Evolution of the YABBY gene family in seed plants

Cédric Finet,^{a,*1,2} Sandra K. Floyd,^{b,1} Stephanie J. Conway,^c Bojian Zhong,^d Charles P. Scutt,^a and John L. Bowman^{b,e,*}

^a Université de Lyon, Laboratoire de Reproduction et Développement des Plantes, UMR 5667, CNRS, INRA, Université Claude Bernard Lyon I, Ecole Normale Supérieure de Lyon, 46 allée d'Italie 69364 Lyon Cedex 07, France

^b School of Biological Sciences, Monash University, Melbourne, Victoria 3800, Australia

^c School of Botany, University of Melbourne, Melbourne, Victoria 3010, Australia

^d College of Life Sciences, Nanjing Normal University, Nanjing, China

^e Section of Plant Biology, University of California Davis, One Shields Ave., Davis, CA 95616, USA

*Authors for correspondence (e-mail: cedric.finet@ens-lyon.org [C. F.]; john.bowman@monash.edu [J. L. B.])

¹These two authors contributed equally to this work.

²Current address: Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY, UK.

SUMMARY Members of the YABBY gene family of transcription factors in angiosperms have been shown to be involved in the initiation of outgrowth of the lamina, the maintenance of polarity, and establishment of the leaf margin. Although most of the dorsal-ventral polarity genes in seed plants have homologs in non-spermatophyte lineages, the presence of YABBY genes is restricted to seed plants. To gain insight into the origin and diversification of this gene family, we reconstructed the evolutionary history of YABBY gene lineages in seed plants. Our findings suggest that

INTRODUCTION

The YABBY gene family encodes a class of transcription factors that are specific to seed plants. In angiosperms, its members play a pivotal role in lamina development in leaves and leaf-derived organs, for example, cotyledons, floral organs, and outer ovule integuments. The *Arabidopsis* genome contains six YABBY genes (Bowman and Smyth 1999; Sawa et al. 1999; Siegfried et al. 1999; Bowman 2000). Four of these (*FILAMENTOUS FLOWER (FIL), YAB2, YAB3,* and *YAB5*) are expressed in leaves and most leaf-derived organs— cotyledons, sepals, petals, stamens and carpels—, whereas expression of *CRABS CLAW (CRC)* and *INNER NO OUTER (INO)* are restricted to specific floral organs that are evolutionarily derived from leaves. In contrast to angiosperm YABBY genes, very little is known about the gene content and the putative role of YABBY genes in gymnosperms.

YABBY genes encode small proteins composed of an N-terminal zinc finger domain and a C-terminal helix-loop-helix motif (the so-called YABBY domain) similar to a high-mobility group box (Bowman and Smyth 1999). Phylogenetic studies resolve angiosperm YABBY genes into five subfamilies,

either one or two YABBY genes were present in the last common ancestor of extant seed plants. We also examined the expression of YABBY genes in the gymnosperms *Ephedra distachya* (Gnetales), *Ginkgo biloba* (Ginkgoales), and *Pseudotsuga menziesii* (Coniferales). Our data indicate that some YABBY genes are expressed in a polar (abaxial) manner in leaves and female cones in gymnosperms. We propose that YABBY genes already acted as polarity genes in the last common ancestor of extant seed plants.

represented in *Arabidopsis* by *FIL* + *YAB3*, *CRC*, *YAB2*, *YAB5* and *INO* (Yamada et al. 2004, 2011; Lee et al. 2005). Each of these five subfamilies includes early-diverging ANA-grade angiosperm sequences, suggesting that at least five YABBY genes were present in the last common ancestor of extant flowering plants (Yamada et al. 2011; Bartholmes et al. 2012). Nevertheless, the relationships between those five subfamilies remain enigmatic due to the lack of an appropriate ougroup.

In *Arabidopsis*, the vegetative YABBY genes are abaxially expressed in lateral organs and have redundant functions in activating laminar programs and repressing shoot apical meristem programs (Sawa et al. 1999; Siegfried et al. 1999; Eshed et al. 2004; Sarojam et al. 2010). Evidence of similar roles for vegetative YABBY genes has been functionally demonstrated in core eudicots (Golz et al. 2004) and monocots (Juarez et al. 2004a,b). In *Arabidopsis*, *CRC* is expressed abaxially in the valves and in placental tissues where it functions to specify carpel identity and establish carpel polarity, in addition to being required for proper nectary development (Bowman and Smyth 1999; Baum et al. 2001). Expression and functional analyses suggest that *CRC* acts as a regulator of carpel development across angiosperms (Yamaguchi et al. 2004; Fourquin et al.

2005; Orashakova et al. 2009) and nectaries in the core eudicots (Lee et al. 2005). Similarly, *Arabidopsis INO* expression is restricted to the outermost cell layer of the outer integument of the ovule and promotes outer integument growth (Villanueva et al. 1999). Abaxial expression has been also reported for *INO* orthologs in eudicots (McAbee et al. 2005), eumagnoliids (Lora et al. 2011), and several representatives of the ANA grade angiosperms (Yamada et al. 2003, 2011).

Previous searches of the complete genomes of land plants failed to identify YABBY genes in mosses and lycophytes, suggesting that the YABBY family found in flowering plants can be traced back only to the origin of seed plants (Floyd and Bowman 2007; Rensing et al. 2008; Banks et al. 2011). Although YABBY gene function has been investigated extensively in angiosperms, the same is not true of gymnosperm YABBY genes due to the paucity of model systems. To begin to address this question, we compiled a dataset of YABBY sequences across gymnosperms from available genomic resources and by experimental cloning. We used the sequences obtained to reconstruct the evolutionary history of seed plant YABBY gene lineages, and also investigated the expression of selected YABBY genes in gymnosperms.

