Author/s:
Flores-Sandoval, E;Eklund, DM;Hong, SF;Alvarez, JP;Fisher, TJ;Lampugnani, ER;Golz, JF;Vázquez-Lobo, A;Dierschke, T;Lin, SS;Bowman, J

Title:
Class C ARFs evolved before the origin of land plants and antagonize differentiation and developmental transitions in Marchantia polymorpha

Date:
2018-06-01

Citation:
Flores-Sandoval, E., Eklund, D. M., Hong, S. F., Alvarez, J. P., Fisher, T. J., Lampugnani, E. R., Golz, J. F., Vázquez-Lobo, A., Dierschke, T., Lin, S. S. & Bowman, J. (2018). Class C ARFs evolved before the origin of land plants and antagonize differentiation and developmental transitions in Marchantia polymorpha. New Phytologist, 218 (4), pp.1612-1630. https://doi.org/10.1111/nph.15090.

DR D. MAGNUS EKLUND (Orcid ID : 0000-0002-0576-7636)

DR JOHN L BOWMAN (Orcid ID : 0000-0001-7347-3691)



Article type : MS - Regular Manuscript

# Class C ARFs evolved before the origin of land plants and antagonize differentiation and developmental transitions in *Marchantia polymorpha*

Eduardo Flores-Sandoval[1], D. Magnus Eklund[1], Syuan-Fei Hong[2], John P. Alvarez[1], Tom J. Fisher[1], Edwin R. Lampugnani[3], John F. Golz[3], Alejandra Vázquez-Lobo[4], Tom Dierschke[1], Shih-Shun Lin[2] and John L. Bowman[1]

[1]School of Biological Sciences, Monash University, Clayton, Melbourne, VIC 3800, Australia; [2]Institute of Biotechnology, National Taiwan University, 81, Chang-Xing ST., Taipei 106, Taiwan; [3]School of BioSciences, University of Melbourne, Parkville, VIC 3010, Australia; [4]CIByC, Universidad Autónoma del Estado de Morelos. Av. Universidad No. 1001. Colonia Chamilpa. CP 62209. Cuernavaca, Morelos, México

Author for correspondence:

*John L. Bowman*

*Tel: +61 3 99020242*

*Email: john.bowman@monash.edu.au*

Received: *8 October 2017*

Accepted: *5 February 2018*



**Summary**

- A plethora of developmental and physiological processes in land plants is influenced by auxin, to a large extent via alterations in gene expression by AUXIN RESPONSE FACTORs (ARFs). The canonical auxin transcriptional response system is a land plant innovation, however, charophycean algae possess orthologs of at least some classes of ARF and AUX/IAA genes, suggesting elements of the canonical land plant system existed in an ancestral alga.
- We reconstructed the phylogenetic relationships between streptophyte ARF and AUX/IAA genes and functionally characterized the solitary class C ARF, Mp*ARF3*, in *Marchantia polymorpha*.
- Phylogenetic analyses indicate that multiple ARF classes, including class C ARFs, existed in an ancestral alga. Loss- and gain-of-function Mp*ARF3* alleles result in pleiotropic effects in the gametophyte, with Mp*ARF3* inhibiting differentiation and developmental transitions in multiple stages of the life cycle. While loss-of-function Mp*arf3* and Mp*mir160* alleles respond to exogenous auxin treatments, strong miR-resistant Mp*ARF3* alleles are auxin insensitive, suggesting class C ARFs act in a context-dependent fashion.

- We conclude that two modules independently evolved to regulate a pre-existing ARF transcriptional network. While the auxin-TIR1-AUX/IAA pathway evolved to repress class A/B ARF activity, miR160 evolved to repress class C ARFs in a dynamic fashion.

**Key words:** auxin, AUXIN RESPONSE FACTOR (ARF), auxin signalling, class C ARF, land plant evolution, *Marchantia*, mir160.

## Introduction

The canonical auxin transcriptional response system was originally characterized in flowering plants. In the absence of auxin, AUXIN/INDOLE-3-ACETIC ACID (AUX/IAA) repressors heterodimerize with AUXIN RESPONSE FACTOR (ARF) transcription factors via the PB1 domain and recruit TOPLESS (TPL) corepressors, thus preventing ARFs from regulating target genes (Kim *et al*., 1997; Ulmasov *et al*., 1997a,b; Szemenyei *et al*., 2008; Vernoux *et al*., 2011). In the presence of auxin, the AUX/IAA proteins are targeted for degradation by an SCF E3 ubiquitin ligase (TRANSPORT INHIBITOR RESPONSE1 (TIR1)/AUXIN SIGNALING F-BOX (AFB)), via auxin dependent interactions between AUX/IAA and TIR1 proteins (Gray *et al*., 2001; Tiwari *et al*., 2001; Dharmasiri *et al*., 2005; Kepinski & Leyser, 2005). Comparison of TIR1, AUX/IAA and ARF orthologs across land plants and charophycean algae indicate that the assembly of the canonical auxin transcriptional response pathway is a land plant innovation (Hori *et al*., 2014; Bowman *et al*., 2017).

Phylogenetic analyses across land plants identify three ARF classes (A, B, C; Ulmasov *et al*., 1999a; Tiwari *et al*., 2003; Finet *et al*., 2013). ARF classes are not equal, with class A acting primarily as transcriptional activators, and classes B and C as transcriptional

repressors (Ulmasov *et al*., 1999a; Tiwari *et al*., 2003). Since auxin application predominantly activates gene expression, it has been argued that class B and C ARFs may act to repress common target genes activated by class A ARFs as all classes can potentially compete for targets as they recognize similar AuxREs (Ulmasov *et al*., 1997a; Ulmasov *et al*., 1999a,b; Tiwari *et al*., 2003). Yeast-two-hybrid studies suggested that activator ARFs interact with AUX/IAAs and that AUX/IAAs may also homodimerize. By contrast, repressor ARFs do not interact efficiently with AUX/IAAs, leaving most class B and C repressors outside this protein network (Vernoux *et al*., 2011). This lack of interaction implies that many ARF repressors regulate gene expression independently of auxin. Thus, it is possible that transcriptional activation by class A ARFs is negatively regulated in an auxin dependent manner by AUX/IAAs and an auxin independent manner by class B and C ARFs.

This system is broadly conserved across land plants, with both the classes and interactions conserved in the moss *Physcomitrella patens* (Prigge *et al*., 2010; Lavy *et al*., 2016) and the liverwort *Marchantia polymorpha* (Flores-Sandoval *et al*., 2015; Kato *et al*., 2015). In *Marchantia*, trans-activation assays suggest the single class A ARF (MpARF1) acts as transcriptional activator, the single class B ARF (MpARF2) acts as a transcriptional repressor and the class C ARF (MpARF3) is not explicitly an activator or a repressor (Kato *et al*., 2015). Loss of Mp*ARF1* results in pleiotropic phenotypes including auxin insensitivity, constitutive tropisms, loss of gemmae dormancy, formation of ectopic meristems and mis-regulated cell division patterns during gemmae development (Sugano *et al*., 2014; Flores-Sandoval *et al*., 2015; Kato *et al*., 2017). However, the morphology of Mp*arf1* mutants does not resemble loss of auxin biosynthesis mutants, where all tissue patterning is lost (Eklund *et al*., 2015). A complete loss of growth and patterning associated with auxin depletion can only be observed by making proteolysis resistant MpIAA (Kato *et al*., 2015) or chimeric MpTPL-PB1 domain fusions using PB1 domains of MpARF1, MpARF2 and MpIAA (Flores-Sandoval *et al*., 2015). Although all MpARFs have the potential to interact with MpIAA

when expressed in yeast (Kato *et al*., 2015), chimeric TOPLESS-PB1$^{MpARF3}$ fusions cannot elicit strong developmental phenotypes *in planta*, suggesting that MpARF3 activity may be independent of AUX/IAA degradation by auxin-TIR1 (Flores-Sandoval *et al*., 2015).

The class C ARF/MIR160 module in angiosperms regulates a range of processes including leaf margin complexity and lamina growth (Mallory *et al*., 2005; Hendelman *et al*., 2012; Ben-Gera *et al*., 2016), nodule formation in soybean roots (Turner *et al*., 2013; Nizampatnam *et al*., 2015), root meristem patterning (Wang *et al*., 2005; Ding & Friml, 2010; Bennett *et al*., 2014) floral determinacy (Liu *et al*., 2010), ovary and fruit patterning, floral organ abscission (Damodharan *et al*., 2016) and seed germination (Liu *et al*., 2007; Liu *et al*., 2013). Class C ARFs also promote the formation of shoot meristems in undifferentiated callus by directly repressing negative regulators of cytokinin signalling (Qiao *et al*., 2012; Liu *et al*., 2016). Auxin response reporters have been shown to be both induced and repressed by class C ARF activity in angiosperms, suggesting a contextual role for these genes in development (Mallory *et al*., 2005; Wang *et al*., 2005; Liu *et al*., 2010; Ben-Gera *et al*., 2016). Irrespective of their function as transcriptional repressors or activators, genetic studies suggest that class C ARFs have the potential to antagonize auxin signalling (Turner *et al*., 2013; Ben-Gera *et al*., 2016). The function of Class C ARFs outside of angiosperms is largely unknown, except that negative regulation by miR160 is conserved across land plants (Axtell *et al*., 2007; Lin *et al*., 2016; Tsuzuki *et al*., 2016). Here we investigate the evolutionary history of the ARF and AUX/IAA gene families and functionally characterize the class C ARF in the liverwort *Marchantia polymorpha*.

## Materials and Methods

Sequence analysis, alignment, and phylogenetic analysis **[Author, please check that the formatting and hierarchy of the headings are correct throughout.]**

Coding nucleotide sequences were manually aligned as amino acid translations (Se-Al v2.0a11; Rambaut, 1996). Sequence alignments and Bayesian and Maximum Likelihood command files for the phylogenetic analyses are provided upon request.

An unambiguous alignment of 78 amino acid characters in 138 PB1 sequences for subsequent Bayesian analysis using Mr. Bayes 3.1.2 (Huelsenbeck & Ronquist, 2001). The fixed rate model option JTT + I was used and analyses were run for 2,500,000 generations, sufficient for convergence of simultaneous runs. To allow for burn-in, 10% of the total saved trees was discarded. Convergence to 0.022 was observed, the relatively high value likely due to polytomies within the IAA, A ARF, and C ARF subclades.

An unambiguous alignment of 1278 nucleotide characters in 99 (or 97) ARF sequences for subsequent Bayesian analysis using Mr. Bayes 3.1.2 (Huelsenbeck and Ronquist, 2001) that were run for 5,000,000 generations, sufficient for convergence of the simultaneous runs. To allow for the burn-in, 10% of the total saved trees was discarded. Convergence to 0.041 (Supporting Information Fig. S1a) and 0.023 (Fig. S1b) were observed, the relatively high value likely due to polytomies and/or low levels of support for some subclades.

**Maximum Likelihood (ML)** Translated protein sequences were analysed against Conserved Domains Database (http://www.ncbi.nlm.nih.gov/cdd/) using a CD-Search. Protein sequences of the three conserved domains were aligned using a multiple alignment algorithm implemented on the online version of MAFFT v. 7. Amino acid alignments were converted to nucleotides using the PAL2NAL on line tool (http://www.bork.embl.de/pal2nal/). A nucleotide matrix using only 1st and 2nd codon positions was used to perform a ML search in RaxML Black box (https://embnet.vital-it.ch/raxml-bb/) using the CAT model of rate heterogeneity.

Plant growth

Plants were grown as described previously (Gamborg *et al*., 1968; Ishizaki *et al*., 2008; Flores-Sandoval *et al*., 2015). For auxin sensitivity assays plants were grown for 30 d in B5 media before transferring to either mock or exogenous auxin plates (2,4-D 10 μM). Ten independent lines were assayed per condition. For transcriptional auxin induction experiments, mature 1 month thalli were transferred to liquid B5 media (2% sucrose, 0.1% Casaminoacids and 0.03% L-glutamine) in either mock (ethanol) or 2,4-D (20 μM) and tissue was harvested after 48 h in continuous white light.

Cloning and plasmids

Genetic nomenclature is as outlined in (Bowman *et al*., 2016). Primers are listed in Table S1.

**gRNA design and cloning** gRNAs targeting genes of interest were co-expressed on independent T-DNAs using plasmids GE010 (Sugano *et al*. (2014) and S. Sugano *et al*. (unpublished)) and GWB301 (Ishizaki *et al*., 2015). Design of specific gRNAs was perfomred such that putative off-targets (*Marchantia* genome assembly 3.1) lacked an associated PAM site. gRNAs were cloned downstream of the Mp*U6* promoter, excluding PAM sites, and were co-expressed with Cas9 (Sugano *et al*., 2014).

**Cloning of Mp*MIR160*** The Mp*MIR160* precursor was amplified using primers EcoRI-MpmiR160-F/KpnI-MpmiR160-R and cloned into pCRII-TOPO. *Eco*RI and *Kpn*I were used to subclone into EF1:BJ36 V2.0 and *Not*I into pHART or pKART (binary vectors conferring Hygromycin or G418 resistance in plants, respectively).

**Cloning and mutagenesis of Mp*ARF3*** Wild-type cDNA from 8-d-old gemmalings was used to amplify a 2.4kb Mp*ARF3* transcript lacking the canonical PB1 domain using nested PCR (primers: MpARF3-F5/MpARF3-F3/MpARF3-R5/MpARF3-R3) and cloned into pCRII-TOPO. This amplicon was fused with an *in vitro* synthesized (Genescript) 592 bp 3'end containing the WT PB1 domain using *Bst*XI and *Hin*dIII. Mp*ARF3* was subcloned into EF1:BJ36 using *Kpn*I and *Hin*dIII and then into HART using *Not*I creating $_{pro}$*EF1:*Mp*ARF3*. Assembly PCR was used to mutagenize Mp*ARF3* in pCRII. Primers M13-R/MpARF10m6-Fwd were used to create fragment 1 (893 bp), and primers MpARF3-F3/MpARF10m6-Rev to create fragment 2 (1651 bp). The assembly used both fragments as templates and primers MpARF10codingKpnI-F/MpARF10codingHindIII-R in a nested PCR and the product was cloned as described for $_{pro}$*EF1:*Mp*ARF3*.

