Independent origins of syringyl lignin in vascular plants

Jing-Ke Weng, Xu Li, Jake Stout, and Clint Chapple*

Department of Biochemistry, Purdue University, 175 South University Street, West Lafayette, IN 47907

Edited by Richard A. Dixon, The Samuel Roberts Noble Foundation, Ardmore, OK, and approved April 1, 2008 (received for review February 20, 2008)

Lycophytes arose in the early Silurian (≈400 Mya) and represent a major lineage of vascular plants that has evolved in parallel with the ferns, gymnosperms, and angiosperms. A hallmark of vascular plants is the presence of the phenolic lignin heteropolymer in xylem and other sclerified cell types. Although syringyl lignin is often considered to be restricted in angiosperms, it has been detected in lycophytes as well. Here we report the characterization of a cytochrome P450-dependent monooxygenase from the lycophyte *Selaginella moellendorffii*. Gene expression data, cross-species complementation experiments, and *in vitro* enzyme assays indicate that this P450 is a ferulic acid/coniferaldehyde/coniferyl alcohol 5-hydroxylase (F5H), and is capable of diverting guaiacyl-substituted intermediates into syringyl lignin biosynthesis. Phylogenetic analysis indicates that the *Selaginella* F5H represents a new family of plant P450s and suggests that it has evolved independently of angiosperm F5Hs.

convergent evolution | DFRC | F5H | P450 | *Selaginella*

Lignin is an aromatic heteropolymer that is deposited most abundantly in the secondary cell walls of vascular plants. It provides structural rigidity to the plant body while enabling individual tracheary elements to withstand the tension generated during water transport; it also serves a defensive role against herbivores and pathogens (1). Lignins are derived mainly from the phenylpropanoid monomers *p*-coumaryl, coniferyl, and sinapyl alcohol, which give rise to *p*-hydroxyphenyl, guaiacyl, and syringyl subunits when incorporated into the lignin polymer (2). In angiosperms, three cytochrome P450-dependent monooxygenases (P450s) are involved in the biosynthesis of lignin monomers, cinnamate 4-hydroxylase (C4H), *p*-coumaroyl shikimate/quinate 3′-hydroxylase (C3′H), and ferulic acid/coniferaldehyde/coniferyl alcohol 5-hydroxylase (F5H) (Fig. 1) (3). C4H and C3′H are responsible for phenylpropanoid 4 and 3-hydroxylation (4, 5), respectively, whereas F5H catalyzes the 5-hydroxylation of coniferaldehyde and coniferyl alcohol, leading to the formation of syringyl lignin (6, 7). Lignin monomer composition has been found to vary among major phyla of vascular plants (2). Generally, ferns and gymnosperms deposit lignins that are derived primarily from guaiacyl monomers together with a small proportion of *p*-hydroxyphenyl units, whereas angiosperm lignins are guaiacyl/syringyl copolymers that also can contain some *p*-hydroxyphenyl monomers. This distribution suggests that F5H may be a relatively recent addition to plants' biochemical repertoire. Nevertheless, there are older reports in the literature in which syringyl monomers have been detected in lignins from lycophytes, including species of *Selaginella* (8–12), by using histochemical reagents and by today's standards relatively crude chemical methods. These results have been verified recently by using more modern techniques (13). How species that diverged from angiosperms >400 Mya (14) acquired the ability to synthesize syringyl lignin is unknown.

Results

Lignin Composition Analysis in Representative Vascular Plants. We used the derivatization followed by reductive cleavage (DFRC) method (15), a procedure specific for β-O-4-linked lignin units, to examine the lignin composition in representative species of major vascular plant taxa. We found that, although guaiacyl lignin derivatives can be detected from all of the species examined, syringyl lignin derivative is only present in the three angiosperm species examined and *S. moellendorffii* (Fig. 2). The lignin of *S. moellendorffii* has a high content of syringyl subunits, with a mole percentage of >70%. Notably, a *Lycopodium* species, which represents another lycophyte lineage, does not deposit syringyl lignin.

Identification and Characterization of *SmF5H* Candidates. To be able to synthesize syringyl lignin, we hypothesized that the *Selaginella* genome encodes a phenylpropanoid 5-hydroxylase capable of diverting guaiacyl lignin precursors to syringyl lignin biosynthesis. To clone *F5H* candidates from the *Selaginella* genome, we initially adopted a nested PCR method by using degenerate primers designed against the regions that are uniquely conserved among angiosperm F5H proteins. In contrast, this approach failed to return any potential *SmF5H* candidates, suggesting that *SmF5H* may be divergent in sequence from the angiosperm *F5H*s. As an alternative approach, we searched for *SmF5H* candidates in a previously published *S. moellendorffii* expressed sequence tag (EST) dataset (16). Although no obvious *SmF5H* homolog was identified at the first glance, we found three P450-encoding ESTs (DN837695, DN837863, and DN839545) with ≈40% similarities to members of the flavonoid hydroxylase (CYP75) and F5H (CYP84) families. Flavonoid hydroxylases and F5Hs represent two closely related plant P450 families (17) and function similarly in their ability to catalyze metahydroxylation reactions on parahydroxylated phenylpropanoids. For this reason, we considered these three ESTs as potential candidates for *SmF5H*. We then isolated the full-length cDNAs corresponding to these three ESTs and used them for further functional analysis.

