

## SUPPORTING INFORMATION

### DNA-binding specificities of plant transcription factors and its potential to define target genes

José M. Franco-Zorrilla, Irene López-Vidriero, José L. Carrasco, Marta Godoy, Pablo Vera and Roberto Solano

### CONTENTS

1. Text S1.....	2
2. Text S2. Material and Methods.....	18
3. Fig. S1 to S36.....	21
4. Table S1.....	57
5. Table S2.....	60
6. Table S3.....	70
7. Table S4.....	71
8. Table S5.....	73
9. References.....	74

## 1. Text S1

### 1. MYB superfamily

The MYB superfamily is one of the largest and functionally diverse families of sequence-specific TFs in Arabidopsis (1, 2). This family is characterized by the presence of up to four repeats of a 52 residues-long domain, the MYB domain. Each conserved MYB domain forms three  $\alpha$ -helices, where the second and third helices adopt a helix–turn–helix (HTH) structure responsible for binding to DNA (2). MYB proteins are classified in relation to their number of conserved adjacent repeats (2).

#### 1.1. MYB R2R3

This family encompasses about 100 members that can be organized in 25 subgroups attending to their similarity in their DNA-binding and C-terminal domains (2, 3). R2R3 MYBs bind directly to DNA through either of their cognate targets: MBSI (CNGTTR), MBSII (GKTWGTTTR), and MBSIIG (GKTWGCGTR; R, A or G; K, G or T; W, A or T; (2, 4). In spite that phylogenetically related proteins show similar binding specificities, different R2R3 proteins can bind to different DNA target motifs (4). We analyzed binding specificities of five R2R3 MYB proteins: MYB52, MYB59, MYB46, MYB111 and MYB55. Top scoring binding motif for MYB52 corresponded to a MBSI element, but evaluation for secondary motifs revealed that it also recognized very efficiently the MBSII sequence. The remaining R2R3 proteins tested recognized the MBSIIG motif (Fig. S1A). Close evaluation of binding motifs for the different proteins revealed specific particularities among them. MYB46, MYB111 and MYB55, although binding to MBSIIG with higher affinity, also recognized MBSII motifs, whereas MYB59 recognized very specifically MBSIIG variants (Fig. S1A).

We scanned these motifs in the promoters of the genes positive and negatively co-regulated. In general, all the motifs were over-represented in the promoter regions of positively co-regulated genes, indicating that they are transcriptional activators (Fig. S1B). Primary MBSIIG motifs for MYB59, MYB46, MYB111 and MYB55 and secondary MBSII (MYB46, MYB111 and MYB55) were over-represented in the promoters of their corresponding co-regulated genes. MBSII but not primary MBSI, was over-represented in the promoters of genes co-regulated with MYB52 (Fig. S1B). In general, these results reflect binding affinity of these TFs for their different cognate sequences. Enrichment of these cognate DNA-sequences in downstream regions of co-regulated genes was in general lower than in promoters, although some were significantly enriched (Fig. S1C). In particular, MBSIIG motif for MYB59 and secondary MBSII motif for the remaining TFs may have a role as regulatory DNA-elements at downstream regions of target genes.

Data observed for co-regulated genes was also confirmed with transcriptomic analysis involving some of the TFs studied. MYB59 has been implicated in the regulation of cell cycle and root growth, presumably by activating the expression of *CYCB1;1* and other genes implicated in the regulation of cell cycle (5). We found an statistical over-representation of the MBSIIG element recognized by MYB59 in the promoters of genes up-regulated in response to MYB59 overexpression whereas the similar non-cognate MBSII was not (Fig. S1D). MYB111 participates -together with MYB11 and MYB12- in flavonoid biosynthesis by activating the expression of the genes coding the key enzymes of this pathway (6). The MBSIIG element recognized by MYB111 is highly over-represented in the promoters of genes down-regulated in the triple mutant *myb11, myb12, myb111* (Fig. S1E), supporting the role of this motif in *cis*-regulation of flavonoid biosynthesis. Secondary MBSII element was not enriched, suggesting that redundant activity of MYB111, MYB11 and MYB12 over flavonoid biosynthesis is exerted mainly though the MBSIIG motif. MYB46 is a master regulator of secondary cell wall biosynthesis (7) and a set of putative targets has been recently described using a transcriptomic approach of transgenic MYB46-HER (human estrogen receptor) fusion (8). MBSIIG and MBSII were over-represented in the promoters of putative targets (Fig. S1F), suggesting that MYB46 actually binds both elements *in vivo*. Finally, overexpression of MYB52 confers drought tolerance at the post-germinative stage (9) and is also implicated in cell wall growth as a target gene of MYB46 (7, 8). Both MBSI and MBSII motifs were over-represented in the promoters of genes up-regulated in response to drought stress (Fig. S1G). In addition, the MBSI variant CCGTTAK and, as expected, the MBSII motifs were over-represented in the promoters of direct targets of MYB46 (Fig. S1G). These observations suggest that the MBSI element, in spite of its

minor role in the activation of co-regulated genes with *MYB52*, has a relevance in the context of *MYB52* biological role. Overall, the results of transcriptomic assays involving direct or indirectly the TFs studied are coincident to those obtained for co-regulated genes, and reinforce the role of the DNA motifs identified as *cis*-regulatory sequences.

## 1.2. MYB-related

MYB-related TFs contain a single R1/R2 repeat, such as CCA1 and RVE1 (1, 2). CCA1 is a core component of the circadian clock and negatively regulates the expression of evening-expressed genes through binding the evening element (EE, AGATATT; 10, 11). RVE1 expression is rhythmically regulated and connects auxin and circadian signaling networks (12). In our assay, we found a perfect EE as the top-scoring 8mer for both proteins and we did not observed remarkable differences in DNA-binding affinity between them (Fig. S2A and S2B). RVE1 is not a core component of the circadian clock but its overexpression disrupts clock-regulated gene expression (12), what could be explained by the similar DNA-binding properties of CCA1 and RVE1. Evaluation of secondary motifs revealed subtle differences in DNA-binding specificity. Although both proteins recognized EE-related elements, CCA1 bound with high affinity to the element AGATATG, whereas RVE1 preferred the sequence AGATWTT (Fig. S2A). In addition to the EE, it has been described the CCA1-binding sequence (CBS, AGATT[G/T]TT) as a positive *cis*-element for activation of morning-specific genes, such as Lhcb1\*3 (13). In our assay, CCA1 recognized with high affinity the CBS variant AGATTTTT (E-score 0.48635) whereas the variant AGATTGTT was poorly recognized by CCA1 (E-score 0.3782; Fig. S2B). Similar affinities were observed for RVE1, but in this case binding affinity to the DNA sequence AGATTTT was higher than that for CCA1 (Fig. S2B), as reflected in RVE1 secondary element.

EE, secondary and CBS motifs were over-represented in the promoters of genes negatively co-regulated with CCA1 and RVE1, pointing to a role of these DNA sequences in *cis*-regulation (Fig. S2C). However, enrichment of EE was higher, in agreement with the central role of CCA1 in the regulation of the circadian clock by repressing gene expression through EE. This was not the case at downstream regions, where only EE for CCA1 and secondary motif for RVE1 were enriched (Fig. S2D). Consistently to co-regulated genes, the EE was highly enriched in the promoters of genes up-regulated in double *cca1/lhy* mutant (14). By contrast, neither CBS or secondary motif play a major role in *cis*-regulation of gene expression dependent on CCA1/LHY under these experimental conditions (Fig. S2E). Similar results were obtained in the case of genes regulated by the structurally similar protein RVE8. Recently, it has been shown that RVE8 directly activates evening-specific genes through the EE (15). In this case, we observed over-representation of the EE in the promoters of direct targets of RVE8, whereas CBS or secondary motif were poorly or no over-represented (Fig. S2F).

Recently, it has been proposed RVE1 as a candidate gene for a QTL responsible of freezing tolerance and suggested that RVE1 is a negative regulator of freezing tolerance in *Arabidopsis* (16). We scanned RVE1 binding motifs in the promoters of genes responsive to cold treatments and found the EE highly enriched in the promoters of genes down-regulated (Fig. S2G).

## 1.3. MYB GARP type-B ARR

GARP family is characterized by a single MYB repeat and two representative classes can be distinguished, type-B ARRs and G2 (1). We have analyzed the binding specificities of two type-B ARRs, ARR11 and ARR14. Both proteins recognized a similar motif containing the sequence AGATWCG (Fig. S3A), that matches the motifs recognized by other type-B ARRs described in the literature, in which the core element AGAT was found to be indispensable for DNA binding (17). Although with the highest affinity, binding of both proteins to the sequence AGATWCG was not exclusive. Evaluation of secondary motifs identified novel DNA elements recognized by both transcription factors, containing AGAT or GGAT core sequences (Fig. S3A). ARR14 binds very efficiently to all the motifs, given their close E-scores (AGATACGG, E=0.49510; AAGATCTT, E=0.4947; CGGATCCG, E=0.4910). By contrast, ARR11 recognized more efficiently the primary motif (AAGATACG, E= 0.4960) than secondary and tertiary elements (CGGATCCG, E=0.4794; AAGATCTT, E= 0.4596).

The elements recognized by ARR11 and ARR14 *in vitro* were poorly or not over-represented in the promoters of their corresponding co-regulated genes, where only AGATACG for ARR11 and GGATCC for ARR14 were slightly over-represented (Fig. S3B). This is not surprising, since type-B ARRs are primary TFs involved in cytokinin signaling and their activation depends on a phosphorelay cascade triggered by cytokinin receptors (18). Actually, the expression of type-B ARR genes is not activated in response to the hormone, what could explain the fact that ARR-binding motifs were not enriched in the promoters of co-regulated genes. Similarly, these DNA sequences are poorly enriched at downstream regions of co-regulated genes (Fig. S3C). Poor enrichment was also observed among down-regulated genes in triple mutant *arr1,10,12* (19), involving three redundant genes in cytokinin signaling, although significant over-representation for the top-scoring motif was observed (Fig. S3D). In addition, the motifs AGATACG and AGATCTT were enriched in the promoters of genes early induced after treatment with cytokinin (Fig. S3E). Finally, the elements AGATACG and GGATCC were also significantly enriched in the promoters of genes down-regulated in a mutant genotype with reduced levels of endogenous cytokinins (quadruple mutant *ipt1,3,5,7*, ref. 20; Fig. S3F). These results concerning enrichment of DNA-motifs recognized by ARR11 and ARR14 *in vitro* support their role as positive regulators of gene expression through these binding elements.

#### 1.4. MYB GARP-G2

GARP-G2 TFs are a different subclass of GARP proteins, in which two different groups can be distinguished, represented by Golden2-like (GLK) and KANADI (KAN) proteins. We analyzed three proteins belonging to the two different groups, GLK1, KAN4 and KAN1. GLK1 showed similar DNA-binding properties than other GARP proteins, recognizing DNA motifs containing the core sequence RGATATCY (being R purine and Y pyrimidine; Fig. S4A). Evaluation of secondary motifs yielded a palindrome containing the core sequence AGATTCT (Fig. S4A), revealing a broader range in DNA binding specificity for this protein. Although structurally related to GLKs, KAN4 (or ATS) bound with the highest affinity to a different cognate sequence, GAATATTC, although DNA variants of the primary sequence revealed that some of the positions of the 8mer indicated are less restrictive (Fig. S4A). KAN1 exhibited a poor DNA-binding capacity in our standard DNA-binding conditions, although a DNA motif containing the KAN4-related sequence GAATAW was obtained with the highest score (Fig. S4A).

The elements found for this group of GARP TFs were enriched in the promoters of co-regulated genes (Fig. S4B). Of particular interest, both primary and secondary motifs found for GLK1 were enriched, supporting their biological relevance (Fig. S4B). It is worth noting that the motif found for KAN1 is related with transcriptional repression. Consistently, KAN1 has been proposed to act as a repressor (21, 22) and the element found resembles to that required in KAN1-dependent transcriptional repression of *ASYMMETRIC LEAVES2* (AGAATAA, underlined G critical residue; 20). GLK1 and GLK2 act redundantly by promoting the expression of a number of genes that are required for chlorophyll biosynthesis and light-harvesting functions (23). In this work, the authors identified by predictive means the element CCAATC as a putative GLK recognition sequence, but they were not able to confirm experimentally this sequence (23). In our binding assays, we have not detected significant binding of GLK1 to that element (top CCAATC-containing element: ATCCAATC, E-score 0.26), suggesting that the GLK1-recognition sequence may be more similar to those of GARP proteins than previously indicated. Primary binding motif found in our *in vitro* experiments was enriched in the promoters of the genes up-regulated in overexpressing *GLK1* and *GLK2* plants (Fig. S4D). The secondary motif was not significantly over-represented, suggesting that this *cis*-element may be involved in regulation of gene expression in a different developmental stage.

## 2. Trihelix

The trihelix family is exclusive from the plant kingdom and is defined by a DNA-binding domain distantly related to the MYB domain. Proteins belonging to this family bind the GT motif (GRWAAW, R: A or G; W: A or T; ref. 24), although members from different clades show specificity for different GT motifs (25). At5g28300 is included in the GT-2 clade and our DNA-binding assays for this protein yielded a DNA element resembling the GT3-box (GGTAAAT; ref. 26) recognized by the protein GT-2 that belongs to the same clade (Fig. S5A). It is worth noting that the proteins included in the clade GT-2 have two trihelix DNA-binding domains located at the N- and C-terminal ends of the proteins. Both domains bind DNA autonomously to different

target sequences (26): the N-terminal DBD preferentially recognizes the GT3-box (GGTAAA) whereas C-terminal prefers GT2-box (GGTAAT). In our *in vitro* experiments, we observed high binding affinity to the GT2-box, in addition to the GT3-box indicated, indicating that the full-length version of the protein tested may recognize both motifs, although binding to GT3 motif was preferred (Fig. S5A and Fig. S5B). Consistently, the elements GT3 and GT2 were over-represented in the promoters of genes negatively co-regulated with At5g28300, indicating that this protein may act as a transcriptional repressor (Fig. S5C). By contrast, these motifs located at downstream regions have a minor regulatory role (Fig. S5D).

The similar trihelix protein GTL1 has also been shown to regulate endoreduplication in trichomes and stomatal development by repressing target genes (27-29). Consistently, GT3 and GT2 elements were over-represented in the promoters of genes up-regulated in a loss-of-function *gtl1* mutant, supporting that related proteins recognize similar DNA-elements (Fig. S5E).

### 3. AP2/EREBP superfamily

The APETALA2/ETHYLENE RESPONSIVE ELEMENT BINDING PROTEIN (AP2/EREBP)-like proteins comprise one of the largest family of TFs in plants, with approximately 150 members in Arabidopsis. This family is characterized by the presence of the AP2-DNA binding domain of about 60 residues long, and can be divided into four major groups: AP2, DREB, ERF and RAV subfamilies (1). In this work we analyzed several proteins belonging to the subfamilies ERF, DREB and AP2

#### 3.1. ERF

Members of this subfamily are usually involved in biotic stress responses and act as regulators of the ethylene signaling pathway through their interaction with the GCC-box (30, 31). In our previous work, we characterized the DNA-binding specificity of ERF1 (32). In this work, we analyzed the DNA-binding specificity of RRTF1, RAP2.3, RAP2.6 and AtERF1. All the proteins tested recognized the GCC-box (Fig. 6SA). Among top-scoring 8mers, we found in some cases sequence variants of the canonical GCC box, with changes at positions 5<sup>th</sup> and 6<sup>th</sup> of the sequence GCCGCC. We performed an “*in silico* mutagenesis” (as in Ref. 32) of the canonical GCC element to analyze the relative binding to different GCC-related sequences. As we previously showed, ERF1 showed marginal binding to other GCC-related sequences (32). In contrast, the proteins analyzed in this work showed high-to-moderate binding affinity for other GCC-related variants. In particular, RAP2.6 showed high binding affinity (median E-score above 0.42) for four sequence variants of the canonical GCC-element (GCCGCA, GCCGAC, GCCGTC, GCCGGC; Fig. 6SB)

GCC-elements were over-represented in the promoters of the genes positively co-regulated with *RRTF1*, *RAP2.3*, *RAP2.6*, *AtERF1* and *ERF1* (Fig. 6SC). In relation to DNA-variants, GCCGTC for RRTF1 and RAP2.6 and GCCGAC for RAP2.3 were over-represented in the promoters of the corresponding co-regulated genes, suggesting that their role in *cis*-regulation of gene expression is less important than canonical GCC-elements. Interestingly, among five DNA elements recognized by RAP2.6, just two have a biological role, indicating that the broader the DNA binding specificity, the lower the representation of *cis*-elements in promoters of co-regulated genes. Regarding downstream regions, only marginal enrichment of particular DNA elements was observed (Fig. 6SD). In our previous work (32), we showed that the GCC-box recognized by ERF1 was highly enriched among promoters of genes up-regulated in response to *ERF1* activation, using an ethanol inducible promoter. In addition, the elements GCCGGC and GCCGTC were also over-represented, but only at longer time points after induction of the expression of *ERF1*, suggesting a cascade of ERF1-related TFs acting downstream of ERF1 (11). We extended the study of ERF1-related binding elements by analyzing all the DNA motifs in the promoters of genes rapidly induced by the activation of *ERF6* overexpression, as candidates of ERF6 target genes (33). In this case, we observed significant enrichment of all the elements but one (GCCGCA) in the promoters of putative targets (Fig. 6SE). In spite of this, enrichment of the GCC-box was higher than the corresponding to the other elements. This results suggest that ERF-related TFs mainly recognize the GCC-box. However, other secondary motifs partially differing from the canonical GCC-element might have a biological role in the context of ERF-related transcriptional regulation.

### 3.2. DREB

The DREB family of AP2/EREBP TFs is one of the best studied, and many of its members are implicated in the regulation of the plant responses to abiotic stress. The first proteins identified for this family, DREB1A and DREB2A, interact specifically with drought-responsive element (DRE) with the sequence [G/A]CCGAC (34). We have analyzed the DNA-binding specificity of 6 DREB proteins: DDF1 (subfamily A-1), DREB2C (subfamily A-2), At1g77200 (subfamily A-4) and ORA47, DEAR3 and DEAR4 (subfamily A-5). DDF1, At1g77200, ORA47 and DEAR4 recognized the DRE motif [G/A]CCGAC, although with particular specificities relating their nucleotide preferences within the core DRE element or in the flanking positions (Fig. S7A). For instance, DEAR4 recognized more specifically an A residue at the first position in the DRE motif, whereas DDF1 preferred G instead of A and ORA47 and At1g77200 did not show preferences for any of these two nucleotides. In relation to the 5<sup>th</sup> position in the DRE-box, DDF1, DEAR4 and At1g77200 showed great dependence of an A, whereas ORA47 was more relaxed. More variable were the binding preferences in relation to nucleotide composition outside the DRE-element, where some proteins recognized more efficiently 3' extended DRE-elements (DDF1, At1g77200, ORA47 and DEAR4) or extended 5' DRE-motifs (DEAR4). Of particular interest, DREB2C and DEAR3 recognized GCC-motifs, similar to those recognized by ERF proteins (Fig. S7A). Moreover, evaluation of secondary motifs revealed that ORA47 and DEAR4 also bound very efficiently to DNA-motifs resembling the GCC-box (Fig. S7A), whereas secondary motifs for DREB2C and DEAR3 contained the expected DRE sequence (Fig. S7A). These results suggest that DNA-binding of DREB proteins may be more complex than initially thought, where different TFs specifically recognize different DRE-elements and, most of them (particularly from A-4 and A-5 subfamilies), possess a broader range of DNA-binding by recognizing GCC-related elements.

We analyzed the incidence of DREB-binding motifs in the promoters of their corresponding co-regulated genes. We observed significant over-representation of the DRE-elements in the promoters of positively co-regulated genes with DDF1, DREB2C, At1g77200 and ORA47 (Fig. S7B). By contrast, over-representation was observed among negatively co-regulated genes with DEAR3 and DEAR4, consistently with their role as transcriptional repressors containing an EAR motif (35). More interestingly, the GCC-like elements observed for DREB2C and ORA47 were also over-represented in the promoters of positively co-regulated genes and, in the case of DEAR3 and DEAR4, in the sets of negatively co-regulated promoters (Fig. S7B). These results indicate that both DRE-like motifs and GCC-like elements are involved in the *cis*-regulation of gene expression directed by these TFs. In relation to downstream regions, only DNA motifs recognized by DEAR4 and the GCC-element corresponding to ORA47 were over-represented (Fig. S7C).

Definition of extended DRE-motifs may be important for defining putative targets of TFs, as in the case of DDF1. Overexpression of DDF1 causes dwarfism by reducing the levels of active GAs, most likely by activating GA2ox7 gene expression through binding to DRE-like motifs, providing a mechanism for adaptation of plants to salt stress (36). Extended DRE-motifs that we obtained *in vitro* were strongly over-represented in the promoters of genes up-regulated in Ox-DDF1 plants (Fig. S7D). However, a “mutated” version of the DRE-box differing at 3'-end extension was not over-represented in the same set of promoters (Fig. S7D). Similarly, extended DRE-boxes may be relevant for At1g77200 protein binding to DNA. In this case, we analyzed gene expression of a gain-of-function allele of the gene *HRD* (37), that belongs to the same subfamily A-4 of DREB TFs. Extended DRE-element is highly over-represented in the promoters of genes up-regulated in *hrd-D* mutant, whereas a corresponding mutant version at the 3'-end was poorly over-represented (Fig. S7E). Finally, ORA47, a positive regulator of JA-mediated signaling (38), also displayed higher affinity for extended DRE-motifs. In this case, both minimum and extended DRE-motifs were over-represented in the promoters of genes up-regulated in response to JA, but the over-representation is proportionally higher for extended motifs (Fig. S7F). In this case, also GCC-elements were over-represented among JA-inducible promoters (Fig. S7F), pointing to ORA47 as a convergence point between abiotic and biotic signaling pathways.

### 3.3. AP2

TOE1 and TOE2, together with AP2, SMZ, SNZ, belong to the same clade targeted by miR172, and they act redundantly in the repression of flowering (39, 40). Both proteins recognized with similar affinities the identical 8mer CCTCGTAC, although some degree of degeneracy is permitted when analyzing secondary motifs (Fig. S8A). Moreover, both proteins showed to recognize with high affinity other DNA-elements differing from the primary ones, suggesting that these proteins recognize different DNA motifs (Fig. S8A). In general, DNA variants of the primary motif, and the sequence AACCTTAA (3ary motif) were significantly enriched in the corresponding negatively co-regulated promoters (Fig. S8B), as expected for transcriptional repressors. In fact, the five proteins belonging to this clade contain EAR motifs and interact with TPL co-repressors, so they can be considered as transcriptional repressors (35, 39, 40). Interestingly, some DNA variants were over-represented at downstream regions of co-regulated genes, indicating that transcriptional regulation of targets may also occur through recognizing specific downstream DNA-sequences (Fig. S8C).

Given the high degree of similarity in DNA-motifs recognized by TOE1 and TOE2, it is expected that AP2 and SMZ proteins recognize similar elements. Consistently, DNA sequences corresponding to primary and secondary motifs were over-represented in the promoters of genes up-regulated in the loss-of-function mutant *ap2* (Fig. S8D). Similarly, the three motifs were also enriched among the promoters of genes down-regulated in the gain-of-function mutant *smzD* (Fig. S8E). These results point to a role of these DNA elements in the repression of flowering exerted through the TOE/AP2 clade of transcriptional repressors.

#### 4. Homeodomain-containing proteins

Plant homeodomain (HD) proteins can be classified in 6 families, according to additional conserved domains beyond the HD DBD. These families include: HD-Zip, PHD-HD, Bell, ZF-HD, NOX and WOX (41). In this work, we analyzed several proteins belonging to WOX and HD-Zip families.

##### 4.1. WOX family

WUS homeobox-containing (WOX) proteins are plant-specific TFs and can be classified into three different clades: the WUS-related, intermediate and ancient clades (42). WOX13 belongs to the ancient clade, and we observed a clear binding preference to DNA containing the motif CAATCA (Fig. S9A). Evaluation of secondary motifs identified the element TAATTA as recognized by WOX13 with lower affinity. These motifs are partially related to that recognized by the protein WUS during transcriptional activation of AGAMOUS (TTAATSS, being S G or C; ref. 43). However, we did not find significant binding of WOX13 to a “perfect” WUS motif, as well as for the G-box (CACGTG), which has been described as a secondary binding element for WUS (Fig. S9B; ref. 44). These results may reflect the diversity and specificity in DNA recognition among WUS-related proteins belonging to different clades.

Both primary and secondary motifs were over-represented in the promoters of positively co-regulated genes with *WOX13* as well as in downstream regions, indicating that both elements are recognized *in vivo* by this TF (Fig. S9C and Fig. S9D). We also analyzed the presence of the same *cis*-elements within genes regulated in response to ethanol-inducible WUS expression (45). As indicated above, WUS belong to the same family, although grouped in a different clade and thus, expected to share some DNA-binding specificities with WOX13. We observed significant over-representation for primary WOX13 binding motif (but not for secondary) in the promoters of genes down-regulated in response to *WUS* overexpression. This result suggests that WUS may also negatively regulate gene expression by recognizing this element, as expected for this transcriptional repressor (Fig. S9E).

##### 4.2. HD-ZIP family

HD proteins containing a leucine zipper domain (HD-ZIP) bind as dimers to a palindromic DNA sequences containing the motif AATNATT, although there is some DNA-binding specificity between proteins in this group. We have analyzed INCURVATA4 (ICU4, CORONA, ATHB15; HD-ZIPIII class), ATHB51 (LATE MERISTEM IDENTITY1, HD-ZIPI class) and ATHB12 (HD-ZIPI class). All three proteins recognized similar elements matching the consensus binding sequence for this group of TFs (46), and showed some specificity in the flanking and central bases (Fig. S10A). We analyzed with more detail the preferred nucleotide at the central position

of the consensus binding motifs for the three proteins, and found that ICU4 and ATHB51 preferred A/T and this position. By contrast, ATHB12 preferentially recognized DNA elements containing G/C at the central position (Fig. S10A and Fig. S10B). These results contrast to previous observations, in which HD-ZIPIII proteins (such as ICU4) prefer G/C at the central positions and HD-ZIPI TFs strongly bind to A/T-containing motifs (41). DNA sequences observed for ICU4 and ATHB12 proteins were over-represented in the promoters of their corresponding co-regulated genes (Fig. S10C). By contrast, DNA motifs for ATHB51 were highly over-represented in the promoters of negatively co-regulated genes (Fig. S10C). Over-representation was more evident in the case of partial sequences, given the low number of genes containing complete binding motifs at their promoters. We obtained small differences regarding the middle nucleotide in the consensus sequence. In particular, DNA-motifs with A/T at the central position were more enriched than G/C ones in the case of ICU4 (two-fold more enriched relative to random distribution in the case of A/T-containing motifs and 65% higher for G/C ones) and ATHB51 (61% A/T versus 46% GC). By contrast, elements containing G/C seemed more relevant in the case of ATHB12 co-expressed genes (58% more enriched G/C-containing sequences versus 26% A/T-ones). These data perfectly reflect DNA-binding preferences observed for these proteins in relation to the residue at the central position. Interestingly, binding motifs were also over-represented (although in lower degree) in downstream regions, indicating that these HD-ZIP proteins may regulate transcription through recognizing *cis*-elements located upstream or downstream the target genes (Fig. S10D).

Previous data regarding ATHB51/LMI1 suggest that this protein acts as an activator of gene expression in concert with LEAFY, in particular of the *CAL* gene (47). Our results indicate that this protein acts mainly as a repressor, given the high over-representation of *cis*-elements in the promoters of negatively co-regulated genes. However, we cannot exclude the possibility of ATHB51/LMI1 acting as an activator in combination with other TFs in a particular niche of cells or stage of development. We analyzed transcriptomic profiles of *Arabidopsis* plants overexpressing the gene *HAHB4* from sunflower (48). Given that closest *Arabidopsis* homologs are *ATHB7* and *ATHB12*, it is expected that they recognize similar DNA-sequences. Both DNA-elements recognized by HD-ZIPI proteins are enriched in the promoters of up-regulated genes (Fig. S10E). Moreover, the DNA motif containing G/C at the middle position was more enriched than the A/T-containing variant, in concordance with the high homology between *HAHB4* and *ATHB12*. These results indicate that similar TFs from different species might recognize identical DNA sequences *in vitro* and share common targets when expressed heterologously *in planta*.

## 5. SQUAMOSA Promoter-Binding Protein-Like

SQUAMOSA Promoter-Binding Protein-Like (SPL) proteins are an evolutionary conserved family of TFs found from unicellular algae to higher plants that participate in numerous developmental and stress-related processes (49). SPL proteins bind DNA through recognition of the SBP-box, TNCGTACAA, where the core GTAC is essential for binding (50, 51). We analyzed the DNA-binding specificity of SPL1 and SPL7 and found that both recognized the SBP-core (Fig. S11A). Although a longer motif was observed for SPL1, similar scores were obtained for other elements containing the SBP-core, in particular the sequence CCGTAC, suggesting that this may be sufficient for recognition (Fig. S11A).

Core SBP-boxes were slightly enriched in the promoters of co-regulated genes with *SPL1* and *SPL7*, although their representation was discrete, particularly for SPL7 (Fig. S11B). Nevertheless, significant enrichment of the element CGTAC was observed in the promoters of co-regulated genes, but not at downstream regions (Fig. S11B and Fig. S11C). SPL7 is a regulator for copper homeostasis in *Arabidopsis* and is required for regulating gene expression in response to copper-limiting conditions (52, 53). However, the expression of *SPL7* is not affected by copper availability and several SPL7-targets are miRNAs that regulate post-transcriptionally some key components for the control of copper homeostasis (52, 53). This additional level of regulation could be obscuring the results of our co-regulation analysis. Consistently, only the sequence CGTAC was enriched in the promoters of genes down-regulated in the mutant *sp7* in copper limiting conditions (Fig. S11D).

Given the similarity in DNA recognition of the SPL proteins studied so far, it is expected that other SPLs recognize similar elements. This might be the case of SPL13, whose coding transcript is a target of miRNA156. Expression *in planta* of a miRNA156-resistant version of *SPL13* (mSPL13) over-accumulates *SPL13* transcripts, what could be considered as a gain-of-function mutation of *SPL13* (54). The analysis of the transcriptomic profiles of mSPL13 plants

showed that SBP boxes were over-represented in the promoters of up-regulated genes, suggesting that SPL13 recognize similar SBP boxes to those for SPL1 and SPL7 (Fig. S11E).

## 6. WRKY

WRKY is one of the largest families of plant-specific transcription factors. They participate in the regulation of plants responses to abiotic and biotic stresses, as well as in developmental processes, such as seed development or senescence (55). All WRKYS previously tested recognize the so-called W box (TTGACY; Y, C or T), although some proteins require specific nucleotides at the 5' end of this box, and this requirement is not group specific (56). We tested four WRKYS grouped in different classes, according to a previous classification (57): WRKY12 (group IIc), WRKY18 (group IIa), WRKY45 (group I) and WRKY38 (group III). All the proteins recognized a perfect W-box, although we observed some specificity in the nucleotides flanking the core W-box (Fig. S12A). For instance, WRKY18 preferred C (instead of T) at 3' end of the W-box, whereas it did not showed any preference for nucleotides at 5'end of the W-element. WRKY12, WRKY38 and mainly WRKY45 recognized more efficiently W-box containing CG at the 5'-end of the minimum element but their preference for C or T at the 3' end was less strict. As described before (56), the preferences for specific flanking nucleotides were not group specific. We have not observed significant binding to other elements different from W-box, revealing a high binding specificity for this motif.

We evaluated the presence of W-boxes in the promoters of genes positive and negatively co-regulated with the genes encoding these transcription factors. W-boxes recognized by WRKY45 and WRKY38 were significantly over-represented in the promoters of genes positively co-regulated (Fig. S12B). WRKY18 and WRKY12 are transcriptional repressors and their cognate DNA motifs were over-represented in the promoters of genes negatively co-regulated (Fig. S12B). It is worth noting that the core W-box was also over-represented in positively co-regulated genes with WRKY18, what could be interpreted as a bi-functional role of this TF. However, a more likely alternative is that over-representation of core W-elements may be the result of a transcriptional cascade involving several WRKYS in response to WRKY18 activity. In line with this, several WRKY-coding genes (*WRKY33*, *WRKY40*, *WRKY46* and *WRKY48*) are co-regulated with *WRKY18*, what could explain enrichment of core elements among co-regulated promoters. Importantly, only extended W-box for WRKY18 was over-represented among negatively co-regulated promoters, supporting the relevance of this extended W-box in WRKY18-dependent gene repression and the importance of the determination of extended motifs. In the case of WRKY45, WRKY38 and WRKY12, both extended and minimal W-boxes were enriched. Of particular interest, extended W-boxes were also over-represented at downstream regions of the genes co-regulated with WRKY45 and WRKY38, indicating that regulation of transcription by these two TFs is not restricted to promoter regions of targets genes (Fig. S12C).

WRKY18, together with redundant WRKY40 and WRKY60, has a negative effect on basal defense to *Pseudomonas syringae* and powdery mildew by negatively regulating SA-mediated defense (58-60). Similarly to data corresponding to co-regulated genes, several W-variants were over-represented in the promoters of genes up-regulated in the double mutant *wrky18/40*, consistently with a role of these proteins as transcriptional repressors (Fig. S12D). Moreover, over-representation was proportionally higher for extended W-elements, consistently to the observed with co-regulated genes. Definition of extended motifs is particularly important for WRK12. This gene and its ortholog in *Medicago truncatula* are negative regulators of secondary wall formation, and their corresponding loss-of-function mutations cause secondary cell wall thickening (61). Furthermore, both ortholog genes act as transcriptional repressors by direct binding to their target promoters (61). We analyzed the presence of different W-boxes in the promoters of mis-regulated genes in two different *wrky12* mutant alleles (61). In this case, only extended motifs were biologically relevant in the context of WRKY12-mediated transcriptional repression (Fig. S12E). WRKY38 is highly induced by both pathogen infection and SA treatment and functions as negative regulator of plant basal defense, participating in a complex network of fine-tuning regulation also involving NPR1 and the transcriptional repressor HDA19 (62). Similarly to examples described above, the W-box is over-represented in the promoters induced after treatment with SA, but this over-representation is more evident with extended W-boxes (Fig. S12F).

## 7. C2H2-type zinc finger

C2H2-type zinc finger proteins are one of the most abundant and heterogeneous family of TFs in *Arabidopsis* (1). C2H2 proteins can be divided into three sets, and each one in different subsets, families and classes (63, 64). One of the most studied families corresponds to C1 set, whose members contain one to five C2H2-type zinc fingers. In particular, ZAT2 (DAZ1) contains three C2H2 domains whereas At4g35610 just one. Both proteins showed very similar binding properties, recognizing a DNA element that contains an AGCT core (Fig. S13A). These binding motifs agree with those observed for the petunia ZPT2-2 protein, whose optimal binding sequences are AGCT and CAGT for the first and second zinc finger domains, respectively (65). Among C1 set, C2H2 proteins with two Zn-finger domains are the most abundant and investigated. Within this group, we analyzed ZAT6, ZAT14 and ZAT18. ZAT14 and ZAT18 recognized the element AGTGCAC, that contains two inverted ACT motifs spaced 2 nucleotides (Fig. S13A). These binding motifs reveal a high degree of complexity in the patterns of C2H2-DNA recognition. Other C2H2 proteins with two zinc finger domains bind to bipartite sequences containing two AGT (or ACT) motifs spaced 3 or 4 nucleotides (66). ZAT6, however, recognized a 8mer DNA sequence that contains just one core AGT motif. Computing for gapped 8mers failed to detect additional AGT core sequences in ZAT6, whereas ZAT14 recognized with similar affinities bipartite AGT core motifs that can be spaced 1 to 4 nucleotides (Fig. S13B).

These proteins, together with most of the proteins belonging to this clade of C2H2 TFs, have EAR domains and are expected to act as transcriptional repressors (35, 67). Consistently, DNA elements recognized by ZAT2 were enriched in the promoters of negatively co-regulated genes (Fig. S13C). In addition, the number of AGCT core elements per promoter is higher in this set of promoters (3.91 sites per promoter) than in positively co-regulated genes or in the complete genome (3.19 and 3.42, respectively). Similarly, DNA-motifs for ZAT6 were also over-represented in the promoters of negatively co-regulated genes (Fig. S13C). In relation to ZAT14 and ZAT18, DNA-binding assays revealed a high degree of flexibility in DNA recognition that may obscure the identification of *cis*-elements in promoters. In fact, cognate DNA-elements were poorly enriched in the promoters of negatively co-regulated genes, although bipartite sequences containing the AGT spaced at three nucleotides were significantly over-represented, supporting their role as repressive *cis*-elements (Fig. S13C). In relation to downstream regions of co-regulated genes, only the element found for ZAT2 was significatively enriched among negatively co-regulated promoters, whereas DNA-motifs corresponding to ZAT6, ZAT14 and ZAT18 were poorly or no over-represented (Fig. S13D).

We extended the analysis of DNA motifs to transcriptomic experiments involving similar TFs. The closest paralog to ZAT6 is ZAT10 that groups together with AZF1 and AZF2 in the same clade (67). The element ACACCA was over-represented in the promoters of genes down-regulated after overexpression of ZAT10, AZF1 and AZF2 (Fig. S13E), as expected for a group of TFs bearing an EAR motif and acting as transcriptional repressors. In relation to ZAT14 and ZAT18, the motifs obtained were scanned in the promoters of genes regulated by ZAT12 included within the same clade. All the bipartite AGT-containing motifs were over-represented in the promoters of down-regulated genes in ZAT12 over-expressing lines, supporting their role as *cis*-elements involved in transcriptional repression (Fig. S13F).

## 8. C2C2-type zinc finger

C2C2-type zinc finger proteins are an heterogeneous group of TFs including the families YABBY, Dof, CO-like and GATA, being the three former families specific to plants (1). In this work, we analyzed several TFs belonging to GATA, YABBY and Dof families.

