Transcriptional and posttranscriptional regulation of transcription factor expression in *Arabidopsis* roots

Ji-Young Lee*, Juliette Colinas*, Jean Y. Wang*, Daniel Mace†, Uwe Ohler†, and Philip N. Benfey*†‡

*Department of Biology and †Institute for Genome Sciences and Policy, Duke University, Durham, NC 27708

Edited by Susan R. Wessler, University of Georgia, Athens, GA, and approved February 17, 2006 (received for review December 9, 2005)

Understanding how the expression of transcription factor (TF) genes is modulated is essential for reconstructing gene regulatory networks. There is increasing evidence that sequences other than upstream noncoding can contribute to modulating gene expression, but how frequently they do so remains unclear. Here, we investigated the regulation of TFs expressed in a tissue-enriched manner in *Arabidopsis* roots. For 61 TFs, we created GFP reporter constructs driven by each TF's upstream noncoding sequence (including the 5′UTR) fused to the GFP reporter gene alone or together with the TF's coding sequence. We compared the visually detectable GFP patterns with endogenous mRNA expression patterns, as defined by a genome-wide microarray root expression map. An automated image analysis method for quantifying GFP signals in different tissues was developed and used to validate our visual comparison method. From these combined analyses, we found that (*i*) the upstream noncoding sequence was sufficient to recapitulate the mRNA expression pattern for 80% (35/44) of the TFs, and (*ii*) 25% of the TFs undergo posttranscriptional regulation via microRNA-mediated mRNA degradation (2/24) or via intercellular protein movement (6/24). The results suggest that, for *Arabidopsis* TFs, upstream noncoding sequences are major contributors to mRNA expression pattern establishment, but modulation of transcription factor protein expression pattern after transcription is relatively frequent. This study provides a systematic overview of regulation of TF expression at a cellular level.

intercellular protein movement | intergenic | microRNA-mediated mRNA degradation | transcriptional regulation

Transcriptional networks mediated by transcription factors (TFs) control the developmental processes of multicellular organisms by specifying when and where genes are expressed (1). Therefore, understanding the regulation of TF expression is essential for building networks. Currently, transcriptional regulation is thought to occur primarily through the binding of transcription factors to cis-regulatory modules within upstream noncoding sequences (or "promoters"). However, cases have been reported in plants (2, 3) and animals (4) where regulatory elements within the transcribed region are necessary for correct mRNA pattern establishment. Furthermore, various posttranscriptional regulatory mechanisms have recently been described, including microRNA-mediated degradation (5, 6), nonsense-mediated mRNA decay (7), and nuclear export control (8). After the establishment of the correct mRNA expression pattern, mechanisms that can alter the expression pattern include control of translation efficiency (9), intercellular protein trafficking (10, 11), and regulated nuclear localization (12, 13). Clearly there is the potential for altering expression patterns mediated by cis-elements within transcribed regions or through posttranscriptional regulation. How often these mechanisms are involved in gene regulation remains unclear.

To address this question, we carried out a systematic and nonbiased investigation of the contribution of upstream noncoding sequences to the establishment of mRNA expression patterns and the influence of coding sequences on the pattern of mRNA and protein expression. We focused on tissue-enriched TFs because these TFs are potential regulators of tissue specification and differentiation. We selected about one-third (61) of all of the TFs expressed in a tissue-enriched manner in any one of five tissues of the *Arabidopsis* root. They were identified from microarray data on sorted GFP-marked cell populations (14, 15). These mRNA expression patterns were compared with the expression patterns conferred by the upstream noncoding region of each TF fused either to GFP alone (transcriptional fusion) or fused to the TF coding sequence (translational fusion). Visual comparison of the radial mRNA expression patterns with GFP expression patterns indicated the following: (*i*) In 80% of the cases, the upstream noncoding region was sufficient to recapitulate the mRNA expression pattern, and (*ii*) adding the coding sequence to the reporter construct affected the expression pattern in 25% of the cases. An automated image analysis program was developed to analyze GFP expression patterns and its results corroborated our visual microarray/GFP comparison method.

Results and Discussion

Transcriptional and Translational GFP Fusions for 61 Tissue-Enriched Root Transcription Factors. TFs expressed in a tissue-enriched manner were identified from the published root mRNA expression map (14–16). The map was generated by using microarray profiling of GFP marked and sorted *Arabidopsis* root tissues (here "tissue" is used interchangeably with "cell-type"). It covered the quiescent center, stele, endodermis, epidermis (atrichoblast), and lateral root cap. Of the 2,033 genes on the ATH1 array currently annotated as TFs, 189 are enriched >2-fold in one of these five tissues. We essentially randomly selected about one-third (61) of these TFs (see *Materials and Methods*), generating transcriptional fusions to GFP for 61 and translational fusions for 59. The constructs were made with the versatile MultiSite Gateway cloning system (details in *Material and Methods*). The transcriptional fusions were constructed by using endoplasmic reticulum-localized GFP (17) driven by 3 kb of upstream noncoding sequence or the entire upstream noncoding region, whichever was shorter. We included the 5′UTR in our constructs because annotation of transcription start sites was not always available. The translational fusions were made by inserting the coding region of each TF in-frame upstream of nontargeted GFP (17) with the same upstream sequences as those used in the transcriptional fusions. At least 15 TF proteins from various families have been shown to remain functional when fused in this fashion to GFP at their C terminus (list of references available in Table 2, which is published as supporting information on the PNAS web site). These transcriptional and translational constructs were transformed into plants, and the GFP expression pattern generated in the root was imaged by using confocal laser-scanning microscopy. The genes analyzed, the length of the upstream sequence used, and the number of lines analyzed for each construct can be found in Table 3, which is published as supporting information on the PNAS web site. Confocal images of roots for each construct are available in the Arex database (www.arexdb.org)

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: TF, transcription factor.

