



LARGE-SCALE BIOLOGY ARTICLE

Deep Conservation of *cis*-Element Variants Regulating Plant Hormonal Responses^[OPEN]

Michal Lieberman-Lazarovich,¹ Chen Yahav, Alon Israeli, and Idan Efroni²

Institute of Plant Sciences and Genetics in Agriculture, The Robert H. Smith Faculty of Agriculture, The Hebrew University, Rehovot 7610001, Israel

ORCID IDs: 0000-0002-1776-8257 (M.L.-L.); 0000-0003-1750-9842 (C.Y.); 0000-0002-3339-2042 (A.I.); 0000-0002-0219-8685 (I.E.)

Phytohormones regulate many aspects of plant life by activating transcription factors (TFs) that bind sequence-specific response elements (REs) in regulatory regions of target genes. Despite their short length, REs are degenerate, with a core of just 3 to 4 bp. This degeneracy is paradoxical, as it reduces specificity and REs are extremely common in the genome. To study whether RE degeneracy might serve a biological function, we developed an algorithm for the detection of regulatory sequence conservation and applied it to phytohormone REs in 45 angiosperms. Surprisingly, we found that specific RE variants are highly conserved in core hormone response genes. Experimental evidence showed that specific variants act to regulate the magnitude and spatial profile of hormonal response in *Arabidopsis* (*Arabidopsis thaliana*) and tomato (*Solanum lycopersicum*). Our results suggest that hormone-regulated TFs bind a spectrum of REs, each coding for a distinct transcriptional response profile. Our approach has implications for precise genome editing and for rational promoter design.

INTRODUCTION

Phytohormones activate coordinated transcriptional responses to regulate plant development and physiology. These responses are mediated by transcription factor (TF) families, such as AUXIN RESPONSE FACTOR (ARF), B-CLASS RESPONSE REGULATORS (RRs), and ABRE binding factors, which mediate the responses to auxin, cytokinin, and abscisic acid, respectively (Choi et al., 2000; Kieber and Schaller, 2018; Roosjen et al., 2018). TFs control the expression of their target genes by binding specific nucleotide sequences, termed response elements (REs), in the regulatory regions of those genes. The sequence preference of each TF is commonly portrayed as a position weight matrix (PWM), or a DNA motif, which presents the contribution of each nucleotide position to the TF binding. Motifs for the main hormone response regulating TFs were derived using biochemical affinity methods or were based on identification of enriched sequences in the promoters of genes induced by hormone application (Bhargava et al., 2013; Zemlyanskaya et al., 2016). These motifs represent a consensus of many sequences and are composed of a short core of invariant nucleotides and several variant nucleotides, which have little effect on TF binding (Boer et al., 2014). The biological role of these variant nucleotides in mediating TF-DNA binding and transcriptional response is unclear. While these motifs were determined by a consensus sequence, hormone-induced transcriptional response of individual genes varies in terms of

magnitude, tissue specificity, and other conditions. This gene-specific information is lost when considering only the commonalities between many regulatory sequences. Indeed, when attention was paid to the different response profiles of different genes, more specific motifs and flanking sequences were identified (Bargmann et al., 2013; Omelyanchuk et al., 2017).

Among plant hormones, auxin serves a key role in plant biology and its core response pathway is conserved from basal to higher plants (Mutte et al., 2018). In its canonical transcriptional pathway, auxin induces the activation of a family of ARF transcriptional factors that bind the auxin response element (auxRE; Roosjen et al., 2018). The auxRE was first defined as the hexamer TGTCTC, and DR5, a synthetic promoter comprised of eight direct repeats of this motif, could grant auxin responsiveness to downstream genes (Ballas et al., 1995; Ulmasov et al., 1995, 1997; Abel et al., 1996). Later biochemical characterization has shown that ARF proteins bind a more permissive TGTCNN motif (Ulmasov et al., 1999; Boer et al., 2014). Similarly to auxin, the hormone cytokinin activates target genes and increases chromatin accessibility through B-class RRs that bind the degenerate RE DGATYN (cytRE; D = A,G,T; Y = C,T; Zürcher et al., 2013; Zubo et al., 2017; Potter et al., 2018; Xie et al., 2018). Abscisic acid regulates the transcription of downstream genes via TFs, such as ABRE binding factor, which bind the degenerate RE BACGTGK (abRE; B = C,G,T; K = G,T; Choi et al., 2000; Hattori et al., 2002; Wu et al., 2018).

Modification of the DR5 element from TGTCTC to TGTCGG produced a stronger response to auxin (Boer et al., 2014; Liao et al., 2015), suggesting that the variant nucleotides in the motif may affect the transcriptional response profile of downstream genes. However, it is unknown whether these nucleotide variants serve any function in a native sequence context or in the plant itself. It is generally accepted that the existence of purifying selection, evidenced by high sequence conservation, is a strong indicator of functionality (Graur et al., 2015). Thus, using the

¹ Current address: Institute of Plant Sciences, Volcani Center, Agricultural Research Organization, Rishon LeZion 7528809, Israel.

² Address correspondence to efroni.idan@mail.huji.ac.il.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Idan Efroni (efroni.idan@mail.huji.ac.il).

^[OPEN]Articles can be viewed without a subscription.

www.plantcell.org/cgi/doi/10.1105/tpc.19.00129

signature of evolutionary conservation has the potential to identify functional nucleotides in REs. However, detecting conservation for specific RE variations is not trivial as rapid structural changes in promoters and the fact that REs exist in multiple copies prevent the alignment of regulatory regions. Furthermore, due to common gene duplications in plants, ortholog assignment in divergent species is often uncertain. Indeed, studies of regulatory region conservation in plants have restricted themselves to individual loci (Ballas et al., 1995; Kim et al., 2006), or focused on ecotypes or a limited number of species (Vandepoele et al., 2009; Korkuc et al., 2014).

Existing methods for detection of binding site conservation are geared toward the de novo identification of motifs (Elemento and Tavazoie, 2005; Kantorovitz et al., 2007; Ivan et al., 2008; Gordân et al., 2010; De Witte et al., 2015). To enable the detection of variant sequence conservation, we developed conservation of motif variants (CoMoVa), an alignment-free method for detection of conserved small degenerate sequences in known motifs over broad evolutionary distances. We applied CoMoVa to detect conservation of hormonal REs in 45 different angiosperms, separated by ~150 million years (Chaw et al., 2004), and identified highly conserved RE variants in core hormonal response genes. We then show that variants of the auxRE can fine-tune the transcriptional response to the hormone.

RESULTS

CoMoVa: An Algorithm for the Identification of RE Variant Conservation

To test the conservation of variant nucleotides in binding motifs, we first focused on the auxRE. We analyzed the sequences 1 kb upstream and 100 bp downstream of the transcription start site (TSS) for all protein-coding genes in 45 angiosperms (Figure 1A; Supplemental Table 1). A mean $91\% \pm 9\%$ of genes carried at least one degenerate TGTCNN auxRE, and among those, there were 6.1 ± 0.8 auxREs per gene (mean \pm standard deviation; sd). To determine possible motif conservation, we first identified candidate orthologs. Using the well-annotated Arabidopsis (*Arabidopsis thaliana*) genome as a reference, a reciprocal BLAST was performed to identify up to eight candidate orthologs from each species (Supplemental Data Set 1). The candidates with the most similar set of motif variants were selected as the likely ortholog.

The auxRE variants in the promoters of all genes were extracted with multiple instances of an auxRE variant treated as one. Thus, the entire promoter sequence was converted into a vector of 16 Boolean values, representing the presence or absence of a specific motif variant in the promoter of a given ortholog. These presence/absence vectors were arranged according to a pre-defined species tree representing known phylogenetic relationships, resulting in a Boolean matrix (Figures 1A to 1D). This analysis could identify genes with little to no conservation of the variant sequences (Figure 1B), some conservation (Figure 1C), or high levels of conservation (Figure 1D), evident by the presence of a specific motif in the promoters of almost all gene orthologs.

To quantify this conservation level, maximum parsimony optimization was applied to compute the number of motif changes

(gain or loss) for each variant along the tree. We defined a conservation score metric for each gene and motif variant combination. This score was computed as the number of orthologs/species having a particular variant, minus the number of changes in the tree. The use of the number of gains/losses of a specific motif along the evolutionary tree enabled to distinguish between common motifs that are rapidly gained or lost and between those that are highly conserved. This conservation metric ranged between 0 to 45, representing no conservation to complete conservation, respectively.

In order to determine the significance of the conservation, the neutral variation rate at promoter positions had to be calculated. Defining a theoretical model for neutral substitution rates at promoters is difficult, as GC content and the distribution of dinucleotides is genome specific (Supplemental Figure 1A; Gentles and Karlin, 2001). Additionally, the available genomes are not equally spaced in terms of evolutionary distance and thus some of the observed conservation may be due to kinship. To correct for both these factors, background rates were determined empirically by calculating the conservation score distribution for the two variant nucleotides 8 bp upstream to the core sequence (NNnnnnnTGTC for the auxRE; Supplemental Figure 1B). auxREs are sometimes found in close proximity to one another (O'Malley et al., 2016; Stigliani et al., 2018), which could affect the distribution of nucleotides at these positions. However, the overall distribution of dinucleotides at position -8 to the auxRE was found to be similar to the overall distribution in promoters, indicating that, genome wide, this effect is negligible (Supplemental Figure 1B). The distribution of conservation scores for the neutral positions was used to compute P-values for significant conservation of variant nucleotide in the auxRE motif.

auxRE Motif Variant Conservation Defines the Core Auxin Response

Using CoMoVa, we detected highly significant conservation for 252 genes compared with the background conservation of the neutral nucleotides at position -8 to the motif. To verify that this arbitrary selection for calculation of background nucleotides does not significantly affect the results, we also ran the analysis using nucleotides at positions -10 , -14 , and -15 bp upstream of the core motif for calculation of neutral rates. The number of significantly detected genes varied only slightly (Supplemental Figure 2A) and of the 252 genes with significantly conserved motif, 224 (89%) were identified as significantly conserved using all four calculations of background distributions (Supplemental Figure 2B).