Complementation of *Arabidopsis* F5H-deficient Mutant by *SmF5H*. To test the function of *Sm*F5H candidates *in planta*, each was introduced into the *Arabidopsis* F5H-deficient *fah1-2* mutant (18, 19) under the control of the constitutive 35S cauliflower mosaic virus promoter. Whereas two of them (DN837695 and DN839545) failed to rescue, one (DN837863) complemented the *Arabidopsis fah1-2* mutant and was then designated as *SmF5H* (Fig. 3). Although all lines appeared normal under white light, when observed under UV light, three of four lines of *fah1-2*-p35S::*SmF5H* T2 transgenic plants show complete complementation of the reduced epidermal fluorescence (*ref*)

Author contributions: J.-K.W. and C.C. designed research; J.-K.W., X.L., and J.S. performed research; and J.-K.W. and C.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. EU032589–EU032601).

*To whom correspondence should be addressed. E-mail: chapple@purdue.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0801696105/DCSupplemental.

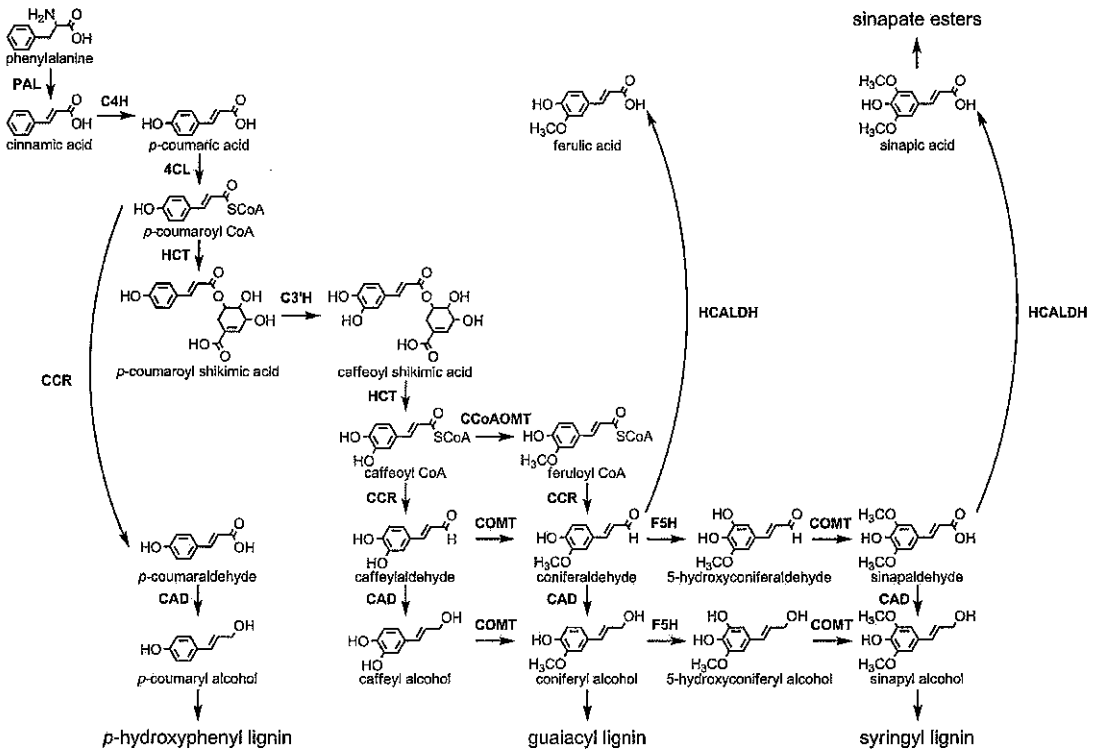

Fig. 1. The plant phenylpropanoid pathway. PAL, phenylalanine ammonia-lyase; 4CL, 4-hydroxy cinnamoyl CoA ligase; C4H, cinnamate 4-hydroxylase; HCT, hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase; C3'H, *p*-coumaroyl shikimate/quinate 3'-hydroxylase; CCoAOMT, caffeoyl CoA *O*-methyl transferase; CCR, (hydroxy) cinnamoyl CoA reductase; HCALDH, hydroxycinnamaldehyde dehydrogenase; F5H, ferulic acid/coniferaldehyde/coniferyl alcohol 5-hydroxylase; COMT, caffeic acid/5-hydroxyferulic acid *O*-methyltransferase; CAD, (hydroxy) cinnamyl alcohol dehydrogenase.

phenotype normally exhibited by the *fah1* mutant, which arises because of a block in sinapoylmalate biosynthesis. The other line shows only partial complementation probably because of a position effect associated with the insertion site of the transgene in the genome (Fig. 3*a*). To evaluate complementation quantitatively, we analyzed leaf methanol extracts of these transgenic plants by HPLC and found that their sinapoylmalate levels ranged from 10% to 100% of the level found in the wild type (Fig. 3*b*), which is consistent with the UV phenotypes observed. The successful complementation of the UV phenotype of *fah1-2* by this *SmF5H* candidate indicates that this gene can take the place of *AtF5H* in *Arabidopsis* soluble phenylpropanoid biosynthesis.