**Mp*ARF3* promoter** Primers MpARF10-pro-NdeI-Fwd/MpARF10proKpnI-Rev were used to amplify a 3.6kb fragment from a genomic PAC clone provided by Takayuki Kohchi. The PCR product was cloned into pCRII-TOPO. *Nde*I and *Kpn*I were used to subclone $_{pro}$Mp*ARF3* into pRITA, a shuttle vector with a *GUS/UidA* reporter gene and a *nos* terminator. *Nde*I and *Kpn*I were used to exchange $_{pro}$*EF1* into $_{pro}$*EF1:*Mp*ARF3*/Mp*ARF3*$^{m}$ to create $_{pro}$Mp*ARF3:*Mp*ARF3* and $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$. The $_{pro}$Mp*ARF3* was exchanged with $_{pro}$Mp*SHI* (Mapoly0049s0083.1) using *Kpn*I/*Hin*dIII into pME159 (Flores-Sandoval *et al*., 2015), creating $_{pro}$*SHI:ARF3* and $_{pro}$Mp*SHI:ARF3*$^{m}$. All constructs were transferred to HART using *Not*I.

**Mp*MIR160* promoter** Wild-type DNA was used to amplify a 4kb fragment using primers 4kbMpMIR160pro-F-pENTR/MpMIR160pro-R, which was cloned into pENTR-D and subsequently recombined into pGWB404 (Ishizaki *et al*., 2015).

**Complementation clones** *Cas9* resistant Mp*ARF3* (Mp*ARF3*$^{Cas9Res}$) was generated by assembly PCR: primer sets M13F/ARF3gRNA5mutR and ARF3gRNA5mutF/M13R were

used to create products then assembled by nested PCR using MpARF3-ATG-KpnI/MpARF3-STP-HindIII. A mutagenized gRNA5 resistant Mp*ARF3* (7118-4) was cloned in pCRII and used as a template for a second round of mutagenesis using ARF3gRNA4mutF/R. Resulting amplicons were cloned into pCRII-TOPO, and subsequently subcloned into $_{pro}$Mp*ARF3*:Mp*ARF3* BJ36 using *Kpn*I/*Hin*dIII and into KART using *Not*I. The class C ARF, At*ARF10*, was amplified with primers AtARF10FsmK5-KpnI/AtARF10R-HindIII. AtARF10FsmK5-KpnI creates a silent mutation erasing the *Hin*dIII site in ARF10-K5 to facilitate cloning into BJ36 (*KpnI/HindII*) and later KART (*NotI*).

RLM-RACE

RACE PCR was performed as described previously (Flores-Sandoval *et al*., 2016) to detect cleavage of Mp*ARF3* transcripts using primers RLM-MpARF3-R11/RLM-MpARF3-R10.

RNAseq

Total RNA isolated from 1-month-old thalli (QIAGEN RNAeasy Plant Kit) from three independent lines per genotype was assessed for quality with a Nanodrop spectrophotometer and a Bioanalyzer 2100 microfluidics system (Agilent). Library preparation substrates were enriched for mRNA (Illumina TruSeq Stranded mRNA technology). Sequencing used the Illumina NextSeq500 in High-Output mode (20 million single-ended 75b reads per sample). Reads were mapped onto the *Marchantia* v3.1 assembly using TopHat Version 2.1.0 for Galaxy (Kim *et al*., 2013). Resulting Bam files were loaded onto IGV 1.3.1 to produce Sashimi Plots (Robinson *et al*., 2011). A count matrix of average raw reads/gene/sample was created using HTseq-code using Galaxy (Anders *et al*., 2015). Raw reads were normalized by total million reads per library (RPM) and transcript lengths in kilobases (RPKM). The

voom/limma, edgeR and DESeq2 packages were used for differential gene expression analysis using Galaxy (Table S2) (Afgan *et al*., 2015). Fastq files are accessible in SRA (NCBI) with the following accessions: SAMN08493813, SAMN08493814, SAMN08493815, SAMN08493816, SAMN08493817, SAMN08493818, SAMN08493819, SAMN08493820, SAMN08493821. BioProject ID: PRJNA433456

qPCR

See Methods S1.

*In situ* hybridization

See Methods S1.

Plant transformation and screening

*Marchantia* sporeling transformation was performed as described by (K. Ishizaki *et al* 2008). Putative CRISPR-gene-editing lines were plated under two successive rounds of Hygromicin/Chlorosulfuron and selected by phenotypic resemblance to loss-or gain-of-function transgenic lines. Co-expression of two gRNAs substantially increased gene-editing frequency.

Complementation assays

Complementation with Mp*ARF3* was obtained by co-expressing gRNAs with Cas9-resistant Mp*ARF3* variants using a third selectable marker (G418), screened with primers MpARF3screenF/MpARF3-screen-R. Mp*MIR160* genotyping was done with primers EcoRI-

MpmiR160-F/KpnI-MpmiR160-R and PCR (TERRA-PCR, Clonetech). Amplification of the endogenous Mp*ARF3* locus in Mp*ARF3*$^{Cas9Res}$ complementation lines employed an intron specific reverse primer (MpARF3-scr-intronR1) in conjunction with MpARF3screenF. Plants were grown under two successive rounds of selection and then transferred to ½ B5 media for another month before analyses.

Area measurements

Area measurements employed Adobe Photoshop® using photos of plants with a conventional ruler to set a 'Costume Measurement Scale'. Two-dimensional area representations were delineated using the 'Color range' function and measurements were done with 'Record Measurement' function.

GUS staining

GUS assays were performed as previously described (Jefferson *et al.*, 1987). Plants were incubated in GUS staining solution (0.5 mM Potassium Ferrocyanid, 0.5mM Potassium Ferricyanide and 1 mM X-Gluc) for 3 h at 37°C and cleared with ethanol.

Microscopy

Scanning electron microscopy was performed as previously described (Flores-Sandoval *et al.*, 2015).

**Results**

Origin and evolution of land plant ARF genes

ARF proteins possess B3 DNA-binding and PB1 protein-protein interaction domains and similar genes are found throughout Streptohphytes (Ju *et al*., 2015; Bowman *et al*., 2017). While IAA proteins possess a PB1 domain, they lack the B3 domain (Hori *et al*., 2014; Bowman *et al*., 2017). However, lack of an appropriate outgroup has hindered resolution of phylogenetic relationships between ARF and IAA genes. Recent analyses of charophycean algal and liverwort sequences revealed another clade of B3 transcription factors, the RAV family, with a PB1 domain suggesting the two domains were already associated during early charophyte evolution (Bowman *et al*., 2017). RAV orthologs from other land plant lineages have lost the PB1 domain. Here we made an assumption that the association of the B3 and PB1 domains in ARF and RAV proteins is homologous and used RAV sequences to root the tree and resolve the position of the IAA genes using genome and transcriptome data from basal land plants and charophycean algae (Rensing *et al*., 2008; Banks *et al*., 2011; Nystedt *et al*., 2013; Wickett *et al*., 2014; Ju *et al*., 2015; Vanneste *et al*., 2015; Cooper & Delwiche, 2016).

Phylogenetic analysis of PB1 amino acid sequences from RAV, ARF, and IAA genes resulted in a well-supported tree, with the three ARF classes forming a clade sister to a clade containing the IAA genes (Fig. 1). Both ARF and RAV clades contain *Mesostigma* sequences, implying their origins occurred before the diversification of extant Streptophytes (Figs 1, S1a). Class C ARFs are sister to all other ARFs, and since charophyte sequences are found in both class C and class A/B clades, the gene duplication giving rise to class C ARFs and the ancestor of class A and B ARFs occurred in a charophycean ancestor. The class A and B ARFs diverged later, with phylogenetic analysis indicating this occurred in the ancestral land plant (Fig. S1a–d). We did not observe a class B ARF in the hornwort *Nothoceros*, although this could be due to sampling bias. The ancestral ARF possessed B3, FD, DD, LFG and PB1 domains, and all three classes retain each of these domains, with

class-specific LFG and PB1 sequence motifs (Figs S2, S3). Interaction of class C ARF transcripts with miR160 evolved in land plants, with the miR160 binding site embedded within a land plant specific CLP-domain (Fig. S4).

Related to class A ARFs are NON-CANONICAL ARF (NCARF) genes, which have lost the B3/FD/DD and LFG domains but retained a PB1 domain and thus, could act as modulators of auxin response via interaction of their PB1 domain with ARF or IAA PB1 domains. Both class A ARFs and NCARF are also characterized by a Q-rich domain. NCARF sequences have been noted previously in *Selaginella* and *Physcomitrella* (Finet *et al*., 2013) and here we find them in liverworts, but not in fern or gymnosperm species surveyed. This phylogenetic distribution implies a gene duplication in the ancestral land plant producing class A ARF and NCARF genes.

IAA genes, together with NON-CANONICAL IAA (NCIAA) genes, form a clade sister to the ARF clade, with the two clades diverging within a charophycean ancestor. This topology implies that the B3 domain was lost in the IAA/NCIAA lineage. Because the well-supported NCIAA clade contains charophyte sequences, the IAA and NCIAA clades also diverged within a charophycean ancestor. Unlike NCARF genes, NCIAA orthologs are broadly distributed from charophycean algae to flowering plants, including Arabidopsis (At5757420/*IAA33*), but may have been lost within moss and lycophyte lineages. Since no charophyte sequence in the IAA/NCIAA clade possesses domains I or II (Bowman *et al*., 2017), these domains I and II were acquired specifically in the IAA ancestor after its divergence from NCIAA. Basal land plant IAA genes are also characterized by two conserved motifs in addition to domains I and II, named here EHDY and KR (Fig. S5); the EHDY domain has been lost in most, if not all, angiosperm genes. Liverwort IAA sequences also contain a Q-rich domain, but its evolutionary relationship with similar domains in A ARF and NCARF genes is unclear.

Wild-type *M. polymorpha* development

As a background for interpreting phenotypes and expression patterns, a brief description of wild type is required. Wild-type *M. polymorpha* plants progress through a series of developmental stages during vegetative growth. Following spore germination, an early cell division produces a rhizoid and a vegetative cell, the latter of which undergoes multiple relatively unordered divisions to produce a small clump of cells (Fig. S6a). A single cell in the clump emerges as an apical cell, which produces lateral derivatives in two planes, forming a two-dimensional sheet of cells called a prothallus that lacks specialized cells (Fig. S6b). The prothallus may branch, but in most cases the apical cell progresses into one that produces derivatives in four planes (two lateral, ventral, and dorsal), producing a three-dimensional complex thallus with air chambers and pores (Fig. S6c). Gemmae cups are produced from the dorsal epidermis with gemmae developing from single cells of the cup base. Dormant gemmae, with two apical meristems, are several cell layers thick, but lack specialized cell types except rhizoid precursors and oil body cells (Fig. S6d).

The Mp*ARF3* expression pattern correlates growth and differentiation

To explore the Mp*ARF3* spatial-temporal expression pattern, we analysed GUS transcriptional fusions in a developmental series of $_{pro}$Mp*ARF3:GUS* (3.6kb upstream of gene model) and $_{pro}$Mp*MIR160:GUS* (4kb upstream of gene model) gemmalings (Fig. 2). Signal for both Mp*ARF3* and Mp*MIR160* was detected in apical notches prior, during and after the first branching event in gemmalings, between days 3 and 6 of gemmaling development (Flores-Sandoval *et al*., 2015; Solly *et al*., 2017). While both are expressed in meristems and the central zone of the gemmae, Mp*ARF3* is expressed at a higher level than Mp*MIR160* (Fig. 2a–d). By day 6, Mp*ARF3* expression is also associated with air pore

differentiation (Fig. 5d), while Mp*MIR160* remains meristematic (Fig. 2e–h). In mature vegetative thalli the $_{pro}$Mp*ARF3*:*GUS* expression pattern (Fig. S7a) follows the mid-rib, similar to the expression pattern observed for the auxin biosynthesis gene Mp*YUC2* (Eklund *et al*., 2015). Finally, $_{pro}$Mp*MIR160:GUS* lines also exhibit staining along the mid-rib and in dormant gemmalings (Fig. S7b). Widespread expression of Mp*ARF3* in the thallus including dormant gemmae, smooth rhizoids and anteridiophores was also observed using *in situ* hybridization (Fig. S7c–h). Low concentrations of exogenous auxin (1 μM 2,4-D) inhibited $_{pro}$Mp*ARF3*:*GUS* signal in the central zone and differentiating air pores after 7 d of growth compared to mock controls (Fig. 2i–l). This was further corroborated by semi-qPCR, where auxin induced tissue had no detectable Mp*ARF3* expression by semi-qRT-PCR (Fig. 2m).

Mp*ARF3* is regulated by miR160

Phylogenetic analysis revealed that the *miR160* binding site evolved in the common ancestor of embryophytes (Figs 1, S4). We confirmed cleavage of the mRNA derived from the single *Marchantia* class C ARF, Mp*ARF3* (Mapoly0043s0098), between bases 10/11 of the miR160 binding site (Fig. 3a), suggesting the conserved miR160 binding site is a cleavage site across all land plants (Fig. 3b). We identified a single Mp*MIR160* (Mapoly0002s0211) genomic locus, encoding a 246 stem loop within a 2.1-kb transcript (Fig. S8) that has previously been used as a backbone to generate artificial miRNAs (Flores-Sandoval *et al*., 2016).

Given these tools, we created loss of Mp*ARF3* activity via ectopic or overexpression of Mp*MIR160* (Fig. S9a) and via CRISPR-Cas9 generated (Fig. S9b) Mp*ARF3* mutant alleles (Table 1; Fig. S10). For constitutive ectopic overexpression of Mp*MIR160* we utilized the EF1-α promoter, which drive expression in most, if not all, cells at high levels (Althoff *et al*., 2014). Conversely, we created gain of Mp*ARF3* activity via CRISPR-Cas9 generated Mp*MIR160* mutant alleles and via ectopic or over-expression of Mp*ARF3*, either a wild-type

allele or a miR160-resistant Mp*ARF3*$^{m}$ transcript (Table 1). We obtained three independent Mp*MIR160*-CRISPR alleles by co-expressing two gRNAs targeting the primary stem loop (Fig. 3d). Deletions affecting stem loop secondary structure (Mp*mir160-2*), miR/miR* sequences (Mp*mir160-3*) or both secondary structure/miR* (Mp*mir160-4*) were obtained. (Figs 3d,k, S11). qRT-PCR detection of mature miR160 levels relative to the highly expressed *miR166*, showed all three mutant lines failed to produce mature *miR160* (Fig. 3e). Predicted stem loop secondary structures (Zuker, 2003) in all mutants were different to wild-type (Fig. S9g).

