTA*) in which the six amino acids were substituted with alanine residues. SOB3-GTA also lost transcriptional activity in yeast (Fig. S7*A*). Similar alleles created for *ESC/AHL27* (*ESC-ΔG* and *ESC-GTA*) also demonstrated loss of transcriptional activity in yeast in contrast with the WT (Fig. S7*B*). Therefore, this six-amino-acid region in the PPC/DUF296 domain is essential for the activation of transcription by AHL proteins.

If transcriptional activation is necessary for biological function, then removal of this six-amino-acid region would alleviate, or even abolish, SOB3/AHL29 suppression of hypocotyl elongation in the light. To test this hypothesis, we overexpressed *SOB3-ΔG* driven by the *CaMV 35S* promoter in WT *Arabidopsis*. Surprisingly, we observed a long-hypocotyl phenotype in multiple independent hemizygous lines (Fig. 5*E*), which is similar to seedlings overexpressing *sob3-6, esc-11,* or the *SOB3 PPC/DUF296* domain alone (Figs. 2 *A* and *C* and 4*E*). Alleles of *sob3-6* and *SOB3-ΔG*, as well as the SOB3 PPC/DUF296 domain alone, are likely to have similar efficacy with regard to promoting hypocotyl growth in the light (Fig. S7*C*). Overexpression of *SOB3-GTA* by the *CaMV 35S* promoter also recapitulated a similar long-hypocotyl phenotype (Fig. S7*D*). Thus, *SOB3-ΔG* and *SOB3-GTA* function as dominant-negative alleles, even though the mutant proteins still bear functional AT-hook motifs and could bind with the AT-rich DNA *PRA2* probe (Fig. S7*E*). This finding suggests that interactions with transcription factors mediated by this Gly-Arg-Phe-Glu-Ile-Leu region in the PPC/DUF296 domain are necessary for the biological function of AHL proteins.

The dependence on this six-amino-acid region in the PPC/DUF296 domain of SOB3/AHL29 for transcriptional activation in yeast, together with the observed interactions with transcription factors identified in this study, implies that this region may play a key role in mediating the physical interaction between SOB3/AHL29 and transcription factor(s) *in planta*. To test this hypothesis, we first used Y2H and BiFC assays to examine the physical interaction with the transcription factor TCP4. TCP4 interacted with WT SOB3/AHL29 in both the BiFC (Fig. 6*A*) and Y2H assays (Fig. S6*C*). However, TCP4 no longer interacted with SOB3-ΔG (Fig. 6*B*). TCP4 also interacted with WT ESC/AHL27 but not with ESC-ΔG or ESC-GTA (Fig. 6 *C* and *D* and Fig. S6*D*), even though all three forms of ESC were expressed at similar levels (Fig. S8*A*). Similarly, we also tested interactions with another transcription factor, TCP13 (previously also known as TCP10). TCP13 has been shown to localize in the nucleus (35), as well as in the chloroplast (36). From the results of our BiFC assay, as well as targeted Y2H analysis, TCP13 interacted with the WT SOB3/AHL29 protein and ESC/AHL27 (Figs. S6 *A* and *B* and S8 *B* and *C*). However, TCP13 no longer interacted with SOB3-ΔG (Fig. S8 *D* and *E*).

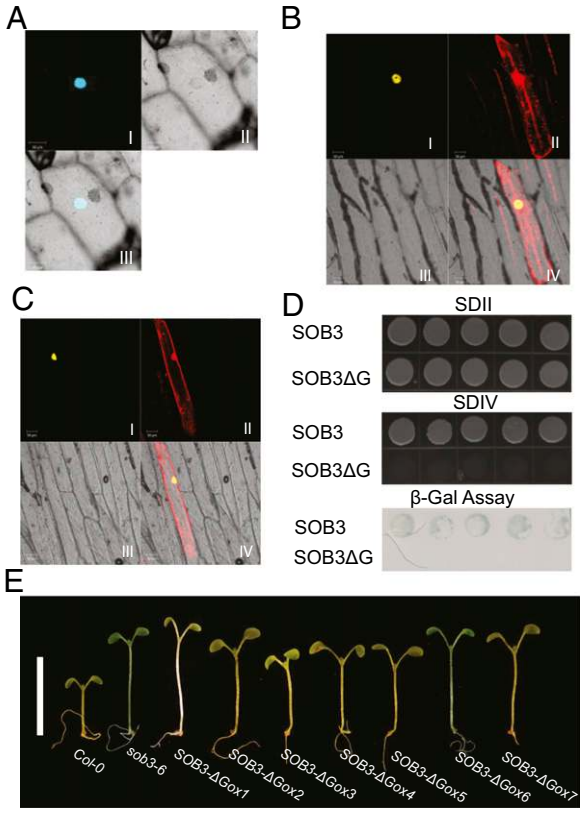

Fig. 5. Transcription activation by SOB3/AHL29 depends on a conserved region in the PPC/DUF296 domain. (*A*) Subcellular localization of SOB3-ΔG protein in onion epidermal cells. Protein-protein interaction analysis of SOB3-ΔG with full-length SOB3/AHL29 in BiFC (*B*). Protein-protein interaction analysis of SOB3-ΔG with the SOB3/AHL29 PPC/DUF296 domain in BiFC (*C*). (*D*) Comparison of transcription activation by SOB3/AHL29- and SOB3-ΔG-containing yeast growing on SDIV plates and in a β-gal assay. Five independent biological replicates are shown. (*E*) Multiple T2 generation, 6-d-old seedlings overexpressing *SOB3-ΔG* under the *CaMV 35S* promoter exhibited a longer hypocotyl phenotype in 23 $\mu mol \cdot s^{-1} \cdot m^{-2}$ white light. (Scale bar, 1 cm.)

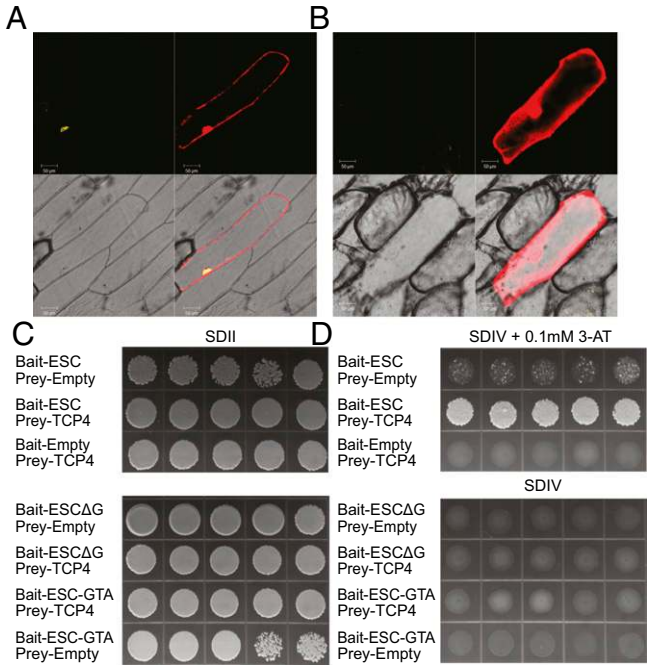

Fig. 6. Interactions of AHLs with transcription factors depends on a six-amino-acid conserved region in the PPC/DUF296 domain. (*A*) SOB3/AHL29 interacted with the transcription factor TCP4, and (*B*) this interaction was abolished when the conserved six amino acids in the PPC/DUF296 domain was removed in the BiFC assay. (*C* and *D*) ESC/AHL27 interacted with TCP4, and the interaction was abolished when the same six-amino-acid region was removed or substituted with alanine residues in a Y2H assay.

Non-AHL Interactors Modulate Hypocotyl Growth in the Light. To understand the roles of the AHL-interacting transcription factors in hypocotyl elongation in the light, we first examined members of the *TCP* gene family. *Jaw-D*, in which several *TCP* genes were specifically knocked down (especially *TCP4*) (37), exhibited a shorter hypocotyl phenotype (Fig. 7*A*), which is consistent with a previous report (38), suggesting *TCP*s play a role in regulating hypocotyl elongation. We also examined roles of other AHL interactors in hypocotyl elongation in the light. We overexpressed the transcription factor ATAF2, one member of the plant-specific NAC gene family, driven by the *CaMV 35S* promoter in WT *Arabidopsis*. The homozygous single-locus insertion overexpression line (*ATAF2ox*) confers a longer hypocotyl than the WT in the light, whereas the loss-of-function line, *ataf2-1*, showed suppressed hypocotyl growth in the light (Fig. 7*B*). Another plant-specific AHL interacting partner that was identified in this study, *Gibberellic Acid-Stimulated Arabidopsis 4* (*GASA4*), also has been implicated in regulating hypocotyl growth in the light (39).

To further understand the biological meaning of the physical interactions between the TCPs and AHLs, we crossed *jaw-D* with the activation-tagged *sob3-6* allele. The *jaw-D sob3-6* double mutant line exhibited a similar hypocotyl length as *jaw-D* (P = 0.13; Fig. 7*A*). This observation suggests that the long-hypocotyl phenotype conferred by the activation-tagged *sob3-6* allele in the light requires the presence of functional TCP proteins. We examined transcript accumulation of *SOB3* (or *sob3-6*) in these lines and observed similar levels of gene expression in each case. These data indicate that the observed genetic epistasis is not due to TCPs affecting transcription accumulation of these alleles of *sob3* (Fig. S9 *A* and *B*). Collectively, these observations demonstrate a direct role for the participation of physical interacting partners in AHL-mediated hypocotyl elongation in the light.

Molecular Model for AHL Proteins. Based on our results, as well as those in the literature, we propose a working model that describes the physical interactions among AHLs, other nuclear proteins, and AT-rich chromosomal DNA (Fig. 8). The AHL proteins form a homo-/hetero-complex via the PPC/DUF296 domain, possibly resulting in a trimer (Fig. 8). The multiple AT-hook motifs of the AHL homo-/hetero-complex then allow it to bind AT-rich DNA regions of the same or different chromosomes. Finally, the AT-rich DNA-anchored AHL complex, aided by the Gly-Arg-Phe-Glu-Ile-Leu region, recruits other nuclear non-AHL proteins, such as transcription factors, to form a DNA-AHL-transcription factor (DNA-AHL-TF) complex, which modulates hypocotyl growth in the light.