‡To whom correspondence should be addressed. E-mail: philip.benfey@duke.edu.

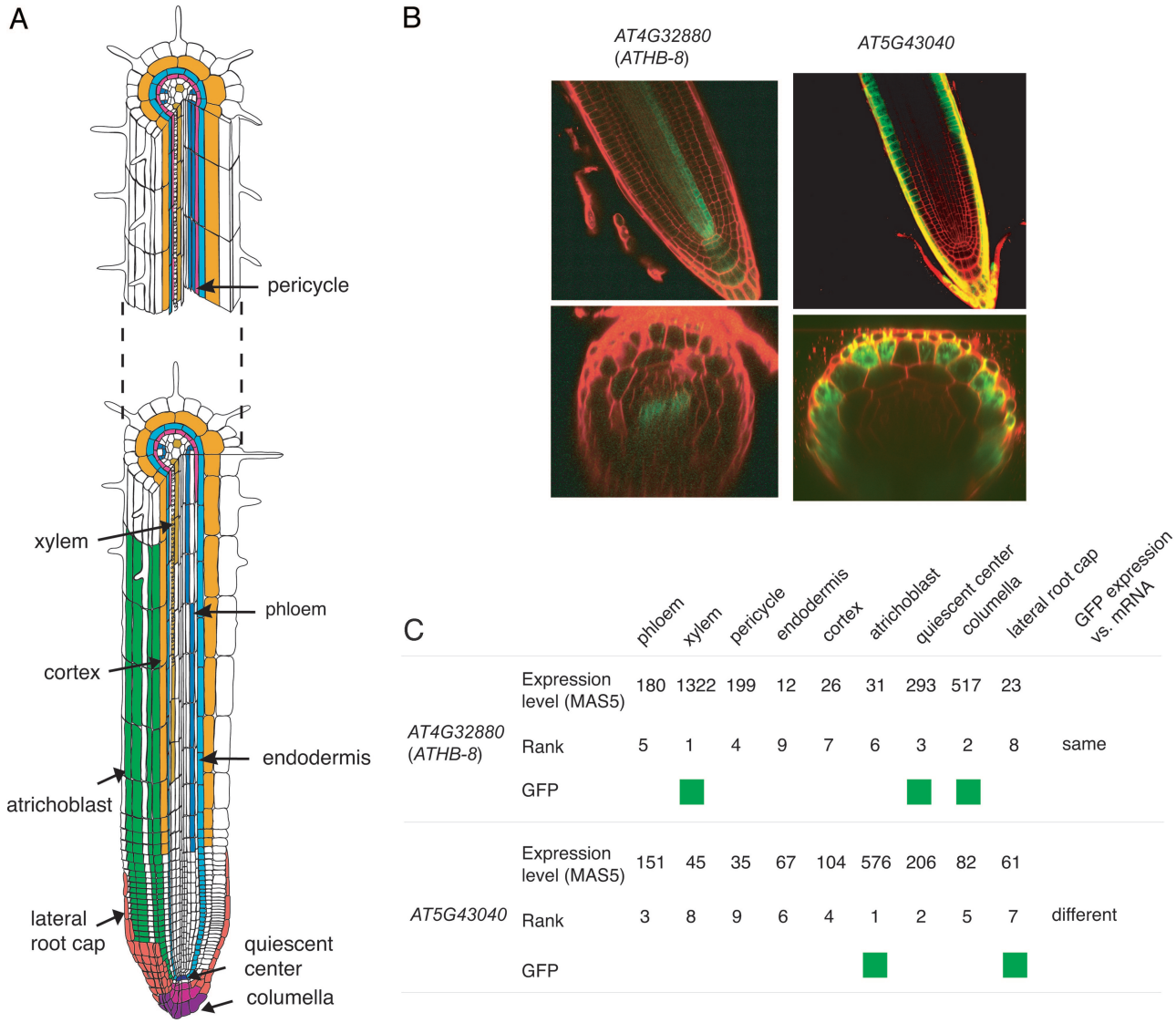

Fig. 1. Comparison of transcriptional fusion GFP expression patterns with mRNA expression patterns from the root expression map by a ranking method. (*A*) The nine root tissues profiled by cell sorting-microarray method to generate the root expression map. (*B* and *C*) Rationale of expression correlation based on the visual rank-based comparison. (*B*) GFP images of transcriptional fusions of two TFs and (*C*) their mRNA expression levels in the nine tissues. Expression levels are ranked based on the mRNA level from the microarray data, and the tissues with GFP expression are marked with green squares. For *ATHB-8*, because GFP is expressed in the three top-ranked tissues (*B Upper*, columella and quiescent center; *B Lower*, xylem), the transcriptional fusion and mRNA patterns are scored as the ''same'' in *C*. For *AT5G43040*, GFP expression was found both in the first-ranked tissue (*B Lower*, atrichoblast) and in the seventh-ranked tissue (*B Upper*, lateral root cap). Therefore the transcriptional fusion and mRNA patterns are scored as ''different'' in *C*.

and in Fig. 4, which is published as supporting information on the PNAS web site. Several tissue-specific GFP lines were produced, specific to the endodermis, quiescent center, phloem, or xylem (Fig. 5, which is published as supporting information on the PNAS web site), and these lines can be used to further refine the root expression map as well as for a wide range of other studies.