Mp*ARF3* facilitates multiple developmental processes

Given the observation that auxin response and auxin biosynthesis mutants have a complete loss of three-dimensional patterning (Eklund *et al*., 2015; Flores-Sandoval *et al*., 2015; Kato *et al*., 2015), and the previously proposed ideas that ARFs work as facilitators and not determinants of developmental transitions (Stewart and Nemhauser, 2010; Bennett and Leyser, 2014; Flores-Sandoval *et al*., 2015), we hypothesized that loss of Mp*ARF3* activity would result in pleiotropic defects in multiple developmental processes. Indeed, nearly every aspect of growth and development was affected when Mp*ARF3* activity was changed.

**Growth** Loss of Mp*ARF3* activity resulted in significant decreases in surface area and apical notches in both $_{pro}$*EF1:*Mp*MIR160* and putative null lines. After 14 d of growth, four independent populations of $_{pro}$*EF1:*Mp*MIR160* lines had grown to *c*. 1/4 of the size of wild-type controls (Figs 3f,i,j, S9c–h). Notch production (branching) is reduced to ~1/2 in $_{pro}$*EF1:*Mp*MIR160* lines after 14 d (Fig. 3g). Similar proportions are observed between wild-type and Mp*arf3* null alleles after ~2 months of growth which will be discussed in detail in the complementation analysis section. Conversely, increased Mp*ARF3* activity in Mp*MIR160* loss-of-function alleles (Mp*mir160*) or expression of a miR160-resistant Mp*ARF3* allele

($_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$), results in proliferative outgrowths where thallus maturation and growth appear decoupled (Fig. 3k,l) and the regular bifurcation-cup-bifurcation branching pattern observed in the wild-type is lacking (Fig. S11a–d). Scanning electron microscopy (SEM) revealed multiple layers of stacked branches in Mp*mir160* alleles compared to the wild-type (Fig. S11e,f).

**Regeneration** Regeneration in complex thalloids has been studied extensively, with the locations of regenerant tissue reflecting that predicted for the higher auxin concentrations (Vöchting, 1885; Dickson, 1932; Binns & Maravolo, 1972; Gaal *et al*., 1982). In *Marchantia*, excision of apical notches triggers *de novo* formation of apical meristems (Fig. 3h). The newly formed thallus, transitions from immature to mature stages and subsequently branches (Fig. 3h). Comparisons of regenerating branches in wild-type, Mp*arf3* and $_{pro}$*EF1*:Mp*MIR160*, demonstrate that loss of Mp*ARF3* delays the regeneration process (Fig. 3i,j). Lines in which a CRISPR-resistant version of Mp*ARF3* (Mp*ARF3*$^{Cas9Res}$) complemented an Mp*arf3* mutant background, showed a capacity to regenerate, grow and branch, similar to wild-type (Fig. 3m), but complementation using the *Arabidopsis* class C orthologue At*ARF10* restored growth but not wild-type levels of regeneration (Fig. 3n). Complementation analysis will be discussed in detail in the following sections.

**Gemmae** A significant reduction in gemmae production was observed in $_{pro}$*EF1*:Mp*MIR160* lines (Fig. 4a,b,e). In putative null Mp*arf3* alleles, no gemmae were produced although gemmae cup production was not supressed, leaving plants with empty cups (Fig. 4c,d). Mp*mir160* alleles produce fewer gemma cups (Fig. S11a–d) and thus, robust measurements of gemmaling production are not feasible in such alleles. This suggests that *MpARF3* is necessary for the production of clonal undifferentiated propagules.

**Differentiation of air chambers and pores** The dorsal surface of wild-type thalli is covered with a regularly spaced pattern of air chambers with a centrally located air pore (Mirbel,

1835; Barnes & Land, 1907; Shimamura, 2016). These structures develop into mature thalli but are not present in the earlier prothallus stage, which is characterized by two-dimensional growth. We quantified thallus differentiation in Mp*ARF3* loss- and gain-of-function alleles by measuring number of air pores per area ($mm^2$). After 14 d of growth (Fig. 4f), wild-type gemmalings produced an average of 406 air pores in a 45 $mm^2$ area, equivalent to ˜9 pores $mm^{-2}$. Mp*miR160-2* plants produced about half as many (˜4 pores $mm^{-2}$, $P = 3.74 \times 10^{-11}$), while $_{pro}$*EF1:*Mp*MIR160* plants produced twice as many as wild-type (˜15 pores $mm^{-2}$, $P = 0.0002$). SEM of $_{pro}$*EF1:*Mp*MIR160* and Mp*arf3* lines revealed smaller cells and air chambers in a qualitative fashion (Figs 4g,h, S12a,b). Finally, examination of mature Mp*mir160-2* epidermis revealed sections without air chambers (Fig. 4i) in comparison to the homogenous pattern observed in wild-type (Fig. 4g). This suggests that *MpARF3* is a repressor of air chamber differentiation.

**Rhizoids and scales** The ventral surface of the *M. polymorpha* mature thallus has two types of rhizoids and two or more rows of scales flanking the midrib and two types of rhizoids (Fig. 4j). Smooth rhizoids are produced along the midrib and act in nutrient uptake, while pegged rhizoids, dead cells at maturity, are associated with the ventral scales and act as capillary water conduits (McConaha, 1941; Duckett *et al*., 2014). The smooth (living) rhizoids are formed at a distance from the apex (SR in Fig. 4j) and thus are mostly absent at the distal end of the thallus (Proust *et al*., 2016; Shimamura, 2016). Conversely, scales and pegged rhizoids (thin and structural) are visible and pronounced in the distal ends of the thallus (asterisk in Fig. 4j). $_{pro}$*EF1:*Mp*MIR160* plants seem to have a decreased number of smooth rhizoids along the midrib, exposing scales and pegged rhizoids (Fig. S12c,d). Moreover, putative null Mp*arf3* alleles, form conspicuous bundles of pegged rhizoids along the ventral midrib and adjacent sites (Fig. 4k,l). Conversely, Mp*mir160-4* alleles showed profuse formation of smooth rhizoids, even visible from a dorsal angle (Fig. 4m).

Mp*arf3* alleles have increased number, size and complexity of meristematic scales protecting the apical notch in comparison to wild-type (Fig. 4n,o). Meanwhile, Mp*mir160-2* did not produce meristematic scales (Fig. 4p). Mp*arf3* plants displayed hyponastic growth (Figs 4q, S12f), possibly due to protrusion of ventral scales visible in the dorsal surface, this could be due to over-growth or over-production (Figs 4q, S12e,f). Qualitative observations of ventral scales revealed that Mp*arf3* alleles have more ventral scales per area than wild-type (Fig. S12g,h), suggesting Mp*ARF3* represses production of scales.

**Gametophores** Wild-type plants subjected to a far-red light (FRL) treatment undergo a developmental transition with some branches differentiating into sexual gametophores: antheridiophores develop antheridia that produce sperm cells; and archegoniophores develop archegonia that produce egg cells. After 1 month of FRL treatment, wild-type plants produced on average ~10 gametophores from a ~40 $cm^2$ thallus **(**Fig. S13a,b,d**).** Both Mp*arf3-2* and Mp*arf3-12* lines produced on average 8 gametophores from a ~10 $cm^2$ thallus, suggesting rates of gametophore production were retained in Mp*arf3* mutants despite growth deficiencies (Fig. S13b,d,e). However, Mp*mir160-3* and Mp*mir160-4* alleles were completely insensitive to FRL treatment with respect to gametophore production from a ~50 $cm^2$ thalli (Fig. S13c–e), suggesting *MpARF3* inhibits the developmental transition from asexual to sexual reproduction. Plants with partial loss of Mp*ARF3* ($_{pro}$*EF1:*Mp*MIR160)* displayed ectopic antheridia protruding from the dorsal surface of the antheridiophore (Fig. 4r,s). In addition, female $_{pro}$*EF1:*Mp*MIR160* plants produced archegoniophores exhibiting varying degrees of stalk fusion (Fig. 4t,u).

Ectopic Mp*ARF3* expression results in less differentiated states

We created gain-of-function Mp*ARF3* alleles expressing wild-type (Mp$ARF3^{WT}$) and miR-resistant versions of Mp*ARF3* (Mp$ARF3^{m}$), using a range of promoters; ectopic expression of

Mp*ARF3*$^{m}$ consistently resulted in stronger phenotypes than Mp*ARF3*$^{WT}$ controls (Fig. 5a). We first used 3.6kb of the putative Mp*ARF3* endogenous regulatory sequences (see Fig. S14). Phenotypes of $_{pro}$Mp*ARF3:*Mp*ARF3* plants ranged from almost wild-type, to folded thalli (Fig. 5a,b), suggesting dosage effects of Mp*ARF3*, or alternatively, the putative promoter lacks some regulatory elements that restrain the wild-type Mp*ARF3* expression pattern. Expression of the miR160 resistant version, $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$, resulted in a prothallus-like phenotype (Fig. 5a,c,d), with thin layers of tissues lacking air chambers and pores (Fig. 5d). qRT-PCR was used to verify increased levels in multiple $_{pro}$Mp*ARF3:*Mp*ARF3* and $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$ lines compared to controls (Fig. 5m). $_{pro}$Mp*ARF3:*Mp*ARF3* lines have the highest levels of Mp*ARF3* expression despite having a weaker phenotype (Fig. 5m).

We next restricted additional Mp*ARF3* activity to a smaller expression domain limited to the apical meristem using the regulatory region of Mp*SHI/STY* (Mapoly0049s0083) (Flores-Sandoval *et al*., 2015). The genotypes $_{pro}$*SHI:*Mp*ARF3* and $_{pro}$*SHI:*Mp*ARF3*$^{m}$ do not appear phenotypically distinct, despite Mp*MIR160:GUS* also showing expression in the apical notch, where Mp*ARF3*$^{WT}$ is a target for cleavage (Fig. 5e–g). Both genotypes develop into highly branched plants that formed air chambers and pores, but no gemmae cups (Fig. 5h).

Constitutive expression of wild-type Mp*ARF3* ($_{pro}$*EF1*:Mp*ARF3*) resulted in plants similar to Mp*mir160* lines, with highly branched convoluted thalli, often lacking air chambers and pores (Fig. 5i,j). Constitutive expression of miR-resistant Mp*ARF3* ($_{pro}$*EF1:*Mp*ARF3*$^{m}$) resulted in a loss of differentiation (Fig. 5i,k,l) similar to wild-type sporeling cells (Fig. S6a) and to *Marchantia* plants lacking the canonical auxin biosynthetic pathway (Eklund *et al*., 2015). Since plants constitutively expressing Mp*ARF3*$^{m}$ resemble those lacking the canonical land plant auxin biosynthetic pathway, the auxin sensitivity of

loss- and gain-of-function Mp*ARF3* alleles was tested by growth on 2,4-D (10 µM). While $_{pro}$Mp*ARF3:*Mp*ARF3* lines are sensitive to exogenous auxin both in terms of growth inhibition and rhizoid production, $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$ lines are insensitive with respect to growth inhibition (Figs 5n, S15a). $_{pro}$*EF1*:Mp*ARF3*$^{m}$ lines show the cleanest auxin resistance when grown in NAA (5 µM), compared to $_{pro}$*EF1*:Mp*ARF3* and wild-type plants after 2 and 4 wk of growth (Fig. S15b). Similar to $_{pro}$Mp*ARF3:*Mp*ARF3* lines, Mp*mir160-3* gemmalings were sensitive to auxin (Fig. S16), suggesting that Mp*ARF3* can antagonize auxin signalling only at the highest doses. Unlike Mp*arf1* loss-of-function alleles (Sugano *et al*., 2014; Kato *et al*., 2017) or $_{pro}$*EF1:amiR*Mp*ARF1* lines (Flores-Sandoval *et al*., 2016), both $_{pro}$*EF1:*Mp*MIR160* (Fig. S17a) and Mp*arf3* (Fig. S17b) lines are sensitive to exogenous auxin (Fig. 5n).

Transcriptional compensatory feedbacks between Mp*MIR160* and Mp*ARF3*

To examine potential feedback regulation between Mp*ARF3* and Mp*MIR160*, we performed RNA-seq on mature 1-month old thallus in three independent WT, Mp*arf3* and Mp*mir160* lines. Mp*ARF3* transcripts are significantly upregulated in Mp*arf3* (~2x) (Fig. 6a; Table S2). Despite this upregulation, no reads were mapped to the deletions sites, suggesting activation of aberrant transcripts (Fig. 6e). As expected, Mp*ARF3* is significantly upregulated in Mp*mir160* mutants (1.78x) (Fig. 6a). Meanwhile, Mp*MIR160* transcripts are significantly downregulated in Mp*arf3* (~0.2x) and upregulated eight-fold in Mp*mir160* mutants (Fig. 6b). Density plots revealed that in wild type, Mp*MIR160* reads accumulate 3' to the miR/miR* region, with lack of 5' reads indicative of *DICER* processing (Fig. 6c). In Mp*mir160* alleles, raw reads are detected 5’ of the miR/miR* site, suggesting a lack of *DICER* processing and accounting for the unexpected transcriptional increase (Fig. 6c). Furthermore, Mp*MIR160* transcription appears dependent on Mp*ARF3* activity, suggesting Mp*ARF3* activates its own

repressor (Fig. 6c). Supporting this, qRT-PCR shows that Mp*miR160* is upregulated in several $_{pro}$Mp*ARF3:*Mp*ARF3* and $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$ lines compared to wild type (Fig. 6d). These results suggest that steady state feedback loops in Mp*ARF3* and Mp*MIR160* act to compensate for the loss of functional transcripts.

Transcriptional upregulation of auxin induced genes in Mp*arf3* and Mp*mir160*

To test whether auxin signalling depends on Mp*ARF3* activity, semi-quantitative RT-PCR was used to asses a set of recently discovered (Kato *et al*., 2017; Mutte *et al*., 2017) auxin-induced transcripts in wild-type, Mp*arf3* and Mp*miR160* alleles. In wild type, we were able to detect upregulation of Mp*EXPANSIN125* (Mapoly0125s0001) after 48 h of auxin treatment (Fig. S18a). Upregulation of Mp*EXPANSIN125* by auxin is not compromised in Mp*arf3-12* and Mp*arf3-22* alleles. Furthermore, Mp*EXPANSIN125* is upregulated in Mp*miR160-4* in the absence of auxin compared to wild-type controls (Fig. S18a). A similar scenario is observed for Mp*C2HDZ* (Mapoly0069s0069) (Fig. S18b). Differential count analysis (DESeq2 and edgeR) using RNAseq revealed that Mp*C2HDZ*, Mp*EXPANSIN125* and Mp*EXPANSIN61* (Mapoly0061s0098) were significantly down-regulated in Mp*arf3* mutants but not significantly changed in Mp*miR160* alleles compared to the wild-type (Fig. S18c). These results suggest that Mp*ARF3* positively regulates auxin-induced genes in an auxin independent fashion.

