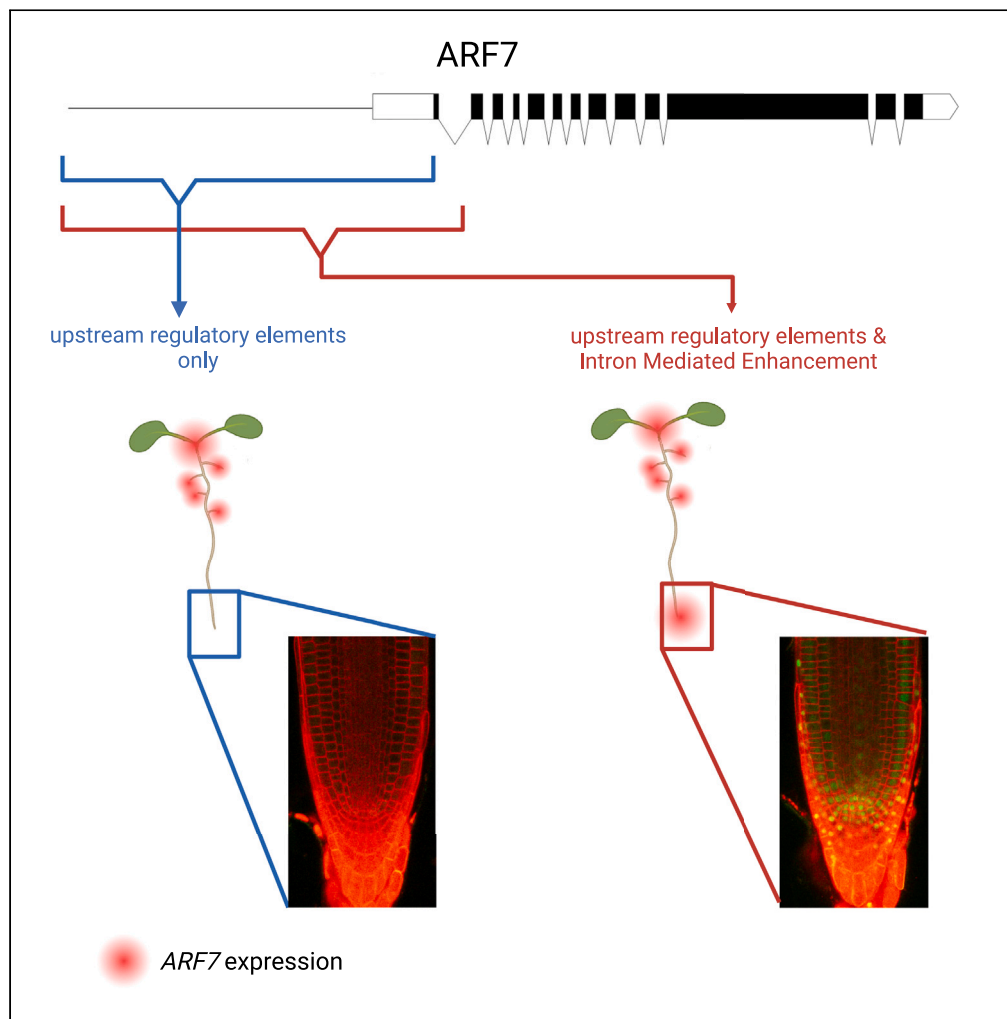


## Article

## The first intron of ARF7 is required for expression in root tips



Jingyi Han (撒静宜),  
Thomas Welch, Ute  
Voß, Teva Vernoux,  
Rahul Bhosale,  
Anthony Bishopp

anthony.bishopp@nottingham.  
ac.uk

## Highlights

The first intron of ARF7 is  
required for expression in  
the root apical meristem

It is not required for  
expression in either lateral  
roots or the shoot apex

We find no binding sites in  
the first intron required for  
expression in root tips

We propose that this intron  
exerts its effects via intron-  
mediated enhancement

Han et al., iScience 27, 109936  
June 21, 2024 © 2024 The  
Author(s). Published by Elsevier  
Inc.  
[https://doi.org/10.1016/  
j.isci.2024.109936](https://doi.org/10.1016/j.isci.2024.109936)

## Article

The first intron of ARF7  
is required for expression in root tipsJingyi Han (撒静宜),<sup>1,2</sup> Thomas Welch,<sup>1</sup> Ute Voß,<sup>1</sup> Teva Vernoux,<sup>3</sup> Rahul Bhosale,<sup>1,4</sup> and Anthony Bishopp<sup>1,5,\*</sup>

## SUMMARY

**Auxin regulates plant growth and development through the transcription factors of the AUXIN RESPONSE FACTOR (ARF) gene family. ARF7 is one of five activators that bind DNA and elicit downstream transcriptional responses. In roots, ARF7 regulates growth, gravitropism and redundantly with ARF19, lateral root organogenesis. In this study we analyzed ARF7 cis-regulation, using different non-coding sequences of the ARF7 locus to drive GFP. We show that constructs containing the first intron led to increased signal in the root tip. Although bioinformatics analyses predicted several transcription factor binding sites in the first intron, we were unable to significantly alter expression of GFP in the root by mutating these. We instead observed the intronic sequences needed to be present within the transcribed sequences to drive expression in the root meristem. These data support a mechanism by which intron-mediated enhancement regulates the tissue specific expression of ARF7 in the root meristem.**

## INTRODUCTION

Auxin plays an essential role in regulating growth and developmental processes in plants, such as the formation and maintenance of shoot apical meristem (SAM) and root apical meristem (RAM), vascular patterning, establishment of phyllotaxy, shoot phototropism, root gravitropism, and female gametophyte development.<sup>1,2</sup> Several gene families comprise the auxin signaling network and mediate transcription of downstream responses genes based on cellular auxin levels. The binding of auxin promotes the interaction between the SCF<sup>TIR1</sup> ubiquitin protein ligase complex and the Aux/IAA co-repressors.<sup>3</sup> The SCF-type complex comprises an F box protein<sup>4</sup> alongside the Skp1 and Cullin proteins.<sup>5</sup> This complex transfers an activated ubiquitin from a ubiquitin-activating enzyme and conjugates the Aux/IAAs.<sup>6,7</sup> The Aux/IAAs themselves act to stabilize auxin binding, leading to them being considered as co-receptors for auxin.<sup>8</sup> Following the ubiquitination of the Aux/IAAs, these proteins are degraded by the 26S proteasome.<sup>9,10</sup> In addition, TIR1/AFB synthesizes cAMP upon AUX/IAA binding, which is essential for a normal auxin response.<sup>11</sup> Transcription Factors (TFs) of the AUXIN RESPONSE FACTOR (ARF) family<sup>12</sup> bind to Auxin response elements within DNA.<sup>13</sup> The Aux/IAA proteins bind some of these ARFs and inhibit their transcriptional activity.<sup>14</sup> Under high auxin concentrations, the Aux/IAA proteins are degraded, thus allowing ARFs to regulate thousands of downstream responses.<sup>15,16</sup>

In Angiosperms, auxin signaling component families are encoded by large gene families. For example, the Arabidopsis genome encodes 23 ARF proteins that control distinct developmental processes.<sup>17</sup> They are divided to 3 clades based on phylogenetic analysis. Clade A has 5 members, including ARF5, ARF6, ARF7, ARF8, and ARF19<sup>18</sup> that have been shown to be transcriptional activators<sup>17</sup> and to be all bound by Aux/IAAs. As activators, they bind in promoter or enhancer regions of target genes to initiate or enhance transcription. Alongside these activators, ARF repressors belonging to the two other clades are thought to bind similar regions to inhibit the expression of auxin target genes.

These five Clade A activator ARFs (A-ARFs) play key specific roles in development of Arabidopsis. To understand how A-ARFs can mediate diverse functions throughout development, the expression of A-ARFs alongside other ARFs has been mapped at the cellular scale.<sup>19–21</sup> In both the root and shoot apical meristems as well as in other organs,<sup>17,20,22,23</sup> each A-ARF shows a unique but partially overlapping expression pattern and it is likely that each ARF or combination of ARFs regulate a different subset of targets. Consistently, over-expression of ARFs causes distinct phenotypes that are not present when other ARFs are overexpressed. For example, overexpression of ARF5 can restore the *arf7* hypocotyl elongation phenotype, but overexpression of ARF7 cannot restore the vascular defects in *arf5*.<sup>24</sup> While there is a certain degree of genetic redundancy within these components, there is evidence that the spatial patterns of ARF response are important for controlling the spatial specificity of auxin responses. Accordingly, either individual ARFs or pairs of ARFs have been associated with specific developmental responses.