There was a marked bias in the conserved variants, with CC, GG, GA, and TC being the most highly conserved motifs, but we also detected conservation for AT, AA, AC, AG, CT, and TG. Many known auxin downstream factors were found among the conserved set, including four AUX/IAA genes (*IAA1*, *IAA2*, *IAA3*, and *IAA4*), *GH3*, and 11 SMALL AUXIN UPREGULATED RNA genes (Figure 2A; Supplemental Data Set 2). These gene families function in the core auxin response and were shown to be auxin induced in multiple species (Abel and Theologis, 1996). Among the genes containing a conserved TGTCGG variant were the cell wall modifier genes *GH9B7* and *GH9B5*, previously shown to act

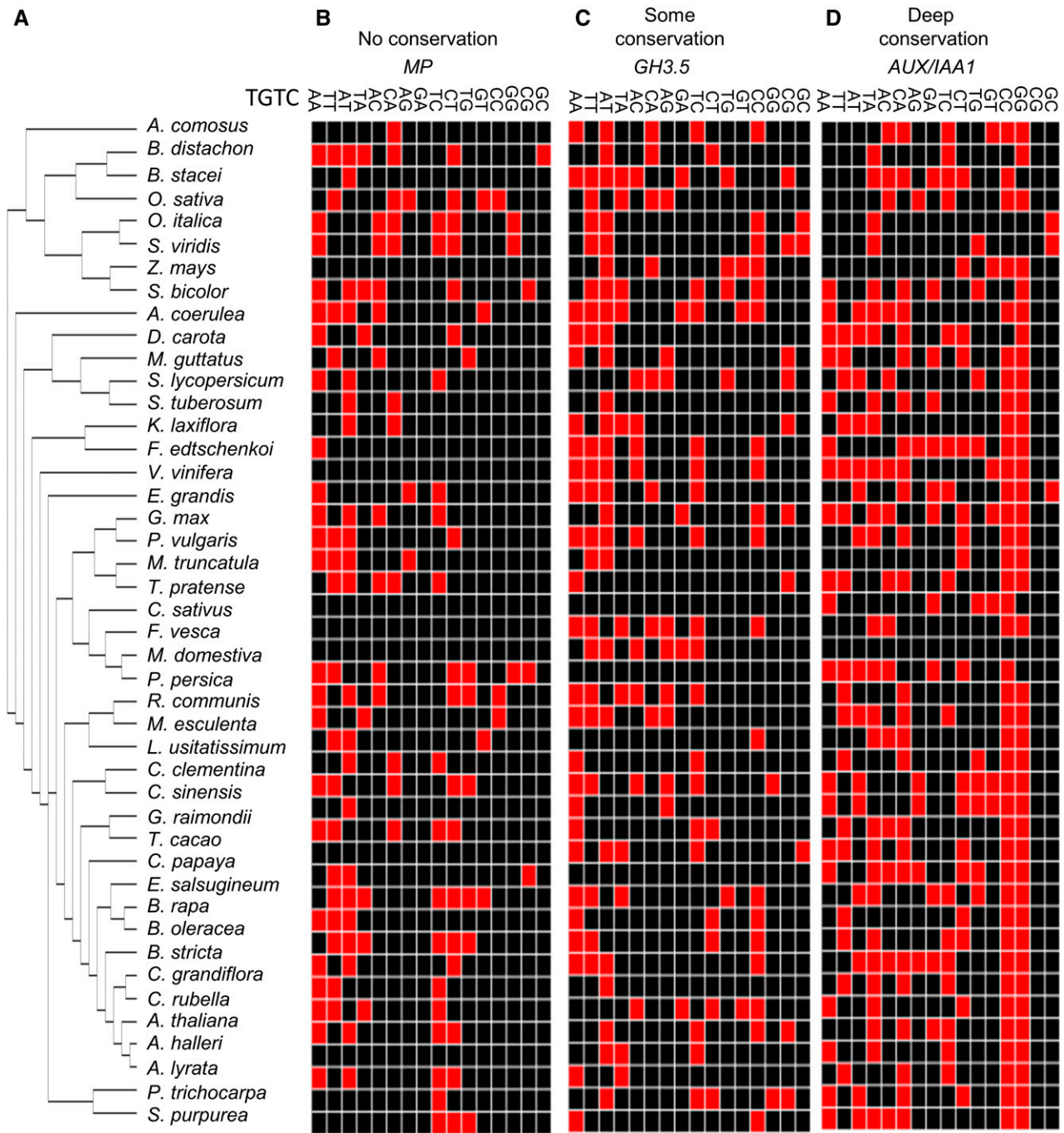


Figure 1. Algorithm for Identification of DNA Motif Variant Conservation.

(A) Phylogenetic tree of the species used in this study. Full latin names of all plants are provided in Supplemental Table 1.

(B) to (D) Occurrences of particular motif variants in orthologs of *MP* **(B)**, *GH3.5* **(C)**, and *AUX/IAA1* **(D)** of the corresponding species of the tree in **(A)**. Red blocks mark the presence of a specific motif variant in the 1-kb gene promoter. Black blocks mark the absence of the variant. Multiple occurrences of the variants are considered as one.

downstream of auxin in lateral root formation (Lewis et al., 2013) as well as two *EXPANSIN* and two mannan biosynthesis genes. TGTC_{CC} was found in association with the TF gene *WUSCHEL*, which acts to maintain the growth of the shoot apical meristem, and *VND2*, which regulates xylem formation (Supplemental Data Set 2; Schoof et al., 2000; Tan et al., 2018). Significant

conservation for several motif variants was found for four *1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE* (*ACS*) genes, key regulators of ethylene biosynthesis, which is known to be auxin induced (Abel et al., 1995).

Gene ontology (GO) term enrichment analysis of genes with conserved variants revealed highly significant enrichment for

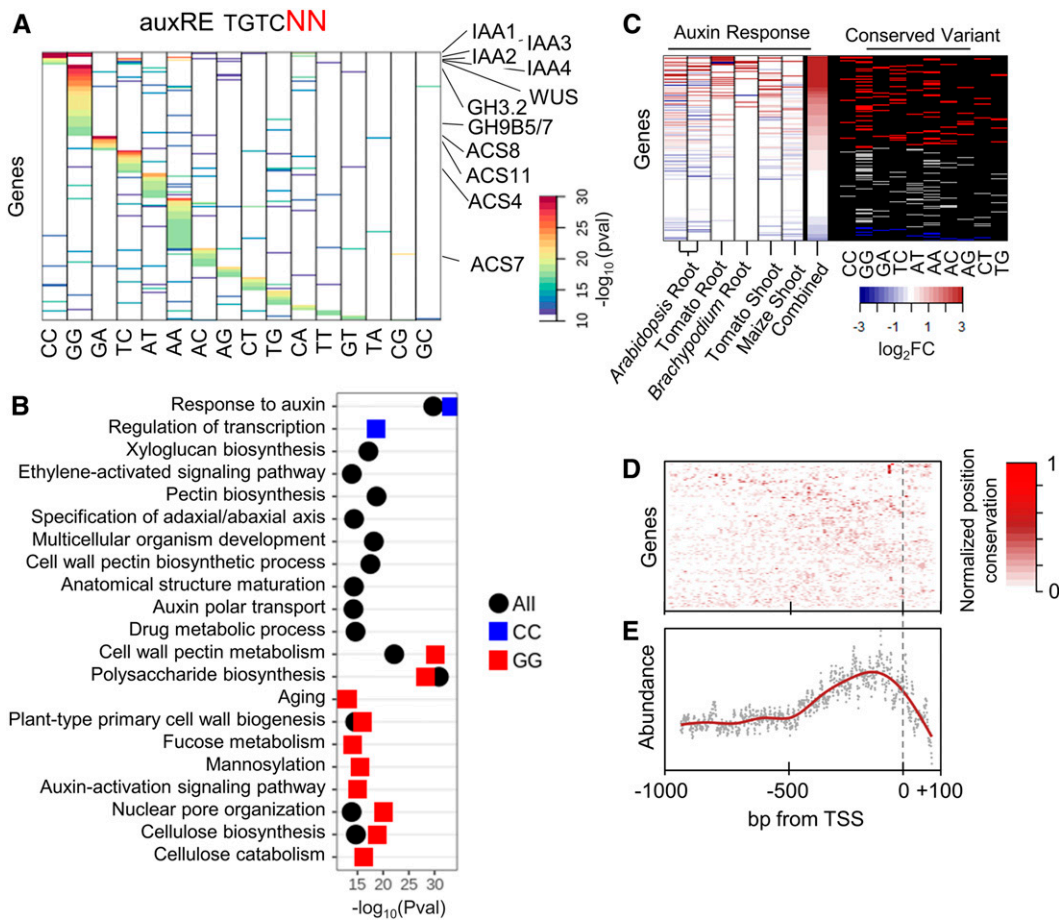


Figure 2. Conservation of auxRE Variants.

(A) Genes and motif variants showing significant conservation. Each column represents a specific dinucleotide variant of the TGTCNN motif, and the color indicates the significance of conservation of the particular variant.

(B) GO term enrichment of genes with a conserved auxRE variant in their promoters. Pval, P-value.

(C) Auxin responsiveness of genes with conserved auxRE variants in different tissues and species. Heatmap on the left shows the \log_2 FoldChange (\log_2FC) of auxin-treated versus control samples. Heatmap on the right shows the particular conserved variants found in the promoter of the matching gene. Red marks auxin induced ($\log_2FC > 1.5$), blue marks auxin repressed ($\log_2FC < 0.75$).

(D) and **(E)** Conservation of the auxRE motif position for each gene for the 45 angiosperms species **(D)** and the mean conservation across all promoters **(E)**. High values (red) represent high conservation of position of the motif.

response to auxin, auxin polar transport, and multiple cell wall modification processes such as xyloglucan, pectin, and cellulose biosynthesis (Figure 2B). The number of auxRE motifs per gene for genes with conserved auxRE motifs was 6.4 ± 1.3 (SD), which was not significantly different than the number for all genes, suggesting that it is the identity of the motif variant that is conserved, rather than it being linked to a higher number of auxREs. Interestingly, there was functional separation between genes enriched for different variants. Specifically, genes associated with a conserved TGTCGG variant were enriched for cell wall synthesis functions, and genes with conserved TGTC~~CGG~~ were associated with response to auxin (Figure 2B).