### 8.1. GATA

GATA-type TFs are found in all eukaryotes from fungi to plants and metazoans. GATA factors from different kingdoms recognize the consensus sequence WGATAR (W, T or A; R, G or A; ref. 68) and this binding specificity is also conserved for several plant GATA proteins regulating the expression of light-responsive genes (69). We analyzed GATA12 and found to recognize very specifically the palindromic motif AGATCT, that contain the core motif GATC (Fig. S14A). Evaluation of secondary motifs always identified GATC-containing elements, while surrounding positions are less informative for protein binding. Binding of GATA12 to the similar GATA-containing motif AGATAT was much lower than to the corresponding GATC-containing sequence (Fig. S14B).

We focused on GATA12-dependent gene expression by analyzing the set of genes positively and negatively co-regulated with *GATA12*. In this case, the GATC-containing element was over-represented in the promoters of genes co-regulated with *GATA12*. The similar GATA-containing element was slightly enriched in promoters of both positive and negatively co-regulated. These results suggest that *GATA12* regulates expression through the cognate GATC-containing element, but we cannot exclude a role of the similar GATA-box (Fig. S14C).

The GATA-type proteins GNC and GNL are repressors of gibberellin signaling downstream from the DELLA and PIF regulators, and they are supposed to act as transcriptional repressors (70). We scanned GATA binding motifs in the promoters of genes de-regulated in plants overexpressing *GNL* and *GNC* genes. In this case, we observed enrichment for both GATC- and GATA-containing motifs in the promoters of down-regulated genes, supporting the role of GNC and GNL as transcriptional repressors (Fig. S14E). Over-representation was slightly higher for GATC-containing motif, and this sequence, but not the GATA-containing, was under-represented in the promoters of genes up-regulated (Fig. S14F), suggesting that GNC and GNL might recognize more specifically GATC sequences, as we found for GATA12.

## 8.2. YABBY

The YABBY family (YAB) of TFs is exclusive from seed plants and is characterized by a zinc finger domain close to the N-terminus and a helix-loop-helix (YABBY) domain at the C-terminal end, partially similar to the HMG domain of high-mobility group non-histone proteins (71). Studies on the molecular mechanisms of YAB function are scarce and controversial. First, YAB proteins are supposed to bind DNA in a non-sequence-specific manner through their YABBY domain (72), but rice YABBY1 recognizes a gibberellin responsive element in a DNA sequence-specific manner (73). Second, several reports suggest that YABs are bifunctional TFs that can act as transcriptional activators but also as repressors through their interaction with LEUNIG and LEUNIG\_HOMOLOG co-repressors (74, 75). We analyzed DNA-binding specificity of two YABs: YAB1 (AFO, FIL) and YAB5. Both proteins displayed similar binding specificity to A/T-rich elements, in agreement with that expected for HMG-like DNA-binding domains (Fig. 15SA). Given the similarity between these recognition elements, we can define tentatively a consensus binding sequence for this family of TFs as WATNATW.

YAB-binding motifs were enriched in the promoters of their corresponding negatively co-regulated genes, indicating that these proteins act as transcriptional repressors (Fig. 15SB). We did not observe particular differences in enrichment of the elements AATNATAA and AATNATTA, what suggests that these variants are recognized with similar affinities by both proteins. One of the variants was also over-represented at downstream regions, indicating that transcriptional repression might also occur through recognizing *cis*-elements located downstream the target genes (Fig. 15SC).

We also analyzed these elements in a transcriptomic assay aimed to identify YAB1 targets, by expressing a YAB1-GR fusion protein (75). The elements recognized by YAB1 and YAB5 were over-represented in the promoters of genes responsive to post-translational activation of YAB1 upon treatment with dexamethasone (Fig. 15SD). This tendency was observed not only for down-regulated genes -as expected for a transcriptional repressor- but also for induced genes upon YAB1 activation. These results agree with the reported bifunctional activity of YAB1 (and most likely YAB5) as an activator or repressor TF (74). Similarly, we scanned YAB1- and YAB5-binding sequences in the promoters of genes de-regulated in double *yab1,yab3* and triple *yab1,yab3,yab5* mutants (76). Binding motif for YAB1 may have a role during both activation and repression of gene expression, given its over-representation in both up- and down-regulated genes (Fig. 15SE). Interestingly, the DNA element identified for YAB5 was enriched only in the promoters of genes up-regulated in the triple mutant *yab1,yab3,yab5* (Fig. 15SE), indicating that YAB5 acts as a transcriptional repressor and that, in spite of their similarity, the DNA-motifs discriminate between transcriptional activation or repression depending on their specificity for different YAB proteins.

## 8.3. DOFs

DOF (DNA-binding with one finger) domain proteins are plant-specific TFs with a conserved DNA-binding domain of the type C2C2 zinc finger, including approximately 40 members in *Arabidopsis* (77). All the DOF proteins studied to date recognize a DNA element AAAG as the

essential core sequence with a limited requirement of the nucleotides flanking this sequence. We analyzed DAG2 and Dof5.7 and both proteins bound to canonical Dof-sequences containing the AAAG core (Fig. 16SA). Outside the conserved core, DAG2 recognized with similar affinities different variants containing the consensus sequence WAAAGT (Fig. 16SA and Fig. 16SB). Sequence requirements for Dof5.7 were more strict, and DNA motifs containing the sequence AAAAAG(G) were preferred (Fig. 16SA and Fig. 16SB). We did not identify other DNA motifs different from the Dof core for DAG2. By contrast, Dof5.7 recognized very efficiently motifs containing a different DNA sequence, what could be considered as a secondary element for this TF (Fig. 16SA and Fig. 16SB).

Several Dof-core containing motifs were enriched in the promoters of co-regulated genes. In the case of DAG2, strong over-representation was observed for positively co-regulated genes, whereas enrichment for Dof5.7 motifs was detected in negatively co-regulated genes, indicating that this TF may act as a transcriptional repressor (Fig. 16SC). These over-representations correlate with the average number of shorter motifs containing the core sequence per promoter. The average of the element AAAAAG per promoter of co-regulated genes was higher than in the total number of promoters (3.764 versus 2.997, p-value ~1.26E-7, t-test) or in the negatively co-regulated genes (3.764 versus 2.758, p-value ~3.26E-7, t-test). Analogously, the average number of motifs per promoter increased in negatively co-regulated genes with Dof5.7 with respect to the complete genome (3.310 versus 2.997, p-value ~0.026, t-test) and relative to co-regulated genes (3.310 versus 2.839, p-value ~0.017, t-test). In relation to the secondary motif found for Dof5.7, we did not observe significant enrichment in the different sets of co-regulated genes (Fig. 16SC) and thus, we could not assign a biological role for this sequence. Some DNA variants were also over-represented at downstream regions of co-regulated genes (Fig. 16SD).

## 9. TCP

TCP is a plant-specific family of TFs characterized by a particular basic helix-loop-helix domain involved in DNA-binding, referred to as the TCP domain (78). TCP genes are grouped in two phylogenetic classes based on particular characteristics of their TCP domains and, hence, on their DNA-binding preferences. With this respect, class I TCPs bind to GGNCCAC sequences, whereas the consensus for class II TCPs is G(T/C)GGNCCC (79). However, these sequences are overlapping and they are not exclusive, since some class II factors bind to class I motifs, and *vice versa* (80). We analyzed binding affinity of three class I TCPs: TCP15, TCP23 and TCP16. In agreement with their close homology, TCP15 and TCP23 recognized exactly the same class I element, with only minor specificities outside the consensus sequence (Fig. S17A). This element (GGNCCAC) was over-represented in the promoters of positively co-regulated genes, but not at downstream regions, indicating its role as a *cis*-regulatory element for gene expression (Fig. S17A and Fig. S17B).

TCP16, a class I TCP but differing in a critical residue in the TCP domain, bound with the highest affinity to class II motif (Fig. S17A). This motif differed from the expected class I for this subgroup of TCP proteins and highlights the relevance of the critical residue in the TCP domain in the determination of DNA-binding specificity. We analyzed with more detail binding of TCP16 to class I and class II motifs. Median signal intensity of the probes containing class II motif was significantly higher than the corresponding to class I or mutant versions of the motif (Fig. S17D). However, TCP16 also bound with high affinity to class I elements, whereas binding to mutant versions was almost negligible, as observed for the distribution of the probe intensities (Fig. S17D). Similar results were obtained by Viola et al. (2012; ref. 80) using a SELEX-based approach with a partially degenerated oligonucleotide.

## 10. Heat shock factors

Heat shock transcription factors (HSFs) are a group of proteins highly conserved in their modular structure and mode of DNA recognition throughout the eukaryotic kingdom (81, 82). Regulation of HSF activity depends on their trimerization *in vivo*, which is prevented by several chaperones under non-inductive conditions. HSF trimers bind a DNA sequence referred to as HSE, that consists in a tandem of inverted repeats of the sequence GAA, generating a perfect HSE, TTCnnGAA<sub>n</sub>TTC (83). We obtained target DNA sequences for two different HSFs: HSFB2a (At5g62020) and HSFC1 (At3g24520). Both proteins showed identical DNA binding properties, recognizing with similar affinities two different 8mers, TTCTAGAA and GAAGCTTC

(inverted GAA repeats underlined), which can be interpreted as incomplete perfect HSEs (Fig. S18A). It is noteworthy that our PBMIs (PBM10 and PBM11) used in this work have the limitation of 10 or 11mer sequences. Therefore, due to these technical limitations, a complete HSE can not be determined. Thus, it is likely that the real sequences recognized by HSFC1 and HSFB2a are the combination of both primary and secondary motifs (GAAGCTTC and TTCTAGAA, respectively). These elements and a perfect HSE were over-represented in the promoters of genes co-regulated with *HSFB2a* and *HSFC1*, but not at their corresponding downstream regions, supporting their role as regulatory *cis*-elements (Fig. S18A and Fig. S18B). Moreover, a tripartite HSE was also enriched in these sets of promoters, indicating that transcriptional regulation of *HSFB2a* and *HSFC1* is exerted through the perfect HSE (Fig. S18B).

As expected for a family of TFs highly conserved among eukaryotes, these DNA-sequences were also enriched in the promoters of genes down-regulated in the double mutant *hsf1,3* (84), indicative of similar recognitions patterns of homolog TFs (Fig. S18D). In a survey of the patterns of expression using *in silico* tools (<http://bar.utoronto.ca/>) the expression of *HSFB2a* is maximum at short time points of a heat stress treatment (fold change 12.32, 1 hour heat treatment). By contrast, *HSFC1* is not induced by heat treatments, but highly responsive to cold stress (fold change 17.65, 3 hours cold treatment). Concurrently with these expression patterns, partial and tripartite HSEs recognized by *HSFB2a* and *HSFC1* were over-represented in the promoters of genes induced under these stress conditions (Fig. S18E and Fig. S18F).

## 11. bZIPs

Basic/leucine zipper (bZIP) TFs are present in all eukaryotes and are characterized by a basic region that binds DNA and a leucine zipper dimerization motif. In *Arabidopsis*, there are about 80 genes encoding bZIP proteins classified into several groups defined on the basis of sequence similarity of the basic region and the presence of additional conserved motifs. Plant bZIPs preferentially bind to the A-box (TACGTA), C-box (GACGTC) and G-box (CACGTG), that share the common ACGT core (85). TGA2 and bZIP60 recognized exactly the same motif, TGACGTCA, a palindromic 8mer that contains the C-box (Fig. S19A). Albeit the palindromic 8mer showed the highest scoring, both bZIPs also recognized with high affinity several variants that contain the core sequence TGACGT, although differing at their 3'-end flanking nucleotide (Fig. S19A). In contrast to other bZIPs (85), these proteins did not show a significant binding to G-box. In general, these elements were enriched in the promoters of positively co-regulated genes (Fig. S19B). In addition, the core sequence TGACGT also have a biological relevance at downstream regions of the genes co-regulated with *bZIP60* (Fig. S19C).

We took advantage of a transcriptomic profiling of the triple mutant *tga2,tga5,tga6* (86). These genes act redundantly during establishment of the salicylic acid-dependent defense response against biotrophic pathogens and for the activation of jasmonic acid- (JA-) and ethylene-dependent defense mechanisms that counteract necrotrophic pathogens (87, 88). The three DNA variants (primary palindromic 8mer, secondary motif and core sequence) were highly over-represented in the promoters of genes down-regulated in the triple mutant in response to the JA precursor OPDA, indicating a prominent role of these DNA motifs as *cis*-elements recognized by TGA proteins (Fig. S19D).

*bZIP60* is activated in response to an endoplasmic reticulum (ER) stress (89). We observed high over-representation of the secondary and the core motifs for *bZIP60* in the promoters of genes down-regulated in the loss of function mutant *bzip60* in response to an ER stress (Fig. S19E). Actually, the secondary motif found for *bZIP60* corresponds to the unfolded protein response element (UPRE, TGACGTGR) described for the mammalian XBP1 protein, the counterpart of the *Arabidopsis* *bZIP60* and the rice *bZIP50* TFs (89, 90). Primary motif for *bZIP60* differs at their two 3'-end terminal nucleotides and does not seem to play a major role during ER stress (Fig. S19E). Given that this motif is enriched in the promoters of co-regulated genes, it is plausible that this primary sequence is involved in *cis*-regulation in different growth conditions or with tissue specificity. In this sense, activation of the *bZIP60* protein also occurs in anthers in the absence of ER stress, suggesting that this TF functions in the normal development of active secretory cells (89).

## 12. B3

The B3 superfamily is exclusive from plants and comprises an heterogeneous group of more than 100 genes in *Arabidopsis*. This superfamily includes four major groups of gene families

(91): LEC2/ABI3 (LEAFY COTYLEDON2–ABSCISIC ACID INSENSITIVE3), ARF (AUXIN RESPONSE FACTOR), RAV (RELATED TO ABI3 and VP1) and REM (REPRODUCTIVE MERISTEM). This heterogeneity is reflected in DNA-binding specificities of the coded proteins: LEC2/ABI3 recognize the so called RY motif (92), CATGCA, whereas ARFs bind quite a different sequence, the AuxRE TGTCTC (93). RAV proteins have a B3 and an AP2 DBD, that interact with DNA independently. In this case, the B3 domain of RAV proteins recognizes the element CACCTG (94). Finally, REM proteins are supposed to be transcription factors (95), but the only protein studied so far is VRN1 that interacts with DNA in a non-specific manner (96). In this work we analyzed two proteins included in this superfamily: ETTIN (ARF family) and REM1 (REM family).

### 12.1. ARF

We analyzed DNA binding affinity of ETTIN (ETT, ARF3) and we observed clear binding preferences to the element TGTCGG when computing for 8mers and 6mers. Search for secondary motifs also identified a palindromic sequence containing the core motif TGTCGA (Fig. S20A). These sequences differ from the AuxRE identified for other members of this family (TGTCTC; ref. 93). Actually, ETT binding affinity to AuxRE is notably lower than to the element that we identified with higher affinity, although similar to that of the secondary motif (Fig. S20B), indicating that ETT binds to a different DNA sequence than that described previously for other members of this family. Primary DNA-motif found for ETT was over-represented in the promoters of negatively co-regulated genes (Fig. S20C). This observation agrees with the fact that ETT/ARF3 is a transcriptional repressor (97). By contrast, secondary motif was not enriched. It is worth noting that canonical AuxREs were enriched in positively co-regulated promoters, as expected for a *cis*-element responsive to auxins and likely regulated by other ARFs. In relation to downstream regions of co-regulated genes, all the elements were over-represented, indicating that ETT-mediated transcriptional repression may be exerted through these sequences located downstream the target genes (Fig. S20D).

We extended the analysis of ETT cognate elements to different regulatory pathways involving other ARFs. In particular, we took advantage of the transcriptomic profiles of the mutant *arf2* (98) and the double mutant *arf7, arf19* (99). ARF2 belongs to the group of ARF repressors, and ETT-binding elements were enriched in the promoters of up-regulated genes in *arf2* mutant, whereas the canonical AuxRE does not seem to play a major role in ARF2-dependent gene expression, at least in etiolated seedlings (Fig. S20E). Similarly, primary ETT-binding motif was enriched in the promoters of down-regulated genes in double mutant *arf7, arf19*, as expected for the loss-of-function of two transcriptional activators (Fig. S20F). Given the well known role of ARF proteins in auxin signaling, we also analyzed the promoters of genes induced in response to an IAA treatment. In this case, both ETT-response elements, as well as the AuxRE were significantly enriched in these sets of promoters (Fig. S20E). Altogether, these data indicate that different ARF proteins might be recognizing the novel elements found for ETTIN, and binding seems to be independent on the activating or repressing activity of the ARFs involved. In addition, these binding motifs have a role during auxin signaling, as AuxRE also does, indicating that DNA binding specificity of ARF may be broader than the initially described.

### 12.2. REM

We determined DNA binding specificity for REM1, included in the REM class. This protein interacts with DNA sequences containing the core motif TGTAG, and evaluation of secondary motifs always contained this core sequence (Fig. S21A). This element represents a novel cognate sequence for B3 proteins that differs from the consensus binding motifs described for other groups belonging to the same superfamily. Moreover, it provides specificity for DNA-binding of a REM domain-containing protein, given that VRN1, the only REM-containing protein studied so far, interacts with DNA in a non-specific manner (96). In relation to REM1 cognate elements, we observed over-representation of all the elements containing the core motif in the promoters of co-regulated genes but not at their downstream regions (Fig. S21B and Fig. S21C), supporting the role of these DNA-motifs as regulatory *cis*-elements and, hence, of REM1 as a sequence-specific transcriptional regulator.

## 13. NAC

NAC family is a plant-specific family of TF that comprises more than 100 members in Arabidopsis, and have been involved in several developmental or stress-related responses (100, 101). NAC TFs bind DNA elements as dimers and their cognate motifs are bipartite palindromic sequences which, in most cases, possess weak base requirements in each position (101). Globally, two different NAC recognition motifs have been discovered to date: NAC recognition site (NACRS; ref. 102), that contains the core sequence CGT[G/A]; and secondary wall NAC binding element (SNBE; ref. 103) with a longer and variable sequence ([T/A]NN[C/T][T/C/G]TN>NN>NNNA[A/C]GN[A/C/T][A/T]). We determined DNA binding specificity for ANAC55, ANAC46 and ANAC58 and obtained similar DNA sequences that contain the conserved core CGT[G/A] (Fig. S22A), with lower base requirements outside this core motif. This is clear in the case of ANAC55, for which we found two motifs with similar affinities differing at their two 3'-end nucleotides (Fig. S22A).

The elements found for these NACs were over-represented in the promoters of their corresponding positively co-regulated genes, suggesting that these TFs act as transcriptional activators (Fig. S22B). Moreover, both motifs found for ANAC55 differing at their 3'-end positions were enriched in these promoters (Fig. S22B). No significant enrichment at downstream regions was observed for any motif (Fig. S22C). ANAC55, together with ANAC19, plays a dual role in regulating JA and ABA responses, acting as components of the crosstalk between these signaling pathways (104, 105). Primary element found for ANAC55 is enriched in the promoters of JA- and ABA-activated genes (Fig. S22D). However, secondary element differing at the 3'-end is related with JA-dependent gene expression and seems to be irrelevant in the ABA signaling pathway (Fig. S22D). Thus, DNA-binding specificity for partially different *cis*-elements may contribute to discriminate these overlapping signaling pathways.

#### 14. AHL

AHL proteins contain a single AT-hook DNA binding domain that characterizes a group of approximately 30 members in Arabidopsis. AHLs are chromatin-associated proteins that bind to A/T-rich stretches in the minor groove of DNA and are thought to co-regulate transcription by modifying the architecture of DNA (106, 107). Several AHLs regulate expression levels of target genes by binding to the *cis*-element on the target promoters, suggesting that binding to A/T-rich motifs is specific. We analyzed three AHL proteins, AHL20, AHL12 and AHL25 and, as expected for this group of TFs, recognized A/T-rich motifs (Fig. S23A). The three proteins showed similar binding affinities for several A/T-containing elements, difficulting the definition of a consensus sequence for this family (Fig. S23A). In spite of this, AHLs may retain specificity for certain A/T-rich elements in the promoters of targets genes. In line with this, the three top scoring 8mers found for AHL20 were enriched in the promoters of co-regulated genes with AHL20, indicating that this protein recognizes specific *cis*-elements in the promoters of target genes (Fig. S23B). In relation to AHL12, top-scoring 8mers were over-represented in the promoters of negatively co-regulated genes, suggesting that AHL12 may contribute to down-regulation of target genes, rather than activation (Fig. S23B).

The AHL protein GIK contributes to down-regulation of target genes, what requires the presence of matrix associated regions in the genomic sequences of the target genes, which serve as binding motifs for the AT-hook DNA binding protein (108). Similarly, AHL25 (also known as AGF1) recognizes specific *cis*-elements in the promoter of *AtGA3ox1* and is responsible for the GA-negative feedback of this gene (109). AHL20 has been implicated in the regulation of plant innate immunity and in JA-response (110 111). In particular, a *cis*-element recognized by several AHL proteins locates in the promoter of one JA-inducible gene and is responsible for quantitative response to the hormone. In line with this, the elements recognized by AHL20 were over-represented in the promoters of genes early induced after treatment with JA, pointing to a role of this TF in the co-activation of target genes (Fig. S23D).

#### 15. bHLH

The basic helix-loop-helix family of TFs is one of the most abundant in plant genomes and is highly conserved in all the kingdoms (112). Broadly, this family of proteins bind the E-box (CANNTG), although most plant bHLHs specifically recognize the so-called G-box, CACGTG. In this work, we analyzed PIF3 and found to recognize a perfect G-box, in which certain nucleotides at flanking positions are preferred (Fig. S24A). Evaluation of secondary motifs

always yielded G-containing elements, suggesting that PIF3 binds with high affinity to all G-box containing sequences whereas flanking positions are less important for recognition. Recently, Zhang et al. (ref. 113) observed that PIF3 binds *in vivo* mainly to G-box but also to PBE-box (CACATG) motifs, although the apparent affinity *in vitro* for the G-box was higher than for the PBE-box. In our binding assays, affinity for G-elements was much higher than for the PBE-box (median E-score 0.486 versus 0.228, p-value 1.809e-12 Wilcoxon test), although we cannot exclude binding of PIF3 to certain PBE-containing motifs (Fig. S24B). Similarly, we did not observe interaction of PIF3 to any of the G-related sequences that have been reported for other bHLH proteins (Fig. S24B; 32).

G-boxes recognized by PIF3 were over-represented in the promoters of co-regulated genes and, to a lesser extent, the PBE-box (Fig. S24C). This indicates that, in spite that PIF3 showed poor binding to PBE-box *in vitro*, this element may participate in PIF3-mediated transcriptional activation *in vivo* in which other PIFs may have an overlapping function (113, 114). We took advantage of the information about PIF3-bound genomic regions (113) and examined the presence of sequence motifs identified *in vitro*. We focused on PIF3-bound fragments with a peak-summit located within 1 kb upstream the transcriptional start site of a gene. All the elements were enriched in this set of promoters but over-representation of the G-box over the random frequency was much higher than that corresponding to the PBE-box (Fig. S24D). In relation to flanking nucleotides of the G-box, the variant found to be recognized by PIF3 with the highest affinity is also the most representative among PIF3-bound fragments (Fig. S24D). These results perfectly reproduce binding affinities found for PIF3 *in vitro*.

In previous works, we determined DNA binding specificity of several bHLH proteins, MYC2, MYC3, MYC4, PIF4 and PIF5 (32, 114, 115). All the proteins recognized as the top-scoring motif the G-box (CACGTG). However, MYC2 and, to a lesser extent, MYC3 recognized with high affinity other G-related elements, whereas DNA binding of MYC4 was more restricted (32, 115). Similarly, PIF4 recognized with high affinity the PBE-box, in addition to the G-box, and DNA-binding of PIF5 was almost exclusively to the G-box (114). We analyzed sequence enrichment of these DNA sequences in the promoters of their corresponding co-regulated genes. In all the cases, the G-box was over-represented in these sets of promoters, demonstrating the prominent role of this motif is *cis*-regulation mediated by these TFs (Fig. S25A). In the case of MYC2, we also obtained significant enrichment of all the DNA elements identified *in vitro* with lower affinity than the G-box (Fig. S25A). MYC3 and MYC4 act redundantly, together with MYC2, in the regulation of gene expression during jasmonic acid-signaling. We obtained significant enrichment of two “secondary” elements (out of four) corresponding to MYC3 and one (out of two) for MYC4 (Fig. S25A). In relation to PIF4, the secondary PBE-box was also over-represented in the promoters of co-regulated genes and this was not observed for PIF5, as expected given its poor binding affinity for this DNA-sequence (Fig. S25A). In relation to downstream regions of co-regulated genes, some of the DNA motifs were enriched (Fig. S25B). This was particularly notable for MYC2, where the top three motifs (G-, PBE- and T/G-elements) were significantly enriched (Fig. S25B). The G-box was also over-represented in data relating PIF5, whereas the secondary motif for MYC4 (PBE-box) also did (Fig. S25B).

## 16. LOB/AS2

The LOB/AS2 domain defines a class of plant-specific transcription factors that contains a motif resembling a zinc finger and another similar to leucine zipper. It has been demonstrated the ability of the proteins LOB, LBD4 and AS2 to bind DNA *in vitro* through the LOB-motif to DNA elements containing the core sequence CGGC (116). We tested DNA-binding properties of LBD16/ASL18 and obtained a palindromic motif slightly different from that described for other LOBs but with some core nucleotides conserved (TCCGGA; Fig. S26A). In a site directed mutagenesis *in silico*, we observed that modifications of the core sequence TCCGG reduce dramatically binding of the protein (Fig. S26B). Moreover, we did not observe significant binding of LBD16 to the LOB-motif described for other LOB proteins (Fig. S26B). These differences in DNA recognition may be explained by particular specificities among proteins of the same group. Alternatively, other residues outside the LOB-domain might be influencing binding to DNA, given that the LOB-motif described for the protein LOB was obtained with a truncated version of the TF containing the LOB-domain (116).

DNA element for LBD16 was slightly over-represented in the promoters of negatively co-regulated genes, suggesting that this protein might act as a transcriptional repressor (Fig.

S26C). By contrast, and consistently with our binding affinities *in vitro*, the LOB-motif described for LOB was not significantly enriched in the same sets of promoters (Fig. S26C). None of the elements tested were enriched at downstream regions of co-regulated genes (Fig. S26D). Data relating co-regulated genes suggest that LBD16 represses transcription. However, LBD16, and the homolog and partially redundant LBD18, activate transcription in both yeast and Arabidopsis protoplasts when fused to the Gal4 DNA-binding domain (117). In addition, overexpression of a chimeric version of LBD16 fused to a transcriptional repressor domain has the opposite effect than overexpression of the native version (118), indicating that LBD16 activates transcription. Given the discrepancy between these results, we cannot exclude the existence of other factors affecting the activity of LBD16 *in planta*, resulting in apparent repression instead of activation of gene expression. We analyzed transcriptomic data involving a translational fusion of LBD18 with the glucocorticoid receptor (GR). In this system, activation of the TF only occurs upon treatment with dexamethasone (118). We obtained a significant enrichment of the DNA motif found for LBD16 (TCCGGA) in the promoters of genes up-regulated upon activation of the TF, whereas the motif identified for LOB protein was not enriched (Fig. S26E). This indicates that LBD18 might be activating gene expression through recognizing the LBD16-congate sequence. Similarly, the motif TCCGGA is over-represented in the promoters of genes up-regulated in plants overexpressing a translational fusion LOB-GR (119; Fig. S26F), indicating that the binding motif that we obtained for LBD16 could be also recognized by other similar proteins.

## 17. SHI/STY

The plant-specific SHI/STY gene family code a group of proteins that contain a zinc-finger similar to RING domains. These genes act redundantly in the regulation of development of the gynoecium, stamen and leaf (120-122). We obtained for STY1 a palindromic motif that contains the sequence CCTAGG (Fig. S27A) and evaluation of subsequent DNA variants yielded DNA motifs containing the core sequence CTAG or even two direct repeats of the same core sequence (CTAGCTAG, Fig. S27A). This core sequence was shown also to be essential for recognition of the protein, given that modifications of this sequence decreased the binding affinity (Fig. S27B). STY1 stimulates the synthesis of auxins via the transcriptional activation of *YUC4*, through an identified putative *cis*-element, ACTCTAC. The motif that we found for STY1 differs considerably from that described previously. However, both elements share the trinucleotide CTA, shown to be essential during transactivation assays in yeast (120).

The gene STY1 is not represented in the ATH1 microarray, but we obtained the genes co-regulated with the close homolog and redundant STY2. DNA motif found to be recognized by STY1 is slightly over-represented in the promoters of genes co-regulated with STY2, as expected for two similar and redundant TFs (Fig. S27C). Similarly, a secondary motif containing two repeats of the core sequence CTAG was also enriched in the promoters of STY2 co-regulated genes. The element ACTCTAC was not enriched in these promoters, indicating that its role as a *cis*-regulatory element may be restricted to a limited number of targets. Primary motif was not over-represented at downstream regions, whereas secondary motif was enriched in the same set of promoters (Fig. S27D).

## 2. Text S2. Material and Methods

**Bacterial expression of tagged-TFs.** A list of full-length TFs were selected from a library of ordered *Arabidopsis* cDNAs cloned in the destination vector pER8 (123) and mobilized into pDONR201 using BP clonase reaction. cDNAs were transferred to destination vector pDEST-TH1 (124) using LR clonase yielding Maltose Binding Proteins (MBP) N-terminal fusions and constructs checked by sequencing. MBP-TF constructs were transformed into BL-21 strain for expression. Induction of bacterial cultures was routinely at 25°C for 6 h with 1 mM Isopropyl β-D-1-thiogalactopyranoside. Expression of recombinant proteins was assessed by PAGE followed by Coomassie staining and parallel gels were transferred to nitrocellulose membranes for probing with an anti MBP antibody (BioLabs).

**Identification of TF-binding motifs using Protein Binding Microarrays (PBMs).** In this work, we used three different designs: PBM11, PBM10 and nPBM11. PBM11 and PBM10 were previously described (32, 125). nPBM11 basically consists on the same design as PBM11 containing 167,773 different oligonucleotide probes as PBM11 but synthesized in an Agilent's SurePrint G3 4x180k format. All the microarrays were manufactured by Agilent Technologies. Recombinant protein extracts were obtained from 25 mL induced *E. coli* cultures as above. Pelleted cultures were resuspended in 1 mL 1x binding buffer, sonicated (2x30 s) and centrifuged twice to obtain cleared extracts of soluble proteins as in Ref. 11. The binding mixture was adjusted to 175 µl and contained 2% milk and 0.89 µg of denatured salmon sperm DNA (ssDNA). Synthesis of double-stranded microarrays, protein incubations and immunological detections of DNA-protein complexes were as described (32, 125). In experiments corresponding to proteins containing Zn-coordinating domains, binding mixtures and washing solutions were supplemented with zinc acetate at 50 µM.

Throughout this work we used two different scanners and quantification softwares. In the case of GenePix 4000B scanner (Axon Instruments), microarrays were scanned at 5 µm resolution and quantified in the GenePix Pro 5.1 software. Protein-bound microarrays were routinely scanned at three PMT conditions, images quantified separately and combined in a common linear scale using Masliner software (126). Most of the microarrays were scanned in a DNA Microarray Scanner at 5 µm resolution and quantified with Feature Extraction 9.0 software (Agilent Technologies). In test experiments, we did not observe noticeable differences in the results obtained with different scanners. Normalization of probe intensities and calculation of E-scores of all the possible 8-mers were carried out with the PBM Analysis Suite (125). Perl scripts were modified to adapt them to different microarray dimensions and different input files generated by GenePix Pro 5.1 or Feature Extraction softwares.

For all the proteins we selected the top-scoring 8-mer as the primary DNA-element recognized by the TF. Secondary or tertiary DNA-motifs were selected if they appeared among top 15 top-scoring motifs and if they differed substantially from their corresponding primary ones. Moreover, a systematic search for secondary motifs was carried out by running the Rerank program in the PBM Analysis Suite (125). In this case, most of the 8-mers initially selected as secondary motifs appeared as top-scoring elements after re-ranking the probes. A list of DNA motifs with E- and Z-scores, their rank positions, PWMs and microarray designs used is in Table S2. Graphical representations of PWMs were obtained with the on-line tool enoLOGOS (<http://www.benoslabs.pitt.edu/cgi-bin/enologos/enologos.cgi>).

**Analysis of TF-coding genes co-expression data.** We obtained the lists of top 200 genes positively co-regulated (positive Pearson's correlation coefficient) and negatively co-regulated (negative Pearson's correlation coefficient) with TF-coding genes from Genevestigator (127). This database contains 8,793 samples ("Samples" dataset), corresponding to over 104 tissue types ("Anatomy" dataset) and 2,213 experimental conditions ("Perturbations" dataset), as in June 2013. Only unambiguous probe sets corresponding to nuclear-coded genes from the top 200 were considered. It has to be noted that since in all comparisons the top 200 correlating genes were selected, Pearson's correlation coefficients vary among different gene sets.

Identification of DNA motifs in the promoter or downstream regions of co-expressed genes was performed by searching DNA patterns with the Patmatch tool (<http://www.arabidopsis.org/cgi-bin/patmatch/nph-patmatch.pl>) in the datasets TAIR10 Loci Upstream Sequences-1,000bp and TAIR10 Loci Downstream Sequences-1000bp. As a control, we searched the same motifs in upstream and downstream regions of all the genes unambiguously represented in the ATH1 microarray. Statistical over-representation of DNA-

motifs was evaluated by comparing the proportion of co-expressed genes containing the motif relative to the corresponding proportion in the complete microarray, following a hypergeometric distribution. For each TF-coding gene, we analyzed co-expressed genes obtained from “Samples”, “Anatomy” and “Perturbations” datasets and chose those giving lowest p-value. Datasets used and the range of Pearson’s correlation coefficients of the co-expressed datasets are shown in Table S1.

Analysis of enrichment of singular Gene Ontology (GO) terms was performed on GeneCodis (128; <http://genecodis.cnb.csic.es/>) using the list of unambiguous genes represented in the ATH1 microarray as the reference list. We selected GO terms significantly (FDR-corrected p-value<0.05, hypergeometric test) enriched in both lists of co-regulated genes (CRGs) and their corresponding subsets of co-regulated genes with element (CRGEs).

**Differential gene expression of TF-involving genotypes.** Gene expression experiments were identified from the literature or from public repositories (GEO: <http://www.ncbi.nlm.nih.gov/geo/>; ArrayExpress: <http://www.ebi.ac.uk/arrayexpress/>). A list of gene expression experiments analyzed can be found in Table S4. Whenever available, raw microarray data were downloaded and reanalyzed for identification of differentially expressed genes. Routinely, microarrays were corrected for background, normalized with RMA and differentially expressed genes identified following a linear model method. All these steps were performed with the Limma package in Bioconductor (129). Identification of DNA motifs in the promoter or downstream regions of differentially expressed genes was performed as indicated above.

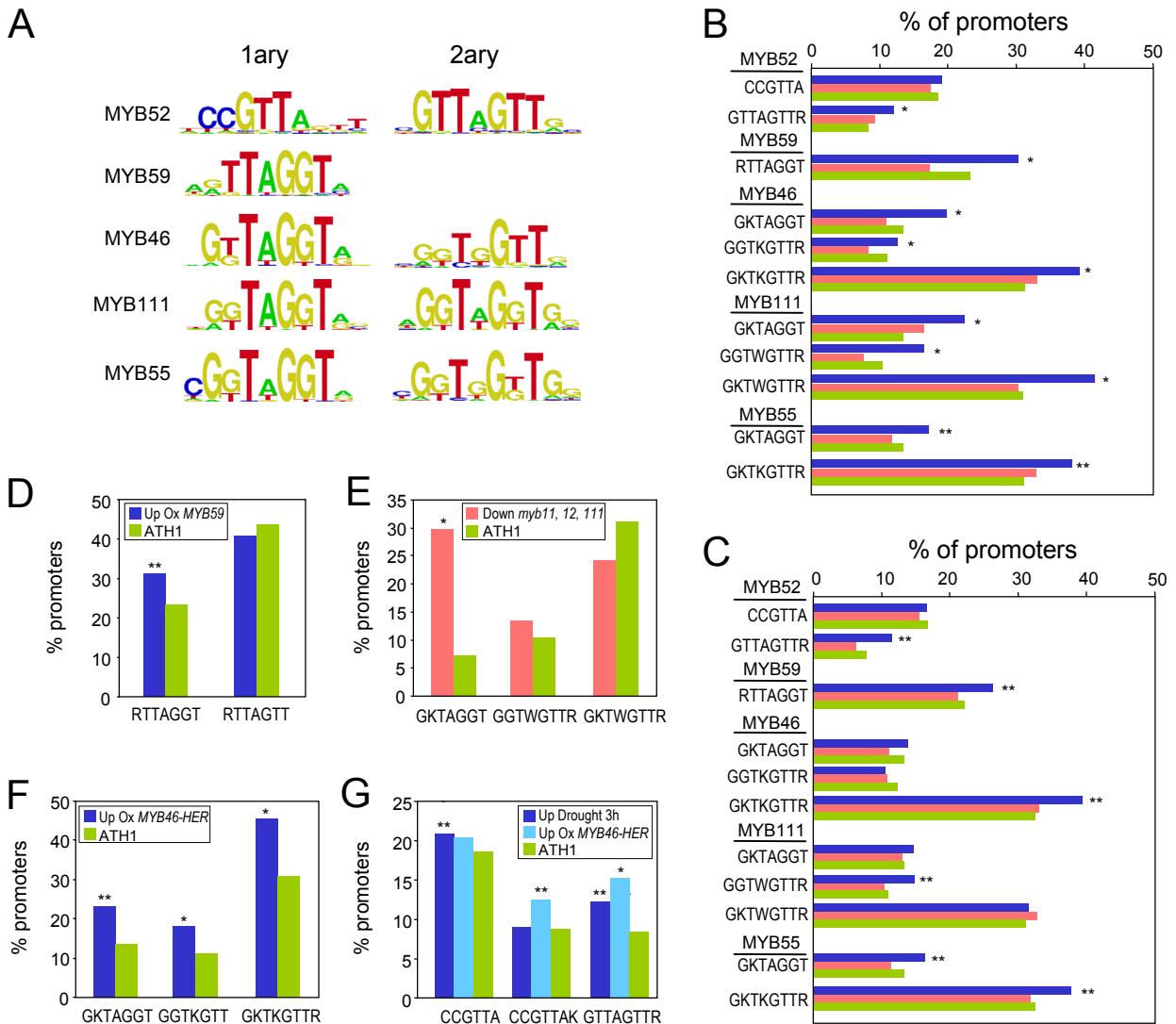
**Mapping of TF-binding matrices.** Genomic DNase I hypersensitive regions (DHS) from leaf and flower tissues were obtained from ref. 130 as BED files. Their corresponding sequences were scanned for the TF-binding matrices in RSAT. The parameters for scanning were pvalue<10e-4 and background sequences inferred from the input with a Markov order 2. As negative controls, we scanned 100 shuffled matrices for each TF-binding motif in the input sequences using the same parameters. For calculation of TF-binding motif densities, mapped motifs were converted to frequency histograms in relation to the number of sequences with that length. In the case of negative controls, the average of the 100 frequency histograms derived from shuffled matrices was obtained. Average histograms for all the TF-binding sites were obtained and plotted with GNUPLOT 4.6 (<http://www.gnuplot.info/>). Upstream and downstream regions of genes represented in the ATH1 microarray were scanned in RSAT using the same parameters as above and frequency histograms obtained for the complete, the positively and the negatively co-regulated gene sets for each motif. Average frequency histograms for all the TF-binding sites were obtained and plotted with GNUPLOT 4.6.