Visual Comparison of the GFP and mRNA Expression Patterns and Validation by Using Automated Image Analysis. The expression patterns of these transcriptional and translational GFP fusions were compared with the *in vivo* mRNA expression pattern of each gene as derived from the root expression map. Root map expression profiles have been shown to accurately reflect published mRNA expression patterns from *in situ* hybridization for 25 of 26 genes and, thus, provide a reliable estimation of mRNA expression patterns *in vivo* (14). To enhance the resolution of the expression map, we profiled three additional radial tissues in the root (xylem, phloem, and cortex; see *Materials and Methods*). Combined with previously described profiles from six other tissues (14, 16, 18), the resulting root expression map covers nine nonoverlapping tissues and represents most of the cell populations found along the radial axis of the root and in the root cap (Fig. 1*A*; all microarray data are available at www.arexdb.org).

The GFP expression patterns were visually compared with the mRNA expression patterns estimated from the microarray data by using a ranking method: The absolute microarray intensity values from the nine tissues of the expression map were ranked, and the two patterns were scored as being the same if the GFP signal was found only in tissues consecutively ranking at the top, otherwise they were scored as different (see Fig. 1 *B* and *C*, and Fig. 6, which is published as supporting information on the PNAS web site). In the cases where the transcriptional GFP and microarray expression patterns were scored as being the same, it was concluded that the upstream noncoding sequence was sufficient to recapitulate the mRNA expression pattern.

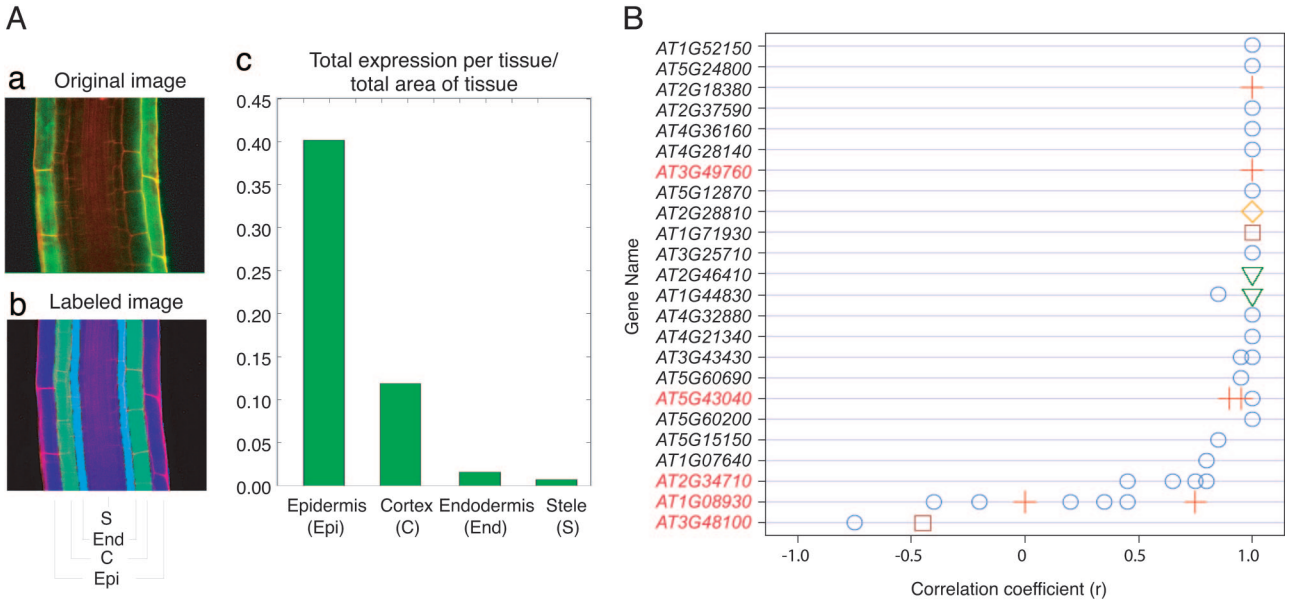

Fig. 2. Comparison of transcriptional fusion expression patterns with mRNA expression patterns from the root expression map by an automated image analysis method. (*A*) Example of quantification of GFP by using automated image analysis. Starting with the original image (*a*), total intensity values of the green channel are quantified by using our labeling scheme (*b*) and allocated to each tissue type. The resulting quantification output of total expression divided by total area (*c*) is used for comparison with data from the root expression map. (*B*) Correlation between automated image analysis data and root expression map data. Pearson correlation values between microarray and numerical GFP expression values obtained from automated image analysis are plotted for multiple lines and genes are sorted by their mean correlation value. Because *r* values are very close for many images, we grouped the correlation values into categories if they were within 0.05 of each other (○, 1 image; +, 2 images; ▽, 3 images; □, 4 images; ◇, 6 images). The genes identified as ''different'' with the visual rank-based comparison are labeled in red.