Complementation of Mp*arf3* reveals functional conservation of embryophyte class C ARFs

We performed complementation assays to (1) corroborate that the Mp*arf3* phenotypes are linked to the observed Mp*ARF3* mutations, (2) test the capacity of the cloned 3.6kb $_{pro}$Mp*ARF3* to drive a natural expression pattern, and (3) evaluate the level of functional

conservation of class C ARF proteins from angiosperms and liverworts. We co-expressed dual gRNAs targeting the endogenous Mp*ARF3* locus with either Cas9-resistant Mp*ARF3*$^{Cas9Res}$ or At*ARF10*, one of the three redundant *Arabidopsis* class C ARF paralogs, using three selectable markers in primary sporeling transformants (see materials and methods). After ~2 months, wild-type plants had grown to an average area of 30 cm$^2$ (Fig. 7a,e), while Mp*arf3* mutants had grown to approx. ¼ that size (Fig. 7b,e). Complementation with $_{pro}$Mp*ARF3:*Mp*ARF3*$^{Cas9Res}$ (Fig. 7c,e) completely restores growth (Student's t test, P = 0.006), while $_{pro}$Mp*ARF3:*At*ARF10* (Fig. 7d,e) complementation lines restored only 60% of the wild-type growth (Student's *t* test, *P* = 0.028). Branching was significantly restored by both $_{pro}$Mp*ARF3:*Mp*ARF3*$^{Cas9Res}$ and $_{pro}$Mp*ARF3:*At*ARF10*, although At*ARF10* complementation lines displayed only a 75% rescue (Fig. 7f). Even without full gemma cup restoration (Fig. 7g), 70% of cups restored gemmae production using Mp*ARF3*$^{Cas9Res}$ (Student's *t* test, *P*=0.048), while only 8% of cups did using At*ARF10* (Fig. 7h). Our results point at a nearly full complementation by Mp*ARF3*$^{Cas9Res}$ but only partial complementation by At*ARF10*.

Other conserved PB1 domain containing proteins can modulate auxin signalling

Two clades of conserved genes distinct from AUX/IAA genes possess PB1 domains, but lack the B3 domain (Fig. 1). One, NCIAA, has its origins in the charophycean algae while the other NCARF, is of more recent origin within land plants. To test whether MpNCARF and MpNCIAA PB1 domains has the potential to interact with MpARFs, we translationally fused PB1$^{MpNCARF}$ and PB1$^{MpNCIAA}$ to the endogenous MpTPL co-repressor. Just as in $_{pro}$*EF1:*Mp*TPL-PB1*$^{MpIAA}$ controls, $_{pro}$*EF1:*Mp*TPL-PB1*$^{Mp\textit{NCARF}/Mp\textit{NCIAA}}$ plants were unable to transition into a three-dimensional thallus, and instead remaining reminiscent of younger developmental prothallus stages. This suggested that unlike MpARF3 (Flores-Sandoval *et al*.,

2015), MpNCARF and MpNCIAA have the capacity to interact with some of the elements of the *M. polymorpha* auxin signalling pathway (Fig. S19).



**Discussion**

Evolution of the ARF and AUX/IAA gene families in streptophytes

While both ARF and AUX/IAA genes encode PB1 domains, the precise evolutionary relationship between these gene families has been enigmatic due to lack of an appropriate PB1-containing outgroup (De Smet *et al*., 2011; Finet *et al*., 2013). Both charophycean algal and liverwort RAV orthologs possess both B3 and PB1 domains, thus providing a potential outgroup. Using the RAV genes as an outgroup, the three ARF classes form a single clade sister to one that contains the AUX/IAA genes and a previously unidentified clade of genes containing PB1 domains but lacking B3 domains, here called NCIAA. The NCIAA clade has persisted in all major lineages of land plants, but have not previously been recognized as distinct from AUX/IAA genes, and have not yet been investigated functionally. The observation that charophycean algal sequences are present in each of the AUX/IAA, NCIAA, C ARF, and A/B ARF+NCARF clades indicates that these gene classes were present in an algal ancestor of land plants. By contrast, given their phylogenetic distribution, the A ARF, B ARF, and NCARF genes may have evolved in an ancestral land plant via gene duplications from an ancestral A/B ARF gene. NCARF genes, which also lack a B3 domain, were lost in the seed plant ancestor. Translational fusions of the PB1 domains from either Mp*NCARF* or Mp*NCIAA* with MpTPL results in similar phenotypes as those observed when the PB1 domains of MpIAA, MpARF1 or MpARF2 are translationally fused to MpTPL (Flores-Sandoval *et al*., 2015), suggesting that both proteins may modulate auxin signalling. The major assumption we have made in our phylogenetic analysis is the choice of RAV sequences as an outgroup. In other approaches lacking this assumption, e.g. midpoint rooting

or an unrooted tree, the relationships between the major clades (RAV, AUX/IAA+NCIAA, ARF) are ambiguous, but their times of origin remain unchanged (Fig. S1e)**.**

Comparative analyses across charophycean algae and land plants indicate that the canonical land plant auxin transcriptional signalling system is a land plant innovation (Hori *et al*., 2014; Bowman *et al*., 2017). This conclusion is based on the observation that charophycean AUX/IAA orthologs encode proteins lacking both domains I and II required for interaction with TIR1, auxin, and the co-repressor TOPLESS, and that furthermore, charophycean TIR1/COI orthologs lack specific amino acids required to interact with AUX/IAA proteins and auxin (Bowman *et al*., 2017). Thus, it was proposed that assembly of the canonical land plant auxin signalling system involved neo-functionalization of both AUX/IAA and TIR1 in the ancestral land plant (Bowman *et al*., 2017). By contrast, at least two classes of ARFs (C, A/B), and two classes of potential ARF regulators (AUX/IAA, NCIAA) already existed in a charophycean algal ancestor before the assembly of the ARFs and AUX/IAAs into an auxin signalling system. This implies that these ARF classes and related proteins acted in an ancestral alga in a transcriptional network whose function is as yet unknown. Finally, the evolutionary scenario presented here implies that auxin and miR160 were integrated as repressors into a pre-existing transcriptional network in the ancestral land plant.

## Mp*ARF3* inhibits differentiation with pleiotropic roles in development

We chose to characterize the class C ARF, Mp*ARF3*, due to its ancient evolutionary origin, i.e. before the evolution of the canonical auxin signalling pathway. Both loss- and gain-of-function alleles of Mp*ARF3* result in pleiotropic defects throughout the gametophytic stages of the life cycle that do not correlate with specific changes in cell or tissue identity. However, a consistent pattern involving Mp*ARF3* as a negative regulator of differentiation or a

promoter of cell proliferation emerges. First, the smaller cell size in Mp*arf3* mutants could be interpreted as a loss of coordination between cell proliferation and differentiation. Second, despite having delayed growth, Mp*arf3* plants differentiate more air pores per area than the wild-type, produce increased numbers of pegged rhizoids and ventral scales, and produce more gametophores per thallus area than wild-type. Both the formation of ectopic antheridia and fused archegoniophores could be interpreted as defects caused by differences in tissue growth rates. Conversely, Mp*mir160* alleles seem to lack proper differentiation of air pores, gemmae cups, meristematic scales and gametophores upon inductive (FRL) conditions. Third, Mp*ARF3* is required for regeneration of thalloid tissue (Fig. 3) and formation of undifferentiated gemmalings (Fig. 4). Finally, increasing levels of ectopic MpARF3 expression transform plants into more juvenile states whose morphology resembles that of wild-type prothalli, sporelings, and eventually a mass of cells lacking tissue patterning (Fig. 5). These observations are consistent with the idea that ARFs facilitate rather than determine developmental processes (Stewart and Nemhauser, 2010; Bennett and Leyser, 2014; Flores-Sandoval *et al*., 2015), adding a role for class C ARFs as specific modulators of differentiation rates and totipotency. One speculative hypothesis is that Mp*ARF3* acts in shifting the balance between meristematic and differentiated states in a cell or tissue type dependent manner. These processes, proliferation and differentiation, occurred in the ancestral alga and might represent reasonable candidates for the function of the ARF transcriptional network in extant charophycean algae. In this scenario, the integration of auxin or miR160 into an ancestral regulatory network may have led to focal growth (meristems) rather than more diffuse growth and differentiation.

Given the antiquity of class C ARFs and their minimal protein-protein interactions with other classes of ARF and AUX/IAAs, it is unknown how, or even whether, class C ARFs function directly in the auxin transcriptional pathway. Mp*arf3* alleles retain the ability

to respond to auxin, indicating that, unlike Mp*ARF1*, auxin signalling does not depend on Mp*ARF3* activity (Figs 5, S15). Furthermore, the auxin-signalling pathway appears to inhibit Mp*ARF3* transcription, and thus, activity (Fig. 2). By contrast, strong gain-of-function miR160 resistant Mp*ARF3* alleles can antagonize auxin signalling: $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$ alleles phenocopy $_{pro}$*EF1:*Mp*TPL-PB1*$^{MpIAA}$ alleles in which repressive regulation of interacting ARF partners is decoupled from auxin signalling (Flores-Sandoval *et al*., 2015), and $_{pro}$*EF1:*Mp*ARF3*$^{m}$ alleles phenocopy $_{pro}$*EF1:amiR*Mp*TAA*$^{MpmiR160}$ alleles in which the IPyA-biosynthetic pathway is compromised (Eklund *et al*., 2015). These observations raise the possibility that most, if not all, of the defects observed in Mp*taa* mutants might be attributed to the loss of auxin transcriptional signaling.

The ability of Mp*ARF3* to repress auxin response in a dose dependent fashion raises the question of whether the different ARF classes may antagonize and share some common downstream targets. Counterintuitively, some conserved auxin-induced genes are upregulated or dependant on Mp*ARF3* activity in certain developmental or environmental conditions (Fig. S18). However, Mp*arf1* loss-of-function mutants do not phenocopy Mp*ARF3* gain-of-function mutants and vice versa (Kato *et al* 2017), suggesting any antagonistic relationship depends on interactions with other transcriptional networks. One hypothesis consistent with these observations is that ARF target genes may be regulated, in conjunction with other ARF-unrelated transcription factors, positively (via Mp*ARF1*) or negatively (via Mp*ARF2*) in an auxin-dependent manner, with Mp*ARF3* antagonizing auxin signaling in an auxin-independent manner. In this scenario the evolution of A and B class ARFs in an ancestor of land plants may have been critical in facilitating both positive and negative auxin-dependent gene regulation.

C class ARF activity is conserved across land plants

The ancestral land plant C class ARF was regulated by miR160, but miR160 target sites are not evident in any of the charophycean algal C class ARF genes. In *M. polymorpha*, both Mp*ARF3* and Mp*MIR160* exhibit feedback regulation, with both genes upregulated in their respective mutant backgrounds. Such compensatory regulation might tune the regulation of Mp*ARF3* activity as tissues differentiate from meristematic regions.

The activity of C class ARF proteins is largely conserved across land plants, as attested by the ability of At*ARF10* to partially complement Mp*arf3* loss-of-function alleles when driven by Mp*ARF3* regulatory sequences. Growth and branching were restored to a large extent when Mp*arf3* mutants were complemented with either Mp*ARF3* or At*ARF10*, but the heterologous gene failed to complement the ability to regenerate or the formation of gemmae. It is of note that the latter two attributes, vegetative propagules and regeneration, have been lost to a great extent in derived land plants.

**Acknowledgements**

Funding from ARC DP130100177, DP160100892 (J.L.B). Monash Micromon provided library preparation, RNA-sequencing and Sanger sequencing. Monash Micro Imaging provided SEM facilities. Takayuki Kohchi provided GE010 and GWB301 plasmids for CRISPR constructs. Kimizune Ishizaki provided pGWB404.

**Author contributions**

E.F-S. and J.L.B. designed the experiments, assembled the figures and wrote the manuscript. E.F-S. and D.M.E. did cloning and production of all lines in figures. S-F.H. and S-S.L. performed qRT-PCR, T.J.F. and E.F-S. worked on reporter lines and SEM. E.R.L and J.F.G performed *in situ* hybridization. T.D. did RLM-RACE. A.V-L. and J.L.B. performed

phylogenetic analysis. E.F-S. and T.J.F. performed RNA-seq analysis. E.F-S and J.P.A performed complementation analysis. All authors analysed the data.



**ORCID**

D. Magnus Eklund http://orcid.org/0000-0002-0576-7636

John L. Bowman http://orcid.org/0000-0001-7347-3691

**References**

Afgan, E., Sloggett, C., Goonasekera, N., Makunin, I., Benson, D., Crowe, M., Gladman, S., Kowsar, Y., Pheasant, M., Horst, R. *et al*. (2015) 'Genomics virtual laboratory: a practical bioinformatics workbench for the cloud', *PLoS ONE* 10(10): e0140829.

Althoff, F., Kopischke, S., Zobell, O., Ide, K., Ishizaki, K., Kohchi, T. and Zachgo, S. (2014) 'Comparison of the MpEF1α and CaMV35 promoters for application in Marchantia polymorpha overexpression studies', *Transgenic Research* 23: 234-244.

Anders, S., Pyl, P. T. and Huber, W. (2015) 'HTSeq-A Python framework to work with high-throughput sequencing data', *Bioinformatics* 31(166–169).

Axtell, M. J., Snyder, J. A. and Bartel, D. P. (2007) 'Common functions for diverse small RNAs of land plants', *The Plant Cell* 19(6): 1750-69.

Banks, J. A. Nishiyama, T. Hasebe, M. Bowman, J. L. Gribskov, M. dePamphilis, C. Albert, V. A. Aono, N. Aoyama, T. Ambrose, B. A. *et al*. (2011) 'The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants', *Science* 332(6032): 960-963.

Barnes, C. R. and Land, W. J. G. (1907) 'Bryological papers I. The origin of air chambers - Contributions from the Hull Botanical Laboratory, 100', *Botanical Gazette* 44: 197-213.