**Clustering analysis of top-ranking 8mers.** E-scores for all the possible 8mers corresponding to the proteins of the MYB and AP2/EREBP structural classes were ranked from 1<sup>st</sup> to 32,896<sup>th</sup> and top 50 ranking 8mers for each protein selected. We obtained two independent matrices corresponding to MYB and AP2/EREBP superfamilies, in which rows corresponded to ranking positions of the 8mers selected and columns to the different proteins analyzed. Ranking matrices were used for hierarchical clustering analysis using Pearson correlation and average linkage. Clustering analysis were performed with MultiExperiment Viewer v4.9 (<http://www.tm4.org/mev.html>). Phylogenetic relationships were assessed by alignment of DNA-binding domains of the TFs belonging to the same structural class using CLUSTALW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Phylogenetic tree (neighbour-joining without distance corrections) and amino-acid identities were as obtained in CLUSTALW2.

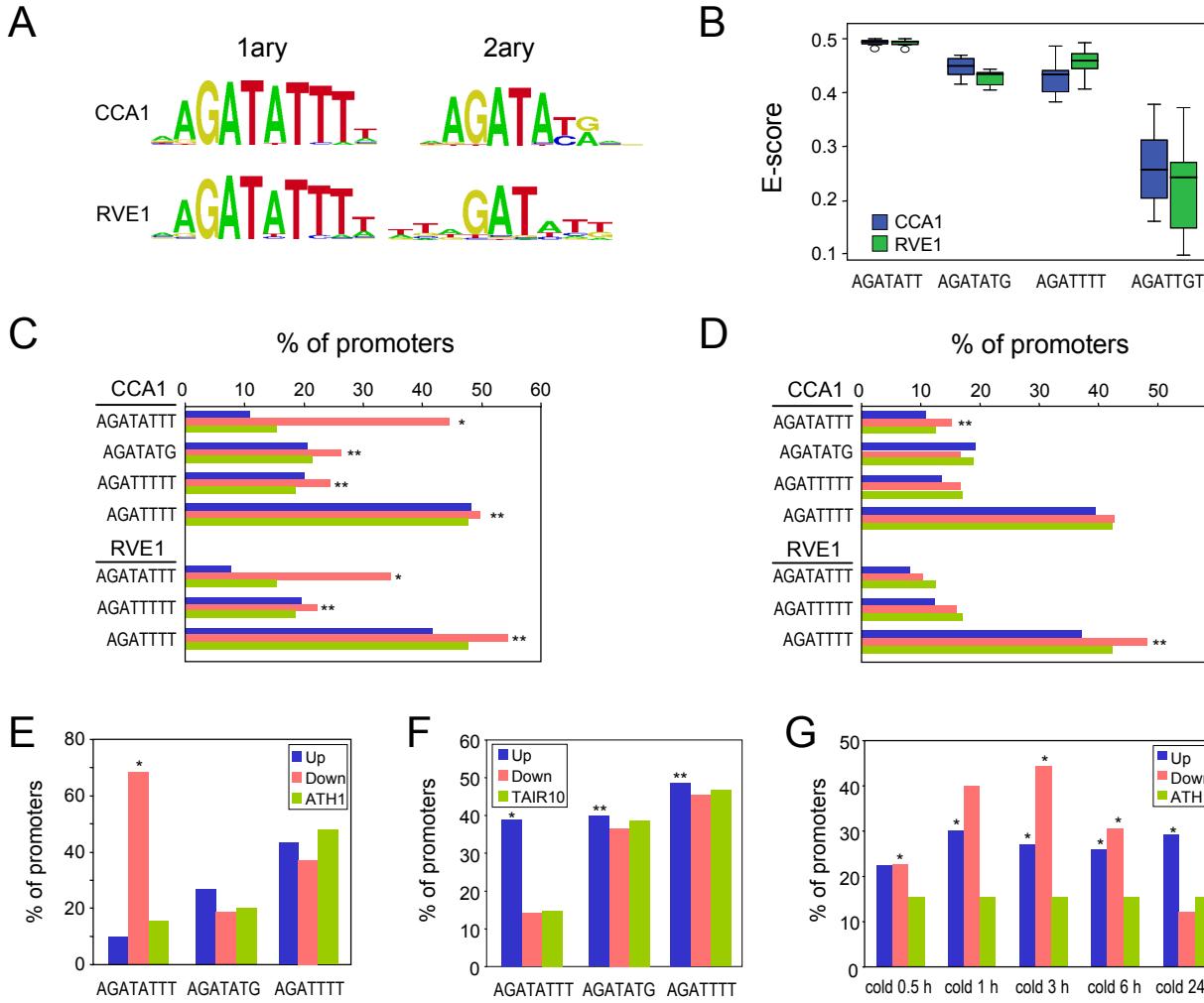
**In planta assays of transcriptional repression.** Promoter regions of one putative target of AHL12, Dof5.7, At5g28300, LBD16 and ATHB51 were transiently expressed in *Nicotiana benthamiana* leaves. Putative targets were selected among negatively co-regulated genes containing at their 5' upstream regions the binding motifs recognized by their regulatory TFs. Promoter regions were PCR-obtained, inserted into pDONR201 (Invitrogen) and then mobilized to the reporter vector pGWB435 (131) to generate promoter-LUC reporter constructs. Effector constructs were obtained by mobilizing the coding regions of the TFs of interest from pDONR201 to destination vector pGWB15 (131) to generate 35S-3xHA-TF fusions. Details of the oligonucleotides used and the constructs are provided in Table S5.

Binary constructs were transformed into the C58C1 strain of *Agrobacterium tumefaciens* and transiently expressed in *N. benthamiana* leaves using the agroinfiltration method as in Ref.

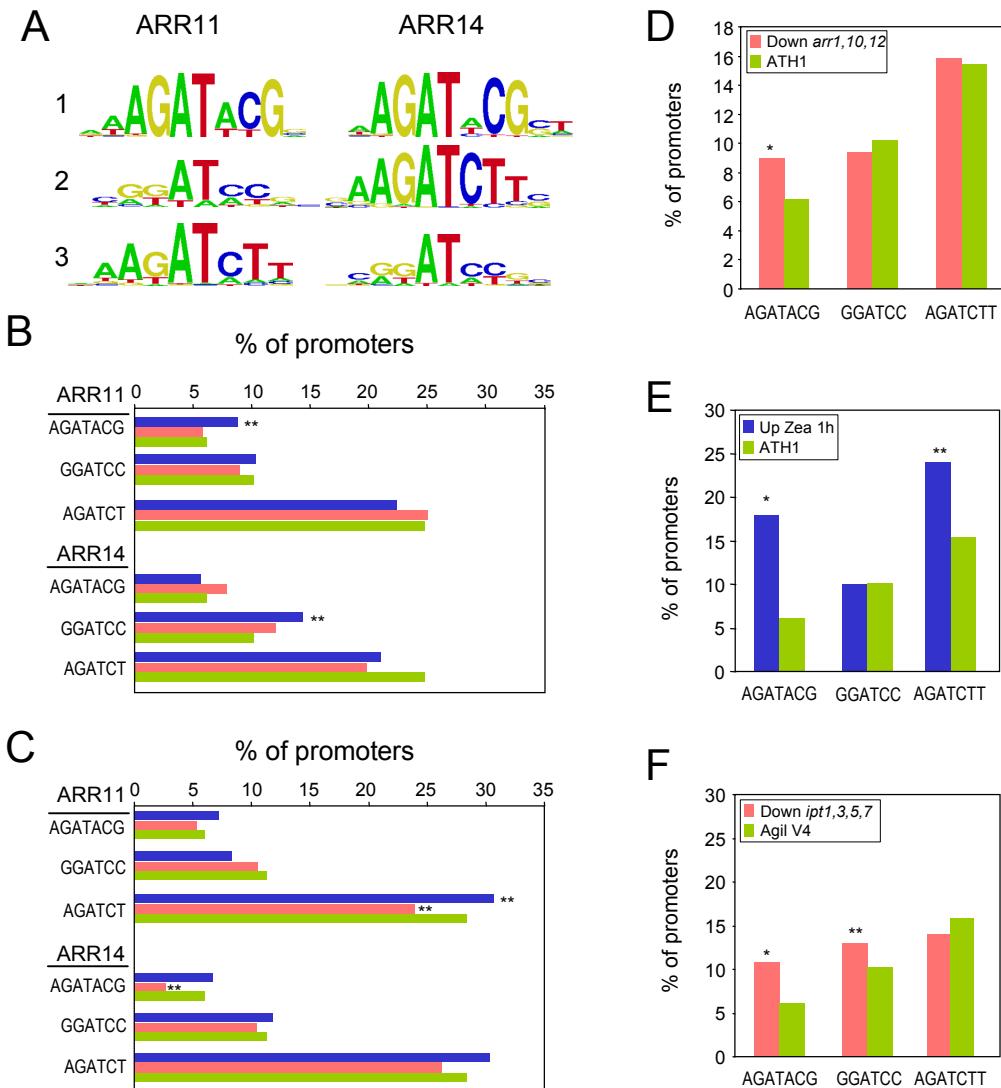
132. Two discs from infiltrated leaves for each of the promoter/TF or promoter/empty vector combinations were obtained 24 hours after infiltration and incubated in 0.02 mg ml<sup>-1</sup> D-luciferin (Sigma-Aldrich) and luciferase activity recorded every hour with a microplate luminometer LB 960 (Berthold Technologies) and further analyzed with the Mikrowin 2000 software. All the experiments were performed in infiltrated leaves from three independent plants.



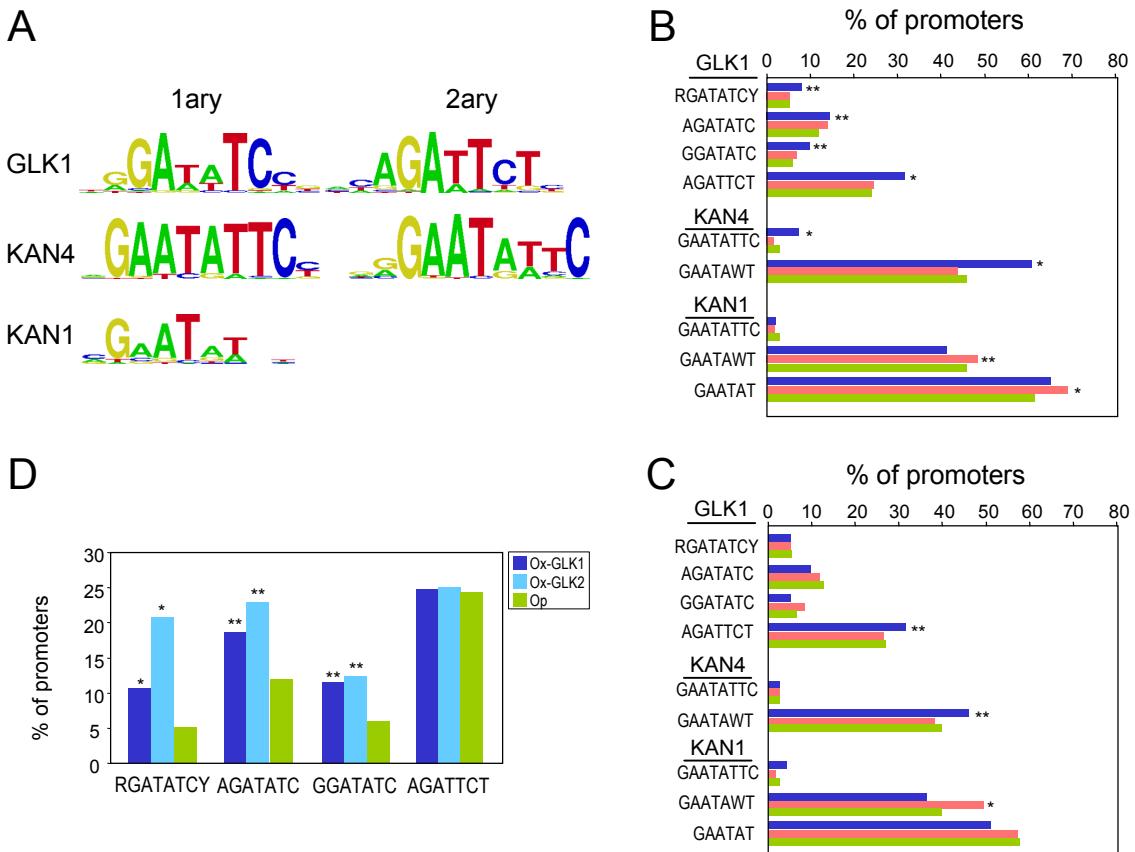
**Fig S1.** DNA-binding specificity of MYB R2R3 TFs. (A) Position weight matrix (PWM) representation of primary and secondary motifs found for the TFs indicated. (B) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with *MYB52*, *MYB59*, *MYB46*, *MYB111* and *MYB55* containing the DNA-elements indicated at their promoter regions (1 kb). Proportions of genes represented in the ATH1 microarray containing the corresponding elements, and thus representing a random distribution, are represented as green bars. Here and after, asterisks indicate the degree of statistical significance in the differences of the proportions indicated (following a binomial distribution) as follows: \* $p$ -value<0.005; \*\* $p$ -value<0.05. Degenerate positions in DNA-elements are as follows: R, A or G; K, G or T; W, A or T. (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with *MYB52*, *MYB59*, *MYB46*, *MYB111* and *MYB55* containing the DNA-elements indicated at their downstream regions (1 kb). Statistical significance is as in (B). (D) Percentage of genes up-regulated in plants overexpressing *MYB59* containing each of the elements indicated at their promoter regions (1 kb). The percentage of genes present in the ATH1 microarray containing the same DNA-sequences at their promoters (1 kb) is represented as green bars. (E) Same as in (D) but data relating genes down-regulated genes in triple mutant *myb11*, *myb12*, *myb111* (red bars). (F) Same as in (D) relating genes up-regulated in *MYB46-HER* overexpressing plants. (G) Same as in (D), relative to genes up-regulated after 3 hours of a drought stress (dark blue bars), and genes up-regulated in *MYB46-HER* overexpressing plants (light blue bars; same data set as in F).



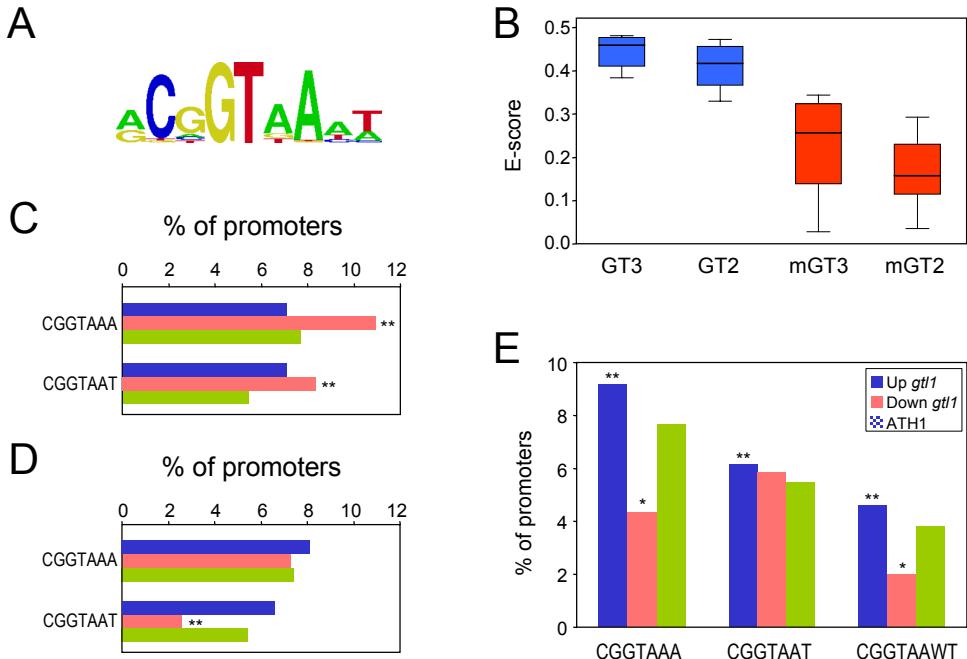
**Fig. S2.** DNA-binding specificity of MYB-related CCA1 and RVE1. (A) PWM representation of primary and secondary motifs found for CCA1 and RVE1. (B) Box-plot of enrichment scores (E-scores) of the elements indicated. Top-scoring motif for both proteins corresponds to the EE (AGATATT). In the graph are also represented secondary motif for CCA1 (AGATATG), secondary for RVE1 (AGATTTT) and a CBS variant (AGATTGT) that is not recognized by either protein. Note that the secondary motif for RVE1 is a sequence variant of CBS (AGATT[T/G]TT). (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with CCA1 and RVE1 containing the DNA-elements indicated at their upstream (C) and downstream (D) regions. Green bars represent percentages of the same elements in the ATH1 microarray. (E) Percentage of genes up-regulated (blue bars) and down-regulated (red bars) in the double mutant *cca1, lhy* containing the elements indicated at their upstream regions. Proportions of genes represented in the ATH1 microarray are indicated with green bars (ATH1). (F) Same as in (E) with data corresponding to direct targets of RVE8 in a transcriptomic analysis of pRVE8:RVE8-GR transgenic plants. (G) Same as in (E) with data relating a time-course experiment of cold-treated plants. Asterisks denote statistical significance as in Fig. S1.



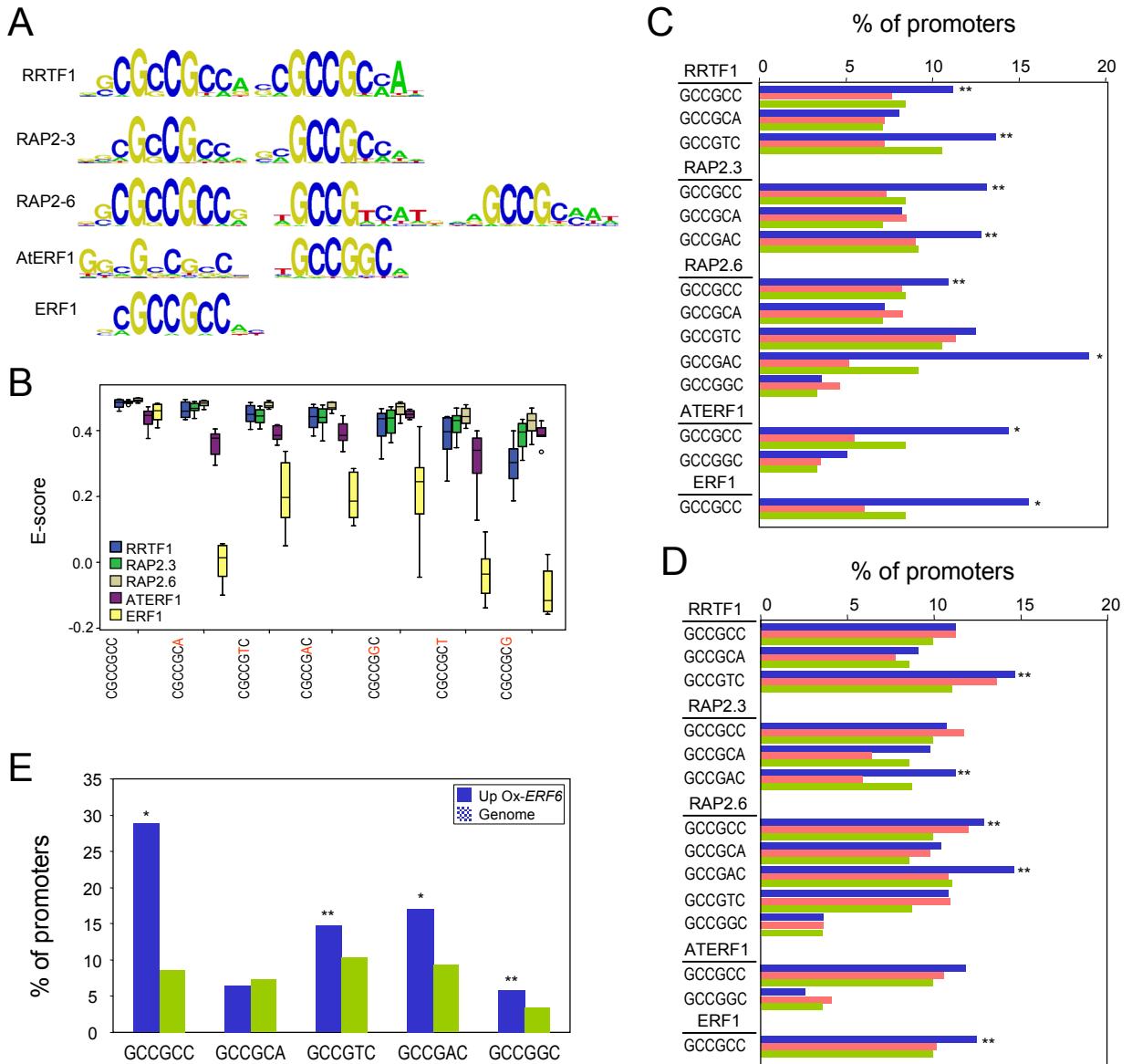
**Fig. S3.** DNA-binding specificity of GARP MYBs ARR11 and ARR14. (A) PWM representation of primary, secondary and tertiary motifs found for ARR11 and ARR14. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with *ARR11* and *ARR14* and genes represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. (D) Percentage of genes down-regulated (red bars) in the triple mutant *arr1*, *arr10*, *arr12* containing the elements indicated at their upstream regions. Proportions of genes represented in the ATH1 microarray are indicated with green bars (ATH1) (E) Percentage of genes up-regulated after an hormonal treatment with zeatin (*Zea*) for 1 hour (blue bars). (F) Same as in (D) with data relating a transcriptomic assay of quadruple *ipt1,3,5,7* mutant, defective in cytokinin biosynthesis. Asterisks denote statistical significance as in Fig. S1.



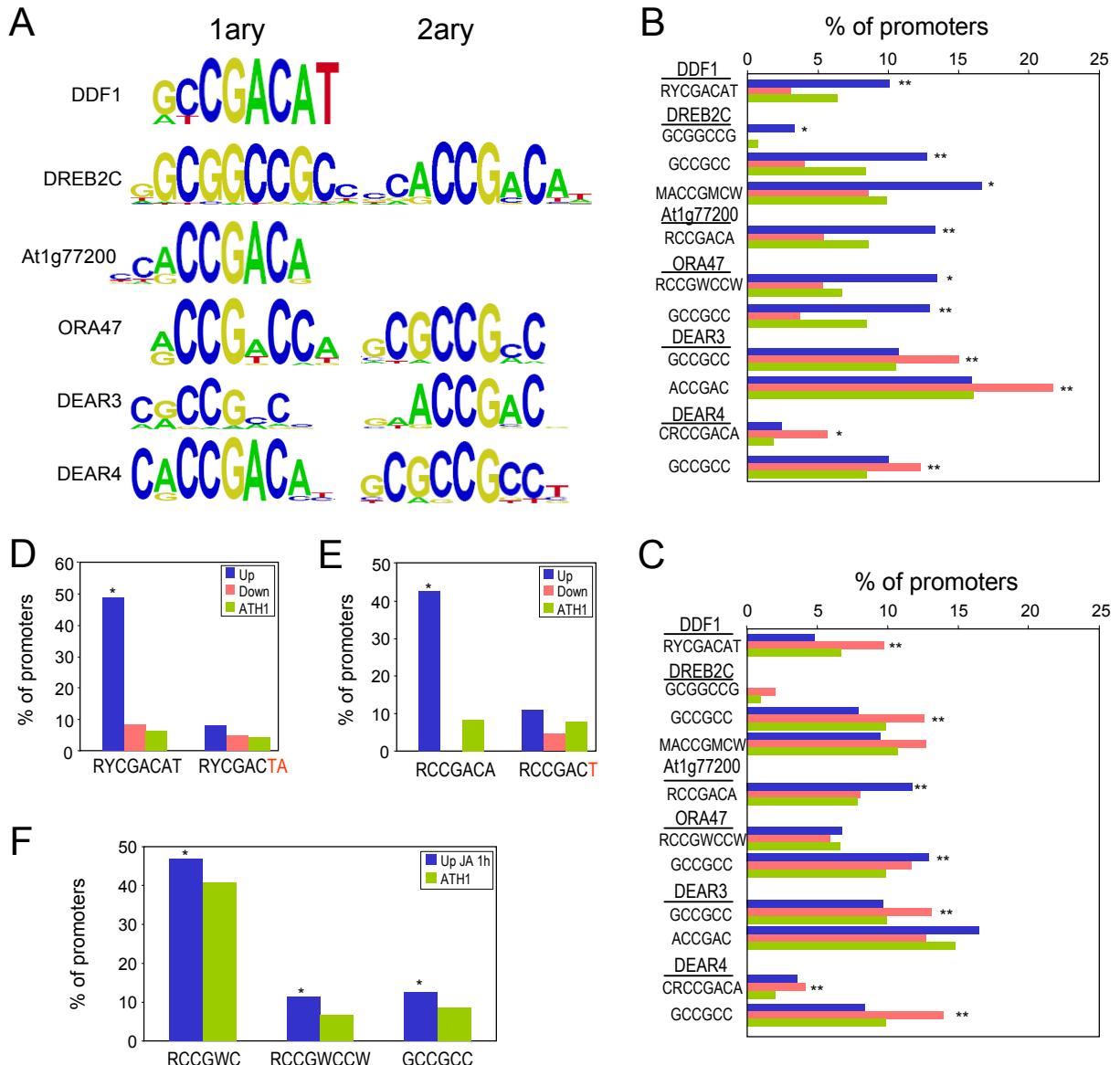
**Fig. S4.** DNA-binding specificity of GARP MYBs GLK1, KAN4 and KAN1. (A) PWM representation of primary and secondary motifs found for these proteins. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with *GLK1*, *KAN4* and *KAN1* and genes represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. (D) Proportions of genes up-regulated in transgenic plants overexpressing either *GLK1* (dark blue) and *GLK2* (light blue) containing the elements indicated at their promoter regions (1kb). Proportions of genes represented in the Operon microarray are indicated with green bars (Op). Asterisks denote statistical significance as in Fig. S1.



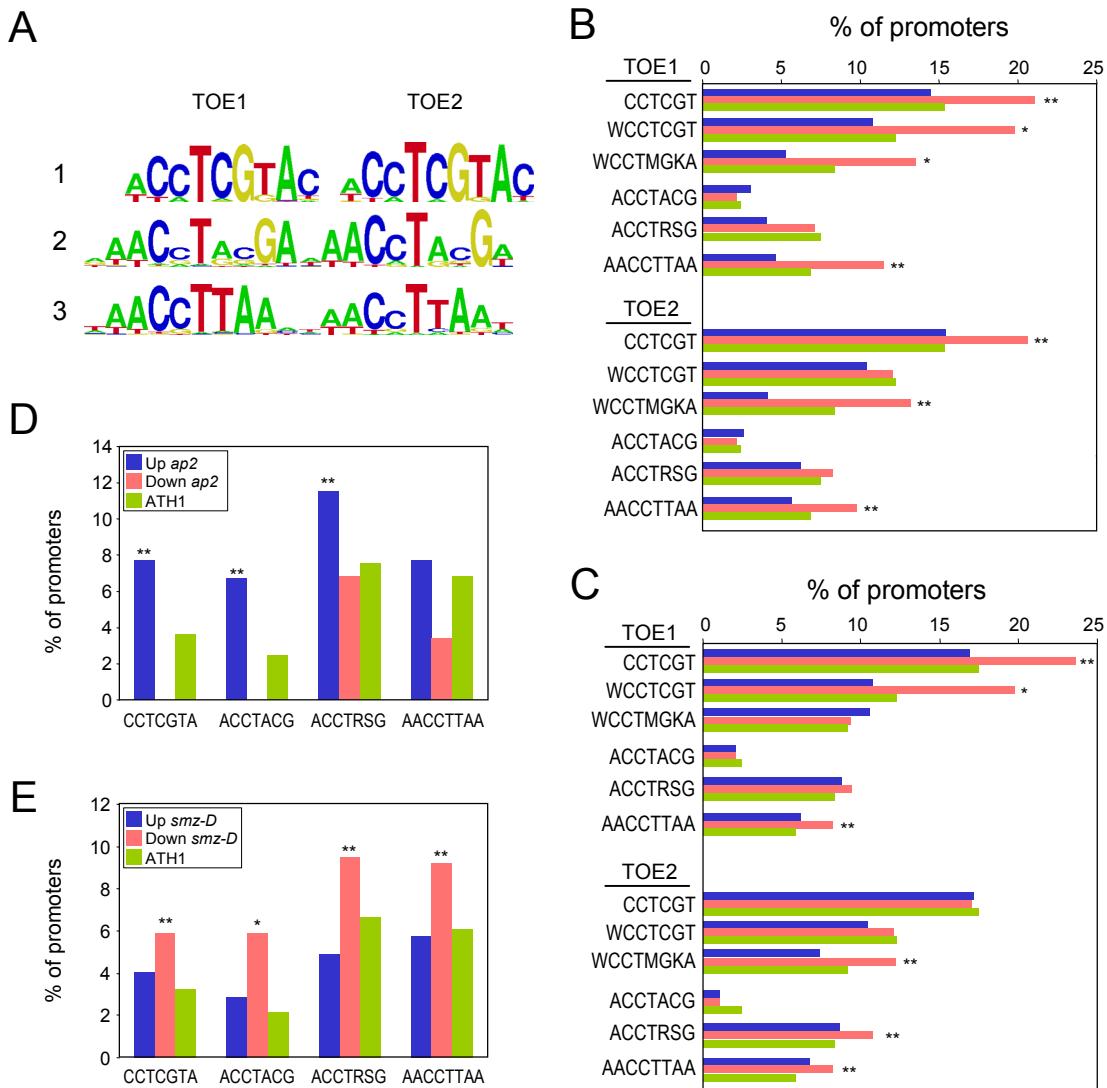
**Fig. S5.** DNA-binding specificity of trihelix TF At5g28300. (A) PWM representation of DNA motif obtained for At5g28300. (B) Box-plot of E-scores of the elements indicated. In the Figure are shown the elements GT3 (CGGTAAA) and GT2 (CGGTAAT), as well as two corresponding mutant derivatives, mGT3 (CGGTATA) and mGT2 (CGGTATT). (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with At5g28300 and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. (E) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in the mutant *gtl1* containing the elements indicated. Proportions of genes represented in the ATH1 microarray are indicated with green bars (ATH1). Asterisks denote statistical significance as in Fig. S1.



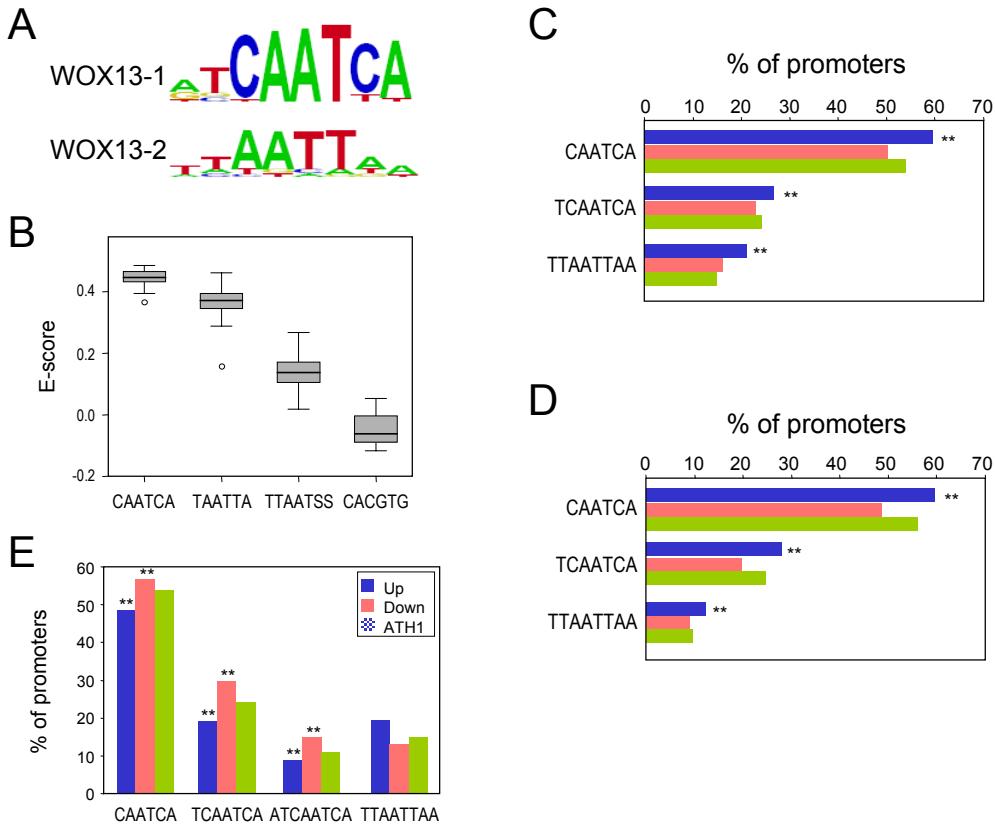
**Fig. S6.** DNA-binding specificity of ERF class of AP2/EREBP TFs. (A) PWM representation of primary and secondary motifs obtained for the proteins RRTF1, RAP2.3, RAP2.6 and ATERF1. DNA-motif shown for ERF1 was previously obtained (Godoy et al., 2011). (B) Box-plot of E-scores of the elements indicated for all the ERF proteins analyzed. Highest binding of all the proteins was observed for the GCC-box (CGCCGCC). Some proteins bound with high affinity to other GCC-related elements. (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. (E) Proportions of genes rapidly up-regulated after *ERF6* induction (blue bars) containing the elements indicated. Proportions of genes represented in the ATH1 microarray are indicated with green bars (ATH1). Asterisks denote statistical significance as in Fig. S1.