ARF7 and ARF19 are well known for their redundant role in regulating lateral root organogenesis via activation of the transcription factors LATERAL ORGAN BOUNDARIES-DOMAIN (LBD)16 and 29<sup>17,25,26</sup>. ARF7 and 19 also control other processes within the root, such as cell wall

<sup>1</sup>School of Biosciences, University of Nottingham, Loughborough, UK

<sup>2</sup>Department of Biosciences, Durham University, Durham, UK

<sup>3</sup>Laboratoire Reproduction et Développement des Plantes, Université de Lyon, ENS de Lyon, CNRS, INRAE, Lyon, France

<sup>4</sup>International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, Telangana, India

<sup>5</sup>Lead contact

\*Correspondence: [anthony.bishopp@nottingham.ac.uk](mailto:anthony.bishopp@nottingham.ac.uk)

<https://doi.org/10.1016/j.isci.2024.109936>



composition and pectin dynamics during root hair tip growth through ERULUS,<sup>27</sup> adventitious root formation via regulating LBD16 and LBD18<sup>28</sup> as well as in tropic responses such as gravitropism.<sup>17,29</sup> In the aerial parts of the plant, besides the hypocotyl aforementioned, ARF7 is expressed in the shoot meristem<sup>21</sup> as well as in veins of maturing leaves, especially in older procambial strands, where it works together with other activating ARFs to control leaf formation.<sup>30</sup> ARF7 also works redundantly with ARF19 to control expansion of leaf cells.<sup>31</sup>

Recently we investigated the regulatory networks regulating the five A-ARFs and found that at these loci chromatin was generally open and that transcription was predominantly controlled by a collection of transcriptional repressors.<sup>20</sup> The genomic structure of the five class A ARFs generally featured a large first intron. As intronic sequences had previously been implicated in regulating expression patterns,<sup>32,33</sup> the expression of all five activating ARFs was investigated in transgenic reporters containing either only sequence 5' to the translational start site or in reporters containing an in-frame fusion of GFP to the second exon.<sup>20</sup> ARF7 showed an interesting expression pattern that deviated to those we observed for the other ARFs. While for ARFs 5, 6, 8, and 19, there was no discernible difference between the two regulatory fragments, there was a clear difference in the ARF7 expression pattern. A broad expression of ARF7 was observed in root tips only in reporters including DNA sequences downstream of translational start, while both reporters produced similar expression patterns in shoots. This study suggested the presence of regulatory elements either in the first intron or first exons guiding expression in roots but not shoots. Given that these regulatory sequences appeared to increase GFP transcription, we hypothesized that transcriptional activators would likely bind these sequences. However, this was in sharp contrast with the majority of components identified within the A-ARF regulatory networks that were predicted to be transcriptional repressors. In this paper, we explore the role that sequences downstream of the translational start site (including the first intron) play in regulating ARF7 expression and identify an organ-specific role for the first intron in boosting expression of ARF7 mRNA in the root meristem.

## RESULTS

### The first intron of ARF7 is required to drive expression in the root apical meristem

Our previous work had shown different expression patterns for ARF7 in the root when driven by a promoter including only sequence 5' of the translational start site to those when the Venus was present as an in-frame fusion to the second exon.<sup>20</sup> To test whether these differences were due to the inclusion of the first intron, rather than a result of either binding sites within the first exon or changes in translatability of the protein, we designed a series of constructs to separate these possibilities. We exploited recent advances in greengate cloning<sup>34</sup> using the dual expression and anatomy lines (DEAL).<sup>35</sup> Alongside the gene of interest, the DEAL vectors introduce a red plasma membrane marker (TdTomato) under the ubiquitously expressed *UBQ10* promoter, and act as a control showing the activity of the transgene. Initially, we created two constructs, one that contained only 5' sequence upstream of the translational start site and the other included an in-frame fusion to the second exon (Figure 1A). Similar to previous reports, the construct containing only 5' sequence (pro:GFP) showed virtually no GFP fluorescence in the RAM, while the construct including the first intron by way of a fusion between the second exon and GFP (pro:EX1-2:GFP) showed broad expression in the RAM, including within the stele, endodermis, cortex, and epidermis (Figures 1B and 1E). Both constructs showed roughly equivalent levels of TdTomato signal.

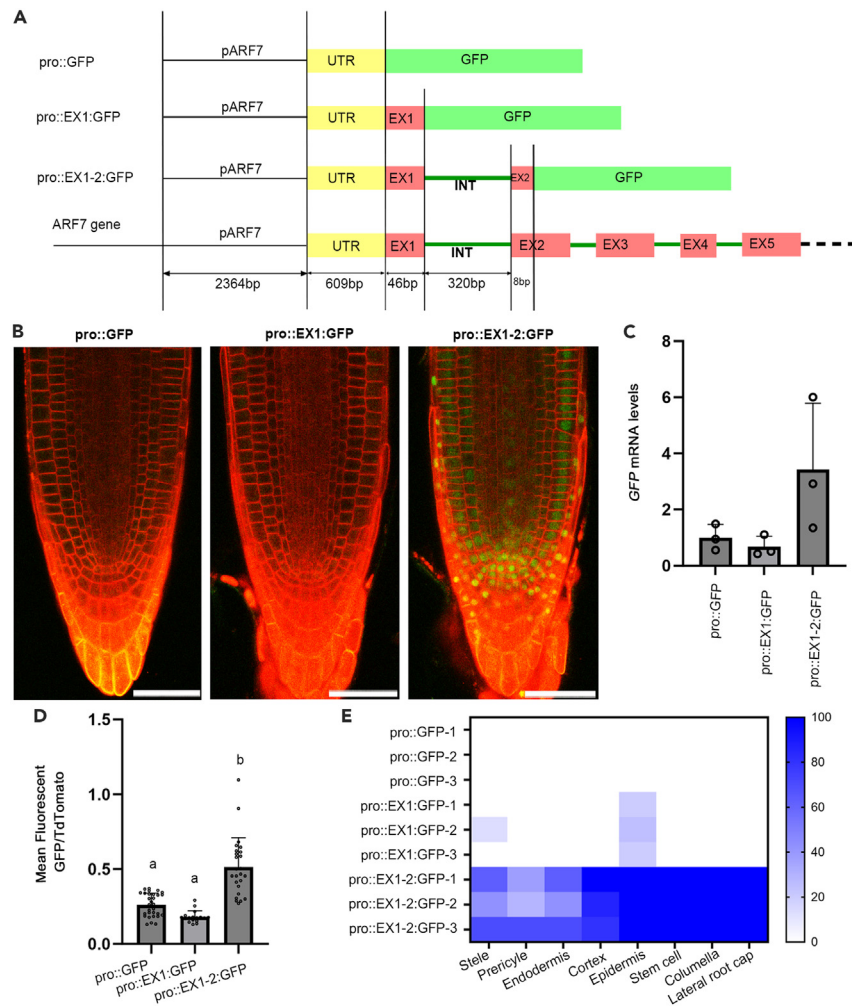
To test whether sequences in the first exon or first intron were responsible for the altered pattern, we created a new reporter line including an in-frame fusion with the first exon (pro:EX1:GFP). Similarly to reporter constructs driven by the 5' sequence alone, we saw little expression within the root meristem (Figure 1B). There was some variation in signal between individual plants. GFP signal was observed in the root tips of a few individual plants, however, in these cases the signal was very weak and restricted to the epidermis (Figure 1E). GFP mRNA levels were quantified by RT-qPCR and were lower in both constructs without the intron compared to the one containing the intron (Figure 1C). As the coding sequence included within pro:EX1:GFP and pro:EX1-2:GFP constructs differed by only 8bp, these results strongly supported the theory that it is the first intron, rather than first exon, that plays an important role in defining ARF7 expression within the RAM.

To test if the first intron affected ARF7 expression in other parts of the root, we observed different tissues with pro:EX1:GFP and pro:EX1-2:GFP. In the elongation zone and maturation zone, pro:EX1-2:GFP showed slightly higher expression in vascular tissue (Figure S1A). However, there was no discernable difference in GFP intensity during lateral root organogenesis (Figure S1B). We also tested levels within the leaves using RT-qPCR and saw only modest changes between the two lines (Figure S1C). These data suggest that the first intron has a specific function in driving expression within the RAM.

### The first intron contains potential cis-elements

Our previous analysis revealed the first intron is vital for ARF7 transcriptional regulation in the root meristem. There are several well-studied examples within the plant kingdom in which transcription factors bind to regulatory elements located within introns to control expression. For instance, the second intron of *AGAMOUS* (AG) contains enhancer elements that are sufficient to cf. AG expression<sup>36</sup> and binding sites within the second intron of *GLABRA3* (*GL3*) regulate its expression in trichomes.<sup>33</sup> To explore whether the ARF7 first intron contained potential regulatory elements, we first examined conservation of the intronic sequence within the 1001 Genomes Sequencing project. The alignment of the sequences of first intron in these lines indicated a high degree of sequence conservation within the first intron, supporting the possibility that this region could have important binding sites.

We next used the PlantRegMap (Plant Transcription Regulatory Map) which provides a high-quality resource of TFs and target genes, including a set of high-quality, non-redundant TF binding motifs derived from experiments and regulatory elements identified from high-throughput sequencing data.<sup>37</sup> When a threshold  $p$  value  $< 1e-4$  was applied, this intron has 33 putative binding sites of 28 transcription factors, including 19 NAC family genes (Figure 2; Table S1). Simultaneously, we examined DNA affinity purification sequencing (DAP-seq) data, showing in-vitro-expressed TFs interacting with genomic DNA.<sup>38</sup> These data identified four transcription factor gene families binding to the



**Figure 1. The first intron of ARF7 is required to drive strong expression within the root apical meristem**

(A) Schematic diagram showing different ARF7 promoter configurations used in this project. Pro, promoter; EX, exon; INT, intron. Please see Supplemental Data for exact plasmid sequences.

(B) Confocal images showing expression of GFP (Green) driven by the ARF7 promoter the in RAM. Cell membranes are visualized with td-Tomato using the DEAL system. Scale bars is 50 $\mu$ m.

(C) GFP expression level quantified by RT-qPCR in 2mm root tips. Error bars are standard deviation of three independent transgenic lines.