Ben-Gera, H., Dafna, A., Alvarez, J. P., Bar, M., Mauerer, M. and Ori, N. (2016) 'Auxin-mediated lamina growth in tomato leaves is restricted by two parallel mechanisms', *Plant Journal* 86(6): 443-457.

Bennett, T. and Leyser, O. (2014) The auxin question: a philosophical overview. In: Zažímalová E, Petrášek J, Benková E. eds. *Auxin and its role in plant development*. Vienna, Austria: Springer, 3-19.

Bennett, T., van den Toorn, A., Willemsen, V. and Scheres, B. (2014) 'Precise control of plant stem cell activity through parallel regulatory inputs', *Development* 141(21): 4055-4064.

Binns, A. N. and Maravolo, N. C. (1972) 'Apical dominance, polarity, and adventitious growth in *Marchantia polymorpha*', *American Journal of Botany* 59(7): 691-696.

Boer, D. R., Freire-Rios, A., van den Berg, W. A. M., Saaki, T., Manfield, I. W., Kepinski, S., Lopez-Vidrieo, I., Franco-Zorrilla, J. M., de Vries, S. C., Solano, R *et al*. (2014) 'Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors', *Cell* 156(3): 577-589.

Bowman, J. L., Araki, T., Arteaga-Vazquez, M. A., Berger, F., Dolan, L., Haseloff, J., Ishizaki, K., Kyozuka, J., Lin, S. S., Nagasaki, H. *et al*. (2016) 'The naming of names: guidelines for gene nomenclature in *Marchantia*', *Plant and Cell Physiology* 57(2): 257-261.

Bowman, J. L. Kohchi, T. Yamato, K. T. Jenkins, J. Shu, S. Ishizaki, K. Yamaoka, S. Nishihama, R. Nakamura, Y. Berger, F. *et al*. (2017) 'Insights into land plant evolution garnered from the *Marchantia polymorpha* genome', *Cell* 171: 287-304.

Cooper, E. and Delwiche, C. (2016) Green algal transcriptomes for phylogenetics and comparative genomics. figshare.https://dx.doi.org/10.6084/m9.figshare.1604778.

Damodharan, S., Zhao, D. Z. and Arazi, T. (2016) 'A common miRNA160-based mechanism regulates ovary patterning, floral organ abscission and lamina outgrowth in tomato', *Plant Journal* 86(6): 458-471.

De Smet, I., Voss, U., Lau, S., Wilson, M., Shao, N., Timme, R. E., Swarup, R., Kerr, I., Hodgman, C., Bock, R. *et al*. (2011) 'Unraveling the evolution of auxin signaling', *Plant Physiology* 155(1): 209-221.

Dharmasiri, N., Dharmasiri, S. and Estelle, M. (2005) 'The F-box protein TIR1 is an auxin receptor', *Nature* 435(7041): 441-445.

Dickson, H. (1932) 'Polarity and the production of adventitious growing points in *Marchantia polymorpha*.', *Annals of Botany* 46(183): 683-701; Pl. XXV.

Ding, Z. J. and Friml, J. (2010) 'Auxin regulates distal stem cell differentiation in Arabidopsis roots', *Proceedings of the National Academy of Sciences, USA* 107(26): 12046-12051.

Duckett, J. G., Ligrone, R., Renzaglia, K. and Silvia, P. (2014) 'Pegged and smooth rhizoids in complex thalloid liverworts (Marchantiopsida): structure, function and evolution', *Botanical Journal of the Linnean Society* 174: 68-92.

Eklund, D. M., Ishizaki, K., Flores-Sandoval, E., Kikuchi, S., Takebayashi, Y., Tsukamoto, S., Hirakawa, Y., Nonomura, M., Kato, H., Kouno, M. *et al*. (2015) 'Auxin produced by the indole-3-pyruvic acid pathway regulates development and gemmae dormancy in the liverwort *Marchantia polymorpha*', *The Plant cell* 27(6): 1650-1669.

Finet, C., Berne-Dedieu, A., Scutt, C. P. and Marletaz, F. (2013) 'Evolution of the ARF gene family in land plants: old domains, new tricks', *Molecular Biology and Evolution* 30: 45-56.

Flores-Sandoval, E., Dierschke, T., Fisher, T. J. and Bowman, J. L. (2016) 'Efficient and inducible use of artificial microRNAs in *Marchantia polymorpha'*, *Plant and Cell Physiology* 57: 281-290.

Flores-Sandoval, E., Eklund, D. M. and Bowman, J. L. (2015) 'A simple auxin transcriptional response system regulates multiple morphogenetic processes in the liverwort *Marchantia polymorpha'*, *PLOS Genetics* 11(5): e1005207.

Gaal, D. J., Dufresne, S. J. and Maravolo, N. C. (1982) 'Transport of C-14-labeled indoleacetic-acid in the hepatic *Marchantia polymorpha'*, *Bryologist* 85(4): 410-418.

Gamborg, O. L., Miller, R. A. and Ojima, K. (1968) 'Nutrient requirements of suspension cultures of soybean root cells', *Experimental Cell Research* 50: 151–158.

Gray, W. M., Kepinski, S., Rouse, D., Leyser, O. and Estelle, M. (2001) 'Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins', *Nature* 414(6861): 271-6.

Guilfoyle, T. J. (2015) 'The PB1 domain in auxin response factor and Aux/IAA proteins – a versatile protein interaction module in the auxin response', *The Plant Cell* 27: 33-43.

Hendelman, A., Buxdorf, K., Stav, R., Kravchik, M. and Arazi, T. (2012) 'Inhibition of lamina outgrowth following *Solanum lycopersicum AUXIN RESPONSE FACTOR 10* (*SlARF10*) derepression', *Plant Molecular Biology* 78(6): 561-576.

Hori, K., Maruyama, F., Fujisawa, T., Togashi, T., Yamamoto, N., Seo, M., Sato, S., Yamada, T., Mori, H., Tajima, N. *et al*. (2014) *'Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation', *Nature Communications* 5:3978

Huelsenbeck, J. P. and Ronquist, F. (2001) 'MRBAYES: Bayesian inference of phylogenetic trees', *Bioinformatics* 17(8): 754-5.

Ishizaki, K., Chiyoda, S., Yamato, K. T. and Kohchi, T. (2008) 'Agrobacterium-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology', *Plant & cell physiology* 49(7): 1084-91.

Ishizaki, K., Nishihama, R., Ueda, M., Inoue, K., Ishida, S., Nishimura, Y., Shikanai, T. and Kohchi, T. (2015) 'Development of Gateway Binary Vector Series with four different selection markers for the liverwort *Marchantia polymorpha*', *PLoS ONE* 10: 1-13; e0138876.

Jefferson, R. A., Kavanagh, T. A. and Bevan, M. W. (1987) 'GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants', *EMBO J.* 6: 3901-3907.

Ju, C. L., Van de Poel, B., Cooper, E. D., Thierer, J. H., Gibbons, T. R., Delwiche, C. F. and Chang, C. R. (2015) 'Conservation of ethylene as a plant hormone over 450 million years of evolution', *Nature Plants* 1(1): 14004.

Kato, H., Ishizaki, K., Kouno, M., Shirakawa, M., Bowman, J. L., Nishihama, R. and Kohchi, T. (2015) 'Auxin-mediated transcriptional system with a minimal set of components is critical for morphogenesis through the life cycle in *Marchantia polymorpha*', *PLoS Genetics* 11(5): e1005084.

Kato, H., Kouno, M., Takeda, M., Suzuki, H., Ishizaki, K., Nishihama, R. and Kohchi, T. (2017) 'The roles of the sole activator-type auxin response factor in pattern formation of *Marchantia polymorpha*', *Plant and Cell Physiology* 58(10): 1642–1651.

Kepinski, S. and Leyser, O. (2005) 'The Arabidopsis F-box protein TIR1 is an auxin receptor', *Nature* 435(7041): 446-451.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R. and Salzberg, S. L. (2013) 'TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions', *Genome Biology* 14: R36.

Kim, J., Harter, K. and Theologis, A. (1997) 'Protein–protein interactions among the Aux/IAA proteins', *Proceedings of the National Academy of Sciences, USA* 94(22): 11786-11791.

Lavy, M., Prigge, M. J., Tao, S., Shain, S., Kuo, A., Kirchsteiger, K. and Estelle, M. (2016) 'Constitutive auxin response in *Physcomitrella* reveals complex interactions between Aux/IAA and ARF proteins', *eLife* 5: e13325

Lin, P. C., Lu, C. W., Shen, B. N., Lee, G. Z., Bowman, J. L., Arteaga-Vazquez, M. A., Liu, L. Y. D., Hong, S. F., Lo, C. F., Su, G. M. *et al*. (2016) 'Identification of miRNAs and their targets in the liverwort *Marchantia polymorpha* by integrating RNA-Seq and degradome analyses', *Plant and Cell Physiology* 57(2): 339-358.

Liu, P. P., Montgomery, T. A., Fahlgren, N., Kasschau, K. D., Nonogaki, H. and Carrington, J. C. (2007) 'Repression of *AUXIN RESPONSE FACTOR10* by *microRNA160* is critical for seed germination and post-germination stages', *Plant Journal* 52(1): 133-146.

Liu, X. D., Huang, J., Wang, Y., Khanna, K., Xie, Z. X., Owen, H. A. and Zhao, D. Z. (2010) 'The role of floral organs in carpels, an *Arabidopsis* loss-of-function mutation in *MicroRNA160a*, in organogenesis and the mechanism regulating its expression', *Plant Journal* 62(3): 416-428.

Liu, X. D., Zhang, H., Zhao, Y., Feng, Z. Y., Li, Q., Yang, H. Q., Luan, S., Li, J. M. and He, Z. H. (2013) 'Auxin controls seed dormancy through stimulation of abscisic acid signaling by inducing ARF-mediated ABI3 activation in *Arabidopsis*', *Proceedings of the National Academy of Sciences, USA* 110(38): 15485-15490.

Liu, Z. H., Li, J., Wang, L., Li, Q., Lu, Q., Yu, Y. C., Li, S., Bai, M. Y., Hu, Y. X. and Xiang, F. N. (2016) 'Repression of callus initiation by the miRNA-directed interaction of auxin-cytokinin in *Arabidopsis thaliana*', *Plant Journal* 87(4): 391-402.

Mallory, A. C., Bartel, D. P. and Bartel, B. (2005) 'MicroRNA-directed regulation of Arabidopsis *AUXIN RESPONSE FACTOR17* is essential for proper development and modulates expression of early auxin response genes', *The Plant cell* 17(5): 1360-75.

McConaha, M. (1941) 'Ventral structures effecting capillarity in the Marchantiales', *American Journal of Botany* 28: 301-306.

Mirbel, M. (1835) 'Researches anatomiques et physiologiques sur le *Marchantia polymorpha*.', *Mém. Acad. Roy. Sc. Inst. France* 13: 337-436.

Mutte, S., Kato, H., Rothfels, C. J., Melkonian, M., Wong, G. K.-S. and Weijers, D. (2017) 'Origin and evolution of the nuclear auxin response system', *bioRxiv*: doi.org/10.1101/220731.

Nizampatnam, N. R., Schreier, S. J., Damodaran, S., Adhikari, S. and Subramanian, S. (2015) *'microRNA160* dictates stage-specific auxin and cytokinin sensitivities and directs soybean nodule development', *Plant Journal* 84(1): 140-153.

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., Vezzi, F., Delhomme, N., Giacomello, S., Alexeyenko, A. *et al*. (2013) 'The Norway spruce genome sequence and conifer genome evolution', *Nature* 497(7451): 579-584.

Prigge, M. J., Lavy, M., Ashton, N. W. and Estelle, M. (2010) *'Physcomitrella patens* auxin-resistant mutants affect conserved elements of an auxin-signaling pathway', *Current Biology* 20(21): 1907-1912.

Proust, H., Honkanen, S., Jones, V. A. S., Morieri, G., Prescott, H., Kelly, S., Ishizaki, K., Kohchi, T. and Dolan, L. (2016) *'RSL* class I genes controlled the development of epidermal structures in the common ancestor of land plants', *Current Biology* 26(1): 93-99.

Qiao, M., Zhao, Z. J., Song, Y. G., Liu, Z. H., Cao, L. X., Yu, Y. C., Li, S. and Xiang, F. N. (2012) 'Proper regeneration from *in vitro* cultured *Arabidopsis thaliana* requires the microRNA-directed action of an auxin response factor', *Plant Journal* 71(1): 14-22.

Rambaut, A. (1996) Se-Al: Sequence alignment editor, http://tree.bio.ed.ac.uk/software/seal/.

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P. F., Lindquist, E. A., Kamisugi, Y. *et al*. (2008) 'The *Physcomitrella genome* reveals evolutionary insights into the conquest of land by plants', *Science* 319(5859): 64-69.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., ., Lander, E. S., Getz, G. and Mesirov, J. P. (2011) 'Integrative genomics viewer', *Nature Biotechnology* 29: 24–26.

Shimamura, M. (2016) *'Marchantia polymorpha*: taxonomy, phylogeny and morphology of a model system', *Plant and Cell Physiology* 57(2): 230-256.

Solly, J. E., Cunniffe, N. J. and Harrison, C. J. (2017) 'Regional growth rate differences specified by apical notch activities regulate liverwort thallus shape', *Current Biology* 27: 1-11.

Stewart, J. L. and Nemhauser, J. L. (2010) 'Do trees grow on money? Auxin as the currency of the cellular economy ', *Cold Spring Harb Perspect Biol.* 2: a001420.

Sugano, S. S., Shirakawa, M., Takagi, J., Matsuda, Y., Shimada, T., Hara-Nishimura, I. and Kohchi, T. (2014) 'CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L.', *Plant and Cell Physiology* 55(3): 475-481.

Szemenyei, H., Hannon, M. and Long, J. A. (2008) 'TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis', *Science* 319(5868): 1384-1386.

Tiwari, S. B., Hagen, G. and Guilfoyle, T. (2003) 'The roles of auxin response factor domains in auxin-responsive transcription', *The Plant cell* 15(2): 533-543.

Tiwari, S. B., Wang, X. J., Hagen, G. and Guilfoyle, T. J. (2001) 'AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin', *The Plant cell* 13(12): 2809-2822.

