

Developmentally distinct MYB genes encode functionally equivalent proteins in *Arabidopsis*

Myeong Min Lee and John Schiefelbein*

Department of Biology, University of Michigan, Ann Arbor, Michigan 48109, USA

*Author for correspondence (e-mail: schiefel@umich.edu)

Accepted 12 February; published on WWW 5 April 2001

SUMMARY

The duplication and divergence of developmental control genes is thought to have driven morphological diversification during the evolution of multicellular organisms. To examine the molecular basis of this process, we analyzed the functional relationship between two paralogous MYB transcription factor genes, *WEREWOLF* (*WER*) and *GLABROUS1* (*GLI*), in *Arabidopsis*. The *WER* and *GLI* genes specify distinct cell types and exhibit non-overlapping expression patterns during *Arabidopsis* development. Nevertheless, reciprocal complementation experiments with a series of gene fusions showed that *WER* and *GLI* encode functionally equivalent proteins, and their unique roles in plant development are entirely due to differences in their *cis*-regulatory sequences. Similar

experiments with a distantly related MYB gene (*MYB2*) showed that its product cannot functionally substitute for *WER* or *GLI*. Furthermore, an analysis of the *WER* and *GLI* proteins shows that conserved sequences correspond to specific functional domains. These results provide new insights into the evolution of the MYB gene family in *Arabidopsis*, and, more generally, they demonstrate that novel developmental gene function may arise solely by the modification of *cis*-regulatory sequences.

Key words: MYB genes, Cell specification, Epidermis, Gene evolution, Morphological diversification, Paralogous genes, *Arabidopsis thaliana*, *WEREWOLF*, *GLABROUS1*

INTRODUCTION

New traits are thought to arise during evolution via gene duplication and divergence. The divergence of gene function may formally result either from changes in regulatory sequences and/or by alterations in protein-coding regions. The relative importance of these mechanisms in generating novel gene functions during evolution is presently unclear (Doebley and Lukens, 1998; Carroll, 2000).

The MYB family of transcriptional regulators in plants provides an excellent model for defining the molecular mechanisms underlying the divergence of gene function during the evolution of multicellular organisms. Although small in animals and fungi, the MYB gene family is large and diverse in flowering plants, and its members regulate a wide array of developmental and biochemical processes (Martin and Paz-Ares, 1997; Rabinowicz et al., 1999; Riechmann et al., 2000). Furthermore, the proliferation of MYB genes during plant evolution coincided with the acquisition of new cellular functions and morphological forms, prompting suggestions that MYB gene divergence played an important role in the diversification of plant developmental mechanisms (Martin and Paz-Ares, 1997; Romero et al., 1998; Rabinowicz et al., 1999).

To explore the factors contributing to morphological diversification, we analyzed two paralogous MYB genes from *Arabidopsis*, *WEREWOLF* (*WER*) and *GLABROUS1* (*GLI*), which participate in different developmental processes. The

WER and *GLI* proteins share 57% amino acid sequence identity, including 91% identity in their MYB domains, and are members of a subclass of R2R3 MYBs in *Arabidopsis* (Fig. 1A; Lee and Schiefelbein, 1999; Oppenheimer et al., 1991). Despite their sequence similarity, the *WER* and *GLI* genes influence distinct aspects of *Arabidopsis* development. The *WER* gene is required to specify the non-hair cell type in the root epidermis and the non-stomatal cell type in the hypocotyl, and, accordingly, *WER* gene expression is limited to a specific subset of developing epidermal cells in the root and hypocotyl (Lee and Schiefelbein, 1999). The *wer* mutations alter the fates of these root and hypocotyl epidermal cells, causing them to adopt the alternate root-hair or stomatal cell fate (Lee and Schiefelbein, 1999). The *GLI* gene specifies the production of trichomes (single-celled hairs) on the surface of leaves and stems, and its expression is limited to the epidermal cells in developing shoot tissue (Koornneef et al., 1982; Larkin et al., 1993). The *gli* mutations inhibit trichome formation, which generates a glabrous (hairless) shoot phenotype, but they do not affect other epidermal cell types (Koornneef et al., 1982; Esch et al., 1994).

In this study, we defined the functional relationship between *WER* and *GLI* by exploiting the distinct mutant phenotypes exhibited by these two related genes. We show that these two genes encode proteins with equivalent functions and that their regulatory sequences are entirely responsible for their distinct roles in plant development. We also identify specific domains

of these proteins that confer common biochemical activities. Further, we find that a third *Arabidopsis* MYB gene, which is distantly related to *WER* and *GL1*, encodes a functionally distinct protein. These results provide direct support for the notion that changes in the *cis*-regulatory sequences of transcription factor genes played a major role in morphological diversification during eukaryotic evolution.

MATERIALS AND METHODS

Arabidopsis strains and growth conditions

The EMS-induced *wer-1* and *gll-1* mutant alleles used in this study have been described previously (Koornneef et al., 1982; Lee and Schiefelbein, 1999). The *wer-1 gll-1* double mutants were obtained by crossing the single mutant lines, generating an F₂ population, and confirming putative double mutants by backcrossing to each single mutant. Plants harboring the *GL2::GUS* reporter construct have been described previously (Masucci et al., 1996). The *GL2::GUS* was introduced into each of the mutant and transgene lines by genetic crosses.

Gene constructs and plant transformation

The selection of the *WER* and *GL1* regulatory sequences used in these constructs was based on previously published work (Larkin et al., 1993; Lee and Schiefelbein, 1999). Constructs designed to encode the *WER*, *GL1*, or *MYB2* proteins were generated by making a gene cassette that included the entire transcriptional unit of each gene from the start to the stop codon. The *WER* regulatory sequences included a 4.0 kb fragment 5' to the start codon and a 1.1 kb fragment 3' to the stop codon. The *GL1* regulatory sequences included a 1.4 kb fragment 5' to the start codon and a 1.8 kb fragment 3' to the stop codon. The 3' region of these genes was included in these constructs because the 3' region of *GL1* is known to be essential for proper expression (Larkin et al., 1993). Details of the transgene constructions are available upon request.