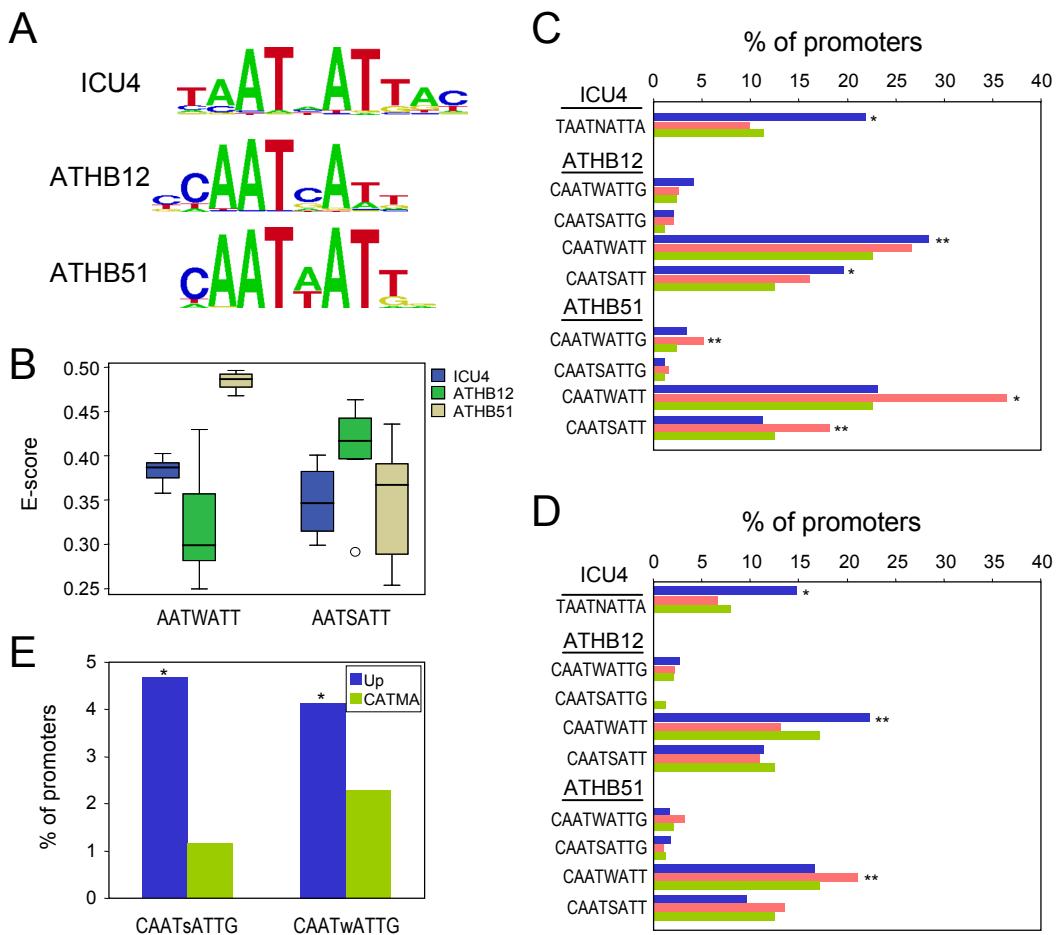
**Fig. S7.** DNA-binding specificity of DREB class of AP2/EREBP TFs. (A) PWM representation of primary and secondary motifs obtained for the proteins DDF1, DREB2C, At1g77200, ORA47, DEAR3 and DEAR4. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. (D) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in *DDF1*-overexpressing plants containing a 3'-end extended DRE (RYCGACAT) or its corresponding mutant version (RYCGACTA) at their promoter regions. (E) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in the gain-of-function mutant *hrd-D*. Extended DRE (RCCGACA), but not a mutant derivative (RCCGACT), is over-represented in the promoters of up-regulated genes. (F) Proportions of genes up-regulated (blue bars) in response to JA-treatment (1 hour) containing the elements indicated. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



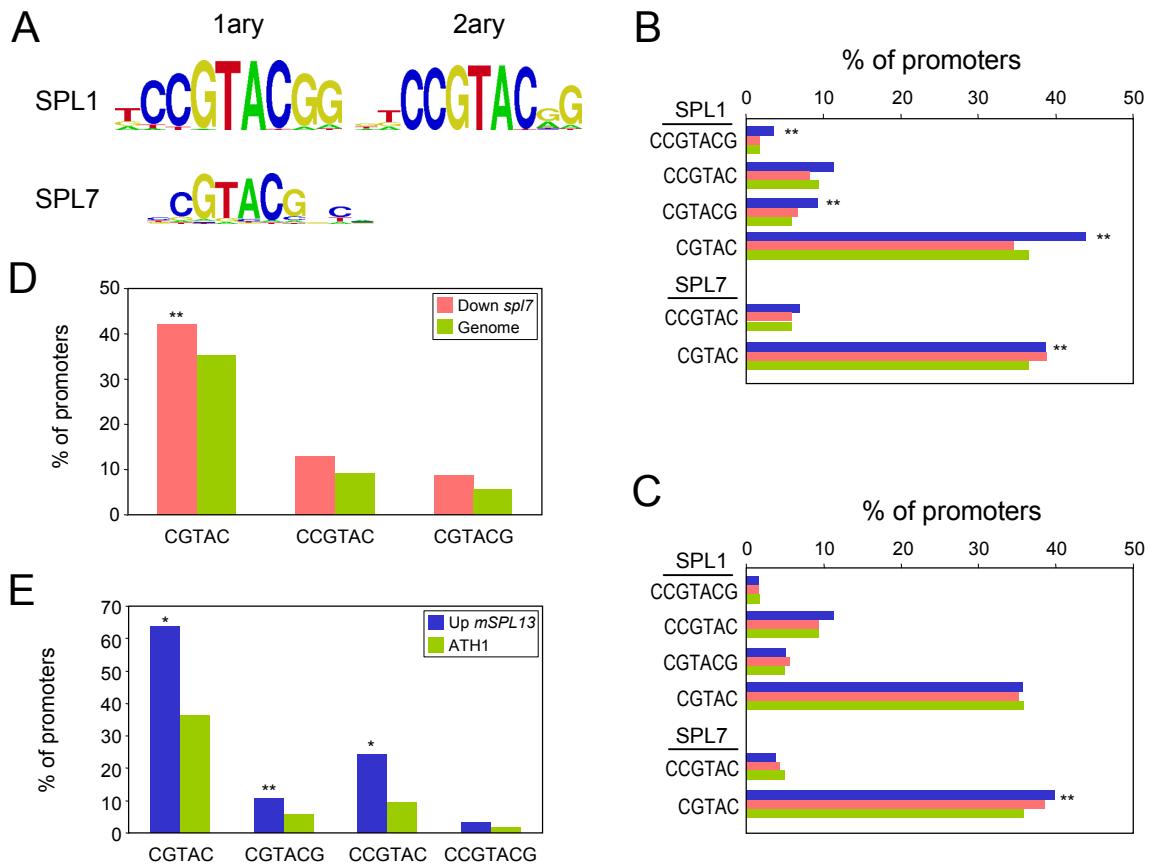
**Fig. S8.** DNA-binding specificity of AP2 class of AP2/EREBP TFs. (A) PWM representation of primary, secondary and tertiary motifs found for TOE1 and TOE2. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. (D) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) the loss-of-function mutant *ap2* containing the elements indicated at their promoter regions. (E) Proportions of genes as in (D) but corresponding to the gain-of-function mutant *smz-D*. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



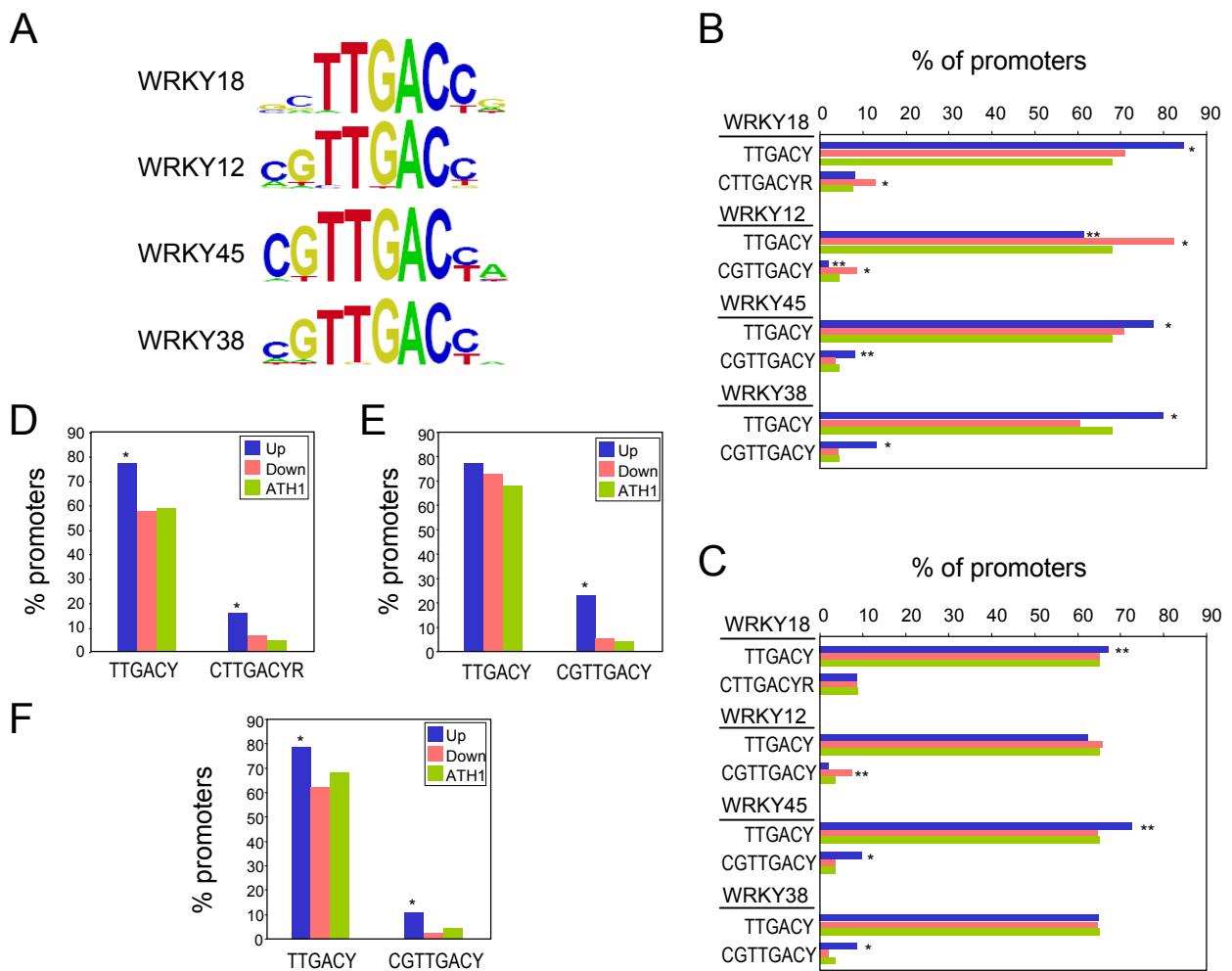
**Fig. S9.** DNA-binding specificity of WUS-like homeodomain protein WOX13. (A) PWM representation of primary and secondary motifs found for WOX13. (B) Box-plot of enrichment scores (E-score) of the elements indicated. These include 1<sup>ary</sup> and 2<sup>ary</sup> motifs found for WOX13 as well as DNA-motifs described for WUSHEL (Ref. 44). (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with WOX13 and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. (E) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in response to ethanol inducible gene expression containing the elements indicated at their promoters. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



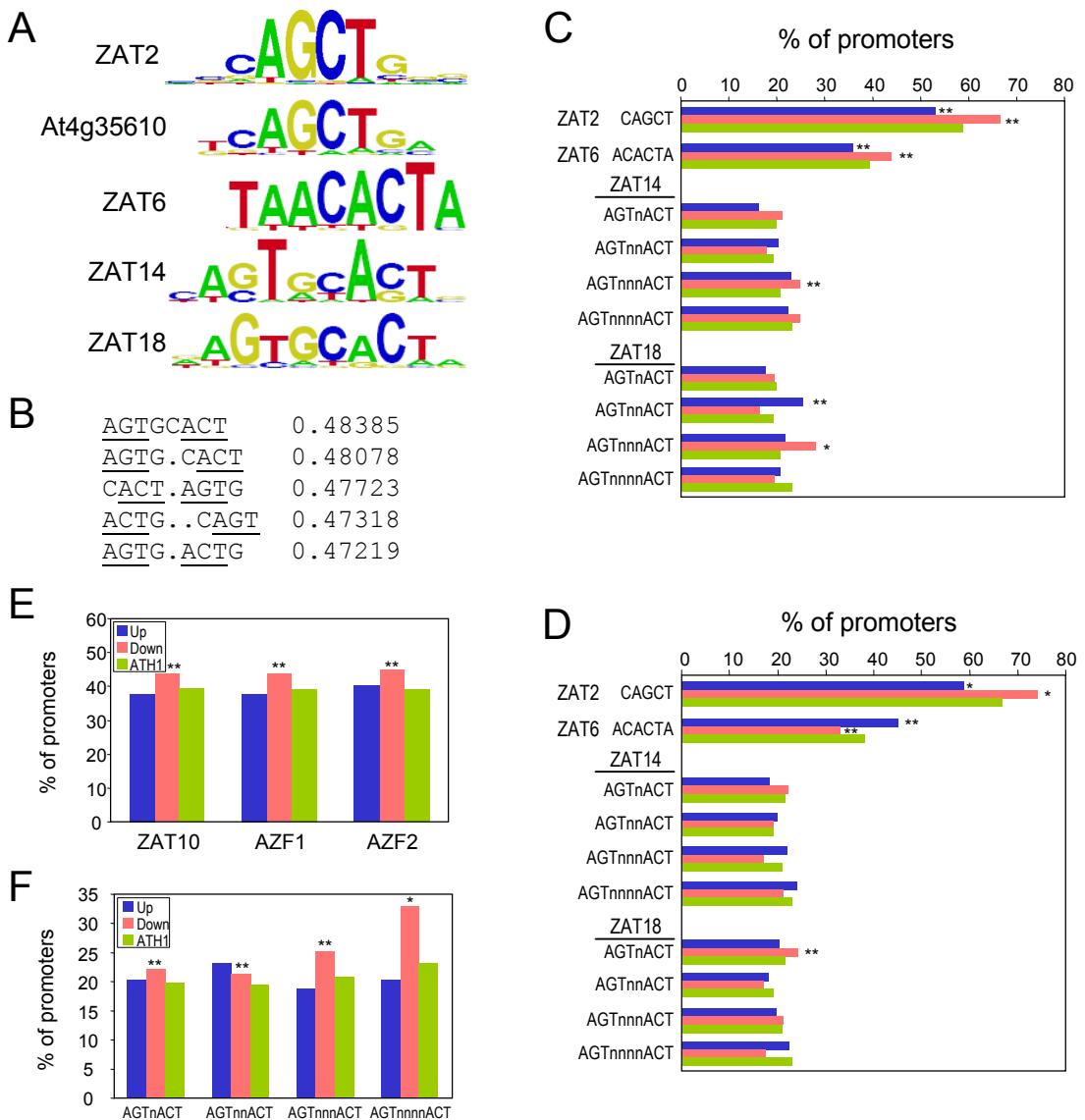
**Fig. S10.** DNA-binding specificity of HD-ZIP class of homeodomain TFs. (A) PWM representation of DNA motifs obtained for the proteins ICU4, ATHB12 and ATHB51. (B) Box-plot of enrichment scores (E-scores) of the elements indicated differing at their central position being W, A or T and S, G or C. (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. (E) Proportions of genes up-regulated (blue bars) in transgenic Arabidopsis plants overexpressing the heterologous gene *HAHB4* from sunflower containing the elements indicated at their promoter regions. Proportions of genes represented in the CATMAv2.2 microarray (CATMA) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



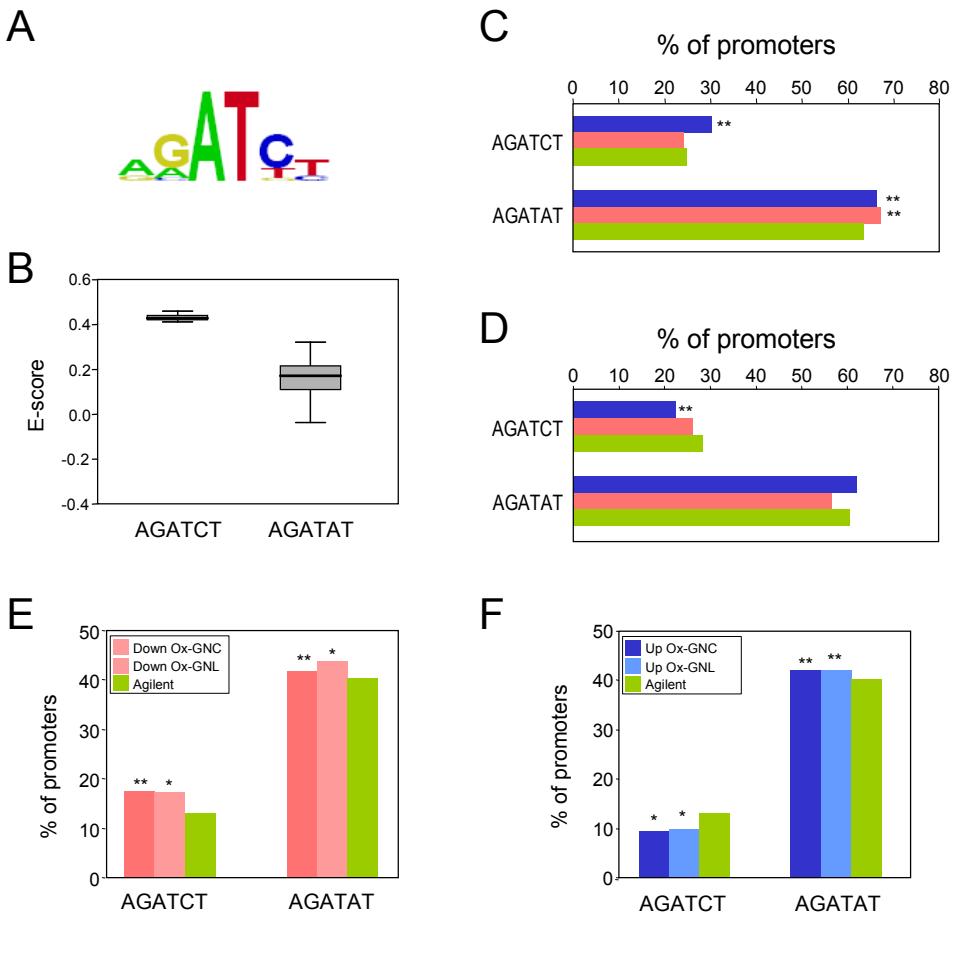
**Fig. S11.** DNA-binding specificity of SPL TFs. (A) PWM representation of DNA motifs obtained for the proteins SPL1 and SPL7. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. (D) Proportions of genes down-regulated (red bars) in the loss-of-function mutant *spl7* in response to copper deficiency containing the elements indicated at their promoter regions. The proportion of the same elements in the complete *Arabidopsis* genome is represented as green bars. (E) Proportions of genes up-regulated (blue bars) in transgenic plants expressing a gain-of-function version of *SPL13* resistant to miR156-guided cleavage containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



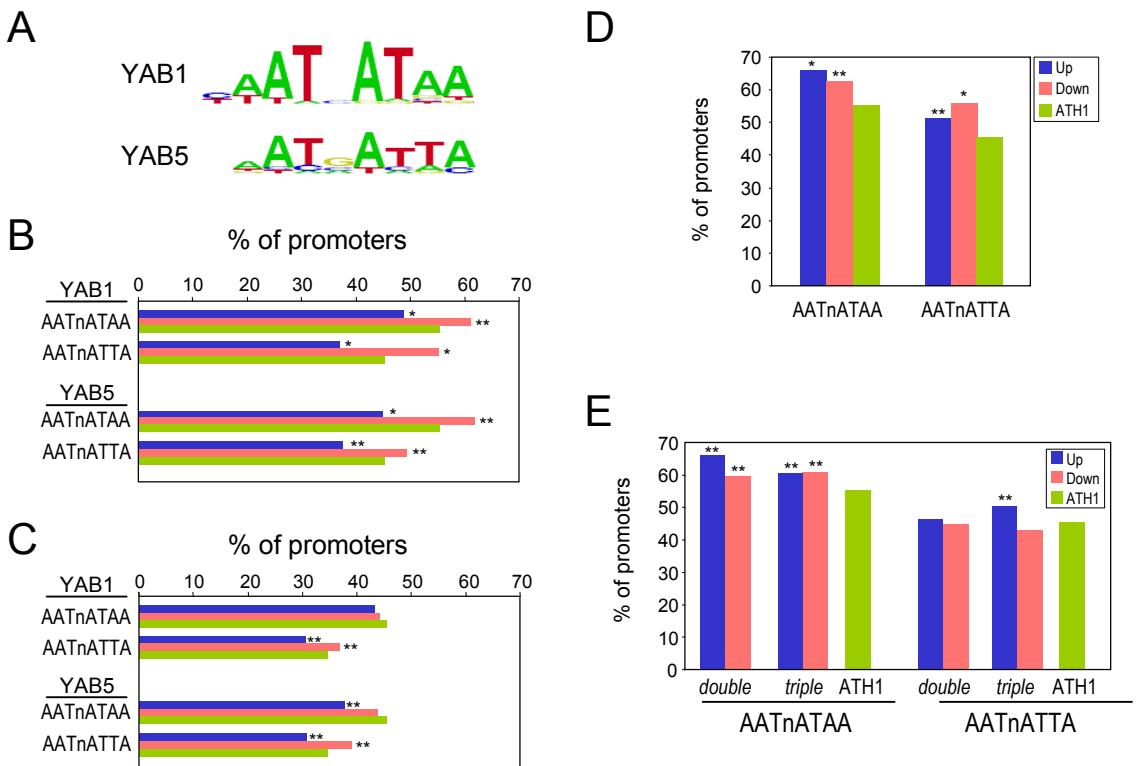
**Fig. S12.** DNA-binding specificity of WRKY TFs. (A) PWM representation of DNA motifs obtained for the proteins WRKY18, WRKY12, WRKY45 and WRKY38. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. (D) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in the double mutant *wrky18*, *wrky40* containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (E) Same as in (D) with transcriptomic data corresponding to *wrky12* loss-of-function mutant. (F) Same as in (D) with transcriptomic data corresponding to a treatment with salicylic acid. Asterisks denote statistical significance as in Fig. S1.



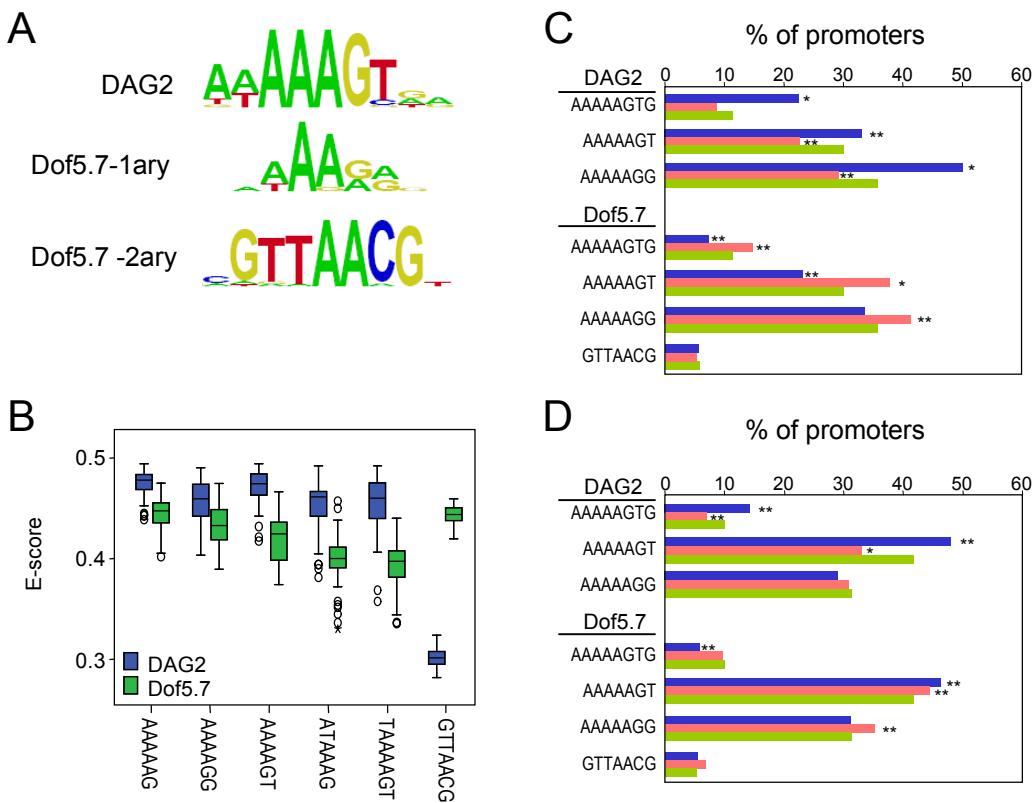
**Fig. S13.** DNA-binding specificity of C2H2 TFs. (A) PWM representation of DNA motifs obtained for the proteins ZAT2, At4g35610, ZAT6, ZAT14 and ZAT18. (B) Enrichment scores of top scoring motifs for ZAT14 showing variable distances between the A[G/C]T modules (underlined). (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. (D) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in transgenic plants overexpressing either ZAT10, AZF1 or AZF2 containing the element ACACTA at their promoter regions. Proportions of genes represented in the ATH1 microarray (background list for ZAT10) or Agilent microarray (background for AZF1 and AZF2) are indicated with green bars. (E) Same as in (D) with transcriptomic data corresponding to ZAT10 overexpressing plants. Asterisks denote statistical significance as in Fig. S1.



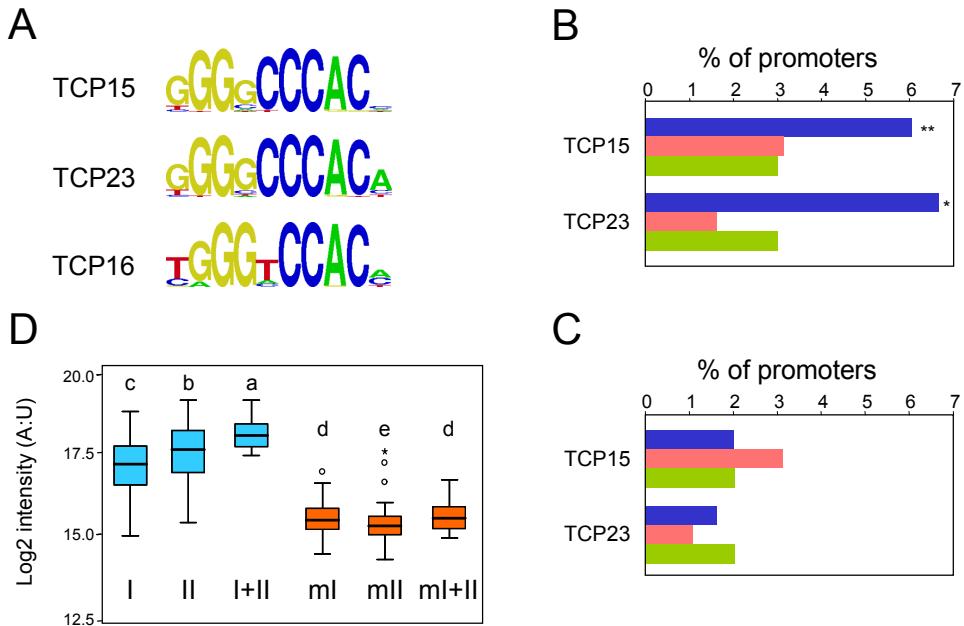
**Fig. S14.** DNA-binding specificity of the C2C2 GATA-type protein GATA12. (A) PWM representation of DNA motifs obtained for GATA12. (B) Box-plot of enrichment scores (E-scores) of the elements indicated corresponding to GATC- and GATA-containing motifs. (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with GATA12 and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. (E) Proportions of genes down-regulated (blue bars) in GNC and GNL overexpressing plants containing the elements indicated at their promoters. (F) Same as in (E) but proportions corresponding to up-regulated genes (blue bars). Proportions of genes represented in the Agilent v4 microarray (Agilent) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



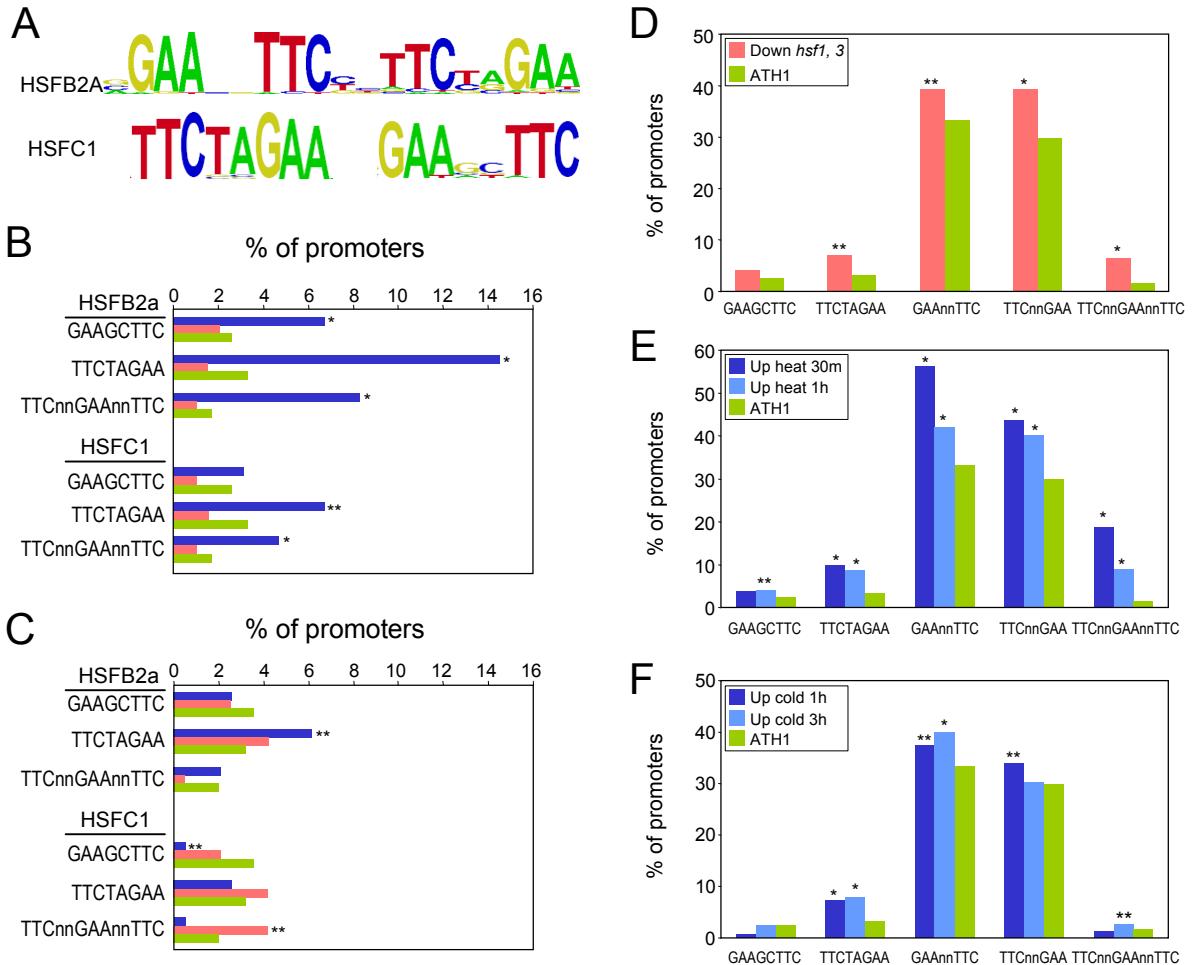
**Fig. S15.** DNA-binding specificity of C2C2 YABBY TFs. (A) PWM representation of DNA motifs obtained for YAB1 and YAB5. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. (D) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in response to YAB1 activation in transgenic plants expressing a YAB1-GR translational fusion containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (E) Same as in (D) but transcriptomic data corresponding to double *yab1,yab3* and triple *yab1,yab3,yab5* mutants. Asterisks denote statistical significance as in Fig. S1.



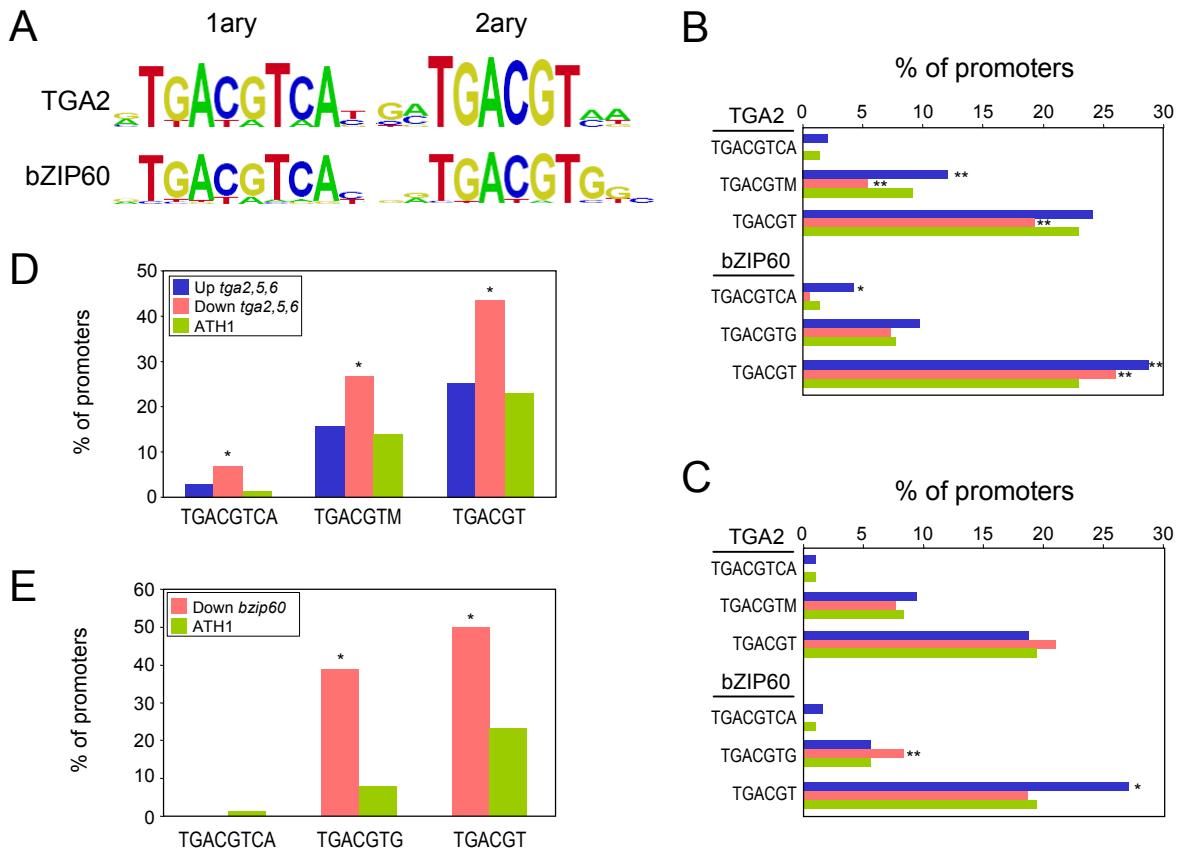
**Fig. S16.** DNA-binding specificity of C2C2 Dof-type TFs. (A) PWM representation of DNA motifs obtained for DAG2 and Dof5.7. (B) Box-plot of enrichment scores (E-scores) of the elements indicated corresponding recognized by DAG2 and Dof5.7. (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. Asterisks denote statistical significance as in Fig. S1.



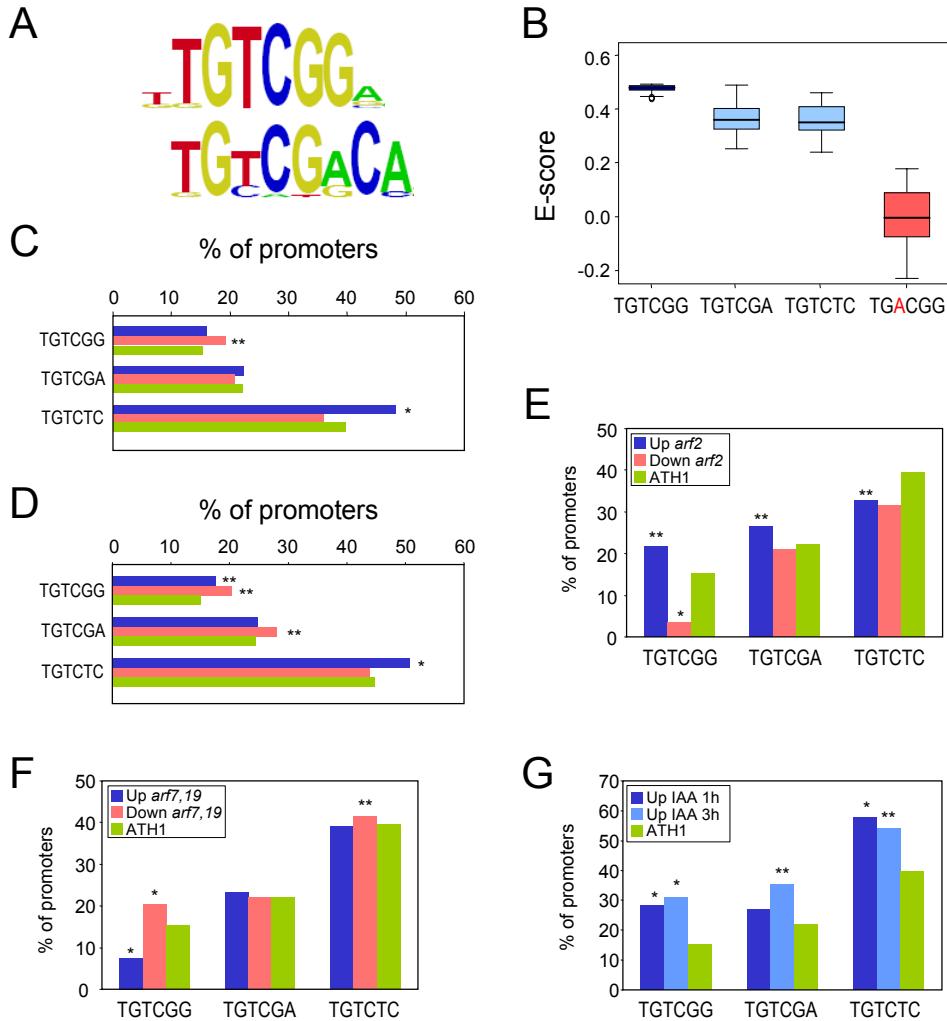
**Fig. S17.** DNA-binding specificity of TCP TFs. (A) PWM representation of DNA motifs obtained for TCP15, TCP16 and TCP23. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars and asterisks denote statistical significance as in Fig. S1. (D) Box-plot of signal intensities (log<sub>2</sub>) of the probes containing the elements indicated after incubation with TCP16. In the Figure are represented site I (I, GGGCCCAC), site II (II, GGCTCCAC) and overlapping site I/site II (I+II, GTGGNCCCCAC) with blue boxes. With red boxes are represented the distributions of their corresponding mutant versions by replacing one critical residue as follows: mutant site I (mI, GCGCCCAC), mutant site II (mII, GGCTCCAC) and mutant overlapping site I/site II (mI+II, GTGCNCCCCAC). Letters above the boxes represent different degrees of statistical significance as follows: difference between a and b, p-value=0.001081; b and c, p-value= 0.0001071; c and d, p-value<1.4e-10; d and e, p-value<0.005 (Wilcoxon rank test).