Tsuzuki, M., Nishihama, R., Ishizaki, K., Kurihara, Y., Matsui, M., Bowman, J. L., Kohchi, T., Hamada, T. and Watanabe, Y. (2016) 'Profiling and characterization of small RNAs in the liverwort, *Marchantia polymorpha*, belonging to the first diverged land plants', *Plant and Cell Physiology* 57(2): 359-372.

Turner, M., Nizampatnam, N. R., Baron, M., Coppin, S., Damodaran, S., Adhikari, S., Arunachalam, S. P., Yu, O. and Subramanian, S. (2013) 'Ectopic expression of *miR160* results in auxin hypersensitivity, cytokinin hyposensitivity, and inhibition of symbiotic nodule development in soybean', *Plant physiology* 162(4): 2042-2055.

Ulmasov, T., Hagen, G. and Guilfoyle, T. J. (1997a) 'ARF1, a transcription factor that binds to auxin response elements', *Science* 276(5320): 1865-1868.

Ulmasov, T., Hagen, G. and Guilfoyle, T. J. (1999a) 'Activation and repression of transcription by auxin-response factors', *Proceedings of the National Academy of Sciences, USA* 96(10): 5844-9.

Ulmasov, T., Hagen, G. and Guilfoyle, T. J. (1999b) 'Dimerization and DNA binding of auxin response factors', *The Plant Journal : for cell and molecular biology* 19(3): 309-19.

Ulmasov, T., Murfett, J., Hagen, G. and Guilfoyle, T. J. (1997b) 'Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements', *The Plant Vell* 9(11): 1963-1971.

Vanneste, K., Sterck, L., Myburg, A. A., Van de Peer, Y. and Mizrachi, E. (2015) 'Horsetails are ancient polyploids: evidence from *Equisetum giganteum'*, *The Plant cell* 27(6): 1567-1578.

Vernoux, T., Brunoud, G., Farcot, E., Morin, V., Van den Daele, H., Legrand, J., Oliva, M., Das, P., Larrieu, A., Wells, D. *et al*. (2011) 'The auxin signalling network translates dynamic input into robust patterning at the shoot apex', *Molecular Systems Biology* 7: 508

Vöchting, H. (1885) 'Über die regeneration der Marchantieen.', *Jahrbücher für wissenschaftliche Botanik* 16: 367-414; Taf. XII-XV.

Wang, J. W., Wang, L. J., Mao, Y. B., Cai, W. J., Xue, H. W. and Chen, X. Y. (2005) 'Control of root cap formation by microRNA-targeted auxin response factors in *Arabidopsis'*, *The Plant Cell* 17(8): 2204-2216.

Wickett, N. J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M. S., Burleigh, J. G., Gitzendanner, M. A. *et al*. (2014) 'Phylotranscriptomic analysis of the origin and early diversification of land plants', *Proceedings of the National Academy of Sciences, USA* 111(45): E4859-E4868.

Zuker, M. (2003) 'Mfold web server for nucleic acid folding and hybridization prediction', *Nucleic Acids Research* 31(13): 3406-3415.



**Supporting Information**

Additional Supporting Information may be found online in the Supporting Information tab for this article:

**Fig. S1** Bayesian phylograms of Streptophyte ARF genes based on nucleotide alignments.

**Fig. S2** LFG motifs.

**Fig. S3** Distinct PB1 domain sequences are found in different clades.

**Fig. S4** The CLP motif in C class ARFs contains the miR160 binding site.

**Fig. S5** Motifs of IAA proteins.

**Fig. S6** Wild-type development.

**Fig. S7** Mp*ARF3:GUS*, Mp*MIR160:GUS* expression lines and *in-situ* hybridization.

**Fig. S8** Mp*MIR160* precursor sequence.

**Fig. S9** Mp*ARF3* loss-of-function mutants have growth deficiencies.

**Fig. S10** Sequences of Mp*ARF3* CRISPR-alleles used in the study.

**Fig. S11** Mp*MIR160* loss-of-function mutants phenotypes.

**Fig. S12** Additional aspects of the Mp*ARF3* loss-of-function phenotype.

**Fig. S13** Far red light induction of Mp*arf3*- and Mp*mir160*- mutants.

**Fig. S14** Promoter sequences of Mp*ARF3* and Mp*MIR160* used in this study.

**Fig. S15** Pharmacological assays in gain-of-function Mp*ARF3* alleles.

**Fig. S16** Pharmacological assays in Mp*miR160-3*.

**Fig. S17** Pharmacological assays in loss-of-function Mp*arf3* alleles.

**Fig. S18** Auxin-responsive genes in Mp*arf3* and Mp*mir160* mutants.

**Fig. S19** MpNCARF and MpNCIAA PB1 domains can interact with auxin signaling components.

**Table S1** Primers used in this study

**Table S2** RNA seq data

**Methods S1** Supplementary materials and methods.



**Fig. 1** Phylogenetic context of AUXIN RESPONSE FACTOR (ARF), IAA and related gene families in Streptophytes. Left panel, a Bayesian phylogram of Streptophyte genes. Tree was constructed using amino acid alignment in supplementary data (Supporting Information Fig. S1f). Numbers at branches indicate posterior probability values; the most relevant are enlarged. Taxa are color coded according to major taxonomic classifications. The evolutionary history of protein domain gain or loss is represented by green or red notation, respectively, on the phylogram. Right panel, protein domain structures encoded by members of the different clades predicted in the ancestral land plant. The B3 and AP2 DNA binding domains and the PB1 protein-protein interaction domain define the ARF and RAV gene families, while domains I and II define the AUX/IAA gene family (Guilfoyle, 2015). The DD domain is involved in dimerization (Boer *et al*., 2014), while the functions of the previously described FD (Guilfoyle, 2015) and LFG (Finet *et al*., 2013) domains are unknown. The CLP, KR, and EHDY domains, the latter two with unknown function, are defined in Figs S4 and S5.

**Fig. 2** Mp*ARF3* spatial expression associates with loss-of-function phenotypes in *Marchantia*. Expression patterns of (a–d) MpARF3 and (e–h) MpMIR160 as assessed by the promoter:GUS reporter gene fusions, of $_{pro}$Mp*ARF3(3.6kb):GUS* and $_{pro}$Mp*MIR160(4kb):GUS* respectively in (a, e) 3-d-old, (b, f) 4-d-old (c, g) 5-d-old and (d, h) 6-d-old gemmalings. Expression is detectable in the apical notches (asterisks) and central zone (cz) for both MpARF3 and MpMIR160 across the developmental series. $_{pro}$Mp*ARF3:GUS* expression is also detected in the differentiating air pores (ap) of 6-d-old gemmalings, a pattern not observed for Mp*MIR160*. (i–l) Gemmalings of $_{pro}$Mp*ARF3(3.6kb):GUS* lines grown in mock (i, j) or 1 μM 2,4-D (k, l) for 7 consecutive days. (m) semiqRT-PCR showing decrease of Mp*ARF3* expression in the wild type upon growth in 20 μM 2,4-D for 48 h. Bars: (a–h), 0.1 mm; (i–l) 0.5 mm.

**Fig. 3** The single class C ARF (Mp*ARF3*) in *Marchantia* promotes growth in the gametophyte. (a) RLM-RACE adapter is bound to Mp*ARF3* transcript between bases 10 and 11 of the miR160-binding site (red arrows), a putative cleavage site conserved (b) in other embryophyte orthologues. (c) Sequencing of Mp*arf3*- alleles show deletions spanning the gRNA binding sites (d) Sequencing of Mp*mir160*- alleles show deletions affecting stem loop, miR and miR* sequences (e) Mp*miR160* expression relative to Mp*miR166* in wild-type and Mp*mir160*- alleles as measured by qRT-PCR. (f) Surface area comparisons in 14-d-old wild-type and four $_{pro}$*EF1:*Mp*MIR160* lines. (g) Apical notch counts of 14-d-old gemmaling populations comparing wild-type with four independent $_{pro}$*EF1:*Mp*MIR160* lines (h–n) Representative images of 30-d-old regenerated *Marchantia polymorpha* plants from a single branch from (h) wild-type, (i) *arf3-51*$^{ge}$ CRISPR knock-out, (j) $_{pro}$*EF1:*Mp*MIR160* knock down, (k) Mp*mir160-3*$^{ge}$ CRISPR knock-out, (l) miR160-resistant Mp*ARF3*

($_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$) overexpressors, (m, n) plants carrying a 229-bp lesion in Mp*ARF3* and co-expressing (m) $_{pro}$Mp*ARF3:*Mp*ARF3*$^{Cas9Res}$-1 or (n) $_{pro}$Mp*ARF3:*At*ARF10*-1. Dashed lines in all plants indicate cut site of subcultured apical notch. 'R' denotes tissue regenerated from the excision as compared to growth derived from the apical notch. White asterisks denote a branch with two or more apical notches each. All plants were grown in ½ B5 media. (f, g) Red asterisks denote *P* values < 0.01 obtained from two-tailed student's *t*-test. (Mp, *Marchantia polymorpha*; Pp, *Physcomitrella patens*; Sm, *Selaginella moellendorffii*; At, *Arabidopsis thaliana*). Error bars: (f and g) ± SD; (e) ± SE. Bars: (i–o) 1 cm.

**Fig. 4** Mp*ARF3* is required for multiple developmental processes in *Marchantia*. Scanning electron micrographs (SEM) depicting the aerial view above gemma cups from (a) wild-type, (b) $_{pro}$*EF1:*Mp*MIR160*, (c) Mp*arf3ko-22*$^{ge}$ and (d) Mp*arf3ko-12*$^{ge}$ lines. Dormant gemma are visible in wild-type (ge), but are either stunted (arrow) or absent (arrowheads) in other genotypes. Quantification of (e) dormant gemmae counts from three independent $_{pro}$*EF1*:*MpMIR160* lines, compared to wild type and (f) the number of air pores per mm$^2$ of 12 gemmalings from wild type, Mp*mir160-2*$^{ge}$ and $_{pro}$*EF1:*Mp*MIR160* after 14 d of growth. SEMs of epidermal air pore distribution in (g) wild-type, (h) Mp*arf3-22* and (i) Mp*mir160-2 c*. 1-month-old plants. Dotted line indicates sections of thallus without air pores. Ventral view of mature wild-type (j), Mp*arf3ko-12* (k) and Mp*arf3ko-2* (l) plants showing smooth rhizoids (SR) and pegged rhizoids (PR) associated with ventral scales (vsc). Mp*arf3* alleles show an increase in formation of pegged rhizoids along the ventral midrib. (m) Dorsal view of Mp*mir160-4* line shows accumulation of smooth rhizoids. (n–p) View of meristematic scales (msc) protecting the (n) wild-type, (o) Mp*arf3-22* and (p) Mp*mir160-2* apical notches. Mp*arf3-22* has more scales with increased complexity, while Mp*mir160-2* does not have meristematic scales (asterisk). (q) Dorsal view of wild-type (left) and *Mparf3ko-22* (right)

thalli show ventral scales (vsc) protrusion in *Mparf3ko-22*. Dorsal view of (r) wild-type and (s) $_{pro}$*EF1:*Mp*MIR160* antheridiophores. Ectopic antheridia are shown (asterisks) and inset of (s). Dorsal view of (t) wild-type and (u) $_{pro}$*EF1:*Mp*MIR160* archegoniophores emerging from an apical notch. Numbers indicate stalks present per apical notch with $_{pro}$*EF1:*Mp*MIR160* lines showing 2 stalks per notch. Inset in (u) shows different degrees of fusion. *P* values are shown using a two-tailed Student's *t*-test comparing wild-type vs mutant populations: **, *P* values <0.01; *, $P < 0.05$. Error bars in (e, f) indicate ± SD. Bars: (a, c, d, g, h, i, q) 0.5 mm; (b) 0.3 mm; (j, k, l, m, s) 1 mm; (n, o) 0.2 mm; (p) 0.25 mm; (r, t, u) 1.2 mm.

**Fig. 5** Enhanced Mp*ARF3* activity results in drastic reversions to undifferentiated cell states in *Marchantia*. (a) $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$ lines have stronger phenotypes than $_{pro}$Mp*ARF3:*Mp*ARF3* after 2 months of growth. (b) Scanning electron micrograph (SEM) of representative $_{pro}$Mp*ARF3:*Mp*ARF3* line showing clear apical notches and branches with air pore differentiation. Inset shows that growth of $_{pro}$Mp*ARF3:*Mp*ARF3* lines is convoluted and reminiscent of other gain-of-function phenotypes. Inset scale, 1 mm. (c) Representative SEM of $_{pro}$Mp*ARF:*Mp*ARF3*$^{m}$ plant showing a collection of prothallus-like outgrowths. Lines have putative notches (asterisks) but no air chambers. (d) Detail of $_{pro}$Mp*ARF:*Mp*ARF3*$^{m}$ lines show single cell layer thick prothalli-like tissue (arrowhead). Mature wild-type (e), $_{pro}$*SHI:*Mp*ARF3* (f), and $_{pro}$*SHI:*Mp*ARF3*$^{m}$ (g) plants grown for 2 months. Experimental lines show ectopic formation of branches. (h) Dorsal SEM of $_{pro}$*SHI:*Mp*ARF3*$^{m}$ shows differentiation of air pores. (i) Left, $_{pro}$*EF1:*Mp*ARF3* line shows a convoluted thallus with poor differentiation; right, $_{pro}$*EF1:*Mp*ARF3*$^{m}$ lines have an undifferentiated callus-like appearance. (j) SEM of $_{pro}$*EF1:*Mp*ARF3* shows that the epidermis lacks air pores. (k) $_{pro}$*EF1:*Mp*ARF3*$^{m}$ lines show no clear establishment of branches. (l) $_{pro}$*EF1:*Mp*ARF3*$^{m}$ lines show a loss of differentiation. (m) Mp*ARF3* expression relative to Mp*ACT1* as measured by

qRT-PCR in *WT*, *Hyg*$^{Res}$, $_{pro}$Mp*ARF3:*Mp*ARF3* and $_{pro}$Mp*ARF:*Mp*ARF3*$^{m}$ lines (n) Auxin sensitivity assay of Mp*ARF3* alleles. *P* values after a two-way student's *t*-test comparing treated vs un-treated are shown. *P* values in red are <0.01. All genotypes had a significant growth decrease in high doses of auxin except for $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$. Error bars in (m, n) indicate ± SD. Bars: (a) 0.075 mm; (b) 0.17 mm; (c, g, i, j, k, m) 1 mm; (d) 0.5 mm; (e) 1 cm; (f, l) 0.6 mm; (h) = 0.3 mm; (n, o) 0.25 mm; (p) 0.06 mm. **[Author, please check text '(n, o) 0.25 mm; (p) 0.06 mm.' as there are no scale bars in (n), and no panels (p) and (o).] [Author, please insert text to explain * and ** in panels (m) and (n).]**