Plant transformation was achieved by electroporating constructs (in the binary vector pBIN19) into the *Agrobacterium* strain GV3101 followed by introduction into *Arabidopsis* using the floral dip method (Clough and Bent, 1998). T₁ seeds were collected in separate pools and transgenic plants selected by plating on medium containing kanamycin (50 µg/ml). More than ten lines were generated for each transgene and all had similar phenotypes, and either two or three independent lines of each were chosen for detailed characterization.

Microscopy

The proportion of cell types in the root epidermis was determined from at least ten 5-day-old seedlings from each strain as previously described (Hung et al., 1998). An epidermal cell was scored as a root-hair cell if any protrusion was visible. Ectopic root-hair cells were identified from transverse hand-cut sections of at least five 4-day-old roots from each strain (>50 hair cells scored per line) as described previously (Galway et al., 1994). 4-day-old roots were embedded in JB-4 resin as described previously (Masucci et al., 1996) and transverse sections were cut.

The distribution of epidermal cell types in the hypocotyl was analyzed as described previously (Hung et al., 1998). Trichome production was determined by counting the total number of trichomes on the surface of the first two leaves of 12-day-old plants. For scanning electron microscopy, specimens from 10-day-old plants were prepared and examined as described previously (Yu et al., 2000).

Histochemical analysis of whole seedlings or plant tissue containing the *GL2::GUS* reporter was performed essentially as described by Masucci et al. (Masucci et al., 1996). 4-day-old roots and hypocotyls were subjected to the GUS histochemical assay,

followed by embedding in JB-4 resin as described previously (Masucci et al., 1996) and transverse sections were cut.

Yeast one-hybrid assay

To assess transcriptional activation, the yeast one-hybrid assay was employed as described by Sadowski et al. (Sadowski et al., 1992). Fragments from the *WER* coding region were fused to the GAL4 DNA-binding domain in the vector pGBT9 and, following transformation into yeast strain HF7c, *lacZ* and *HIS3* gene expression was assessed.

RESULTS

WER and *GL1* possess non-overlapping roles in *Arabidopsis* development

Although their distinct gene expression patterns and mutant phenotypes indicate independent roles for *WER* and *GL1* in plant development, it was conceivable that they possess some degree of functional overlap. Therefore, as a first step toward understanding the relationship between *WER* and *GL1* gene function, we generated plants homozygous for mutations in each gene and examined them for the presence of novel phenotypes. These *wer-1 gll-1* plants displayed the phenotype expected from the simple additive effect of each single mutation; that is, they lack trichomes in the shoot epidermis and produce a small number of non-hair cells in the root (Table 1). Aside from these attributes, the plants are vigorous, fertile

Table 1. Production of epidermal cell types in roots and leaves of *Arabidopsis*

Strain	% Non-hair cells in root	Number of trichomes/leaf
Wild type		
Columbia ecotype	58.7±6.6	19.5±3.0
WS ecotype	59.2±4.1	28.7±3.8
<i>gll</i>	61.0±5.7	0±0
<i>wer</i>	6.8±5.0	21.4±6.4
<i>gll wer</i>	6.5±6.2	0±0
<i>WER::WER wer</i>		
Line 1	60.0±2.4	20.0±4.2
Line 2	60.1±6.4	21.8±3.5
<i>WER::GL1 wer</i>		
Line 1	58.1±7.4	20.7±2.6
Line 2	56.4±5.4	20.7±3.3
Line 3	59.7±2.0	20.7±3.8
<i>GL1::GL1 gll</i>		
Line 1	57.6±4.2	29.6±7.3
Line 2	59.7±3.7	28.2±4.5
<i>GL1::WER gll</i>		
Line 1	59.3±2.0	28.3±4.4
Line 2	57.7±5.1	24.8±3.8
Line 3	59.2±5.1	22.9±11.4
<i>WER::MYB2 wer</i>		
Line 1	2.0±1.2	22.9±3.8
Line 2	3.3±1.4	25.2±3.5
Line 3	4.6±4.0	25.6±5.5
<i>GL1::MYB2 gll</i>		
Line 1	59.5±4.1	0±0
Line 2	59.1±7.2	0±0
Line 3	59.5±4.1	0±0

Values represent mean ± s.d.

and exhibit no visible abnormalities (data not shown). The lack of a synergistic effect between these mutations confirms that *WER* and *GL1* have non-overlapping functions in *Arabidopsis*.

The GL1 protein can functionally substitute for WER

To define the relationship between the *WER* and *GL1* functions, we wished to examine the effect of each protein when controlled by the other gene's regulatory sequences. As a first step, we identified the *cis*-regulatory sequences necessary for *WER* gene function. Based on results from a prior study (Lee and Schiefelbein, 1999), we first tested a construct harboring the *WER* protein-coding region fused to a 4.0 kb fragment 5' to the start codon and a 1.1 kb fragment 3' to the stop codon. This construct (designated *WER::WER*; Fig. 1B) was introduced into the *wer-1* mutant background, and the resulting transgenic plants were able to produce a normal proportion of epidermal cell types in the root (Fig. 2A; Table 1). The ability of this *WER::WER* construct to rescue the *wer-1* mutant phenotype indicates that the *cis*-regulatory sequences employed in this construct are sufficient to direct normal *WER* gene function.