This visual rank-based comparison is sensitive to our ability to detect GFP. To corroborate this method, we developed an automated image analysis program by which the GFP signals in the different tissues are converted into numerical values. We trained the algorithm to identify four tissues of the root elongation zone (stele, endodermis, cortex, and atrichoblast) for which we have direct microarray data (see Fig. 2*A*). We were able to analyze 60 images from 24 transcriptional fusion constructs for which relative green fluorescence levels were quantified for the four cell types (described in *Materials and Methods*). Pearson correlation coefficients between the relative expression levels, determined by the automated image analysis method and the microarray data, were calculated and plotted for each image and each gene (Fig. 2*B*). Among the 24 genes analyzed, 5 had been found to have a different pattern with the visual rank-based method (Fig. 2*B*). Three of these (*AT3G48100*, *AT1G08930*, and *AT2G34710-PHABULOSA*) had the lowest average correlation coefficients. The other two (*AT3G49760* and *AT5G43040*) had high positive correlation coefficients, but, in both cases, the difference in expression patterns was based on expression in tissues not identified by the automated imaging algorithm (phloem vs. pericycle, and phloem vs. lateral root cap, respectively). Thus there were no false ''different'' calls by the visual nonparametric ranking method on this set of 24 genes.

The automated image analysis method confirmed that our visual approach reliably identifies differences between GFP and microarray expression patterns. Also, digital images of RNA *in situ* hybridizations or GFP reporter constructs are emerging as alternative data to assess tissue-specific expression at single-cell resolution and have been used to reconstruct regulatory pathways in the development of *Drosophila* and *Arabidopsis* (examples are in refs. 19 and 20). The technique developed here to analyze root images is currently limited to analyzing expression in the elongation zone, but could be extended to the rest of the root. This development will constitute a valuable resource for transcription network reconstruction studies. As such it could also be used to study subtle changes in expression patterns, which are not visually detectable.

Upstream Regulatory Regions Are Generally, but Not Always, Sufficient to Reconstitute mRNA Expression Patterns. The results of the visual rank-based comparison of the GFP and microarray expression patterns were used to assess the overall ability of upstream noncoding sequences to recapitulate mRNA expression patterns (results listed in Table 3). Among the 61 transcriptional fusions generated, we were not able to detect GFP visually in any root tissue for 13 of them (21%). It is possible that some of these transcriptional fusions did not generate detectable GFP because the mRNA is naturally expressed at a low level. Indeed, for these genes the maximum microarray expression value in the nine profiled tissues was significantly lower on average than for genes whose transcriptional fusion recapitulated the mRNA expression pattern (average 419 vs. 1,434; Student's *t* test, $P = 0.00003$). In any case, we excluded these 13 fusions from further analysis because the ability of an upstream regulatory sequence to recapitulate an mRNA expression pattern cannot be assessed without detectable GFP. We also excluded four other fusions with visible GFP expression in tissues not covered by the microarray expression map. Of the remaining 44 transcriptional fusions, 35 (80%) were scored as having the same mRNA and GFP transcriptional fusion expression patterns (Table 1). Therefore, our analysis suggests that the noncoding sequence within 3 kb upstream of a TF is sufficient for driving the endogenous mRNA expression pattern in 80% of the cases.

Nine of the 44 TFs (20%) exhibited different transcriptional GFP and mRNA expression patterns based on the microarray data. These differences seem to reflect real discrepancies between the endogenous mRNA pattern and the GFP expression pattern. Indeed, only one gene (*AT3G48100*) exhibited a dramatic difference between its transcriptional GFP (procambium and columella) and mRNA expression domains from the root expression map (enriched in atrichoblast). All of the other genes showed differences

Table 1. Summary comparison of transcriptional and translational fusion expression patterns with mRNA expression patterns from the root expression map

Translational vs. transcriptional	mRNA pattern = transcriptional	mRNA pattern ≠ transcriptional	Total
Same	**18**	2	20
Expanded and nuclear localized	**4**	2	6
Reduced	**1**	**1**	2
Unclear*	6	1	7
No GFP	5	2	7
No data†	1	1	2
Total	35 (80%)	9 (20%)	44

Among the 61 transcription factors (TFs) for which transcriptional GFP was analyzed, only 44 are listed in the table as their GFP patterns could be visually compared with the mRNA expression patterns based on the microarray data. Bolded numbers represent those 24 genes for which the transcriptional fusion likely recapitulates the endogenous mRNA expression pattern.

*More lines should be analyzed for definitive judgment (5/7) or the translational expression pattern is expanded but not nuclear-localized (2/7).

†Cloning of the translational fusion failed due to cDNA cloning issues.

among neighboring cell types (e.g., phloem vs. pericycle), and these differences are unlikely to be because of contamination of neighboring tissues in generating the microarray data because the expression levels in the neighboring tissues are dramatically different. Furthermore, independent confirmation of the mRNA expression pattern from *in situ* hybridization data has been published for one of them (*DAG1 - AT3G61850*) (21). An incomplete upstream noncoding sequence limited to 3 kb had been used for four of these nine transcriptional fusions. However, we think it is unlikely that this incomplete upstream noncoding sequence is the reason for the failure to recapitulate the mRNA expression pattern because we find a similar proportion of incomplete noncoding sequences among the 35 transcriptional fusions that gave the same expression pattern (17/35). Thus, a more likely explanation is that regulatory elements downstream of the start site are necessary to modify the expression pattern conferred by the upstream noncoding sequence. Consistent with this hypothesis, for one of these nine genes the mRNA expression pattern was restored in the translational fusion (*PHABULOSA*), probably mediated by microRNA-mediated mRNA degradation as described in the next section (Fig. 3*A*).

