R2R3-MYB GENE EVOLUTION IN PLANTS, INCORPORATING FERNS INTO THE STORY

Brenda Hernández-Hernández,* Rosalinda Tapia-López,† Barbara A. Ambrose,‡ and Alejandra Vasco2,§

*Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad de México DF 04510, México; †Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad de México, México; ‡New York Botanical Garden, 2900 Southern Boulevard, Bronx, New York 10458-5126, USA; and §Botanical Research Institute of Texas, 1700 University Drive, Fort Worth, Texas 76107-3400, USA

Editor: Patrick S. Herendeen

Premise of research. MYB transcription factors are one of the largest families of genes in land plants. They are involved in many regulatory processes, such as primary and secondary metabolism, cell fate and identity, developmental processes, and responses to abiotic and biotic stresses. Phylogenetic studies have mainly focused on R2R3-MYB evolution in seed plants; however, a comprehensive sampling across land plants is lacking, as none have included ferns.

Methodology. To better understand the evolution of R2R3-MYB genes in land plants, we surveyed the R2R3-MYB gene sequences from six genomes belonging to the main lineages of land plants (bryophytes, lycophytes, ferns, angiosperms, and gymnosperms). In addition, we searched transcriptome sequences of selected fern and lycophyte transcriptomes. We assembled a nucleotide matrix of the MYB domain and conducted maximum likelihood analyses to infer phylogenetic relationships and gene-tree reconciliation analyses to infer gene duplications. We labeled the main clades with the known MYB functions onto the resulting tree. To detect reported as well as new conserved motifs, we ran protein motif analyses through MEME.

Pivotal results. Our results support previous studies indicating that R2R3-MYBs in seed plants are more diverse than in any other lineage of vascular plants. Most of the well-supported clades recovered are inferred to be already present in the ancestor of land plants. We found that ferns have numerous copies of R2R3-MYBs, although not as many as seed plants. Protein motif analyses revealed that R2R3-MYB motifs are highly conserved across land plants, suggesting that most R2R3-MYB orthologues, including those of ferns, might have DNA-binding capabilities and could be involved in regulatory processes that are similar to those of their angiosperm homologues. Our analyses showed that there are no fern orthologues of the angiosperm MYB leaf developmental genes ARP, raising interesting questions about the evolution and origin of leaves in ferns and seed plants.

Conclusions. Our results provide the phylogenetic context for research on the genetic and functional evolution of an important gene family of developmental and metabolic regulators across the plant tree of life.

Keywords: ASYMMETRIC LEAVES1, gene expression, leaf development, MYB genes, transcription factors.

Online enhancements: appendixes.
et al. 2013). MYB proteins are classified into four major groups based on the number of adjacent MYB repeats (1R, 2R, 3R, and 4R). The three repeats of the prototypic animal MYB protein C-MYB are referred to as R1, R2 and R3; repeats from other MYB proteins are named according to their similarity to these (Dubos et al. 2010).

Surprisingly, land plants (embryophytes) have the highest diversity of MYB proteins and contain all four classes of MYBs (Stracke et al. 2001). Among these, the R2R3-MYB proteins are the most common in vascular plants (Kelemen et al. 2015). R2R3-MYB proteins make up one of the largest TF groups in seed plants, with more than 170 members reported in Arabidopsis thaliana (Riechmann et al. 2000; Stracke et al. 2001; Du et al. 2015), 157 in Zea mays (Du et al. 2012), 192 in Populus trichocarpa (Wilkins et al. 2009), and 159 in the gymnosperm Picea abies (Du et al. 2015), among others.

In seed plants, functional and expression studies have shown that R2R3-MYBs control many important processes (Jin and Martin 1999; Bomal et al. 2008; Ambawat et al. 2013; Palmer et al. 2013; Pandey et al. 2015; He et al. 2016). These have been classified into four categories: primary and secondary metabolism, cell fate and identity, developmental processes, and responses to biotic and abiotic stresses (Dubos et al. 2010). Comparative and evolutionary studies have shown that R2R3-MYBs are also diverse in other lineages of embryophytes; for example, 50 R2R3-MYBs have been reported in the bryophyte Physcomitrium patens (Du et al. 2015), 21 in the liverwort Marchantia polymorpha (Bowman et al. 2017), and 19 in the lycophyte Selaginella moellendorfii (Du et al. 2015). To date, little is known about R2R3-MYB functions beyond seed plants.