To determine whether the GL1 protein can functionally substitute for the WER protein, *wer-1* mutant plants were transformed with a construct (designated *WER::GL1*) containing the *GL1* coding region fused to the same *WER* *cis*-regulatory sequences employed in the *WER::WER* construct (Fig. 1B). These *WER::GL1 wer-1* plants were able to generate a normal proportion of epidermal cell types in the root (Fig. 2A, Table 1), providing the first evidence for functional equivalence between the GL1 and WER proteins.

The two epidermal cell types in the *Arabidopsis* root are normally distributed in a position-dependent pattern. Epidermal cells located outside a single underlying cortical cell (i.e. outside a periclinal cortical cell wall) are specified as non-hair cells, whereas cells located outside two cortical cells (i.e. outside an anticlinal cortical cell wall) are specified as hair cells (Dolan et al., 1994; Galway et al., 1994; Berger et al., 1998). This pattern is disrupted in the *wer-1* mutant, which produces 47% of its hair cells in the non-hair position (termed ectopic hair cells; Lee and Schiefelbein, 1999; Fig. 2A; Table

Table 2. Location of epidermal cell types in the root and hypocotyl of *Arabidopsis*

Strain	% Ectopic root hair cells in the root*	% Ectopic stomata in the hypocotyl‡
Wild type (Columbia)	2.0±2.0	3.6±1.3
<i>wer</i>	47.1±5.2	32.5±12.4
<i>WER::WER wer</i> Line 1	1.7±1.5	3.3±2.2
<i>WER::GL1 wer</i> Line 1	2.1±2.3	4.2±2.9
<i>WER::MYB2 wer</i> Line 1	n.d.	28.9±8.9

*Percentage of the root-hair bearing cells that are located outside a periclinal cortical cell wall in the root epidermis. Values represent mean ± s.d.
‡Percentage of the stomatal complexes that are located outside a periclinal cortical cell wall in the hypocotyl epidermis. Values represent mean ± s.d.

2). In the *WER::GL1 wer-1* plants, we discovered that the normal cell pattern was restored, with a low frequency of ectopic hair cells that is similar to the wild type and to the *WER::WER wer-1* plants (Fig. 2A, Table 2). This shows that, when controlled by *WER* regulatory sequences, the GL1 protein is able to direct root epidermal patterning in the same manner as WER.

In addition to its effect on root epidermis development, the *WER* gene influences cell-type specification in the hypocotyl, where it contributes to a position-dependent pattern of stomatal and non-stomatal cells that is similar to the hair/non-hair

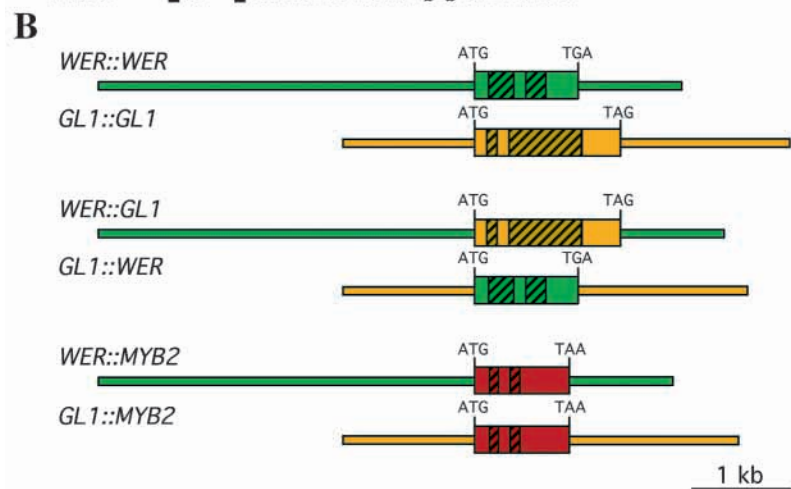
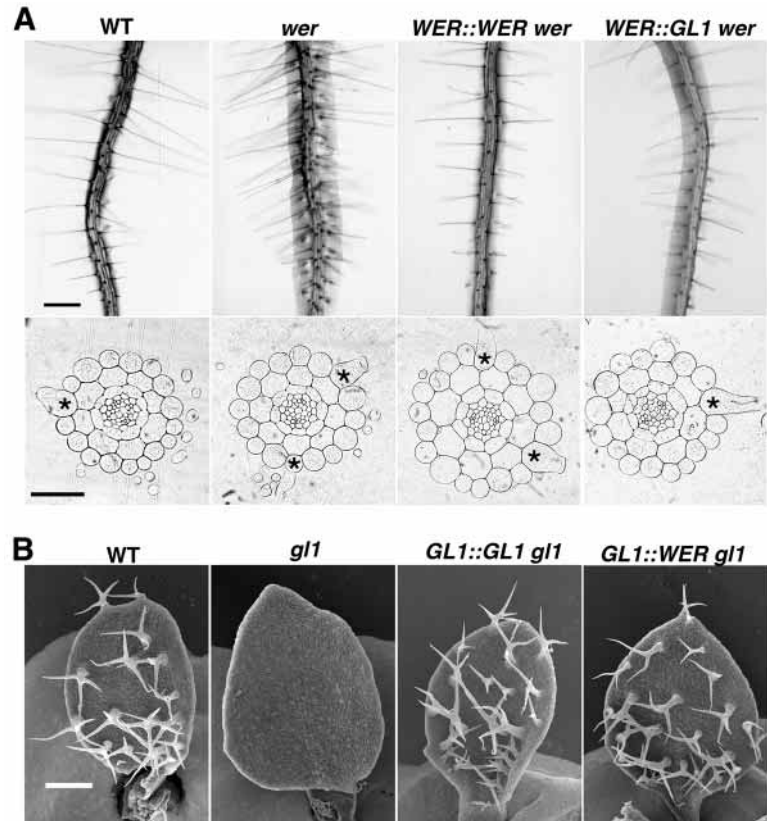


Fig. 1. MYB genes and constructs in *Arabidopsis*.