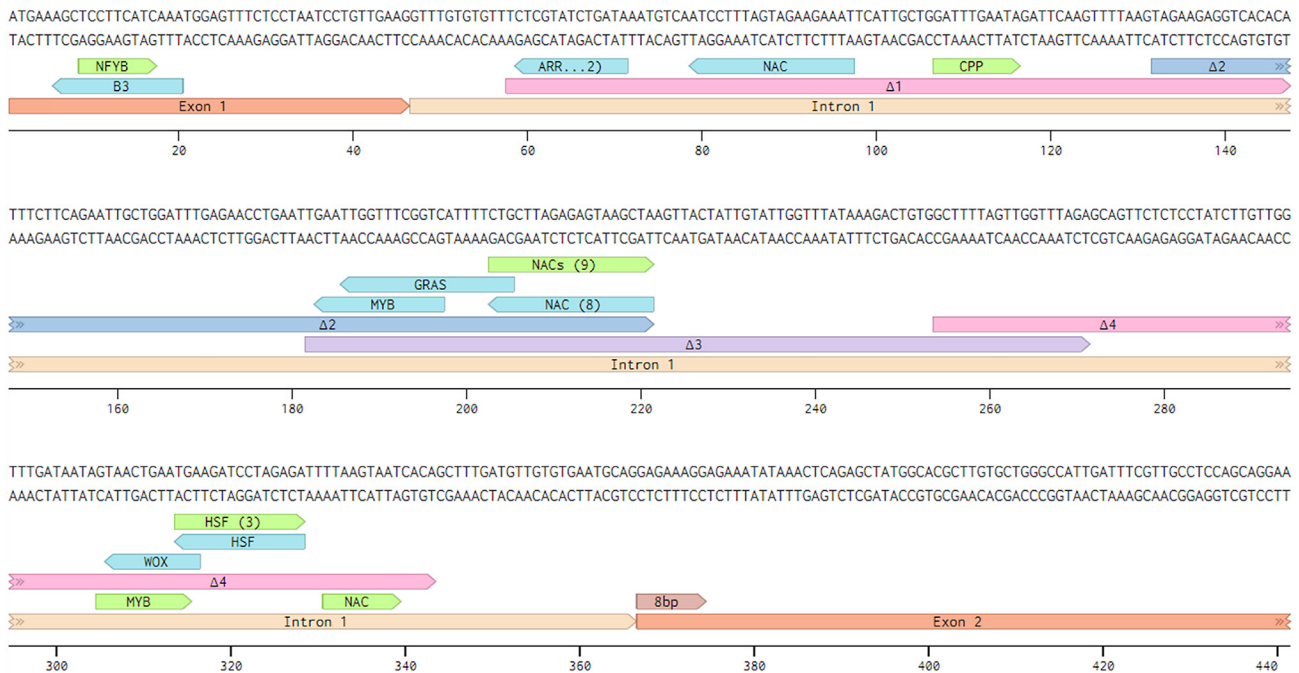
(D) ARF7 expression level quantified in the proximal meristem zone by the ratio of mean fluorescent GFP/td-Tomato. All constructs were measured based on 3 independent transgenic lines, for pro::GFP each line included 10, 8, and 12 individuals, for pro::EX1:GFP each line included 5, 5 and 7 individuals, for pro::EX1-2:GFP each line included 8, 10, and 7 individuals. Error bars are standard deviation of samples. Significant test was done by ANOVA with Tukey HSD.  $p < 0.05$

(E) Heatmap showed the percentage of roots displaying clear nuclear localized GFP in each cell-type in the RAM. The scale is %.

region of intron, these are bHLH, C2H2, NAC, and WRKY (Figure S2). All these NAC binding sites aligned with putative binding sites previously identified using the PlantRegMap analysis. To understand whether these selected binding sites could be unique and essential we studied their conservation across different *Arabidopsis thaliana* genetic variants. In this context, we first identified sequences from the first intron of ARF7 from Arabidopsis 1001 genomes. Next, we MAFFT aligned all sequences, overlaid selected binding sites and visualized using Benchling to reveal that these TF binding sites were highly conserved.

### No single transcription factor binding site within the first intron is required to drive transcription in the primary root meristem

In order to functionally dissect the ARF7 intron and identify the location of key binding sequences, the intron was divided into four overlapping parts of approximately 90 bp, with each part containing several of the previously identified putative binding sites (Figure 2; Table S2). Each section was sequentially deleted in pro::EX1-2:GFP to give four constructs, named int $\Delta$ 1-4, respectively (Figure 3A).



**Figure 2. The first intron of *ARF7* contains potential NAC and MYB transcription factor binding sites**

Genomic *ARF7* sequence surrounding the 5' end of the first exon, including the entire first intron and first 75bp of second exon. The 8bp of second exon was used in constructs. The Intron is shaded in the lighter shade. Putative transcription factor binding motifs are annotated in this sequence (The green annotations are motifs binding to forward strand and the blue annotations are motifs binding to reverse). The number in brackets shows the number of members of each transcription factor family predicted to bind to these sites. Δ1 to Δ4 shows the 4 sections of intronic DNA deleted in the various constructs (see main text). Image generated using Benchling.

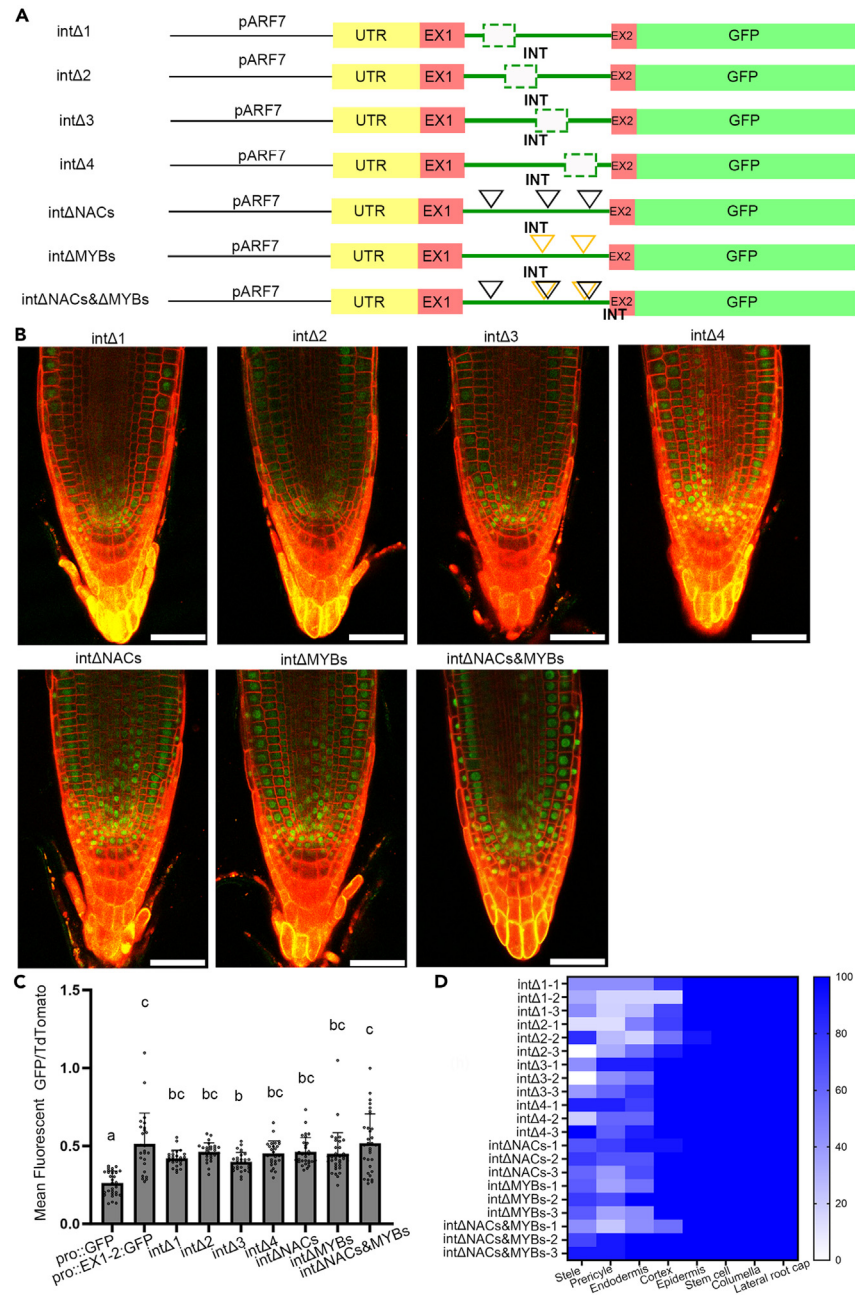
Plants transformed with these deletion constructs showed GFP signal within the RAM with strong expression around the stem cell niche (Figures 3B and 3D). Quantification of the GFP signal showed that all four deletion constructs were statistically distinct from the pro:GFP line (Figure 3B–3D).

These results showed no single binding site was sufficient to fully explain the role that the first intron has in mediating expression in the root tip, as no single deletion could remove GFP expression at the root tip. However, in our bioinformatic analysis we identified some classes of transcription factors with multiple binding sites. We focused on the transcription factors predicted to bind to multiple locations within the intron, namely the NAC and MYB (Myeloblastosis) transcription factors (Figure 2). The three NAC binding sites were in Δ1, the overlap of Δ2 and Δ3, and Δ4. The two MYB sites were located in Δ2 and Δ4 deletions.