To test whether *Sm*F5H can rescue the syringyl lignin-deficient phenotype of *fah1-2*, we transformed into the mutant a construct in which the *SmF5H* gene was under the control of the *Arabidopsis C4H* promoter, a strategy previously reported to efficiently target transgene expression to vascular tissue (6). Stem cell wall samples were prepared from the 17 independent *fah1-2*-p*AtC4H*::*SmF5H* T1 transgenic lines and were analyzed for lignin composition by using the DFRC method. Syringyl lignin derivatives were detected from all of the transgenic lines examined, with syringyl/(syringyl plus guaiacyl) mole percentages ranging from 13% to 70% [supporting information (SI) Table S1], similar or much higher than that observed in *Arabidopsis* wild-type plants (6). When brought to homozygosity in the T2 generation, one line showed a mol% syringyl value of >80% (Fig. 3*c*). These results indicate that *SmF5H* is the functional equivalent of *AtF5H* in lignin biosynthesis.

Considering that *Sm*F5H also shows a level of sequence similarity to flavonoid hydroxylases, we transformed *SmF5H* into the *Arabidopsis transparent testa 7* (*tt7*) mutant, a mutant defective in its flavonoid 3′-hydroxylase (F3′H) gene (20), to test the hypothesis that *Sm*F5H may be a bifunctional enzyme that also can hydroxylate flavonoids. Seeds from 10 independent transgenic lines were examined at T2 generation. None of them shows complementation of the seed *transparent testa* phenotype, a phenotype caused by a decreased level of condensed tannin accumulation in seed coat (data not shown). This result suggests that *Sm*F5H does not possess F3′H activity.

Kinetic Analysis of *Sm*F5H Substrate Specificity. To assess the substrate specificity of *Sm*F5H *in vitro*, we expressed recombinant *Sm*F5H in the WAT11 yeast strain and prepared the microsomal protein for kinetic assays. In assays conducted by using ferulic acid as a substrate, we found that the K_m and V_{max} of *Sm*F5H for this substrate were 0.3 mM and 2.3 pkat·mg^{-1}, respectively. In contrast, assays conducted with coniferaldehyde demonstrated that the K_m and V_{max} of *Sm*F5H were 0.6 μM and 2.5 pkat·mg^{-1}, respectively, and the corresponding values for coniferyl alcohol were 1.1 μM and 1.9 pkat·mg^{-1} (Fig. 4). These data indicate that coniferaldehyde and coniferyl alcohol are the preferred substrates for *Sm*F5H *in vitro*, which is similar to what has been reported for angiosperm F5Hs (7, 21). Yeast expressed *Sm*F5H also was assayed against naringenin, a substrate for F3′H. However, no activity was detected (data not shown). These *in vitro* data confirm the previous *in vivo* results which suggested that *Sm*F5H does not have flavonoid hydroxylase activity.

Tissue Specificity of Syringyl Lignin Accumulation and *SmF5H* Expression. It has been shown previously that syringyl lignin biosynthesis is developmentally regulated and its deposition is restricted to the cells of the sclerified parenchyma in *Arabidopsis* (6). To investigate the tissue specificity of syringyl lignin accu-

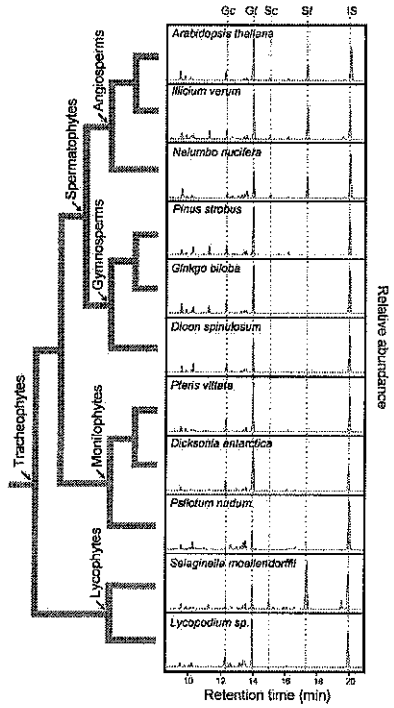

Fig. 2. DFRC GC analysis of lignin in representative plant taxa. A simplified version of the plant phylogenetic tree is used to indicate the phylogenetic relationships between the plant species whose lignin monomer composition was evaluated by using DFRC lignin analysis. Authentic standards of hydroxycinnamyl alcohol diacetates were used to identify the monomer specific peaks. G/S, guaiacyl/syringyl lignin derivative; *c/t*: *cis/trans*; IS, internal standard.