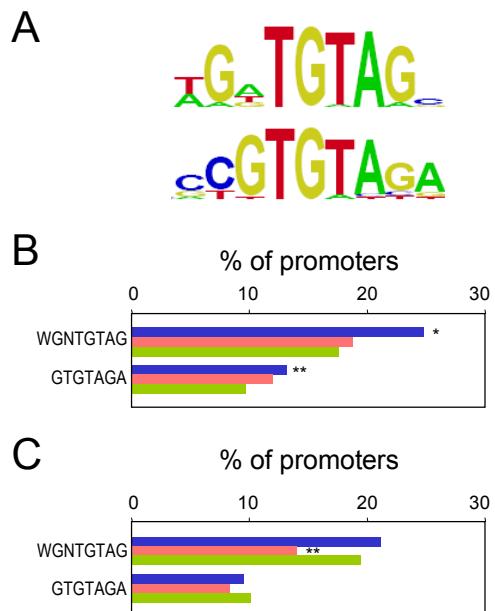
**Fig. S18.** DNA-binding specificity of HSF TFs. (A) PWM representation of DNA motifs obtained for HSFB2a and HSFC1. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (D) Proportions of genes down-regulated in double mutant *hsf1,hsf3* in response to a heat treatment containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (E) Same as in (D) with transcriptomic data relating up-regulated genes in response to heat treatments. (F) Same as in (D) with transcriptomic data relating up-regulated genes in response to cold treatments. Asterisks denote statistical significance as in Fig. S1.



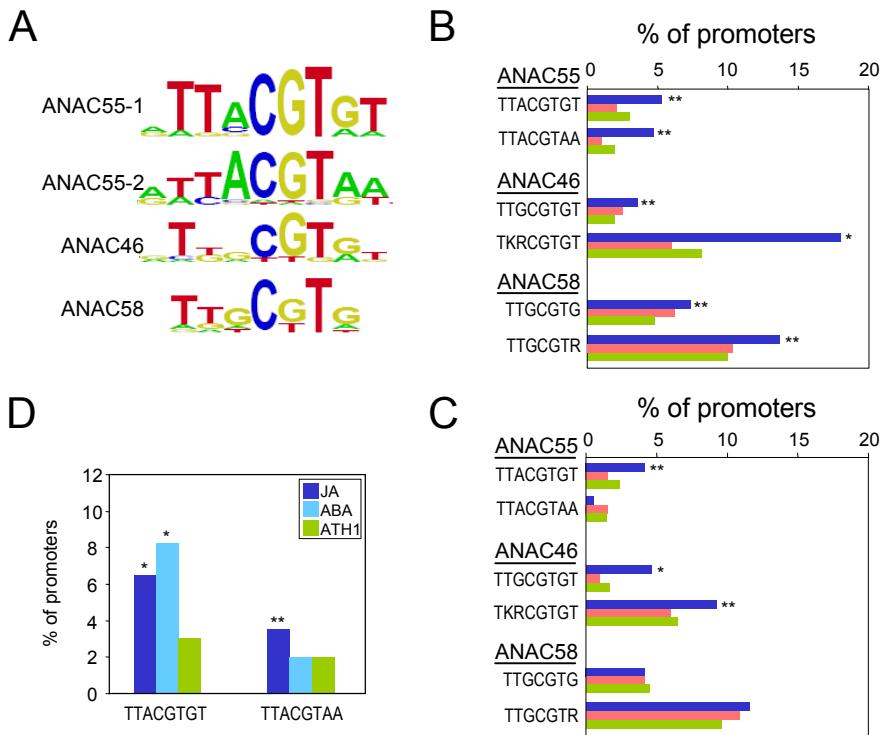
**Fig. S19.** DNA-binding specificity of bZIP TFs. (A) PWM representation of DNA motifs obtained for TGA2 and bZIP60. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (D) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in triple mutant *tga2,5,6* containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (E) Same as in (D) with transcriptomic data relating down-regulated genes in the loss-of-function *bzip60* mutant Asterisks denote statistical significance as in Fig. S1.



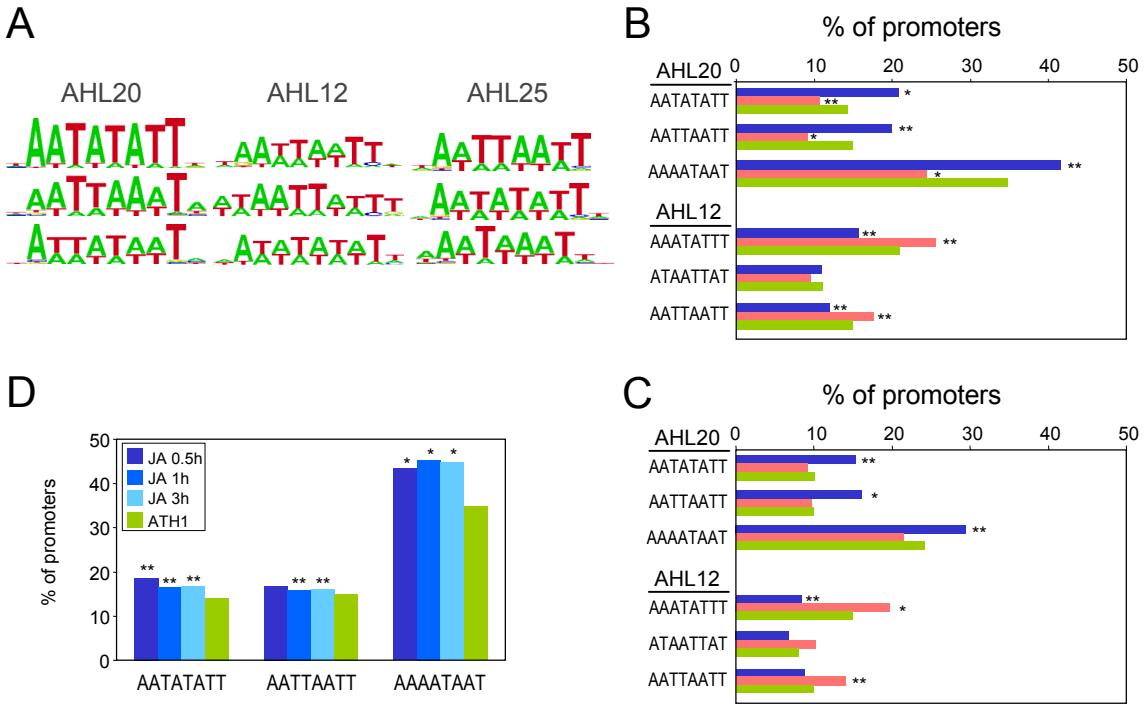
**Fig. S20.** DNA-binding specificity of the B3 ARF-type ETTIN protein. (A) PWM representation of primary and secondary motifs obtained for ETTIN. (B) Box-plot of enrichment scores (E-scores) of the elements indicated. In the Figure are shown distribution of E-scores for the elements obtained for ETTIN (primary TGTCGG, and secondary TGTCGA), as well as canonical AuxRE identified for other ARF proteins (TGTCTC). Is also shown the distribution of E-scores of a mutant derived from primary sequence by replacing one critical residue (TGACGG, mutated nucleotide in red). (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with *ETTIN* and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (E) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in the loss-of-function mutant *arf2* containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (F) Same as in (E) with transcriptomic data relating double mutant *arf7, 19*. (G) Same as in (E) with transcriptomic data relating up-regulation of genes in response to treatment with the auxin IAA. Asterisks denote statistical significance as in Fig. S1.



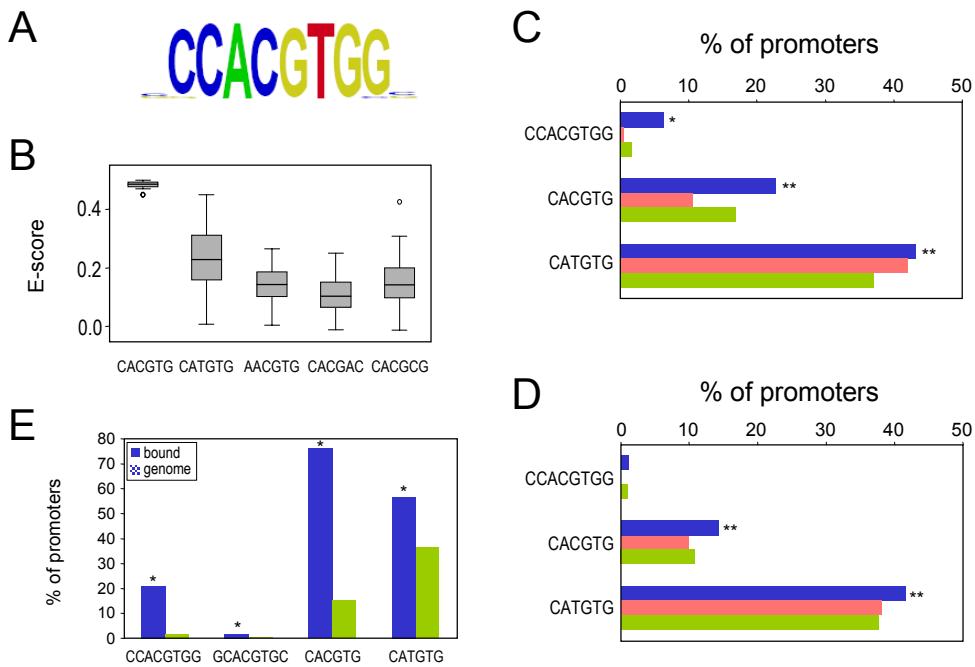
**Fig. S21.** DNA-binding specificity of the B3 REM-type REM1 protein. (A) PWM representation of primary and secondary motifs obtained for REM1. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with *REM1* and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. W represents A or T. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



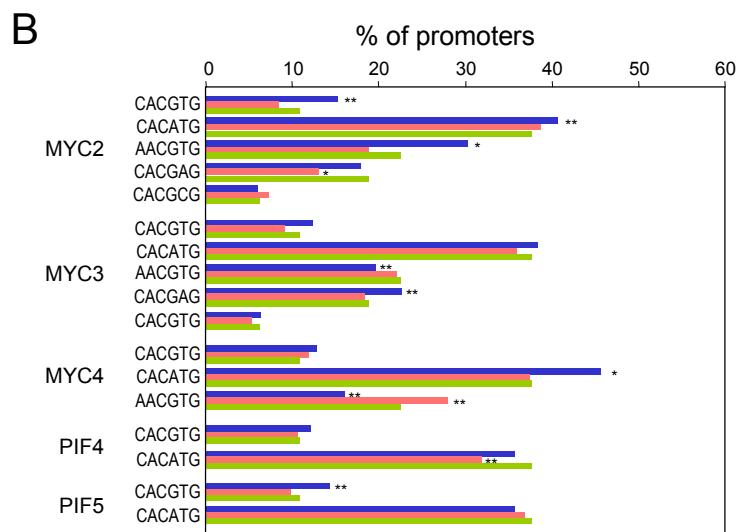
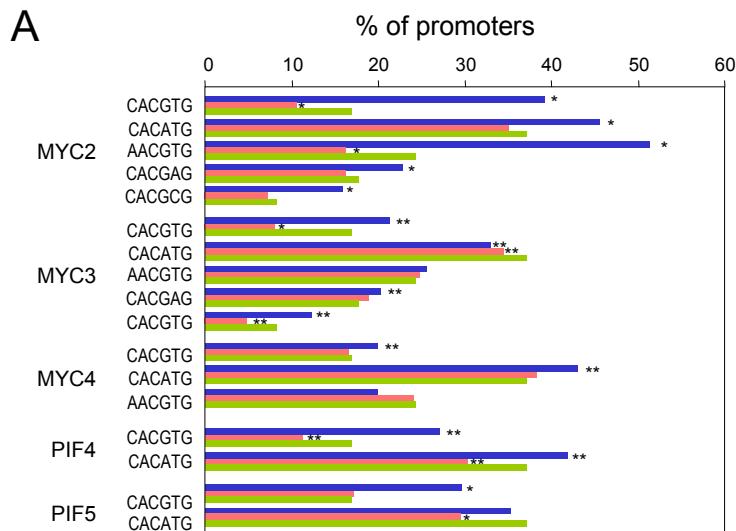
**Fig. S22.** DNA-binding specificity of NAC TFs. (A) PWM representation of primary and secondary motifs obtained for ANAC55, ANAC46 and ANAC58. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (D) Proportions of genes up-regulated in response to JA and ABA treatments containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



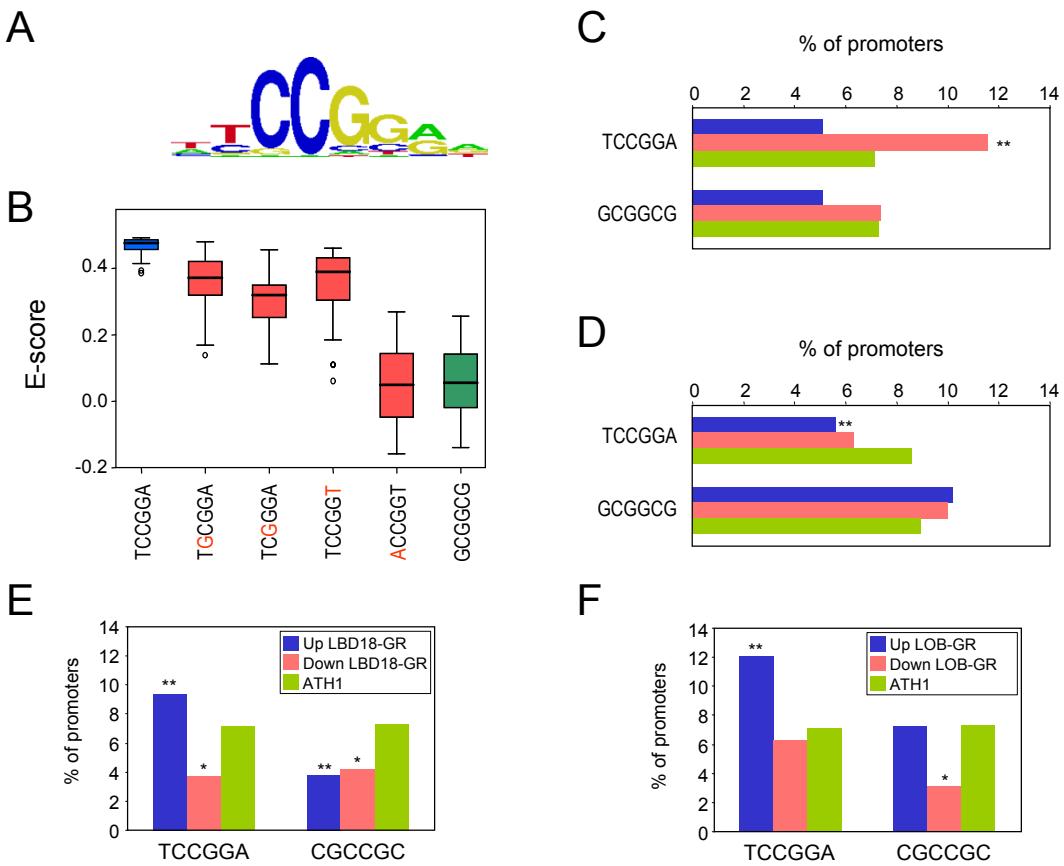
**Fig. S23.** DNA-binding specificity of AHL TFs. (A) PWM representation of primary and secondary motifs obtained for AHL20, AHL12 and AHL25. (B) and (C) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (B) and downstream (C) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (D) Proportions of genes up-regulated at several time points (0.5 hours, 1 hour and 3 hours) in response to JA treatment containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



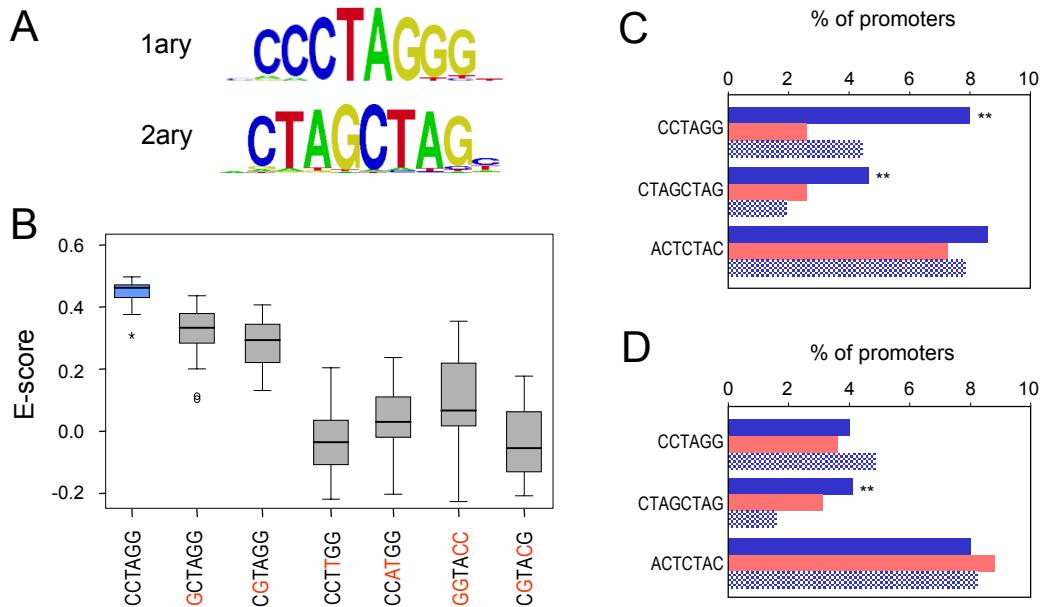
**Fig. S24.** DNA-binding specificity of the bHLH protein PIF3. (A) PWM representation of primary obtained for PIF3. (B) Box plot representation of the distribution of enrichment scores (E-scores) corresponding to different G- and G-like elements indicated. PIF3 recognized *in vitro* almost exclusively the G-box (CACGTG). (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with *PIF3* and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (E) Proportions of PIF3 targets containing the elements indicated at their promoter regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.



**Fig. S25.** Enrichment of bHLH binding sites in co-regulated promoters. (A) and (B) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with the genes indicated and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (A) and downstream (B) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1. Recognition sites for these proteins were previously determined (Godoy et al., 2011; Fernández-Calvo et al., 2012; Hornitschek et al., 2012).

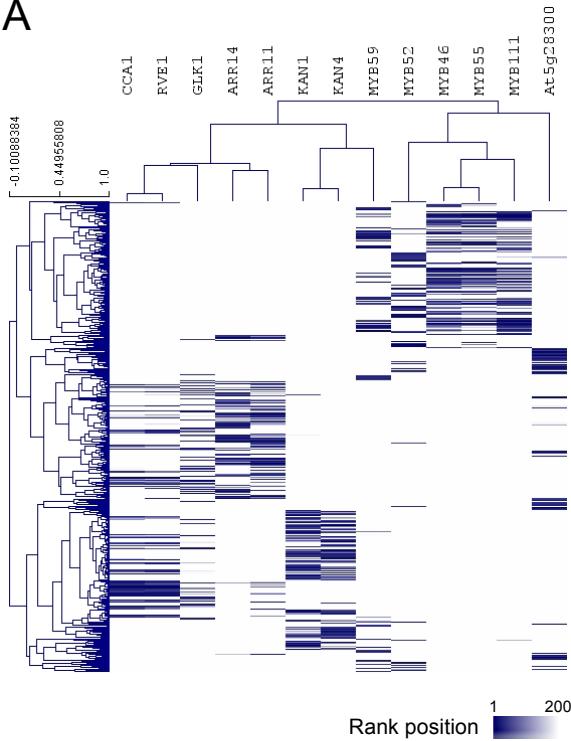


**Fig. S26.** DNA-binding specificity of the LOB/AS2 protein LBD16. (A) PWM representation of primary obtained for LBD16. (B) Box plot representation of the distribution of enrichment scores (E-scores) of DNA motif obtained for LBD16 (TCCGGA, blue box) as well as other mutant derivatives revealing the relevance of the central core sequence CCGG (red boxes). In green is represented the distribution of E-scores observed for LOB proteins (GCGGGCG; Husbands et al., 2007). (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with *LBD16* and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (E) Proportions of genes up-regulated (blue bars) and down-regulated (red bars) in response to activation of LBD18-GR. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. (F) Same as in (E) but data corresponding to activation of LOB. Asterisks denote statistical significance as in Fig. S1.

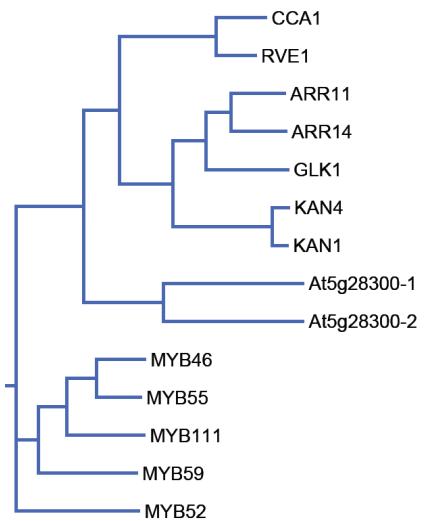


**Fig. S27.** DNA-binding specificity of the STY protein STY1. (A) PWM representation of primary obtained for SY1. (B) Box plot representation of the distribution of enrichment scores (E-scores) of DNA motif obtained for STY1 (CCTAGG, blue box) as well as other mutant derivatives revealing the relevance of the central core sequence (CTAG). (C) and (D) Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated with STY2 and represented in the ATH1 microarray (green bars) containing the elements indicated at their upstream (C) and downstream (D) regions. Proportions of genes represented in the ATH1 microarray (ATH1) are indicated with green bars. Asterisks denote statistical significance as in Fig. S1.

A



B

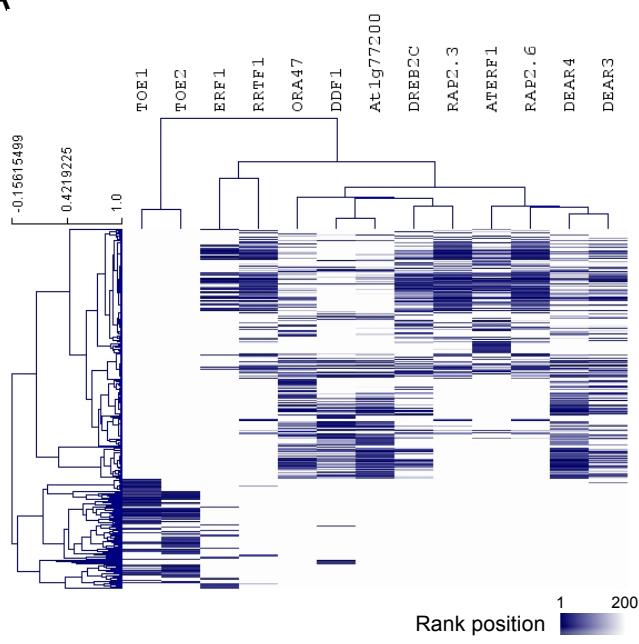


C

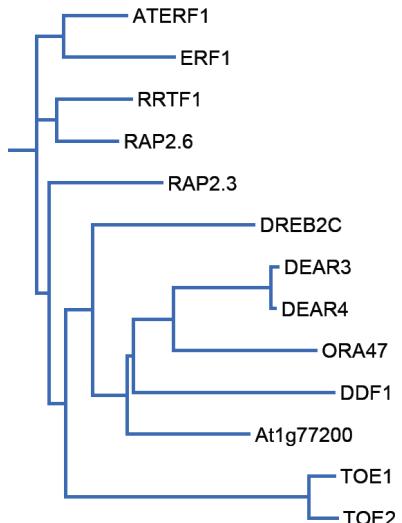
	CCA1	RVE1	ARR11	ARR14	GLK1	KAN4	KAN1	MYB46	MYB55	MYB111	MYB59	MYB52	At5g28300-1	At5g28300-2
CCA1	100.0	80.0	26.7	29.6	31.8	31.1	33.3	13.3	15.6	17.8	20.0	20.0	9.7	16.3
RVE1		100.0	31.1	31.8	31.8	35.6	35.6	17.8	20.0	22.2	22.2	22.2	9.7	13.9
ARR11			100.0	75.9	63.5	48.1	46.3	13.2	15.1	15.1	17.0	2.6	2.6	13.7
ARR14				100.0	65.4	49.1	47.2	13.5	17.3	11.5	15.4	11.5	5.4	10.0
GLK1					100.0	50.0	53.9	11.3	13.2	11.3	15.1	15.1	2.6	7.8
KAN4						100.0	92.6	7.6	9.4	18.9	15.1	9.4	10.5	9.8
KAN1							100.0	7.6	9.4	17.0	15.1	11.3	10.5	9.8
MYB46								100.0	79.4	63.9	56.7	45.8	10.0	8.6
MYB55									100.0	68.0	56.7	44.8	11.4	5.4
MYB111										100.0	53.5	44.8	12.9	8.4
MYB59											100.0	46.9	10.0	9.5
MYB52												100.0	10.0	10.7
At5g28300-1													100.0	38.6
At5g28300-2														100.0

**Fig. S28.** DNA-binding specificity of the MYB superfamily of TFs. (A) Heat map of hierarchical clustering analysis of 468 different 8mers recognized by TFs belonging to the MYB superfamily of TFs. The 469 8mers were among top 50 ranking motifs for at least one of the proteins. (B) Phylogenetic relationships among DNA-binding domains of MYB-proteins analyzed. (C) Amino-acid identities among different MYB proteins tested.

A



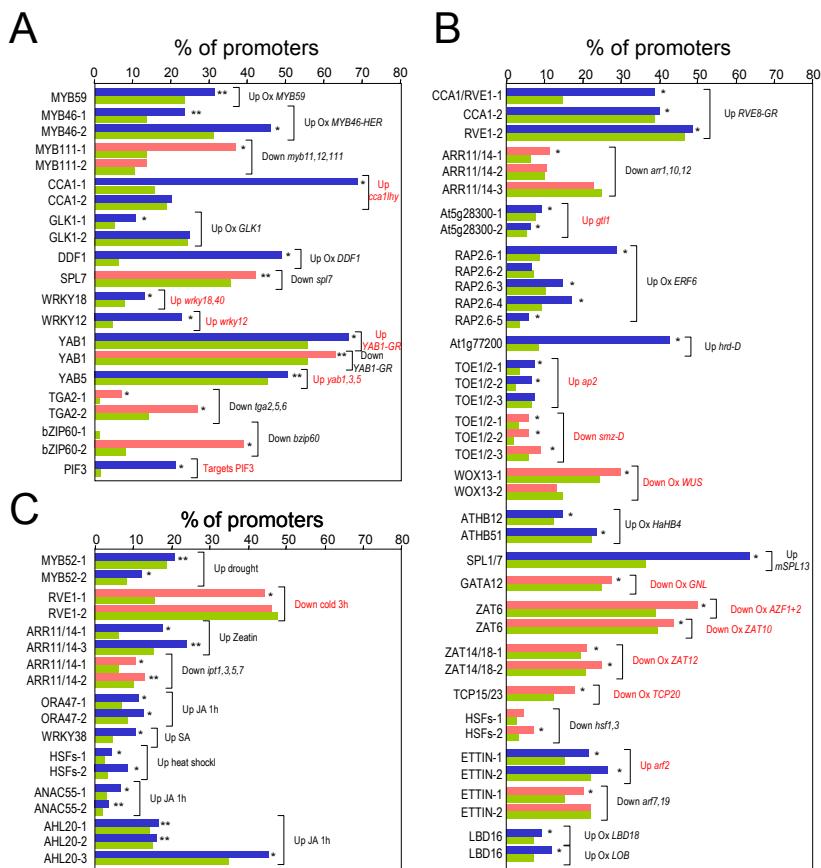
B



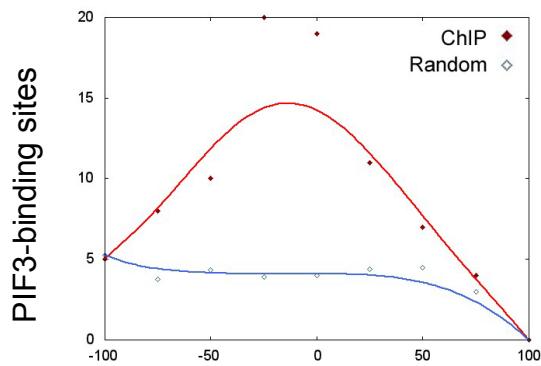
C

	ATERF1	RRTF1	ERF1	RAP2.6	RAP2.3	DREB2C	DEAR3	DEAR4	ORA47	DDF1	At1g77200	TOE1	TOE2
ATERF1	100.0	82.8	78.0	77.6	70.7	62.1	55.2	55.2	51.7	46.6	56.9	50.0	50.0
RRTF1		100.0	69.0	82.8	70.7	56.9	56.9	51.7	48.3	60.3	53.6	53.6	51.8
ERF1			100.0	69.0	69.0	60.3	51.7	51.7	46.6	46.6	55.2	46.4	46.4
RAP2.6				100.0	75.9	58.6	60.3	58.6	51.7	51.7	65.5	53.6	53.6
RAP2.3					100.0	62.1	58.6	58.6	51.7	53.4	62.1	48.2	48.2
DREB2C						100.0	55.2	56.9	46.6	48.3	62.1	41.1	41.1
DEAR3							100.0	98.3	69.0	58.6	65.5	39.3	39.3
DEAR4								100.0	69.0	58.6	67.2	39.3	39.3
ORA47									100.0	50.0	60.3	42.9	41.1
DDF1										100.0	58.6	35.5	35.7
At1g77200											100.0	42.9	42.9
TOE1												100.0	93.0
TOE2													100.0

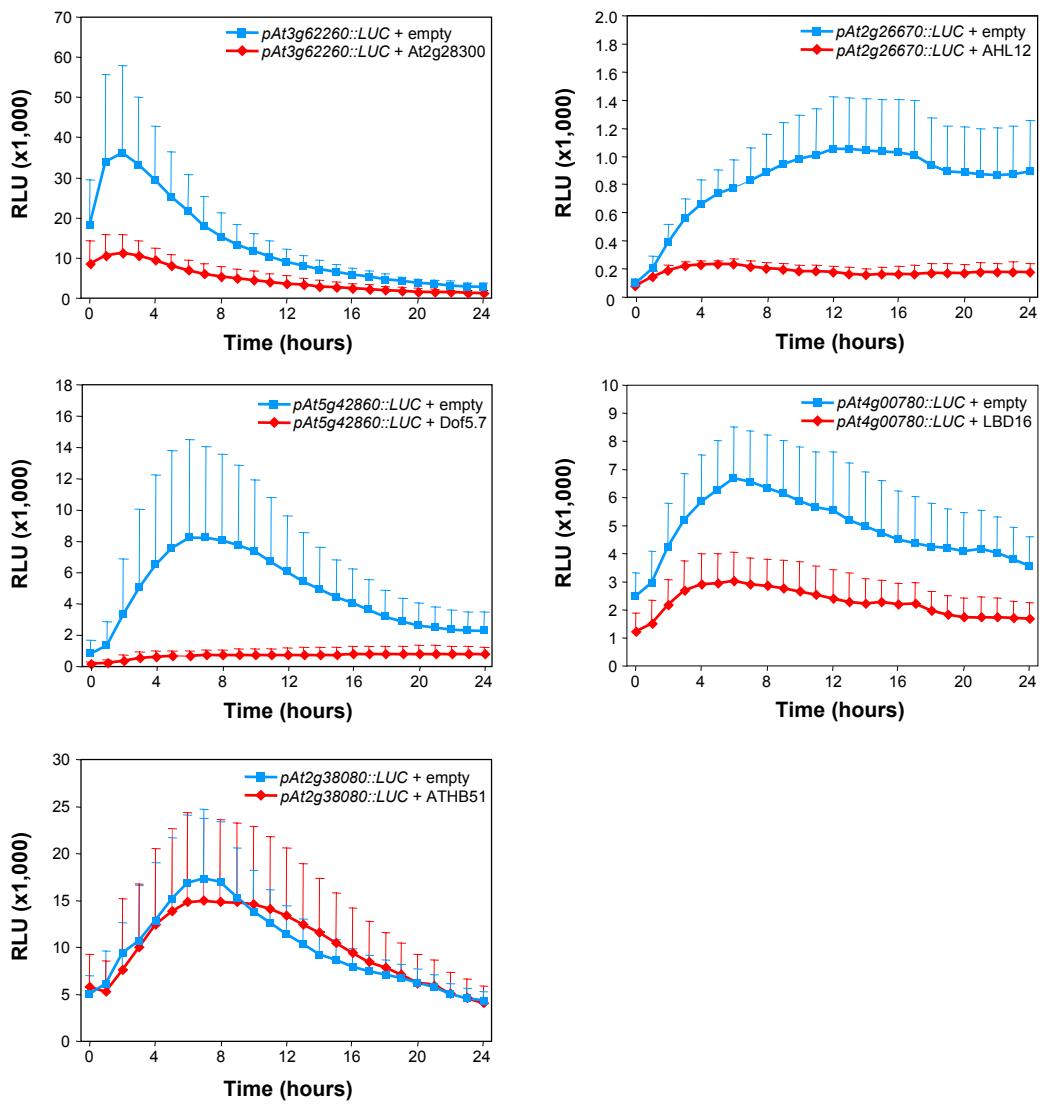
**Fig. S29.** DNA-binding specificity of the AP2/EREBP superfamily of TFs. (A) Heat map of hierarchical clustering analysis of 325 different 8mers recognized by TFs belonging to the AP2/EREBP superfamily of TFs. The 325 8mers were among top 50 ranking motifs for at least one of the proteins. (B) Phylogenetic relationships among DNA-binding domains of AP2/EREBP-proteins analyzed. (C) Amino-acid identities among different MYB proteins tested.



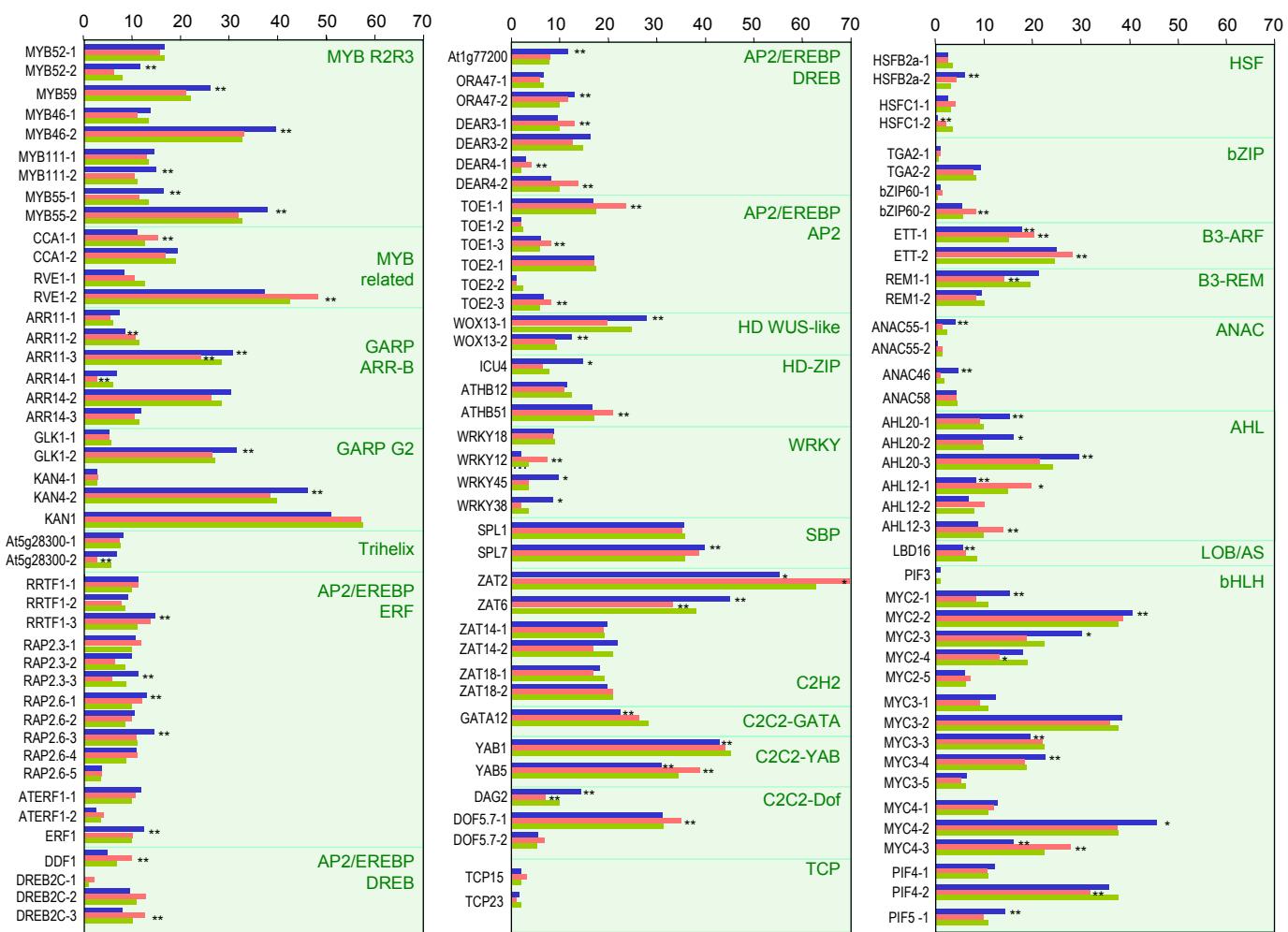
**Fig S30.** Evaluation of biological relevance of binding sites from transcriptomic data. (A) DNA-motifs obtained *in vitro* were scanned in the promoters (1 kb) of the genes de-regulated in different mutant or overexpressing genotypes involving the TF under study according to publicly available transcriptomic data. In the Figure are represented the proportions (in %) of genes up-regulated (blue bars) in a given genotype that contain the DNA-element indicated at their promoters. Red bars represent the same proportions corresponding to genes down-regulated. The proportions of the total number of genes containing the same elements, and thus representing a random distribution, are represented with green bars. Several gene sets are shown in the Figure, depending on the genotype and transcriptional activity of the TF: I) Up-regulated genes in transgenic plants overexpressing transcriptional activators (e.g. Ox *MYB59*; Ox *MYB46-HER*; Ox *GLK1*; Ox *DDF1*). II) Down-regulated in loss-of-function mutants involving transcriptional activators (*myb11,12,11*; *spl7*; *tga2,5,6*; *bzip60*). III) Up-regulated in loss-of-function mutants of a transcriptional repressor (*cca1,hy*; *wrky18,40*; *wrky12*; *yab1,3,5*). In the case of Ox *YAB1-GR*, both up- and down-regulated genes are represented, given its possible activity as a bifunctional TF. Gene set for PIF3 corresponds to target genes identified by ChIP-seq. (B) Same as in (A) with transcriptomic data involving structurally similar TFs. In this case, genes sets are as follows: I) Up-regulated genes after post-translational activation of the TF (*RVE8-GR*). II) Up-regulated genes in transgenic plants overexpressing transcriptional activators (Ox *ERF6*; *mSPL13*; Ox *HAHB4*; Ox *LBD18*; Ox *LOB*). III) Down-regulated in loss-of-function mutants involving transcriptional activators (*arr1,10,12*; *hsf1,3*; *arf7,19*). IV) Up-regulated in loss-of-function mutants of transcriptional repressors (*gtl1*; *ap2*; *arf2*). V) Down-regulated in transgenic plants overexpressing transcriptional repressors (Ox *WUS*; Ox *GNL*; Ox *AZF1+AZF2*; Ox *ZAT10*; Ox *ZAT12*; Ox *TCP20*). VI) Up-regulated in a gain-of-function mutant of an activator (*hrd-D*). VII) Down-regulated in a gain-of-function mutant of a repressor (*smz-D*). (C) Same as in (a) with transcriptomic data involving experimental conditions related with the TFs under study: drought, cold, zeatin (*Zea*), jasmonic acid (JA), heat shock, salicylic acid (SA). *ipt1,3,5,7* is a mutant genotype defective in cytokinins synthesis. Asterisks indicate the degree of statistical significance in the differences of the proportions indicated (following a binomial distribution) as follows: \*p-value<0.005; \*\*p-value<0.05.