(A) Sequence alignment of the WER, GL1 and MYB2 proteins. Shaded letters indicate identical residues. The two MYB DNA binding domains present in each of these proteins is indicated by arrows. Percentage amino acid identities: WER-GL1: 57%; WER-MYB2: 33%; GL1-MYB2: 34%. Single-letter abbreviations for amino acid residues: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (B) Representation of the MYB genes and constructs used in the transgenic plant experiments. Thin lines represent regulatory sequences and thick bars indicate protein coding (colored) or intron (hatched) sequences. Sequences derived from the *WER* gene are shown in green, *GL1* sequences are in yellow, and *MYB2* sequences are in red.

Fig. 2. Root and leaf epidermis phenotypes in wild-type, mutant and transgenic plants. (A) Mature root in side view (top row) and transverse section (lower row) from 4-day-old plants showing root hair production from the epidermis of wild type (WT), *wer-1*, *WER::WER* in *wer-1* and *WER::GL1* in *wer-1*. The asterisks mark the cells with visible root hairs in these sections. Note that although all cells outside anticlinal cortical cell walls produce root hairs, the extended longitudinal length of the cells and the small thickness of the sections prevents all hairs from being observed in each section. Bar, 300 μ m for the top row and 50 μ m for the lower row. (B) Leaves from 10-day-old plants showing trichome production in wild type (WT), *gll-1*, *GL1::GL1* in *gll-1*, and *GL1::WER* in *gll-1*. Bar, 200 μ m.



pattern in the root epidermis (Lee and Schiefelbein, 1999). The *wer-1* mutations alter the normal pattern of hypocotyl cell types, causing a large proportion of epidermal cells outside a periclinal cortical cell wall to differentiate into stomatal complexes (termed ectopic stomata; Table 2). We found that the *WER::GL1* transgene restores the normal pattern of hypocotyl cell types to the *wer-1* mutant (Table 2), which shows that GL1 (when controlled by the *WER* regulatory sequences) is able to provide the *WER* function necessary for appropriate epidermal development in the hypocotyl.

To further define the molecular basis for the effect of the *WER::GL1* transgene on root and hypocotyl development, we examined the expression pattern of a presumptive target gene, *GLABRA2* (*GL2*), which encodes a homeodomain protein (Rerie et al., 1994). The *WER* gene is known to be required for appropriate position-dependent regulation of *GL2* gene expression, directing *GL2* expression in differentiating epidermal cells of the root and hypocotyl that are located outside periclinal cortical cell walls (Fig. 3; Lee and Schiefelbein, 1999). Using a *GL2* promoter::*GUS* reporter construct, we analyzed *GL2* expression during root and hypocotyl development in the *wer-1*, *WER::WER wer-1*, and *WER::GL1 wer-1* lines. In the root, we discovered that *GL2* was expressed in an appropriate temporal manner (during the period of cell differentiation) and spatial pattern (within epidermal cells located outside periclinal cortical cell walls) in the *WER::GL1 wer-1* and *WER::WER wer-1* plants (Fig. 3A). Similarly, the normal *GL2* expression pattern was restored in the hypocotyl of the *WER::GL1 wer-1* plants (as well as the *WER::WER wer-1* plants), relative to the *wer-1* mutant plants (Fig. 3B). Together, these results indicate that the GL1 protein (produced by the *WER::GL1* transgene) utilizes the same molecular mechanisms as *WER* to direct *GL2* expression and epidermal cell fate in an appropriate position-dependent manner during root and hypocotyl development.

The *WER* protein can functionally substitute for GL1

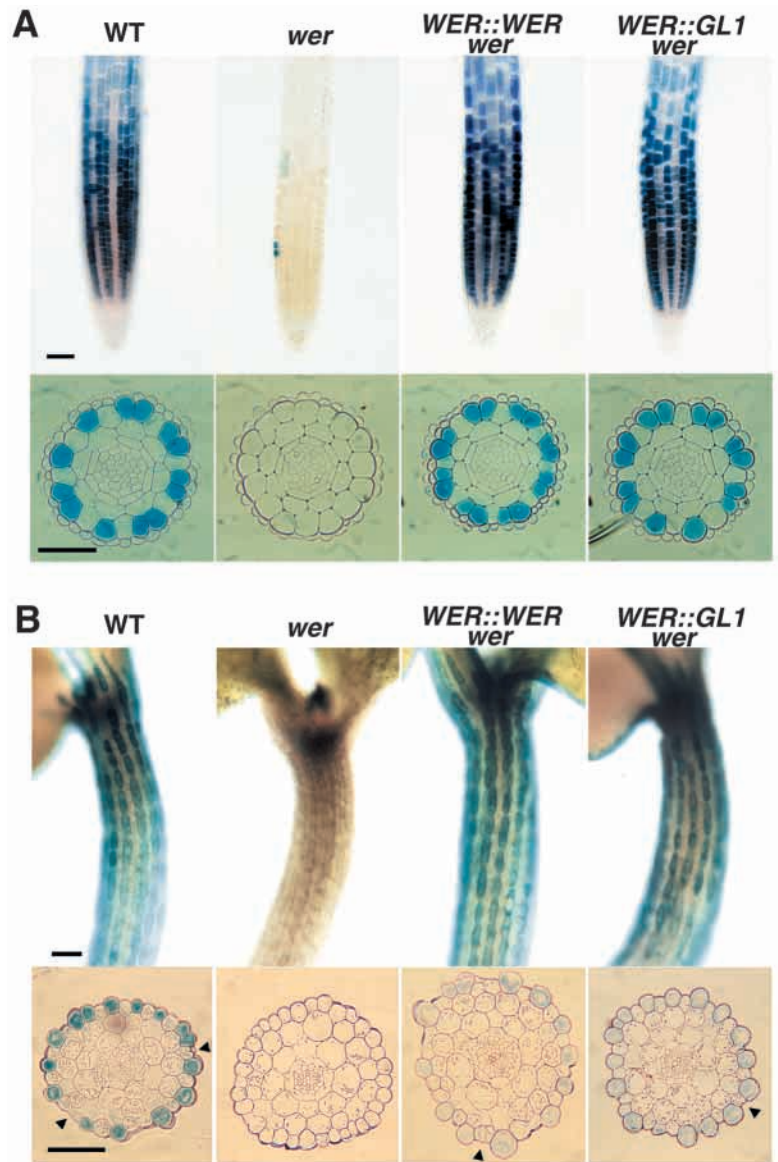
In a set of experiments that are complementary to those described above, we sought to determine whether the *WER* protein can functionally substitute for GL1. We first generated a construct harboring the *GL1* protein-coding region fused to a 1.4 kb fragment flanking the 5' end and a 1.8 kb fragment flanking the 3' end of the *GL1* gene. This construct (designated *GL1::GL1*; Fig. 1B) was introduced