Three of the highly conserved variants (TGTCGG/CC/TC) were previously identified as enriched in sets of auxin-responsive genes

in Arabidopsis (Ulmasov et al., 1995; Bargmann et al., 2013; Mironova et al., 2014; Zemlyanskaya et al., 2016). To test for auxin responsiveness of genes across multiple species, we compiled five rapid auxin response-related transcriptional activity data sets in Arabidopsis, maize (*Zea mays*), and *Brachypodium*. To add additional dicots, we also generated RNA sequencing profiles of tomato (*Solanum lycopersicum*) roots and shoots at 3 h following auxin application. While auxin response differed between organs and species, most of the genes with conserved variants exhibited auxin response in at least one sample (Figure 2C).

We next sought to determine whether the position of the auxRE relative to the TSS is also conserved. Despite large differences in the length of intergenic regions between species, and in agreement with previous studies (O'Malley et al., 2016;

Zemlyanskaya et al., 2016; Galli et al., 2018), auxREs were enriched in the +50- to -250-bp region relative to the TSS across all tested angiosperms (Figures 2D and 2E). Furthermore, for a subset of genes, we could identify high conservation of the position of the auxRE relative to the TSS in their orthologs (Figure 2E).

Promoter sequences vary in length between species, and RE may be found further than 1 kb upstream to the TSS. However, a limitation of alignment-free algorithms such as CoMoVa is that longer sequences reduce the ability to identify signal in the data due to higher background. To test the ability to detect conserved variants at different distances from the TSS, we applied CoMoVa to identify auxRE conservation in sequences of 500 bp, 2 kb, and 3 kb upstream to the TSS. Consistent with the higher background, there was a reduction in the number of genes with significantly conserved auxRE for longer promoters. Additionally, analysis of the shorter 500-bp promoter resulted in the identification of fewer conserved REs, probably due to loss of data ($n = 127, 215,$ and 157 for the 500-bp, 2-kb, and 3-kb regions, respectively; Supplemental Figure 3A). Regardless of the reduced number of genes with conserved REs, significant enrichment was still detected for the TGTCCC/GG/TC/AC variants (Supplemental Figure 3B; Supplemental Data Set 3), and there was large overlap between the genes with conserved auxRE between the different promoter lengths, although some were unique to specific promoter lengths (Supplemental Figure 3C). GO enrichment test for genes with conserved auxRE in longer promoters resulted in enrichment for similar terms as those detected for the 1-kb promoters (Supplemental Data Set 3), suggesting that there are functional conserved auxRE located at a longer distance from the TSS. However, it should be noted that we cannot rule out that this apparent conservation of distant RE is due to the mis-annotation of the TSS in the available angiosperm genomes.

Overall, the consistent biological role of genes with conserved auxRE variants, which included well-known factors such as the negative auxin feedback genes, *AUX/IAA* and *GH3* genes (Bargmann and Estelle, 2014), suggests that CoMoVa identified a conserved regulatory module acting downstream of auxin. Furthermore, the conservation of specific auxRE variants across vast evolutionary distances suggests the presence of a strong selection pressure on the variable dinucleotides, indicating that their identity is important for gene function.

Motif Variant Conservation Defines Core Response for Cytokinin and Abscisic Acid

If the conservation of variable nucleotides in promoters of core response genes is a general phenomenon, we hypothesized that applying the same algorithm to other hormonal REs will identify core target genes for these hormones. To test this hypothesis, we applied CoMoVa to characterize the conservation of the cytRE (DGATYN), which mediates the cytokinin response and appears in $93.3\% \pm 0.9\%$ of the 1-kb gene promoters in all species (14.58 ± 1.4 mean \pm SD REs per promoter). Similarly to the auxRE analysis, we used matching adjacent sequence (DnnnnnnnnGATCnnnnnnnnYN) to compute the background conservation rates. In agreement with our hypothesis, CoMoVa identified conservation in 10 motifs (out of the possible 24) for 803 genes. Among these were highly

conserved AGATTT and TGATTT variants that were found in 7 of the 10 A-class RRs (Figure 3A). These negative feedback regulators of the cytokinin response are known targets of the B-class *ARABIDOPSIS RESPONSE REGULATORS*, as shown for multiple species (Kieber and Schaller, 2018). Multiple *HISTONE* genes had a strongly conserved GGATCG motif, while the cell cycle regulators *CYCA2;1*, *CYDB2;2*, and *CYCD3;3* had conserved AGATTT, TGATTT, and GGATCT motifs, respectively (Supplemental Data Set 4). Apart from response to cytokinin, the conserved gene list was enriched for regulators of cell proliferation, nucleosome assembly and position, and carbohydrate and amino acid metabolism (Figure 3B). There was a weak preference for the RE to be located approximately -100 bp to the TSS, but this was much less apparent than for the auxRE (Figures 3C and 3D).

As an additional test, we computed the conservation of the abRE motif, which mediates the abscisic acid response and appears in $29\% \pm 7\%$ of all genes in all species (motif, BACGTGK; background, BnnnnnnnnACGTGnnnnnnnnK). CoMoVa identified motif conservation in 146 genes (Figure 4A; Supplemental Data Set 5) including in known abscisic acid-inducible genes, such as the abscisic acid biosynthetic enzyme *NCED* (Barrero et al., 2006; Yang and Tan, 2014), members of the ABI Five Binding Protein (AFP) family (Garcia et al., 2008), and *HIGHLY ABSCISIC ACID-INDUCED PPC GENES (HAI)* (Hirayama and Shinozaki, 2007). GO term analysis showed enrichment for negative regulation of abscisic acid signaling pathway, photosynthesis, response to cold, and regulation of seed germination (Figure 4B). Moreover, despite large differences in intergenic region length between the species, conserved abRE variants were very highly enriched at -150 bp from the TSS, suggesting that this position is important for this RE function (Figures 4C and 4D). Overall, we conclude that motif variant conservation is prevalent and that core hormonal response targets are enriched with specific variants.

Deep Conservation of Motif Flanking Sequences

TF binding sites are widely distributed in the genome, but only a subset of them is bound by TF in vivo. These TF-bound sites are characterized by sequence conservation in regions flanking the core binding site (Dror et al., 2015). Indeed, promoter mutagenesis has shown that nucleotides flanking the auxRE are critical for generating proper auxin response (Ballas et al., 1995; Ulmasov et al., 1997; Galli et al., 2018). We therefore sought to identify whether regions flanking the auxREs also show signs of conservation. To this end, the flanking 20 bp up- and downstream for each ortholog group were aligned, centered on the conserved motif. A PWM was calculated, and the information content for each position served as a measure of conservation at a specific position. As predicted, additional conserved nucleotides were identified in the vicinity of the auxRE (Figures 5A and 5B; Supplemental Data Set 6), and, in some cases, were quite long, such as the stem cell regulator *WUS* that had an 18-bp conserved region flanking the TGTCCC motif and a TCCCTTCTA sequence 20 bp upstream (Figure 5C).

For many genes, conservation outside the auxRE among all angiosperms was weak, suggesting these sequences evolved faster than the variant nucleotides in the motif (Figure 5B). As many aspects of gene expression profiles may be specific to more

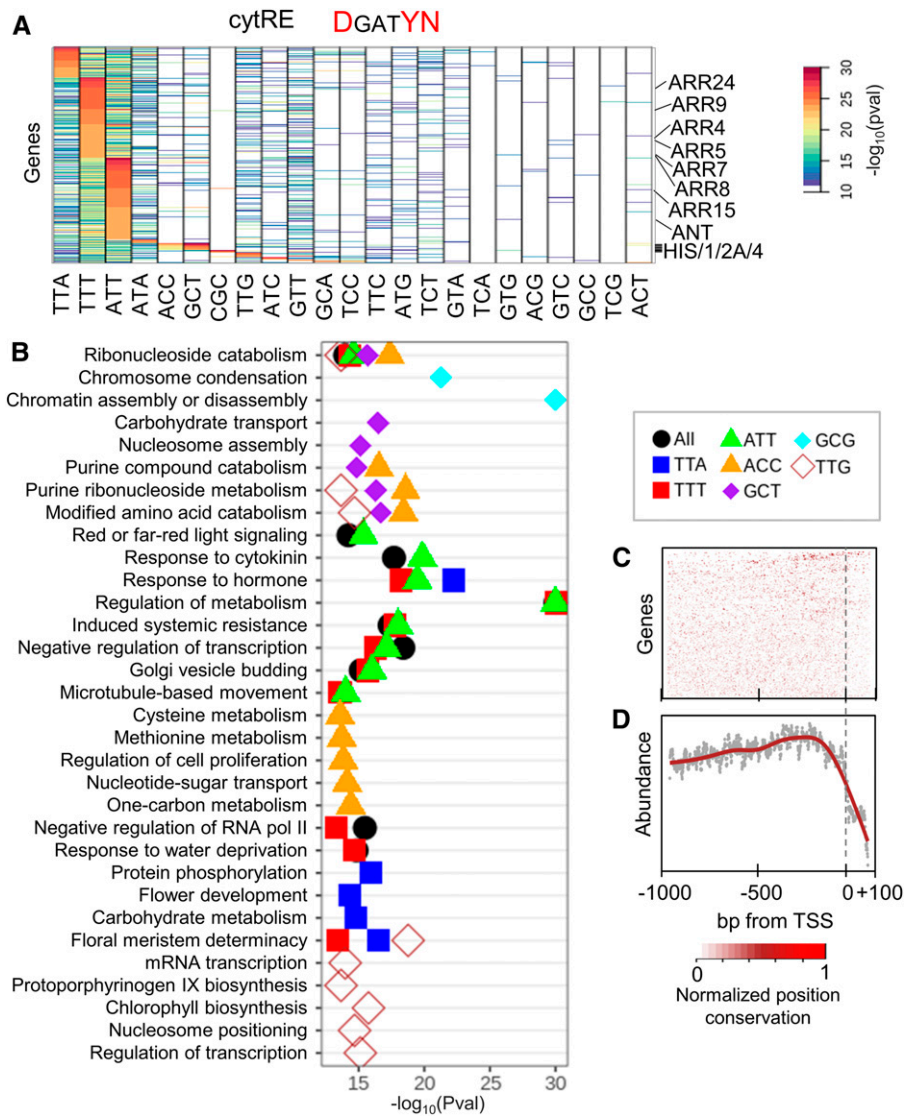


Figure 3. Conservation of cytRE Variants.

