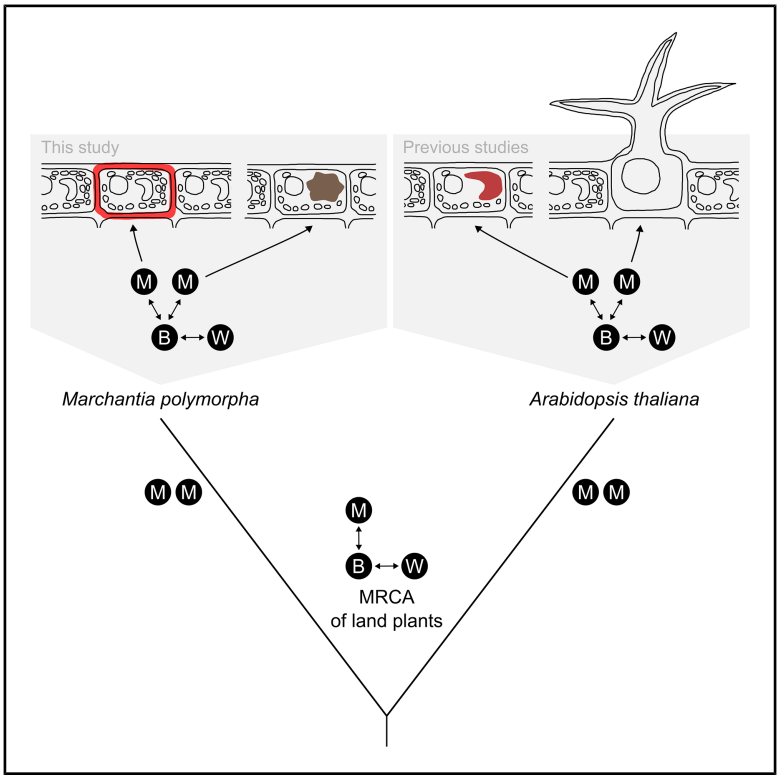
Article

Current Biology

Replicated repurposing of an ancestral transcriptional complex in land plants

Graphical abstract



Authors

Thea E. Kongsted, Facundo Romani, Chiara A. Airoldi, Jim Haseloff, Beverley J. Glover

Correspondence

bjg26@cam.ac.uk

In brief

Kongsted et al. find that, while the red pigments produced in liverworts and seed plants are distinct, they are regulated by homologous MYB-bHLH transcriptional complexes. They also find that in liverworts, multiple paralogous functionally specialized MYB proteins interact with the same pleiotropic bHLH partner, a replicated pattern across lineages.

Highlights

- *Marchantia* homologs of seed plant MBW complexes form homologous interactions
- A shared function is the regulation of red pigments from the flavonoid pathway
- The *Marchantia* homologs also regulate metabolites in a lineage-specific organelle
- Specialized MYB paralogs with pleiotropic bHLH partners are a replicated pattern

Kongsted et al., 2026, Current Biology *36*, 2593–2605
May 18, 2026
https://doi.org/10.1016/j.cub.2026.04.031

CellPress

Current Biology

CellPress
OPEN ACCESS

Article

Replicated repurposing of an ancestral transcriptional complex in land plants

Thea E. Kongsted,[1,2] Facundo Romani,[1] Chiara A. Airoldi,[1,3] Jim Haseloff,[1] and Beverley J. Glover[1,4,*]
[1]Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK
[2]Present address: Gregor Mendel Institute of Molecular Plant Biology, 1030 Vienna, Austria
[3]Present address: Syndex Bio, Cambridge CB22 6SA, UK
[4]Lead contact
*Correspondence: bjg26@cam.ac.uk
https://doi.org/10.1016/j.cub.2026.04.031

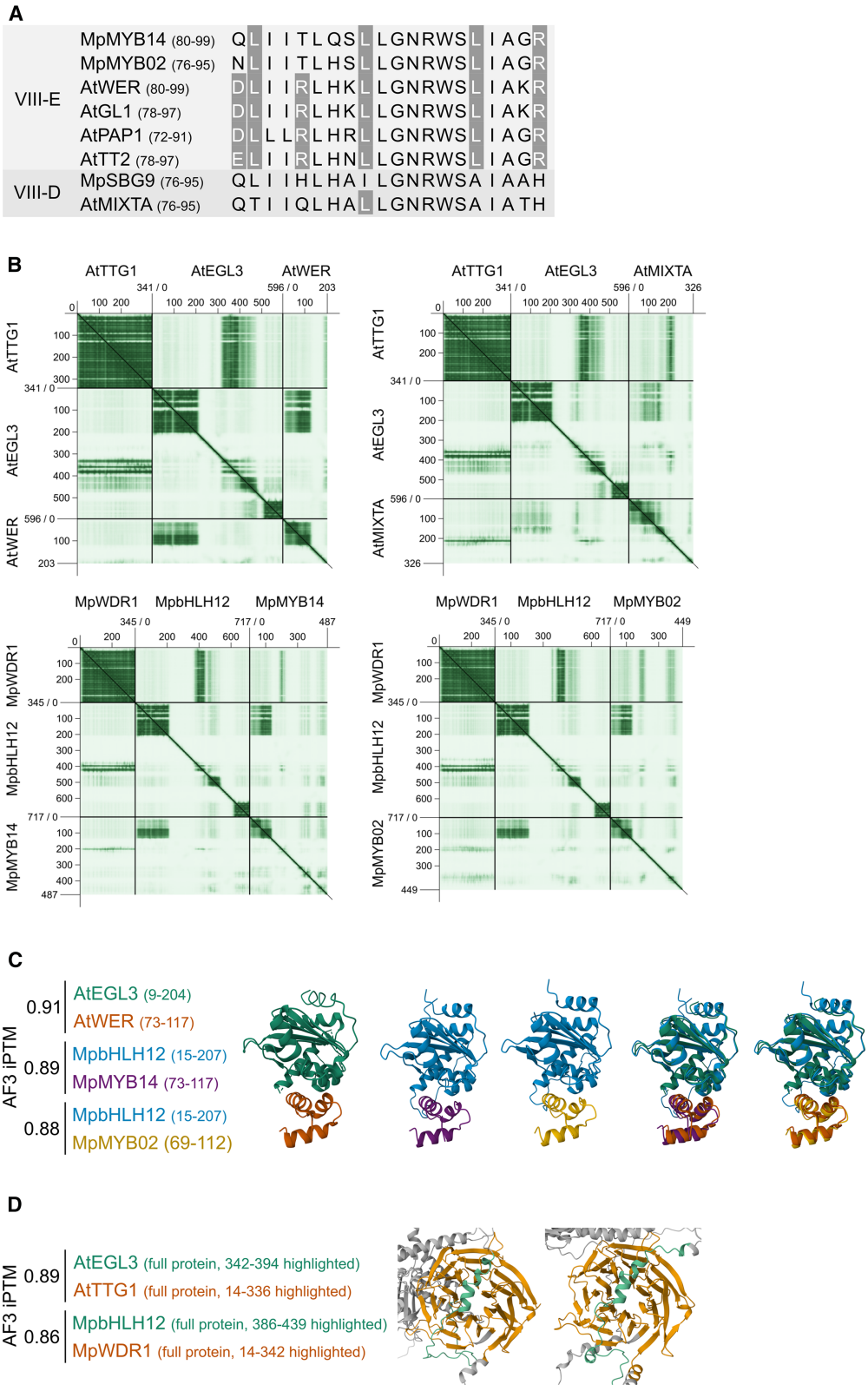
SUMMARY

In flowering plants, transcriptional complexes with a common composition have been found to regulate multiple pathways in the epidermal cell layer, including the production of flavonoid pigments and the differentiation of trichomes and root hairs. These complexes are composed of transcription factors from the MYB and bHLH families along with a WDR scaffold protein (MBW complexes). The MYB member has been found to be the most pathway-specific component of the complex, and modifications to these *MYB* genes are overrepresented in studies investigating the genetic basis of differences in pigmentation phenotypes. Here, we investigated the origin of this complex and its independent functional diversification in the liverwort lineage. We found evidence that homologous transcriptional complexes form in *Marchantia polymorpha*, indicating that they are ancestral to land plants. The single orthologous *bHLH* gene, Mp*bHLH12*, regulates both pathways controlled by the two orthologous *MYB* genes, Mp*MYB14* and Mp*MYB02*. Moreover, MpbHLH12 interacts with two of the TTG1-like WDR proteins in *M. polymorpha*, although their functional role has yet to be determined. We propose that two transcriptional complexes with alternative *MYB* paralogs in *M. polymorpha* represent an ancestral function, regulation of the flavonoid pathway, and a derived function, maturation of liverwort-specific oil bodies. Notably, we find independently in the liverworts, as has been observed in angiosperms, that functional diversification is associated with duplication of the *MYB* member and co-option of its interaction partners.

INTRODUCTION

Across taxa, some loci appear to be overrepresented in studies investigating the genetic basis of phenotypic differences. For instance, a few recurrent loci have been found to be associated with the evolution of diverse wing color patterns across divergent lineages of butterflies and moths.[1] Systematic identification of such "hotspots" of evolution may give broader insights into how the developmental processes that give rise to phenotypes structure evolutionary change.[2] In the flowering plants, red, purple, and blue colors are most commonly caused by anthocyanin pigments derived from the flavonoid pathway. The evolution of pigmentation phenotypes in both natural and cultivated populations has been found to show a higher association with changes to the activating R2R3-MYB transcription factors than with any other component of this pathway.[3–5] One such *R2R3-MYB* gene (*COLORED ALEURONE1*) encodes the first transcription factor identified in a plant and was the site of integration when transposable elements were first discovered by Barbara McClintock in variegated *Zea mays* lines.[6] These MYB proteins regulate the transcription of enzymes in the flavonoid biosynthetic pathway as part of higher-order transcriptional complexes known as MYB-bHLH-WDR (MBW) complexes. In the model flowering plant *Arabidopsis thaliana*, MBW complexes have also been identified that regulate other epidermal phenotypes, including the *de novo* patterning of trichomes on aerial organs,[7] the position-dependent spacing of both root hairs[8] and hypocotyl stomata,[9] and the differentiation of the seed coat.[10] In each case, the MYB member of the complex is pathway-specific.[11] Overlapping subsets of MYB proteins form complexes with each of four basic helix-loop-helix (bHLH) proteins, which thus show several functions and partial redundancy.[7] Finally, a single tryptophan-aspartate repeat (WDR) scaffold protein, TRANSPARENT TESTA GLABRA1 (TTG1), is required for the functions of all of these complexes.[12] So far, MBW complexes have been studied extensively in the angiosperms (reviewed in Ramsay and Glover[11]), and one report found that they also regulate the flavonoid pathway in a gymnosperm.[13] As such, their origin can be traced back at least as far as the stem lineage of seed plants. The participating gene families are found across eukaryotes,[14–16] so it is not clear when the complexes first evolved.

The ca. 20 MYB proteins that participate in MBW complexes in *A. thaliana* share a common motif in the R3 MYB repeat, modification of which was found to perturb interaction with the bHLH partner.[17,18] In a phylogenetic analysis of *R2R3-MYB* genes in land plants and their closest relatives, the Zygnematophycean algae, Jiang and Rao[19] found that this motif is restricted to and broadly conserved across the VIII-E clade, which originated by



(legend on next page)

duplication in the stem lineage of land plants. The motif was identified in sequences from seed plants, but also in sequences from *Selaginella moellendorffii* (a member of the lycophyte clade, which is sister to seed plants and ferns within the vascular plants) and *Physcomitrium patens* (a member of the bryophyte clade, which is sister to the vascular plants within land plants). The most parsimonious scenario would appear to be an origin of the motif at the base of this clade with several shallower secondary losses. Another bryophyte, the liverwort *Marchantia polymorpha*, encodes two clade VIII-E *R2R3-MYB* genes,[19] both of which show only partial conservation of the motif. One of these genes, Mp*MYB14*, is known to activate the production of flavonoids, including red auronidin pigments,[20,21] in response to different environmental challenges, including mineral nutrient insufficiency,[20] high light,[20] and oomycete infection.[22] The other, Mp*MYB02*, is required for the maturation of oil bodies.[23,24] Oil bodies are membrane-bound organelles that originated in the liverwort lineage.[25,26] They accumulate cytotoxic terpenoids and phenylpropanoids, including liverwort-specific bis-bibenzyl compounds,[27] which serve as a defense against herbivory.[28] The limited conservation of the interaction motif in the R3 MYB repeats of the *M. polymorpha* orthologs has previously been taken to imply that they act independently of an MBW complex.[29]

The bHLH proteins that participate in MBW complexes in seed plants all belong to subclass IIIf, whereas a member of subclass III(d + e) has been shown not to interact with the MYB partners.[17] Recently, we found that these two subclasses both derive from early land plant duplications and that they are sisters to one another.[30] This suggests an origin of MBW participation for bHLH proteins coincident with that inferred above for R2R3-MYB proteins in the stem lineage of land plants. *M. polymorpha* encodes a single member, Mp*bHLH12*.[30,31]

For the scaffold protein AtTTG1, highly sequence-conserved putative orthologs can be identified across eukaryotes.[32] The At*TTG1* gene derives from a duplication at the base of seed plants,[33] which also gave rise to the progenitor of the Brassicaceae-level paralogs At*LWD1* and *2*, which redundantly regulate circadian rhythmicity.[34] Three homologs have been identified in *M. polymorpha*, two of which (either Mp*WDR1* or Mp*WDR2*, but not Mp*WDR3*) can complement both the defects in flavonoid pigmentation and epidermal cell fate and in circadian rhythmicity when heterologously expressed in an *A. thaliana ttg1 lwd1 lwd2* triple mutant.[33] The *A. thaliana* paralogs cannot complement one another's functions.[33] This suggests that the seed plant duplication was followed by subfunctionalization and that an ancestral TTG1-like protein in the land plants was capable of scaffolding MBW complexes.[33]

Given that the genes whose protein products form MBW complexes in seed plants were present in the most recent common ancestor of land plants, we hypothesized that their capacity to form complexes is ancestral. To test this hypothesis, we investigated the orthologs in *M. polymorpha*, which diverged from seed plants at the first split within land plants. Indeed, we found evidence that homologous complexes form in *M. polymorpha*, although the functional necessity of the WDR proteins could not be tested. We propose that one complex in *M. polymorpha* retains the ancestral function, shared with seed plants and conserved since stem land plants, whereas another arose within the liverwort lineage. Notably, we find that the functional diversification event in the liverwort lineage was associated with a *MYB* duplication, recapitulating what occurred independently in the flowering plants.