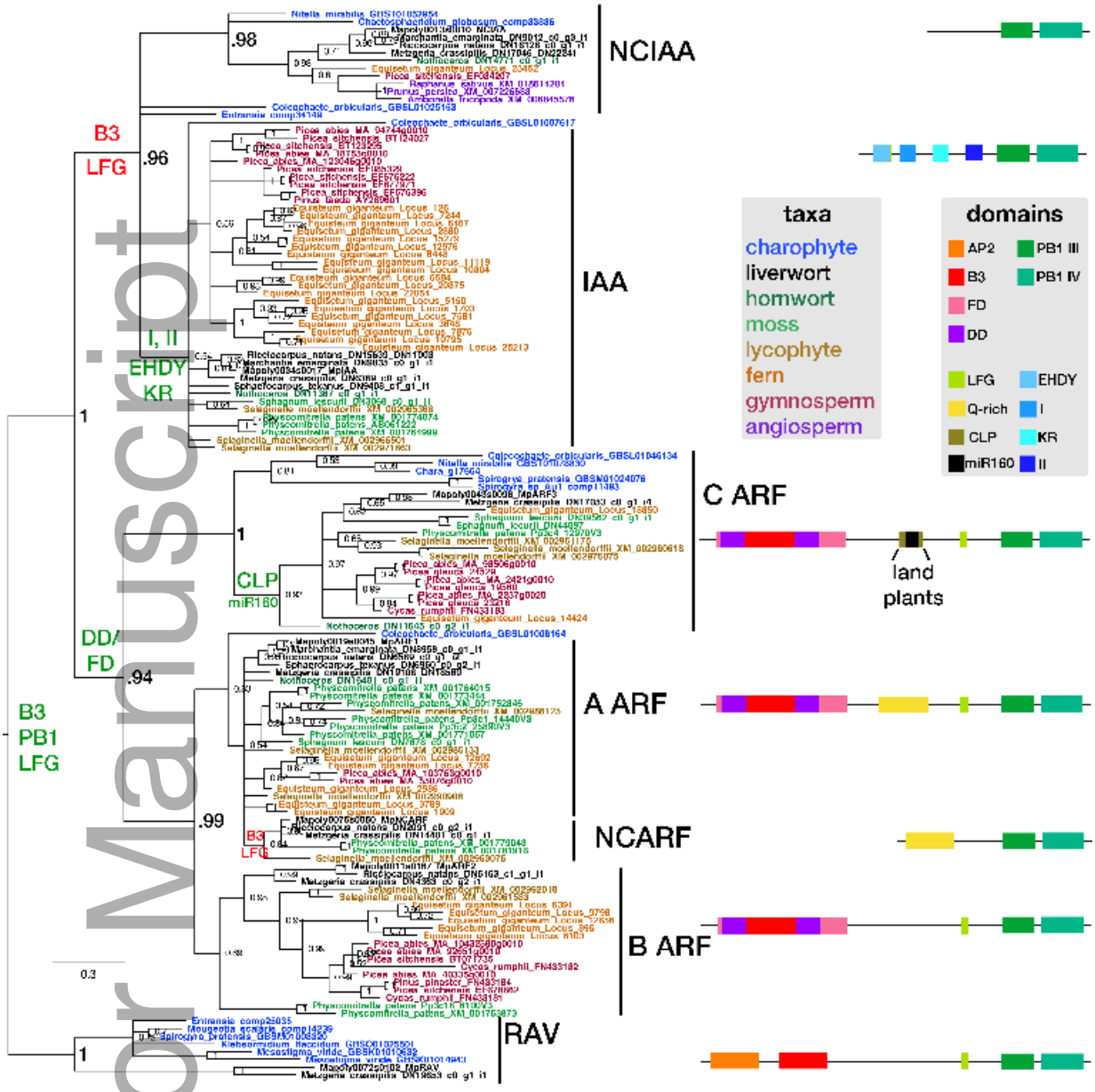
**Fig. 6** Transcriptional regulatory feedbacks associated with loss of Mp*ARF3* and Mp*MIR160* in *Marchantia.* (a) RPKM values of Mp*ARF3* loci in three independent *WT*, Mp*arf3*- and Mp*mir160* alleles. (b) RPKM values of Mp*MIR160* precursor in three independent *WT*, Mp*arf3*- and Mp*mir160* alleles. *P* values for (a) and (b) come from the voom/limma pipeline. (c) Raw RNA-seq density reads (Sashimi plots) of Mp*MIR160* in all replicates. miR/miR* sites are indicated with a red arrowhead. Reads are undetectable at the primary stem loop site (dotted rectangle) in the *WT*, suggesting processing by DICER. Mp*arf3* alleles (-12, -22 & -2) show loss of Mp*MIR160* transcription, while Mp*mir160*- alleles (-2, -3 & -4) show increases in raw reads throughout the locus and particularly at putative processing site (red asterisks). (d) Relative expression levels of mature *Mpmir160* transcripts to *MpmiR166 using* qRT-PCR, in *Hyg*$^{Res}$ controls, and multiple $_{pro}$Mp*ARF3:*Mp*ARF3* and $_{pro}$Mp*ARF3:*Mp*ARF3*$^{m}$ lines. Levels of significance indicated above columns using a Student *t*-test. (e) Raw RNA-seq density reads (Sashimi plots) of Mp*ARF3* loci in all replicates. Mp*arf3-12* and Mp*arf3-2* show transcriptional gaps at deletion sites (red asterisks), producing aberrant transcripts. Mp*arf3-22* has a 100-bp insertion and did not show production of aberrant transcripts. All alleles have increased transcription of Mp*ARF3* as a putative feedback mechanism. Blue

arrows indicate direction of transcription. Junction reads are displayed as arks connecting exons. Error bars indicate: (a, b) ± SD; (d) ± SE. Scale = 0-500 raw reads. **[Author, please insert text to explain * and ** in panels (a), (b) and (d).]**

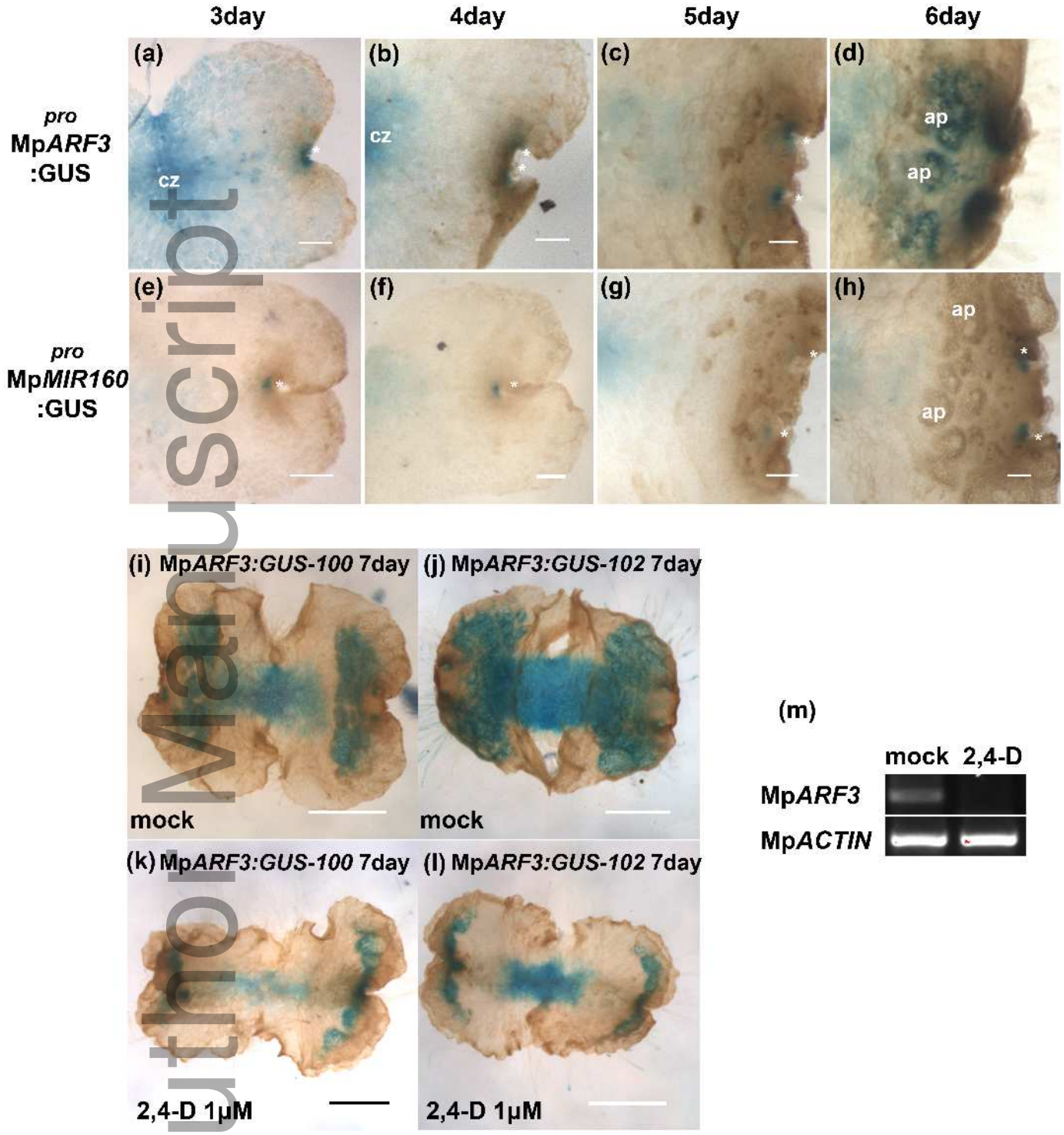
**Fig. 7** Mp*ARF3*$^{Cas9Res}$ and At*ARF10* complement Mp*arf3* mutations in *Marchantia*. (a–d) Aerial view of three representative *c*. 2-month-old mature plants from (a) wild-type, (b) Mp*arf3* lines, co-expressing gRNA4/5 targeting Mp*ARF3*, (c) plants co-expressing $_{pro}$Mp*ARF3:*Mp*ARF3*$^{Cas9Res}$ and dual gRNAs (4/5) targeting Mp*ARF3* show a wild-type growth rescue, (d) plants co-expressing $_{pro}$Mp*ARF3:*At*ARF10* as well as dual gRNAs 4/5 targeting Mp*ARF3* with confirmed mutations in the endogenous gene. Quantification of various morphological features of plant growth of the plants shown in (a–d) indicating (e) thallus area ($cm^2$) measurements, (f) apical notch counts, (g) the number of gemmae cups and (h) the proportion of cups with dormant gemmae. *P* values shown are comparisons of means of all genotypes vs Mp*arf3* mutants (*Mparf3-42, -43, -44*) using a two-tailed Student's *t*-test. *P* values shown in red are lower than 0.01). Error bars indicate ± SD. Bars: (a–d) 1 cm

**Table 1** Mutant alleles used in this study

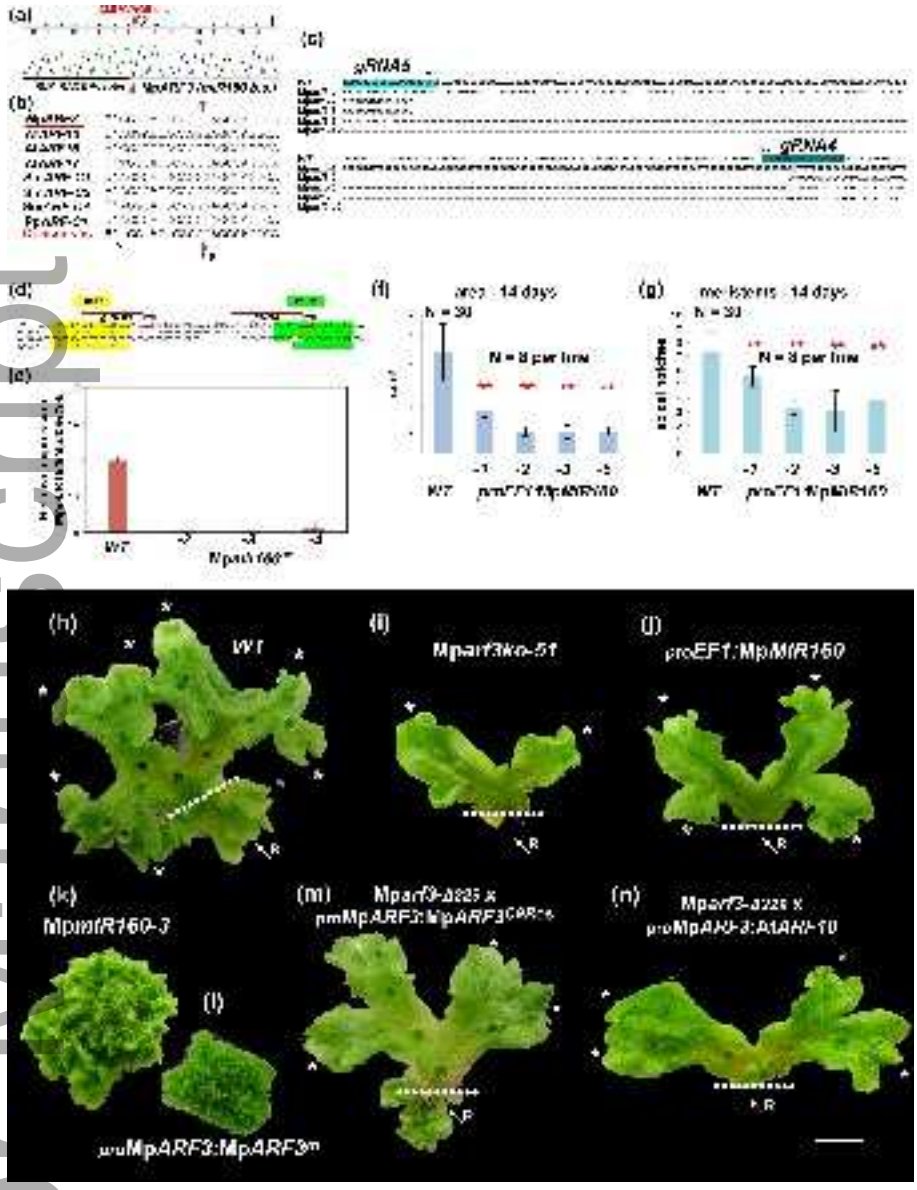
***[Typesetters, please see separate Excel file for Table 1]***