We next investigated the importance of the three NAC binding sites by making three constructs in which one of NAC sites was mutated, termed intΔNAC1-3 (Figure S3A). However, no differences were observed in the GFP levels in these constructs (Figure S3B). We then made two constructs in which either both the MYB sites (intΔMYBs) or all three NAC sites (intΔNACs) were mutated and used this to drive GFP (Figure 3A). Again, we did not see significant changes in the GFP signal (Figures 3B and 3C). Finally, to test whether the NACs and MYB work together to promote transcription of *ARF7*, we made a new construct in which both MYB sites and all three NAC sites were mutated and used this to drive GFP. We still did not see significant changes in GFP signal (intΔNACs&MYBs) (Figures 3B and 3C).

### Intron-mediated enhancement may regulate *ARF7* expression at the root tip

Despite our extensive search for *cis*-regulatory elements within the first intron that may mediate expression of *ARF7* in roots, we were unable to identify any binding sites that were required to drive expression at the RAM. This does not conclusively prove that no such elements exist but led us to consider alternative mechanisms. There are several examples in which introns play an important role in regulating transcription.<sup>32,33</sup> In some cases, introns can increase the level of transcription, and in these examples greater quantities of mRNA are seen in constructs incorporating introns.<sup>39</sup> This effect has been described as “intron-mediated enhancement” or IME.<sup>40</sup> The role that mRNA-increasing introns play is not well understood, and the sequences associated with such introns are not well defined. However, there are a general set of rules for expression stimulating introns; these should be no more than 1kb from the transcriptional start site and must be in a transcribed region. For introns providing an IME, deletion experiments do not provide a good way to identify regulatory elements as in cases where



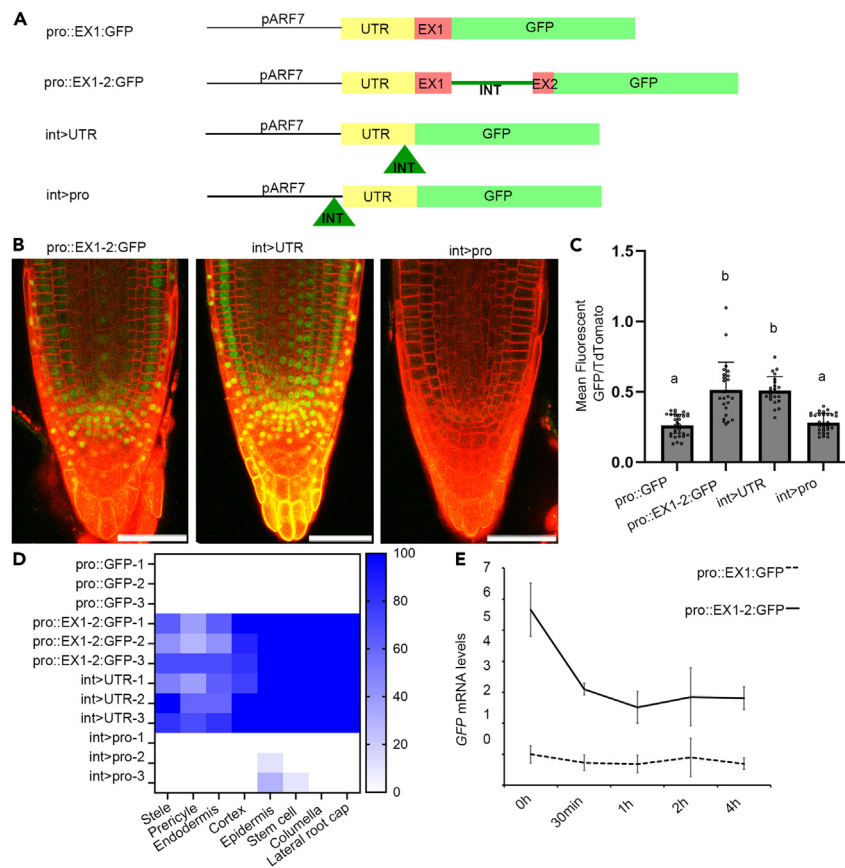
**Figure 3. Functional Dissection of first intron reveals that no single regions is required to drive expression in the root apical meristem**

(A) Schematic illustration of intron dissection constructs. The deleted sequences are shown by a green dashed box. Black triangle boxes show the modified NAC binding sites and yellow triangle boxes show the modified MYB binding sites.

(B) Confocal images showed GFP expression within the meristem of all the constructs. Scale bar is 50μm.

(C) GFP level quantified by mean fluorescent GFP/Td-Tomato. All constructs measured 3 independent transgenic lines, for intΔ1 each line including 9, 6, and 11 samples, for intΔ2 each line including 8, 11, and 9 samples, for intΔ3 each line including 9, 9, and 10 individual plants, for intΔ4 each line including 8, 11, and 8 samples, for intΔNACs each line including 12, 12, and 9 samples, for intΔMYBs each line including 11, 13, and 11 samples, for intΔNACs&MYBs each line including 9, 11, and 12 samples. Error bars are standard deviation of samples. Significant test was done by ANOVA with Tukey HSD.  $p < 0.05$  (D) Heatmap showed the ARF7 expression frequency in different cell types in the RAM.

this has been tried (e.g., UBCQ10) it is likely that active sequences are distributed through the intron.<sup>41</sup> The IMEter computational tool<sup>42</sup> allows the degree of similarity of a given intron to other promoter-proximal introns to be calculated. The rationale is that only promoter-proximal introns are likely to play a role in IME. When the first intron of ARF7 is run through this tool it comes out with a strong enrichment for



**Figure 4. The first intron only affected expression when it was present in the transcribed mRNA**

(A) Schematic illustration of different intron location, the green triangle showed first intron insertion.

(B) Confocal images showed *ARF7* expression within the meristem of all the constructs. Scale bar is 50 μm.

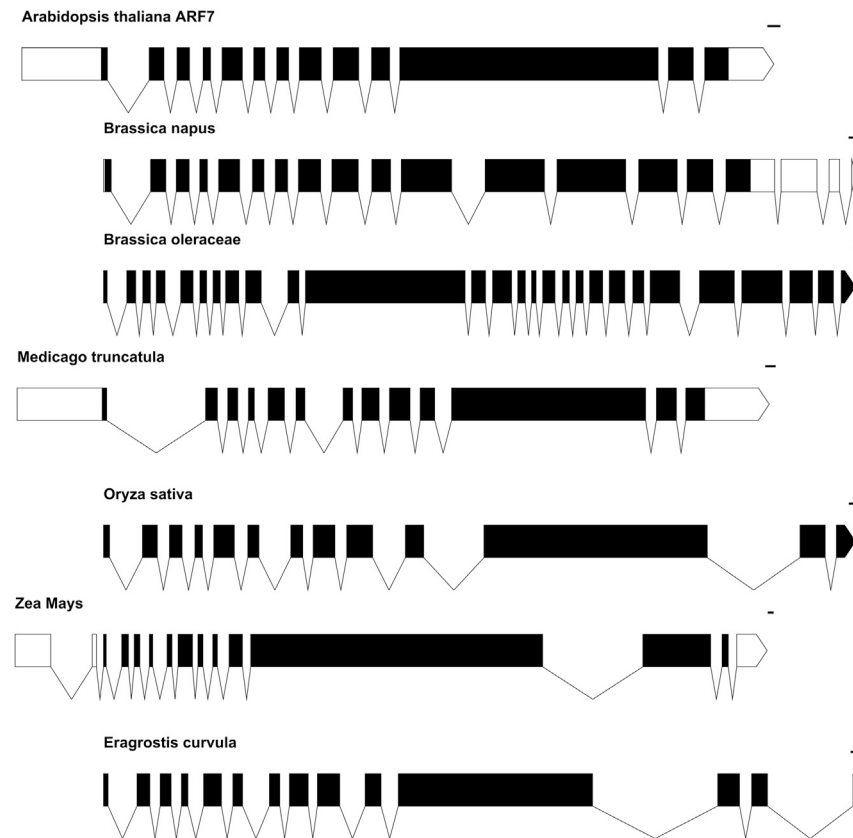
(C) *ARF7* expression level quantified by mean fluorescent GFP/TdTomato. All constructs measured 3 independent transgenic lines, for int>UTR each line including 8, 5 and 10 samples, for int>pro each line including 11, 9 and 10 samples. Error bars are standard deviation of samples. Significant test was done by ANOVA with Tukey HSD.  $p < 0.05$

(D) Heatmap showed the *ARF7* expression frequency in different cell type in RAM.

(E) *ARF7* expression level quantified by RT-qPCR in 2mm root tips after Actinomycin D treatment. Error bars are standard deviation of three independent transgenic lines.

promoter-proximal elements, ranking within the top 94th percentile of Arabidopsis introns. IMETER scores in the 90<sup>th</sup> percentile and higher are generally considered to be strong candidates for IME. For example, a study of the most active genes in soybean revealed that all the genes identified as producing the highest amount of mRNA contained an intron with an IMETER score in the 92<sup>nd</sup> percentile or higher.<sup>40,43</sup> Given the prediction with IMETER, we next considered whether IME could play a role in controlling *ARF7* expression in the primary root meristem.