RESULTS

Structural analysis predicts that the *M. polymorpha* orthologs of the MBW components form complexes

A motif in the R3 MYB repeat is conserved across those of the *A. thaliana* clade VIII-E R2R3-MYB proteins (e.g., AtWER, GL1, PAP1, and TT2; Figure 1A), which form MBW complexes. This motif is not conserved in members of the closest paralogous clade, VIII-D (e.g., AtMIXTA and MpSBG9, Figure 1A). The two clade VIII-E R2R3-MYB proteins encoded by *M. polymorpha* show partial conservation of the motif, lacking conservation of the D/E at the first position of the motif and the R/K at the fifth position (Figure 1A). Neither of these two positions was found to be among those with a strong effect on interaction strength when mutated in the study of Zimmermann et al.[17] As a first step in investigating whether these motifs mediate protein-protein interaction in *M. polymorpha*, we performed *in silico* structural predictions with Alphafold3. First, for reference, we examined regions of low predicted aligned error (PAE), representing predicted interaction interfaces, in a structural prediction of the MBW complex formed by AtWER, AtEGL3, and AtTTG1. This prediction recapitulated experimental findings that the N-terminal region of AtEGL3

Figure 1. The interaction interfaces of the MBW complex are predicted to be structurally conserved in the *Marchantia polymorpha* orthologs, despite partial sequence conservation

(A) Amino acid sequence alignment of a region of the R3 MYB repeat of representative R2R3-MYB proteins from *M. polymorpha* and *Arabidopsis thaliana*. Conserved residues of the motif shown to mediate interaction ofclade VIII-E R2R3-MYB proteins with subclass IIIf bHLH proteins in *A. thaliana* are highlighted.[17]

(B) Predicted aligned error plots for trimer predictions for the *A. thaliana* MBW complex WER-EGL3-TTG1, where WER is replaced by the paralogous MIXTA, or where all components are replaced by orthologous proteins from *M. polymorpha*. The predicted error in position of the scored residue (*x* axis) relative to the aligned residue (*y* axis) is shown with a color scale ranging from 0 to 30 Å (dark to light). Each protein is aligned to itself and to each of the other proteins in the prediction.

(C) The solved crystal structure of the interaction interface between AtWER and AtEGL3[36] (PDB: 7FDL) includes the N-terminal region of AtEGL3 (aa 9–204) and the R3 MYB repeat of AtWER (aa 73–117). Alphafold3 predictions were carried out for the homologous regions of the two hypothesized *M. polymorpha* dimers (MpbHLH12 aa 15–207, MpMYB14 aa 73–117, and MpMYB02 aa 69–112) and of the *A. thaliana* dimer for comparison of the interface predicted template modeling (iPTM) scores for predictions of the known and hypothesized dimers. The predicted structures of the *M. polymorpha* dimers were also superimposed on the AtEGL3-AtWER crystal structure for comparison.

(D) Homologous regions were predicted by Alphafold3 to form interaction interfaces exhibiting low predicted aligned error between TTG1-like WDR proteins and IIIf bHLH proteins from *M. polymorpha* (MpbHLH12 aa 386–439 and MpWDR1 aa 14–342 highlighted) as in *A. thaliana* (AtEGL3 aa 342–394 and AtTTG1 aa 14–336 highlighted). The iPTM values are given for dimer predictions using the full protein sequences.

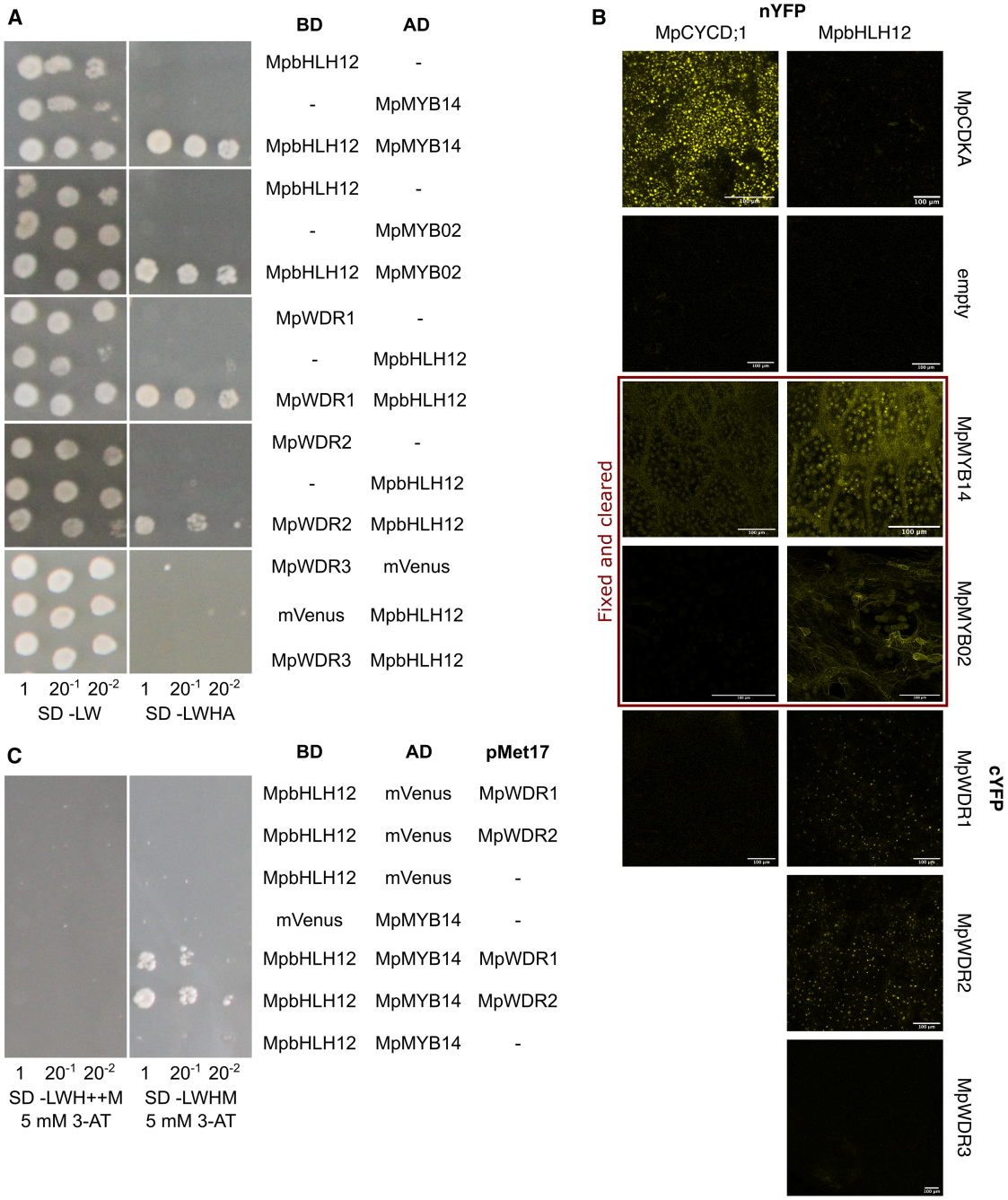


Figure 2. The single subclass IIIf bHLH protein of *Marchantia polymorpha* can form protein complexes with the two clade VIII-E R2R3-MYB proteins and with two of the three TTG1-like WDR proteins

(A) Yeast two-hybrid tests for interaction between the single *M. polymorpha* subclass IIIf bHLH protein and all clade VIII-E R2R3-MYB and TTG1-like WDR proteins from *M. polymorpha*. In each case, the two autoactivation controls and the interaction test are shown for one biological replicate out of four; the rest are given in Figure S1. For each replicate, two serial dilutions of the eluted cells were performed, as indicated. Growth on medium without leucine and tryptophan (−LW) confirms that the yeast clones carry both plasmids with the genes of interest in translational fusion with either the binding domain (BD) or activation domain (AD) of the yeast transcription factor GAL4 (indicated in the table on the right). Interaction between the proteins of interest allows a reconstituted GAL4 to drive expression of reporter genes from the GAL4 upstream activation sequence. This reporter gene activity can complement two independent auxotrophic mutations in histidine and adenine biosynthesis genes in the background genotype, allowing the yeast to grow in the absence of these amino acids (−HA). Further tests for interaction strength using 3-amino-1,2,4-triazole, a competitive inhibitor of the gene product of the histidine biosynthesis reporter, are shown in Figure S1.

(B) Bimolecular fluorescence complementation tests were performed for the single *M. polymorpha* subclass IIIf bHLH protein and all clade VIII-E R2R3-MYB and TTG1-like WDR proteins from *M. polymorpha* by stable transgenic expression of combinations of proteins of interest fused to the N-terminal region or the C-terminal region of YFP. Protein-protein interaction between the proteins of interest results in fluorescence of reconstituted YFP, which was visualized by

(legend continued on next page)

mediates its interaction with both AtTTG1 and AtGL1 (shown to be functionally interchangeable with AtWER[35]).[7] Within AtEGL3, the predicted interaction interface with AtTTG1 was found to be C-terminal to that with AtWER, just N-terminal to the bHLH domain (Figure 1B). We compared this with trimer predictions for the orthologous proteins from *M. polymorpha* and found that interaction interfaces were predicted to form between homologous regions (Figure 1B). This was found with both MpMYB14 and MpMYB02 (Figure 1B). In contrast, if a clade VIII-D R2R3-MYB protein, e.g., AtMIXTA, was instead included in the prediction, the PAE of the homologous residues rose markedly (Figure 1B).

We then investigated more closely the structures of these putative conserved interaction interfaces. The crystal structure of the AtWER-AtEGL3 interaction interface has been solved.[36] Alphafold3 dimer predictions for the regions of AtEGL3 and AtWER included in the solved structure and for the homologous regions of the *M. polymorpha* orthologs showed similar overall interface predicted template modeling (iPTM) scores (Figure 1C), and the predicted *M. polymorpha* dimer structures aligned closely to the AtEGL3-AtWER crystal structure (Figure 1C). No crystal structure has been reported for the bHLH-WDR interaction interface. However, the iPTM scores in dimer predictions for the full protein sequences were similarly high for AtTTG1-AtEGL3 (0.89) and MpbHLH12 with each of MpWDR1 (0.86), MpWDR2 (0.79), or MpWDR3 (0.85). The predicted interaction interfaces in the AtTTG1-AtEGL3 and MpWDR1-MpbHLH12 dimer predictions are highlighted in Figure 1D. Taken together, structural analysis predicts that the interaction interfaces within the MBW complex are conserved in *M. polymorpha*.

The *M. polymorpha* orthologs of the MBW components form complexes *in vivo*

To test the predictions *in vivo*, we expressed our proteins of interest in *Saccharomyces cerevisiae* AH109 in translational fusion with domains of the *S. cerevisiae* transcription factor GAL4 to perform yeast two-hybrid assays (Figure 2A). Consistent with the predictions, we found that protein-protein interaction between MpbHLH12 and each of MpMYB14, MpMYB02, and MpWDR1 (all four biological replicates) and between MpbHLH12 and MpWDR2 (three of four biological replicates) drove sufficient GAL4 reporter activity to complement the strain's auxotrophy for both histidine and adenine (Figure 2A). The strength of these interactions was sufficient to overcome the addition of at least 1 mM of 3-amino-1,2,4-triazole, a competitive inhibitor of the histidine biosynthesis reporter gene (Figure S1). No protein-protein interaction was detected between MpbHLH12 and the third TTG1-like WDR protein, MpWDR3. These data demonstrate that stable protein-protein interactions can form between MpbHLH12 and each of MpMYB14, MpMYB02, MpWDR1, and MpWDR2 in *S. cerevisiae*.

To test whether these protein-protein interactions can also take place in *M. polymorpha*, we performed bimolecular fluorescence complementation using stable transgenic lines. We used MpCYCD;1-nYFP and MpCDKA-cYFP as positive controls for interaction[37] and to test proteins of interest for non-specific interaction. Lines expressing MpbHLH12-nYFP in combination with MpWDR1-cYFP or MpWDR2-cYFP, but not with MpWDR3-cYFP, showed fluorescence of reconstituted YFP in the dorsal thallus surface (Figure 2B). Ubiquitous expression of MpMYB14 and MpMYB02 for this assay resulted in background fluorescence, associated with the production of flavonoids, as reported before.[23] Because of this, samples had to be fixed and cleared prior to imaging. After this treatment, lines expressing MpbHLH12-nYFP and MpMYB14-cYFP showed strong fluorescence in assimilatory filaments (Figure 2B), which was not seen when MpbHLH12-nYFP was swapped for MpCYCD;1-nYFP. No such signal could be clearly distinguished from the background in the lines expressing MpbHLH12-nYFP and MpMYB02-cYFP. In summary, using this method, we could validate *in planta* all interactions found in yeast except for the MpbHLH12-MpMYB02 interaction.

To test whether the *M. polymorpha* proteins that form pairwise bHLH-MYB and bHLH-WDR dimers also form tripartite MBW complexes, we modified our yeast assay to allow us to interrogate the effect of inducible expression of a third protein on the strength of interaction between two others. Noisy growth under the highly restrictive conditions in this assay prevented reliable interpretation for some combinations of proteins tested, but we identified a signal indicating that the presence of either MpWDR1 or MpWDR2 enhances the MpbHLH12-MpMYB14 interaction (Figure 2C). This was consistent across all four biological replicates for MpWDR2 but varied between replicates for MpWDR1 (Figure S1).

The tissue expression domains of *M. polymorpha* orthologs of all three MBW components overlap

Having found that complexes could form when the proteins were experimentally co-expressed, we sought to determine whether they are expressed in the same tissue context. For this purpose, we generated *M. polymorpha* lines carrying transcriptional reporters for Mp*bHLH12*, Mp*MYB02*, Mp*MYB14*, Mp*WDR1*, and Mp*WDR2*. Plants growing from gemmae were imaged by confocal microscopy (Figure 3). The signal from the Mp*MYB14* reporter was widely distributed around the notch regions (Figure 3A). The Mp*MYB02* reporter showed specific expression in oil body cells (Figure 3B), as has been reported independently.[24] Notably, we did not observe signal from the Mp*MYB14*

confocal microscopy of the dorsal thallus surface. Scale bars, 100 μm. The *M. polymorpha* cyclin MpCYCD1 and its interaction partner MpCDKA were used as controls for the method,[37] as they were expected to interact with one another but not with any of the proteins of interest. Due to background fluorescence associated with flavonoid production in the lines expressing the R2R3-MYB proteins ubiquitously, these lines were fixed and cleared before imaging, as indicated.

(C) Conditional expression assay to test the effect of the presence of a third protein on the interaction strength of a dimer. One biological replicate out of four is shown in each instance; the rest are given in Figure S1. In the presence of methionine (++M), the *Met17* promoter is inactive, and growth on -LWH of the yeast carrying AD-MpMYB14 and BD-MpbHLH12 could be abolished with the addition of 5 mM of 3-AT. In the absence of methionine (−M), the *MET17* promoter is active and drives the expression of a *TTG1*-like *WDR* gene. For Mp*WDR1* and Mp*WDR2*, this allowed growth at ≥5 mM 3-AT (although variation was observed between biological replicates for MpWDR1; see Figure S1), whereas the absence of methionine had no effect on the growth of the yeast carrying AD-MpMYB14 and BD-MpbHLH12 without the *MET17* cassette.