into the *gll-1* mutant background, and the resulting transgenic plants were found to produce a normal number of trichomes (Fig. 2; Table 1). This shows that a sufficient amount of 5' and 3' *cis*-regulatory sequences were employed in this construct to enable normal *GL1* gene function during leaf epidermis development.

A *GL1::WER* fusion construct was generated, which included the *WER* coding sequence fused with the *GL1* regulatory sequences used in the *GL1::GL1* construct (Fig. 1B). Following transformation of *gll-1* mutant plants with this *GL1::WER* construct, we discovered that the *GL1::WER gll-1* plants possess a wild-type number of morphologically normal trichomes on their leaves and stems (Fig. 2B; Table 1, data not shown). Further, although trichome spacing is not determined by a position-dependent mechanism like the epidermal cells of the root and hypocotyl, we observed a general distribution of trichomes (and a lack of trichome clusters) in the *GL1::WER gll-1* leaves that is similar to the non-random distribution found in the wild-type and *GL1::GL1 gll-1* plants (Fig. 2B, data not shown).

The *GL1* gene is a positive regulator of *GL2*, specifically directing *GL2* expression in differentiating trichome cells (Fig. 4; Szymanski et al., 1998). Using the *GL2::GUS* reporter construct, we monitored the spatial and temporal expression of *GL2* in the *GL1::WER gll-1* plants during leaf development. We discovered that the *GL1::WER* construct (like the *GL1::GL1* construct) restores the normal pattern of *GL2::GUS* expression in the developing leaf, relative to the *gll-1* mutant (Fig. 4). This demonstrates that the *WER* protein (when directed by *GL1* *cis*-regulatory sequences) can replace the GL1 function and appropriately direct *GL2*

Fig. 3. Analysis of *GL2* gene expression in the root and hypocotyl of plants bearing the *GL2::GUS* reporter construct. Proper *GL2::GUS* expression, limited to epidermal cells located outside anticlinal cortical cell walls, is restored by the *WER::WER* and *WER::GL1* constructs in the *wer-1* mutant background. (A) Root tip (top row) and transverse section of roots (bottom row) from 4-day-old *GL2::GUS* plants showing GUS accumulation in the developing epidermal cells from wild type (WT), *wer-1*, *WER::WER* in *wer-1*, and *WER::GL1* in *wer-1*. Bars, 50 μ m for all panels. (B) Hypocotyl side-view (top row) and transverse section of hypocotyl (bottom row) from 4-day-old *GL2::GUS* plants showing GUS accumulation in epidermal cells from wild type (WT), *wer-1*, *WER::WER* in *wer-1*, and *WER::GL1* in *wer-1*. Triangles indicate the location of stomata visible in some of the transverse sections. The apparent quantitative difference in GUS staining in some of the transverse sections is not reproducible and reflects section to section variation. Bars, 100 μ m for all panels.



expression and trichome specification during leaf development.

In addition to the epidermal phenotypes described above, we did not observe any abnormalities in the *WER::GL1 wer-1* and *GL1::WER gli-1* plants (data not shown). Taken together, our analysis of these plants shows that the WER and GL1 proteins are functionally equivalent. Either one can appropriately specify the non-hair cell type in the root and the non-stomatal cell type in the hypocotyl (when expressed under control of *WER* regulatory sequences) or the trichome cell type in the shoot (when expressed under the control of the *GL1* regulatory sequences).