Discussion

A unique missense allele, *sob3-6*, caused by an amino acid change in the core of the AT-hook motif, confers a dramatic long-hypocotyl phenotype in the light (16). In this study, we determined that *sob3-6* functions as a dominant-negative allele due to abolished AT-rich DNA-binding ability (Fig. 2). Overexpression of the SOB3/AHL29 PPC/DUF296 domain also led to a similar long-hypocotyl phenotype in light-grown seedlings through a dominant-negative process (Fig. 4). These dominant-negative alleles demonstrated the importance of the AT-hook

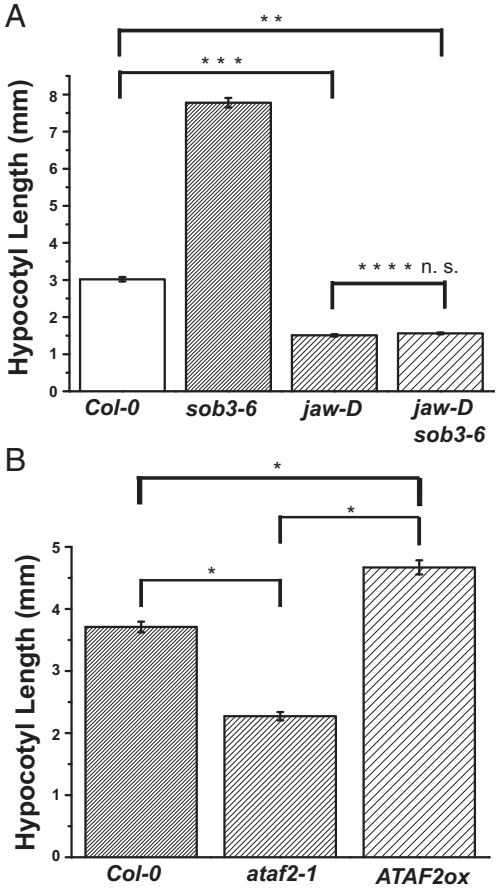

Fig. 7. Hypocotyl analysis of seedlings grown in continuous white light. (*A*) Six-day-old seeedlings were grown in 20 $\mu mol \cdot m^{-2} \cdot s^{-1}$ of white light at 25 °C. *Col-0*, n = 147. *sob3-6*, n = 63. *jaw-D*, n = 70. *jaw-D sob3-6*, n = 232. (*B*) Seedlings (3.5 d old) were grown in 10 $\mu mol \cdot m^{-2} \cdot s^{-1}$ of white light at 25 °C. *Col-0*, n = 48. *ataf2-1*, n = 48. *ATAF2ox*, n = 48. The error bar denotes SEM. In a Student t test (unpaired two-tailed t test with unequal variance): *P < 1.6E-9. **P < 1.7E-54. ***P < 1.3E-55. ****Not significant, P = 0.13.

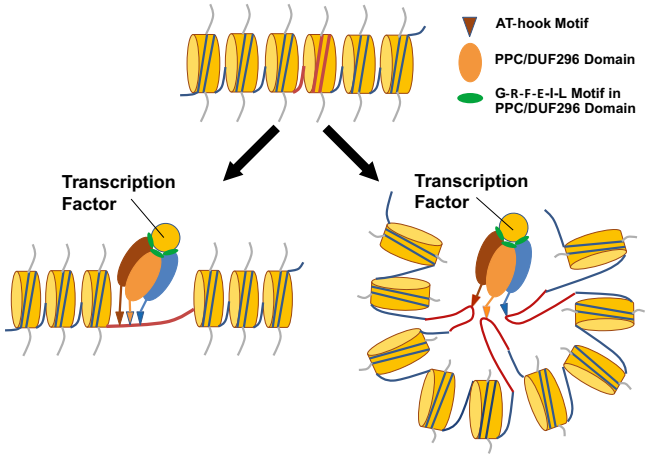

Fig. 8. The *AHL* gene family regulates plant growth and development by forming an AHL complex. The AHL proteins associate with themselves and each other to form a homo-/hetero-trimeric complex via the PPC/DUF296 domains. The complex uses its AT-hook motifs to anchor itself to AT-rich DNA regions (red-color region) and recruit either transcription factors or other non-AHL proteins in order to regulate plant growth and development. In the proposed molecular model, the hypocotyl-growth-promoting functions of associated transcription factors are rendered nonfunctional by their physical interaction with the DNA-AHL-TF complex.

motif(s) in AHL proteins and implied a biological function for the PPC/DUF296 domain. The severity of the phenotype conferred by these dominant-negative mutations also suggests that other *AHL* genes are involved in regulating hypocotyl growth in *Arabidopsis* seedlings.

Our previous study showed that *SOB3/AHL29* and *ESC/AHL27* redundantly suppress hypocotyl elongation in light-grown *Arabidopsis* seedlings (16). As a part of this study, we characterized hypocotyl growth in four triple-null lines, each in which a different *AHL* gene (*AHL5, AHL6, AHL15*, or *AHL22*) was removed from the *sob3-4 esc-8* double mutant background (Fig. 2*D*). We determined that the removal of *AHL5* (unlike the other three *AHL* genes) did not have an impact on hypocotyl elongation. Of the other three *AHL*s, removal of *AHL6* had the greatest impact, demonstrating that *AHL* genes have different quantitative contributions to seedling elongation in the light. The *sob3-4 esc-8 ahl6 ahl22* quadruple-null line exhibited an even longer hypocotyl than the relevant triple-null controls, although it was still shorter than the dominant-negative *sob3-6* mutant (Fig. 2*D*). This difference suggests that there are additional proteins, possibly AHLs, involved in this developmental process.

Given that not all *Arabidopsis AHL*s contribute to light-grown hypocotyl elongation (e.g., *AHL5*), and the probability that some may act redundantly, a dominant-negative approach may be useful for revealing the biological role of other members in this gene family. It is important to note, however, that it would be best if the dominant-negative proteins were expressed according to their endogenous pattern, using the native promoter (possibly with enhancer elements), or via TILLING (40), TALEN (41), or CRISPR/Cas (42) approaches. Such genetic approaches could be very powerful in the case of genomes with even larger *AHL* gene families, such as *Glycine max* (43) and *Brassica rapa* (44), which each have >50 *AHL*s.

In our proposed molecular model, we suggest that the genetic redundancy among AHL proteins is not only due to interactions among themselves, mediated by the PPC/DUF296 domain, but also due to shared, nuclear-localized, non-AHL interactors (Fig. 8). The predicted tertiary structure of the PPC/DUF296 domain of SOB3/AHL29 suggests trimeric oligomerization, which is similar to that observed in seven determined crystal structures of this domain from Bacteria and Archaea [Protein Data Bank (PDB) ID: 2H6L, 2DT4, 2HX0, 2NMU, 2P6Y, 3HTN, and 3HWU] (14, 15). Trimer formation in a proposed AHL complex is also supported by the physical interactions observed among AHLs in this study (Figs. 3 and 4 and Fig. S5). The 29 *AHL* genes encoded by the *A. thaliana* genome allow flexibility in complex formation and subsequent regulation of the related biological processes in plant growth and development.

According to our proposed model, a trimeric AHL complex would have three or more (up to six) AT-hook motifs, depending on the composition of its subunits. The number of AT-hook motifs, as well as their types (type 1 or 2), would determine the AHL complex's affinity for AT-rich DNA. It is important to note that the AT-hook motifs in AHL proteins are identical to the ones in human HMGA proteins. The human genome encodes three HMGA proteins, each of which contains three AT-hook motifs (45). Among the three AT-hook motifs in an HMGA protein, the middle one is identical to the AT-hook motif in SOB3/AHL29 (Fig. S10). These AT-hook motifs also confer the highest affinity to AT-rich DNA (12, 46). The flanking AT-hook motifs of HMGAs are similar to the type 2 AT-hook motif of AHLs. These flanking HMGA AT-hook motifs both have decreased DNA-binding affinity compared with type 1 (12). With these different affinities in mind, trimeric AHL complexes could possess a variety of combinations of AT-hook motifs that in turn could impart to the complex a wide range of binding affinities for AT-rich chromosomal regions. Such AHL complexes could regulate a variety of biological processes.

The human HMGA proteins have been proposed to promote the assembly of an enhanceosome by binding to the AT-rich promoter region of genes and further recruiting other nuclear proteins, such as chromatin remodeling machinery and transcription factors, to positively or negatively regulate gene expression (47–49). In the case of virus-induced transcription of the human *IFN-β* gene, HMGA has been proposed to first bind AT-rich DNA in a 1:1 (protein:DNA) ratio and subsequently alter chromatin status, as well as serve as a platform for the recruitment of eight transcription factors. This process leads to the assembly of the nuclear macromolecular complex (50, 51). The 1:1 HMGA:DNA binding ratio indicates that three AT-hook motifs of one HMGA are necessary to target HMGA to the corresponding chromosome region and induce assembly of the enhanceosome.

Similar promoter binding capabilities have been reported for several AHL proteins. AHL15 and AHL25 bind to the promoter of *GA 3-oxidase 1* in *Arabidopsis* (23). AHL21 binds to the promoter of the *ETTIN/auxin response factor 3* gene (20). AHL22 binds to the promoter of the *Flowering Locus T* gene (18). In *Catharanthus roseus*, five AHLs bind to the promoter of the APETALA2 (AP2) transcription factor *ORCA3* gene (24). These observations suggest that trimeric AHL complexes could also be targeted to promoter regions through their AT-hook motifs.

The three HMGA AT-hook motifs necessary for AT-rich DNA binding correspond to the three motifs (at minimum) in a proposed trimeric AHL complex. This molecular model explains the dominant-negative long-hypocotyl phenotypes conferred by the *sob3-6* allele (which does not bind AT-rich DNA) or expression of the PPC domain alone (which lacks an AT-hook motif). Both dominant-negative alleles encode proteins containing an intact PPC/DUF296 domain, which could still physically interact with WT AHL proteins and form an AHL complex with attenuated DNA-binding activity (Figs. 2 and 4). A similar dominant-negative phenomenon has also been observed with a human HMGA protein bearing a nonfunctional AT-hook motif (32), which resembles the *sob3-6* allele in this study.

We demonstrated that the conserved Gly-Arg-Phe-Glu-Ile-Leu region of the PPC/DUF296 domain is necessary for physical interaction with the transcription factors TCP4 and TCP13 (Fig. 6 and Figs. S6 *A–D* and S8). We also showed that expression of

SOB3 alleles that are mutated in this region confers a dominant-negative long-hypocotyl phenotype (Fig. 5*E* and Fig. S7*D*). Homology modeling of the SOB3/AHL29 PPC/DUF296 domain suggests that this conserved region is located at the exterior of each monomer and could combine to form an exposed quaternary domain after trimerization (Fig. S3*D*). This prediction suggests that the quaternary domain composed of these three six-amino-acid regions could serve as the site of physical interaction between the AHL complex and transcription factors (Fig. 8). However, the exact biological importance of this domain with regard to physical interactions with transcription factors is still unknown.

Some members of the *AHL* gene family are negative regulators of hypocotyl elongation in the light (16, 17) (Fig. 2*D*). In contrast, we identified AHL-interacting transcription factors (ATAF2 and TCPs) that are positive regulators of this event (37, 38) (Fig. 7). In the proposed molecular model, the hypocotyl growth-promoting function of associated transcription factors is rendered nonfunctional by their physical interaction in the DNA-AHL-TF complex. For example, the overexpression of WT *AHL*s (e.g., *SOB3-D*, *SOB3ox,* and *ESCox*) leads to an increased accumulation of DNA-AHL-TF complexes resulting in a short hypocotyl phenotype (16) (Fig. 2). When the relevant *AHL* genes are knocked out in higher-order combinations (e.g., *sob3-4 esc-8 ahl6 ahl22*), or lose their ability to bind DNA (e.g., *sob3-6* and *esc-11*), this leads to a decreased accumulation of DNA-AHL-TF complexes resulting in a long hypocotyl phenotype (16) (Fig. 2 *A* and *C*), possibly by releasing the transcription factors from the complex and allowing them to promote hypocotyl elongation.

Similarly, mutating the conserved Gly-Arg-Phe-Glu-Ile-Leu region of the PPC/DUF296 domain (e.g., *SOB3-ΔG* and *SOB3-GTA*) disrupts interactions with certain transcription factors and also leads to a decreased accumulation of DNA-AHL-TF complexes resulting in a long hypocotyl phenotype via growth promotion from the free transcription factors (Fig. 5*E* and Fig. S7*D*). The *jaw-D sob3-6* double mutant demonstrates the importance of some TCP transcription factors in contributing to the long-hypocotyl phenotype conferred by the *sob3-6* disruption of the DNA-AHL-TF complexes (Fig. 7*A*), presumably due to the lack of free TCPs for promoting hypocotyl elongation. Hence, the *sob3-6* phenotype is gone and the double mutant exhibits a similar hypocotyl length as observed in the *jaw-D* single mutant.