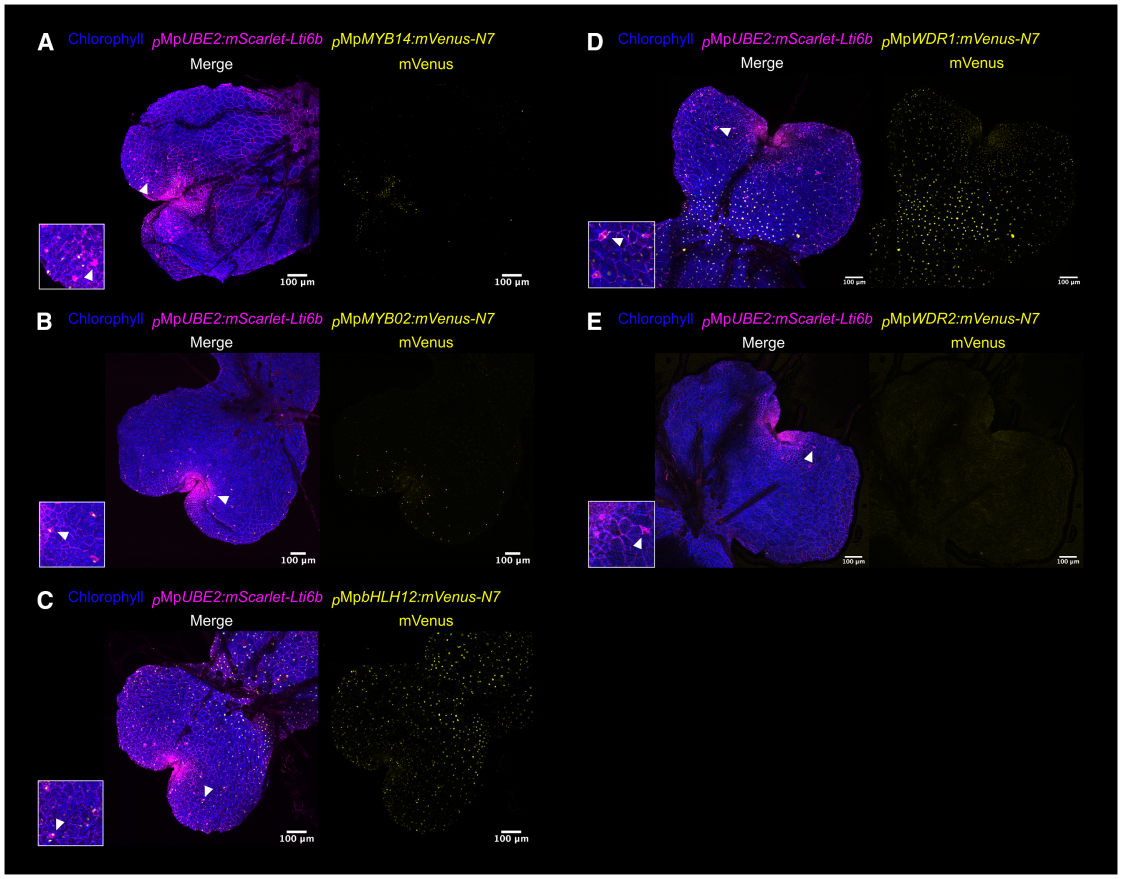


Figure 3. Mp*bHLH12* and Mp*WDR1* are expressed in both of the distinct tissue domains of the two VIII-E *MYB* genes

Confocal micrographs of transcriptional reporter lines for Mp*MYB14* (A), Mp*MYB02* (B), Mp*bHLH12* (C), Mp*WDR1* (D), and Mp*WDR2* (E) at three days after gemma plating. Scale bars, 100 µm. Insets show oil body cells. Gemmae were imaged of lines derived from Cam-1 × Cam-2 spores with two independent transgenes integrated, an mScarlet plasma membrane marker (shown in magenta) and the 5′ cis-regulatory region of each of the genes of interest driving the expression of nuclear-localized mVenus (shown in yellow). Chlorophyll autofluorescence was also captured (shown in blue). Maximum-intensity projections of z stacks were prepared for all three channels and merged. Images are also shown for only the mVenus channel. A translational reporter line was also generated for MpbHLH12 (see Figure S2).

reporter in the oil body cells (Figure 3A), indicating that the expression patterns of these two *R2R3-MYB* paralogs may be mutually exclusive.

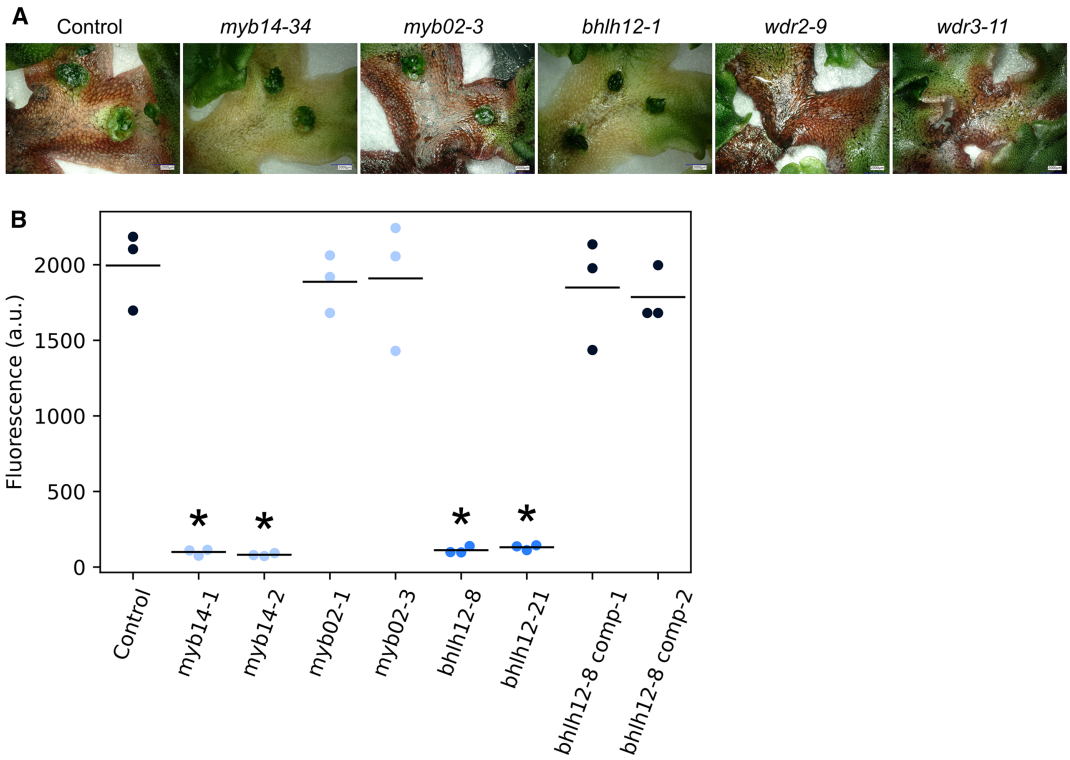
RNA-sequencing has found Mp*bHLH12* to be highly expressed in thallus tissue.[38] Using both a transcriptional (Figure 3C) and a translational reporter (Figure S2), we found that it has a broad domain of expression inclusive of both of the more specific *R2R3-MYB* expression domains, encompassing both the notch regions and oil body cells. Based on these observations, it is possible that MpbHLH12 could dimerize with both MpMYB14 and MpMYB02 in each of their distinct expression domains in *M. polymorpha*.

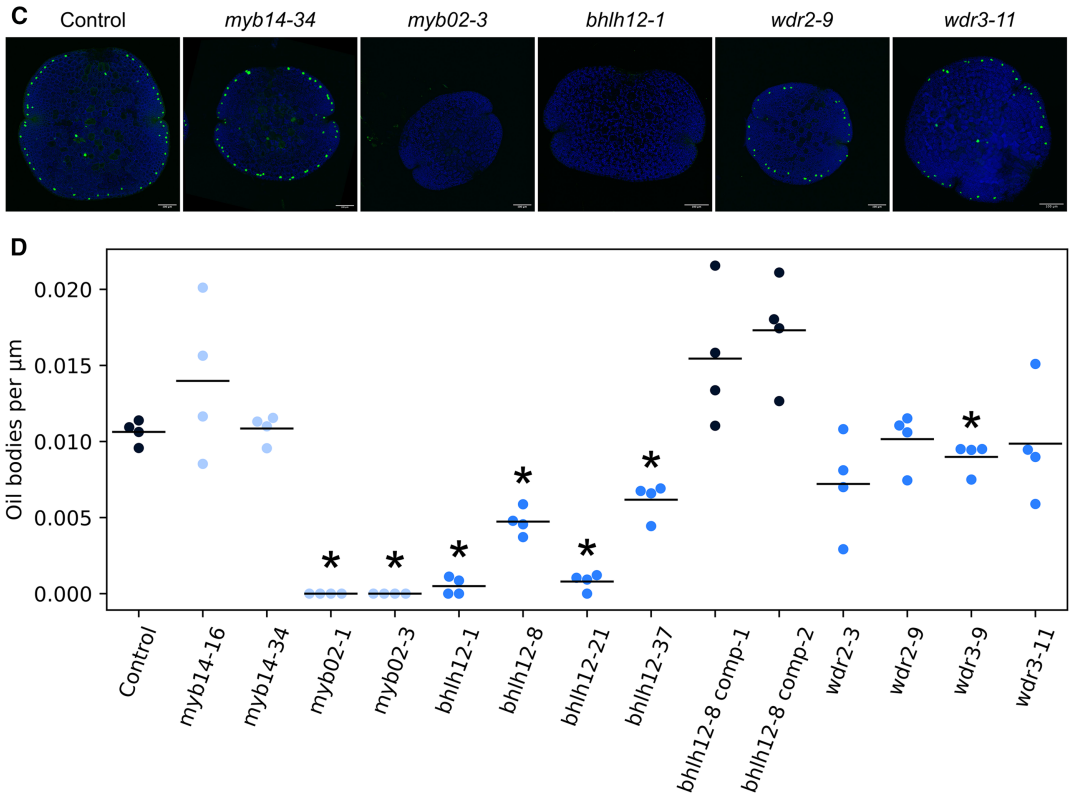
The Mp*WDR1* reporter was strongly and ubiquitously active, including in the notch regions and in oil body cells (Figure 3D). Mp*WDR2* reporter activity was much weaker and could not be clearly discerned in oil body cells (Figure 3E). These observations are consistent with RNA-sequencing data, which additionally found the third homolog, Mp*WDR3*, to be expressed at very low levels in thallus.[38] The Mp*WDR1* expression pattern indicates that MBW trimers could form in *M. polymorpha*.

The single IIIf *bHLH* gene in *M. polymorpha* regulates both pathways regulated by each of the two VIII-E *MYB* genes

As noted, previous studies have demonstrated the requirement for Mp*MYB14* and Mp*MYB02* for auronidin pigment production and oil body maturation, respectively. To test our hypothesis that each of the R2R3-MYB proteins they encode forms a transcriptional complex with MpbHLH12 to carry out each of these functions, we generated gene knockout lines using CRISPR-Cas9 (Figure S3) and scored these phenotypes.

Consistent with reports from other accessions,[20,22] we could induce strong auronidin pigmentation in plants derived from a cross between the Cam-1 and Cam-2 accessions by transferring four-week-old plants to minimal medium, and we found that this response was abolished in several independent Mp*myb14* mutant lines. Figure 4A shows images of these pigmentation phenotypes, measurements carried out of auronidin content in thallus pieces cultured on minimal medium are given in Figure 4B, and images of additional independent lines are given in Figure S4. Pigmentation was not affected in mutants for the





(legend on next page)

paralogous Mp*myb02* (Figures 4A, 4B, and S2). However, auronidin induction was abolished in four independent Mp*bhlh12* mutant lines (Figures 4A, 4B, and S2). Furthermore, auronidin induction could be rescued by the introduction of the $_{pro}$Mp*bHLH12*:Mp*bHLH12-mVenus* transgene to an Mp*bhlh12* mutant line (Figures S2 and 4B). In conclusion, Mp*bHLH12* is required for auronidin pigmentation, as is Mp*MYB14*, whereas Mp*MYB02* does not regulate this pathway.

To assess the maturation of oil bodies in these lines, we stained dormant gemmae with the fluorescent lipid stain BODIPY and imaged them using confocal microscopy. No BODIPY-staining oil bodies were found in our Mp*myb02* mutant lines (Figures 4C and 4D), consistent with the reported lack of mature oil bodies in Mp*myb02* loss-of-function mutants.[24] Oil body counts in mutants for the paralogous Mp*MYB14* did not differ from the control (Figures 4C and 4D). Mp*bhlh12* mutant gemmae exhibited a reduction in the number of mature oil bodies, in some cases lacking them entirely. The severity of this defect varied between mutant lines (Figures 4C and 4D). Introduction of the $_{pro}$Mp*bHLH12*:Mp*bHLH12-mVenus* transgene rescued the number of BODIPY-positive oil bodies in an Mp*bhlh12* mutant line (Figures 4D and S2). In conclusion, Mp*bHLH12* promotes the Mp*MYB02*-dependent maturation of oil bodies. Additionally, we conclude that the two VIII-E *R2R3-MYB* paralogs regulate distinct pathways. Overall, this is consistent with the hypothesis that MpbHLH12 forms transcriptional complexes with each of MpMYB14 and MpMYB02 to promote their activities.

To test whether TTG1-like WDR scaffold proteins are also required for these activities, we set out to generate loss-of-function mutants by targeting the three *M. polymorpha* homologs with CRISPR-Cas9 (Figure S3). Given the data presented here and those of Airoldi et al.,[33] we reasoned that Mp*WDR1* and Mp*WDR2* may redundantly scaffold MBW complexes, whereas the third homolog, Mp*WDR3*, is likely divergent in function. Based on this, we devised strategies to target the Mp*WDR1* and *2* loci simultaneously, all three loci simultaneously, and only Mp*WDR3* (see STAR Methods and Figure S3). We recovered predicted loss-of-function single mutants for Mp*WDR2* and Mp*WDR3*, but not for Mp*WDR1*. Among 80 primary transformants genotyped at the Mp*WDR1* locus, two edited alleles were identified. One harbored a 3-bp deletion, likely to be functionally neutral. Another exhibited a 1 bp deletion causing an early frameshift (Figure S3). This lone putative loss-of-function allele was identified in a line in which a predicted loss-of-function mutation was also detected at Mp*WDR2* and which showed severe growth defects (Figure S4). This Mp*wdr1,2* double mutant line was also the only higher-order mutant line recovered. Since only a single such line was recovered, we could not investigate the effects of *wdr1* mutation independent of genetic background.

The single mutant lines recovered, Mp*wdr2* and Mp*wdr3*, exhibited no marked impairment in auronidin pigmentation nor oil body maturation (Figures 4A, 4C, and 4D). A small, statistically significant reduction in the number of mature oil bodies in gemmae was observed in only one of the two independent *wdr3* mutant lines (Figure 4D). These pathways also putatively appeared to be functional in the lone Mp*wdr1,2* line recovered, which exhibited visible auronidin pigmentation (Figure S4) and BODIPY-staining structures, although these could not be clearly identified in the misshapen putative gemmae sampled from the dorsal surface (Figure S4). In conclusion, we found that neither Mp*WDR2* nor Mp*WDR3* was singly required for the activities of Mp*MYB14* or Mp*MYB02*. For the final homolog, Mp*WDR1*, no predicted loss-of-function single mutants were recovered despite extensive screening. No triple mutants for all homologs and only a lone unreplicated line for one double-mutant combination were recovered. For this reason, we could not evaluate the individual effect of mutation of Mp*WDR1* nor address whether there are functional interactions between the homologs.