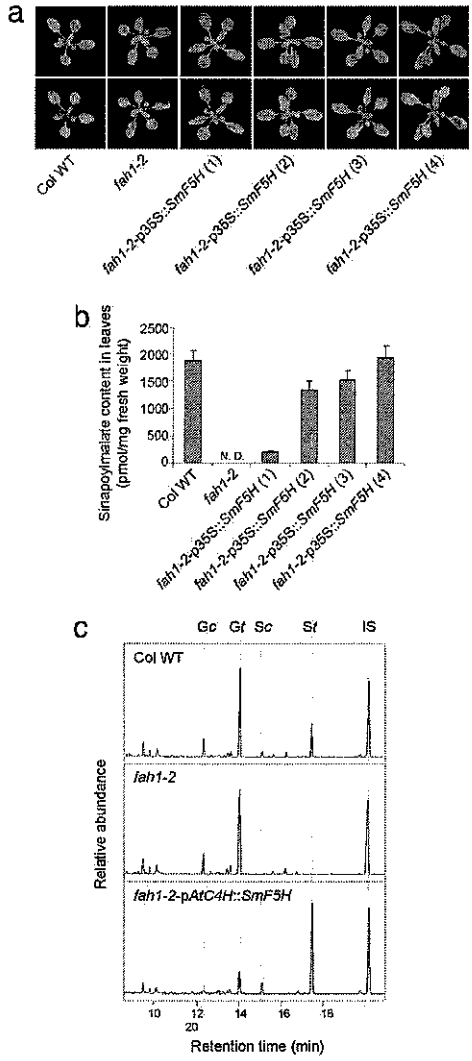

Fig. 3. Complementation of the *Arabidopsis fah1-2* mutant by *Sm*F5H. (*a*) *Arabidopsis* Columbia WT (Col WT), *fah1-2*, and four independent T2 transgenic lines of *fah1-2*-p35s::*SmF5H* rosette-stage plants were photographed under visible light (*Upper*) and UV light (*Lower*). Blue fluorescence reflects the presence of sinapoylmalate and complementation of the *fah1* phenotype. (*b*) Restoration of the leaf sinapoylmalate production in *fah1-2*-p35s::*SmF5H* transgenic plants quantified by HPLC. Error bars represent 1 SD of triplicate samples (ND, not detectable). (*c*) GC chromatograms of the DFRC lignin analysis in Col WT, *fah1-2*, and a representative of *fah1-2*-p*AtC4H*::*SmF5H* T2 transgenic plant. G/S, guaiacyl/syringyl lignin derivative; *c/t*: *cis/trans*; IS, internal standard.

mulation in *S. moellendorffii*, we performed Mäule histochemical staining on *S. moellendorffii* stem cross-sections. Red staining was observed only in the cortex and not in the xylem, indicating that syringyl lignin is predominantly deposited in this tissue (Fig. 5*a*). Consistent with these results, *in situ* hybridization to detect *SmF5H* mRNA accumulation clearly indicated expression in cortical cells (Fig. 5 *b* and *c*), where syringyl lignin is deposited. Hybridization signal also was observed in the phloem cells surrounding the xylem, suggesting that F5H also may be involved in the biosynthesis of secondary metabolites other than lignin.

Phylogenetic Analysis of *Sm*F5H. Multiple sequence alignment analysis of *Sm*F5H revealed that *Sm*F5H contains all of the signature motifs that are conserved among P450s (Fig. S1). *Sm*F5H shows <40% sequence identity to previously identified P450s and thus defines a new P450 family according to the P450 nomenclature (22). The P450s that are the most closely related to *Sm*F5H are CYP75 and CYP84 members, with their sequence identities to *Sm*F5H at ≈37% (Fig. S1). Considering that *Selaginella* CYP73 (C4H) and CYP98 (C3′H) homologs share >60% sequence identity with their angiosperm counterparts, this result suggests that *Sm*F5H is not likely to be orthologous to angiosperm F5Hs. It is noteworthy that the recent availability of the *S. moellendorffii* whole genome sequence allowed us to identify 10 *Selaginella* P450s that may be related phylogenetically to *Sm*F5H, and these proteins show sequence identities to *Sm*F5H ranging from 39–44%. These related *Selaginella* P450s may give hints to the evolutionary history of *Sm*F5H and thus were included in the phylogenetic analysis described below.

To more rigorously infer the phylogeny of *Sm*F5H, we performed a Bayesian phylogenetic analysis that includes *Sm*F5H, related P450s from representative plant taxa, and the 10 *Sm*F5H-related P450s from the *S. moellendorffii* genome (Fig. 6). Although CYP73 and CYP98 families appear to be conserved from mosses to flowering plants, *Sm*F5H is not clustered with the angiosperm F5H clade (CYP84 family), but belongs to a unique clade of *Selaginella* P450s that is distinct from all of the known P450s, suggesting an independent origin of F5H in *Selaginella*. Similar results also can be inferred from a phylogenetic tree generated independently by using the Neighbor-Joining method (Fig. S2). Two P450s in the *Sm*F5H-containing clade (DN837695 and DN839545) are the two *Sm*F5H candidates that failed to complement *fah1-2*. To test the possibility that the other *Selaginella* P450s in this clade also may possess F5H activity, we expressed them in yeast and assayed their activities toward F5H

PLANT BIOLOGY

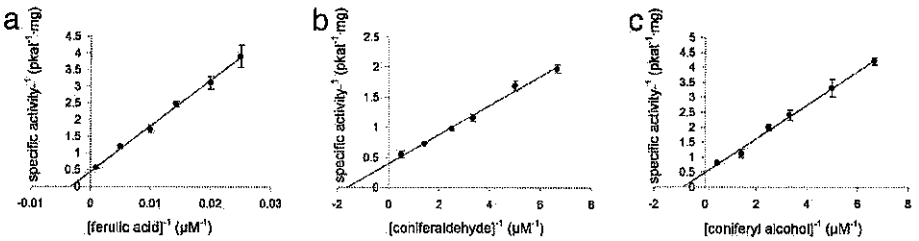

Fig. 4. Kinetic analysis of *Sm*F5H-catalyzed substrate 5-hydroxylation reactions. (*a*–*c*) Assays of ferulic acid (*a*), coniferaldehyde (*b*), and coniferyl alcohol (*c*) are conducted by using recombinant *Sm*F5H. The error bars represent 1 SD for triplicate assays.

substrate coniferyl alcohol (Fig. S3). Whereas *Sm*F5H and *At*F5H, as positive controls, show complete conversion of coniferyl alcohol to 5-hydroxyconiferyl alcohol, no such activity was detected for any of the 10 *Sm*F5H-related P450s, indicating that the enzyme activity of 5-hydroxylating guaiacyl-substituted lignin intermediates is unique for *Sm*F5H and not shared by the other *Selaginella* P450s in this clade. The angiosperm F3′H substrate naringenin also was tested as a substrate for these enzymes in parallel assays. Although yeast expressing *At*F3′H completely converted naringenin to eriodictyol, none of the 10 *Sm*F5H-related P450s did so (data not shown), indicating that these *Selaginella* enzymes are not F3′H analogs.