MATERIALS AND METHODS

Plant material

Material of *Ephedra distachya* L. and *Gnetum gnemon* L. was collected from plants cultivated in Lyon Botanic Garden (Lyon, France). Material of *Cycas rumphii* Miq. was collected from plants cultivated in the New York Botanical Garden. Material of *Ginkgo biloba* L. was collected from plants cultivated in the Arnold Arboretum (Harvard University). Material of *Podocarpus reichei* J. Buchholz & N.E. Gray was collected by Francisco Vergara-Silva from plants cultivated at Universidad Nacional Autónoma de México (Mexico City, Mexico). *Taxus globosa* Schltdl. was wild-collected by Francisco Vergara-Silva at El Chico (Hidalgo, Mexico). Material of *Pseudotsuga menziesii* (Mirb.) Franco was collected from plants cultivated at UC Davis.

Data collection: Molecular work

Molecular cloning of YABBY cDNAs from fresh material of *C. rumphii*, *G. biloba*, *G. gnemon*, *P. menziesii*, *P. reichei*, and *T. globosa* was performed by PCR amplification. Total RNA was isolated from vegetative or reproductive tissues by the method of Chang et al. (1993). First strand cDNA was synthesized using RevertAidTM M-MuLV Reverse Transcriptase (Fermentas, Burlington, Canada). Partial fragments of YABBY cDNAs were amplified by using the degenerate primers 5'-TTGGATACAGTGACAGTGAAATGYGGNCAI TG and 5'-TGCCCAATTTT-TTGCTGCTGC. Alternatively,

two nested PCR reactions were performed (using first strand cDNA as template) with one forward primer (5'-GTIACIG-TIMGITGYGGICAYTG) and two reverse primers (outer 5': 5'-GCCCARTTYTTIGCIGC and inner 5': 5'-AAICKRAAR-TAIGCIGAIGGIAC) (Floyd et al. 2006). cDNA sequences were completed by 5' and 3' RACE using the MARATHON cDNA amplification kit (Clontech, Palo Alto, CA) or the SMART RACE kit (Clontech). For *E. distachya* and *G. biloba*, the cloning of YABBY sequences was performed by screening reproductive tissue cDNA libraries (Finet et al. 2013).

Data collection: Bioinformatic work

YABBY genes were identified in available transcriptomes by tBLASTn using a set of selected A. thaliana YABBY genes as a probe. The accession numbers of all the sequences used in this study are indicated in Supplementary Table S1. Species of origin are indicated by the following prefixes Aas: Asparagus asparagoides, Aba: Abies balsamea, Ach: Annona cherimola, Afi: Aristolochia fimbriata, Afo: Aquilegia formosa, Ala: Abies lasiocarpa, Ama: Antirrhinum majus, Asg: Annona squamosa, Ath: Arabidopsis thaliana, Atr: Amborella trichopoda, Cat: Cedrus atlantica, Cca: Cabomba caroliniana, Cch: Cephalotaxus harringtonia, Cia: Cryptomeria japonica, Cmi: C. micholitzii, Cru: C. rumphii, Cse: Chloranthus serratus, Eca: Eschscholzia californica, Edi: E. distachva, Gbi: G. biloba, Ggn: G. gnemon, Ghi: Gossypium hirsutum, Gro: Grevillea robusta, Ltu: Liriodendron tulipifera, Mgr: Magnolia grandiflora, Nad: Nuphar advena, Nal: Nymphaea alba, Nala: Nepenthes alata, Nco: Nymphaea colorata, Nja: Nuphar japonica, Osa: Oryza sativa, Pam: Persea americana, Pgl: Picea glauca, Phy: Petunia hybrida, Pla: Pinus lambertiana, Pma: P. macrophyllus, Pme: P. menziesii, Ppi: Pinus pinaster, Pre: P. reichei, Psi: Picea sitchensis, Pta: Pinus taeda, She: Saruma henryi, Sly: Solanum lycopersicum, Sve: Sciadopitys verticillata, Tba: T. baccata, Tcu: T. cuspidata, Tgl: T. globosa, Vvi: Vitis vinifera, Wmi: Welwitschia mirabilis, Zva: Zamia vazquezii. The novel sequences generated for this analysis have been deposited in the EMBL database with the following accession numbers: LN871570 (CruYABB), LN871571 (EdiYABB), LN871572 (EdiYABD), LN871573 (GbiYAB1B), LN871574 (GbiYAB2B), LN871575 (GbiYABC), LN871576 (PmeYABA), LN871577 (PmeYABC), LN871578 (PreYABB), LN871579 (PreYABD), LN871580 (TglYAB1A), LN871581 (TglYAB1C), LN871582 (ZvaYABB).