DISCUSSION

A heteromeric transcriptional complex is ancestral to land plants

MBW transcriptional complexes regulate a diverse set of pathways in the elaboration of the epidermis of seed plants.[11,13] The *R2R3-MYB* member of the complex has been highlighted as a "hotspot" locus underlying the evolution of pigmentation phenotypes across flowering plants.[3–5] Here, we investigated the origin of the complex and its functional diversification in an independent lineage. It has been known for some time that orthologous genes are encoded by bryophytes, which diverged from seed plants at the first bifurcation within land plants, but incomplete conservation of the canonical interaction motif in the *R3 MYB* repeat[17,18] had been taken to imply that these do not form MBW

Figure 4. Gene knockout of the single IIIf *bHLH* gene causes both the lack of auronidin pigment and the impaired oil body maturation seen in mutants for each of the two VIII-E *MYB* genes

Phenotypes of CRISPR-Cas9 knockout lines are shown. Details about the edits introduced are given in Figure S3. A Cam-1 × Cam-2-derived line bearing an Mp*TOC1*:*LUC* reporter was used as a non-edited transgenic control line, except for (B), where Cam-1 was used. Additionally, lines designated as "comp" carry the $_{pro}$Mp*bHLH12*:Mp*bHLH12-mVenus* transgene with a guide-resistant version of Mp*bHLH12*. Images of the phenotypes of these complemented lines are given in Figure S2. Data points are colored according to genotype, with lines carrying wild-type alleles in dark blue, *bHLH* and *WDR* mutants in medium blue, and *MYB* mutants in light blue.

(A) Auronidin pigmentation was scored in lines grown under nutrient deprivation (see STAR Methods). Scale bars are 2 mm. The same scoring for additional independent lines is shown in Figure S4.

(B) Auronidin content was estimated in thallus pieces grown under nutrient deprivation (see STAR Methods). Each line was compared with the control line using the Mann-Whitney U test, and those with lower values at $p \leq 0.05$ are marked with an asterisk. All data points are shown and the mean values are indicated with horizontal lines. The raw data are given in Data S2.

(C) Oil bodies were counted in dormant gemmae stained with the fluorescent lipid stain BODIPY and imaged with a confocal microscope. One example gemma is shown for each genotype. Scale bars are 100 μm.

(D) The number of BODIPY-staining oil bodies was counted in four gemmae for each line. For each gemma, the count was normalized by the perimeter length to account for differences in gemma size. Each line was compared with the control line using the Mann-Whitney U test, and those with lower values at $p \leq 0.05$ are marked with an asterisk. All data points are shown and the mean values are indicated with horizontal lines. The raw counts and measurements of gemma perimeters and areas are given in Data S3.

complexes.[29] Contrary to this, structural analysis predicted that MpMYB14 and MpMYB02, the orthologous R2R3-MYB proteins in the bryophyte *M. polymorpha*, form a conserved interaction interface with the single orthologous bHLH protein, MpbHLH12. Whereas the predicted MpbHLH12-MpMYB02 dimer was only validated by heterologous expression in yeast, the MpbHLH12-MpMYB14 dimer could also be detected in *M. polymorpha*. Additionally, we validated both in yeast assays and in *M. polymorpha* that MpbHLH12 forms a conserved interaction interface with the TTG1-like WDR proteins MpWDR1 and MpWDR2, but not with the third homolog, MpWDR3. Conditional expression of Mp*WDR1* or *2*increased the interaction strength between MpbHLH12 and MpMYB14 in a yeast assay, indicating that these proteins may engage in cooperative tripartite complexes, as shown for the AtTT2-AtTT8-AtTTG1 MBW complex.[10] Airoldi et al.[33] found that the same homologs, Mp*WDR1* and *2*, but not *3*, could complement *ttg1* mutant phenotypes when heterologously expressed in *A. thaliana*, suggesting that they can also participate in MBW complexes in this context. Taking our new data from *M. polymorpha* into account, we can trace back the origin of the MBW interaction interfaces all the way to the origin of the clade VIII-E *R2R3-MYB* genes and the subclass IIIf *bHLH* genes in the stem lineage of land plants.[19,30]

Using CRISPR-Cas9 gene knockout, we found that Mp*bHLH12* was fully required for activation of the auronidin pigmentation pathway, which requires Mp*MYB14*,[20,21] and partially required for oil body maturation, which requires Mp*MYB02*.[24] This is consistent with the hypothesis that two alternative transcriptional complexes form in each *R2R3-MYB* expression domain, an MpbHLH12-MpMYB14 complex, which regulates the flavonoid pathway, and an MpbHLH12-MpMYB02 complex, which regulates the maturation of oil bodies. A single subclass IIIf *bHLH* gene associated both with the production of flavonoids and bis-bibenzyls, the latter of which accumulate in oil bodies, has been identified in another member of the order Marchantiales, *Plagiochasma appendiculatum*.[39] This indicates that the two proposed transcriptional complexes may be conserved in this lineage. Mp*MYB14* overexpression causes strong accumulation of auronidin pigments,[22,23] but the same is not seen in Mp*bHLH12* overexpression lines.[31] However, the expression of Mp*MYB14* in thallus grown under standard conditions is very low,[38] whereas Mp*bHLH12* is abundant.[38] This is consistent with the hypothesis that Mp*MYB14* and Mp*bHLH12* are reciprocally dependent, but that Mp*MYB14* is the rate-limiting factor under standard conditions. In further support of this, Mp*MYB14* overexpression does not lead to accumulation of auronidin pigments in a *bhlh12* mutant background (Figure S2).

In *A. thaliana*, *TTG1* is required for the activities of the MBW complexes.[12] We found that both Mp*WDR2* and *3* are individually dispensable for auronidin pigmentation and oil body maturation. Mp*WDR1* was the most strongly expressed in the thallus and the only TTG1-like WDR gene for which we could confirm overlap in tissue expression domain with the *bHLH* and *R2R3-MYB* genes. Only a single line with a mutation predicted to disrupt Mp*WDR1* function was recovered, which exhibited a severe growth defect and an additional mutation at Mp*WDR2*. Given that no other *wdr1* allele was identified, we could not test its effect independent of genetic background, including whether mutations at other loci affected the phenotypes analyzed. The triple mutant in *A. thaliana* (*ttg1 lwd1 lwd2*), in addition to a loss of pigments and mispatterning of epidermal cell types, shows complete arrhythmicity of the circadian clock, late flowering, and an aberrant leaf morphology, and Mp*WDR1* and *2* could complement these defects.[33] It seems likely that pleiotropic effects of mutating these genes could cause reduced viability in *M. polymorpha*, hindering the recovery of mutants. In summary, although conservation of the bHLH-WDR interaction interfaces across almost half a billion years of evolution suggests that they have functional importance, their functional role in *M. polymorpha* has yet to be determined.

Other instances of transcriptional complexes conserved since the most recent common ancestor of land plants have been found, including bHLH heterodimers regulating the establishment of the germline[40] and the development of stomata.[41,42] Even more ancient is the reported conserved KNOX-BELL TALE homeodomain heterodimer involved in development in the diploid generation across the Viridiplantae.[43] Binding of the GRAS transcription factor DELLA to members of the same transcription factor families across land plant lineages has also been demonstrated, including interaction of MpDELLA with both MpMYB14[44] and MpMYB02.[45] However, to our knowledge, this study is the first example of ortholog-scale analysis that has identified a conserved transcriptional complex incorporating members of multiple different protein families, highlighting that transcriptional complexes between very distantly related proteins may also be conserved over long evolutionary timescales.

A homologous transcriptional complex regulates convergent flavonoid pigments

Our work demonstrates that homologous MYB-bHLH transcriptional complexes regulate the flavonoid pathway in *M. polymorpha* and in the seed plants. Red pigments are produced from the flavonoid pathway in all major lineages of land plants, with the exception of hornworts.[46] We could not detect any instances of an $LX_6LX_6LX_3R$ R3 MYB motif in the genome of the hornwort *Anthoceros agrestis*,[47] and *A. agrestis* also lacks the IIIf subclass of bHLH proteins.[30] Together, the data are consistent with a single ancestral origin of transcriptional regulation of the flavonoid pathway by a MYB-bHLH complex in land plants, which has been conserved in both seed plants and liverworts, and the correlated loss of these regulators and red pigmentation in the hornwort lineage. However, the red flavonoid pigments produced across lineages are not homologous. While the early steps of the flavonoid biosynthetic pathways are shared, distinct later steps lead to the production of 3-hydroxyanthocyanins in seed plants, 3-deoxyanthocyanins in ferns and mosses, and auronidins in liverworts.[46] As such, the ancestral condition is not yet clear, but, as has been posited elsewhere,[48] it appears likely that a red pigment was produced from the flavonoid pathway in stem land plants. Regardless of the ancestral red pigmentation pathway, if any, all the steps of the flavonoid pathway required for the production of UV-screening flavones have been inferred to have been present in the last common ancestor of land plants.[46]

Lineage-specific repurposing has occurred by an independently repeated mechanism

In addition to the flavonoid pathway, MBW complexes that regulate lineage-specific traits have been identified in multiple taxa.

For instance, in the rosid lineage of angiosperms, MBW complexes regulate the patterning of unicellular trichomes on aerial organs.[49] Here, we find evidence that orthologs in liverworts are involved in the maturation of oil bodies, an innovation in the liverwort lineage. In both cases, the gain of a new function is associated with a lineage-specific duplication of the *R2R3-MYB* member and co-option of its *bHLH* interaction partner. In angiosperms, a duplication gave rise to clade VIII-E *R2R3-MYB* subgroups 6 and 15,[19] and the bHLH protein AtEGL3 regulates both anthocyanin biosynthesis and trichome development by interacting with R2R3-MYB proteins from either subgroup (Zhang et al.[7]). The duplication that gave rise to Mp*MYB14* and Mp*MYB02* appears to be liverwort-specific,[19] and we found that MpbHLH12 regulates both of their functions. As a result of this pattern, the *R2R3-MYB* genes remain pathway-specific, whereas their partners act across multiple pathways. This pleiotropy may constrain their evolution and explain the observed bias toward changes to the *R2R3-MYB* member within the complex.[3–5]

Other examples have been reported of MBW orthologs that may have lineage-specific functions in a number of angiosperms. For example, specific expression of clade *VIII-E R2R3-MYB* genes has been identified in idioblasts of bright eyes (*Catharanthus roseus*)[50] and *Citrus* oil glands.[51] Interestingly, trichomes, oil glands, and oil bodies play analogous roles as secretory structures that allow the accumulation of toxic compounds to defend against herbivory, but they have evolved independently in each lineage.[26] Further, clade VIII-E *R2R3-MYB* genes with conservation of the interaction motif within the R3 MYB repeat are found across land plants, notably in the moss *Physcomitrium patens* and in the lycophyte *Selaginella moellendorffii*.[19] These species also encode subclass IIIf *bHLH* genes[30] and *TTG1*-like *WDR* genes.[33] As such, there is ample opportunity across phylogenetic contexts to test whether the observed patterns in the evolution of these transcriptional complexes are independently repeated.

RESOURCE AVAILABILITY

Lead contact

Requests for resources and further information should be directed to and will be fulfilled by the lead contact, Beverley J. Glover (bjg26@cam.ac.uk).

Materials availability

All plasmids and transgenic lines generated in this study are available from the lead contact upon request.

Data and code availability

Raw phenotyping data related to Figure 4 are given in Data S2 and S3. All data generated in this study are available from the lead contact upon request. This paper does not report any original code. Any additional information required to reanalyze the data is available from the lead contact upon request.

ACKNOWLEDGMENTS

We are grateful to Eva Herrero Serrano for extensive discussions about comparative analysis of protein functions and for establishing in the lab the yeast assays we use in this manuscript. We thank Ignacy Bonter for help with fixing and clearing *M. polymorpha* samples, Aurore Lormet for help with auronidin extractions and for additional help with protocols, we thank Susana Sauret-Güeto, Jenna Rever, Davide Annese, and David Hoey. For discussion of the project as it developed, we want to thank Nathanael Walker-Hale, Sam Brockington, Alex Webb, Hanna Marie Schilbert, Julia Davies, and Sebastien Andreuzza. We are grateful to Matthew Dorling and Qi Wang for support in the lab and on the computing cluster, respectively. This work was funded by the UKRI Natural Environment Research Council doctoral training program studentship 2277410 to T.E.K., the Biotechnology and Biological Sciences Research Council grant BB/T007117/1 to J.H., and the BBSRC/EPSRC OpenPlant Synthetic Biology Research Centre BB/L014130/1 to J.H. F.R. is a Leverhulme Early Career Fellow.

AUTHOR CONTRIBUTIONS

T.E.K. led the project, analyzed the data, and wrote the paper with input from all authors. T.E.K. and F.R. performed the experiments. C.A.A. cloned genes for expression in yeast, performed preliminary experiments, and provided training and supervision to T.E.K. J.H. shared protocols and resources and acquired funding. B.J.G. supervised the project and acquired funding.

DECLARATION OF INTERESTS

B.J.G. is a member of the *Current Biology* advisory board.

STAR★METHODS

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

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Analysis of protein sequences and structures
  - Cloning genes from *M. polymorpha*
  - Plasmid construction for expression in *S. cerevisiae*
  - Genetic transformation of *S. cerevisiae*
  - *S. cerevisiae* spot testing
  - Genetic transformation of *M. polymorpha*
  - Plasmid construction for bimolecular fluorescence complementation in *M. polymorpha*
  - Bimolecular fluorescence complementation in *M. polymorpha*
  - Generation and imaging of *M. polymorpha* reporter lines
  - Plasmid construction for genome editing in *M. polymorpha*
  - Phenotypic analysis of *M. polymorpha*
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2026.04.031.

Received: December 19, 2024
Revised: February 4, 2026
Accepted: April 14, 2026
Published: May 8, 2026

REFERENCES

1. Orteu, A., and Jiggins, C.D. (2020). The genomics of coloration provides insights into adaptive evolution. Nat. Rev. Genet. *21*, 461–475. https://doi.org/10.1038/s41576-020-0234-z.
2. Stern, D.L., and Orgogozo, V. (2009). Is genetic evolution predictable? Science *323*, 746–751. https://doi.org/10.1126/science.1158997.
3. Sobel, J.M., and Streisfeld, M.A. (2013). Flower color as a model system for studies of plant evo-devo. Front. Plant Sci. *4*, 321. https://doi.org/10.3389/fpls.2013.00321.
4. Wheeler, L.C., Walker, J.F., Ng, J., Deanna, R., Dunbar-Wallis, A., Backes, A., Pezzi, P.H., Palchetti, M.V., Robertson, H.M., Monaghan, A., et al. (2022). Transcription factors evolve faster than their structural gene targets

in the flavonoid pigment pathway. Mol. Biol. Evol. *39*, msac044. https://doi.org/10.1093/MOLBEV/MSAC044.