Discussion

Lycophytes today comprise ≈1,200 species in the three extant orders Lycopodiales, Selaginellales, and Isoetales, accounting for only a small and inconspicuous group of living vascular plants. In contrast, the ancestors of these plants once dominated the Earth's flora during the Carboniferous period and can be traced back to ≈420 Mya, 280 million years earlier than the emergence of angiosperms (23). The distribution of syringyl lignin in the plant kingdom suggested two possible models for the evolution of F5H. First, the enzyme could have arisen early in plant evolution, was lost in ferns and gymnosperms, but was not lost in angiosperms or Selaginella. Alternatively, F5H could have evolved independently in lycophyte and angiosperm lineages after they had diverged. Our results suggest that the second model is correct and that F5H from *Selaginella* is functionally equivalent to, but phylogenetically independent from, angiosperm F5Hs. This conclusion is further supported by the observation that syringyl lignin derivatives are not detected in extant members of the Lycopodiaceae (11, 24) and have not been found in fossils of the extinct lycophyte *Sigillaria ovata* (order Lepidodendrales) (24). Taken together, these data suggest that the Selaginellales may be the only lycophyte order that acquired the ability to synthesize syringyl lignin, although, if confirmed, early reports of syringyl lignin in Isoetes and Huperzia (11, 25) may indicate that the enzymatic activities required for syringyl lignin biosynthesis are more widespread within the lycophytes.

Although independent occurrence of identical enzyme function in distinct lineages is not commonly observed, similar cases have been presented in the literature. For example, limonene synthase, a plant terpenoid synthase, has been characterized from both angiosperm species and one gymnosperm species, *Abies grandis* (26). Despite their functional resemblance, phylogenetic analysis suggests that the genes that encode limonene synthase in angiosperms and gymnosperms evolved independently (27). In gibberellin biosynthesis, *ent*-kaurene oxidase and *ent*-kaurenoic acid oxidase from higher plants are encoded by P450s from CYP701 and CYP88 families, respectively (28, 29), whereas the analogous enzymes in fungus *Gibberella fujikuroi* are encoded by very distinct P450s from CYP503 and CYP68 families (30, 31). This phenomenon also has been attributed to be the result of convergent evolution (32).

It is interesting to consider what evolutionary advantages may have led to the independent invention of syringyl lignin in two lineages of vascular plants. For example, in angiosperms, syringyl lignin is often associated with fiber cells that have an important role in mechanical support. This correlation has led to the hypothesis that syringyl lignin may be superior to guaiacyl lignin in its ability to strengthen cell walls into which it is incorporated (33). Our study shows that, in *Selaginella*, syringyl lignin accumulates primarily in the sclerified cortical cells, suggesting that these cells may play an important role in support of the plant body. Alternatively, a recent study of resistance responses of wheat to pathogen attack revealed that syringyl lignin was hyperaccumulated in the plant cell wall in response to fungal penetration, suggesting that syringyl lignin also may provide a selective advantage in defense against pathogens (34).

In conclusion, we identified and characterized a unique cytochrome P450 from *Selaginella* that is capable of diverting guaiacyl-substituted intermediates into syringyl lignin biosynthesis. Our phylogenetic analysis suggested that the occurrences of syringyl lignin in lycophytes and angiosperms might be independent. The gene identified in this article also adds a potential tool for engineering lignin biosynthesis in gymnosperms where syringyl lignin is absent.

Materials and Methods

Plant Materials. *S. moellendorffii* was obtained from Plant Delights Nursery and grown in a local greenhouse under 50% shade cloth. *A. thaliana* was grown under a 16-h light/8-h dark photoperiod at 100 μE·m^{-2}·s^{-1} at 22°C.

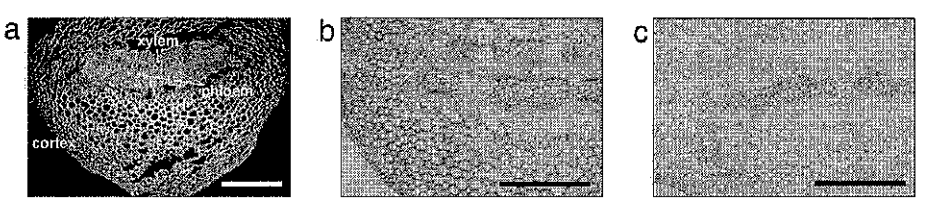

Fig. 5. Tissue-specific deposition of syringyl lignin in *S. moellendorffii* is associated with the expression pattern of *Sm*F5H. (*a*) Mäule staining of transverse sectioned *S. moellendorffii* stem. Red color indicates the presence of syringyl lignin, and brown color indicates the presence of guaiacyl lignin. (*b* and *c*) *In situ* hybridization of *SmF5H* mRNAs in *Selaginella* transverse sections using antisense (*b*) and sense (*c*) *SmF5H* probes. Purple color indicates the presence of the *SmF5H* mRNA. (Scale bars: 200 μm.)