How exactly the physical interaction with the DNA-AHL complex impairs the functions of their associated transcription factors remains to be explored. It is possible that the association with DNA-AHL complex interferes with the normal binding of transcription factors to the promoter regions of the genes that they regulate, subsequently impairing their biological functions. It is also possible that the association with the DNA-AHL complex alters posttranslational modifications associated with the transcription factors, in turn rendering the associated transcription factors nonfunctional. On the other hand, in the *sob3-6* and *esc-11* dominant-negative lines, the abolishment of DNA-binding ability for the AHL complex possibly destabilizes the association between AHLs and transcription factors; therefore, transcription factors that would normally be part of the DNA-AHL-TF complexes are released and function to promote hypocotyl growth in the light.

Through a combination of overexpression and loss-of-function analysis, we determined that multiple members of the *AHL* gene family suppress hypocotyl elongation redundantly in light-grown seedlings. The physical interactions among the AHL proteins via the PPC/DUF296 domain, as well as their interactions with other common nuclear proteins, such as transcription factors, may be important for their biological functions. However, biological functions of each *AHL* gene in the *Arabidopsis* genome still need to be determined. Multiple lines of evidence suggest that functional similarities exist between the HMGA and AHL proteins in that they both form nuclear macromolecular complexes on AT-rich chromatin regions. Further experiments should be conducted to reveal binding sites of AHL proteins in the *Arabidopsis* genome. The biological importance of the protein-protein interactions between the AHL complex and other nuclear proteins also needs to be further determined.

Materials and Methods

Detailed descriptions of plant materials, seedling growth conditions and hypocotyl measurement, plasmid construction, phylogenetic analysis, protein expression and purification, EMSA, bimolecular fluorescent complementation assay, and yeast two-hybrid assays are provided in *SI Materials and Methods*. Briefly, all *Arabidopsis* lines are in *Columbia* background. Sequences of the *AHL* genes and the transfer DNA insertion lines were obtained from the *Arabidopsis* Biological Research Center. Surface-sterilized *Arabidopsis* seeds were sown on media containing 1.0% (wt/vol) phytagel, 1.5% (wt/vol) sucrose, and 0.5× Linsmaier and Skoog–modified basal medium and incubated for 5 d at 25 °C in continuous white light under the indicated intensity. EMSA was performed using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific).

ACKNOWLEDGMENTS. We thank Dr. Raymond Reeves (Washington State University) and the members in the M.M.N. laboratory for comments on the manuscript. We thank the Franceschi Microscopy and Imaging Center at Washington State University for technical support, Drs. Hanjo Hellmann and Sutton Mooney (Washington State University) for the Y2H plasmids and Y2H library, Dr. Ian Street (Dartmouth College) for developing the SOB3/AHL29 antibody used in this study, and Dr. Amit Dhingra (Washington State University) for using the PDS-1000/He system (Bio-Rad). This work was supported by Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Sciences, US Department of Energy Grant DE-PS02-09ER09-02 (to M.M.N.). Analysis of ATAF2 was supported by the United States National Science Foundation, Division of Integrative Organismal Systems Grant 0758411 (to M.M.N.).

1. Lynch M, Conery JS (2003) The origins of genome complexity. *Science* 302(5649): 1401–1404.
2. Bowman JL, Floyd SK (2007) The ancestral developmental tool kit of land plants. *Int J Plant Sci* 168(1):1–35.
3. Bowman JL, Floyd SK, Sakakibara K (2007) Green genes-comparative genomics of the green branch of life. *Cell* 129(2):229–234.
4. Arabidopsis Interactome Mapping C; Arabidopsis Interactome Mapping Consortium (2011) Evidence for network evolution in an Arabidopsis interactome map. *Science* 333(6042):601–607.
5. Rensing SA, et al. (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319(5859):64–69.
6. Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408(6814):796–815.
7. Yu J, et al. (2002) A draft sequence of the rice genome (*Oryza sativa L. ssp. indica*). *Science* 296(5565):79–92.
8. Goff SA, et al. (2002) A draft sequence of the rice genome (*Oryza sativa L. ssp. japonica*). *Science* 296(5565):92–100.
9. Paterson AH, et al. (2009) The *Sorghum bicolor* genome and the diversification of grasses. *Nature* 457(7229):551–556.
10. Tuskan GA, et al. (2006) The genome of black cottonwood, *Populus trichocarpa* (*Torr. & Gray*). *Science* 313(5793):1596–1604.
11. Aravind L, Landsman D (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res* 26(19):4413–4421.
12. Huth JR, et al. (1997) The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat Struct Biol* 4(8):657–665.
13. Fujimoto S, et al. (2004) Identification of a novel plant MAR DNA binding protein localized on chromosomal surfaces. *Plant Mol Biol* 56(2):225–239.
14. Lin LY, et al. (2007) Crystal structure of *Pyrococcus horikoshii* PPC protein at 1.60 A resolution. *Proteins* 67(2):505–507.
15. Lin LY, et al. (2005) Crystallization and preliminary X-ray crystallographic analysis of a conserved domain in plants and prokaryotes from *Pyrococcus horikoshii OT3*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61(Pt 4):414–416.
16. Street IH, Shah PK, Smith AM, Avery N, Neff MM (2008) The AT-hook-containing proteins SOB3/AHL29 and ESC/AHL27 are negative modulators of hypocotyl growth in *Arabidopsis*. *Plant J* 54(1):1–14.
17. Xiao C, Chen F, Yu X, Lin C, Fu YF (2009) Over-expression of an AT-hook gene, *AHL22*, delays flowering and inhibits the elongation of the hypocotyl in *Arabidopsis thaliana*. *Plant Mol Biol* 71(1-2):39–50.

18. Yun J, Kim YS, Jung JH, Seo PJ, Park CM (2012) The AT-hook motif-containing protein AHL22 regulates flowering initiation by modifying *FLOWERING LOCUS T* chromatin in *Arabidopsis*. *J Biol Chem* 287(19):15307–15316.
19. Lim PO, et al. (2007) Overexpression of a chromatin architecture-controlling AT-hook protein extends leaf longevity and increases the post-harvest storage life of plants. *Plant J* 52(6):1140–1153.
20. Ng KH, Yu H, Ito T (2009) AGAMOUS controls GIANT KILLER, a multifunctional chromatin modifier in reproductive organ patterning and differentiation. *PLoS Biol* 7(11):e1000251.
21. Zhou J, Wang X, Lee JY, Lee JY (2013) Cell-to-cell movement of two interacting AT-hook factors in Arabidopsis root vascular tissue patterning. *Plant Cell* 25(1):187–201.
22. Jiang C (2004) US Patent 6,717,034 B2.
23. Matsushita A, Furumoto T, Ishida S, Takahashi Y (2007) AGF1, an AT-hook protein, is necessary for the negative feedback of *AtGA3ox1* encoding GA 3-oxidase. *Plant Physiol* 143(3):1152–1162.
24. Vom Endt D, Soares e Silva M, Kijne JW, Pasquali G, Memelink J (2007) Identification of a bipartite jasmonate-responsive promoter element in the *Catharanthus roseus ORCA3* transcription factor gene that interacts specifically with AT-Hook DNA-binding proteins. *Plant Physiol* 144(3):1680–1689.
25. Rashotte AM, Carson SD, To JP, Kieber JJ (2003) Expression profiling of cytokinin action in *Arabidopsis*. *Plant Physiol* 132(4):1998–2011.
26. Lu H, Zou Y, Feng N (2010) Overexpression of *AHL20* negatively regulates defenses in *Arabidopsis*. *J Integr Plant Biol* 52(9):801–808.
27. Gallavotti A, et al. (2011) *BARREN STALK FASTIGIATE1* is an AT-hook protein required for the formation of maize ears. *Plant Cell* 23(5):1756–1771.
28. Jin Y, et al. (2011) An AT-hook gene is required for palea formation and floral organ number control in rice. *Dev Biol* 359(2):277–288.
29. Obayashi T, Hayashi S, Saeki M, Ohta H, Kinoshita K (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. *Nucleic Acids Res* 37(Database issue):D987–D991.
30. Obayashi T, Nishida K, Kasahara K, Kinoshita K (2011) ATTED-II updates: Condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. *Plant Cell Physiol* 52(2):213–219.
31. Gordon BR, et al. (2011) Structural basis for recognition of AT-rich DNA by unrelated xenogeneic silencing proteins. *Proc Natl Acad Sci USA* 108(26):10690–10695.
32. Himes SR, et al. (2000) The role of high-mobility group I(Y) proteins in expression of IL-2 and T cell proliferation. *J Immunol* 164(6):3157–3168.
33. Citovsky V, Gafni Y, Tzfira T (2008) Localizing protein-protein interactions by bimolecular fluorescence complementation in planta. *Methods* 45(3):196–206.
34. Citovsky V, et al. (2006) Subcellular localization of interacting proteins by bimolecular fluorescence complementation in planta. *J Mol Biol* 362(5):1120–1131.
35. Suzuki T, Sakurai K, Ueguchi C, Mizuno T (2001) Two types of putative nuclear factors that physically interact with histidine-containing phosphotransfer (Hpt) domains, signaling mediators in His-to-Asp phosphorelay, in Arabidopsis thaliana. *Plant Cell Physiol* 42(1):37–45.
36. Baba K, Nakano T, Yamagishi K, Yoshida S (2001) Involvement of a nuclear-encoded basic helix-loop-helix protein in transcription of the light-responsive promoter of psbD. *Plant Physiol* 125(2):595–603.
37. Palatnik JF, et al. (2003) Control of leaf morphogenesis by microRNAs. *Nature* 425(6955):257–263.
38. Schommer C, et al. (2008) Control of jasmonate biosynthesis and senescence by miR319 targets. *PLoS Biol* 6(9):e230.
39. Chen IC, Lee SC, Pan SM, Hsieh HL (2007) *GASA4*, a GA-stimulated gene, participates in light signaling in Arabidopsis. *Plant Sci* 172(6):1062–1071.
40. Till BJ, Zerr T, Comai L, Henikoff S (2006) A protocol for TILLING and Ecotilling in plants and animals. *Nat Protoc* 1(5):2465–2477.
41. Christian M, et al. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 186(2):757–761.
42. Feng Z, et al. (2013) Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res* 23(10):1229–1232.
43. Schmutz J, et al. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463(7278):178–183.
44. Wang X, et al.; *Brassica rapa* Genome Sequencing Project Consortium (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat Genet* 43(10):1035–1039.
45. Reeves R (2004) HMGA proteins: Isolation, biochemical modifications, and nucleosome interactions. *Methods Enzymol* 375:297–322.
46. Dragan AI, Liggins JR, Crane-Robinson C, Privalov PL (2003) The energetics of specific binding of AT-hooks from HMGA1 to target DNA. *J Mol Biol* 327(2):393–411.
47. Panne D, Maniatis T, Harrison SC (2007) An atomic model of the interferon-beta enhanceosome. *Cell* 129(6):1111–1123.
48. Lomvardas S, Thanos D (2002) Modifying gene expression programs by altering core promoter chromatin architecture. *Cell* 110(2):261–271.
49. Martinez Hoyos J, et al. (2004) Identification of the genes up- and down-regulated by the high mobility group A1 (HMGA1) proteins: Tissue specificity of the HMGA1-dependent gene regulation. *Cancer Res* 64(16):5728–5735.
50. Dragan AI, Carrillo R, Gerasimova TI, Privalov PL (2008) Assembling the human IFN-beta enhanceosome in solution. *J Mol Biol* 384(2):335–348.
51. Fonfría-Subirós E, et al. (2012) Crystal structure of a complex of DNA with one AT-hook of HMGA1. *PLoS ONE* 7(5):e37120.