(A) Genes and motif variants of the cytRE showing significant conservation.

(B) GO term enrichment of genes with a conserved cytRE variant in their promoter.

(C) and **(D)** Conservation of the cytRE motif position for each gene for the 45 angiosperms species **(C)** and the mean conservation across all promoters **(D)**. High values (red) represent high conservation of position of the motif. pval, P-value.

closely related species, we applied the same algorithm separately to the dicots, monocots, and Brassicaceae to test whether we could find broader signs of conservation within shorter evolutionary distances. Significantly higher levels of conservation were detected when the subtrees were analyzed, with clear divergence between monocots and dicots (Figures 5A to 5C; Supplemental Data Set 6). The Brassicaceae family was represented by 10 species in our sample, with ~16 million years between some (Franzke et al., 2011). Despite this, sequence conservation around the auxRE motifs was very high. The conservation followed gene-specific patterns, but notably, declined with increasing distance from the motif, indicating nucleotides closer to the motif are more

highly conserved (Figure 5B). Similar to the auxRE, sequence conservation for regulatory elements extended beyond the conserved RE for genes containing either cytRE or abRE, and declined away from it (Supplemental Data Sets 7 and 8). Overall, our analysis revealed that while the general response to auxin may be defined by the short auxRE motif, the broad sequence context also tends to be conserved. Conservation beyond the motif was more pronounced within plant families and while part of this apparent conservation may be due to kinship, the fact the nucleotides closer to the core motif are more highly conserved than nucleotides further away from it suggests that these sequences are selected to serve family-specific functionality.

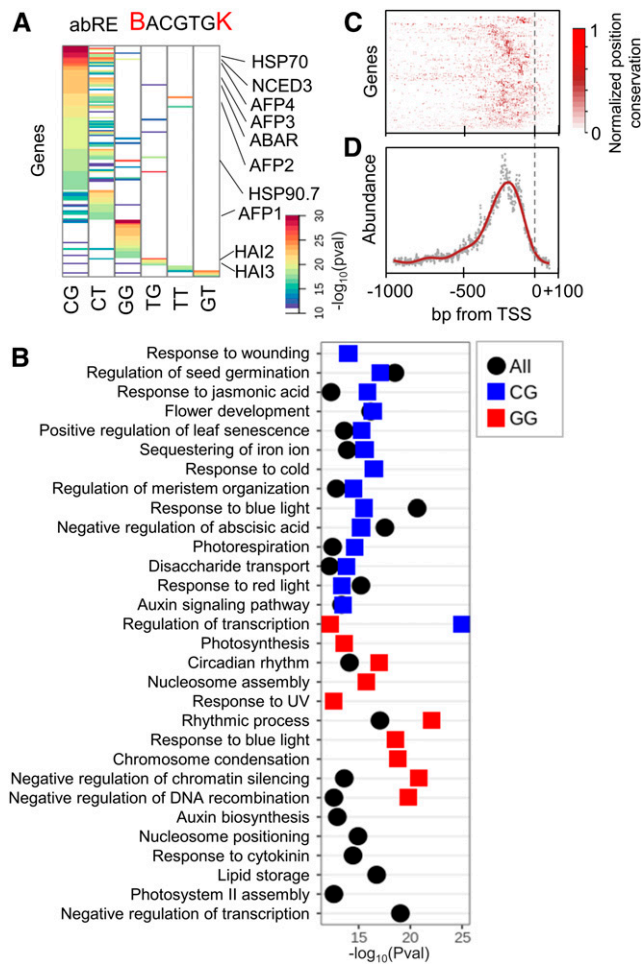


Figure 4. Conservation of abRE Variants.

(A) Genes and motif variants of the abRE showing significant conservation. pval, P-value.

(B) GO term enrichment of genes with a conserved abRE variant in their promoter. Pval, P-value.

(C) and **(D)** Conservation of the abRE motif position for each gene for the 45 angiosperms species **(C)** and the mean conservation across all promoters **(D)**. High values (red) represent high conservation of position of the motif.

auxRE Motif Variants Control the Transcriptional Response Profile in Synthetic Promoters

To identify a possible role of the RE variable nucleotides in planta, we generated four versions of the canonical DR5 synthetic promoter containing different variants of the auxRE motif: the conserved TGTCCC, TGTCTC, and TGTCAC and the non-conserved TGTCGC (Supplemental Figure 4A). Promoters were cloned upstream of the fluorescent protein *VENUS*. Arabidopsis root protoplasts were transfected with the construct and then treated with auxin for 6 h. Fluorescence was measured using a flow cytometer. While all four variants responded to auxin in a dose-dependent manner, there was a significant difference in response magnitude, with TGTCCC eliciting the strongest response and TGTCAC the weakest response (Figure 6A). The

attenuated response of the TGTCAC motif was unexpected, as ARFs bind this sequence in vitro (Boer et al., 2014). To test whether ARF proteins can activate this motif in planta, we overexpressed two A-class ARFs, *ARF7* and *ARF8*, together with our reporter. *ARF* overexpression successfully restored the auxin responsiveness of the TGTCAC motif to a similar level as that recorded for TGTCTC, suggesting that in planta, a higher ARF concentration can compensate for lower activity associated with specific variant nucleotides (Figure 6B). Despite the TGTCAC motif producing a weak response to auxin, 18 genes were shown to have this conserved motif in their promoters, including the auxin transporter gene *ABCB1* (Bailey et al., 2008) and two *MYB* genes (Supplemental Data Set 2). However, we could not identify an obvious common functionality for these genes.

To test how the variant nucleotides affect transcriptional response in a developmental context, we generated stably transformed Arabidopsis lines of the different *DR5* variants fused to GFP and selected a representative line (out of at least 10 independent transformation events). At 5 d after germination, seedlings were treated with auxin for 6 h. Reporter expression and RNA levels, as measured by RT-qPCR, recapitulated the transient expression results (Figures 6C and 6D). Interestingly, we observed a difference in the spatial expression pattern, where the response in TGTCAC lines was restricted to the root elongation zone, as opposite to the almost ubiquitous response of other variants. Notably, our results show that the magnitude of transcriptional response mediated by the motif variants did not correlate with its conservation level, suggesting that the conserved motifs are not merely the strongest activators, but mediate a specific response profile.

Conserved auxRE Variant Fine-Tunes the Auxin Transcriptional Response in Native Promoters

To test the impact of the motif variants in native promoters, we modified the conserved TGTCCC auxRE motif to the weak-response variant TGTCAC in the promoters of two auxin-responsive genes, *IAA2* and *IAA4*. Both promoters had a second TGTCCC site with a 5-bp spacer, a configuration preferred by ARF dimers (O'Malley et al., 2016; Stigliani et al., 2018). As this pair arrangement was perfectly conserved within the Brassicaceae, we mutated both motifs (Figures 7A to 7F).

Roots of *IAA2:GFP* plants exhibited strong fluorescence in the root tip and in the stele, which was upregulated by auxin, consistent with previous reports (Swarup et al., 2001; Grieneisen et al., 2007). By contrast, roots carrying the mutated *pIAA2_{c2a}:GFP* had weak-to-no *IAA2* expression in the root tip and weak stele expression, with auxin triggering a mild increase in stele expression (Figure 7G). Roots of *pIAA4_{c2a}:GFP* exhibited reduced fluorescence and auxin responsiveness compared with *pIAA4:GFP*, but the response was stronger than *pIAA2_{c2a}:GFP* (Figure 7G), possibly due to an additional TGTCCC RE in their promoter (Figure 7B). To verify that these changes in expression were not mediated by position effect, we screened multiple insertion lines and found consistent effects of the C-to-A change in the auxRE (Figure 7H).