A discriminating characteristic of IME is that it requires the intron to be transcribed, whereas enhancers do not need to be present in the mRNA.<sup>44</sup> To test this, we constructed two new reporters in which we moved the intronic sequences either to the 5' UTR (int>UTR) or upstream of the UTR (int>pro). Our rationale was that if the intron mediated its effect on *ARF7* transcription via an enhancer sequence then both constructs would be sufficient to drive expression in the RAM. However, if the intron mediated its effect on expression through IME, then only the construct in which the intronic sequences are located in the transcribed region should drive expression at the root meristem. We designed the first construct such that the intron was inserted –46 nt upstream from the ATG translational start site (Figure 4A). To ensure correct splicing of this intron, we also inserted exon-intron junctions on both sides. To verify that this new intron was correctly spliced, we used RT-PCR with a pair of primers located within the 5'UTR and the GFP coding sequence (Figure S4A). We compared the size of amplified band between several lines. In plants with the pro::EX1-2:GFP construct, the size of the fragment with splicing or without splicing should be 332bp or 652bp. For the int>UTR lines, the fragment spliced or without splicing should be 312bp or 632bp. A fragment of around 300bp was present in transgenic lines harboring either the pro::GFP or pro::EX1:GFP constructs, thus showing correct transcript splicing of both constructs (Figure S4B). These results were confirmed by Sanger sequencing of amplified fragments. Plants harboring the int>UTR showed expression patterns in the RAM that were indistinguishable from (pro::EX1-2:GFP) (Figure 4B). In the second construct (int>pro) the intronic sequences were located –622 nt upstream from the ATG translational start site outside of the translated region (and 13 bp from the transcriptional start site). Plants harboring the int>pro showed similar levels of GFP fluorescence in the root tip (Figure 4B) but retained expression in lateral roots (Figure S1B).



**Figure 5. Putative ARF7 orthologues in six diverse species all contain a sizable first intron 40–46bp from the translational start site, related to main text** Exon-intron structures of putative ARF7 orthologues in diverse flowering plants. Arabidopsis ARF7 was subjected to a BLAST search and a sample of species that contained one-to-one orthologues were selected. In this selection of three monocots and three dicots, the first intron in the coding sequence was larger than average and located at an identical place. Zea mays has two introns within the 5' UTR so that the conserved third intron shows corresponds with the first in other species. Schematic diagrams have been aligned based on the translational start site. The scale bars show 100 bp.

Collectively the fact that the intronic sequences needed to be present in a region of the promoter that was transcribed, supports a model in which the intron exerts its effect on ARF7 expression via IME rather than through the existence of specific enhancer elements, as the enhancers were present in both constructs. The importance of the first intron in Arabidopsis ARF7 led us to question whether this feature was conserved in other plant species. We took a sample of three monocot and three other dicot species where we found one to one ARF7 orthologues and investigated the intron-exon structure in these lines. All these species contained a sizable intron with a highly conserved junction 40–46 bp downstream of the translational start (Figures 5, S5, and S6, Tables S3 and S4).

### The first intron does not affect mRNA stability

The mechanism by which IME operates is yet to be fully determined. In the case of ARF7, our results show greater mRNA levels in lines containing the first intron. We therefore reasoned that the first intron may affect either transcriptional initiation or mRNA stability. To test the latter, we treated the pro:EX1:GFP and pro:EX1-2:GFP lines with Actinomycin D. This prevents the syntheses of new mRNA.<sup>45</sup> We then quantified the levels of GFP mRNA via RT-qPCR over a period of 4 h. As expected, the intron containing line (pro:EX1-2:GFP) had a higher initial mRNA level than the line without intron (pro:EX1:GFP) (Figure 4E). For the line without intron, the mRNA degraded slowly, however, for the line with intron, there was a rapid reduction in mRNA level within 30 min but this then became more stable over longer periods. These dynamics do not support a hypothesis in which the first intron enhances mRNA stability but rather suggests the intron increases the production rate of ARF7 mRNA in the root.

## DISCUSSION

In this study we developed a series of constructs which collectively indicated that the first intron of ARF7 is required for broad expression in RAM. It was previously shown that sequences in the 5' end of the ARF7 mRNA located between the first exon and the start of the second exon play an important role in transcription regulation in the root.<sup>20</sup> We show that the sequences exerting this effect lie within the first intron and are



unlikely linked to either the first exon or the start of the second exon. We generated multiple constructs affecting either putative or demonstrated transcription factors binding sites within this intron or harboring partial deletions. In both cases the mutated sequences were capable of driving expression at the root tip. While it is possible that we have missed non-predicted regulatory sequences acting as enhancers within the intron, the extent of our constructs supports a hypothesis in which there is no single enhancer element within the intron that solely drives expression at the root tip.

This led us to consider an alternative regulatory mechanism by which the first intron could influence the transcription of *ARF7* via intron-mediated enhancement. As the position of the intron plays a crucial role in determining transcript levels in cases of intron-mediated enhancement,<sup>40</sup> we explored whether the first intron exerted the same effect when it was present either within or upstream of the transcribed sequence. We observed a broad expression of *ARF7* in the RAM when the intron is present in the UTR but not if it is upstream of the transcriptional start. We also showed that when incorporated within the UTR the intron was correctly spliced out, thus ruling out a stabilization of the mRNA through addition of the first intron sequence to the final mRNA. Collectively our results are consistent with a mechanism in which the intron needs to be present in the mRNA to boost expression in the root. This is consistent with the idea that the first intron may exert its role on *ARF7* expression via IME rather than by the presence of specific enhancer elements. It also raises the intriguing question of whether the *ARF7* intron would exert a similar role if it was introduced to a different gene.

The effect of IME on boosting gene expression has been documented on many occasions.<sup>32,33,46</sup> There are examples of introns causing increased transcription,<sup>47</sup> enhancing mRNA stability,<sup>48</sup> and boosting translation.<sup>49</sup> To test whether *ARF7* first intron had a role in regulating mRNA stability we measured the transcript levels of GFP in two constructs (one with and one without the first intron). After an initial decrease in the level of mRNA of the intron containing line, we did not see a clear difference in mRNA degradation between the two constructs. This suggests that in this context IME does not play a role in regulating mRNA stability. Our results are consistent with a role for the *ARF7* first intron in boosting transcription levels, perhaps through transcriptional re-initiation. While there are numerous examples of IME within the plant kingdom these generally involve genes expressed at high levels with broad expression patterns, including genes such as *UBQ10*.<sup>41</sup> Our study of *ARF7* provides a rare example of what could be tissue-specific IME. The observation that *ARF7* orthologues in diverse flowering plants also share a similar intron-exon structure at the 5'-end of the gene hints at the possibility of a conserved role of introns in regulating *ARF* expression in a wider range of species. Previously, the effect that IME has on tissue-specific promoters had been investigated by including the first *UBQ10* intron in a suite of tissue specific and non-specific promoters. Incorporation of this intron causing tissues-specific promoters to be expressed more widely.<sup>50</sup> If *ARF7* expression is controlled by IME, this raises intriguing questions regarding how it shapes expression in some cell types and not others. Specifically, why the intron exerts this effect in primary root tips, but not in other tissues such as lateral root primordia and shoots.

### Limitations of the study

The role of the first intron has been investigated in the context of its regulation of GFP expression. A limitation of this study is that these promoter/intron variants have not been used to drive *ARF7* cDNA and look for complementation of *ARF7* mutants. This was our intention, but while cloning and transformation of the GFP reporters was highly efficient, we repeatedly failed to transform constructs including the *ARF7* cDNA into plants. Transforming *ARF7* is clearly possible, as 35S::*ARF7* lines have been published.<sup>26</sup> However, we tried the transformation 3 times, and failed to recover independent transgenics. This leads us to suspect that the presence of *ARF7* cDNA is either inhibitory to the insertion of the transgene into *Arabidopsis* or often lethal to plant growth.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Plant material and growth conditions
- METHOD DETAILS
  - Constructs of transcriptional reporter lines
  - Confocal analysis
  - Expression analysis by qRT-PCR
  - Methodology for motif analysis
  - Sequences and alignment
- QUANTIFICATION AND STATISTICAL ANALYSIS

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2024.109936>.

## ACKNOWLEDGMENTS

A.B. was funded by a Royal Society University Fellowship (RGF/EA/180308). U.V. received a Nottingham Research Fellowship and R.B. a Discovery Research Fellowship (BB/S011102/1) and a New Investigator grant (BB/X014843/1) from the Biotechnology and Biological Sciences Research Council.