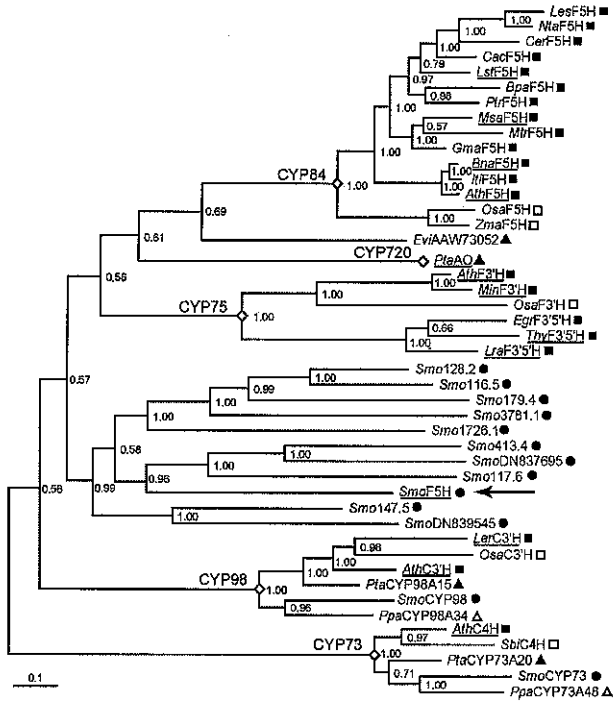

Fig. 6. Bayesian inference of the phylogenetic relationship between *Sm*F5H and other related land plant P450s. Bayesian posterior probabilities are indicated on the right of branches. The scale measures evolutionary distance in substitutions per amino acid. The details of the P450 sequences used in this tree were summarized in Table S2. An arrow is used to point to *Sm*F5H. Proteins for which functions have been confirmed genetically and/or biochemically are underlined. The taxonomy information of the sequences is indicated by the symbol at the right of the gene name (filled square, dicot; open square, monocot; filled triangle, gymnosperm; filled circle, lycophyte; open triangle, bryophyte).

Transgenic *Arabidopsis*. The *SmF5H* ORF was cloned by RT-PCR using a gene-specific primer pair cc1560 (5′-tcactcagtcagtcatgaatc-3′) and cc1559 (5′-ccttttgtttggatcaagcttgatagagatg-3′), and A-T cloned into pGEM T-Easy (Promega) to generate pCC0819. To generate the p35S::*SmF5H* construct, the *SmF5H* ORF was released by EcoRI and SpeI digestion and ligated into EcoRI/SpeI-digested pCC0790, a pCAMBIA1390-derived binary vector. To generate the p*AtC4H*::*SmF5H* construct, the *SmF5H* ORF was released from pCC0819 by EcoRI digestion and ligated into EcoRI-digested pCC0916, a pBI101-based vector containing a 2,977-bp fragment of the Arabidopsis C4H promoter. Arabidopsis transformation was performed by using the floral dip method (35).

Sinapoylmalate Analysis. Three-week-old *Arabidopsis* rosette leaves were extracted in 50% methanol and analyzed by reverse-phase HPLC as previously described (36).

Lignin Analysis. Mäule staining of lignin in microtome sections of *S. moellendorffii* stem was conducted as described (18). Stem cell wall samples were prepared as previously described (6), and DFRC lignin analysis was performed essentially according to Lu and Ralph (15).

Yeast Expression of *Sm*F5H and Enzyme Assays. To generate the pYeDP60-*SmF5H* construct, the *SmF5H* ORF without the start codon was PCR amplified by using the primers 5′-ccggaattcaatctctcctcgatcatgggg-3′ and 5′-cggggtacct-ggatcaagcttgatagagatg-3′, digested with EcoRI and KpnI, and subjected to a three-way ligation in the presence of BamHI/KpnI-digested pYeDP60 and a FLAG tag linker with 5′-*Bam*H1 and 3′-*Eco*R1 overhangs. The resulting pYeDP60-*SmF5H* plasmid was transformed into the WAT11 yeast strain. Yeast growth, preparation of yeast microsomal extracts, and enzyme kinetics assays were conducted essentially as described (7).

***In Situ* Hybridization.** To study the localization of *SmF5H* mRNA in *Selaginella* stem tissue, 8-μm sections of *Selaginella* stem were subjected to *in situ* hybridization as previously described (37). To generate *SmF5H* antisense or sense probes, pCC0819 was linearized with NcoI or SpeI and transcribed from the SP6 promoter or the T7 promoter, respectively, by using the SP6/T7 transcription kit (Roche Applied Science).

Phylogenetic Analysis. The sequences and their corresponding GenBank accession numbers used in the analysis are summarized in Table S2. The amino acid alignment of plant P450s was created by using T-Coffee (38). The Bayesian phylogenetic tree was inferred by using MRBAYES Version 3.1.1 (39). Bayesian analysis of amino acid alignments invoked a comparable model (aamodelpr, mixed; nset, 6; rates, invgamma). The Neighbor-joining tree was constructed by using MEGA Version 4.0 (40).

ACKNOWLEDGMENTS. We thank J. Banks (Purdue University) for providing a variety of plant materials, J. Ralph for providing the standards for DFRC lignin analysis, M. Zanis for advice on the phylogenetic analysis, and J. Ogas for critical review of the manuscript. This work was supported by National Science Foundation Grant IOB-0450289. This article is journal paper no. 2008-18334 of the Purdue University Agricultural Experiment Station.