Supporting Information

Zhao et al. 10.1073/pnas.1219277110

SI Materials and Methods

Sequence Alignment and Phylogenetic Analysis. The coding sequences of the 29 *Arabidopsis thaliana AT-HOOK MOTIF CONTAINING NUCLEAR LOCALIZED* (*AHL*) genes were downloaded from The *Arabidopsis* Information Resource (TAIR10). The coding sequence of the plant and prokaryote conserved (PPC)/domain of unknown function #296 (PPC/DUF296) domain from *Escherichia coli* was obtained from National Center for Biotechnology Information (NCBI). The sequences were aligned with the MAFFT program (1). We used MrBayes 3.1.2 to run six independent chains with 10 million generations (2, 3). The General Time Reversible model was used with trees sampled every 1,000 generations (the first 25,000 trees were used as burn-in). After 10 million generations, the average SD of split frequencies was 0.004437. The consensus tree was drawn using MEGA 5.05 (4).

Plant Materials. All *Arabidopsis* lines are in the *Columbia* background. *A. thaliana* transfer DNA (T-DNA) insertion lines were obtained from the *Arabidopsis* Biological Research Center. Each T-DNA line has been back-crossed two times to the WT to eliminate other possible T-DNA insertions and further crossed with the *SUPPRESSOR OF PHYTOCHROME B-4 #3* (*SOB3*/*AHL29*) or *ESCAROLA* (*ESC*/*AHL27*) *sob3-4 esc-8* double null to create a corresponding triple-null line. Homozygous F_3 lines were tested by PCR and used in the study. The quadruple T-DNA insertion line was generated by crossing two triple-null lines.

Antibody Generation. Primary antibody was generated against the synthetic peptide RGN MSG YDQ FAG DPH L, from SOB3/AHL29, using the 2-rabbit-90-d protocol by Open Biosystems.

Seedling Growth Conditions and Hypocotyl Measurement. Surface-sterilized *Arabidopsis* seeds were sown on media containing 1.0% (wt/vol) phytagel (Sigma-Aldrich), 1.5% (wt/vol) sucrose, and 0.5× Linsmaier and Skoog–modified basal medium, incubated at 4 °C in the dark for 5 d, and treated with 24-h red light to synchronize germination at 25 °C. Plates were then incubated for 5 d at 25 °C under 22 $\mu mol \cdot m^{-2} \cdot s^{-1}$ continuous white light in a Percival E-30B growth chamber. Plated seeds were checked after 2 d, and only those with synchronized germination were included in phenotypic analysis. Six-day-old seedlings were then scanned into a TIFF format at 1,200 dpi using an HP ScanJet3500 flatbed scanner (Hewlett-Packard). Hypocotyl lengths were measured using National Institutes of Health (NIH) ImageJ 1.29J (5) and analyzed using Excel software.

Plasmid Construction. The Gateway compatible entry plasmids containing the coding sequences of the *Arabidopsis AHL* genes were obtained from the *Arabidopsis* Biological Research Center. All of the clones have been sequenced to confirm absence of mutations. The coding sequence of *SOB3*/*AHL29* has been cloned and fused into *pENTR*/*D-TOPO* vector via Gateway BP reactions (Invitrogen) as described in Street et al. (6). For the *esc-11* allele, a fragment of the *ESC* coding sequence was amplified by PCR using primers, 5′-CGT CCA CGT GGA CAT CCA CCA GG-3′ and 5′-CGT CTG GAG TGA CCG GTC AGG G-3′. The amplified fragment was ligated into the *pGEM-T Easy* vector (Promega) and confirmed by sequencing to be error free. The *ESC* coding sequence bearing the R77→H mutation was generated by the ligation of PmlI and AgeI digested fragments of the resulting *pGEM-T Easy* plasmid and *pENTR223-Sfi-AT1G20900* (ABRC Stock# G23750). The resulting *esc-11* allele, which bears the R→H mutation, was confirmed by sequencing. To clone the PPC/DUF296 domain from the SOB3/AHL29 protein, primer pairs (5′-CAT GCC ATG GAC CTT AGA TCT CAT GTT CTT GAA G-3′ and 5′-TTG GCG CGC CCT AGA AAG TTG CAT TAG AGA ACG ATG C-3′) were used with the *pENTR-D-TOPO-SOB3* plasmid as template. The fragment amplified was further fused into *pGEM-T Easy* vector. The resulting *pGEM-TEasy-SOB3-PPC* plasmid and *pENTR*/*D-TOPO-SOB3* were digested with NcoI and AscI, and the fragments were ligated together to generate the final *pENTR*/*D-TOPO-SOB3-PPC* plasmid. To generate the *SOB3-ΔG* allele, the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) was used to generate the G-R-F-E-I-L deletion in the *pENTR-D*/*TOPO-SOB3* plasmid using primer pair 5′-GGA GGT GTT GTG GCT CTA CAT TCC CTC ACA GG-3 and 5′-CCT GTG AGG GAA TGT AGA GCC ACA ACA CCT CC-3′. The pENTR-D/TOPO-SOB3-GTA was generated using primer pair 5′-GAA CCG GAG GTG TTG TGG CTC TAC ATG CAG CGG CTG CGG CAG CTT CCC TCA CAG GTA CGG TGT TGC CGC C-3′ and 5′-GGC GGC AAC ACC GTA CCT GTG AGG GAA GCT GCC GCA GCC GCT GCA TGT AGA GCC ACA ACA CCT CCG GTT C-3′. The resulting plasmids were sequenced to confirm absence of errors.

Protein Expression and Purification. The plasmid *pMAL2-c2x* was a generous gift from Douglas Cole (University of Idaho, Moscow, ID). The coding sequence of *SOB3*/*AHL29* was amplified by the primer pair (5′-ACG CGT CGA CGA CGG TGG TTA CGA TCA ATC CGG AG-3′ and 5′-AAC TGC AGC TAA AAG GCT GGT CTT GGT GGT G-3′) and cloned into the SalI and PstI restriction sites of the *pMAL2-c2x* plasmid to generate the final plasmid. Recombinant maltose binding protein (MBP)-tagged SOB3 was extracted from transformed *E. coli* (*Rosetta2 DE3*) after 6 h of incubation at 30 °C following induction with 1 μM isopropyl β-D-1-thiogalactopyranoside. The recombinant MBP-SOB3 protein was purified using amylose resin (New England Biolabs).

EMSA. A 39-bp oligonucleotide in the pea *PRA2* promoter region was used as a probe. The forward (5′-TAA CAC ATA TTT TGA TAA ATT TAT TAC TAA AAC TAT TTT-3′) and reverse-complement (5′-AAA ATA GTT TTA GTA ATA AAT TTA TCA AAA TAT GTG TTA-3′) forms of the probe were ordered from Sigma-Aldrich with 5′-biotin modification. The EMSA probe was generated by annealing the two oligonucleotides at 95 °C for 5 min followed by a 1 °C decrease per minute to 25 °C. The cold (unlabeled) form of the probe was also generated in the same way. EMSA was performed using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific). Seven micrograms of MBP-tagged protein was incubated in a 20-μL binding reaction [50 ng/μL poly(dI-dC), 10% (vol/vol) glycerol, 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM EDTA, 10 mM Tris, and 1 mM DTT, pH 7.5] for 20 min at room temperature. The EMSA reaction system was then resolved by using the 5% Mini-PROTEAN TBE precast gel (Bio-Rad). The chemiluminsecence signal was detected by using the chemiluminescent nucleic acid detection module (Thermo Scientific).

Bimolecular Fluorescence Complementation and Bimolecular Fluorescence Complementaion-FRET Assay. *pSAT BiFC* plasmids were obtained from ABRC. Coding sequences of *SOB3*, *ESC,* and *HRC* were cloned into the bimolecular fluorescence complementation (BiFC)

vectors (*pSAT4-DEST-nEYFP-C1* and *pSAT5-DEST-cEYFP-C1*) to be expressed in-frame with either the N- or C-terminal half of YFP. Empty vectors and other nuclear localized proteins were used as negative controls. Pairs of BiFC plasmids together with the *pSAT6-mRFP* plasmid encoding a red fluorescent protein (300 ng each) were cobombarded into onion epidermal cells using a PDS-1000/He Biolistic transformation system (Bio-Rad). Equivalent amounts of *pSAT4-cyan fluorescent protein* (*CFP*)-*SOB3-PPC* or *pSAT4-CFP-SOB3-ΔG* were used to transform into onion epidermal cells for localization analysis. For BiFC assay, reconstructed fluorescence was examined after 40 h of incubation in the dark at room temperature by a Zeiss LSM 510 META confocal microscope. Observation of monomeric red fluorescence was used to identify successfully transformed onion cells and successful expression of delivered plasmids. BiFC-FRET assay was performed according to the Leica FRET AB Wizard by the Leica TCS SP5 confocal microscope. Transformed onion epidermal cells with a medium red fluorescent intensity were used for analysis.

Agrobacterium-Mediated Plant Transformation (Floral Dipping Method). Coding sequences were fused 3′ of the *CaMV 35S* promoter via Gateway LR reactions (Invitrogen) into the binary vector *pED15* (7). Transgenic *Arabidopsis* plants expressing these clones were generated using the floral dipping method (8). Transgenic seeds were screened using media containing 1.0% (wt/vol) phytagel (Sigma-Aldrich), 1.5% (wt/vol) sucrose, and 0.5× Linsmaier and Skoog–modified basal medium supplemented with 30 ng/mL Basta (glufosinate-ammonium) under 30 $\mu mol \cdot s^{-1} \cdot m^{-2}$ white light at 25 °C in a Percival E-30B growth chamber.

Yeast Two-Hybrid Assay and Library Screen. A GAL4-based yeast two-hybrid (Y2H) system was used to test protein-protein interactions in yeast. The reporter yeast strain *L40ccU3,* bait vector (*pBTM116-GW-D9*) with *TRP1* reporter marker, and prey vector (*pACT2-GW*) with *LEU* reporter marker were obtained from Hanjo Hellmann's laboratory (Washington State University, Pullman, WA) and used for Y2H analysis. For targeted Y2H assays, transformed yeast cells were first plated on synthetic dropout II (SDII) media, which lacked tryptophan and leucine, to identify transformants. Five individual colonies were picked and replated simultaneously on both SDII, as well as synthetic dropout IV (SDIV) media, which lacked tryptophan, leucine, histidine, and uracil and contained the predetermined level of 3-amino-1, 2, 4-triazol (3-AT) required to suppress autoactivation for protein-protein interaction analysis. Yeast were incubated at 30 °C in a growth chamber and checked for growth after 3–7 d. For the Y2H library screen, a root tissue–derived cDNA library, cloned into the *pACT2-GW* vector, was used. Yeast containing *pBTM-GW-D9-AHLs* were transformed with the cDNA library in *pACT2-GW* and plated on SDIII media, which lacked tryptophan, leucine, and histidine, and contained 3-AT. After incubation for 4 d at 30 °C, colonies were transferred to SDIV media containing 3-AT. After incubation for 3 d at 30 °C, a filter lift X-gal staining assay was used to confirm interactions. The cDNA portion of the *pACT2-cDNA* plasmids were amplified from individual blue colonies via colony PCR using MATCHMAKER Insert Screening Primers (Clontech). PCR products were purified via gel electrophoresis and sequenced.