5. Marin-Recinos, M.F., and Pucker, B. (2024). Genetic factors explaining anthocyanin pigmentation differences. BMC Plant Biol. *24*, 627. https://doi.org/10.1186/s12870-024-05316-w.
6. Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P.A., and Saedler, H. (1987). The regulatory c1 locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. *6*, 3553–3558. https://doi.org/10.1002/j.1460-2075.1987.tb02684.x.
7. Zhang, F., Gonzalez, A., Zhao, M., Payne, C.T., and Lloyd, A. (2003). A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. Development *130*, 4859–4869. https://doi.org/10.1242/dev.00681.
8. Bernhardt, C., Lee, M.M., Gonzalez, A., Zhang, F., Lloyd, A., and Schiefelbein, J. (2003). The bHLH genes GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3) specify epidermal cell fate in the Arabidopsis root. Development *130*, 6431–6439. https://doi.org/10.1242/dev.00880.
9. Berger, F., Linstead, P., Dolan, L., and Haseloff, J. (1998). Stomata patterning on the hypocotyl of Arabidopsis thaliana is controlled by genes involved in the control of root epidermis patterning. Dev. Biol. *194*, 226–234. https://doi.org/10.1006/dbio.1997.8836.
10. Baudry, A., Heim, M.A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in Arabidopsis thaliana. Plant J. *39*, 366–380. https://doi.org/10.1111/j.1365-313X.2004.02138.x.
11. Ramsay, N.A., and Glover, B.J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci. *10*, 63–70. https://doi.org/10.1016/j.tplants.2004.12.011.
12. 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*, 1337–1350. https://doi.org/10.1105/tpc.11.7.1337.
13. Nemesio-Gorriz, M., Blair, P.B., Dalman, K., Hammerbacher, A., Arnerup, J., Stenlid, J., Mukhtar, S.M., and Elfstrand, M. (2017). Identification of Norway spruce MYB-bHLH-WDR transcription factor complex members linked to regulation of the flavonoid pathway. Front. Plant Sci. *8*, 305. https://doi.org/10.3389/fpls.2017.00305.
14. Roussel, M., Saule, S., Lagrou, C., Rommens, C., Beug, H., Graf, T., and Stehelin, D. (1979). Three new types of viral oncogene of cellular origin specific for haematopoietic cell transformation. Nature *281*, 452–455. https://doi.org/10.1038/281452a0.
15. Murre, C., McCaw, P.S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777–783. https://doi.org/10.1016/0092-8674(89)90682-X.
16. Smith, T.F., Gaitatzes, C., Saxena, K., and Neer, E.J. (1999). The WD repeat: A common architecture for diverse functions. Trends Biochem. Sci. *24*, 181–185. https://doi.org/10.1016/S0968-0004(99)01384-5.
17. Zimmermann, I.M., Heim, M.A., Weisshaar, B., and Uhrig, J.F. (2004). Comprehensive identification of Arabidopsis thaliana MYB transcription factors interacting with R/B-like BHLH proteins. Plant J. *40*, 22–34. https://doi.org/10.1111/j.1365-313X.2004.02183.x.
18. 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 *97*, 13579–13584. https://doi.org/10.1073/pnas.250379897.
19. Jiang, C.-K., and Rao, G.-Y. (2020). Insights into the Diversification and Evolution of R2R3-MYB Transcription Factors in Plants. Plant Physiol. *183*, 637–655. https://doi.org/10.1104/pp.19.01082.
20. Albert, N.W., Thrimawithana, A.H., McGhie, T.K., Clayton, W.A., Deroles, S.C., Schwinn, K.E., Bowman, J.L., Jordan, B.R., and Davies, K.M. (2018). Genetic analysis of the liverwort Marchantia polymorpha reveals that R2R3MYB activation of flavonoid production in response to abiotic stress is an ancient character in land plants. New Phytol. *218*, 554–566. https://doi.org/10.1111/nph.15002.
21. Berland, H., Albert, N.W., Stavland, A., Jordheim, M., McGhie, T.K., Zhou, Y., Zhang, H., Deroles, S.C., Schwinn, K.E., Jordan, B.R., et al. (2019). Auronidins are a previously unreported class of flavonoid pigments that challenges when anthocyanin biosynthesis evolved in plants. Proc. Natl. Acad. Sci. USA *116*, 20232–20239. https://doi.org/10.1073/pnas.1912741116.
22. Carella, P., Gogleva, A., Hoey, D.J., Bridgen, A.J., Stolze, S.C., Nakagami, H., and Schornack, S. (2019). Conserved biochemical defenses underpin host responses to oomycete infection in an early-divergent land plant lineage. Curr. Biol. *29*, 2282–2294.e5. https://doi.org/10.1016/J.CUB.2019.05.078.
23. Kubo, H., Nozawa, S., Hiwatashi, T., Kondou, Y., Nakabayashi, R., Mori, T., Saito, K., Takanashi, K., Kohchi, T., and Ishizaki, K. (2018). Biosynthesis of riccionidins and marchantins is regulated by R2R3-MYB transcription factors in Marchantia polymorpha. J. Plant Res. *131*, 849–864. https://doi.org/10.1007/s10265-018-1044-7.
24. Wang, L., Wan, M.-C., Liao, R.-Y., Xu, J., Xu, Z.-G., Xue, H.-C., Mai, Y.-X., and Wang, J.-W. (2023). The maturation and aging trajectory of Marchantia polymorpha at single-cell resolution. Dev. Cell *58*, 1429–1444.e6. https://doi.org/10.1016/j.devcel.2023.05.014.
25. Kanazawa, T., Morinaka, H., Ebine, K., Shimada, T.L., Ishida, S., Minamino, N., Yamaguchi, K., Shigenobu, S., Kohchi, T., Nakano, A., et al. (2020). The liverwort oil body is formed by redirection of the secretory pathway. Nat. Commun. *11*, 6152. https://doi.org/10.1038/s41467-020-19978-1.
26. Romani, F., Flores, J.R., Tolopka, J.I., Suárez, G., He, X., and Moreno, J.E. (2022). Liverwort oil bodies: diversity, biochemistry, and molecular cell biology of the earliest secretory structure of land plants. J. Exp. Bot. *73*, 4427–4439. https://doi.org/10.1093/jxb/erac134.
27. Tanaka, M., Esaki, T., Kenmoku, H., Koeduka, T., Kiyoyama, Y., Masujima, T., Asakawa, Y., and Matsui, K. (2016). Direct evidence of specific localization of sesquiterpenes and marchantin A in oil body cells of Marchantia polymorpha L. Phytochemistry *130*, 77–84. https://doi.org/10.1016/j.phytochem.2016.06.008.
28. Romani, F., Banić, E., Florent, S.N., Kanazawa, T., Goodger, J.Q.D., Mentink, R.A., Dierschke, T., Zachgo, S., Ueda, T., Bowman, J.L., et al. (2020). Oil body formation in Marchantia polymorpha is controlled by MpC1HDZ and serves as a defense against arthropod herbivores. Curr. Biol. *30*, 2815–2828.e8. https://doi.org/10.1016/j.cub.2020.05.081.
29. Davies, K.M., Jibran, R., Zhou, Y., Albert, N.W., Brummell, D.A., Jordan, B.R., Bowman, J.L., and Schwinn, K.E. (2020). The evolution of flavonoid biosynthesis: a bryophyte perspective. Front. Plant Sci. *11*, 7. https://doi.org/10.3389/fpls.2020.00007.
30. Kongsted, T.E., and Glover, B.J. (2023). Phylogenetic analysis of bHLH classes III and IV in land plants and their algal relatives. New Phytol. *240*, 1717–1721. https://doi.org/10.1111/nph.19202.
31. Arai, H., Yanagiura, K., Toyama, Y., and Morohashi, K. (2019). Genome-wide analysis of MpBHLH12, a IIIf basic helix-loop-helix transcription factor of Marchantia polymorpha. J. Plant Res. *132*, 197–209. https://doi.org/10.1007/s10265-019-01095-w.
32. De Vetten, N., Quattrocchio, F., Mol, J., and Koes, R. (1997). The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev. *11*, 1422–1434. https://doi.org/10.1101/gad.11.11.1422.
33. Airoldi, C.A., Hearn, T.J., Brockington, S.F., Webb, A.A.R., and Glover, B.J. (2019). TTG1 proteins regulate circadian activity as well as epidermal cell fate and pigmentation. Nat. Plants *5*, 1145–1153. https://doi.org/10.1038/s41477-019-0544-3.

34. Wu, J.F., Wang, Y., and Wu, S.H. (2008). Two new clock proteins, LWD1 and LWD2, regulate arabidopsis photoperiodic flowering. Plant Physiol. *148*, 948–959. https://doi.org/10.1104/pp.108.124917.

35. Lee, M.M., and Schiefelbein, J. (2001). Developmentally distinct MYB genes encode functionally equivalent proteins in Arabidopsis. Development *128*, 1539–1546. https://doi.org/10.1242/dev.128.9.1539.

36. Wang, B., Luo, Q., Li, Y., Du, K., Wu, Z., Li, T., Shen, W.-H., Huang, C.-H., Gan, J., and Dong, A. (2022). Structural insights into partner selection for MYB and bHLH transcription factor complexes. Nat. Plants *8*, 1108–1117. https://doi.org/10.1038/s41477-022-01223-w.

37. Romani, F., Bonter, I., Rebmann, M., Takahashi, G., Guzman-Chavez, F., De Batté, F., Hirakawa, Y., and Haseloff, J. (2026). A simple cell-cycle control system in Marchantia polymorpha provides a framework for understanding plant cell proliferation. Plant Cell, koag103. https://doi.org/10.1093/plcell/koag103.

38. Kawamura, S., Romani, F., Yagura, M., Mochizuki, T., Sakamoto, M., Yamaoka, S., Nishihama, R., Nakamura, Y., Yamato, K.T., Bowman, J.L., et al. (2022). MarpolBase Expression: A web-based, comprehensive platform for visualization and analysis of transcriptomes in the liverwort Marchantia polymorpha. Plant Cell Physiol. *63*, 1745–1755. https://doi.org/10.1093/pcp/pcac129.

39. Wu, Y.F., Zhao, Y., Liu, X.Y., Gao, S., Cheng, A.X., and Lou, H.X. (2018). A bHLH transcription factor regulates bisbibenzyl biosynthesis in the liverwort Plagiochasma appendiculatum. Plant Cell Physiol. *59*, 1187–1199. https://doi.org/10.1093/pcp/pcy053.

40. Saito, M., Momiki, R., Ebine, K., Yoshitake, Y., Nishihama, R., Miyakawa, T., Nakano, T., Mitsuda, N., Araki, T., Kohchi, T., et al. (2023). A bHLH heterodimer regulates germ cell differentiation in land plant gametophytes. Curr. Biol. *33*, 4980–4987.e6. https://doi.org/10.1016/j.cub.2023.09.020.

41. Chater, C.C., Caine, R.S., Tomek, M., Wallace, S., Kamisugi, Y., Cuming, A.C., Lang, D., MacAlister, C.A., Casson, S., Bergmann, D.C., et al. (2016). Origin and function of stomata in the moss Physcomitrella patens. Nat. Plants *2*, 16179. https://doi.org/10.1038/nplants.2016.179.

42. Harris, B.J., Harrison, C.J., Hetherington, A.M., and Williams, T.A. (2020). Phylogenomic evidence for the monophyly of bryophytes and the reductive evolution of stomata. Curr. Biol. *30*, 2001–2012.e2. https://doi.org/10.1016/j.cub.2020.03.048.

43. Dierschke, T., Flores-Sandoval, E., Rast-Somssich, M.I., Althoff, F., Zachgo, S., and Bowman, J.L. (2021). Gamete expression of TALE class HD genes activates the diploid sporophyte program in Marchantia polymorpha. eLife *10*, e57088. https://doi.org/10.7554/eLife.57088.

44. Hernández-García, J., Serrano-Mislata, A., Lozano-Quiles, M., Úrbez, C., Nohales, M.A., Blanco-Touriñán, N., Peng, H., Ledesma-Amaro, R., and Blázquez, M.A. (2024). DELLA proteins recruit the Mediator complex subunit MED15 to coactivate transcription in land plants. Proc. Natl. Acad. Sci. USA *121*, e2319163121. https://doi.org/10.1073/pnas.2319163121.

45. Briones-Moreno, A., Hernández-García, J., Vargas-Chávez, C., Blanco-Touriñán, N., Phokas, A., Úrbez, C., Cerdán, P.D., Coates, J.C., Alabadí, D., and Blázquez, M.A. (2023). DELLA functions evolved by rewiring of associated transcriptional networks. Nat. Plants *9*, 535–543. https://doi.org/10.1038/s41477-023-01372-6.

46. Davies, K.M., Landi, M., van Klink, J.W., Schwinn, K.E., Brummell, D.A., Albert, N.W., Chagné, D., Jibran, R., Kulshrestha, S., Zhou, Y., et al. (2022). Evolution and function of red pigmentation in land plants. Ann. Bot. *130*, 613–636. https://doi.org/10.1093/aob/mcac109.

47. Li, F.W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N., Barker, M.S., Bennett, T., Blázquez, M.A., et al. (2020). Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat. Plants *6*, 259–272. https://doi.org/10.1038/s41477-020-0618-2.

48. Bowman, J.L. (2022). The origin of a land flora. Nat. Plants *8*, 1352–1369. https://doi.org/10.1038/s41477-022-01283-y.

49. Serna, L., and Martin, C. (2006). Trichomes: different regulatory networks lead to convergent structures. Trends Plant Sci. *11*, 274–280. https://doi.org/10.1016/j.tplants.2006.04.008.

50. Li, C., Colinas, M., Wood, J.C., Vaillancourt, B., Hamilton, J.P., Jones, S.L., Caputi, L., O'Connor, S.E., and Buell, C.R. (2025). Cell-type-aware regulatory landscapes governing monoterpene indole alkaloid biosynthesis in the medicinal plant Catharanthus roseus. New Phytol. *245*, 347–362. https://doi.org/10.1111/nph.20208.

51. Wang, H., Ren, J., Zhou, S., Duan, Y., Zhu, C., Chen, C., Liu, Z., Zheng, Q., Xiang, S., Xie, Z., et al. (2024). Molecular regulation of oil gland development and biosynthesis of essential oils in Citrus spp. Science *383*, 659–666. https://doi.org/10.1126/science.adl2953.

52. Delmans, M., Pollak, B., and Haseloff, J. (2017). MarpoDB: An open registry for Marchantia polymorpha genetic parts. Plant Cell Physiol. *58*, e5. https://doi.org/10.1093/pcp/pcw201.

53. Sauret-Güeto, S., Frangedakis, E., Silvestri, L., Rebmann, M., Tomaselli, M., Markel, K., Delmans, M., West, A., Patron, N.J., and Haseloff, J. (2020). Systematic tools for reprogramming plant gene expression in a simple model, Marchantia polymorpha. ACS Synth. Biol. *9*, 864–882. https://doi.org/10.1021/acssynbio.9b00511.

54. Tse, S.W., Annese, D., Romani, F., Guzman-Chavez, F., Bonter, I., Forestier, E., Frangedakis, E., and Haseloff, J. (2024). Optimizing promoters and subcellular localization for constitutive transgene expression in Marchantia polymorpha. Plant Cell Physiol. *65*, 1298–1309. https://doi.org/10.1093/pcp/pcae063.

55. Romani, F., Sauret-Güeto, S., Rebmann, M., Annese, D., Bonter, I., Tomaselli, M., Dierschke, T., Delmans, M., Frangedakis, E., Silvestri, L., et al. (2024). The landscape of transcription factor promoter activity during vegetative development in Marchantia. Plant Cell *36*, 2140–2159. https://doi.org/10.1093/plcell/koae053.

56. Čermák, T., Curtin, S.J., Gil-Humanes, J., Čegan, R., Kono, T.J.Y., Konečná, E., Belanto, J.J., Starker, C.G., Mathre, J.W., Greenstein, R.L., et al. (2017). A multipurpose toolkit to enable advanced genome engineering in plants. Plant Cell *29*, 1196–1217. https://doi.org/10.1105/tpc.16.00922.

57. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

58. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

59. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w.