The MYB2 protein is not functionally equivalent to WER or GL1

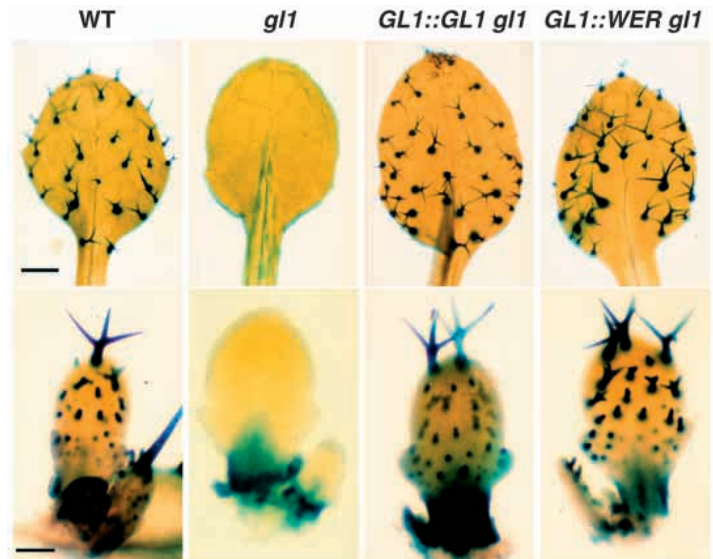
More than 100 MYB genes of the R2R3 type exist in the *Arabidopsis* genome (Romero et al., 1998; Riechmann et al., 2000). One extreme possibility raised by our findings with WER and GL1 is that all of these MYB genes may encode functionally equivalent proteins. To test this possibility, we determined whether a distantly related R2R3 MYB protein from *Arabidopsis* (*MYB2*) can functionally substitute for the WER or GL1 protein. The *MYB2* protein is a member of a different MYB subclass than WER and GL1, it possesses approximately 34% amino acid identity with WER and GL1, and it is associated with plant hormone action (Fig. 1A) (Urao et al., 1993). We fused the coding region of the *MYB2* gene to the *cis*-regulatory regions of the *WER* gene and the *GL1* gene to generate the *WER::MYB2* and *GL1::MYB2* constructs (Fig. 1B). After introducing these constructs into appropriate mutant backgrounds (*WER::MYB2* into the *wer-1* mutant and *GL1::MYB2* into the *gli-1* mutant), we discovered that the phenotypes of the transgenic plants were indistinguishable from the phenotypes of the original mutant lines (Table 1). The reciprocal experiment, testing the ability of WER or GL1 to functionally substitute for *MYB2*, was not possible because no known *myb2* mutant line exists. Nevertheless, the inability of the *WER::MYB2* or *GL1::MYB2* constructs to rescue the phenotype of the *wer-1* or *gli-1* mutant indicates that the

MYB2 protein is not functionally equivalent to either WER or GL1.

A C-terminal domain is required for transcriptional activation

The functional equivalence of WER and GL1 means that these proteins must possess the same biochemical activities. We sought to define the domain that confers one of these biochemical activities, transcriptional activation, by employing the yeast one-hybrid assay. A series of *WER*-derived segments were fused to the GAL4 DNA-binding domain and GAL4-responsive gene expression was assayed. We determined that a protein segment containing the C-terminal 24 amino acids of WER is sufficient for transcriptional activation in yeast (Fig. 5). Interestingly, this 24-residue segment represents the portion of the protein, outside the MYB domains, that is most similar between WER and GL1 (14 of the 24 residues are identical in WER and GL1; Fig. 1A). This suggests that WER and GL1 employ an evolutionarily conserved C-terminal domain to activate gene transcription.

Fig. 4. Analysis of *GL2* gene expression during leaf development in plants bearing the *GL2::GUS* reporter construct. Appropriate *GL2::GUS* expression is restored by the *GL1::GL1* and *GL1::WER* constructs in the *gl1-1* mutant background. Mature leaves (top row) and developing leaves (bottom row) from 12-day-old *GL2::GUS* plants showing GUS accumulation in the epidermal cells from wild type (WT), *gl1-1*, *GL1::GL1* in *gl1-1*, and *GL1::WER* in *gl1-1*. Bar, 300 μ m in the top row and 100 μ m in the bottom row.



DISCUSSION

WER and GL1 are functionally equivalent proteins

This study exploited the unique availability of two related genes that have distinct developmental roles, to define the functional relationship between members of an important family of transcriptional regulators. The most striking finding is that the *WER* and *GL1* genes, which specify distinct cell types during *Arabidopsis* development, actually encode functionally equivalent MYB proteins. Either protein can fully direct the cell specification events normally controlled by the other when expressed under the other gene's regulatory controls. This provides an elegant demonstration of the importance of *cis*-regulatory sequences in specifying developmental gene function. Together with previous comparative evolutionary studies that correlated changes in regulatory sequences of genes with changes in developmental features (Doebley and Lukens, 1998; Wang et al., 1999; Carroll, 2000), our functional analysis of these MYB genes supports the view that modification of *cis*-regulatory sequences of transcriptional regulators has played a key role in morphological diversification during the evolution of multicellular organisms.

To our knowledge, this represents the first direct *in vivo* demonstration of functionally equivalent proteins in plants. Furthermore, unlike studies in animal systems that uncovered functionally equivalent proteins encoded by partially redundant genes or genes with overlapping expression patterns (Bouchard et al., 2000; Greer et al., 2000; Malynn et al., 2000), the present study employed genes with entirely non-overlapping expression domains, which uniquely demonstrates that functional equivalence can be maintained even after complete divergence of regulatory elements.

Conserved domains in WER and GL1 confer common biochemical activities

The ability of the *WER* and *GL1* proteins to fully substitute for one another in different developmental processes shows that, despite sharing only 57% amino acid sequence identity, they possess common biochemical activities. The common activities for these transcription factor proteins are likely to include equivalent DNA binding specificities, equivalent partner protein interactions, and equivalent transcriptional activation properties. As outlined below, each of these activities appears to be conferred by specific domains conserved between the *WER* and *GL1* proteins.

The DNA-binding activity of MYB proteins is known to be determined by the Myb repeats themselves, which form α -helices that contact DNA in a sequence-specific manner (Thompson and Ramsay, 1995). Consistent with the notion that *WER* and *GL1* possess equivalent DNA binding

properties, these proteins exhibit a high degree of amino acid sequence identity (91%) within their two (R2R3) Myb repeats (Fig. 1A).

The view that *WER* and *GL1* possess equivalent partner protein interactions is further supported by previous studies demonstrating that each protein physically interacts with a basic helix-loop-helix (bHLH) transcription factor related to the maize R protein (Larkin et al., 1997; Lee and Schiefelbein, 1999). In a separate study, the specific interaction between the R protein and an R2R3 MYB protein from maize (C1) was shown to be mediated by residues in the first helix of the second Myb repeat of C1 (Grotewold et al., 2000). Significantly, the *WER* and *GL1* protein sequences are identical within this 12-amino-acid first-helix region of their second Myb repeats, implying that these two proteins possess the same bHLH interaction domain.