Coexpression Analysis of the *AHL* Gene Family in *A. thaliana*. Genes that are coexpressed with each of the 29 *AHL* genes in *A. thaliana* were examined using the ATTED-II v6.0 server (9, 10). The coexpression networks of the *AHL* genes were constructed based on the coexpression coefficients of each coexpressed pair.

Secondary Structure Prediction of the PPC/DUF296 Domains of the AHL Proteins in *A. thaliana*. The amino acid sequences of AHL proteins were downloaded from TAIR10 (www.arabidopsis.org). Amino acid sequences of the PPC/DUF296 domain in each AtAHL protein were identified. The secondary structure of each identified sequence was examined with the PSIPRED Protein Structure Prediction Server (PSIPRED v3.0) (11, 12).

Tertiary Structure Prediction of SOB3 PPC/DUF296 Domain and Quality Estimation. The tertiary structure of the SOB3/AHL29 PPC/DUF296 domain was predicted using three different programs: CPHmodel3.0 (13), ESyPred3D (14), and SWISS_MODEL server (15). The quality of resulting predictions was then evaluated using the QMEAN Server (16–18).

T-DNA Insertions. Genomic DNA was extracted from 10-d-old seedlings. In addition to the universal T-DNA *LBb1-3* primer (5′-ATT TTG CCG ATT TCG GAA C-3′), T-DNA line-specific primers were used to identify specific T-DNA insertions (SALK_018866, 5′-GAA AAG ACT CGT GAC TGT TTG C-3′ and 5′-AAG ACG GTG ACA CTT TCC ATG-3′; SALK_037803c, 5′-GAG GAA GAC CGA GGA AGT ACG-3′, and 5′-GAT CGA ACC AGT TGC AGA AAG-3′; SALK_123590c, 5′-AGA GCC ATG ACA TTG ACT TG-3′ and 5′-TCT GGT GAA TGA AGA AGG TGG-3′; SALK_040729c, 5′-CCC TGT TGG GTC TAT CTT TCT C-3′ and 5′-CAA CTT GAC CTT GAG CTC CAG-3′; *ataf2-1*, SALK_136355c, 5′-GTA TAT ATC CAG CTG GCT GCG-3′ and 5′-AAA AGC TTC CAA AGC ATC CTC-3′).

RNA Extraction, RT-PCR, and Quantitative Real-Time PCR. Total RNA was extracted from the 6-d-old *ahl* triple-null lines that were used in the hypocotyl-length measurement experiment, using the Qiagen RNeasy Mini Kit (Qiagen). cDNAs were further generated using the RNA-to-cDNA EcoDry Premix (Oligo dT) kit (Clontech). Identical amounts of cDNA input were used in PCR reactions using either a tubulin primer pair (5′-GAA GTC GAC CGT GAG ATT CTT CAC ATC CAG GGT GGT C-3′ and 5′-CGC GGA TCC CAT AGT AGC AGA AAT CAA GTG GTT CAA A-3′) or *AHL*-specific primers. For the *sob3-4 esc-8 ahl5* (DN123590) line, the primer pair of 5′-TCT GTG AAT TCG ATG GGA GAG AAG CTA TGG CA-3′ and 5′-TAG AAG CTT TCA CCC GCG AGT CAA GTC-3′ was used. For the *sob3-4 esc-8 ahl6* (DN037803) line, the primer pair of 5′-GAG GAA GAC CGA GGA AGT ACG-3′ and 5′-GAT CGA ACC AGT TGC AGA AAG-3′ was used. For the *sob3-4 esc-8 ahl15* (DN040729) line, the primer pair of 5′-TCA GCT CCT TCT TTG CAC CAC-3′ and 5′-ATA CGA AGG AGG AGC ACG AGG-3′ was used. For the *sob3-4 esc-8 ahl22* (DN018866) line, the primer pair of 5′-CCA CAC CAT CAA TTC CAG CAT-3′ and 5′-CAT AAG ATT CGG AGG CAA CCC-3′ was used. The transcript accumulation levels of sob3-6ox, SOB3-ΔGox, and SOB3-GTAox lines were examined by primer pairs, 5′-CTT AGA TCT CAT GTT CTT GAA G-3′ and 5′-GAA AGT TGC ATT AGA GAA CGA TGC-3′.

RNA was extracted from yeast used for Y2H analysis, which had been growing on SDII media for ∼1.5 d at 30 °C. Using the MasterPure Yeast RNA Purification Kit (Epicentre), total RNA was isolated from yeast transformed with two plasmids: a prey plasmid (*pACT2-GW*) containing the *TCP4* gene and a bait plasmid (*pBTM116-GW-D9*) containing either the WT *ESC* gene or one of two versions of the *ESC* gene containing modifications to the conserved G-R-F-E-I-L domain (ESC-ΔG or ESC-GTA). Subsequently, cDNA was generated using the iScript cDNA Synthesis Kit (Bio-Rad). Equal quantities of cDNA input were added in PCR reactions using either a control primer pair, 5′-TCG GTG GTT CCA TTC TTG CT-3′ and 5′-GCT TTT TAA GCC TTT GAT CTT GAG AG-3′) or primers specific to the *ESC* gene (5′-CCG G AAT TC T ATG GAA GGC GGT TAC GAG CAA-3′ and 5′-CGC GGATCC TTA AAA AGG TGG TCT TGA AGG-3′).

For quantitative real-time PCR (real-time qPCR), total RNA was extracted from 6-d-old seedlings of corresponding genotypes. cDNA were further generated by using the RNA-to-cDNA EcoDry Premix (Oligo dT) kit (Clontech). Real-time qPCR were performed using SsoFast EvaGreen Supermix with Low ROX (Bio-RAD) on the 7500 Fast Real-time PCR System (Applied Biosystems) with the comparative $\Delta\Delta C_T$ method. Sequences of the real-time qPCR primers are as follows—for *SOB3*/*sob3-6*: 5′-AAA TGG TGG AAC CGG AGG TGT TG-3′ and 5′-AAC ACC GTA CCT GTG AGG GAA AG-3′. For *TCP4*, 5′-ACC GAT ACA GGA AAC GGA GGA-3′ and 5′-AGG ATC AAA CCA AGC ACG GAT-3′. For *UBQ10*, 5′-TCT TCG TGG TGG TTT CTA AAT CTC G-3′ and 5′-AAA GAG ATA ACA GGA ACG GAA ACA TAG T-3′.

1. Katoh K, Asimenos G, Toh H (2009) Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol* 537:39–64.
2. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8):754–755.
3. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572–1574.
4. Tamura K, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731–2739.
5. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.
6. Street IH, Shah PK, Smith AM, Avery N, Neff MM (2008) The AT-hook-containing proteins SOB3/AHL29 and ESC/AHL27 are negative modulators of hypocotyl growth in *Arabidopsis*. *Plant J* 54(1):1–14.
7. Turk EM, et al. (2005) BAS1 and SOB7 act redundantly to modulate Arabidopsis photomorphogenesis via unique brassinosteroid inactivation mechanisms. *Plant J* 42(1):23–34.
8. Clough SJ, Bent AF (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* 16(6):735–743.
9. Obayashi T, Hayashi S, Saeki M, Ohta H, Kinoshita K (2009) ATTED-II provides coexpressed gene networks for Arabidopsis. *Nucleic Acids Res* 37(Database issue):D987–D991.
10. Obayashi T, Nishida K, Kasahara K, Kinoshita K (2011) ATTED-II updates: Condition-specific gene coexpression to extend coexpression analyses and applications to a broad range of flowering plants. *Plant Cell Physiol* 52(2):213–219.
11. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292(2):195–202.
12. Buchan DW, et al. (2010) Protein annotation and modelling servers at University College London. *Nucleic Acids Res* 38(Web Server issue):W563-8.
13. Nielsen M, Lundegaard C, Lund O, Petersen TN (2010) CPHmodels-3.0–remote homology modeling using structure-guided sequence profiles. *Nucleic Acids Res* 38(Suppl 2):W576–W581.
14. Lambert C, Léonard N, De Bolle X, Depiereux E (2002) ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* 18(9):1250–1256.
15. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* 22(2):195–201.
16. Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27(3):343–350.
17. Benkert P, Künzli M, Schwede T (2009) QMEAN server for protein model quality estimation. *Nucleic Acids Res* 37(Web Server issue):W510-4.
18. Benkert P, Tosatto SC, Schomburg D (2008) QMEAN: A comprehensive scoring function for model quality assessment. *Proteins* 71(1):261–277.
19. Lin LY, et al. (2007) Crystal structure of *Pyrococcus horikoshii* PPC protein at 1.60 A resolution. *Proteins* 67(2):505–507.
20. The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC.

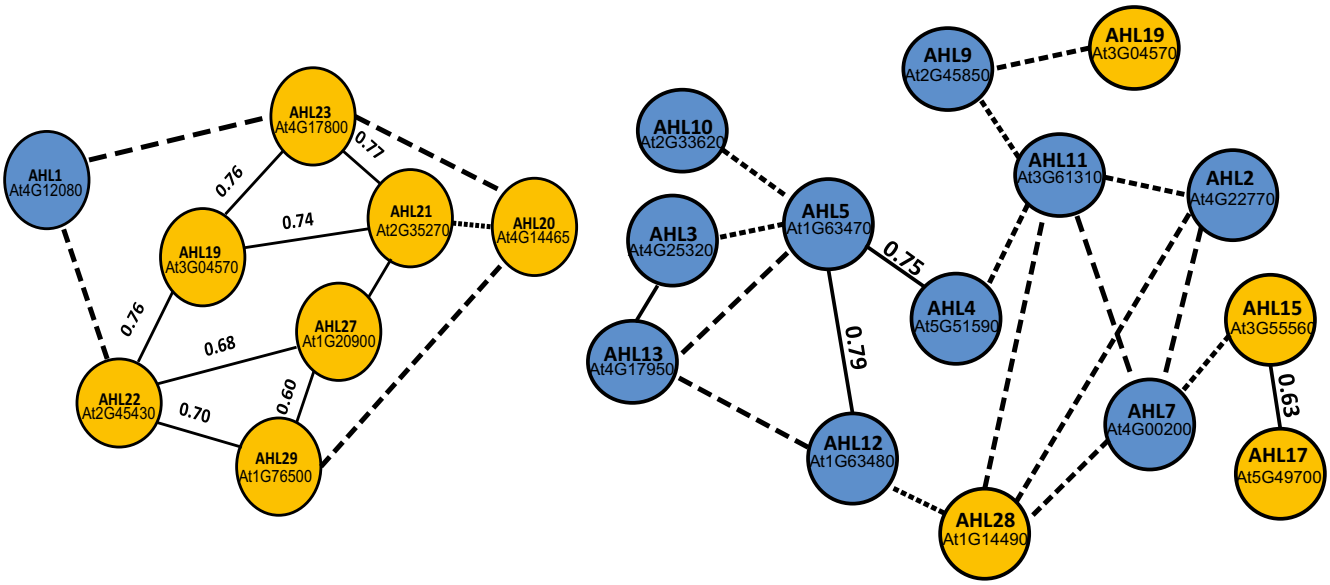

Fig. S1. Coexpression analysis of the *AHL* gene family in *A. thaliana* using the ATTED-II (v6.0) server (9, 10). For each pair of *AHL* genes within a coexpression network, the correlation coefficient of coexpression is labeled. Solid lines between pairs of *AHL* genes indicate direct coexpression relationship between these two *AHL*s. Dotted lines between pairs of *AHL* genes indicate indirect coexpression relationships with at least one non-*AHL* gene that bridges the direct co-expression relationship between these two *AHL*s. Orange, clade A *AHL*s; blue, clade B *AHL*s.