Overall, our results suggest that for ≈80% of *Arabidopsis* TFs, regulatory regions sufficient for reconstituting mRNA expression patterns reside within 3 kb of the upstream noncoding region. Interestingly, ≈85% of known transcription factor binding sites in humans are found within 3 kb upstream of the transcription start site (22). However, for the remaining 20%, sequences not contained within our transcriptional fusions such as the 3′UTR, introns, or the coding region, are probably required for modifying the mRNA pattern conferred by the upstream sequence. This modification results in a modest expansion or reduction of the expression domain in most cases. Indeed, a recent study in *Drosophila* suggests that microRNA-mediated regulation is an important means of modulating the boundaries of the expression domain (23). Such effects might not have been easily recognized by using traditional RNA *in situ* hybridization data. Thus, refining the microarray expression map further is likely to reveal more genes regulated by sequences other than the upstream noncoding sequences.

Transcription Factor Localization Can Be Affected by Sequences Within the Coding Region. Gene expression patterns can be affected by various posttranscriptional events, including RNA degradation and intra- and intercellular protein trafficking. From the 48 TFs for which the transcriptional fusion yielded detectable GFP, we obtained translational fusion data for 47. Of these translational fusions, 8 yielded no visible GFP, and 5 had expression that was inconsistent between independently transformed lines. We found notably different translational and transcriptional expression patterns for 10 of the 34 remaining genes, reduced in breadth in 2 and expanded in 8 cases (Table 3). To estimate how often coding sequences affect expression patterns, we compared these pattern changes to the endogenous expression pattern obtained from the microarray data.

The translational pattern was reduced in breadth compared with the transcriptional pattern for *PHABULOSA* and *ATHB-8 (AT4G32880)*. *PHABULOSA* is a known target of microRNA 165/166 in shoots (Fig. 2*A*) (24, 25), and the closest paralog of *ATHB-8* was shown to be regulated by the same microRNA (26). Translational GFP expression was suppressed in the quiescent center for both genes and in the endodermis and cortex for *PHABULOSA* (Fig. 3*A*), matching the mRNA expression pattern from the microarray data. Therefore, addition of the coding sequence for this gene allows the mRNA expression pattern to be recapitulated. For *ATHB-8*, the transcriptional and mRNA expression patterns had been scored as the same (Table 1). However, this scoring is still consistent with mRNA degradation occurring because the mRNA expression in the tissue where it is putatively suppressed (quiescent center) ranked lower than in other tissues with visible GFP in the transcriptional fusion (see Fig. 1 *B* and *C*). In summary, among the 24 TFs for which the transcriptional fusion (the translational fusion for *PHABULOSA*) recapitulated the mRNA expression pattern (bolded numbers in Table 1), 2 (8%) are likely to be posttranscriptionally regulated by microRNA-mediated mRNA degradation. It is noteworthy that in both cases the effects take place in the quiescent center, a tissue involved in suppressing differentiation of the surrounding initial cells (27). In an animal study (22), stem cell differentiation was shown to be suppressed by the down-regulation of genes inducing differentiation, this effect partly taking place via microRNAs. Consistent with this finding, expression profiling also showed enrichment of genes involved in mRNA silencing in the quiescent center (16).

We found the translational pattern to be increased in breadth compared with the transcriptional pattern for eight TFs. Movement between cell layers has been described as functional for other plant TFs (*SHR*, *CPC*, *LFY*, *KNOTTED*) based on differences in protein and mRNA expression patterns and nuclear localization in the tissues into which the protein has moved (10). Here, for six TFs, expression was localized to the nucleus in the tissues in which the pattern was expanded. The transcriptional fusion recapitulated the mRNA expression pattern for four of these six TFs (*AT4G00940*, *AT4G27410*, *AT4G37940*, and *CPC-AT2G46410*) (Table 1). Because the expansion in expression pattern must occur after the establishment of the mRNA expression pattern in these cases, it is likely that it occurred via cell-to-cell protein movement, as shown for *CPC* in ref. 28. Therefore, among the 24 TFs for which the reporter mRNA is expressed in the appropriate tissues (bolded numbers in Table 1), 4 (17%) are likely to undergo posttranscriptional modification via intercellular protein trafficking. For the other 2 TFs (*DAG1* and *AT2G22850*), because the transcriptional and mRNA expression patterns were scored as different, the protein might not be expressed in its native tissue. Nevertheless, because the mRNA and translational expression patterns are also different, the expansion of their pattern is unlikely to be because of mRNA movement.

Whether TF protein movement between cells is an active (targeted) or passive (nontargeted) process is still debated (29). Our results are not consistent with a passive process based on three observations. First, for passively moving proteins, smaller proteins were shown to move more effectively than larger ones, presumably dependent on the size exclusion limits of plasmodesmata (30). Here, the size distribution of moving and nonmoving TFs overlapped. Second, it has been suggested from studies of LEAFY movement that proteins move unless they are retained in a specific