Phylogenetic analysis

Nucleotide sequences were aligned with MUSCLE (Edgar 2004), manually adjusted, and selected blocks were used for phylogenetic reconstruction. Maximum-likelihood (ML) searches were performed using RAxML v.8.1 (Stamatakis 2014) under the site-homogeneous $GTR+\Gamma$ model. One

hundred bootstrap replicates were conducted for support estimation. To detect potential impact of long-branch attraction, we also employed the site-heterogeneous CAT-GTR model implemented in PhyloBayes v.3.3 (Lartillot et al. 2009). We ran two independent chains for at least 10,000 cycles, discarded the first 1000 cycles as burn-in, and recovered no major topological changes with regard to ML tree. We also evaluated the potential influence of compositional heterogeneity in the dataset by a posterior predictive test based on the PhyloBayes analysis. Taxa with high compositional heterogeneity (Z-score >2) were removed for the 83-sequence analysis. For the 84-sequence analysis, coding nucleotide sequences were manually aligned as amino acid translations using Se-Al v2.0 (Rambaut 2002). We excluded ambiguously aligned sequence to produce an alignment of 113 amino acid characters in 84 YABBY gene sequences. The corresponding 339 nucleotides were used for subsequent Bayesian analysis. Bayesian phylogenetic analysis was performed using Mr. Bayes 3.2.1, run on multiple parallel processors (Ronquist et al. 2012). The characters were partitioned by position in codon allowing for estimation of different substitution rates at each position within the codons with a general time reversible model applied. The Bayesian analyses were run for 10,000,000 generations, which was sufficient for convergence of the two simultaneous runsconvergence took 653,000 generations. For the burn-in phase, 25% of the total number of saved trees was discarded. The different nucleotide sequence alignments and tree files are downloadable from TreeBASE (http://purl.org/phylo/treebase/ phylows/study/TB2:S18258).

Histology

Ephedra tissues were fixed in FAA (formaldehyde-acetic acid-ethanol) or 1% paraformaldehyde (PFA) overnight and then dehydrated through an ethanol series to 100% ethanol. The ethanol was gradually replaced with Histo-Clear II (National Diagnostics, Summerville, NJ, USA). The Histoclear was gradually replaced with Paraplast X-tra (Fisher Scientific, Pittsburgh, PA, USA) at 60°C. Embedded specimens were sectioned at 8-10 µm. Tissues were hydrated through an ethanol series to H₂0 and the slides were stained in 0.1% toluidine blue. Images were captured by using a Leica MZ12 binocular equipped with a Leica DFC320 digital camera. Shoot apices (including young leaves) of Ginkgo and Pseudotsuga were fixed in a solution of 1.5% glutaraldehyde, 1% paraformaldehyde, and 4% acrolein in Pipes buffer (84 mM Pipes, 8.4 mM EGTA, and 1.6 mM MgSO₄) at pH 6.8. Specimens were left in fixative a minimum of 24 h, then rinsed in Pipes buffer and dehydrated through an ethanol series to 95% ethanol. Specimens were then infiltrated with catalyzed monomer A of the JB-4 embedding kit (Polysciences, Warrington, PA, USA), and embedded in an oxygen-free environment following the basic protocol provided with the

kit. Blocks were serially sectioned at 5 μ m on an HM 355S (Microm, Thermo Scientific, Waltham, MA, USA) rotary microtome using glass knives. Slides were stained in 0.1% toluidine blue, examined, and photographed on a Zeiss Axioskop microscope equipped with a Zeiss Axiocam digital camera using bright field microscopy.

In situ hybridization

Tissues were fixed in FAA or PFA and subsequently embedded and sectioned as detailed above. Digoxigenin-labeled antisense and sense RNA probes were prepared from full-length or partial cDNA clones. Our prehybridization, hybridization, and posthybridization procedures were based on those of Jackson (1991) with some modifications. Whole mount in situ hybridizations were carried out as described by Hejátko et al. (2006). Slides were examined and photographed on a Leica MZ12 binocular equipped with a Leica DFC320 digital camera (*Ephedra*), a Zeiss Axioskop microscope equipped with a Zeiss Axiocam digital camera (*Ginkgo* and *Pseudotsuga*), or a Zeiss Lumar stereomicroscope equipped with a Zeiss Axiocam digital camera (whole-mount *Ginkgo*).

RESULTS AND DISCUSSION

The ancestral YABBY complement

Resolution of YABBY family evolution is essential to understanding the contribution that YABBY genes have made to seed plant evolution. However, the paucity of phylogenetically informative characters in YABBY genes and the high degree of similarity between family members has made it difficult to reconstruct robust phylogenies of the YABBY family across large evolutionary distances. In order to address this issue and to better understand YABBY diversification across all seed plants, we compiled a dataset of 43 YABBY sequences from across gymnosperms.

Our phylogenetic analyses indicate that YABBY homologs in gymnosperms are resolved into four distinct clades. The clades include sequences from most of the major lineages of gymnosperm taxa (Fig. 1A). Although clades A, C, and D are well supported, the branch leading to the clade B has only weak support. However, that clade B consists of sequences from each of the major gymnosperm clades and reflects their accepted phylogenetic relationships (Lu et al. 2014) provides additional support for the clade. The grouping of Coniferales I, Coniferales II, and Gnetales within clade B is additionally supported by a 15-bp deletion in the region encoding the YABBY domain (Fig. 1B). High variability in the relative position of Cycadales and Ginkgoales at the base of clade B could explain the poor statistical values observed. These four clades indicate that at least four YABBY genes were present in the last common ancestor of extant gymnosperms.



Fig. 1. Phylogeny of YABBY genes in gymnosperms. (A) Phylogram of the 43-sequence analyses. RAxML maximum-likelihood analyses and PhyloBayes Bayesian analyses were conducted under the $GTR+\Gamma$ and the CAT-GTR model, respectively. Support values obtained after 100 bootstrap replicates (BP) and Bayesian posterior probabilities (PP) are shown for all branches. Scale bar indicates number of changes per site. (B) Alignment of the YABBY domain showing the 15-bp deletion in clade B.