60. Elfmann, C., and Stülke, J. (2023). PAE viewer: a webserver for the interactive visualization of the predicted aligned error for multimer structure predictions and crosslinks. Nucleic Acids Res. *51*, W404–W410. https://doi.org/10.1093/nar/gkad350.

61. Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. *16*, 276–277. https://doi.org/10.1016/S0168-9525(00)02024-2.

62. Bittrich, S., Segura, J., Duarte, J.M., Burley, S.K., and Rose, Y. (2024). RCSB protein Data Bank: exploring protein 3D similarities via comprehensive structural alignments. Bioinformatics *40*, btae370. https://doi.org/10.1093/bioinformatics/btae370.

63. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

64. Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. *47*, W636–W641. https://doi.org/10.1093/nar/gkz268.

65. Solow, S.P., Sengbusch, J., and Laird, M.W. (2005). Heterologous protein production from the inducible MET25 promoter in Saccharomyces cerevisiae. Biotechnol. Prog. *21*, 617–620. https://doi.org/10.1021/bp049916q.

66. Annese, D., Romani, F., Grandellis, C., Ives, L., Frangedakis, E., Buson, F.X., Molloy, J.C., and Haseloff, J. (2025). Semi-automated workflow for high-throughput Agrobacterium-mediated plant transformation. Plant J. *122*, e70118. https://doi.org/10.1111/tpj.70118.

67. Frangedakis, E., Tomaselli, M., Rebmann, M., and Sauret-Gueto, S. (2020). Marchantia genotyping (quick and dirty genomic DNA extraction). Protocols.io. https://doi.org/10.17504/protocols.io.bcmwiu7e.

68. Pollak, B., Cerda, A., Delmans, M., Álamos, S., Moyano, T., West, A., Gutiérrez, R.A., Patron, N.J., Federici, F., and Haseloff, J. (2019). Loop assembly: a simple and open system for recursive fabrication of DNA circuits. New Phytol. *222*, 628–640. https://doi.org/10.1111/nph.15625.

69. Sakamoto, Y., Ishimoto, A., Sakai, Y., Sato, M., Nishihama, R., Abe, K., Sano, Y., Furuichi, T., Tsuji, H., Kohchi, T., et al. (2022). Improved clearing method contributes to deep imaging of plant organs. Commun. Biol. *5*, 12. https://doi.org/10.1038/s42003-021-02955-9.

70. Patron, N.J., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., Raitskin, O., Leveau, A., Farré, G., Rogers, C., et al. (2015). Standards for plant synthetic biology: a common syntax for exchange of DNA parts. New Phytol. *208*, 13–19. https://doi.org/10.1111/nph.13532.

71. Lormet, A., Haseloff, J., and Romani, F. (2026). Engineering auronidin production in Marchantia polymorpha enables a new class of plant-derived textile pigments. Preprint at bioRxiv. https://doi.org/10.64898/2026.02.07.704544.

72. Kluyver, T., Ragan-Kelley, B., P., Granger Brian, Bussonnier Matthias, Frederic Jonathan, Kelley Kyle, Hamrick Jessica, Grout Jason, Corlay Sylvain, et al. (2016). Jupyter Notebooks – a publishing format for reproducible computational workflows. In Positioning and Power in Academic Publishing: Players, Agents and Agendas (IOS Press), pp. 87–90. https://doi.org/10.3233/978-1-61499-649-1-87.

73. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2.

74. Waskom, M. (2021). seaborn: statistical data visualization. JOSS *6*, 3021. https://doi.org/10.21105/joss.03021.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
Escherichia coli DH5α	N/A	N/A
Agrobacterium tumefaciens GV3101	N/A	N/A
Chemicals, peptides, and recombinant proteins		
Gamborg B5 medium including vitamins	Duchefa Biochemie	Cat# G0210
Phytoagar	Duchefa Biochemie	Cat# P1003
Phusion High-Fidelity DNA polymerase	New England Biolabs	Cat# M0530
KOD Xtreme Hot Start Polymerase	Sigma	Cat# 71975-M
Q5 High-Fidelity DNA polymerase	New England Biolabs	Cat# M0491S
Restriction enzymes	New England Biolabs	N/A
iTOMEi-D	Tokyo Chemical Industry	# T3940
BODIPY 493/503	Invitrogen	Cat# D3922
Critical commercial assays		
PureLink Quick Gel Extraction Kit	Invitrogen	Cat# K210012
Gateway BP Clonase II Enzyme Mix	Invitrogen	Cat# 11789020
Gateway LR Clonase II Enzyme Mix	Invitrogen	Cat# 11791020
HiFi DNA Assembly Master Mix	New England Biolabs	Cat# E2621
GeneJet PCR purification kit	Thermo Scientific	Cat# K0701
Monarch PCR & DNA Cleanup Kit	New England Biolabs	Cat# T1030
Experimental models: Organisms/strains		
Saccharomyces cerevisiae AH109	Clontech	N/A
Saccharomyces cerevisiae Y2HGold	Takara Bio	Cat# 630498
Marchantia polymorpha Cam-1	Delmans et al.[52]	N/A
Marchantia polymorpha Cam-2	Delmans et al.[52]	N/A
Oligonucleotides		
See Data S1	This paper	N/A
Recombinant DNA		
OpenPlant kit	Sauret-Güeto et al.[53]	#1000000272
pDonr201	Invitrogen	Cat# 11789-014
p406MET25	Addgene	#17421
pAS2	Clontech	N/A
pACT2	Clontech	N/A
pUAP1	Addgene	#63674

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pICSL50003	Addgene	#245577
pICSL50002	Addgene	#245576
pBy_12	Tse et al.[54]	N/A
pBy_20	Tse et al.[54]	N/A
pBy_01	Romani et al.[55]	N/A
pBy_10	Tse et al.[54]	N/A
Csy4-P2A	Čermák et al.[56]	N/A
pMODB2112	Čermák et al.[56]	#91064
Software and algorithms		
MAFFT v7	Katoh and Standley[57]	N/A
Jalview v2.11.2.7	Waterhouse et al.[58]	N/A
Alphafold 3	Abramson et al.[59]	N/A
PAE viewer	Elfman and Stülke[60]	N/A
EMBOSS NEEDLE	Rice et al.[61]	N/A
RCSB PDB Pairwise Structure Alignment	Bittrich et al.[62]	N/A
Fiji for ImageJ 2.9.0	Schindelin et al.[63]	N/A
Other		
384-well plate, polypropylene	Greiner	Cat# 781209

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Marchantia polymorpha plants deriving from crosses of the accessions, Cam-1 (male) and Cam-2 (female) were maintained on Gamborg plates (2 g/L Gamborg B5 including vitamins, pH 5.8, 5 g/L Phytoagar) in a growth room in long-day conditions at 22-23°C, except where otherwise stated.

METHOD DETAILS

Analysis of protein sequences and structures

R2R3-MYB protein sequences were aligned with MAFFT[57] and visualised with Jalview.[58] Structural predictions of protein complexes were generated with Alphafold 3[59] and predicted aligned error plots were generated with PAE viewer.[60] Homologous regions were identified by alignment with EMBOSS NEEDLE.[61,64] Structural alignments were performed with the RCSB PDB Pairwise Structure Alignment tool[62] with default settings.

Cloning genes from *M. polymorpha*

Mp*bHLH12* (MarpolBase: Mp2g05070), Mp*MYB14* (Mp5g19050), Mp*MYB02* (Mp3g07510), Mp*WDR1* (Mp5g03900), Mp*WDR2* (Mp5g01830) and Mp*WDR3* (Mp5g06230) were amplified from cDNA by Phusion HF two-step PCR (primers AC11-14, 39-40, 72-75, 191, 192), and attB sites were added by another round of Phusion HF two-step PCR (primers AC87, 88). Amplicons were gel purified and integrated into the Gateway entry vector pDonr201. Presence of the desired insert was confirmed by Sanger sequencing (primers attL1 and attL2).

Plasmid construction for expression in *S. cerevisiae*

The genes of interested cloned in pDonr201 were recombined into Gateway-modified versions of the destination vectors pAS2 and pACT2. To allow inducible expression of a third protein, pAS2-Mp*bHLH12* was modified as follows. Mp*WDR1*, *2* and *3* were amplified from the pDonr201 entry clones by Phusion HF two-step PCR (primers TK9-14), and the methionine-repressible *MET17* promoter[65] and the *CYC1* terminator from p406 (primers TK15-18). pAS2-Mp*bHLH12* was digested with AflII and MluI-HF and gel purified. The parts were integrated into the digested vector by Gibson assembly with the HiFi Assembly Master Mix. The final assemblies were verified by restriction enzyme digestion and by Sanger sequencing (primers TK50 and M13F).

Genetic transformation of *S. cerevisiae*

Yeast media were prepared according to the Yeast Protocols Handbook (2009, Clontech Laboratories Inc.), except as stated. *Saccharomyces cerevisiae* strain AH109 was co-transformed with two plasmids, a pAS2 destination clone and a pACT2 destination

clone, by small-scale LiAC yeast transformation (as in Yeast Protocols Handbook (2009, Clontech Laboratories Inc.), except YPDA with higher adenine content (40 mg/L) was used for the overnight cultures, and in step 8 pellets were resuspended directly in TE/LiAC). Transformants were selected by growth on -LW dropout medium at 30°C for 5-6 days.

S. cerevisiae spot testing

When yeast colonies had grown to a size of ~1-2 cm on the selective plates, spot tests were performed as follows. Each colony picked was inoculated in 200 μL sterile 1X TE in wells of a sterile microtitre plate. 10 μL was transferred from the inoculated wells to clean wells of 200 μL sterile TE and mixed, and 10 μL was transferred from each of these wells to clean wells of 200 μL sterile TE, to obtain two serially diluted replicates for each colony. 3 μL was plated from each well onto the amino acid dropout media plates. The media used were -LW, -LWH, - LWHA, and -LWH and where applies were supplemented with 3-amino-1,2,4-triazole (3-AT) for the desired concentration. Spot test plates were grown at 30°C for 4-5 days and imaged with a Canon EOS 500D. As our lab-maintained AH109 strain was found to require methionine for growth, a fresh stock of Takara Matchmaker Y2HGold was maintained on modified -M medium (with 40 mg/L adenine and 100 mg/L aspartic acid and pH adjusted to 5.8), for use with the methionine-repressible expression system. For this, co-transformation was carried out as above, with the following exceptions. In preparation for transformation, the strain was grown on the modified -M medium for ~7 days. ~10-18 colonies (each ~1-2 mm in diameter) were inoculated in liquid modified -M medium and grown at 30°C and 180 rpm for ~24 hours. Transformants were selected on -LWM (with 40 mg/L adenine and 100 mg/L aspartic acid and pH adjusted to 5.8). Four colonies were picked from these plates, streaked out individually on the same selective medium and grown for six days before spot tests. The media used for spot tests in this case were -LWHM (with 40 mg/L adenine and 100 mg/L aspartic acid), and -LWH++M (with 40 mg/L adenine and 300 mg/L methionine). These spot test plates were grown at 30°C for 6 days before imaging with a Canon EOS 500D.

Genetic transformation of *M. polymorpha*

Transformation of *M. polymorpha* sporelings, deriving from silica-dried archegoniophores carrying Cam-1 x Cam-2 spores, was carried out according to Annese et al.[66] Sporelings were selected on Gamborg plates supplemented with 100 μg/mL cefotaxime and 0.5 μM chlorsulfuron and/or 20 μg/mL hygromycin. Primary transformants were transferred to new selective medium with 0.5% sucrose to encourage gemma cup production. Isogenic lines were established by culture of gemmae on selective medium. DNA extractions for genotyping were performed according to Frangedakis et al.[67] Samples were stored at 4°C and 5 μL was used as the template for PCR amplification with KOD Xtreme Hot Start polymerase (primers TK80-81, 83-85, 90-94, 98, 99, 102, 103, 134-136). For gel assessment only, a positive plasmid template control and a negative water template control were included. For sequencing, the remaining PCR product was purified with the GeneJet PCR purification kit, eluted in 20 μL DI water and sent for Sanger sequencing.

Plasmid construction for bimolecular fluorescence complementation in *M. polymorpha*

Mp*WDR1*, *2* and *3* were amplified from the pDonr201 entry clones. For compatibility with LOOP assembly,[68] CDS12 overhangs were added by PCR using the Q5 High-Fidelity DNA polymerase and the Monarch PCR & DNA Cleanup Kit. Mp*MYB02* and Mp*MYB14* were synthesised as CDS12 parts directly and cloned in pUAP1 by Genewiz. Parts were cloned into the pBy_12 backbone as in Tse et al.[54] with *cYFP* (pICSL50003) or in the pBy_20 backbone with OP-49 (PROM5_35S), OP-53 (3TERM_Nos), and *nYFP* (pICSL50002). Mp*CYCD;1* and Mp*CDKA* CDS12 parts were used as positive and negative controls.[37]

Bimolecular fluorescence complementation in *M. polymorpha*

M. polymorpha sporelings were transformed with two *A. tumefaciens* strains harbouring each plasmid and transformants were selected by growth on plates containing hygromycin and chlorsulfuron after two weeks. >12 independent primary transformants were screened for fluorescence of reconstituted YFP. Representative images were taken using a Leica SP8 confocal microscope with excitation at 515 nm and emission 522-540 nm for YFP and 687-739 nm for chlorophyll fluorescence. For lines expressing Mp*MYB02* or Mp*MYB14*, autofluorescence caused by pigment production overlaps with the YFP signal, creating strong background. For this reason, these samples were fixed and pigments cleared using iTOMEi-D.[69,70]

Generation and imaging of *M. polymorpha* reporter lines

pMp*MYB14*, pMp*MYB02*, pMp*bHLH12* and 5UTR_Mp*bHLH12* were cloned in the pBy_01 binary vector as described in Romani et al.[55] For pMp*WDR1* and pMp*WDR2*, 3 kb of sequence 5′ to the ATG, domesticated to remove internal BsaI and SapI sites, was synthesised by Genewiz and cloned in a pUAP1 backbone, then subcloned in the pBy_01 binary vector using BsaI as described in Romani et al.[55] For genetic complementation and to generate a translational reporter, a gRNA-resistant and domesticated version of the Mp*bHLH12* CDS was synthesised by Genewiz and cloned in a pUAP1 backbone, then subcloned into a pBy_10 backbone using BsaI[54] together with pMp*bHLH12*, 5UTR_Mp*bHLH12*, OP-044 (CTAG_mVenus) and OP-053 (3TERM_Nos). The final assemblies were confirmed by restriction enzyme digestion and Sanger sequencing. All constructs were introduced into Cam-1 x Cam-2 sporelings, and $_{pro}$Mp*bHLH12*:Mp*bHLH12-mVenus*, as well as p*35S*:Mp*MYB14-mVenus* (pBy_10)[71], also to the Mp*bhlh12-8* line by thallus transformation (Kubota et al., 2013), and lines were screened for mVenus fluorescence. Lines were imaged with a Leica SP5 confocal

microscope, and the following three channels were acquired, mVenus (excitation 514 nm, emission 527-552 nm), mScarlet (561 nm, 595-620 nm) and chlorophyll (633 nm, 687-739 nm). Maximum intensity projections were prepared and merged for the Z-planes collected from the channels using Fiji.