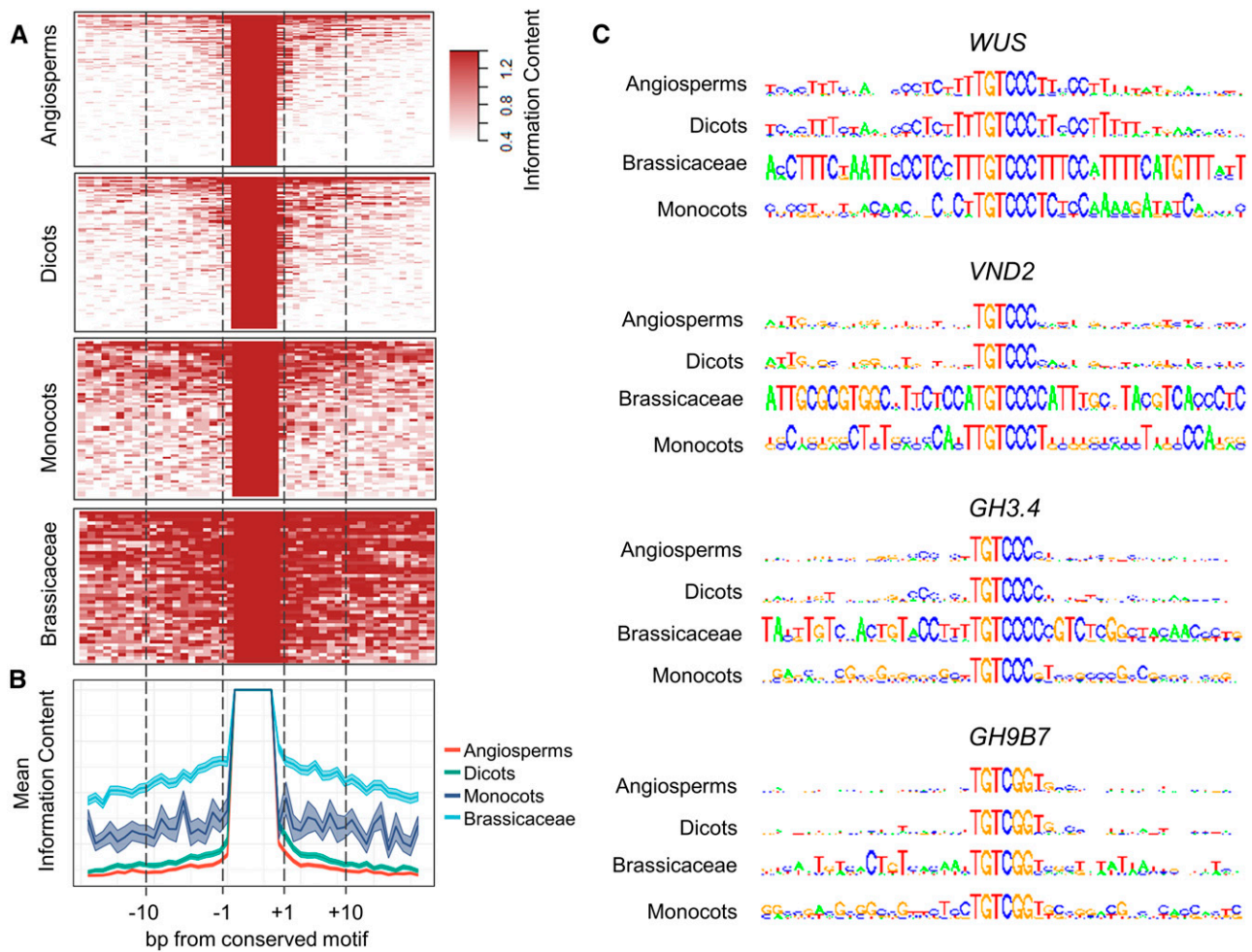


Figure 5. Deep Conservation of Sequence Context for Conserved auxRE Variants.

(A) Information score values computed from PWM for sequences flanking the conserved auxRE.

(B) Mean conservation of each position in the flanking sequences of conserved motifs described in **(A)**. Note the gradual reduction in conservation away from the conserved motif.

(C) PWM for sequences around the highly conserved auxRE variant for selected genes.

auxRE Conservation Enables Rational Design of Synthetic Reporters in Multiple Species

Following our observation that the TGTCCC motif variant in the promoter of the *AUX/IAA* genes is an important determinant of the magnitude of transcriptional response to auxin, we asked whether we can use this motif to generate a new synthetic reporter for auxin. To maintain conserved position of the motif at -220 to -280 bp to the TSS and to avoid introduction of repetitive sequences that are prone to silencing, we cloned three ~ 50 -bp regions of conserved TGTCCC-containing motifs from three different *Arabidopsis* and *Populus trichocarpa* promoters and placed them upstream of a minimal 35S promoter (pIAAmotif; Supplemental Figure 5). *Arabidopsis* plants carrying this reporter exhibit broad expression in the root, especially in the root cap, in a region that overlapped but expanded the original DR5 (Figures 8A and 8B). Following treatment with auxin, expression of the

reporter was rapidly induced in the epidermis, cortex and internal stele tissues in a broader pattern than DR5 (Figures 8C and 8D).

In order to determine the evolutionary conservation of these responses, we also generated tomato plants driving *RED FLUORESCENT PROTEIN* from the same pIAAmotif reporter. Similar to *Arabidopsis*, the pIAAmotif reporter exhibited overlapping but broader expression domains compared with the classical DR5 in both roots and shoots (Figures 8E to 8I), supporting the notion that the designed high-sensitivity promoters can be used in broad evolutionary contexts.

DISCUSSION

How promoters encode specific transcriptional responses is still far from fully understood. The presence of particular DNA motifs contributes to TF binding but cannot explain the complex

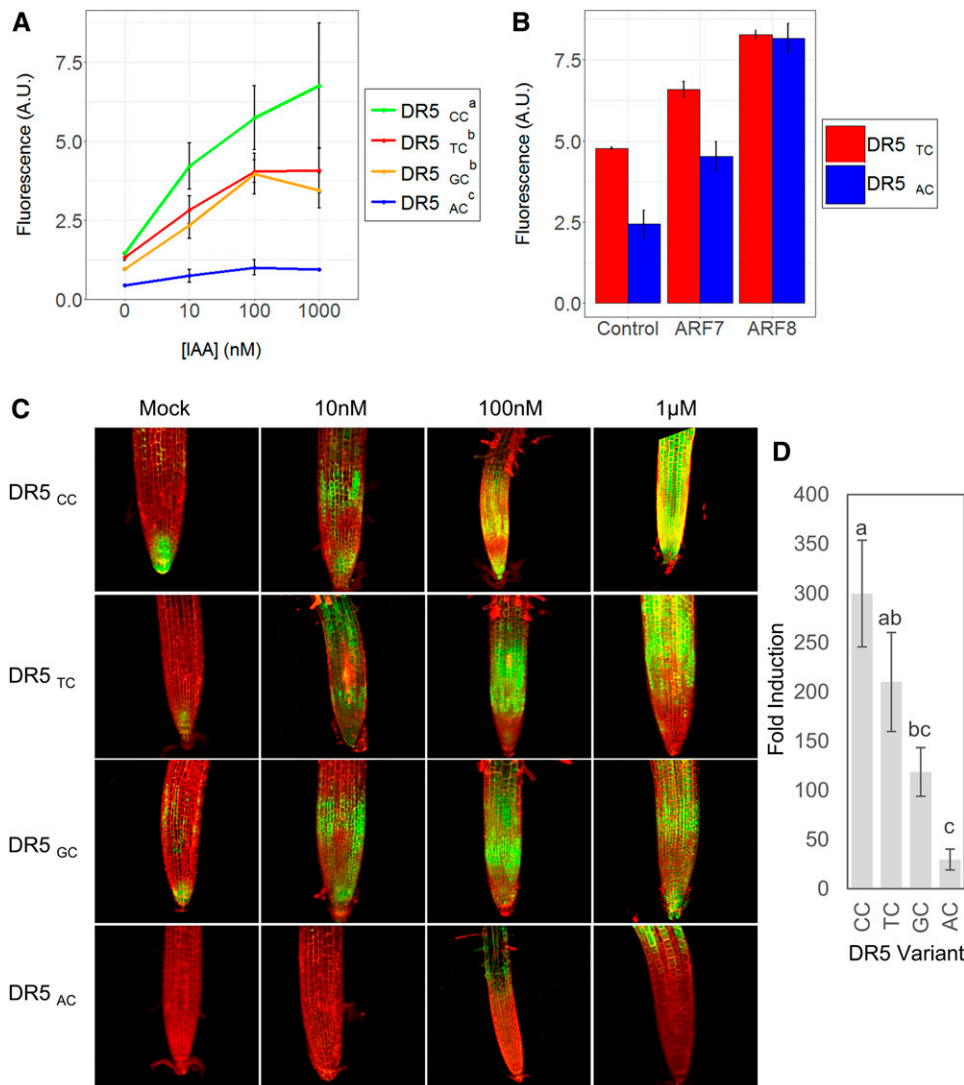


Figure 6. Auxin Responsiveness of Different auxRE Variants.

(A) and (B) Mean normalized fluorescence of protoplasts transfected with variants of DR5 and treated with auxin for 6 h (A) or with DR5 variants and an activating ARF (B). Experiments were performed in triplicates as independent transformation events. Superscript in the legend (A) indicates significance as computed by a Tukey post hoc test ($\alpha = 0.05$). Induction of DR5_{AC} was significant at $P < 0.05$ (Student's *t* test). Error bars represent se. A.U., arbitrary units. (C) Maximal projection of confocal images of roots carrying GFP driven by different auxRE variants. Roots were treated with auxin for 6 h.

(D) Fold-change of the GFP transcript 6 h following auxin application to roots carrying different DR5 variants, as measured by qPCR. Error bars represent se. $n = 4$. Letters indicate statistical significance (Tukey post hoc test, $\alpha = 0.05$).

transcriptional response of the gene. Many sequence features contribute to TF-DNA binding specificity and ability to activate transcription of downstream genes. ARF proteins, for example, form complexes and preferentially bind RE pairs with specific spacing (Boer et al., 2014; O'Malley et al., 2016; Galli et al., 2018). In this work, we highlight the contribution of specific variant nucleotides in the vicinity of the core motif to the promoter transcriptional response profile.

Identification of regulatory sequences in promoters is statistically challenging due to the large size of promoters. Methods that can reduce the sequence search space are required to address

this problem. By confining the search just to two or three nucleotides in specific sequence contexts, CoMoVa was able to identify significant conservation of many REs even in relatively long sequences. The ability to identify such conservation is increased with a larger number of genomes and is reduced by a larger sequence search space (longer promoters) or by high degeneracy of core motif. With the current number of genomes, the most effective promoter length for the method is ~ 1 kb, although we could identify significant conservation even when 2-kb and 3-kb promoters were used. Furthermore, while the method may not currently be well suited for TFs with highly degenerate

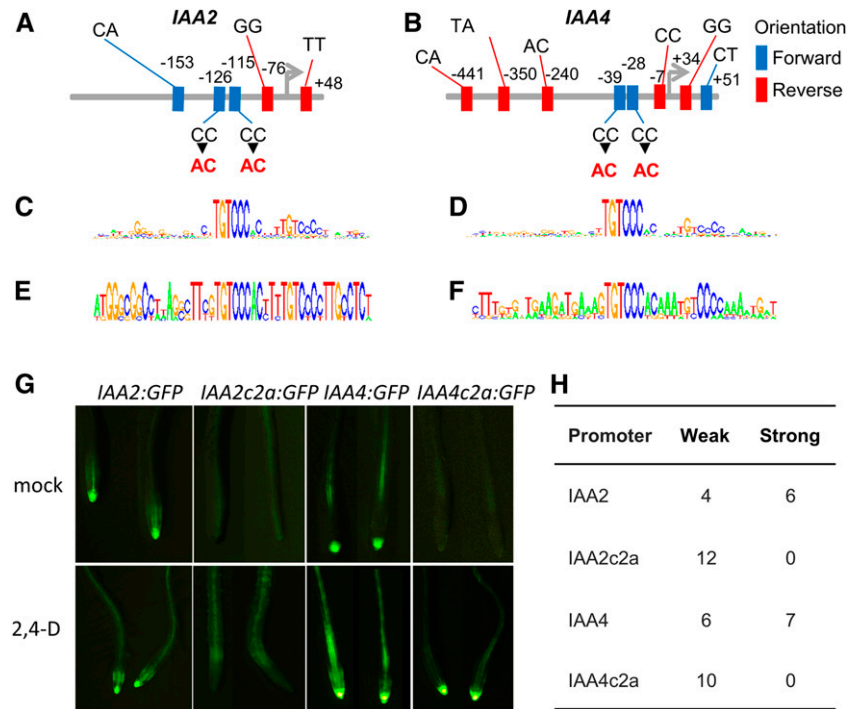


Figure 7. auxRE Variants Generate Predictable Auxin Response Profiles in Native and Synthetic Promoters.