1. Campbell MM, Sederoff RR (1996) Variation in lignin content and composition (mechanisms of control and implications for the genetic improvement of plants). *Plant Physiol* 110:3–13.
2. Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. *Annu Rev Plant Biol* 54:519–546.
3. Humphreys JM, Chapple C (2002) Rewriting the lignin roadmap. *Curr Opin Plant Biol* 5:224–229.
4. Franke R, *et al.* (2002) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *Plant J* 30:33–45.
5. Teutsch HG, *et al.* (1993) Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. *Proc Natl Acad Sci USA* 90:4102–4106.
6. Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. *Proc Natl Acad Sci USA* 95:6619–6623.
7. Humphreys JM, Hemm MR, Chapple C (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. *Proc Natl Acad Sci USA* 96:10045–10050.
8. Logan KJ, Thomas BA (1985) Distribution of lignin derivatives in plants. *New Phytol* 99:571–585.
9. Faix O, Gyzas E, Schweers W (1977) Comparative investigations on different fern lignins. *Holzforschung* 31:137–144.
10. Erickson M, Miksche GE (1974) Characterization of pteridophyta lignins by oxidative-degradation. *Holzforschung* 28:157–159.
11. Towers GHN, Gibbs RD (1953) Lignin chemistry and the taxonomy of higher plants. *Nature* 172:25–26.
12. White E, Towers GHN (1967) Comparative biochemistry of lycopods. *Phytochemistry* 6:663–667.
13. Jin ZF, *et al.* (2005) Proof of the presence of guaiacyl-syringyl lignin in *Selaginella tamariscina*. *J Wood Sci* 51:424–426.
14. Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. *Nature* 389:33–39.
15. Lu F, Ralph J (1998) The DFRC method for lignin analysis. 2. Monomers from Isolated Lignins. *J Agric Food Chem* 46:547–552.
16. Weng JK, Tanurdzic M, Chapple C (2005) Functional analysis and comparative genomics of expressed sequence tags from the lycophyte *Selaginella moellendorffii*. *BMC Genomics* 6:85.
17. Nelson DR, Schuler MA, Paquette SM, Werck-Reichhart D, Bak S (2004) Comparative genomics of rice and *Arabidopsis*. Analysis of 727 cytochrome P450 genes and pseudogenes from a monocot and a dicot. *Plant Physiol* 135:756–772.
18. Chapple CC, Vogt T, Ellis BE, Somerville CR (1992) An *Arabidopsis* mutant defective in the general phenylpropanoid pathway. *Plant Cell* 4:1413–1424.
19. Meyer K, Cusumano JC, Somerville C, Chapple CC (1996) Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. *Proc Natl Acad Sci USA* 93:6869–6874.
20. Koornneef M, Luiten W, de Vlaming P, Schram AW (1982) A gene controlling flavonoid-3′-hydroxylation in *Arabidopsis*. *Arabidopsis Inf Serv* 19:113–115.
21. Osakabe K, *et al.* (1999) Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. *Proc Natl Acad Sci USA* 96:8955–8960.
22. Werck-Reichhart D, Feyereisen R (2000) Cytochromes P450: A success story. *Genome Biol* 1: REVIEWS3003.
23. Stewart WN, Rothwell GW (1993) *Paleobotany and the Evolution of Plants* (Cambridge Univ Press, New York).
24. Logan KJ, Thomas BA (1987) The distribution of lignin derivatives in fossil plants. *New Phytol* 105:157–173.
25. Towers GHN, Maass WSG (1965) Phenolic acids and lignins in the lycopodiales. *Phytochemistry* 4:57–66.

26. Bohlmann J, Steele CL, Croteau R (1997) Monoterpene synthases from grand fir (*Abies grandis*). cDNA isolation, characterization, and functional expression of myrcene synthase, (−)-(4S)-limonene synthase, and (−)-(1S,5S)-pinene synthase. *J Biol Chem* 272:21784–21792.
27. Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: Molecular biology and phylogenetic analysis. *Proc Natl Acad Sci USA* 95:4126–4133.
28. Helliwell CA, *et al.* (1998) Cloning of the *Arabidopsis* ent-kaurene oxidase gene GA3. *Proc Natl Acad Sci USA* 95:9019–9024.
29. Helliwell CA, Chandler PM, Poole A, Dennis ES, Peacock WJ (2001) The CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc Natl Acad Sci USA* 98:2065–2070.
30. Tudzynski B, Hedden P, Carrera E, Gaskin P (2001) The P450–4 gene of *Gibberella fujikuroi* encodes ent-kaurene oxidase in the gibberellin biosynthesis pathway. *Appl Environ Microbiol* 67:3514–3522.
31. Rojas MC, Hedden P, Gaskin P, Tudzynski B (2001) The P450–1 gene of *Gibberella fujikuroi* encodes a multifunctional enzyme in gibberellin biosynthesis. *Proc Natl Acad Sci USA* 98:5838–5843.
32. Hedden P, Phillips AL, Rojas MC, Carrera E, Tudzynski B (2001) Gibberellin biosynthesis in plants and fungi: A case of convergent evolution? *J Plant Growth Regul* 20:319–331.
33. Li L, *et al.* (2001) The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. *Plant Cell* 13:1567–1586.
34. Menden B, Kohlhoff M, Moerschbacher BM (2007) Wheat cells accumulate a syringyl-rich lignin during the hypersensitive resistance response. *Phytochemistry* 68:513–520.
35. Weigel D, Glazebrook J (2002) *Arabidopsis: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY).
36. Franke R, *et al.* (2002) Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of *Arabidopsis*. *Plant J* 30:47–59.
37. Vielle-Calzada JP, *et al.* (1999) Maintenance of genomic imprinting at the *Arabidopsis medea* locus requires zygotic DDM1 activity. *Genes Dev* 13:2971–2982.
38. Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302:205–217.
39. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755.
40. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599.