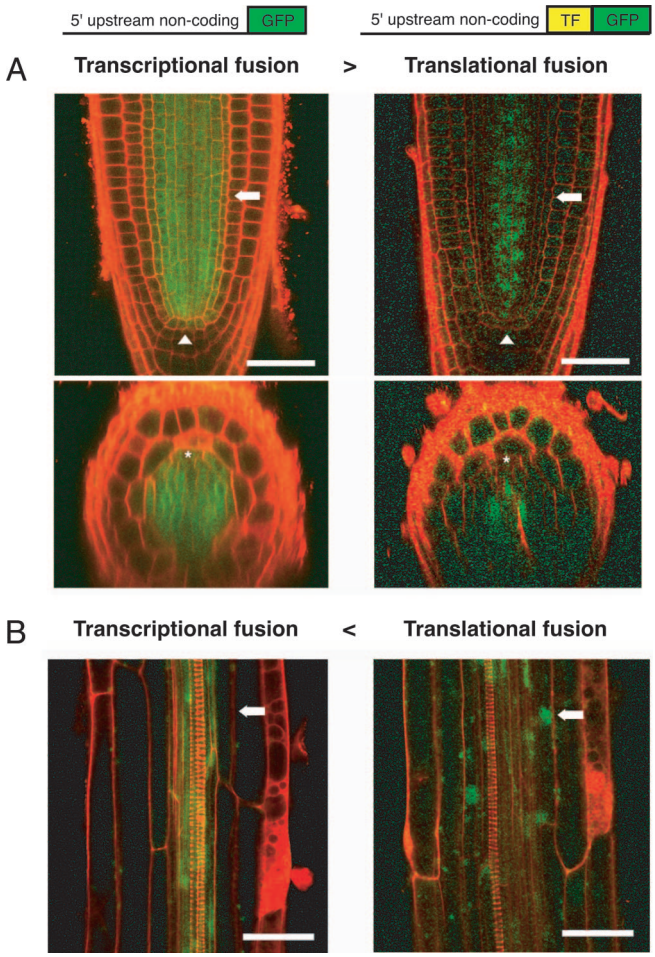

Fig. 3. Examples of coding sequences affecting the localization of transcription factors. (*Left*) Expression patterns of the transcriptional fusion constructs. (*Right*) Translational fusion patterns. (*A*) Example of reduction in expression breadth most likely because of microRNA-mediated mRNA degradation for *PHABULOSA*. In the translational fusion line, GFP expression is lost in the endodermis (arrows in *Upper*; ∗ in *Lower*), the endodermis/cortex initials, and the quiescent center (arrowheads in *Upper*). (*B*) Example of increase in the expression breadth most likely because of intercellular protein movement for *DAG1*. The GFP expression domain has expanded one cell layer from the stele to the endodermis (arrows). (Scale bar: 40 μm.)

subcellular compartment (31). Partially supporting this idea, both cytoplasmic and nuclear localization was found for translational GFP movement within the stele (see *CPC* and *DAG1* in Fig. 4). However, among all moving (8) and nonmoving (26) TFs, a similar proportion of moving TFs was found for those that were not nuclear localized (2/9) and those that were (6/25). Our observation suggests that cytoplasmic localization is necessary but not sufficient for intercellular protein movement. Third, for the three genes for which the expansion occurred within the vasculature (*DAG1*,*AT2G22850*, and *CPC*), expansion occurred throughout the entire tissue (i.e., multiple cell layers). In contrast, in the cases where expansion occurred to nonvascular tissues (including *CPC* in the epidermis), it only occurred to the adjacent cell layer. This observation provides further evidence that the ability of a TF to move can depend on the cell type in which it is expressed (32). In summary, our results suggest that movement of TFs is more likely an actively regulated rather than a simple size-dependent passive process (33).

Besides regulation of intercellular movement, TF activity might also be regulated via modulation of subcellular localization, as has been previously reported in plants (34, 35) and animals (12, 13). For 10 of the 39 (25%) translational fusions that had detectable GFP, the GFP-fusion protein was not found to be nuclear-localized in any of the root tissues (Table 3). Because only two of these GFP-fusion proteins had different transcriptional and mRNA expression patterns, ectopic protein expression is unlikely to be the reason for this lack of nuclear localization. Assuming that the annotation of these ten genes as transcription factors is correct, it is thus possible that this regulatory mechanism is relatively common for transcription factors.

In summary, by comparing mRNA expression patterns with transcriptional and translational reporter patterns, we have obtained an estimate of the relative importance of transcriptional and posttranscriptional mechanisms in regulating TF expression patterns in the *Arabidopsis* root. We conclude that 5′ upstream noncoding sequences control the major patterns of transcription, and that movement of TF proteins is relatively common. These data will provide valuable information for future gene regulatory network analyses.

Materials and Methods

Tissue-Enriched Transcription Factor Selection. A list of transcription factors represented on the Affymetrix ATH1 array was compiled from three TF databases (http://rarge.gsc.riken.jp/rartf and two databases in refs. 36 and 37). This list was then used to search the microarray data for TFs with a two-fold enrichment in one tissue as compared with all of the others. Two-fold enrichment was chosen based on "spike-in" experiments (38). Also, 150 was set as the minimum MAS5 level in the enriched tissue because 99% of the genes with MAS5 expression levels below that value were called absent by the Affymetrix software. All of the TFs found enriched in the endodermis (14), stele (27), and atrichoblasts (15) were cloned whenever this cloning was successful, and TFs of interest to our laboratory were selected for the quiescent center (6) and the lateral root cap (2). Since these 61 TFs had been selected, the total *Arabidopsis* transcription factor list has been expanded to 2,033 genes. Among these 2,033 TFs, 189 TFs were found enriched by using our query criteria. Therefore, this selection represents an essentially random subset of about one-third of the TFs expressed in a tissue-enriched manner in *Arabidopsis* roots.