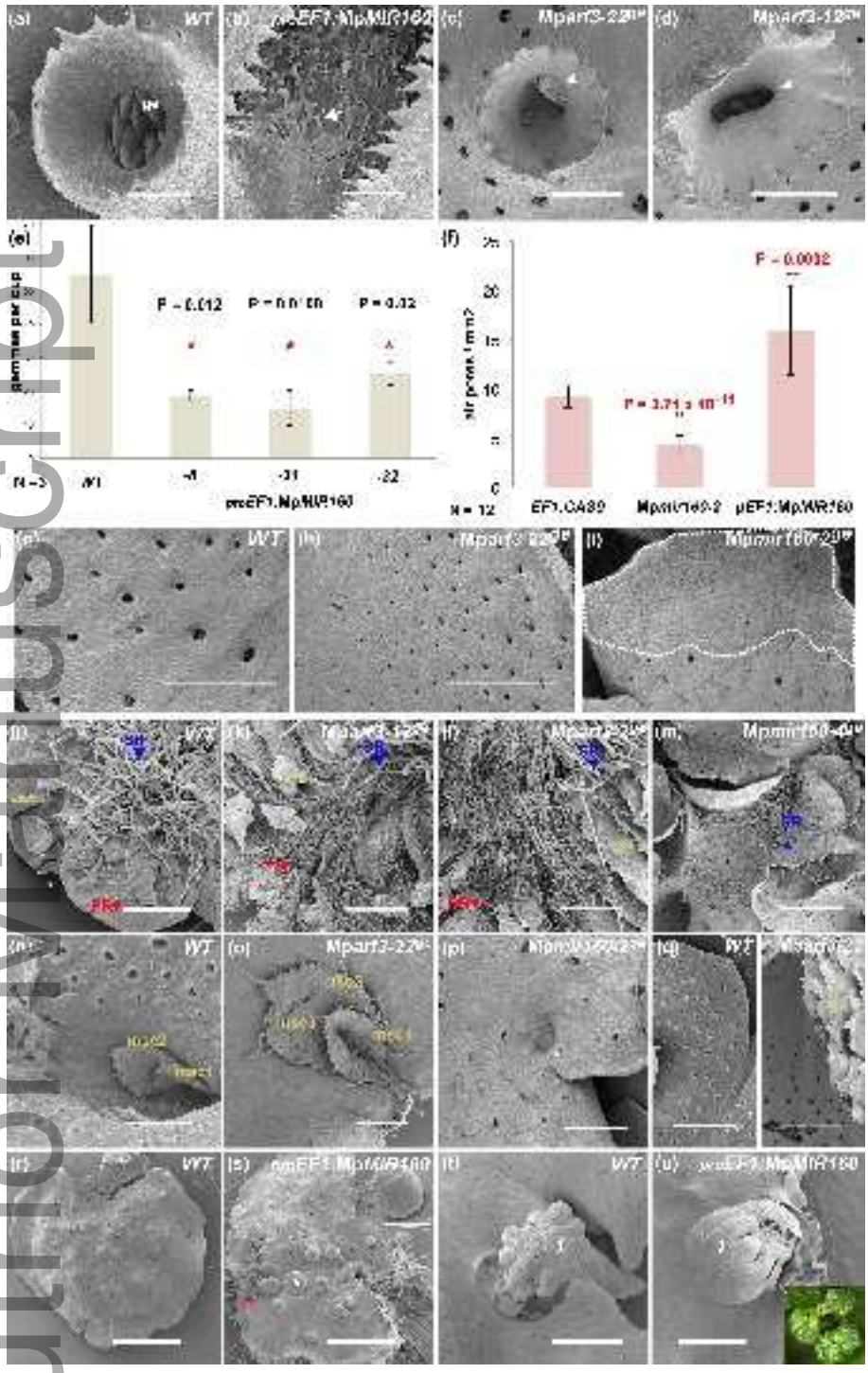

nph_15090_f1.tif



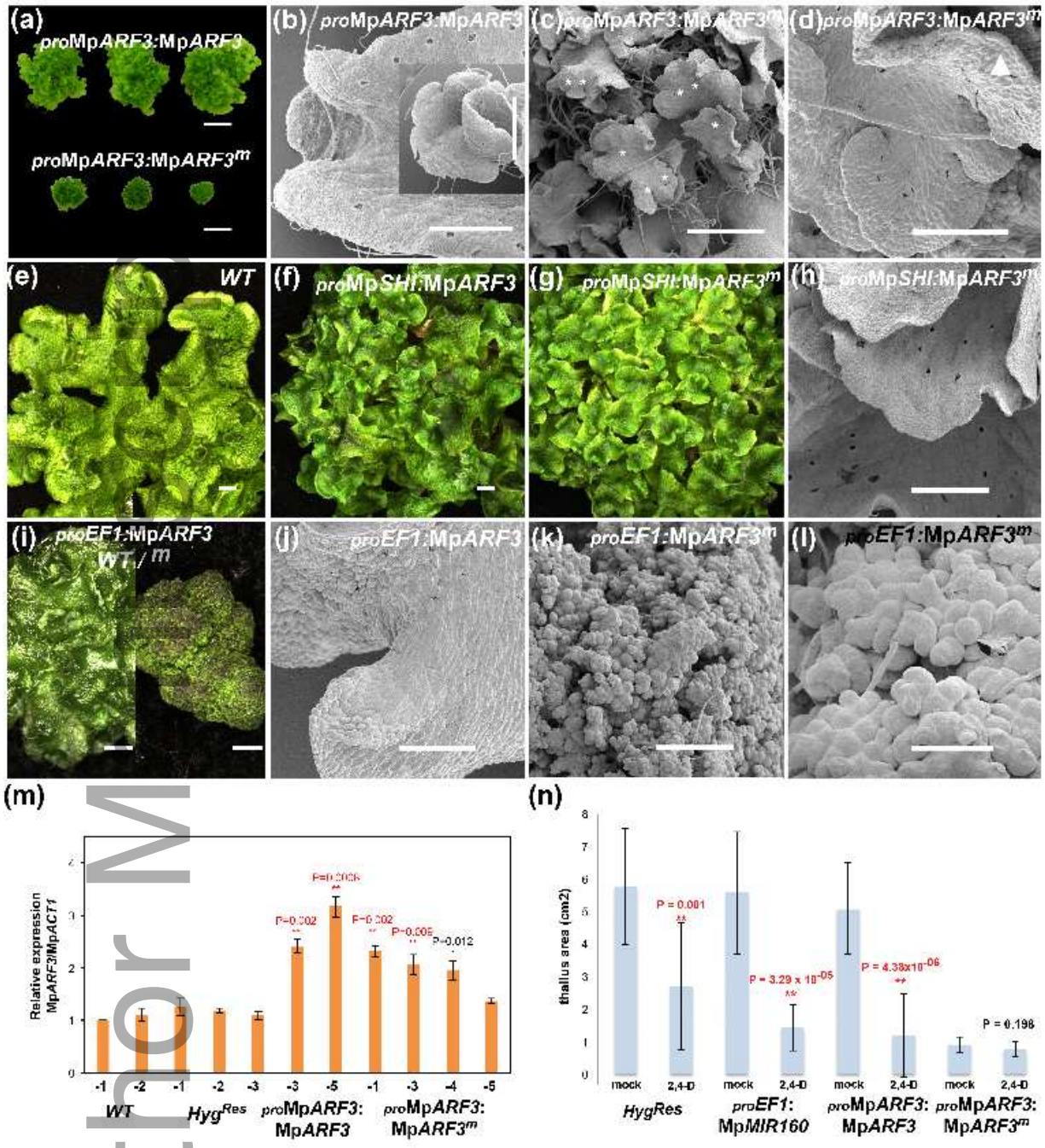

nph_15090_f2.tif



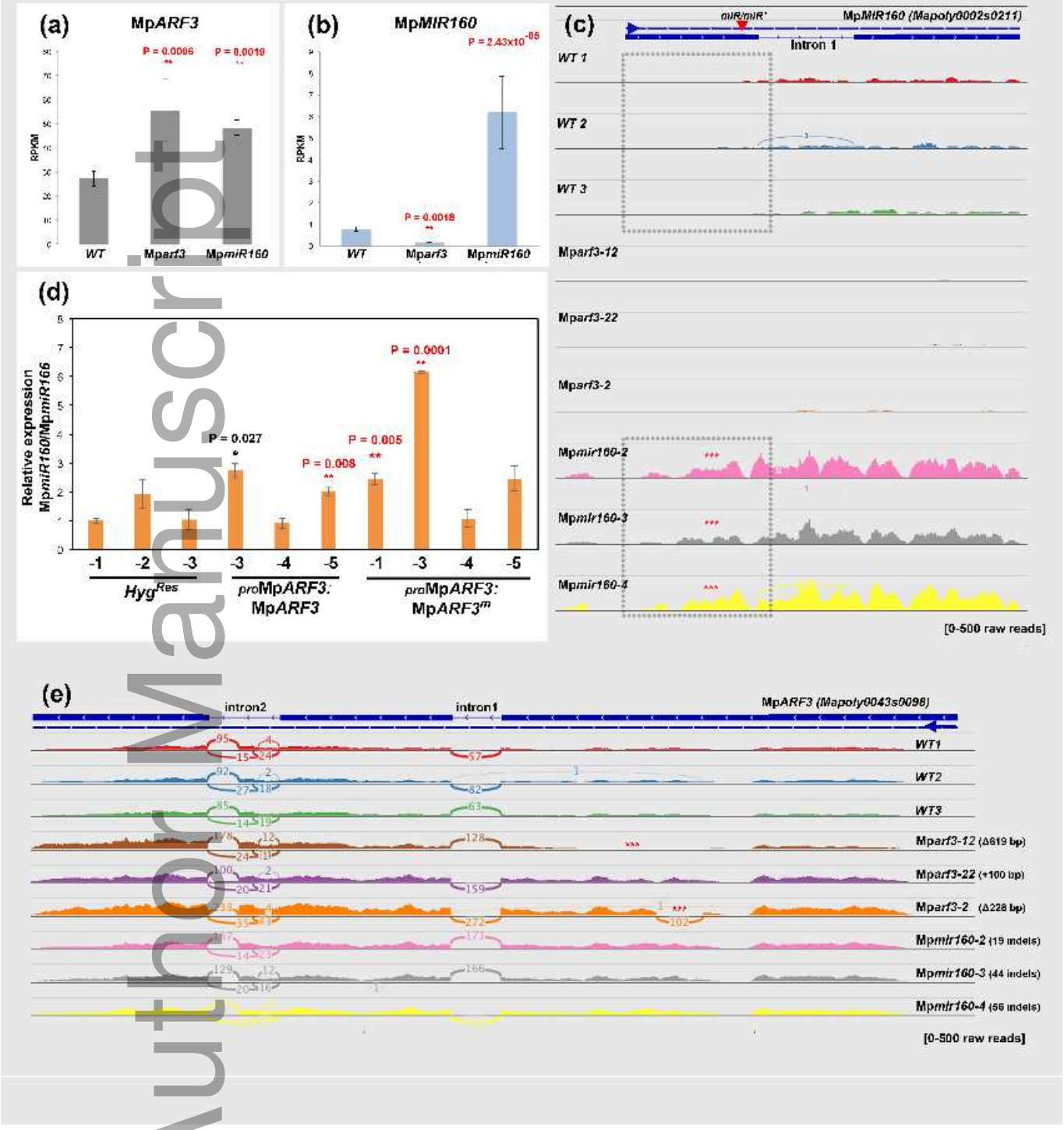

nph_15090_f3.tif



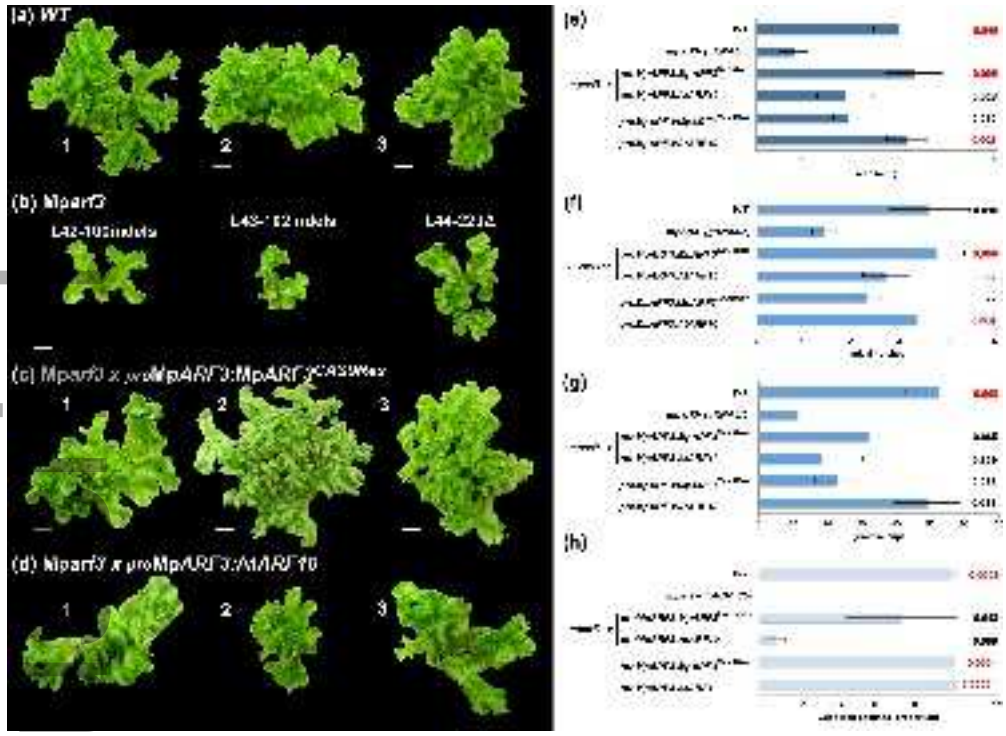
nph_15090_f4.tif




nph_15090_f5.tif




nph_15090_f6.tif




nph_15090_f7.tif