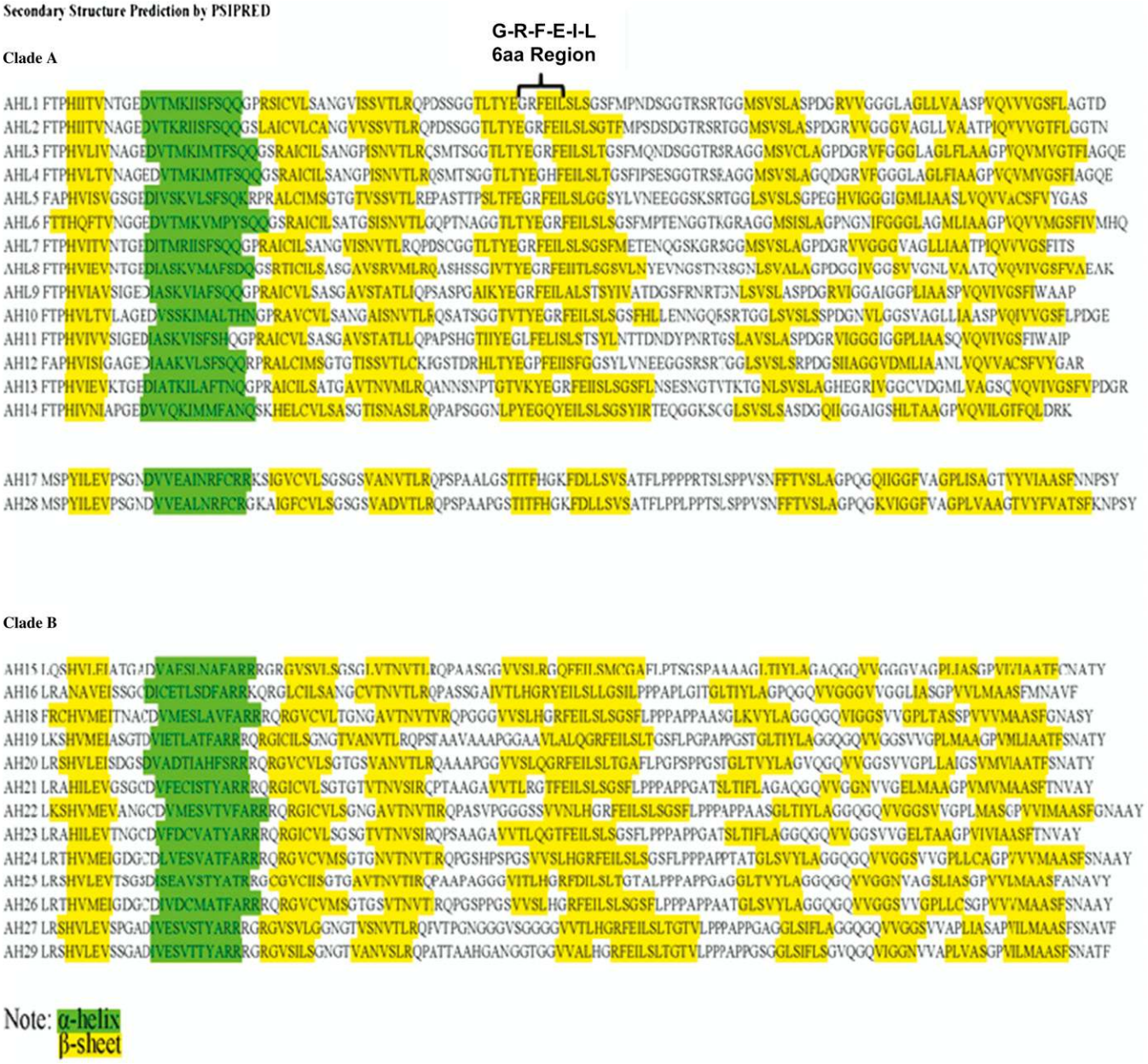

Fig. S2. Prediction of secondary structure of AHL PPC/DUF296 domains in *A. thaliana*. The amino acid sequences were downloaded from TAIR. The secondary structure was predicted using the PSIPRED server (12). The six conserved amino acid region is shown in AHL1.

A

```
                1                                                  50
AHL1_PPC        FTPHIITVNTGEDVTMKIISFSQQGPRSICVLSANGVISSVTLRQPDSS-
AHL5_PPC        FAPHVISVGSGEDIVSKVLSFSQKRPRALCIMSGTGTVSSVTLREPAST-
AHL6_PPC        FTTHQFTVNGGEDVTMKVMPYSQQGSRAICILSATGSISNVTLGQPTNA-
AHL9_PPC        FTPHVIAVSIGEDIASKVIAFSQQGPRAICVLSASGAVSTATLIQPSAS-
ESC_PPC         LRSHVLEVSPGADIVESVSTYARRRGRGVSVLGGNGTVSNVTLRQPVTPG
SOB3_PPC        LRSHVLEVSSGADIVESVTTYARRRGRGVSILSGNGTVANVSLRQPATTA
                51                                                 100
AHL1_PPC        --------GGTLTYEGRFEILSLSGSFMPNDSGGTRSRTGGMSVSLASPD
AHL5_PPC        --------TPSLTFEGRFEILSLGGSYLVNEEGGSKSRTGGLSVSLSGPE
AHL6_PPC        --------GGTLTYEGRFEILSLSGSFMPTENGGTKGRAGGMSISLAGPN
AHL9_PPC        --------PGAIKYEGRFEILALSTSYIVATDGSFRNRTGNLSVSLASPD
ESC_PPC         NGGGVSGGGGVVTLHGRFEILSLTGTVLPPP---APPGAGGLSIFLAGGQ
SOB3_PPC        AHGANGGTGGVVALHGRFEILSLTGTVLPPP---APPGSGGLSIFLSGVQ
                101
AHL1_PPC        GRVVGGGLAGLLVAASPVQVVVGSFLAGTD
AHL5_PPC        GHVIGGGIG-MLIAASLVQVVACSFVYGAS
AHL6_PPC        GNIFGGGLAGMLIAAGPVQVVMGSFIVMHQ
AHL9_PPC        GRVIGGAIGGPLIAASPVQVIVGSFIWAAP
ESC_PPC         GQVVGGSVVAPLIASAPVILMAASFSNAVF
SOB3_PPC        GQVIGGNVVAPLVASGPVILMAASFSNATF
```

B

	QMEAN Score	Z-score
CPHmodel_SOB3_PPC.pdb	**0.635**	**-1.12**
EsyPred3D_SOB3_PPC.pdb	**0.518**	**-2.42**
SWISS_MODEL_SOB3_PPC.pdb	**0.587**	**-1.68**

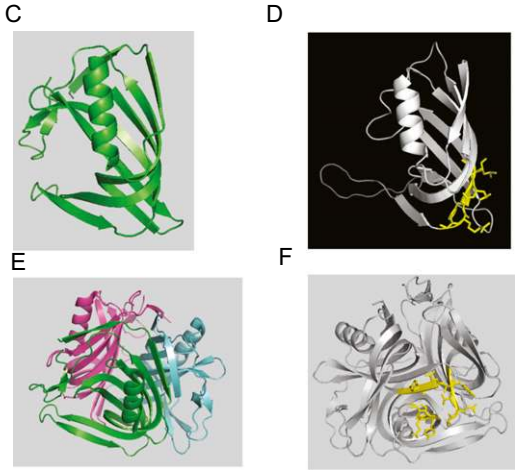

Fig. S3. Predicted structure of the PPC/DUF296 domain of SOB3/AHL29. (*A*) Alignment of amino acid sequences from the PPC/DUF296 domains of selected AtAHL proteins. The Gly-Arg-Phe-Glu-Ile-Leu region is underlined and in red. (*B*) Evaluation of the quality of the predicted tertiary structure of the SOB3 PPC/DUF296 domain by the QMEAN server (16–18). QMEAN score and the estimated absolute quality Z-score are shown. (*C*) Predicted monomer structure of SOB3/AHL29 PPC/DUF296 domain. (*D*) The monomer structure of PhPPC/DUF296 protein from the PDB file (2DT4) (19) with the conserved Gly-Arg-Phe-Glu-Ile-Leu shown in stick form in yellow. (*E* and *F*) Predicted trimer formation of PPC/DUF296 domain of SOB3/AHL29. The model shown was generated by CPHmodels 3.0 server (13). The figure was prepared using PyMol version 1.3 (20).

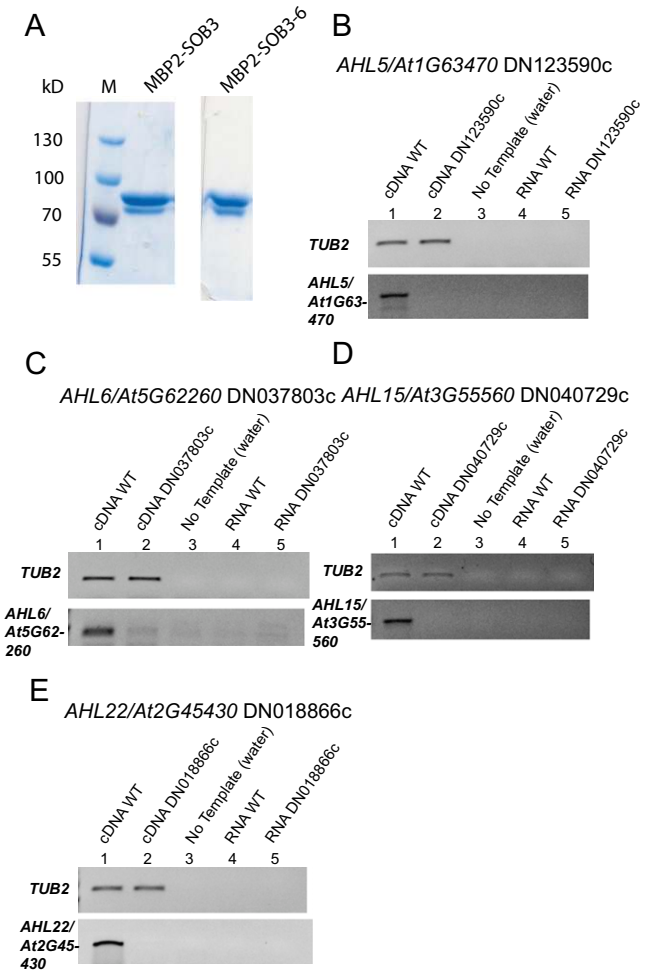

Fig. S4. Protein loading in the EMSA assay and examination of transcript accumulation of *AHL* genes in the triple-null lines. (*A*) The same amounts of recombinant protein used in the EMSA assay were examined by SDS/PAGE followed by Coomassie blue staining. (*B–E*) PCR amplification using RNA as template are also shown to demonstrate absence of genomic DNA contamination. *TUB2*, tublin 2. *DN*, *sob3-4 esc-8* double-null background.