We next investigated the relationships between angiosperm and gymnosperm YABBY clades. We assembled a dataset of 107 YABBY sequences, including the list of 43 gymnosperm sequences (Fig. 1A) compiled in the present work and sequences spanning the phylogenetic diversity of angiosperms. Phylogenetic analyses of the 107-sequence dataset identify gymnosperm and angiosperm YABBY genes as two distinct monophyletic groups (Figs. 2A and S1A). Because composition biases among synonymous substitutions have been shown to potentially generate conflicting phylogenies (Cox et al. 2014), we identified and removed sequences having such a bias. Phylogenetic analysis of the resulting 83-sequence dataset supports monophyly of gymnosperm YABBY genes (Fig. S1B). We also observed a few sequences with faster rates of molecular evolution. However, phylogenetic analysis of a dataset in which sequences with longer branches have been removed supports an alternative topology, in which gymnosperm YABBY genes were not resolved as monophyletic. In this analysis, a gymnosperm clade composed of clades B and D is sister to an angiosperm clade composed of CRC and FIL/YAB3, whereas the two gymnosperm clades A and C form a polytomy with the angiosperm clade composed of the YAB5, YAB2, and INO clades (Figs. 2B and S1C).

Given there is not an appropriate outgroup sequence to root the trees, we rooted the trees based on parsimony, minimizing the numbers of gene gains and losses required, assuming both angiosperm and gymnosperm taxa are monophyletic groups. Based on this rooting, the two alternative topologies lead to different evolutionary scenarios. Considering gymnosperm YABBYs as paraphyletic, with gymnosperm sequences that group with two distinct angiosperm clades (Fig. 2B), would suggest the presence of two YABBY genes in the last common ancestor of extant seed plants. Conversely, the monophyly of gymnosperm YABBY genes (Fig. 2A) would suggest that all spermatophyte YABBY genes derived from a single gene in the common ancestor of extant seed plants and that diversification of the gene family proceeded independently in both angiosperms and gymnosperms. Previous analyses with very limited gymnosperm datasets also suggested that gymnosperm YABBY genes could be either monophyletic (Bartholmes et al. 2012) or paraphyletic (Yamada et al. 2011). It has been proposed that prior to the divergence of angiosperms and gymnosperms, a whole genome duplication occurred in the ancestral seed plant and another in the ancestral angiosperm (Jiao et al. 2011). In addition, analyses of individual gene families that have included homologs from all major lineages of land plants suggest that multiple paralogs arose in the seed-plant common ancestor after the divergence of monilophytes: ARFs (Finet et al. 2013), class III HD-Zips (Floyd et al. 2006; Prigge and Clark 2006), class IV HD-Zips (Zalewski et al. 2013), LBDs (Chanderbali et al. 2015), PINs (Bennett et al. 2014). There is less evidence for any recent whole genome duplications within the gymnosperm lineage, although analyses to date have focused only on the conifers (Pavy et al. 2012; Nystedt et al. 2013). Either of the two alternative topologies proposed could be consistent with this history of genome duplication, although if gymnosperm and angiosperm YABBY genes are both monophyletic, it would imply a gene loss in the seed-plant common ancestor and following the hypothesized genome duplication in the seed plant common ancestor.

The relationship among the angiosperm YABBY genes is very robust in our analyses with regard to any common phylogenetic bias we tested for. We can distinguish a first cluster composed of the clades CRC and FIL, and a second cluster in which the clades YAB5, YAB2, and INO group together. Noticeably, *CRC* and *FIL* share the common feature of being "strong" patterning genes by promoting abaxial cell fate (Eshed et al. 1999; Sawa et al. 1999; Siegfried et al. 1999), whereas *YAB2, YAB5*, and *INO* seem to have a milder effect on patterning but a stronger role in tissue growth (Siegfried et al. 1999; Eshed et al. 2004; Golz et al. 2004; Sarojam et al. 2010). Furthermore, the two genes directing the development of angiosperm specific innovations, the carpel (*CRC*) and ovule outer integument (*INO*), were derived from different YABBY gene ancestors.



Fig. 2. Phylogeny of YABBY genes in seed plants. Phylogenetic analyses generate two alternative topologies. (A) Gymnosperm YABBY genes form a monophyletic clade, suggesting a single YABBY gene in the ancestor of extant seed plants. (B) Gymnosperm YABBY genes are paraphyletic, suggesting at least two YABBY genes in the common ancestor of extant seed plants.

However, this evolutionary scenario of the YABBY family within angiosperms is not congruent with conclusions from other phylogenetic studies (Yamada et al. 2011; Bartholmes et al. 2012), where a clade composed of FIL, CRC, and INO was observed. Given the inherent instability in tree topology depending upon inclusion/exclusion of taxa and mode of analysis, the phylogenetic relationships between the gymnosperm and angiosperm YABBY genes presented here remain hypotheses for future investigation.