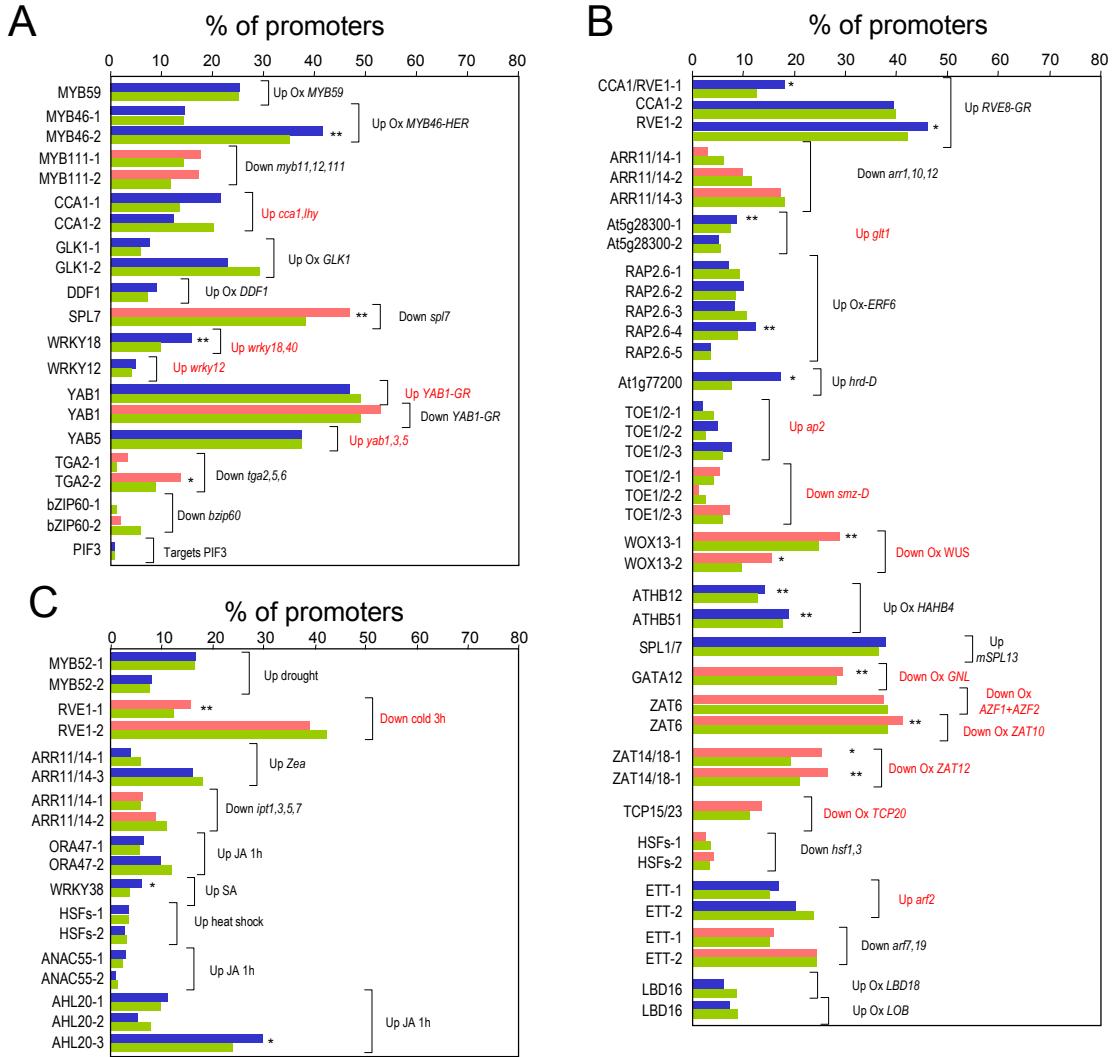
**Fig. S31.** PIF3-binding sites are enriched in PIF3-bound DNA *in vivo*. The plot represents the number of PIF3-binding sites (red diamonds and line) found in PIF3-bound genomic DNA previously obtained by chromatin immunoprecipitation (ChIP; Ref. 114). In blue is represented the average number for 100 randomized PWMs. Ssites are recorded along bound fragments (200 bp, centered at the middle position).



**Fig. S32.** Relative luminiscence units (RLU) from transient expression studies of the indicated promoters fused to the *LUC* reporter gene co-transformed with their corresponding effector constructs (At2g28300, AHL12, Dof5.7, LBD16 and ATHB51) or empty vector. Constructs were transiently expressed in *N. benthamiana* leaves and luciferase activity measured at 1 hour interval for 24 hours.

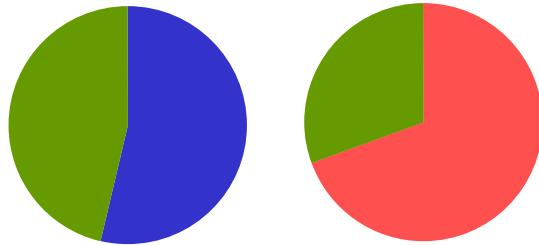


**Fig. S33.** Evaluation of TF-binding sites located downstream the co-regulated genes. Lists of co-regulated genes with the TF-coding genes were scanned for the presence of DNA-motifs obtained *in vitro* at their downstream regions (1 kb). Frequencies (in %) of genes positively (blue bars) and negatively (red bars) co-regulated genes containing the indicated DNA-elements are shown. Proportions of genes represented in the ATH1 microarray containing the corresponding elements, and thus representing a random distribution, are represented as green bars. Asterisks indicate the degree of statistical significance in the differences of the proportions indicated (following a binomial distribution) as follows: \*p-value<0.005; \*\*p-value<0.05.



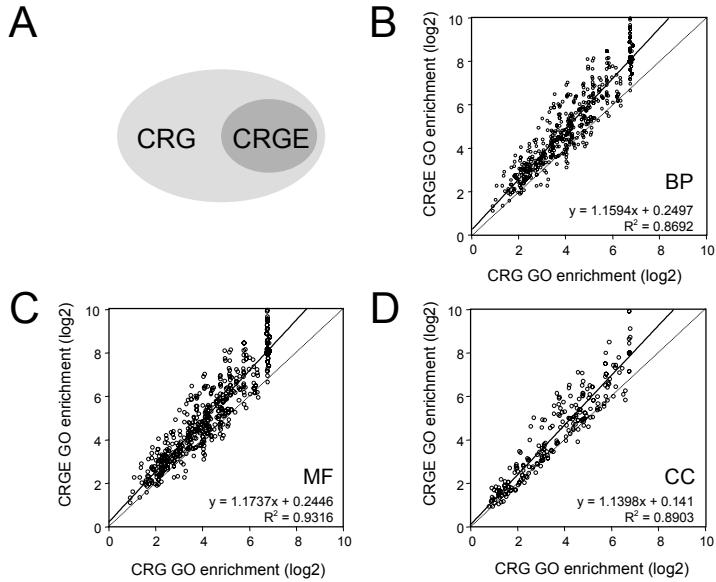
**Fig. S34.** Evaluation of TF-binding sites located downstream the genes de-regulated in mutant or over-expressing genotypes. (A) DNA-motifs obtained *in vitro* were scanned at downstream regions (1 kb) of the genes de-regulated in different mutant or overexpressing genotypes involving the TF under study according to publicly available transcriptomic data. Blue bars represent the frequency (in %) of genes up-regulated in a given genotype that contain the DNA-element indicated. Red bars represent the same proportions corresponding to genes down-regulated. The proportions of the total number of genes containing the same elements, and thus representing a random distribution, are represented with green bars. Gene sets are as in Fig. S28A. (B) Same as in (A) with transcriptomic data involving structurally similar TFs. Gene sets are as in Fig. S28B. (C) Same as in (A) with transcriptomic data involving experimental conditions related with the TFs under study. Gene sets are as in Fig. S28C.

## Activators      Repressors



- Downstream enriched motifs corresponding to activators
- Downstream enriched motifs corresponding to repressors
- Downstream non enriched enriched motifs

**Fig. S35.** Transcriptional repressors binding sites are more enriched at downstream regions than activators binding sites



**Fig. S36.** Subsets of co-regulated genes containing DNA-motifs at their promoters are enriched in particular Gene Ontology (GO) terms. (A) Co-regulated genes with each TF (CRG) and their subsets containing TF-binding sequences at their promoters (CRGE) were scanned for GO term enrichment. (B) Plot of enrichment of coincident Biological Process (BP) GO terms between CRG and CRGE lists. (C) Plot of enrichment of coincident Molecular Function (MF) GO terms between CRG and CRGE lists. (D) Plot of enrichment of coincident Cellular Component (CC) GO terms between CRG and CRGE lists.

**4. Table S1.** TFs analyzed by PBM in this study.

	TAIR ID	Family (subfamily)	Element <sup>a</sup>	Sequence <sup>a</sup>	Dataset <sup>b</sup>	r range <sup>c</sup>
MYB52	At1g17950	MYB (R2R3)	MYB52-1 MYB52-2	CCGTTA GTTAGTTR	Anatomy Perturbations	0.897:0.531 0.612:0.286
MYB59	At5g59780	MYB (R2R3)	MYB59	RTTAGGT	Samples	0.693:0.502
MYB46	At5g12870	MYB (R2R3)	MYB46-1 MYB46-2	GKTAGGT GKTKGTR	Anatomy Perturbations	0.939:0.576 0.738:0.341
MYB111	At5g49330	MYB (R2R3)	MYB111-1 MYB111-2	GKTAGGT GGTWGTTT	Samples Samples	0.549:0.367 0.549:0.367
MYB55	At4g01680	MYB (R2R3)	MYB55-1 MYB55-2	GKTAGGT GKTKGTR	Anatomy Samples	0.668:0.483 0.683:0.426
CCA1	At2g46830	MYB (Myb-related)	CCA1-1 CCA1-2	AGATATT AGATATG	Samples Samples	-0.573:-0.291 -0.573:-0.291
RVE1	At5g17300	MYB (Myb-related)	RVE1-1 RVE1-2	AGATATT AGATTTC	Samples Samples	-0.547:-0.362 -0.547:-0.362
ARR11	At1g67710	MYB (GARP, B type-ARR)	ARR11-1 ARR11-2 ARR11-3	AGATACG GGATCC AGATCT	Perturbations Perturbations Perturbations	0.413:0.283 0.413:0.283 0.413:0.283
ARR14	At2g01760	MYB (GARP, B type-ARR)	ARR14-1 ARR14-2 ARR14-3	AGATACG AGATCT GGATCC	Samples Samples Samples	0.736:0.610 0.736:0.610 0.736:0.610
GLK1	At2g20570	MYB (GARP, G2)	GLK1-1 GLK1-2	RGATATCY AGATTCT	Perturbations Perturbations	0.610:0.474 0.610:0.474
KAN4	At5g42630	MYB (GARP, G2)	KAN4-1 KAN4-2	GAATATTC GAATAWT	Perturbations Perturbations	0.652:0.328 0.652:0.328
KAN1	At5g16560	MYB (GARP, G2)	KAN1	GAATAT	Anatomy	-0.530:-0.383
At5g28300	At5g28300	Trihelix	At5g28300-1 At5g28300-2	CGGTAAA CGGTAAT	Perturbations Perturbations	-0.458:-0.312 -0.458:-0.312
RRTF1	At4g34410	AP2/EREBP (ERF)	RRTF1-1 RRTF1-2 RRTF1-3	GCCGCC GCCGCA GCCGTC	Samples Samples Samples	0.831:0.429 0.831:0.429 0.831:0.429
RAP2.3	At3g16770	AP2/EREBP (ERF)	RAP2.3-1 RAP2.3-2 RAP2.3-3	GCCGCC GCCGCA GCCGAC	Anatomy Anatomy Anatomy	0.624:0.450 0.624:0.450 0.624:0.450
RAP2.6	At1g43160	AP2/EREBP (ERF)	RAP2.6-1 RAP2.6-2 RAP2.6-3 RAP2.6-4 RAP2.6-5	GCCGCC GCCGCA GCCGTC GCCGAC GCCGGC	Samples Samples Samples Samples Samples	0.679:0.482 0.679:0.482 0.679:0.482 0.679:0.482 0.679:0.482
ATERF1	At4g17500	AP2/EREBP (ERF)	ATERF1-1 ATERF1-2	GCCGCC GCCGGC	Perturbations Perturbations	0.768:0.561 0.768:0.561
ERF1 <sup>e</sup>	At3g23240	AP2/EREBP (ERF)	ERF1	GCCGCC	Perturbations	0.601:0.377
DDF1	At1g12610	AP2/EREBP (DREB)	DDF1	RYCGACAT	Samples	0.831:0.363
DREB2C	At2g40340	AP2/EREBP (DREB)	DREB2C-1 DREB2C-2 DREB2C-3	GC GGCCG MACCGMCW GCCGCC	Samples Samples Samples	0.690:0.399 0.690:0.399 0.690:0.399
At1g77200	At1g77200	AP2/EREBP (DREB)	At1g77200	RCCGACA	Perturbations	0.414:0.211
ORA47	At1g74930	AP2/EREBP (DREB)	ORA47-1 ORA47-2	RCCGWCCW GCCGCC	Samples Samples	0.739:0.375 0.739:0.375
DEAR3	At2g23340	AP2/EREBP (DREB)	DEAR3-1	GCCGCC	Anatomy	-0.809:-0.613

			DEAR3-2	ACCGAC	Anatomy	-0.809:-0.613
DEAR4	At4g36900	AP2/EREBP (DREB)	DEAR4-1	CRCCGACA	Anatomy	-0.752:-0.589
			DEAR4-2	GCCGCC	Anatomy	-0.752:-0.589
TOE1	At2g28550	AP2/EREBP (AP2)	TOE1-1	CCTCGT	Perturbations	-0.567:-0.433
			TOE1-2	ACCTACG	Samples	-0.574:-0.385
			TOE1-3	AACCTAA	Samples	-0.574:-0.385
TOE2	At5g60120	AP2/EREBP (AP2)	TOE2-1	CCTCGT	Perturbations	-0.482:-0.318
			TOE2-2	ACCTACG	Samples	-0.559:-0.415
			TOE2-3	AACCTAA	Samples	-0.559:-0.415
WOX13	At4g35550	HD (WUS-like)	WOX13-1	TCAATCA	Perturbations	0.463:0.276
			WOX13-2	TTAATTAA	Perturbations	0.463:0.276
ICU4	At1g52150	HD (HD-ZIPIII)	ICU4	TAATnATTA	Perturbations	0.742:0.381
ATHB12	At3g61890	HD (HD-ZIPI)	ATHB12	CAATsATT	Perturbations	0.736:0.365
ATHB51	At5g03790	HD (HD-ZIPI)	ATHB51	CAATwATT	Perturbations	-0.273:-0.162
WRKY18	At4g31800	WRKY	WRKY18	CTTGACYR	Anatomy	-0.641:-0.489
WRKY12	At2g44745	WRKY	WRKY12	CGTTGACY	Samples	-0.513:-0.363
WRKY45	At3g01970	WRKY	WRKY45	CGTTGACY	Perturbations	0.697:0.495
WRKY38	At5g22570	WRKY	WRKY38	CGTTGACY	Perturbations	0.756:0.403
SPL1	At2g47070	SBP	SPL1	CGTAC	Samples	0.545:0.391
SPL7	At5g18830	SBP	SPL7	CGTAC	Anatomy	0.710:0.487
ZAT2	At2g17180	Zn finger (C2H2)	ZAT2	CAGCT	Samples	-0.569:-0.459
At4g35610 <sup>d</sup>	At4g35610	Zn finger (C2H2)				
ZAT6	At5g04340	Zn finger (C2H2)	ZAT6	ACACTA	Samples	-0.606:-0.477
ZAT14	At5g03510	Zn finger (C2H2)	ZAT14-1 ZAT14-2	AGTnnACT AGTnnnACT	Perturbations Perturbations	-0.360:-0.207 -0.360:-0.207
ZAT18	At3g53600	Zn finger (C2H2)	ZAT18-1 ZAT18-2	AGTnnACT AGTnnnACT	Perturbations Perturbations	-0.549:-0.404 -0.549:-0.404
GATA12	At5g25830	Zn finger (C2C2-GATA)	GATA12	AGATCT	Perturbations	0.482:0.286
DAG2	At5g65590	Zn finger (C2C2-DOF)	DAG2	AAAAAGTG	Samples	0.644:0.386
DOF5.7	At5g65590	Zn finger (C2C2-DOF)	DOF5.7-1 DOF5.7-2	AAAAGGG GTAAACG	Anatomy Anatomy	-0.732:-0.575 0.842:0.652
YAB1	At2g45190	Zn finger (C2C2-YABBY)	YAB1	AATnATAA	Samples	-0.633:-0.511
YAB5	At2g26580	Zn finger (C2C2-YABBY)	YAB5	AATnATTA	Anatomy	-0.736:-0.548
TCP15	At1g69690	TCP	TCP15	GGNCCCAC	Samples	0.732:0.559
TCP23	At1g35560	TCP	TCP23	GGNCCCAC	Perturbations	0.252:0.357
TCP16 <sup>d</sup>	At1g69690	TCP				
HSFB2a	At5g62020	HSF	HSFB2a-1 HSFB2a-2	GAAGCTTC TTCTAGAA	Perturbations Perturbations	0.671:0.479 0.671:0.479
HSFBC1	At3g24520	HSF	HSFBC1-1 HSFBC1-2	TTCTAGAA GAAGCTTC	Perturbations Perturbations	0.694:0.343 0.694:0.343
TGA2	At5g06950	bZIP	TGA2-1 TGA2-2	TGACGTCA TGACGTM	Perturbations Perturbations	0.560:0.404 0.560:0.404
bZIP60	At1g42990	bZIP	bZIP60-1 bZIP60-2	TGACGTCA TGACGTG	Perturbations Anatomy	0.785:0.628 0.811:0.645
ETTIN	At2g33860	B3 (ARF)	ETT-1 ETT-2	TGTCGG TGTCGA	Perturbations Perturbations	-0.643:-0.489 -0.643:-0.489
REM1	At4g31610	B3 (REM)	REM1-1 REM1-2	WGNTGTAG CGTGTAG	Perturbations Perturbations	0.555:0.189 0.555:0.189
ANAC55	At3g15500	NAC	ANAC55-1 ANAC55-2	TTACGTGT TTACGTAA	Samples Samples	0.805:0.502 0.805:0.502

ANAC46	At3g04060	NAC	ANAC46	TTGCGTGT	Samples	0.741:0.540
ANAC58	At3g18400	NAC	ANAC58	TTGCGTG	Samples	0.876:0.664
AHL20	At4g14465	AT-Hook	AHL20-1	AATATATT	Anatomy	0.828:0.582
			AHL20-2	AATTAATT	Perturbations	0.410:0.242
			AHL20-3	AAAATAAT	Perturbations	0.410:0.242
AHL12	At1g63480	AT-Hook	AHL12-1 AHL12-2 AHL12-3	AAATATTT ATAATTAT AATTAATT	Samples Samples Samples	-0.677:-0.490 -0.677:-0.490 -0.677:-0.490
AHL25 <sup>d</sup>	At4g35390	AT-Hook				
LBD16	At2g42430	LOB/AS2	LBD16	TCCGGA	Samples	-0.668:-0.537
PIF3	At1g09530	bHLH	PIF3	CCACGTGG	Samples	0.595:0.391
STY1 <sup>d</sup>	At3g51060	SRS				
MYC2 <sup>e</sup>	At1g32640	bHLH	MYC2-1 MYC2-2 MYC2-3 MYC2-4 MYC2-5	CACGTG CATGTG AACGTG CACGAG CACGCG	Perturbations Perturbations Perturbations Samples Samples	0.730:0.350 0.730:0.350 0.730:0.350 0.632:0.405 0.632:0.405
			MYC3-1 MYC3-2 MYC3-3 MYC3-4 MYC3-5	CACGTG CATGTG AACGTG CACGAG CACGCG	Samples Samples Samples Samples Samples	0.572:0.439 0.572:0.439 0.572:0.439 0.572:0.439 0.572:0.439
			MYC4-1 MYC4-2 MYC4-3	CACGTG CATGTG AACGTG	Samples Samples Samples	0.502:0.333 0.502:0.333 0.502:0.333
PIF4 <sup>g</sup>	At2g43010		PIF4-1 PIF4-2	CACGTG CATGTG	Perturbations Perturbations	0.587:0.381 0.587:0.381
PIF5 <sup>g</sup>	At3g59060		PIF5	CACGTG	Perturbations	0.577:0.348

<sup>a</sup> Corresponds to the elements shown in Fig. 3.

<sup>b</sup> Corresponds to Genevestigator's dataset for co-expressed genes.

<sup>c</sup> *r* range indicates the minimum and maximum Pearson's coefficients of co-expressed genes in the chosen dataset.

<sup>d</sup> Genes not represented in ATH1 microarrays and for which there is no co-regulation data.

<sup>e</sup> ERF1 and MYC2 binding motifs were previously determined (32).

<sup>f</sup> MYC3 and MYC4 binding motifs were previously determined (115).

<sup>g</sup> PIF4 and PIF4 binding motifs were previously determined (114).

**5. Table S2.** E-scores, rank positions and frequency PWMa corresponding to top-scoring elements for the TFs determined. Last column indicated the PBM design employed for each TF.

	Sequence	Rank position	E-score	Z-score	Frequency PWM												Design											
MYB52	CCGTTAGT	1°	0.46437	6.9707	A:	0.6416	0.4932	0.2375	0.1039	0.9249	0.8660	0.0169	0.0891	0.1294	0.4551	PBM10												
	GTTAGTTG	4°	0.43818	6.5272	C:	0.0781	0.3006	0.5198	0.1067	0.0134	0.0970	0.9237	0.0530	0.0225	0.1844	G:	T:	0.1681	0.1471	0.1025	0.0352	0.0318	0.0139	0.0312	0.7888	0.7878	0.0919	0.2686
MYB59	GTTAGGTA	1°	0.44624	5.6001	A:	0.0788	0.0281	0.0211	0.0138	0.7329	0.0221	0.0128	0.0185	0.2358	0.1511	nPBM11												
	GTTAGGTA	1°	0.49364	9.3332	C:	0.4849	0.0219	0.0089	0.0257	0.1518	0.0313	0.0414	0.0266	0.0506	0.4545	G:	T:	0.2624	0.9307	0.0478	0.0133	0.0305	0.9168	0.0327	0.0650	0.6223	0.1663	0.2281
MYB46	GTTAGGTA	1°	0.47897	9.3618	A:	0.5918	0.3573	0.0171	0.0308	0.8737	0.0166	0.0240	0.0386	0.5522	0.3282	nPBM11												
	GGTAGGTTG	9° 1° reranked	0.47897	9.3618	C:	0.0481	0.0143	0.0071	0.0143	0.0390	0.0079	0.0108	0.0404	0.2673	0.2796	G:	T:	0.1611	0.5484	0.1141	0.0058	0.0166	0.9564	0.9475	0.0046	0.0228	0.1561	0.2361
MYB111	GGTAGGTA	1°	0.47205	14.6751	A:	0.4193	0.0809	0.0033	0.0065	0.8362	0.0036	0.0047	0.0021	0.5802	0.2557	PBM10												
	GGTAGTTG	11°	0.44870	10.7159	C:	0.1593	0.0027	0.0015	0.0070	0.0299	0.0035	0.0027	0.0328	0.0238	0.2227	G:	T:	0.2170	0.9111	0.3874	0.0014	0.0062	0.9911	0.8979	0.0013	0.3584	0.3726	0.1490
MYB55	CGGTAGGT	1°	0.49297	13.2622	A:	0.2839	0.2289	0.0046	0.0031	0.8743	0.0178	0.0062	0.0059	0.5253	0.0729	nPBM11												
	GGTAGGTTG	13°	0.47949	10.2486	C:	0.0899	0.0045	0.0020	0.0117	0.0067	0.0039	0.0082	0.0039	0.0302	0.3242	G:	T:	0.2829	0.7566	0.6364	0.0041	0.0099	0.9762	0.8576	0.0049	0.3897	0.4556	0.1472
CCA1	AGATATTG	1°	0.49952	16.1177	A:	0.1124	0.9800	0.0484	0.0024	0.2266	0.9788	0.2667	0.0022	0.1185	0.2491	PBM10												
					C:	0.3593	0.0013	0.9460	0.9894	0.0490	0.0008	0.7299	0.9579	0.0538	0.2766	G:	T:	0.0504	0.0156	0.0020	0.0024	0.0231	0.0161	0.0010	0.0011	0.7562	0.3004	0.1739
					A:	0.0812	0.0824	0.0169	0.0081	0.1056	0.0068	0.0061	0.0186	0.1005	0.1413	nPBM11												
					C:	0.5390	0.0090	0.0094	0.0765	0.0314	0.0041	0.0126	0.0047	0.0466	0.1121	G:	T:	0.1405	0.8910	0.6977	0.0059	0.6911	0.9809	0.4448	0.0129	0.8280	0.6142	0.1323
					A:	0.2393	0.0175	0.2760	0.9096	0.1719	0.0082	0.5365	0.9638	0.0249	0.0004	0.1010	0.0010	0.0011	0.9599	0.0021	0.9856	0.0038	0.0908	0.2054	0.1130	0.5104		

	AGATATGA	56° 1° reranked	0.47128	7.8426	A: C: G: T:	0.4585 0.1203 0.2095 0.2117	0.9135 0.0151 0.0438 0.0276	0.0128 0.0063 0.9532 0.0277	0.9703 0.0083 0.0088 0.0126	0.0133 0.0050 0.0136 0.9680	0.9326 0.0338 0.0042 0.0294	0.0183 0.3447 0.0129 0.6241	0.4690 0.0092 0.5101 0.0117	0.4060 0.2577 0.2310 0.1053	0.2968 0.1364 0.3476 0.2193	
RVE1	AGATATTT	1°	0.49973	17.5660	A: C: G: T:	0.6476 0.0449 0.0599 0.2476	0.9110 0.0046 0.0238 0.0606	0.9237 0.0017 0.0684 0.0062	0.9750 0.0046 0.0079 0.0125	0.0651 0.0006 0.0382 0.0016	0.9845 0.0126 0.0014 0.9862	0.0021 0.0107 0.0010 0.0016	0.0105 0.0107 0.0010 0.0129	0.0171 0.0762 0.0005 0.0129	0.0888 0.1036 0.0427 0.4338	PBM10
	TAGATATT	15° 1° reranked	0.47860	8.2920	A: C: G: T:	0.2038 0.3055 0.2886 0.2020	0.5464 0.3092 0.0369 0.1075	0.6523 0.0891 0.1565 0.1020	0.2328 0.0267 0.1511 0.5894	0.9090 0.0410 0.0306 0.0194	0.0251 0.0261 0.0367 0.9120	0.0599 0.8901 0.0364 0.0136	0.2974 0.1988 0.0462 0.4576	0.6087 0.1645 0.1319 0.0950	0.4951 0.1575 0.0650 0.2825	
ARR11	AAGATACG	1°	0.49603	10.3663	A: C: G: T:	0.4779 0.1149 0.2161 0.1911	0.6250 0.0510 0.0548 0.2692	0.9270 0.0113 0.0294 0.0323	0.0047 0.0022 0.0042 0.0110	0.9894 0.0021 0.0042 0.0044	0.0074 0.0032 0.0019 0.9874	0.7317 0.0222 0.0036 0.2425	0.0034 0.8329 0.0018 0.1618	0.0135 0.0053 0.9684 0.0128	0.1877 0.2175 0.5042 0.0906	nPBM11
	CGGATCCG	19°	0.47938	5.7700	A: C: G: T:	0.2918 0.1950 0.2703 0.2429	0.1649 0.5099 0.0870 0.2381	0.2260 0.1050 0.6368 0.0322	0.0940 0.0160 0.5562 0.3339	0.8331 0.0110 0.0118 0.1442	0.1442 0.0118 0.0110 0.8331	0.3339 0.5562 0.6368 0.0940	0.0322 0.6368 0.1050 0.2260	0.2381 0.0870 0.5099 0.1649	0.1991 0.4236 0.1617 0.2155	
	AAGATCTT	55° 1° reranked	0.45964	12.1170	A: C: G: T:	0.3441 0.1337 0.2184 0.3037	0.6794 0.0658 0.1166 0.1382	0.8353 0.0193 0.0659 0.0795	0.0492 0.0164 0.0193 0.1998	0.9523 0.0081 0.0193 0.0203	0.0203 0.0193 0.0081 0.9523	0.1998 0.7345 0.0164 0.0492	0.0795 0.0659 0.0193 0.8353	0.1382 0.1166 0.0658 0.6794	0.2204 0.1972 0.3066 0.2758	
ARR14	AGATACGG	1°	0.49510	13.6586	A: C: G: T:	0.5318 0.0938 0.1234 0.2510	0.9454 0.0051 0.0257 0.0238	0.0021 0.0008 0.9842 0.0129	0.9906 0.0010 0.0033 0.0050	0.0020 0.0028 0.0021 0.9931	0.5065 0.1538 0.0010 0.3387	0.0012 0.9595 0.0011 0.0381	0.0095 0.0047 0.9596 0.0262	0.0830 0.4840 0.4006 0.0324	0.1940 0.1161 0.0822 0.6078	nPBM11
	CGGATCCG	8°	0.49098	11.6216	A: C: G: T:	0.1847 0.2602 0.3836 0.1715	0.2553 0.5444 0.0790 0.1213	0.3570 0.0623 0.5704 0.0103	0.1290 0.0056 0.6270 0.2384	0.9237 0.0059 0.0124 0.0581	0.0581 0.1270 0.0056 0.9237	0.2384 0.5704 0.0623 0.1290	0.0103 0.5704 0.3570 0.2553	0.1213 0.0790 0.5444 0.1749	0.1572 0.4783 0.1895 0.1749	
	AAGATCTT	6° 1° reranked	0.49473	13.8795	A: C: G: T:	0.1044 0.3838 0.4199 0.0919	0.7710 0.0770 0.1363 0.0157	0.9027 0.0099 0.0834 0.0040	0.0325 0.0006 0.0140 0.0273	0.9812 0.0017 0.0140 0.0032	0.0032 0.0140 0.0017 0.9812	0.0273 0.9396 0.0006 0.0325	0.0040 0.0834 0.0099 0.9027	0.0157 0.1363 0.0770 0.7710	0.0966 0.4766 0.3777 0.0492	
GLK1	GGATATCC	1°	0.47608	6.7367	A: C: G: T:	0.2504 0.4519 0.2138 0.0839	0.2497 0.0357 0.6268 0.0879	0.0231 0.0373 0.9198 0.0198	0.9490 0.0100 0.0361 0.0049	0.2621 0.0038 0.0853 0.6487	0.6487 0.0853 0.0038 0.2621	0.0049 0.0361 0.0100 0.9490	0.0198 0.9198 0.0373 0.0231	0.0879 0.6268 0.0357 0.2497	0.3774 0.1311 0.1900 0.3016	PBM10
	AGAATCTG	11° 1° reranked	0.44726	11.4129	A: C: G: T:	0.3020 0.1706 0.4508 0.0767	0.8273 0.1176 0.0444 0.0107	0.1362 0.0318 0.7657 0.0663	0.9198 0.0112 0.0229 0.0461	0.7755 0.0059 0.0045 0.2141	0.0081 0.0135 0.0137 0.9647	0.0249 0.9371 0.0078 0.0303	0.0309 0.1018 0.0410 0.8263	0.2205 0.0498 0.5366 0.1931	0.2399 0.0880 0.2541 0.4179	
KAN4	GAATATTTC	1°	0.49936	14.2634	A: C: G: T:	0.3646 0.1105 0.5008 0.0240	0.0106 0.0101 0.9652 0.0141	0.9408 0.0120 0.0262 0.0210	0.9053 0.0015 0.0015 0.0918	0.0024 0.0027 0.0027 0.8933	0.8933 0.0015 0.0106 0.0024	0.0918 0.0129 0.0015 0.9053	0.0210 0.0262 0.0120 0.9408	0.0141 0.9652 0.0101 0.0106	0.1133 0.2786 0.2198 0.3883	nPBM11

	GGAATATT	2°	0.49392	8.5828	A: C: G: T:	0.0106 0.0101 0.9652 0.0141	0.7560 0.0463 0.1269 0.0708	0.6644 0.0035 0.0028 0.3293	0.0145 0.1933 0.0064 0.7858	0.9531 0.0028 0.0408 0.0033	0.0214 0.0058 0.0022 0.9706	0.0360 0.0193 0.0105 0.9342	0.0081 0.9732 0.0061 0.0126	0.0091 0.5744 0.1228 0.2937	0.1326 0.4408 0.1288 0.2978	
KAN1	CGAATAT.T	1°	0.39225	2.2231	A: C: G: T:	0.1504 0.2683 0.3342 0.2471	0.4331 0.1293 0.2373 0.2003	0.2591 0.2133 0.2851 0.2425	0.6176 0.0160 0.0433 0.3231	0.1279 0.1918 0.0575 0.6228	0.8982 0.0140 0.0682 0.0196	0.0543 0.0868 0.1469 0.8396	0.0357 0.0868 0.1469 0.6480	0.0633 0.8758 0.1933 0.0435	0.0545 0.4716 0.2806	nPBM11
At5g28300	ACGGTAAA	1°	0.48111	10.6604	A: C: G: T:	0.2393 0.3763 0.2090 0.1755	0.5761 0.0143 0.3786 0.0311	0.0112 0.9317 0.0260 0.0311	0.0900 0.0871 0.7478 0.0751	0.0055 0.0069 0.9826 0.0050	0.0062 0.0073 0.0032 0.9832	0.7493 0.0128 0.1279 0.1100	0.9507 0.0127 0.0178 0.0188	0.6082 0.1545 0.0419 0.1953	0.2881 0.0527 0.0193 0.6398	nPBM11
RRTF1	GGCGGCCGC	1°	0.49501	17.2081	A: C: G: T:	0.2417 0.1469 0.4593 0.1521	0.0385 0.2871 0.0600 0.0132	0.0499 0.9357 0.9891 0.0064	0.0031 0.0046 0.1685 0.0032	0.0027 0.0046 0.0025 0.0022	0.0128 0.8266 0.9785 0.0029	0.0040 0.9818 0.0134 0.0050	0.0382 0.0126 0.0134 0.0865	0.2066 0.8620 0.0025 0.0178	0.6909 0.7732 0.2100 0.0783	PBM11
	CGCCGCAA	3°	0.49436	16.4920	A: C: G: T:	0.0705 0.4599 0.4311 0.0385	0.0991 0.8672 0.0215 0.0121	0.0034 0.9744 0.9870 0.0035	0.0047 0.0127 0.0166 0.0044	0.0127 0.9804 0.0038 0.0031	0.0042 0.0053 0.9849 0.0057	0.0278 0.9078 0.0177 0.0467	0.4442 0.5424 0.0037 0.0096	0.9027 0.0185 0.0273 0.0515	0.2655 0.1702 0.0831 0.4812	
	CGCCGTCA	13°	0.48474	12.4036	A: C: G: T:	0.0602 0.3015 0.5973 0.0409	0.0524 0.9192 0.0181 0.0103	0.0026 0.0037 0.9891 0.0047	0.0043 0.9676 0.0253 0.0028	0.0187 0.9723 0.0136 0.0063	0.0085 0.0084 0.9785 0.0046	0.0624 0.8110 0.0136 0.1130	0.1709 0.7661 0.0043 0.0588	0.8095 0.0401 0.0769 0.0736	0.2143 0.0811 0.0319 0.6727	
RAP2.3	GCGCCGCC	1°	0.49319	12.9932	A: C: G: T:	0.2537 0.1326 0.4639 0.1497	0.1186 0.2315 0.5426 0.1073	0.1470 0.7614 0.0333 0.0583	0.0066 0.0065 0.9822 0.0047	0.0059 0.7465 0.2438 0.0038	0.0164 0.9770 0.0038 0.0028	0.0058 0.0028 0.9765 0.0149	0.1088 0.7802 0.0257 0.0853	0.2572 0.6640 0.0056 0.0731	0.4874 0.0945 0.1682 0.2499	nPBM11
	CGCCGCAA	14°	0.48723	11.0046	A: C: G: T:	0.0934 0.2870 0.5730 0.0465	0.3417 0.5656 0.0363 0.0564	0.0038 0.0034 0.9886 0.0042	0.0046 0.9773 0.0278 0.0076	0.0098 0.0137 0.0091 0.0039	0.0057 0.0137 0.9724 0.0083	0.0240 0.8660 0.0300 0.0800	0.2807 0.6692 0.0052 0.0450	0.6224 0.0703 0.1078 0.1994	0.3836 0.1667 0.0925 0.3572	
	CGCCGACA	26°	0.48339	10.4428	A: C: G: T:	0.1356 0.1638 0.6508 0.0498	0.2073 0.7029 0.0284 0.0614	0.0048 0.0019 0.9897 0.0036	0.0023 0.9751 0.0124 0.0101	0.0817 0.9145 0.0017 0.0021	0.0067 0.2288 0.7584 0.0061	0.1570 0.6985 0.0402 0.1043	0.0778 0.9037 0.0017 0.0168	0.7283 0.0364 0.0601 0.1752	0.3114 0.1704 0.1804 0.3378	
RAP2.6	GCGCCGCC	1°	0.49787	23.8053	A: C: G: T:	0.1209 0.1856 0.4810 0.2125	0.0345 0.2709 0.6809 0.0137	0.0271 0.9579 0.0070 0.0080	0.0071 0.0023 0.9889 0.0017	0.0015 0.8940 0.1035 0.0009	0.0119 0.9849 0.0020 0.0011	0.0041 0.0032 0.9900 0.0027	0.0198 0.9468 0.0137 0.0198	0.0662 0.9213 0.0025 0.0100	0.3025 0.0252 0.6112 0.0612	PBM10
	CCGCCGCA	12° 2° reranked	0.48941	12.2946	A: C: G: T:	0.3771 0.3623 0.0726 0.1880	0.5628 0.0577 0.2905 0.0890	0.0526 0.0043 0.9352 0.0079	0.0159 0.9461 0.0173 0.0208	0.0121 0.9360 0.0448 0.0071	0.0324 0.0050 0.9546 0.0080	0.0517 0.7740 0.0710 0.1033	0.4908 0.4565 0.0135 0.0393	0.6756 0.1515 0.0118 0.1610	0.3390 0.2354 0.0109 0.4147	
	CGCCGTCA	10° 1° reranked	0.49232	15.5485	A: C: G: T:	0.2195 0.4987 0.0878 0.1940	0.6904 0.0947 0.0794 0.1355	0.1341 0.0242 0.1296 0.7121	0.1421 0.0800 0.0697 0.1341	0.6801 0.9754 0.0037 0.1702	0.0103 0.0084 0.9603 0.0106	0.0031 0.0084 0.9689 0.0282	0.0105 0.0101 0.9620 0.0106	0.0069 0.9620 0.0030 0.0281	0.4988 0.1678 0.0172 0.3162	