PURDUE
UNIVERSITY

DEPARTMENT OF BIOCHEMISTRY

March 30, 2009

Colleen L. Gabauer, Ed.D.
Director of Interdisciplinary Graduate Programs
The Graduate School
Purdue University
Ernest C. Young Hall, Rm. B-85
155 S. Grant Street
West Lafayette, IN 47907

Dear Colleen:

I'm pleased to have the opportunity to write a letter of recommendation on behalf of my student Jing-Ke Weng as the Plant Biology Training Group's nominee for the 2009 PULSe Paper of the Year Award.

Jing-Ke's research focuses on the isolation of genes encoding cytochrome P450-dependent monooxygenases involved in phenylpropanoid biosynthesis from *Selaginella moellendorffii*. Selaginella deposits syringyl lignin, a type of lignin that is normally thought of as being associated with flowering plants. This observation is noteworthy in that over 400,000,000 years of evolutionary time separates Selaginella (a spore-forming club moss) from flowering plants, and that all of the taxa that branched off the plant family tree between the lycopsids (the clade to which Selaginella belongs) and the angiosperms (flowering plants) deposit only guaiacyl lignin. According to our current understanding of lignin biosynthesis, the ability to synthesize syringyl lignin requires two additional enzymes over and above those required for guaiacyl lignin synthesis: a 5-hydroxylase and an *O*-methyltransferase. Jing-Ke has succeeded in isolating the genes encoding both of these Selaginella enzymes. The paper with which Jing-Ke is applying for this award (*Weng, J-K, Li X, Stout J, Chapple C (2008) Independent origins of syringyl lignin in vascular plants. Proc. Natl. Acad. Sci. U.S.A.* ***105****: 7887-7892*) focuses on the identification of the 5-hydroxylase, which he demonstrated is a cytochrome P450-dependent monooxygenase (SmF5H). Jing-Ke accomplished this feat by completing a Selaginella EST project (*Weng JK, Tanurdzic M, Chapple C. (2005) Functional analysis and comparative genomics of expressed sequence tags from the lycophyte Selaginella moellendorffii. BMC Genomics* ***6****:85*) and using candidate SmF5H genes to complement the corresponding mutant of Arabidopsis. The most important finding from this paper is that SmF5H and angiosperm F5Hs are of independent origin and are thus examples of convergent evolution.

An aspect of our efforts to publish this work highlighted illustrated an important aspect of Jing-Ke's character and work ethic that might be of interest to the award selection committee. When we submitted this work to PNAS, we received two supportive reviews, but the editor declined the paper. Because the genome sequence for Selaginella is now available, Jing-Ke included in his P450 phylogenetic tree the ten Selaginella P450 sequences most closely related to SmF5H and commented in the manuscript that all of these proteins are distantly related to flavonoid hydroxylases. The editor's decision letter focused on the fact that we had tested "only" three P450 candidates, one of which had complemented the Arabidopsis *fah1* mutant, and question what the other eight might do, and whether any of the proteins are also flavonoid hydroxylases. To be honest, Jing-Ke and I thought it was outrageous that in order to publish on this one protein, we should have to characterize ten others for their activity in two different pathways, and I indicated to Jing-Ke that we should just submit the manuscript elsewhere. Jing-Ke felt strongly that this manuscript belonged in PNAS, and his response to my suggestion was that "If we do that, the editor will have 'won'." I really admire that kind of perseverance. Instead of giving up or trying to rebut what we thought was an unreasonable barrier to publication, Jing-Ke teamed up with a post doctoral fellow in my lab, and together, expressed all of the proteins in yeast, and tested them for activity against all relevant substrates! None of the proteins turned out to have flavonoid hydroxylase activity toward the substrates

we employed, and no other protein had F5H activity. Having addressed the editors concerns, the manuscript was accepted without further delay, and was even highlighted on the NSF website http://www.nsf.gov/news/news_summ.jsp?cntn_id=111597.

What is particularly relevant to this award nomination is that Jing-Ke completed all of the experiments (save for the previously mentioned collaborative experiments on SmF5H substrate specificity and some early experiments on Selaginella lignin composition performed by another student) with virtually no input from me. He composed all of the figures, and wrote the entire manuscript on his own. It required only minor editorial changes from me and was entirely logical in its flow. In other words, it did not need the reordering of paragraphs or corrections to the scientific logic that manuscripts written by students so often require. Jing-Ke had referenced all of the appropriate literature, including a number of papers dating back to the 1950s and 1960s. Simply put, it was an extremely impressive effort. I have since learned that this was no fluke. Jing-Ke's next paper from his Ph.D. research, similarly written almost entirely by him, is currently undergoing full editorial review at Science.

In conclusion, I feel comfortable stating that Jing-Ke's paper is one of the two top student-written papers to have ever been published from my laboratory. If you have any questions about Jing-Ke or this paper, please don't hesitate to contact me.

Yours sincerely,

Clint Chapple
Distinguished Professor of Biochemistry and Head