Although previous evolutionary studies focusing on R2R3-MYB genes have included several lineages of vascular plants, none have included a comprehensive sampling of ferns, a key lineage of about 10,500 species sister to the seed plants (PPG I 2016). To better understand the evolution of R2R3-MYB genes in land plants, we comprehensively surveyed all available genomic data for ferns. We obtained R2R3-MYB sequences, including the first two genomes of ferns published (Li et al. 2018) and one genome of each major land plant lineage (bryophytes, lycophytes, gymnosperms, and angiosperms), for phylogenetic reconstruction. We also include transcriptomes from selected ferns and lycophytes.

Our results support previous studies indicating that R2R3-MYBs in seed plants are more diverse than in any other lineage of vascular plants. Most of the well-supported clades we recovered include at least one representative of each of the land plant lineages we investigated, indicating that those were already present in the ancestor of embryophytes and that the high number of seed plant R2R3-MYB genes is in part due to several different expansions in seed plant subclades. Although they are not as numerous as in seed plants, we found that ferns also have numerous copies of R2R3-MYBs. Our results provide a phylogenetic context for studying the genetic and functional evolution of this important protein family of developmental and metabolic regulators across the plant tree of life.

**Material and Methods**

We conducted similarity searching with BLAST (Altschul et al. 1990), first using all R2R3-MYB genes reported from Arabidopsis thaliana (Stracke et al. 2001; Du et al. 2015; app. A; apps. A–C are available online) and then all the fern and lycophyte sequences obtained from the first round of searches. Our sampling included one species per lineage of land plants, except for in ferns, where two genomes were mined. The databases and species queried are listed in table 1.

Our preliminary phylogenetic analyses recovered a clade composed only of sequences from seed plants and lycophytes (clade 6). These sequences belong to a well-studied developmental gene family: ASYMMETRIC LEAVES1/ROUGH SHEATH2/PHANTASTICA (ARP). To better understand the evolution of this clade and make sure that there were no additional lycophyte or fern sequences recovered as part of it, we conducted additional BLAST searches in all the lycophyte and fern transcriptomes available in the oneKP database (http://sites.google.com/a/ualberta.ca/onekp; Matasci et al. 2014; Wickett et al. 2014; Leebens-Mack et al. 2019), using as queries the genes ASYMMETRIC LEAVES1 (AS1) from Arabidopsis, ROUGH SHEATH2 from Zea mays, and PHANTASTICA from Antheridium majus and the SkARP1 homologue from Selaginella kraussiana (app. A).

**Table 1**

| Lineage, species | Name | Database | URL |
|------------------|------|----------|-----|
| Bryophytes: | Phytozone | http://phytozone.jgi.doe.gov/pz/portal.html |
| *Physcomitrium patens* (Hedw.) Mitt. | | |
| Lycophytes: | Phytozone | http://phytozone.jgi.doe.gov/pz/portal.html |
| *Selaginella moellendorfii* Hieron. | NCBI/GenBank; TAIR | http://arabidopsis.org |
| Angiosperms: | | |
| *Arabidopsis thaliana* (L.) Heynh. | | |
| Gymnosperms: | | |
| *Picea abies* (L.) H. Karst. | ConGenIE | http://congenie.org/ |
| Ferns: | | |
| *Azolla filiculoides* Lam. | FernBase | http://fernbase.org/ |
| *Ceratopteris richardii* Brongn. | NCBI | http://ncbi.nlm.nih.gov/genome/?term=txid49495 |
| *Salvinia cucullata* Roxb. | FernBase | http://fernbase.org/ |

Note. NCBI = National Center for Biotechnology Information; TAIR = The Arabidopsis Information Resource; ConGenIE = Conifer Genome Integrative Explorer.
Nucleotide sequences were aligned using MAFFT version 7 (Katoh and Standley 2013). Alignments were refined by hand using Mesquite version 3.61 (Maddison and Maddison 2019). Sequences with an incomplete MYB domain were excluded from the analyses.