In the present study, we showed that a 24-residue C-terminal domain, which is similar in the *WER* and *GL1* proteins, confers transcriptional activation properties on GAL4 fusion proteins in yeast. Therefore, equivalent transcriptional activation by *WER* and *GL1* is likely to be mediated by this conserved C-terminal domain.



Fig. 5. Identification of the transcriptional activation domain of *WER* in yeast. The indicated portions of the *WER* protein were fused to the GAL4 DNA-binding domain, expressed in yeast cells, and tested for their ability to activate GAL4-responsive *lacZ* and *HIS* gene expression (indicated by a + or -). Numbers indicate amino acid sequence positions.

Similarities in epidermal cell patterning in the root, hypocotyl and leaf of *Arabidopsis*

Our identification of functional equivalence between WER and GL1 extends our understanding of the close relationship between the molecular mechanisms controlling cell specification in the root, hypocotyl, and leaf of *Arabidopsis*. Each of these processes is dependent upon an identical WD-repeat protein (TRANSPARENT TESTA GLABRA; Koornneef, 1981; Galway et al., 1994; Hung et al., 1998; Walker et al., 1999), a similar, if not identical, bHLH transcription factor related to the maize R protein (Lloyd et al., 1992; Galway et al., 1994; Hung et al., 1998; Payne et al., 2000), and a functionally equivalent MYB transcription factor (WER or GL1; this study). Together, these findings suggest that the same transcriptional machinery may be employed to specify epidermal cell fate in each of these three organs.

The ability of these common transcriptional components to specify different cell types (non-hair cells, non-stomatal cells, and trichome cells) implies that as yet unidentified organ-specific factors also influence epidermal cell fate in the root, hypocotyl, and leaf. One general possibility is that organ-specific regulatory proteins exist and act within the MYB/bHLH pathway to generate distinct sets of target proteins in the three differentiating epidermal cell types. An alternative hypothesis is that the MYB/bHLH pathway is identical in each of the cells and generates the same suite of proteins, but the presence of additional organ-specific proteins cause these three differentiating cells to complete distinct developmental programs.

The MYB family in *Arabidopsis* consists of functionally equivalent and functionally dissimilar proteins

In addition to identifying WER and GL1 as functionally equivalent proteins, we have shown that the distantly related MYB2 protein is not equivalent to either WER or GL1. This demonstrates that alterations in protein-coding sequences has also played a role in generating distinct gene functions during the evolution of the *Arabidopsis* MYB family.

The inability of MYB2 to functionally substitute for WER or GL1 may be due to differences in its DNA binding activity, partner protein interactions, or transcriptional activation properties. The Myb repeats in MYB2 exhibit only 58% identity with the WER Myb repeats, which suggests that these proteins have different DNA binding properties. Also, the first α -helix of the second Myb repeat in MYB2 possesses only 5 of the 12 residues present in the equivalent region of WER, implying that MYB2 is unlikely to interact with an R-like bHLH protein. Finally, MYB2 differs significantly from WER and GL1 in its C-terminal domain (only 6 of 24 residues are identical in MYB2 and WER), which implies that MYB2 may lack the transcriptional activation ability of WER and GL1.

Considering the functional and structural relationships we have discovered between the WER, GL1 and MYB2 proteins, an attractive possibility is that the extensive proliferation of MYB genes that occurred during plant evolution generated specific subclasses of MYB proteins (like the GL1/WER subclass) composed of functionally similar (or equivalent) members, due in part to constraints placed on their divergence by their common biochemical activities. Clearly, the functional characterization of additional MYBs in *Arabidopsis* and other

plant species will be necessary to rigorously test this hypothesis. From this perspective, it is significant to note that among the more than 100 R2R3 MYBs in *Arabidopsis* (Romero et al., 1998; Riechmann et al., 2000) there is a third member of the GL1/WER subclass. This third protein, AtMYBrtf (GenBank accession number Z68158), has an unknown function in *Arabidopsis* but possesses a high level of overall sequence identity to WER and GL1 (approximately 63%), including extensive identity within the Myb repeats (95% identical to WER), a perfect match within the putative bHLH interaction domain described above (the same 12 residues as in WER and GL1), and a similar C-terminal domain (16 of the 24 residues identical to WER). Thus, it will be interesting to determine whether AtMYBrtf is functionally equivalent to WER and GL1 and whether it represents a third MYB that participates in epidermal cell specification during *Arabidopsis* development.

This research was supported by the US National Science Foundation (grant IBN-0077900). The authors thank Steve Clark and Michael Frohlich for helpful comments.