## AUTHOR CONTRIBUTIONS

Conceptualization, J.H., T.V., and A.B.; investigation, J.H.; formal analysis, J.H., T.W., and R.B.; writing—original draft, J.H. and A.B.; writing—review and editing, T.V., U.V., and R.B.; funding acquisition, J.H. and A.B.; supervision, U.V. and A.B.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 26, 2024

Revised: April 3, 2024

Accepted: May 6, 2024

Published: May 8, 2024

## REFERENCES

- Roychoudhry, S., and Kepinski, S. (2022). Auxin in Root Development. *Cold Spring Harb. Perspect. Biol.* 14, a039933. <https://doi.org/10.1101/cshperspect.a039933>.
- Shi, B., and Vernoux, T. (2022). Hormonal control of cell identity and growth in the shoot apical meristem. *Curr. Opin. Plant Biol.* 65, 102111. <https://doi.org/10.1016/j.pbi.2021.102111>.
- Tan, X., Calderon-Villalobos, L.I.A., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M., and Zheng, N. (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* 446, 640–645. <https://doi.org/10.1038/nature05731>.
- Hayashi, K.I. (2012). The interaction and integration of auxin signaling components. *Plant Cell Physiol.* 53, 965–975. <https://doi.org/10.1093/pcp/pcs035>.
- Smalle, J., and Vierstra, R.D. (2004). The ubiquitin 26S proteasome proteolytic pathway. *Annu. Rev. Plant Biol.* 55, 555–590. <https://doi.org/10.1146/annurev.arplant.55.031903.141801>.
- Dharmasiri, N., Dharmasiri, S., and Estelle, M. (2005). The F-box protein TIR1 is an auxin receptor. *Nature* 435, 441–445. <https://doi.org/10.1038/nature03543>.
- Kepinski, S., and Leyser, O. (2005). The Arabidopsis F-box protein TIR1 is an auxin receptor. *Nature* 435, 446–451. <https://doi.org/10.1038/nature03542>.
- Tiwari, S.B., Wang, X.J., Hagen, G., and Guilfoyle, T.J. (2001). AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. *Plant Cell* 13, 2809–2822. <https://doi.org/10.1105/tpc.010289>.
- Gagne, J.M., Downes, B.P., Shiu, S.H., Durski, A.M., and Vierstra, R.D. (2002). The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in Arabidopsis. *Proc. Natl. Acad. Sci. USA* 99, 11519–11524. <https://doi.org/10.1073/pnas.162339999>.
- Dharmasiri, N., and Estelle, M. (2004). Auxin signaling and regulated protein degradation. *Trends Plant Sci.* 9, 302–308. <https://doi.org/10.1016/j.tplants.2004.04.003>.
- Qi, L., Kwiatkowski, M., Chen, H., Hoermayer, L., Sinclair, S., Zou, M., Del Genio, C.I., Kubes, M.F., Napier, R., Jaworski, K., and Friml, J. (2022). Adenylate cyclase activity of TIR1/AFB auxin receptors in plants. *Nature* 611, 133–138. <https://doi.org/10.1038/s41586-022-05369-7>.
- Cho, H., Ryu, H., Rho, S., Hill, K., Smith, S., Audenaert, D., Park, J., Han, S., Beeckman, T., Bennett, M.J., et al. (2014). A secreted peptide acts on BIN2-mediated phosphorylation of ARFs to potentiate auxin response during lateral root development. *Nat. Cell Biol.* 16, 66–76. <https://doi.org/10.1038/ncb2893>.
- Ulmasov, T., Hagen, G., and Guilfoyle, T.J. (1999). Dimerization and DNA binding of auxin response factors. *Plant J.* 19, 309–319. <https://doi.org/10.1046/j.1365-313x.1999.00538.x>.
- 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. <https://doi.org/10.1105/tpc.9.11.1963>.
- Zenser, N., Ellsmore, A., Leasure, C., and Callis, J. (2001). Auxin modulates the degradation rate of Aux/IAA proteins. *Proc. Natl. Acad. Sci. USA* 98, 11795–11800. <https://doi.org/10.1073/pnas.211312798>.
- Worley, C.K., Zenser, N., Ramos, J., Rouse, D., Leyser, O., Theologis, A., and Callis, J. (2000). Degradation of Aux/IAA proteins is essential for normal auxin signalling. *Plant J.* 21, 553–562. <https://doi.org/10.1046/j.1365-313x.2000.00703.x>.
- Okushima, Y., Overvoorde, P.J., Arima, K., Alonso, J.M., Chan, A., Chang, C., Ecker, J.R., Hughes, B., Lui, A., Nguyen, D., 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, 444–463. <https://doi.org/10.1105/tpc.104.028316>.
- Finet, C., Berne-Dedieu, A., Scutt, C.P., and Marlétaz, F. (2013). Evolution of the ARF gene family in land plants: old domains, new tricks. *Mol. Biol. Evol.* 30, 45–56. <https://doi.org/10.1093/molbev/mss220>.
- Rademacher, E.H., Möller, B., Lokerse, A.S., Llavata-Peris, C.I., van den Berg, W., and Weijers, D. (2011). A cellular expression map of the Arabidopsis AUXIN RESPONSE FACTOR gene family. *Plant J.* 68, 597–606. <https://doi.org/10.1111/j.1365-313X.2011.04710.x>.
- Truskina, J., Han, J., Chrysanthou, E., Galvan-Ampudia, C.S., Lainé, S., Brunoud, G., Macé, J., Bellows, S., Legrand, J., Bågman, A.M., et al. (2021). Publisher Correction: A network of transcriptional repressors modulates auxin responses. *Nature* 589, E7. <https://doi.org/10.1038/s41586-020-03066-x>.
- Vernoux, T., Brunoud, G., Farcot, E., Morin, V., Van den Daele, H., Legrand, J., Oliva, M., Das, P., Larrieu, A., Wells, D., et al. (2011). The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Mol. Syst. Biol.* 7, 508. <https://doi.org/10.1038/msb.2011.39>.
- Krogan, N.T., Marcos, D., Weiner, A.I., and Berleth, T. (2016). The auxin response factor MONOPTEROS controls meristem function and organogenesis in both the shoot and root through the direct regulation of PIN genes. *New Phytol.* 212, 42–50. <https://doi.org/10.1111/nph.14107>.
- Nagpal, P., Ellis, C.M., Weber, H., Ploense, S.E., Barkawi, L.S., Guilfoyle, T.J., Hagen, G., Alonso, J.M., Cohen, J.D., Farmer, E.E., et al. (2005). Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* 132, 4107–4118. <https://doi.org/10.1242/dev.01955>.
- Hardtke, C.S., Ckurshumova, W., Vidaurre, D.P., Singh, S.A., Stamatidou, G., Tiwari, S.B., Hagen, G., Guilfoyle, T.J., and Berleth, T. (2004). Overlapping and non-redundant functions of the Arabidopsis auxin response factors MONOPTEROS and NONPHOTOTROPIC HYPOCOTYL 4. *Development* 131, 1089–1100. <https://doi.org/10.1242/dev.00925>.
- Harper, R.M., Stowe-Evans, E.L., Luesse, D.R., Muto, H., Tatematsu, K., Watahiki, M.K., Yamamoto, K., and Liscum, E. (2000). The NPH4 locus encodes the auxin response factor ARF7, a conditional regulator of differential growth in aerial Arabidopsis tissue. *Plant Cell* 12, 757–770. <https://doi.org/10.1105/tpc.12.5.757>.