Microarray Data. Microarray profiling of the phloem, xylem, and cortex was performed by using lines containing promoter-GFP constructs from the genes *AT1G79430*, *AT5G12870*, and *AT1G09750* respectively, with two biological replicates for the phloem and three for the other two. The plants were grown on 1% sucrose nutrient agar and profiled as described in ref. 15. Other microarray profiling data are from Birnbaum *et al.* (14) (atrichoblast, lateral root cap, stele, and endodermis), Nawy *et al.* (16) (columella and quiescent center), and Levesque *et al.* (18) (pericycle). All microarray data were globally normalized to 250 of the average expression value with MAS5 software (Affymetrix).

Cloning, Plant Transformation, and Imaging. Transcriptional and translational GFP reporter fusion constructs were generated by using the MultiSite Gateway Three-Fragment Vector Construction system (Invitrogen). Upstream noncoding sequences were isolated by using nested PCR: for the first reaction, by using primers outside the region of interest and Columbia ecotype genomic DNA as template, and, for the second reaction, by using nested primers with recombination sites and matching sequences to the region of interest, and the first PCR product as template (sequences of nested primers available in Table 3). These PCR products were then cloned by recombination into the appropriate vector. Clones were sequenced and those with no more than one indel mutation per kilobase were selected. For translational fusion constructs, for each gene a full-length cDNA without the stop codon was isolated by reverse transcription–PCR (except for *AT4G32710*; genomic clone including introns) and directionally cloned into pENTR/D-TOPO

by using the pENTR Directional TOPO Cloning kit (Invitrogen). To directly perform the final recombination into a plant transformation vector, we cloned an attR4:ccdB:CmR:attR3 fragment from pDEST R4-R3 into HindIII and XhoI sites of pGreenII 0229 (39) and replaced the selection marker with spectinomycin resistance. This modified vector was then recombined with the upstream noncoding sequence clone and either endoplasmic-reticulum localized GFP for the transcriptional fusions or the clone of a coding region and GFP with no localization signals for the translational fusions. The final constructs in *Agrobacterium* were transformed into Columbia ecotype plants by using the floral dip method (40). Transformed plants (T1) were selected with BASTA (Farnam, Phoenix) on soil and 5- to 7-day-old T2 seedlings were grown and imaged by using confocal laser scanning microscopy as described in ref. 16. An average of six independently transformed lines (for *AT5G43030*, *AT5G15150*, and *AT2G46410*, only two lines were examined) were analyzed for each TF.

Automated Image Analysis. Images (111) from transcriptional fusions showing expression in the elongation zone of the root were selected and automatically analyzed to retrieve GFP signal values in the stele, endodermis, cortex, and epidermis (see an example in Fig. 7, which is published as supporting information on the PNAS web site). These GFP signal values were then aligned to the microarray data values, and both were normalized for contrast enhancement by subtracting 150 (the microarray data absence call value) and setting resulting negative values to zero. Eighty-four images remained after removal of those with no detectable expression after this enhancement. The GFP signal and microarray data values were then converted to a relative intensity level, and for each image a Pearson correlation coefficient was calculated between the microarray and the GFP signal values in the four tissues. The resulting 84 correlation coefficients were plotted with respect to their gene. Images with low correlation coefficients (<0.7) were investigated, and 24 of these outliers were found to be from incorrectly mapped tissue types derived from the image analysis and were removed from the analysis.

The automated image analysis was performed as follows. After each GFP image was separated into red (cell walls), green (GFP), and a blank blue channel, and analyzed in three phases: (*i*) noise removal on the green channel, application of a series of morphological filters and partial subtraction of the red channel to remove refraction noise and cross-talk of the red stain (41) for cleaner separation of tissue layers and contrast enhancement; (*ii*) labeling of the tissue types, determination of the outer cell boundaries of the root by using a contour-tracking algorithm (41) and labeling the tissues by using a template reference image of a typical cross-section; (*iii*) quantification and normalization of the expression values in each tissue type. First, GFP intensity values for all pixels in each tissue were summed. The sum was divided by the total area of each tissue type, to provide relative expression values for the tissues similar to the gene expression map data, and to normalize variation in magnification. All image processing and analysis was performed by using the software program MATLAB IMAGE PROCESSING TOOLBOX.

We thank Yka Heraliutta, Martin Bonke, and Sari Tähtiharju for providing the pAPL::GFP line for gene expression profiling and Paul Magwene, Ken Birnbaum, Siobhan Brady, Jose Dinneney, Terri Long, Hong Chang Cui, and Jee Jung for helpful comments on the manuscript. This work was supported by National Science Foundation Grant 0209754 (to P.N.B.).