REFERENCES

- Berger, F., Haseloff, J., Schiefelbein, J. and Dolan, L. (1998). Positional information in root epidermis is defined during embryogenesis and acts in domains with strict boundaries. *Curr. Biol.* **8**, 421-430.
- Bouchard, M., Pfeffer, P. and Busslinger, M. (2000). Functional equivalence of the transcription factors Pax2 and Pax5 in mouse development. *Development* **127**, 3703-3713.
- Carroll, S. B. (2000). Endless forms: the evolution of gene regulation and morphological diversity. *Cell* **101**, 577-580.
- 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.
- Dolan, L., Duckett, C., Grierson, C., Linstead, P., Schneider, K., Lawson, E., Dean, C., Poethig, S. and Roberts, K. (1994). Clonal relations and patterning in the root epidermis of *Arabidopsis*. *Development* **120**, 2465-2474.
- Doebley, J. and Lukens, L. (1998). Transcriptional regulators and the evolution of plant form. *Plant Cell* **10**, 1075-1082.
- Esch, J. J., Oppenheimer, D. G. and Marks, M. D. (1994). Characterization of a weak allele of the GL1 gene of *Arabidopsis thaliana*. *Plant Mol. Biol.* **24**, 203-207.
- Galway, M. E., Masucci, J. D., Lloyd, A. M., Walbot, V., Davis, R. W. and Schiefelbein, J. W. (1994). The TTG gene is required to specify epidermal cell fate and cell patterning in the *Arabidopsis* root. *Dev. Biol.* **166**, 740-754.
- Greer, J. M., Puetz, J., Thomas, K. R. and Capecchi, M. R. (2000). Maintenance of functional equivalence during paralogous Hox gene evolution. *Nature* **403**, 661-664.
- Grotewold, E., Sainz, M. B., Tagliani, L., Hernandez, J. M., Bowen, B. and Chandler, V. L. (2000). Identification of the residues in the Myb domain of maize C1 that specify the interaction with the bHLH cofactor R. *Proc. Natl. Acad. Sci. USA* **25**, 13579-13584.
- Hung, C.-Y., Lin, Y., Zhang, M., Pollock, S., Marks, M. D. and Schiefelbein, J. (1998). A common position-dependent mechanism controls cell-type patterning and GLABRA2 regulation in the root and hypocotyl epidermis of *Arabidopsis*. *Plant Physiol.* **117**, 73-84.
- Koornneef, M. (1981). The complex syndrome of ttg mutants. *Arabidopsis Inf. Serv.* **18**, 45-51.
- Koornneef, M., Dellaert, L. W. M. and Van der Veen, J. H. (1982). EMS- and radiation-induced mutation frequencies at individual loci in *Arabidopsis thaliana*. *Mutat. Res.* **93**, 109-123.
- Larkin, J. C., Oppenheimer, D. G., Pollock, S. and Marks, M. D. (1993). *Arabidopsis GLABROUS1* gene requires downstream sequences for function. *Plant Cell* **5**, 1739-1748.
- Larkin, J. C., Marks, M. D., Nadeau, J. and Sack, F. (1997). Epidermal cell fate and patterning in leaves. *Plant Cell* **9**, 1109-1120.

- Lee, M. M. and Schiefelbein, J. (1999). WEREWOLF, a MYB-related protein in *Arabidopsis*, is a position-dependent regulator of epidermal cell patterning. *Cell* **99**, 473-483.
- Lloyd, A. M., Walbot, V. and Davis, R. W. (1992). Arabidopsis and Nicotiana anthocyanin production activated by maize regulators R and C1. *Science* **258**, 1773-1775.
- Malynn, B. A., Moreno de Alboran, I., O'Hagan, R.C., Bronson, R., Davidson, L., DePinho, R. A. and Alt, F. W. (2000). N-myc can functionally replace c-myc in murine development, cellular growth, and differentiation. *Genes Dev.* **14**, 1390-1399.
- Martin, C. and Paz-Ares, J. (1997). MYB transcription factors in plants. *Trends Genet.* **13**, 67-73.
- Masucci, J. D., Rerie, W. G., Foreman, D. R., Zhang, M., Galway, M. E., Marks, M. D. and Schiefelbein, J. W. (1996). The homeobox gene GLABRA2 is required for position-dependent cell differentiation in the root epidermis of *Arabidopsis thaliana*. *Development* **122**, 1253-1260.
- Oppenheimer, D. G., Herman, P. L., Sivakumaran, S., Esch, J. and Marks, M. D. (1991). A *myb* gene required for leaf trichome differentiation in *Arabidopsis* is expressed in stipules. *Cell* **67**, 483-493.
- Payne, C. T., Zhang, F. and Lloyd, A. M. (2000). *GL3* encodes a bHLH protein that regulates trichome development in *Arabidopsis* through interaction with GL1 and TTG1. *Genetics* **156**, 1349-1362.
- Rabinowicz, P. D., Braun, E. L., Wolfe, A. D., Bowen, B. and Grotewold, E. (1999). Maize R2R3 Myb genes: sequence analysis reveals amplification in the higher plants. *Genetics* **153**, 427-444.
- Rerie, W. G., Feldmann, K. A. and Marks, M. D. (1994). The GLABRA2 gene encodes a homeodomain protein required for normal trichome development in *Arabidopsis*. *Genes Dev.* **8**, 1388-1399.
- Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O. J., Samaha, R. R. et al. (2000). Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. *Science* **290**, 2105-2110.
- Romero, I., Fuertes, A., Benito, M. J., Malpica, J. M., Leyva, A. and Paz-Ares, J. (1998). More than 80 R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J.* **14**, 273-284.
- Sadowski, I., Bell, B., Broad, P. and Hollis, M. (1992). GAL4 fusion vectors for expression in yeast or mammalian cells. *Gene* **118**, 137-141.
- Szymanski, D. B., Jilk, R. A., Pollock, S. M. and Marks, M. D. (1998). Control of *GL2* expression in *Arabidopsis* leaves and trichomes. *Development* **125**, 1161-1171.
- Thompson, M. A. and Ramsay, R. G. (1995). Myb: An old oncoprotein with new roles. *BioEssays* **17**, 341-350.
- Urao, T., Yamaguchi-Shinozaki, K., Urao, S. and Shinozaki, K. (1993). An *Arabidopsis* myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. *Plant Cell* **5**, 1529-1539.
- Wang, R. L., Stec, A., Hey, J., Lukens, L. and Doebley, J. (1999). The limits of selection during maize domestication. *Nature* **398**, 236-239.
- Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., Esch, J. J., Marks, M. D. and Gray, J. C. (1999). The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *Plant Cell* **11**, 1337-1349.
- Yu, L. P., Simon, E. J., Trotochaud, A. E. and Clark, S. E. (2000). POLTERGEIST functions to regulate meristem development downstream of the CLAVATA loci. *Development* **127**, 1661-1670.