There are reports of YABBY genes in other lineages of eukaryotes, such as within the Chlorophyta (Worden et al. 2009). We searched GenBank for sequences similar to YABBY genes in species other than land plants and found similar sequences in both the Chlorophyta and fungi. However, the region of similarity was restricted to the helix-loop-helix portion of the YABBY domain, and in none of the sequences was a C2C2 zinc finger domain found, either amino or carboxyl to the region of similarity within the helix-loop-helix domain. Furthermore, the putative helixloop-helix domains were not conserved across taxa within the Chlorophyta, in which there are 31 pairwise amino acid differences between the sequences in Ostreococcus and Micromonas (Fig. S2A). In contrast, other than CRC and INO, which have evolved specialized functions in the carpel and ovule of angiosperms, all the other seed plant YABBY genes have accumulated fewer than 10 mutations (and most five or less) from the predicted ancestral seed plant YABBY amino acid sequence (Fig. S2A). Phylogenetic analyses place the fungal and Chlorophyta sequences approximately equidistant between seed plant YABBY sequences and canonical HMG box sequences (Fig. S2B). Although it remains possible that the fungal and Chlorophyta sequences may be more directly evolutionarily related to seed plant YABBY genes, it seems equally, or perhaps more likely, that they represent independent derivations from ancestral HMG box sequences.

YABBY genes and the evolution of seed plant leaves

Although most genes that direct leaf development in angiosperms predate the evolutionary origin of seed-plant leaves, the YABBY genes originated along the lineage leading to seed plants (Floyd and Bowman 2007; Sarojam et al. 2010). In *Arabidopsis*, simultaneous loss-of-function of the four vegetative YABBY genes (*FIL*, YAB2, YAB3, and YAB5) results in polarity defects, loss of lamina expansion, and reactivation of shoot meristem genetic programs (Sarojam et al. 2010). Those findings led some authors to propose that YABBY genes act to transform an ancestral shoot-specific network into a leafspecific one (Floyd and Bowman 2010; Sarojam et al. 2010).

We performed in situ hybridization experiments to investigate expression patterns of YABBY genes in vegetative tissues in *Ginkgo* and *Pseudotsuga*. In *Pseudotsuga*, expression of *PmeYABC* does not show any clear evidence of polar patterning. The staining is visible throughout the provascular strand in both leaf primordia and young leaves (Fig. 3, D and E). In contrast to PmeYABC expression, in Ginkgo, strong expression of GbiYABC is evident on the abaxial side of leaf primordia (Fig. 4, B and C). Staining from the GbiYABC probe is also visible in the leaf primordium provasculature (Fig. 4B). The expression pattern of Ginkgo gene GbiYAB1B is very similar to GbiYABC (Fig. 4, F and G), being abaxial in young primordia and maintaining a similar, though slightly weaker pattern in older primordia. Expression of GbiYAB2B is also restricted to the abaxial side of the youngest leaf primordia (Fig. 4, D and E). In older primordia, its expression remains abaxial but is restricted to distal, marginal positions (Fig. 4E). In older leaf primordial, staining of GbiYAB1B and GbiYAB2B appears to be complementary, with the former present in the basal-abaxial domain and the later present in the distal-abaxial domain. Phylogenetic analysis identified GbiYAB1B and GbiYAB2B as paralogs that originated from a gene duplication event unique to the Ginkgo lineage (Fig. 1A). In situ expression data are consistent with subfunctionalization following that duplication event.

The three Ginkgo genes residing in gymnosperm clades B and C appear to be expressed abaxially in leaves, similar to expression of YABBY genes in the angiosperm YAB2, YAB5, and FIL/YAB3 clades. Based on parsimony, this phylogenetic distribution suggests that this pattern may represent the ancestral expression pattern of the entire YABBY family, irrespective of the two alternative evolutionary scenarios presented, implying that other expression patterns observed are derived. Interestingly, expression pattern of GbiYABC, which stains the outside (abaxial domain) of the bifurcating lamina and petiole (Fig. 4B), appears to be complementary to the expression domain of the Ginkgo class III HD-zip gene, GbC3HDZ1, which stains in the inside (adaxial domain) of the notch of the bifurcation (Floyd and Bowman 2006). This observation suggests that the mutual antagonism between class III HD-ZIP and YABBY genes characterized in angiosperms could have arisen prior to the split between gymnosperms and angiosperms.

There are some interesting differences in YABBY gene expression between *Pseudotsuga*, *Ginkgo*, and *Arabidopsis*. In contrast to *Arabidopsis* and Gingko, YABBY expression in the conifer *Pseudotsuga* (*PmeYABC*) was not clearly polar in young leaf primordia. *Pseudostuga* leaves, like other Pinaceae, are linear and needle-like. The needle-like leaves of *Pseudotsuga* differ from leaves of *Ginkgo* and flowering plants in that they lack laminar expansion and clear dorsoventral organization of mesophyll and epidermal tissues (Fig. 3C). Based upon the parsimony criterion, the needle-like leaves of conifers were likely derived from ancestors with more broadly laminate leaves. The evolution of conifer leaves was likely the result of numerous changes to the leaf developmental program. Our data suggest that at least one such change may have been a loss of YABBY expression polarity during leaf development.



Fig. 3. In situ hybridization of YABBY gene in *Pseudotsuga*. Purple stain indicates the localized hybridization of the antisense RNA probe. Brown staining indicates the presence of oxidized tannins. (A) Longitudinal section through the *P. menziesii* shoot apex. (B) Transverse section through apex at the level of leaf initiation. Leaf primordia are labeled P0 (youngest) through P4 (oldest). (C) Cross section of leaf. (D, E) *PmeYABC* is expressed in the provascular strand in both leaf primordial and young leaves. endo: endodermis; If: leaf; lp: leaf primordium; pv: provascular tissue; res: resin duct; sam: shoot apical meristem; and vb: vascular bundle.