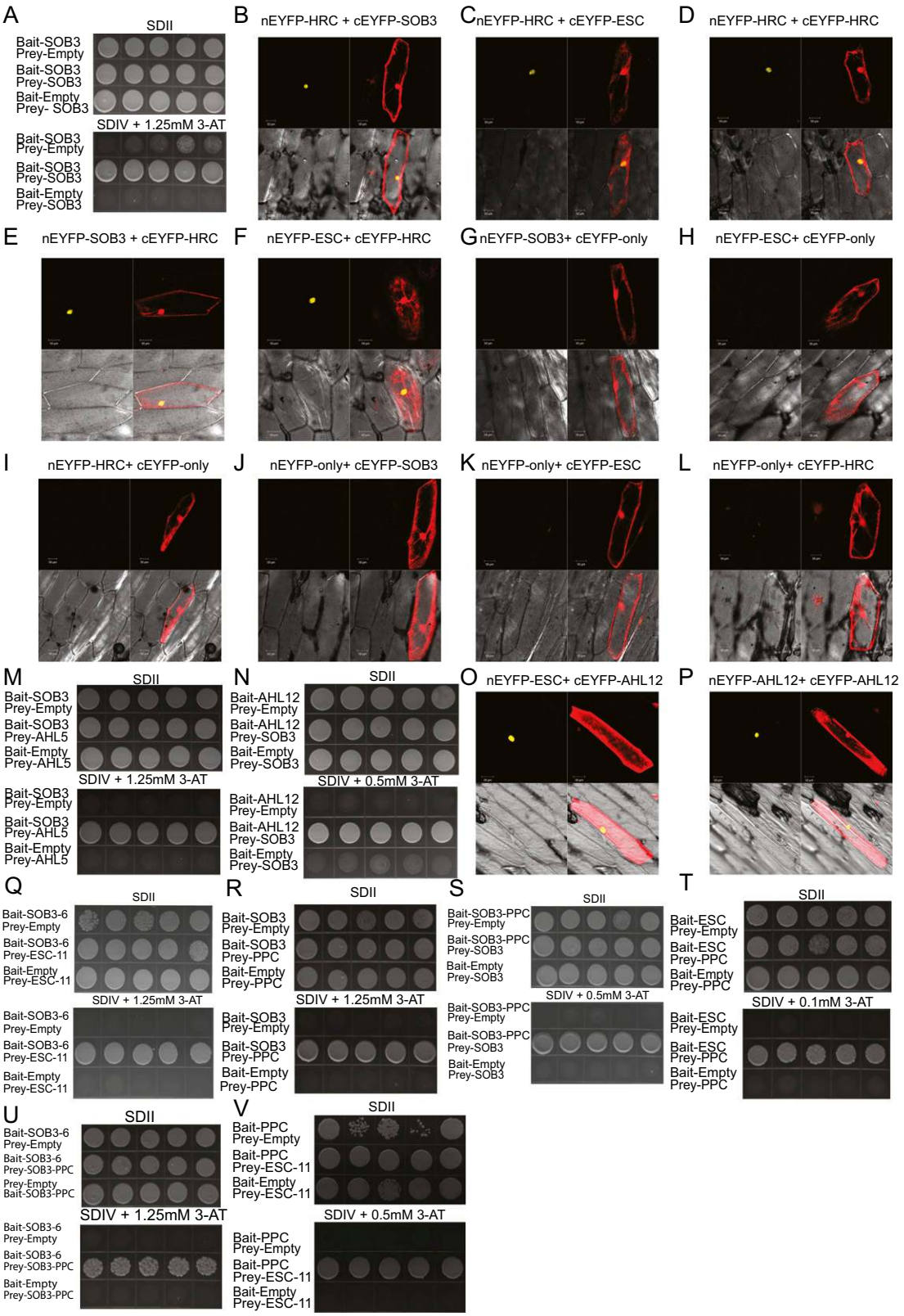

Fig. S5. AHL proteins interact with each other. (*A*) SOB3/AHL29 interacts with itself in the yeast two-hybrid assay. Five independent biological replicates are shown. (*B–F*) SOB3/AHL29, ESC/AHL27, and HRC/AHL25 interacted in the BiFC assay. Negative controls using empty vector are shown in *G–L*. SOB3 interacted with AHL5 (*M*) and AHL12 (*N*) in Y2H assays. AHL12 interacted with ESC (*O*) and with itself (*P*) in BiFC assays. (*Q*) SOB3-6 and ESC-11 interact with each other. The PPC/DUF296 domain of SOB3/AHL29 interacted with SOB3/AHL29 (*R* and *S*), ESC/AHL27 (*T*), SOB3-6 (*U*), and the mutant ESC-11 (*V*) proteins in yeast two-hybrid assays. Five independent biological replicates are shown.

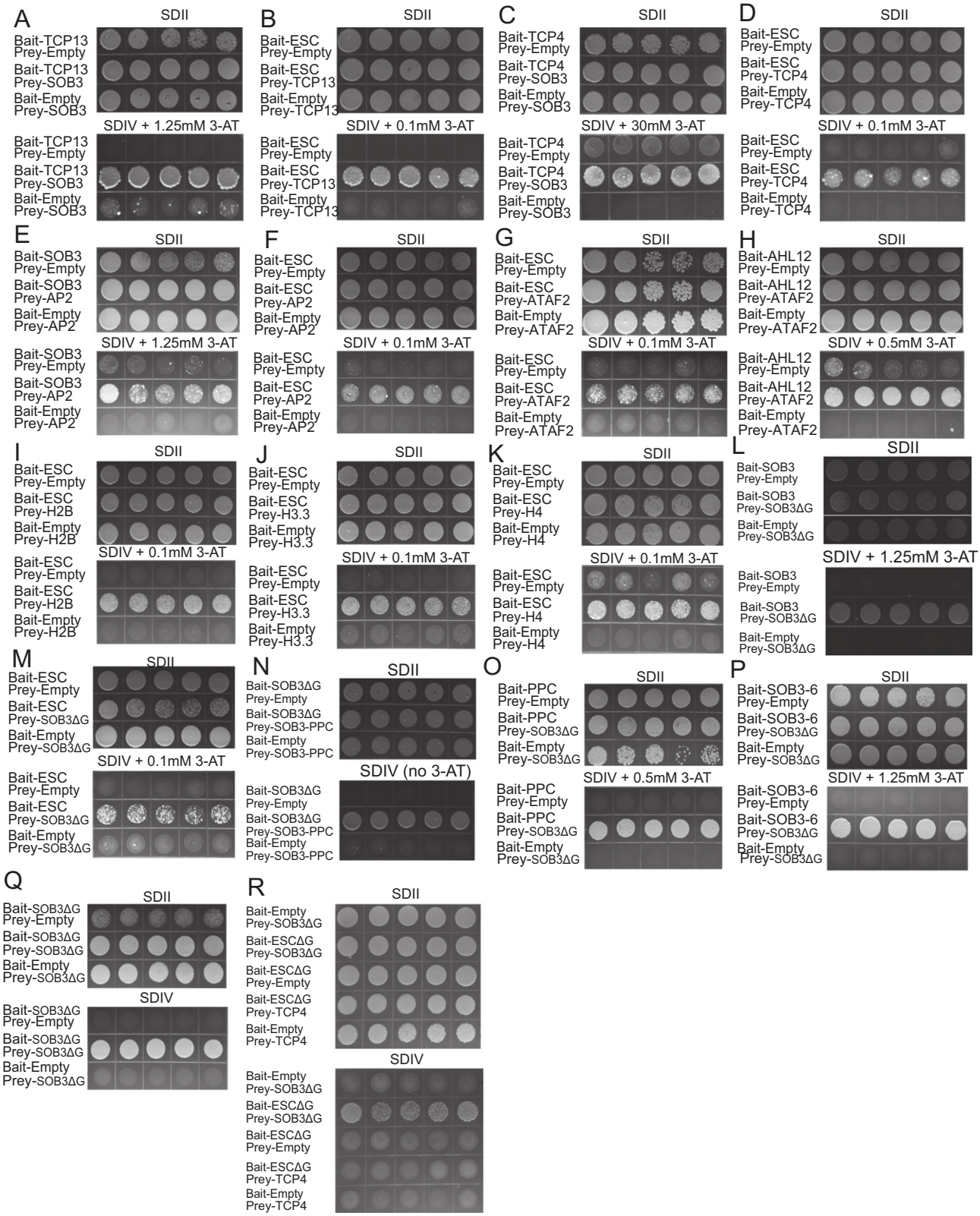

Fig. S6. Physical interaction between AHLs and non-AHL proteins. AHL proteins interact with transcription factors (*A*–*H*) and histone proteins (*I*–*K*) in yeast two-hybrid assays. Removal of the conserved six-amino-acid sequence in the PPC/DUF296 domain of SOB3/AHL29 did not affect its physical interaction with AHL proteins (*L*–*R*).

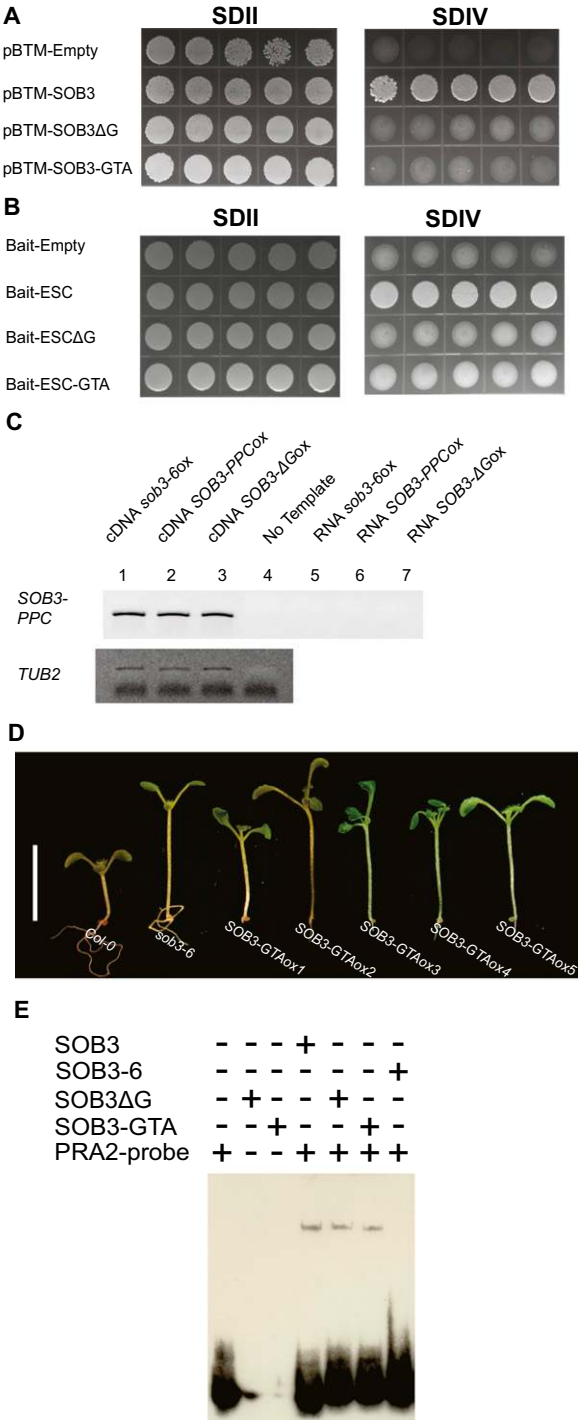

Fig. S7. Disturbance of the conserved six-amino-acid region disrupts the transcriptional activity of AHLs. (*A* and *B*) Yeast transformed with each indicated plasmids were grown on SDII and SDIV media. Five biological replicates are shown. (*C*) Transcript accumulation examined in lines overexpressing *sob3-6*, *SOB3-PPC,* and *SOB3-ΔG*. (*D*) Overexpression of *SOB3-GTA* allele leads to long-hypocotyl phenotype in the light. Multiple independent T3 generation seedlings are shown. Seedlings were grown in 20 $\mu mol \cdot m^{-2} \cdot s^{-1}$ of white light at 25 °C. (Bar, 1 cm.) (*E*) Deletion or alanine substitution of the six conserved amino acid region does not abolish the AT-rich DNA binding capacity of SOB3/AHL29 protein.