(A) and (B) Diversity of auxRE motifs at the promoters of Arabidopsis *IAA2* (A) and *IAA4* (B).

(C) to (F) PWM of the conserved TGTCC₂ auxRE for *IAA2* (C) and (E) and *IAA4* (D) and (F) in all angiosperms (C) and (D) or within the Brassicaceae (E) and (F).

(G) Response of promoters and mutated promoters to auxin.

(H) Number of independent transformation lines exhibiting the same expression as in (G).

binding sites, an increase in the number of sequenced plant genomes can compensate for the low information in the motif and expand the usability of the method.

Conservation of the variant nucleotides in the core binding sequence was common in genes making up the core feedback response for the hormone tested. This raises the hypothesis that motif conservation can be used to identify TF targets. While this is limited only to conserved targets of the TF, it may help uncover the principal activities of different factors. For example, cytokinins have long been tightly associated with promotion of cell division, but the mechanism by which cytokinins mediate their control of the cell cycle has been elusive (Schaller et al., 2014). Interestingly, apart from the A-class *RRs*, which are known downstream factors of cytokinins, we also identified deep conservation of specific cytREs in promoters of key histones and cell cycle genes, suggesting that *RRs* may have direct transcriptional control on the cell cycle machinery. While the focus of this work was on transcriptional response to hormonal transcriptional responses, a degenerate RE was defined to many TFs (Franco-Zorrilla et al., 2014; O'Malley et al., 2016) and a similar approach can be applied to them.

Motif variant conservation can be used to apply rational promoter design. For example, the recently developed 6xABRE synthetic abscisic acid reporter is based on repeats of the TACGTGTC abRE variant (Wu et al., 2018). However, motif

conservation analyses showed that this RE is poorly conserved, while the ATnnAACACGTGG variant (Figure 4A; Supplemental Data Set 7) is strongly conserved in highly responsive *NCED* genes. It would be interesting to assess whether such variants can be used to generate reporters with different degrees of sensitivity to the hormones. Furthermore, the demonstration that position relative to the TSS is highly conserved for some REs, warrants maintenance of this relative position in synthetic promoters.

The advent of gene editing techniques has opened possibilities for generation of new alleles in important crops by directed mutagenesis of gene promoters. We propose that RE conservation can be used as a heuristic tool for identification of gene editing targets in order to generate new quantitative alleles for breeding purposes (Rodríguez-Leal et al., 2017).

METHODS

Plant Growth and Imaging

Arabidopsis (*Arabidopsis thaliana*) seeds were planted on half strength Murashige and Skoog agar plates, kept for 48 h in the dark at 4°C, and transferred to a growth chamber at 21°C in continuous light. Tomatoes (*Solanum lycopersicum* cv M82) were germinated in a growth chamber and moved to a temperature-controlled greenhouse with natural light (day, 26°C; night, 18°C). Transgenic *Arabidopsis* were generated using floral

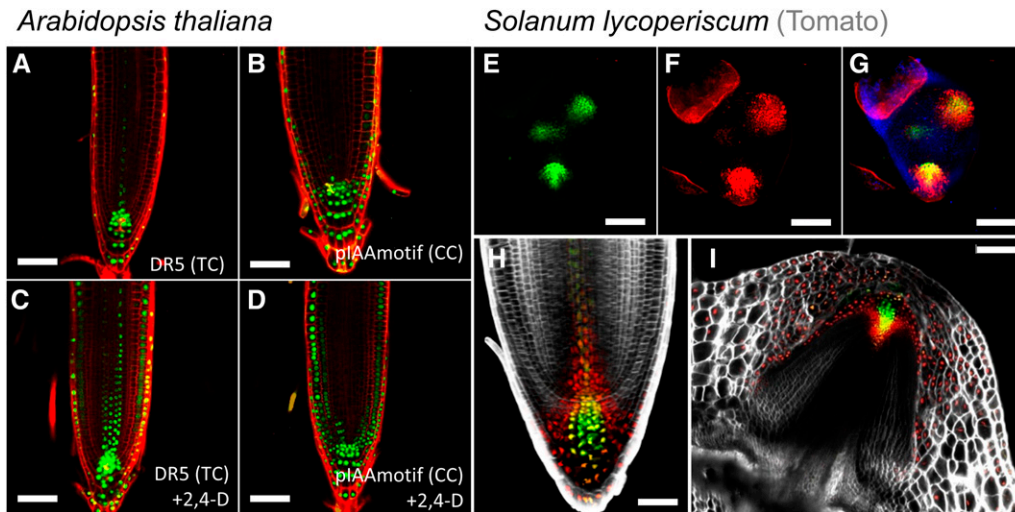


Figure 8. New Synthetic Sensitive Auxin Reporter.

(A) to (D) Roots of 7-d-old *Arabidopsis* plants carrying *DR5* and *pIAAmotif:mNeonGreen-NLS* treated with mock ([A] and [B]) or $1 \mu\text{M}$ 2,4-D ([C] and [D]). (E) to (I) Shoot meristem ([E] to [G]), main root ([H]), and stem-borne root ([I]) of tomato carrying *DR5:3xVENUS-N7* (green) and *pIAAmotif->opRFP* (red). Blue in (G) is chlorophyll autofluorescence. Cell wall stained with propidium iodide ([A] to [D]) or SR2200 ([H] and [I]).

dipping (Clough and Bent, 1998) and phenotyped at T3. Transgenic tomatoes were generated using *Agrobacterium tumefaciens* (GV3103)-mediated cotyledon transformation (Van Eck et al., 2006). Following one backcross, they were crossed with *DR5:3xVENUS-N7* tomatoes and examined in the F1 generation. Images were taken using a SMZ18 fluorescence stereoscope (Nikon), with identical settings for all images. Confocal microscopy was performed using a SP8 microscope (Leica). For *Arabidopsis* roots, propidium iodide was used to stain cell walls. Tomato lines carrying the *pIAAmotif:LhG4* driver were crossed to plants carrying *op:mRFP*. Tomato roots and adventitious roots were fixed and cleared using ClearSee (Kurihara et al., 2015), and cell wall staining was performed using SR2200 (Renaissance Chemicals) prior to mounting and visualization using 405, 488, and 561 nm lasers. To image the shoot meristem, live tomato shoots were dissected, mounted in gel, and imaged using a Lightsheet Z.1 microscope (Zeiss).

Construct Generation

For the protoplast assay, DNA for the four *DR5* variants was synthesized and cloned, using Gateway (Invitrogen), upstream to a minimal 35S promoter fused to Venus. To generate transgenic *Arabidopsis*, the same variants were cloned, using Gateway, to the pKGWFS7 plasmid, upstream to a GFP- β -glucuronidase fusion protein. For ARF overexpression, the coding sequence was cloned, using Gateway, into a p2GW7.0 plasmid. To generate the *pIAA2* and *pIAA4* promoter, the sequences 1200 and 1500 bp upstream, respectively, were amplified and cloned into pENTR using Gateway. Site-directed mutagenesis was used to replace the TGTCCC motifs with TGTCAC. The fragments were then shuttled to the pKGWFS7 vector. Cloning primers are listed in Supplemental Table 2. The synthetic auxin reporter *pIAA* motif was constructed by cloning the TGTCCC-containing regulatory regions of the *IAA1* ortholog from *Populus trichocarpa*, *IAA1* from *Arabidopsis*, and *IAA2* from *Arabidopsis*, upstream to a minimal 35S promoter. For *Arabidopsis* lines, the three segments were cloned into a level 0 MoClo part using Golden Gate and then fused to *mNeonGreen-N7* and a heat shock protein terminator to form a level 1 construct. The level 1 was transferred to a level 2, together with a kanamycin resistance cassette. For tomato transformation, the synthetic

promoter was cloned upstream to *LhG4* using restriction cloning and then subcloned into pART27.

Protoplast Assay

Arabidopsis plants were grown on agar plates as described in the section "Plant Growth and Imaging." At 7 d after transfer to light, protoplasts were extracted and transfected with $10 \mu\text{g}$ of plasmid carrying the synthetic reporter fused to VENUS, together with a transfection control of $10 \mu\text{g}$ of pMon plasmid, carrying *35S:RFP*, according to published protocols (Bargmann and Birnbaum, 2009). Cell aliquots were treated with $1 \mu\text{M}$ 2,4-D or mock treated for 6 h. Mean yellow fluorescent protein and RFP fluorescence of at least 10,000 cells was measured using a BD Accuri 6 Plus analyzer. The mean RFP signal was used to normalize the yellow fluorescent protein signal. Experiments were performed in triplicates of independent transformation events.

qPCR Analysis

Arabidopsis seedlings at 10 d after transfer to light were transferred to agar plates containing $1 \mu\text{M}$ 2,4-D, or DMSO (mock), and incubated for 6 h. RNA was extracted from ~ 100 seedlings using TRI REAGENT (MRC) according to the manufacturer's instructions, treated with DNAase (TURBO DNA-free kit, Invitrogen) to remove DNA remains, and cDNA was synthesized using SuperScript II reverse transcriptase (Invitrogen) with $1 \mu\text{g}$ of total RNA according to the manufacturer's instructions. RT-qPCR was performed using Absolute Blue qPCR SYBR Green (AB-4162, Thermo Fisher Scientific) in a Rotor-Gene 6000 cyclor (Corbett Research) with three technical repeats for each sample. Four biological replicates, from independently grown seedlings, were used. *ACTIN* was used as a reference gene. Primers are listed in Supplemental Table 2.