Phylogenetic relationships were inferred from the nucleotide data of the MYB domain-aligned matrix using maximum likelihood (ML). Analyses were performed on CIPRES (http://www.phylo.org; Miller et al. 2010). ML best tree and bootstrap searches were performed simultaneously with 1000 replicates using the GTR+CAT model in RAxML version 7.2.8 (Stamatakis 2006; Stamatakis et al. 2008). The most likely tree was depicted and color edited using FigTree 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). The tree was rooted in one MYB Selaginella moellendorfii sequence that lacked the main motifs of the R2 repetition. The history of duplications of the R2R3-MYB genes in land plants was reconstructed by comparing our gene tree with recent molecular phylogenetic hypotheses of the relationships among the lineages included (Leebens-Mack et al. 2019) using gene-tree reconciliation analysis in Notung 2.9 (Vernot et al. 2008). For the gene-tree reconciliation analysis, we considered only four clades in both the gene tree and the species tree: bryophytes, lycophytes, ferns, and seed plants. Duplications were inferred in nodes with bootstrap supports of >50%.

To better understand the evolution of the recovered well-supported clades, we noted the presence or absence of each of the lineages included (bryophytes, lycophytes, ferns, and seed plants) in each clade and inferred its common ancestor. Clades were labeled on the basis of the Arabidopsis R2R3-MYB sequences recovered in each clade and the functions known for them (app. B) and with the MYB subfamilies previously defined by Du et al. (2015). To detect reported as well as new conserved motifs, our matrix was permanently translated and uploaded as aa to the online MEME server (http://meme-suite.org/tools/meme; Bailey and Elkan 1994) and run under these parameters: classical mode of motive discovery, zero or one occurrence per sequence, and a maximum number of 15 motifs between 10 and 15 aa.

Results

Our final aligned matrix included 335 sequences, of which 30 belong to mosses, 50 to lycophytes, 102 to ferns, 66 to gymnosperms, and 107 to angiosperms (app. A). The aligned matrix had 345 nucleotide characters. Phylogenetic relationships found with ML analyses are presented as an ML phylogram (fig. 1B). The final data set is available in figshare (http://figshare.com/s/42d421fb6c4d547bab6e; nucleotides 2589–2934 were the ones used for the phylogenetic analysis). In general, support across the phylogeny backbone is low, likely because of the fact that the domains used to infer the tree are short compared with the number of sequences included (fig. 1B).

Our phylogeny recovered 19 well-supported (bootstrap supports, >70%) major clades of R2R3-MYB orthologues (fig. 1B). On the basis of the presence/absence of the included lineages (bryophytes, lycophytes, ferns, and seed plants) in each clade, we inferred that 10 of these well-supported clades evolved in the ancestor of land plants (clades 1–5, 7, 13, 15–17, 19), one in the ancestor of vascular plants (clade 6), two in the ancestor of euphyllophytes (clades 12, 18), one in the ancestor of seed plants (clade 9), one in the ancestor of angiosperms (clade 14), one in the ancestor of gymnosperms (clade 8), one in the ancestor of ferns (clade 11), and one in the ancestor of lycophytes (clade 10).

We found 43, 23, and 36 copies of R2R3-MYBs in the ferns Azolla filiculoides, Ceratopteris richardii, and Selvinia cucullata, respectively, and 25 copies in the lycophyte Selaginella moellendorfii (fig. 1B; app. A). Fern copies are recovered in 13 of the 19 well-supported major clades, whereas lycophyte copies are recovered in 12 clades. Clade 6 is inferred to have evolved in vascular plants, but it lacks fern sequences. Additional searches in lycophyte and fern transcriptome databases focusing on this clade resulted in only six lycophyte sequences being recovered as part of clade 6, all of them belonging to the genus Selaginella (fig. 1B; app. A). Gene-tree reconciliation analyses recovered four major duplications that likely happened before the evolution of land plants (fig. 1B, yellow stars).

The protein motif analyses included the region of the R2R3 motifs of our previously aligned matrix. We found that the R2 and R3 MYB domains are highly conserved in all embryophyte sequences, including those of ferns (fig. 2; app. C). In our analyses, R2 corresponds to motifs 1, 3, and 6, while R3 corresponds to motifs 2, 4, and 5. Motifs 1 and 2 have the best e-values and form the third helix of each repetition. Of the 355 analyzed sequences, motifs 1 and 2 are present in 354 and 320 sequences, respectively. In general, most R2R3-MYBs have the same motif structure, and there are no exclusive motifs for each clade (fig. 2, app. C), except for clade 6, which has an exclusive motif in the carboxyl terminal region (fig. 2, motif 7).