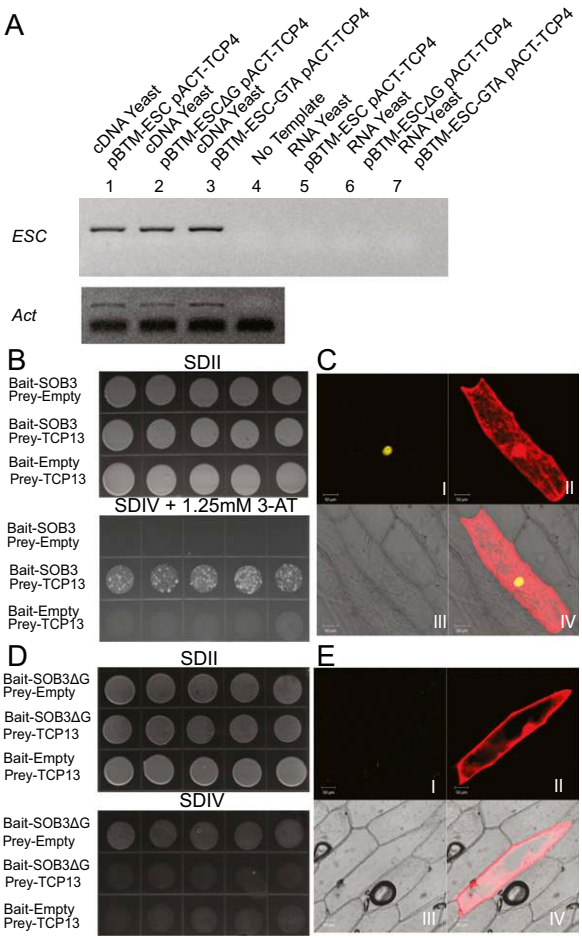

Fig. S8. SOB3/AHL29 interacted with non-AHL proteins. (*A*) Levels of *ESC* transcript accumulation were examined in yeast used in an Y2H assay. The transcription factor TCP13 interacts with SOB3 (*B* and *C*), and this interaction was abolished when the conserved six amino acids in PPC/DUF296 domain were removed (*D* and *E*).

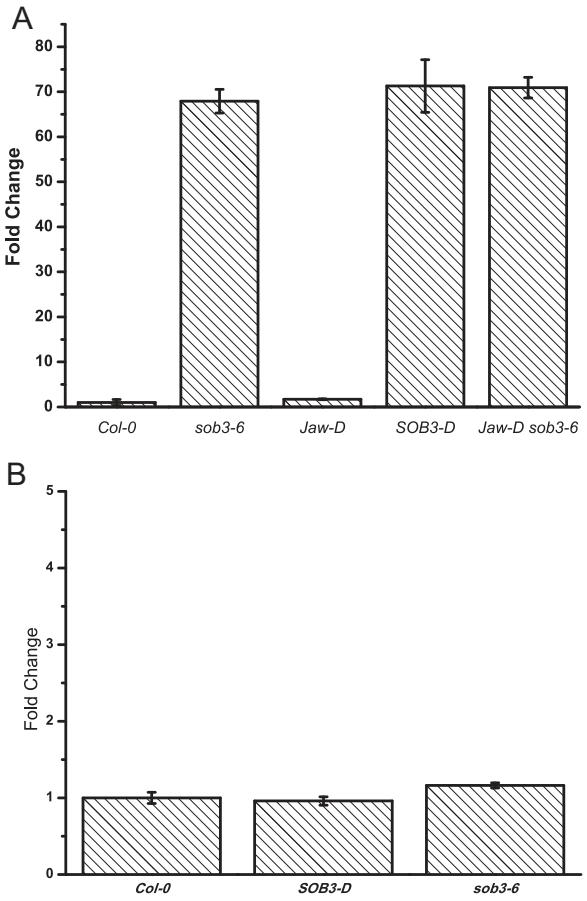

Fig. S9. Examination of transcript accumulation in 6-d-old seedlings grown in the light by quantitative real-time PCR. (*A*) *SOB3* (or *sob3-6*) transcript accumulation in different genotypes. (*B*) *TCP4* transcript accumulation in different genotypes. Fold changes were normalized with internal control *UBQ10* and compared with WT *Col-0*.

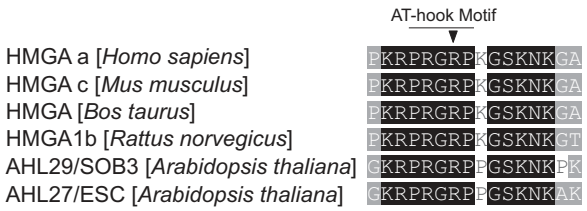

Fig. S10. Alignment of the AT-hook motifs in mammalian high mobility group A (HMGA) proteins and the type 1 AT-hook motif in the AHL proteins. The second AT-hook motifs from HMGA proteins and the ones from SOB3/AHL29 and ESC/AHL27 are used. The arrowhead points out the arginine that was mutated in the *sob3-6* allele (*R→H*).

Table S1. Interactors identified in yeast two-hybrid library screen and targeted yeast two-hybrid assay

Bait	Annotation of interactor	Interactor AGI number
SOB3/AHL29	40S ribosomal protein S17	AT2G05220
	Lipid transfer protein (PR-14)	AT5G01870
	Glyceraldehyde-3-phosphate dehydrogenase	AT3G04120
	40S ribosomal protein S20	AT5G62300
	Glycine-rich protein	AT2G05510
	60S ribosomal protein L30	AT3G18740
	Ribosomal protein S6	AT4G31700
	Arabidopsis thaliana drought-induced 21	AT4G15910
	40S ribosomal protein S14	AT3G11510
	Cytosolic ribosomal protein RPS15	AT1G04270
	Ribosomal protein 26Se	AT3G56340
	SEC14 cytosolic family protein	AT4G35750
	Ribosomal L38e protein family	AT2G43460
	Unknown protein	AT1G11240
	AXR4	AT1G54990
	Lipid transfer protein 2	AT2G38530
	RUB2	AT2G35635
	KIWI transcriptional activator	AT5G09250
	Unknown protein	AT5G42110
	Ribosomal protein L34e	AT1G26880
	Dynein light chain type 1	AT1G23220
	Maternal effect embryo arrest 60	AT5G05950
	Unknown protein	AT5G52420
	Zinc finger family protein	AT2G36930
	Protein disulfide isomerase 7	AT4G27080
	DNA-binding enhancer protein-related, F5E6.6	AT3G06610
	Unknown protein	AT5G47920
	IAA carboxy-methyltransferase	AT5G55250
	Polyamine oxidase	AT2G43020
	ORS1	AT2G41230
	Light-regulated WD1 (LWD1)	AT1G12910
	HOP3	AT4G12400
	Ribosomal protein L7Ae/L30e/S12e	AT5G20160
	Ribosomal protein L6 family protein	AT1G74060
	SAUR-like auxin-responsive protein	AT1G75580
	PPPDE	AT3G07090
	Aldolase superfamily protein	AT3G52930
	SecY transport family protein	AT2G34250
	Integrase type DNA-binding superfamily	AT1G22190
	Ferredoxin hydrogenase	AT4G16440
	Protein kinase superfamily protein	AT4G17660
PPC/DUF296 of SOB3/AHL29 only	Unknown protein	AT2G38370
	SOB3	AT1G76500
	Leucine-rich repeat protein	AT5G19680
	F-Box family protein	AT4G33160
	AHL3	AT4G25320
	PHD zinc finger domain protein	AT1G09520
	Phosphoglycerate/biphosphoglycerate mutase	AT1G78050
SOB3ΔG	B12D protein	AT3G29970
	Histone H3	AT5G10980
	Actin depolymerizing factor 3	AT5G59880
ESC/AHL27	60S ribosomal protein L22-2	AT3G05560
	Unknown protein	AT5G54970
	AP2-domain containing transcription factor	AT5G61890
	AHL20	AT4G14465
	Putative peroxidase	AT5G39580
	Unknown protein	AT5G58990
	AHL23	AT4G17800
	RSH2, RELA-SPOT HOMOLOG2	AT3G14050
	GASA4	AT5G15230
	GLY2	AT3G10850
	SADHU3-2	AT3G42658
	WLIM1	AT1G10200

Table S1. Cont.

Bait	Annotation of interactor	Interactor AGI number
	Histone H3	AT5G10980
AHL12	Glutamine amidotransferase type 1 family protein	AT5G57890
	Ribosomal protein S21e	AT3G53890
	Ribosomal protein L2 family	AT2G18020
	adenosyl-L-methionine dependent methyltransferase	AT5G10830
	Splicing factor 3B subunit 5/RDS3 complex subunit 10	AT4G14342
	Histone H4	AT3G53730
	EIF3G1 eukaryotic translation initiation factor	AT3G11400
	HVA22D	AT4G24960
	UDP apiose/UDP xylose synthase I	AT2G27860
	Profilin 2	AT4G29350
	40S ribosomal protein SA	AT1G72370
	Peroxidase superfamily protein	AT4G21960
	Unknown protein	AT1G68680
	ATAF2 transcription factor	AT5G08790
	Uridine kinase-like 2	AT3G27190
AHL1		
	AHL8	AT5G46640
	Histone H3	AT5G10980
	Histone H3	AT4G40040
AHL5		
	AHL3	AT4G25320
Additional targeted interactions tested		
SOB3/AHL29	Histone H2B (bait-SOB3, prey-histone H2B)	AT1G07790
	AP2 transcription factor (bait-SOB3, prey-AP2)	AT5G61890
	TCP4 transcription factor (bait-TCP4, prey-SOB3)	AT3G15030
	TCP13 transcription factor (both configurations)	AT3G02150
	AHL12 (bait-AHL12, prey-SOB3)	
	SOB3/AHL29	
	ESC/AHL27 (both configurations)	
	AHL5 (bait-SOB3, prey-AHL5)	
	TCP14 transcription factor (both configurations)	AT3G47620
ESC/AHL27	ESC/AHL27	
	TCP4 transcription factor (bait-ESC, prey-TCP4)	AT3G15030
	TCP13 transcription factor (bait-ESC, prey-TCP13)	AT3G02150
	Histone H4 (bait-ESC, prey-histone H4)	AT3G53730
	Histone H2B (bait-ESC, prey-histone H2B)	AT1G07790
	TCP14 transcription factor (both configurations)	AT3G47620
AHL1	AHL12 (bait-AHL1, prey-AHL12)	
AHL5	AHL12 (bait-AHL5, prey-AHL12)	