Ortholog Calling

For each species, a reciprocal blast was performed between all proteins and the *Arabidopsis* protein sequences using a minimum E-value of $1\text{E}-10$. Orthologs were ranked based on E-value. An ortholog quality

score for each gene was computed using a custom R script (buildOrthologDB.r), which computed the combined rank in the reciprocal blast, with 2 being the highest score. Gene pairs were sorted based on ortholog quality score, and the genes with lowest scores (up to eight genes) were considered as putative orthologs. The mean number and sd of putative orthologs per species, for each Arabidopsis gene was 2.61 ± 0.34 genes.

Computational Identification of Motif Conservation

Genome sequences and annotation gff3 files were obtained from Phytozome 12, and a perl script (buildpromoterDB.perl) was used to generate upstream sequences and identify the motifs variants for each gene. Upstream sequences (1 kb upstream and 100 bp downstream, unless specified otherwise) were isolated from the genome file based on the GFF annotation. If an annotated gene was found within the promoter region, its sequences were masked and not used for identification of RE. A database of all occurrences of a given motif in the gene promoters was generated using a pattern matching script written in Perl. Calculation of conservation scores was performed as follows. For each Arabidopsis gene, the orthologs for all species were obtained, as described in the "Ortholog Calling" section. Of the putative orthologs for a given species, the one with a motif variant composition most similar to the Arabidopsis gene was selected as an ortholog. The motif variants for each gene were arranged on a predefined species tree. The maximum parsimony method was used to derive the motif variants on the internal nodes. The conservation score for each motif variant was calculated as the number of occurrences of the variant in the tree minus the number of changes (gain or loss) this variant had across the tree. To determine the background distribution, conservation scores of neutral nucleotides, as defined by the background search motif (NNnnnnnTGTC, DnnnnnnnGATCnnnnnnnYN, and BnnnnnnnACGTGnnnnnnnK for AuxRE, CytRE, and abRE, respectively), were calculated for all genes. A negative binomial distribution was fitted to the overall conservation score for the neutral variants. This fitted distribution was used to compute P-values for the conservation of the motif variants. Significance cutoffs for the $-\log_2(p\text{-value})$ were set to 17, 17, and 38 for the auxRE, abRE, and cytRE motifs, respectively. All R and perl scripts are deposited at <https://github.com/idanefroni/CoMoVa>.

GO Term Analysis

GO term analysis was performed for the Arabidopsis orthologs of the genes with conserved motifs, using the topGO package with the "weight01" algorithm. P-values were calculated using the hypergeometric test without further correction to multiple testing. Enrichment testing was performed for each variant individually and then for the entire list of genes with conserved variants.

Identification of Auxin-Responsive Genes

For the root response, tomato seedlings were germinated on half strength Murashige and Skoog plates and grown in a growth chamber at 22°C under continuous light. Four days after germination, seedlings were transferred to plates containing 1 μM of the auxin analog 2,4-D (D7299, Sigma-Aldrich) for 3 h, followed by excision of 1 cm of root tips and flash freezing in liquid nitrogen. For shoot samples, 3-week-old M82 plants were sprayed with either water (mock) or 1 μM of the auxin analog picloram (P5575, Sigma-Aldrich). Three hours post-application, 15 young leaf primordia (leaf number 5 at the P5 stage) were collected and flash frozen. RNA extraction was performed using Qiagen RNeasy Micro kit, and sequencing libraries were prepared using Lexigen 3' Quant-Seq kit, according to the manufacturer's instructions. Experiments were performed in duplicates or

triplicates. Single-end sequencing was performed using an Illumina NextSeq 500 system. Expression calling was performed using Salmon 0.8.2, by aligning to the tomato genome ITAG3.2. As the library is 3' biased, we extended all genes to cover 500 additional base pairs downstream. Expression data were normalized using Deseq2, and genes induced at false discovery rate < 0.01 were selected. Processed expression values for the Arabidopsis and *Brachypodium* root auxin response were obtained from analysis previously published by Bargmann et al., (2013) and Lewis et al., (2013). Maize (*Zea mays*) auxin response data (Galli et al., 2018) were obtained from Gene Expression Omnibus (GSE111792), aligned to the maize transcriptome (AGPv3, obtained from Phytozome), and auxin-responsive genes were identified using a similar pipeline as for the tomato genes.

Accession Numbers

Tomato auxin response data are deposited in Gene Expression Omnibus (series GSE126372).

Supplemental Data

Supplemental Figure 1. Genome properties of the angiosperms used in this study.

Supplemental Figure 2. Effect of different nucleotide selection for calculation of background distribution.

Supplemental Figure 3. Effect of promoter length on the number of identified conserved motifs.

Supplemental Figure 4. Sequence and response of DR5 variants.

Supplemental Figure 5. Synthetic pIAA motif high-sensitivity auxin reporter.

Supplemental Table 1. List of genomes used in this study.

Supplemental Table 2. Primers used in this study.

Supplemental Data Set 1. Ortholog candidates for all genes in the species used in this study.

Supplemental Data Set 2. List of genes with conserved auxRE motifs.

Supplemental Data Set 3. Gene lists and enriched GO terms for genes with conserved auxRE in varying promoter lengths.

Supplemental Data Set 4. List of genes with conserved cytRE motifs.

Supplemental Data Set 5. List of genes with conserved abRE motifs.

Supplemental Data Set 6. Sequence conservation of sequence flanking conserved auxRE motifs.

Supplemental Data Set 7. Sequence conservation of sequence flanking conserved cytRE motifs.

Supplemental Data Set 8. Sequence conservation of sequence flanking conserved abRE motifs.

ACKNOWLEDGMENTS

We thank Sigal Savaldi-Goldstein and Yuval Eshed for their comments on the work, Ziva Amsellem for help with tomato transformation, and Yossi Capua for assistance with microscopy. Support for this study was provided by the Israel Science Foundation (ISF966/17) and the Howard Hughes Medical Institute (55008730).

AUTHOR CONTRIBUTIONS

M.L.-L., C.Y., and A.I. performed the experiments. I.E. designed the research and coded the computational tools. M.L.-L. and I.E. wrote the article.

Received February 25, 2019; accepted August 27, 2019; published August 29, 2019.