Plasmid construction for genome editing in *M. polymorpha*

Parts from the OpenPlant kit and the assembly protocols reported in Sauret-Güeto et al.[53] were used, with modifications as noted. CRISPR-Cas9 constructs singly expressing two gRNAs from separate Mp*U6* promoters were generated for each of Mp*bHLH12*, Mp*MYB14*, Mp*MYB02* and Mp*WDR3* to target the locus with two gRNAs, and for Mp*WDR1* and *2* to target them simultaneously with one gRNA for each locus. For these constructs, gRNA DNA oligos were annealed and integrated in L1_*lacZgRNA*-Ck2 or L1_*lacZgRNA*-Ck3. Additionally, we used the *Csy4* CRISPR-associated ribonuclease system for polycistronic expression of four gRNAs (two for Mp*WDR1* and two for Mp*WDR2*) or six gRNAs (two for each of Mp*WDR1*, *2* and *3*). The *Csy4* coding sequence was amplified by Phusion HF two-step PCR from pAC_06_L0_Csy4P2A with TK23 and 24, cloned in pUAP4 and assembled in an L1 construct (L1_Csy4-Ck3) with the Mp*UBE2* promoter (OP-047) and the *Nos-35S* terminator (OP-054), and confirmed by restriction enzyme digestion and Sanger sequencing (primer TK71). Additionally, a version of the *5UTR_MpUBE2* part was prepared with a 3′ *Csy4* recognition sequence by Phusion HF two-step PCR (primer TK19-22) and cloned in pUAP4. Parts for the polycistronic gRNA arrays were amplified from this part and from pMODB2112 by Phusion HF two-step PCR (primers TK28-29,34-37,52-59) and assembled directly into pUAP4 using a custom SapI syntax and verified by Sanger sequencing. These were incorporated into L1 constructs (L1_polycistronic_gRNA_array-Ck2) as above, with the Mp*UBE2* promoter (OP-047) and the *Nos-35S* terminator (OP-054). L2 assemblies were generated always with L1_CsR-Ck1 (OP-062) and L1_Cas9-Ck4 (OP-073) and with either L1_gRNA1-Ck2 and L1_gRNA2-Ck3 or L1_polycistronic_gRNA_array-Ck2 and L1_Csy4-Ck3. L3 assemblies were carried out to add a ubiquitous *mTurquoise* for initial screening for transformation success, by combining L2_UBE2:mT-N7-CsA, the CRISPR-Cas9 L2 plasmids generated (all in CsC) and spacers in pCsB and E (OP-014, OP-016). The final assemblies were confirmed by restriction enzyme digestion and Sanger sequencing (with p_CF, p_CR, AC188, TK41, TK45, TK49).

Phenotypic analysis of *M. polymorpha*

Induction of auronidin pigmentation was carried out according to Albert et al.[20] Four-week old plants were transferred to minimal medium, with the exception of the complemented lines (Figure S2), which were transferred at two weeks of age. 14–15 days after transfer to minimal medium, pigmentation status was scored visually and plants were imaged with a Keyence VHX Digital Microscope. Auronidin content was quantified in thallus pieces cultured on minimal medium for 11 days, as follows. Approximately 100 mg of fresh thallus tissue was disrupted with a plastic pestle in 1 mL 80:19:1 methanol:water:formic acid,[20] vortexed for 10 seconds and centrifuged for 2 minutes at 13,300 rpm. To perform absorbance measurements with a BMG ClariostarPlus plate reader, 50 μL of the supernatant was transferred to a 384-well polypropylene microplate. Excitation at 490–15 nm and emission 570-8 nm (gain: 1512) were used, values were collected as 1 mm orbital average with the fluorescence intensity end-point mode, adjusted by subtracting the value of the blank and normalised by fresh weight. Oil bodies were analysed in gemmae stained with 500 nM BODIPY 493/503. Gemmae were stained for 10 minutes, rinsed twice in water with 0.1% Triton-X100 and imaged with a Leica SP5 Confocal Microscope (excitation 496 nm, emission 503-533nm). Maximum intensity projections were prepared and merged for the Z-planes collected from the two channels and stained oil bodies were counted in Fiji. In addition to oil bodies, staining was observed in rhizoid initials and slime papillae, and these structures were excluded from the count. The length of the perimeters of the gemmae were measured by Fiji on masks generated using thresholding (15-255) on maximum intensity projections of the Z-planes collected from the chlorophyll channel.

QUANTIFICATION AND STATISTICAL ANALYSIS

To evaluate the structural predictions of protein complexes, we used conventional confidence scores, the predicted aligned error (PAE) for each position relative to each other position in the prediction and the interface predicted template modelling score (iPTM) for the overall prediction. We used a one-tailed Mann-Whitney U test to test, for each of several independent lines for each genotype, whether they showed a significant reduction relative to the control line in auronidin content and normalised oil body counts respectively. Plotting and statistical analyses were carried out with the Python libraries *pandas, matplotlib*, *scipy* and *seaborn* using the Jupyter Notebook platform.[72–74]

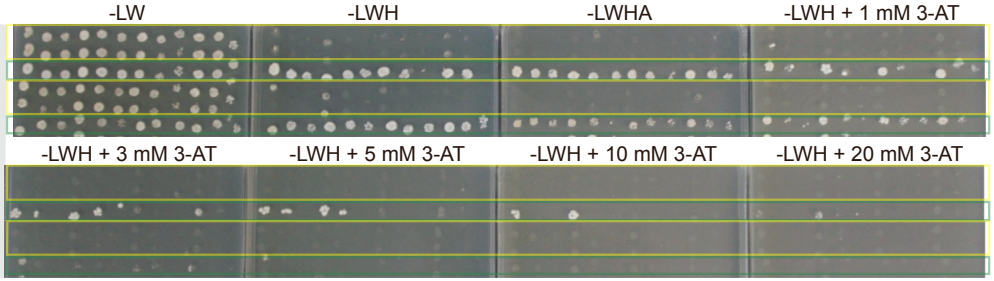
Current Biology, Volume *36*

Supplemental Information

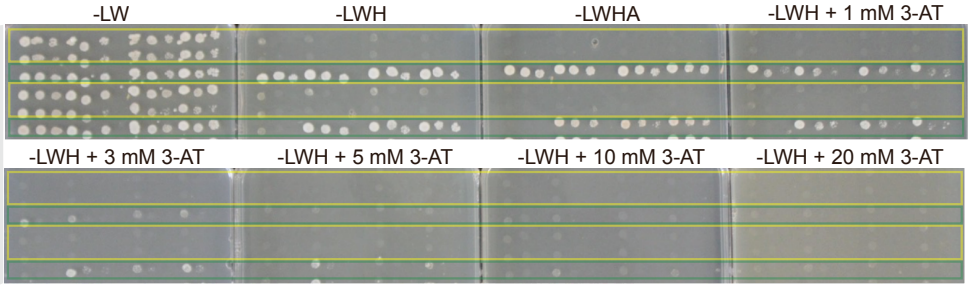
Replicated repurposing of an ancestral transcriptional complex in land plants

Thea E. Kongsted, Facundo Romani, Chiara A. Airoldi, Jim Haseloff, and Beverley J. Glover

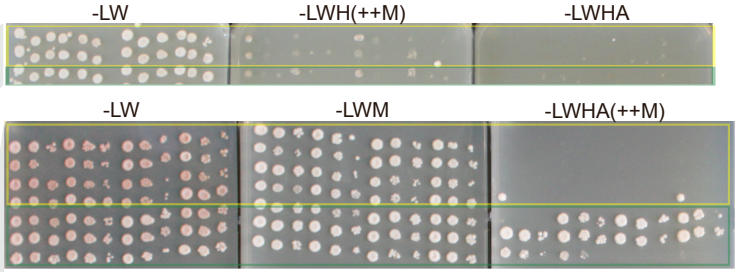
BD	AD
MpbHLH12	-
-	MpMYB02
MpbHLH12	MpMYB02
MpWDR1	-
-	MpbHLH12
MpWDR1	MpbHLH12



BD	AD
MpbHLH12	-
-	MpMYB14
MpbHLH12	MpMYB14
MpWDR2	-
-	MpbHLH12
MpWDR2	MpbHLH12



BD	AD
MpWDR3	mVenus
mVenus	MpbHLH12
MpWDR3	MpbHLH12



BD	*pMet17*	AD
MpbHLH12	MpWDR1	mVenus
MpbHLH12	MpWDR2	mVenus
MpbHLH12	-	mVenus
mVenus	-	MpMYB14
MpbHLH12	MpWDR1	MpMYB14
MpbHLH12	MpWDR2	MpMYB14
MpbHLH12	-	MpMYB14

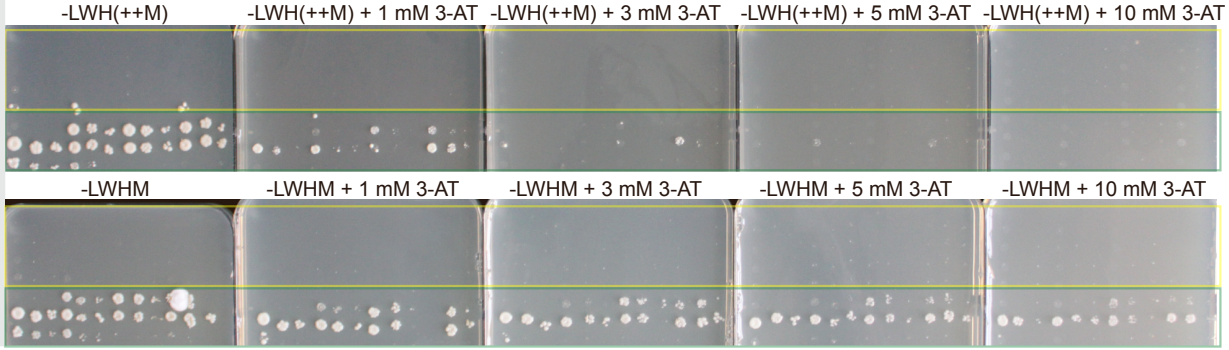


Figure S1. Additional replicates of yeast two-hybrid tests, related to Figure 2

On each plate, each row corresponds to a different combination of plasmids, indicated in the tables, used for co-transformation. The columns going left to right represent one replicate, followed by two 20-times serial dilutions, followed by the next replicate, and so on. Four biological replicates were performed for each combination, one of which is given in Figure 2. Autoactivation tests are highlighted in yellow and interaction tests between proteins of interest are highlighted in green. Growth on medium without leucine and tryptophan (-LW) confirms that the colonies carry both plasmids with the gene indicated in the table in translational fusion with either the Binding Domain (BD) or Activation Domain (AD) of the yeast transcription factor GAL4. Protein-protein interaction allows a reconstituted GAL4 to drive expression of reporter genes from the GAL4 Upstream Activation Sequence (UAS). Reporter gene expression complements auxotrophic mutations in histidine and adenine biosynthesis genes in the background genotype, allowing the yeast to grow in the absence of these amino acids (-HA). 3-amino-1,2,4-triazole (3-AT) is a competitive inhibitor of the gene product of the histidine reporter. In one experiment, a *MET17* cassette integrated on the BD plasmid drives methionine-repressible (-M condition vs ++M condition) expression of a third gene.

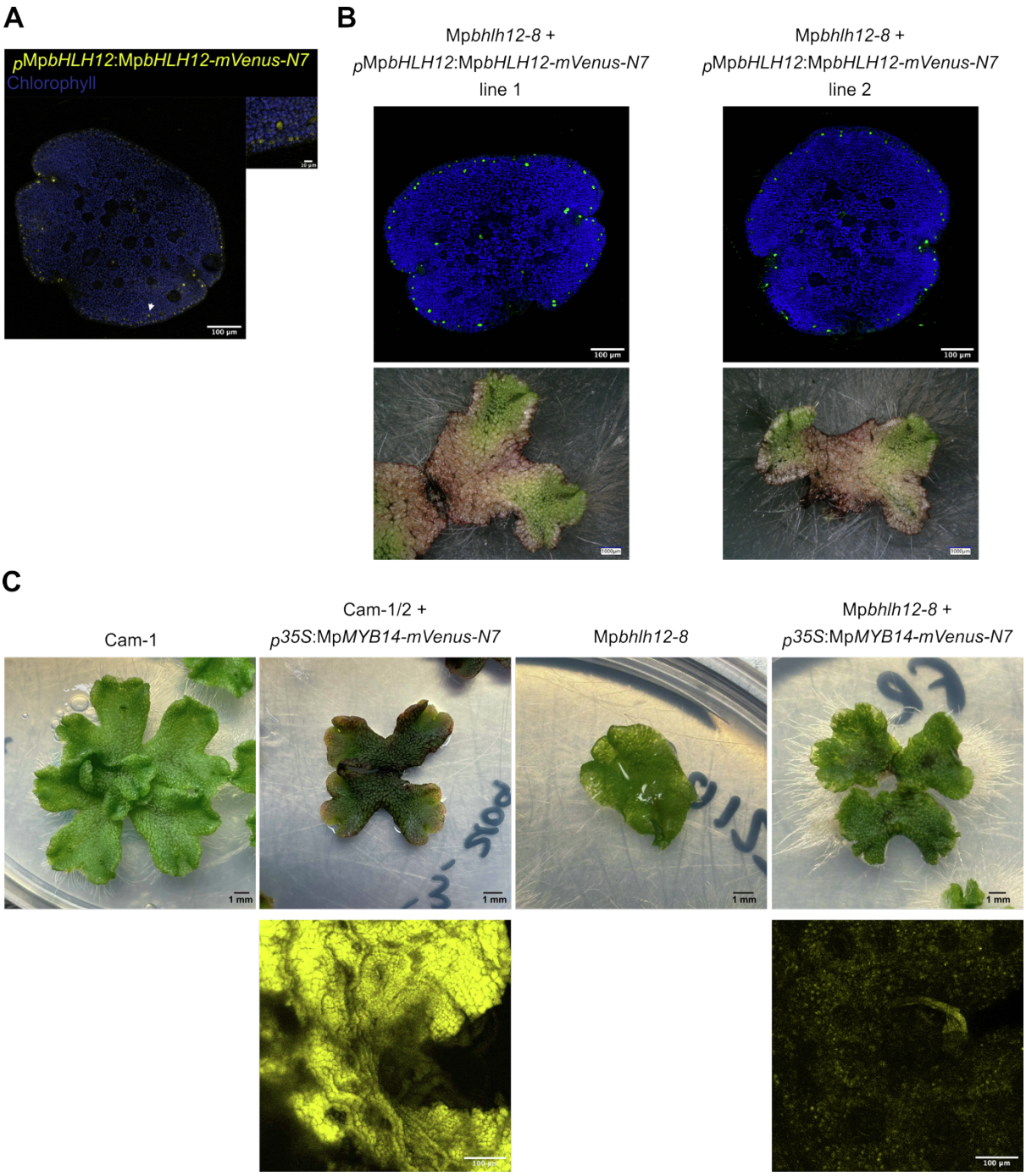
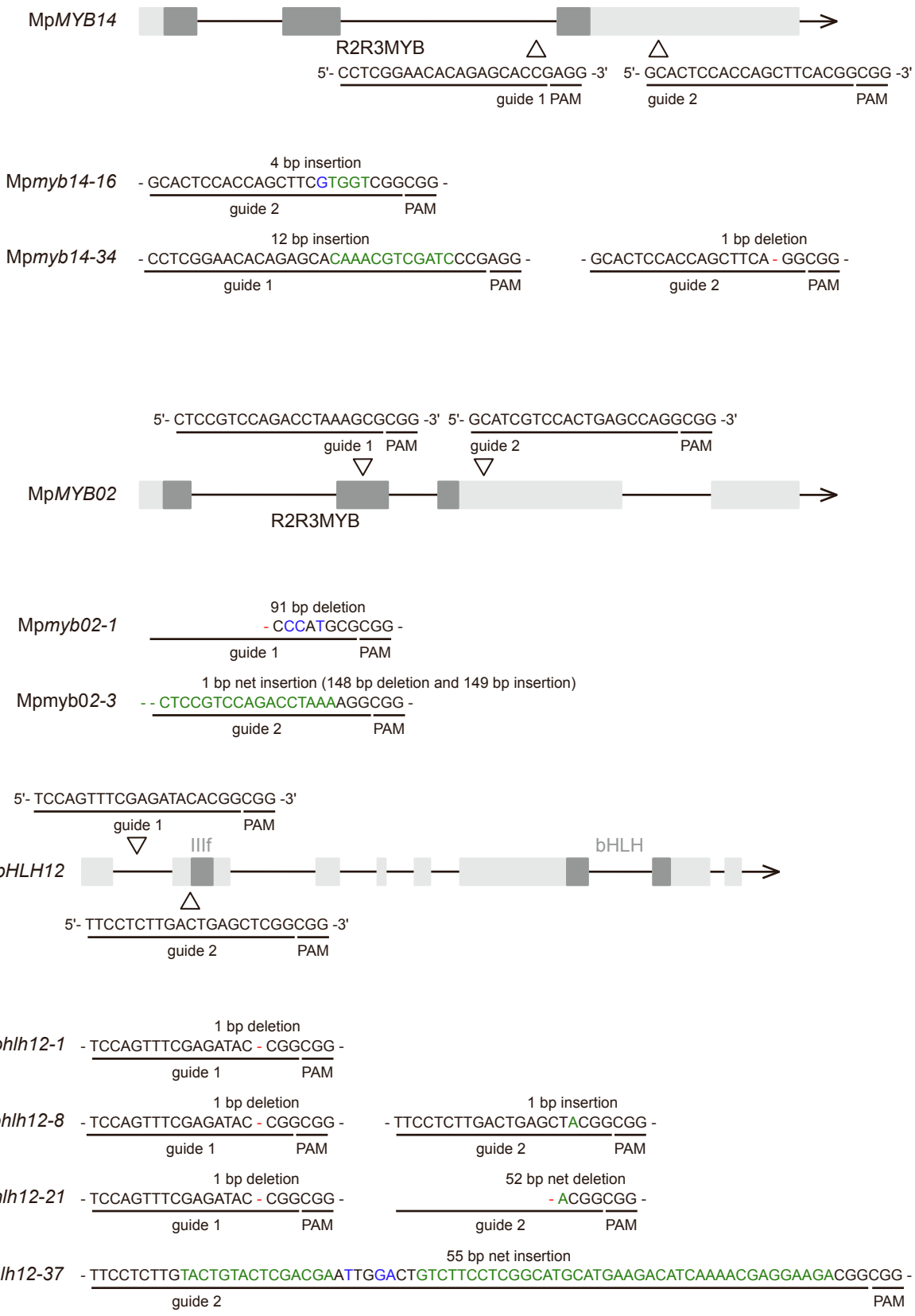