In *Arabidopsis*, FIL/YAB3 expression in leaf primordia becomes restricted to marginal positions reflecting the early differentiation of the distal end of the developing leaf (Siegfried et al. 1999). In contrast, *Ginkgo* leaves continue to grow throughout their development by distal meristematic activity (Bower 1884; Mundry and Stützel 2004). Expression of *GBiYAB2B* is limited to abaxial-distal regions in older leaf primordia. This expression pattern is unlike that of FIL/YAB3 and suggests that YABBY expression is associated with laminar expansion in gymnosperm leaves with distal (apical) and angiosperm leaves with marginal meristematic activity.

YABBY genes and the evolution of seed plant ovules

We performed in situ hybridization experiments to investigate expression patterns of YABBY genes in female reproductive structures in *Ephedra*. *EdiYABD* transcripts were detected in ovule integuments and at the base of sterile bracts (Fig. 5B). The staining is throughout the innermost bract but is restricted to the abaxial side in outer bracts (Fig. 5C). The other gene investigated, *EdiYABB*, is expressed throughout the ovule integument, with a clear enrichment in the outermost cell layer of the integument (Fig. 5, D and E). *EdiYABB* is also weakly expressed at the base of sterile bracts and may be more strongly expressed in the abaxial side.

Gymnosperm ovules have a single integument with the micropyle positioned opposite the stalk (orthotropous). Many angiosperms, including all ANA-grade taxa, have two integuments, and the ovule is commonly bent over (anatropous), positioning the micropyle close to the stalk. The second integument is not simply a reiteration of the first; the two clearly have different attributes and evolutionary origins. The inner integument most likely originated directly from the fusion of telomes (Zimmermann 1952) or sporangiophores (Benson 1904; Kenrick and Crane 1997). However, the origin of the second integument remains one of the cryptic and controversial questions concerning the origin of the flowering plants. It has been proposed that the outer integument evolved later, from a cupule, a leaf-like structure surrounding one or more ovules



Fig. 4. In situ hybridization of YABBY genes in Ginkgo. Purple stain indicates the localized hybridization of the antisense RNA probe. Brown staining indicates the presence of oxidized tannins. (A) Longitudinal section through the G. biloba shoot apex and leaf. (B) Leaf primordia showing GbiYABC expression restricted to abaxial side. (C) In situ hybridization of whole-mount apices confirms the abaxial expression of GbiYABC in leaf primordia labeled P0-P3 from youngest to oldest. (D) GbiYAB2B expression is restricted to abaxial side of leaf primordia. (E) Expression of GbiYAB2B is still abaxially restricted in older leaf primordia, vet weaker than in youngest primordia. (F) Expression patterns of GbiYAB1B are very similar to those of GbiYAB2B at the shoot apex. ab: abaxial; ad: adaxial; lf: leaf; lp: leaf primordium; pv: provascular tissue; and sam: shoot apical meristem.

(Stebbins 1974; Gasser et al. 1998). Consistent with an origin from a leaf, in angiosperms, the YABBY gene *INO* is a molecular marker specific to the outer integument of ovules (Villanueva et al. 1999) and development of the inner integument does not require YABBY expression or function. In the present study, we show that the two genes *EdiYABB* and *EdiYABD* are strongly expressed in the ovule integument in *Ephedra*. Our finding of YABBY expression in the single integument of a gymnosperm is surprising in the light of the classical view that the gymnosperm integument is homologous to the inner integument in angiosperms. However, Gnetales female cones are known to show highly derived characters. Investigation of gene expression in ovules of other nonflowering plants is needed to determine whether YABBY genes are universally associated with the gymnosperm integument. These data should be informative with respect to the evolution of ovule integuments during seed plant evolution.

In summary, either one or two YABBY genes were present in the last common ancestor of extant seed plants. That some YABBY genes are expressed in a polar (abaxial) manner in leaves and female cones in gymnosperms suggests that in the last common ancestor of seed plants at least one member of the YABBY gene family was likely expressed in a polar manner. We propose that YABBY genes already acted as polarity genes



in the last common ancestor of extant seed plants, and as polarity is intimately linked to lamina expansion, that YABBY gene function is associated with the evolution of laminar leaves in seed plants.

Acknowledgments

We thank Jean-Pierre Grienay (Lyon Botanic Garden), Andrew Groover, and Dave Johnson (USDA Institute of Forest Genetics, Placerville) for assistance with obtaining plants used in this study. We thank Eric Brenner (formerly of New York Botanical Garden) for providing *C. rumphii* cDNA libraries and Francisco Vergara-Silva for providing RNA samples of *T. globosa* and *P. reichei*. We thank Judith Lucas and Karen Yip for assistance with cloning YABBY cDNAs from *P. menziesii* and *G. biloba*. We thank LinHua Sun for assistance with phylogenetic analyses. S. K. F. was supported, in part, by the UC Davis

Fig. 5. In situ hybridization of YABBY genes in Ephedra. Purple stain indicates the localized hybridization of the antisense RNA probe. Brown staining indicates the presence of oxidized tannins. (A) Longitudinal section through the E. distachya female cone. The female cone usually bears whorls of bracts, among which only the uppermost whorl is fertile. Each ovule has a leathery outer envelope (chlamys) and a membranous inner envelope (integument). (B) Biovulate female cone showing EdiYABD expression in integument. (C) Sterile bracts showing EdiYABD expression in successive whorls of bracts. Staining throughout the innermost bract becomes restricted to the abaxial side in outer bracts. (D) Female cone showing EdiYABB expression in integument. (E) At higher magnification, EdiYABB shows a stronger expression in the outermost cell layer of integument. ab: abaxial; ad: adaxial; br: bract; ch: chlamys; int: integument; and nu: nucellus.