REFERENCES

- Abel, S., Ballas, N., Wong, L.M., and Theologis, A. (1996). DNA elements responsive to auxin. *BioEssays* **18**: 647–654.
- Abel, S., Nguyen, M.D., Chow, W., and Theologis, A. (1995). ASC4, a primary indoleacetic acid-responsive gene encoding 1-aminocyclopropane-1-carboxylate synthase in *Arabidopsis thaliana*. *J. Biol. Chem.* **270**: 19093–19099.
- Abel, S., and Theologis, A. (1996). Early genes and auxin action. *Plant Physiol.* **111**: 9–17.
- Bailly, A., Sovero, V., Vincenzetti, V., Santelia, D., Bartnik, D., Koenig, B.W., Mancuso, S., Martinoia, E., and Geisler, M. (2008). Modulation of P-glycoproteins by auxin transport inhibitors is mediated by interaction with immunophilins. *J. Biol. Chem.* **283**: 21817–21826.
- Ballas, N., Wong, L.M., Ke, M., and Theologis, A. (1995). Two auxin-responsive domains interact positively to induce expression of the early indoleacetic acid-inducible gene PS-IAA4/5. *Proc. Natl. Acad. Sci. USA* **92**: 3483–3487.
- Bargmann, B.O.R., and Birnbaum, K.D. (2009). Positive fluorescent selection permits precise, rapid, and in-depth overexpression analysis in plant protoplasts. *Plant Physiol.* **149**: 1231–1239.
- Bargmann, B.O.R., and Estelle, M. (2014). Auxin perception: In the IAA of the beholder. *Physiol. Plant.* **151**: 52–61.
- Bargmann, B.O.R., Vanneste, S., Krouk, G., Nawy, T., Efroni, I., Shani, E., Choe, G., Friml, J., Bergmann, D.C., Estelle, M., and Birnbaum, K.D. (2013). A map of cell type-specific auxin responses. *Mol. Syst. Biol.* **9**: 688.
- Barrero, J.M., Rodríguez, P.L., Quesada, V., Piqueras, P., Ponce, M.R., and Micol, J.L. (2006). Both abscisic acid (ABA)-dependent and ABA-independent pathways govern the induction of NCED3, AAO3 and ABA1 in response to salt stress. *Plant Cell Environ.* **29**: 2000–2008.
- Bhargava, A., Clabaugh, I., To, J.P., Maxwell, B.B., Chiang, Y.-H., Schaller, G.E., Loraine, A., and Kieber, J.J. (2013). Identification of cytokinin-responsive genes using microarray meta-analysis and RNA-seq in *Arabidopsis*. *Plant Physiol.* **162**: 272–294.
- Boer, D.R., Freire-Rios, A., van den Berg, W.A., Saaki, T., Manfield, I.W., Kepinski, S., López-Vidriero, I., Franco-Zorrilla, J.M., de Vries, S.C., Solano, R., Weijers, D., and Coll, M. (2014). Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors. *Cell* **156**: 577–589.
- Chaw, S.M., Chang, C.C., Chen, H.L., and Li, W.H. (2004). Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *J. Mol. Evol.* **58**: 424–441.
- Choi, H., Hong, J., Ha, J., Kang, J., and Kim, S.Y. (2000). ABFs, a family of ABA-responsive element binding factors. *J. Biol. Chem.* **275**: 1723–1730.
- Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**: 735–743.
- De Witte, D., Van de Velde, J., Decap, D., Van Bel, M., Audenaert, P., Demeester, P., Dhoedt, B., Vandepoele, K., and Fostier, J. (2015). BLSSpeller: Exhaustive comparative discovery of conserved cis-regulatory elements. *Bioinformatics* **31**: 3758–3766.
- Dror, I., Golan, T., Levy, C., Rohs, R., and Mandel-Gutfreund, Y. (2015). A widespread role of the motif environment in transcription factor binding across diverse protein families. *Genome Res.* **25**: 1268–1280.
- Elemento, O., and Tavazoie, S. (2005). Fast and systematic genome-wide discovery of conserved regulatory elements using a non-alignment based approach. *Genome Biol.* **6**: R18.
- Franco-Zorrilla, J.M., López-Vidriero, I., Carrasco, J.L., Godoy, M., Vera, P., and Solano, R. (2014). DNA-binding specificities of plant transcription factors and their potential to define target genes. *Proc. Natl. Acad. Sci. USA* **111**: 2367–2372.
- Franzke, A., Lysak, M.A., Al-Shehbaz, I.A., Koch, M.A., and Mummenhoff, K. (2011). Cabbage family affairs: The evolutionary history of Brassicaceae. *Trends Plant Sci.* **16**: 108–116.
- Galli, M., Khakhar, A., Lu, Z., Chen, Z., Sen, S., Joshi, T., Nemhauser, J.L., Schmitz, R.J., and Gallavotti, A. (2018). The DNA binding landscape of the maize AUXIN RESPONSE FACTOR family. *Nat. Commun.* **9**: 4526.
- Garcia, M.E., Lynch, T., Peeters, J., Snowden, C., and Finkelstein, R. (2008). A small plant-specific protein family of ABI five binding proteins (AFPs) regulates stress response in germinating *Arabidopsis* seeds and seedlings. *Plant Mol. Biol.* **67**: 643–658.
- Gentles, A.J., and Karlin, S. (2001). Genome-scale compositional comparisons in eukaryotes. *Genome Res.* **11**: 540–546.
- Gordán, R., Narlikar, L., and Hartemink, A.J. (2010). Finding regulatory DNA motifs using alignment-free evolutionary conservation information. *Nucleic Acids Res.* **38**: e90.
- Graur, D., Zheng, Y., and Azevedo, R.B.R. (2015). An evolutionary classification of genomic function. *Genome Biol. Evol.* **7**: 642–645.
- Grieneisen, V.A., Xu, J., Marée, A.F.M., Hogeweg, P., and Scheres, B. (2007). Auxin transport is sufficient to generate a maximum and gradient guiding root growth. *Nature* **449**: 1008–1013.
- Hattori, T., Totsuka, M., Hobo, T., Kagaya, Y., and Yamamoto-Toyoda, A. (2002). Experimentally determined sequence requirement of ACGT-containing abscisic acid response element. *Plant Cell Physiol.* **43**: 136–140.
- Hirayama, T., and Shinozaki, K. (2007). Perception and transduction of abscisic acid signals: Keys to the function of the versatile plant hormone ABA. *Trends Plant Sci.* **12**: 343–351.
- Ivan, A., Halfon, M.S., and Sinha, S. (2008). Computational discovery of cis-regulatory modules in *Drosophila* without prior knowledge of motifs. *Genome Biol.* **9**: R22.
- Kantorovitz, M.R., Robinson, G.E., and Sinha, S. (2007). A statistical method for alignment-free comparison of regulatory sequences. *Bioinformatics* **23**: i249–i255.
- Kieber, J.J., and Schaller, G.E. (2018). Cytokinin signaling in plant development. *Development* **145**: dev149344.
- Kim, D.W., Lee, S.H., Choi, S.-B., Won, S.-K., Heo, Y.-K., Cho, M., Park, Y.-I., and Cho, H.-T. (2006). Functional conservation of a root hair cell-specific cis-element in angiosperms with different root hair distribution patterns. *Plant Cell* **18**: 2958–2970.
- Korkuc, P., Schippers, J.H.M., and Walther, D. (2014). Characterization and identification of cis-regulatory elements in *Arabidopsis* based on single-nucleotide polymorphism information. *Plant Physiol.* **164**: 181–200.
- Kurihara, D., Mizuta, Y., Sato, Y., and Higashiyama, T. (2015). ClearSee: A rapid optical clearing reagent for whole-plant fluorescence imaging. *Development* **142**: 4168–4179.
- Lewis, D.R., Olex, A.L., Lundy, S.R., Turkett, W.H., Fetrow, J.S., and Muday, G.K. (2013). A kinetic analysis of the auxin transcriptome reveals cell wall remodeling proteins that modulate lateral root development in *Arabidopsis*. *Plant Cell* **25**: 3329–3346.

- Liao, C.-Y., Smet, W., Brunoud, G., Yoshida, S., Vernoux, T., and Weijers, D. (2015). Reporters for sensitive and quantitative measurement of auxin response. *Nat. Methods* **12**: 207–210, 2, 210.
- Mironova, V.V., Omelyanchuk, N.A., Wiebe, D.S., and Levitsky, V.G. (2014). Computational analysis of auxin responsive elements in the *Arabidopsis thaliana* L. genome. *BMC Genomics* **15**: S4.
- Mutte, S.K., Kato, H., Rothfels, C., Melkonian, M., Wong, G.K.-S., and Weijers, D. (2018). Origin and evolution of the nuclear auxin response system. *eLife* **7**: e33399.
- O'Malley, R.C., Huang, S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cis-trome and epicistrome features shape the regulatory DNA landscape. *Cell* **165**: 1280–1292.
- Omelyanchuk, et al. (2017). Auxin regulates functional gene groups in a fold-change-specific manner in *Arabidopsis thaliana* roots. *Sci. Rep.* **7**: 2489.
- Potter, K.C., Wang, J., Schaller, G.E., and Kieber, J.J. (2018). Cytokinin modulates context-dependent chromatin accessibility through the type-B response regulators. *Nat. Plants* **4**: 1102–1111.
- Rodriguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E., and Lippman, Z.B. (2017). Engineering quantitative trait variation for crop improvement by genome editing. *Cell* **171**: 470–480.
- Roosjen, M., Paque, S., and Weijers, D. (2018). Auxin response factors: Output control in auxin biology. *J. Exp. Bot.* **69**: 179–188.
- Schaller, G.E., Street, I.H., and Kieber, J.J. (2014). Cytokinin and the cell cycle. *Curr. Opin. Plant Biol.* **21**: 7–15.
- Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F., Jürgens, G., and Laux, T. (2000). The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* **100**: 635–644.
- Stigliani, A., Martin-Arevalillo, R., Lucas, J., Bessy, A., Vinos-Poyo, T., Mironova, V., Vernoux, T., Dumas, R., and Parcy, F. (2018). Capturing auxin response factors syntax using DNA binding models. *Mol. Plant* **12**: 822–832.
- Swarup, R., Friml, J., Marchant, A., Ljung, K., Sandberg, G., Palme, K., and Bennett, M. (2001). Localization of the auxin permease AUX1 suggests two functionally distinct hormone transport pathways operate in the *Arabidopsis* root apex. *Genes Dev.* **15**: 2648–2653.
- Tan, T.T., Endo, H., Sano, R., Kurata, T., Yamaguchi, M., Ohtani, M., and Demura, T. (2018). Transcription factors VND1–VND3 contribute to cotyledon xylem vessel formation. *Plant Physiol.* **176**: 773–789.
- Ulmasov, T., Hagen, G., and Guilfoyle, T.J. (1999). Dimerization and DNA binding of auxin response factors. *Plant J.* **19**: 309–319.
- Ulmasov, T., Liu, Z.B., Hagen, G., and Guilfoyle, T.J. (1995). Composite structure of auxin response elements. *Plant Cell* **7**: 1611–1623.
- Ulmasov, T., Murfett, J., Hagen, G., and Guilfoyle, T.J. (1997). Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9**: 1963–1971.
- Vandepoele, K., Quimbaya, M., Casneuf, T., De Veylder, L., and Van de Peer, Y. (2009). Unraveling transcriptional control in *Arabidopsis* using cis-regulatory elements and coexpression networks. *Plant Physiol.* **150**: 535–546.
- Van Eck, J., Kirk, D.D., and Walmsley, A.M. (2006). Tomato (*Lycopersicon esculentum*). In *Agrobacterium Protocols*, K. Wang, ed (Totowa: Humana Press), pp. 459–474.
- Wu, R., Duan, L., Pruneda-Paz, J.L., Oh, D.-H., Pound, M., Kay, S., and Dinneny, J.R. (2018). The 6xABRE synthetic promoter enables the spatiotemporal analysis of ABA-mediated transcriptional regulation. *Plant Physiol.* **177**: 1650–1665.
- Xie, M., Chen, H., Huang, L., O'Neil, R.C., Shokhirev, M.N., and Ecker, J.R. (2018). A B-ARR-mediated cytokinin transcriptional network directs hormone cross-regulation and shoot development. *Nat. Commun.* **9**: 1–13.
- Yang, Y.Z., and Tan, B.C. (2014). A distal ABA responsive element in AtNCED3 promoter is required for positive feedback regulation of ABA biosynthesis in *Arabidopsis*. *PLoS One* **9**: e87283.
- Zemlyanskaya, E.V., Wiebe, D.S., Omelyanchuk, N.A., Levitsky, V.G., and Mironova, V.V. (2016). Meta-analysis of transcriptome data identified TGTCNN motif variants associated with the response to plant hormone auxin in *Arabidopsis thaliana* L. *J. Bioinform. Comput. Biol.* **14**: 1641009.
- Zubo, Y.O., et al. (2017). Cytokinin induces genome-wide binding of the type-B response regulator ARR10 to regulate growth and development in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **114**: E5995–E6004.
- Zürcher, E., Tavor-Deslex, D., Lituiev, D., Enkerli, K., Tarr, P.T., and Müller, B. (2013). A robust and sensitive synthetic sensor to monitor the transcriptional output of the cytokinin signaling network in planta. *Plant Physiol.* **161**: 1066–1075.