Figure S2. Additional transgenic lines, related to Figures 3 and 4 and the Discussion

A. Maximum intensity projection of the mVenus and chlorophyll channels acquired by confocal microscopy of a dormant gemma from a translational reporter line for Mp*bHLH12*. MpbHLH12 protein accumulated around the periphery, including at notch regions, and in oil body cells (see arrow and inset). **B.** The gRNA-resistant version of the Mp*bHLH12* CDS encoded by the translational reporter could genetically complement the defects in pigmentation and oil body maturation in a *bhlh12* knockout line. See quantification of the phenotypes in Figure 4. Here,

the images show pigmentation in lines transferred to minimal medium at two weeks of age and grown for a further two weeks. **C.** Overexpression of Mp*MYB14-mVenus* in the wild type (Cam-1/2) and the Mp*bhlh12-8* mutant background. Representative bright field images are shown alongside maximum intensity projections of the mVenus channel acquired by confocal microscopy of equivalent lines. In the wild-type background, auronidins accumulated and generated autofluorescence in the mVenus channel. In the mutant background, this accumulation was not observed, and the nuclear localised MpMYB14-mVenus could be clearly distinguished.


MpMYB14
R2R3MYB
5'- CCTCGGAACACAGAGCACCGAGG -3'
guide 1 PAM
5'- GCACTCCACCAGCTTCACGGCGG -3'
guide 2
PAM
Mpmyb14-16
4 bp insertion
- GCACTCCACCAGCTTCGTGGTCGGCGG -
guide 2
PAM
Mpmyb14-34
12 bp insertion
- CCTCGGAACACAGAGCACAAACGTCGATCCCGAGG -
guide 1
PAM
1 bp deletion
- GCACTCCACCAGCTTCA - GGCGG -
guide 2
PAM
5'- CTCCGTCCAGACCTAAAGCGCGG -3'
guide 1 PAM
5'- GCATCGTCCACTGAGCCAGGCGG -3'
guide 2
PAM
MpMYB02
R2R3MYB
Mpmyb02-1
91 bp deletion
- CCCATGCGCGG -
guide 1
PAM
Mpmyb02-3
1 bp net insertion (148 bp deletion and 149 bp insertion)
- - CTCCGTCCAGACCTAAAAGGCGG -
guide 2
PAM
5'- TCCAGTTTCGAGATACACGGCGG -3'
guide 1
PAM
IIIf
bHLH
MpbHLH12
5'- TTCCTCTTGACTGAGCTCGGCGG -3'
guide 2
PAM
Mpbhlh12-1
1 bp deletion
- TCCAGTTTCGAGATAC - CGGCGG -
guide 1
PAM
Mpbhlh12-8
1 bp deletion
- TCCAGTTTCGAGATAC - CGGCGG -
guide 1
PAM
1 bp insertion
- TTCCTCTTGACTGAGCTACGGCGG -
guide 2
PAM
Mpbhlh12-21
1 bp deletion
- TCCAGTTTCGAGATAC - CGGCGG -
guide 1
PAM
52 bp net deletion
- ACGGCGG -
guide 2
PAM
Mpbhlh12-37
55 bp net insertion
- TTCCTCTTGTACTGTACTCGACGAATTGGACTGTCTTCCTCGGCATGCATGAAGACATCAAAACGAGGAAGACGGCGG -
guide 2
PAM

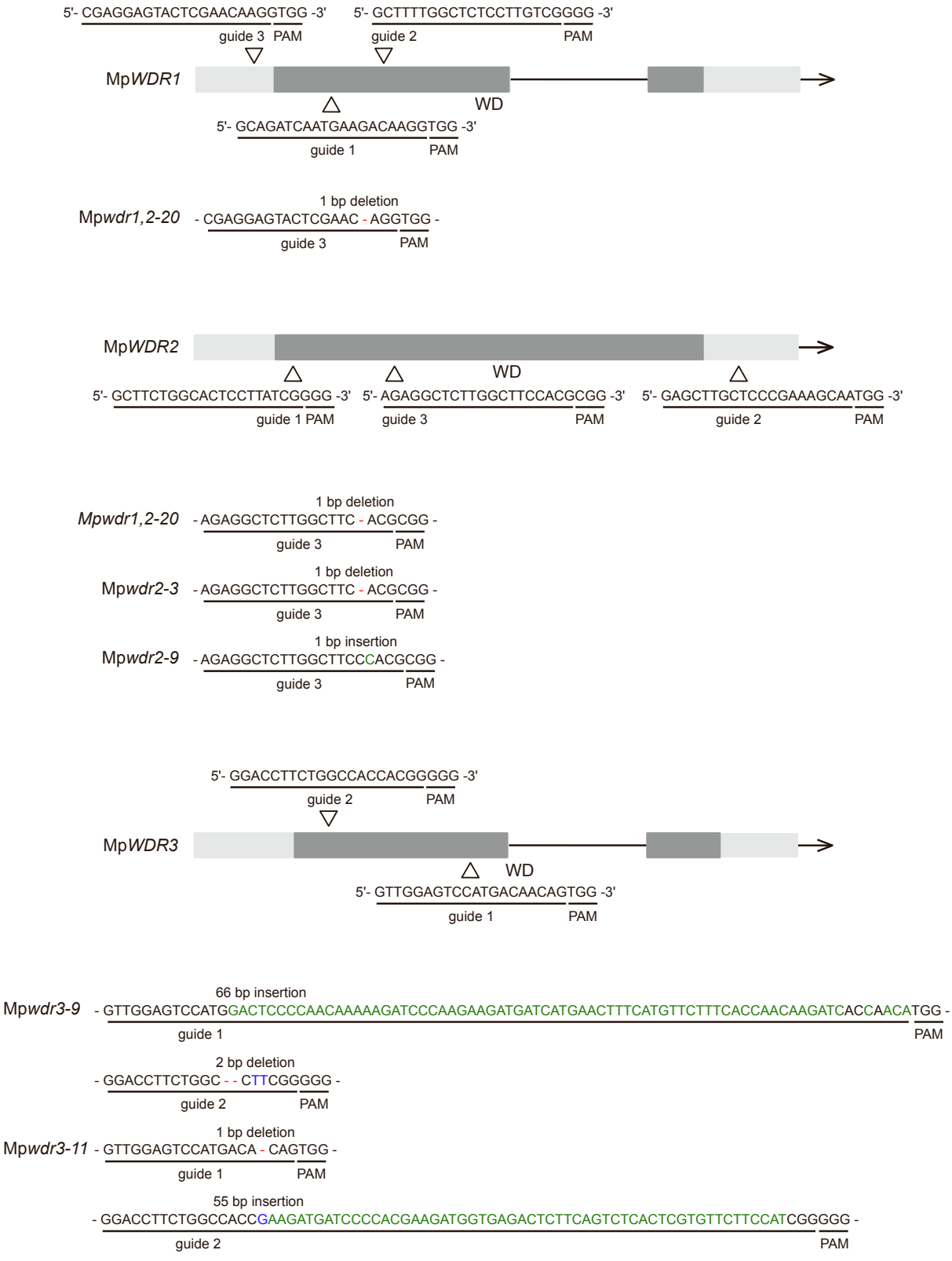


Figure S3. Overview of CRISPR/Cas9 mutants, related to Figure 4

The positions of CRISPR/Cas9 guide target sequences on the gene models (not to scale) are shown, and the edits introduced are given for each line analysed.

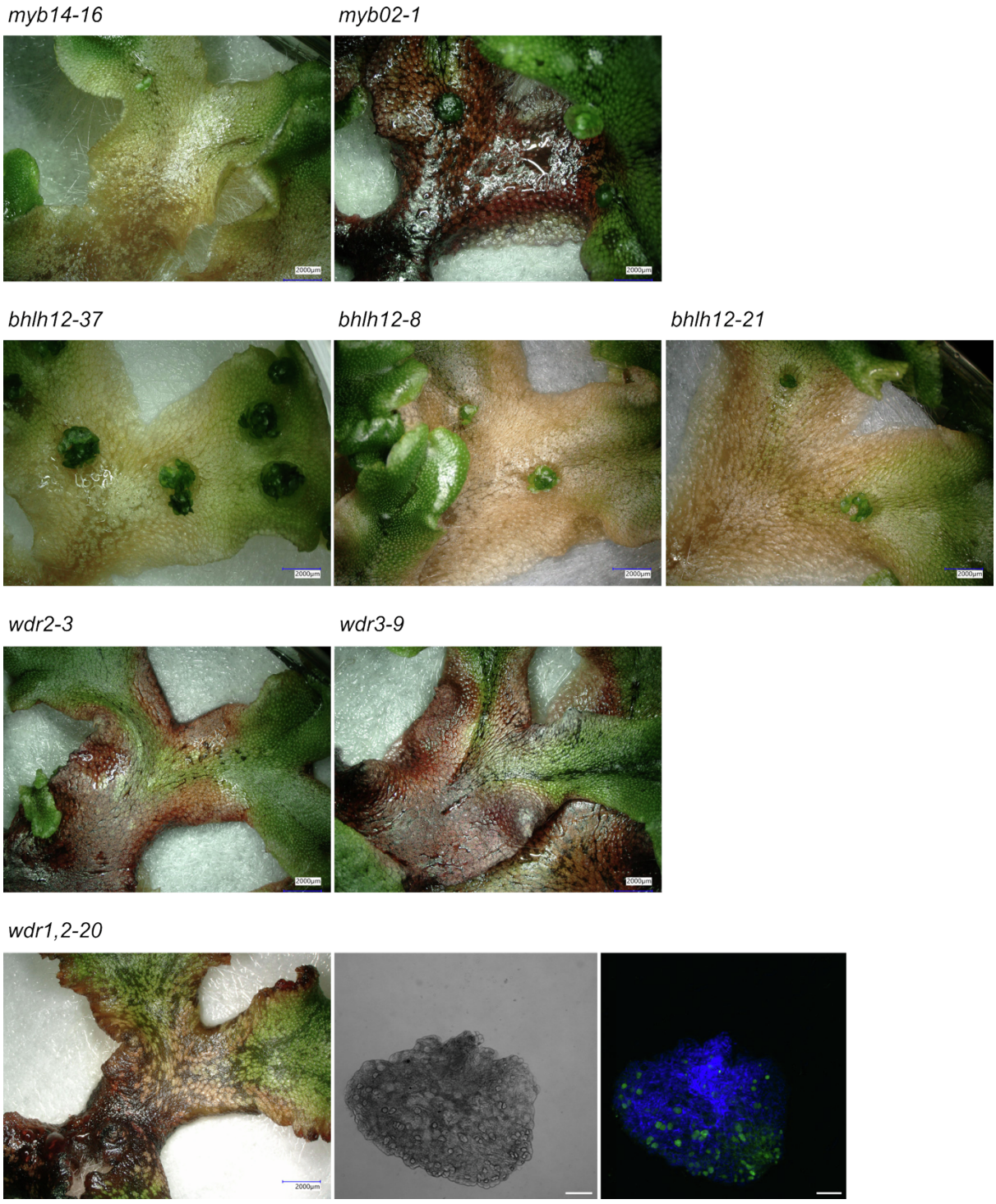


Figure S4. Phenotypic scoring for additional independent transgenic lines, related to Figure 4

Pigmentation under nutrient deprivation is shown for an additional independent CRISPR/Cas9 knockout line to that in the main figure for each of Mp*MYB14*, Mp*MYB02*, Mp*WDR2* and Mp*WDR3*. The scale bar is 2 mm. For Mp*bHLH12*, three additional lines are shown. Only one *wdr1,2* double mutant line was identified, shown here, and no other *wdr1* mutant alleles were

recovered. Visual scoring indicated that this line showed pigmentation under nutrient deprivation, and structures recovered from the dorsal surface of the thallus, which may represent anatomically aberrant gemmae, showed BODIPY staining in unidentified areas. The brightfield and merged fluorescence channels are shown for an example of such a structure, here the scale bar is 100 μm. Note that the phenotypes here associated with a *wdr1,2* genotype could not be validated by examination of another independent line.