ATERF1	GCGCCGCC	1°	0.47224	4.7899	A: C: G: T:	0.0466 0.0280 0.8395 0.0859	0.0428 0.1456 0.7099 0.1017	0.1867 0.6024 0.1468 0.0641	0.0315 0.0177 0.9330 0.0177	0.0398 0.4832 0.3773 0.0998	0.0583 0.8480 0.0503 0.0434	0.0648 0.0922 0.7340 0.1090	0.0789 0.5150 0.3472 0.0589	0.0658 0.8004 0.0968 0.0370	0.0921 0.3692 0.2905 0.2482	PBM10	
	TGCCGGCA	2°	0.47284	4.6237	A: C: G: T:	0.1350 0.3323 0.2375 0.2952	0.1404 0.1474 0.1157 0.5965	0.0439 0.0025 0.9417 0.0119	0.0290 0.8983 0.0597 0.0131	0.0094 0.9455 0.0071 0.0380	0.0380 0.0071 0.9455 0.0094	0.0131 0.0071 0.8983 0.0290	0.0119 0.9417 0.0025 0.0439	0.5965 0.1157 0.1474 0.1404	0.3522 0.1934 0.2918 0.1627		
ERF1	GCGCCGCC	1°	0.48175	7.7086	A: C: G: T:	0.3760 0.1163 0.2803 0.2273	0.0646 0.3291 0.4613 0.1450	0.1465 0.8250 0.0093 0.0193	0.0103 0.0062 0.9743 0.0092	0.0071 0.9578 0.0253 0.0097	0.0380 0.9511 0.0036 0.0073	0.0178 0.0023 0.9711 0.0088	0.0478 0.0144 0.0230 0.0275	0.0144 0.9704 0.0052 0.0100	0.5834 0.0518 0.2430 0.1218	PBM1	
DDF1	GCCGACAT	1°	0.49956	33.8394	A: C: G: T:	0.1900 0.3400 0.2657 0.2044	0.9782 0.0020 0.0152 0.0045	0.0089 0.0150 0.0034 0.9726	0.0055 0.0009 0.9907 0.0030	0.0031 0.0027 0.0018 0.9924	0.0027 0.9917 0.0013 0.0044	0.0023 0.0016 0.9921 0.0041	0.1694 0.0005 0.8285 0.0016	0.0008 0.7901 0.0012 0.2079	0.2764 0.3059 0.2897 0.1281	nPBM11	
DREB2C	GCGGCCGC	1°	0.49420	19.3286	A: C: G: T:	0.1400 0.0448 0.6717 0.1436	0.0271 0.0155 0.9310 0.0263	0.0045 0.9599 0.0050 0.0305	0.0058 0.0473 0.9426 0.0042	0.0216 0.0044 0.9719 0.0020	0.0020 0.9719 0.0044 0.0216	0.0042 0.9426 0.00473 0.0058	0.0305 0.0050 0.9599 0.0045	0.0263 0.9310 0.0155 0.0271	0.1255 0.6703 0.0699 0.1343	nPBM11	
	CACCGACA	2°	0.49263	16.0199	A: C: G: T:	0.0902 0.5152 0.2713 0.1232	0.1097 0.7566 0.1209 0.0128	0.8088 0.0034 0.1851 0.0027	0.0042 0.9847 0.0030 0.0080	0.0074 0.9856 0.0032 0.0038	0.0083 0.0034 0.9841 0.0041	0.7758 0.1583 0.0470 0.0190	0.0053 0.9773 0.0048 0.0126	0.8063 0.0503 0.1123 0.1167	0.3039 0.0893 0.1123 0.4945		
At1g77200	CACCGACA	1°	0.49734	25.2793	A: C: G: T:	0.0617 0.5572 0.1512 0.2300	0.0801 0.7551 0.0401 0.1247	0.7462 0.9927 0.2516 0.0010	0.0019 0.9933 0.0011 0.0043	0.0027 0.9933 0.0012 0.0028	0.0037 0.0013 0.9930 0.0021	0.9888 0.0024 0.9930 0.0075	0.0022 0.9922 0.0013 0.0047	0.8879 0.0025 0.0010 0.0086	0.2519 0.3235 0.1010 0.2176	nPBM11	
ORA47	ACCGACCA	1°	0.49732	22.5761	A: C: G: T:	0.1309 0.6570 0.0287 0.1835	0.5810 0.0008 0.4123 0.0059	0.0012 0.9928 0.0022 0.0038	0.0064 0.9904 0.0019 0.0013	0.0024 0.0017 0.9934 0.0024	0.7182 0.0083 0.0652 0.2082	0.0028 0.9914 0.0010 0.0048	0.0437 0.9431 0.0009 0.0122	0.7353 0.0104 0.0698 0.1844	0.6369 0.1337 0.0533 0.1762	PBM10	
	GCGCCGCC	40° 1° reranked	0.48761	16.3467	A: C: G: T:	0.1742 0.0548 0.3041 0.4669	0.0949 0.1301 0.7284 0.0466	0.0148 0.9315 0.0091 0.0447	0.0676 0.0064 0.9207 0.0054	0.0044 0.9624 0.0295 0.0037	0.0131 0.9761 0.0083 0.0026	0.0112 0.0020 0.9748 0.0120	0.2500 0.6505 0.0481 0.0514	0.0403 0.9265 0.0084 0.0248	0.2057 0.2499 0.1266 0.4177		
DEAR4	CACCGACA	1°	0.49913	26.9326	A: C: G: T:	0.1834 0.3999 0.2742 0.1425	0.0221 0.9554 0.0110 0.0114	0.7950 0.0004 0.2022 0.0024	0.0023 0.9932 0.0023 0.0008	0.0182 0.9804 0.0006 0.0031	0.0013 0.0029 0.9927 0.0083	0.9849 0.0062 0.0009 0.0080	0.0038 0.9872 0.0004 0.0087	0.8483 0.1218 0.0122 0.0177	0.1218 0.3247 0.0735 0.4800	PBM10	
	GCGCCGCC	25° 1° reranked	0.49714	18.5208	A: C: G: T:	0.2883 0.2214 0.3659 0.1244	0.0327 0.1275 0.7566 0.0832	0.0218 0.9419 0.0140 0.0223	0.1166 0.0050 0.8716 0.0068	0.0057 0.9119 0.0089 0.0043	0.0041 0.9806 0.0555 0.0064	0.0092 0.0069 0.9520 0.0319	0.0079 0.8381 0.0851 0.0690	0.0331 0.8593 0.0158 0.0919	0.0365 0.2076 0.1467 0.6092		
DEAR3	CGCCGCC	1°	0.46169	9.7497	A: C: G: T:	0.2013 0.2267 0.3265 0.2455	0.1479 0.7871 0.0241 0.0408	0.3349 0.0146 0.6335 0.0170	0.0273 0.8887 0.0578 0.0262	0.0390 0.9336 0.0192 0.0081	0.0656 0.0435 0.8649 0.0260	0.3201 0.5238 0.0555 0.1007	0.0695 0.8276 0.0685 0.0344	0.1001 0.4170 0.3292 0.1537	0.3182 0.2342 0.1508 0.2968	PBM10	

	GAACCGAC	2°	0.45890	9.5059	A: C: G: T:	0.1117 0.3343 0.2753 0.2786	0.1192 0.1644 0.5961 0.1203	0.5923 0.0948 0.1139 0.1989	0.8815 0.0111 0.0930 0.0144	0.0133 0.9626 0.0097 0.0145	0.0272 0.9559 0.0083 0.0086	0.0315 0.0179 0.9405 0.0101	0.8210 0.0770 0.0620 0.0400	0.0179 0.9573 0.0045 0.0202	0.1554 0.3183 0.3998 0.1266	
TOE2	CCTCGTAC	1°	0.48690	6.1535	A: C: G: T:	0.6132 0.0232 0.0388 0.3248	0.0023 0.9134 0.0023 0.0820	0.1288 0.8455 0.0171 0.9848	0.0047 0.0049 0.0056 0.0086	0.0551 0.9222 0.0082 0.0145	0.0065 0.0050 0.0054 0.0054	0.0038 0.0034 0.9831 0.8876	0.9764 0.0110 0.1052 0.0105	0.0265 0.8328 0.0211 0.1272	0.1856 0.2395 0.0136 0.3379	nPBM11
	AACCTACG	6° 1° reranked	0.46711	5.0064	A: C: G: T:	0.4441 0.1099 0.0915 0.3545	0.8051 0.0187 0.0065 0.1697	0.8704 0.0026 0.0056 0.1214	0.0076 0.9413 0.0033 0.0479	0.1228 0.7900 0.0742 0.0130	0.0247 0.0058 0.0085 0.9611	0.7641 0.0238 0.1888 0.0233	0.0792 0.6816 0.1527 0.0865	0.0060 0.0047 0.9785 0.0108	0.6260 0.0433 0.0262 0.3045	
	AACCTTAA	7° 2° reranked	0.45869	4.9739	A: C: G: T:	0.2900 0.2318 0.0793 0.3989	0.6375 0.0282 0.0139 0.3204	0.7277 0.8692 0.0271 0.2369	0.0202 0.7285 0.0072 0.1034	0.1066 0.7285 0.1340 0.0309	0.0391 0.0098 0.0106 0.9405	0.0839 0.0961 0.0066 0.8134	0.9194 0.0134 0.0192 0.0480	0.7600 0.0106 0.1402 0.0892	0.3011 0.1561 0.0321 0.5107	
	CCTCGTAC	1°	0.48620	6.2366	A: C: G: T:	0.6858 0.0184 0.0173 0.2786	0.0065 0.8760 0.0015 0.1160	0.1287 0.8363 0.0185 0.0166	0.0074 0.0094 0.0082 0.9751	0.0464 0.9287 0.0165 0.0084	0.0072 0.0053 0.9785 0.0089	0.0112 0.0019 0.1830 0.8040	0.9639 0.0177 0.0016 0.0167	0.0467 0.8072 0.0207 0.0990	0.2337 0.3103 0.0471 0.2490	
TOE1	AACCTACG	10°	0.45083	4.3737	A: C: G: T:	0.4218 0.0498 0.3232 0.2052	0.7497 0.0299 0.0360 0.1845	0.7918 0.9169 0.6393 0.1865	0.0339 0.0125 0.0210 0.0282	0.1364 0.9169 0.2075 0.0168	0.0404 0.0270 0.0079 0.9247	0.6784 0.0165 0.2619 0.0432	0.0758 0.6300 0.2535 0.0407	0.0315 0.0090 0.9273 0.0322	0.9078 0.0251 0.0096 0.0576	PBM10
	AACCTTAA	19° 2° reranked	0.44985	9.2107	A: C: G: T:	0.2591 0.1355 0.1017 0.5036	0.7430 0.0261 0.0167 0.2153	0.8466 0.9282 0.0266 0.1159	0.0094 0.9282 0.1308 0.0358	0.0167 0.8310 0.1775 0.0215	0.0354 0.0293 0.0097 0.9374	0.0321 0.0293 0.0206 0.9180	0.9390 0.0319 0.0147 0.0144	0.8352 0.0112 0.1050 0.0486	0.4120 0.3454 0.0745 0.1681	
	ATCAATCA	1°	0.48472	13.2465	A: C: G: T:	0.3201 0.1563 0.1505 0.3731	0.5385 0.0278 0.1234 0.1169	0.0295 0.1234 0.0077 0.7237	0.0244 0.9329 0.0062 0.0350	0.9510 0.0149 0.0142 0.0279	0.9583 0.0076 0.0095 0.0199	0.0097 0.0125 0.0142 0.9683	0.0097 0.8644 0.0142 0.1116	0.9016 0.0275 0.0189 0.0519	0.3993 0.1536 0.2071 0.2400	
WOX13	TTAATTAA	15° 1° reranked	0.44159	9.7613	A: C: G: T:	0.2162 0.2314 0.1867 0.3656	0.1603 0.1549 0.1142 0.5707	0.1959 0.1660 0.0438 0.5943	0.8073 0.0588 0.0279 0.1060	0.8070 0.0224 0.0929 0.0776	0.0776 0.0929 0.0224 0.8070	0.1060 0.0279 0.0588 0.8073	0.5943 0.0438 0.1660 0.1959	0.5707 0.1142 0.1549 0.1603	0.2675 0.3382 0.2067 0.1875	nPBM11
ICU4	TAAT.ATTA	1°	0.44550	8.9370	A: C: G: T:	0.2033 0.0677 0.6407 0.0882	0.0932 0.1719 0.0596 0.6753	0.7218 0.1927 0.0325 0.0284	0.9259 0.0115 0.0165 0.9277	0.0461 0.0165 0.2823 0.3812	0.2743 0.0165 0.0096 0.0461	0.9277 0.0154 0.0325 0.9259	0.0262 0.0154 0.0325 0.7218	0.0284 0.0571 0.1927 0.0932	0.6753 0.0596 0.1719 0.0932	PBM10
ATHB12	CAATCATT	1°	0.49446	11.0272	A: C: G: T:	0.2163 0.1780 0.2353 0.3704	0.5788 0.1907 0.1066 0.1239	0.6677 0.0463 0.0529 0.2331	0.0183 0.0352 0.0055 0.9411	0.0484 0.2006 0.6331 0.1178	0.9651 0.0066 0.0152 0.0131	0.0092 0.0081 0.0171 0.9656	0.0297 0.0105 0.0193 0.9405	0.1281 0.0184 0.0193 0.0877	0.2440 0.1448 0.7658 0.0530	nPBM11
ATHB51	CAATAATT	1°	0.49835	12.5885	A: C: G: T:	0.1394 0.3797 0.1608 0.3201	0.6914 0.2533 0.0042 0.0511	0.9849 0.0047 0.0037 0.0067	0.0025 0.0078 0.0013 0.9884	0.4109 0.0062 0.0034 0.5795	0.9901 0.0018 0.0052 0.0029	0.0100 0.0016 0.0041 0.9843	0.0046 0.0183 0.0122 0.9649	0.1046 0.0176 0.0122 0.1005	0.3248 0.2581 0.2083 0.2088	PBM10

SPL1	CCGTACGG	1°	0.49943	27.2144	A: C: G: T:	0.3126 0.1336 0.2853 0.2686	0.0307 0.8802 0.0080 0.0811	0.0100 0.9239 0.0129 0.0532	0.0138 0.0020 0.9819 0.0023	0.0016 0.0083 0.0056 0.9845	0.9845 0.0056 0.0083 0.0016	0.0023 0.9819 0.0020 0.0138	0.0532 0.0129 0.9239 0.0100	0.0811 0.0080 0.8802 0.0307	0.6005 0.2152 0.0102 0.1741	nPBM11	
	CGTACGGA	2	0.49892	26.0242	A: C: G: T:	0.0307 0.8802 0.0080 0.0811	0.0538 0.6276 0.0470 0.2716	0.0148 0.0010 0.9832 0.0010	0.0017 0.0056 0.0017 0.9911	0.9915 0.0038 0.0038 0.0009	0.0027 0.9928 0.0018 0.0027	0.0069 0.0044 0.9855 0.0032	0.0131 0.0021 0.9788 0.0061	0.6005 0.2152 0.0102 0.1741	0.3361 0.4048 0.0936 0.1655		
SPL7	CGTACGAC	1°	0.40973	5.0465	A: C: G: T:	0.0800 0.4329 0.2425 0.2446	0.0509 0.7772 0.1227 0.0492	0.0419 0.0262 0.9120 0.0198	0.0429 0.0121 0.0894 0.8557	0.8854 0.0617 0.9219 0.0307	0.0387 0.0157 0.1057 0.0250	0.0651 0.2860 0.2079 0.0312	0.2860 0.1178 0.6138 0.1415	0.1178 0.3832 0.2532 0.2272	0.3832 0.2532 0.2433 0.1203	PBM10	
WRKY18	CGGTCAAG	1°	0.49100	12.3185	A: C: G: T:	0.2786 0.2169 0.1415 0.3630	0.1353 0.5755 0.0691 0.2201	0.1527 0.0171 0.0073 0.0043	0.0055 0.0072 0.0038 0.0124	0.0072 0.9835 0.0089 0.9846	0.0076 0.0089 0.0085 0.0056	0.9754 0.0898 0.0898 0.0111	0.9201 0.1057 0.0136 0.0578	0.0660 0.2079 0.5991 0.1395	0.1441 0.4955 0.2842 0.0762	PBM10	
WRKY12	CGTTGACC	1°	0.43367	6.2052	A: C: G: T:	0.2093 0.3519 0.1830 0.2558	0.1720 0.7070 0.0755 0.0456	0.0699 0.0327 0.8189 0.0785	0.0199 0.0619 0.0054 0.9128	0.0103 0.0153 0.0289 0.9455	0.0081 0.0057 0.0058 0.0326	0.9706 0.0958 0.9494 0.0136	0.0289 0.7224 0.0554 0.0163	0.0249 0.2071 0.1012 0.1514	0.2071 0.1866 0.3270 0.2793	nPBM11	
WRKY45	CGTTGACC	1°	0.47165	8.8638	A: C: G: T:	0.3009 0.1173 0.3194 0.2624	0.0601 0.8752 0.0251 0.0396	0.0152 0.0036 0.8921 0.0891	0.0065 0.0177 0.0018 0.9740	0.0042 0.0032 0.0111 0.9815	0.0041 0.0034 0.9853 0.0072	0.9883 0.0052 0.0023 0.0042	0.0058 0.9865 0.0027 0.0051	0.0077 0.5312 0.0111 0.4500	0.6446 0.5312 0.1191 0.1071	PBM10	
WRKY38	CGTTGACC	1°	0.44537	8.1772	A: C: G: T:	0.1874 0.3050 0.3137 0.1938	0.1515 0.6722 0.0727 0.1036	0.0686 0.0275 0.8491 0.0547	0.0255 0.0305 0.0207 0.9233	0.0080 0.0229 0.0391 0.9300	0.0093 0.0157 0.0111 0.0137	0.9546 0.0169 0.9613 0.0183	0.0159 0.9507 0.0102 0.0202	0.0150 0.6794 0.0132 0.2887	0.4030 0.1752 0.0169 0.1220	nPBM11	
ZAT2	CCAGCTGG		0.49238	8.0555	A: C: G: T:	0.2174 0.4302 0.1939 0.1586	0.0996 0.4434 0.2412 0.2158	0.0892 0.6896 0.1497 0.0715	0.0890 0.0024 0.0175 0.0812	0.0189 0.0308 0.9468 0.0035	0.0035 0.9468 0.0308 0.0189	0.0812 0.0175 0.0024 0.8990	0.0715 0.1497 0.6896 0.0892	0.2158 0.2412 0.4434 0.0996	0.1674 0.3759 0.3886 0.0680	PBM10	
At4g35610	TCAGCTGA	1°	0.41835	4.5065	A: C: G: T:	0.3061 0.1438 0.3628 0.1872	0.1252 0.0627 0.2582 0.5539	0.0766 0.6854 0.1285 0.1095	0.8215 0.0545 0.0217 0.1023	0.0312 0.0209 0.8727 0.0751	0.0751 0.8727 0.0209 0.0312	0.1023 0.0217 0.0545 0.8215	0.1095 0.1285 0.6854 0.0766	0.5539 0.2582 0.0627 0.1252	0.3809 0.3829 0.1468 0.0894	PBM10	
ZAT6	TAACACTA	1°	0.44627	3.7946	A: C: G: T:	0.1677 0.2411 0.1429 0.4482	0.0385 0.0421 0.0561 0.8633	0.8283 0.0367 0.0342 0.1008	0.8500 0.0387 0.0241 0.0872	0.0098 0.8916 0.0439 0.0547	0.8795 0.0088 0.0157 0.0961	0.0113 0.8908 0.0895 0.0984	0.0240 0.0179 0.0113 0.9469	0.8666 0.0529 0.0419 0.0386	0.2897 0.2851 0.2192 0.2060	PBM10	
Zat14	AGTGCACT	1°	0.48144	3.2324	A: C: G: T:	0.1228 0.5604 0.1105 0.2062	0.7407 0.0510 0.0557 0.1526	0.0124 0.3207 0.6471 0.0198	0.0468 0.0239 0.0245 0.9165	0.2777 0.0245 0.6246 0.0731	0.0731 0.6246 0.0245 0.2777	0.9165 0.0028 0.6471 0.0468	0.0198 0.1526 0.0557 0.0124	0.1526 0.3074 0.0510 0.0905	0.2163 0.3074 0.0510 0.0905	nPBM11	
ZAT18	AGTGCACT	1°	0.40066	4.2761	A: C: G: T:	0.3237 0.0267 0.3440 0.3057	0.7097 0.0383 0.0773 0.1747	0.0329 0.0534 0.8848 0.0288	0.0320 0.1833 0.0648 0.7200	0.0956 0.0562 0.8064 0.0418	0.0418 0.8064 0.0562 0.0956	0.7200 0.0648 0.1833 0.0320	0.0288 0.8848 0.0534 0.0329	0.1747 0.0773 0.0383 0.7097	0.5041 0.1223 0.2490 0.1246	PBM10	

GATA12	TAGATCTA	1°	0.46065	5.6236	A: C: G: T:	0.1502 0.2090 0.2219 0.4188	0.1896 0.2825 0.1459 0.3820	0.5992 0.1119 0.2299 0.0590	0.1889 0.0673 0.7383 0.0055	0.9685 0.0024 0.0219 0.0072	0.0072 0.0219 0.0024 0.9685	0.0055 0.7383 0.0673 0.1889	0.0590 0.2299 0.1119 0.5992	0.3820 0.1459 0.2825 0.1896	0.2834 0.2804 0.1801 0.2561	nPBM11
YAB1	AATCATAA	1°	0.46507	5.4610	A: C: G: T:	0.0886 0.4654 0.1067 0.3393	0.6410 0.0335 0.0427 0.2828	0.8559 0.0297 0.0243 0.0900	0.0441 0.0193 0.0361 0.9005	0.2909 0.4245 0.1306 0.1540	0.9114 0.0072 0.0468 0.0345	0.0462 0.0297 0.0428 0.8813	0.7291 0.0297 0.1138 0.0744	0.7719 0.0503 0.0825 0.0952	0.1430 0.1875 0.1015 0.5680	PBM10
YAB5	AATGATTA	1°	0.44299	3.1541	A: C: G: T:	0.2784 0.2292 0.2222 0.2702	0.5271 0.0800 0.1840 0.2089	0.7770 0.0862 0.0546 0.0823	0.0715 0.1632 0.0292 0.7362	0.1751 0.2371 0.5365 0.0512	0.7831 0.0451 0.0255 0.1462	0.1050 0.1794 0.0535 0.6621	0.1502 0.0407 0.1000 0.7091	0.7320 0.1656 0.0626 0.0399	0.2919 0.2579 0.1562 0.2939	PBM10
DAG2	AAAAAGTG	1°	0.49472	16.8533	A: C: G: T:	0.1637 0.4127 0.3138 0.1097	0.7560 0.0255 0.0112 0.0364	0.5978 0.0012 0.0008 0.3889	0.9903 0.0008 0.0008 0.0054	0.9858 0.0008 0.0014 0.0058	0.7438 0.0077 0.0014 0.0013	0.0040 0.0077 0.0009 0.0181	0.0077 0.2382 0.1985 0.0181	0.2382 0.1767 0.0445 0.1601	0.3865 0.1767 0.0445 0.1276	nPBM11
Dof5.7	AAAAAGGG	1°	0.47446	9.0255	A: C: G: T:	0.3949 0.2086 0.2712 0.1253	0.4484 0.0865 0.2240 0.2410	0.6837 0.0123 0.0608 0.2432	0.9705 0.0043 0.0096 0.0156	0.8989 0.0034 0.0787 0.0190	0.3631 0.0424 0.5889 0.0056	0.4847 0.0610 0.4434 0.0109	0.1018 0.1912 0.4601 0.2469	0.1740 0.2147 0.4079 0.2035	0.2946 0.2179 0.3072 0.1803	nPBM11
	ACGTTAAC	12° 1° reranked	0.47579	8.0289	A: C: G: T:	0.4414 0.0832 0.2360 0.2394	0.4706 0.1475 0.1264 0.2556	0.0249 0.9378 0.0181 0.0192	0.0160 0.0027 0.0196 0.0378	0.0160 0.0093 0.0196 0.9551	0.0120 0.0168 0.0158 0.9554	0.8868 0.0257 0.0311 0.0564	0.8641 0.0533 0.0314 0.0512	0.0440 0.8963 0.0082 0.0515	0.1187 0.1074 0.5249 0.2491	
TCP15	GGGCCAC	1°	0.49928	26.1363	A: C: G: T:	0.0126 0.0517 0.8068 0.1289	0.0056 0.0024 0.9781 0.0140	0.0047 0.0034 0.9894 0.0025	0.0601 0.1385 0.7714 0.0301	0.0085 0.9496 0.0058 0.0361	0.0020 0.9901 0.0013 0.0066	0.0031 0.9861 0.0020 0.0087	0.9488 0.0121 0.0365 0.0026	0.0058 0.9619 0.0204 0.0119	0.2656 0.4522 0.1172 0.1649	PBM10
TCP23	GGGCCAC	1°	0.49152	16.3507	A: C: G: T:	0.0226 0.0457 0.7723 0.1593	0.0050 0.0028 0.9657 0.0265	0.0096 0.0051 0.9805 0.0048	0.0367 0.0876 0.8350 0.0406	0.0072 0.9708 0.0085 0.0134	0.0036 0.9843 0.0083 0.0038	0.0115 0.9795 0.0029 0.0061	0.9543 0.0104 0.0307 0.0046	0.0095 0.9586 0.0106 0.0213	0.7386 0.1354 0.0282 0.0978	PBM10
TCP16	GGGTCCAC	1°	0.49950	23.3583	A: C: G: T:	0.0180 0.2454 0.0329 0.7037	0.1206 0.0035 0.8660 0.0099	0.0021 0.0016 0.9909 0.0054	0.0015 0.0025 0.9923 0.0037	0.1182 0.1159 0.0271 0.7388	0.0036 0.9922 0.0014 0.0028	0.0045 0.9913 0.0200 0.0021	0.9796 0.0026 0.0167 0.0012	0.0022 0.9837 0.0120 0.0090	0.4559 0.4154 0.0051 0.1166	nPBM11
HSFB2A	GAACGTTC	1°	0.49721	13.6581	A: C: G: T:	0.2699 0.3525 0.3683 0.0093	0.0278 0.0145 0.9445 0.0131	0.9337 0.0115 0.0407 0.0142	0.9526 0.3576 0.2722 0.0228	0.1316 0.2722 0.3576 0.2386	0.2386 0.0228 0.0298 0.1316	0.0228 0.0137 0.0109 0.9526	0.0142 0.0407 0.0115 0.9337	0.0131 0.9445 0.0145 0.0278	0.0047 0.4632 0.1953 0.3368	PBM10
	TTCTAGAA	3°	0.49548	13.7222	A: C: G: T:	0.0927 0.3871 0.2058 0.3144	0.0668 0.0465 0.0584 0.8283	0.0344 0.0680 0.0120 0.8857	0.0294 0.9253 0.1378 0.5915	0.0028 0.2679 0.2679 0.0028	0.5915 0.0298 0.0294 0.0294	0.0298 0.8857 0.0154 0.0344	0.8857 0.8283 0.0120 0.0668	0.2849 0.2643 0.0584 0.1567		
HSFC1	TTCTAGAA	1°	0.48371	10.8566	A: C: G: T:	0.0759 0.3853 0.2204 0.3185	0.0265 0.0179 0.0187 0.9369	0.0070 0.0274 0.0049 0.9607	0.0037 0.9856 0.0062 0.0046	0.0011 0.0472 0.0514 0.9003	0.9003 0.0514 0.0472 0.0011	0.0046 0.0062 0.0049 0.0037	0.9607 0.0049 0.0179 0.0070	0.9369 0.4524 0.2418 0.1619	PBM10	

	GAAGCTTC	2°	0.48177	8.1512	A: C: G: T:	0.3416 0.2943 0.3500 0.0141	0.0174 0.0214 0.0100 0.0143	0.9333 0.0338 0.0109 0.0229	0.9124 0.0246 0.0109 0.0521	0.1628 0.1675 0.0619 0.0678	0.0678 0.0109 0.1675 0.1628	0.0521 0.0109 0.0246 0.0124	0.0229 0.0338 0.0100 0.0333	0.0143 0.0469 0.0214 0.0174	0.0240 0.3911 0.2461 0.3389	
TGA2	TGACGTCA	1°	0.49977	22.8079	A: C: G: T:	0.5015 0.1481 0.3263 0.0240	0.0156 0.0207 0.0057 0.9579	0.0055 0.0060 0.8662 0.1222	0.9667 0.0013 0.0114 0.0206	0.0009 0.8511 0.0045 0.1435	0.1435 0.0045 0.8511 0.0009	0.0206 0.0114 0.0013 0.9667	0.1222 0.8662 0.0060 0.0055	0.9579 0.0057 0.0207 0.0156	0.0270 0.4125 0.1660 0.3945	PBM11
	GATGACGT	2°	0.49778	17.3646	A: C: G: T:	0.0328 0.0878 0.3770 0.5024	0.9612 0.0012 0.0243 0.0133	0.0010 0.9799 0.0029 0.0162	0.0235 0.0022 0.9736 0.0007	0.0062 0.0048 0.0008 0.9883	0.0273 0.9697 0.0019 0.0011	0.9890 0.0016 0.0066 0.0028	0.0019 0.1096 0.3801 0.5084	0.0781 0.7036 0.1607 0.0576	0.4873 0.1430 0.2035 0.1663	
bZIP60	TGACGTCA	1°	0.48883	9.6150	A: C: G: T:	0.3278 0.1511 0.4246 0.0965	0.0183 0.0643 0.0241 0.8933	0.0265 0.0594 0.0167 0.0744	0.8915 0.7993 0.0428 0.0491	0.0129 0.7993 0.0128 0.1750	0.1750 0.0128 0.7993 0.0129	0.0491 0.0428 0.0167 0.8915	0.0744 0.8397 0.0594 0.0265	0.8933 0.0241 0.0643 0.0183	0.0494 0.4931 0.1161 0.3414	nPBM11
	TGACGTGG	2°	0.45417	6.3058	A: C: G: T:	0.1870 0.1033 0.5290 0.1807	0.1978 0.6444 0.0880 0.0698	0.0458 0.8356 0.0796 0.0390	0.9100 0.283 0.0291 0.0326	0.0218 0.8993 0.0287 0.0502	0.0704 0.0308 0.0176 0.0164	0.0316 0.0337 0.0176 0.9171	0.0481 0.8810 0.0383 0.0326	0.9020 0.0326 0.0423 0.0232	0.0568 0.4284 0.1159 0.3989	
ETT	TGTCGGAA	1°	0.49101	9.4343	A: C: G: T:	0.0988 0.0830 0.2400 0.5783	0.0015 0.0042 0.0771 0.9172	0.0025 0.0033 0.0016 0.0033	0.0014 0.0164 0.0014 0.9807	0.0025 0.9903 0.0014 0.0058	0.0023 0.0075 0.9797 0.0105	0.0027 0.0235 0.9640 0.0097	0.6178 0.1240 0.2122 0.0460	0.5751 0.0948 0.1907 0.1393	0.3438 0.2846 0.1114 0.2602	PBM10
	TGTCGACA	4° 1° reranked	0.48840	9.0160	A: C: G: T:	0.2507 0.2749 0.1147 0.3596	0.0451 0.0636 0.2316 0.6598	0.0691 0.0210 0.9002 0.0098	0.0291 0.0071 0.0303 0.9335	0.0870 0.8831 0.0209 0.0090	0.0090 0.0303 0.8831 0.0870	0.9335 0.0303 0.0071 0.0291	0.0098 0.9002 0.0210 0.0691	0.6598 0.2316 0.0636 0.0451	0.0857 0.0583 0.6595 0.1965	
REM1	TGATGTAG	1°	0.46371	7.1746	A: C: G: T:	0.1178 0.1984 0.5304 0.1534	0.0092 0.8971 0.0223 0.0714	0.0175 0.0060 0.0077 0.9689	0.9571 0.9815 0.0051 0.0339	0.0050 0.0039 0.0033 0.0102	0.9816 0.0038 0.0067 0.0079	0.2424 0.2334 0.0043 0.5200	0.0102 0.9030 0.0158 0.0709	0.5292 0.0329 0.0218 0.4161	0.3201 0.2479 0.1536 0.2785	PBM10
	CGTGTAGA	3° 1° reranked	0.45864	6.4386	A: C: G: T:	0.1042 0.6298 0.2279 0.0382	0.0374 0.7598 0.0043 0.1984	0.0335 0.0096 0.9157 0.0412	0.0184 0.0203 0.0088 0.9524	0.0205 0.0104 0.0241 0.0121	0.0759 0.0103 0.0241 0.8897	0.8947 0.0480 0.0140 0.0433	0.0404 0.1104 0.7631 0.0860	0.7585 0.0564 0.1199 0.0652	0.1382 0.4315 0.3680 0.0623	
ANAC55	TTACGTGT	1°	0.47972	6.0861	A: C: G: T:	0.3553 0.1860 0.1630 0.2957	0.7959 0.0264 0.0442 0.1335	0.0193 0.7738 0.0184 0.1885	0.9697 0.9595 0.0064 0.0092	0.0225 0.0162 0.9485 0.0116	0.0157 0.1165 0.1336 0.0196	0.0135 0.0804 0.0044 0.7364	0.8783 0.0094 0.0041 0.0369	0.8659 0.3972 0.1055 0.4099	0.0873 0.3228 0.0566 0.4099	PBM11
	TTACGTAA	2°	0.47415	6.1824	A: C: G: T:	0.3947 0.1137 0.3012 0.1903	0.2728 0.0308 0.0573 0.6391	0.0289 0.2255 0.0439 0.0157	0.8860 0.0543 0.0439 0.0458	0.0501 0.8811 0.0231 0.0501	0.0458 0.0231 0.0439 0.0501	0.0157 0.0439 0.0377 0.8860	0.7079 0.0377 0.0573 0.0289	0.6391 0.0573 0.3228 0.5542	0.0566 0.0663 0.0663 0.5542	
ANAC46	TTGCGTGT	1°	0.42525	5.4579	A: C: G: T:	0.1800 0.1499 0.0414 0.6286	0.5444 0.2166 0.0798 0.1593	0.0655 0.6299 0.0378 0.2668	0.8128 0.1199 0.0432 0.0241	0.1017 0.8581 0.0245 0.0157	0.0938 0.0489 0.8122 0.0451	0.1427 0.5648 0.1124 0.1801	0.4588 0.4021 0.0481 0.0909	0.8054 0.0395 0.0960 0.0591	0.1221 0.4561 0.1235 0.2983	PBM10