Discussion

We found that R2R3-MYB genes are very diverse in land plants. After adding fern R2R3-MYB genes to the story, we still find that seed plants are the lineage with more copies of R2R3-MYB genes, supporting previous phylogenetic studies that showed that R2R3-MYBs are very diverse in land plants and underwent a rapid expansion during the evolution of seed plants (Riechmann et al. 2000; Stracke et al. 2001; Dubos et al. 2010; Ambawat et al. 2013; Du et al. 2015; Bowman et al. 2017). The expansion of gene families contributes to the evolution of metabolic, regulatory, and signaling networks and to the diversification of gene function (Force et al. 1999). We found that seed plants have more R2R3-MYB copies than ferns, and ferns have more copies than lycophytes. R2R3-MYB genes are known to control many different plant-specific processes in seed plants (Jin and Martin 1999; Stracke et al. 2001; Bomal et al. 2008; Dubos et al. 2010; Ambawat et al. 2013; Palmer et al. 2013; Du et al. 2015; Pandey et al. 2015; He et al. 2016); the inclusion of fern and lycophyte copies in the R2R3-MYB phylogeny will aid in the further annotation of lycophyte and fern genomes and guide further functional and expression analyses. Interestingly, a previous study has found that many TF families present in land plants are more numerous in seed plants than in bryophytes, lycophytes, or ferns (Leebens-Mack et al. 2019).

From the 19 clades recovered in our phylogeny, 13 were inferred to have evolved in the ancestor of land plants, vascular plants, or euphyllophytes. The origin of these clades suggests that the high number of R2R3-MYB copies in seed plants is due to a combination of lineage-specific expansions that have occurred in land plants, vascular plants, and seed plants, as has been suggested before (Stracke et al. 2001; Dubos et al. 2010;
Fig. 1  Phylogeny of R2R3-MYB genes in land plants. A, Evolutionary relationships of lineages included in the phylogenetic analyses; lineages and arrowhead colors correspond to the taxon and/or group in B. B, Maximum likelihood hypotheses for the evolution of R2R3-MYB genes in land plants. Arrowheads correspond to major well-supported clades; the color of the arrowheads corresponds to the inferred common ancestor of each clade. Stars correspond to major gene duplications. Clades are labeled on the basis of the Arabidopsis R2R3-MYB sequences recovered in each clade and the functions known for them (defense, development [dev.], differentiation [diff.], metabolism [met.]; app. B; apps. A–C are available online). S numbers correspond to the MYB subfamilies defined by Du et al. (2015) and listed in appendix B. Thick lines correspond to bootstrap values higher than 70%; asterisks correspond to bootstrap values of 100%.
Fig. 2  Conserved motifs of R2R3-MYB proteins in clades 5 and 6 identified through a MEME analysis. A, Schematic representation of the conserved motifs of clade 5, which is typical of most R2R3-MYBs in land plants, and clade 6 (for a complete set of sequences, see app. C; apps. A–C are available online). Each motif is represented by a colored box displayed and numbered in B. The repeats R2 (here represented by motifs 1, 3, and 6) and R3 (here represented by motifs 2, 4, and 5) are highly conserved across all land plant R2R3-MYB proteins. B, Sequences of the conserved motifs detected in the R2R3-MYB homologues across land plants. The relative font size for each residue represents sequence conservation, with larger fonts representing higher conservation. Arabic numbers on the X-axis represent the position of every residual within the motif.
Du et al. 2015). The four major duplications that we found gave rise to most of the R2R3-MYB well-supported clades we recovered and suggest that there was a first set of expansions (including at least four duplications) in this gene family before the main lineages of land plants evolved.

Within the 19 clades recovered, sequences of each lineage tend to be retrieved in their own subclades, suggesting that the different number of R2R3-MYB genes in lycophytes, ferns, and seed plants can be explained in part by independent expansions that occurred after the divergence of each of these lineages. Previous phylogenetic studies focusing on seed plants have suggested that R2R3-MYBs underwent a huge expansion in this lineage or in its ancestor (Du et al. 2015). Our results show that the huge expansion of R2R3-MYBs is exclusive to seed plants and likely happened after ferns and seed plants diverged.

Previous studies have suggested that R2R3-MYB phylogenies recovered clades of Arabidopsis sequences that have similar functions (Dubos et al. 2010; Du et al. 2015; Bowman et al. 2017). When contrasting the phylogeny with the protein motif analyses, we did not find characteristic motifs for each clade, with the exception of clade 6 (fig. 2). Motifs 1 and 2, which have been shown to be important highly conserved regions for TF binding (Du et al. 2012), are present in more than 90% of the sequences included here (fig 2; app. C). This suggests that most of the R2R3-MYB orthologues reported here, including those of ferns, might have DNA-binding capabilities and thus could possibly have functions similar to the ones reported for their Arabidopsis homologues (fig. 1B; app. B). Moreover, the high diversity of fern R2R3-MYBs and their presence in 13 different clades suggest that fern R2R3-MYB proteins likely play several different important functions and control many important processes that are similar to those reported for Arabidopsis homologues (app. B). Expression and functional studies will be necessary to test this hypothesis.