26. Okushima, Y., Fukaki, H., Onoda, M., Theologis, A., and Tasaka, M. (2007). ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in Arabidopsis. *Plant Cell* 19, 118–130. <https://doi.org/10.1105/tpc.106.047761>.
27. Schoenaers, S., Balcerowicz, D., Breen, G., Hill, K., Zdanio, M., Mouille, G., Holman, T.J., Oh, J., Wilson, M.H., Nikonorova, N., et al. (2018). The Auxin-Regulated CrRLK1L Kinase ERULUS Controls Cell Wall Composition during Root Hair Tip Growth. *Curr. Biol.* 28, 722–732.e6. <https://doi.org/10.1016/j.cub.2018.01.050>.
28. Lee, H.W., Cho, C., Pandey, S.K., Park, Y., Kim, M.J., and Kim, J. (2019). LBD16 and LBD18 acting downstream of ARF7 and ARF19 are involved in adventitious root formation in Arabidopsis. *BMC Plant Biol.* 19, 46. <https://doi.org/10.1186/s12870-019-1659-4>.
29. Weijers, D., Benkova, E., Jäger, K.E., Schlereth, A., Hamann, T., Kientz, M., Wilmoth, J.C., Reed, J.W., and Jürgens, G. (2005). Developmental specificity of auxin response by pairs of ARF and Aux/IAA transcriptional regulators. *EMBO J.* 24, 1874–1885. <https://doi.org/10.1038/sj.emboj.7600659>.
30. Schuetz, M., Fidanza, M., and Mattsson, J. (2019). Identification of Auxin Response Factor-Encoding Genes Expressed in Distinct Phases of Leaf Vein Development and with Overlapping Functions in Leaf Formation. *Plants* 8, 242. <https://doi.org/10.3390/plants8070242>.
31. Wilmoth, J.C., Wang, S., Tiwari, S.B., Joshi, A.D., Hagen, G., Guilfoyle, T.J., Alonso, J.M., Ecker, J.R., and Reed, J.W. (2005). NPH4/ARF7 and ARF19 promote leaf expansion and auxin-induced lateral root formation. *Plant J.* 43, 118–130. <https://doi.org/10.1111/j.1365-3113.2005.02432.x>.
32. Bradnam, K.R., and Korf, I. (2008). Longer first introns are a general property of eukaryotic gene structure. *PLoS One* 3, e3093. <https://doi.org/10.1371/journal.pone.0003093>.
33. Friede, A., Zhang, B., Herberth, S., Pesch, M., Schrader, A., and Hülskamp, M. (2017). The Second Intron Is Essential for the Transcriptional Control of the Arabidopsis thaliana GLABRA3 Gene in Leaves. *Front. Plant Sci.* 8, 1382. <https://doi.org/10.3389/fpls.2017.01382>.
34. Lampropoulos, A., Sutikovic, Z., Wenzl, C., Maegele, I., Lohmann, J.U., and Forner, J. (2013). GreenGate—a novel, versatile, and efficient cloning system for plant transgenesis. *PLoS One* 8, e83043. <https://doi.org/10.1371/journal.pone.0083043>.
35. Kumpers, B.M.C., Han, J., Vaughan-Hirsch, J., Redman, N., Ware, A., Atkinson, J.A., Leftley, N., Janes, G., Castiglione, G., Tarr, P.T., et al. (2022). Dual expression and anatomy lines allow simultaneous visualization of gene expression and anatomy. *Plant Physiol.* 188, 56–69. <https://doi.org/10.1093/plphys/kiab503>.
36. Deyholos, M.K., and Sieburth, L.E. (2000). Separable whorl-specific expression and negative regulation by enhancer elements within the AGAMOUS second intron. *Plant Cell* 12, 1799–1810. <https://doi.org/10.1105/tpc.12.10.1799>.
37. Jin, J., Tian, F., Yang, D.C., Meng, Y.Q., Kong, L., Luo, J., and Gao, G. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. <https://doi.org/10.1093/nar/gkw982>.
38. O'Malley, R.C., Huang, S.S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cistrome and Epistrome Features Shape the Regulatory DNA Landscape. *Cell* 166, 1598. <https://doi.org/10.1016/j.cell.2016.08.063>.
39. Callis, J., Fromm, M., and Walbot, V. (1987). Introns increase gene expression in cultured maize cells. *Genes Dev.* 1, 1183–1200. <https://doi.org/10.1101/gad.1.10.1183>.
40. Rose, A.B. (2018). Introns as Gene Regulators: A Brick on the Accelerator. *Front. Genet.* 9, 672. <https://doi.org/10.3389/fgenet.2018.00672>.
41. Rose, A.B., Elfers, T., Parra, G., and Korf, I. (2008). Promoter-proximal introns in Arabidopsis thaliana are enriched in dispersed signals that elevate gene expression. *Plant Cell* 20, 543–551. <https://doi.org/10.1105/tpc.107.057190>.
42. Parra, G., Bradnam, K., Rose, A.B., and Korf, I. (2011). Comparative and functional analysis of intron-mediated enhancement signals reveals conserved features among plants. *Nucleic Acids Res.* 39, 5328–5337. <https://doi.org/10.1093/nar/gkr043>.
43. Zhang, N., McHale, L.K., and Finer, J.J. (2015). Isolation and characterization of "GmScream" promoters that regulate highly expressing soybean (Glycine max Merr.) genes. *Plant Sci.* 241, 189–198. <https://doi.org/10.1016/j.plantsci.2015.10.010>.
44. Zabidi, M.A., and Stark, A. (2016). Regulatory Enhancer-Core-Promoter Communication via Transcription Factors and Cofactors. *Trends Genet.* 32, 801–814. <https://doi.org/10.1016/j.tig.2016.10.003>.
45. Gil, P., and Green, P.J. (1996). Multiple regions of the Arabidopsis SAUR-AC1 gene control transcript abundance: the 3' untranslated region functions as an mRNA instability determinant. *EMBO J.* 15, 1678–1686.
46. Gallegos, J.E., and Rose, A.B. (2019). An intron-derived motif strongly increases gene expression from transcribed sequences through a splicing independent mechanism in Arabidopsis thaliana. *Sci. Rep.* 9, 13777. <https://doi.org/10.1038/s41598-019-50389-5>.
47. Samadder, P., Sivamani, E., Lu, J., Li, X., and Qu, R. (2008). Transcriptional and post-transcriptional enhancement of gene expression by the 5' UTR intron of rice rubi3 gene in transgenic rice cells. *Mol. Genet. Genom.* 279, 429–439. <https://doi.org/10.1007/s00438-008-0323-8>.
48. Narsai, R., Howell, K.A., Millar, A.H., O'Toole, N., Small, I., and Whelan, J. (2007). Genome-wide analysis of mRNA decay rates and their determinants in Arabidopsis thaliana. *Plant Cell* 19, 3418–3436. <https://doi.org/10.1105/tpc.107.055046>.
49. Akua, T., and Shaul, O. (2013). The Arabidopsis thaliana MXH gene includes an intronic element that boosts translation when localized in a 5' UTR intron. *J. Exp. Bot.* 64, 4255–4270. <https://doi.org/10.1093/jxb/ert235>.
50. Emami, S., Arumainayagam, D., Korf, I., and Rose, A.B. (2013). The effects of a stimulating intron on the expression of heterologous genes in Arabidopsis thaliana. *Plant Biotechnol. J.* 11, 555–563. <https://doi.org/10.1111/pbi.12043>.
51. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>.
52. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. <https://doi.org/10.1038/nmeth.1318>.
53. Stormo, G.D. (2013). Modeling the specificity of protein-DNA interactions. *Quant. Biol.* 1, 115–130. <https://doi.org/10.1007/s40484-013-0012-4>.
54. Khan, A., Fornes, O., Stigliani, A., Gheorghie, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Chêneby, J., Kulkarni, S.R., Tan, G., et al. (2018). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* 46, D260–D266. <https://doi.org/10.1093/nar/gkx1126>.
55. 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. <https://doi.org/10.1046/j.1365-3113.1998.00343.x>.
56. Harrison, S.J., Mott, E.K., Parsley, K., Aspinall, S., Gray, J.C., and Cottage, A. (2006). A rapid and robust method of identifying transformed Arabidopsis thaliana seedlings following floral dip transformation. *Plant Methods* 2, 19. <https://doi.org/10.1186/1746-4811-2-19>.
57. Jin, J., He, K., Tang, X., Li, Z., Lv, L., Zhao, Y., Luo, J., and Gao, G. (2015). An Arabidopsis Transcriptional Regulatory Map Reveals Distinct Functional and Evolutionary Features of Novel Transcription Factors. *Mol. Biol. Evol.* 32, 1767–1773. <https://doi.org/10.1093/molbev/msv058>.
58. Tian, F., Yang, D.C., Meng, Y.Q., Jin, J., and Gao, G. (2020). PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res.* 48, D1104–D1113. <https://doi.org/10.1093/nar/gkz1020>.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE  | SOURCE                    | IDENTIFIER  |
|--|---------------------------|---|
| <b>Bacterial and virus strains</b>                             |                           |   |
| <i>E. coli</i> DH5 $\alpha$                                    | In house stock            | N/A   |
| Agrobacterium GV3101   | In house stock            | N/A   |
| <b>Chemicals, peptides, and recombinant proteins</b>           |                           |   |
| Murashige & Skoog Medium, basal salt mixture                   | Sigma-Aldrich             | M5524   |
| Agar   | Sigma-Aldrich             | A1296   |
| Hygromycin B (50 mg/mL)  | Thermo Fisher             | Cat#10687010  |
| Actinomycin D  | Sigma-Aldrich             | <a href="#">A1410</a>   |
| Silwet L-77  | De Sangosse               | N/A   |
| <b>Critical commercial assays</b>                              |                           |   |
| Gibson Assembly® Master Mix                                    | NEW ENGLAND BioLabs       | E2611L  |
| Golden Gate Assembly Kit (Bsal-HF® v2)                         | NEW ENGLAND BioLabs       | E1601L  |
| Q5® Site-Directed Mutagenesis Kit                              | NEW ENGLAND BioLabs       | E0552S  |
| GreenGate Cloning System                                       | Addgene                   | Kit #1000000036   |
| RNeasy Plant Mini Kit  | QIAGEN                    | Cat# 74904  |
| RevertAid First Strand cDNA Synthesis Kit                      | Thermo Fisher             | K1621   |
| SensiMix™ SYBR® Hi-ROX Kit                                     | Bioline                   | QT605-05  |
| <b>Deposited data</b>  |                           |   |
| Microscopy data  | This paper; Mendeley Data | <a href="https://doi.org/10.17632/4t9twsz9bg.2">https://doi.org/10.17632/4t9twsz9bg.2</a> |
| <b>Experimental models: Organisms/strains</b>                  |                           |   |
| <i>Arabidopsis Thaliana</i> Columbia-0                         | In house stock            | N/A   |
| <i>Arabidopsis Thaliana</i> : pro:GFP                          | This study                | NASC ID: N2112003   |
| <i>Arabidopsis Thaliana</i> : pro:EX1:GFP                      | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : pro:EX1-2:GFP                    | This study                | NASC ID: N2112004   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ 1                   | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ 2                   | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ 3                   | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ 4                   | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ NACs                | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ MYBs                | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ NACs& $\Delta$ MYBs | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ NAC1                | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ NAC2                | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int $\Delta$ NAC3                | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int>pro                          | This study                | N/A   |
| <i>Arabidopsis Thaliana</i> : int>UTR                          | This study                | NASC ID: N2112005   |
| <b>Oligonucleotides</b>  |                           |   |
| Primers (Table S5)   | This study                | N/A   |
| <b>Recombinant DNA</b>   |                           |   |
| pGGA-ARF7proS  | This study                | Addgene 218563  |
| pGGA-ARF7UTRINT  | This study                | Addgene 218564  |

(Continued on next page)