ANAC58	TTGCGTGC	1°	0.40886	5.1465	A: C: G: T:	0.2576 0.2304 0.2825 0.2295	0.3699 0.1245 0.3438 0.1618	0.0948 0.6736 0.0547 0.1769	0.9127 0.0343 0.0219 0.0311	0.2571 0.6940 0.0258 0.0231	0.0265 0.0375 0.0980 0.0280	0.1159 0.6927 0.0716 0.1198	0.6277 0.2453 0.0198 0.1072	0.7697 0.0881 0.0265 0.1158	0.2138 0.2698 0.1617 0.3547	PBM11
AHL20	AATATATT	1°	0.49590	10.5728	A: C: G: T:	0.4377 0.1650 0.1121 0.2852	0.9509 0.0084 0.0137 0.0271	0.8770 0.0023 0.0070 0.1137	0.0769 0.0051 0.0099 0.0981	0.8661 0.0031 0.0025 0.1282	0.1282 0.0025 0.0031 0.8661	0.9081 0.0099 0.0051 0.0769	0.1137 0.0070 0.0023 0.8770	0.0271 0.0137 0.0084 0.9509	0.4361 0.1036 0.2083 0.2520	PBM10
	ATTTAATT	2°	0.48871	8.1840	A: C: G: T:	0.2787 0.2689 0.0948 0.3576	0.7742 0.0655 0.0727 0.0876	0.7534 0.0062 0.0051 0.2352	0.1205 0.0044 0.0132 0.8619	0.2499 0.0056 0.0062 0.7383	0.7741 0.0045 0.0036 0.2179	0.8426 0.0080 0.0014 0.1480	0.6484 0.0116 0.0028 0.3371	0.0545 0.0282 0.0176 0.8998	0.6056 0.0873 0.1038 0.2033	
	ATTATAAT	3°	0.48833	8.5613	A: C: G: T:	0.1716 0.1974 0.2083 0.4228	0.8898 0.0459 0.0299 0.0344	0.4032 0.0102 0.0080 0.5785	0.3705 0.0053 0.0148 0.6094	0.6923 0.0240 0.0032 0.2804	0.2804 0.0056 0.0032 0.6923	0.6094 0.0116 0.0028 0.3705	0.5785 0.0282 0.0176 0.4032	0.0344 0.0299 0.0102 0.8898	0.4836 0.2236 0.0459 0.1865	
AHL12	AAATATTT	1°	0.49268	6.4852	A: C: G: T:	0.3651 0.0848 0.1309 0.4191	0.7719 0.0176 0.1464 0.0641	0.8431 0.0353 0.0236 0.0979	0.4837 0.0144 0.0047 0.4972	0.3315 0.0061 0.0029 0.6596	0.6596 0.0029 0.0047 0.3315	0.4972 0.0144 0.0144 0.4837	0.0979 0.0236 0.0353 0.8431	0.0641 0.1464 0.0353 0.7719	0.3922 0.1060 0.0176 0.3654	PBM10
	ATAATTAT	2°	0.48975	6.2513	A: C: G: T:	0.3438 0.0940 0.4166 0.1457	0.5527 0.0233 0.0245 0.3994	0.3738 0.0187 0.0215 0.5859	0.7858 0.0039 0.0114 0.1990	0.7384 0.0022 0.0076 0.2519	0.2519 0.0076 0.0022 0.7384	0.1990 0.0114 0.0187 0.7858	0.5859 0.0215 0.0039 0.3738	0.3994 0.0245 0.0233 0.5527	0.0601 0.1751 0.0280 0.7368	
	AATTAATT	3°	0.48844	6.4436	A: C: G: T:	0.4156 0.1117 0.1106 0.3621	0.8348 0.0231 0.0407 0.1014	0.4456 0.0054 0.0118 0.5372	0.5507 0.0123 0.0071 0.4299	0.3551 0.0039 0.0052 0.6358	0.6358 0.0071 0.0123 0.3551	0.4299 0.0071 0.0123 0.5507	0.5372 0.0118 0.0054 0.4456	0.1014 0.0407 0.0231 0.8348	0.2546 0.1762 0.0480 0.5212	
AHL25	AATTAATT	1°	0.49556	9.0256	A: C: G: T:	0.2978 0.1685 0.2822 0.2515	0.8725 0.0257 0.0527 0.0491	0.6155 0.0037 0.0057 0.3752	0.2073 0.0036 0.0026 0.7865	0.2246 0.0033 0.0018 0.7704	0.7704 0.0018 0.0033 0.2246	0.7865 0.0026 0.0057 0.2073	0.3752 0.0026 0.0057 0.6155	0.0491 0.0527 0.0257 0.8725	0.3944 0.1578 0.0901 0.3578	nPBM11
	AATATATT	2°	0.49421	8.3794	A: C: G: T:	0.4921 0.1421 0.1320 0.2338	0.8609 0.0371 0.0582 0.0439	0.6444 0.0045 0.0073 0.3438	0.3487 0.0028 0.0135 0.6350	0.6954 0.0051 0.0019 0.2976	0.2976 0.0019 0.0051 0.6954	0.6350 0.0135 0.0028 0.3487	0.3438 0.0073 0.0045 0.6444	0.0439 0.0582 0.0371 0.8609	0.3577 0.1575 0.1204 0.3644	
	ATTTAATT	3°	0.49160	7.6173	A: C: G: T:	0.3654 0.1765 0.1647 0.2933	0.5420 0.0860 0.1480 0.2240	0.7389 0.0049 0.0057 0.2504	0.3851 0.0046 0.0038 0.6057	0.2852 0.0046 0.0028 0.7075	0.4878 0.0017 0.0017 0.5088	0.8268 0.0024 0.0064 0.1643	0.3752 0.0057 0.0037 0.6155	0.0824 0.0313 0.0551 0.8311	0.2796 0.0960 0.0651 0.5594	
PIF3	CCACGTGG	1°	0.49924	48.4142	A: C: G: T:	0.1834 0.3664 0.3657 0.0844	0.0044 0.9566 0.0331 0.0059	0.0090 0.9853 0.0034 0.0023	0.9866 0.0060 0.0050 0.0024	0.0010 0.9681 0.0025 0.0284	0.0284 0.0025 0.9681 0.0010	0.0024 0.0050 0.0060 0.9866	0.0023 0.0050 0.0060 0.0090	0.0059 0.0331 0.0566 0.0044	0.0912 0.4031 0.3482 0.1576	nPBM11
LBD16	TCCGGAAA	1°	0.49258	9.0121	A: C: G: T:	0.3147 0.1973 0.0603 0.4277	0.0617 0.2736 0.0705 0.5942	0.0176 0.9004 0.0780 0.0040	0.0146 0.9605 0.0128 0.0121	0.0439 0.0745 0.8509 0.0306	0.0372 0.1823 0.6860 0.0945	0.4675 0.0515 0.4385 0.0424	0.4430 0.0650 0.2689 0.2231	0.4819 0.1278 0.1035 0.2868	0.4050 0.1357 0.0847 0.3746	nPBM11

STY1	CCCTAGGG	1°	0.49699	14.8468	A: C: G: T:	0.1769 0.4493 0.1857 0.1881	0.0556 0.8699 0.0551 0.0195	0.0481 0.9057 0.0146 0.0316	0.0082 0.9491 0.0307 0.0120	0.0094 0.0063 0.0104 0.9740	0.9740 0.0104 0.0063 0.0094	0.0120 0.0307 0.9491 0.0082	0.0316 0.0146 0.9057 0.0481	0.0195 0.0551 0.8699 0.0556	0.0995 0.2696 0.2383 0.3927	PBM10
	CTAGCTAG	4° 1° reranked	0.49424	16.3197	A: C: G: T:	0.4084 0.2337 0.2010 0.1568	0.0304 0.8889 0.0540 0.0267	0.0499 0.0147 0.0239 0.9114	0.9133 0.0037 0.0327 0.0504	0.0207 0.0087 0.9480 0.0226	0.0226 0.9480 0.0087 0.0207	0.0504 0.0327 0.0037 0.9133	0.9114 0.0239 0.0147 0.0499	0.0267 0.0540 0.8889 0.0304	0.1059 0.4955 0.0840 0.3146	

**6. Table S3.** List of transcriptional repressors according to co-expression data in Figure 3.

TF	TAIR ID	Proposed repressor	EAR domain	Interaction co-repressors
CCA1	At2g46830	11		
RVE1	At5g17300	12		TPL, TPR2, TPR4 (22)
KAN1	At5g16560	21		TPL (22, 133)
At5g28300	At5g28300			
DEAR3	At2g23340		YES (35)	
DEAR4	At4g36900		YES (35)	
TOE1	At2g28550	39, 40	YES (35)	TPL, TPR1-4, TPR3 (22, 133)
TOE2	At5g60120	39, 40	YES (35)	TPL, TPR1-4 (22, 133)
ATHB51	At5g03790			
WRKY18	At4g31800	59	YES (35)	
WRKY12	At2g44745	61		
ZAT2	At2g17180		YES (35, 67)	
ZAT6	At5g04340		YES (35, 67)	
ZAT14	At5g03510		YES (35, 67)	
ZAT18	At3g53600		YES (35, 67)	
Dof5.7	At5g65590			
YAB1	At2g45190	74		
YAB5	At2g26580	74		
ETT	At2g33860	97		TPR1 (22)
AHL12	At1g63480			
LBD16	At2g42430			

**7. Table S4.** Summary of gene expression experiments

Experiment	Series ID	Filtering	Platform	Reference	Comments
Ox <i>MYB46-HER</i>	n.a.		ATH1	8	Gene list from publication
<i>myb11,12,111</i>	E-ATMX-5		ATH1	6	Gene list from publication
Ox <i>MYB59</i>	n.a.		ATH1	5	Gene list from publication
Drought 3h	GSE5624	Fold-change 2 FDR Limma< 0.05	ATH1	134	
<i>cca1, lhy</i>	GSE19263	Fold-change 2 FDR Limma< 0.05	ATH1	14	
RVE8-GR	GSE38879		RNA-seq	15	Gene list from publication
Cold stress	GSE5621	Fold-change 2 FDR Limma< 0.05	ATH1	134	
<i>arr1,10,12</i>	E-MEXP-1573	Fold-change 2 FDR Limma< 0.05	ATH1	19	
Zeatin 1h	GSE39384	Fold-change 2 pvalue Limma< 0.05	ATH1	135	
<i>ipt1,3,5,7</i>	GSE32087	Fold-change 3 FDR Limma< 0.05	Agilent V4	20	
Ox <i>GLK1/Ox GLK2</i>	E-MEXP-1794		Operon 30k	23	Gene list from publication
<i>gtl1</i>	n.a.		ATH1	29	Gene list from publication
Ox <i>ERF6</i>	GSE45830		AGRONOMICS Tiling Array	33	Gene list from publication
Ox <i>DDF1</i>	n.a.		ATH1	36	Gene list from publication
<i>hardy-D</i>	GSE8936	Fold-change 2 FDR RP< 0.05	ATH1	37	
JA 1h	GSE39384	Fold-change 2 FDR Limma< 0.05	ATH1	135	
<i>ap2</i>	GSE5633	Fold-change 2 FDR Limma< 0.05	ATH1	136	
<i>smz-D</i>	E-MEXP-2040	Fold-change 2 FDR RP< 0.05	ATH1	137	
Ox <i>WUS</i>	E-MEXP-432	Fold-change 2 FDR Limma< 0.05	ATH1	45	
Ox <i>HaHB4</i>	GSE8878	Fold-change 2 FDR Limma< 0.05	CATMA_v2.2	48	
<i>spl7</i>	GSE24696		RNA-seq	53	Gene list from publication
<i>mSPL13</i>	GSE10414	Fold-change 2 FDR Limma< 0.05	ATH1	54	
<i>wrky18,40</i>	E-MEXP-2371	Fold-change 2 FDR RP< 0.05	ATH1	60	
<i>wrky12</i>	E-MEXP-2792	Fold-change 2 FDR RP< 0.05	ATH1	61	
SA 3h	GSE39384	Fold-change 2 FDR Limma< 0.05	ATH1	135	
Ox <i>ZAT10</i>	E-ATMX-20	Fold-change 2 FDR Limma< 0.05	ATH1	138	
Ox <i>ZAT12</i>	GSE5742	Fold-change 2 FDR Limma< 0.05	ATH1		
Ox <i>AZF1; Ox AZF2</i>	E-MEXP-2922	Fold-change 2 FDR Limma< 0.05	ATH1	67	
Ox <i>GNC; Ox GNL</i>	GSE21256		Agilent V4	70	Gene list from publication
Ox <i>FIL-GR</i>	n.a.		ATH1	75	Gene list from publication
<i>yab1,3; yab1,3,5</i>	GSE21705	Fold-change 3 FDR Limma< 0.05	ATH1	76	
TCP20-GR	GSE29012		Affymetrix Tiling 1.0R	139	Gene list from publication
<i>hsf1,3</i>	E-MEXP-98	Fold-change 2 pvalue Limma< 0.005	ATH1	84	
Heat stress	GSE5628	Fold-change 2 FDR Limma< 0.05	ATH1	134	
<i>tga2,5,6</i>	GSE10732	Fold-change 2 FDR Limma< 0.05	ATH1	86	
<i>bzip60</i>	E-MEXP-1753		Agilent V2	89	Gene list from publication
IAA	GSE39384	Fold-change 2	ATH1	135	

		FDR Limma< 0.05			
<i>arf2</i>	GSE11216	Fold-change 2 FDR Limma< 0.05	ATH1	98	
<i>arf7,19</i>	GSE627	Fold-change 2 FDR Limma< 0.05	ATH1	99	
ABA 1h	GSE39384	Fold-change 2 FDR Limma< 0.05	ATH1	135	
PIF3 targets	GSE39217		ChIP-seq	113	Gene list from publication
LBD18-GR	E-MEXP-1400	Fold-change 2 FDR Limma< 0.05	ATH1	117	

**8. Table S5.** Oligonucleotides and details on the constructs used in promoter-Luciferase/effector constructs. Sequences in lower letters correspond to Gateway recombination sites.

Name	Sequence (5'-3')	Size (bp)	Target	Effector
AT3G62260-B1 AT3G62260-B2	ggggacaagttgtacaaaaaaaggcaggctAATACCATGTCAAATGTCGGT ggggaccacttgcataagaaagctgggtCAGTCTTTCTTTTCACA	1,896	At3g62260	At5g28300
AT2G26670-B1 AT2G26670-B2	ggggacaagttgtacaaaaaaaggcaggctGTGCCCACATGAAGGAATATG ggggaccacttgcataagaaagctgggtTGGTTTGATCGGAATAGAAAAA	1,707	At1g68200	AHL12
AT5G42860-B1 AT5G42860-B2	ggggacaagttgtacaaaaaaaggcaggctCAAGCCTCAGTCACAAGACA ggggaccacttgcataagaaagctgggtCTTGAGATTGGAATGGAAGATG	1,913	At5g42860	Dof5.7
AT2G38080-B1 AT2G38080-B2	ggggacaagttgtacaaaaaaaggcaggctCATGCCAAATAAGCGACCA ggggaccacttgcataagaaagctgggtCTCCCTCTATCTTCTCTTC	2,023	At2g38080	ATHB51
AT4G00780-B1 AT4G00780-B2	ggggacaagttgtacaaaaaaaggcaggctCAAGTGAGCAGAGGAGACAA ggggaccacttgcataagaaagctgggtTTGTAAGAAGCTACAAATGATATCT	1,943	At4g00780	LBD16

## 9. References

1. Riechmann JL, et al. (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290(5499): 2105-2110.
2. Dubos C, et al. (2010) MYB transcription factors in Arabidopsis. *Trends Plant Sci* 15(10):573-581.
3. Romero I, et al. (1998) More than 80 R2R3-MYB regulatory genes in the genome of Arabidopsis thaliana. *Plant J* 14(3):273–284.
4. Prouse MB, Campbell MM (2012) The interaction between MYB proteins and their target DNA binding sites. *Biochim Biophys Acta* 1819(1):67-77.
5. Mu RL, et al. (2009) An R2R3-type transcription factor gene AtMYB59 regulates root growth and cell cycle progression in Arabidopsis. *Cell Res* 19(11):1291-1304.
6. Stracke R, et al. (2007) Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the Arabidopsis thaliana seedling. *Plant J* 50(4):660-677.
7. Zhong R, Richardson EA, Ye Z-H (2007) The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. *Plant Cell* 19(9):2776-2792.
8. Zhong R, Ye ZH (2012) MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol* 53(2):368-380.
9. Park MY, Kang JY, Kim SY (2011) Overexpression of AtMYB52 confers ABA hypersensitivity and drought tolerance. *Mol Cells* 31(5):447-454.
10. Alabadi D, et al. (2001) Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. *Science* 293(5531):880–883.
11. Harmer SL, Kay SA (2005) Positive and negative factors confer phase-specific circadian regulation of transcription in Arabidopsis. *Plant Cell* 17(7):1926-1940.
12. Rawat R, et al. (2009) REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways. *Proc Natl Acad Sci USA* 106(39):16883-16888.
13. Wang Z-Y, et al. (1997) A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell* 9(4): 491–507.
14. Pokhilko A, et al. (2012) The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. *Mol Syst Biol* 8:574.
15. Hsu PY, Devisetty UK, Harmer SL (2013) Accurate timekeeping is controlled by a cycling activator in Arabidopsis. *Elife* 30:e00473.
16. Meissner M, et al. (2013) Mapping quantitative trait loci for freezing tolerance in a recombinant inbred line population of Arabidopsis thaliana accessions Tenela and C24 reveals REVEILLE1 as negative regulator of cold acclimation. *Plant Cell Environ* 36(7):1256-1267.
17. Hosoda K, et al. (2002) Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the Arabidopsis response regulators. *Plant Cell* 14(9):2015-2029.
18. Sheen J (2002) Phosphorelay and transcription control in cytokinin signal transduction. *Science* 296(5573):1650-1652.
19. Argyros RD, et al. (2008) Type B response regulators of Arabidopsis play key roles in cytokinin signaling and plant development. *Plant Cell* 20(8):2102-2116.
20. Nishiyama R, et al. (2012) Transcriptome analyses of a salt-tolerant cytokinin-deficient mutant reveal differential regulation of salt stress response by cytokinin deficiency. *PLoS One* 7(2):e32124.
21. Wu G, et al. (2008) KANADI1 regulates adaxial-abaxial polarity in Arabidopsis by directly repressing the transcription of ASYMMETRIC LEAVES2. *Proc Natl Acad Sci USA* 105(42):16392-16397.
22. Causier B, Ashworth M, Guo W, Davies B (2012) The TOPLESS interactome: a framework for gene repression in Arabidopsis. *Plant Physiol* 158(1):423-438.
23. Waters MT, et al. (2009) GLK transcription factors coordinate expression of the photosynthetic apparatus in Arabidopsis. *Plant Cell* 21(4):1109-28.
24. Zhou DX (1999) Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. *Trends Plant Sci* 4(6):210-214.
25. Kaplan-Levy RN, Brewer PB, Quon T, Smyth DR (2012) The trihelix family of transcription factors light, stress and development. *Trends Plant Sci* 17(3):163–171.

26. Dehesh K, Hung H, Tepperman JM, Quail PH (1992) GT-2: a transcription factor with twin autonomous DNA-binding domains of closely related but different target sequence specificity. *EMBO J* 11(11): 4131- 4144.
27. Breuer C, et al. (2009) The trihelix transcription factor GTL1 regulates ploidy-dependent cell growth in the *Arabidopsis* trichome. *Plant Cell* 21(8):2307-2322.
28. Yoo CY, et al. (2010) The *Arabidopsis* GTL1 transcription factor regulates water use efficiency and drought tolerance by modulating stomatal density via transrepression of SDD1. *Plant Cell* 22(12):4128-4141.
29. Breuer C, et al. (2012) Transcriptional repression of the APC/C activator CCS52A1 promotes active termination of cell growth. *EMBO J* 31(24): 4488-4501.
30. Ohme-Takagi M, Shinshi H (1995) Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* 7(2):173-212.
31. Solano R, Stepanova A, Chao Q, Ecker JR (1998) Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes Dev* 12(23):3703-3714.
32. Godoy M, et al. (2011) Improved protein-binding microarrays for the identification of DNA-binding specificities of transcription factors. *Plant J* 66(4):700-711.
33. Dubois M, et al. (2013) The ETHYLENE RESPONSE FACTOR 6 Acts as Central Regulator of Leaf Growth under Water Limiting Conditions in *Arabidopsis thaliana*. *Plant Physiol* 162(1):319-332.
34. Shinozaki K and Yamaguchi-Shinozaki K. 2010. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr Opin Plant Biol* 3(3):217-223.
35. Kagale S, Links MG, Rozwadowski K (2010) Genome-wide analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in *Arabidopsis*. *Plant Physiol* 152(3):1109-1113.
36. Magome H, Yamaguchi S, Hanada A, Kamiya Y, Oda K (2008) The DDF1 transcriptional activator upregulates expression of a gibberellin-deactivating gene, GA2ox7, under high-salinity stress in *Arabidopsis*. *Plant J* 56(4):613-626.
37. Karaba A, et al. (2007) Improvement of water use efficiency in rice by expression of HARDY, an *Arabidopsis* drought and salt tolerance gene. *Proc Natl Acad Sci USA* 104(39):15270-15275.
38. Pauwels L, et al. (2008) Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. *Proc Natl Acad Sci USA* 105(4):1380-1385.
39. Jung JH, et al. (2007) The GIGANTEA-regulated microRNA172 mediates photoperiodic flowering independent of CONSTANS in *Arabidopsis*. *Plant Cell* 19(9):2736-2748.
40. Yant L, et al. (2010) Orchestration of the floral transition and floral development in *Arabidopsis* by the bifunctional transcription factor APETALA2. *Plant Cell* 22(7):2156-2170.
41. Ariel FD, Manavella PA, Dezar CA, Chan RL (2007) The true story of the HD-Zip family. *Trends Plant Sci* 12(9):419-426.
42. van der Graaff E, Laux T, Rensing SA (2009) The WUS homeobox-containing (WOX) protein family. *Genome Biol* 10(12):248.
43. Lohmann JU, et al. (2001) A molecular link between stem cell regulation and floral patterning in *Arabidopsis*. *Cell* 105(6):793-803.
44. Busch W, et al. (2010) Transcriptional control of a plant stem cell niche. *Dev Cell* 18(5):849-861.
45. Leibfried A, et al. (2005) WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* 438(7071):1172-1175.
46. Sessa G, Steindler C, Morelli G, Ruberti I (1998) The *Arabidopsis* Athb-8, -9 and -14 genes are members of a small gene family coding for highly related HD-ZIP proteins. *Plant Mol Biol* 38(4):609-622.
47. Saddic LA, et al. (2006) The LEAFY target LMI1 is a meristem identity regulator and acts together with LEAFY to regulate expression of CAULIFLOWER. *Development* 133(9):1673-1682.
48. Manavella PA, et al. (2006) Cross-talk between ethylene and drought signalling pathways is mediated by the sunflower Hahb-4 transcription factor. *Plant J* 48(1):125-137.
49. Chen X, et al. (2010) SQUAMOSA promoter-binding protein-like transcription factors: star players for plant growth and development. *J Integr Plant Biol* 52(11):946-951.

50. Cardon G, et al. (1999) Molecular characterization of the *Arabidopsis* SBP-box genes. *Gene* 237(1):91–104.
51. Birkenbihl RP, Jach G, Saedler H, Huijser P (2005) Functional dissection of the plant-specific SBP-domain: Overlap of the DNA-binding and nuclear localization domains. *J Mol Biol* 352(3):585–596.
52. Yamasaki H, Hayashi M, Fukazawa M, Kobayashi Y, Shikanai T (2009) SQUAMOSA Promoter Binding Protein-Like7 Is a central regulator for copper homeostasis in *Arabidopsis*. *Plant Cell* 21(1):347-361.
53. Bernal M, et al. (2012) Transcriptome sequencing identifies SPL7-regulated copper acquisition genes FRO4/FRO5 and the copper dependence of iron homeostasis in *Arabidopsis*. *Plant Cell* 24(2):738-761.
54. Martin RC, et al. (2010) The microRNA156 and microRNA172 gene regulation cascades at post-germinative stages in *Arabidopsis*. *Seed Science Research* 20:79-87.
55. Rushton PJ, Somssich IE, Ringler, P, Shen, QJ (2000) WRKY transcription factors *Trends Plant Sci* 5(5):199-206.
56. Ciolkowski I, Wanke D, Birkenbihl RP, Somssich IE (2008) Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Mol Biol* 68(1-2):81-92.
57. Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. *Trends Plant Sci* 5(5):199-206.
58. Xu X, Chen C, Fan B, Chen Z (2006) Physical and functional interactions between pathogen-induced *Arabidopsis* WRKY18, WRKY40, and WRKY60 transcription factors. *Plant Cell* 18(5):1310-1326.
59. Shen QH, et al. (2007) Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science* 315(5815):1098-1103.
60. Pandey SP, Roccaro M, Schön M, Logemann E, Somssich IE (2010) Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of *Arabidopsis*. *Plant J* 64(6):912-23.
61. Wang H, et al. (2010) Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proc Natl Acad Sci USA* 107(51):22338-22343.
62. Kim KC, Lai Z, Fan B, Chen Z (2008) *Arabidopsis* WRKY38 and WRKY62 transcription factors interact with histone deacetylase 19 in basal defense. *Plant Cell* 20(9):2357-2371.
63. Englbrecht CC, Schoof H, Bohm S (2004) Conservation, diversification and expansion of C2H2 zinc finger proteins in the *Arabidopsis thaliana* genome. *BMC Genomics* 5(1):39.
64. Ciftci-Yilmaz S, Mittler R (2008) The zinc finger network of plants. *Cell Mol Life Sci* 65(7-8):1150-1160.
65. Yoshioka K, Fukushima S, Yamazaki T, Yoshida M, Takatsujii H (2001) The plant zinc finger protein ZPT2-2 has a unique mode of DNA Interaction. *J Biol Chem* 276(38):35802-35807.
66. Sakamoto H, et al. (2004) *Arabidopsis* Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. *Plant Physiol* 136(1):2734- 2746.
67. Kodaira KS, et al. (2011) *Arabidopsis* Cys2/His2 zinc-finger proteins AZF1 and AZF2 negatively regulate abscisic acid-repressive and auxin-inducible genes under abiotic stress conditions. *Plant Physiol* 157(2):742-756.
68. Lowry JA, Atchley WR (2000) Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding Domain. *J Mol Evol* 50(2):103-115.
69. Teakle GR, Manfield IW, Graham JF, Gilmartin PM (2002) *Arabidopsis thaliana* GATA factors: organisation, expression and DNA-binding characteristics. *Plant Mol Biol* 50(1):43-57.
70. Richter R, Behringer C, Müller IK, Schwechheimer C (2010) The GATA-type transcription factors GNC and GNL/CGA1 repress gibberellin signaling downstream from DELLA proteins and PHYTOCHROME-INTERACTING FACTORS. *Genes Dev* 24(18):2093-2104.
71. Bowman JL (2000) The YABBY gene family and abaxial cell fate. *Curr Opin Plant Biol* 3(1):17-22.
72. Kanaya E, Nakajima N, Okada K (2002) Non-sequence-specific DNA binding by the FILAMENTOUS FLOWER protein from *Arabidopsis thaliana* is reduced by EDTA. *J Biol Chem* 277(14):11957-11964.

73. Dai M, et al. (2007) The rice YABBY1 gene is involved in the feedback regulation of gibberellin metabolism. *Plant Physiol* 144(1):121-133.
74. Stahle MI, Kuehlich J, Staron L, von Arnim AG, Golz JF (2009) YABBYs and the transcriptional corepressors LEUNIG and LEUNIG\_HOMOLOG maintain leaf polarity and meristem activity in *Arabidopsis*. *Plant Cell* 21(10):3105–3118.
75. Bonaccorso O, Lee JE, Puah L, Scutt CP, Golz JF (2012) FILAMENTOUS FLOWER controls lateral organ development by acting as both an activator and a repressor. *BMC Plant Biol* 12:176.
76. Sarojam R, et al. (2010) Differentiating *Arabidopsis* shoots from leaves by combined YABBY activities. *Plant Cell* 22(7):2113-2130.
77. Yanagisawa S (2004) Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant Cell Physiol* 45(4):3863-3891.
78. Martín-Trillo M, Cubas P (2010) TCP genes: a family snapshot ten years later. *Trends Plant Sci* 15(1):31-39.
79. Kosugi S, Ohashi Y (2002) DNA binding and dimerization specificity and potential targets for the TCP protein family. *Plant J* 30(3):337–348.
80. Viola IL, Reinheimer R, Ripoll R, Manassero NG, Gonzalez DH (2012) Determinants of the DNA binding specificity of class I and class II TCP transcription factors. *J Biol Chem* 287(1):347-356.
81. Åkerfelt M, Morimoto RI, Sistonen L (2010) Heat shock factors: integrators of cell stress, development and lifespan. *Nat Rev Mol Cell Biol* 11(8):545-555.
82. Scharf KD, Berberich T, Ebersberger I, Nover L (2012) The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. *Biochimica et Biophysica Acta* 1819(2):104-119.
83. Perisic O, Xiao H, Lis JT (1989) Stable binding of *Drosophila* heat shock factor to head-to-head and tail-to-tail repeats of a conserved 5 bp recognition unit. *Cell* 59(5):797–806.
84. Busch W, Wunderlich M, Schöffl F (2005) Identification of novel heat shock factor-dependent genes and biochemical pathways in *Arabidopsis thaliana*. *Plant J* 41(1):1-14.
85. Jakoby M, et al. (2002) bZIP transcription factors in *Arabidopsis*. *Trends Plant Sci* 7(3):106-111.
86. Mueller S, et al. (2008) General detoxification and stress responses are mediated by oxidized lipids through TGA transcription factors in *Arabidopsis*. *Plant Cell* 20(3):768-785.
87. Zhang Y, Tessaro MJ, Lassner M, Li X (2003) Knockout analysis of *Arabidopsis* transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *Plant Cell* 15(11):2647-2653.
88. Zander M, La Camera S, Lamotte O, Métraux JP, Gatz C (2010) *Arabidopsis thaliana* class-II TGA transcription factors are essential activators of jasmonic acid/ethylene-induced defense responses *Plant J* 61(2):200-210.
89. Iwata Y, Fedoroff NV, Koizumi N (2008) *Arabidopsis* bZIP60 is a proteolysis-activated transcription factor involved in the endoplasmic reticulum stress response. *Plant Cell* 20(11):3107-3121.
90. Hayashi S, Takahashi H, Wakasa Y, Kawakatsu T, Takaiwa F (2013) Identification of a cis-element that mediates multiple pathways of the endoplasmic reticulum stress response in rice. *Plant J* 74(2):248-257.
91. Swaminathan K, Peterson K, Jack T (2008) The plant B3 superfamily. *Trends Plant Sci* 13(12): 647–655.
92. Reidt W, et al. (2000) Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. *Plant J* 21(5):401-408.
93. Ulmasov T, Hagen G, Guilfoyle TJ (1997) ARF1, a transcription factor that binds to auxin response elements. *Science* 276(5320):1865–1868.
94. Kagaya Y, Ohmiya K, Hattori T (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res* 27(2):470-478.
95. Franco-Zorrilla JM, et al. (2002) AtREM1, a member of a new family of B3 domain-containing genes, is preferentially expressed in reproductive meristems. *Plant Physiol* 128(2):418-427.
96. Levy YY, Mesnage S, Mylne JS, Gendall AR, Dean C (2002) Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time control. *Science* 297(5579):243–246.
97. Guilfoyle TJ, Hagen G. (2007) Auxin response factors. *Curr Opin Plant Biol* 10(5):453-460.

98. Vert G, Walcher CL, Chory J, Nemhauser JL (2008) Integration of auxin and brassinosteroid pathways by Auxin Response Factor 2. *Proc Natl Acad Sci USA* 105(28):9829-9834.
99. Okushima Y, et al. (2005) Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in *Arabidopsis thaliana*: unique and overlapping functions of ARF7 and ARF19. *Plant Cell* 17(2):444-463.
100. Olsen AN, Ernst HA, Leggio LL, Skriver K (2005) NAC transcription factors: structurally distinct, functionally diverse. *Trends Plant Sci* 10(2):79-87.
101. Puranik S, Sahu PP, Srivastava PS, Prasad M (2012) NAC proteins: regulation and role in stress tolerance. *Trends Plant Sci* 17(6):369-381.
102. Tran LS, et al. (2004) Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter. *Plant Cell* 16(9):2481-2498.
103. Zhong R, Lee C, Ye ZH (2011) Global analysis of direct targets of secondary wall NAC master switches in *Arabidopsis*. *Mol Plant* 3(6):1087-1103.
104. Bu Q, et al. (2008) Role of the *Arabidopsis thaliana* NAC transcription factors ANAC019 and ANAC055 in regulating jasmonic acid-signaled defense responses. *Cell Res* 18(7):756-767.
105. Jiang H, Li H, Bu Q, Li C (2009) The RHA2a-interacting proteins ANAC019 and ANAC055 may play a dual role in regulating ABA response and jasmonate response. *Plant Signal Behav* 4(5):464-466.
106. Aravind L, Landsman D (1998) AT-hook motifs identified in a wide variety of DNA-binding proteins. *Nucleic Acids Res* 26(19):4413-4421.
107. 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.
108. 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):1000251.
109. 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.
110. Vom Endt D, Soares-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.
111. Lu H, Zou Y, Feng N (2010) Overexpression of AHL20 negatively regulates defenses in *Arabidopsis*. *J Integr Plant Biol* 52(9):801-808.
112. Toledo-Ortiz G, Huq E, Quail PH (2003) The *Arabidopsis* basic/helix-loop-helix transcription factor family. *Plant Cell* 15(8):1749-1770.
113. Zhang Y, et al. (2013) A Quartet of PIF bHLH Factors Provides a Transcriptionally Centered Signaling Hub That Regulates Seedling Morphogenesis through Differential Expression-Patterning of Shared Target Genes in *Arabidopsis*. *PLoS Genet* 9(1):e1003244.
114. Hornitschek P, et al. (2012) Phytochrome interacting factors 4 and 5 control seedling growth in changing light conditions by directly controlling auxin signaling. *Plant J* 71(5):699-711.
115. Fernández-Calvo P, et al. (2011) The *Arabidopsis* bHLH transcription factors MYC3 and MYC4 are targets of JAZ repressors and act additively with MYC2 in the activation of jasmonate responses. *Plant Cell* 23(2):701-715.
116. Husbands A, Bell EM, Shuai B, Smith HM, Springer PS (2007) LATERAL ORGAN BOUNDARIES defines a new family of DNA-binding transcription factors and can interact with specific bHLH proteins. *Nucleic Acids Res* 35(19):6663-6671.
117. Lee HW, Kim MJ, Kim NY, Lee SH, Kim J (2013) LBD18 acts as a transcriptional activator that directly binds to the EXPANSIN14 promoter in promoting lateral root emergence of *Arabidopsis*. *Plant J* 73(2):212-224.
118. Goh T, Joi S, Mimura T, Fukaki H (2012) The establishment of asymmetry in *Arabidopsis* lateral root founder cells is regulated by LBD16/ASL18 and related LBD/ASL proteins. *Development* 139(5):883-893.
119. Bell EM, et al. (2012) *Arabidopsis* lateral organ boundaries negatively regulates brassinosteroid accumulation to limit growth in organ boundaries. *Proc Natl Acad Sci USA* 109(51):21146-21151.
120. Eklund DM, et al. (2010) The *Arabidopsis thaliana* STYLISH1 protein acts as a transcriptional activator regulating auxin biosynthesis. *Plant Cell* 22(2):349-363.

121. Ståldal V, et al. (2012) The *Arabidopsis thaliana* transcriptional activator STYLISH1 regulates genes affecting stamen development, cell expansion and timing of flowering. *Plant Mol Biol* 78(6):545-59.
122. Baylis T, Cierlik I, Sundberg E, Mattsson J (2013) SHORT INTERNODES/STYLISH genes, regulators of auxin biosynthesis, are involved in leaf vein development in *Arabidopsis thaliana*. *New Phytol* 197(3):737-750.
123. Papdi C et al. (2008) Functional identification of *Arabidopsis* stress regulatory genes using the controlled cDNA overexpression system. *Plant Physiol* 147(2):528-542.
124. Hammarström M, Hellgren N, van den Berg S, Berglundm H, Härd T (2002) Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Sci* 11(2): 313-321.
125. Berger MF, Bulyk ML (2009) Universal protein-binding microarrays for the comprehensive characterization of the DNA-binding specificities of transcription factors. *Nat Protoc* 4(3):393-411.
126. Dudley A, Aach J, Steffen M, Church GM (2002) Measuring absolute expression with microarrays using a calibrated reference sample and an extended signal intensity range. *Proc Natl Acad Sci* 99(11):7554-7559.
127. Hruz T, et al. (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics* 420747.
128. Tabas-Madrid D, Nogales-Cadenas R, Pascual-Montano A (2012) GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res* 40 (Web Server issue):478-483.
129. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3:Article3.
130. Zhang W, Zhang T, Wu Y, Jiang J (2012) Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. *Plant Cell* 24(7):2719-2731.
131. Nakagawa T, Ishiguro S, Kimura T. 2009. Gateway vectors for plant transformation. *Plant Biotechnology* 26: 275–284.
132. Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, Leyva A, Weigel D, García JA, Paz-Ares J. 2007. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet*. 39(8):1033-1037.
133. Braun P, Carvunis AR, Charlotteaux B, Dreze M, Ecker JR, Hill DE, Roth FP, Vidal M, Galli M, Balumuri P, et al. 2011. Evidence for network evolution in an *Arabidopsis* interactome map. *Science* 333(6042): 601-607.
134. Kilian J et al. (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J* 50(2):347-363.
135. Goda H et al. (2008). The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. *Plant J* 55(3):526-542.
136. Schmid M et al. (2005) A gene expression map of *Arabidopsis thaliana* development. *Nat Genet* 37(5):501-506.
137. Mathieu J, Yant LJ, Mürdter F, Küttner F, Schmid M (2009) Repression of flowering by the miR172 target SMZ. *PLoS Biol* 7(7) e1000148.
138. Rossel JB et al. (2007) Systemic and intracellular responses to photooxidative stress in *Arabidopsis*. *Plant Cell* 19(12): 4091-4110.
139. Danisman S et al. (2012). *Arabidopsis* class I and class II TCP transcription factors regulate jasmonic acid metabolism and leaf development antagonistically. *Plant Physiol*. 159(4): 1511-1523.