Katherine Esau postdoctoral fellowship and the Australian Research Council Future Fellowship FT100100763. This work was also supported by the National Natural Science Foundation of China (31570219), the Natural Science Foundation of Jiangsu Province (BK20150971), and the Priority Academic Program Development of Jiangsu Higher Education Institutions to B. Z., and Australian Research Council Discovery Project DP110100070 to J. L. B.

REFERENCES

Banks, J. A., et al. 2011. The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. Science 332: 960–963.

Bartholmes, C., Hidalgo, O., and Gleissberg, S. 2012. Evolution of the YABBY gene family with emphasis on the basal eudicot *Eschscholzia californica* (Papaveraceae). *Plant Biol.* 14: 11–23.

- Baum, S. F., Eshed, Y., and Bowman, J. L. 2001. The Arabidopsis nectary is an ABC-independent floral structure. Development 128: 4657–4667.
- Bennett, T., et al. 2014. Paralogous radiations of PIN proteins with multiple origins of noncanonical PIN structure. *Mol. Biol. Evol.* 31: 2042–2060.
- Benson, M. 1904. Telangium scotti, a new species of Telangium (Calymmatotheca) showing structure. Ann. Bot. 78: 161–177.
- Bower, F. 1884. On the comparative morphology of the leaf in the vascular cryptograms and gymnosperms. *Philos. Trans. R. Soc. B* 175: 565–615.
- Bowman, J. L. 2000. The YABBY gene family and abaxial cell fate. Curr. Opin. Plant Biol. 3: 17–22.
- Bowman, J. L., and Smyth, D. R. 1999. CRABS CLAW, a gene that regulates carpel and nectary development in *Arabidopsis*, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development* 126: 2387–2396.
- Chanderbali, A. S., He, F., Soltis, P. S., and Soltis, D. E. 2015. Out of the water: origin and diversification of the LBD gene family. *Mol. Biol. Evol.* 32: 1996–2000.
- Chang, S., Puryear, J., and Cairney, J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* 11: 113–116.
- Cox, C. J., Li, B., Foster, P. G., Embley, T. M., and Civan, P. 2014. Conflicting phylogenies for early land plants are caused by composition biases among synonymous substitutions. *Syst. Biol.* 63: 272–279.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792–1797.
- Eshed, Y., Baum, S. F., and Bowman, J. L. 1999. Distinct mechanisms promote polarity establishment in carpels of *Arabidopsis. Cell* 99: 199–209.
- Eshed, Y., Izhaki, A., Baum, S. F., Floyd, S. K., and Bowman, J. L. 2004. Asymmetric leaf development and blade expansion in *Arabidopsis* are mediated by KANADI and YABBY activities. *Development* 131:2997–3006.
- Finet, C., Berne-Dedieu, A., Scutt, C. P., and Marlétaz, F. 2013. Evolution of the ARF gene family in land plants: old domains, new tricks. *Mol. Biol. Evol.* 30: 45–56.
- Floyd, S. K., and Bowman, J. L. 2006. Distinct developmental mechanisms reflect the independent origins of leaves in vascular plants. *Curr. Biol.* 16: 1911–1917.
- Floyd, S. K., and Bowman, J. L. 2007. The ancestral developmental tool kit of land plants. *Int. J. Plant Sci.* 168: 1–35.
- Floyd, S. K., and Bowman, J. L. 2010. Gene expression patterns in seed plant shoot meristems and leaves: homoplasy or homology? *J. Plant Res.* 123: 43–55.
- Floyd, S. K., Zalewski, C. S., and Bowman, J. L. 2006. Evolution of class III homeodomain-leucine zipper genes in streptophytes. *Genetics* 173: 373–388.
- Fourquin, C., Vinauger-Douard, M., Fogliani, B., Dumas, C., and Scutt, C. P. 2005. Evidence that crabs claw and tousled have conserved their roles in carpel development since the ancestor of the extant angiosperms. *Proc. Natl. Acad. Sci. USA* 102: 4649–4654.
- Gasser, C. S., Broadhvest, J., and Hauser, B. A. 1998. Genetic analysis of ovule development. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49: 1–24.
- Golz, J. F., Roccaro, M., Kuzoff, R., and Hudson, A. 2004. Graminifolia promotes growth and polarity of Antirrhinum leaves. Development 131: 3661–3670.
- Hejátko, J., Blilou, I., Brewer, P. B., Friml, J., Scheres, B., and Benková, E. 2006. In situ hybridization technique for mRNA detection in whole mount *Arabidopsis* samples. *Nat. Protoc.* 1: 1939–1946.
- Jackson, D. (1991). In situ hybridisation in plants. In D. Bowles, S. Gurr, and and M. McPherson (eds.). *Molecular Plant Pathology, A Practical Approach*. Oxford University Press, UK.
- Jiao, Y., et al. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–100.
- Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A., and Timmermans, M. C. P. 2004a. MicroRNA-mediated repression of rolled leaf1 specifies maize leaf polarity. *Nature* 428: 84–88.
- Juarez, M. T., Twigg, R. W., and Timmermans, M. C. P. 2004b. Specification of adaxial cell fate during maize leaf development. *Development* 131: 4533–4544.
- Kenrick, P., and Crane, P.(1997) The Origin and Early Diversification of Land Plants: A Cladistic Study. Smithsonian Institution Press, Washington, DC.