Only sequences of the genus Selaginella and sequences of seed plants are recovered in clade 6 (fig. 1B), which corresponds to the well-studied ARP gene lineage of angiosperms (Sun et al. 2002). This result indicates that orthologues of ARP were lost in ferns. ARP proteins in angiosperms have been shown to act as regulators for plant responses: an overview. Physiol Mol Biol Plants 19:307–321. In ferns, ARP expression was studied by immunolocalization of the maize ARP protein in Osmunda regalis and was found in the SAM and leaf primordia (Harrison et al. 2005). Although this comparative study had suggested that the ability to repress CIKNOX expression during leaf initiation might be an ARP function that is ancestral to vascular plants (Harrison et al. 2005), the study did not include native ARP function in ferns.

Our results show that ARP genes are absent from ferns, and thus the CIKNOX/ARP genetic module does not exist. The absence of ARP genes in ferns supports previous hypotheses stating that the mutually exclusive CIKNOX/ARP expression may have occurred only in the ancestor of seed plants (Harrison et al. 2005; Plackett et al. 2018). Furthermore, the absence of ARP genes in ferns and in two lineages of lycophytes suggests that other mechanisms, perhaps asymmetric leaves2 orthologues (Sun et al. 2012) or auxin-related processes (Reinhardt et al. 2003; Zhao et al. 2010), might be responsible for the downregulation of CIKNOX in fern leaf primordia. This further supports paleobotanical data indicating that leaves of ferns and seed plants have different origins (Galtier 2010; Corvez et al. 2012; Vasco et al. 2013). Further phylogenetic analyses of these factors, additional lycophyte and fern genomes, and functional analyses will help clarify the regulation of CIKNOX in the different lineages of vascular plants and will shed light on the evolution and origin of leaves in vascular plants.

Note Added in Proof

A recent publication presented phylogenetic analyses of plants’ R2R3-MYB TFs including a dense sampling of streptophyte algae (Jiang and Rao 2020). They found that many R2R3-MYBs are not exclusive to land plants (our clades 1, 4–19). These data further support our hypothesis that most of the R2R3-MYB clades recovered were already present in the common ancestor of land plants.

Acknowledgments

This work is part of the undergraduate senior thesis of B. Hernández-Hernández at the Universidad Nacional Autónoma de México and was supported in part by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, Universidad Nacional Autónoma de México (IA201416 to A. Vasco). We thank the Ceratopteris Genome Project for the Ceratopteris richardii transcriptome sequences. Thanks to Juan David Pardo and Elena Yepes for help with the digital files for figure 1. A. Vasco, R. Tapia-López, and B. A. Ambrose were responsible for conceptualization of the article; A. Vasco, B. Hernández-Hernández, and R. Tapia-López were responsible for developing the methodology and performing the analyses; B. Hernández-Hernández was responsible for data curation; and A. Vasco wrote the original draft and was responsible for visualization and funding acquisition. All authors reviewed, edited, and approved the final version of the article.

Literature Cited

Altschul SF, W Gish, W Miller, EW Myers, D Lipman 1990 Basic local alignment search tool. J Mol Biol 215:403–410.

Ambawat S, P Sharma, NR Yadav, RC Yadav 2013 MYB transcription factor genes as regulators for plant responses: an overview. Physiol Mol Biol Plants 19:307–321.

Bailey TL, C Elkan 1994 Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2:28–36.

Banks JA, T Nishiyama, M Hasebe, JL Bowman, M Gribskov, CW DePamphilis, VA Albert, et al 2011 The Selaginella genome identifies
genetic changes associated with the evolution of vascular plants. Science 332:960–963.

Bomal C, F Bedon, S Caron, SD Mansfield, C Levasseur, JEK Cooke, S Blais, et al 2008 Involvement of Pinus taeda MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative in planta analysis. J Exp Bot 59:3925–3939.

Bowman JL, T Kohchi, KT Yamato, J Jokims, S Shu, K Ishizaki, S Yamaoka, et al 2017 Insights into land plant evolution garnered from the Marchantia polymorpha genome. Cell 171:287–304.