1. Levine, M. & Davidson, E. H. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 4936–4942.
2. Ito, T., Sakai, H. & Meyerowitz, E. M. (2003) *Curr. Biol.* **13,** 1524–1530.
3. Hong, R. L., Hamaguchi, L., Busch, M. A. & Weigel, D. (2003) *Plant Cell* **15,** 1296–1309.
4. Kleinjan, D. A., Seawright, A., Childs, A. J. & van Heyningen, V. (2004) *Dev. Biol.* **265,** 462–477.
5. Wienholds, E. & Plasterk, R. H. (2005) *FEBS Lett.* **579,** 5911–5922.
6. Kidner, C. A. & Martienssen, R. A. (2005) *Curr. Opin. Plant Biol.* **8,** 38–44.
7. Conti, E. & Izaurralde, E. (2005) *Curr. Opin. Cell Biol.* **17,** 316–325.
8. Sommer, P. & Nehrbass, U. (2005) *Curr. Opin. Cell Biol.* **17,** 294–301.
9. Mignone, F., Gissi, C., Liuni, S. & Pesole, G. (2002) *Genome Biol.* **3,** REVIEWS0004.
10. Kim, J. Y. (2005) *Curr. Opin. Plant Biol.* **8,** 45–52.
11. Gallagher, K. L. & Benfey, P. N. (2005) *Genes Dev.* **19,** 189–195.
12. Ziegler, E. C. & Ghosh, S. (2005) *Sci. STKE* **2005,** re6.
13. Schuller, C. & Ruis, H. (2002) *Results Probl. Cell Differ.* **35,** 169–189.
14. Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W. & Benfey, P. N. (2003) *Science* **302,** 1956–1960.
15. Birnbaum, K., Jung, J. W., Wang, J. Y., Lambert, G. M., Hirst, J. A., Galbraith, D. W. & Benfey, P. N. (2005) *Nat. Methods* **2,** 615–619.
16. Nawy, T., Lee, J.-Y., Colinas, J., Wang, J. Y., Thongrod, S. C., Malamy, J. E., Birnbaum, K. & Benfey, P. N. (2005) *Plant Cell* **17,** 1908–1925.
17. Haseloff, J., Siemering, K. R., Prasher, D. C. & Hodge, S. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 2122–2127.
18. Levesque, M. P., Vernoux, T., Busch, W., Cui, H., Wang, J. Y., Blilou, I., Hassan, H., Nakajima, K., Matsumoto, N., Lohmann, J. U., *et al.* (2006) *PloS Biology*, in press.
19. Reddy, G. V. & Meyerowitz, E. M. (2005) *Science* **310,** 663–667.
20. Jaeger, J., Surkova, S., Blagov, M., Janssens, H., Kosman, D., Kozlov, K. N., Manu, Myasnikova, E., Vanario-Alonso, C. E., Samsonova, M., *et al.* (2004) *Nature* **430,** 368–371.
21. Papi, M., Sabatini, S., Altamura, M. M., Hennig, L., Schafer, E., Costantino, P. & Vittorioso, P. (2002) *Plant Physiol.* **128,** 411–417.
22. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., *et al.* (2005) *Cell* **122,** 947–956.
23. Stark, A., Brennecke, J., Bushati, N., Russell, R. B. & Cohen, S. M. (2005) *Cell* **123,** 1133–1146.
24. Kidner, C. A. & Martienssen, R. A. (2004) *Nature* **428,** 81–84.
25. Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A. & Timmermans, M. C. (2004) *Nature* **428,** 84–88.
26. Kim, J., Jung, J. H., Reyes, J. L., Kim, Y. S., Kim, S. Y., Chung, K. S., Kim, J. A., Lee, M., Lee, Y., Narry Kim, V., *et al.* (2005) *Plant J.* **42,** 84–94.
27. van den Berg, C., Willemsen, V., Hendriks, G., Weisbeek, P. & Scheres, B. (1997) *Nature* **390,** 287–289.
28. Wada, T., Kurata, T., Tominaga, R., Koshino-Kimura, Y., Tachibana, T., Goto, K., Marks, M. D., Shimura, Y. & Okada, K. (2002) *Development (Cambridge, U.K.)* **129,** 5409–5419.
29. Zambryski, P. & Crawford, K. (2000) *Annu. Rev. Cell Dev. Biol.* **16,** 393–421.
30. Kim, I., Kobayashi, K., Cho, E. & Zambryski, P. C. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 11945–11950.
31. Wu, X., Dinneny, J. R., Crawford, K. M., Rhee, Y., Citovsky, V., Zambryski, P. C. & Weigel, D. (2003) *Development (Cambridge, U.K.)* **130,** 3735–3745.
32. Sena, G., Jung, J. W. & Benfey, P. N. (2004) *Development (Cambridge, U.K.)* **131,** 2817–2826.
33. Gallagher, K. L., Paquette, A. J., Nakajima, K. & Benfey, P. N. (2004) *Curr. Biol.* **14,** 1847–1851.
34. Stacey, M. G., Hicks, S. N. & von Arnim, A. G. (1999) *Plant Cell* **11,** 349–364.
35. Chen, M., Tao, Y., Lim, J., Shaw, A. & Chory, J. (2005) *Curr. Biol.* **15,** 637–642.
36. Davuluri, R. V., Sun, H., Palinkiswamy, S. K., Matthews, N., Molina, C., Kurtz, M. & Grotewald, E. (2003) *BMC Bioinformatics* **4,** 25.
37. Guo, A., He, K., Liu, D., Bai, S., Gu, X., Wei, L. & Luo, J. (2005) *Bioinformatics* **21,** 2568–2569.
38. Choe, S. E., Boutros, M., Michelson, A. M., Church, G. M. & Halfon, M. S. (2005) *Genome Biol.* **6,** R16.
39. Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S. & Mullineaux, P. M. (2000) *Plant Mol. Biol.* **42,** 819–832.
40. Clough, S. J. & Bent, A. F. (1998) *Plant J.* **16,** 735–743.
41. Gonzalez, R. C., Woods, R. E. & Eddins, S. L. (2003) *Digital Image Processing Using* MATLAB (Prentice–Hall, Upper Saddle River, NJ).