- Lartillot, N., Lepage, T., and Blanquart, S. 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25: 2286–2288.
- Lee, J.-Y., Baum, S. F., Oh, S.-H., Jiang, C.-Z., Chen, J.-C., and Bowman, J. L. 2005. Recruitment of CRABS CLAW to promote nectary development within the eudicot clade. *Development* 132: 5021–5032.
- Lora, J., Hormaza, J. I., Herrero, M., and Gasser, C. S. 2011. Seedless fruits and the disruption of a conserved genetic pathway in angiosperm ovule development. *Proc. Natl. Acad. Sci. USA* 108: 5461–5465.
- Lu, Y., Ran, J.-H., Guo, D.-M., Yang, Z.-Y., and Wang, X.-Q. 2014. Phylogeny and divergence times of gymnosperms inferred from singlecopy nuclear genes. *PLoS ONE* 9:e107679.
- McAbee, J. M., Kuzoff, R. K., and Gasser, C. S. 2005. Mechanisms of derived unitegmy among impatiens species. *Plant Cell* 17: 1674– 1684.
- Mundry, M., and Stützel, T. 2004. Morphogenesis of leaves and cones of male short-shoots of *Ginkgo biloba* L. Flora 199: 437–452.
- Nystedt, B., et al. 2013. The Norway spruce genome sequence and conifer genome evolution. *Nature* 497: 579–584.
- Orashakova, S., Lange, M., Lange, S., Wege, S., and Becker, A. 2009. The CRABS CLAW ortholog from California poppy (*Eschscholzia californica*, Papaveraceae), EcCRC, is involved in floral meristem termination, gynoecium differentiation and ovule initiation. *Plant J.* 58: 682–693.
- Pavy, N., Pelgas, B., Laroche, J., Rigault, P., Isabel, N., and Bousquet, J. 2012. A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. *BMC Biol.* 10: 84.
- Prigge, M. J., and Clark, S. E. 2006. Evolution of the class III HD-Zip gene family in land plants. *Evol. Dev.* 8: 350–361.
- Rambaut, A. 2002. Se-Al v. 2.0a11: sequence alignment program. http:// tree.bio.ed.ac.uk/software/seal/
- Rensing, S. A., et al. 2008. The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319: 64–69.
- Ronquist, F., et al. 2012. Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61: 539–542.
- Sarojam, R., et al. 2010. Differentiating Arabidopsis shoots from leaves by combined YABBY activities. *Plant Cell* 22: 2113–2130.
- Sawa, S., Ito, T., Shimura, Y., and Okada, K. 1999. Filamentous flower controls the formation and development of *Arabidopsis* inflorescences and floral meristems. *Plant Cell* 11: 69–86.
- Siegfried, K. R., Eshed, Y., Baum, S. F., Otsuga, D., Drews, G. N., and Bowman, J. L. 1999. Members of the YABBY gene family specify abaxial cell fate in *Arabidopsis*. *Development* 126: 4117– 4128.
- Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312– 1313.
- Stebbins, G. L. (1974) Flowering plants: evolution above the species level. Harvard University Press, Cambridge, Massachusetts.
- Villanueva, J. M., Broadhvest, J., Hauser, B. A., Meister, R. J., Schneitz, K., and Gasser, C. S. 1999. Inner no outer regulates abaxialadaxial patterning in *Arabidopsis* ovules. *Genes Dev.* 13: 3160–3169.
- Worden, A. Z., et al. 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas. Science* 324: 268–272.
- Yamada, T., Ito, M., and Kato, M. 2003. Expression pattern of Inner no outer homologue in *Nymphaea* (water lily family, Nymphaeaceae). *Dev. Genes Evol.* 213: 510–513.
- Yamada, T., Ito, M., and Kato, M. 2004. YABBY2-homologue expression in lateral organs of *Amborella trichopoda* (Amborellaceae). *Int. J. Plant Sci.* 165: 917–924.
- Yamada, T., Yokota, S., Hirayama, Y., Imaichi, R., Kato, M., and Gasser, C. S. 2011. Ancestral expression patterns and evolutionary diversification of YABBY genes in angiosperms. *Plant J.* 67: 26–36.
- Yamaguchi, T., Nagasawa, N., Kawasaki, S., Matsuoka, M., Nagato, Y., and Hirano, H.-Y. 2004. The YABBY gene drooping leaf regulates carpel specification and midrib development in *Oryza sativa*. *Plant Cell* 16: 500–509.

Finet et al.

Zalewski, C. S., Floyd, S. K., Furumizu, C., Sakakibara, K., Stevenson, D. W., and Bowman, J. L. 2013. Evolution of the class IV HD-Zip gene family in streptophytes. *Mol. Biol. Evol.* 30: 2347–2365.

Zimmermann, W. 1952. Main results of the telome theory. *Paleobotanis* 1: 456–470.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Fig. S1. Phylogram of the 107-sequence analysis. RAxML maximum-likelihood analysis was conducted under the GTR+ Γ model.

Fig. S2. YABBY-like sequences from taxa other than seed plants.

Table S1. List of YABBY sequences selected for phylogenetic analyses.