Byrne ME, R barley, M Curtis, JM Arroyo, M Dunham, A Hudson, RA Martienssen 2000 Asymmetric leaves1 mediates leaf pattern-}

and stem cell function in Arabidopsis. Nature 408:967–971.

Corvez A, V Barriel, J-Y Dubuisson 2012 Diversity and evolution of the megaphyll in Euphyllophytes: phylogenetic hypotheses and the problem of foliar organ definition. C R Palevol 11:403–418.

Dh H, BR Feng, SS Yang, Y-B Huang, YX Tang 2012 The R2R3-MYB transcription factor gene family in maize. PLoS ONE 7:1–12.

Dh H, Z Liang, S Zhao, M-G Nan, L-SP Tran, K Lu, Y-B Huang, J-N Du 2012 The R2R3-MYB-gene family. Plant Mol Biol 41:577–585.

Evkaikina AI, L Berke, MA Romanova, E Proux-Wéra, AN Ivanova, C Rydin, K Pawlowski, VO Voitekhovskaja 2017 The Huperzia selago shoot tip transcriptome sheds new light on the evolution of leaves. Genome Biol Evol 5:11037.

Force A, M Lynch, FB Proulx, L Carretero-Paulet, S Cheng, J de Vries, P-M Plackett ARG, S Conway, EH Hewett Hazelton, KD Rabbinowitsch, NJ Long JA, EL Moan, JI Medford, MK Barton 1996 A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of Arabidopsis. Nature 379:66–69.

Galtier J 2010 The origins and early evolution of the megaphyllous leaves. Genome Biol Evol 2:393–405.

Harrison CJ, SB Corley, EC Moylan, DL Alexander, RW Scotland, JA Langdale 2005 Independent recruitment of a conserved developmental mechanism during leaf evolution. Nature 434:509–514.

He Q, DC Jones, W Li, F Xie, J Ma, R Sun, Q Wang, S Zhu, B Zhang 2016 Genome-wide identification of R2R3-MYB genes and expression analyses during abiotic stress in Gossypium raimondii. Sci Rep 6:1–14.

Jackson D, B Veit, S Hake 1994 Expression of maize KNOTTED1 related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. Development 120:405–413.

Jiang C-K, G-Y Rao 2020 Insights into the diversification and evolution of R2R3-MYB transcription factors in plants. Plant Physiol 183:637–655.

Jin H, C Martin 1999 Multifunctionality and diversity within the plant MYB-gene family. Plant Mol Biol 41:577–585.

Katoh K, DM Standley 2013 MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780.

Kleemen Z, A Sebastian, W Xu, D Grain, F Salsac, A Avon, N Berger, et al 2015 Analysis of the DNA-binding activities of the Arabidopsis R2R3-MYB transcription factor family by one-hybrid experiments in yeast. PLoS ONE 10:1–22.

Klempnauer KH, TJ Gonda, J Michael Bishop 1982 Nucleotide se-

quence of the retroviral leukemia gene c-myc and its cellular progen-}

itor c-myc: the architecture of a transduced oncogene. Cell 31:453–463.

Leebens-Mack JH, MS Barker, MK Deholsos, MA Greimendanner, SW Graham, I Grosse, et al 2019 One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574:679–685.

Li F-W, P Brouwer, L Carretero-Paulet, S Cheng, J de Vries, P-M Delaux, A Eily, et al 2018 Fern genomes elucidate land plant evolution and cyanobacterial symbioses. Nat Plants 4:460–472.

Lincoln C, J Long, J Yamaguchi, K Serikawa, S Hake 1994 A knotted1-like homeobox gene in Arabidopsis is expressed in the vegetative meri-
Vernot B, M Stolzer, A Goldman, D Durand 2008 Reconciliation with non-binary species trees. J Comput Biol 15:981–1006.

Wickett NJ, S Mirarab, N Nguyen, T Warnow, E Carpenter, N Matasci, S Ayyampalayam, et al 2014 Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc Natl Acad Sci USA 111:E4859–E4868.

Wilkins O, L Waldron, H Nahal, NJ Provart, MM Campbell 2009 Genotype and time of day shape the Populus drought response. Plant J 60:703–715.

Zhao Z, SU Andersen, K Ljung, K Dolezal, A Miotk, SJ Schultheiss, JU Lohmann 2010 Hormonal control of the shoot stem-cell niche. Nature 465:1089–1092.