**Continued**

| REAGENT or RESOURCE                             | SOURCE     | IDENTIFIER     |
|---|------------|----------------|
| pGGA-ARF7INTUTR                                 | This study | Addgene 218565 |
| pGGB-EX1  | This study | Addgene 218566 |
| pGGB-EX1-2                                      | This study | Addgene 218567 |
| pGGB-EX1-2 (int $\Delta$ 1)                     | This study | Addgene 218568 |
| pGGB-EX1-2 (int $\Delta$ 2)                     | This study | Addgene 218569 |
| pGGB-EX1-2 (int $\Delta$ 3)                     | This study | Addgene 218570 |
| pGGB-EX1-2 (int $\Delta$ 4)                     | This study | Addgene 218571 |
| pGGB-EX1-2 (int $\Delta$ NACs)                  | This study | Addgene 218572 |
| pGGB-EX1-2 (int $\Delta$ MYBs)                  | This study | Addgene 218573 |
| pGGB-EX1-2 (int $\Delta$ NACs& $\Delta$ MYBs)   | This study | Addgene 218574 |
| pGGB-EX1-2 (int $\Delta$ NAC1)                  | This study | Addgene 218575 |
| pGGB-EX1-2 (int $\Delta$ NAC2)                  | This study | Addgene 218576 |
| pGGB-EX1-2 (int $\Delta$ NAC3)                  | This study | Addgene 218577 |
| pro:GFP   | This study | Addgene 218563 |
| pro:EX1:GFP                                     | This study | Addgene 218564 |
| pro:EX1-2:GFP                                   | This study | Addgene 218565 |
| pro:EX1-2(int $\Delta$ 1):GFP                   | This study | Addgene 218566 |
| pro:EX1-2(int $\Delta$ 2):GFP                   | This study | Addgene 218567 |
| pro:EX1-2(int $\Delta$ 3):GFP                   | This study | Addgene 218568 |
| pro:EX1-2(int $\Delta$ 4):GFP                   | This study | Addgene 218569 |
| pro:EX1-2(int $\Delta$ NACs):GFP                | This study | Addgene 218570 |
| pro:EX1-2(int $\Delta$ MYBs):GFP                | This study | Addgene 218571 |
| pro:EX1-2(int $\Delta$ NACs& $\Delta$ MYBs):GFP | This study | Addgene 218572 |
| pro:EX1-2(int $\Delta$ NAC1):GFP                | This study | Addgene 218573 |
| pro:EX1-2(int $\Delta$ NAC2):GFP                | This study | Addgene 218574 |
| pro:EX1-2(int $\Delta$ NAC3):GFP                | This study | Addgene 218575 |
| int>pro:GFP                                     | This study | Addgene 218576 |
| int>UTR::GFP                                    | This study | Addgene 218577 |

**Software and algorithms**

|                        |                                |   |
|------------------------|--------------------------------|---|
| ImageJ                 | Schneider et al. <sup>51</sup> | <a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a>   |
| Leica LAS X            | Leica                          | <a href="https://www.leica-microsystems.com/products/microscope-software/p/leica-las-x-ls">https://www.leica-microsystems.com/products/microscope-software/p/leica-las-x-ls</a> |
| Benchling              | Benchling                      | <a href="https://www.benchling.com">https://www.benchling.com</a>   |
| GraphPad Prism 10      | Graphpad                       | <a href="https://www.graphpad.com/features">https://www.graphpad.com/features</a>   |
| IBM SPSS Statistics 29 | IBM                            | <a href="https://www.ibm.com/products/spss-statistics">https://www.ibm.com/products/spss-statistics</a>   |
| MEGA-X                 | megasoftware                   | <a href="https://megasoftware.net/">https://megasoftware.net/</a>   |
| Jalview                | Jalview                        | <a href="https://www.jalview.org/">https://www.jalview.org/</a>   |

**RESOURCE AVAILABILITY**

**Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Anthony Bishopp ([Anthony.bishopp@nottingham.ac.uk](mailto:Anthony.bishopp@nottingham.ac.uk)).

**Materials availability**

- Destination plasmids generated in this study have been deposited to Addgene. All intermediate modules are available on request from the [lead contact](#).

- The most useful Arabidopsis lines generated in this study have been deposited to NASC. All other lines are available on request from the [lead contact](#).

#### Data and code availability

- The microscopy data used for quantification is deposited on Medeley Data <https://data.mendeley.com/datasets/4t9twsz9bg/1>. All other datasets reported in this work are available from the [lead contact](#) on request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### Plant material and growth conditions

All lines generated in this study were in the Arabidopsis thaliana Col-0. The seedlings were grown at 21°C. The growth conditions for individual experiments are listed in the methods.

## METHOD DETAILS

### Constructs of transcriptional reporter lines

All constructs were produced using the DEAL vector system.<sup>35</sup> This is a modified version of the GreenGate cloning system<sup>34</sup> to include a red plasma membrane marker to highlight anatomy and serve as an internal control to ensure that all transgenes were expressed. All modules and primers are shown in [Tables S5](#) and [S6](#).

For reporter lines, the promoter and intron were cloned respectively into A and B entry modules. In line int>UTR and int>pro lines, the promoter A modules were edited to insert the intronic sequences within the UTR or upstream of the UTR using the Gibson cloning system.<sup>52</sup> The Site-Directed Mutagenesis Kit (NEW ENGLAND BioLabs) was used to introduce polymorphisms affecting specific binding sites within intron constructs by editing the EX1-2 B module. The DNA sequences knocking out multiple MYB and NAC binding sites were synthesized by Eurofins. NAC and MYB mutations were mutated by exchanging 2bp within the core binding motif, which is shown in [Table S2](#). 2bp in the core TF binding sequence were altered based on the TFs position weight matrices and TFs flexible models provided by JASPAR.<sup>53,54</sup> Combining the position information from JASPAR and the sequences from PlantRegMap, we selected motifs to target to modulate the interaction of NAC and MYB proteins with the intronic sequence. Entry modules recombined in order, with GFP:NLS (C module) and D-dummy, RBCS terminator and HygromycinR (for plant selection) into the DEAL vector.

All destination plasmids were assembled by Golden Gate Assembly Kit (NEW ENGLAND BioLabs). These constructs were transformed in *Agrobacterium tumefaciens* GV3101 via electroporation and transformed into Col-0 via the floral dip method,<sup>55</sup> which is dipping Arabidopsis inflorescences for a 1min into a 5% sucrose solution containing 0.05% Silwet L-77 and resuspended *Agrobacterium*. Transgenic T0 seeds were plated on 1% agar containing 1/2 MS medium and 20 µg/L Hygromycin B. Only T2 lines with single locus insertions were analyzed, as judged by a 3:1 resistance to hygromycin. After a 2days stratification period, seeds were grown in 21°C with 6 h light, 48 h dark and 24 h light.<sup>56</sup>

### Confocal analysis

For the primary root microscopy, 3 independent T2 lines were grown on 1/2 MS medium supplemented with 1% agar in 16h light/8h dark conditions for 5 days (for primary roots) or 8 days (for lateral roots). The roots were examined in the TCS-SP8 confocal microscope (Leica) with excitation at 488 nm and emission at 496–550 nm for GFP, excitation at 560nm and emission at 590–700 nm for tdTomato. Fluorescent signal was quantified in ImageJ for over 5 lines (including both hetero and homozygous individuals, the numbers of lines were indicated in figure legends) by selecting a rectangle (130 × 190 micrometers) with the longest side parallel to the root axis. Fluorescence was measured as mean fluorescence intensity of GFP and calculated as a ration of GFP to TdTomato.

### Expression analysis by qRT-PCR

For reporter lines analysis, 3 independent T2 plants were grown on 1/2 MS medium supplemented with 1% agar in 16h light/8h dark conditions for 5 days (for root meristem) or 10 days (for leaf). The root meristem sample of each T2 line was obtained by dissecting and pooling 2mm root tips from 50 individual plants. The leaf sample of each line contained 3 plant leaves. Each line including 3 technical replicates. For mRNA stability analysis, 5 days roots were treated with 100µg/ml Actinomycin D in 1/2 MS medium, for 30min, 1h, 2h and 4h. RNA was collected from T3 homozygous lines. Each line has 3 biological repeats and each sample including 50 2mm root tips. Each line including 3 technical replicates. The Qiagen RNeasy Plant Mini Kit was used for RNA extraction. The cDNA was synthesized using RevertAid First Strand cDNA Synthesis kit (Thermo Scientific) with 200ng RNA and diluted 3:20 for qRT-PCR reaction. The SensiMix SYBR Hi-Rox Mastermix (BIOLINE) was used for qRT-PCR reaction. The control gene was *UBC21* (At5g25760). Primers were showed in [Table S5](#).

### Methodology for motif analysis

We first used AT5G20730.1 as a representative model to obtain the 1st intron sequence (Chromosome 5: 7021458.7021139; minus strand; reverse complemented) of ARF7. Next, we scanned this input sequence using PlantRegMap tool<sup>37,57,58</sup> for the presence of any manually curated 674 non-redundant and high-quality binding motifs derived from various experiments, literature and ChIP-seq datasets. We specifically used the Binding Site Enrichment tool (with settings: species = *Arabidopsis thaliana* and Threshold  $p$ -value  $\leq 1e-4$ ) that uses modified FIMO (Find Individual Motif Occurrences) to search TF binding sites in the input sequence.

To cross-validate the above predictions, we used recent DNA affinity purification sequencing data (DAP-seq) to scan input region (Chromosome 5: 7021458.7021139) for peaks (with FRiP  $\geq 5\%$  i.e., fraction of reads in peaks) of 46 different transcription factors. The overlap between predicted and Dap-Seq derived TF binding sites (termed as selected binding sites) were considered for further analysis.

### Sequences and alignment

To illustrate conservation of regions in the intron and first two exon sequences of ARF7 across different plant species, known orthologues were selected from the orthologue table of AT5G20730.3 in EnsemblPlants. These selections were made to include primarily agriculturally important species, and representation of both monocots and dicots. Multiple sequence alignment of these sequences were performed using MUSCLE in MEGA-X (version 10.2.1), with alignment figure constructed using Jalview software (version 2.11.3.2).

### QUANTIFICATION AND STATISTICAL ANALYSIS

T test statistical analyses were performed in Excel. ANOVA statistical analyses were performed in SPSS 29. The precise statistical tests can be found in the corresponding Figure legend.